Modelos, Simulación y Teoría de la Decisión

Ing. Alejandro Roberti
Ing. Gustavo Chijani — Ing Esteban Gidekel
2021

Diseño de portada: **aero@optimiza.org**

Versión digital disponible (e-book) **http://optimiza.org**

No es fácil preguntar.

— *No veo porqué. Bastaría con preguntar: ¿quién me serviría mejor como Primer Ministro? Solo eso.*
— *Si, pero el Rey no sabe que significa “me serviría mejor”. Podría querer decir que el hombre elegido entregara el Valle de Orgoreyn, o que se exiliará, o asesinará al Rey. Podría querer decir muchas cosas que el Rey no esperaría ni aceptaría nunca.*
— *La pregunta tendría que ser muy precisa.*
— *Si, pero serían necesarias muchas preguntas, y también el Rey ha de pagar su precio.*

…

— *Si, solo una pregunta tiene respuesta, Genry, y yo conocemos la respuesta… La vida solo es posible a causa de esa permanente e intolerable incertidumbre: no conocer lo que vendrá*

Ursula K. Le Guin
“La mano izquierda de la oscuridad”

ÍNDICE

Presentación

La serie optimiza comenzó en el año 1999 con Optimiza1, que cubría los contenidos de las asignaturas: Investigación Operativa e Ingeniería de Procesos.

Optimiza fue agregando ediciones hasta llegar, en 2011 a Optimiza9, último volumen con los contenidos completos de ambas asignaturas.

A partir de 2013, Optimiza10, se dividió en dos libros, el principal, pasó a cubrir solamente los contenidos de Investigación Operativa y la mayor parte de Ingeniería de Procesos, mientras que el capítulo “Diseño de Experimentos” pasó a integrar el Libro 2.

En esta edición, Optimiza12, se presenta en seis libros, los temas Investigación Operativa e Ingeniería de Procesos (libros 1 a 4), Ingeniería de Procesos (Diseño de Experimentos) (libro 5) y Modelos, Simulación y Teoría de la Decisión (libro 6)

Este último volumen, editado fuera de la serie desde 2013, se completa, actualiza e incorpora como último aporte a ella en esta Edición, por lo cual la tomaremos como la primera formal de este tema. Creemos que cubre las expectativas de los estudiantes de la Licenciatura de Sistemas de Información, Ingeniería en Informática y la Licenciatura en Informática de las Universidades Nacionales de Luján y del Noroeste de la Provincia de Buenos Aires.

Alejandro Roberti, enero 2020-octubre 2021

CAPITULO 1

TEORÍA GENERAL DE SISTEMAS

Sistemas

Un sistema es un modelo conceptual en sí mismo, creado para comprender asociaciones de objetos, reales o virtuales, que tienen una funcionalidad específica que los convierten en una colección o conjunto.

Obviamente existen muchas maneras de definir el concepto "*sistema*". Una de ellas sería a partir de su propia *función*:

> *"Conjunto de funciones, virtualmente referenciada sobre ejes reales o abstractos"*

Esta definición nos sugiere que debemos contar con una descripción documentada de lo que se pretende que sea necesario para lograr cierto comportamiento (función general o patrón de comportamiento) en un posible aspecto (eje). Para ello, hacen falta un conjunto de funciones integradas en un todo (sistema)

Alternativamente, otra manera sería definir a partir de los *componentes* del sistema:

> *"Conjunto de **elementos** dinámicamente relacionados que conforman una actividad para **alcanzar un objetivo**, para lo cual deben operar sobre datos, energía y/o materia y, eventualmente, proveer información"*

En esta segunda definición aparecen dos conceptos: los términos "***elementos***" y "***alcanzar un objetivo***".

Esos *elementos* pueden ser físicos o abstractos. Es correcto, y corriente, llamarlos *objetos*, los objetos son aquellos elementos que pueden ser observados y estudiados y hasta aprendidos cuando se representan abstractamente en la mente mediante un proceso de generalización.

En ese plano, los ***objetos*** se caracterizan por algunos parámetros importantes. Por ejemplo,

- La IDENTIDAD. Es una propiedad, o un conjunto de propiedades varias, que permiten a un objeto diferenciarse de otros. Esto se aplica a individuos o colecciones. Pérez (individuo humano) es diferente de González (individuo humano también); los humanos (especie, colección de individuos) son diferentes de los equinos (que también es una especie)
- El COMPORTAMIENTO. Está relacionado con su funcionalidad y determina lo que puede realizar o de qué manera puede responder ante estímulos enviados por otros o a otros objetos.

- El ESTADO. Es el conjunto de los valores que asumen cada uno de sus atributos en un instante de tiempo dado.

Ambiente, suprasistemas y subsistemas

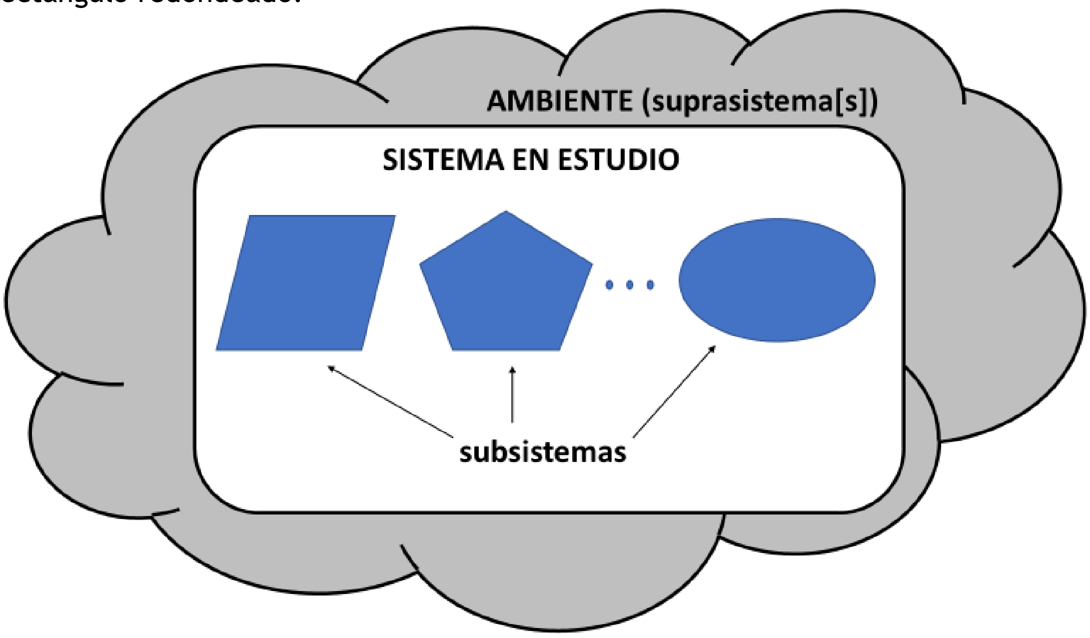
Un sistema cualquiera siempre existe en un ambiente. Según la relación que tenga con el ambiente, el sistema puede ser clasificado como:

- SISTEMA ABIERTO
- SISTEMA CERRADO
- SISTEMA AISLADO

El ambiente es, siempre, o un suprasistema del sistema estudiado o un conjunto de sistemas, alguno de los cuales puede ser un suprasistema del estudiado.

Debemos tener en cuenta que la definición de ambiente es circunstancial y que lo dicho en el párrafo anterior no es privativo del ambiente. Por ejemplo, el sistema que estamos estudiando, es probable que sea simultáneamente, el "ambiente" de uno o varios de los subsistemas que lo integran, lo que será válido en el caso que sea uno de ellos el que estudiemos o tengamos en cuenta.

En la figura que sigue, se ejemplifica lo que decimos en el párrafo anterior: El rectángulo redondeado tiene uno o varios ambientes (nube), pero si estudiamos una parte de él, un subsistema, por ejemplo el pentágono, el ambiente pasa a ser el rectángulo redondeado.

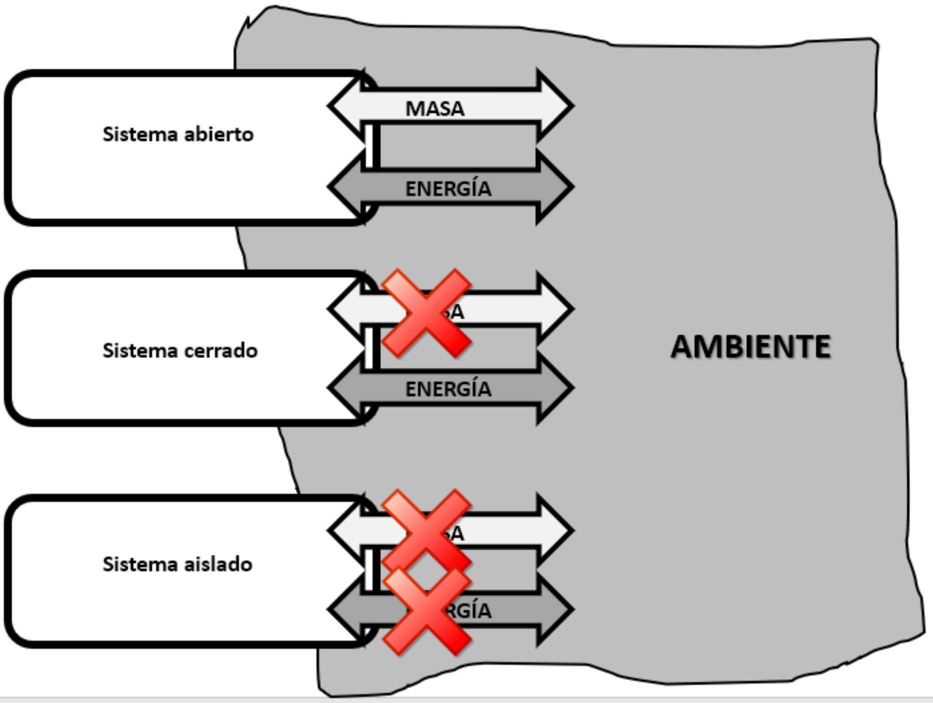


Tipos de sistemas

Es posible tipificar los sistemas en función de su relación con el ambiente. Como ya dijimos, esa tipificación es:

- **Sistema abierto**, es aquel que puede compartir materia y/o energía con su ambiente. Es la aproximación mayor a la realidad. Cualquier organismo vivo, en cuanto sistema, es un sistema abierto. Cualquier sistema mecánico o electrónico, también lo es.

- **Sistema cerrado** cuando no puede compartir materia, pero si puede compartir energía con su ambiente. Es una primera capa de abstracción. Permite protegerlo de amenazas externas, pero requiere energía para, por ejemplo, funcionar.
- **Sistema aislado** es el que se define como el sistema que no puede compartir ni energía ni materia con su ambiente. Es la mayor capa de abstracción. Si bien es difícil encontrar ejemplos reales, son muy utilizados para comprender fenómenos complejos aislando las variables de ruido o irrelevantes. El caso más conocido es el de los modelos conceptuales que se aplican en física, por ejemplo, movimiento rectilíneo uniforme, émbolos perfectamente aislados en termodinámica, el modelo de sistema planetario solar en astronomía, etc.



Caracterización de un sistema

La caracterización de cualquier sistema comienza, generalmente, a partir de la función. Como ya señalamos, el sistema es una conceptualización usada para explicar o comprender una funcionalidad. De hecho, si aceptamos que un sistema es una colección de objetos agrupados con el criterio de compartir un objetivo funcional, o — de otra manera — por tener un objetivo común, entonces deberíamos poder aceptar que esos mismos objetos o algunos de ellos pueden formar parte, simultáneamente o no, de otro sistema pensado en función de otros objetivos.

Por lo tanto, la primera característica de un sistema será el *propósito* (u objetivo).

- **Propósito u objetivo**. Los elementos (u objetos) y sus relaciones, definen una distribución que trata siempre de alcanzar un objetivo. Esa distribución a la que nos referimos será, en cada objeto, el conjunto de funciones y la forma de relacionarse con los demás objetos.

La segunda característica será el globalismo, o la manera en que las interrelaciones predefinidas entre los objetos constituyentes se manifiestan cuando uno de esos objetos cambia de estado.

- **Globalismo o totalidad**: un cambio en un objeto puede producir cambios en los otros. El efecto final es un ajuste de todo el sistema.

Se trata de una relación causa/efecto que deriva en comportamientos del sistema en general (ya que cambios individuales se propagan estructuralmente en todo el conjunto) y que son significativos en dos aspectos, que son:

- **Entropía**: que es la tendencia de los sistemas a desgastarse, a desintegrarse, relajar los estándares y aumentar la aleatoriedad. La entropía aumenta con el tiempo. Si aumenta la información, disminuye la entropía, porque es la base de la configuración y del orden. Obviamente el proceso entrópico, en tanto se trata de un proceso termodinámico, implica que hay un consumo irreversible de energía e información para mantener la integridad del sistema.

- **Homeostasis**: que es el equilibrio dinámico entre las partes del sistema. Los sistemas tienen una tendencia a adaptarse con el fin de alcanzar un equilibrio interno frente a los cambios externos del entorno. Forma parte indisoluble de la funcionalidad del sistema, para que esto ocurra**, una porción del sistema** (y de sus recursos) **debe estar dirigida a la homeostasis.**

Hacia una teoría de sistemas

Con lo expuesto hasta acá estamos en condiciones de enunciar premisas fundamentales para abordar un estudio de los sistemas.

La primera de esas premisas en realidad ya la hemos enunciado un poco más arriba cuando dijimos que los sistemas existen dentro de otro u otros sistemas más abarcativos que generalmente denominamos *suprasistema*.

La segunda premisa será que, en términos reales, consideramos que los sistemas son abiertos. De esta manera, cada sistema que se examine recibe algo y descarga algo de y en los otros sistemas, generalmente en los contiguos. Los sistemas abiertos se caracterizan por un proceso de cambio incesante con su entorno, que son los otros sistemas. Si el intercambio cesara, el sistema se desintegraría, ya que perdería sus fuentes de energía o de información.

Por último, debemos tener presente que existe una relación estrecha entre la funcionalidad del sistema y la estructura de este. El término *estructura* se refiere a la manera y forma física en que los objetos que lo integran se relacionan entres sí, teniendo en cuenta que esta relación solo se puede dar para contribuir a lograr el objetivo del sistema. En otras palabras: de qué manera “funciona” el sistema para poder lograr su objetivo.

Para los sistemas biológicos y mecánicos esta afirmación es intuitiva. Sobre todo, en una máquina, donde todo el sistema *máquina* se ensambla para su función. (Pensemos en el sistema de frenos de un suprasistema automóvil) En biología, si estudiáramos el sistema circulatorio de un suprasistema animal mamífero, por ejemplo, veríamos que cada componente de este (cada objeto) está claramente diferenciado para la función. Sin embargo, aun así, hay subsistemas y sistemas que son comunes y funcionales a otros sistemas: el sistema reproductivo de las células epiteliales del estómago es similar al de las células de la epidermis, por ejemplo.

Parámetros

Los parámetros más importantes presentes en un sistema son:

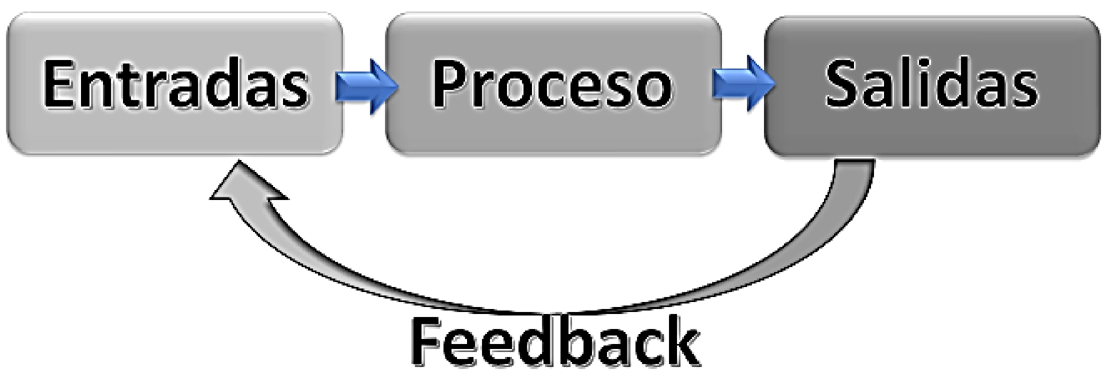
a. **Entradas**, *inputs*, insumos, impulso. Son las diversas denominaciones que reciben aquellos elementos que producen una perturbación de manera funcional en el sistema. Si esa perturbación es desconocida, imprevista o no es funcional, será tratada como **ruido**.

b. **Salidas**. *Outputs*, productos, resultados. Son las diversas denominaciones que reciben aquellos elementos u objetos que presentan un cambio de estado medible u observable que indica que la perturbación de la entrada fue tratada de determinada manera (procesada) por el sistema.

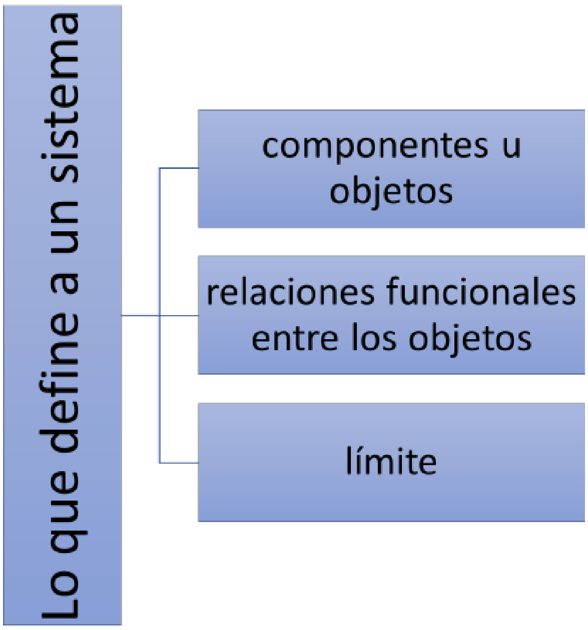
c. **Procesamiento**. *Throughput,* procesador, transformador. Son las diversas denominaciones que recibe la funcionalidad intrínseca del sistema. Es el comportamiento del sistema frente a las perturbaciones de entrada. Puede ocurrir que este procesamiento no se estudie sino que se lo considere o valore como una caja negra.

d. **Retroacción**. *Feedback*, retroalimentación, retroinformación. Son las denominaciones que recibe la funcionalidad que permite que información de salida se utilice como elemento de entrada con el objetivo de contribuir a la función o a la estabilidad de la función del sistema y ocasionalmente con su homeostasis.

e. **Ambiente**. *Environment,* medio, entorno. (erróneamente es común que se utilice el vocablo *medio ambiente*, que, como vemos, está compuesto por dos palabras sinónimas). Son las denominaciones que reciben todos los conjuntos externos a los límites fijados en el sistema en estudio. En ese ámbito se localizan objetos o sistemas que pueden ser recursos, que pueden ser amenazas o que pueden ser indiferentes.



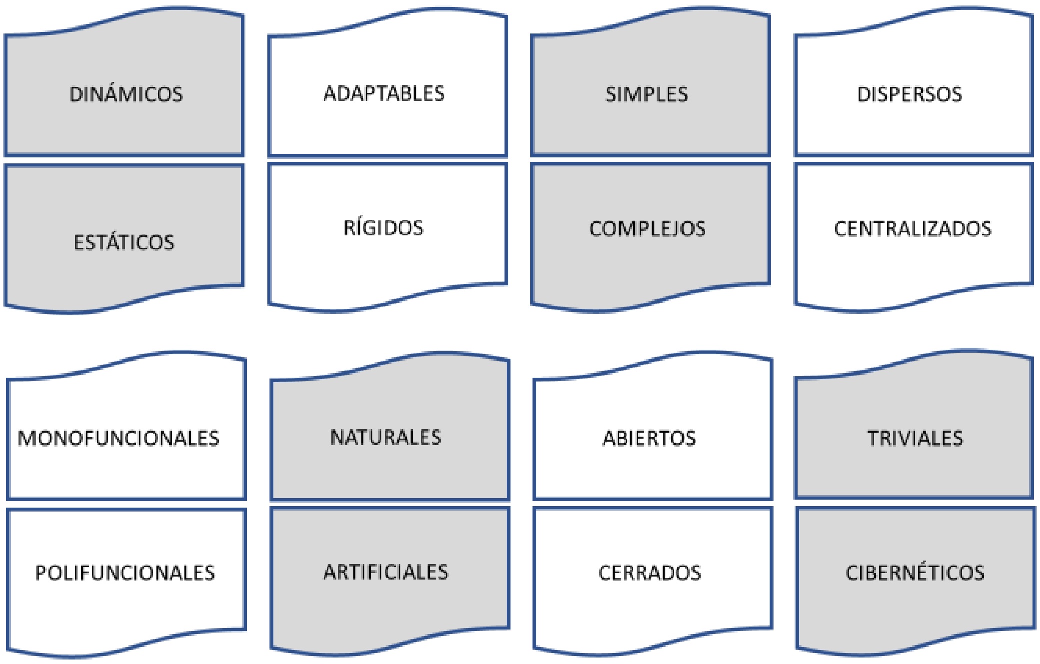
En definitiva, si consideramos que un sistema con las características señaladas arriba, es un conjunto de entradas a un procesador que nos ofrece otro conjunto de salidas, podemos conceptualizarlo con el esquema de la figura.



El estudio del sistema se basará en tres aspectos:

a) Las relaciones funcionales entre los componentes. No nos interesa cualquier relación, nos interesan aquellas que hacen a la función del sistema.
b) La transferencia de información. Intra y extra sistémica.
c) Las relaciones de orden

Así, los sistemas pueden ser:



SISTEMA DINÁMICO. Se caracteriza por mostrar estados que varían en función de un parámetro determinado. El más común es el tiempo. Un ejemplo es el modelo de universo en expansión.
SISTEMA ESTÁTICO es aquel cuya estructura o dinámica permanece invariable respecto a un parámetro. Por ejemplo, el modelo de sistema solar copernicano, que no contempla transiciones de estados (ni origen de los movimiento ni fuerzas perturbadoras de éstos)

SISTEMA ADAPTABLE. Generalmente se trata de sistemas que poseen un elevado número de objetos con relaciones complejas entre ellos, por ejemplo sociedades humanas o colonias de insectos, ecosistemas, redes neuronales, sistemas macroeconómicos.

SISTEMA RÍGIDO Es el antepuesto al anterior, en el cual la función, cantidad y relaciones entre los objetos se encuentra predeterminada. Por ejemplo, el sistema de frenos de un automóvil. El sistema solar.

SISTEMA COMPLEJO es el que está compuesto por varios objetos o subsistemas cuyos vínculos crean información adicional no visible por quien estudia el sistema. No debería confundirse con un sistema grande o complicado, ya que en esos casos no necesariamente los vínculos producen información adicional. En los sistemas complejos existen variables desconocidas. En los sistemas **SIMPLES**, en contrapartida, e independientemente de que sean grandes o no, se conocen todas las variables y las relaciones entre objetos no producen información más que la funcional.

SISTEMAS DISPERSO Y CENTRALIZADO. Se refiere a la ubicación del objeto que reacciona frente a una perturbación o cambio en la información. Generalmente se refiere a sistemas complejos. Si la reacción se produce en cualquier objeto que reciba información de otros objetos y con propósito de mantener coherente la función del todo, entonces significa que el control se encuentra altamente disperso. Lo contrario, cuando la reacción se produce en un único objeto o subsistema que está dedicado a ello, el sistema es centralizado.

SISTEMA ABIERTO. es aquel que tiene interacción constante con el medio, es decir intercambia energía, materiales o información y, por lo tanto, influye sobre el medio o es influido por él.

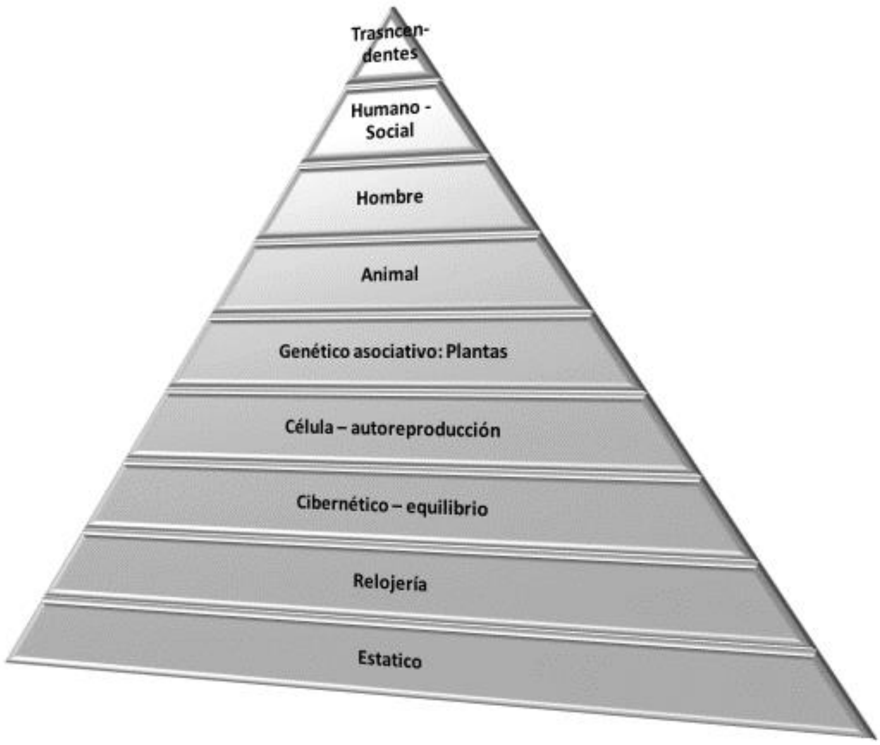
SISTEMA CERRADO, por contrario, no recibe ni envía energía, materiales o información al medio, aunque puede degradarse por el fenómeno de entropía. Suelen ser sistemas conceptuales termodinámicos, aunque existen otros casos.

SISTEMA TRIVIAL es aquel que es predecidle, que responde con una misma salida cuando recibe una determinada entrada. No modifica su comportamiento con la experiencia debido a que en estos sistemas no hay codificación y decodificación de la actividad que enfrentan; el sistema no interpreta su medio ni su actividad, simplemente actúa.

SISTEMA CIBERNÉTICO, por contrapartida, es aquel que sí es capaz de decodificar información tanto de entrada como de salida o del medio circundante. Se pueden describir dos tipos de sistemas cibernéticos: 1) los que tienen capacidad de realizar procesos de control o de comunicación (generalmente asociados a sistemas complejos, como los biológicos o maquinarias de alta complejidad como procesos industriales o vehículos como naves marítimas, aéreas o espaciales) y 2) aquellos que son un conjunto de procedimientos relacionados que llevan a un objetivo concreto (ejemplo: la heladera familiar mantiene la temperatura interna en un rango especificado de temperatura. El mismo sistema tiene un subsistema capaz de informarle la temperatura en tiempo real, compararla con consignas introducidas en el sistema y provocar la realización de acciones en consecuencia)

Jerarquías de los sistemas

La Teoría General de Sistemas, como veremos más adelante, fue desarrollada en un proceso que se prolongó alrededor de una década. Para poder comprenderla mejor, hemos elegido comenzar con la propuesta de jerarquías de los sistemas, hecha por un economista, Kennet ***Boulding***, en 1954, quien la describió en un artículo que tituló "*La teoría general de sistemas y la estructura científica*"[1]. Boulding, propone una clasificación de los sistemas en una estructura jerárquica por niveles de complejidad y de posibilidades de dominio del hombre:



Nivel 1. Estructura Estática

El **primer nivel** es el de la estructura estática. Boulding también lo llama "*nivel de marcos de referencia*".

Todo lo que hoy es el conocimiento básico del universo puede ser incluido en este nivel inicial: los electrones alrededor de un núcleo, la disposición de los

[1] Kenneth Boulding (18 de enero 1910 — 1993) Publicado originalmente en Manager Science, 2, 3 (Abr. 1956) PP- 197 – 208. Reimpreso en General Systems Yearbook of the Society of General System Research, vol. 1, 1956. Consultado en http://www.panarchy.org/boulding/systems.1956.html, (en inglés), enero 2020.

átomos en una fórmula molecular, o en un cristal, la constitución de los genes, las células, las plantas, un animal, los mapas de la tierra, el sistema solar, el universo astronómico.

La descripción precisa de estos marcos de referencia es el comienzo del conocimiento teórico organizado en casi cualquier campo, de lo contrario no es posible realizar teorías aceptables. Por ejemplo, el modelo de Copérnico fue realmente el descubrimiento de un nuevo “marco estático” para el sistema solar que permitió una descripción más simple de su funcionamiento (de su dinámica).

En este nivel se ve la necesidad de simplificación para obtener modelos. Con el mismo ejemplo del sistema solar: el modelo de Ptolomeo[2] no solo no describía con precisión el sistema, sino que era demasiado complejo: recordemos que incluía esferas y subesferas para poder explicar el movimiento de los planetas. Nicolás Copérnico[3] — en cambio — plantea un modelo sumamente simple, cuyo primer principio establece el concepto de “movimientos uniformes, circulares y eternos” que no necesitó de explicaciones previas sino, por el contrario, fue explicado matemática, astronómica y físicamente por Keppler, Galileo[4] y Newton, respectivamente[5].

---

[2] **Claudio Ptolomeo** (Egipto, 100 – 170), astrólogo y astrónomo, concibió, mediante el empirismo, un modelo de sistema solar basado en el geocentrismo. Este modelo fue aceptado por varios siglos, aun por la jerarquía teológica, a pesar de que tenía dos aspectos que no podía resolver: la retrogradación o cambios de dirección en el movimiento de los planetas y sus cambios en el brillo y las diferencias en la duración de los períodos de revolución sideral.

[3] **Nicolás Copérnico** (Polonia, 1473 – 1543). Enuncia el modelo que se llamó heliocéntrico en 1533, lo dio a conocer bajo presión de amigos y colegas, era muy renuente a hacerlo, ya que comprendía las implicancias teológicas y filosóficas de su teoría. (ver referencia 5, más abajo)

[4] **Galileo Galilei** (Pisa, 1564 – 1642). Científico italiano, uno de los precursores del método científico. Al descubrir y observar los satélites de Júpiter, da un espaldarazo al modelo Copernicano de sistema solar, ya que verifica que el sol es el centro del sistema. Posteriormente observa las fases de Venus, que ratifican esas conclusiones. También describe las manchas solares y sus ciclos de 11 años. Esto es expuesto ante la jerarquía eclesiástica y aceptado científicamente por los jesuitas. Poco después se inicia un conflicto entre Galileo y la Iglesia, que – según Bertrand Russell – se basó en la antinomia del método inductivo seguido por Galileo y el razonamiento deductivo, o apelación a la autoridad, que pretendía la iglesia. Esta dicotomía produce una profunda división de opiniones y el mundillo científico de Italia se divide por varios años en dos sectores: lo copernicanos y los ptolomeístas. Paradójicamente el Papa Urbano VII protege a Galileo, aunque en 1633, finalmente, la Inquisición, bajo amenaza de tortura, condena a un Galileo enfermo y agotado a prisión perpetua, que cumple en su casa, aunque – mediante contrabando – sigue publicando en el exterior algunas obras. Queda ciego y finalmente muere a los 77 años.

[5] El modelo de Copérnico, al principio bien aceptado por la Iglesia Católica, fue incluido en el Index durante el proceso a Galileo y generó una importante controversia teológica, ya que fue enunciado en el momento en que aparece la crisis eclesiástica provocada por Lutero.

Isaac Newton (1642 — 1727), casi un siglo después que Copernico se decidiera a publicar el modelo “heliocéntrico” al enunciar la Ley de Gravitación Universal (1685), da una base metodológica y comprobable desde la disciplina que hoy llamamos física.

Johanes Keppler (1571 — 1630), con el fin específico de describir matemáticamente el modelo de Copernico en 1609 y 1619 enunció sus tres leyes: 1) Los cuerpos celestes tienen movimientos elípticos alrededor del Sol, estando éste ubicado en uno de los dos focos; 2) Las áreas barridas por los cuerpos celestes son proporcionales al tiempo usado por esos cuerpos en recorrer los perímetros de las área y 3) el cuadrado de los períodos de las órbitas [de los planetas] es proporcional al cubo de la distancia de

Por ello, el modelo copernicano instala un sencillo marco de referencia y produce una revolución en el pensamiento científico, o, mejor dicho, en el método científico.

Nivel 2. Mecánico o de relojería

El **segundo nivel** de análisis sistemático es el del ***sistema dinámico simple con movimientos predeterminados y necesarios***. Podríamos denominarlo *nivel de relojería*. El sistema solar en sí mismo es, por supuesto, el gran reloj del universo desde el punto de vista del hombre, y las predicciones exactas de los astrónomos son un testimonio del excelente reloj que estudian.

Las máquinas simples, como la palanca y la polea, incluso las máquinas bastante complicadas, como las máquinas de combustión externa o interna, se incluyen principalmente en esta categoría.

Veamos dos casos especiales:

Los sistemas simples de equilibrio caen en la categoría dinámica, ya que cada sistema de equilibrio debe considerarse como un caso límite de un sistema dinámico, y su estabilidad no puede determinarse excepto a partir de las propiedades de su sistema dinámico padre.
Los sistemas dinámicos estocásticos que conducen a equilibrios, a pesar de su complejidad, también se incluyen en este grupo de sistemas; tal es la vista moderna del átomo e incluso de la molécula, cada posición o parte del sistema se da con un cierto grado de probabilidad, sin embargo, el conjunto exhibe una estructura determinada.

Aquí son importantes dos tipos de métodos analíticos, que podemos llamar ***estadísticas comparativas*** y ***dinámicas verdaderas***.

- En ***estadística comparativa***, comparamos dos posiciones de equilibrio del sistema bajo diferentes valores para los parámetros básicos. Estas posiciones de equilibrio generalmente se expresan como la solución de un conjunto de ecuaciones simultáneas.
  El método de la estadística comparativa es comparar las soluciones cuando se cambian los parámetros de las ecuaciones.
  La mayoría de los problemas mecánicos simples se resuelven de esta manera.

- En la ***dinámica real***, por otro lado, exhibimos el sistema como un conjunto de ecuaciones diferenciales, que luego se resuelven en forma de una función explícita de cada variable con el tiempo. Un sistema así puede alcanzar, o no, una posición de equilibrio estacionario. Hay ejemplos de sistemas explosivos y dinámicos, uno muy simple es el crecimiento de una suma a interés compuesto. De hecho, la mayoría de las reacciones físicas y químicas y la mayoría de los sistemas sociales exhiben una tendencia al equilibrio; de lo contrario, el mundo se habría desintegrado hace mucho tiempo.

---

ellos al Sol. Esta última ley fue enunciada posteriormente a las anteriores (diez años después), ya que sin ella no lograba relacionar las órbitas de los planetas entre sí.

La mayor parte de la estructura teórica de la física, la química, e incluso de la economía entra en esta categoría. Sobre este tema volveremos más adelante, al discutir el **método *cartesiano***.

Nivel 3. Cibernético o de equilibrio

El siguiente nivel es el del mecanismo de control o sistema cibernético, que podría denominarse ***nivel del termostato***.

Difiere del simple sistema de equilibrio estable principalmente en el hecho de que la transmisión e interpretación de la información es una **parte esencial** del sistema. Como resultado, la posición de equilibrio no está determinada simplemente por las ecuaciones del sistema, sino que el sistema reaccione para permitir el mantenimiento de cualquier equilibrio dado, dentro de límites.

De esta manera, el termostato mantendrá cualquier temperatura a la que pueda ajustarse, por lo tanto, la temperatura de equilibrio del sistema no está determinada únicamente por sus ecuaciones.

Lo fundamental en este nivel es que la variable esencial del sistema dinámico es la diferencia entre un valor "observado" o "registrado" de la variable mantenida y su valor "ideal" o de referencia. Si esa diferencia no es nula, el sistema se mueve para disminuirla; así, el quemador del horno recibe mayor caudal de gas cuando la temperatura registrada es *"más fría comparada con la ideal o consigna*" y recibe menos o se apaga cuando la temperatura registrada es "*más caliente comparada con la ideal o consigna*".

El modelo de homeostasis, que es importante en fisiología, es un ejemplo de mecanismo cibernético, y tales mecanismos existen en todo el mundo empírico del biólogo y el científico social. El término cibernético está utilizado en su acepción más restringida en teoría de control[6].

Nivel 4. Estructura de auto reproducción o de célula.

El cuarto nivel es el del "***sistema abierto***" o estructura **autosuficiente**. Este es el nivel en el que la vida comienza a diferenciarse de la no vida: podría llamarse el nivel de la célula.

Obviamente, existe algo parecido a un sistema abierto incluso en los sistemas de equilibrio fisicoquímico; las estructuras atómicas se mantienen en medio de un intercambio y procesos de electrones, las estructuras moleculares se mantienen en medio de procesos de átomos.

Las llamas y los ríos también son sistemas esencialmente abiertos de un tipo muy simple. Sin embargo, a medida que pasamos la escala de complejidad de la organización hacia los sistemas vivos, la propiedad de auto mantenimiento

---

[6] La palabra cibernética proviene del griego: *timonel,* el que dirige la embarcación, (*kubernites*). El término se usa ampliamente en teoría de sistemas en cuanto a la estructura funcional y se usa restrictivamente en teoría del control y de la comunicación.

de la estructura en medio de un proceso de producción de material adquiere una importancia dominante.

Podemos admitir que un átomo o una molécula pueden existir sin procesos, pero jamás podremos concebir la existencia del organismo vivo más simple sin ingestión, excreción e intercambio metabólico.

Muy relacionada con la propiedad del auto mantenimiento está la propiedad de la auto reproducción. De hecho, puede ser que la auto reproducción sea un sistema más primitivo o de "nivel inferior" que el sistema abierto, y que el gen y el virus, por ejemplo, puedan reproducirse sin ser sistemas abiertos. Quizás no sea una pregunta importante en qué punto de la escala de complejidad creciente comienza la "vida". Lo que está claro, sin embargo, es que cuando llegamos a los sistemas que se reproducen y se mantienen en medio de procesos de materia y energía, tenemos algo a lo que sería difícil negar el título de "vida".

Nivel 5. Genético asociativo o nivel de las plantas.

El quinto nivel podría llamarse el nivel genético — social. Está tipificado por las plantas y es el mundo empírico del botánico.

Las características sobresalientes de estos sistemas son, en primer lugar, una división del trabajo entre las células para formar una sociedad celular con partes diferenciadas y mutuamente dependientes (raíces, hojas, semillas, etc.) y, en segundo lugar, una clara diferenciación entre el genotipo y el fenotipo[7], asociado con el fenómeno del crecimiento hacia un equilibrio final o "planificado". En este nivel no hay órganos sensoriales altamente especializados y los receptores de información son difusos e incapaces de obtener mucha información: es dudoso que un árbol pueda distinguir mucho más que la luz de la oscuridad, los días largos de los días cortos, el frío del calor.

Nivel 6. Mundo animal.

En este nivel se incluyen los sistemas complejos con objetivos más amplios que la propia estabilidad. Es un nivel caracterizado por comportamiento dirigido a una finalidad u objetivo. Pueden tener conciencia y receptores especializados (o desarrollados) de información relevante (ojos, oídos, olfato) que conduce a una gran capacidad de procesar información.

El nivel animal está caracterizado por una mayor movilidad, comportamiento teleológico[8] y autoconciencia. El desarrollo de receptores de información especializados conduce a un enorme aumento en la capacidad de recibir

---

[7] El genotipo es la información genética (genoma) de un organismo determinado. El fenotipo, junto al genotipo y los factores circunstanciales dan para cada individuo, las características genéticas. Dicho más sencillamente: el genotipo sería el conjunto de genes de un organismo y el fenotipo el conjunto de rasgos de un organismo.

[8]. Teleología es la doctrina de las causas finales. O, dicho en términos simples, rama de la metafísica que se refiere al estudio de los fines o propósitos de algún objeto o algún ser. Atribución de una finalidad u objetivo a procesos concretos

información. Se encuentra un subsistema nervioso complejo, que conduce en última instancia al cerebro, como organizador de la entrada de información en una estructura de conocimiento o "imagen".

Cuanto más ascendemos en la escala de la vida animal, el comportamiento deja de ser la respuesta a un estímulo específico, y pasa a ser una estructura de conocimiento o visión del entorno en su conjunto. Esta imagen está determinada por la información recibida en el organismo. La relación entre la recepción de información y la construcción de una imagen es compleja. No es solamente acumulación de información recibida, aunque esto sucede con frecuencia, sino la posibilidad de transformar esa información en algo diferente.

Las dificultades en la predicción del comportamiento de estos sistemas surgen en gran medida debido a la manera en que interviene la imagen entre el estímulo y la respuesta.

Nivel 7. Capacidad de procesar información a gran velocidad.

Se refiere al ser humano individual considerado como un sistema. Además de todas, o casi todas, las características de los sistemas animales, el hombre posee autoconciencia, que es diferente de la mera conciencia. Su imagen, además de ser mucho más compleja que la de los animales superiores, tiene una cualidad autorreflexiva: no solo lo sabe, sino que sabe que lo sabe.

Esta propiedad probablemente esté relacionada con el fenómeno del lenguaje y el simbolismo. Lo que más claramente distingue al hombre de sus hermanos animales es la capacidad de hablar, la capacidad de producir, absorber e interpretar símbolos, en lugar de simples signos como el grito de advertencia de un animal. El hombre se distingue de los animales también por una imagen mucho más elaborada del tiempo y la relación con acontecimientos.

El ser humano tiene la capacidad de prever consecuencias de acciones que aún no ocurrieron (a esto lo llamamos “imaginar”, que literalmente significa crear una imagen de algo que aún no es). El hombre es probablemente la única organización que sabe que muere, que contempla en su comportamiento una vida entera y más que una vida. El hombre existe no solo en el tiempo y el espacio, sino también en la historia, y su comportamiento se ve profundamente afectado por su visión del proceso de tiempo en el que se encuentra.

Nivel 8. Sistemas sociales humanos.

Las imágenes simbólicas y el comportamiento basado en ellas son de importancia vital para el hombre-individuo. Por eso no es fácil separar claramente el nivel del organismo humano individual del nivel de las organizaciones sociales.

El hombre aislado de sus congéneres es prácticamente desconocido. Tan esencial es la imagen simbólica en el comportamiento humano que se podría afirmar que el hombre verdaderamente aislado no sería "humano" en el sentido generalmente aceptado, aunque sería *potencialmente* humano.

Sería conveniente, en determinadas circunstancias, diferenciar al humano individual como subsistema de los sistemas sociales que lo rodean, y de esta manera afirmar que las organizaciones sociales constituyen otro nivel de organización.

El componente u objeto básico de estos sistemas quizás no sea la persona humana (el individuo como tal), sino el "*rol*" que este cumple en el sistema, que es esa parte de la persona que se ocupa de la organización o situación en cuestión. Así es tentador definir organizaciones sociales, o casi cualquier sistema social, como un *conjunto de roles* unidos con canales de comunicación.

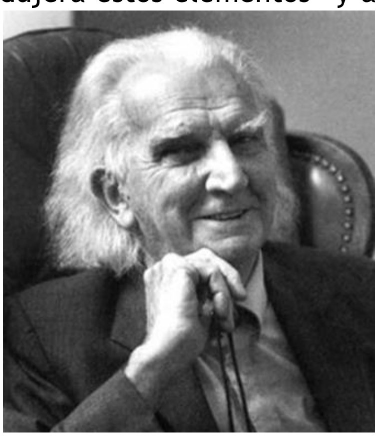
Sin embargo, las interrelaciones del rol y la persona nunca se pueden ignorar totalmente ya que, por ejemplo, si a una persona determinada que no tiene vocación para determinada tarea o papel, se le asignara igual esa tarea, seguramente podrá hacerla y adaptarse completamente, bastante o nada, y la hará de manera que sea poco diferente o muy diferente, según las circunstancias. La percepción de un rol, o sea, la manera en que los demás ven o imaginan ese rol, se ve afectada por las personalidades de aquellos que lo ocuparon antes.

En este nivel, son importantes tanto el contenido como el significado de los mensajes, la naturaleza y las dimensiones de los sistemas de valores, la transcripción de imágenes en un registro histórico, las simbolizaciones del arte, la música y la poesía, y la compleja gama de emociones humanas.

El universo empírico (que es el suprasistema) aquí es la vida humana y la sociedad en toda su complejidad y riqueza.

Nivel 9. Sistemas trascendentes

No referimos a aquellos sistemas complejos, que la mente humana es capaz de describir someramente o, si no es capaz de describirlos, al menos podría tener la capacidad de suponer que existen, pero que — paradójicamente — no puede comprenderlos o relatarlos acabadamente. Estos sistemas derivan de las capacidades de sistematizar el conocimiento del ser humano y, en general, de deducciones lógicas del tipo "si el fenómeno observable A existe, solo se puede explicar si existiera un sistema tal que produjera estos elementos" y a continuación hay una serie de parámetros y especificaciones conjeturales que describen dicho sistema.



La razón por la cual comenzamos con esta clasificación de Boulding es, según sus propias palabras[9], para tener "*… una idea de las brechas actuales en el conocimiento tanto teórico como empírico. Los modelos teóricos adecuados se extienden hasta aproximadamente el cuarto nivel, y no mucho más allá. El conocimiento empírico es deficiente en prácticamente todos los niveles. Por lo tanto, a nivel de la estructura estática,*

[9] Ver llamada 1, más arriba.

hay disponibles modelos descriptivos bastante adecuados para geografía, química, geología, anatomía y ciencias sociales descriptivas. Sin embargo, incluso en este nivel más simple, el problema de la descripción adecuada de estructuras complejas aún está lejos de resolverse."

[…] "*La catalogación de eventos, ideas, teorías, estadísticas y datos empíricos apenas ha comenzado. Sin embargo, la multiplicación misma de registros a medida que pasa el tiempo nos obligará a sistemas de catalogación y referencia mucho más adecuados que los que tenemos ahora. Este es quizás el principal problema teórico no resuelto a nivel de la estructura estática.*"

"*En el campo empírico todavía hay grandes áreas donde las estructuras estáticas son muy imperfectamente conocidas, aunque el conocimiento avanza rápidamente, gracias a nuevos dispositivos de sondeo como el microscopio electrónico. Sin embargo, la anatomía de esa parte del mundo empírico que se encuentra entre la molécula grande y la célula todavía es oscura en muchos puntos. […] es precisamente esta área, que incluye, por ejemplo, el gen y el virus, la que guarda el secreto de la vida, y hasta que se aclare su anatomía, la naturaleza de los sistemas funcionales involucrados será inevitablemente oscura.*"

"*El nivel del «mecanismo de relojería» es el nivel de las ciencias naturales «clásicas», especialmente física y astronomía, y es probablemente el nivel más completamente desarrollado en el estado actual del conocimiento, especialmente si ampliamos el concepto para incluir la teoría de campo y modelos estocásticos de física moderna. Incluso aquí, sin embargo, hay brechas importantes, especialmente en los niveles empíricos más altos. Todavía queda mucho por saber sobre la mecánica pura de las células y los sistemas nerviosos, de los cerebros y de las sociedades. Más allá del segundo nivel, los modelos teóricos adecuados se vuelven más escasos." [aunque se] "han visto grandes desarrollos en los niveles tercero y cuarto." [*Como sistemas de control y sistemas autosuficientes, o abiertos]

Queda, dice Boulding, mucho por aprender en sistemas tales como la cibernética de los genes y los sistemas genéticos, por ejemplo, y sobre los mecanismos de control involucrados en el mundo mental y social. Se está lejos en cualquier sistema automantenido que implique avanzar más allá del cuarto nivel.

"*Ante los sistemas vivos estamos casi indefensos; ocasionalmente podemos cooperar con sistemas que no entendemos: ni siquiera podemos comenzar a reproducirlos. El estado ambiguo de la medicina, que se cierne inquietantemente entre la magia y la ciencia, es un testimonio del estado del conocimiento sistemático en esta área. A medida que avanzamos en la escala, la ausencia de los sistemas teóricos apropiados se hace cada vez más notable. Difícilmente podemos concebirnos construyendo un sistema que sea, en algún sentido reconocible, "consciente", mucho menos consciente de sí mismo.*" […] "*El tipo de conocimiento y habilidad que tenemos en el nivel simbólico es muy diferente del que tenemos en los niveles más bajos, es como, digamos, el «conocimiento» del gen en comparación con el conocimiento del biólogo. Sin embargo, es un tipo de conocimiento real y es la fuente de los logros creativos del hombre como artista, escritor, arquitecto y compositor.*"

Teoría General de Sistemas[10]

La Teoría General de Sistemas (TGS) fue enunciada por Ludwin von Bertalanffy[11] durante un período que va desde 1945 (fin de la Segunda Guerra Mundial) a 1968.

El objetivo perseguido por Bertalanffy era el intento de reunir el estudio del metabolismo, del crecimiento, de la morfogénesis y de la fisiología en un enfoque que un fuera estático, sino que incluyera los cambios y movimientos necesarios y utilizando sistemas abiertos.

De esta manera, si describimos un organismo bajo la mirada de la TGS, al ser un sistema abierto veremos que tiene como meta lograr un estado de armonía con el entorno.

Bertalanffy estableció dos principios:

1) Los organismos tienden a mantenerse en condiciones de no — equilibrio
2) La estructura de los sistemas tiene una organización jerárquica

Es un intento de unificación de varias teorías y de encontrar un punto de partida común para las ciencias en todas sus ramas, tanto naturales como sociales.

De esta manera, la TGS tiene en cuenta — o incluye a posteriori — las siguientes teorías:

1) **La de conjuntos**[12]
2) **La de Redes de Anatol Rapoport**[13]. En la década de los 1950 Rapoport comenzó a aplicar al campo social los estudios de redes, particularmente, en

---

[10] En el Anexo al presente, se incluye el texto original de Teoría General de Sistemas de Ludwig von Bertalanffy

[11] Nacido en 1901 en Austria y fallecido en 1972 en Nueva York. Filósofo, Biólogo. La vida académica de Bertalanffy se centró sobre la biología y el abordaje de esta con una visión orgánica. Su teoría se publicó en 1969 aunque fue apareciendo y desarrollándose durante los 20 años de post Segunda Guerra Mundial. Inclusive, la clasificación ya vista de Boulding sobre los sistemas se refiere críticamente a la TGS y fue publicada en 1956.

[12] La teoría de conjuntos es una rama de las matemáticas que estudia las propiedades de colecciones abstractas de objetos como una herramienta básica en la formulación de algunas teorías matemáticas y es lo suficientemente rica como para construir el resto de los objetos y estructuras matemáticas de interés como números, funciones, figuras geométricas, etc. Se acepta que el conjunto de axiomas de la teoría es suficiente para desarrollar toda la matemática.
El origen de la teoría de conjuntos se atribuye a Georg Cantor, en la segunda mitad del siglo XIX, en base a ideas de Bernhard Bolzano. Posteriormente aparecieron trabajos de Bertrand Russell, Ernst Zermelo, Abraham Fraenkel y otros a principios del siglo XX.

[13] Rapoport (Matemático Ucraniano, 1911 — 2007) contribuyó a la teoría general de sistemas, a la biología matemática y al modelo matemático de la interacción social y los modelos estocásticos de contagio. Combinó su experiencia matemática con conocimientos psicológicos sobre el estudio de la teoría de juegos, las redes sociales y la semántica. Rapoport amplió estos conceptos en estudios de conflictos psicológicos, tratando el desarme nuclear y la política internacional. Su autobiografía, ***Certainties and Doubts: A Philosophy of Life***, se publicó en 2001.

1957, definió los lazos fuertes y débiles – si A conoce a B y a C, lazo fuerte, es probable que B conozca a C, pero C podría no conocer a B, lazo débil[14]

3) **La de los autómatas de Turing**[15]. La máquina de Turing (1936) es un modelo general de autómata finito. Expresado en su manera más simple, un autómata de Turing es una máquina abstracta capaz de imprimir (o borrar) las marcas “1” y “0” en una cinta de longitud infinita. Es demostrable que cualquier proceso, de la complejidad que sea, puede ser simulado por una máquina, si este proceso es expresable mediante un número finito de operaciones lógicas.
4) **La teoría de juegos de von Neumann**[16] que establece un modelo de las especulaciones estratégicas de dos contendientes que poseen la misma información y compiten por un objetivo excluyente (por ejemplo, lo que gana uno de ellos es lo que necesariamente debe perder el otro).

Bases filosóficas y características de la TGS

Las bases de la TGS se pueden vincular con algunos aspectos filosóficos, los cuales, obviamente, escapan a los alcances de esta publicación, que podemos resumir en tres enfoques: ontológico, empírico y valorización filosófica.

1. La **ontología** de la TGS propone la definición de cada sistema y de su ubicación.

   El término **ontología,** en filosofía, hace referencia a la rama que se ocupa del estudio del ser[17]. Pero, en informática, se toma esta palabra por analogía para describir la formulación de un riguroso esquema conceptual dentro de dominios dados con la finalidad de facilitar la comunicación y el intercambio de información entre diferentes sistemas y entidades.

   Un uso común tecnológico actual del concepto de ontología, en este sentido, es la inteligencia artificial y la representación del conocimiento.

2. En la TGS se propende a diferenciar ésta del **empirismo** clásico, entendido como un modelizado “causa-efecto” (modelo físico) enunciando que, en

---

[14] En el lenguaje cotidiano se ha utilizado la idea de "red social" durante más de un siglo para conjuntos complejos de relaciones entre miembros de los sistemas sociales, desde el ámbito interpersonal hasta el internacional. En 1954, el antropólogo J. A. Barnes comenzó a utilizar el término para mostrar patrones de lazos, abarcando los conceptos utilizados por los científicos sociales: grupos delimitados (p.e., tribus, familias) y categorías sociales (p.e., género, etnia). Anatol Rapoport expandió el uso del análisis de redes sociales sistemático y propuso un algoritmo (toma y daca) escrito en Basic para desarrollar la mejor estrategia conocida del “Dilema del Prisionero”, a fines de los 1980 (Ver capítulo 12 del Libro 2 de la Serie Optimiza12)

[15] Alan Turing, es un matemático, criptógrafo, filósofo e informático inglés (1912 – 1954) considerado uno de los fundadores de la informática. Se desarrollará más sobre él en siguientes capítulos.
[16] Ver el Capítulo 11 en el Libro 2 de la Serie Optimiza12.
[17] En Wikipedia (extracto): “La **ontología** (ciencia, estudio, teoría del ser) es una parte de la metafísica que estudia lo que existe. Muchas preguntas tradicionales de la filosofía pueden ser entendidas como preguntas de ontología: ¿Existe Dios? ¿Existen entidades mentales, como ideas y pensamientos? ¿Existen entidades abstractas, como los números?
Además, la ontología estudia la manera en que se relacionan las entidades que existen. Por ejemplo, la relación entre un general o universal (verde) y un particular que "lo tiene" (esta hoja).

realidad, esas relaciones no siempre son lineales. Se profundiza este aspecto un poco más adelante (Ver "Diferencias entre mecanicismo y TGS").

3. Se establece una **filosofía de valores**, que establece la relación entre los seres humanos y los sistemas que lo rodean.

Las características de los sistemas, desde la óptica de la TGS, son varias y hay diferentes propuestas. Por ejemplo, y siendo seguramente redundantes, podemos señalar:

1.- **Interrelación** e interdependencia. Los objetos (entidades) no deben ser incluidos en un sistema si no poseen vinculaciones específicas.

2.- **Totalidad**. No debe estudiarse un sistema en forma parcial o por partes sin tener en cuenta la totalidad. Cuando abordamos un sistema debemos hacerlo bajo esta óptica y aunque podamos trabajar con cada parte por separado, antes debemos al menos comprender el todo.

3.- Los **objetivos** del sistema deben estar claramente definidos

4.- Todo sistema requiere **entradas** para su existencia. Las entradas pueden ser materia, energía o información.

5.- Las **salidas** o productos de un sistema comunican (o no) a éste con el medio en forma análoga a lo que se produce con las entradas. La diferencia entre unas y tras son las transformaciones funcionales del sistema.

6.- Las **transformaciones** funcionales son el proceso que transforma las entradas en salidas, o simplemente, la función que cumple el sistema.

Entrada ⊗ Proceso = Salida.

7.- Los sistemas tienen un componente importante: la **entropía** o tendencia a la desorganización o degradación en función del tiempo. Esa degradación puede llevar a su desaparición.

Para describir la entropía de un sistema hace falta conocer la manera de relacionarse con los demás sistemas del entorno

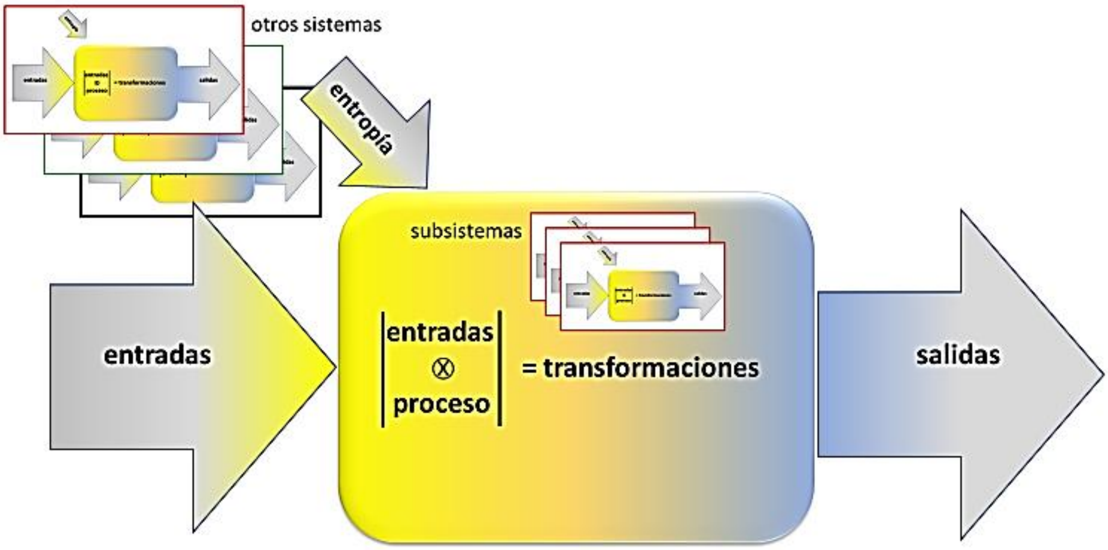
8.- Según el grado de complejidad, cada sistema tiene una cantidad de **subsistemas**. Las relaciones y orden jerárquico entre ellos permiten regular y obtener funciones.

La **regulación** es, entonces, una consecuencia de lo anterior: si hay jerarquías debe haber una manera de manejar las interacciones con el fin de que las metas y objetivos se cumplan satisfactoriamente.

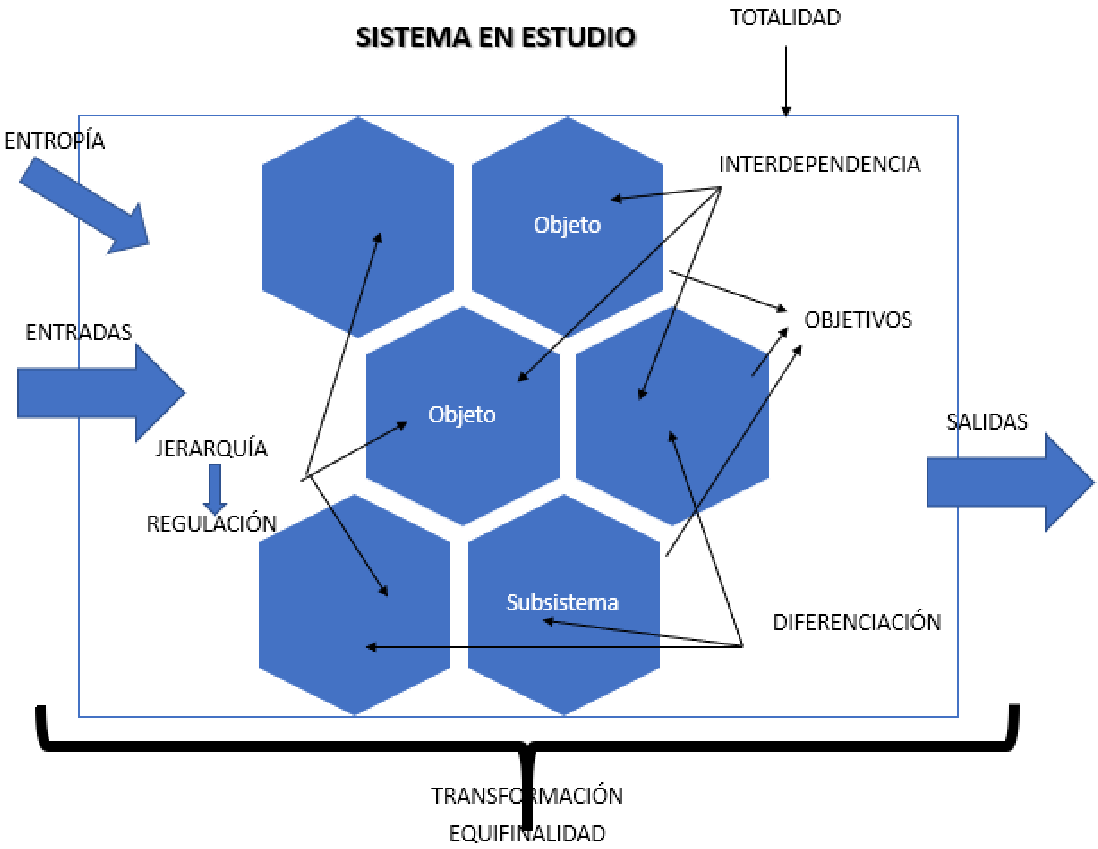
9.- En todos los sistemas debe existir una clara **diferenciación** entre los distintos subsistemas. Esas diferencias serán funcionales o estructurales, y existirán para lograr los objetivos más eficientemente, para una mejor adaptación al medio y para poder aprovechar los insumos (entradas) que el medio proporciona.

10.- La capacidad de los sistemas de lograr sus objetivos a partir de diferentes estados iniciales y por procesos diferentes se denomina ***equifinalidad***. **Es una**

diferenciación muy importante respecto al *empirismo*, que plantea que toda causa tiene su efecto y el camino o proceso es siempre el mismo[18].



[18] Para comprender este pequeño párrafo tomemos como ejemplo un dilema de la física moderna: ¿Por qué el Universo es como lo vemos? Para llegar a este aspecto debieron haberse dado condiciones iniciales muy particulares. La pregunta crucial es: hay un solo conjunto de condiciones iniciales que producen este resultado o hay un número muy alto de otras condiciones que *podrían* haber conducido al mismo resultado. Si pensamos un poco esta sencilla pregunta implica una profundidad importante, ya que, según la respuesta se involucra al azar o a Dios en la creación del Universo.



Diferencias entre mecanicismo y TGS

Hemos señalado antes, en dos oportunidades, que existen diferencias entre el empirismo clásico y la TGS, por lo que creemos que es útil dedicar un apartado para discutir esas diferencias.

Partimos de la base de que éste no es un trabajo sobre el método científico lo que nos permite tomar alguna licencia y simplificar un poco diciendo que, a lo largo de la historia de la humanidad existieron varias maneras de describir las organizaciones y los fenómenos naturales. Nos encontramos con que cada una de esas maneras, según las circunstancias históricas, pudo haber sido denominada "científicamente correcta", "método científico" o, simplemente, reconocida como una forma válida de describir lo que estamos estudiando, sea lo que fuere.

En uso de esas licencias que nos tomamos y simplificando las cosas, podemos afirmar que, en la era moderna, de todas esas maneras perduran, válidamente, dos visiones para describir los fenómenos naturales: La visión **mecanicista** y la visión **sistémica**.

Visión Mecanicista

El enfoque llamado actualmente "mecanicista", fue una propuesta de Descartes[19], por lo que generó un enfoque metodológico aplicado al método científico, que es

[19] René Descartes (1596 – 1650), fue un filósofo francés, que escribía su apellido en una forma latinizada (Cartesius). Entre sus muchos aportes a la física, la matemática y la filosofía, propició un racionalismo

conocido como “***método cartesiano***”. Se fundamenta en especulaciones filosóficas enunciadas por Galileo Galilei, Huygens[20], Boyle[21], y otros.

La **ontología cartesiana** es un tema extenso y complejo, pero para los fines que acá estamos persiguiendo, lo podríamos simplificar con dos hipótesis básicas:

1) El mundo y todo lo que él contiene es una máquina y se comporta como una máquina
2) Todo lo real es físico

Si bien estos dos puntos pueden parecer demasiado simplificadores, nos explican claramente la base del conocimiento (gnoseología) cartesiano mecanicista porque permitieron (y permiten) que se terminen con las explicaciones sobrenaturales[22] o incomprobables (punto 2) y establece la repetitividad de las observaciones y experimentos (a cada causa corresponden los mismos efectos: la máquina).

Esto significa que los fenómenos naturales se reducen a sus partes físicas y a las interacciones mecánicas entre ellas.

En definitiva, la visión cartesiana mecanicista se caracteriza por principios tales como:

- **Los sistemas son cerrados.**
    - El campo de experimentación se remite al objeto en estudio sin tener en cuenta su entorno. Las relaciones sujeto — objeto siempre se producen a distancia y por fuerzas determinadas. Ejemplo: el modelo que se utiliza para describir el movimiento rectilíneo uniforme emplea un móvil sin rozamiento, es decir, sin interacción con el medio, y sin un elemento disparador.

- **Los sistemas son estáticos.**

científico riguroso, publicado en 1637 bajo el nombre “Discurso del método”, que buscaba terminar con métodos – a menudo basados en consideraciones teológicas – complicados para poder explicar las cosas. Se basaba en modelos muy sencillos y pocas reglas básicas. Formuló el célebre principio *cogito ergo sum* ("pienso, luego existo"), elemento esencial del racionalismo occidental.

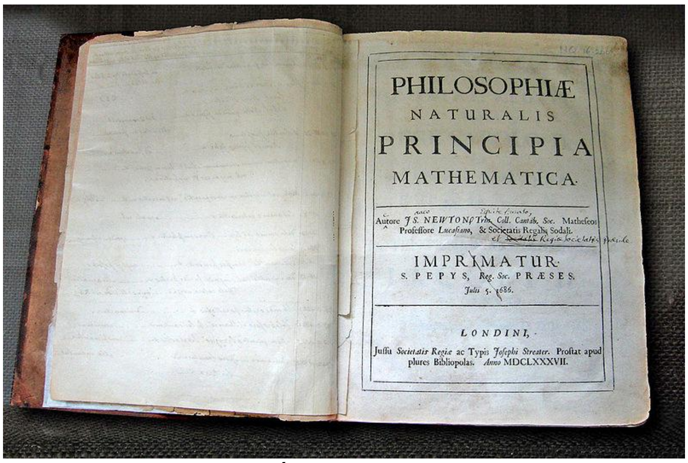
[20] Christiaan Huygens (1629 – 1695) fue un astrónomo y físico holandés. Si bien conoció a Descartes en sus primeros años como científico, existió una interacción en el interés de aplicar un método racional científico.

[21] Robert Boyle (1627 – 1691) fue un químico e inventor holandés. También era teólogo, por lo cual existió cierta influencia en el desarrollo del método cartesiano, ya que en sus trabajos uso métodos racionales, sin necesariamente estar convencido de la obra de Descartes.

[22] Quizá este párrafo no se entienda si no se recuerda que durante toda la edad media (y, por supuesto antes también) era válido explicar sucesos naturales o sociales con apelaciones a la magia, la voluntad de Dios o del demonio, a la influencia de fenómenos astronómicos como cometas o meteoritos, apelaciones a la autoridad del que afirmaba algo, etc. Descartes nació cuando finalizaba la edad media y esas ideas todavía tenían vigencia o se empleaban. Incluso propiciadas por la Iglesia basada en principios teológicos con varios siglos de vigencia.

- el método es solo cuantitativo y no cualitativo, por lo tanto encierra una visión estática del mundo y sus objetos. Ejemplos: el modelo de sistema solar describe con precisión el movimiento de los planetas alrededor del sol, pero no explica cómo se puso en movimiento y que acontecimiento podría alterarlo. El modelo de movimiento rectilíneo uniforme del punto anterior no incluye condiciones de inicio y de final.
- **Los sistemas se abordan utilizando métodos cuantitativos.**
  - La separación sujeto — objeto permite crear relaciones matemáticas que describen esa distancia y esa interacción. Ejemplo: el modelo de sistema solar permite enunciar relaciones matemáticas descriptoras de la relación entre los objetos como la ley de la gravitación universal de Newton o de su movimiento como las leyes de Kepler[23].
- **Las explicaciones funcionales son matemáticas.**
  - Ejemplos: Las leyes de Kepler, para el sistema solar. El modelo matemático e = v / t para el movimiento uniforme.

El método cartesiano fue puesto en práctica por Isaac Newton[24], quien en todos sus trabajos sobre fuerzas gravitatorias, naturaleza de la luz, cálculo, etc. empleó rigurosamente el método cartesiano: proponiendo hipótesis y demostrando teorías usando razonamiento y experimentación.



Con su ley de gravitación, Newton fue el primer científico que enunció una teoría del universo basado en la matemática y no solo en la observación o experimentación

John Locke[25], basándose en el método cartesiano y las ideas newtonianas, efectuó un intento de aplicarlas al desarrollo de estudios sobre organizaciones humanas. Enunció que "… el ser humano es autónomo y alejado del entorno…" con lo que concluía que cada grupo social era el resultado del comportamiento de los individuos como tales sin tener en cuenta el todo. Concluía que cada ser humano es un universo y que la finalidad es la eficiencia. Si bien esta cita puede parecer pueril a los efectos de la discusión, se intenta mostrar la manera en que un pensador liberal (sostenía Locke, por ejemplo, que la religión es un asunto privado y que Dios era el gran relojero) hace converger esa idea mecanicista hasta el punto de que – sostiene – todos los comportamientos sociales son posiblemente abordados y entendidos mediante simples relaciones causa y efecto.

Visión sistémica

[23] Johannes Kepler (1571-1630), astrónomo y matemático alemán; enunció sus leyes sobre el movimiento de los planetas en su órbita alrededor del Sol. Colaborador de Tycho Brahe, a quien sustituyó como matemático imperial de Rodolfo II, emperador Astro Húngaro.

[24] Isaac Newton (1642 – 1727), fue un físico, filósofo y matemático inglés.

[25] John Locke (1631 – 1704), fue un pensador inglés. Se lo considera padre del liberalismo moderno.

La propuesta de Berthalanfy, puede ser vista – en este contexto – como algo opuesto de la visión mecanicista, por eso vamos a remarcar y a tener en cuenta que **no la excluye**. Y no la excluye porque siempre tendremos en cuenta que es deseable encontrar modelos más simples y eficaces que los que propuso, por ejemplo, Newton para física y para óptica.

Es en la modelización de sistemas reales complejos (jerarquías superiores) donde notaremos esa oposición: acá encontraremos la descripción de sistemas abiertos de carácter cíclico, analizados con un pensamiento integrador e incorporando elementos dinámicos y cualitativos.

- **Sistemas abiertos**, en cuanto que en esta visión, no solo se tiene en cuenta el entorno, sino que éste resulta indispensable en la existencia del mismo sistema *sub examine* ya sea porque ese ambiente es necesario para que exista o para que funcione nuestro sistema.

- El **carácter cíclico** de los sistemas lo determina la interconexión entre las entidades que lo componen, de esta manera los cambios en cada parte pueden condicionar el todo mediante propagación y pueden cambiar el sistema.

- Solamente con un **pensamiento integrador** podremos abarcar el sistema, ya que este no se aísla y nos obliga a analizarlo en conjunto con las relaciones con sus subsistemas y su entorno o suprasistemas.

- El **carácter dinámico** proviene de aceptar que el propio desarrollo en el tiempo de los sucesivos estados del sistema (¿funcionamiento?), va a producir modificaciones estructurales o funcionales en los componentes o en el entorno y hasta en el mismo comportamiento del sistema.

Los conceptos básicos de la visión sistémica son:

1) <u>La visión de la realidad</u>, que determina que el observador tenga presente la complejidad del sistema observado. Tendremos en cuenta que encontraremos estos elementos conceptuales:
    - **Eventos:** cualquier acción del entorno sobre el sistema
    - **Patrón de conducta:** explicación simple del observador sobre las reacciones que provocan los eventos
    - **Estructura:** define los patrones de conducta de un sistema

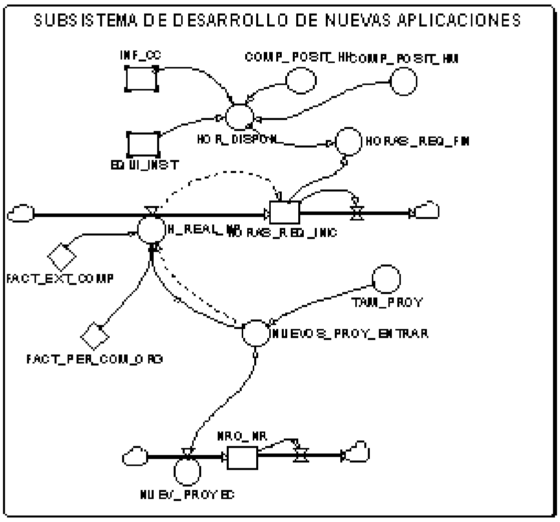
2) <u>Las leyes del pensamiento</u>, que pueden sintetizarse con estos enunciados:
    a. La estructura define el comportamiento
    b. Pensar en todo el sistema pero actuar en sus componentes
    c. Una intervención del observador siempre empeorará las cosas en vez de mejorarlas
    d. Los problemas de hoy fueron las soluciones de ayer

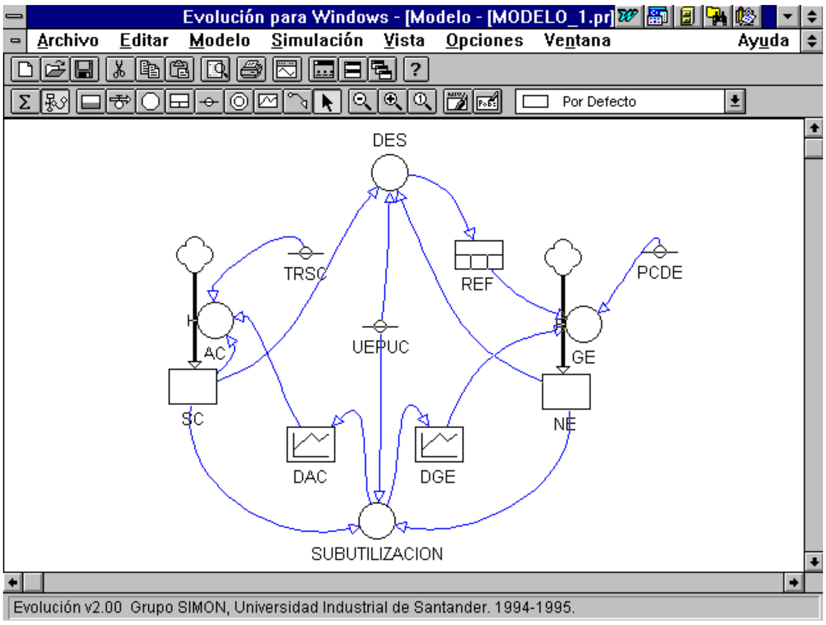
3) <u>La dinámica de los sistemas</u>, que se constituye en la base conceptual para comprender la relación entre estructura y comportamiento.

Las herramientas de la visión sistémica son los diagramas de influencia y los diagramas de Forrester.

Los diagramas de influencia se construyen siguiendo unos pocos pasos importantes, tales como, una vez seleccionado el tema, fijar un límite de tiempo, identificar el comportamiento de las variables en función del tiempo, fijar límites en el sistema, establecer los niveles de detalle.

Los diagramas de Forrester[26] parten de diagramas de influencia. Para obtener una descripción básica de estos diagramas, ver Anexo I de este volumen.





[26] Jay Forrester (1918) es un ingeniero estadounidense, considerado padre de la dinámica de los sistemas. Participó en el desarrollo de la memoria de acceso aleatorio (RAM).

CAPITULO 1 – SEGUNDA PARTE

APLICACIONES DE LA TGS

Ingeniería de sistemas

Entenderemos por *Ingeniería de Sistemas* al conjunto de técnicas interdisciplinarias que se utilizan para la creación de sistemas o al mejoramiento de los existentes.

La herramienta conceptual básica para desarrollar cualquier tarea en este campo es el *Análisis de Sistemas*.

Análisis de sistemas

El análisis de sistemas, como herramienta conceptual, presenta características fundamentales que lo definen y que le dan su valor metodológico. Las características principales son:

- **Enfoque sistémico**: punto de partida fundamental, ya que consiste en considerar como un sistema complejo a cualquier tipo de organización a examinar.

- **Enfoque dinámico**: en el mismo sentido que el punto anterior, se parte de la base que las organizaciones no son estáticas o se desarrollan en estado estacionario.

- **Múltiples niveles**: Para estudiar un sistema de estas características y llegar a comprenderlo es necesario analizar todos sus niveles

- **Multidisciplinario:** el punto de partida, considerando que estamos en un ambiente de avance veloz de las tecnologías, es que los sistemas tienden a ser progresivamente más complejos y, por tanto, para su estudio, resulta imprescindible el abordaje multidisciplinario en cuanto aporta visiones desde otras ramas del conocimiento

- **Descriptivo:** se persigue la capacidad de realizar una descripción completa no solo de todo el sistema sino también de sus componentes

- **Adaptativo:** deriva de considerar como premisa que los sistemas son abiertos. Al estudiar las organizaciones como sistemas abiertos se hace necesaria la posibilidad de adaptarlos para lograr un acople con el entorno y con las posibles variaciones que en ese entorno se produzcan.

Las etapas del proceso de análisis son tres:

- Definición,
- Modelización
- Evaluación.

Para concretar y llevar a la práctica esas etapas necesitamos efectuar una serie de pasos, que en términos generales son[27]:

- Identificar los objetivos y las limitaciones
- Interpretar el funcionamiento
- Describir las posibles acciones de mejora
- Construir el modelo
- Detectar y planificar las acciones apropiadas
- Implementar las acciones

Existen diversas técnicas a utilizar para realizar concretamente el análisis de sistemas, para sus procedimientos y metodologías. Algunos son:

- **Análisis estructural:** que se basa en el conocimiento, lo más profundo que sea posible, de todos los componentes del sistema para diagnosticarlo. El principio es conocer lo más que podamos la estructura del sistema que estudiamos.

- **Estudio de viabilidad:** es la herramienta que permite al analista disponer de elementos para determinar de qué forma las soluciones propuestas pueden afectar al sistema y muchas veces, a su entorno.

- **Simulación en computadora:** modelos de sistemas programados en software para predecir comportamientos

- **Tormenta de ideas:** Grupo de personas que, trabajando sobre el problema, proponen, discuten y evalúan alternativas de solución

- **Árbol de relevancia:** análisis sobre las relaciones entre componentes del sistema, realizado con teoría de grafos y seleccionando los niveles de influencia sobre el resultado final

- **Método Delphi:** Desarrollado por la Corporación RAND[28] en los años 1950, se basa en consultas a expertos por etapas iterativas a fin de lograr consensos sobre las soluciones emergentes

En definitiva, el análisis de sistemas debe estar caracterizado por:
- Una correcta definición del sistema
- Visión integral del sistema
- Esfuerzo multidisciplinario
- Encaminado a todo el ciclo de vida del sistema

Esto que presentamos nos indica la dinámica constitutiva interna del análisis de sistemas, pero debemos tener en cuenta que también se relaciona con otras

[27] Cfr. Libro 1 de esta serie, "Construcción de modelos para programación lineal"

[28] RAND, siglas de "Research and Development", Corporación originalmente llamada "Proyecto RAND", que se dedica a desarrollos e investigaciones para uso del gobierno. Sin embargo hay muchos aportes al medio civil que provienen de ese organismo. Es uno de los principales desarrolladores del Análisis de sistemas, de la computación, de la teoría de juegos y muchos otros campos. La página del organismo (www.rand.org) se describe a sí misma como "*RAND Corporation proporciona servicios de investigación objetiva y análisis de políticas públicas*"

disciplinas, con algunas de ellas de manera muy completa. Algunos de esos campos del conocimiento con los que se relaciona el análisis de sistemas, son:

Sistemas de Información (SI):
Es un conjunto de técnicas y procesos apropiados que interactúan entre sí con el fin de apoyar las actividades de una organización. No siempre un SI debe estar automatizado, en caso de que lo estuviera sería un sistema informático.

Un sistema de información realiza cuatro actividades básicas: **entrada**, **almacenamiento**, **procesamiento** y **salida** de información.

Investigación Operativa (IO)
Es el conjunto de técnica que permite la optimización de un proceso arbitrario bajo múltiples restricciones

Ingeniería de Sistemas Cognitivos
Rama que trata los entes cognitivos, sean humanos o no, como un tipo de sistemas capaces de tratar información y de utilizar recursos cognitivos como la percepción, la memoria o el procesamiento de información. Utiliza tanto psicología cognitiva[29] como ingeniería de sistemas.

Organizaciones
Cuando aplicamos la Ingeniería de sistemas o el análisis de sistemas a sistemas de organización humana (instituciones, empresas, negocios, control social, etc.) nos resulta útil describir el organismo que estudiamos como la organización de un sistema. De esta manera, podremos describir, por ejemplo, una institución a partir de su organización y con un enfoque sistémico. Así, encontramos que una organización dada, tomada como un sistema, tiene subsistemas importantes. Los subsistemas más relevantes son:

- **Subsistema psicosocial**, que comprende las interacciones entre los individuos o grupos o sectores que conforman la organización. Existe interacción intrínseca (entre sectores) y extrínseca (sectores con el entorno)

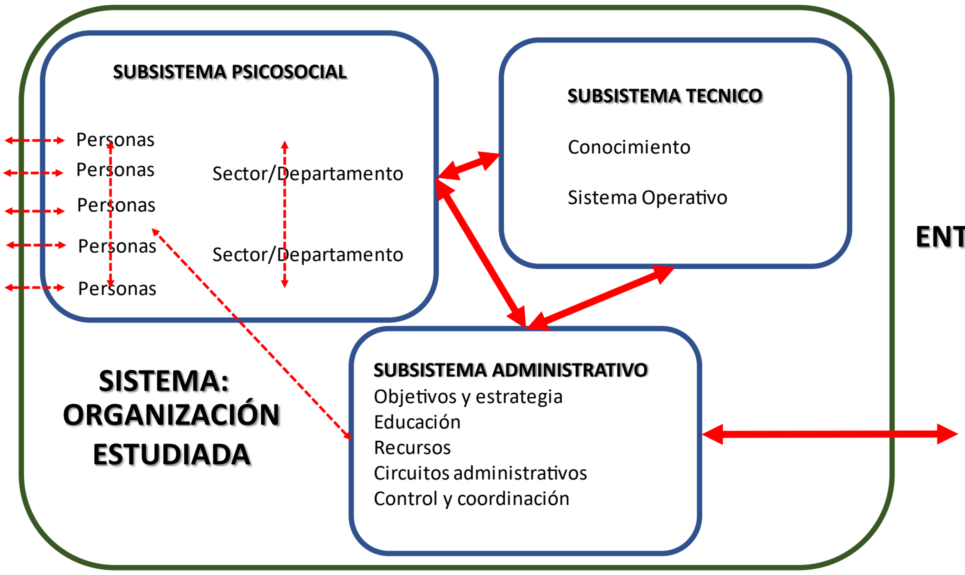
- **Subsistema técnico**: es el conjunto de conocimientos y herramientas operativas que hacen que la organización funcione, cumpla con sus objetivos, entregue sus productos. También conocida como *know how* o sistema operativo.

- **Subsistema Administrativo**: es la relación con el ambiente, por lo cual establece objetivos, planes de integración, estrategias, operación, diseño de estructura y mecanismos de control. Incluye la posibilidad de tomar decisiones sobre la funcionalidad de los otros subsistemas y sobre sí mismo [por ejemplo, se decide un cambio drástico en

---

[29] Psicología cognitiva (cognitivismo) se refiere al estudio de los procesos mentales implicados en la elaboración del conocimiento: la percepción, la memoria y el aprendizaje, la formación de conceptos y el razonamiento lógico. El término ***cognitivo*** se refiere a las acciones de almacenar, recuperar, reconocer, comprender, organizar y usar la información recibida a través de los sentidos. Está interrelacionada con la neurociencia, la inteligencia artificial, la psicología, la lingüística, la antropología y la filosofía.

tecnología. Ello afectará el subsistema técnico pero también al psicosocial (educación para actualización técnica, nuevo personal, etc.) y probablemente produzca también, cambios en los circuitos administrativos — ventas, publicidad, proveedores, etc. — en este subsistema]

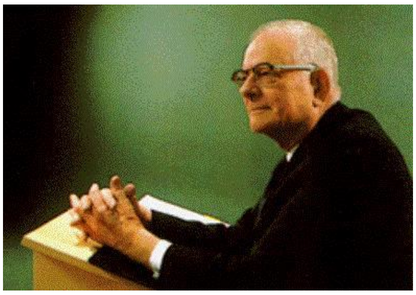


Por otra parte, el análisis de los sistemas de este tipo propende a lograr una evolución favorable y adaptativa a lo largo del tiempo, la cual puede planificarse en base a dos teorías

- Calidad Total
- Quinta disciplina

Calidad total

Concepto formulado por William Deming (1903 – 1993) se basa en el concepto *de El ciclo PHVA*, (planear, hacer, verificar, actuar) se denomina "circulo Deming" en su honor, aunque fue realmente propuesto por Shewhart.



Deming fue contratado por el gobierno japonés en 1947 para cambiar la economía y la base empresarial del país que terminaba de ser vencido en la más terrible y destructora guerra de la historia.

Los estadounidenses, ante el empuje de la industria japonesa de posguerra, recuperan estos conceptos que, paradójicamente, no habían tenido en cuenta, ignorando al propio Deming entre otros.

El concepto de calidad total está fundado sobre una actitud comercial y empresarial que comprende a todos los niveles de la organización, no solo a los jerárquicos y que se pueden resumir en 14 puntos:

1. Crear constancia en la mejora de productos y servicios, para ser competitivo, mantenerse en el negocio, y generar puestos de trabajo.

2. Adoptar una nueva filosofía de cooperación enseñándola a los empleados, clientes y proveedores.

3. Dejar de buscar calidad mediante la práctica de inspección a gran escala de los productos terminados. En su lugar, de buscarse mejorar el proceso y calidad en el producto desde el comienzo.

4. Terminar con la práctica de comprar a los más bajos precios. En lugar de esto, se debe minimizar el costo total con un criterio menos inmediato. Buscar un solo proveedor para cada ítem, basándose en una relación a lo largo del tiempo basada en confianza y lealtad mutua.

5. Mejorar constantemente los sistemas de producción, servicio y planeamiento de cualquier actividad, lo que permite una mejora de la calidad y la productividad, bajando los costos constantemente.

6. Establecer programas de capacitación y entrenamiento dentro del ambiente de trabajo.

7. Identificar líderes, reconociendo sus diferentes habilidades, capacidades y aspiraciones. El objetivo de la supervisión debería ser ayudar a la gente, máquinas y dispositivos a realizar su trabajo.

8. Eliminar el miedo y fomentar confianza en sí mismos de las personas, de manera tal que el resultado sea que la tarea individual sea realizada más eficientemente.

9. Borrar las barreras entre los departamentos. Abolir la competencia y construir un sistema de cooperación basado en el mutuo beneficio que abarque toda la organización.

10. Eliminar eslóganes, exhortaciones y metas pidiendo cero defectos o nuevos niveles de productividad. Esto solo crea rivalidad, la principal causa de la baja calidad y productividad reside en el sistema y no en el poder de la fuerza de trabajo.

11. Eliminar cuotas numéricas y la gestión por objetivos.

12. Remover barreras para apreciar la mano de obra y los elementos que privan a la gente de la alegría en su trabajo. Esto incluye eliminar las evaluaciones anuales o el sistema de méritos que da rangos a la gente y crean competencia y conflictos.

13. Instituir un programa vigoroso de educación y auto mejora.

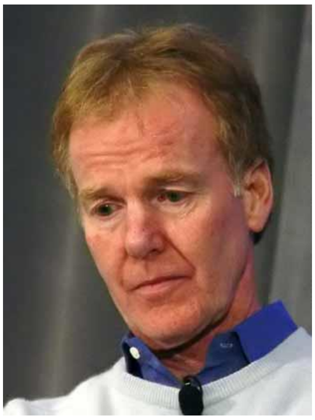
14. Poner a todos a trabajar para llevar a cabo la transformación. La transformación es trabajo de todos.

En paralelo, Dening había señalado lo que él llamaba “las siete enfermedades mortales de una organización”:

1. No definir los objetivos o propósitos generales de la empresa claramente o cambiarlos a menudo.
2. Buscar la obtención rápida de ganancias
3. Evaluar al personal y a los integrantes de la empresa por rendimiento, o revisión periódica de resultados
4. Cambiar a menudo y con gastos elevados a los ejecutivos
5. Administrar basándose solamente en las cifras visibles o costos tangibles
6. No tener adecuados planes de salud y prevención para el personal, que implica un gasto excesivo en medicina.
7. Costo elevado en garantías financieras y físicas

Quinta Disciplina

El concepto de "Quinta disciplina" fue establecido por Peter Senge, nacido en 1947 y graduado en ingeniería en la Universidad de Stanford. Hizo un Master en *Social Systems Modeling* en el MIT. Actualmente (consultado en abril 2020) es el director del Centro para el Aprendizaje Organizacional del mismo MIT. En los años 1990 publicó su libro *The Fifth Discipline* donde desarrolla la noción de organización como un sistema (desde el punto de vista de la Teoría General de Sistemas), en el cual expone un cambio de mentalidad profesional.



Se basa en el concepto de organizaciones que deben aprender continuamente para lograr mejores resultados. Ese proceso de aprendizaje se basa en cinco disciplinas. Hay que señalar en este punto, que el concepto que utiliza Senge para el término "disciplina" NO es el "rama del conocimiento" que generalmente encontramos en el ambiente universitario, sino el que se refiere al comportamiento de las personas o grupos humanos (por ejemplo: "en este equipo mantendremos la disciplina")

Como los empleados deben tener libertad para aprender, mejorar y proponer mejoras, se parte del concepto de que la organización es un sistema abierto que tiene – o sus componentes tienen – las siguientes capacidades:

1. Se puede aprender tanto de hechos internos como externos
2. Se deben buscar y evaluar nuevas prácticas en todas partes
3. Se debe poner el aprendizaje como centro de las reuniones a través de líderes
4. Se debe enfrentar a los críticos y escucharlos atentamente para aprender de ellos
5. Se debe propiciar la creación de relaciones estratégicas con clientes, proveedores y competidores para aumentar la capacidad de aprendizaje
6. Es necesario y conveniente establecer comunicación de ida y vuelta entre jefes y empleados
7. Hay que escuchar las opiniones de clientes y accionistas

Los cinco modos de comportamiento, **las cinco disciplinas**, se pueden sintetizar de la siguiente manera, que se focaliza en el concepto de retroalimentación (*feed back*)

1. Continuamente cada uno profundiza y desarrolla la visión personal, poniendo el énfasis en tener más paciencia, en ver la realidad de manera objetiva y de disponer de tiempo para reflexionar sobre su entorno.

2. Cada individuo debe aprender a identificar y desarrollar los modelos mentales de manera tal que pueda comprender como se entiende el mundo y se ejecutan las acciones
3. Cada persona debe impulsar la visión compartida como fuente de inspiración y productividad en la que creen todos los empleados y los motiva para participar y no como una obligación a cumplir
4. Se debe fomentar el trabajo en equipo para que todos sus miembros asuman lo que los demás están pensando a fin de generar un auténtico pensamiento grupal
5. Se debe generar el pensamiento sistémico. Esta forma de ver la organización es la Quinta disciplina y la que integra los resultados de las otras cuatro.

Aplicaciones de la Teoría General de Sistemas

Si bien existen varias disciplinas y metodologías que derivan de la Teoría General de Sistemas, podemos arriesgarnos a afirmar que, entre ellas, las más conocidas e importantes son: Investigación Operativa, Cibernética y Teoría de la Comunicación.

Investigación Operativa[30]

Se trata de un conjunto de técnicas orientadas a la tarea de recopilación y análisis de información sobre operaciones generales para optimizar, tomar decisiones y planificar en el corto, medio y largo plazo.

Comenzó su desarrollo a fines de la década de 1930 en el Reino Unido y en algunas empresas de USA (principalmente Bell y CBS). A fines de la década de 1930 se creó la sección de Investigación Operativa de la RAF (*Royal Air Force,* Fuerza Aérea Real) y luego en el comando de defensa naval, ambas instituciones de Gran Bretaña creadas frente a la posibilidad de una guerra generalizada en Europa, la cual se concretó en 1939 cuando Alemania invadió Polonia y, posteriormente, comenzó su avance hacia el oeste europeo. Gran Bretaña se vio amenazada de invasión marítima, sufriendo rápidamente ataques frecuentes por vía aérea y el bloqueo, por parte de la flota de submarinos alemanes, de sus suministros enviados desde el continente americano.

Los métodos que se usaron en el sistema de alertas tempranas, intercepción para los ataques aéreos enemigos y en el desarrollo de la táctica de ataques a los submarinos alemanes que bloqueaban los suministros a partir de 1939, fueron perfilando lo que hoy se conoce universalmente con la denominación de Investigación Operativa, y, en ese momento se refería a las operaciones táctica a emplear en ambos casos.

Estas acciones se inscriben en el comienzo de un conflicto armado de grandes proporciones que luego se denominó Segunda Guerra Mundial y que se desarrolló entre 1939 y 1945. Involucró, directa o indirectamente a más de 80 países y combatieron efectivos de alrededor de 50 de esos países.

Con estas técnicas se planificó el mantenimiento y la selección de tipos de aviones y demás armamento adecuados a cada misión y, con buen éxito, destrucción de submarinos alemanes optimizando el empleo de nuevas tecnologías de detección y utilización óptima del armamento disponible.

[30] Ver Libro 1 de Optimiza

Hubo dos éxitos fundamentales en el comienzo de la guerra: La defensa de Londres en la Batalla de Inglaterra fue el campo de experimentación. Menos popular pero más efectivo fue el éxito en la batalla del Atlántico contra las “manadas de lobos” como se llamaron a los grupos de submarinos alemanes dedicados a destruir convoyes de aprovisionamiento, en forma conjunta y coordinada, con resultados letales.

Para los británicos (y sus aliados), atacar a un submarino alemán requería disponer de unas 34.000 horas hombre con una probabilidad del 2 al 3% de destruirlo, con lo cual, en promedio, la destrucción concreta de una unidad insumía alrededor de 1.500.000 horas hombre.

El ataque a un submarino se realizaba cuando se detectaba o se tenían indicio de su presencia, navegando en círculos sobre la zona sospechosa y arrogando una cierta cantidad de bombas que estallaban a una profundidad prefijada (se denominan “cargas de profundidad”). La respuesta ofrecida por la investigación operativa consistió en encarar el problema tomando en cuenta seis variables de decisión, que en términos simplificados eran:

1. Profundidad o tiempo de explosión de las cargas
2. Radio de acción de las cargas
3. Orientación del ataque respecto al probable rumbo del submarino
4. Separación entre las cargas lanzadas
5. Miras
6. Errores de puntería al disparar

Los resultados fueron demostrativos del éxito de la técnica, junto con el desarrollo de otras ayudas, tales como un sistema de ecos sonoros (SONAR) para detección en inmersión y el incipiente uso de ecos electromagnéticos (RADAR) para detección en superficie. Si se analiza como variable de salida la cantidad de submarinos destruidos hasta 1941 fue

1939: 9 unidades
1940: 24 unidades
1941: 35 unidades

Al comenzar a aplicarse la técnica de investigación de operaciones, a fines de 1941, los resultados fueron:

1942: 87 unidades
1943: 237 unidades
1944: 242 unidades
1945: 151 unidades (fin de la guerra en mayo)

TOTAL: 785 submarinos perdidos (88,0%)
Supervivientes: **108 unidades** (12,0%)

En cuanto a la defensa territorial frente a los ataques aéreos, el resultado fue espectacular, al punto que algunos historiadores creen que Alemania desistió de la invasión al Reino Unido al no poder doblegar a su fuerza aérea. Estas acciones reciben el nombre general de "Batalla de Inglaterra" y fueron un conjunto de operaciones aéreas que ocurrieron entre el 10 de julio de 1940 y el 17 de septiembre de ese año, con algunos ataques posteriores hasta mayo de 1941 y bombardeos aislados de Londres hasta 1945).

Se recuerda, particularmente, el día 15 de septiembre de 1940 como el punto de inflexión de la batalla. Ese día dos olas de bombarderos protegidos por escoltas alemanes fue efectivamente interceptada por una cantidad mucho menor de cazas piloteados por ingleses y polacos operando coordinadamente según las técnicas que nos ocupan. A partir de ese día, oficialmente, Alemania cesó su plan de invasión a las islas británicas. La diferencia en el número de aviones, al comienzo de la batalla era enorme: 3600 alemanes contra 870 ingleses.

Cibernética

El vocablo "Cibernética" significa "Hombre que conduce" y es la ciencia que estudia *el control y comunicación en los animales y máquinas* (fuente: RAE). Es un término introducido por Norbert Wiener (**Cibernética** o **el control y comunicación en animales y máquinas**, 1948).

Se parte de la concepción de que los sistemas complejos afectan y luego se adaptan a su ambiente externo. Esta relación, en términos técnicos, se centra en funciones de control y comunicación: ambos fenómenos son simultáneamente externos e internos del/al sistema. Esta capacidad es natural en los organismos vivos y se ha imitado en máquinas y organizaciones. Merece una especial mención la retroalimentación y los conceptos y subsistemas que provienen de ella. Incluso podemos encontrar sistemas donde el observador forma parte de él.

A partir de la cibernética se desarrollaron varios campos conceptuales y disciplinas de investigación: Entre ellas podemos señalar:

- **Biónica**, que se refiere al reemplazo, sustitución o emulación de partes animales por mecanismos, artefactos o sustituciones robóticas.
- **Robótica**, es el desarrollo, construcción y operación de artefactos con sensores capaces de receptar información de entrada y enviarla a un procesador para realizar determinadas operaciones mecánicas, electrónicas o combinadas.
- **Inteligencia Artificial**, ciencia cuyo objetivo es ayudar a las máquinas a encontrar soluciones a problemas complejos imitando procesos mentales, particularmente relacionados con aprendizaje y toma de decisiones.

Teoría de la Información

La *Teoría de la Información* también es llamada **Teoría Matemática de la Comunicación** *(Mathematical Theory of Communication)* es una propuesta teórica presentada por Claude E. Shannon y Warren Weaver en 1948. Está relacionada con las leyes matemáticas que rigen la transmisión y el procesamiento de la información. Se ocupa de la medición de la información, de su representación y de la capacidad de los sistemas para transmitir y procesar información. Actualmente es una rama de la matemática y de las ciencias de la computación que estudia la información y todo lo relacionado con ella: canales, compresión de datos, criptografía, almacenamiento, clasificación y recuperación de mensajes y temas relacionados.

Componentes de un sistema de información:

FUENTE
Componente humano o mecánico que determina el tipo y complejidad del mensaje a transmitir

TRANSMISOR
Componente encargado de transformar el mensaje original en señales apropiadas para su transmisión

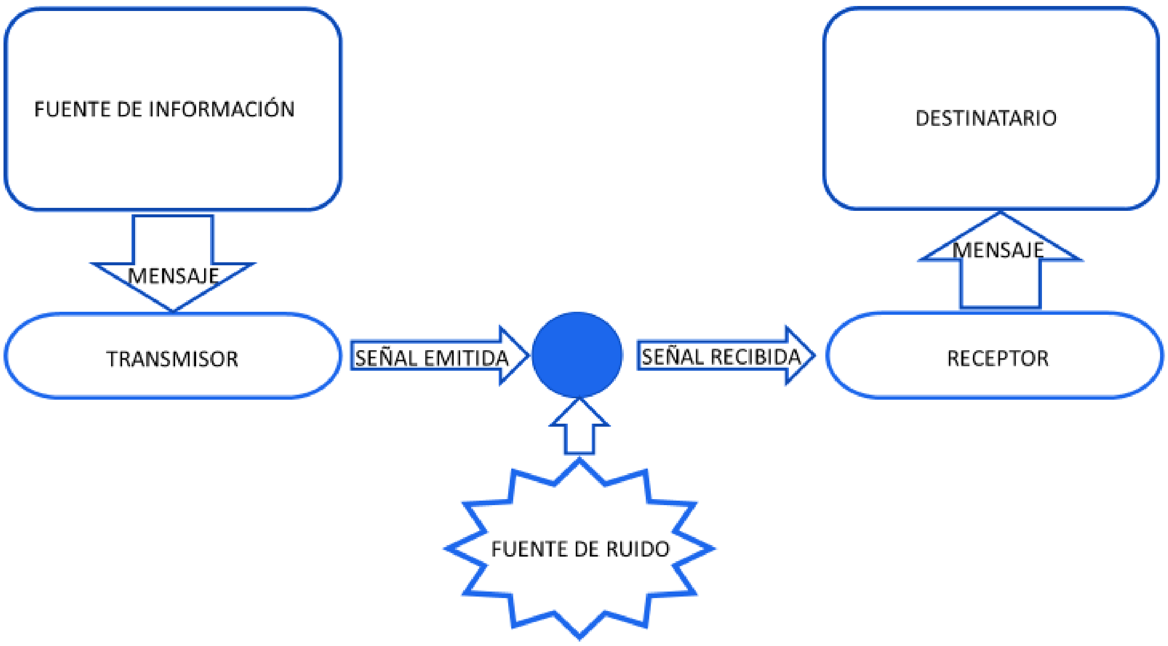
CANAL
Medio que se encarga de transportar las señales en el espacio. Función de Transporte

RECEPTOR
Medio técnico encargado de traducir la señal recibida al formato y lenguaje del mensaje original enviado por la fuente

DESTINATARIO
Componente al cual fue emitido el mensaje por la fuente

RUIDO
Elementos que se agregan aleatoriamente a la señal y que puede influir en la correcta interpretación del mensaje.

CODIFICACIÓN
Forma física en la que se transmite el mensaje, la cual depende del medio o soporte en que se transmite, emite y recibe (por ejemplo modulado sobre una onda portadora) o del nivel de seguridad que se desea frente a posibles desvíos o intercepciones (criptografía).



Niveles de la Teoría de la Comunicación

Weaver enunció que, en todo sistema de comunicación, existen tres aspectos conceptuales o ***niveles*** que deben ser abordados. Ellos son: el nivel técnico, el nivel semántico y el nivel pragmático.

- **Nivel Técnico**

Son los aspectos que se relacionan con la fidelidad o deformación en que los mensajes pueden ser transportados desde la fuente al destino. La falta de fidelidad puede ocasionar distorsiones, cortes o fragmentación, elementos espurios (ruido)

- **Nivel Semántico**

Se refiere a los problemas relacionados con el significado y/o la interpretación de los mensajes. Se busca que la codificación (lenguaje) sea un acuerdo entre emisor y receptor.

- **Nivel Pragmático**

Aborda el comportamiento que el receptor adquiere a partir del significado del mensaje. Es la reacción ante el contenido del mensaje, que, en este nivel adquiere el significado de estímulo sobre el receptor.

Seguridad

Todo sistema de información es pasible de ser interceptado por un tercer componente diferente al emisor y destinatario, que son los elementales. Por ello en casos determinados se emplean diferentes niveles de protección a fin entorpecer, dificultar dicha intercepción.

Hay dos niveles de protección:

1) el básico es proteger la vía de transmisión de posibilidades de apertura o desvío, lo cual es muy simple en sistemas de transporte físico (correos) y muy complicado en sistemas de transporte naturalmente abiertos (radio, cable, medios físicos públicos, etc.). La diferencia estiba en que en los transportes físicos se trata solamente de blindar adecuadamente el medio de transporte y protegerlo físicamente. En cambio, cuando se transmite por radio (cuyas señales pueden ser captadas y decodificadas con facilidad, sean estas de baja, alta, ultra alta frecuencia o microondas), por cable en distancias medias y largas o con correos que no son propios o posibles de proteger individualmente, la protección del transporte es muy complicada.
2) el nivel criptográfico o codificado, que se aplica a cualquier sistema de transmisión, y que consiste en modificar la naturaleza del mensaje con un acuerdo previo y privado entre emisor y receptor sobre la naturaleza y manera de modificar y decodificar el mensaje. Generalmente se emplean técnicas criptográficas (***criptografía***: arte de escribir con clave secreta o de un modo enigmático, según la RAE), algunas de las cuales describiremos más adelante.

De todas maneras el tema de seguridad de comunicaciones es un tema central en la teoría de comunicaciones, ya que introduce un fuerte factor de ruido o error en el proceso si no hay seguridad de interferencias desconocidas o no deseadas, lo que produce errores o incertidumbres en la interpretación del mensaje recibido.

CAPÍTULO 2

MODELOS

modelo.

(Del it. *modello*).
1. m. Arquetipo o punto de referencia para imitarlo o reproducirlo.
2. m. En las obras de ingenio y en las acciones morales, ejemplar que por su perfección se debe seguir e imitar.
3. m. Representación en pequeño de alguna cosa.
4. m. Esquema teórico, generalmente en forma matemática, de un sistema o de una realidad compleja, como la evolución económica de un país, que se elabora para facilitar su comprensión y el estudio de su comportamiento.
5. m. Objeto, aparato, construcción, etc., o conjunto de ellos realizados con arreglo a un mismo diseño. (Auto modelo 1976. Lavadora último modelo).
6. m. Vestido con características únicas, creado por determinado modista, y, en general, cualquier prenda de vestir que esté de moda.
7. m. En empresas, u. en aposición para indicar que lo designado por el nombre anterior ha sido creado como ejemplar o se considera que puede serlo.(Empresa modelo. Granjas modelo).
8. m. Esc. Figura de barro, yeso o cera, que se ha de reproducir en madera, mármol o metal.
9. m. Cuba. impreso (hoja con espacios en blanco).
10. com. Persona de buena figura que en las tiendas de modas se pone los vestidos, trajes y otras prendas para que las vean los clientes.
11. com. Esc. Y Pint. Persona u objeto que copia el artista.

"Diccionario de la Lengua Española", 23ª. ed. [versión 23.2 en línea] <https://dle.rae.es>

Como mostramos en el recuadro, disponemos de varias acepciones para el término "modelo", de las cuales, para esta asignatura – y en la mayor parte de las acciones de la vida profesional cotidiana – prestaremos especial atención a la 3 y a la 4.

Mediante la utilización de *modelos* es posible representar un sistema, resaltando los elementos componentes (u objetos) y las relaciones entre ellos, pero limitando la representación solo a aquellos componentes o relaciones que interesan al observador o que éste considera relevantes.

Hay que tener en cuenta que un modelo es un sistema en sí mismo, que generalmente se caracteriza por poseer menos interrelaciones y elementos que el sistema que representa.

Podríamos intentar clasificar los modelos según varios criterios. Esa clasificación solamente es útil para comprender mejor el concepto de lo que es un modelo y la manera en que, cotidianamente, los utilizamos, muchas veces sin ser conscientes de ello. Veamos algunos ejemplos de una supuesta clasificación conceptual de modelos que conocemos y usamos:

- el *plano*, o *lay out*, de una planta es un *Modelo icónico*[31]. Estos modelos son de una gran utilidad y los construimos o utilizamos casi a diario. Imaginemos el ***modelo mental*** que nos formamos de una casa. Queremos realizarla. Podremos explicarle a un arquitecto nuestra idea, él construirá un ***modelo icónico*** que generalmente llamamos "plano", en el cual "veremos" plasmada nuestra o idea o podremos entender modificaciones son necesarias para que sea satisfactoria. Quizá debamos hacer varios ajustes, pero siempre a escala de modelo, sin mover un solo ladrillo. Ese modelo, en manos del constructor, será utilizado para erigir nuestra idea.

- Los modelos *analógicos*[32] son aquellos que explican las relaciones mediante la aplicación de lo conocido o con fenómenos comparables: el modelo de la resistencia eléctrica de Ohm se utiliza para los fenómenos de transporte de energía térmica, los economistas hablan de flujo de dinero, en analogía con un modelo hidráulico: taponamiento, drenaje de capitales, suba de nivel de tasas, etc.

- Los *modelos matemáticos*, son capaces de presentar mediante números abstractos todo lo ocurrido durante el mes, por ejemplo, en los ingresos y egresos de la cuenta de caja de ahorro (por otra parte la terminología "caja de ahorro" es un modelo analógico en sí misma).

- Los modelos *simbólicos* representan ciertas relaciones sistémicas mediante símbolos, como el que utilizan los químicos, que a un fenómeno tan complejo como la reacción de ionización de la sal gema en agua lo representan con

$$ClNa + H_2O \leftrightarrow Cl^- + Na^+ + HO^- + H^+,$$

  Este mismo texto es una colección de símbolos (letras) cada una de las cuales representa un sonido y que es diferente según la cultura. Un fonema chino, por ejemplo, puede parecer un dibujo más o menos bonito para un argentino.

---

[31] La palabra "*ícono*" proviene del griego y significa imagen o representación. (*εἰκών*, *eikon*, "imagen") Las religiones, en general, son proclives a utilizar íconos para representar ideas muy abstractas. Los sistemas operativos los utilizan en sus interfaces de usuario para representar comandos más o menos complejos o con muchas sentencias.

[32] Debe distinguirse el uso del término "analógico". Podríamos ensaya tres aplicaciones: la informática, donde se usa – en contraposición a "digital" – para designar señales variables en forma continua o no-discreta; la que se refiere a mecanismos configurados por analogía a las leyes matemáticas y el uso lingüístico original que se refiere a comparaciones en aspectos particulares, que es el que se emplea en estos modelos.

Además, y de manera independiente de lo anterior se pueden clasificar de acuerdo a la funcionalidad:

- Los modelos *deterministas* son aquellos en los cuales los componentes están relacionados entre sí por funciones conocidas y perfectamente predecibles.
- Lo contrario a ello, son los modelos *estocásticos*, en los que las funciones pueden ser desconocidas o conocidas en términos probabilísticos.

Independientemente de la propuesta de tipificación anterior, también se puede separar por su comportamiento en función del tiempo:

- Los modelos *estacionarios* son los que describen sistemas y sus comportamientos independientemente de los fenómenos que ocurren en el inicio o en el fin de las relaciones (por ejemplo, el modelo de universo heliocéntrico describe a la tierra girando alrededor del sol, sin tener en cuenta el resto del universo ni lo que ocurrió en el pasado, si tuvo un origen o no, su evolución – cómo llegó el sol al estado actual o si se trata de un fenómeno sin evolución (inmutable) – y su futuro)
- Por último se mencionan los modelos *dinámicos*, que son aquellos cuyo propio comportamiento se modifica a medida que es utilizado como representación, ya que recibe y emite información desde y al sistema que representa a fin de adaptarse a la realidad. Generalmente se van resolviendo secuencialmente.

MODELOS

En 1945 Rosenblueth y Wiener señalaron la importancia de un modelo:
"Ninguna parte sustancial del universo es lo suficientemente simple como para que pueda ser comprendida y controlada sin abstracción, la que consiste en remplazar la parte del universo bajo consideración por un modelo de estructura similar pero más simple. Los modelos constituyen un elemento central del proceso científico"

El uso de modelos simbólicos y matemáticos es el más difundido, pues se construyen con elementos baratos y simples de adaptar. El ejemplo del plano de la casa que mencionábamos más arriba, muestra claramente la simpleza de realizar un modelo icónico para ensayar ajustes y cambios.

En cualquier caso, el modelo será una simplificación, por lo tanto contendrá solamente *algunas* variables y pocas veces todas (si las tuviera sería la realidad o un prototipo y no un modelo conceptual). Decidir qué variables se usan y cuales se descartan implica el uso de hipótesis[33] y de un modelo de control[34].

---

[33] Una hipótesis puede caracterizarse como una solución provisional (tentativa) para un problema dado. El nivel de verdad que se le asigne a tal hipótesis dependerá de la medida en que los datos empíricos recogidos apoyen lo afirmado en ella. Esto es lo que se conoce como contraste o proceso de validación de la hipótesis. Este proceso puede realizarse de uno de dos modos: mediante confirmación o mediante verificación.

[34] Según Fayol, 1961, "el control consiste en asegurarse de que todo lo que ocurra esté de acuerdo con las reglas establecidas y las instrucciones dadas". El Modelo de Control parte del estudio del entorno y el análisis de la estructura. Tal análisis permite establecer las amenazas y oportunidades que ofrece el entorno y también las debilidades y fortalezas. Luego, se está en condiciones de realizar un diagnóstico, de proponer soluciones, de

Estudiar directamente la realidad puede ser muy caro, o requerir mucho tiempo; incluso puede ser imposible, porque hay casos en que el sistema no existe (es hipotético) o, si existe, pero en realidad es imposible prácticamente experimentar con él o no se puede cambiar arbitrariamente ciertas actividades o propiedades sin causar graves trastornos irreversibles en el sistema dado.

Pero aún tenemos otra manera de clasificar los modelos. Es la siguiente:

- **Físicos**: Se asemejan al sistema en estudio. Son representaciones de sistemas físicos y se describen por variables medibles.
    - Modelos a escala (aviones, autos, reactores, edificios, planta piloto, etc.)
    - Modelos de imitación
    - Modelos de analogía
    - Prototipos. (Como se señaló en párrafos anteriores, los prototipos son modelos muy singulares, ya que generalmente tienen todas las variables y complejidad del objeto estudiado).
- **Mentales**: son modelos heurísticos o intuitivos sin existencia más que en la mente del modelizador. En general es privativo del ser humano, ya que, entra en esta categoría el modelo de mundo que cada ser humano tiene en función de su experiencia. Estos modelos suelen derivar en modelos simbólicos para permitir difundirlos y sortear la dificultad de comunicación. (Mozart podía tener un modelo mental del aria de una ópera, pero si no escribía las partituras (pasarlo a modelo simbólico) ninguna orquesta podría interpretarla)
- **Simbólicos**: son los que realizan operaciones matemáticas o lógicas. Son más baratos y fáciles de construir que los físicos. Se dividen en dos grupos:
    - Matemáticos
    - No Matemáticos
        - Verbales
        - Gráficos
        - Esquemáticos

Pasos para formular un modelo:

Podemos describir la metodología básica para construir un modelo con mucho detalle, en cuyo caso tendremos que prever varias alternativas que se adapten a diferentes circunstancias, o, al contrario, sin detalles y solo en términos muy generales, que se va a adaptar mejor a más casos y que deberemos completar con las particularidades que hagan falta en cada caso.

En ese caso, una guía metodológica básica sería:

i) Estudiar el problema, identificar las variables de control y las no controlables. Identificar los parámetros. Evitar los prejuicios.
ii) Tratar de expresar el problema matemáticamente. (Símbolos)
iii) Definir el campo de aplicación y tratar de aislarlo de las interacciones.
iv) Verificar si el problema tiene una solución y si el modelo puede usarse en otros problemas.
v) Probar la solución: verificar el grado de descripción de la realidad.
vi) Ajustar el modelo y volver a empezar.

---

analizar y mejorar los sistemas funcionales, planteando la situación actual, el objetivo deseado y como llegar a él". (Dianelys Nogueira Rivera – Universidad de Matanzas, Cuba.)

En ese caso, hay que tener en cuenta algunos aspectos importantes, que detallaremos seguidamente.

Definición del Problema

Es el primer paso obligado y quizá sea el más importante, pues tiene que estar bien hecho. Consiste en la identificación y la capacidad de describir con precisión el problema enfrentado. Implica conocer el objetivo global, las limitaciones fundamentales y las limitaciones que es conveniente introducir. Muchas veces no se tiene conocimiento de estos elementos, lo cual requiere de un profundo análisis y – generalmente – negociaciones con los miembros del equipo de trabajo.

Este paso es importante porque puede haber puntos potencialmente conflictivos aunque no se perciban en un análisis simplista. Tomemos uno de los varios puntos en que puede existir divergencias de opinión. La definición del *objetivo* puede ser uno de esos puntos de conflicto. En un mismo caso, puede ocurrir que nos encontremos con más de un objetivo deseable y que éstos, por añadidura, sean opuestos. Por ejemplo, podemos establecer como objetivo deseable construir una PC para juegos de máxima prestación, y que otro objetivo sea construirla con el menor costo en componentes. Ambos objetivos son válidos y deseables, pero contrapuestos. Es difícil elegir uno sobre otro: el objetivo de minimizar los costos de producción suele ser antagónico con el objetivo de maximizar la calidad del producto terminado. ¿Cuál es el objetivo global? Si hay una decisión política o empresarial sobre los costos de producción, entonces habrá que establecer una restricción de nivel de calidad. Y en todo caso ¿cuáles serán las unidades de medida? ¿cuáles son los puntos de compromiso entre calidad y costo? ¿cuáles son los mínimos aceptables en cada caso?

Claramente responder a estas preguntas no es posible usando un algoritmo o un método rígido, sino que requiere opiniones, acuerdos entre representantes de diversos enfoques, búsqueda de caminos alternativos y un largo etcétera.

También puede ocurrir que no se perciba con claridad cuál es el verdadero problema o el meollo del asunto que se quiere abordar, sino que se tenga una imagen diferente de la realidad. Un modelo construido sobre esa base no brindará elementos de juicio y esto ocurre con frecuencia ante problemas mal definidos o con objetivos irreales.

Como definir el problema

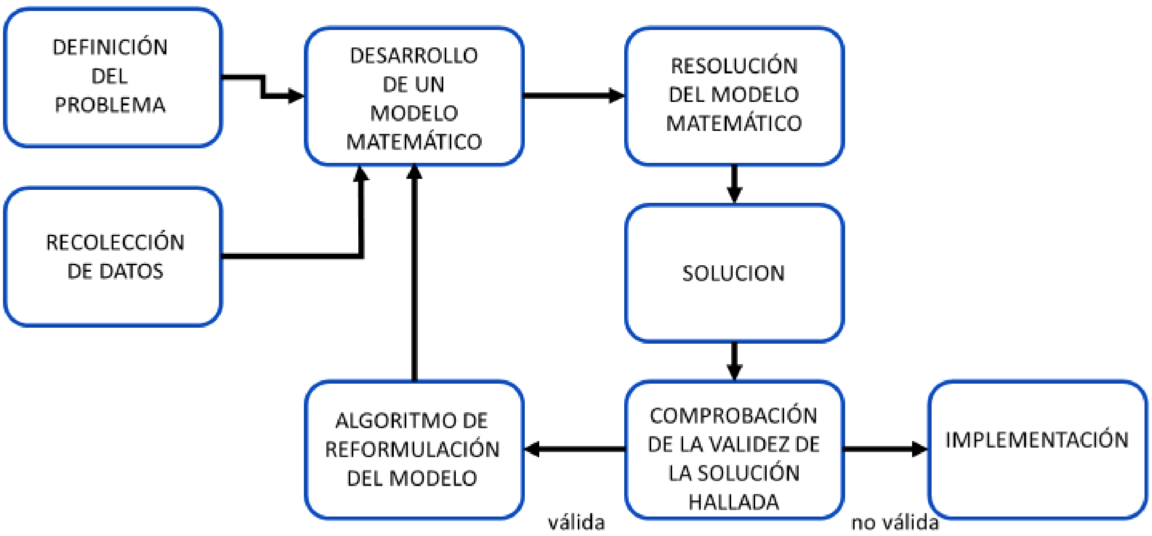
Si bien puede haber otras formas para definir un problema, podríamos estables el siguiente proceso general:

1) Establecer el objetivo o los objetivos para satisfacer los cuales estamos abordando el problema.
2) Conocer o adaptar la política empresarial o institucional para armonizar estos objetivos con ella y entre ellos.
3) Buscar los consensos necesarios. Este aspecto suele ser denominado “político” y conduce a la efectiva enunciación de un problema.
4) Reconocer las limitaciones o recursos limitados para lograr el objetivo.

Podemos decir, sin temor a equivocarnos, que esta es la etapa más importante en el desarrollo de un modelo. Pensemos solamente en el significado del término “problema”. Hemos usado esa palabra para, por ejemplo, ejercicios a resolver en un examen o en un práctico. Sin embargo, si lo analizamos con nuestra óptica actual, veremos que, en realidad, no eran problemas.

Esos llamados *problemas* eran anunciados previamente. Tenían títulos tales como problema N° 14 o cosas similares. En nuestra vida profesional, al contrario, no sabremos nunca cuando y como se presenta un problema. La otra característica de esos problemas es que tenían un enunciado y los datos necesarios para resolverlos. En el mundo real, los datos no están. Hay que buscar cuales son los necesarios y como se consiguen e interpretan.

Los problemas enunciados venían precedidos de un contexto claro: teníamos las herramientas conceptuales necesarias para desarrollarlos definidas. Si estábamos cursando movimiento rectilíneo uniforme en la asignatura física, el problema que debíamos abordar se relacionaba con ese tema (y no con termometría, por ejemplo y mucho menos con dinámica de grupos de estudios) y habíamos visto recientemente los modelos necesarios para resolverlo, a los que llamábamos “fórmulas a aplicar”.



Esto jamás nos ocurrirá en nuestra labor: nunca estaremos muy seguro si nuestro problema es de, por ejemplo, física elemental de la electricidad o de física termodinámica o simplemente de humedad en el gabinete que contiene ese circuito que se recalienta.

Por último, la mayoría de los problemas tenían solución, que era un conjunto de valores o conclusiones que debían coincidir con los valores o conclusiones aceptados por quien nos había solicitado que resolviéramos el problema. En la vida profesional a veces no encontraremos resultados claros y estos, cuando los tengamos no siempre podremos compararlos con valores preestablecidos o con las "respuestas a problemas seleccionados" que a veces aparecían al final de los textos y deberemos tomar decisiones sobre la validez de las conclusiones a las que arribemos.

En cuanto a la necesidad de ***establecer claramente los objetivos***, debemos tener en cuenta que probablemente hay que estar seguros de que estamos todo de acuerdo con ellos y con la política general de la empresa o institución donde estamos aplicando estos métodos. Por ejemplo, tenemos como objetivo hacer el mejor producto. Este objetivo no es claro y no significa lo mismo para todos. Para control de calidad, será el que no presente fallas, para ventas será el que más impacte sobre el consumidor y brinde la mayor ganancia, para producción el más fácil de ensamblar, para el analista de costos el que lleve menos costos de componentes y mano de obra. Esto de armoniza logrando acuerdos entre todos ellos y ajustándolo a la política de la empresa, que puede ser buscar máxima calidad o ventas masivas por bajo costo…

También es importante conocer de antemano las limitaciones que tenemos para desarrollar tanto el modelo como para implementar la futura solución que estamos buscando. Estas limitaciones, que aplican en ambos aspectos, son, comúnmente, mano de obra, capacidad técnica o científica, acceso a tecnología, recursos materiales, económicos, tiempos, etc. Por otra parte, sin alguna o todas de estas limitaciones, tampoco sería posible definir un problema. Imaginemos por un instante que alguien nos plantee el siguiente problema: "Quiero construir este equipo. Haga el diseño". Y que agregue que no hay límites de tiempo para hacerlo, que tenemos todos los recursos que queramos disponibles, que el presupuesto es ilimitado, que la mano de obra es toda la que necesitemos, que podemos usar libremente la última tecnología que esté disponible (estado del arte), etc. Obviamente no tendríamos nada que decir y probablemente poco que hacer, porque a medida que lo vamos haciendo lo iremos mejorando y rediseñando… eternamente.

Desarrollo de un modelo matemático DETERMINÍSTICO y recolección de datos.

Veremos un caso muy simple con la intención de introducirnos en forma sencilla en la metodología para formular un modelo, tal como la describimos precedentemente:

*Nuestra pequeña planta de armado de tablets tiene una capacidad de producción fija, por lo cual hemos decido fabricar y vender dos versiones, los que llamaremos Mark **A** y Mark **B** y usar el total de recursos que dispongamos luego de solicitar un fondo para inicio de producción.*

Disponemos de la siguiente información:
*a-La ganancia promedio esperada del Mark **A** es del 10% por encima de los fondos aportados, que se recuperan y devuelven totalmente.*
*b-La ganancia promedio esperada del Mark **B** es del 6%, pero con mejores probabilidades de venta.*

c-Establecemos que no vamos a aportar más del 75% en ninguna de las dos alternativas, para mantener la imagen de la marca con al menos dos versiones.
*d-Por seguridad los fondos a otorgar para Mark **A** no deberán exceder el doble de lo que aportaremos en Mark **B**.*

Deberemos calcular

¿Cuántas posibilidades de inversión hay?
¿Cuál sería el modelo que permite tomar decisiones?
¿Cuál sería la mejor combinación de aportes de fondos?

Variables de decisión y función objetivo

Para llevar este problema a una forma de modelo matemático comenzaremos por definir que es aquello sobre lo cual podemos tomar decisiones. En la terminología de la Investigación Operativa se emplea el término:

IDENTIFICAR LAS VARIABLES DE DECISIÓN

En este caso, las *variables de decisión* serán dos y a cada una de ellas se les asigna, un símbolo arbitrario y – cuando corresponde – las unidades correspondientes:

tA = fracción del fondo a aportar a Mark A,
tB = fracción del fondo a aportar a Mark B

El enunciado del problema nos obliga a encontrar valores para *tA* y para *tB* que nos retribuyan una ganancia máxima y que, simultáneamente, cumplan con todos los requisitos del problema.

Esos valores serán buscados para cumplir el ***objetivo*** que es "**maximizar la ganancia esperada satisfaciendo las *restricciones* impuestas**".

Como ya tenemos claro nuestro objetivo, lo que sigue es plantear ese objetivo como un modelo matemático.

Se espera que cada unidad destinada a Mark A tenga un rendimiento de 10 centavos y de 6 centavos en Mark B. Si llamamos ***tA*** a la fracción del fondo que vamos a destinar al Mark A, entonces la ganancia será 0,10*tA*.

De esta manera, la expresión matemática que refleja el objetivo de este modelo deberá tener en cuenta la contribución de cada una de las posibilidades: *0,10 tA* y *0,06 tB* (donde ***tB*** será la fracción del dinero aportado a Mark **B**). Con estos elementos se podrá expresar el objetivo global como una ***función*** que se puede expresar así

*Maximizar 0,10 t*A + *0,06 t*B

Como vemos, la función consta de dos partes:
a) la intención (maximizar) que se denomina **criterio** y que es el objetivo del modelo: maximizar las ganancias.
b) la función operativa – que se denomina **función objetivo** – que tiene coeficientes asociados a variables de decisión, lo cual muestra qué manera o la política empleada para lograr el objetivo y satisfacer el criterio.

Restricciones

Pero con esto solo no nos alcanza para definir el problema como un modelo matemático: existen las *limitaciones* o *restricciones*, que debemos incorporar utilizando las mismas variables de decisión en forma coherente y exhaustiva:

Repasemos, una por una, las restricciones que tenemos. La primera restricción es: Ningún aporte de fondos puede superar el 75% y la podemos escribir de esta manera:

$$tA \leq 0{,}75$$
$$tB \leq 0{,}75$$

De esta manera, vemos que ambas desigualdades representan los límites superiores de inversión en Mark A y Mark B, respectivamente.

b) La segunda es: El aporte para Mark A (mayor ganancia) no debe superar el doble del que destinamos a Mark B, y la representamos así:

$$tA \leq 2tB$$

aunque hay otras maneras más cómodas, por ejemplo esta:

$$tA - 2tB \leq 0$$

c) La tercera es: No pueden haber aportes negativos, en contraposición con la primera restricción (donde establecíamos los límites superiores) ahora definimos el límite inferior de aportes para cada alternativa, aclarando, además, que está permitido el cero, que significaría que la decisión sería no hacer aportes, lo que es perfectamente válido.

$$tA \geq 0$$
$$tB \geq 0$$

d) Por último tenemos que la fracción de dinero a aportar debe ser coherente. Tanto esta restricción como la anterior no eran explícitas en la presentación del caso, pero, si nos olvidáramos cometeríamos un error importante y al usar el modelo nos podría dar resultados incoherentes como, por ejemplo, aconsejarnos aportar el 75% del fondo en Mark A y el 75% en Mark B, que es algo imposible de hacer. Estas restricciones, como números no negativos, proporciones imposibles, u otras de este tipo, suelen denominarse "*restricciones lógicas*".

En este caso es así:

$$tA + tB = 1$$

El modelo, ahora completo, se escribe uniendo todas estas partes:

$$\begin{aligned}
& Z = 0{,}10tA + 0{,}06tB \equiv Maximo \\
& \text{sujeto a :} \\
& \qquad tA \qquad\quad \leq 0{,}75 \\
& \qquad\qquad\quad tB \leq 0{,}75 \\
& \qquad tA - 2{,}00tB \leq 0 \\
& \qquad tA + \quad tB = 1 \\
& \qquad tA \qquad\quad \geq 0 \\
& \qquad\qquad\quad tB \geq 0
\end{aligned}$$

En este caso tenemos un modelo compuesto por una función objetivo y seis restricciones. Tanto la función objetivo como las restricciones están presentadas en función de variables de decisión (cada una de ellas representa el aporte de capital a uno u otro modelo) y de informaciones conocidas (**datos**): cuánto rinden cada uno de los equipos y cuánto como máximo y como mínimo se permite invertir en cada uno de ellos: estas son variables y/o parámetros sin control.

En la práctica es difícil encontrar un problema como éste. Generalmente hay que estimar los datos. Y el modelo va a ser tan bueno como la estimación de los datos, nunca mejor.

Resolución del modelo matemático.

En este caso, resolver el modelo es hallar los valores para *tA* y para *tB* que hagan máxima la expresión de la función objetivo y que, en forma simultánea, cumplan con todas y con cada una de las restricciones. Para resolver estos modelos, es decir para encontrar valores satisfactorios para las variables de decisión, hay dos tipos de técnicas:

a) **Optimización**: se encuentran los mejores valores para las variables, que hagan máximo el objetivo y satisfagan simultáneamente todas las limitaciones-

b) **Heurísticas:** Producen valores aceptables para la función objetivo y satisfacen las restricciones.

Usando cualquiera de esas técnicas podríamos encontrar alguna solución aceptable. Aceptable significa que el valor hallado para cada variable es válido, cumple, con todas y cada una de las restricciones. Una solución aceptable para este problema es:

$$tA = 0{,}6667$$
$$tB = 0{,}3333$$

Cuyo significado es que por cada unidad monetaria de capital aportado habrá una ganancia de

$$0{,}6667 \times 0{,}10 + 0{,}3333 \times 0{,}06 = 0{,}08667$$

ya que se destinan 2/3 al Mark A y 1/3 al Mark B.

Comprobación de la validez de la solución hallada.

Puede ocurrir que el modelo no esté del todo bien planteado, por ejemplo, porque no hemos previsto todas las restricciones, o que omitiéramos o simplificáramos algún parámetro importante, que los datos estuvieran mal estimados o cargados. En todos los casos es conveniente comprobar si la solución hallada es válida. Para eso existen varios métodos, aunque a veces basta con un análisis preliminar simple.

Si el modelo no funciona hay que revisarlo, incluir nuevas restricciones, eliminar restricciones o replantear con más exactitud las limitaciones reales. A veces es necesario usar la experiencia para adaptar las soluciones dadas por el modelo. Otras veces, aun estando todo bien y con resultados válidos, las soluciones son impracticables. Es difícil, por ejemplo, incluir en un modelo datos que tienen que ver con conductas. Otra fuente de disturbios son los datos que cambian con el transcurso del tiempo: las acciones pueden caer en el futuro, no dejando de ser válido el dato de su evolución anterior.

Usos de los modelos

Los modelos ayudan a tomar dos tipos de decisiones:

a) **Estratégicas**: son decisiones de única vez, afectan intervalos de tiempo largos.
Son del tipo:

> ¿Se debe abrir una nueva línea de producción?
> ¿Debería hacer inventarios en período regular o solo cuando caigan los niveles?

Este tipo de decisiones implican impactos grandes, por lo cual se justifica invertir esfuerzo y tiempo en sus modelos y resoluciones.

b) **Operacionales**: afectan procesos en curso y períodos más cortos de tiempo.
Son del tipo:

> ¿Cómo reprogramar el trabajo de la semana cada semana?
> Sobre los tres productos de la fábrica, en estas condiciones de mercado y con estos recursos de materias primas, ¿cuál es la producción a realizar este mes de cada uno de ellos?

Estos modelos se usan repetidamente, por lo cual también se justifica invertir en ellos.

Las ventajas de los modelos, en general, son:

1. Lograr objetivos con recursos escasos
2. Evaluar el impacto y costo de un cambio sin necesidad de gastar en producirlo previamente.
3. Se puede evaluar la fortaleza de los resultados propuestos mediante el *análisis de sensibilidad* o análisis de tipo ¿qué sucedería si...?. Por ejemplo ¿qué pasaría si las ventas del Mark A diera una ganancia del 8% en vez del 10%?

Técnicas para construir modelos matemáticos DETERMINÍSTICOS

Veremos ahora los pasos a seguir y algunas técnicas a emplear para formular modelos *determinísticos*. Debemos tener presente que no pretendemos decir que hay un *método único*. Simplemente es una propuesta metodológica básica, aplicada a un ejemplo simple. Cada uno, conforme gana experiencia, emplea métodos perfeccionados o adecuados a lo que busca.

Para mostrar paso a paso el desarrollo de un modelo usaremos este caso que está en el Libro 1 de esta colección:

> **Planeamiento de la producción en Alcoholes Argentinos (*ALAR*).**
>
> ALAR produce dos marcas de alcohol, “*Alas*” (Catálogo AA01) y “*Biguá*” (Catálogo AA02), en una planta chica recién adquirida, que funciona en San Nicolás.
>
> El sector de **producción** opera 40 horas semanales empleando a cinco trabajadores de tiempo completo y a dos que trabajan quince horas semanales.

Una vez terminado el producto, éste pasa al sector de **rectificado**, que tiene equipos operados por seis empleados de tiempo completo y uno de 10 horas semanales.

ALAR no tiene problemas de materias primas para ambos productos. Puede vender todo lo que quiera de "*Alas*" pero tiene una demanda limitada a 120.000 litros semanales de "*Biguá*".

¿Qué cantidad de cada producto debe producir ALAR para maximizar las ganancias si se sabe que "*Alas*" deja una ganancia neta de 3 $ por mil litros y que "*Biguá*" de 5 $/1.000 litros?

Nos proponemos formular el problema en forma de modelo matemático siguiendo cuatro etapas. Con este desarrollo queremos disponer de una metodología apta para aplicar luego a cualquier caso.

Primera etapa: Identificación de las variables de decisión

Tanto el personal de producción como el de rectificación deben saber al principio de la semana cual es el plan de producción, por tanto hay que informarles lo que deben hacer ¿Qué información necesitan? A ellos hay que informarles cuántos kilolitros/semana deben producir de "Alas" y de "Biguá". Esas serán, entonces, las variables de decisión, porque tendremos que decidir esos valores, o sea, dar valor a las variables.

En primer lugar daremos un nombre simbólico a las variables de decisión, que sea fácil de identificar, por ejemplo, *P* por Producción. Así sería:

P_1: producción de kilolitros semanales de "Alas"
P_2: producción de kilolitros semanales de "*Biguá*".

Definimos estas variables con precisión, sin dejar ambigüedades: P_1 será la producción semanal de "Alas" expresada en mil litros/semana [kl/sem] y con la misma unidad para "Biguá"

Para saber cómo identificar las variables de decisión deberemos preguntarnos y ser capaces de identificar:

a) los elementos que afectan los costos, las ganancias o lo que represente el objetivo global del problema.

b) los elementos que podemos elegir y/o controlar con libertad.

c) las decisiones que tenemos que tomar

d) la información que deberá disponer quien quiere llevar adelante la solución propuesta cuando le demos el informe resultante de haber resuelto el modelo.

Segunda etapa: Identificación de datos del problema

Para determinar las cantidades reales a producir con el objetivo de maximizar las ganancias necesitamos saber:

- Número de horas-hombre de trabajo disponibles en el sector de elaboración, o, dicho en otros términos, la capacidad de producción semanal de ese sector.
- Número de horas-hombre disponibles en el sector de rectificación, o sea, la capacidad de producción semanal del sector.
- Ganancias por la venta de cada uno de los productos.

Como se trata de un problema determinístico, necesitamos conocer estos datos o tener acceso concreto a ellos para poder desarrollar el modelo. Si, por el contrario, se tratara de un problema estocástico, en lugar de datos concretos usaríamos probabilidades o estimaciones.

Supongamos que cuando averiguamos o nos informan como funciona la planta. Dicho en otras palabras accedemos a los siguientes datos:

Producción:	5 hombres x 40 hs/hombre semana =	200 hs/hombre semana
	2 hombres x 15 hs/hombre semana =	30 hs/hombre semana
	Total en producción =	**230 hs/hombre semana**
Rectificado:	6 hombres x 40 hs/hombre semana =	240 hs/hombre semana
	1 hombre x 10 hs/hombre semana =	10 hs/hombre semana
	Total en rectificado =	**250 hs/hombre semana**

- El margen de ganancia es de 3 $/kl "Alas" y de 5 $/kl de "Biguá".

Como vemos, la diferencia entre las variables y los datos estiba en que el operador no puede controlar los valores de los datos. No se podría (al menos no en este enfoque) cambiar la capacidad de trabajo de un hombre o de una máquina, pero sí se podría afectar a más o menos hombres o máquinas. Deberíamos prestar atención a este detalle. También ocurre que, cuando se avanza en el planteo del problema, puede ser necesario ampliar datos, obtener valores más completos u otros que no parecieran ser necesarios al principio e, incluso, descartar información que no es pertinente.

Tercera etapa: Formulación de la función objetivo.

Esta etapa nos permite el planteo del problema en forma matemática. Recordamos que cuando queríamos saber cuáles son las variables de decisión preguntábamos "¿qué hay que informar al jefe de producción?", ahora, entonces, para poder desarrollar la función objetivo deberemos hacernos la siguiente pregunta:

"¿Qué esperamos obtener con este modelo?"

La respuesta será la información que daremos a – por ejemplo – quien nos encargó el análisis, supongamos el dueño de la fábrica, y esa respuesta se relaciona con el objetivo global, la decisión estratégica, no los detalles operacionales.

Notamos que con las respuestas a las preguntas anteriores generamos dos reportes: con la primera, destinada al nivel operativo, les informamos **cuánto hay que fabricar de cada producto por semana** (propuesta de plan de producción); la segunda, dirigida al nivel superior, informa **cuál será la ganancia** que se obtiene con el plan de producción propuesto, lo que es una respuesta estratégica.

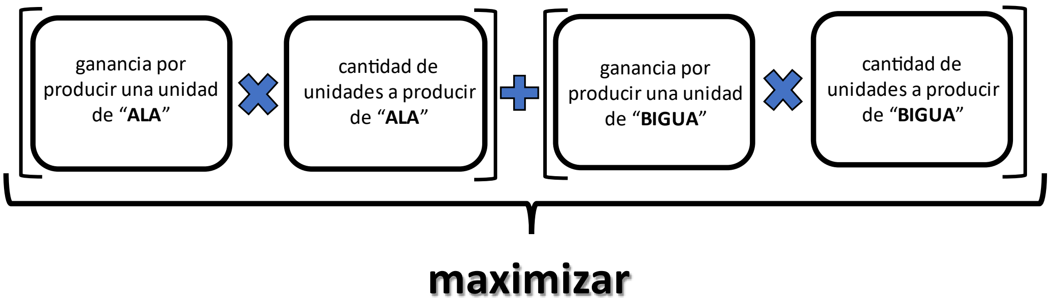
Para formular el objetivo podemos usar el siguiente método, aunque, por supuesto, hay varias maneras de hacerlo:

1) Establecemos el objetivo en forma verbal, que en este caso podría ser algo así como "*Queremos, buscamos, perseguimos el objetivo de…*"

Maximizar la ganancia semanal del total de la producción de "Alas" y de "Biguá", lo que lograremos determinando las cantidades a producir de cada uno de los alcoholes.

2) Si es posible, descomponemos el objetivo en operaciones aritméticas básicas (suma, resta o producto) de cantidades individuales:

Maximizar
=
(Ganancia por producir 1 unidad de "Alas") x (cantidad de unidades a producir de "Alas")
+
(Ganancia por producir 1 unidad de "Biguá") x (cantidad de unidades a producir de "Biguá")



3) Expresamos las cantidades individuales usando las variables, unidades y datos del problema.

Para poder concretar la tercera etapa puede ser útil elegir algunos valores supuestos y así poder saber cómo usar la función objetivo. Esta técnica puede ser denominada **trabajo con un ejemplo específico** (o **especificación del modelo**). Se elige un valor cualquiera – razonable – para comenzar:

En este caso, podemos *suponer* una producción semanal de 10 mil litros de "Alas" y de 20 mil de "Biguá":

Ganancia de "Alas" x Producción semanal de "Alas" = 3 $/kl x 10 kl/sem = 30 $/sem
Ganancia de "Biguá" x Producción semanal de "Biguá" = 5 $/kl x 20 kl/sem =100 $/sem
Ganancia total 130 $/sem

Este análisis no resuelve el problema, pero sirve para mostrar como plantearlo. Ahora podemos transformarlo fácilmente, usando la definición que hemos hecho antes de las variables de decisión P_1 y P_2:

Ganancia de "Alas" x Producción semanal de "Alas" →	$3\ P_1$
Ganancia de "Biguá" x Producción semanal de "Biguá" →	$5\ P_2$
Ganancia total	$= 3\ P_1 + 5\ P_2$

Por tanto, ya podemos ver la función objetivo que será

$$\boxed{\text{Maximizar } 3\ P_1 + 5\ P_2}$$

Cuarta etapa: Identificación de las restricciones.
Es evidente que los problemas de este tipo no tienen solución si no se plantean las restricciones o límites, porque, si no existen límites siempre podemos encontrar un valor de P_1 y/o de P_2 que supere a los valores previos y por lo tanto den más ganancias. Estos límites o restricciones surgen, en general, de:

- Limitaciones físicas. (horas de trabajo, capacidad de producción, cantidad de materia prima, etc.)
- Limitaciones de ventas (por ejemplo compromisos con un cliente, políticas de ventas, preferencias de mercado, competencia, pedidos en firme, etc.)
- Limitaciones externas (por ejemplo inestabilidad económica, oportunidades, huelgas, enfermedades, etc.)
- Relaciones entre variables (por ejemplo la imposibilidad que la suma de variables sea superior a un número fijo, lo que vale en porcentajes, tantos por uno o partes de una unidad productiva)
- Restricciones lógicas (por ejemplo no puede producirse litros negativos de alcohol o no pueden fabricarse 6,35 pianos)

Volviendo al caso de la producción de alcoholes:

1. Limitación física: límite de horas/hombre semanales disponibles para producción y rectificado

En forma verbal: las horas totales semanales en elaboración no pueden superar las 230.

Descomposición: las horas usadas para producir *"Alas"* + las horas usadas para producir *"Biguá"* no pueden superar las 230

Matemática: Al llegar a este punto nos damos cuenta de que nos faltan datos, tales como cuánto tiempo se emplea en la elaboración de cada producto. Lo que notamos es que, a diferencia de los "problemas" que se solucionan en clase (donde el enunciado incluye los datos) en los casos de la vida real, comenzamos con el método de solución para recién después averiguar cuáles son los datos que necesitamos.

Supongamos, entonces, que en la planta de producción, recolectamos e incorporamos en una planilla los siguientes datos:

Horas/hombre por 1000 litros producidos

	“Alas” AA01	“Biguá” AA02
Sector Producción	2	1
Sector Rectificado	1	2

2. Con estos datos ya podremos escribir la restricción en forma genérica, para cualquier valor de las variables establecidas al principio:

horas hombre para producir 1 unidad de “Alas” x unidades producidas de “Alas”

+

horas hombres para producir 1 unidad de “Biguá” x unidades de “Biguá” producidas

=

horas hombres usadas en el sector.

3. Limitación física 1: límite de horas/hombre disponibles en el sector Elaboración, el tiempo necesario para producir una unidad del *Alas* multiplicado por las unidades de *Alas* producidas, más el tiempo que lleva producir una unidad de *Biguá* por la unidades de *Biguá* producidas no debe superar el tiempo total disponible en el sector. Numéricamente lo expresamos así:

$$2\,P_1 + 1P_2 \leq 230$$

dimensionalmente

[hh/kl] x [kl/sem] = [hh/sem])

Se lee: *gastando a razón de 2 Hh en cada kl elaborado de AA01 multiplicando por la cantidad de kl/sem de AA01 que se va a elaborar más 1 Hh/kl de AA02 por la cantidad de kl/sem de AA02 a elaborar no se deben superar las 230 Hh/sem disponibles en el sector*

4. Limitación física 2: límite de horas/hombre disponibles para rectificación

$$1\,P_1 + 2P_2 \leq 250$$

5. Limitación externa: restricción de límite de producción.
No pueden venderse más de 120 mil litros/semana de “Biguá”:

$$P_2 \leq 120$$

6. Limitaciones lógicas: restricción de no negatividad

$$P_1 \geq 0$$
$$P_2 \geq 0$$

Formulación matemática completa del problema:

$$\boldsymbol{Z = 3P_1 + 5P_2 \equiv Maximo}$$ **Objetivo**: máxima ganancia

sujeto a:

$$\boldsymbol{2P_1 + 1P_2 \leq 230}$$ **Restric. 1**: disponibilidad Hh/sem Producción

$$\boldsymbol{1P_1 + 2P_2 \leq 250}$$ **Restric. 2**: disponibilidad Hh/sem Rectificado

$$\boldsymbol{0P_1 + 1P_2 \leq 120}$$ **Restric. 3**: demanda max de *Biguá*

$$\boldsymbol{1P_1 + 0P_2 \geq 0}$$ **Restric. 4**: no negatividad Prod. *Alas*

$$\boldsymbol{1P_1 + 0P_2 \geq 0}$$ **Restric. 5**: no negatividad Prod. *Biguá*

Para verificar la validez del modelo hallado podemos hacer las siguientes pruebas de coherencia:

1) ¿es posible decidir no fabricar nada? Sería lo mismo que los valores de las variables sean nulos, $P_1 = 0$, $P_2 = 0$.

   Si, es posible, porque se pueden escribir valores nulos en todo el modelo y se verifican bien las desigualdades y restricciones, aunque en ese caso será Z = 0 (difícil que sea un máximo) pero todas las condiciones se satisfacen simultáneamente.

2) ¿es posible fabricar $P_2 = 120$ kl/sem y no fabricar P_1?

   Sí, es posible. Sería emplear el criterio de fabricar todo lo que se puede del alcohol AA02 que tiene la máxima ganancia. Sería de esta manera:
   $P_1 = 0$; $P_2 = 120$. Z = 600,
   $120 < 230$ (Restricción 1)
   $240 < 250$ (Restricción 2)
   $120 = 120$ (Restricción 3)
   Lo que nos muestra que todos los requisitos del modelo quedan satisfechos.

3) ¿es posible fabricar $P_1 = 200$ y $P_2 = 300$?.

   No es posible, si bien da un Z alto (2100), no se satisfacen las inecuaciones o restricciones.

Por tanto el modelo construido hasta acá parece lógico y funcional

Problemas de redes

Con el objetivo de mostrar otros casos en los cuales construimos modelos similares, vamos a plantear en este apartado y en el que sigue ejemplos de problemas que abordan relaciones entre grupos de nodos.

En el primer caso, veremos nodos que vamos a agrupar bajo la denominación convencional de “plantas” que se relacionan mediante vínculos con nodos agrupados bajo la denominación “centros de distribución”.

Este tipo de problemas se conoce con el nombre de “Problemas de Transporte”, debido a que de esta manera es más fácil comprenderlos. Sin embargo, la aplicación de estos modelos no se restringe solo al transporte, sino que alcanza a varias otras situaciones.

Para poder avanzar debemos conocer algunos datos, como, por ejemplo:

1. demandas de cada nodo destino (centro de distribución, en este caso)
2. capacidad de elaboración de cada nodo de origen (planta)
3. costo de transporte de cada planta a cada destino.

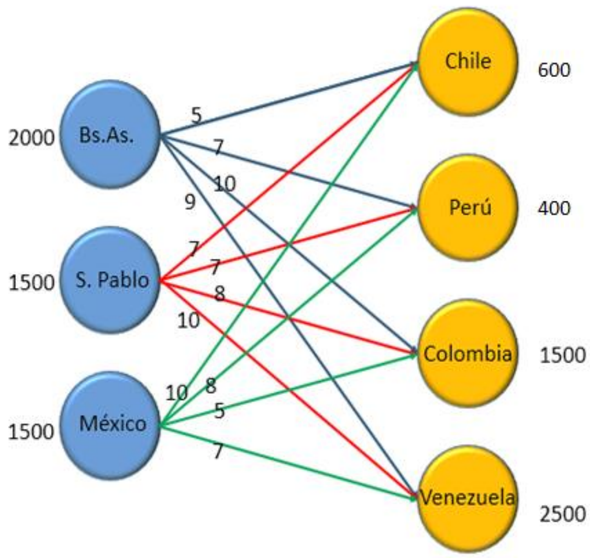
CASO:

Una empresa automotriz tiene plantas en Buenos Aires, San Pablo y Ciudad México en las cuales produce un modelo que se comercializa, además de localmente, en Chile, Perú, Colombia y Venezuela. La Planta Buenos Aires produce, para exportación, 2000 unidades/mes *y las otras dos 1500 unidades/mes cada una de las dos plantas restantes. Con esos vehículos hay que satisfacer las demandas de Chile que es de 600, de Perú que necesita 400, de Colombia que son 1500 y de Venezuela de 2500 unidades/mes.*

La tabla que sigue nos muestra los costos de embarque de cada unidad desde cada planta a cada centro

PLANTAS	DISTRIBUIDORES			
	CHILE	PERU	COLOMBIA	VENEZUELA
BUENOS AIRES	5	7	10	9
SAN PABLO	7	7	8	10
MEXICO DF	10	8	5	7

En este tipo de problemas es útil hacer un diagrama en redes con nodos y arcos antes de abordar la expresión matemática. Cada elemento de salida o llegada (planta o centro de distribución) es un **nodo**, y se representa con un círculo. Cada línea indica la posibilidad de llevar un vehículo entre esos dos nodos. Junto a cada nodo se coloca la oferta o demanda, según sea emisor o receptor y en cada línea (llamada arco o ruta de transporte) indicamos el costo de transporte.



Identificación de las variables de decisión

Nuevamente, hay que identificar las variables y lo hacemos teniendo en cuenta:
a) los elementos que afectan los costos, las ganancias o lo que represente el objetivo global del problema.
b) los elementos que puede elegir y/o controlar con libertad.
c) las decisiones que hay que tomar
d) la información que deberá disponer quien quiere llevar adelante la solución propuesta cuando se resuelva el problema.

Al analizar los datos de este problema vemos que hay doce variables de decisión: cada una de ellas es la cantidad (número) de autos que hay que transportar desde cada centro a cada distribuidor. Efectivamente hay doce rutas posibles fábrica—distribuidor. Estas variables las podemos identificar de cualquier manera: x_1, x_2, ... o bien x_{11}, x_{12}, x_{13}, x_{14}, x_{21},.... o bien más explícitamente x_{BA-CH} , x_{BA-PE}....

Además, como en el caso anterior, podemos expresar la función objetivo y las variables que ésta contiene en forma literal y descomponiendo la expresión en detalle:

<u>En forma verbal y generalizada</u>: "*Minimizar los costos de transporte desde*… [todas las plantas] *a*… [todos los distribuidores]".

Descomposición: *"Minimizar costo de transporte desde **Buenos Aires** hasta Chile, hasta Perú, hasta Colombia y hasta Venezuela, más el costo de transporte desde **San Pablo** hasta Chile, hasta Perú, hasta Colombia y hasta Venezuela, más el costo de transporte desde **México** hasta Chile, hasta Perú, hasta Colombia y hasta Venezuela"*.

Entendiendo *costo de transporte* como el costo de transportar una unidad entre un origen y un destino multiplicado por el número de unidades transportadas entre esos extremos.

Si usamos este ejemplo específico y pasamos a una expresión matemática, nos queda:

$$\text{Minimizar} \quad \begin{aligned} &(5x_{BA/CH} + 7x_{BA/PE} + 10x_{BA/CO} + 9x_{BA/VE}) + \\ &(7x_{SP/CH} + 7x_{SP/PE} + 8x_{SP/CO} + 10x_{SP/VE}) + \\ &(10x_{ME/CH} + 8x_{ME/PE} + 5x_{ME/CO} + 7x_{ME/VE}) \end{aligned}$$

Identificación de restricciones.

Para incorporar las restricciones vamos a averiguar qué elementos son los que nos impiden asignar a las variables cualquier valor. Por ejemplo, para minimizar costos de transporte, lo ideal es no transportar nada en ninguna ruta. ¿Se puede elegir el cero? Si, podemos elegir no transportar nada de, por ejemplo, Buenos Aires a Perú, pero debemos asegurarnos de que todos los autos que hay en Buenos Aires vayan por lo menos a un destino y no quede ninguno

Las restricciones surgen de considerar:
a) el embarque total, con rumbo a todos los destinos, de cada planta no debe exceder ni ser inferior a la cantidad que la planta dispone para enviar
b) el recibo de cada centro, suma de lo que llega de cada fábrica, debe ser lo que necesita para satisfacer la que ese destino necesita. Pero debemos asegurarnos que no se puede enviar más autos que los que pidieron. Parece que este no es el caso, pero no puede haber ambigüedades.
c) el envío debe ser de números positivos enteros, (no es posible enviar, por ejemplo, 0,345 vehículos).

Todas estas restricciones deberán ser convertidas a expresiones matemáticas, por eso vamos a tener en cuenta que en cada nodo de salida no se puede superar la capacidad de producción de la planta: desde Buenos Aires., por ejemplo, el número total de unidades despachadas será igual a la suma de las unidades despachadas a Chile, a Perú, a Colombia y a Venezuela. Pero hay una restricción: esa suma no puede ser superior a 2000, aunque, por la manera en que se plantea el problema tampoco podría ser inferior ya que no se encuentra en el planteo frases como "*el sobrante se destina a…*" o "*se enviarán como máximo*"

La expresión matemática de esto sería diferente en un caso u otro,

$$x_{ME/CH} + x_{ME/PE} + x_{ME/CO} + x_{ME/VE} \leq 2000$$

en el caso de que se admitan sobrantes (*se enviarán como máximo…*)

$$x_{ME/CH} + x_{ME/PE} + x_{ME/CO} + x_{ME/VE} = 2000$$

en el caso que tengamos que embarcar todo lo producido

Lo mismo para cada una de las otras dos terminales. Luego tendremos las restricciones de demanda: cada distribuidor recibirá lo que necesita (lo que solicitó en firme): Las unidades recibidas en Chile serán la suma de las despachadas desde Buenos Aires más las de San Pablo más las de México, y debe ser igual a su demanda (1000) y valen las mismas consideraciones que hicimos con las de oferta.

$$X_{BA/CH} + x_{JSP/CH} + x_{ME/CH} = 1000$$

y así para los otros tres distribuidores.

Finalmente incorporaremos las restricciones de no negatividad y de enteros. El resultado final será:

Minimizar $(10x_{ME/CH} + 8x_{ME/PE} + 5x_{ME/CO} + 7x_{ME/VE}) +$
$(7x_{SP/CH} + 7x_{SP/PE} + 8x_{SP/CO} + 10x_{SP/VE}) +$
$(5x_{BA/CH} + 7x_{BA/PE} + 10x_{BA/CO} + 9x_{BA/VE})$

Sujeto a:

$x_{ME/CH} + x_{ME/PE} + x_{ME/CO} + x_{ME/VE} = 2000$	Capacidad Buenos Aires.
$x_{SP/CH} + x_{SP/PE} + x_{SP/CO} + x_{SP/VE} = 1500$	Capacidad San Pablo
$x_{BA/CH} + x_{BA/PE} + x_{BA/CO} + x_{BA/VE} = 1500$	Capacidad México
$x_{ME/CH} + x_{SP/CH} + x_{BA/CH} = 1000$	Demanda Chile
$x_{ME/PE} + x_{SP/PE} + x_{BA/PE} = 500$	Demanda Perú
$x_{ME/CO} + x_{SP/CO} + x_{BA/CO} = 1500$	Demanda Colombia
$x_{ME/VE} + x_{SP/VE} + x_{BA/VE} = 1200$	Demanda Venezuela
$x_{ij} \geq 0$ y entero para todo i j	

Otro tipo de problemas de redes será aquel que corresponda a tomar decisiones con alternativas excluyentes. En el ejercicio de la ingeniería muchas veces nos encontramos con casos de toma de decisión que implican estrategias o acciones binarias tales como ***hacemos esto*** o ***no lo hacemos***: Por ejemplo, ampliación de una línea de producción, realización de un proyecto, renovación de un equipo por un modelo más rendidor, cambio de un proveedor, presentar una segunda marca, etc.

Generalmente estas decisiones se denominan "**binarias**" porque el valor que pueden asumir las variables de decisión está restringido a solamente dos estados:

"proceder" (hacer, fabricar, comprar, ejecutar…)	"no proceder" (no hacer, no fabricar, no comprar, no ejecutar…)

generalmente, a los únicos valores posibles de las variables de decisión se les asigna los números ***0 y 1***, y se trata de una variable entera que solamente puede asumir esos valores.

*El grupo de inversores "AERO Capital" (**AC**) recibe cuatro propuestas de inversión las cuales fueron analizadas y por ello se sabe que tienen un rendimiento alto con un riesgo comparable entre ellas y aceptable. También hay una buena estimación del retorno esperado en un horizonte de 4 años. Los proyectos requieren de un programa de inversiones pautado en esos cuatro años. **AC** dispone de fondos a invertir y de un cronograma de desembolsos para cubrir esos cuatro años de los proyectos. Por otro lado, la inversión en telecomunicaciones y en electrónica comprende algunos aspectos tecnológicos que se superponen, por lo tanto se ha decidido invertir en uno u otro rubro pero no en ambos simultáneamente.*

Los fondos que no se usen un año no están disponibles en el año siguiente porque se destinan a otros fines.

Todos los datos mencionados se sintetizan en la siguiente tabla:

PROYECTO	AÑO 1	AÑO 2	AÑO 3	AÑO 4	RETORNO
FARMACÉUTICO	60	10	10	10	250
TELECOMUNICACIONES.	35	35	35	35	375
ELECTRÓNICA	10	50	50	10	275
SUPERMERCADOS	15	10	10	40	140
FONDOS DISPONIBLES	90	80	80	50	

Identificación de las variables de decisión
La pregunta es *¿que podemos controlar libremente en este problema?*, y la respuesta es: puede aceptarse o rechazarse la posibilidad de invertir en cada una de las propuestas.

Por tanto generamos variables de decisión, que serán las siguientes
$F \rightarrow$ inversión en farmacia. Los valores posibles de F serán: F= 1, invertir; F = 0 no invertir
$T \rightarrow$ inversión en telecomunicaciones, T = 1, invertir; T = 0, no invertir
$E \rightarrow$ inversión en electrónica, E = 1, invertir; E = 0, no invertir
S $\rightarrow$ inversión en supermercados, S = 1, invertir; S = 0, no invertir

Estas variables de decisión son diferentes a las anteriores porque el valor que pueden asumir ahora está restringido fuertemente a dos números enteros positivos, mientras que antes eran positivos (miles de litros de alcohol) o positivos enteros (autos a transportar).

Identificación de la función objetivo
Como el rendimiento total será la suma de los rendimientos en cada una de las cuatro compañías, nos haremos la siguiente pregunta: ¿Cuánto rendirá lo invertido en — por ejemplo — farmacia? La respuesta es que el rendimiento si invertimos en farmacia será el retorno esperado y si NO invertimos el rendimiento será cero. Es lo mismo que multiplicar el retorno esperado por uno (decisión: invertimos) o multiplicar por cero (decisión: no invertimos). En forma numérica sería así

Rendimiento en farmacia = 250 F
(si la decisión, F, vale 1 el retorno es 250 x 1, caso contrario es 250 x 0)

Entonces el rendimiento total, para todas las posibilidades, será

Rendimiento total = 250 F + 375 T + 275 E + 140 C

Identificación de las restricciones
En este caso tendremos que trabajar con las siguientes limitaciones:

1) Los fondos disponibles para invertir en cada año.

2) La imposibilidad de invertir simultáneamente en electrónica y comunicaciones.
3) Las limitaciones lógicas, como por ejemplo, asegurarnos que no invertiremos fondos negativos.

La limitación de fondos disponibles por año impide seleccionar los cuatro proyectos a la vez. Como la cifra es variable, se necesita una restricción de este tipo por año:

Los fondos para invertir en todos los proyectos seleccionados en el primer año deberán ser, como máximo, de 90

Y, por otro lado, cada proyecto aporta acumulativamente al total:

Los fondos totales para invertir en el primer año en todos los proyectos **=**
invertido en farmacia
+
invertido en electrónica
+
invertido en telecomunicaciones
+
invertido en supermercado.

Si esto lo escribimos con símbolos matemáticos para el primer año, podemos obtener una expresión donde la suma de lo que se invierte en cada proyecto multiplicado por la variable de decisión (cuyo valor será 0 o 1) dará la inversión total para ese año, la cual no puede superar el capital disponible de ese mismo año:

$$60F + 35T + 10E + 15S \leq 90 \quad \text{(restricción del año 1)}$$

con los datos de la tabla podemos completar las restricciones correspondientes a los demás años:

$$10F + 35T + 50E + 10S \leq 80 \quad \text{(año 2)}$$
$$10F + 35T + 50E + 10S \leq 80 \quad \text{(año 3)}$$
$$10F + 35T + 10E + 40S \leq 50 \quad \text{(año 4)}$$

2) Ahora trabajamos con la restricción de imposibilidad de invertir simultáneamente en electrónica y en telecomunicaciones, que es sencilla y la podemos escribir de dos maneras.

A) Invertir en electrónica (0 o 1) multiplicado por invertir en telecomunicaciones (0 o 1) debe ser igual a cero o

B) Invertir en E + invertir en T no pude superar 1. Ambas son equivalentes, y se escriben así:

A) $T\,E = 0$
B) $T + E \leq 1$

Preferimos la variante B, porque tiene el formato de las demás restricciones (sumas del lado izquierdo y desigualdad)

3) Restricciones lógicas. Similares a los casos anteriores, pero ahora hay una restricción lógica muy fuerte: los valores que pueden asumir las variables de decisión está limitados a 0 y 1:

F, T, E y S son 0 o 1 y enteros.

Planteo del modelo completo:

Maximizar:					
Rendimiento total	250 *F*	375 *T*	275 *E*	140 S	
Sujeto a:					
Restr. Año 1	60 *F* +	35 *T* +	10 *E* +	15 *S*	≤ 90
Restr. Año 2	10 *F* +	35 *T* +	50 *E* +	10 *S*	≤ 80
Restr. Año 3	10 *F* +	35 *T* +	50 *E* +	10 *S*	≤ 80
Restr. Año 4	10 *F* +	35 *T* +	10 *E* +	40 *S*	≤ 50
Restr. T E no simultaneas		*T* +	*E* +		≤ 1
Restr. lógicas	*[F, T, E y S]* = 0 o 1				

Modelos lineales no determinísticos con datos observados.

En los modelos que vimos hasta acá, teníamos la función objetivo y las condiciones de contorno (datos operativos), nuestra tarea consistió en averiguar los valores que asumen las variables en la función objetivo. En otras palabras, sabíamos de antemano el comportamiento y queríamos optimizarlo. Pero en la realidad, cuando necesitamos hacer un modelo, no siempre es en estas condiciones que estuvimos viendo.

Hay muchas chances de que los modelos deban elaborarse a partir de datos experimentales, o sea, de valores de las variables obtenidos por la experiencia y no buscados por nosotros como en los anteriores casos. Cuando esto ocurre se dice que tenemos datos experimentales que debemos ajustar a un modelo y, una vez ajustado, tener una función matemática descriptiva del proceso, que eventualmente, podemos optimizar. Como podemos darnos cuenta es casi un camino inverso.

Ahora, para abordar estos casos nuevos, vamos a necesitar forzosamente el aporte de herramientas estadísticas, cualquiera sea el origen de los datos que vamos a usar (planillas de operación, ensayos experimentales factoriales o rotables, resultados de laboratorio o de planta piloto, etc.).

Veamos primero como elaborar un modelo de este tipo, comenzando por el más simple. Se trata de ajustar los datos que tenemos, o que obtuvimos por experimentación, a un modelo lineal.

$$z = b_0 + b_1 x$$

nuestra tarea es determinar los valores de b_0 y de b_1 Lo podemos hacer a partir de dos ensayos, uno en un nivel[35] y el otro en otro nivel para la variable $\boldsymbol{x}$. Sin embargo, no nos alcanzaría con un ensayo en cada nivel, sino que debemos repetirlos para poder discriminar el error experimental.

Comenzaremos diciendo que el modelo general de arriba también puede ser expresado de una manera equivalente, como la siguiente

$$z = \mu + \alpha_i + e_{ij}$$

donde μ es el valor medio del ensayo,
α es la incidencia del factor utilizado
e es el error experimental
siendo i = 1, 2, ..., m (réplicas o repeticiones)
j = 1, 2, ..., n (niveles de las variables[35])

Si llamamos γ_{ij} al resultado de una corrida en el nivel *j* durante la repetición *i*

y llamando γ_{**} al resultado promedio de todos los ensayos *n.m*

y γ_{i*} al resultado promedio de todos los m ensayos realizados en el nivel *j*

tendremos una tercera manera de ver el modelo lineal:

$$\gamma_{ij} = \gamma_{**} + (\gamma_{i*} - \gamma_{**}) + (\gamma_{ij} - \gamma_{i*})$$

donde cada uno de los términos representa, respectivamente a μ, α_i y e_{ij}.

Así podemos obtener un **modelo de regresión** sobre la base de una serie de datos obtenidos de ensayos o de observaciones sobre procesos existentes.

Veamos un ejemplo:

Hacemos ensayos para determinar la eficiencia de un procesador en función de un solo factor, temperatura, a un nivel bajo (105ºC, que lo codificamos como nivel 0) y a nivel alto (110ºC, codificado como nivel 1). Cada ensayo lo repetimos tres veces. Los resultados obtenidos están en la siguiente tabla:

		j = 1	j = 2
		Nivel 0	Nivel 1
i=1	Repetición 1	90	79
i=2	Repetición 2	91	80
i=3	Repetición 3	89	81

En primer lugar reordenamos los datos a fin de calcular el modelo lineal y su coeficiente de regresión:

Nivel	Xij
0	90
0	91
0	89
1	79
1	80
1	81

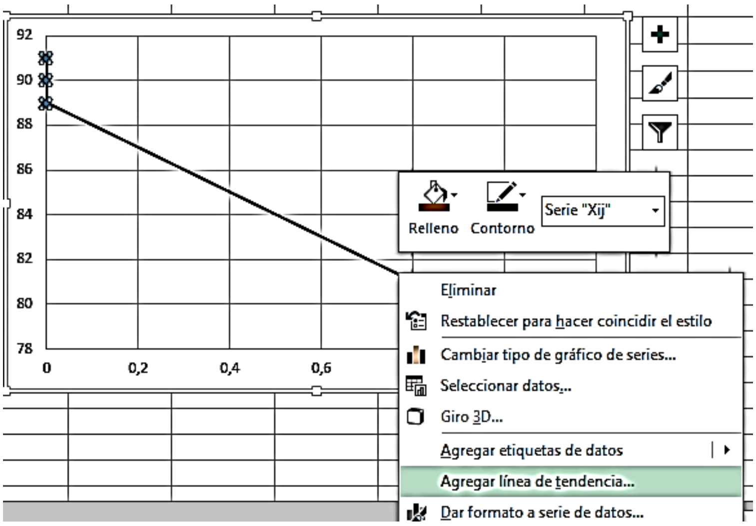
Lo llevamos a una planilla de cálculo y los escribimos de las dos maneras. Así, si usamos la segunda manera, podemos obtener un gráfico al que se adicionamos el análisis de regresión correspondiente. (La curva de regresión calculada aparece en línea gruesa).

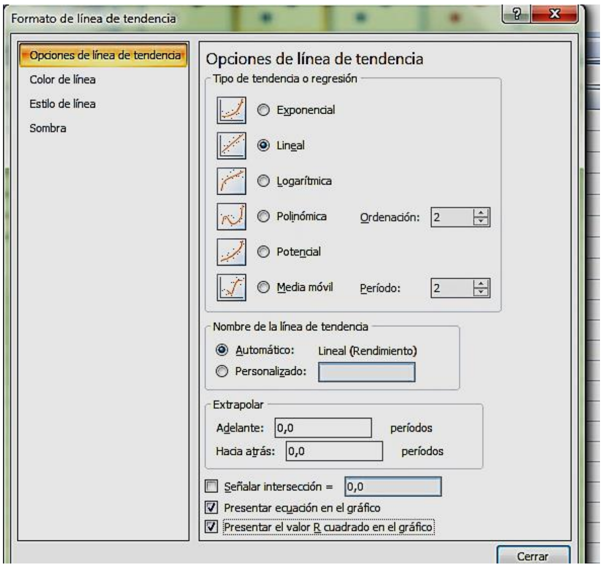
El procedimiento paso a paso es así:

1. cargamos los datos en una hoja, en las dos formas vistas arriba
2. usamos la forma mostrada en la última tabla (Columna "Niveles" y columna X_{ij})

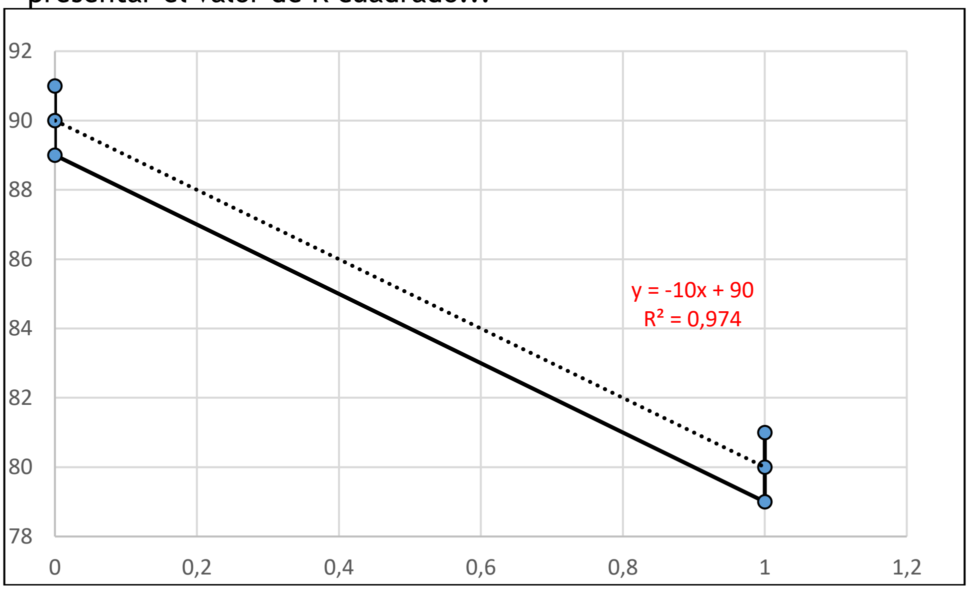
[35] Por ejemplo, son dos niveles: alto-bajo, frio-caliente, concentrado-diluido, pH 8 – pH 8.5, 120° — 180°, etc.)

3. Construimos el gráfico de dispersión XY, con el formato “puntos unidos por líneas”.
4. Editamos el gráfico (doble clic en el área de gráfico) y haciendo clic sobre la línea que une los puntos seleccionamos la línea.
5. Con la línea seleccionada pedimos el menú emergente y optamos por “Agregar línea de tendencia...”
6. Optamos por “Regresión Lineal” en la solapa “Opciones…” o en la barra derecha emergente, según la versión de hoja usada.





6. Y por último marcamos las casillas "presentar ecuación en el gráfico" y "presentar el valor de R cuadrado..."



Con estos pasos obtenemos, además del gráfico, el modelo lineal, donde podemos ver que

$b_0 = 90$

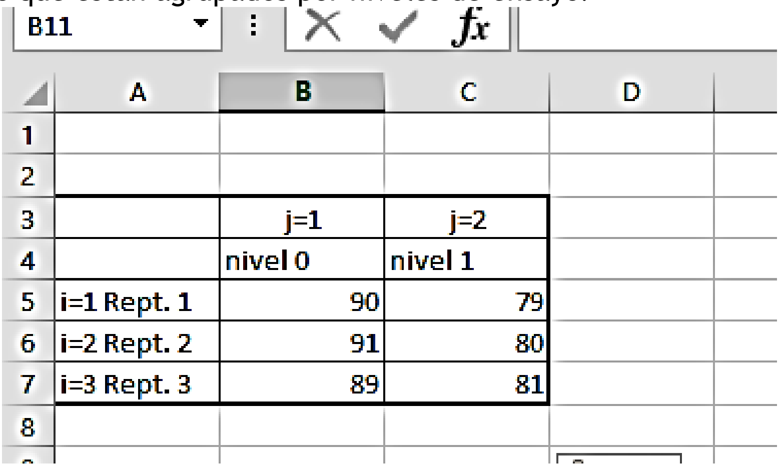
$b_1 = -10$

Prueba estadística del factor de temperatura

Ahora tenemos que probar, mediante análisis de varianza, que la variabilidad del factor α_i es significativa frente a la variabilidad del error e_{ij}.

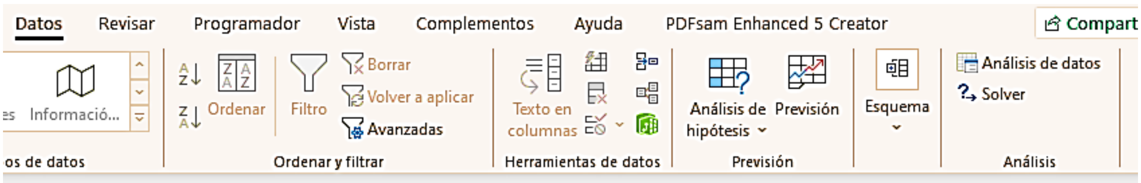
Mediante análisis de la regresión probaremos que la pendiente b_1 = — 10 no admite el cero como solución ($\beta \neq 0$).

Análisis de varianza
Para efectuar el análisis de varianza con Excel, partiremos de los datos cargados en la planilla, los que están agrupados por niveles de ensayo:

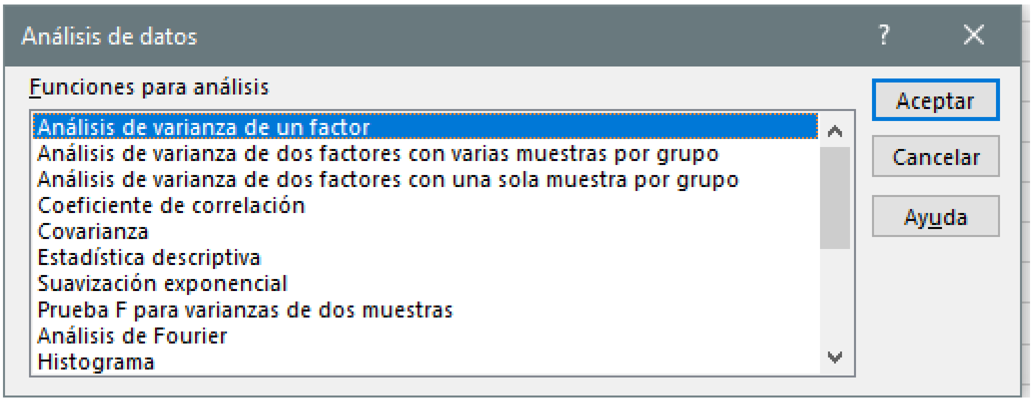


	A	B	C	D
1				
2				
3		j=1	j=2	
4		nivel 0	nivel 1	
5	i=1 Rept. 1	90	79	
6	i=2 Rept. 2	91	80	
7	i=3 Rept. 3	89	81	
8				

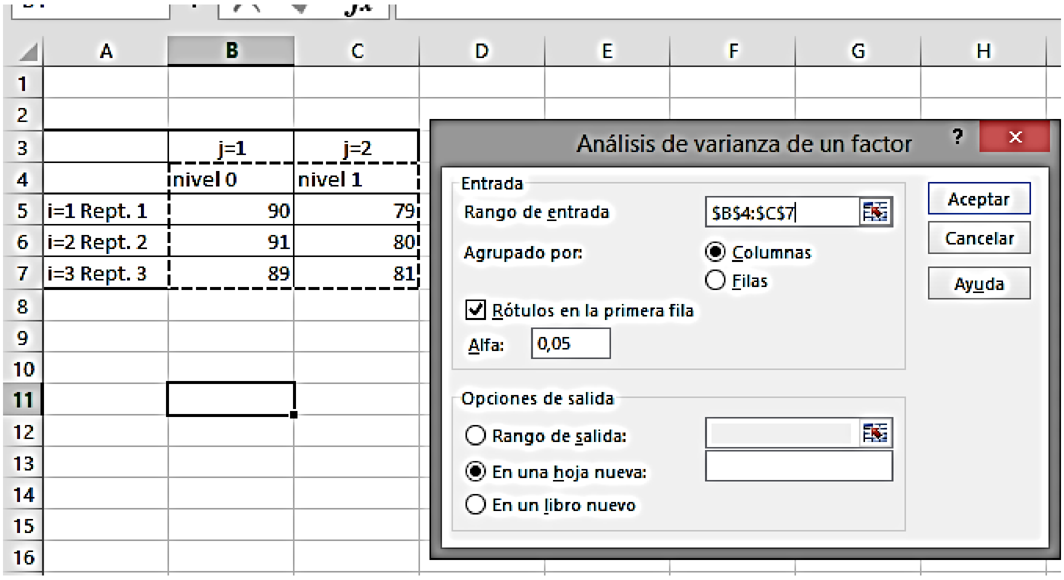
1. seleccionamos en la pestaña “Datos” y en el grupo “Análisis” el botón “Análisis de datos...”



2. Seleccionamos la opción Análisis de varianza para un solo factor.



3. Por último seleccionamos el rango de datos (desde celda que contiene “Nivel 0” a celda con “81”; B4 a C7, en el caso de la figura), seleccionando “datos agrupados en columnas con rótulos en la primera fila” y que la salida será en una hoja nueva:



Con este procedimiento veremos que se crea una nueva hoja, en la que aparece la siguiente información:

	A	B	C	D	E	F	G	H
1	**Análisis de varianza de un factor**							
2								
3	RESUMEN							
4	*Grupos*	*Cuenta*	*Suma*	*Promedio*	*Varianza*			
5	nivel 0	3	270	90	1			
6	nivel 1	3	240	80	1			
7								
8								
9	ANÁLISIS DE VARIANZA							
10	*n de las varia*	*na de cuadra*	*ados de libert*	*lio de los cua*	*F*	*Probabilidad*	*lor crítico para F*	
11	Entre grupos	150	1	150	150	0,00025522	7,70864742	
12	Dentro de lo	4	4	1				
13								
14	Total	154	5					
15								

La expresión "entre grupos" indica lo que se ha denominado α (variabilidad del factor). "Dentro de los grupos" indica el error, la variabilidad del error.
La suma de los cuadrados será, entonces,

la expresión $(\gamma_{i*} - \gamma_{**})^2$ para la variabilidad del factor (Entre los grupos)[36]

y la expresión $(\gamma_{ij} - \gamma_{i*})^2$ para el error (Dentro de los grupos)[37]

Al ser 6 ensayos, disponemos de 5 grados de libertad. El factor tiene uno y los restantes grados de libertad son asignados al error.

[36] Significa algo así como "cuando se detectan valores distintos entre los grupos (hay un grupo de datos en cada nivel) es porque el factor (que cambió de nivel) es el responsable"
[37] Acá, en cambio, los cambios están dentro de cada grupo, al no cambiar el factor, el responsable es el error.

Los cuadrados medios surgen de dividir la suma de cuadrados por los grados de libertad.

El estimador F se calcula dividiendo los cuadrados medios del factor por los del error. Este valor es clásicamente obtenido en tablas de tres entradas: grados de libertad del numerador, grados de libertad del denominador y probabilidad.

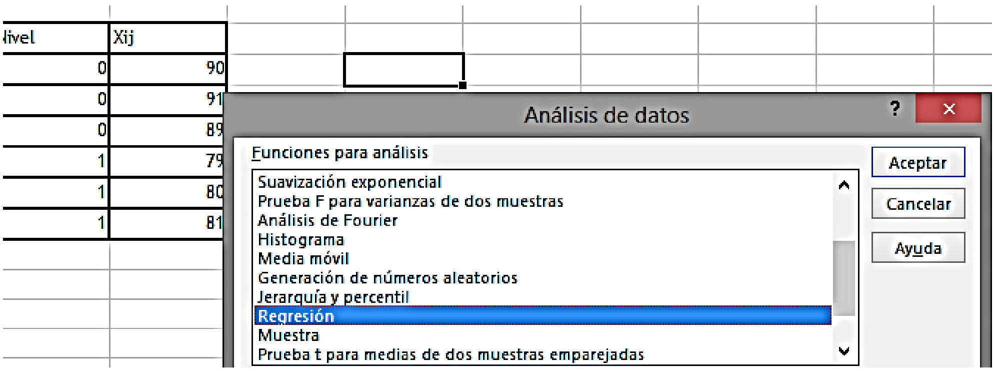
La hipótesis nula (H0) es que *las medias de ambos grupos son iguales*, por lo tanto la hipótesis a prueba es que *no lo son*, para eso, deberíamos encontrar que el valor esperado del numerador sea mayor que el del denominador.

Esto es, para H0, falso, debe haber un valor grande de F. Llamando a la probabilidad de error ***p*** y haciendo ***p*** = 1 — α. Decimos que la diferencia entre medias es *poco significativa* si α = 0,05 o menos; *significativa,* si α = 0,01 o menos y *muy significativa* si α = 0,001 o menos.

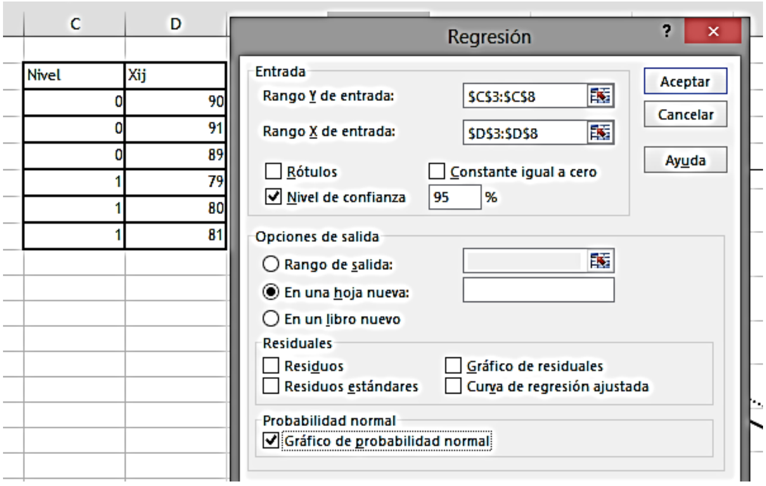
En este caso, la tabla de Análisis de Varianza presenta un valor de F de 150 (150/1), con un valor crítico para F de 7,7 y una probabilidad (de error) de 0,0002552, que es menor a α = 0,001, por lo que el valor de F resulta *muy significativo*, o, en otras palabras, la probabilidad de la hipótesis nula es muy baja.

Análisis de regresión

Por otro lado, podemos realizar el análisis de regresión a fin de demostrar que la pendiente b_1 no admite el cero como solución (lo que quiere decir es que **SI** hay una dependencia entre la temperatura y la eficiencia del procesador) en forma similar al análisis de varianza. Para eso simplemente seleccionamos la opción “Regresión” en el cuadro de diálogo que vemos al elegir “análisis de datos...”



Los datos que vamos a usar son los que están en la tabla con dos columnas (“Niveles” y “Xij”) que pusimos en primer lugar.



Obtenemos una nueva hoja con valores, algunos de ellos ya estaban presentes desde el comienzo del ejemplo, pero no tiene importancia porque todo el cálculo puede realizarse en conjunto.

Hay tres tablas en esta hoja. La primera nos indica los coeficientes de regresión, el número de observaciones, y el error típico.

La segunda tabla es la que más nos interesa: vemos que es similar a la de análisis de varianza ya vista, pero diferente en las fuentes de variación.

1	Resumen								
2									
3	*Estadísticas de la regresión*								
4	Coeficiente de correlación múltiple	0,98692754							
5	Coeficiente de determinación R^2	0,97402597							
6	R^2 ajustado	0,96753247							
7	Error típico	0,09869275							
8	Observaciones	6							
9									
10	ANÁLISIS DE VARIANZA								
11		*Grados de libert*	*a de cuadra*	*io de los cua*	*F*	*alor crítico de F*			
12	Regresión	1	1,46103896	1,46103896	150	0,00025522			
13	Residuos	4	0,03896104	0,00974026					
14	Total	5	1,5						
15									
16		*Coeficientes*	*Error típico*	*Estadístico t*	*Probabilidad*	*Inferior 95%*	*Superior 95%*	*nferior 95,0%*	*uperior 95,0%*
17	Intercepción	8,77922078	0,67719521	12,9640916	0,00020424	6,89902545	10,6594161	6,89902545	10,6594161
18	Variable X 1	-0,0974026	0,00795289	-12,2474487	0,00025522	-0,11948336	-0,07532184	-0,11948336	-0,07532184
19									

En efecto, la primera fila ahora se denomina “Regresión” y representa a la Hipótesis en prueba (“La pendiente es diferente de cero”). Otra vez la probabilidad de error merece alta confianza (H0 es improbable)

Conclusión

Con los elementos trabajados encontramos un modelo que predice con una confianza del 95% el comportamiento de un procesador en función de la temperatura, que se ajusta a la siguiente expresión:

$$\eta = 90 - 10\,T$$

donde η es la eficiencia del procesador y T la temperatura codificada de operación.

Pero aún nos falta encontrar la manera de "decodificar" la temperatura. En esa expresión vemos que la eficiencia es de 90 cuando T = 0. Pero cuando comenzó el desarrollo del modelo habíamos establecido un nivel de temperatura "**bajo**" llamado arbitrariamente "0", pero que en realidad equivalía a 105°C, y un nivel "**alto**" de 110°C, por ello la pendiente de — 10 deberá ser reemplazada por la relación entre pendientes original.

Recordemos que habíamos usado dos valores codificados en 0 y 1 con el que obtenemos un

$$\Delta T = 1 - 0 = 1$$

si no codificamos, el ΔT sería calculado de la siguiente manera:

$$\Delta T = 110 - 105 = 5.$$

Y ahora hacemos una regla de tres simple, de esta manera:

> <u>**Si**</u> para un ΔT en los valores reales de 5 corresponde un cambio de la eficiencia **η** de — 10 (recordando que **η** es la variable dependiente),
> <u>**entonces**</u>, para un cambio de 1 corresponderá una nueva pendiente **x***:*

$$\frac{5}{-10} = \frac{1}{x} = -2$$

de esta manera la nueva pendiente será —2.

Luego simplemente usamos la expresión η = m — 2 T y calculamos la ordenada con

[2 x 105 + 90]

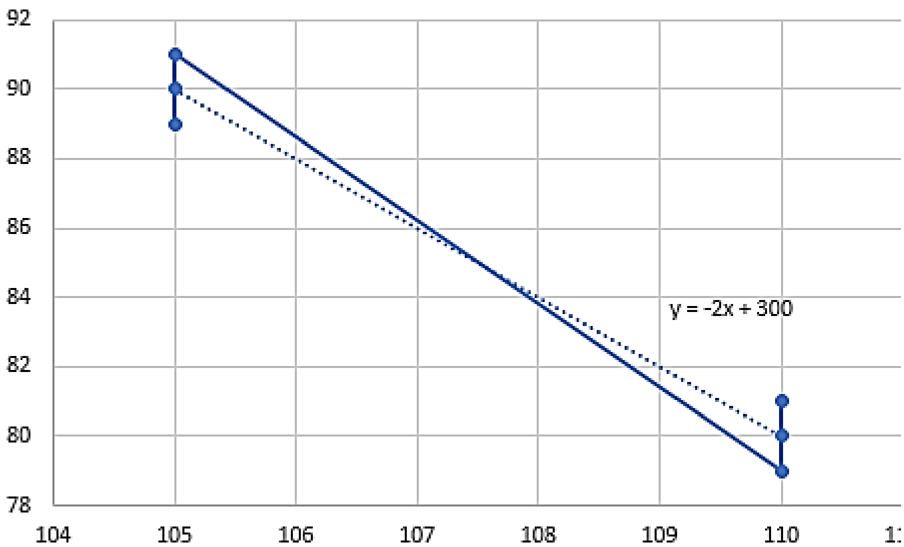
o con

[2 x 110 + 80]

donde lo que hicimos fue reemplazar un par cualquiera de valores η;T (105;90 o 110;80) y despejamos m.

$\eta = 300 - 2\,T$

Con esta función podemos calcular la eficiencia a cualquier temperatura. En un rango de temperaturas T desde 100 a 115 °C tendríamos



Temperatura (°C)	Eficiencia	

90	120	Zona fuera del modelo
100	100	
110	80	Zona de operación segura
120	60	
130	40	
140	20	
150	0	Zona fuera del modelo Procesador inoperable
160	-20	

Lo que nos muestra que a temperaturas inferiores a 110°C no hay influencia de la temperatura y a más de 130 °C comienzan serios desperfectos hasta un cese de funcionamiento encima de los 140 °C.

Modelo no lineal y no determinístico con datos observados.

Supongamos ahora que tenemos un recipiente de agua que se llena usando una bomba y una electro-válvula de corte. Nos piden diseñar un sistema automatizado PID (Proporcional, Integral y Derivativo), para lo cual necesitamos conocer, en primer lugar, el comportamiento del llenado del tanque. Haciendo varios ensayos hemos medido la tasa de llenado en porcentaje del volumen completado en función del tiempo. Son los siguientes valores:

Tiempo (segundos)	Nivel de agua (%)
X	**P(x)**
5	19
10	61
15	82

Vamos a construir un modelo de interpolación polinómico, buscando una función (modelo) del tipo:

$$P(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots + a_nx^n$$

Por la cantidad de datos que tenemos en este caso será suficiente usar un polinomio de orden 2:

$$P(x) = a_0 + a_1x + a_2x^2$$

Ahora construimos una expresión P(x) para cada observación (x) que tenemos en la tabla de arriba:

$$19 = a_0 + a_1 5 + a_2 5^2$$
$$61 = a_0 + a_1 10 + a_2 10^2$$
$$82 = a_0 + a_1 15 + a_2 15^2$$

Y a continuación lo planteamos como un sistema de tres ecuaciones con tres incógnitas:

$$a_0 + 5a_1 + 25a_2 = 19$$
$$a_0 + 10a_1 + 100a_2 = 61$$
$$a_0 + 15a_1 + 225a_2 = 82$$

Vamos a encontrar las raíces, para eso hacemos los siguientes pasos. Primero construimos una matriz, que queda así:

$$\left|\begin{matrix} 1 & 5 & 25 \\ 1 & 10 & 100 \\ 1 & 15 & 225 \end{matrix}\right|\begin{matrix} 19 \\ 61 \\ 82 \end{matrix}$$

Usamos el Método de Gauss Jordan, para encontrar las raíces

A cada fila de esta primera matriz la llamamos A_i, donde *i* será el número de cada una de las *m=3* ecuaciones

A1➔	1	5	25	19
A2➔	1	10	100	61
A3➔	1	15	225	82

El primer paso será forzar a que el elemento a_{11} de la matriz sea un 1. En este caso, ese elemento ya es un 1. Construimos una segunda matriz (B) en blanco con la misma cantidad de filas y columnas que la matriz A y llamamos Bj a cada nueva fila. Como el primer elemento de la fila A1 ya es un uno, a esa fila íntegra la copiamos igual en la fila B1 (**B1 = A1**)[38]

B1➔	1	5	25	19
B2➔				
B3➔				

El segundo paso consiste en que las filas B2 y B3 contengan un cero (0) abajo del 1 recién "logrado", para eso hacemos la siguiente operación elemento a elemento:
B2 = A2 – B1:

B1➔	1	5	25	19
B2➔	0	5	75	42
B3➔				

En la fila B3, también buscamos un cero (0) debajo del uno, y entonces hacemos:
B3 = A3 – B1

B1➔	1	5	25	19
B2➔	0	5	75	42
B3➔	0	10	200	63

Todo esto lo hacemos porque nuestro objetivo es construir una matriz identidad con su diagonal de 1 en a_{11}, a_{22} y a_{33}.

El tercer paso es construir una tercera matriz, C. En esta nueva matriz debemos convertir el 5 de la fila B2 en un 1 en la fila C2 de la nueva matriz. Lo hacemos dividiendo toda la fila B2 entre 5 y copiando los resultados en la fila C2 (**C2 = B2 / 5)**

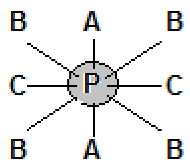
---

[38] Si hubiésemos tenido que convertir a_{11} en 1, tendríamos que dividir toda la fila por el mismo a_{11}. En ese caso lo que copiamos a la fila B1 es el resultado de la división. Por ejemplo si hubiese sido la fila original A1 ➔ 2; 8; 4; 16, hubiésemos dividido todo entre 2 y nos quedaba B1 ➔ 1; 4; 2; 8

C1➔
C2➔ 0 1 15 8,4
C3➔

A partir de acá, y sabiendo que encima y debajo de celda que contiene un 1 calculado por nosotros, completamos los elementos usando el siguiente diagrama con la siguiente terminología:

P (por Punto Pivote) es el “5” que por división se transformó en “1” para la nueva matriz — la buscada — de ahí su nombre de pivote o eje
A su ortogonal vertical a la altura de **B**,
C su ortogonal horizontal, a la altura de **B**,
B es el elemento que vamos a transformar y que en la nueva matriz será **B’**



Utilizando: $B' = B - \frac{A\,C}{P}$

Así, observando la fila B1, en lugar del “25” ¿Qué escribimos en C1?

B’ = $25 - \frac{75 \times 5}{5}$ = 25 - 75 = - 50

B1➔	1	**5**	25	19		C1➔			**-50**	
B2➔	0	**5**	**75**	42		C2➔	0	1	15	8,4
B3➔	0	10	200	63		C3➔				

En lugar del “19” quedará:

B’ = $19 - \frac{42 \times 5}{5}$ = 19 - 42 = - 23

C1➔ -50 **-23**
C2➔ 0 1 15 8,4
C3➔

En el renglón B3, el 200 será

B’ = $200 - \frac{75 \times 10}{5}$ = 50

C1➔ -50 -23
C2➔ 0 1 15 8,4
C3➔ **50**

Y el “63”

B’ = $63 - \frac{42 \times 10}{5}$ = - 21

C1➔ -50 -23

C2➔	0	1	15	8,4
C3➔			50	-21

Para el "1" que hay en B1 y que ya sabemos que debe seguir teniendo ese valor en C1, el cálculo lo ratifica, porque es

$$B' = 1 - \frac{0 \times 5}{5} = 1$$

aunque éste y el de los dos ceros encima y debajo del 1 de C2, se podrían obviar porque ya sabemos los valores que en esos lugares siempre tienen que aparecer[39]
Queda:

C1➔	1	0	-50	-23
C2➔	0	1	15	8,4
C3➔	0	0	50	-21

La siguiente matriz, será

D1➔	1	0	0	-44
D2➔	0	1	0	14,7
D3➔	0	0	1	0,42

Por lo tanto hemos encontrado las raíces del polinomio, que son

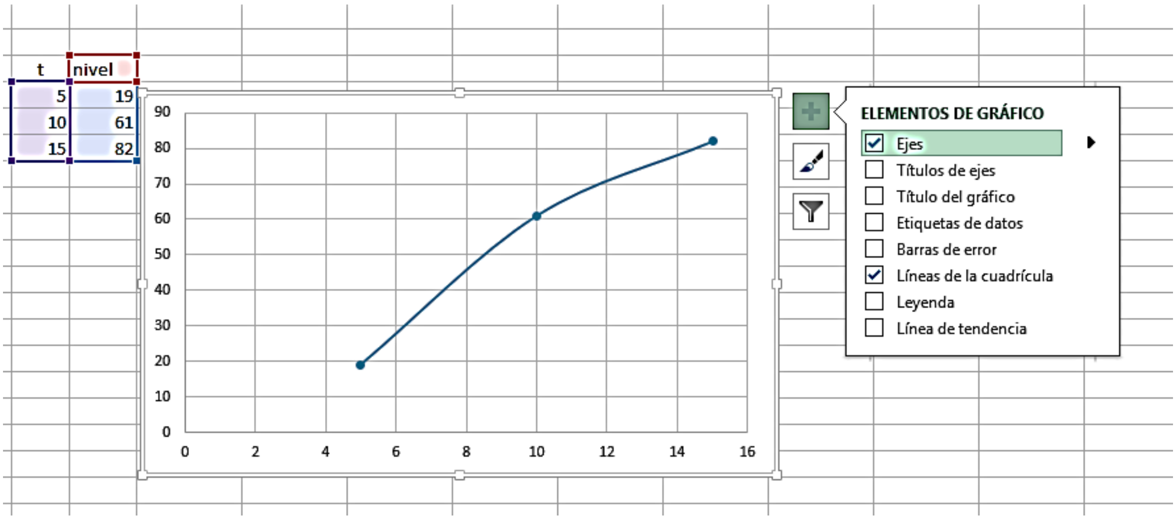
a_0 = -44
a_1 = 14,7
a_2 = 0,42

y ya podemos construir el modelo en función del tiempo (que está en segundos), y que será:

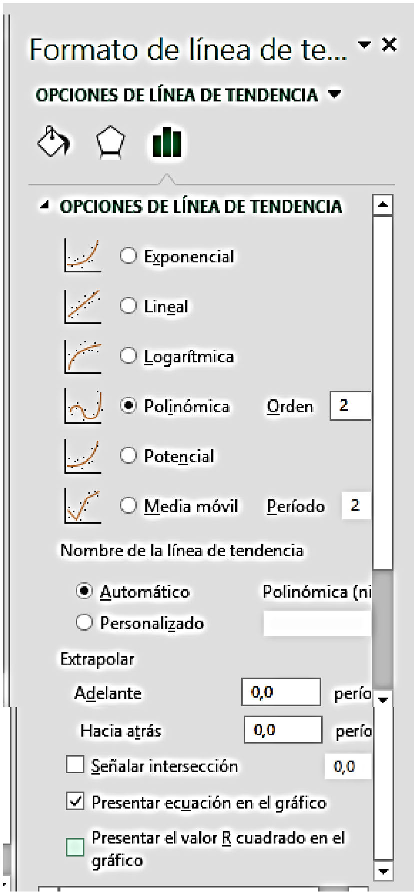
$$\boxed{P(x) = -44 + 14{,}7\,x + 0{,}42\,x^2}$$

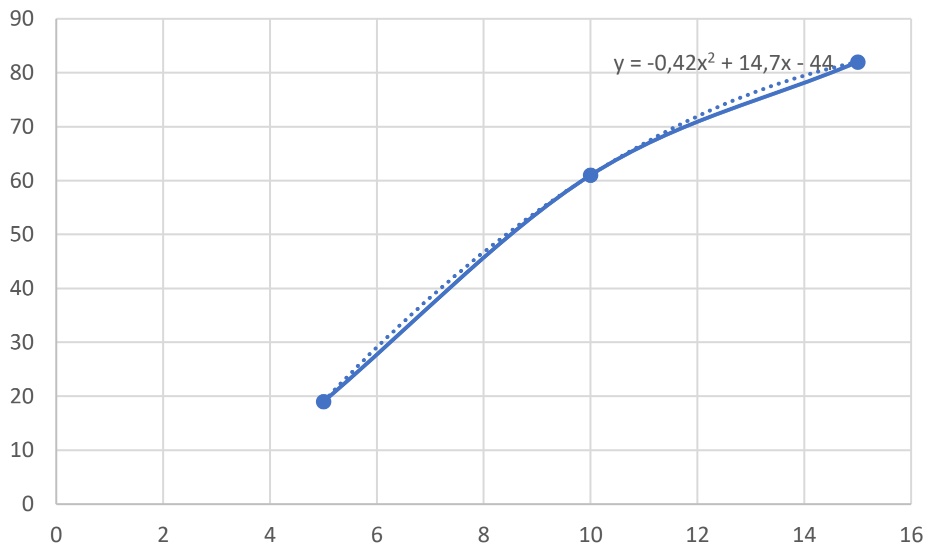
Alternativamente, en vez de estos pasos, podríamos resolverlo usando una hoja de cálculo, cargando los datos y graficándolos como dispersión XY.

[39] Observamos que, cuando hay que reemplazar los valores ortogonales verticales del "punto pivote" (del "1" en la nueva matriz), siempre deben ser ceros.
En esos casos, (ver la figura de análisis) A será igual a B y C = P, por lo tanto B' = B – (A . C / P) reemplazando, se convierte en B' = B – (B . P / P) = B – B = 0



Agregamos línea de tendencia, (clic derecho sobre la línea del gráfico), luego seleccionamos "polinómica" (en este caso de orden 2, porque así se desarrolló el ejemplo) y "agregar ecuación":

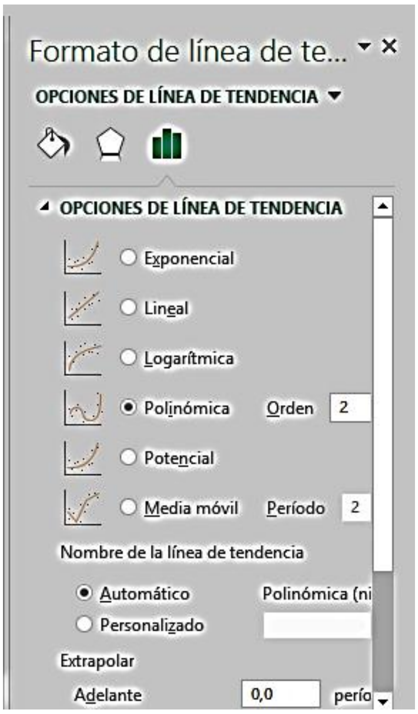




Como vemos, muy simplemente, obtuvimos un modelo matemático de los datos dados. ¿con que nivel de ajuste? Seleccionamos el "Presentar el valor R cuadrado en el gráfico", que en este caso nos dará $R^2 = 1$.

¿Cómo saber de antemano que modelo será el mejor?

Al contar con las facilidades que nos ofrecen las hojas de cálculo podemos presentar los datos y realizar el gráfico, como se indicó más arriba, hasta llegar a la ventana:



En ella vemos que podemos optar por varias alternativas para línea de tendencia: exponencial, lineal, logarítmica, polinómica y potencial (no corresponde el ajuste de media móvil). Vamos a seleccionar esas opciones, una por una, y registrar los resultados obtenidos, que en este caso los transcribimos en la planilla de abajo.

Modelo	Función	R^2
Exponencial	$P(x) = 10,6\ e^{0,15x}$	0,894
Lineal	$P(x) = 6,3\ x - 9$	0,964
Logarítmica	$P(x) = 57,7\ \ln(x) - 73,3$	0,998
Polinómica orden 2	$P(x) = -0,42x^2+14,7x-44$	1
Potencial	$P(x) = 2,23\ x^{1,37}$	0,968

Ahora estamos en condiciones de elegir el modelo que represente el mejor ajuste, que será aquel que presente un R^2 más cerca de "1".

CAPÍTULO 3

TOMA DE DECISIONES

El arte de tomar decisiones ha sido valorado desde tiempos antiguos y es — aun hoy — considerado como la habilidad o capacidad de algunas personas para evaluar en forma muy rápida la mayor cantidad de alternativas posibles y las consecuencias que acarrearía cada elección, y, de esta manera poder elegir aquella que más favorezca a sus intereses.

A medida que el método científico fue desarrollándose, desde la óptica cartesiana, forzosamente ese "arte" debía ser puesto en números y parámetros y, de esta manera se podría abordar el fenómeno racionalmente. Así se evitaba pensar que los grandes tomadores de decisión tuvieran, efectivamente, intuición, habilidades, dones u otros esquemas de razonamiento inconmensurables.

Las primeras especulaciones fueron hechas tomando como modelo cualquier juego en el cual dos contrincantes tuvieran a su disposición un número finito de estrategias posibles idénticas para cada uno de ellos, que ambos contrincantes también tuvieran igual información y un campo de resultados "de suma cero", es decir, que lo que ganaba uno era lo que perdía el otro.

En este capítulo desarrollaremos algunos aspectos de la toma de decisiones[40] de manera tal que podamos comprender el significado de algunas premisas y entender de qué manera la simulación puede ayudarnos a encontrar respuestas satisfactorias en ciertos casos.

Las técnicas de toma de decisiones como aproximaciones que pueden llevarse a modelos, se dividen en dos grandes grupos: ***Toma de decisiones bajo certidumbre*** y ***bajo condiciones de incertidumbre***.

La primera de ellas se refiere a aquellos casos en que los datos y las condiciones en que pueden ocurrir los fenómenos externos son determinísticos.

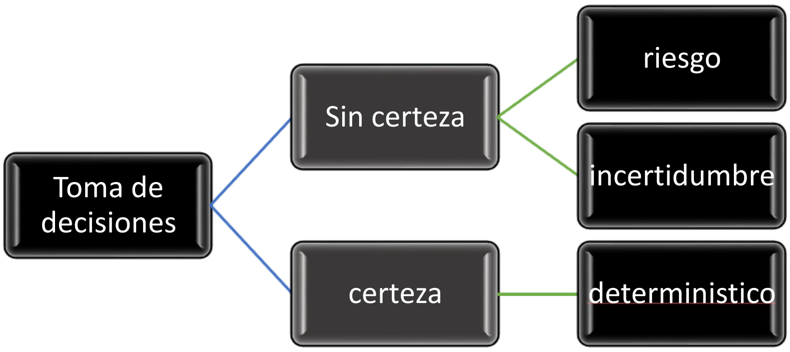
En el segundo grupo, los acontecimientos pueden ocurrir de dos maneras:
1. en forma aleatoria y sin posibilidades prácticas de conocer ni tan solo la probabilidad de que ocurran ni de asignarle nosotros una probabilidad subjetiva. En estos casos se habla de ***condiciones de incertidumbre***.
2. En forma probabilística, cuando se conoce o se puede asignar una probabilidad a la ocurrencia, entonces son ***condiciones de riesgo***. En estos casos hay que tener presente que, si llamamos N_j a una serie finita desde j= 1 a n de estados N, que las probabilidades de que ocurran son mutuamente excluyentes

[40] La teoría de juegos es, desde hace una décadas, uno de los puntos a los que más atención se presta en las ciencias económicas, a partir de trabajos de Von Neumann en los años 1940 y de John Nash en los 1950. Sin embargo hay teorías clásicas anteriores a estos autores, como la de Pascal, entre otros. Puede encontrar una aproximación en Optimiza, Libro 2 (Teoría Clásica y Von Neumann y Teoría de Nash)

$$\sum_{j=1}^{n} P(N_j) = 1$$

donde $P(N_j)$ son probabilidades mutuamente excluyentes.

[por ejemplo si tenemos una probabilidad de que llueva = 0,15, la llamaremos $P(N_1)$ = 0,15. Entonces tendremos que será $P(N_2)$ = probabilidad de que no llueva = 0,85]



Modelos de toma de decisiones

Para construir un modelo de toma de decisiones comenzaremos por ordenar la información en una **matriz de beneficios**, en donde A_i se refiere a las acciones o alternativas a seguir y N_i a los distintos estados posibles o estados de la naturaleza. Cada b_{ij} nos dice que beneficio o resultado obtendríamos en caso de tomar una acción A_i para cada estado N_i. Estos beneficios o resultados podemos expresarlos de diferentes maneras, según el objetivo que nos hemos impuesto. Por ejemplo, en términos económicos, pueden ser los beneficios asociados a cada acción A_i en caso que ocurra N_j .

A_i \ N_j	N_1	N_2	…	N_n
A_1	b_{11}	b_{12}		b_{1n}
A_2	b_{21}	b_{22}		b_{2n}
…				
A_m	b_{m1}	b_{m2}		bmn

Ejemplo: Una empresa considera la posibilidad de incorporar un nuevo servidor, y debe decidir un modelo de acuerdo con la capacidad de manejo de almacenamiento. Para simplificar, disponemos de tres tipos de servidores para elegir uno de ellos:

Grande (G), Mediano (M) o Chico (C).

El éxito o fracaso de la inversión depende de las condiciones económicas de la próxima temporada, que pueden ser:

MUY BUENAS, BUENAS, REGULARES o MALAS.

Construimos la siguiente *matriz de beneficios* expresada en unidades arbitrarias:

A_i \ N_j	Muy buenas	Buenas	Regular	Malas
G	2,5	0	-2,5	-4,1
M	1,6	0,3	-1,2	-2,1
C	0,6	0,4	-0,3	-0,3
NADA	-0,4	0	0	0

Como no conocemos las probabilidades para cada N_j ni podemos inventarlas subjetivamente, amos a decidir que modelo comprar usando alguno de los **modelos para condiciones de incertidumbre** que veremos a continuación y que presentaremos en orden a como fueron apareciendo históricamente, por lo tanto, el último es el que, generalmente, se usa actualmente.

Modelos para condiciones de incertidumbre

Criterio de Wald[41] (maximin), (pesimista)

Consiste en buscar el mínimo beneficio de cada acción y en recomendar la acción que tenga el máximo de esos mínimos. Significa que para cada A_j debemos encontrar el mínimo b_{i*}. Se basa en un criterio pesimista, es decir, algo así como razonar de la siguiente manera: "*las cosas siempre van a ir mal, si nos preparamos para elegir lo que mejor resulte si todo va mal, entonces cualquier otra situación siempre será favorable. Por eso elegimos lo mejor de todo lo peor*".

En este ejemplo, es:

$$\Omega = \max_i (\min_j (b_{ij}))$$

	MIN
G	-4,1
M	-2,1
C	-0,3
NO	-0,4

Como el máximo de estos mínimos es -0,3, la acción recomendada corresponde a la fila j = 3, que es comprar el servidor más pequeño. Este método, en la práctica, no se utiliza en acciones contra la naturaleza.

Criterio Maximax, (optimista)

Es el método opuesto al de Wald: buscamos ahora para cada fila A_j los máximos b_{i*}.y recomendamos el máximo de todos ellos. La idea subyacente es que todo va a ir bien, y tenemos que apostar por la mejor alternativa:

$$\Omega = \max_i (\max_j (b_{ij}))$$

	MAX
G	2,5
M	1,6
C	0,6
NO	0

Por consiguiente recomendaríamos la acción asociada con 2,5: comprar el servidor grande. Huelga decir que este criterio tampoco se usa mucho.

[41] Abraham Wald (1902 – 1950) matemático Rumano, creador del análisis secuencial.

Criterio de Hurwicz[42]
A diferencia de los anteriores, este criterio se basa en considerar que cada día, quien toma las decisiones, tiene un "estado de ánimo" o una "intuición" de como van a ir las cosas. Introduce un cierto grado de subjetividad, similar a *"cuan bien me va a ir hoy en lo que haga"*

Para utilizar este criterio es necesario crear e introducir un *coeficiente de optimismo* ω que establece, en forma subjetiva, el grado de optimismo que en particular tiene el tomador de decisiones cuando aborda un problema. Ese coeficiente expresa en una escala del 0 a 1 cuan optimista nos sentimos frente al problema que analizamos, siendo $\omega = 0$ totalmente pesimistas y $\omega = 1$ totalmente optimistas.

De esta manera, para cada acción vamos a calcular un beneficio resultante de la combinación de optimismo y pesimismo, de esta manera:

$$H_i = \omega \, MAX(\boldsymbol{b_{i*}}) + (1 - \omega) \, MIN(\boldsymbol{b_{i*}})$$

luego elegimos el máximo de estos H_i.

Por ejemplo si creemos que hoy nuestro coeficiente de optimismo es relativamente bajo, creemos que somos un 40% de optimistas, establecemos que $\omega = 0,4$, buscamos en cada acción cual es el H resultante:

$$H_1 = 0,4 * 2,5 + (1 - 0,4) * - 4,1 = - 1,46$$
$$H_2 = 0,4 * 1,6 + (1 - 0,4) * - 2,1 = - 0,62$$
$$H_3 = 0,4 * 0,6 + (1 - 0,4) * - 0,3 = 0,06$$
$$H_4 = 0,4 * 0 + (1 - 0,4) * - 0,4 = - 0,24$$
$$\Omega = max_i \, H_j = 0,06$$

Por tanto, nuestra recomendación será la <u>acción 3</u>, comprar el servidor chico.

Criterio de Savage[43], (minimax)

Es el criterio que se aplica en la actualidad. Se basa en que, cuando definimos el objetivo (que es a lo que nos enfocamos, o el interés de la toma la decisión, por ejemplo, ganar lo más posible) asumimos que tendremos que tomar una decisión y realizar una acción.

Sabemos que los estados de la naturaleza Nj son aleatorios, por tanto la acción que tomemos tendrá un grado de éxito mayor (todo ocurrió en la natura según lo esperado) o menor (no ocurrió según lo esperado).

Cuanto mayor sea el éxito, o, dicho de otra manera, cuanto más beneficio obtengamos si todo salió de acuerdo con lo que esperábamos, mayor será nuestra pérdida o costo si ocurre lo contrario.

---

[42] Leonid Hurwicz (1917 – 2008) matemático estadounidense de origen ruso, premio nobel en 2007 por trabajos de economía relacionados con teoría de juegos.
[43] Leonard Jimmie Savage (1917 – 1971) matemático estadounidense, colaboró con von Neumann. El concepto a que se alude (arrepentimiento) fue enunciado en 1951 y publicado en 1954.

Savage plantea que lo que interesa es que, después de que todo pase, quien decide no tenga que lamentarse por la decisión tomada, o, al menos, tenga que lamentarse lo menos posible. Para esto definió una función monetaria a la que llamó *el arrepentimiento*.

Apliquemos la función a nuestro problema: supongamos que finalmente compramos el servidor C y que luego, en la naturaleza, ocurre que la situación económica fue MUY BUENA.

Ganamos 0,6 pero lo mejor hubiera sido, para ese estado de la natura (MUY BUENA), que hubiéramos comprado un servidor grande. Evidentemente nos arrepentimos de no haber comprado uno grande. ¿Cuánto nos arrepentimos? Podemos darle valor a nuestro arrepentimiento por haber comprado uno chico en vez de uno grande simplemente viendo la diferencia entre lo que hubiésemos ganado con uno grande y lo que realmente ganamos con uno chico:

Arrepentimiento = Máximo beneficio posible — lo realmente ganado

Si para cada uno de los beneficios posibles podemos calcular el arrepentimiento, que implica respecto a lo mejor para ese estado de la natura, entonces estamos en condiciones de construir una *matriz de arrepentimientos*. La hacemos simplemente calculando el máximo beneficio para cada estado de la naturaleza y luego haciendo la diferencia con el beneficio correspondiente a cada una de las acciones para ese estado. La matriz de arrepentimientos que obtendremos sería la siguiente:

A_i \ N_j	Muy buenas	Buenas	Regular	Malas
G	0	0,4	2,3	4,1
M	0,9	0,1	1,2	2,1
C	1,9	0	0,3	0,3
NADA	2,9	0,4	0	0

de todo lo que tenemos para arrepentirnos para cada acción a tomar, buscaremos arrepentirnos lo **menos posible**. Buscaremos el mínimo de los máximos arrepentimientos:

$$\Omega = \min_i (\max_j (Arr_{ij}))$$

	MAX Arr
G	4,1
M	2,1
C	1,9
NO	2,9

Este criterio nos indica que el menor arrepentimiento lo tendremos si nos decidimos por la acción 3, comprar el servidor chico.

Este criterio es el que normalmente se utiliza en la práctica, a veces también si se disponen de probabilidades como veremos más adelante.

EJEMPLO

Una empresa decide lanzar al mercado un producto que tiene un costo de 32 y se recibe por él 47. No se sabe nada sobre la demanda, salvo que hay una mínima de 200 unidades y además se sabe que la planta puede producir desde ese mínimo de 200 unidades hasta un máximo de 2000. Las demandas insatisfechas no provocan pérdidas.

La matriz por construir sería el beneficio correspondiente a cada unidad producida, desde 200 hasta 2000 y a cada unidad demandada, desde 200 a 2000:

P \ X	200	201	...	2000
200	3000	3000		3000
201	2968	3015		3015
...				
2000	-54600	-54553		30000

de esta manera, la primera celda indica el beneficio de haber producido 200 con una demanda de 200, en la segunda celda de ese mismo renglón, como seguimos produciendo 200 tenemos el mismo beneficio, a pesar de que la demanda ahora es de 201, pero sólo hemos producido 200 y por eso obtenemos la misma ganancia.

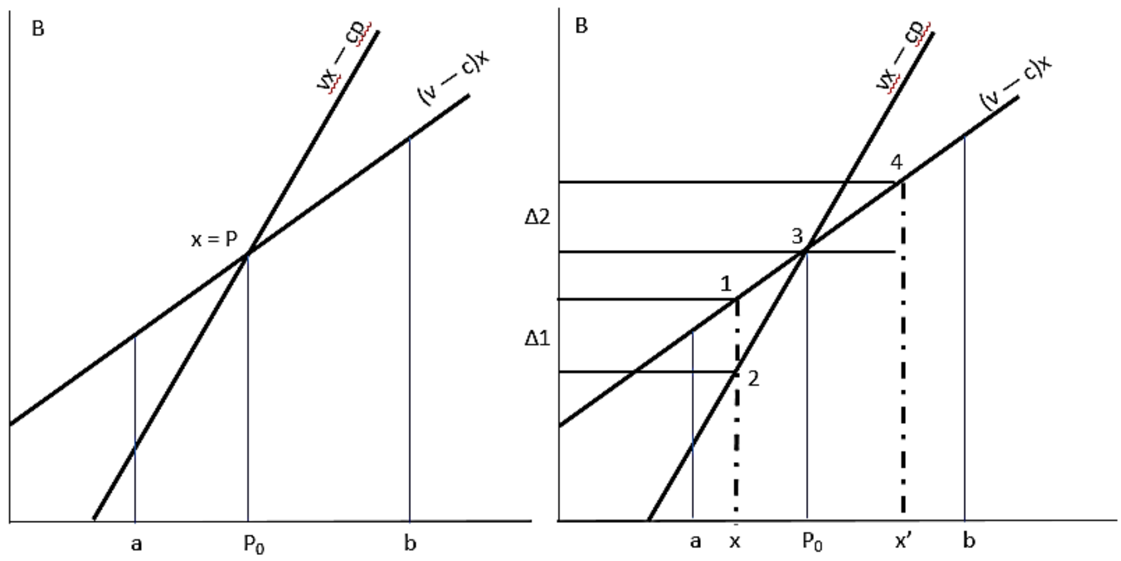
Por el contrario, el segundo renglón, que corresponde a una producción de 201, nos muestra que si la demanda es sólo 200, nos sobra una unidad y ganamos menos, pero desde una demanda de 201 en adelante ganamos el máximo para esa producción, y así renglón a renglón.

Sería muy complicado hacer esta matriz completa, porque se trataría de una tabla de 1800 columnas y 1800 filas. A pesar de ello, con lo que hemos hecho ya podemos advertir que la enorma matriz resultante va a tener una diagonal que refleja el valor máximo de ganancia para la producción de ese renglón.

A este valor lo llamaremos "**mejor acto**" porque refleja la ganancia de haber producido exactamente lo que se demandaba, ni más ni menos.

Como no vamos a escribir la tabla completa, nos conviene reescribir todo en forma de ecuaciones, usando dos expresiones, una para cuando la demanda es inferior a la producción y la otra para cuando la demanda supera la producción:

$x \leq P \qquad B = vx - cP$
$x > P \qquad B = (v - c)\,P$



El **punto 1** es el mejor acto cuando la demanda es igual a lo que producimos o sea, cuando x = P

El **punto 2** es la realidad de x < P, luego, Δ_1 es el arrepentimiento para este caso, que podremos denominar Ar_1
El **punto 3** es el mejor acto.
El **punto 4** es la realidad de x > P (mayor demanda que lo que se produce) y Δ_2 es el arrepentimiento de este caso, que podremos llamar Ar_2

Fijado un arrepentimiento a la derecha y a la izquierda, tendremos

$$x \leq P \quad Ar_1 = (v - c)x - (vx - cP) = c(P - x) \qquad \Delta_1$$

$$x > P \quad Ar_2 = (v - c)x - (v - c)P = (v - c)(x - P) \qquad \Delta_2$$

Vemos que hay arrepentimientos máximos en los extremos *a y b*.
Vamos cambiando los valores de P entre a y b, y cuando el arrepentimiento por derecha es igual al arrepentimiento por izquierda, se trata del óptimo.

$$c\ (P - a) = (v - c)\ (b - P)$$

reemplazando los valores de c, v, a y b, tenemos:

$$32\ (P\text{-}200) = (47\text{-}32)\ (2000\text{-}P)$$

$$P = 774$$

Siendo P = 774 el nivel óptimo de producción.

Modelos para condiciones de riesgo

Los modelos para condiciones de riesgo se usan cuando conocemos las probabilidades asociadas a cada estado de la naturaleza, el criterio empleado es el de *máximo beneficio esperado*, y consiste en calcular la **esperanza matemática** del beneficio de cada acción y elegir aquella que es máxima.

Vamos a escribir de nuevo la matriz de beneficios inicial, pero esta vez le agregaremos una nueva columna para las esperanzas matemáticas, *E(B),* y una nueva fila en el encabezado donde anotamos las probabilidades *$P(x_j)$* para cada estado de la natura:

$P_{(xj)} \rightarrow$		0,2	0,2	0,4	0,2	
$A_i \downarrow$	$N_j \rightarrow$	Muy buenas	Buenas	Regulares	Malas	E(B)
Grande G		2,5	0	–2,5	–4,1	–1,32
Mediano M		1,6	0,3	–1,2	–2,1	–0,52
Chico C		0,6	0,4	–0,3	–0,3	0,02
NADA		–0,4	0	0	0	–0,08

Para calcular la nueva columna E(B) usamos la esperanza matemática de cada fila *i* y luego extraeremos el máximo de esa columna.

$$E(B) = \sum_{j}^{n} P(x_j) b_{ij}$$

$$\Omega = \max E(B_{ij})$$

El máximo de las *E(B)* es 0,02, asociado con la compra de un servidor pequeño. Si quisiéramos usar este método con una matriz de arrepentimientos, entonces deberíamos calcular de esta manera:

$$E(Arr) = \sum P(x_j) Arr_{ij}$$

$$\Omega = \min E(Arr_{ij})$$

Predicción de los estados de la naturaleza.

Puede ocurrir que tuviéramos la posibilidad de contratar a alguien o a algún servicio que sea capaz de pronosticar el futuro, de darnos información sobre lo que ocurrirá con los estados de la naturaleza. Es el equivalente a saber cuál es el subíndice *j* que prevalecerá en los N_j estados posibles.

Obviamente la capacidad de predicción del futuro varía desde el extremo de no poder predecir nada hasta la predicción perfecta (o, en términos realistas, la predicción más confiable). La predicción se perfecciona con investigación, por lo tanto aparece un costo asociado a la predicción: este costo aumenta con el grado de perfección, pues aumenta la necesidad de investigación asociada con ese grado de perfección.

Para llegar a una aproximación definiremos el *Beneficio esperado con información perfecta* (*BEIP*), que es la esperanza matemática obtenida usando las probabilidades que nos dan cuando usamos asesoramiento para predicción de futuro

$$BEIP = \sum_{j}^{n} P(x_j) \max B_{ij}$$

que podemos aplicar al caso que estamos analizando, en cuyo caso encontraremos que:

BEIP = 2,5 x 0,2 + 0,4 x 0,2 + 0 x 0,4 + 0 x 0,2 = 0,58

Como nosotros sabemos el beneficio de la esperanza matemática sin ninguna información adicional, calculado como

$\max E(B_{ij})$

podemos ahora encontrar que diferencia hay con el beneficio calculado con la información perfecta (*BEIP):*

$$BEIP - \max E(B_{ij}) = VEIP$$

A esta diferencia la llamamos *valor esperado con información perfecta*, siendo en este caso:

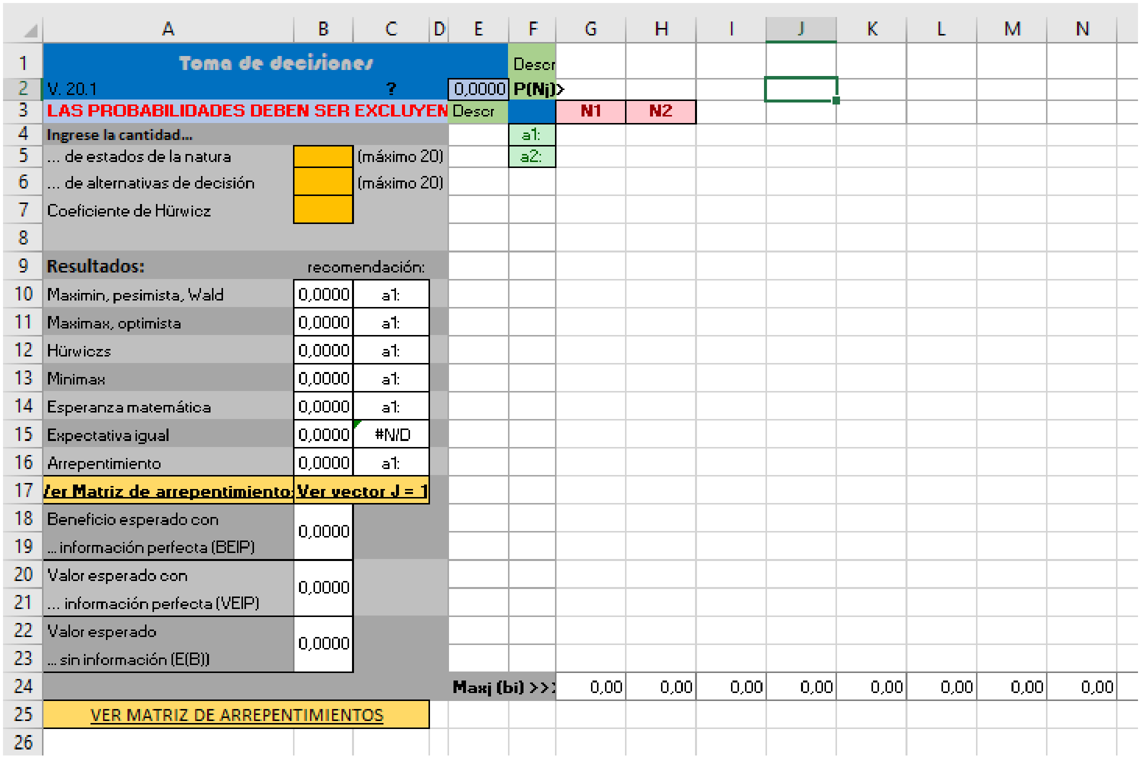
$$0{,}58 - 0{,}02 = 0{,}56$$

Como vemos, es la diferencia entre lo que se esperamos obtener con pronóstico y lo que esperamos sin ese pronóstico, por lo tanto, este valor VEIP nos indica lo máximo que podemos pagar por información (pronóstico) sobre el futuro.

El *VEIP* es proporcional a σ, si asumimos que trabajamos con información con distribución normal.

Resolución del caso con Excel.
Hemos preparado una plantilla[44] para poder resolver casos de toma de decisión de hasta 20 estados de la natura y hasta 20 acciones posibles.

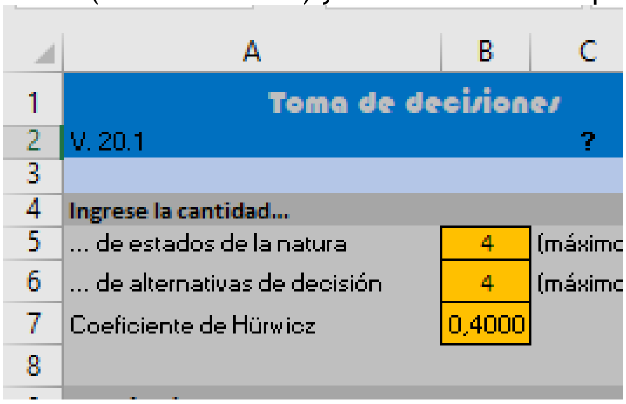
Usando esa plantilla, comenzamos por cargar los datos generales del caso que estamos desarrollando. La figura siguiente muestra la parte izquierda de la planilla antes de introducir los datos. Como vemos, aparece una matriz de 2 x 2 en blanco (tamaño mínimo) y tres celdas en amarillo:



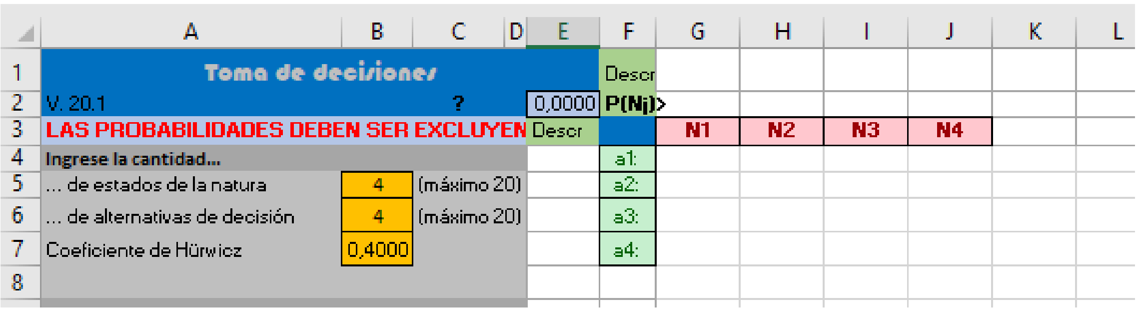
[44] Consulte la manera de obtener una copia de la plantilla Excel/decisiones, lista para usar, en www.optimiza.org

Esta plantilla tiene un código de colores: Las celdas amarillas se destinan a ingreso de datos. Veamos ahora como los cargamos y los pasos que deberíamos seguir:

1) introducimos la cantidad de estados de la natura (cuatro, en este caso), las acciones a decidir (también cuatro) y el coeficiente de optimismo (0,4):

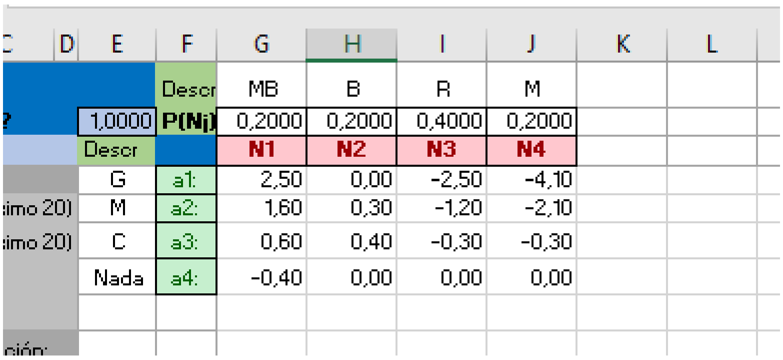


Esto genera una matriz en blanco destinada a la carga de datos y nombre de las acciones y estados de la natura:



Que completamos con los datos del caso que estamos siguiendo:

a) El nombre descriptivo de los estados de la natura (primera fila),
b) las probabilidades de los estados de la natura, que pueden ser nulos o no. En caso de que la suma de probabilidades no sea igual a uno, aparecerá una advertencia como se ve en la figura anterior.
c) El nombre descriptivo de las acciones posibles en la primera columna
d) la matriz de beneficios propiamente dicha



2) Una vez completada la carga, veremos el marginal inferior de la matriz con el máximo de cada columna

Maxj (bi) >>:	2,50	0,40	0,00	0,00	0,00	0,0

y en el marginal derecho el vector columna correspondiente a cada uno de los criterios junto a la acción recomendada, según corresponda, señalada en rojo para cada uno de ellos:

Z	AA	AB	AC	AD	AE	AF	AG	AH	AI
	Columnas marginales (j=5)								
	Acción	Wald	optimista	Hurwicz	Minimax	E(B)	Equi-P()	Acción	
	a1:	-4,10	2,50	-1,46	2,50	-1,32	-1,03	G	
	a2:	-2,10	1,60	-0,62	1,60	-0,52	-0,35	M	
	a3:	-0,30	0,60	0,06	0,60	0,02	0,10	C	
	a4:	-0,40	0,00	-0,24	0,00	-0,08	-0,10	Nada	

3) Volviendo a la izquierda, debajo de las celdas destinadas a carga de datos encontraremos un resumen de todos los resultados

	A	B	C	D	E	F	G	H
1	Toma de decisiones					Descr	MB	B
2	V. 20.1		?		1,0000	P(Nj)	0,2000	0,200
3					Descr		N1	N2
4	Ingrese la cantidad...				G	a1:	2,50	0,0
5	... de estados de la natura	4	(máximo 20)		M	a2:	1,60	0,3
6	... de alternativas de decisión	4	(máximo 20)		C	a3:	0,60	0,4
7	Coeficiente de Hürwicz	0,4000			Nada	a4:	-0,40	0,0
8								
9	Resultados:		recomendación:					
10	Maximin, pesimista, Wald	-0,3000	a3: C					
11	Maximax, optimista	2,5000	a1: G					
12	Hürwiczs	0,0600	a3: C					
13	Minimax	0,0000	a4: Nada					
14	Esperanza matemática	0,0200	a3: C					
15	Expectativa igual	0,1000	a3: C					
16	Arrepentimiento	1,9000	a3: C					
17	/er Matriz de arrepentimiento:	Ver vector J = 5						
18	Beneficio esperado con	0,5800						
19	... información perfecta (BEIP)							
20	Valor esperado con	0,5600						
21	... información perfecta (VEIP)							
22	Valor esperado	0,0200						
23	... sin información (E(B))							
24					Maxj (bi) >>:		2,50	0,4
25	VER MATRIZ DE ARREPENTIMIENTOS							

como vemos en la parte inferior, aparecen los valores de BEIP, BEIP y valor de la E(B). Finalmente, encontramos enlaces a la matriz de arrepentimientos, que nos llevan a ella:

AJ	AK	AL	AM	AN	AO	AP	AQ	AR	A	BD	BE	BF
									Matri			
	P(Nj)>	0,2000	0,2000	0,4000	0,2000							
volver		N1	N2	N3	N4							
	a1	0,00	0,40	2,50	4,10							4,10
	a2	0,90	0,10	1,20	2,10							2,10
	a3	1,90	0,00	0,30	0,30							1,90
	a4	2,90	0,40	0,00	0,00							2,90
												0,00
												0,00
												0,00
		2,50	0,40	0,00	0,00	0,00	0,00	0,00	0,0C			

Toma de decisiones bajo riesgo con variable continua.

En estos casos el criterio que generalmente se utiliza es el de máximo beneficio esperado. En ese caso tendremos que calcular la *esperanza matemática* del beneficio para cada alternativa, trabajando bajo la hipótesis de que conocemos el modelo probabilístico al que puede ser ajustado el sistema.

Ejemplo I

Una empresa proveedora de tecnología ha establecido un negocio para la instalación de una planta procesadora de leche, llave en mano, por un monto de \$ 600.000 con un plazo de entrega de 210 días.

La obra tiene un costo fijo de \$ 330.000 y un costo variable de 500 \$/día. En el contrato se establece una penalización de 2.500 \$/día de atraso y un premio de 1.500 \$/día de adelanto.

Mediante la técnica PERT encontramos que el desarrollo del proyecto tiene un tiempo medio de duración μ = 195 días con σ = 25. La variable aleatoria es el tiempo de duración de la obra, con una distribución normal.

Vamos a calcular el beneficio esperado, usando los siguientes parámetros

x = tiempo de duración de la obra
μ = tiempo medio de duración de la obra = 195 días
σ = 25
plazo de entrega = 210 días

Si tenemos en cuenta las penalizaciones, premios y costos podemos establecer dos expresiones, una para finalización antes del plazo de entrega y otra para después:

Para x ≤ 210

B = 600000 – 330000 – 500x + 1500 (210 – x)

Para x > 210

B = 600000 – 330000 – 500x – 2500(x – 210)

que corresponde a

B = 585000 – 2000x para x ≤ 210

B = 795000 – 3000x para x > 210

Podemos calcular la esperanza matemática del plazo de entrega B:

$$E(B) = \int_{-\infty}^{210} (585000 - 2000x) f(x)dx + \int_{210}^{\infty} (795000 - 3000x) f(x)dx$$

$$E(B) = 585000\int_{-\infty}^{210} f(x)dx - 2000\int_{-\infty}^{210} xf(x)dx + 795000\int_{210}^{\infty} f(x)dx - 3000\int_{210}^{\infty} xf(x)dx$$

que puede ser escrita, reemplazando

$$\int_{-\infty}^{x} f(x)dx = F(x) = \Phi(z)$$

$$E(B) = 585000F(x) - 2000\int_{-\infty}^{210} xf(x)dx + 795000G(x) - 3000\int_{210}^{\infty} xf(x)dx$$

Esta es la expresión de la esperanza matemática del beneficio. Aparece aquí una novedad: existe una variable de corte en la expresión funcional del beneficio, lo que da lugar a dos nuevas definiciones,

Expectativa Parcial Izquierda

$$\int_{-\infty}^{x} xf(x)dx = H(x)$$

Expectativa Parcial Derecha

$$\int_{x}^{\infty} xf(x)dx = J(x)$$

de manera tal que se cumpla la condición que la suma de ambas expectativas parciales sea la media del proyecto:

$$H(x) + J(x) = \mu = \int_{-\infty}^{\infty} xf(x)dx$$

Luego

$$H(x) = \int_{-\infty}^{x} xf(x)dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma} \right)^2 \right] dx$$

Desarrollando la integral y efectuando cambio de variable:

$$z = \frac{x - \mu}{\sigma}$$

$$x = \mu + \sigma z$$

$$dx = \sigma dz \Rightarrow dz = \frac{dx}{\sigma}$$

resolviendo queda

$$H(x) = \mu\Phi(z) - \frac{\sigma}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}$$

volviendo al cálculo de $E_{(B)}$

$E_{(B)}$ = 585000 F(210) – 2000 H(210) + 795000 G(210) – 3000 J(210)

$$z = \frac{x-\mu}{\sigma} = \frac{210-195}{25} = 0{,}6$$
$$\Phi(0{,}6) = F(210) = 0{,}7257$$
$$H(x) = H(210) = 195 \times 0{,}7257 - \frac{25}{\sqrt{2\pi}} \times e^{\frac{-0{,}6^2}{2}} = 133{,}18$$
$$J(210) = \mu - H(210) = 195 - 133{,}18 = 61{,}82$$
$$G(210) = 1 - F(210) = 1 - 0{,}7257 = 0{,}2743$$
$$E(B) = 190784$$

Optimización con demanda aleatoria

Para resolver estos casos, necesitamos conocer el modelo de distribución de probabilidad de la variable aleatoria. Generalmente es difícil conocer estas distribuciones, entonces lo que haremos comúnmente es ajustar a una distribución conocida. Usualmente, para períodos cortos de tratamiento la distribución puede ser la distribución *gamma* y para períodos largos la distribución *normal*.

Básicamente necesitaremos dos tipos de informaciones:

1) **Información Económica**. Determinística.
   a) c — Costo unitario
   b) v — Precio de venta unitario
   c) v_R — Valor residual de las unidades sobrantes
   d) cs — Costo de escasez de las unidades faltantes
2) **Información Estadística**
   a) Modelo de distribución
      i) μ
      ii) σ

Se trata de hallar el óptimo con un concepto de decisión de **máximo beneficio esperado.**

Ejemplo II

Un producto perecedero con demanda aleatoria tiene etiquetado un plazo de vida útil. Pasado el tiempo puede tener cierto valor residual diferente de cero. Existe un costo de escasez. Hallar la cantidad óptima de producto a fabricar, suponiendo un modelo normal de media μ y varianza σ

Si el lote a producir es *P*, el problema consiste en hallar *P* óptimo. Una vez más plantearemos dos expresiones, una para la variable X (demanda aleatoria) inferior o igual a *P y* otra superior a *P:*

$x \leq P$ B = v_x – cP + v_R (P – x) = (v – v_R) x– (c – v_R) P
$x > P$ B = v_P – cx – cs (x – P) = (v c + cs) P – cs x

$$E(B)=\int_{-\infty}^{P}\left[(v-v_R)x-(c-v_R)P\right]f(x)dx+\int_{P}^{\infty}\left[(v-c+cs)P-csx\right]fxdx$$
$$E(B)=(v-v_R)H(P)-(c-v_R)PF(P)+(v-c+cs)P(1-F(P))-cs(\mu-H(P))$$
$$E(B)=y(P)\therefore\frac{dy(P)}{dP}=0\Rightarrow P=Poptimo\text{ si y solo si }\frac{dy(P)^2}{dP^2}<0\text{ dado que en P hay}$$
un maximo
$$\frac{dE(B)}{dP}=(v-v_R)H'(P)-\left[(c-v_R)PF'(P)+(c-v_R)F(P)\right]+$$
$$+(v-c+cs)(1-F(P))-(vc+cs)PF'(P)+csH'(P)$$

pero como

$$F(x)=\int_{-\infty}^{x}f(x)dx\Rightarrow\frac{dF(x)}{dx}=\frac{d}{dx}\int f(x)dx=f(x)$$
$$H(x)=\int_{-\infty}^{x}f(x)dx\Rightarrow\frac{dH(x)}{dx}=\frac{d}{dx}\int xf(x)dx=xf(x)$$

resulta:

$$\frac{dE(B)}{dP}=(v-c+cs)-(v-v_R+cs)F(P)=0$$
$$F(P)=\frac{(v-c+s)}{(v-v_R+cs)}$$
$$\frac{d^2E(B)}{dP^2}=-(v-v_R+cs)F'(P)<0$$

luego *P* es óptimo, maximizando el beneficio esperado. Reemplazando F(P) en E(B)

$$maxE(B)=(v-v_R+cs)H(P)-cs\mu$$
$$donde$$
$$H(P)=\mu\Phi(z)-\frac{\sigma e^{\frac{-z^2}{2}}}{\sqrt{2\pi}}$$

Para una distribución normal N. ¿Cuál es el mejor acto?
Cuando x = P, entonces B = (v – c)x

E((v – c)x) = BEIP = v – c E(x) = (v – c)μ

BEIP = (v – c)μ

Entonces:

$$VEIP = BEIP - E(B)$$
$$VEIP = (v-c)\mu - (v - v_R + cs)H(P) - cs\mu$$
$$VEIP = (v-c)\mu - \left[(v - v_R + cs)\left((\ )\mu\Phi(z) - \frac{\sigma e^{-z^2/2}}{\sqrt{2\pi}} \right) \right] - cs\mu$$
$$VEIP = (v - v_R + cs)\frac{\sigma e^{-z^2/2}}{\sqrt{2\pi}}$$

Si σ aumenta, aumenta el *VEIP*: cuando mayor es la dispersión de la información mayor es el costo a pagar por ella

Ejemplo III
Una panificadora desarrolla un nuevo producto a lanzar al mercado. El producto tiene un costo unitario de 12 y un valor de venta de 20. El producto no vendido se destina a pan rallado y tiene un valor de 8. La demanda sigue una distribución normal estable de σ = 3233 y de μ = 22750 unidades semanales.
1. Determinar la cantidad óptima a producir semanalmente
2. El beneficio que arrojaría esa producción
3. El *VEIP*

$$F(P) = \frac{v - c + cs}{v - v_R - cs} = \frac{20-12}{20-8} = 0{,}66$$
$$z = \frac{x-\mu}{\sigma} \Rightarrow x = \mu + \sigma z$$
$$F(P) = 0{,}66 = \Phi(z) \text{ por tabla z = 0,4316}$$

luego x = P_{opt} = 22750 + 0,4316 * 3233 = 24144 unidades

Arboles de decisión. Toma de decisiones multinivel

Supongamos un caso de toma de decisiones, en el cual, basándose en la experiencia que tenemos sobre el negocio, somos capaces de estimar las siguientes probabilidades para la aceptación de un nuevo producto: 0,4 fracaso; 0,4 para éxito y 0,2 para un gran éxito. Debemos decidir entre una inversión Baja (B), Media (M) o Alta (A).

Utilizando el criterio de la esperanza matemática (probabilístico) obtendremos la decisión de inversión Media:

	RESULTADOS			
	FRACASO - F	**ÉXITO - E**	**GRAN ÉXITO - G**	
DECISIÓN ↓	0,4	0,4	0,2	←PROBABILIDAD
BAJA – B	-2	5	8	
MEDIA – M	-5	10	12	
ALTA – A	-8	6	15	

Estos datos pueden resumirse en un árbol que comienza con un nodo cero que representa el momento en el cual tomamos una decisión: elegimos una de las tres alternativas de inversión. Cada alternativa tiene tres resultados posibles. Así, el nodo cero conecta con el nodo 1, con el 2 y con el 3. Si se llegó al nodo 1, hay tres arcos, uno al nodo 4, otro al 5 y el último al 6.

Si se llegó al 4 se obtuvo una “ganancia” de –2 (Inversión baja con resultado Fracaso). La probabilidad de encontrar este resultado es:

P(fracaso | inversión baja) = P(F|B) = 0,4.

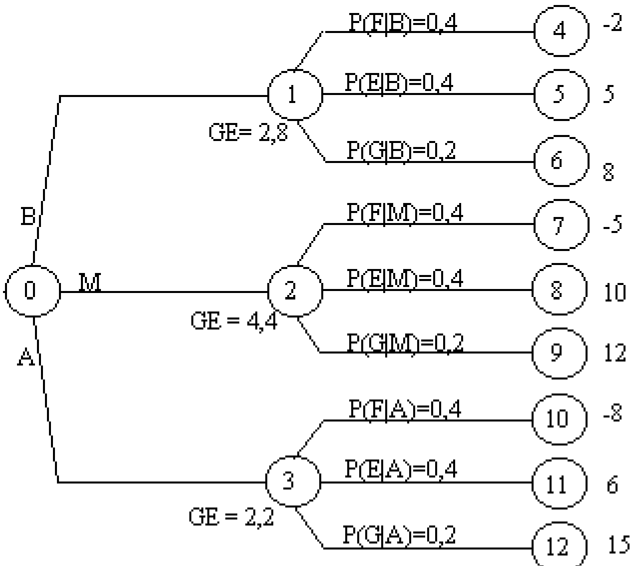
Para utilizar el árbol se debe comenzar con calcular la ganancia esperada con cada alternativa de los nodos 1, 2 y 3. Por ejemplo:

Ganancia Esperada En el nodo 1	=	Ganancia Esperada En el nodo 4	x	Probabilidad Asociada con el Arco 1–4	+	Ganancia esperada en el Nodo 5	x	Probabilidad Asociada con el Arco 1-5	+	Ganancia esperada en el Nodo 6	x	Probabilidad Asociada con el Arco 1–6

Ganancia Esperada En el nodo 1	=	-2	x	0,4	+	5	x	0,4	+	8	x	0,2

Ganancia Esperada (GE) en el nodo 1	= 2,8

Este valor calculado para cada alternativa se incorpora al árbol:



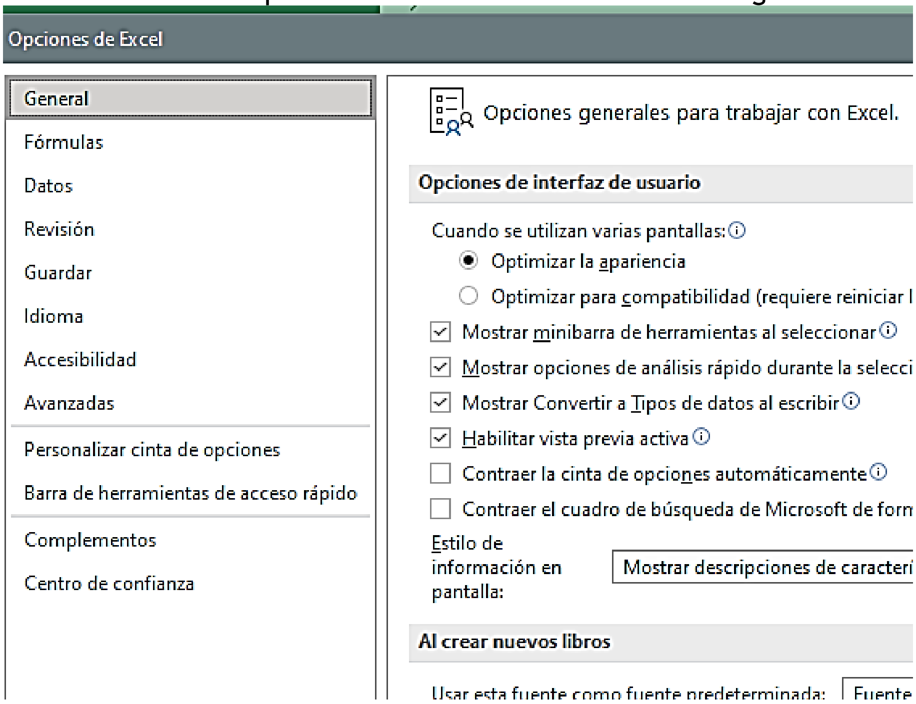
Ganancia Esperada: GE

La mejor alternativa resulta ser la inversión moderada, que promete una ganancia de 4,4. Es un resultado congruente con los obtenidos, ya que se usan los mismos cálculos. Sin embargo este método permite calcular tomas de decisiones más complejas, como aquellas en que las probabilidades de los resultados se afectan por la decisión tomada.

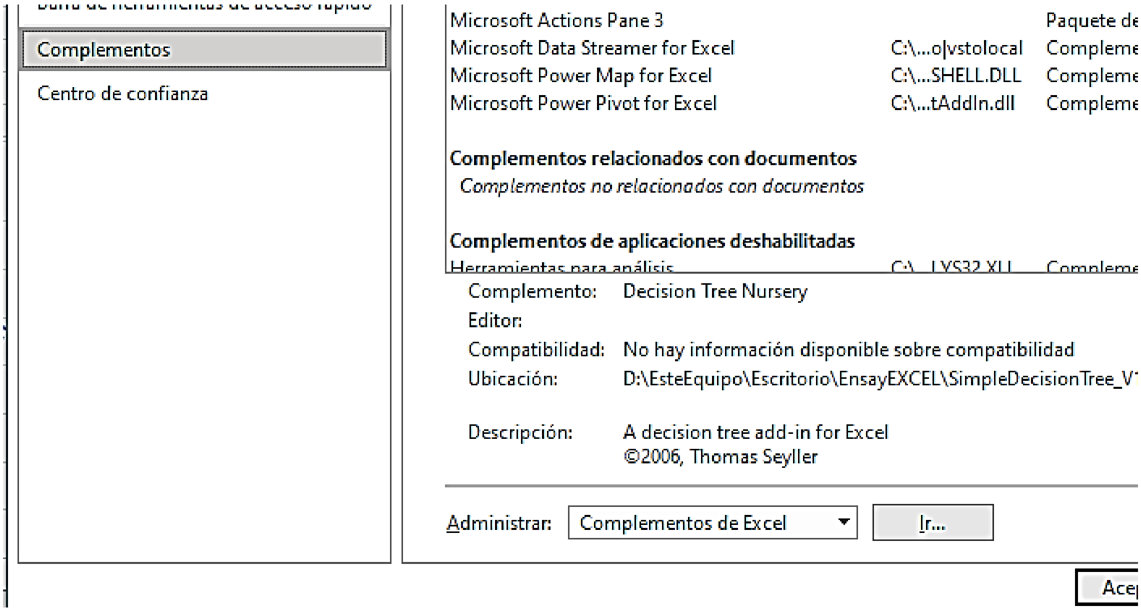
Desarrollo en Excel

Para calcular y construir árboles de decisión en nuestra PC tenemos varias alternativas que arrancan desde las opciones más profesionales como *Precision Tree* que tiene un costo de u$s 1.095, pasando por TreePlan con versiones gratuitas para estudiantes (versión 1.7), gratis por 30 días y Pro de u$s 59, hasta la más sencilla *SimpleDecision Tree 1.4* gratuita, que es la que analizamos acá.

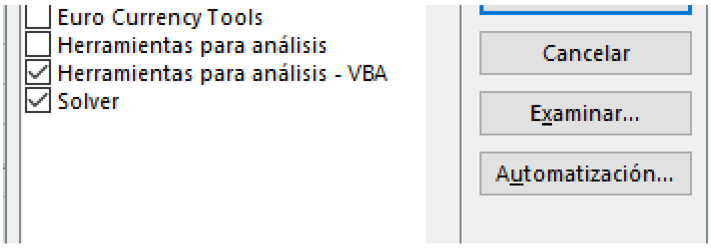
Es un complemento de Excel muy sencillo de usar, que, como dijimos, se descarga en forma gratuita (consulte en *optimiza.org* las direcciones de descarga). Se descarga como un archivo con extensión ".xla" que es la de los complementos de Excel. El archivo descargado debe instalarse y para eso abrimos Excel, libro nuevo, vamos al menú "Archivos", que nos despliega una lista de posibilidades de la que seleccionamos la inferior "Opciones". Con esto accedemos a la siguiente ventana.



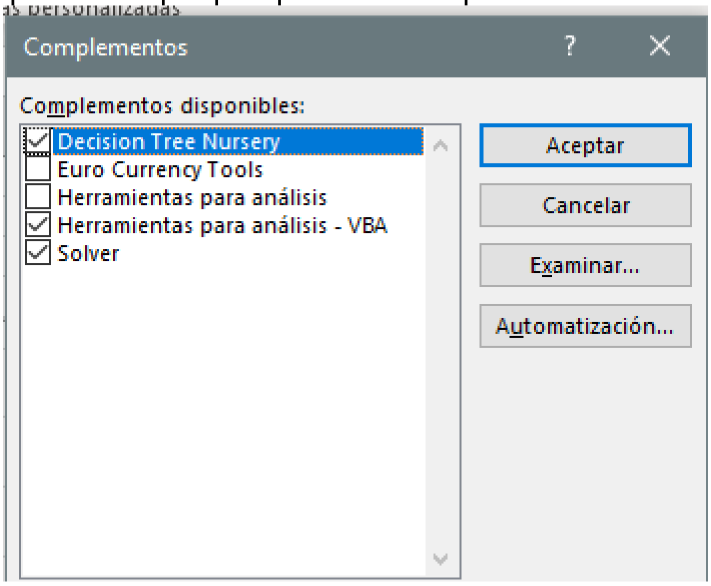
Seleccionaremos la alternativa "Complementos" y veremos que surge una nueva ventana en la cual buscamos el botón "Ir…" en el sector "Administrar [Complementos de Excel]"



Veremos la ventana de complementos instalados (entre ellos, Solver) que pueden habilitarse o deshabilitarse. Usando el botón "Examinar..." buscamos el archivo que hemos descargado en el primer paso y que tiene extensión ".xla"

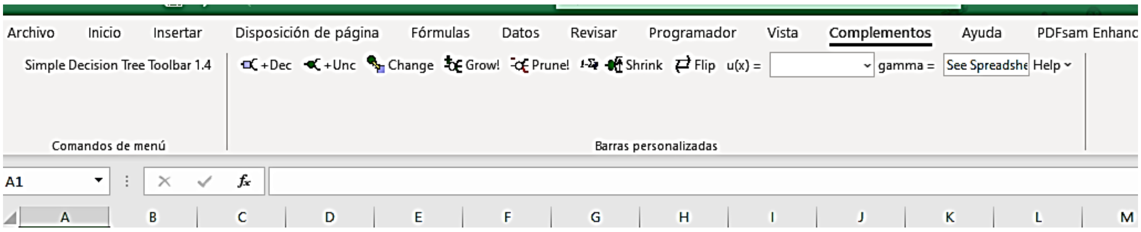


Una vez encontrado y pulsado el botón "Abrir", nos remite nuevamente a la lista de complementos, que cambió porque aparece el complemento "Decision Tree":



Aceptamos y ahora lo único que debemos hacer es reiniciar Excel, cerrándolo y abriéndolo nuevamente.

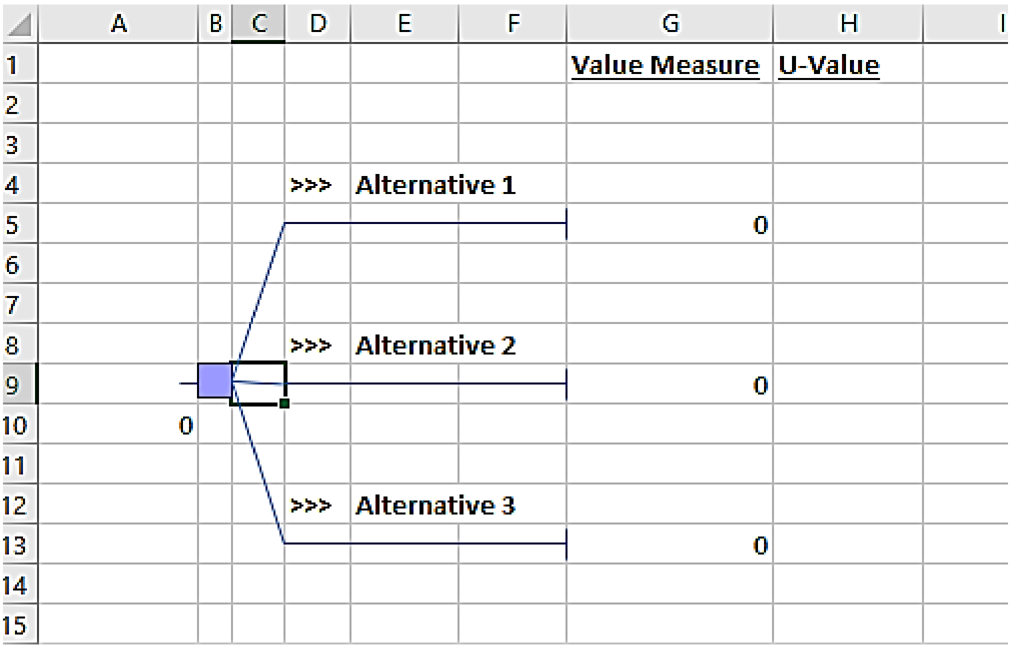
Cuando hemos reiniciado, por ejemplo en un libro nuevo, aparecerán en la pestaña “Complementos”, cuando la seleccionamos, estos componentes:



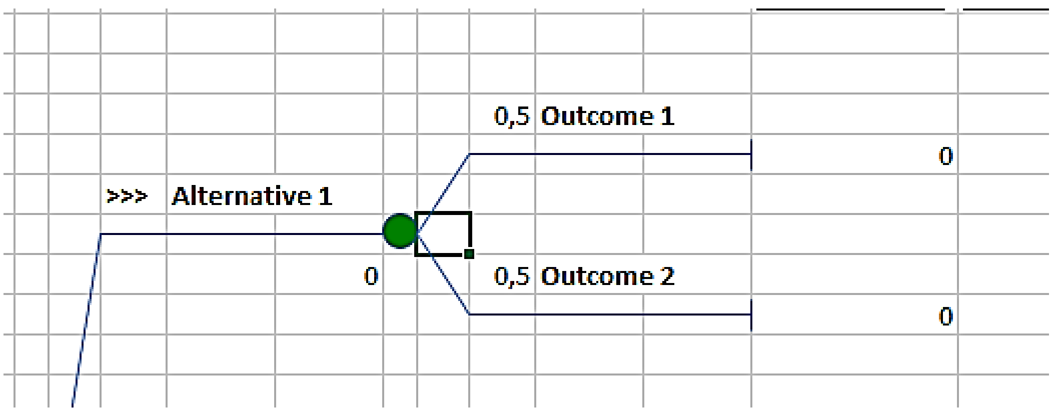
Como vemos en la figura son comandos aptos para construir árboles de decisión.

La utilización es sencilla. Para crear un nuevo nodo de decisión disponemos del botón “+Dec”

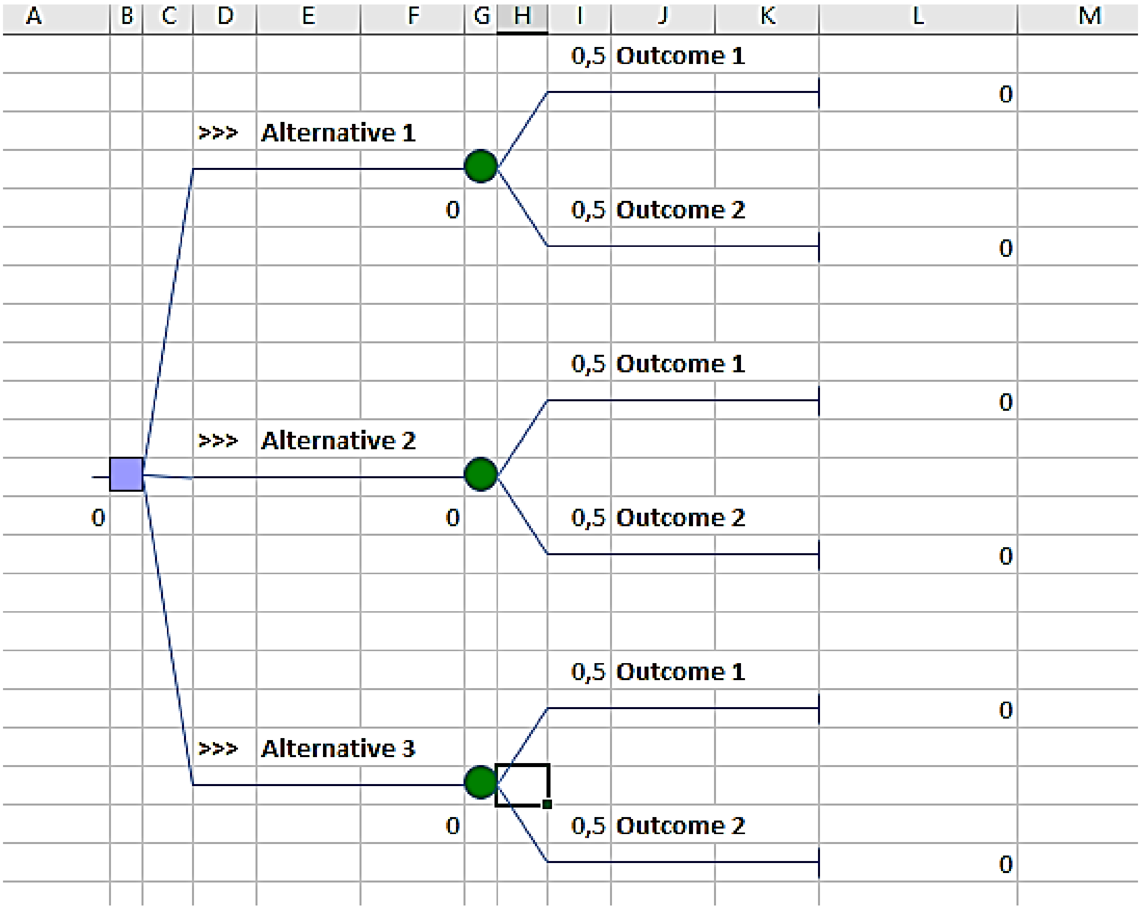
Si lo usamos nos aparecen dos alternativas, que en nuestro caso no nos sirve, ya que tenemos tres alternativas. Por tanto seleccionamos una celda lo más cerca posible del nodo de decisión y pulsamos el botón “Grow!”, que incrementará en una la cantidad de decisiones, así hasta la cantidad que necesitemos (en este caso una sola).



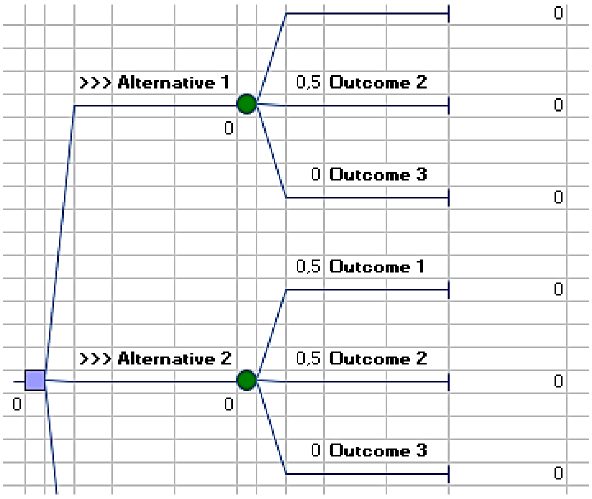
Ahora corresponde agregar los nodos de riesgo, usando el botón “+Unc” (abreviatura de incertidumbre en inglés), habiendo seleccionado la celda más próxima al fin de la alternativa que vamos a tratar.



Hacemos lo mismo en las demás alternativas.



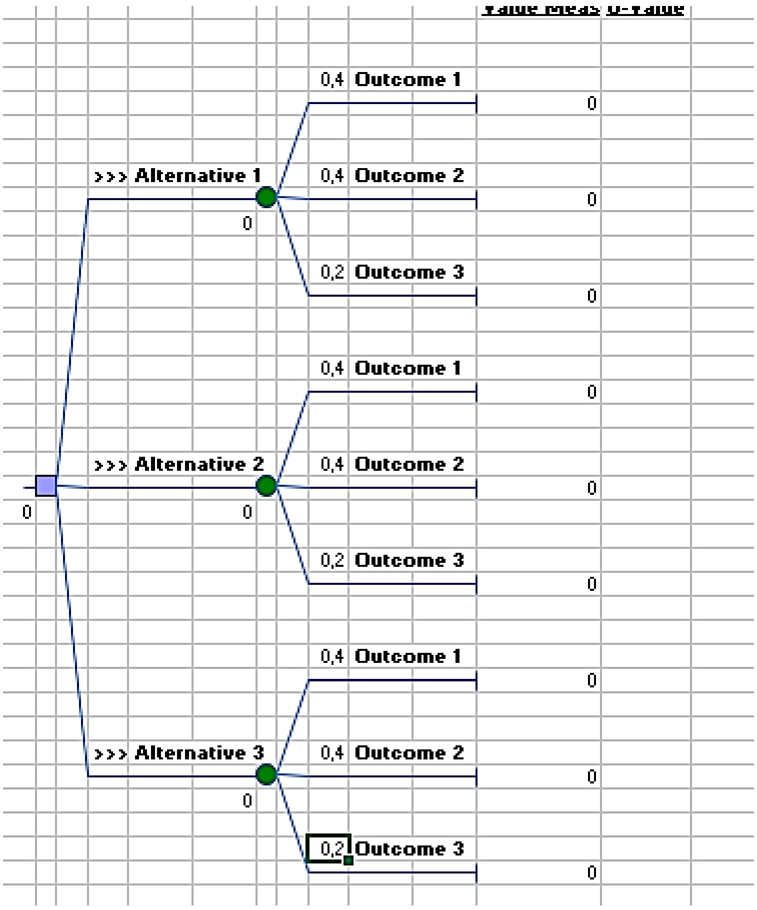
Como vemos, nos han quedado dos opciones de riesgo, y necesitamos en cada rama tres. Las agregamos seleccionando una vez más la celda más cercana al nodo y pulsando una vez más “Grow!”



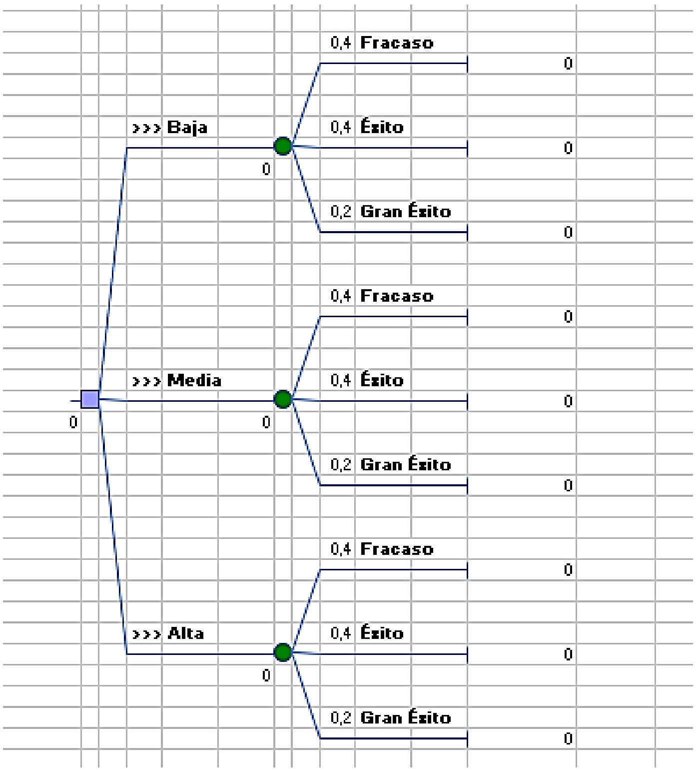
Corregimos las probabilidades por aparecieron por defecto en cada alternativa, simplemente seleccionando y escribiendo la que corresponda. La tercera alternativa de cada decisión la dejamos para ajuste. Quiere decir que, por ejemplo en la Alternativa 1, escribimos "0,4" en lugar del primer y del segundo "0,5". Seleccionamos la tercera, que dice "0" y en vez de escribir 0,2, pulsamos el botón

"1 – Σ"

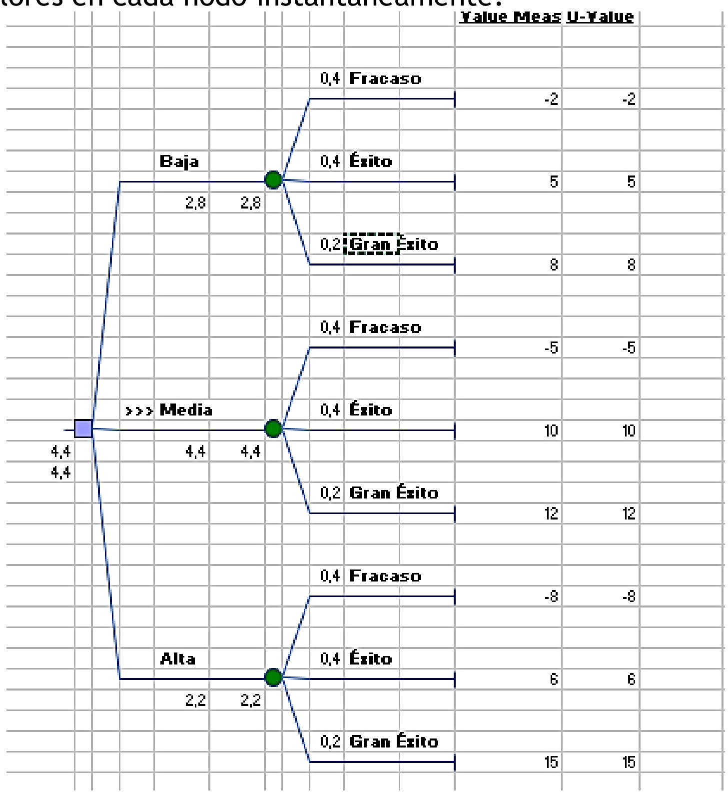
Esto nos asegura no cometer errores y que las probabilidades sean siempre excluyentes. Lo haremos en cada alternativa.



Ahora completamos el resto del árbol. Sustituimos “Alternativa x” por “Baja”, “Media” y “Alta” y “Salida x (outcome x)” por “Fracaso”, “Éxito” y “Gran Éxito”



Solo nos queda valorizar las terminales en la primera columna, con eso van a aparecer los valores en cada nodo instantáneamente.



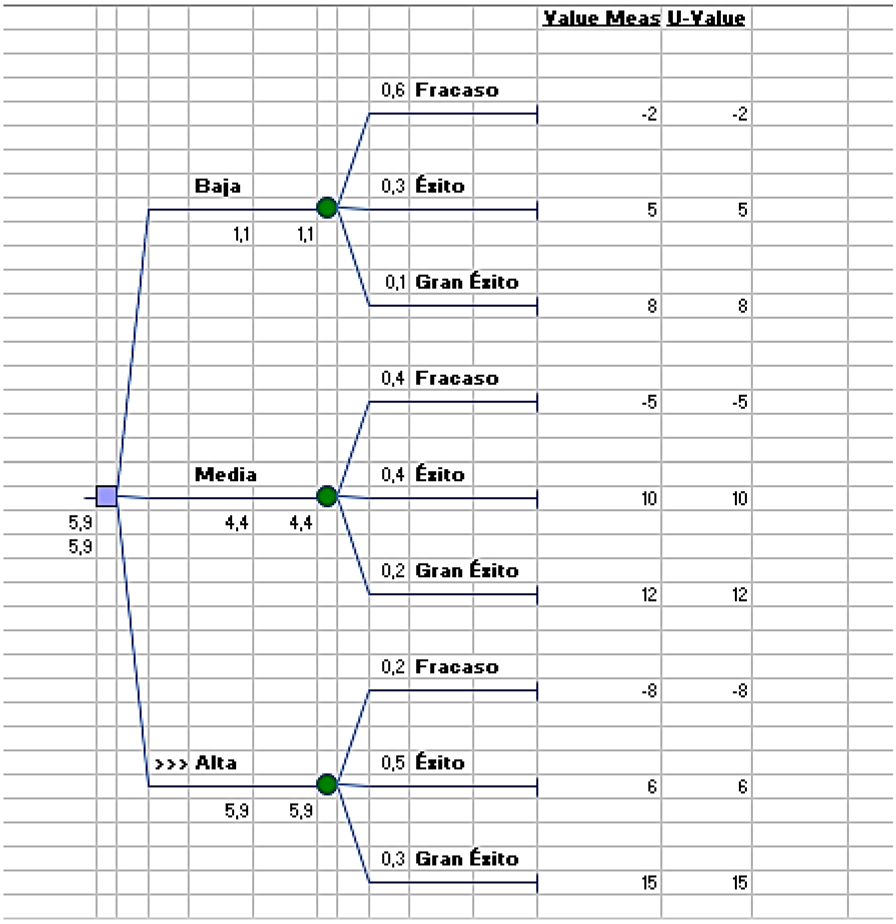
Se puede analizar el mismo problema desde otro punto de vista: suponer que si se realiza una inversión alta las probabilidades de éxito son mayores que si se realizara una baja inversión. En ese caso rediseñamos la tabla inicial con probabilidades que están relacionadas con el nivel de inversión:

	RESULTADOS		
DECISIONES↓	**FRACASO - F**	**ÉXITO - E**	**GRAN ÉXITO - G**
BAJA – B	0,6	0,3	0,1
MODERADA – M	0,4	0,4	0,2
ALTA – A	0,2	0,5	0,3

Ahora la ganancia esperada para el nodo 1 pasa a ser

$$-2 \times 0,6 + 5 \times 0,3 + 8 \times 0,1 = 1,1.$$

Lo único que haremos en la hoja de cálculo será cambiar las 4 probabilidades que corresponden a las alternativas 1 y 3, respectivamente (la tercera probabilidad de cada alternativa se va a ajustar sola) y mientras las cambiamos vemos que se ajustan todos los valores en los nodos:



En este caso vemos que lo mejor es la inversión alta.

Árboles Multinivel

Podemos trabajar en árboles de varios niveles, que nos permiten tomar decisiones en varios puntos del tiempo, de manera tal que el resultado de la decisión tomada en un momento dado influye en las decisiones que tomaremos en momentos posteriores.

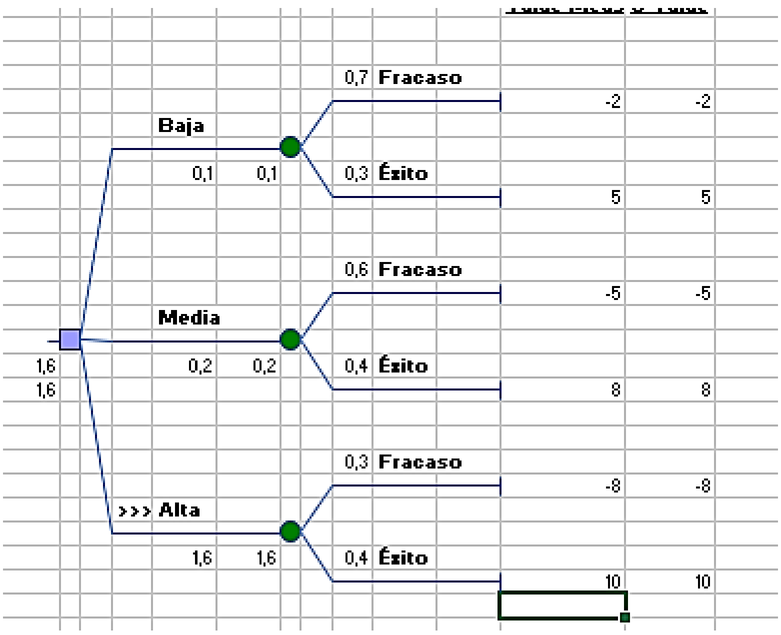
Para tomar un ejemplo vamos a replantear el problema anterior, suponiendo ahora que solo pueden existir dos resultados: Fracaso (F) o éxito (E), con las siguientes estimaciones de ganancias:

	Ganancias de los Resultados	
	Fracaso - F	**Éxito - E**
Baja - B	-2	5
Moderada - M	-5	8
Alta - A	-5	10

Con las siguientes probabilidades condicionadas para cada resultado:

	Probabilidades de los Resultados	
	Fracaso - F	*Éxito - E*
Baja - B	0,7	0,3
Moderada - M	0,6	0,4
Alta - A	0,3	0,7

El árbol de decisiones correspondiente lo obtenemos seleccionando la celda más próxima al fin de cada alternativa “Gran Éxito” y pulsando el botón “Prune!”. Luego corregimos la primera probabilidad en cada rama y ajustamos con sumatoria la otra, en cada alternativa. Por ultimo cambiamos los valores de la columna de la derecha. El árbol se reconfigura automáticamente:



Supongamos ahora que luego de lanzar el producto se esperan 8 semanas para tomar una segunda decisión sobre el resto de la inversión que usaremos con fines publicitarios para completar el año.

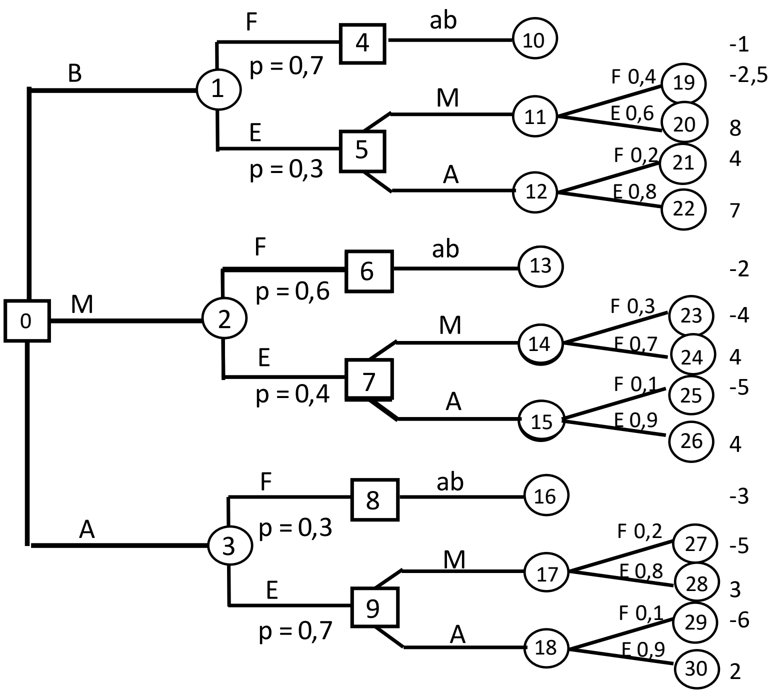
Esta segunda decisión depende de cuál fue la primera y de qué resultado se obtuvo al cabo de las ocho semanas, por ejemplo, si la primera decisión fue hacer una inversión alta, las alternativas para la segunda decisión serán:

Primera decisión		**Segunda Decisión**
Decisión Tomada	**Resultado**	**Alternativas**
A A	F E	Abandonar Seguir con inversión en publicidad Moderada M o Seguir con inversión Alta A

Esta tabla se completa cuando se agregan todas las posibilidades, quedando:

Primera Decisión		**Segunda Decisión**			
Decisión	**Resultado**	**Decisión**	**Resultado**	**Probabilidad**	**Ganancia**
Baja	Fracaso	Abandonar	-	-	1
Baja	Éxito	Seguir con M	Fracaso	0,4	-2,5
			Exito	0,6	8
		Seguir con A	Fracaso	0,2	-4
			Éxito	0,8	7
Moderada	Fracaso	Abandonar	-	-	-2
Moderada	Éxito	Seguir con M	Fracaso	0,3	-4
			Éxito	0,7	4
		Seguir con A	Fracaso	0,1	-5
			Éxito	0,9	4
Alta	Fracaso	Abandonar	-	-	-3
Alta	Éxito	Seguir con M	Fracaso	0,2	-5
			Éxito	0,8	3
		Seguir con A	Fracaso	0,1	-6
			Éxito	0,9	2

Con ella podemos construir un árbol de decisiones, que por ahora lo mostramos sin usar Excel:



El objetivo es calcular la ganancia en cada nodo como la ganancia óptima esperada de ahí en adelante. Sea por ejemplo el nodo 11, que es un nodo conectado con los nodos terminales del árbol, los nodos 19 y 20. También están conectados a nodos terminales los nodos 12, 14, 15 17 y 18.

La ganancia en el nodo 11 es cero, por lo que la ganancia esperada será:

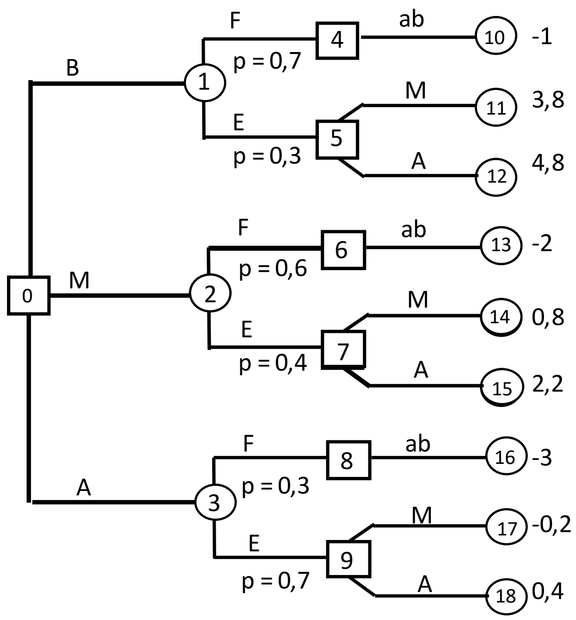
Ganancia total esperada en el nodo 11	=	Ganancia En el nodo 11	+	Ganancia Futura en el Nodo 11	=

Ganancia total esperada en el nodo 11	=	0	+	Probabilidad Del arco 11 – 19	x	Ganancia En el Nodo 19	+	Probabilidad Del arco 11 – 20	x	Ganancia En el Nodo 20

= 0 + 0,4 x -2,5 + 0,6 x 8 = 3,8

Este valor puede incluirse en el árbol y calcularse en forma similar para cada uno de los demás nodos.

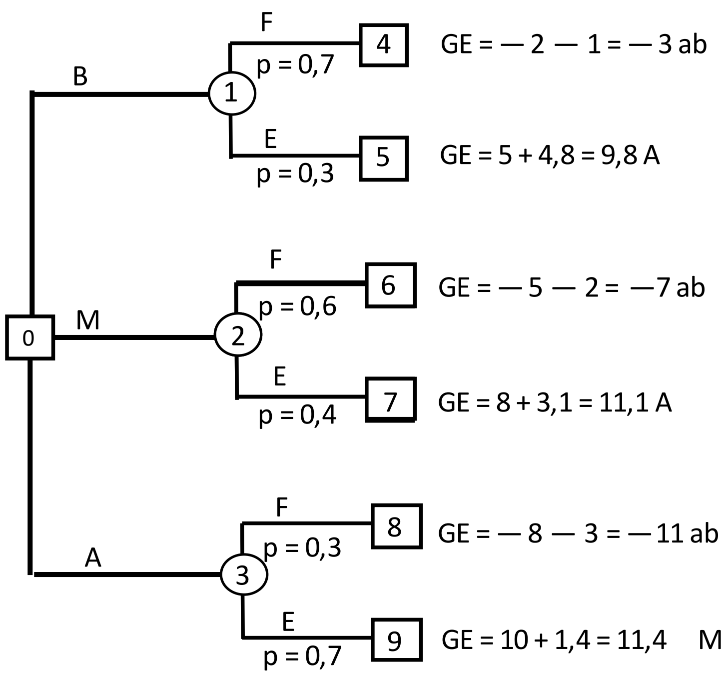
El árbol, en esta etapa, representa las GANANCIAS FUTURAS ESPERADAS (GFE)



Procediendo hacia atrás en el tiempo, en este punto es necesario calcular las ganancias totales esperadas en los nodos 4 a 9, utilizando las ganancias de los nodos 10 a 18. Por ejemplo, el nodo 5, que es un punto de decisión de segundo nivel, representa una inversión baja que dio un resultado exitoso.

En ese punto se pueden realizar cualquiera de las siguientes acciones: Invertir en publicidad moderadamente, con una ganancia esperada de 3,8 o invertir en un nivel alto, con una ganancia esperada de 4,8 (nodos 11 y 12, respectivamente) Se debería elegir el mejor resultado esperado, 4,8, el cual sumado a la ganancia propia del nodo, 5, arroja un total esperado de 9,8.

El conjunto de los nodos se analiza de la misma manera y queda el árbol que sigue:

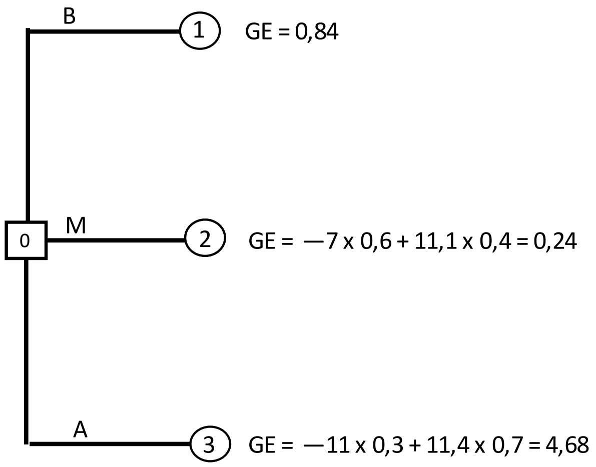


Seguimos retrocediendo en el tiempo y analizamos ahora los nodos 1 a 3. Por ejemplo, para el nodo 1, la ganancia futura esperada es:

Ganancia total Esperada en el Nodo 1	=	Ganancia en el nodo 1	+	Ganancia Futura en el Nodo 1	=

Ganancia total Esperada en el Nodo 1	=	0	+	Probabilidad Del arco 1 - 4	x	Ganancia En el Nodo 4	+	Probabilidad Del arco 1 - 5	x	Ganancia En el Nodo 5

GE N1 = 0 + 0,7 x (–3) + 0,3 x 9,8 = 0,84



El último paso será calcular el valor del nodo cero, que significa evaluar los nodos 1, 2 y 3:

Con lo cual el árbol queda:

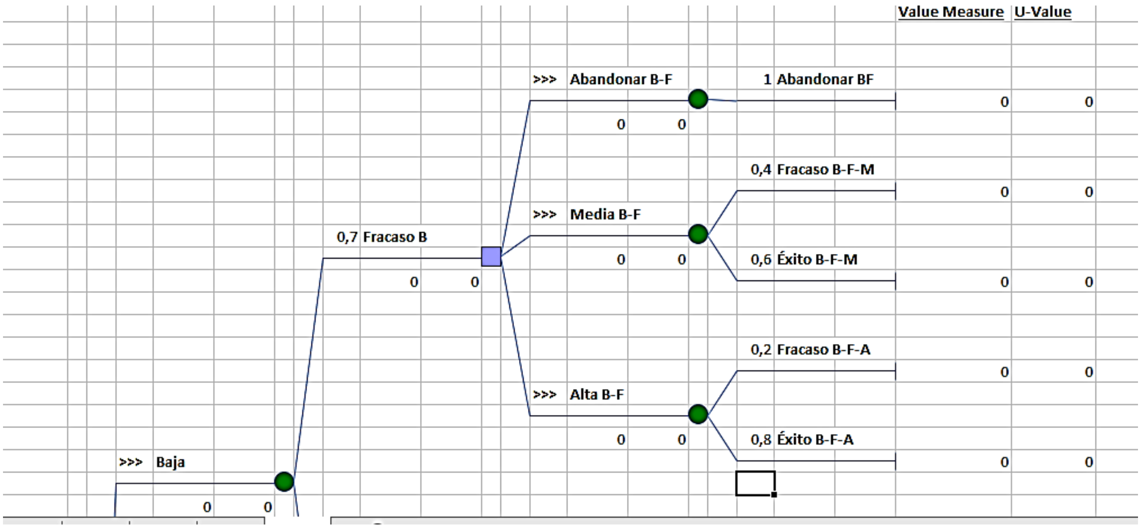
a) Inversión Baja, B, con una ganancia esperada de 0,84 ➔ Nodo 1
b) Inversión moderada, M, con una ganancia esperada de 0,24 ➔ Nodo 2
c) Inversión alta, A, con una ganancia esperada de 4,68 ➔ Nodo 3

Significa que la primera decisión, el primer nivel de decisiones, será la inversión alta, y conduce al nodo 3. En ese momento la nueva decisión se tomará en función de los resultados obtenidos, que conduce a los nodos 8 o 9: si fue un fracaso, ir al nodo 8, abandonar, con una ganancia de –11 (pérdida de 11). Si fue éxito, ir al nodo 9, inversión publicitaria moderada, con una ganancia de 11,4.

Resolución con Excel

Tendremos que construir el árbol con 31 nodos, de los cuales hay 15 que son nodos terminales.

Vemos la figura correspondiente a la primera rama. Parte del nodo de decisión 0 (de ahí parten las decisiones “Baja”, “Media” y “Alta”). Vemos como construimos la rama “Baja”:



Para agregar los nodos de alternativas (círculos, en la figura) usamos el botón "**+Unc**" y los de nodos de decisión el botón "**+Dec**".

Para agregar una alternativa o una chance más en cualquiera de esos nodos usamos el botón "**Grow!**".

Si nos equivocamos y queremos convertir un nodo de Decisión en uno de chance, o viceversa, pulsamos el botón "**Change**" (en todos los casos la celda seleccionada debe ser la más próxima al lugar donde vamos a producir el ingreso o cambio).

Para borrar un ramal o un nodo pulsamos el botón "**Prune!**", por ejemplo en el ramal *Abandonar*, agregamos el nodo de chance con una sola alternativa, pero se dibujan 2. Seleccionamos la celda al fin de un ramal y pulsamos "**Prune!**" y ese ramal desaparece, quedando solo uno.

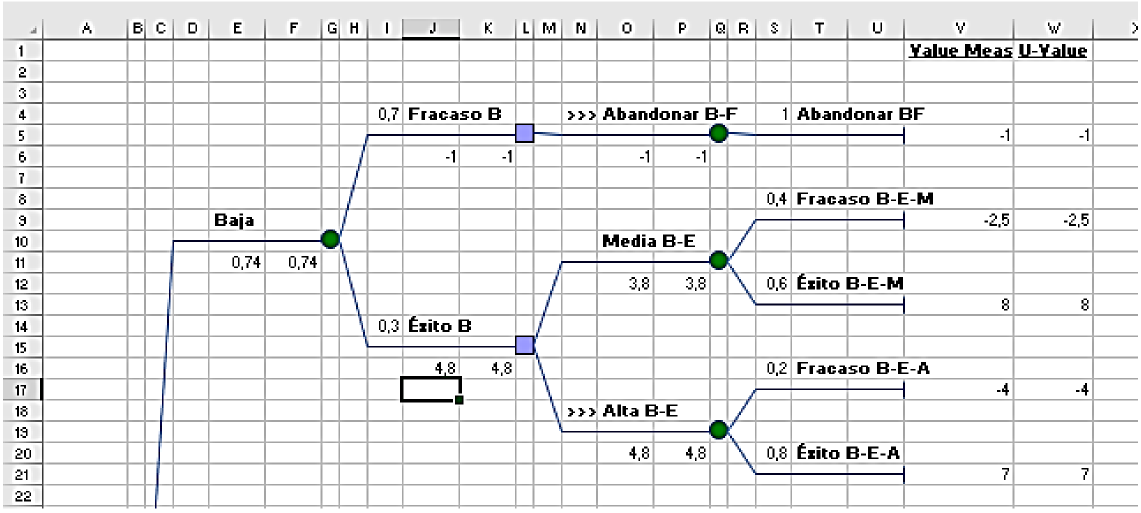
Hemos puesto los nombres de los ramales con el agregado de las letras para no confundir los caminos, así que el ramal Baja, sigue en todas las bifurcaciones con una letra B a la que se le agregan las iniciales de las bifurcaciones: La ruta Baja, Fracaso, Media, Éxito tendrá en su último tramo (Éxito) las letras B-F-M, así "Éxito B-F-M".

Las probabilidades deben ser agregadas en las etiquetas que no tienen el símbolo "**>>>**" y siempre reservando la última alternativa para usar el botón "**1-Σ**", incluso en el ramal "Abandonar" en la que hay una sola via, esa debe completarse con dicho botón.

De esta manera completamos las tres ramas de decisión y luego agregamos las probabilidades asociadas.

Ahora debemos hacer una operación un poco más complicada: agregar los beneficios de los nodos intermedios. Veremos como ejemplo solamente la primera decisión.

La figura siguiente nos muestra lo que hemos hecho hasta ahora.



El siguiente paso es completar celdas con la ganancia debajo del resultado parcial de la rama. En la figura de arriba es la celda J17, la que está señalada. La otra de esa rama es la J7. Allí escribiremos las ganancias de ese nodo (J7=—2; J17=5). Pero antes vamos a modificar el cálculo.

Vemos que la celda K6 tiene la función [=MAX($P$6)]. La vamos a convertir en [=J7+MAX($P$6)]
La celda K16, que tiene la función [=MAX($P$12:$P$20)], pasará a tener [=J17+MAX($P$12:$P$20)]
ATENCION. Estas direcciones son ejemplos. Hay que revisar las direcciones reales en cada caso.

Y así completamos las otras tres ramas. El aspecto debería ser el de la figura de la página siguiente.

Toma de decisión determinística

En estos casos lo usual es trabajar a partir de uno de los modelos presentados en el capítulo anterior como, por ejemplo, usando programación lineal que es una de las técnicas más empleadas.

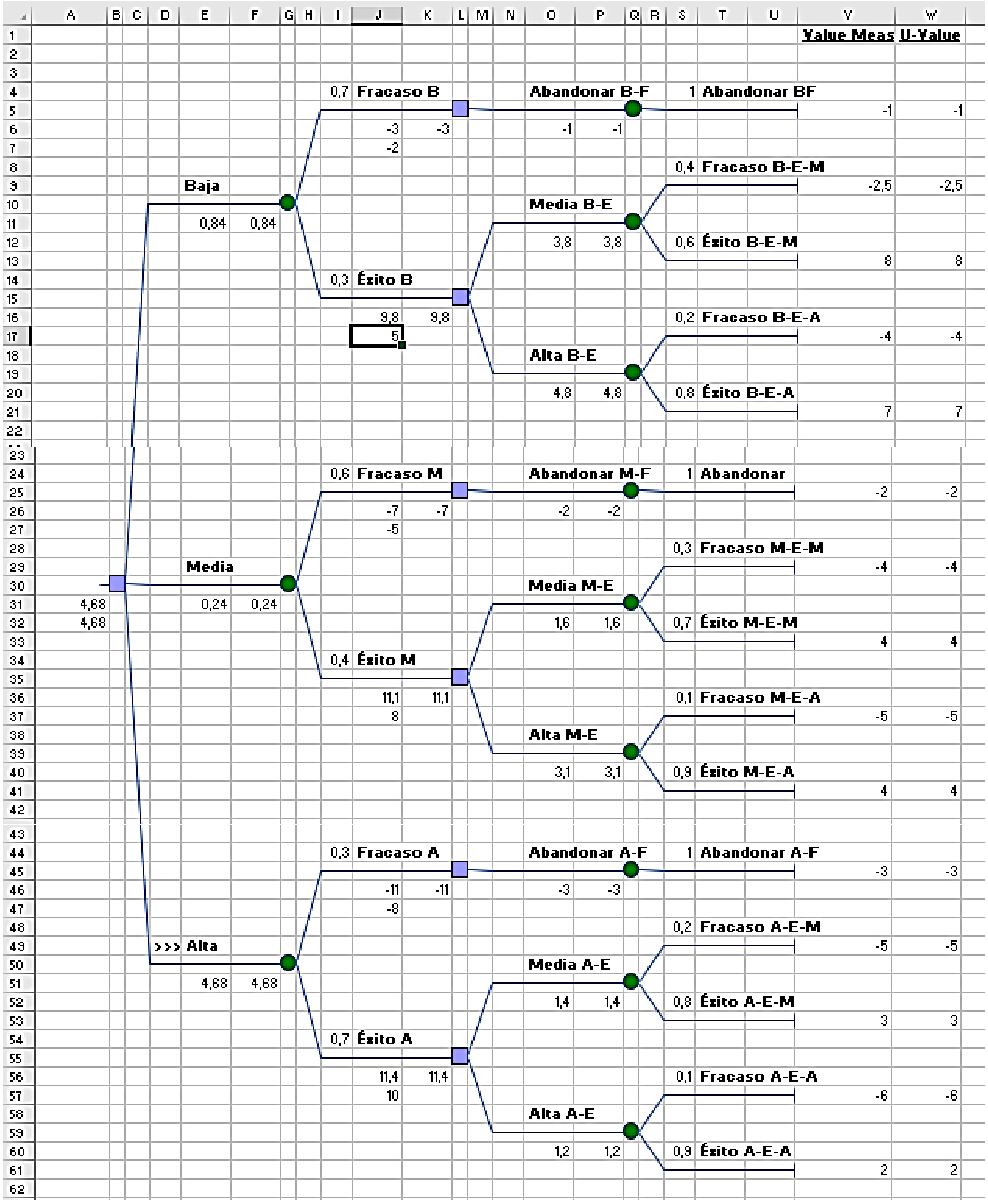
Para estos modelos hay métodos de resolución heurísticos y analíticos, que describiremos en esta sección.

Método heurístico – Modelo Gráfico

Vamos a ver como se usa este método en base a un ejemplo. Supongamos que tenemos el siguiente caso:

Planeamiento de la producción en TFH.

Se pretenden producir dos tipos de equipos para oficina: el AVANZADO y el BÁSICO. Ambos equipos pasan por el sector ARMADO, que cuenta con 31 horas/semanales, que se usan a razón de 2 hs por cada AVANZADO y 1 h por cada BÁSICO
Una vez armado el equipo pasa al sector TERMINACIÓN, (carga, encendido y empaque), donde se le instala el SO y *float ware*. Allí se disponen de 35 horas semanales que se usan a razón de 1 h por cada AVANZADO y 2 hs por cada BÁSICO.
No se pueden vender más de 15 equipos semanales del modelo BASICO, no hay límite para el AVANZADO. Las ganancias que aportan son 260 $ el equipo AVANZADO y $ 410 el BÁSICO.
Planificar la producción semanal con el objetivo de maximizar las ganancias.



El modelo hallado con las técnicas discutidas en el capítulo 2 tiene dos variables de decisión (A que representa la cantidad de equipos AVANZADOS a producir y B que representa la cantidad de equipos BÁSICOS) con tres restricciones, a saber: cantidad de recursos de armado expresada en horas semanales, cantidad de recursos en terminado, también en horas semanales y capacidad de colocación de equipos Básicos semanales.

$$Z = 260\ A + 410\ B\ \equiv max$$

$sujeto\ a$

$$\begin{aligned} &2A + 1B \leq 31, \qquad sector\ armado \\ &1A + 2B\ \leq 35, \qquad sector\ terminado \\ &0A + 1B \leq 15, \qquad demanda \\ &A \geq 0\ y\ B \geq 0 \end{aligned}$$

La decisión que hay que tomar es cuánto producir de cada equipo en función de las restricciones existentes. En este caso es una decisión determinística, bajo condiciones de certidumbre.

Este planteo es un caso tipo de Programación Lineal. Existen varios métodos para resolverlo e interpretarlo, pero su desarrollo excede los alcances de este texto. Para ver un desarrollo muy completo, incluyendo solución mediante software consulte el Capítulo 2 de Optimiza — Libro 1.

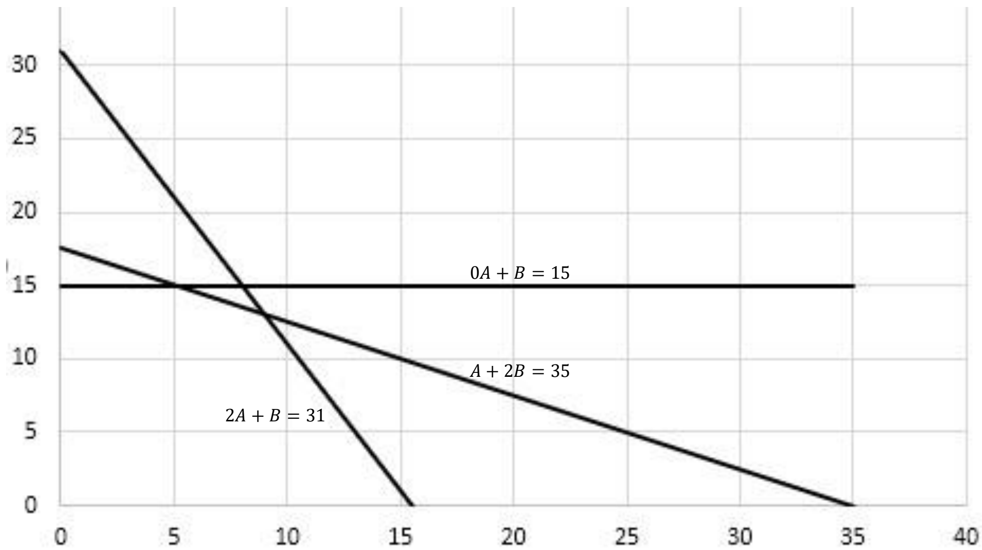
Acá veremos solamente una primera aproximación a la resolución mediante el denominado método gráfico.

Programación Lineal. Método gráfico

Comenzamos por definir un par de ejes de coordenadas cartesianas ortogonales, donde *B* será el eje de las ordenadas y *A* el de las absisas.

Luego vamos a considerar una restricción por vez. Para graficar cada restricción comenzamos ignorando el signo de la desigualdad y tratando la función como si fuera una igualdad. Luego incluiremos la desigualdad.

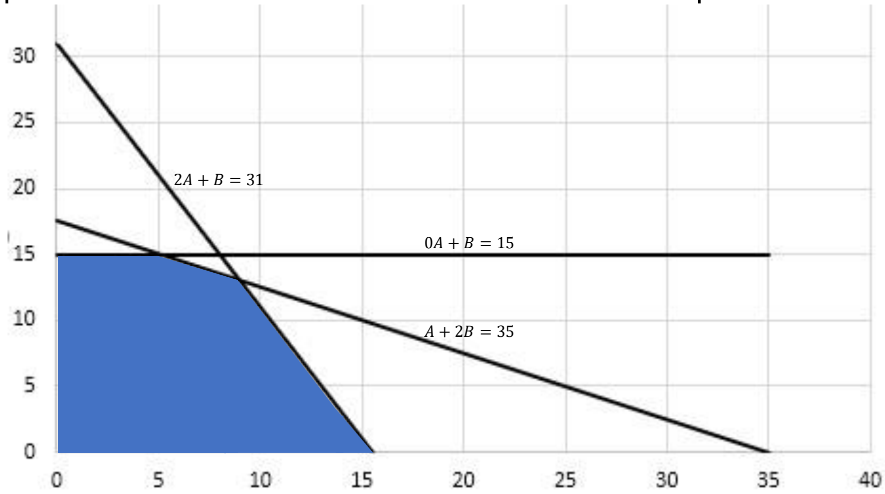
Las últimas dos restricciones coinciden con los ejes de coordenadas y, si tenemos en cuenta el sentido de la desigualdad, determinan como espacio factible solamente al primer cuadrante. Si graficamos ahora todas las restricciones como igualdades tendremos un gráfico similar al que vemos en la figura siguiente:



En el siguiente paso determinaremos, para cada restricción, cuáles son los valores factibles: aquellos valores permitidos por la inecuación. Dicho de otra manera, vamos ahora a tener en cuenta las desigualdades.

Para la primera restricción son aquellos que cumplen con $2A + B \leq 31$, lo cual significa que la recta graficada ($2A + B = 31$), es uno de los conjuntos de valores factibles, pero también los son todos los puntos ubicados a la izquierda de dicha recta.

Una vez que fueron consideradas **todas** las restricciones queda determinada un *área factible* que es la intersección de las sub áreas determinadas por cada restricción:



Ese polígono encierra una superficie que contempla exactamente todas las restricciones: el eje B a la derecha es la última restricción, el eje A hacia arriba es la anteúltima y así sucesivamente. Significa que cualquier solución al problema debe ubicarse forzosamente dentro del área sombreada o en sus límites, ya que este conjunto determina lo que es admitido por las restricciones.

Resulta evidente que pueden graficarse para cada valor de B y de A una curva

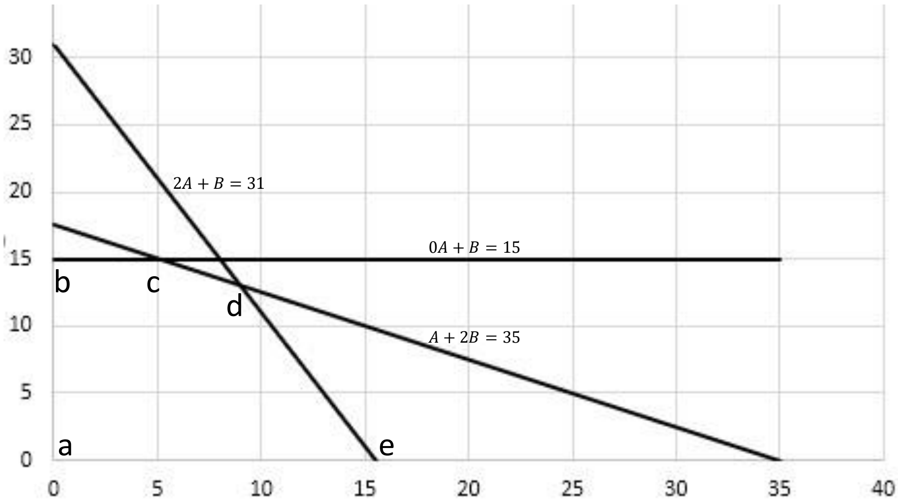
$$260A_1 + 410B,$$

a pendiente constante e igual a $-260/410$ ($=-26/41 = -0{,}6341$) y a ordenada al origen variable proporcional a Z.

Obviamente será la solución óptima aquella que pase por alguno de los puntos extremos del área sombreada, entendiendo como “extremo” a aquel punto que se encuentre lo más alejado posible del origen.

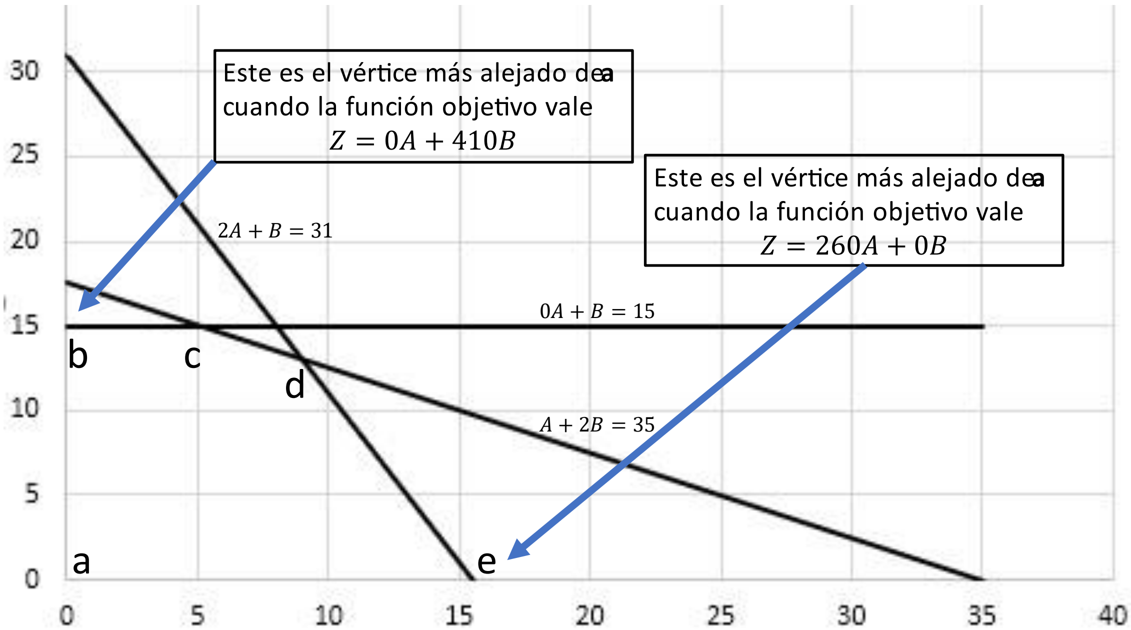
¿De qué depende que un vértice sea el más “alejado”. Obviamente de la pendiente de la función objetivo.

Los puntos extremos – o más alejados del origen – más el mismo origen son los vértices de ese polígono y se designan con las letras a, b, c, d y e.



Si el equipo A no nos diera ninguna ganancia o muy poca en relación al equipo B, la pendiente en vez de ser $-26/41$ podría ser $-0/41 = 0$ y una recta con pendiente nula nos haría ir directamente al punto ***b*** como el más alejado de ***a***. Espe punto significa no fabricar nada del equipo B y fabricar todo lo que puedo del A.

Al contrario, si la ganancia en vez de 260 fuera enorme, desproporcionada en relación a los 410 $ del Básico, la pendiente tendería a infinito, sería una recta vertical, y el punto más alejado de ***a*** sería el ***e*** (no fabricar nada de Básico y todo de Avanzado)



Observando la figura se ve que si la pendiente de Z fuera nula, la recta sería horizontal y el eje más lejano sería el “b”. En ese caso la ganancia de la Avanzada sería cero, entonces ¿Por qué fabricarla?

Por otro lado, si la ganancia de Básico pasara a un valor nulo, la curva Z tendría pendiente infinita, sería vertical y entonces el punto más alejando sería el vértice “e” (Obvio: si no se gana nada por producir Básica, entonces hay que hacer todo Avanzada)

Esto es lógico: si el origen significa “no hacer nada”, entonces, para maximizar las ganancias hay que fabricar lo más posible ya que eso permite lograr el objetivo. Cada uno de los vértices es una solución posible (en el sentido que está permitida) y son parte de un polígono.

Para analizar el caso actual con la pendiente (—26/41) no hace falta graficar la función objetivo. Basta con estudiar el valor de la función en cada uno de los vértices. Los valores de estos vértices para las coordenadas *B* y *A*, y el valor de la función objetivo son:

Vértice	A	B	Función Objetivo (Z)
a	0	0	0
b	0	15	6150
c	5	15	7450
d	9	13	7670
e	0	15,5	6355

Como el problema es maximizante seleccionaremos aquel vértice que corresponde al mayor valor de la función objetivo. El vértice “d” presenta una respuesta máxima de la función objetivo de 7670, indica, por tanto, que lo óptimo es programar una producción de 9 y 13 Avanzadas y Básicas, respectivamente.

La solución de este problema se encuentra resolviendo el sistema de ecuaciones que corresponde a las dos restricciones involucradas en cada vértice, con lo que se obtienen las coordenadas del mismo. Particularmente en el caso del “d” el sistema sería

$$\begin{cases} 2A + B = 35 \\ A + 2B = 31 \end{cases}$$

Para encontrar el valor de cada uno de los vértices, entonces, solo debemos ver en el gráfico cuales son las rectas que se cruzan en él y buscar la raíz del sistema de ecuaciones que configuran. Así:

Vértice a:

$$\begin{cases} A = 0 \\ B = 0 \end{cases}$$

Vértice b:

$$\begin{cases} A = 15 \\ B = 0 \end{cases}$$

Vértice c:

$$\begin{cases} A = 15 \\ A = 35 - 2B \end{cases}$$

Vértice d:

$$\begin{cases} A = 35 - 2B \\ A = \dfrac{31}{2} - \dfrac{1}{2}B \end{cases}$$

Vértice e:

$$\begin{cases} A = \dfrac{31}{2} - \dfrac{1}{2}B \\ A = 0 \end{cases}$$

Finalmente, la expresión del resultado del modelo es:

Producción semanal de AVANZADA: 9
Producción semanal de BASICA: 13
Ganancia semanal: Z = 260 x 9 + 410 x 13 = $ 7670

Método Analítico – Método SIMPLEX DANTZSIG

El método comienza por reelaborar el modelo original:

$$z = 260\,A + 410\,B \equiv max$$

Sujeto a:

$$2A + 1B \leq 35, disponibilidad\ en\ armado$$
$$1A + 2B \leq 31, disponibilidad\ en\ terminado$$
$$0A + 1B \leq 15, demanda\ de\ equipos\ Básicos$$
$$A \geq 0; B \geq 0, no\ puede\ haber\ equiposnegativos$$

A un sistema que no tenga desigualdades, para lo cual se crean “variables de ajuste”, o variables que indican, para cada valor de A y de B cuanto falta para llegar al valor del lado derecho.

Por lo tanto deben crearse tres variables de ajuste, S1, S2 y S3, una para cada restricción, con lo cual el modelo anterior se convierte al siguiente:

$$Z = 260 + 410 + 0S_1 + 0S_2 + 0S_3 \equiv max$$

Sujeto a:

$$2A + 1B + 1S_1 + 0S_2 + 0S_3 = 35$$
$$1A + 2B + 0S_1 + 1S_2 + 0S_3 = 31$$
$$0A + 1B + 0S_1 + 0S_2 + 1S_3 = 15$$
$$A \geq 0; B \geq 0$$

Esto nos permite plantear un sistema de tres ecuaciones (restricciones) con cinco incógnitas (variables) que se puede resolver poniendo dos de las incógnitas en cero en todas las combinaciones posibles.

Como el número de operaciones a realizar en ese caso sería muy grande, Dantzig propuso un método evolucionario, muy simple, basado en reducción de matrices. No vamos a describirlo en este Libro, pero pude ser consultado en Optimiza — Libro1, donde está detallado paso a paso.

CAPÍTULO 4

INTRODUCCIÓN A SIMULACIÓN

Comenzamos viendo conceptualmente el término “simulación” para así encarar los aspectos relevantes de esta disciplina que, en la actualidad, se vincula profundamente con la informática.

Un buen punto de partida para describir el significado de *simulación* es la definición que nos aporta R.E. Shannon[45]

"*La simulación es el proceso de diseñar un modelo de un sistema real y llevar a término experiencias con él, con la finalidad de comprender el comportamiento del sistema o evaluar nuevas estrategias dentro de los límites impuestos por un cierto criterio — o un conjunto de ellos — para el funcionamiento del sistema*".

Como vemos, se trata de un sistema usado como modelo que representa a otro sistema y que tiene, como modelo, criterio y límites. Podemos encontrar en nuestro entorno varias derivaciones de esta definición: los juegos, por ejemplo, permiten simular estrategias con diversos grados de complejidad. Los deportes extremos o juegos de supervivencia que a veces se ven en TV, los juegos de guerra en los colegios militares, o la teoría de juegos aplicada en toma de decisiones discutida en otros capítulos, los túneles de viento en fábricas de aviones o de coches de fórmula 1, los simuladores de partidos de fútbol, y un largo etc.

También podemos usar el concepto aplicado a equipos que permiten analizar comportamientos en situaciones de alto riesgo sin poner en peligro la vida o la propiedad: los simuladores (entrenadores) de vuelo, de conducción o manejo de equipos, de entrenamiento (como la bicicleta fija), etc.

Por último se pueden describir los simuladores de realidades mucho más complejas, como los simuladores de desarrollos sociales (un ejemplo comercial y muy conocido es el software, que se vende como un juego para PC que es capaz de simular un desarrollo social cambiante ante hechos o decisiones determinadas).

---

[45] **Claude Elwood Shannon** (1916 - 2001), ingeniero electrónico y matemático estadounidense, se conoce como el padre de la teoría de la información. El número de Shanonn (10^{120}) indica el número posible de posiciones legales en una partida de ajedrez. Como medida de comparación se supone que el número de átomos en el universo es del orden de 10^{72} hasta 10^{87} átomos, sin considerar la materia oscura. Como curiosidad, en 1938, el matemático Kasner definió un número inimaginablemente grande, al que su sobrino bautizó googol (gúgol, en español), que es 70! (10^{100}), 333 bits, del que toma su nombre, con un cambio en el tipeo mediante, una famosa empresa dedicada a informática.

La posibilidad de realizar este tipo de modelos ha dado origen a varios abordajes teóricos, a conceptos básicos y a profundizaciones de la teoría de sistemas.

Para hacer un análisis conceptual de la simulación vamos a necesitar clasificar los sistemas en tres grupos basándonos en el punto de vista de su comportamiento global en el tiempo:

- ***Sistemas estables***: aquellos que tienden, en el tiempo, a una posición puntual o espacial (órbita): estos sistemas reconocen un atractor. Ej.: péndulo, sistema solar
- ***Sistemas inestables:*** aquellos que no reconocen o escapan de los atractores. Ej. Reacción en cadena
- ***Sistemas caóticos***: en los que se encuentran los dos comportamientos: tiene elementos que responden a un atractor y hay otros que lo alejan de él.

Una de las mayores características de un sistema **inestable** es que tiene una gran independencia de las condiciones iniciales. Si estudiamos un sistema del que se conocen sus ecuaciones características, y con unas condiciones iniciales fijas, entonces podremos conocer exactamente su evolución en el tiempo. En el caso de los sistemas **caóticos**, una mínima diferencia en esas condiciones, aunque sea en una sola de ellas, hace que el sistema evolucione de manera totalmente distinta. Ejemplos de tales sistemas incluyen la atmósfera terrestre, los fluidos en régimen turbulento y los crecimientos de población.

Sin embargo debemos tener cuidado con este tipo de clasificación porque no siempre resulta tan sencillo establecer estas propiedades. Veamos unos ejemplos para explicar mejor lo que queremos decir:

El sistema solar es, claramente, un sistema estable: no tenemos dudas que su comportamiento se refiere a atractores e inmutabilidad a lo largo del tiempo. Sin embargo se podría argumentar que debe de haber habido un comienzo y que esto se ve reflejado en haber comprendido que el combustible que alimenta las reacciones que se producen el sol en algún momento se agotará, lo que sugiere que en algún momento esas reacciones comenzaron. Además que hay religiones y tradiciones humanas que establecen que existió un Creador o, al menos, un comienzo.

El modelo que describe el Universo, del cual el sistema solar forma parte, establece, a partir de trabajos de Hubble[46] y otros a principios del Siglo XX, que, efectivamente, el Universo debería tener un origen ya que actualmente se observa que está en expansión, lo que significa que, retrospectivamente, esa expansión comenzó en un punto, y ese es otro indicio de comienzo semejante al del combustible solar.

Este modelo requiere describir las condiciones iniciales, como dijimos más arriba, y ahí es donde pueden surgir dudas. El instante inicial podría haber sido perfectamente determinado o cualquier otro, con lo cual el modelo será perfectamente inestable o caótico. Las leyes de la naturaleza, que son modelos que nos explican el sistema

---

[46] Edwin Powell Hubble (1889 — 1953) fue uno de los más importantes astrónomos del siglo XX, principalmente por haber demostrado en 1929 la expansión del universo midiendo el corrimiento al rojo de galaxias distantes. Aunque comenzó estudiando abogacía en Oxford, en el Observatorio Yerkes de la Universidad de Chicago, obtuvo el doctorado en física en 1917. Al volver de la Primera Guerra Mundial, en 1919, trabajó en el entonces nuevo observatorio del Monte Wilson, con un telescopio de 254 centímetros, que en esa época era el más potente del mundo. En el observatorio, trabajó junto a Milton Humason. Actualmente lleva su nombre como homenaje el telescopio espacial Hubble.

funcionan bien y de acuerdo con lo que observamos, pero no funcionan de ninguna manera ni en el instante inicial ni en instantes anteriores a él, sea lo que fuere que signifique “instante anterior”, ya que, según nuestras propias leyes y observaciones, el tiempo se inició junto con el universo y, por tanto, no puede haber tiempo anterior a él.

Por tanto acá es necesario simplificar aún más el modelo, diciendo que lo podemos utilizar si al usarlo ponemos la premisa de no definir las condiciones iniciales, por lo cual pasa a ser un tipo especial de modelo estable[47].

Por otro lado están los modelos caóticos. Para ellos podemos usar otro ejemplo que surge en 1963, cuando *Lorenz*[48] propone un modelo meteorológico basado en tres ecuaciones diferenciales bien definidas. Trabajaba sobre la idea de un que se podía crear un modelo de sistema inestable que funcionaría si conociéramos las condiciones iniciales y que con él se podría predecir el clima.

Sin embargo, se descubrió que el sistema atmosférico es un sistema caótico porque es imposible conocer con exactitud todos los parámetros que constituyen las condiciones iniciales (en cualquier sistema de medición, por definición, siempre se comete un error, por pequeño que éste sea y muchas veces originado por el mismo sistema de medición[49]) hace que, aunque se conozca el modelo, éste se aleja de la realidad pasado un cierto tiempo. Por otra parte, el modelo atmosférico es teórico y puede no ser perfecto, y el determinismo en el que se basa, es también teórico.

Volvamos a la hipótesis del *big bang*. Si aceptáramos la teoría que el universo estaba concentrado en un punto perfectamente ordenado de energía infinita y ninguna ley física vigente en el “momento” (entrecomillado porque no existía el concepto de tiempo, según acabamos de postular) en que se produce el estallido inicial. Menos de una millonésima de nano segundo después rigen las leyes de la naturaleza y el universo comienza a expandirse. Las mismas leyes de la naturaleza, ya vigentes, predicen que esa expansión será homogénea en todas direcciones, porque las condiciones iniciales eran perfectamente homogéneas. Sin embargo observamos que hay galaxias, materia oscura y espacios aparentemente vacíos, lo que no habla de homogeneidad inicial. Por lo tanto el sistema sería un sistema caótico ya que no podemos definir condiciones iniciales.

Estas consideraciones muy simples[50] llevarían a suponer que cuando decimos *Caos* hablamos de una teoría. No necesariamente es una teoría sino que puede entenderse

---

[47] Esta afirmación de que un “tipo especial” de modelo estable corre por cuenta del autor de estas líneas y debe tratarse, a su vez, como una licencia didáctica. No hay literatura que la avale, entre otras cosas porque hay evidencia que el ritmo o velocidad de expansión aumenta y no es constante.

[48] Edward Norton Lorenz (1917- 2008) fue un matemático y meteorólogo estadounidense, desarrolló ideas innovadoras sobre la rotación de los fluidos y contribuciones en dinámica atmosférica y predicciones climatológicas. Fue pionero en el desarrollo de la teoría del caos. Fue quien introdujo el concepto de atractores extraños y acuñó el término efecto mariposa. (Fte. Wikipedia, abril 2020)

[49] Por ejemplo, un termómetro de mercurio sumergido en agua a 90° siempre va a enfriar el agua y calentar el termómetro hasta que ambas temperaturas estén en equilibrio. El hecho de que la variación de temperatura pueda ser despreciable no significa que no existe.

[50] Debemos tomarlas como ejemplos muy simplificados. En realidad el estudio del tiempo y origen del universo es un tema de muchísima más complejidad que este ejemplo. Para tener una explicación más profunda se recomiendan dos textos del físico inglés Stephen Hawkins: “Breve historia del tiempo” (A Brief History of Time: From the Big Bang to Black Holes, Bantam Books, RU, 1991) y “El Universo en una

como un campo de investigación, que abarca diferentes líneas de pensamiento. Lorenz señalaba que el término *Caos* adquiere el significado de "*ausencia de determinado orden*"

El caos está entendido como cierto tipo de orden de características impredecibles, pero descriptibles en forma concreta y precisa. Es decir: un tipo de orden de movimiento impredecible.

La idea de la que parte la Teoría del Caos es simple: en determinados sistemas naturales, pequeños cambios en las condiciones iniciales conducen a grandes discrepancias en los resultados

En conclusión, podemos definir como "Sistema caótico" a aquel sistema que es sensiblemente dependiente de cambios en las condiciones iniciales.

"Una propiedad esencial del comportamiento caótico es que estados próximos entre sí terminan por divergir sin importar lo pequeñas que puedan ser las diferencias iniciales"... "La definición de equilibrio inestable tiene mucho en común con la dependencia sensible: ambas suponen la amplificación de diferencias inicialmente pequeñas".[51]

De esta manera en un sistema con dependencia sensible, será imposible poder realizar predicciones perfectas. Un ejemplo es el pronóstico meteorológico donde los resultados no son precisos y pueden variar de un momento a otro. Los sistemas dinámicos son capaces de entrar y salir del caos.

Los atractores extraños, según Lorenz, pueden ser un conjunto de estados:

"Un conjunto de estados de un sistema que se dan una y otra vez, o que son aproximadamente los mismos una y otra vez, cada vez más próximos entre sí

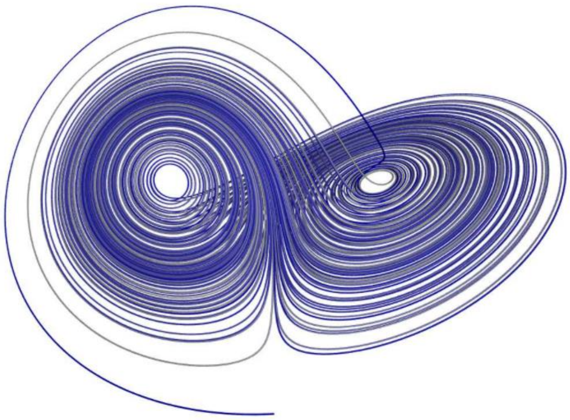
Cáscara de Nuez" (The Universe in a Nutshell, Bantam Spectra, RU, 2001). Ambos con versiones e-book en español disponibles.

[51] N., Lorenz, Edward (1995). La esencia del caos: un campo de conocimiento que se ha convertido en parte importante del mundo que nos rodea (1a. ed edición). Debate. ISBN 8474448700. OCLC 431231741.

pertenecen al conjunto de los atractores". Se le llama atractor extraño al que consiste en un número infinito de curvas, superficies o compuesto de más dimensiones. Es el corazón del sistema caótico.

"En los sistemas dinámicos un cambio casi imperceptible en una constante producirá un cambio cualitativo en el comportamiento del sistema: de estacionario a periódico, de estacionario o periódico a cuasi periódico, o de estacionario, periódico o cuasi periódico a caótico. Hasta el caos puede cambiar abruptamente a un caos de mayor complicación y, por supuesto, cada uno de estos cambios puede darse en dirección opuesta. Semejantes cambios se llaman bifurcaciones".

Efecto Mariposa



El término *"efecto mariposa"* se usa para referirse a la sensibilidad de un sistema respecto de las condiciones iniciales. Su nombre proviene de un proverbio chino: "*el aleteo de las alas de una mariposa se puede sentir al otro lado del mundo*", por un lado, y por la realización de lo que se denomina "*diagrama de fases*" para definir, en cada sistema, lo que es un atractor. Si bien, como se ve en la figura este diagrama tiene una forma similar a la de una mariposa, no tiene nada que ver con la creación de la frase por parte de Lorenz. Un sistema estable en reposo tiene un diagrama de fases que es un punto. Un sistema con movimiento, como un péndulo, tendrá un diagrama de fases como una circunferencia, aunque en el movimiento cíclico de un péndulo existen otros atractores no importantes denominados "*extraños*". Lorenz descubrió que su modelo tenía estos atractores.

La idea es que, dadas unas condiciones iniciales de un determinado sistema natural, la más mínima variación en ellas puede provocar que el sistema evolucione en formas totalmente diferentes.

Consecuencia de este efecto es el abandono de la idea determinista de que la naturaleza tiene comportamientos sujetos a las leyes de la mecánica. Por el contrario, la naturaleza no se asemeja a modelos previsibles y determinados, sino que existe un orden aparentemente aleatorio en los acontecimientos, que incapacita al hombre y su saber científico a predecir y controlar la realidad.

Ray Bradbury[52] escribió en 1952 un cuento titulado El ruido de un trueno (*A sound of Thunder*). En él, unos cazadores viajan en el tiempo, a la prehistoria y sin darse cuenta, matan a una mariposa. En consecuencia y debido a ello, cuando vuelven al presente se dan cuenta que el mundo en que se encuentran es diferente al que habían dejado cuando partieron. Esa muerte no histórica habría provocado un efecto en cadena de dimensiones inmensurables a lo largo de los milenios. Curiosa

[52] Ray Bradbury (1920 – 2012) es conocido como autor de ciencia ficción, aunque — a pesar de títulos tales como "Crónicas Marcianas" — él mismo declaró que no era escritor de ciencia ficción sino de fantasía y que su única novela de ciencia ficción es "Fahrenheit 451"

coincidencia de la apelación a una mariposa para describir el mismo efecto: Proverbio chino, Bradbury y Lorenz…

Cuando Lorenz presentó su trabajo en la Academia de Ciencias de Nueva York, tratando de predecir el clima a través de modelos matemáticos que relacionaban variables, no logró el nivel de predicción esperado. Al revisar los datos se advierte que, haciendo pequeñísimos cambios en ellos, se lograban resultados absolutamente diferentes. (La divergencia del modelo meteorológico se originó cuando usó una impresora programada para imprimir cifras redondeadas con 3 decimales en lugar de 5 para ahorrar espacio. Esas cifras eran las usadas para cargar los datos iniciales)

Diferencias entre métodos numéricos, métodos analíticos y simulación

Es común que un ingeniero esté frente a un sistema cuya funcionalidad deba conocer a fondo a fin de mejorar su rendimiento, repararlo, replicarlo, etc.

Muchas veces es posible experimentar con los sistemas mismos, pero también es muy frecuente que la observación directa, la recopilación de datos estadísticos y la introducción de variantes en los parámetros de funcionamiento no sean posibles de hacer y, por tanto, no sea posible buscar mejoras o cambios. Lo que ocurre es que la experimentación con el sistema real puede plantear problemas éticos o económicos. En esos casos, puede construirse una versión simplificada, un *prototipo* del sistema.

Por ejemplo, una fábrica de carrocerías de ómnibus que desea lanzar un nuevo modelo con doble piso apto para viajes de larga distancia. Antes del diseño e implementación completa de la línea de montaje, podría probar montar partes o versiones simplificadas que orienten sobre la factibilidad del proyecto o las direcciones a seguir en su desarrollo.

Sin embargo, muchas veces no es necesario, o no es posible, o no es económico, construir un prototipo (en este caso de las carrocerías, por ejemplo, habría que disponer de un bastidor comercial que es costoso), pero en cambio es posible construir un *modelo* lógico-matemático que describa mediante ecuaciones y relaciones el comportamiento básico de la línea de producción, de los componentes y los requerimientos, tomados como sistema.

El estudio, análisis, validación y verificación del modelo puede hacerse en muchos casos por *métodos analíticos*, por ejemplo, mediante la resolución de un sistema de ecuaciones diferenciales que describa la evolución de una población; o por *métodos numéricos*, por ejemplo, utilizando un algoritmo de programación no lineal para resolver un problema de optimización relativo al diseño de la trayectoria de un proyectil.

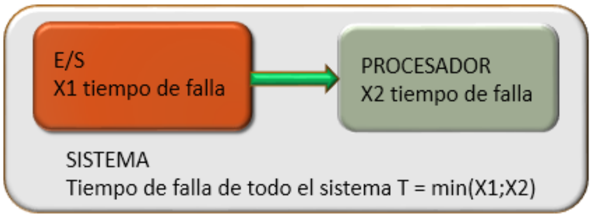
Cuando se estudian modelos más realistas o detallados, es posible llegar a situaciones en las que no se puede resolver analítica o numéricamente el problema. Entonces podemos utilizar simulación. Esencialmente, la simulación consistirá en construir un programa de computadora que describa el comportamiento del sistema de interés, o refleje el modelo que lo representa, y proceder a experimentar con el programa para llegar a conclusiones que apoyen la toma de decisiones.

Usaremos un caso para ilustrar y poder comparar los distintos métodos de tratamiento de modelos. Se trata de un ejemplo de un proceso simplificado que está compuesto por un sistema de entrada—salida (E/S) y un procesamiento o caja negra (procesador). El sistema falla cuando alguno de los dos componentes falla.

Expresado matemáticamente, X_1 representa el tiempo hasta que falla el sector de E/S, X_2 es el tiempo hasta que falla el procesador, y T será el tiempo hasta que falla todo el sistema. De esta manera el tiempo de falla de todo el sistema será:

$T = \min(X_1, X_2)$.

Que se lee:
"El tiempo de falla T será el tiempo de falla X_1 o el tiempo de falla X_2, lo que ocurra primero"



Como hay incertidumbre sobre los tiempos anteriores, consideramos que X_1, X_2 y, consecuentemente, T son aleatorios. Podremos calcular el tiempo esperado E(T) hasta que se produce una falla, como parte de un estudio destinado a mejorar este sistema.

Para simplificar, establecemos que X_i tiene una distribución exponencial de parámetro λ_i, para i = 1,2, y que ambas variables aleatorias son independientes. En este caso, se tiene que

$$E(T) = \iint_0^{\infty} \min(x_1, x_2)\lambda_1 e^{-\lambda_1 x_1}\, \lambda_2 e^{-\lambda_2 x_2}\, dx_1\, dx_2$$

Método analítico
Para llevar adelante el método analítico aplicamos cálculo de probabilidades, teniendo en cuenta la independencia de las variables y la distribución exponencial utilizada,

$$P(T > t) = P(X_1 > t, X_2 > t) =$$
$$= P(X_1 > t) \times P(X_2 > t) =$$
$$= e^{-\lambda_1 t} e^{-\lambda_2 t} = e^{-(\lambda_1 t + \lambda_2 t)}$$

Observamos que *T* tiene distribución exponencial de parámetro $\lambda_1 + \lambda_2$
Entonces

$$E(T) = \frac{1}{\lambda_1 + \lambda_2}$$

Como vemos esta es una expresión sencilla. Si, por ejemplo λ_1 fuera 1 falla/unidad de tiempo y λ_2 fuera 2 fallas por unidad de tiempo, el resultado sería E(T) = 1/3.

Como en esta aproximación hemos usado las hipótesis de independencia (P()) y de distribución exponencial de las variables, resulta que hay sensibilidad a cualquiera de esas dos hipótesis, y que el cálculo exacto puede resultar mucho más complicado e, incluso, imposible de obtener analíticamente.

Método numérico

Otra alternativa es apelar a algún método de integración numérica. Por ejemplo reparametrizar el problema, de manera tal que la región de integración sea acotada.

Una forma sería hacer el cambio de variables siguiente

$$X_i = \frac{-(\ln(1-\lambda_i))}{\lambda_i}$$

Con lo que nos queda

$$E(T) = \iint_0^1 \min\left(-\frac{-(\ln(1-\lambda_1))}{\lambda_1}, -\frac{-(\ln(1-\lambda_2))}{\lambda_2}\right) d\lambda_1 d\lambda_2$$

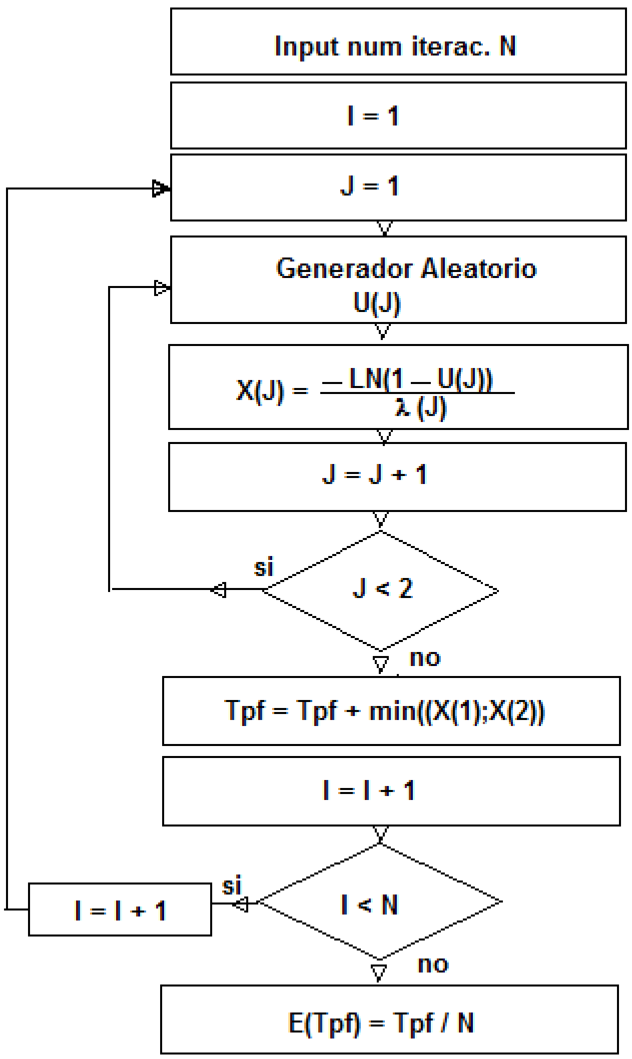
Aplicando cualquier método de integración disponible podríamos obtener un valor que representará el tiempo esperado de falla del sistema.

¿Cuál es la diferencia? Probamos dando valores a las tasas de ocurrencia de las fallas. Sigamos con el supuesto de que esas tasas de ocurrencia valen $\lambda_1 = 1$ y $\lambda_2 = 2$, en el primer caso encontramos más arriba que E(T) = 1/3 (0,3333…). En este segundo caso, obtendremos un valor cercano a 0,32

Simulación

Pero tenemos la alternativa de construir un simulador para poder "ver" cómo se desarrollan estas fallas. Un simulador sencillo estaría construido con un algoritmo con una estructura que podría ser la siguiente:

1. Tiempo para la falla, Tpf = 0 ‘Reiniciamos el valor de T a cero
2. Ingresar el número de iteraciones (n) ‘Establecemos el número de ensayos
3. I = 1
4. J = 1
5. Generar un número aleatorio U(J) entre 0 y 1
6. Calcular $x(j) = -\frac{\ln(1-U(j)}{\lambda(j)}$ ➔X(J)
7. J = J + 1
8. Si J < 2 ir a punto 5
9. Tpf = Tpf + mínimo |X(1),X(2)|
10. I = I + 1
11. Si I < n hacer i = i +1 e ir a punto 4 ‘Hacemos un nuevo ensayo
12. E(Tpf) = Tpf / n



Este algoritmo genera dos valores aleatorios para la variable X, en un vector (X1, X2) mediante un método que se discutirá más adelante.

Esta operación se repetirá n veces, y luego extraeremos el promedio de los tiempos esperados en cada vez.

Como podemos ver, hemos hecho un algoritmo que tiene ciertos bloques que adelantan los componentes básicos de cualquier simulador:

Un generador de eventos que constituyen las entradas, una función que transformación de las entradas en salidas, una recopilación de estados y una contabilidad y análisis estadístico de los datos obtenidos (en este caso, el análisis estadístico es un simple promedio)

Entre los casos de simulación puede darse la situación de aquellos en que los cambios de estado del sistema ocurren discretamente.

Por ejemplo, tomemos un modelo de colas M/M/1 donde el número de clientes promedio en el sistema[53] es

$$L = \frac{\lambda}{(\lambda - \mu)}$$

Si consideramos que en un instante dado existe la probabilidad *p* de que llegue un cliente y que, simultáneamente, si el servidor estaba ocupado hay una probabilidad *q* de que un cliente salga del sistema. La expresión anterior, expresada en términos de número *esperado* promedio de clientes en el sistema, siempre que *q* > *p*, será:

[53] Para quien no tenga presente los modelos de colas, recordaremos que λ representa la tasa de llegadas promedio y μ la de salidas, ambas con distribución exponencial. (Ver Libro 4).
En este mismo libro, en el Capítulo 7 presentamos el desarrollo de un simulador de colas M/M/1

$$L = \frac{p(1-p)}{(q-p)}$$

Por ejemplo, si la probabilidad de que en un momento dado de tiempo llegue un cliente es de p = 0,4 y que un cliente salga es q = 0,6, entonces:

L = 0,4 x 0,6 / 0,2 = 1,2

Si queremos simular este caso, podemos seguir los siguientes pasos:

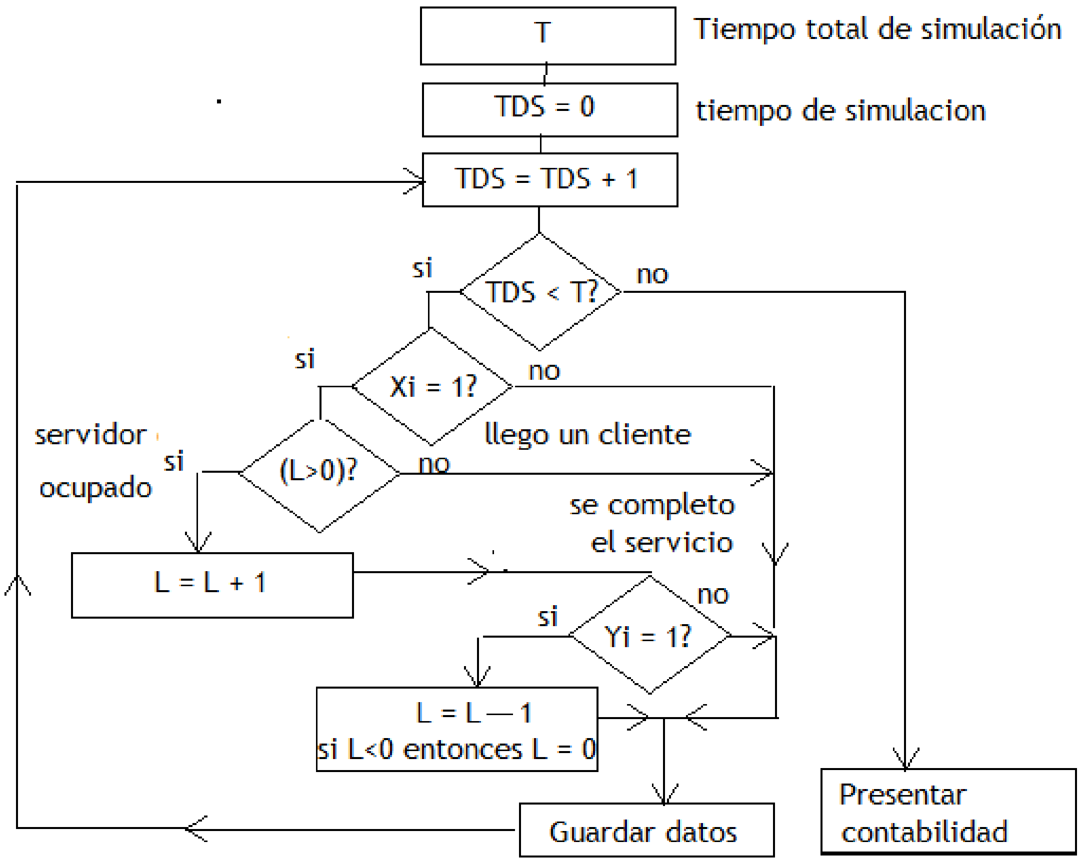
a. Usaremos una tabla de números aleatorios, supongamos que decidimos usar la tabla de 100 números que vemos más abajo. Esa tabla es muy sencilla de construir.
b. En la tabla, seleccionaremos, al azar, 50 valores para las llegadas y otro conjunto del mismo tamaño para las salidas
c. En el primer conjunto, ***llegadas***, buscamos al azar un número, i. Arbitrariamente establecemos que si el número que seleccionamos es < 40, entonces Xi será 1, lo que indica “se produjo una llegada”. En cambio, si el número aleatorio i de la serie mayor o igual a 40, entonces Xi = 0 “no se produjo una llegada”
d. De manera similar, en el segundo conjunto, ***salidas***, usamos un valor Yi = 1 (se produjo una salida) si el número aleatorio i de la serie es < 60. Si es mayor o igual a ese valor, entonces Yi = 0 (no se produjo una salida)

Tabla de 100 números aleatorios generada con Excel, buscando valores entre 0 y 100. Los cinco primeros renglones se asignan a entradas y los cinco últimos corresponden a las salidas:

25	52	20	19	13	98	60	59	100	59
64	55	37	97	13	28	5	96	19	62
98	41	22	95	30	99	39	69	50	91
31	85	58	34	64	84	68	48	0	99
23	51	65	40	24	69	48	0	17	55
97	11	13	83	66	22	19	20	41	93
5	72	27	26	43	94	55	82	96	82
77	57	70	97	39	42	19	38	89	13
45	99	73	43	72	13	20	51	81	5
69	16	100	18	62	51	72	66	33	67

El algoritmo que proponemos debería tener esta estructura:

1. Ingresar un tiempo total de simulación T
2. Tiempo de simulación TDS = 0
3. TDS = TDS + 1
4. Si TDS < T ir a 5. Caso contrario ir a 9
5. Si Xi = 1 (llega un cliente) y si el servidor está ocupado incrementar la cola
6. Si Yi = 1 (se completó el servicio) decrementar la cola en uno o poner bandera de servidor libre
7. Guardar datos
8. Ir a 3
9. Presentar contabilidad



Como ejemplo mostramos una tabla con los primeros 5 eventos:

(i) Evento N°	***Aleat(i)*** entrada	***Xi***	***Aleat(i)*** salida	***Yi***	***L***
1	25	1	97	0	1
2	52	0	11	1	0
3	20	1	83	0	1
4	19	1	53	1	1
5	13	1	66	0	2

Así construiremos una tabla con los 50 eventos. Si se suma la columna L (Clientes en el sistema) y se divide entre 50, obtendremos el número promedio de clientes en el sistema

El fundamento del algoritmo es el siguiente: al generar una tabla de 100 números aleatorios entre 0 y 100 cada uno de ellos, estamos indicando que cada uno tuvo la misma probabilidad de "aparecer". Esa probabilidad es de 1/100. Para el caso de las llegadas, al seleccionar entre 0 y 40 para llegar y 40 y 100 para no llegar, hemos convertido las probabilidades entre 1/40 y 1/60, respectivamente. En el caso de la salida, la distribución inicial uniforme de 1/100 la pasamos a 1/60 y 1/40.

Como vemos, podemos hacer casi todo esto "manualmente" en una simple hoja de papel, salvo la generación de los números aleatorios. Generamos esos números usando hoja de cálculo, pero podríamos usar un bolillero o una tabla de números aleatorios de un texto o cualquier método. Veamos este mismo caso, pero usando ahora una hoja de cálculo:

	A	B	C	D	E	F	G	H
1								
2		Numeros aleatorios						
3		Entradas	Salidas		Xi	Yi	L	Lmedia
4		42	95		0	0	0	0,86
5		82	63		0	0	0	Lmax
6		32	20		1	1	0	3
7		46	85		0	0	0	
8		0	24		1	1	0	
9		30	36		1	1	0	
10		53	58		0	1	0	
11		[illegible]	[illegible]		1	1	0	

Con las siguientes funciones:

	A	B	C	D	E	F	G	H
1								
2		Numeros aleatorios						
3		Entradas	Salidas		Xi	Yi	L	Lmedia
4		=ALEATORIO.ENTRE(0;100)	=ALEATORIO.ENTRE(0;100		=SI(B4<40;1;0)	=SI(C4<60;1;0)	=SI((E4-F4)<0;0;E4	=PROMEDIO(G4:G53)
5		=ALEATORIO.ENTRE(0;100)	=ALEATORIO.ENTRE(0;100		=SI(B5<40;1;0)	=SI(C5<60;1;0)	=SI((G4+E5-F5)<0;	Lmax
6		=ALEATORIO.ENTRE(0;100)	=ALEATORIO.ENTRE(0;100		=SI(B6<40;1;0)	=SI(C6<60;1;0)	=SI((G5+E6-F6)<0;	=MAX(G4:G53)
7		=ALEATORIO.ENTRE(0;100)	=ALEATORIO.ENTRE(0;100		=SI(B7<40;1;0)	=SI(C7<60;1;0)	=SI((G6+E7-F7)<0;	
8		=ALEATORIO.ENTRE(0;100)	=ALEATORIO.ENTRE(0;100		=SI(B8<40;1;0)	=SI(C8<60;1;0)	=SI((G7+E8-F8)<0;	
9		=ALEATORIO.ENTRE(0;100)	=ALEATORIO.ENTRE(0;100		=SI(B9<40;1;0)	=SI(C9<60;1;0)	=SI((G8+E9-F9)<0;	
10		=ALEATORIO.ENTRE(0;100)	=ALEATORIO.ENTRE(0;100		=SI(B10<40;1;0)	=SI(C10<60;1;0)	=SI((G9+E10-F10)<	

Las columnas B y C las utilizamos para generar un número aleatorio entero entre 0 y 100 usando la función Excel “ALEATORIO.ENTRE()” de la siguiente manera:

=ALEATORIO.ENTRE(0;100)

La columna E la usamos para convertir, renglón a renglón, la distribución uniforme de números aleatorios que habíamos generado en la columna A. La convertimos de uniforme continua a una distribución binaria que presenta una probabilidad 0,4 para un “1”, que indicaría una llegada. Para lograr este resultado usamos primero el reglón 4 (el primero en nuestro ejemplo) y simulamos una llegada expresando que si el número aleatorio allí generado es menor que 40, entonces escribimos un “1”, que significa llegada, de lo contrario un “0” que significa que no hubo llegadas. Así:

=SI(B4<40;1;0)

Del mismo modo, en la columna F utilizamos una probabilidad de 0,6 para hacer la simulación de salidas:

=SI(C4<60;1;0)

Solo nos resta autocompletar estas dos columnas E y F hasta cubrir todos los renglones en que generamos llegadas/salidas.

Ahora trabajaremos con la columna L donde representamos el número de clientes en la cola, mediante dos pasos:

1) en el primer renglón (en este caso el 4) hacemos la diferencia entre entradas E y salidas F. Como existe la posibilidad que el número que obtengamos sea negativo, averiguamos si es y en ese caso escribimos “0”:

=SI((E4—F4)<0;0;E4—F4)

2) Desde el segundo renglón y hasta el final, sumamos entradas y restamos salidas, pero teniendo en cuenta cuantos clientes había en la cola en el renglón anterior y cuidando siempre que no quede un número negativo de clientes en la cola. De esta manera el segundo renglón y los siguientes hasta el final, serán:

=SI((G4+E5—F5)<0;0;G4+E5—F5)

Por último, debemos hacer la contabilidad final y presentar los resultados, para lo cual, usando la función “PROMEDIO()” presentamos la cantidad media de clientes en la cola.

=PROMEDIO(G4:G53)

y con la función “MAXIMO()” el número máximo de clientes en la cola:

=MAX(G4:G53)

En el capítulo 7, más adelante, veremos otra manera de simular en Excel este mismo caso.

Metodología de la Simulación

Si bien puede haber diferentes maneras de diseñar y operar un sistema de simulación, varios autores proponen esquemas de pasos a observar. Uno de ellos es el Schmeisser (1990) simplificado en Ríos Insua y otros (Simulación, Métodos y Aplicaciones, 2009):

Paso 1

- A partir de un generador de números aleatorios se obtienen "observaciones" o datos o eventos

Paso 2

- Esas observaciones se transforman en entradas aptas para el modelo

Paso 3

- Se opera el modelo de manera tal que transforme las entradas en salidas

Paso 4

- Se realiza la "contabilidad" de las salidas y se usa estadísticas para describir el compartimiento del sistema y tomar decisiones

Generación de Números Aleatorios

Requisitos para una verdadera aleatoriedad

Los estudios de generación de números aleatorios se basan en el desarrollo de la criptografía, inicialmente más como necesidad diplomática, militar o comercial que científica. Uno de los métodos más "seguros" y analizados desde el origen de la criptografía actual (principios del siglo XX) fue y es la "***libreta de un solo uso***" (LSU), que fue objeto de estudio de varios matemáticos, que originan conceptos de seguridad criptográfica que podemos dividir en dos aspectos:

- El primero de esos aspectos es verificar la seguridad teórica del sistema comenzando con la LSU, cuya seguridad fue demostrada por Shannon y que está sujeta a la disponibilidad práctica de los números aleatorios (A partir de esto se definió lo que se conoce como ***Seguridad de Shanonn)***
- El segundo es la seguridad ofrecida por los cifrados de punta (por ejemplo, el AES) diseñados con los principios aprendidos durante la larga historia de la rotura de códigos y sujetos al testeo intensivo en un proceso de estandarización, ya sea público o de un servicio de seguridad de excelencia. Este proceso de pruebas es lo que se conoce como seguridad empírica. No está demostrada pero recibe la confianza de la mayoría de los grandes usuarios para proteger sus secretos más vitales.

Métodos que pueden ofrecer seguridad empírica pero no tienen seguridad de Shannon

> **¿Qué diferencia hay entre "*seguridad teórica*" y "*seguridad empírica de Shanonn*"?**
> La *seguridad teórica* es aquella en que matemáticamente se puede demostrar que la probabilidad de predicción de un número cualquiera de la serie tiende a cero. La empírica, en tanto, si bien eso no se puede demostrar, tiene una probabilidad razonablemente baja como para que, en la práctica, resulte muy seguro.

Para comprender lo que sigue supongamos el siguiente escenario: queremos generar un cifrado y para hacerlo necesitamos una serie de números aleatorios. ¿Dónde buscamos esos números? Usamos cualquier programa que tenga un generador de números aleatorios de lo mejor. Pero es un algoritmo y, por lo tanto es determinístico.

Si la clave la genera un programa determinístico, entonces no es aleatoria ni se puede afirmar que el sistema de cifrado ofrezca la seguridad teórica de la libreta de un solo uso.

Esta manera de cifrar a partir de un generador de números aleatorios se conoce como ***cifrado de flujo*** en referencia al flujo de números aleatorios que entrega el programa. Generalmente estos algoritmos utilizan una clave pequeña que se usa como semilla para un flujo pseudoaleatorio largo, que luego se combina con el mensaje empleando algún mecanismo como los de la libreta de un solo uso (por ejemplo, XOR[54]). Los cifrados en flujo pueden ser seguros en la práctica, pero no pueden ser absolutamente seguros en el mismo sentido demostrable de la libreta de un solo uso.

[54] XOR significa "O exclusivo" en lógica proposicional. Es una operación básica junto a "AND" y "OR" ("Y" y "O")

Los cifrados usados por el ejército alemán en la Segunda Guerra Mundial resultaron ser cifrados en flujo inseguros, y no eran las útiles libretas de un solo uso automatizadas como pretendían sus diseñadores[55].

Sin embargo, si se utiliza un buen generador de números pseudoaleatorios criptográficamente seguro y moderno, se pueden lograr los cimientos de un cifrado en flujo ***empíricamente seguro***. Hay muchos diseños bien probados en el dominio público, que varían desde la simplicidad del RC4 al uso de un cifrado en bloque como el AES. Parecería que hay pocos motivos para inventar nuevos cifrados en flujo, pero se cree desde hace tiempo que la NSA y otras agencias emplean un esfuerzo considerable en los cifrados en flujo.

Métodos que no ofrecen seguridad empírica ni seguridad de Shannon

La similitud entre los cifrados en flujo y las libretas de un solo uso lleva a menudo a que se inventen cifrados en flujo inseguros bajo la creencia falsa de haber desarrollado una versión práctica de la libreta de un solo uso. Una versión especialmente insegura son los generadores de números aleatorios que se distribuyen en muchas (quizás la mayoría) de las bibliotecas accesorias de los lenguajes de programación o en forma de llamadas al sistema operativo. Normalmente producen secuencias que pasan pruebas estadísticas, pero sin embargo se pueden romper con técnicas criptoanalíticas.

Hace un tiempo, el ANSI C estándar restringía la salida de la rutina de números aleatorios del lenguaje C a un entero de precisión simple, 16 bits para la mayoría de las implementaciones, dando 32768 valores distintos antes de repetirse. Esto es completamente inseguro y fácil de vulnerar por fuerza bruta (pensemos en una vieja computadora con un reloj de 1 GHz que tarde 10.000 ciclos de reloj en comprobar cada salida del ciclo RNG — que es un número de ciclos exageradamente grande — tardaría menos de un tercio de segundo en comprobar todas las salidas posibles.

Los generadores estándar de números aleatorios no sirven para propósitos criptográficos, concretamente para la libreta de un solo uso. En particular, el relativamente reciente algoritmo ***Tornado*** (Twister) de Mersenne[56], muy usado, y que es lo bastante “aleatorio” (o “bueno”) para la mayoría de los usos de simulación o investigación, mejor que la mayoría de los generadores de su mismo tipo, y

---

[55] Destinado a aquellos que se interesen en el tema, en el Anexo I se describe muy brevemente como funcionaba la máquina de cifrado “Enigma” utilizada desde 1930 a 1960 y en el anexo II un fragmento de la novela “Criptonomicon” que explica, en términos de aritmética modular su principio de funcionamiento

[56] El “Mersenne twister” es un Generador de números pseudoaleatorios desarrollado en 1997 por Matsumoto y Nishimura

Su nombre proviene de que la longitud del periodo corresponde a un número primo de Mersenne. Existen al menos dos variantes de este algoritmo, distinguiéndose únicamente en el tamaño de primos Mersenne utilizados. El más reciente y utilizado es el Mersenne Twister MT19937, con un tamaño de palabra de 32-bit. Existe otra variante con palabras de 64 bits, el MT19937-64 que genera otra secuencia.

Desde la versión 2010 lo utiliza — según la página de soporte de Microsoft — la hoja de cálculo Excel (cfr. artículo “Ayuda sobre la función ALEATORIO”) (https://support.office.com/es-es/article/aleatorio-función-aleatorio/)

también bastante rápido, no es apto para generar claves de libreta de un solo uso. El algoritmo es determinista y no fue diseñado para la seguridad criptográfica.

Además, los valores conocidos públicamente como los dígitos finales de los tiempos de las carreras, los precios de cierre de valores de bolsa, por muy poco conocidos que sean, las temperaturas o presiones atmosféricas diarias, etc., aunque aparentemente aleatorios, son predecibles después de que se produzca el hecho. En realidad, tampoco pueden usarse secuencias verdaderamente aleatorias que hayan sido publicadas, ya que si se identifican son predecibles. Un ejemplo es la publicación de una tabla de un millón de números aleatorios hecha en 1950 por la RAND[57]; ha pasado todas las pruebas estadísticas de aleatoriedad hasta ahora, y se cree que es verdaderamente aleatoria. Pero, al haberse publicado, es completamente predecible. También lo son los dígitos de *pi*, *e* y otros números irracionales o trascendentales; puede que las secuencias sean aleatorias (lo cual es algo discutible), pero son completamente predecibles.

Ejemplo de libreta de un solo uso

Para comprender como funciona la "libreta de un solo uso" veremos un ejemplo simplificado. El emisor del mensaje dispone, cada día, del código correspondiente. Ese código tendrá la misma longitud del mensaje. En la Segunda Guerra Mundial, los británicos usaban grupos de cuatro letras en una página con numerosos grupos. El receptor dispone ese día esa misma página, o en el mismo mensaje se informa, en clave simple, cual es la página de uso para ese mensaje. En cualquier caso, esa página no se usa nunca más.

Por otro lado, a cada letra del alfabeto se le asigna un número (por ejemplo: A=0, B=1…Y=24, Z=25 o cualquier otra secuencia). La clave del día debe tener la misma longitud que el mensaje.

Supongamos que la clave del día es **PTZB** (para un mensaje de 4 letras) y el texto a emitir en limpio es "ROMA". Para encriptar se comienza por sumar la clave al texto convertido en secuencia de números del alfabeto. Luego, usando módulo 26, se resta a cada valor el número 26 (letras del alfabeto): así el mensaje ROMA se encriptó como "GHLB", de acuerdo con el siguiente procedimiento:

ENCRIPTADO

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

CLAVE DEL DIA	P	T	Z	B
NUMERO DE LETRA DE LA CLAVE	15	19	25	1
TEXTO CLARO	**R**	**O**	**M**	**A**
NUMERO LETRA DE TEXTO CLARO	17	14	12	0
SUMA	15+17=32	19+14=33	25+12=37	1+0=1
MODULO 26	32-26=6	33-26=7	37-26=11	1
TEXTO **CIFRADO**	**G**	**H**	**L**	**B**

DESENCRIPTADO

[57] RAND, siglas de "***Research and Development***" (Investigación y Desarrollo). Corporación de Estados Unidos utilizada como laboratorio de ideas y vinculada con las fuerzas armadas de ese país.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

CLAVE DEL DIA	P	T	Z	B
NUMERO DE LETRA DE LA CLAVE	15	19	25	1
TEXTO CIFRADO	**G**	**H**	**L**	**B**
NUM LETRA DE TEXTO CIFRADO	6	7	11	1
SUMAR 26	26+6=32	26+7=33	26+11=37	26+1=27
RESTAR CLAVE	32-15=17	33-19=14	37-25=12	27-1=26
TRUNCAR EXCESO 25 (RESTAR 26 SI >25)	17	14	12	26-26=0
TEXTO DESCIFRADO	**R**	**O**	**M**	**A**

En la siguiente figura mostramos una libreta de un solo uso casera hecha con hoja de cálculo de manera muy sencilla. En esa figura se pueden ver las dos fuentes de aleatoridad de este método:

1) la que proviene de la elección al azar de los grupos de 4 letras (o grupos de cantidades variables o de otras cantidades) y de su propia secuencia y cantidad indeterminada.
2) La que proviene del orden del alfabeto, que no necesariamente tiene que ser el convencional y de la asignación numérica de ese orden, que no necesariamente debe comenzar en A=0 o incluir o no la "Ñ" (en este caso la incluye)

CLAVE DEL DIA																
	W	U	L	M	A	D	S	L	G	M	O	U	W	B	E	I
	23	21	11	12	0	3	19	11	6	12	15	21	23	1	4	8
TEXTO CLARO ENTRADA																
	H	O	L	A	S	O	Y	A	L	E	J	A	N	D	R	O
	7	15	11	0	19	15	25	0	11	4	9	0	13	3	18	15
TEXTO ENCRIPTADO SALIDA																
	4	10	22	12	19	18	18	11	17	16	24	21	10	4	22	23
	E	K	V	M	S	R	R	L	Q	P	X	U	K	E	V	W
CLAVE SECUNDARIA PERMANENTE																
A	0	A														
B	1	B														
C	2	C														
D	3	D														
E	4	E														

TEXTO ENCRIPTADO ENTRADA																
	E	K	V	M	S	R	R	L	Q	P	X	U	K	E	V	W
	4	10	22	12	19	18	18	11	17	16	24	21	10	4	22	23
TEXTO CLARO SALIDA																
	7	15	11	0	19	15	25	0	11	4	9	0	13	3	18	15
	H	O	L	A	S	O	Y	A	L	E	J	A	N	D	R	O

Libreta de un solo uso de 10 grupos				
Grupo 1	Y	G	U	P
Grupo 2	V	G	Y	N
Grupo 3	S	G	P	T
Grupo 4	A	L	C	Ñ
Grupo 5	T	G	U	J
Grupo 6	S	A	U	K
Grupo 7	O	R	Q	C
Grupo 8	A	I	H	Ñ
Grupo 9	I	Q	J	B
Grupo 10	Ñ	P	X	O

Podíamos completarla colocando un generador de claves, o libreta de clave del día, utilizando alguna de las funciones de números aleatorio de Excel, pero en ese caso sería mejor hacerla mediante una extracción de números no volátiles, al menos hasta que el operador necesite generar una nueva libreta. La figura muestra un ejemplo de libreta de un solo uso de 10 grupos hecha con las funciones =ALEATORIO.ENTRE(0;26) y BUSCARX.

Aleatoriedad con seguridad de Shannon

Para conseguir la seguridad de Shannon

necesitaremos una fuente de datos realmente aleatoria, o, dicho en otras palabras, que esos datos sean perfectamente impredecibles[58]. Una base teórica para la existencia física de la impredecibilidad es la mecánica cuántica, que aún está sujeta a la comprobación experimental. Otra base es la teoría de los sistemas dinámicos inestables y la teoría del caos. Estas teorías sugieren que incluso en el mundo determinista de la mecánica newtoniana, los sistemas reales evolucionan de maneras que no se pueden predecir en la práctica porque haría falta conocer las condiciones iniciales con una precisión que crece exponencialmente con el tiempo de la predicción.

Los datos buscados deben mostrar una aleatoriedad perfecta. En la práctica, la mayoría de las fuentes muestran alguna imperfección o desvíos. La calidad de la aleatoriedad se mide por entropía. Un bit perfectamente aleatorio tiene una entropía uno. Una idea, propuesta por Von Neumann, es utilizar un algoritmo para combinar varios bits aleatoriamente imperfectos, con una entropía menor que uno, para producir un bit con entropía igual a uno. Este proceso se llama destilación de entropía o ***blanqueamiento de Von Neumann***, y permite generar en la práctica datos aleatorios adecuados para su uso en una secuencia. El blanqueamiento de entropía de Von Neumann es uno de los pilares matemáticos de la física cuántica estadística y su tratamiento excede los alcances de este texto y nuestros cursos.

Generadores en Linux

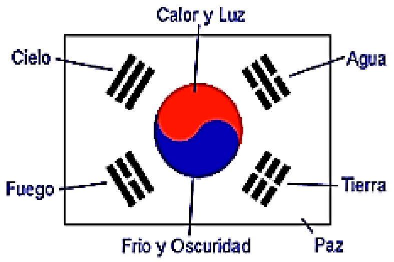
En ***Linux*** (y otros sistemas derivados de ***Unix***), el generador de números aleatorios del kernel **(/dev/random)** utiliza el ruido ambiental para generar datos aleatorios y es mejor que algunos diseños basados en llamadas al sistema. Intenta estimar la cantidad de entropía que recoge y se bloquea si el fondo de entropía se agota. Pretende ser mejor que muchos generadores parecidos, y si fuera así, estaría muy cerca de ser satisfactoriamente aleatorio, a pesar de que es un proceso lento en sistemas que tienen pocas fuentes de ruido utilizables. Sin embargo, puede alimentarse con entropía adicional leyendo de un dispositivo generador de ruido.

Linux también proporciona **/dev/urandom**, que emplea un algoritmo determinista para generar los datos cuando no hay ruido ambiental disponible. Existen diseños mejorados, como el algoritmo de Yarrow[59].

---

[58] Como dato anecdótico, los británicos distribuían en la población civil, a manera de "trabajo" voluntario para el esfuerzo de guerra, una bolsa de tela con 26 bolas con las letras del alfabeto. Los voluntarios tenían la consigna, varias veces por día, de extraer las bolas de a cuatro y anotar en un cuaderno las letras extraídas en el orden en que lo habían hecho. Así durante algunos días al cabo de los cuales devolvían esos cuadernos con los que se confeccionaban las libretas. Las páginas de esos cuadernos se mezclaban con todas las recibidas y luego se extraían al azar las necesarias para confeccionar las libretas de un solo uso. Como se ve, un procedimiento totalmente impredecible.

[59] El algoritmo Yarrow ("algoritmo de la milenrama") es llamado así en homenaje a un Micólogo (David Yarrow) que describió la planta llamada milenrama, que se usa como fuente aleatoria en el proceso de consulta del ***I Ching*** o "Libro de los cambios" aunque también se usan monedas u otros elementos.
El algoritmo fue diseñado por Schneier, Kelsey y Ferguson. Es un algoritmo libre, no se requiere licencia para utilizarlo y está incorporado en Mac OS X y FreeBSD para sus dispositivos /dev/random.



La bandera de la República de Corea presenta cuatro trigramas obtenidos del I Ching. Tradicionalmente, quien consulta el *Libro de los Cambios,* utiliza tallos de milenrama para obtener los hexagramas o trigramas en forma aleatoria.

Aleatoriedad, reproducibilidad y simulación

Existe una amplia variedad de generadores de números aleatorios y pseudo aleatorios, tanto que es un campo de investigación donde se emplean recursos importantes, porque la calidad de estos generadores aumenta a medida que los requerimientos de métodos de encriptación, juegos y simuladores también aumentan.

Nos puede parecer poco importante que, para discutir simulación, necesitemos conocer sobre la generación de números aleatorios, sin embargo, un conocimiento básico sobre ellos no va a permitir entender las limitaciones y posibilidades, incluido los riesgos, cuando los usamos en un simulador.

Los generadores de números aleatorios pueden ser ***físicos*** o ***algorítmicos***.

Los ***generadores físicos*** incluyen una variedad de métodos: desde ruletas o lanzadores de monedas (o extraer bolillas de una bolsa, como ya dijimos), hasta decodificadores de ruido electrónico blanco. Los inconvenientes de estos métodos derivan, precisamente de su aleatoriedad: es imposible o al menos muy difícil la repetición de secuencias, al menos que las series se almacenen y guarden, lo que lleva a uso importante de recursos.

Ahora bien, surge la pregunta obligatoria ¿por qué es necesario que podamos reproducir series de números pseudoaleatorios tal como fueron obtenidas en otro momento? Es lo que se conoce como ***reproductividad***. Supongamos que estamos desarrollando un simulador y sospechamos que uno de los pasos tiene un error que enmascara las salidas. Queremos hacer cambios en el paso sospechoso y comparar resultados, haciendo una corrida antes de hacer los cambios y otra idéntica después. Si no disponemos de la misma serie de números aleatorios idénticos en las dos corridas no podremos determinar si la variación del resultado es por la variación en la serie o por el cambio realizado. De allí la necesidad de contar con generadores a los que se les pueda pedir que reproduzcan exactamente la misma serie de números "aleatorios" (que, obviamente, ya no los son)

A partir de los años 1950, con las primeras computadoras, comenzaron a desarrollarse los ***generadores algorítmicos***, que son generadores de números aparentemente aleatorios o pseudo aleatorios. La propuesta de Von Neumann consiste en producir una serie de números a partir de una función y de un número inicial llamado "semilla" de manera tal que, una vez repetida esa semilla, la sucesión de números obtenida es idéntica en cada generación (reproductividad, como dijimos más arriba).

Estos conceptos básicos nos permiten definir claramente la idea de número aleatorio.

Según la obra citada de Ríos Insúa y otros tendremos cuatro definiciones básicas:

1. La de ***Kolmogorov*** (1987), "*Una sucesión de números es aleatoria si no puede producirse eficientemente mediante un programa más corto que la propia serie*". El análisis de esta definición puede llevar a un criterio similar al de Turing para reconocer inteligencia artificial[60]
2. Una sucesión de números es aleatoria si nadie que utilice recursos computacionales razonables en tiempos razonables puede distinguir entre la serie y una sucesión verdaderamente aleatoria de una forma mejor que tirando una moneda para distinguir cual es cual. Esta definición conduce a la de ***L'Ecuyer*** (1990), generadores PT-perfectos, de uso en criptografía.
3. De la ***paradoja de imprevisibilidad***: Si una función determinística es impredecible, es difícil probar cosas sobre ella, en particular que es impredecible.

Más allá de estas disquisiciones académicas, donde la última definición es particularmente poco aplicable en la práctica, tomaremos como base que un generador de números aleatorios puede ser definido como un conjunto finito de estados u observaciones U_i

Ese conjunto, en particular, está constituido por una serie finita de símbolos que son la salida de un estado inicial o semilla y una función de transición.

La distribución de probabilidad de la salida debe ser uniforme (donde uniforme significa que cada uno de los elementos de salida de la serie tiene la misma probabilidad de aparición que cualesquiera de los demás), acotada entre dos extremos, generalmente 0 y 1 y cada valor obtenido se espera que se comporte como una variable aleatoria independiente e idénticamente distribuida.

Hay autores, como *David Kahanner*, que definen cuatro características sobre las que se debería valorar cualquier generador de números aleatorios (también llamados GNU):

- **Calidad**. Se deben satisfacer las pruebas estadísticas adecuadas. Antes de la repetición de la secuencia debiera haber un periodo lo suficientemente largo.
- **Eficiencia**. El generador debiera ser rápido. Debe requerir la menor cantidad de recursos computacionales posible (uso de procesador, memoria, almacenamiento, etc.).
- **Repetitividad.** El método debe depender de una semilla, para permitir que un experimento se pueda repetir.
- **Transportabilidad**. Si un método dado se implementa sobre un sistema o plataforma diferente debe producir los mismos resultados.

[60] Turing se preguntó si se podía afirmar que las máquinas, que él había diseñado teóricamente, pensaban. Aún más, expuso un criterio bien sencillo y práctico para dilucidarlo, la llamada prueba **de Turing**. En esencia consiste en poner una persona y una computadora en dos habitaciones aisladas con los que un operador puede comunicarse mediante un teletipo. Si, planteando algunas preguntas y cuestiones, el operador no es capaz de distinguir en qué habitación está la computadora y en cuál la persona, es que la máquina es capaz de realizar operaciones inteligentes iguales a las humanas. Su opinión era que las máquinas superarían la prueba y, por tanto, deberían ser consideradas inteligentes. Para Turing, como buen matemático, pensar consiste en realizar operaciones lógicas manejando símbolos de manera adecuada. Es indiferente de que esa tarea la realicen unas neuronas encerradas en una caja craneana de hueso, o se haga por otros dispositivos dispuestos dentro de una caja de metal o de plástico. Por eso, concluye: "Creo que al final de este siglo -el XX- nuestras ideas y opiniones habrán evolucionado lo suficiente como para poder hablar sin rubor de máquinas pensantes".

Generadores congruenciales

Son los generadores *pseudo aleatorios* clásicos, (Lehmer, 1951) donde cada miembro de la serie se calcula según:

$$x_{n+1} = (ax_n + b) \bmod m$$
$$u_n = \frac{x_n}{m}$$

Donde *a* es el multiplicador, *b* es el término aditivo de manera tal que tanto *a* como *b* sean menores que *m*, que se busca que sea un número primo y *a* una raíz primitiva[61]. El término inicial, x_0 es la semilla.

En general, estos generadores:

1. Tienen ciclos. Significa que después de una serie de extracciones, la sucesión de extracciones se repetirá en forma idéntica.
2. La longitud del ciclo depende de los parámetros *a, b* y *m*
3. Dentro de la selección de parámetros que conducen a longitudes de ciclo iguales, algunas series parecen más aleatorias que otras

Estas consideraciones llevan a definir un ***generador tipo*** basado en la disponibilidad de arquitecturas de 32 bits, lo que nos da un mínimo estándar: período máximo, salida que parezca aleatoria y que pueda implementarse con facilidad en 32 bits. Esto condujo a $m = 2^{31} - 1$ con un multiplicador $a = 7^5 = 16807$ que permite un período máximo al ser 7 una raíz primitiva de *m*.

Generadores recursivos múltiples

Los *generadores congruenciales* son lineales, pero pueden generalizarse desde

$$x_{n+1} = (ax_n + b) \bmod m$$

A

$$x_{n+1} = (a_1 x_n + \cdots + a_k x_{n+1-k}) \bmod m$$

Que es válido para $k >= 1$ pero definiendo que *m* debe ser un número primo.

Observemos que si $m = 2$ podemos hablar de un generador aleatorio de bits fácilmente realizable con hardware. Se buscaron valores de los parámetros necesarios para obtener ciclos máximos. Por ejemplo, $k = 521$ produce un ciclo de $2^{521} - 1$ elementos.

Los generadores de este tipo son varios y conocidos por sus nombres propios. Estos son algunos ejemplos:

---

[61] Una raíz primitiva de un primo p es un número entero g de tal manera que g(mod p) tiene orden multiplicador p−1
(Fuente: https://www.i-ciencias.com/)
Como ejemplo, la raíz primitiva de 5 es 3, porque:

$$3^1 = 3$$
$$3^2 = 9 \equiv 4(mod5)$$
$$3^3 = 3^2 \times 3 \equiv 9 \times 3 = 27 \equiv 2(mod5)$$
$$3^4 = 3^3 \times 3 \equiv 27 \times 3 = 81 \equiv 1(mod5)\ \ hasta\ acá\ es\ p-1$$
$$3^5 = 3^4 \times 3 = 243 \equiv 3(mod5)\ que\ repite\ 3^1$$

y obtenemos todos los valores posibles 1, 2, 3 y 4 hasta 5. (El cero se omite)

- Generador tornado de Mersenne
- Generador Taustworthe[62]
- Generadores conocidos como "propios"

Para mejorar la generación de números, particularmente aumentar el período, se realizan generadores híbridos, que en definitiva son combinaciones de ambos tipos. Se usan, en general, dos operaciones para combinarlos: ***bajadura*** y ***combinación***.

La ***bajadura*** consiste en generar bloques fijos de números aleatorios de determinada longitud a cada uno de los cuales se le aplica una permutación, en algunos casos sacando las posiciones a permutar en forma aleatoria.

La ***composición***, introducida en 1988 por L'Ecuyer, consiste en un algoritmo con combinación de los dos métodos. Esto dio lugar a uno de los generadores más usados (MRG32k3a) que combina dos generadores múltiples de orden 3, con un período 2^{191}

Es importante tener en cuenta que si el período aumenta los requerimientos de memoria operativa necesarios son mayores.

Generadores aleatorios en las PC

Casi todas las aplicaciones de software aplicado a simulación, estadística, matemática o simplemente programación, incorporan algún tipo de generador de números aleatorios, los que, a su vez, son catalogados como "poco confiables" por los matemáticos teóricos, aunque para el uso práctico y cotidiano son, en realidad, bastante buenos y confiables.

Vamos a mencionar algunos como ejemplo, y debemos tener en cuenta que se han desarrollado bastante mejoras aprovechando las arquitecturas de 64 bits, que permiten extender ciclos y tamaños.

El lenguaje de simulación ***SIMSCRIPT***[63] implementa su generador con $m = 2^{31} - 1$ y a = 630360016. Es uno de los generadores más probados y, a la vez, más renovados, ya que el lenguaje es un clásico que se usa en varias aplicaciones, incluido ARENA.

El programa de simulación ***ARENA***, que se usa habitualmente en asignaturas relacionadas con simulación y en simulación de procesos industriales o generales, utiliza el algoritmo MRG32k3a, junto con otros programas específicos, entre ellos el soporte de SIMSCRIPT.

En los programas estándar de aplicación general, hojas de cálculo, lenguajes de programación, paquetes estadísticos, sistemas operativos, etc., se utilizan generadores menos potentes (aunque, se insiste, no por ello menos útiles a efectos de uso habitual).

Ejemplos:

[62] Si al lector le interesa experimentar con Taustworthe, sugerimos visitar: http://www.paconet.org/tausworthe.html

[63] SIMSCRIPT fue desarrollado en la corporación RAND con la participación de Herb Karr y Harry Markowitz en 1962, siendo el primer lenguaje de programación de simulación, luego pasó a la empresa de Markowitz CACI (California Analisys Center Inc.).

- La biblioteca ***UNIX*** usa:
  $m = 2^{48}$, $a = 25214903917$ y $b = 11$, con una salida que es $u_n = \frac{x_n}{2^{48}}$

- ***Java***, tiene una clase llamada ***java.util.Random*** que usa el generador anterior pero la salida es

$$u_n = \left[2^{27}\frac{x_{2n}}{2^{22}} + \frac{x_{2n+1}}{2^{21}}\right]/2^{53}$$

- ***Visual Basic***, inclusive la versión VBA, puede ser usada con semilla o usa por defecto como semilla un registro del reloj del sistema. Utiliza
  $m = 2^{24}$, $a = 1140671485$, $b = 12820163$ y la salida es igual a la de Unix, pero dividiendo por 2^{24}.

- ***Excel*** hasta 97, inclusive, $m = 1$; $a = 9821$; $b = 0{,}211327$, usando la salida directamente, sin dividir (el divisor es 1)

- ***Excel***, en las versiones 2003 y 2007, intentó usar el generador de *Wichman & Hill* pero, tuvo que corregir importantes defectos mediante parches, entre ellos que podía generar números negativos. Tiene tres generadores primarios:

$$x_{n+1} = 171x_n \; mod \; 30269$$
$$y_{n+1} = 172y_n \; mod \; 30307$$
$$z_{n+1} = 170 \; mod \; 30323$$

Cuya salida es

$$u_n = \left[\frac{x_n}{30269} + \frac{y_n}{30307} + \frac{z_n}{30323}\right] mod \; 1$$

Según la página de Microsoft *https://support.office.com/es-es/aleatorio-funcion-aleatorio* desde 2010 usa el algoritmo ***Mersenne Twister*** que, con mucho, es el generador pseudo aleatorio de propósito general más utilizado. Su nombre deriva del hecho de que la duración de su período se elige para ser un número primo Mersenne. Fue desarrollado en 1997 por Makoto Matsumoto y Takuji Nishimura. Fue diseñado específicamente para rectificar la mayoría de los defectos encontrados en los PRNG más antiguos.

Generación de números aleatorios usando hoja de cálculo.

Si bien nos referimos a cualquier hoja de cálculo disponible en la actualidad, por razones de simplicidad usaremos la sintaxis y funciones de *Excel* para describir las diversas maneras de obtener números aleatorios: En realidad, hay diferencias muy pequeñas, cuando las hay, con las demás hojas.

=ALEATORIO()
Esta función genera un número aleatorio con distribución uniforme entre 0 y 1 y usa un registro de reloj interno para generar la secuencia.

Ejemplo:
=ALEATORIO(), podría generar 0,27649582

Se trata de una función volátil. Es decir, que si ingresamos **=ALEATORIO()** en una celda, al pulsar "*Enter*" aparecerá un número aleatorio en esa celda. Pero si por algún motivo se recalcula la hoja (por ejemplo, por haber escrito un valor en otra celda cualquiera), se va a generar otro número aleatorio en la celda con la función.

Si queremos volver a generar un número aleatorio diferente sin escribir ni borrar nada en la hoja, entonces pulsaremos la tecla "*F9*", lo que provoca un recálculo de toda la hoja, incluyendo, por supuesto, la celda con la función Aleatorio(), por eso cambiará el número.

Si, al contrario, lo que queremos es evitar que esto ocurra, lo que debemos hacer es escribir en la celda elegida la función **=ALEATORIO()** y pulsar la tecla "***F9***" en lugar de "***Enter***", con lo cual el número que se obtendrá quedará fijo, aunque se recalcule toda la hoja. Esto ocurre porque en la celda donde escribimos la función **=ALEATORIO(),** al pulsar **F9** en lugar de "***Enter***", la función se borrará y será reemplazada por el número obtenido.

=ALEATORIO()*(máximo — mínimo) + mínimo

Es exactamente igual a la anterior y vale todo lo que dijimos, respecto a volatilidad y modo de ingreso. La diferencia es que genera un número con decimales y distribución uniforme entre ***mínimo*** y ***máximo***. Por ejemplo:

=ALEATORIO()*(50 – 10) + 10, podría generar 38,6782135

=ALEATORIO.ENTRE(mínimo; máximo)

Genera un número aleatorio ENTERO con distribución uniforme entre ***mínimo*** y ***máximo*** y valen todas las consideraciones hechas para los casos anteriores respecto a la volatilidad y modo de ingreso. Por ejemplo:

=ALEATORIO.ENTRE(10;50), podría generar 38

{=MATRIZALEAT(filas; columnas; inferior; superior;entero/decimal)}

Genera una matriz de **filas** x **columnas** entre un número **inferior** y otro **superior**, ambos números enteros y mayores de cero**.** Si pulsamos las teclas "**Control**"+"**Mayúsculas**"+"**Enter**" dará una matriz volátil. Si esa combinación de teclas la reemplazamos por pulsar solamente **F9** nos dará una matriz fija. Con el parámetro **entero/decimal** puesto en **VERDADERO** o "**1**" dará una matriz de números enteros, en **FALSO** o "**0**", de números decimales. Ejemplos:

```
=MATRIZALEAT(2;3;2;60;VERDADERO
MATRIZALEAT([filas]; [columnas]; [min]; [max]; [entero])
```

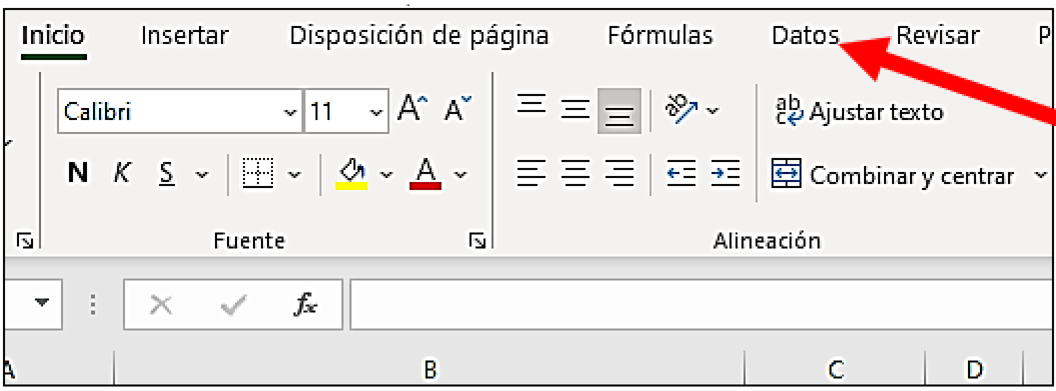
53	38	9
49	15	15

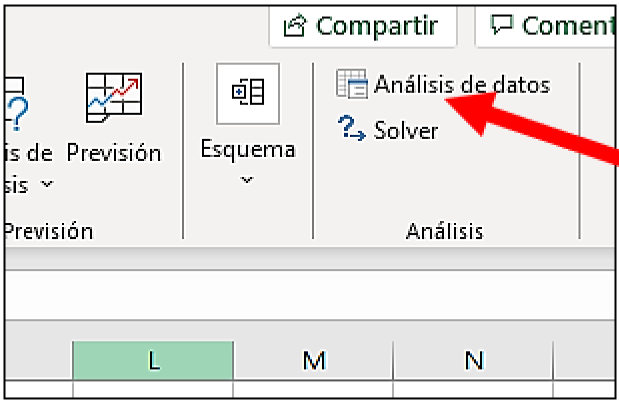
```
=MATRIZALEAT(2;3;5;60;0
MATRIZALEAT([filas]; [columnas]; [min]; [max]; [entero])
```

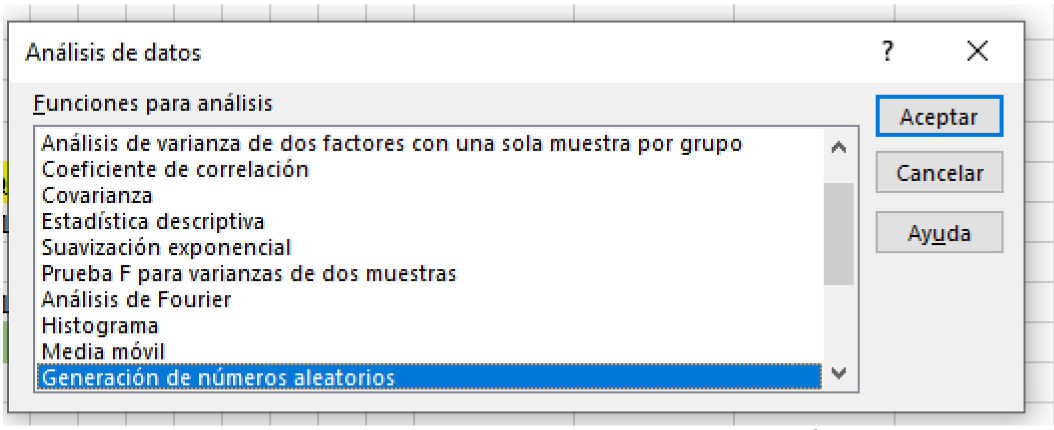
37,4030203	34,9223992	16,0909604
36,1353192	29,1752625	35,256989

Existen otros trucos para obtener estos tipos de números aleatorios. Son simples y dependen de cada usuario o circunstancia. Por ejemplo, una secuencia de números aleatorios con decimales entre 0 y 100 vimos arriba que se puede obtener con **=ALEATORIO(100 — 0) + 0** pero también podemos usar con **=ALEATORIO()*100**

Generador en “Análisis de datos”.
En la pestaña “Datos” de Excel, encontraremos el botón “Análisis de datos”[64], que muestra una lista de funciones, entre ellas “Generación de números aleatorios”



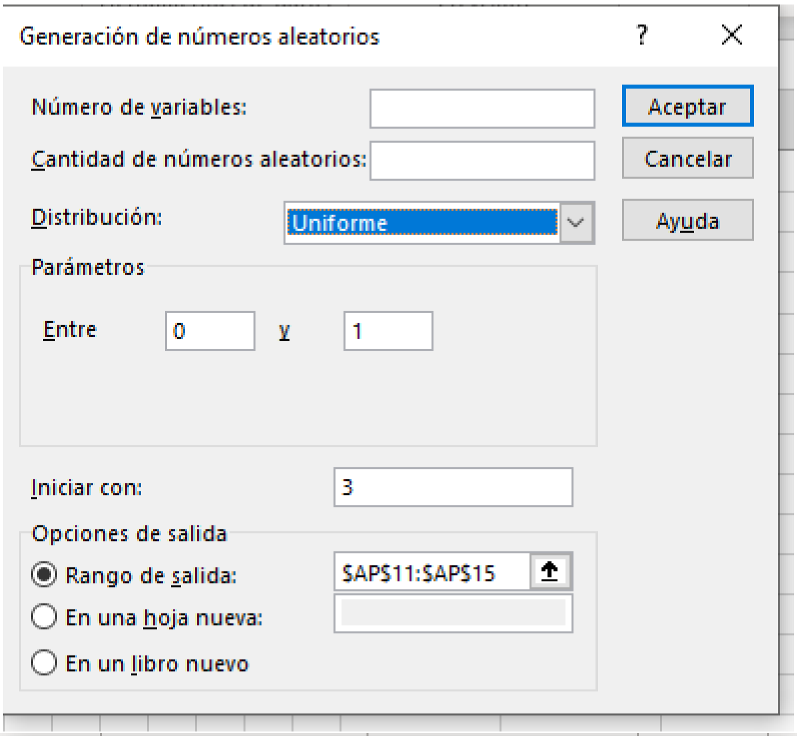


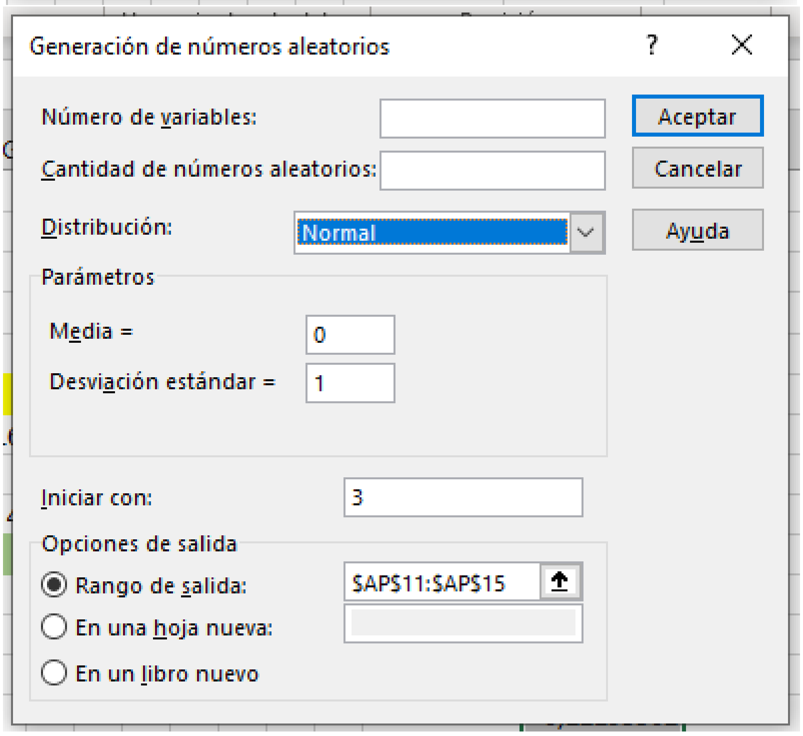


Al dar clic en aceptar podemos obtener un conjunto de números aleatorios en un rango de celdas y con distintas distribuciones. Si se elige “*Uniforme*” aparecerá la posibilidad de que varíe entre 0 y 1 u otros valores a elegir y la cantidad de números independientes a obtener. Si queremos convertir a una distribución normal, debemos

[64] Requiere instalación igual que Solver, ya que la instalación típica básica de Excel por defecto no la incluye. Ver capítulo 4 en el Libro 1 de esta serie.

proporcionar la media y la varianza a la que se deben ajustar los números que obtendremos.





Generador RND de Visual Basic for Applications

Si estamos trabajando con macros o en un programa basado en Visual Basic, como es el caso de *Visual Basic for Aplications* que se puede usar tanto en Excel como en Calc, la función disponible es ***RND(Número)*** que podemos usar de varias maneras:

a) ***a** = RND* ‘Guarda en *a* un valor aleatorio entre 0 y 1 con distribución uniforme

b) ***a = RND(—x)*** ´donde ***x*** es un número, guarda en *a* un valor aleatorio que se repetirá en cada llamada, en este caso ***x*** también la semilla
c) ***a = RND(x)*** ´donde ***x*** es un número, guarda en *a* un valor aleatorio que será el siguiente de la secuencia con semilla ***x*** en cada llamada.

Podemos usarlo precedido de las funciones ***RANDOMIZE*** y ***RANDOMIZE(x)*** donde ***x*** es un número o una función, que inicializa el generador de la siguiente manera, según su sintaxis:

1. **NO SE USA RANDOMIZE**. En este caso, cada llamada a ***Rnd*** (sin otro parámetro) usará el reloj del sistema como nueva semilla la primera vez y el último número generado las veces subsiguientes.
2. **RANDOMIZE(número)**. Inicializa la secuencia de llamadas **Rnd.** Usar el mismo número no significa que la secuencia se repetirá. En lugar de **número** se puede escribir una función, por ejemplo **timer** que llamará al reloj del sistema.
3. **Secuencia Rnd(—número)** y luego **RANDOMIZE(número)**. En este caso se generará una secuencia que se repetirá idéntica en cada corrida del programa.

Ejemplos:

```
Range("P1").Select
    ActiveCell.FormulaR1C1 = Rnd
```

Devuelve (por ejemplo) 0,70514755 en la celda P1 y un valor diferente cada vez que se corra la macro.

```
Range("P1").Select
    ActiveCell.FormulaR1C1 = Rnd(-35)
```

Devuelve 0,55133802 en la celda P1 cada vez que se corra la macro. En este caso el valor 35 es la semilla que se usó y que siempre devolverá 0,55133802.

```
Dim numero(10)
 Randomize
  For a = 0 To 9
    numero(a) = Rnd(-35)
    Next a
End Sub
```

Crea un vector llamado ***número*** con 10 elementos y guardará en ella diez veces repetidos el mismo número aleatorio, originado con ***Randomize***, entre 0 y 1 que será igual en cada corrida. Una vez más, 35 es la semilla y el primer número de los 10 que devolverá será el del ejemplo anterior.

```
Randomize
    For i = 0 To 9
    Cells(i + 3, 6) = Rnd(-35)
    Next
```

Escribe una columna con 10 números iguales. Es el equivalente del ejemplo anterior, pero ahora los números se guardan en una columna de la hoja de cálculo.

```
   Cells(3, 6) = Rnd(-35)
    For i = 1 To 9
    Cells(i + 3, 6) = Rnd(35)
```

```
    Next
```
Escribe la misma columna cada vez que se ejecuta. Observemos que el primer elemento de la columna es el de los ejemplos anteriores. Los siguientes serán una serie de números aleatorios que se repetirá idéntica en cada corrida porque Rnd(35) indica que genere a partir del número anterior. Si en lugar de Rnd(35) escribimos Rnd o Rnd(), el resultado no varía.

```
Dim numero(10)
 Randomize
  For a = 0 To 9
    numero(a) = Rnd()
    Next a
End Sub
```
Crea un vector llamado ***número*** con 10 elementos y guardará en ella diez números aleatorios, originados con ***Randomize***, entre 0 y 1 que serán diferentes en cada corrida.

```
For i = 0 To 9
    Cells(i + 3, 6) = Rnd(35)
    Next
```
Crea en la hoja una columna igual que la del ejemplo anterior. Rnd(35) no tiene efectos diferentes que Rnd(), salvo que los valores serán generados desde el número anterior, lo que no tiene ingerencia en este ejemplo.

```
Rnd (-2)
  Randomize (2)
    For a = 0 To 9
    Cells(a + 2, 1) = Rnd
    Next
```
Crea en la hoja una columna de 10 números aleatorios que se va a repetir exactamente igual en todas las corridas del programa. En el ejemplo se usaron los números — 2 y 2, pero podría usarse cualquier valor sin necesidad de repetirlos, con la condición de que el primero sea negativo, por ejemplo: —33267 y 23

Transformación del número aleatorio en variable aleatoria

Si bien es cierto que algunos paquetes ofrecen variables aleatorias con diversas distribuciones, partiremos de la suposición de que lo que obtuvimos al correr un generador aleatorio convencional es una serie con distribución uniforme independiente, y que necesitamos transformarla en una serie con una determinada distribución, por ejemplo, normal, beta, exponencial[65].

Llamaremos *X* a la variable aleatoria continua, para la función de densidad *f(x)* y para la función de distribución *F(x)=P(x<=x)*.

Para lograr estas conversiones disponemos de varios métodos, por ejemplo:

1. Método de inversión
2. Método de rechazo

[65] En este punto se recomienda la lectura del Capítulo 17, Libro 4 de esta serie.

Método de inversión

Este método requiere que conozcamos la función explícita a la cual queremos convertir los valores, de manera tal que $F_X(x)=F(x)=F^{-1}(U)$.

Veamos un ejemplo. Si tenemos una función de densidad de una distribución exponencial, como

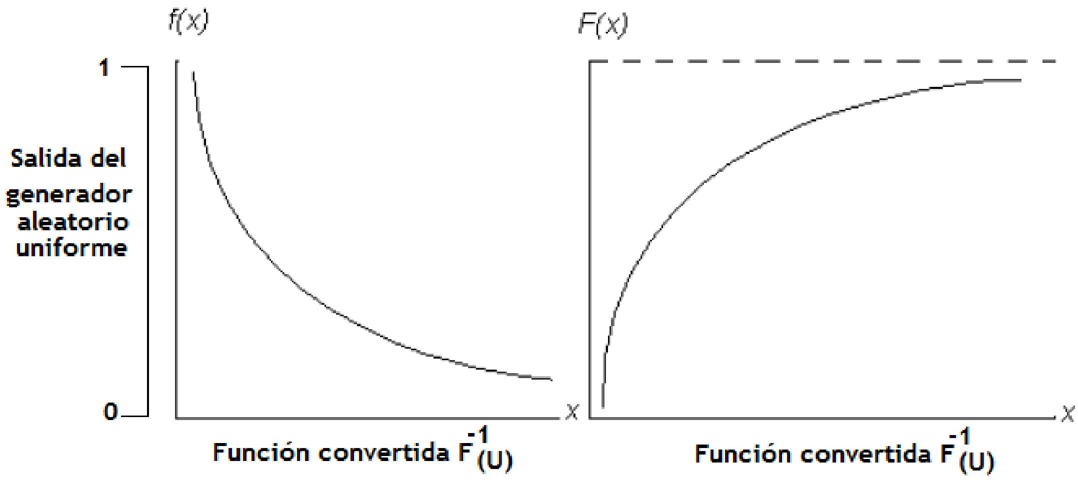
$$f(x) = \frac{1}{\lambda} e^{-\lambda t}$$

tendremos la función de distribución

$$F(T) = P(T \leq t) = 1 - e^{-\lambda t}$$

$$f(t) = \frac{1}{\lambda} e^{-\lambda t}$$

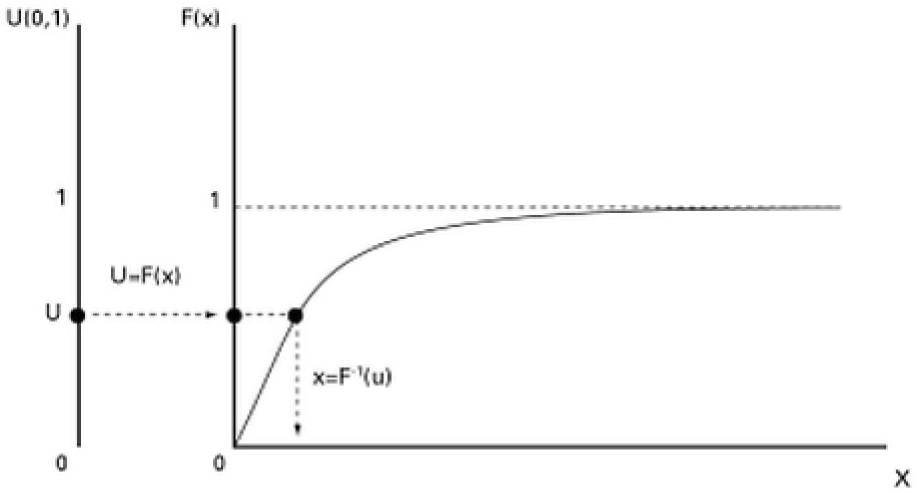
$$F(T) = P(T \leq t) = 1 - e^{-\lambda t}$$

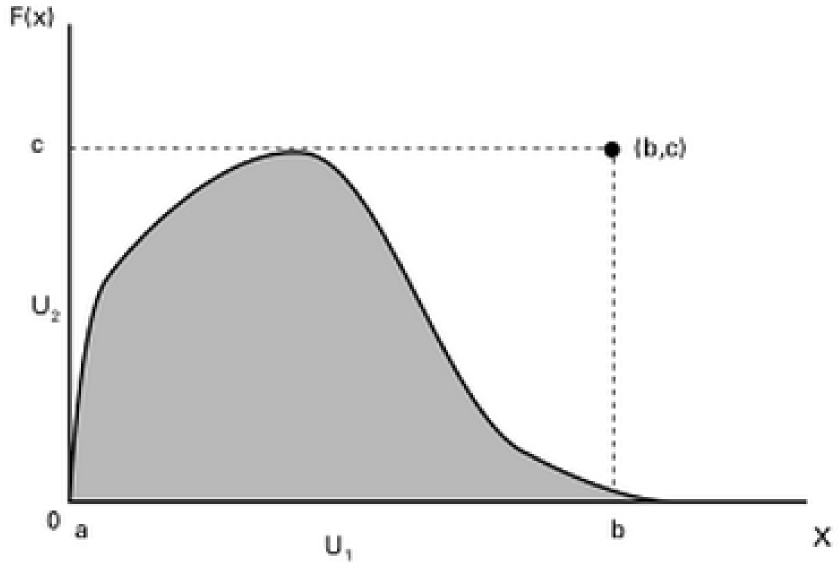


Método de rechazo

Este método puede ser aplicado cuando:

1) La función de densidad de probabilidad de la variable tiene forma compleja.
2) La variable toma valores en el intervalo *a*, *b*
3) La función de densidad *f(x)* está acotada a un valor conocido *c*.





Si las condiciones 2 y 3 se cumplen, entonces se aplican los siguientes pasos:

a) Generamos dos números aleatorios con distribución uniforme U_1 en el intervalo |*a*, *b*| y U_2 en el intervalo |0, *c*|, de esta manera obtendremos un punto P cuyas coordenadas podrán caen dentro del rectángulo definido para la función de densidad *f(x)*.
b) Si el punto obtenido cae efectivamente dentro del área que se definió con *f(x)*, entonces aceptamos $x = U_2$ y rechazamos en caso contrario.

Otros métodos

Existen otros métodos que se aplican cuando la solución por el método anterior se vuelve muy compleja o, directamente, no puede aplicarse. Necesitaremos definir herramientas tales como funciones de densidad y/o cantidad de iteraciones. En capítulos siguientes se pueden consultar algunos ejemplos.

CAPÍTULO 5

APLICACIÓN DEL MÉTODO DE MONTECARLO

Se conoce como *técnica* o *método de Montecarlo* a un procedimiento básico en la simulación de sistemas en los que hay objetos cuyo comportamiento puede ser aleatorio o con distribuciones de probabilidad no conocidas. Dicho en otras palabras, se trata de una técnica que nos permitirá simular "observaciones" del comportamiento de variables estocásticas que, en realidad fueron generadas por nosotros.

El método se originó en 1940 a partir de trabajos (que en esos momentos eran extremadamente secretos) realizados por Von Neumann y Stanislaw Ulman[66] con el objetivo de resolver problemas de difusión nuclear. La técnica matemática, que era conocida desde muchos años antes y fue revivida por estos trabajos, adquirió ese nombre y rápidamente se aplicó en otros campos, como la simulación[67].

El uso de la técnica de Montecarlo es útil en simulación probabilística. Aunque también lo es en ciertos modelos completamente deterministas que no pueden ser resueltos analíticamente. Por ejemplo: calcular una integral doble sin primitiva en una región del plano, puede hacerse con la generación de números al azar en una zona que contiene la región de integración (veremos ejemplos más adelante).

La Técnica

Las variables estocásticas deben tener una función de distribución de probabilidad asociada (continua o discreta). Esa función, que muchas veces no es explícita, puede ser aproximada en base a:

- Datos empíricos derivados del pasado.
- Experimentos recientes.
- Una distribución teórica conocida, que sea apropiada.

---

[66] **John Von Neumann**, fue un matemático nacido en 1903 en Budapest y fallecido en 1957 en Washington. Hizo importantes aportes a la física cuántica, teoría de juegos, computación, estadística y otros campos, al punto que se lo considera uno de los matemáticos más importantes de la historia.
Stanislaw Ulam, nacido en 1904 y fallecido en 1984, de origen polaco, fue matemático y trabajo en varios campos, algunos aplicados a la física atómica

[67] Hay una versión sobre el nombre del método. Ulam y otros, trabajaban en el Laboratorio Nacional de Los Álamos, que era en ese momento muy secreto y estaba destinado a desarrollar y producir la primera bomba atómica. Debían resolver una serie de integrales complejas para estudiar la difusión nuclear y decidieron usar números al azar. Algunas fuentes aseguran que Ulam, jugando a las cartas, tuvo la idea de que un buen generador de números aleatorios podía ser una ruleta. La más famosa del mundo está ubicada en Montecarlo, principado de Mónaco, y de ahí surgió el nombre del método.

En general se busca asociar una cierta distribución de probabilidad a una variable estocástica antes que transformarla en una constante.

Si no sabemos qué función de distribución de probabilidad tiene una determinada variable continua estocástica, asumiremos que tiene una distribución uniforme en un rango acotado de valores posibles. En cuyo caso la probabilidad es igual para todos los puntos; *f(x)=k*; y tendremos que la función de probabilidad acumulada es *F(x)*,

Como *k* es la probabilidad, y la distribución es uniforme en un intervalo cualquiera, supongamos desde *a* hasta *b*, la probabilidad total será igual a uno y podemos expresarla así:

$$k = \int_a^b k dx = k(b-a) = 1$$

por lo tanto, podemos despejar:

$$k(b-a) = 1$$

$$k = \frac{1}{b-a}$$

Ahora calcularemos la probabilidad acumulada en el intervalo *a* hasta la variable *x*:

$$F(x) = \int_a^x \frac{1}{b-a} dt = \frac{x-a}{b-a}$$

Caso 1

Supongamos que la variable aleatoria es: "*número de accidentes que requieren atención en un día*", y que tenemos información histórica que nos permite calcular con que probabilidad ocurrieron. Esa información está en la siguiente tabla:

N° de accidentes/dia	**Probabilidad**	**P acumulada**	**Intervalo asociado**
0	0,35	0,35	[0 , 0,35)
1	0,40	0,75	[0,35 , 0,75)
2	0,15	0,90	[0,75 , 0,90)
3	0,05	0,95	[0,90 , 0,95)
4	0,02	0,97	[0,95 , 0,97)
5 o más	0,03	1,00	[0,97 , 1)

Nos proponemos simular lo que podría ocurrir en una semana cualquiera de 7 días. Lo haremos por simulación y comenzamos generando para cada día el número de "accidentes que se producen ese día y que requieren atención".

Usamos una tabla de números aleatorios uniformemente distribuidos entre 0 y 1, y, con los ojos cubiertos marcamos ("obtenemos") siete números *U(0,1)*, uno para cada día, que, son los siguientes:

0,36 0,83 0,42 0,32 0,49 0,93 0,16

Ahora a cada número obtenido le asignamos dos significados: 1) Ordinal: el orden en que fue obtenido será el día de la semana de 1 a 7. 2) cardinal: El valor numérico lo haremos corresponder a la probabilidad histórica de ese día, según la segunda columna de la tabla anterior.

De esta manera ahora podemos obtener para cada día la cantidad de accidentes asociada con esa probabilidad, usando la primera y segunda columna.

Resultado:

Día de la semana	Probabilidad simulada	N° accidentes de tabla I
1	0,36	1
2	0,83	2
3	0,42	1
4	0,32	0
5	0,49	1
6	0,93	3
7	0,16	0

El 1er día, se produce: 1 accidente
El 2° día se producen: 2 accidentes
El 3er día se produce: 1 accidente
El 4° día se produce: 0 accidente
El 5° día se produce: 1 accidente
El 6° día se producen: 3 accidentes
El 7° día se produce: 0 accidente

Con esta información podemos continuar simulando más semanas, y/o sacar conclusiones estadísticas respecto del comportamiento de la variable.

Caso 2
Queremos generar las llegadas de clientes a un local comercial durante una hora. Durante unas semanas hemos medido la afluencia de clientes y por eso sabemos que en 5 minutos la probabilidad de número de clientes que pueden llegar es:

N° de clientes	Probabilidad	Prob, Acumulada	Intervalo de probab, asignado
0	0,25	0,25	[0 , 0,25)
1	0,40	0,65	[0,25 , 0,65)
2	0,20	0,85	[0,65 , 0,85)
3	0,15	1,00	[0,85 , 1)

Buscamos o generamos doce números al azar uniformemente distribuidos en [0;1) para usarlos con criterio similar al anterior, pero la gran diferencia en este caso es que ahora vamos a ver en que intervalo de probabilidad asignado (última columna) cae cada número aleatorio obtenido:

Números aleatorios obtenidos:
1) 0,492 2) 0,871 3) 0,753 4) 0,122 5) 0,333 6) 0,677
7) 0,469 8) 0,010 9) 0,905 10) 0,507 11) 0,646 12) 0,745

Reemplazamos ahora cada número aleatorio por el número de clientes cada 5 minutos que le corresponde según la primera columna de la tabla, con lo cual tenemos la llegada de clientes cada doce grupos de cinco minutos, o sea, durante una hora:

1, 3, 2, 0, 1, 2, 1, 0, 3, 1, 1, 2.

Pasos para aplicar la técnica de Montecarlo

Podemos resumir la metodología que usamos cuando empleamos la técnica de Montecarlo en cinco pasos:

1. Obtener y, eventualmente, graficar la curva de frecuencias acumuladas con los valores de la variable en las abscisas, y la probabilidad acumulada en las ordenadas. Los valores están en el rango comprendido entre 0 y 1.

2. Obtener o elegir un número decimal ***U*** entre 0 y 1 uniformemente distribuido con tantos decimales como se desee, por medio de un generador de números aleatorios.

3. Obtener la imagen ***X*** de ese número ***U*** por medio de la función de frecuencia acumulada. Es decir, obtener ***X*** tal que
$$P[X \leq x] = U$$

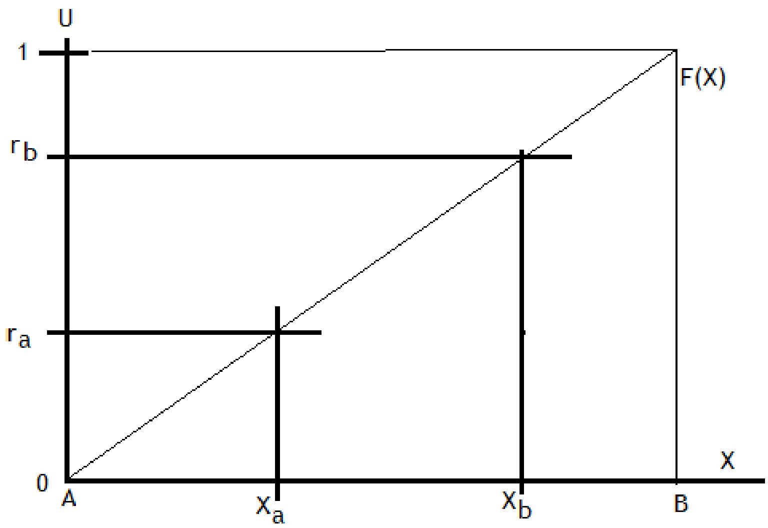
4. Ese valor ***X*** obtenido es el valor muestreado convertido a partir del número generado uniformemente.

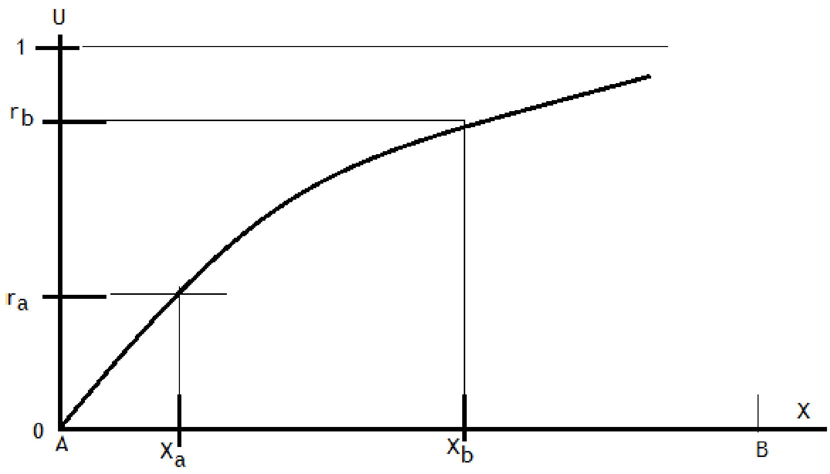
5. Repetir los pasos 2 y 3 hasta generar el número de observaciones que se desee.

Base de método:

El método de muestreo de Montecarlo se basa en el supuesto siguiente:
"La probabilidad de que un número uniformemente distribuido entre 0 y 1 caiga en el intervalo $[r_a , r_b)$ es: $r_b - r_a$".

En los gráficos siguiente vemos el rango de probabilidad del número generado (U), en el eje *y* vs la salida correspondiente al rango que buscamos (*x*). Se obtiene un número entre *a* y *b*, entramos por el eje ***y***, operamos con la función *F(x)* y leemos en el eje ***x***.





$$P_{(x_a \leq X \leq x_b)} = P_{(X \leq x_b)} - P_{(X \leq x_a)} = r_b - r_a$$

Así, con números uniformes en [0, 1) podremos generar valores de la variable X conociendo su función de probabilidad acumulada, obtenida a partir de su función de densidad de probabilidad.

Distribución exponencial
Tomemos como primer caso una distribución cuya función conocemos. (Ver Libro 4, Líneas de Espera y Simulación)

$$f(x) = \lambda e^{-\lambda x}$$

Su función de probabilidad acumulada ***F(x)*** es:

$$F(x) = \int_0^x \lambda e^{-\lambda t}\, dt = \frac{-e^{-\lambda x}}{0^x} = 1 - e^{-\lambda x}$$

cuando $x \to +\infty$, la probabilidad acumulada tiende a 1 (ver figura anterior)

En este caso también podemos obtener una expresión de términos de ***U*** para el valor de ***X*** muestreado.

$$U = 1 - e\text{-}\lambda x$$
$$e\text{-}\lambda x = 1 - U$$
$$X = -\frac{\ln(1 - U)}{\lambda}$$

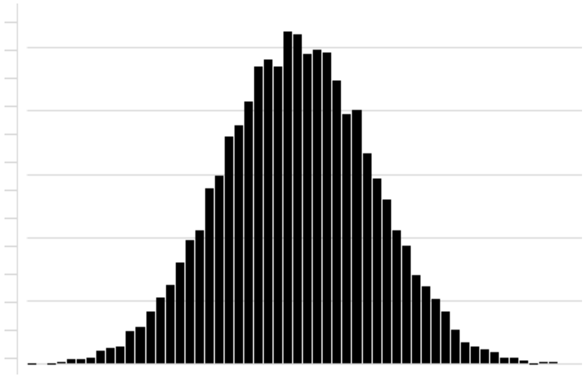
que nos da el valor que adquiere *X* para un *r* = ***U*** dado obtenido con un generador de números aleatorios.

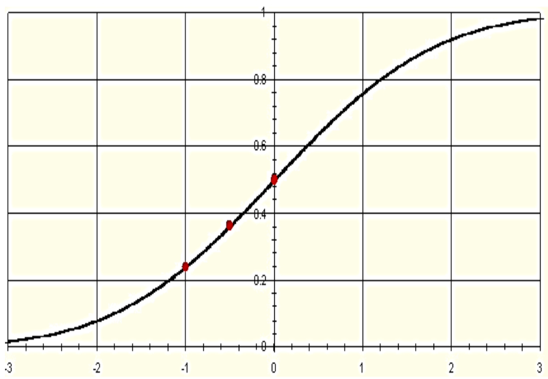
Ejemplo: obtenemos el número

U = 0,456789

significa que obtuvimos X = 0,61025 para λ = 1

Distribución normal
Si la distribución que usamos fuera un caso de distribución normal





$$f(x) = N(\mu, \sigma^2)$$

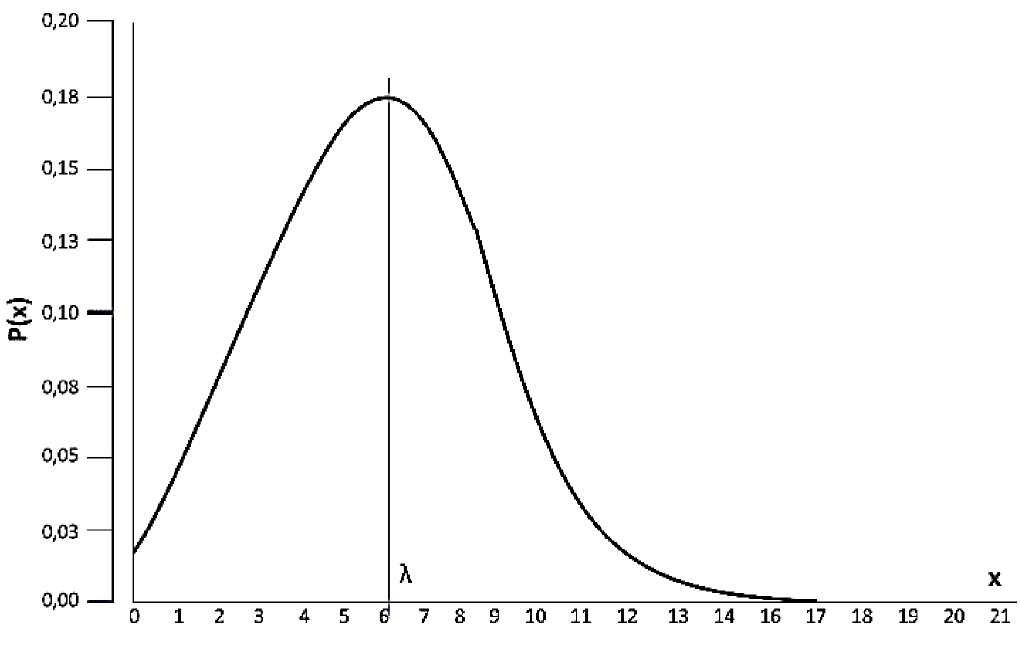
$$N(0,1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{Z^2}{2}} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$F(x) = \int_{-\infty}^{x} f(x)dt$$

Distribución de Poisson.

$$P(x; \lambda) = \frac{e^{-\lambda}\lambda^x}{x!}$$

Obtenemos la probabilidad de que se presenten exactamente x eventos en un intervalo de tiempo dado. λ es el promedio de sucesos que se presentan en ese tiempo.



Otras distribuciones o distribuciones desconocidas

Si tenemos una función continua sin inversa para la función de probabilidad acumulada, o de muy difícil expresión algebraica y/o de difícil manejo, podremos usar tablas que contengan los valores de F o bien, recurrir a discretizar esa función; como ocurre con la función de probabilidad acumulada de la curva normal.

Integrar funciones con el método de Montecarlo

Supongamos que tenemos una variable *X* continua uniforme aleatoria como las que nos provee cualquier generador de números pseudoaleatorios en un rango [0, 1]. La función de densidad de esa variable, como dijimos más arriba será:

$$f(X) = \frac{1}{1-0} = 1 \; sii \; X \in [0,1]$$

$$f(X) = \frac{1}{1-0} = 0 \; sii \; X \notin [0,1]$$

Ahora establecemos otra función cualquiera Y=G(x), cuya media o esperanza matemática podremos definirla así:

$$E[G(X)] = \int_0^1 G(X)f(x)dx$$

Como *f(X)* = 1 en el rango [0,1], la expresión anterior queda así.

$$\int_0^1 G(X)f(x)dx = \int_0^1 G(X)dx = E[G(X)]$$

Si generamos y usamos una cantidad M de números aleatorios uniformes como puntos de muestra, podemos calcular la esperanza matemática aproximadamente como la sumatoria del valor de la función en cada punto de la muestra dividido entre el total de puntos (o sea, la media):

$$E[G(X)dx \cong \frac{1}{M}\sum_{i=1}^{M} G(X_i)$$

Con esta expresión podríamos calcular cualquier integral en el intervalo [0, 1]. Es difícil que necesitemos hacer una integral en ese intervalo, entonces, para integrar entre cualquier límite fuera de ese intervalo tendremos que hacer un cambio de variable.

Si llamamos *a* al límite de integración **inferior** y *b* al límite **superior**, entonces llamaremos

$$x = a + (b-a)u$$

donde *u* será un número cualquiera correspondiente al intervalo [0,1]. Esta expresión puede ser derivada en ambos lados, y nos queda

$$dx = (b-a)du$$

La integral se convierte en:

$$\int_a^b G(X)dx = \int_0^1 G(a+(b-a)u)(b-a)du =$$

$$(b-a)\int_0^1 G(a+(b-a)u)du$$

entonces, usando la ley de los grandes números, podemos decir que

$$\int_a^b G(X) \cong \frac{b-a}{M}\sum_{i=1}^{M} G(a+(b-a)u_i)$$

Si esos números u_i provienen de un generador de números aleatorios, podremos resolver la integral.

Ejemplo 1: Calculo del área bajo la curva $\boldsymbol{y = x}$ usando hoja de cálculo
Comenzamos estableciendo límites de integración. Supongamos que vamos a integrar entre 0 y 10:

$$f(x) = \int_0^{10} xdx$$

Podemos preparar una hoja que luego nos sirva para otros valores o límites de integración, como la que vemos abajo.

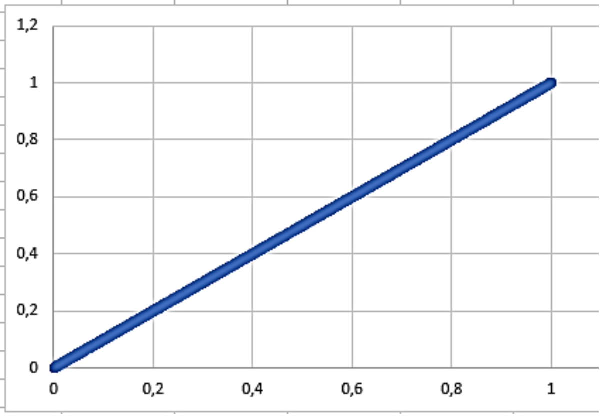
	A	B	C	D	E	F	G	H
1	GENERADOR DE NUMEROS ALEATORIOS					Limite inferior X	0	
2						Limite superior X	10	
3								
4						Limite inferior f(x)	0	
5						Limite superior f(x)	10	
6								
7								

Luego procedemos a construir un generador de números aleatorios, en la columna A, que nos represente la variable X, destinamos la columna B para convertir ese valor de X en f(x) (en este ejemplo son iguales porque f(x)=X. Generaremos, por ejemplo, 3000 números aleatorios.

	A	B	C	D	E	F	G	H
1	GENERADOR DE NUMEROS ALEATORIOS					Limite inferior X	0	
2	X	f(X)				Limite superior X	10	
3	0,9243464	0,9243464						
4	0,41359482	0,41359482				Limite inferior f(x)	0	
5	0,48218013	0,48218013				Limite superior f(x)	10	
6	0,08367833	0,08367833						
7	0,07637622	0,07637622						
8	0,14625625	0,14625625						
9	0,75355718	0,75355718						
10	0,31794079	0,31794079						

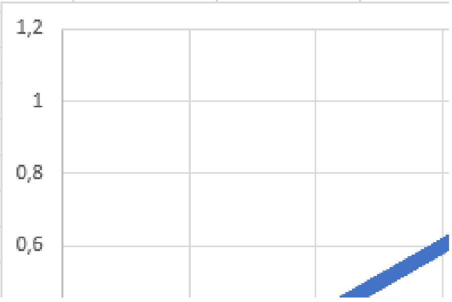
Podemos ahora graficar X vs f(X),

	A	B	C	D	E	F	G	H	I	J
1	GENERADOR DE NUMEROS ALEATORIOS					Limite inferior X	0			
2	x	f(x)				Limite superior X	10			
3	0,18806031	0,18806031								
4	0,2819631	0,2819631				Limite inferior Y	0			
5	0,44091455	0,44091455				Limite superior y	10			
6	0,6053476	0,6053476								
7	0,12361295	0,12361295								
8	0,15226755	0,15226755								
9	0,59784602	0,59784602								
10	0,6317284	0,6317284								
11	0,7405947	0,7405947								
12	0,86364434	0,86364434								
13	0,01723128	0,01723128								
14	0,83912571	0,83912571								
15	0,59386566	0,59386566								
16	0,74985065	0,74985065								
17	0,61933957	0,61933957								
18	0,28888623	0,28888623								
19	0,85466673	0,85466673								
20	0,92951859	0,92951859								



Generamos ahora una tercera columna “Y simulado” destinada a producir valores aleatorios para comparar entre los de la función f(x).

	A	B	C	D	E	F	G	H
1	GENERADOR DE NUMEROS ALEATORIOS					Limite inferior X	0	
2	x	f(x)	y simulado			Limite superior X	10	
3	0,02218047	0,02218047	0,41496774					
4	0,95434575	0,95434575	0,95939158			Limite inferior Y	0	
5	0,91000504	0,91000504	0,88905398			Limite superior y	10	
6	0,29170617	0,29170617	0,67967961					
7	0,09781156	0,09781156	0,70333832					
8	0,63200931	0,63200931	0,77242066					
9	0,01676765	0,01676765	0,86311812					
10	0,15097541	0,15097541	0,79892512					
11	0,9244085	0,9244085	0,80170452					
12	0,45888801	0,45888801	0,49533166					
13	0,5137718	0,5137718	0,94002569					
14	0,41248924	0,41248924	0,49773761					

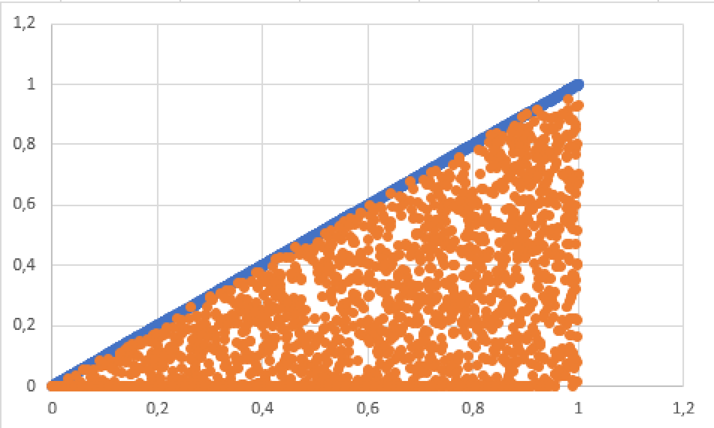


A continuación averiguaremos si el punto simulado está en el área debajo de la curva o no, comparando, en este caso, simplemente, si el valor simulado es menor o no que el correspondiente a la f(x), usando la función SI

A	B	C	D	E
GENERADOR DE NUMEROS ALEATORIOS				Limit
x	f(x)	y simulado		Limite
0,02218047	0,02218047	0,41496774	=Si(C3<B3;C3;0)	
0,95434575	0,95434575	0,95939158		Limit
0,91000504	0,91000504	0,88905398		Limite
0,29170617	0,29170617	0,67967961		

Incluimos estos nuevos puntos filtrados con la función “SI” en el gráfico anterior como una nueva serie

(x)	y simulado			
			Limite superior X	10
0,58018964	0,35050275	0,35050275		
0,08118648	0,02876586	0,02876586	Limite inferior Y	0
0,78816231	0,89237161	0	Limite superior y	10
0,92523353	0,32905081	0,32905081		
0,81715678	0,39358991	0,39358991		
0,30728104	0,34083937	0		
0,30341741	0,28795452	0,28795452		
0,87060265	0,0005026	0,0005026		
0,31159281	0,79818561	0		
0,88248673	0,07707728	0,07707728		
0,88894001	0,82147428	0,82147428		
0,12675652	0,87615957	0		
0,78084296	0,70100676	0,70100676		
0,37327757	0,21753074	0,21753074		
0,82157518	0,92400753	0		
0,44430944	0,72285792	0		
0,15698057	0,04483686	0,04483686		
0,30532605	0,3504399	0		
0,38007277	0,55333534	0		

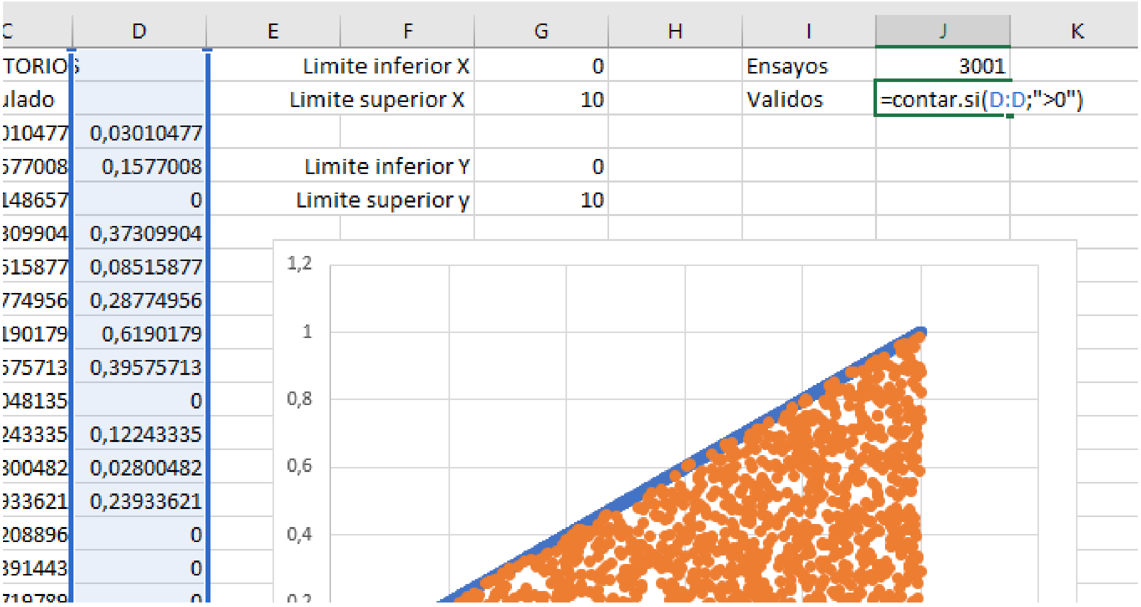


Y ahora solo nos queda realizar la contabilidad y presentación de resultados. Contamos cuantos puntos graficamos:

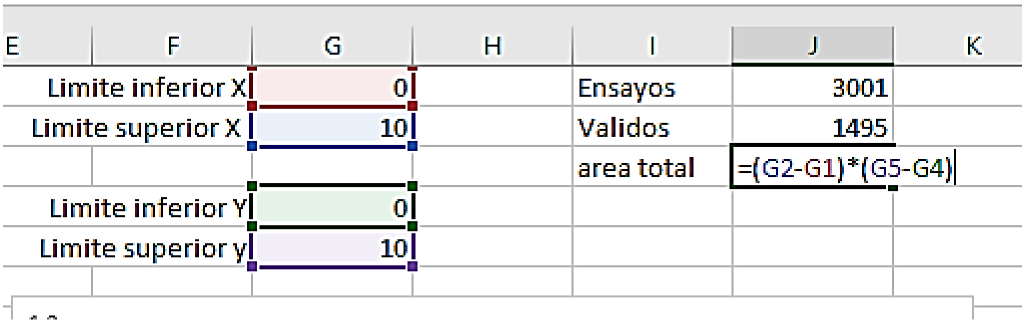
fx =contar(A:A)

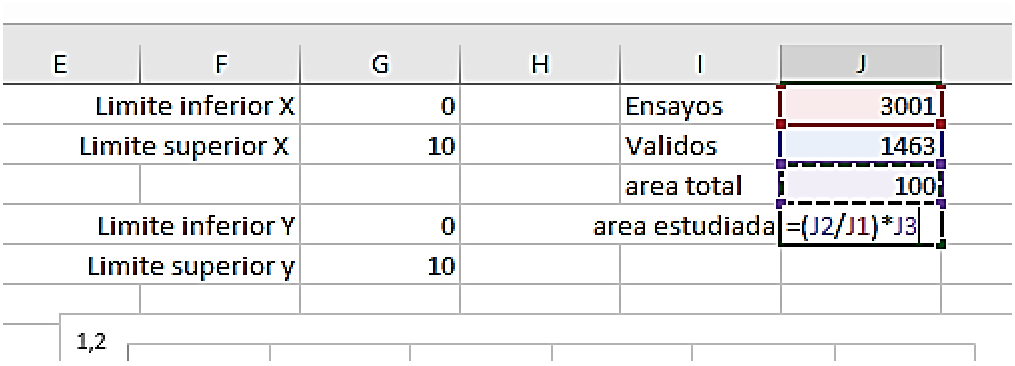
C	D	E	F	G	H	I	J
ATORIOS			Limite inferior X	0		Ensayos	=contar(A:A)
ıulado			Limite superior X	10			
1927873	0						
4060192	0		Limite inferior Y	0			
2039457	0,12039457		Limite superior y	10			

Y contabilizamos cuántos de ellos son válidos, o sea, caen debajo de la f(x).

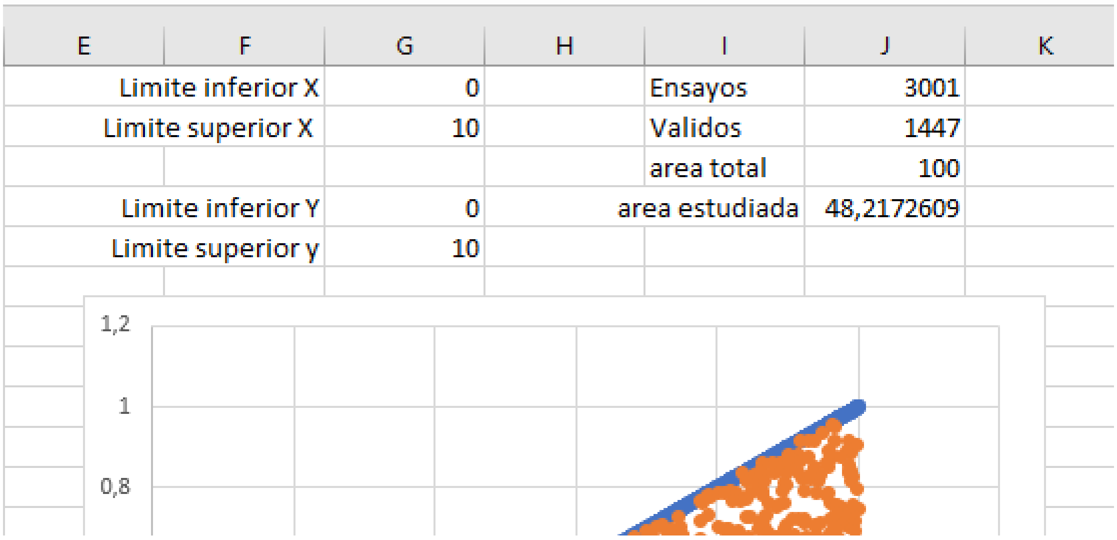


El área bajo la curva será el área graficada de todos los puntos simulados, un rectángulo de base ***límite superior de x — límite inferior de x*** y de altura ***límite superior de y — límite inferior de y***, y lo calculamos como área total.



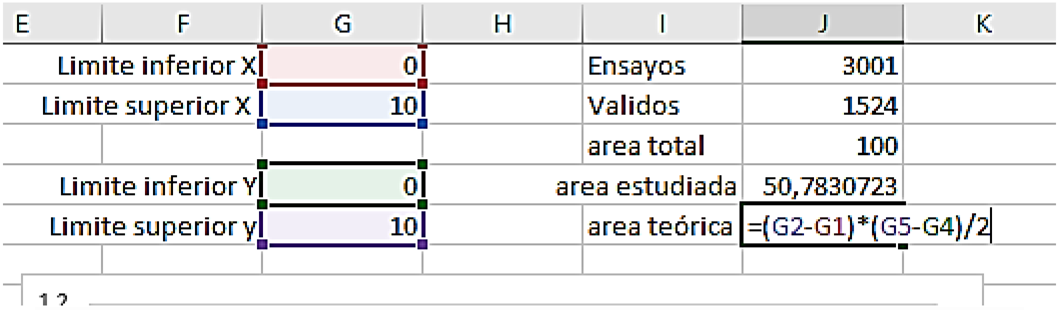


Resultando que el área encerrada bajo la f(x) es igual a la relación entre puntos válidos y totales multiplicado por el área total del rectángulo que tiene como lados la base del rectángulo y su altura.

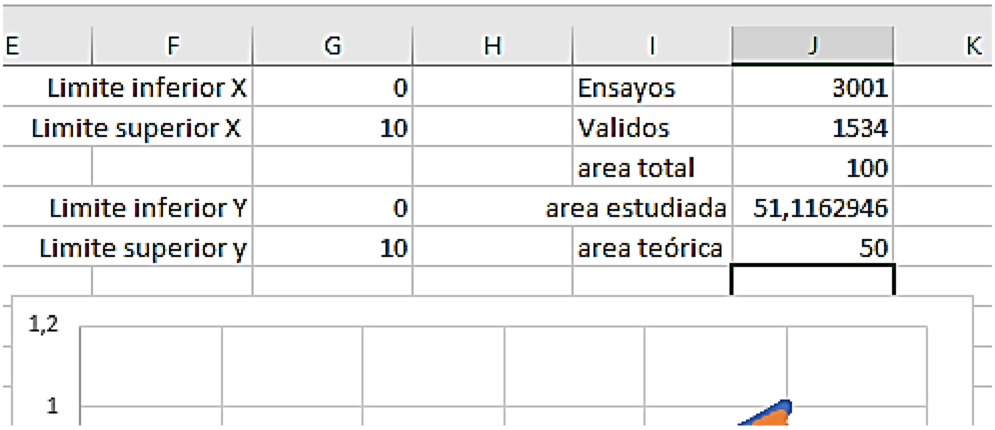


E	F	G	H	I	J	K
	Limite inferior X	0		Ensayos	3001	
	Limite superior X	10		Validos	1447	
				area total	100	
	Limite inferior Y	0		area estudiada	48,2172609	
	Limite superior y	10				

Podemos intentar ensayos cambiando los límites inferiores y superiores entre cualquier valor y así veremos que obtenemos rápidamente el área correspondiente, la cual se verificaría simplemente con el valor teórico del área del triángulo correspondiente [b*h/2]



E	F	G	H	I	J	K
	Limite inferior X	0		Ensayos	3001	
	Limite superior X	10		Validos	1524	
				area total	100	
	Limite inferior Y	0		area estudiada	50,7830723	
	Limite superior y	10		area teórica	=(G2-G1)*(G5-G4)/2	



E	F	G	H	I	J	K
	Limite inferior X	0		Ensayos	3001	
	Limite superior X	10		Validos	1534	
				area total	100	
	Limite inferior Y	0		area estudiada	51,1162946	
	Limite superior y	10		area teórica	50	

Vamos a probar ahora nuestro trabajo cambiando los límites a un valor cualquiera, por ejemplo 7 y 22 como límites inferior y superior, respectivamente.

E	F	G	H	I	J
	Limite inferior X	7		Ensayos	3001
	Limite superior X	22		Validos	1524
				area total	225
	Limite inferior Y	7		area estudiada	114,261913
	Limite superior y	22		area teórica	112,5

Ejemplo 2: cálculo del área de una circunferencia.

En este caso trataremos de hacer "por fuerza bruta" una hoja de Excel donde representaremos una circunferencia. Usamos para un radio cualquiera arbitrario que anotamos en la celda G1. Vamos a comenzar escribiendo radio 1 aunque puede ser cualquier otro, pero al poner 1estamos facilitando los cálculos.

Debajo multiplicamos el contenido de la celda G1 por 2, para obtener el lado de un supuesto cuadrado que inscribe el círculo de radio G1. Y debajo calculamos el área de ese cuadrado.

E	F	G	H
	Radio	1	
	Lado	2	
	Area	4	

Ahora, en la columna A, escribimos los ángulos en grados correspondientes a un círculo, (360°) usando un paso de 10 en 10. Cuanto más chico sea el paso, menos error obtendremos en la medición del área. En la columna B, multiplicaremos el valor de los ángulos de la columna A por π/180 para convertir los grados en radianes. (la primera fila de la columna B será "**=A2*Pi()/180**" y así las restantes con A3, A4, etc).

Lo siguiente, en la columna C, será crear un rango con el cálculo de la coordenada en X correspondiente al ángulo inicial (0°), que será el radio multiplicado por el seno del ángulo correspondiente. El radio lo tenemos en la celda $G$1 (como vemos en párrafo anterior), entonces queda [=$G$1*SENO(B2)].

La columna D, de manera semejante, será destinada a la coordenada en Y de ese mismo ángulo [=$G$1*COSENO(B2)].

Completamos la tabla hasta los 360°. Insistimos con lo que dijimos más arriba, puede ser que vayamos de 0 a 360 de a 10°, de a 5°, de a 1°. Variará la precisión del resultado.

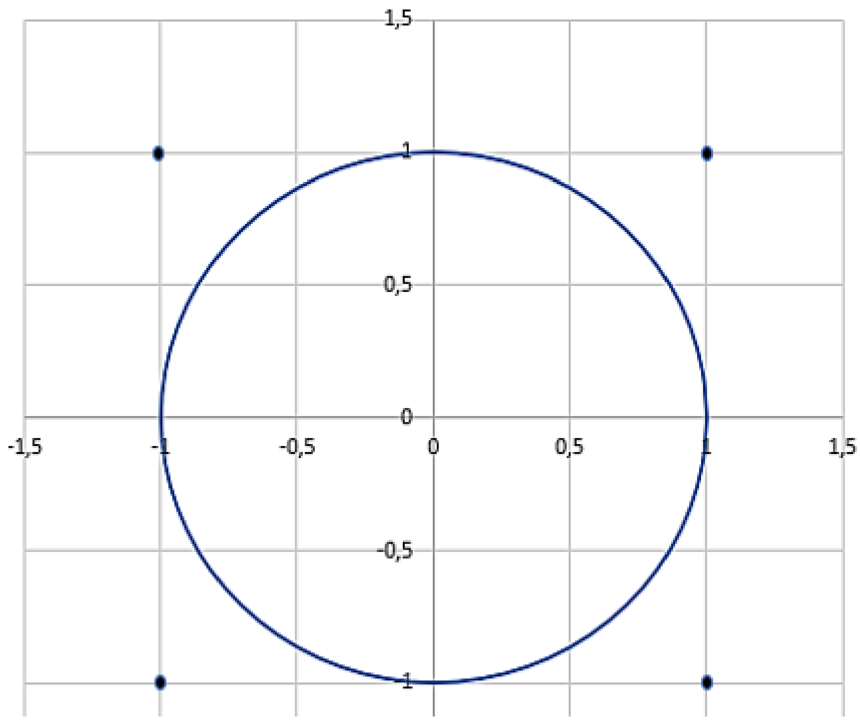
	A	B	C	D
1	Angulo	AnguloRad	**Eje X**	**Eje Y**
2	0	0	0	3
3	10	0,17453293	0,52094453	2,954423259
4	20	0,34906585	1,02606043	2,819077862
5	30	0,52359878	1,5	2,598076211
6	40	0,6981317	1,92836283	2,298133329
7	50	0,87266463	2,29813333	1,928362829
8	60	1,04719755	2,59807621	1,5
9	70	1,22173048	2,81907786	1,02606043
10	80	1,3962634	2,95442326	0,520944533
11	90	1,57079633	3	1,83772E-16
12	100	1,74532925	2,95442326	-0,52094453
13	110	1,91986218	2,81907786	-1,02606043
14	120	2,0943951	2,59807621	1,5

A continuación en otro sector de la hoja ponemos los limites siguientes y calculamos algunas áreas teóricas:

1) Ya dijimos antes que el radio de la circunferencia está en la celda $G$1
2) También que el lado del cuadrado que inscribe la circunferencia (2*radio ➔ $G$1*2) en la celda $G$2
3) El cálculo de las áreas teóricas del cuadrado y de la circunferencia (lado^2 ➔ $G$2^2) y (πr^2 ➔ PI()*$G$1^2) respectivamente.

F	G
Radio	3
Lado	6
Area cuadr	36
Area teorica	28,2743339

Graficamos las columnas “Eje X” vs “Eje Y” y obtendremos el círculo. También graficamos los cuatro vértices correspondientes al cuadrado, solo los vértices para simplificar el dibujo.



Generamos ahora valores aleatorios para cada ángulo, o sea, en una nueva columna para cada renglón que representará el valor “simulado del eje Y.

Usaremos la función

[=$G$1*COS(ALEATORIO())]

para los renglones que corresponden al primer y cuarto cuadrante.

La función

[=$G$1*SENO(-1*ALEATORIO())]

para los renglones correspondientes al segundo y tercer cuadrante

Con estas funciones lo que buscamos es convertir el valor obtenido en positivo o negativo, según el cuadrante donde nos encontremos, porque hemos elegido trabajar con una circunferencia cuyo centro está en las coordenadas [0 ; 0] y determina el origen de los cuatro cuadrantes donde centraremos el cuadrado de referencia.

Solo nos resta averiguar, para todos y cada uno de los puntos simulados, si cae dentro del círculo.

En el primer cuadrante (0° a 90°), haremos la prueba de manera tal que si el punto “simulado” es menor que el correspondiente al círculo en ese ángulo (en la ordenada), lo escribimos, de lo contrario dejamos la celda en blanco. Por ejemplo, para el renglón 2 (que es el primer renglón y le corresponde un ángulo de 0°) la función que debemos introducir sería

[=SI(L2<D2;L2;"")].

Esa misma, cambiando el número de renglón 2 por el que corresponda (L3, L4, etc.) será copiada hasta llegar a 90°.

Para el segundo y tercer cuadrante (91° a 180° y 181° a 270°), usaremos, por ejemplo en el renglón 12:

[=SI(L12>D12;L12;"")]

Y para el cuarto cuadrante (271° a 360°) lo mismo que para el primer cuadrante.

Solo nos queda contar el número total de ensayos con la función CONTAR en la columna donde "simulamos" los valores de Y, y contar el número de ensayos válidos, contando en la columna de valores validados. Luego simplemente calculamos la relación entre la ***cantidad de ensayos válidos/cantidad total de ensayos*** y multiplicamos por el área total del cuadrado. Eso nos dará el área de la circunferencia obtenida por simulación.

	E	F	G	H
28			Ensayos:	=CONTAR(L:L)
29			Válidos:	=CONTAR(M:M)
30			area total cuadrado	=G2*G2
31			area estudiada	=(H29/H28)*H30
32			rea teorica circunferencia	=(PI())*G1^2
33				

Ensayos:	37
Válidos:	24
area total cuadrado	36
area estudiada	23,3513514
area teorica circunferencia	28,2743339

Podemos ver que el error es importante, fundamentalmente por los pocos ensayos que realizamos, pero si hacemos todo este procedimiento con un paso de un grado, el cambio sería muy importante.

Ejemplo 3: Integral de la función $y = e^{-x^2}$

Vamos a integrar con el método de Montecarlo usando un muy sencillo programa, que en este caso realizamos en ***Visual Basic for App***, aprovechando la facilidad de usar una hoja de cálculo Calc o Excel.

Comenzaremos con la construcción de un ciclo con M acontecimientos.

```
For i = 0 To M
```

Luego generamos una variable (***equis***, en este caso) donde almacenamos el exponente (variable) de la función a integrar (e^{-x^2}): El exponente lo definimos así

$$x^2 = (a + (b - a)u_i)^2 = equis$$

donde el término u_i es un número aleatorio uniforme entre 0 y 1. Con esta expresión convertimos el generador de números aleatorios entre 0 y 1 en un generador de números aleatorios entre ***a*** y ***b***.

Ahora, a esta expresión, la escribiremos con la sintaxis de VBA, de esta manera

```
equis = (a + (b - a) * Rnd()) ^ 2
```

Por último solo tenemos que calcular la función completa (e^{-x^2}) y acumular su valor en cada ciclo usando una variable que llamamos, en ese ejemplo, ***acumula***

```
acumula = acumula + Exp(-equis)
    Next i
```

Para presentar luego en la celda F6 (columna 6, renglón 6) los resultados calculando la esperanza matemática:

```
Cells(6, 6) = ((b - a) / M) * acumula
```

Abajo vemos el código completo en **Visual Básic** aplicado en una hoja de Excel, presenta en la celda F6 la integral entre 1 y 3 luego de generar 1000 puntos.

```
Dim M As Integer, a As Double, b As Double
Randomize
 M = 1000
 a = 1
 b = 3

    For i = 0 To M
         equis = (a + (b - a) * Rnd()) ^ 2
         acumula = acumula + Exp(-equis)
    Next i
    Cells(6, 6) = ((b - a) / M) * acumula
```

En Python:

```
import numpy as np
def G(x):
    return np.exp(-x**2)
def monte(G,a,b,M):
    acu=0
    for i in range(M):
        acu+=G(a+(b-a)*np.random.uniform(0,1,1))
        return ((b-a)/M)*acu(0)
```

Obviamente M puede asumir cualquier valor e, incluso, incorporar un prompt para que el operador pueda ajustar la cantidad de ciclos en cada corrida, además de otras mejoras que podrían introducirse.

Reducción de varianza

Las técnicas de reducción de varianza se utilizan para aumentar la precisión de las estimaciones muestrales a fin de hacer más preciso el cálculo de los parámetros poblacionales. Por ejemplo, una variable tiene como función de densidad de probabilidad $f(x) = e^{-x}$, de donde su función de probabilidad acumulada ***F(x)*** es:

$$F(x) = \int_0^x e^{-t} dt = 1 - e^{-x}$$

Su media es 1, pero supongamos que no la conocemos y queremos estimarla.

La técnica de Montecarlo, como vimos, se usa en simulación directa, generando observaciones aleatorias a partir de la distribución usada y luego el promedio de esta muestra para estimar la media.

Al aplicar Montecarlo obtenemos:

$$U = 1 - e^{-x}$$
$$X = -Ln\ (\ 1 - U\ ) = -Ln\ (1 - U_i)$$

Con esto, encontramos un valor alejado de la media poblacional, pero al aplicar la técnica de **números aleatorios complementarios**, generamos un número aleatorio, y calculamos su número complementario, en otra secuencia. Si hacemos la sumatoria de ambas secuencias y sacamos la media de ambas sumas, obtendremos

i	Número aleatorio (μ_i)	Observación Obtenida $x_i = -Ln(1-\mu_i)$	Número complementario del aleatorio (μ_{ci})	Observación aleatoria complementaria $x_i = -Ln(1-\mu'_i)$
1	0.495	0.684	0.505	0.702
2	0.335	0.408	0.665	1.092
3	0.791	1.568	0.209	0.234
4	0.469	0.633	0.531	0.756
5	0.279	0.328	0.721	1.275
6	0.698	1.199	0.302	0.359
7	0.013	0.014	0.987	4.305
8	0.761	1.433	0.239	0.272
9	0.290	0.343	0.710	1.236
10	0.693	1.183	0.307	0.366
		$\sum = 7.793$ x = 0.7793		$\sum = 10.597$ x = 1.060

$$x = \frac{1}{2}(0{,}7793 + 1{,}060) = 0{,}920$$

*Dado un número aleatorio "u" de una función de probabilidad dada, se llama **número aleatorio complementario** de "u" a un número "u_c" tal que el promedio de los 2 sea el promedio de distribución de la variable.*

Por ejemplo, si obtenemos el número aleatorio uniformemente distribuido entre [0,1), u_c= 0,78 entonces el número aleatorio complementario de u es $u_{c'}$=0,22 porque el promedio de ellos es 0,5 que es el promedio de los números de 0 a 1, uniformemente distribuidos.

Si es una distribución normal, se obtendría u = 0,78 (número aleatorio normalmente distribuido con promedio 0 y desviación σ), entonces el número aleatorio complementario de u es $u_c = -$ 0,78 porque el promedio de ellos es 0.

Si u = 3,78 obtenido de una función de distribución normalmente distribuida con promedio 5 y desviación σ, entonces la altura aleatoria complementaria es u_c = 6,22 porque el promedio de u y u_c es 5.

Una tercera alternativa es usar un ***muestreo estratificado***. Para ello, dividimos la distribución de probabilidades acumulada *F(x)* en partes, y cada una de ellas se muestra por separado, con lo que obtenemos un muestreo desproporcionado, pero más denso en los estratos críticos. Se debe determinar tamaño de la muestra en cada estrato, deducir el número aleatorio del intervalo del estrato y determinar una ponderación de la muestra.

Ejemplo 1
Transformaremos una variable continua en discreta. Una variable poblacional (supongamos la altura de un individuo en determinado grupo), se distribuye normalmente con μ = 180 y σ = 5. Sabemos que la probabilidad de altura, con esos parámetros:

Intervalo de altura de 1 individuo	Probabilidad
[-∞ , 170)	0,0228
[170 , 175)	0,1359
[175 , 180)	0,3413
[180 , 185)	0,3413
[185 , +∞]	0,1587

Calculamos la probabilidad acumulada y asignamos un intervalo.

Intervalo de altura de 1 individuo	Probabilidad	Probabilidad Acumulada	Intervalo asignado
[-∞ , 170)	0,0228	0,0228	[0 , 0,0228)
[170 , 175)	0,1359	0,1587	[0,0228 ; 0,1587)
[175 , 180)	0,3413	0,5000	[0,1587 ; 0,5000)
[180 , 185)	0,3413	0,8413	[0,5000 ; 0,8413)
[185 , +∞]	0,1587	1,0000	[0,8413 ; 1]

Generamos 10 números uniformes entre 0 y 1 y obtenemos, por ejemplo:

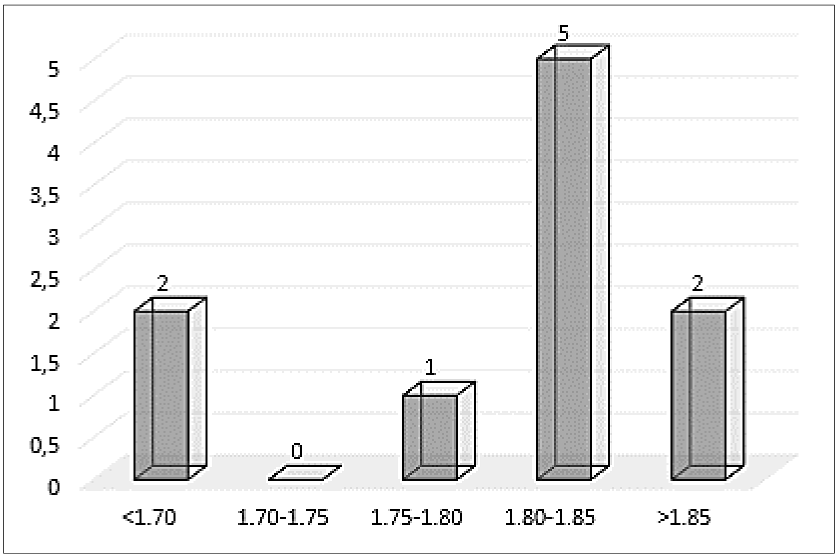
U(0,1)=	0,871	0,57	0,75	0,47	0,019	0,67	0,016	0,64	0,81	0,923
Individuo	1°	2°	3°	4°	5°	6°	7°	8°	9°	10°

Consultando la tabla de arriba comprobamos que:

1er. individuo mide sobre 1,85 m
2do individuo mide entre 1,80 y 1,85 m
3er individuo mide entre 1,80 y 1,85 m
4to individuo mide entre 1,75 y 1,80 m
5to individuo mide bajo 1,70 m
6to individuo mide entre 1,80 y 1,85 m
7mo individuo mide bajo 1,70 m
8vo individuo mide entre 1,80 y 1,85 m
9no individuo mide entre 1,80 y 1,85 m
10mo individuo mide sobre 1,85 m

Completamos ahora la tabla con los valores encontrados:

Intervalo de altura de 1 individuo	Probabilidad	Probabilidad Acumulada	Intervalo asignado	Resultado
[-∞ , 170)	0,0228	0,0228	[0 , 0,0228)	2 individuos
[170 , 175)	0,1359	0,1587	[0,0228 ; 0,1587)	0 individuos
[175 , 180)	0,3413	0,5000	[0,1587 ; 0,5000)	1 individuo
[180 , 185)	0,3413	0,8413	[0,5000 ; 0,8413)	5 individuos
[185 , +∞]	0,1587	1,0000	[0,8413 ; 1]	2 individuos



Ejemplo 2
Este es un caso de colas del que sabemos que las llegadas tienen una tasa promedio de 10 clientes diarios y con distribución Poisson,

llegadas:	Probabilidad Acumulada	Intervalo asociado
0	0,00004	[0 , 0,00004)
1	0,001	[0,00004 , 0,001)
2	0,003	[0,001 , 0,003)
3	0,010	[0,003 , 0,010)
4	0,029	[0,010 , 0,029)
5	0,067	[0,029 , 0,067)
6	0,130	[0,067 , 0,130)
7	0,220	[0,130 , 0,220)
8	0,333	[0,220 , 0,333)
9	0,458	[0,333 , 0,458)
10	0,583	[0,458 , 0,583)
11	0,697	[0,583 , 0,697)
12	0,792	[0,697 , 0,792)
13	0,864	[0,792 , 0,864)
14	0,917	[0,864 , 0,917)
15	0,951	[0,917 , 0,951)
16	0,973	[0,951 , 0,973)
17	0,986	[0,973 , 0,986)
18	0,993	[0,986 , 0,993)
19	0,999	[0,993 , 0,999)
20	1,000	[0,999 , 1,000)
.....		

Para simular vamos a generar las llegadas diarias durante una semana. Para eso obtenemos siete números aleatorios, por ejemplo, los siguientes:

0,945	0,485	0,552	0,208	0,709	0,081	0,748

Ahora usando la tabla de arriba con estos números, convertimos y tenemos, para los siete días, las siguientes llegadas diarias:

15	10	10	7	12	6	12

Con esto ya podemos calcular el promedio de llegadas por día: 72/7 = 10,28

Ejemplo 3

Supongamos que tenemos tres variables aleatorias continuas:

- x_1 con distribución normal de media μ = 1,75 y σ = 0.5
- x_2 con distribución uniforme entre 10 y 50
- x_3 convertida en discreta en 4 valores con las probabilidades que vemos en la siguiente tabla:

x_3	P [$x_3 = x_i$]	P | $x_3 \leq x_i$]
10	0.15	0.15
20	0.35	0.50
30	0.45	0.95
40	0.05	1.00

Esas variables se relacionan entre sí mediante la función

$$x_4 = x_1 + x_2^2 + x_1 * x_3$$

Usamos el método de Montecarlo para generar 6 observaciones para la función *f(x_4)*.

Para la variable *x_1*:
Con un generador aleatorio de distribución normal con los parámetros *μ* = 0 y σ = 1 (distribución normal estándar), generamos algunos valores. Supongamos que obtuvimos seis números:

−1,02 1,82 − 0,42 − 0,38 − 0,51 0,03

que, como son estándar, los convertimos a la distribución normal que necesitamos en este ejemplo (*μ* = 1,75 y σ = 0.5)
1,75 + 0,5 * (−1,02) = 1,24
1,75 + 0,5 * (1,82) = 2,66
1,75 + 0,5 * (−0,42) = 1,54
1,75 + 0,5 * (−0,38) = 1,56
1,75 + 0,5 * (−0,51) = 1,50
1,75 + 0,5 * (0,03) = 1,77

Para la variable *x_2*
Obtenemos 6 números aleatorios con distribución uniforme entre 0 y 1:
0,59 — 0,63 — 0,77 — 0,82 — 0,81 — 0,36

y con ellos calculamos:

$$10 + 40 * 0,59 = 33,6$$
$$10 + 40 * 0,63 = 35,2$$
$$10 + 40 * 0,77 = 40,8$$
$$10 + 40 * 0,82 = 42,8$$
$$10 + 40 * 0,81 = 12,4$$
$$10 + 40 * 0,36 = 24,4$$

Para la variable x_3

Generando los 6_números de distribución uniforme entre 0 y 1:

0,81 — 0,58 — 0,74 — 0,28 — 0,12 — 0,22

Buscamos en la tabla correspondiente a esta variable (ver más arriba) y vemos que corresponden a:

30 — 30 — 30 — 20 — 10 — 20

Cálculo de x_4

Reemplazamos la función en cada valor obtenido y ordenado de las variables x_i

$$1,24 + 33,6^{2} + 1,24 * 30 = 1167,4$$
$$2,66 + 35,2^{2} + 2,66 * 30 = 1321,5$$
$$1,54 + 40,8^{2} + 1,54 * 30 = 1712,4$$
$$1,56 + 42,8^{2} + 1,56 * 20 = 1864,6$$
$$1,50 + 42,4^{2} + 1,50 * 10 = 1814,3$$
$$1,77 + 24,4^{2} + 1,77 * 20 = 632,5$$

Luego podremos informar el valor de la $f(x_4)$ luego de haber usando algunos elementos estadísticos básicos, como por ejemplo la media y varianza u otros parámetros para que nuestro informe sea completo y tenga validez.

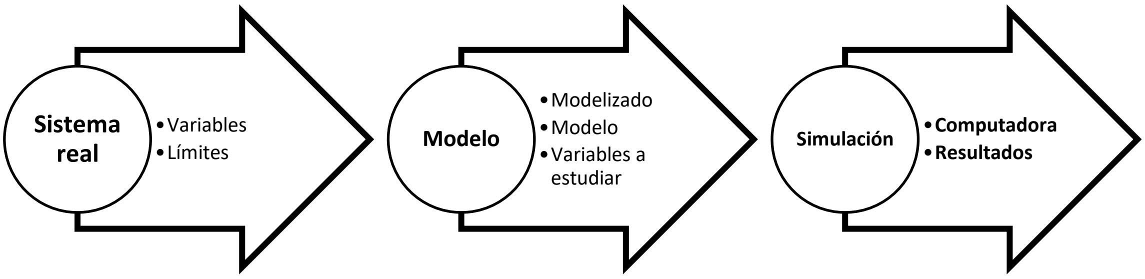
CAPÍTULO 6

CONSTRUCCION DE SIMULADORES

Los conceptos “*Sistema*”, “*Modelo*” y “*Simulación*” están íntimamente vinculados desde la perspectiva de las técnicas de la simulación. A su vez, diremos que lo que ahora conocemos como *simulación* es un conjunto de procedimientos, métodos y técnicas que se originan en los trabajos de Student[68] para determinar la distribución de la variable ***t*** y reafirmada por Von Neumann con la aplicación del *Método de Montecarlo* al estudio de distribución de neutrones en el desarrollo de las bombas atómicas que se emplearon en la Segunda Guerra Mundial.

La relación entre conceptos que mencionamos arriba se basa en asumir que para comprender un *sistema real* es útil construir un *sistema conceptual* llamado *modelo*, a partir del cual se pueden realizar experiencias acotadas o aisladas, mediante la *simulación* y con una computadora, por ejemplo, como herramienta a fin de entender el comportamiento en condiciones específicas.

En términos reales, la simulación no requiere de una computadora, aunque en la práctica, sobre todo después de la década de 1960, están íntimamente relacionados



Como ya vimos, los sistemas pueden ser estudiados por experimentación directa, construyendo un prototipo, a escala o natural, o mediante el diseño de un sistema de relaciones lógicas o matemáticas que lo describa (modelo).

De todas maneras, es común que se use el término “solución”: habrá soluciones: no

Es importante saber que los modelos y la simulación no brindarán “soluciones” en el sentido estricto del término, sino que darán elementos válidos para la toma de decisiones o para el conocimiento mejor del sistema objeto.

[68] Lo curioso es que Student no existe. La famosa distribución “
en 1908 por William Gosset, empleado en la cervecería Guinne
empleados publicar trabajos científicos para cuidar sus secreto
sobre la distribución con el pseudónimo “Student”

factibles (ver Programación Lineal), factibles, óptimas y satisfactorias, que son aquellas que, sin ser óptimas, están razonablemente cerca de ella.

Modelos en Simulación

Es importante tener en cuenta que los modelos y la simulación no brindarán "soluciones" en el sentido estricto del término, sino que darán elementos válidos para la toma de decisiones o para el conocimiento mejor del sistema objeto.

Los modelos matemáticos de simulación (de acá en adelante los denominaremos "*Simuladores*", usando una expresión general, abarcativa) suelen ser tipificados por las condiciones de prueba, algunas de esas condiciones suelen ser denominadas como "*Qué pasaría si..*" "*y si…*" *(what if…)*

Estos modelos se caracterizan por poseer parámetros de entrada y de salida, y lo correcto es que se el simulador se "ponga en marcha" o se "desarrolle", aunque en la jerga propia de la simulación se usa la expresión "correr el simulador". Tenemos que tener presente que no podemos usar el término "resolver la simulación". Cuando "corremos" el simulador obtendremos, por lo tanto, una respuesta y no una solución óptima.

No hay posibilidad de obtener respuestas satisfactorias de un simulador si no se disponen de herramientas estadísticas adecuadas

Lo más importantes es que permiten analizar el comportamiento del sistema bajo las condiciones especificadas por el usuario.

No existe un modelo único para un sistema debido a que cada modelo estará determinado por los objetivos que se planteen cuando se desarrolle o construya. Unos tendrán mayor validez que otros. **No es posible definir "*el*" modelo de un sistema**.

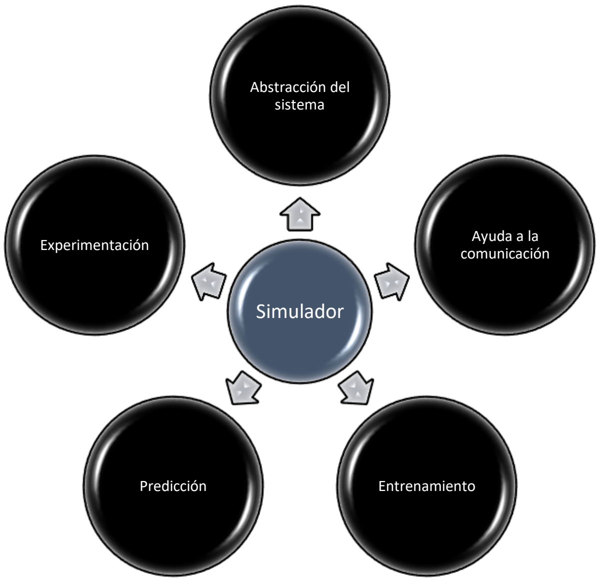
Si bien un modelo puede estar construido para que funcione con la ayuda de una computadora, o presentarse y correrse de algún otro modo, es conveniente trabajar sobre los primeros para obtener herramientas sólidas que ayuden a representar la realidad y a simular el paso del tiempo y el azar.

Usos comunes de los modelos de simulación

Nos parece útil hacer una pequeña referencia al tipo de uso que podemos encontrar en los modelos de simulación, es decir, ¿Por qué puede ser importante invertir esfuerzo en adquirir habilidades para construir o desarrollar modelos e interpretar sus resultados?

Los "simuladores" suelen ser útiles, entre otras cosas, como:

1. Una ayuda a la abstracción del sistema.
2. Una ayuda a la comunicación.
3. Entrenamiento e instrucción.
4. Una herramienta de predicción.
5. Una ayuda a la experimentación.



1. **Como elemento de ayuda para la comprensión abstracta del sistema**. El modelo construido, o a construir, fuerza a organizar, evaluar y examinar la validez de los conceptos o concepciones que el observador tiene de ese sistema y de sus subsistemas.
2. **Como ayuda a la comunicación**. El modelo sirve para expresarnos de la mejor manera y transmitir conceptos. El arquitecto concibe una obra, construye el modelo "Plano de la obra" y es la mejor manera de transmitirlo al constructor.
3. **Como herramienta educativa** (entrenar o enseñar). Las personas necesitan practicar. El modelo le muestra eventualidades antes que le ocurran en la práctica. Modelos de naves espaciales, de conducción de vehículos, de juegos y de negocios, de crecimiento, etc.
4. **Como herramienta de predicción**. Quizá sea uno de los más importantes usos prácticos. Históricamente es y ha sido empleado para predecir el comportamiento característico de la entidad modelada. Prueba de diseño para aviones, barcos, autos, misiles, tiempo de vida útil en medicamentos y alimentos, reacciones sociales, impacto de nuevos productos, corrosión, cargas de estructuras, etc.
5. **Como ayuda a la experimentación**. Los modelos permiten experimentos controlados en sistemas, lo que de otra manera podría ser impracticable, peligroso o de costos prohibitivos.

Propósitos de los modelos de simulación

Un modelo sirve a uno de los dos principales propósitos siguientes:
Descriptivo: para explicar y/o entender el sistema.
Prescriptivo: predecir y/o replicar el comportamiento característico del sistema.

Etapas de la simulación

Como ya hemos visto cuando discutimos la construcción de modelos, el proceso de simulación tiene etapas definidas y similares, que son las siguientes:

1. Identificación del problema.

Es la etapa más importante, ya que puede resolverse un problema equivocado, y obtener un modelo que no represente el sistema a estudiar[69]. Para encontrar una

[69] "La formulación apropiada del problema fue más esencial que su solución" (Albert Einstein).

solución aceptable o solución óptima, primero debemos saber exactamente cuál es el problema.

Tendremos como meta clarificar los objetivos que estableceremos para el simulador. No puede haber un modelo sin objetivos. Tendremos claro lo que pretendemos que el modelo responda, los alcances del estudio, que aspectos se van a analizar.

Como sabemos que construir un simulador implica costos y beneficios, buscaremos compensar los costos logrando, en lo posible, el máximo beneficio derivado del empleo del simulador.

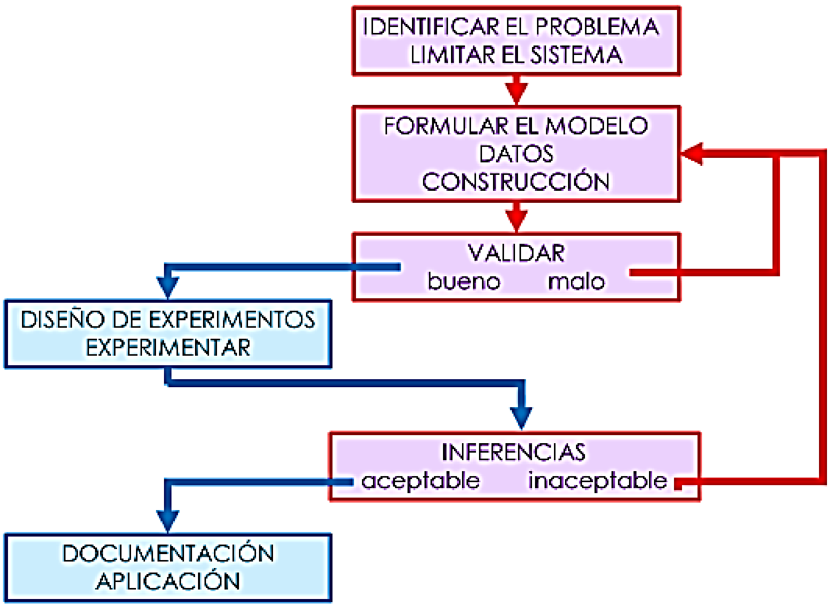
Tendremos que determinar el nivel de agregación y de detalle en la información que entregará el simulador. Tiene que ver con el nivel de decisión que pretende apoyar el modelo. Un nivel de agregación alto en la información de salida del simulador, indica que el simulador apoyará decisiones de nivel superior, estratégicas o tácticas. Un nivel de agregación bajo es indicativo de que el simulador apoyará decisiones de nivel operativo.

Deberemos tener en cuenta las restricciones de tiempo y de presupuesto para el diseño y construcción del modelo de simulación. También para el diseño de experimentos y las corridas del simulador en su etapa de explotación.

2. Establecer los límites del sistema

Es necesario que sepamos definir el sistema y su entorno. En el sistema estarán los entes y el objeto de la simulación. En el entorno estarán las cajas negras. En el resto del universo estarán los sistemas a los que no se les reconoce influencia sobre el sistema en estudio; es por esto por lo que debemos precisar los límites.

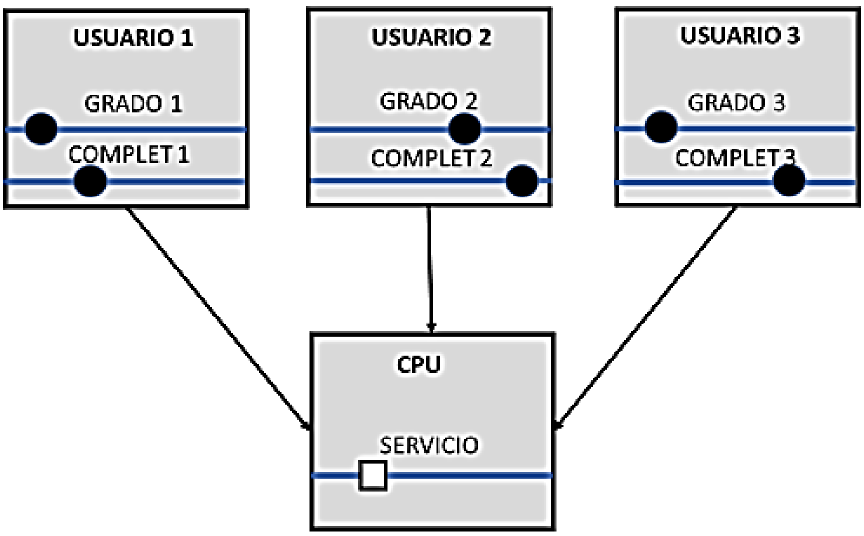
En esta etapa es cuando debemos dividir el sistema que estudiamos en elementos componentes o subsistemas, cuyo comportamiento sea importante predecir. Debemos describir las actividades o procesos que caracterizan la evolución del sistema. Estas actividades podrán ser endógenas y exógenas.



- **Identificación del problema**. Implica que debemos tener definidos los objetivos que pretendemos alcanzar, determinado el nivel de decisión a apoyar y tener en cuenta las restricciones de tiempo.
- **Definición del Sistema**. En esta etapa es necesario determinar los límites, restricciones y medidas o indicadores de la efectividad (logro del objetivo) a ser usadas en la definición del sistema a ser estudiado.
- **Formulación del modelo**. Se produce la reducción o abstracción del sistema real a un diagrama de flujo lógico, con su adecuada documentación.
- **Preparación de datos**. Es la identificación de los datos que necesita el modelo que se formuló, y su reducción a una forma apropiada para el ingreso al modelo.
- **Construcción del modelo**. Tendremos que construir el modelo en un lenguaje apto para el computador u otro elemento físico que será usado. Es hacer el programa computacional y probarlo.
- **Validación**. Consiste en llevar a un nivel aceptable la confiabilidad de una inferencia efectuada desde el modelo sobre el sistema real. Debemos cuidad de que esta inferencia sea correcta. Se requieren herramientas estadísticas adecuadas a los objetivos del modelo. (adecuadas no significa, necesariamente, complejas sino de un nivel acorde a lo que buscamos o necesitamos)
- **Diseño de experimentos**. Deberemos ser capaces de planificar un experimento cuyos pasos nos brinden la información deseada. Se trata de diseñar cómo se realizará, y cómo se correrán los experimentos planificados y los indicadores buscados. Tiene una etapa de planificación estratégica y táctica.
- **Ejecución de experimentos** (Experimentación) Esta etapa consiste en ejecutar la simulación de acuerdo con lo que hemos diseñado para generar los datos deseados y realizar un análisis de sensibilidad.
- **Interpretación**. En esta etapa haremos inferencias sobre el sistema que estamos estudiando, a partir de los datos generados por la simulación. Usaremos herramientas tales como estadística y matemática para formalizarlas.
- **Documentación**. Volcaremos todo lo concerniente al modelo y a su uso, especificaciones y restricciones en un manual, también el proyecto de actividades futuras que planifiquemos.

Representación gráfica de los modelos de simulación

Hay varias formas de representar mediante diagramas los modelos, aunque la más utilizada son los diagramas de influencia. Por ejemplo, un diagrama de influencia para un sistema que comparte un procesador es el siguiente:



Este diagrama representa cuatro componentes: Un procesador central (**CPU**) y tres terminales (**USUARIO1**, **USUARIO2** y **USUARIO3**) cada componente tiene sus parámetros y variables (Discretas representadas por □ y continuas representadas por ●).

El bloque **CPU** contiene la variable *SERVICIO* cuyos valores son 1, 2 o 3, mientras que los componentes llamados **USUARIOx** tienen las variables *COMPLETx* que puede valer 0 a 1 según se completa la fracción de su tarea. Además, un parámetro *GRADOx* que indica la tasa de cumplimiento de **USUARIOx**.

Este tipo de diagrama evolucionó a los diagramas de Forrester, ya mencionados, y a otros como los de bucles causales.

Simulación de sucesos discretos

Los sistemas continuos generalmente se representan en forma de ecuaciones diferenciales. Los modelos discretos son aquellos que cambian en instantes determinados. Un caso que puede tomarse como modelo general es el sistema de colas donde se busca estimar medidas representativas de largo de la cola o tiempo medio de espera. Como podemos ver, en un sistema de este tipo no hay "soluciones" concretas. No podremos afirmar cosas como "a los 16 minutos vamos a tener una cola de cinco clientes" ni nada parecido. Solamente podremos encontrar estimadores generales del número de clientes o de probabilidad de arribos o probabilidad de que el servidor esté desocupado.

Estos casos se caracterizan por que hay momentos en que se producen cambios en el sistema: ***sucesos*** en el sistema. Como estos sucesos se dan en instantes determinados tendremos que simularlos, entonces, como discretos.

En la simulación de sucesos discretos, **SSD**, encontraremos diferentes tipos de variables:

1. **Variables de tiempo**, la más importante es la que se refiere al tiempo ***T*** de simulación, aunque hay otros tiempos, como el de sucesos, de jornada laborable, llegadas, etc.
2. **Variables de contabilidad** del número de veces que se producen los sucesos, o intervalos o cualquier parámetro que sea importante registrar y contabilizar.
3. **Variables de estado** que describen el sistema en cada uno de los instantes, con lo cual se conoce el estado en cada instante ***t*** y permiten predecir el futuro.

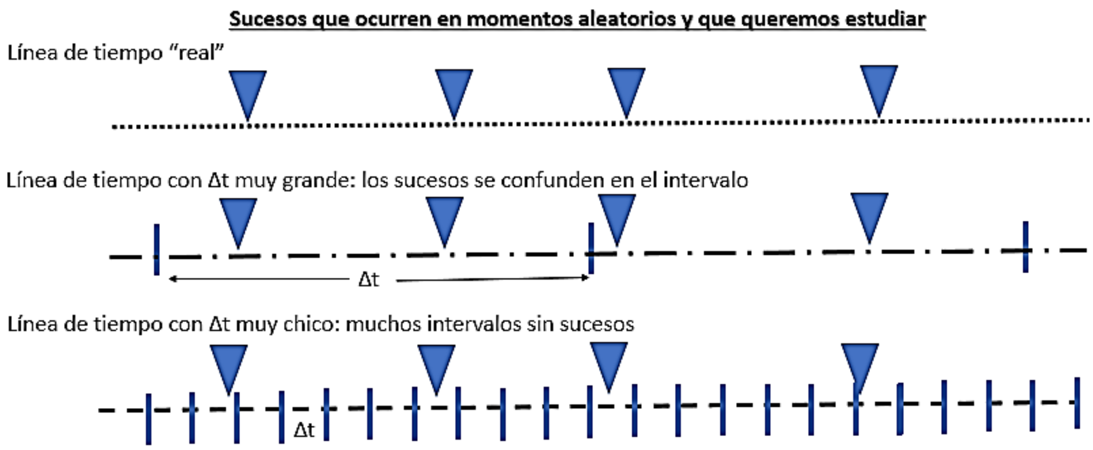
Reloj y estrategias de simulación

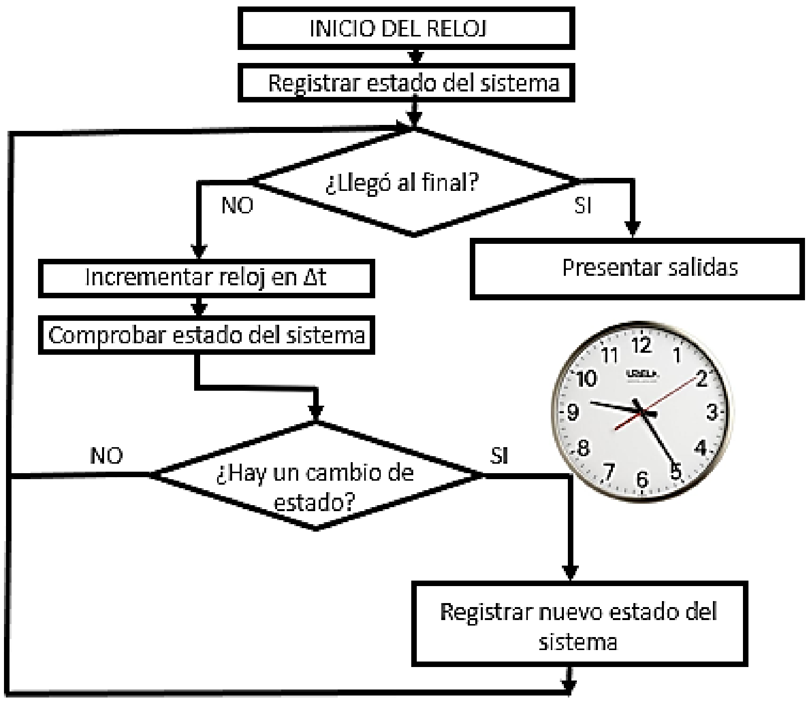
Llamaremos **mecanismo de reloj**, o **reloj de simulación**, al objeto que define la evolución del simulador, instante a instante.

Según el mecanismo de reloj que empleemos dispondremos de dos *estrategias* de simulación, que son "Simulación sincrónica" y "Simulación asincrónica":

- **Simulación sincrónica**, también llamada ***orientada a intervalos*** que es aquella en que el reloj avanza en intervalos prefijados Δt hasta cumplir alguna condición prestablecida. Lo importante en este tipo de simulación es la habilidad que tengamos en el momento de elegir Δt, ya que si ese intervalo es muy grande (en términos relativos) pueden saltearse algunos sucesos que se inician y terminan dentro del período de tiempo que hemos elegido, y por lo tanto, no los percibimos (por ejemplo, un Δt muy grande englobaría dos sucesos que ocurren en distintos instantes y los "veríamos" como que ocurrieron en un único instante y ese instante sería computado erróneamente como el final del intervalo Δt) Por otro lado, si elegimos un Δt muy chico, tendríamos muchos intervalos de tiempo sin que ocurran sucesos dando lugar a ineficiente uso del simulador.

•





RELOJ SINCRÓNICO

- **Simulación asincrónica**: es la que mayormente se utiliza, está orientada a sucesos: el reloj avanza hasta el momento en que ocurre un nuevo suceso y es entonces cuando se actualiza el estado del sistema, se registra el tiempo y se actualiza la contabilidad. El procedimiento continúa hasta cumplir alguna condición preestablecida. Requiere llevar un registro de sucesos y los instantes en que se producen

RELOJ ASINCRÓNICO

Estudio de caso. Construcción de un simulador M/M/1 en Excel[70]

Trabajaremos sobre un modelo de colas M/M/1. Vamos a construir un simulador en hoja de cálculo. Es un sistema elemental que nos permite comprender una posibilidad de diseño del simulador y entender su funcionamiento. Este ejemplo es abierto, en el sentido que podemos agregarle elementos que nos permitan obtener más datos.

Queremos evaluar

- El tiempo que pasan los clientes en el sistema y en la cola
- El tiempo que transcurre entre ***T*** y el instante en que se termina de atender al último cliente.

Para lograr lo que nos proponemos vamos a desarrollar el modelo de simulación, siguiendo los siguientes pasos:

Determinaremos de qué manera avanzará nuestro simulador. Vamos a elegir el siguiente método, basado en un reloj asincrónico:

a) se produce una primera llegada. Nuestro reloj "del sistema" estará en cero, pues aún no habíamos comenzado a simular. Al producirse la primera llegada hacemos avanzar el reloj hasta la hora de esa llegada. ¿Por qué hacemos esto? Porque nuestro sistema es un sistema de llegadas y salidas y si no hay una primera llegada nunca habrá una salida y el sistema permanecerá inmutable.

[70] En el capítulo 7 más adelante se vuelve a este simulador en términos comparativos y también se encuentra una descripción teórica del modelo M/M/1 con cierto grado de detalle (Capítulo 7 Simulación de Inventarios y Colas / Modelo de colas M/M/1)

Como la llegada es un suceso (el primero en realidad) el reloj avanzará a la hora de ese suceso, como lo predice la definición de reloj asincrónico que vimos arriba.

b) Sucesivamente se producirán llegadas y salidas hasta el fin de la simulación.
c) El fin de la simulación puede ser por tiempo, por ejemplo, simular 8 horas de trabajo o por número máximo de clientes que llegan, por ejemplo, detener la simulación cuando llegaron 30 clientes. En esta oportunidad elegiremos la segunda alternativa, simularemos un número determinado de clientes. Sin embargo es muy simple transformar el código para que ambas opciones estén disponibles.

Para desarrollar este modelo de simulación vamos a establecer condiciones de simulación y de contorno.

1) Nuestro modelo correrá en forma asincrónica, es decir, cuando se produce un suceso, se registra la hora de ocurrencia. Los eventos pueden ser de dos tipos. Uno es la *llegada* de un cliente y el otro es la *salida* de un cliente porque terminó su atención.
2) ¿Cómo simulamos el suceso "*llegada*"? Si conocemos la tasa media de arribos de clientes, podemos deducir, a partir de la probabilidad, el momento en que ocurre. Veamos:

   Si llamamos λ a la tasa media de arribos de clientes/unidad de tiempo

   [λ] = cl/t

   y suponemos que esta sigue una distribución exponencial, tendremos que la probabilidad de que un cliente llegue en un instante *t* será:

   $$P() = 1 - e^{-\lambda t}$$

   así, si por ejemplo, λ=20 cl/h y queremos saber la probabilidad de que un cliente llegue en los próximos 10 minutos, será, cuidando que las unidades sean todas en horas (10 min será expresado como 1/6 de hora):

   $$P() = 1 - e^{-20\left(\frac{1}{6}\right)} = 1 - e^{-\left(\frac{20}{6}\right)} = 0{,}964$$

   Pero, inversamente, si nuestro dato es un valor que representa la probabilidad del instante, por ejemplo 0,964, podremos saber a qué instante corresponde:

   $$t = -\frac{\ln(1 - P())}{\lambda} = -\frac{\ln(1 - 0{,}964)}{20} = 0{,}1667\ h \equiv 10\ min$$

   Esto significa que para simular una llegada, solo debemos generar un número aleatorio de distribución uniforme entre cero y uno, al cual, arbitrariamente, le asignamos el valor de la probabilidad para saber en qué instante se produjo una llegada, usando λ. En la expresión anterior, el número aleatorio generado sería 0,964 y el instante obtenido 0,1667.

3) Una vez que obtenemos el tiempo de llegada simplemente acumulamos para cada cliente el tiempo obtenido con el tiempo del cliente anterior, que a su vez fue acumulado. Así tenemos una función de acumulación que nos permite tener horas de llegadas consecutivas. Nuestra tarea será disponer de las **horas** a las que "van a llegar" cada uno de los *n* clientes que vamos a procesar. Observamos que calculamos "tiempo" de llegada y lo convertimos a "hora" de

llegada. El primero es un intervalo, el segundo es lo que marca el reloj cuando ese intervalo se cumplió. Por ejemplo: el cliente 1 llegó justo a las 15:00:00, el cliente 2 llegó con un tiempo (de demora) de 1:10 (un minuto diez segundos después). La hora de llegada del cliente 2, por lo tanto, será 15:00:00 + 00:01:10, o sea 15:01:10.

4) Lo mismo aplicaremos para los tiempos de atención de cada uno de los clientes, usando, en este caso, una tasa de atención o salidas **μ**. En este caso no acumularemos los tiempos y estos, por ahora no serán más que lo que vamos a tardar en atender a cada uno de los clientes. En nuestra jerga, a esto le llamamos “tiempo de atención”, mientras que en el punto anterior eran “hora de llegada”. La diferencia está en las palabras “tiempo” (demora) y “hora” (reloj). Como en el caso anterior, para el cliente 2, si obtenemos un tiempo de atención de 2:30, los datos del cliente 2 serán: Hora de llegada 15:01:10, tiempo de atención 02:30

5) A pesar de ello necesitaremos un reloj que acumule la **hora** de la salida del cliente. Esta hora se calcula sumando la hora de llegada (punto 3 de arriba), más el tiempo de atención (punto 4) más un nuevo tiempo que debemos calcular que es el tiempo en que estuvo en la cola esperando a que el servidor se desocupara. En definitiva, cada uno de los clientes va a tener tres parámetros de tiempo registrados, que son:

   a. la hora reloj de llegada al sistema,
   b. el tiempo que estuvo esperando (lapso),
   c. el tiempo en que fue atendido (lapso) y
   d. la hora reloj en que se fue del sistema

   Veamos como se aplica esto al hipotético cliente que llegó a las 15:00:00

   a. la hora de llegada al sistema es 15:01:10
   b. el tiempo que estuvo esperando (lapso) es 0:00
   c. el tiempo en que fue atendido (lapso) es 2:30
   d. la hora reloj en que se fue del sistema es
      15:01:10 + 0:00 + 2:30 = 15:03:40

6) El único tiempo que nos queda para computar es el tiempo en que, eventualmente, el servidor estuvo desocupado y esperando la llegada de un cliente. Este tiempo no puede ser registrado individualmente, sino que será un registro global de “servidor desocupado”, con lapsos acumulados.

7) Usaremos un reloj del sistema para ir computando el avance del tiempo a medida que se producen los sucesos llegada o salida

8) Definimos entonces las variables del modelo que usaremos:

cl	contador de clientes
RSIST	Reloj del sistema. En esta variable llevamos registro de hora del último evento
HLL(cl)	Matriz de registro de la **hora** de llegada de cada uno de los clientes
TESP(cl)	registro del **tiempo** de espera a ser atendido de cada uno de los clientes
TATE(cl)	registro del **tiempo** de atención de cada uno de los clientes

HOSA(cl)	registro de la **hora** de salida de cada uno de los cliente
Numax	número máximo de clientes a simular (máximo 1000)
mu	tasa de salidas media exponencial de clientes
lambda	tasa de llegadas media exponencial de clientes
LQ	Largo instantáneo de la cola
proximocliente	Número del primer cliente que está en la cola

Los pasos que seguiremos serán:

1) Preparamos la hoja de cálculo con los siguientes elementos:

Colas M/M/1

DATOS DE ENTRADA	
SERVIDORES:	1
UNIDAD DE TIEMPO	Hora
TASA DE SERVICIO POR Hora (μ)	66
TASA DE LLEGADAS POR Hora (λ)	60
Número de clientes a procesar	896
(máximo 1000)	

Optimiza Versión 2.0.0
Simulador de un sistema M/M/1
Capacidad Máxima: 1000 clientes

Julio 2021

SIMULACION		
# Clientes	896	
Tiempo total	10:30:19	
Lq medio	12,2734	
L medio	13,2734	
W	0,1289	
Wq	0,1139	
Serv Libre [%]	13,1147	%
Serv Ocupado (τ)* - Pw	86,8853	%
Po	13,1147	%
Lq máxima	36	
*Equivale a factor de tráfico		
NUMERICO		
Serv Ocupado (τ)*	0,9091	90,91%
Po	0,0909	9,09%
Lq	9,0909	
L	10,0000	
Wq	0,1515	
W	0,1667	
Pw	0,9091	90,91%

Cliente#	H Llegada	T espera	T atención	H Salida	T Serv Des	Lq	L
1	0,00234	0,00000	0,01542	0,01776	0,00234	0	[illegible]
2	0,03630	0,00000	0,00661	0,04291	0,01854	0	[illegible]
3	0,03727	0,00563	0,00822	0,05113	0,01854	2	[illegible]
4	0,03912	0,01201	0,00279	0,05392	0,01854	2	[illegible]
5	0,04508	0,00885	0,03151	0,08543	0,01854	1	[illegible]
6	0,05433	0,03110	0,00739	0,09282	0,01854	5	[illegible]
7	0,06164	0,03118	0,02768	0,12050	0,01854	4	[illegible]
8	0,07013	0,05037	0,00194	0,12244	0,01854	5	[illegible]
9	0,07986	0,04258	0,04902	0,17145	0,01854	4	[illegible]
10	0,08234	0,08911	0,02787	0,19932	0,01854	9	[illegible]
11	0,10312	0,09620	0,00155	0,20087	0,01854	12	[illegible]
12	0,10552	0,09535	0,03363	0,23450	0,01854	11	[illegible]
13	0,12573	0,10876	0,04679	0,28129	0,01854	12	[illegible]
14	0,13239	0,14889	0,00438	0,28567	0,01854	12	[illegible]
15	0,13298	0,15269	0,05796	0,34362	0,01854	11	[illegible]
16	0,14300	0,20063	0,01220	0,35583	0,01854	13	[illegible]
17	0,15070	0,20513	0,01031	0,36614	0,01854	15	[illegible]
18	0,16546	0,20068	0,00653	0,37267	0,01854	14	[illegible]
19	0,18139	0,19128	0,00154	0,37420	0,01854	13	[illegible]

A pesar de los numerosos elementos que se ven en la imagen anterior (que corresponde a una captura con los resultados de una corrida) lo que nos interesa es que destinamos las celdas **D5**, **D6** y **D7** a ingresar los únicos tres datos que necesitamos, que son: valor de la tasa de salida, de la tasa de llegadas y cantidad de clientes a simular, respectivamente. Todo lo demás son simplemente etiquetas para poder interpretar los valores que escribiremos en sus adyacencias.

DATOS DE ENTRADA	
SERVIDORES:	1
UNIDAD DE TIEMPO	Hora
TASA DE SERVICIO POR Hora (μ)	66
TASA DE LLEGADAS POR Hora (λ)	60
Número de clientes a procesar	896
(máximo 1000)	

*Equivale a factor de tráfico		
NUMERICO		
Serv Ocupado (τ)*	0,9091	90,91%
Po	0,0909	9,09%
Lq	9,0909	
L	10,0000	
Wq	0,1515	
W	0,1667	
Pw	0,9091	90,91%

El recuadro "SIMULACIÓN" estará destinado a presentar los datos obtenidos por simulación, mientras que "NUMERICO" presenta, para los parámetros de entrada el cálculo con el método M/M/1 tradicional, empleando las siguientes funciones:

Servidor ocupado	$\tau = \lambda/\mu$
Prob. serv. libre	$P_0 = 1 - \lambda/\mu$
Largo cola	$Lq = \tau^2/P_0$
Largo sistema	$L = Lq + c$
Espera en cola	$Wq = Lq/\lambda$
Espera en sistema	$W = Wq + (1/\mu)$
Prob. espera	$Pw = \tau$

2) Guardamos el libro habilitándolo para macros

3) Abrimos el editor de *VisualBasic* y comenzamos a escribir el código de la macro que describimos a continuación paso a paso.

a) vamos a diseñar esta macro para operar con un máximo de 1000 clientes, ya que en esta instancia hemos definido que nuestro parámetro de control será el número de clientes, aunque como ya señalamos, podría ser el tiempo de simulación.
b) Primero vamos a dimensionar las matrices y las demás variables. Observemos que cuando dimensionamos los arreglos o matrices fijamos, indirectamente, el máximo de clientes a procesar:

```
Sub Macro1()
' Macro1 Macro
' Acceso directo: CTRL+w
'
Dim cl As Integer
Dim RSIST As Double
Dim HLL(1 To 1000) As Double
Dim TESP(1 To 1000) As Double
Dim TATE(1 To 1000) As Double
Dim HOSA(1 To 1000) As Double
Dim Numax As Integer
Dim mu As Double
Dim lambda As Double
Dim LQ as Double
Dim proximocliente as Integer
Range("J3:R1003").ClearContents
Range("H3:H10").ClearContents
```

c) luego preparamos el ingreso de los datos del problema y la verificación que en las celdas destinadas a ingresar los valores correspondientes a μ y λ, existan valores numéricos. En caso de que no sea así interrumpe la ejecución del programa.

```
'Ingreso de datos
If Cells(5, 4) = 0 Or Cells(5, 4) = "" Then
MsgBox "Ingrese un valor válido en Tasa de Servicio", , "Colas 2021"
GoTo findetodo
Else
mu = Cells(5, 4)
End If
If Cells(6, 4) = 0 Or Cells(6, 4) = "" Then
MsgBox "Ingrese un valor válido en Tasa de Llegadas", , "Colas 2021"
GoTo findetodo
Else
lambda = Cells(6, 4)
End If
```

d) si quien ejecute el programa omite ingresar el número de clientes, lo fijamos en 100. Si, por el contrario, hubiera ingresado un valor superior a la capacidad del programa, lo fijamos en 1000.

```
If Cells(7, 4) = 0 Or Cells(7, 4) = "" Then Cells(7, 4) = 100
If Cells(7, 4) > 1000 Then Cells(7, 4) = 1000
Numax = Cells(7, 4)
```

e) la segunda parte del programa se dedica a generar todos los tiempos de llegada y tiempos de atención para todos y cada uno de los clientes, (cuya cantidad, recordemos, está en Numax). Los tiempos de llegada los acumulamos para obtener un reloj con horas consecutivas de llegada, no así los tiempos de atención, que los dejamos como tales. Cada una de las horas de llegada y cada tiempo de atención se guardan para cada cliente en particular.

```
'carga de los parámetros de llegada y atención de los clientes
Randomize
For i = 1 To Numax
HLL(i) = -((Log(1 - Rnd)) / lambda)
TATE(i) = -((Log(1 - Rnd)) / mu)
Next i
'acumulamos tiempos de llegada
For i = 2 To Numax
HLL(i) = HLL(i) + HLL(i - 1)
Next i
```

f) Ahora vamos a procesar al primer cliente e imprimimos un renglón con todos sus datos

```
' Procesamiento del primer cliente
cl = 1
RSIST = HLL(cl) + TATE(cl)
HOSA(cl) = RSIST

'impresión del primer renglón, cliente 1=1
Cells(3, 10) = cl
Cells(3, 11) = HLL(cl)
Cells(3, 12) = TESP(cl)
Cells(3, 13) = TATE(cl)
Cells(3, 14) = HOSA(cl)
Cells(3, 15) = HLL(1)
Cells(3, 16) = 0
Cells(3, 17) = 1
Cells(3, 18) = Cells(3, 14) - Cells(3, 11)
cl = cl + 1
proximocliente = 2
```

g) Lo que sigue, tercera parte, es el procesamiento de todos los clientes después del primero

```
'atención de clientes Siguientes, cliente i>1
rutinadeatencion:
```

a. Si el reloj del sistema muestra una hora mayor que la llegada del próximo cliente (o sea ya llegó y está esperando), debemos aumentar la cola, verificando cuántos clientes llegaron después de él y hasta el momento.

```
If RSIST > HLL(cl) Then
    i = proximocliente
chico:
    If RSIST > HLL(i) Then
        If HLL(i) = 0 Then GoTo grande
        LQ = LQ + 1
        i = i + 1
        proximocliente = i
        If i > Numax Then GoTo grande
        GoTo chico
    End If
```

b. Si ya actualizamos la cola, atendemos al primero

```
grande:
    TESP(cl) = RSIST - HLL(cl)
```

```
    RSIST = RSIST + TATE(cl)
    HOSA(cl) = RSIST
    LQ = LQ - 1
    If LQ < 0 Then LQ = 0
```

c. Si, por el contrario, no había llegado nadie, “esperamos” la llegada y atendemos al que llegue. Esto simplemente se hace colocando en el reloj del sistema la hora de llegada del próximo cliente.

```
Else
    TEDES = TEDES + (HLL(cl) - RSIST)
    RSIST = HLL(cl) + TATE(cl)
    HOSA(cl) = RSIST
    TESP(cl) = 0
    LQ = LQ - 1
    If LQ < 0 Then LQ = 0
End If
```

h) Para cada cliente imprimimos un renglón con todos sus datos.

```
'impresion de los renglones subsiguientes
Cells(cl + 2, 10) = cl
Cells(cl + 2, 11) = HLL(cl)
Cells(cl + 2, 12) = TESP(cl)
Cells(cl + 2, 13) = TATE(cl)
Cells(cl + 2, 14) = HOSA(cl)
Cells(cl + 2, 15) = TEDES
Cells(cl + 2, 16) = LQ
Cells(cl + 2, 17) = LQ + 1 'Cálculo de L es simple por ser sistema M/M/1
Cells(cl + 2, 18) = Cells(cl + 2, 14) - Cells(cl + 2, 11)
```

i) Avanzamos al próximo cliente, verificando si no hemos atendido el cupo que establecimos, en cuyo caso iremos a la rutina final del programa.

```
cl = cl + 1
If cl <= Numax Then GoTo rutinadeatencion
```

j) La cuarta parte es la rutina final del programa presenta los cálculos resumen del proceso

```
findetodo:
Cells(3, 8) = cl - 1
Reloj = Format(RSIST, "hh:mm:ss")
Cells(4, 8) = Reloj
Cells(5, 8) = "=SUM(P3:P" & Format(Numax + 2) & ")/" & Format(Numax)
Cells(6, 8) = Cells(5, 8) + 1
Cells(7, 8) = "=SUM(R3:R" & Format(Numax + 2) & ")/" & Format(Numax)
Cells(8, 8) = "=SUM(L3:L" & Format(Numax + 2) & ")/" & Format(Numax)
Cells(9, 8) = (TEDES / RSIST) * 100
Cells(10, 8) = 100 - Cells(9, 8)
Cells(11, 8) = Cells(9, 8)
Cells(12, 8) = "=Max(P:P)"

End Sub
```

Completar el modelo de acuerdo con los requerimientos.
Como podemos ver, el modelo de simulación que presentamos no está completo, ya que le falta cumplir la segunda parte de la especificación del caso. Recordemos el enunciado…

Queremos evaluar • *El tiempo que pasan los clientes en el sistema y en la cola* • *El tiempo que transcurre entre **T** y el instante en que se termina de atender al último cliente.*

Para cumplir solo agregamos una pequeña porción de código que nos permita obtener resultados como los que mostramos en las siguientes capturas de pantalla:

Primero, hemos puesto una lista desplegable que nos permite la opción entre “Tiempo” o “Número de clientes”

DATOS DE ENTRADA		
SERVIDORES:		1
UNIDAD DE TIEMPO		Hora
TASA DE SERVICIO POR Hora (μ)		66
TASA DE LLEGADAS POR Hora (λ)		60
		1000
Tiempo máximo simulado(*)		0,13
*Ingresar en formato decimal (Ej.: 1:30 es 1,5)		
VOY A SIMULAR POR	TIEMPO	MAXIMO

DATOS DE ENTRADA		
SERVIDORES:		1
UNIDAD DE TIEMPO		Hora
TASA DE SERVICIO POR Hora (μ)		66
TASA DE LLEGADAS POR Hora (λ)		60
Número de clientes a procesar(*)		1000
		0,13
*Maximo: 1000		
VOY A SIMULAR POR	CLIENTES	MAXIMO

Luego le agregamos dos parámetros más a la lista de datos de salidas, que podrán visualizarse o no, de acuerdo con lo elegido en la lista desplegable “Tiempo/Clientes”. Las capturas siguientes están efectuadas después de sendas corridas, una en “Tiempo” y la otra en “Clientes”:

DE ENTRADA		
		1
		Hora
Hora (μ)		66
Hora (λ)		60
		1000
do(*)		0,13
j.: 1:30 es 1,5)		
	TIEMPO	MAXIMO

ón 2.0.0

e un sistema M/M/1

áxima: 1000 clientes

Optimiza

SIMULACION		
# Clientes	5	
Tiempo total	0,1660	Hora
Lq medio	0,0000	
L medio	1,0000	
W	0,0001	Hora
Wq	0,0000	Hora
Serv Libre [%]	44,7001	%
Serv Ocupado (τ)(*) - Pw	55,2999	%
Po	44,7001	%
Lq máxima	0	
L al momento de corte	0	
Tiempo adicional(**)	0,02526661	
*Equivale a factor de tráfico		
** hasta que se atiende el ultimo cliente		

DE ENTRADA				SIMULACION		
		1		# Clientes	10	
		Hora		Tiempo total	0,2165	Hora
Hora (μ)		66		Lq medio	1,5000	
Hora (λ)		60		L medio	2,5000	
rocesar(*)		10		W	0,0156	Hora
		0,13		Wq	0,0066	Hora
				Serv Libre [%]	39,5528	%
	CLIENTES	MAXIMO		Serv Ocupado (τ)(*) - Pw	60,4472	%
				Po	39,5528	%
				Lq máxima	4	
ón 2.0.0						
e un sistema M/M/1						
áxima: 1000 clientes				*Equivale a factor de tráfico		

Finalmente, el aspecto completo del simulador es el siguiente, considerando que se le agregó un botón para correrlo:

Colas M/M/1

DATOS DE ENTRADA		
SERVIDORES:		1
UNIDAD DE TIEMPO		Hora
TASA DE SERVICIO POR Hora (μ)		66
TASA DE LLEGADAS POR Hora (λ)		60
Número de clientes a procesar(*)		10
		0,13
*Maximo: 1000		
VOY A SIMULAR POR	CLIENTES	MAXIMO

COLAS Versión 2.0.0
Simulador de un sistema M/M/1
Capacidad Máxima: 1000 clientes
Optimiza
Julio 2021

Start!!

	AYUDA	
1) Entradas	2) Salidas	3) Numérico

SIMULACION		
# Clientes	10	
Tiempo total	0,2165	Hora
Lq medio	1,5000	
L medio	2,5000	
W	0,0156	Hora
Wq	0,0066	Hora
Serv Libre [%]	39,5528	%
Serv Ocupado (τ)(*) - Pw	60,4472	%
Po	39,5528	%
Lq máxima	4	

*Equivale a factor de tráfico

NUMERICO		
Serv Ocupado (τ)*	0,9091	90,91%
Po	0,0909	9,09%
Lq	9,0909	

Cliente#	H Llegada	T espera	T atención	H Salida	T Serv Des	Lq	L	W
1	0,04099	0,00000	0,00414	0,04513	0,04099	0	1	0,00414
2	0,09691	0,00000	0,00203	0,09894	0,05178	0	1	0,00203
3	0,10324	0,00000	0,01664	0,11988	0,05608	0	1	0,01664
4	0,10386	0,01602	0,00159	0,12147	0,05608	3	4	0,01761
5	0,10520	0,01627	0,01320	0,13466	0,05608	2	3	0,02946
6	0,15393	0,00000	0,01622	0,17015	0,07534	1	2	0,01622
7	0,18042	0,00000	0,01058	0,19100	0,08562	0	1	0,01058
8	0,18416	0,00684	0,00143	0,19243	0,08562	4	5	0,00828
9	0,18701	0,00543	0,01763	0,21006	0,08562	3	4	0,02306
10	0,18856	0,02150	0,00641	0,21647	0,08562	2	3	0,02791

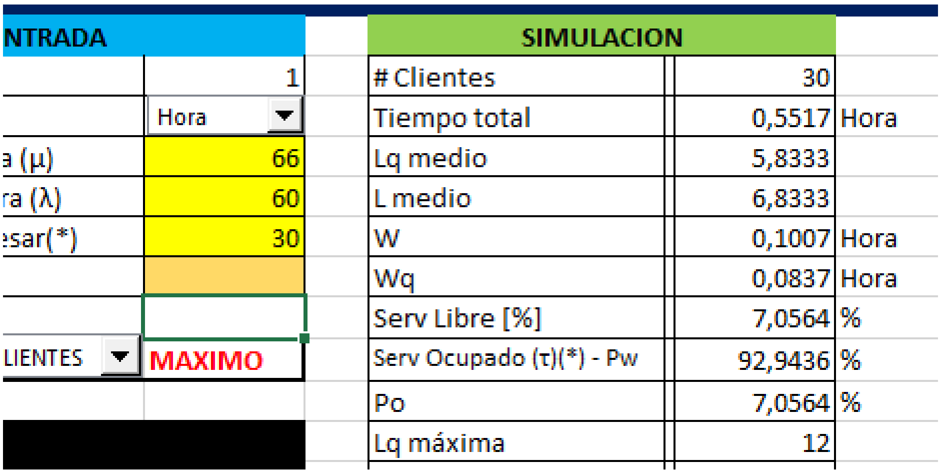
Análisis y comparación de resultados

La base de análisis la estableceremos en el modelo numérico teórico de un sistema M/M/1, que predice los siguientes resultados:

1) para un sistema con λ = 60 cl/h y μ = 66:

NUMERICO		
Serv Ocupado (τ)*	0,9091	90,91%
Po	0,0909	9,09%
Lq	9,0909	
L	10,0000	
Wq	0,1515	
W	0,1667	
Pw	0,9091	90,91%

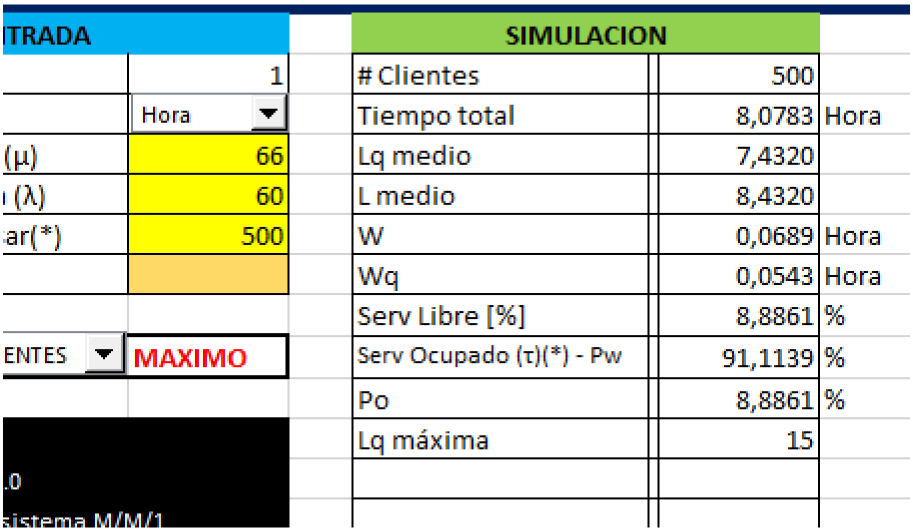
Si lo simulamos para 30 clientes obtenemos corridas parecidas a la siguiente:



NTRADA	
	1
	Hora
a (μ)	66
ra (λ)	60
esar(*)	30
LIENTES	MAXIMO

SIMULACION		
# Clientes	30	
Tiempo total	0,5517	Hora
Lq medio	5,8333	
L medio	6,8333	
W	0,1007	Hora
Wq	0,0837	Hora
Serv Libre [%]	7,0564	%
Serv Ocupado (τ)(*) - Pw	92,9436	%
Po	7,0564	%
Lq máxima	12	

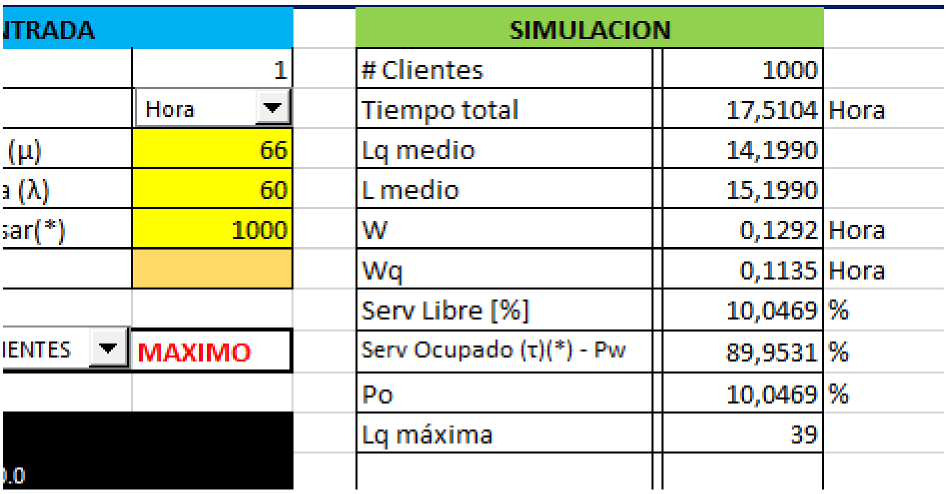
Como vemos los valores están en los órdenes previstos por el método teórico. Si ahora probamos simular 500 clientes, tendremos valores que, en términos generales convergen más hacia los previstos por el método numérico:



ITRADA	
	1
	Hora
(μ)	66
(λ)	60
ar(*)	500
ENTES	MAXIMO
.0	
sistema M/M/1	

SIMULACION		
# Clientes	500	
Tiempo total	8,0783	Hora
Lq medio	7,4320	
L medio	8,4320	
W	0,0689	Hora
Wq	0,0543	Hora
Serv Libre [%]	8,8861	%
Serv Ocupado (τ)(*) - Pw	91,1139	%
Po	8,8861	%
Lq máxima	15	

y lo mismo ocurre cuando lo llevamos al máximo de capacidad de 1000 clientes:



ITRADA	
	1
	Hora
(μ)	66
a (λ)	60
sar(*)	1000
IENTES	MAXIMO
.0	

SIMULACION		
# Clientes	1000	
Tiempo total	17,5104	Hora
Lq medio	14,1990	
L medio	15,1990	
W	0,1292	Hora
Wq	0,1135	Hora
Serv Libre [%]	10,0469	%
Serv Ocupado (τ)(*) - Pw	89,9531	%
Po	10,0469	%
Lq máxima	39	

Ahora vamos a limitar el tiempo de simulación. Comenzaremos con una simulación de 1 hora y con ello sabremos cuantos clientes están sin atender al cabo de esa hora y cuanto tardaríamos en “despacharlos” si cerráramos el acceso de nuevos clientes. Este análisis no lo podríamos hacer con el método numérico.

Una de las corridas nos indica, como vemos en la siguiente captura, que al cabo de una hora tenemos 2 clientes en la cola y necesitamos un tiempo adicional de 0,045 hora para atenderlos. Otras corridas similares que hemos hecho muestran que estos valores varían entre un mínimo de ningún cliente a un máximo de ocho, siendo lo más frecuente una franja que se ubica entre 2 y 3.

ENTRADA	
	1
	Hora
ra (μ)	66
ɔra (λ)	60
	1000
(*)	1
1:30 es 1,5)	
TIEMPO	MAXIMO
2.0.0	
un sistema M/M/1	
xima: 1000 clientes	
Optimiza	

SIMULACION		
# Clientes	53	
Tiempo total	1,0076	Hora
Lq medio	0,1260	
L medio	1,1260	
W	0,0017	Hora
Wq	0,0009	Hora
Serv Libre [%]	20,5189	%
Serv Ocupado (τ)(*) - Pw	79,4811	%
Po	20,5189	%
Lq máxima	6	
L al momento de corte	2	
Tiempo adicional(**)	0,04545574	
*Equivale a factor de tráfico		
** hasta que se atiende el ultimo cliente		

FORMULACION DE MODELOS DE SIMULACION

Diseñar o formular la simulación de un sistema que deseamos estudiar o comprender conlleva la habilidad en determinar la estructura en términos de un modelo lógico, o físico, que mejor represente las características que deseamos observar de ese sistema.

De esta manera podemos pensar en dos actitudes básicas cuando queremos plantear un diseño:

El ***diseño modular***, que consiste en diseñar y construir modelos para cada uno de los procesos o elementos o subsistemas simples que contiene el sistema estudiado. De esta manera la relación entre estos modelos sencillos nos permitirá obtener el modelo que buscamos para el sistema. Generalmente, en estos casos se usan con pocas variables de entrada y pocas de respuesta. Una de las ventajas de este tipo de diseño es la facilidad para aislar errores y corregirlos sin alterar el resto del diseño.

El ***diseño monolítico***, en cambio, está construido con un solo bloque, generalmente grande, que puede tener como inconveniente la dificultad de aislar errores y que es difícil modificar algo sin alterar todo el simulador. La ventaja es que es más sencillo tener coherencia metodológica y un funcionamiento más armónico e independiente del operador.

Criterios de diseño

- El simulador estará al servicio de los objetivos del estudio en que se inserta. Por eso no existe un único modelo de simulación para un sistema real dado, pues puede haber varias alternativas, una o más para cada aspecto u objetivo que queremos abordar en determinado momento.

- Un modelo no debe simplificar mucho el sistema que se desea simular, pero tampoco hacerlo tan complejo que resulte que su construcción o desarrollo sea muy caro o demasiado lento. No debe estar formado por muchas variables ya que crecerá el tiempo para cargarlo y las dificultades de interpretación o enmascaramiento de efectos.

- Cuando formulamos el simulador siempre hay que darle un estado inicial al sistema.

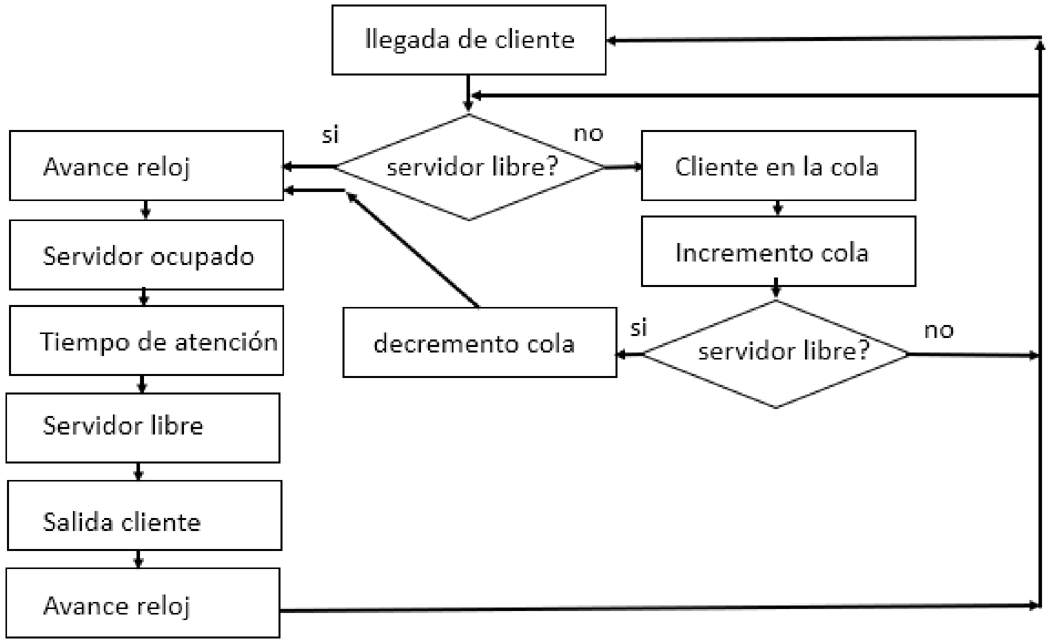
- Hay que tomar decisiones como el orden de la secuencia de cambio de los atributos del sistema y la unidad de tiempo del reloj del modelo.

- Se debe decidir también si será ***sincrónico*** o ***asincrónico***, es decir si el tiempo simulado se incrementará de manera fija, un minuto, un día, un mes, o un año, o bien en incrementos de tiempo variable, por ejemplo que el tiempo avance según la ocurrencia de un evento, como lo hemos discutido unas páginas antes.

Aplicación y comentarios sobre el modelo de Colas M/M/1 visto en la sección anterior.
Cuando se diseña un simulador de este tipo, el diseñador tiene en mente el orden de suceso del sistema natural, que es aproximadamente así: llega un cliente (hora); se observa si el servidor está libre (estado); se procede a atender al cliente (tiempo de

atención); el cliente atendido abandona el sistema (hora de salida); el servidor se desocupa (estado).

Transcribimos a un diagrama de flujo este proceso y nos queda:



Pero tenemos que implementar este proceso de manera tal que los sucesos “llegada”, “salida”, “incremento” y “decremento” ocurran con dependencia parcial entre ellos y no en una secuencia suceso a suceso, sino tiempo de un suceso a tiempo de otro suceso, sea cual fuere. Concretamente, la secuencia real no necesariamente será llegada—salida, sino que puede ser llegada—llegada—llegada—salida—llegada—salida—salida o cualquier otro orden dado por el reloj.

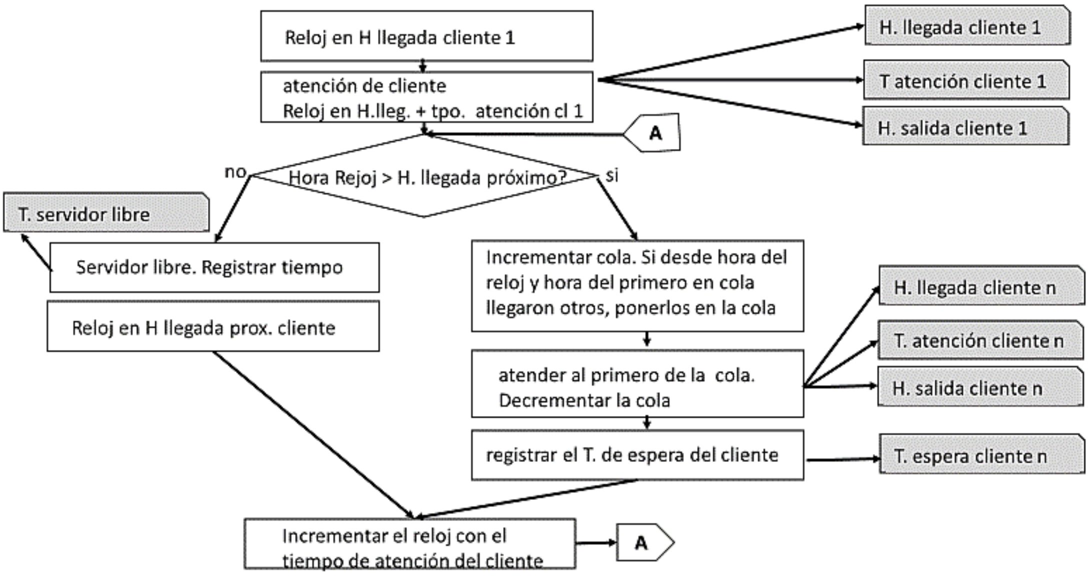
Por eso, hemos generado primero en la hoja de cálculo la hora de llegada de todos los clientes y los tiempos de atención de cada uno de ellos.

Recién luego de ese paso, comienza la secuencia de simulación al poner en marcha el reloj del sistema, que asume como hora inicial la hora de llegada del primer cliente y el tiempo transcurrido hasta ese momento como el primer tiempo de servidor desocupado. El reloj avanza leyendo el tiempo de atención del primer cliente y sumándolo a su hora. Esa hora es la hora de salida del primer cliente.

Interín, mientras el servidor estaba ocupado, pudo haber llegado un nuevo cliente, varios o ninguno. Ese evento configura el largo de la cola. Por lo tanto en los siguientes clientes el servidor examina si hay clientes en la cola, elige al primero y lo atiende igual que antes. Si la cola está vacía espera que llegue uno nuevo.

Por tanto, el proceso real de diseño del simulador se deberá ajustar a un nuevo diagrama de flujo, que contemple acciones básicas como la secuencia de proceso y los puntos de salida de información.

La figura siguiente representa ese diagrama, en el que se omitieron detalles como la finalización del proceso, la contabilidad final del modelo y otros, pero que refleja los aspectos básicos.



Las variables

Si buscamos los componentes fundamentales de un simulador nos encontramos que uno de ellos es lo que denominamos ***variables***. Son elementos que asumen valores a lo largo de una corrida y que debemos interpretar al final del procedimiento, por eso es muy importante definirlas cuidadosamente porque representan elementos, componentes o subsistemas del sistema en consideración.

Podemos clasificar las variables según el rol que cumplan en el simulador:

- las de ***entrada***, que representan el estado inicial del sistema,
- las de ***salida*** (o información) que representan el estado del sistema al final del proceso de simulación y que se relacionan (y a veces coinciden) con las variables
- de ***proceso*** que representan el estado del sistema durante la simulación.

Cualquiera de ellas puede ser estocástica (o de comportamiento probabilístico, rara vez incertidumbre) o determinísticas.

Cualquiera de estas, también, puede ser exógena o endógena, según represente componentes externos al sistema o internos a él, respectivamente.

Por último, y fuera de la clasificación anterior, se encuentran las ***variables de estado*** que caracterizan la condición en que se encuentra el modelo del sistema en un punto del tiempo simulado.

Cuando abordamos un proyecto de simulación debemos tener en cuenta algunos aspectos importantes relacionados con las variables.

- Buscaremos la manera de definir variables para las que sea sencilla la carga de sus valores y con datos en cantidad y costo adecuados a las restricciones que tengamos.

- Respecto de las ***variables de entrada*** debemos asegurarnos de que

    - nos permitan disponer de un rango amplio de control sobre el sistema, para predecir las respuestas del modelo ante variación de datos e inferir lo más ampliamente sobre los aspectos interesantes del sistema real.

    - estén definidas de manera tal que seamos capaces de responder a múltiples alternativas que determine el usuario.

    - deben ser pocas, la menor cantidad posible, pero suficientes para permitir obtener todas las variables de salida necesarias.

- Si las variables están bien definidas, hay coherencia entre ellas y armonizan con el simulador en su conjunto, entonces el modelo no debería dar respuestas absurdas en el rango que el analista determinó. Ni esas respuestas irán contra lo que es observable en el sistema real, ni en contra de lo que universalmente se acepta como verdadero en el sistema.

Ejemplos de simulación en dos juegos.
Vamos a plantear dos juegos, uno con solución analítica obvia, (Juego de la raspadita) y el otro sin posibilidad de tener respuestas analíticas sencillas. De esta manera tendremos, en el primer caso, la posibilidad de comparar resultados (entre las salidas de la simulación y las respuestas analíticas) y, en el segundo, de ver enfoques diferentes al momento de plantear un modelo.

Juego de raspadita

Una empresa de golosinas decide lanzar una promoción de sus pastillas "*Quicky*" Cada paquete se acompañará de una tarjeta que tiene tres filas con dos casillas cada una. De esta manera en cada fila una de las casillas tendrá un valor oculto de 1 y la otra uno de 5, aleatoriamente una u otra. El jugador debe raspar solamente una de cada renglón. Si obtiene los tres renglones con números iguales se gana la cantidad de paquetes indicado (uno, cinco o ninguno si los números son diferentes). La pregunta es

¿Cuál es la mínima cantidad que se debe incrementar el precio por unidad de Quicky que incluya una tarjeta para compensar el costo de los premios y no tener pérdidas?

Asumiremos que el costo unitario por paquete es de $ 1, (de esta manera tomamos el juego como ganancias o pérdidas de dinero)

Como el problema es simple, podríamos aplicar la teoría de probabilidades y construir el modelo que permita obtener la respuesta: la cantidad a cobrar debe ser al menos igual a las ganancias esperadas por los consumidores, que es lo que la empresa debe pagarles. La ganancia esperada se calcula como la esperanza matemática:

Ganancia esperada =
$ 1/paquete x *(probabilidad de ganar 1 paquete)*
+
$ 5/paquete x *(probabilidad de ganar 5 paquetes)*

La probabilidad que tiene un consumidor de ganar 1 o 5 paquetes ($1 o $5) es la probabilidad de que la casilla raspada en las tres filas tenga 1 (o 5):

P($1) = P(fila1,1) x P(f2,1) x P(f3,1) = ½ x ½ x ½ = 1/8 ,
P($5) = P(fila1,5) x P(f2,5) x P(f3,5) = ½ x ½ x ½ = 1/8 ,

por tanto

Ganancia esperada = $ 1 x (1/8) + $ 5 x (1/8) = $ 0,75

Como vemos, ese importe sería el costo estadístico (teórico) de la promoción para la empresa ya que para simplificar no tenemos en cuenta otros costos.

Es un modelo aceptable. Pero ¿cuánto nos conforma? ¿es aceptable? Analicemos esto con un ejemplo también sencillo: supongamos que tenemos un dado y queremos saber si al arrojarlo va a salir un as. Podemos afirmar, con el criterio recién visto, que tirándolo varias veces, y sabiendo que la probabilidad de que salga un as es de uno de cada seis tiros, tendremos el as luego de algo así como seis tiros. Eso es formalmente correcto, pero, en la realidad, tirando realmente un dado, el as ¿va a salir en el primer tiro y en el séptimo? ¿Va a haber un solo as en seis tiros? Si se hace la prueba a veces en el primer tiro sale y otras veces se tiran diez o doce veces el dado y no sale ninguno, o sale cinco veces. Estos resultados que parecen erráticos, sin embargo, no lo son, ya que tirando muchas más veces se obtendrá una proporción de salidas que, efectivamente, tiende a ser de 1 en 6.

Sería interesante saber, por un lado lo que el modelo predice: un as cada seis tiros, basado en que, en muchos tiros se va a dar esa proporción, por otro lado, frente a un tirador determinado, en un número acotado de tiros, ¿Cuántos ases van a salir?

Acá podemos construir un sencillo simulador de un dado y verificar cuantos ases salen en, digamos, 15 tiros.

	A	B	C	D	E	F
1	Simulador de dado					
2						
3	1	15	Cantidad de tiros			
4	4	4	Cantidad de ases			
5	2	0,26666667	Relación - Relación esperada	0,16666667		
6	3					
7	4					
8	1					
9	2					
10	4					
11	6					
12	5					
13	2					
14	2					
15	3					
16	1					
17	1					
18						

En la columna A escribimos la función =ALEATORIO.ENTRE(1;6) con lo cual generamos la cantidad de eventos que necesitemos, arrastrando hasta A17 o hasta A5000.

En B3 escribimos =CONTAR(A:A) con lo cual comenzamos la contabilidad de los eventos. Acá contamos todos los eventos realizados. Al usar "A:A" lo hacemos independiente de cuantos eventos decidimos probar, el simulador servirá igual para dos que para 10000.

En B4 escribimos =CONTAR.SI(A:A;1). Acá contamos cuantas veces apareció un as (o el número que queramos, usando el mismo criterio anterior respecto a la cantidad de eventos.

En B5 escribimos =B4/B3. Lo que hacemos en esta celda es presentar una media de ases respecto a los tiros totales. Teóricamente debería tender a 0,166667. Pulsando la tecla F9 vemos que este valor es errático.

En esta figura vemos otra corrida con F9

	A	B	C	D	E
1	Simulador de dado				
2					
3	1	15	Cantidad de tiros		
4	2	2	Cantidad de ases		
5	5	0,13333333	Relación - Relación esperada	0,16666667	
6	2				
7	3				
8	6				

En la figura siguiente vemos un resultado simulando 2693 tiros.

A	B	C
Simulador de dado		
6	2693	Cantidad de tiros
2	418	Cantidad de ases
6	0,15521723	Relación - Relación
2		

En el caso de los paquetes de pastillas querríamos conocer lo que ocurriría en la realidad y para eso vamos a suponer que tenemos las tarjetas impresas y se las entregamos, en carácter de prueba, a una cierta cantidad de personas. Esta sería una simulación física, realizada imprimiendo, por ejemplo 200 tarjetas, dándoselas a determinadas personas para que raspen y tabular los resultados. Usando las ganancias obtenidas se puede calcular un promedio y una desviación. Obtendríamos así una estimación de ganancia y podemos usar esos datos para tomar decisiones.

¿Porqué queremos hacer esto? Porque la solución analítica no nos dice nada sobre cosas vagas e imposibles de definir como “la buena (mala) suerte del que raspa” “la corazonada” “una buena (mala) racha” y otras cosas por el estilo que tratan de explicar los desvíos puntuales y en números acotados de ensayos que se apartan del modelo. Pero, por otra parte, imprimir tarjetas y repartirlas entre 200 personas y registrar los resultados no resulta muy cómodo ni sencillo ni barato.

Otra manera es mediante simulación analógica: por ejemplo lanzando al aire monedas (recordando que cada cara de una moneda tiene una probabilidad de aparecer p = ½). Establecemos que tres lanzamientos cara son 1 paquete, tres cruces son 5 paquetes, los demás son cero paquetes. De esta manera u otra similar, evitamos imprimir tarjetas, pero no nos simplifica la tarea de recopilar datos y llevar la contabilidad.

La simulación por computadora es ideal para estos problemas, ya que se pueden arrojar las monedas, tabular los resultados y calcular lo necesario mediante una sola operación simple.

Todos los lenguajes de programación son capaces de generar números aleatorios con distribución uniforme. Veremos a continuación un pequeño programa en Basic capaz de hacerlo:

```
RANDOMIZE
FOR i=1 TO 200
        a = RND
        b = RND
        c = RND
        IF a > .499 THEN a = 1 ELSE a = 5
        IF b > .499 THEN b = 1 ELSE b = 5
        IF c > .499 THEN c = 1 ELSE c = 5
        IF a = b AND b = c THEN
                gano = gano + a
                aciertos = aciertos + 1
```

```
        END IF
NEXT i
‘
PRINT “Aciertos  : “; aciertos
PRINT “% de ac.  : “; aciertos * 100/200
PRINT
PRINT “Ganancias : “; gano
PRINT “Promedio  : “; gano / 200
```

La siguiente es una salida obtenida con una semilla de 2

```
Aciertos  : 53
% de ac.  : 26.5

Ganancias : 153
Promedio  : 0.765
```

La ventaja de este programa es que se puede simular para, por ejemplo, 11200 intentos. Los resultados con la misma semilla son:

```
Aciertos  : 2744
% de ac.  : 24.5

Ganancias : 8140
Promedio  : 0.7268
```

O una de 100000 tarjetas, igual semilla:

```
Aciertos  : 24938
% de ac.  : 24.93

Ganancias : 74390
Promedio  : 0.744
```

Por último se puede intentar correr una simulación de 1000000 apuestas:

```
Aciertos  : 249759
% de ac.  : 24.976

Ganancias : 746287
Promedio  : 0.7463
```

Como vemos, un simple programa nos permite hacer cualquier tipo de ensayo y, aún disponer de distintos resultados. Podemos, además, concluir que a medida que avanza el número de apostadores hay una tendencia a un 25% de aciertos y a un promedio de costo para el fabricante (ganancia para el apostador) por tarjera de 0,75, que son los valores que se obtendrían por aplicación lisa y llana del modelo estadístico.

Simulando “la raspadita” en Planilla de Cálculo

Podemos realizar este modelo de simulación de forma mucho más sencilla todavía si empleamos una hoja de cálculo. Por ejemplo:

1. A partir de un reglón cualquiera (en este ejemplo del renglón 7) simulamos una tarjeta. Así el renglón 7 será la primera tarjeta, el 8 la segunda y el 107 la centésima.

2. La columna A, representa el resultado, para cada renglón, de la primera raspada, la B la de la segunda y la C la de la tercera, pero en formato de número aleatorio de distribución uniforme entre 0 y 1.

3. Las columnas D, E y F adoptan valores de "1" y de "5" en función de los resultados de las A, B y C respectivamente.

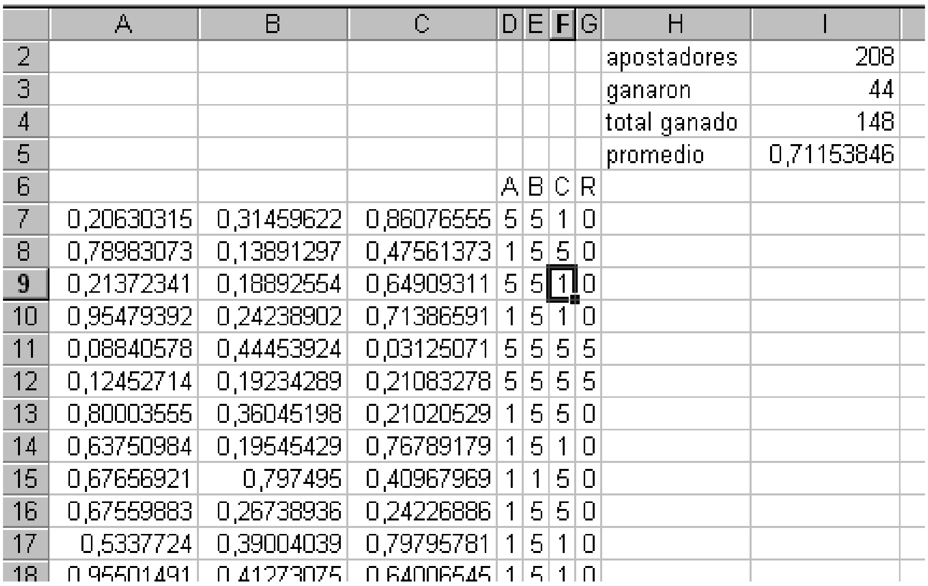
4. La columna G le da un valor al juego (0, si no hay coincidencias, 1, si las coincidencias son de unos y 5 si las coincidencias son de cincos):

CELDA	FUNCIÓN	COMENTARIOS
A7	=ALEATORIO()	Calcula un nº aleatorio entre 0 y 1
B7	=ALEATORIO()	Ídem
C7	=ALEATORIO()	Ídem
D7	=SI(A7<0,5;1;5)	Si el nº en A7 es menor que 0,5 escribe 1, si es mayor escribe 5
E7	=SI(B7<0,5;1;5)	Ídem para B7
F7	=SI(C7<0,5;1;5)	Ídem para C7
G7	=SI(Y(D7=E7;E7=F7);SI(D7=1;1;5);0)	Si los tres son iguales y si valen 1 escribe 1, valen 5, escribe 5 si son distintos escribe 0
I2	=CONTAR(A:A)	Cuenta el número de ensayos. Esta función solo cuenta las celdas que tienen valores numéricos.
I3	=CONTAR.SI(G:G);"<>0")	Cuenta solo las ganadoras (celdas de la columna G distintas de cero)
I4	=SUMA(G:G)	Cálculo del importe total ganado por los apostadores
I5	=PROMEDIO(G:G)	Cálculo del valor medio (premio medio) de todas las tarjetas

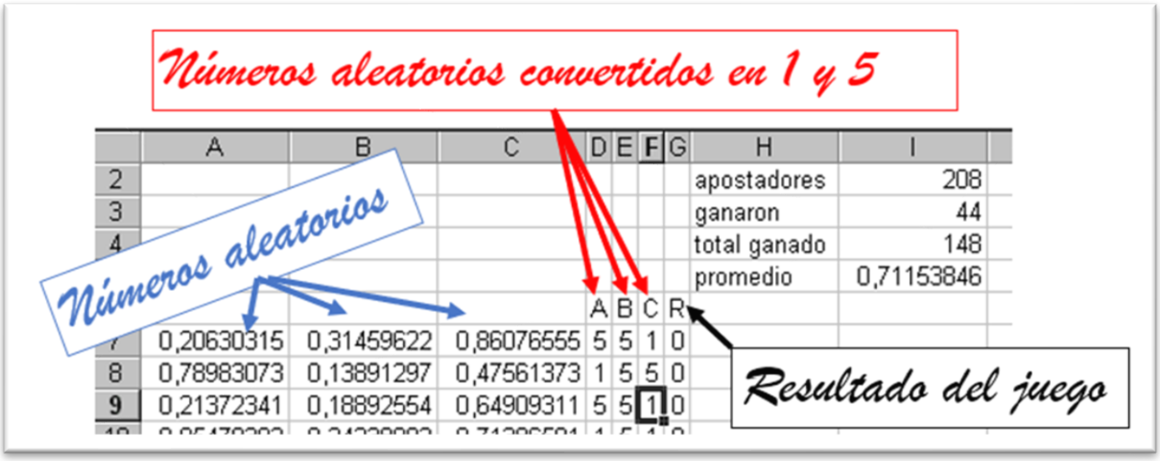
Usando "autocompletar" esta planilla se realiza en pocos segundos, ya que solo se escribe la fila 7 en todas las columnas A, B, C. D, E, F y G, y luego se arrastra hasta el renglón deseado a fin de tener el número de renglones equivalentes a la cantidad de tarjetas que se simulan.

En las celdas I2, I3, I4 e I5 presentaremos los resultados de la contabilidad del modelo.

Una vez terminada la planilla deberíamos de obtener un resultado como el que se muestra en las figura siguientes. Recordemos que la función ALEATORIO() es volátil, por lo cual los valores cambiarán cada vez que hagamos alguna operación.



	A	B	C	D	E	F	G	H	I
2								apostadores	208
3								ganaron	44
4								total ganado	148
5								promedio	0,71153846
6				A	B	C	R		
7	0,20630315	0,31459622	0,86076555	5	5	1	0		
8	0,78983073	0,13891297	0,47561373	1	5	5	0		
9	0,21372341	0,18892554	0,64909311	5	5	1	0		
10	0,95479392	0,24238902	0,71386591	1	5	1	0		
11	0,08840578	0,44453924	0,03125071	5	5	5	5		
12	0,12452714	0,19234289	0,21083278	5	5	5	5		
13	0,80003555	0,36045198	0,21020529	1	5	5	0		
14	0,63750984	0,19545429	0,76789179	1	5	1	0		
15	0,67656921	0,797495	0,40967969	1	1	5	0		
16	0,67559883	0,26738936	0,24226886	1	5	5	0		
17	0,5337724	0,39004039	0,79795781	1	5	1	0		
18	0,95501491	0,41273075	0,64006545	1	5	1	0		



Uso de macros en Hoja de Cálculo

También podemos escribir una macro en Visual Basic para Office o Calc, usando el lenguaje de macros de OpenOffice o LibreOffice.

A continuación vemos el código desarrollado para el problema de la raspadita. Debemos recordar que hay varias maneras de hacer lo mismo. Este es uno de esos métodos. Cada uno puede probar otros.

```
Sub Raspadita2()
'
' 1er. Paso, limpiamos todos los contenidos de la hoja de trabajo:
  Sheets("Hoja1").Select
    Cells.Select
    Selection.ClearContents

'2do: Escribimos el título de la hoja en la celda A1
  Range("A1").Select
    Selection.Font.Bold = True
    ActiveCell.Value = "PROBLEMA DE LA RASPADITA EN HOJA DE
CÁLCULO"
```

‘3ro. Creamos un cuadro de diálogo que le pregunte al operador cuantas tarjetas desea simular.
‘ la cantidad que ingrese la guardamos en la variable **numtarj**. Si no se ingresa nada
‘ o se cancela termina el programa

```
    numtarj = InputBox("Ingresar el numero de tarjetas a
simular", "ModeliZa ", 100)
    If numtarj = "" Then End
```

‘4to. Comienza un ciclo que se repite tantas veces como tarjetas se simulen.
‘ En este ciclo se trabaja igual que en el ejemplo visto en qBasic

```
  For i = 1 To numtarj
    a = Rnd
    b = Rnd
    c = Rnd
    If a <= 0.5 Then a = 1 Else a = 5
    If b <= 0.5 Then b = 1 Else b = 5
    If c <= 0.5 Then c = 1 Else c = 5
    If a = b And a = c Then
        gano = gano + 1
        sumagana = sumagana + a
        If a = 1 Then gana1 = gana1 + 1 Else gana5 = gana5 + 1
        End If
    Next i
```

‘5to. Terminadas de “jugar” las numtarj tarjetas, imprimimos en la hoja los
‘ resultados obtenidos

```
  Range("C3").Value = " Tarjetas jugadas"
    Range("C4").Value = " Tarjetas ganadoras (" + Format(gano *
100 / numtarj) + " %)"
    Range("C5").Value = " Tarjetas ganadoras de $1 (" +
Format(gana1 * 100 / numtarj) + " %)"
    Range("C6").Value = " Tarjetas ganadoras de $5 (" +
Format(gana5 * 100 / numtarj) + " %)"
    Range("C7").Value = " Importe total ganado"
    Range("C8").Value = " Promedio Ganado"
    Range("B3").Value = numtarj
    Range("B4").Value = gano
    Range("B5").Value = gana1
    Range("B6").Value = gana5
    Range("B7").Value = Format(sumagana) + " $"
    Range("B8").Value = Format(sumagana / numtarj) + " $/Trj"
End Sub
```

	A	B	C	D	E	F
1	PROBLEMA DE LA RASPADITA EN HOJA DE CÁLCULO					
2						
3		100	Tarjetas jugadas			
4		26	Tarjetas ganadoras (26 %)			
5		16	Tarjetas ganadoras de $1 (16 %)			
6		10	Tarjetas ganadoras de $5 (10 %)			
7		66 $	Importe total ganado			
8		0.66 $/Trj	Promedio Ganado			
9						
10						
11						
12						

En esta captura de pantalla vemos el resultado de una corrida de la macro.

Juego de las tres diferencias
Este juego, al contrario del anterior, no tiene solución analítica. Consiste en lanzar una moneda repetidas veces hasta que la diferencia entre el número de caras y número de secas que hayan salido sea tres.

La reglas son sencillas y consisten en que el jugador debe pagar $1 cada vez que lanza la moneda, juega lanzando la moneda hasta que logra la diferencia de 3, con lo que termina la partida y recibe $8. El jugador no puede abandonar el juego cuando ya comenzó, lo que significa que debe seguir tirando la moneda (y pagando $1) hasta que obtenga la diferencia de 3.

Por ello, el jugador gana dinero si el número necesario de lanzamientos es menor que 8, pero pierde si se tiene que lanzar la moneda más de 8 veces.

¿Cómo se podría calcular la esperanza matemática de este juego y, por lo tanto, decidir si hay esperanza de obtener ganancias?
Alternativas:

1. Intentar una resolución analítica, que no es obvia.
2. Intentar una resolución empírica, no segura. Pasar un tiempo relativamente largo lanzando la moneda y anotando los resultados y computar al final las ganancias y pérdidas.
3. Realizar una simulación en computadora.

Vamos a optar por la tercera alternativa mediante la generación de dígitos aleatorios para obtener el resultado del juego y entenderlo bien.

Utilizaremos un generador común de distribución uniforme entre 0 y 1, por eso — al igual que en el juego anterior — si el número aleatorio que obtenemos está comprendido entre 0 y 0,49 equivaldrá a decir que es “cara” (C) y de 0,5 a 1 corresponde a “seca” (S) (o “cruz”).

El siguiente código está escrito en VBA en Excel. El nombre del proyecto es “tres”:

```
Sub _tres()
' sub_tres Macro
' Acceso directo: CTRL+q
```

‘Primera parte: iniciación de variables y limpieza de hoja de cálculo

```
Dim juegostotales, serie, numeroanterior, cara As String
Dim evento, repetido, a, juegos As Long
Dim cuentacara, cuentaseca, diferencia As Long
```

‘Borramos las celdas que usaremos y nos posicionamos en la celda A1

```
Range("A1:G10000").Clear
Range("A1").Activate
```

‘Colocamos el título en letras rojas bold y de los títulos de las celdas con las salidas

```
ActiveCell.Value = "LAS 3 MONEDAS - PULSE CTRL q"
```

```
ActiveCell.Font.Bold = True
ActiveCell.Font.Color = vbRed
Range("A2").Activate
Range("A3").Value = "Juego#"
Range("B3").Value = "Tiros"
Range("C3").Value = "Serie obtenida"
Range("E1").Value = "JUEGOS"
Range("E2").Value = "RESULTADO MEDIO"
```

'<u>Segunda parte:</u> ingreso de cantidad de intentos a simular y activación del generador de números aleatorios. Si no ingresamos nada, fin del programa.

```
juegostotales = InputBox("Cantidad de juegos a simular",
"ModeliZa", 10)
Randomize
    If juegostotales = 0 Then End
```

'<u>Tercera parte:</u> simulación de juego. Un loop hasta la cantidad deseada que está en juegostotales

```
   For juegos = 1 To juegostotales
```

'comenzamos poniendo las variables en cero o vaciándolas

```
      evento = 0
      serie = ""
      numeroanterior = ""
      repetido = 0
      cuentacara = 0
      cuentaseca = 0
      diferencia = 0
```

'creamos la etiqueta "game" y generamos la primera moneda, averiguamos si es Cara o Seca. Contamos que hubo un evento y la cantidad de caras o secas

```
game:
      a = Rnd
      If a < 0.5 Then cara = "C" Else cara = "S"
      evento = evento + 1
      serie = serie & cara
      If cara = "C" Then
                  cuentacara = cuentacara + 1
            Else
                  cuentaseca = cuentaseca + 1
            End If
```

'Averiguamos si la diferencia entre Caras y secas es mayor que 3. Si no es así volvemos a "game". Si si, contabilizamos las ganancias y volvemos a un nuevo juego hasta llegar al total

```
      diferencia = Abs(cuentacara - cuentaseca)
      If diferencia < 3 Then GoTo game

      Range("B" + Format(juegos + 3)).Value = evento
      evento = 0
      Range("A" + Format(juegos + 3)).Value = juegos
      Range("C" + Format(juegos + 3)).Value = serie
      serie = ""
   Next juegos
```

```
'Cuarta parte: presentamos la contabilidad y termina.

contabilidad:
      Range("E1").Value = "JUEGOS"
      Range("G1").Value = juegos - 1
      Range("E2").Value = "RESULTADO MEDIO"
      Range("G2").Value = "=SUM(B4:B" + Format(juegos + 3) +
")/G1"

End Sub
```

Vemos capturas con ejemplos de pantallas obtenidas con 1 simulación, con 10 simulaciones, con 100 simulaciones y con 1000 simulaciones:

las3monedas.xlsm

	A	B	C	D	E	F	G
1	LAS 3 MONEDAS - PULSE CTRL q				JUEGOS		1
2					RESULTADO MEDIO		5
3	Juego#	Tiros	Serie obtenida				
4	1	5	CCSCC				
5							
6							

las3monedas.xlsm

	A	B	C	D	E	F	G
1	LAS 3 MONEDAS - PULSE CTRL q				JUEGOS		10
2					RESULTADO MEDIO		8
3	Juego#	Tiros	Serie obtenida				
4	1	25	CSSSCSCSCCCCSSSCSSCCSSCSS				
5	2	5	SSCSS				
6	3	3	CCC				
7	4	7	SCCSSSS				
8	5	7	CSCCSCC				
9	6	3	SSS				
10	7	5	CCSCC				
11	8	17	SCCCSSSCCCSSSCSSS				
12	9	5	CSSSS				
13	10	3	CCC				
14							

las3monedas.xlsm

	A	B	C	D	E	F	G	H
1	LAS 3 MONEDAS - PULSE CTRL q				JUEGOS		100	
2					RESULTADO MEDIO		8,84	
3	Juego#	Tiros	Serie obtenida					
4	1	5	SCSSS					
5	2	9	CSCCSCSCC					
6	3	11	CSSSCCCCSCC					
7	4	3	SSS					
8	5	3	CCC					
9	6	3	SSS					
10	7	17	SCSCCSCSCCSSSCCCC					
11	8	5	SCSSS					
12	9	5	SCCCC					

las3monedas.xlsm

	A	B	C	D	E	F	G
1	LAS 3 MONEDAS - PULSE CTRL q				JUEGOS		1000
2					RESULTADO MEDIO		8,952
3	Juego#	Tiros	Serie obtenida				
4	1	9	SCSCCSCCC				
5	2	11	SCCSCCSCSCC				
6	3	9	CCSSCCSCC				
7	4	11	CSCCSSCCSCC				
8	5	13	SSCCSCSCCCSCC				
9	6	15	CSSCSSCCCSSCCCC				

Como vemos, con una sola corrida se obtuvo una ganancia de 3 ya que se invirtieron 5 y se recibió un premio de 8, sin embargo en sucesivas jugadas, en promedio, no hay ganancia.

Obviamente estas salidas mostradas son puro azar, jugar una sola vez podría haber dado también una inversión de 25 (ver primer juego de la figura que representa 10 juegos) con una pérdida neta de 17.

Estos promedios de lanzamientos por jugada son estimaciones de la media poblacional verdadera de la distribución de probabilidad teórica.

Los errores más comunes en el uso de la simulación ocurren cuando las conclusiones se basan en muestras muy pequeñas, o porque se basan en un análisis estadístico inadecuado, o porque simplemente no hicimos ningún análisis estadístico.

En el caso de los 10 juegos (promedio 8), figura segunda, la desviación estándar hallada fue 7,26

En este caso el desvío medio de la muestra es:

$$\frac{\sum|x_i - x|}{N} = 5,2$$

Por ello, aun cuando se suponga que la distribución de probabilidad del número de lanzamientos requeridos en una jugada es una distribución normal (que es una suposición, ya que la distribución real es sesgada), cualquier intervalo de confianza razonable para la media verdadera de esta distribución va más allá del 8.

Así, antes de concluir, se requiere un tamaño de muestra más grande.

Recuerde que en la distribución normal, el área que abarcamos representa la probabilidad de ocurrencia, tomando 1, 2 o 3 sigmas en torno al promedio, y es:

1 sigma ➔ 68,34%
2 sigma ➔ 95,44%
3 sigma ➔ 99, %

1 sigma significa que hay un 68,3% de probabilidades de que un valor de la variable tomado al azar esté en el intervalo (promedio — 1 sigma , promedio + 1 sigma).

En este caso, ocurre que la media verdadera de la variable "Número de lanzamientos que se requieren en una jugada", es 9[71]. Así, se perdería $1 cada vez que se participa en el juego.

Relaciones funcionales

Las relaciones funcionales son las que describen el comportamiento de las variables consideradas durante el tiempo de simulación.

Las relaciones, como las variables mismas, pueden ser determinísticas o estocásticas. Si una de las variables fuera aleatoria probabilística, entonces la relación funcional que la vincula con las demás también sería probabilística.

Al inicio del proceso de diseño debemos decidir la probabilidad de que las variables asuman determinados valores en determinados intervalos de tiempo (que, pueden ser intervalos fijos o no). Generalmente esas probabilidades que asignamos corresponden a distribuciones de probabilidad conocidas. Luego de esto aplicaremos la relación funcional entre las variables.

Puede ocurrir también que todas las variables sean determinísticas y que creamos conveniente tener un margen aleatorio para el resultado de la relación funcional. En ese caso la relación tendrá la forma:

$$var = f(var1, var2, ...) + k*aleatorio$$

mientras que en los demás es:

$$var = f(var1, var2, ...)$$

Así, en el primer caso, la relación entre todas las variables es igual a la del segundo, pero afectada por la adición de un término aleatorio, *k*aleat*.

Las relaciones funcionales pueden ser:

---

[71] Este valor lo obtuvimos por otros medios. Hay que tener presente la ventaja de disponer de un simulador tan simple y poderoso que permite realizar un enorme número de tiradas, cuyo único límite es el tamaño de la hoja (y eso porque decidimos mostrar cada jugada, de lo contrario el límite sería aún mayor). El lector puede hacer ensayos modificando muy poco el código y verificar que a un número grande de ensayos el promedio tiende a 9. El autor obtuvo en tres simulaciones sucesivas de 60.000 tiradas los promedios de 8,9954 – 9,006 – 9,054, respectivamente.

- **Estáticas o de transformación**, con este tipo de relaciones podremos analizar el valor que asume una variable pero no la manera en que cambia en el tiempo. Estas relaciones generan el valor de la variable dependiente en forma determinística o estocástica y en función del tiempo simulado.

- **Dinámicas o de transición**, en este caso es posible analizar los valores que asumen las variables, el proceso del sistema, o sea, como evoluciona en el tiempo y el comportamiento de las variables. Describen el comportamiento de la variable dependiente durante todo el tiempo simulado.

Resumen

En resumen, las etapas de simulación son:

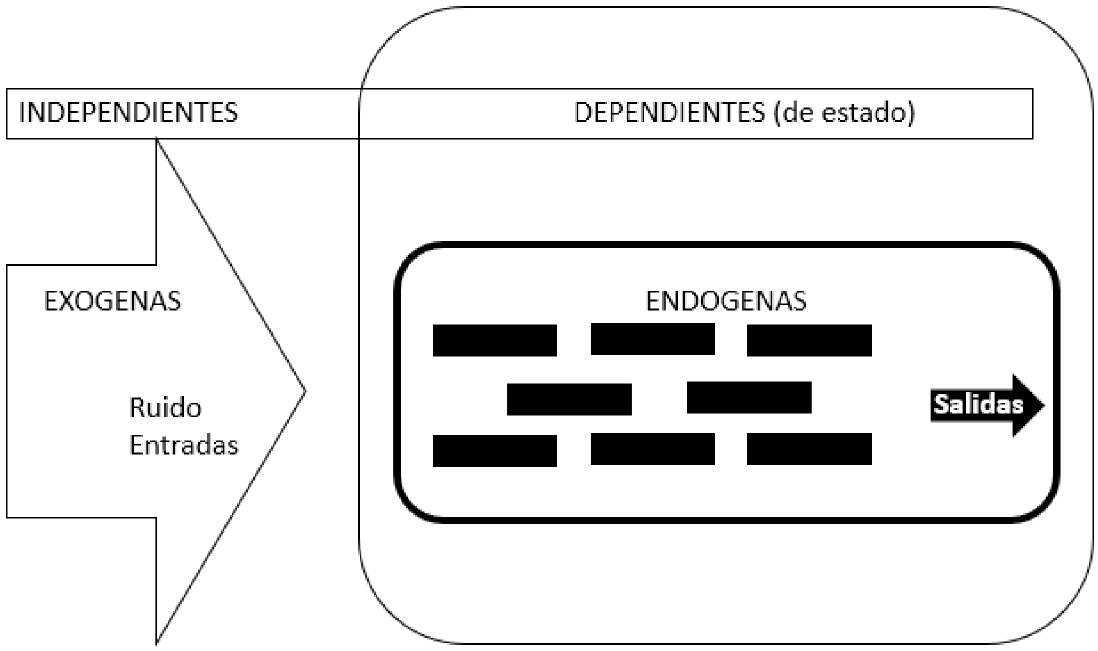
1. Identificación del problema.
2. La definición del sistema, en que se establecen las variables
3. Formulación del modelo, en que se establece la o las funciones.

Los modelos de simulación consisten en combinaciones de los siguientes elementos:

- **Componentes, objetos, entidades**. Son las partes que en conjunto constituyen el sistema. Estos componente u objetos los podemos entender como los elementos o los subsistemas del sistema en estudio.

- **Variables**. Son representaciones simbólicas o numéricas de los objetos, componentes, subsistemas u otros elementos del sistema **real** en estudio.

  - Si son **exógenas** representan entradas al sistema, entrada al modelo o ruido. Significa que su valor no depende ni es función del mismo sistema.
  - Si son **endógenas** representan los objetos del sistema. El valor que asumen se produce dentro del sistema o modelo. Son **variables de estado**. Si se consideran para describir estados finales, son **variables de salida**
  - Puede denominarse variables **independientes** a las variables exógenas y variables **dependientes** a las variables endógenas.



- **Parámetros**. Valores que definen el estado de objetos determinados.

- **Relaciones funcionales**. Presenta variables y parámetros que describen su comportamiento **dentro** de un objeto componente, o **entre** objetos y componentes del sistema.
  - **Relaciones funcionales determinísticas**: Son identidades o definiciones que relacionan ciertas variables o parámetros. En estos casos cada salida del proceso es siempre la misma para entradas iguales.
  - **Relación funcional Estocástica**: Se dan estas relaciones cuando la salida es incierta para una entrada fija dada.
  - Ambos tipos de relaciones funcionales usualmente toman la forma de una ecuación matemática relacionando las variables endógenas o de estado con las variables exógenas. Frecuentemente esas relaciones pueden inferidas luego de efectuarse un análisis matemático y/o estadístico.

- **Restricciones**: Son limitaciones que tienen los valores que pueden asumir las variables o la manera en que los recursos pueden disponerse, distribuirse o usarse. Las restricciones pueden ser definidas por el diseñador o impuestas naturalmente o lógicamente por el sistema.

- **Función**. Es una expresión explícita de los objetivos o metas establecidos para el simulador del sistema y además determina de qué manera será evaluado el grado de cumplimiento de esos objetivos o metas, mediante el ***criterio***.

- **Criterio**. Forma parte de la función e indica cómo se medirán los logros del o de los objetivos. Ni la función ni el criterio deben ser ambiguos.

Respecto al **criterio** y a los **objetivos** es necesario tener presente que existen tanto para el sistema en estudio como para el propio simulador del sistema. Por eso es importante tener mucha claridad en las medidas de desempeño que definiremos porque tienen una influencia alta en el diseño y manipulación del modelo. Una afirmación errónea del criterio de desempeño conduce a conclusiones equivocadas.

El **funcional** y su **criterio** (función y medida del desempeño) se consideran habitualmente como parte del modelo, y todas las operaciones y trabajos que se hacen en el modelo tienen como destino optimizar o satisfacer el criterio establecido.

Cuando se construye el modelo se hacen simplificaciones del sistema original, por eso esas simplificaciones deben ser tenidas en cuenta cuando se formulan las medidas de desempeño para no exigir más de lo que se diseñó; y cumplir con lo que el usuario del simulador define como mínimo.

> *Hacer un modelo es un arte que consiste en la capacidad y habilidad de analizar un problema, aislarlo en sus hechos esenciales, seleccionar y modificar supuestos básicos que caracterizan el sistema, y así elaborar y enriquecer el modelo hasta una razonable aproximación a los resultados.*

En general, cuando construimos un modelo siempre seguimos pasos similares. La siguiente es una guía de algunos procedimientos básicos para construir un modelo. Por supuesto no es exhaustiva, pero sirve en la mayoría de los caos.

1. Desglosar el *sistema — problema* en *sistemas — problemas* simples.

2. Establecer los límites del sistema real y del sistema modelo
3. Definir las variables que son relevantes y su tipo (determinísticas, estocásticas, de estado, etc.)
4. Establecer claramente los objetivos del estudio y del simulador.
5. Definir las restricciones del sistema real y del modelo
6. Buscar analogías con el sistema en estudio.
7. Considerar instancias o ejemplos – numéricos, mecánicos, etc. – del problema.
8. Establecer un glosario (nombre de las variables, de los parámetros, etc.)
9. Secuenciar los módulos.

La bondad del modelo se mide en términos del mayor a menor realismo con que entregue indicadores respecto del sistema real, recordando siempre que, a diferencia de los problemas matemáticos habituales, modelar no es resolver,

Tendremos en cuenta que utilizamos el modelo construido para entender el sistema original y ser capaces de tomar decisiones sobre los aspectos o variables que nos interesan.

Cualquier buen modelo de simulación:
1. Debe ser fácil de entender por el usuario.
2. Debe tener metas y propósitos claros y declarados.
3. Debe ser seguro y confiable, no debe dar salidas absurdas.
4. Debe ser robusto, es decir que debe funcionar sin sorpresas, sin “colgarse”, sin errores y capaz de detectar fallas en la carga de parámetros o datos.
5. Debe ser fácil de usar, controlar y manipular.
6. Debe ser sencillo la comunicación entre el usuario y el simulador.
7. Debe ofrecer salidas que cubran todos los aspectos que consideramos importantes cuando lo construimos.
8. Debe ser adaptativo, permitiendo modificar el modelo frente a cambios de paradigmas.
9. Debe ser evolutivo, partiendo de una forma simple pueda llegar a ser más complejo en la interacción con el usuario.

CAPITULO 7

SIMULACIÓN DE INVENTARIOS Y COLAS

Modelo con distribución de probabilidades específica

Veremos un caso específico en el cual presentaremos una importante variación de la manera de provocar los acontecimientos.

En una ciudad regenteamos nuestro emprendimiento de alquiler de consolas de juego de alta gama "The Gamer Club" que se dedica exclusivamente a alquilar consolas y juegos de última generación y exclusivamente recién lanzados al mercado. Compramos juegos en BD a $US 25 la copia y los alquilamos a $ 3 por día, luego un mes las pasamos a otra sucursal que se destina a todos los títulos y no a novedades exclusivas ubicada en otro barrio a un precio residual de $5 la copia. Las probabilidades de demanda diaria de los juegos son:

COPIAS	PROBABILIDAD
0	0,15
1	0,25
2	0,45
3	0,10
4	0,05

¿Cuántas copias deben comprarse?

Si llamamos ***n*** al número de copias adquiridas, podemos formalizar el problema de esta manera:

Ganancia esperada
=
ingresos mensuales esperados
+
precio de venta de las copias
–
costo de las copias

Ganancia esperada = ingresos diarios esperados x 30 días + $5 x n – $25 x n

a su vez, los ingresos diarios esperados serán (si tenemos las 4 copias):

ingresos diarios esperados =
= Σ(ingreso por juego) x P(número de juegos demandadas) =
= $0 x P(0 pedidos) + $3 x P(1 pedido) + $6 x P(2) + $9 x P(3) + $12 x P(4) =

= $0 x 0,15 + $3 x 0,25 + $6 x 0,45 + $9 x 0,10 + $12 x 0,05 = $ 4,95

con lo cual tendremos:

Ganancia esperada = $4,95 x 30 días + $5 x n – $25 x n = 68,5 =
= 30.f(n) – 20n

Recordamos que ***n*** es el número de copias, y ***f(n)*** será los ingresos diarios esperados de acuerdo con el número de copias, que calculamos de la siguiente manera:

0 copias	$0 x *0,15* + $0 x *0,25* + $0 x *0,45* + $0 x *0,10* + $0 x *0,05* = *$ 0*
1 copias	$0 x *0,15* + $3 x *0,25* + $3 x *0,45* + $3 x *0,10* + $3 x *0,05* = *$ 2,55*
2 copias	$0 x *0,15* + $3 x *0,25* + $6 x *0,45* + $6 x *0,10* + $6 x *0,05* = *$ 4,35*
3 copias	$0 x *0,15* + $3 x *0,25* + $6 x *0,45* + $9 x *0,10* + $9 x *0,05* = *$ 4,80*
4 copias	$0 x *0,15* + $3 x *0,25* + $6 x *0,45* + $9 x *0,10* + $12 x *0,05* = *$ 4,95*

Generamos una tabla con valores de número de copias n de 0 a 4:

n	*Ganancia esperada*
0	0
1	56,50
2	90,50
3	84,00
4	68,50

Podemos comprobar que con 2 copias se obtiene un máximo en la ganancia esperada.

Ahora resolveremos este caso construyendo un modelo de simulación de pedidos de clientes, para cada uno de los días del mes. Vamos a realizar los siguientes pasos:

a. Generaremos la demanda diaria de acuerdo con la distribución de probabilidad que vimos cuando describimos el caso.
b. Calcularemos el ingreso por alquiler basado en que decidimos disponer de dos copias del juego.
c. Calcularemos la ganancia total de esta manera:

Ganancia esperada total
=
ingresos totales por alquiler
+
ingresos por reventa
–
costo de compra
=
= *ingresos totales por alquiler* + 2 x 5 – 2 x 25 =
= *ingresos totales por alquiler* – 40

para obtener los ingresos totales por alquiler generamos una tabla de números aleatorios uniformes, pero dividiendo los intervalos con la probabilidad del número de copias:

```
a = RND

      IF a < .15 THEN a = 0 : Goto final
      IF a < .40 THEN a = 1
      IF a < .85 THEN a = 2
      IF a < .95 THEN a = 3
      IF a < 1   THEN a = 4
```

```
final:
```

como vemos, el intervalo uniforme de los números aleatorios fue dividido en intervalos proporcionales a la distribución enunciada: si obtenemos, por ejemplo, el número aleatorio 0,34342, la demanda diaria que le corresponde es 1 copia, ya que cae entre 0,15 y 0,40. Dicho de otra manera, el número aleatorio lo utilizamos para establecer que probabilidad tiene un evento y proceder según esa probabilidad.

A continuación escribimos un programa que cubra los 30 días.

1. Generamos el número aleatorio para establecer la probabilidad de eventos del día
2. Determinamos la demanda asociada
3. Calculamos los ingresos diarios basados en las dos copias disponibles.

Para simular lo que ocurre con las ganancias en el caso de poseer DOS copias, podemos escribir simplificadamente en VBA:

```
RANDOMIZE
FOR i = 1 TO 30
       a = RND
       IF a < .15 THEN a = 0: GOTO final
       IF a < .4 THEN a = 1
       IF a < .85 THEN a = 2
       IF a < .95 THEN a = 3
       IF a < 1 THEN a = 4
final:
       IF a = 0 THEN diaria = 0
       IF a = 1 THEN diaria = 3
       IF a >= 2 THEN diaria = 6
       acumula = acumula + diaria
NEXT i

Range("XY").Value = "Ingresos totales : " & acumula

Range("XY").Value = "Ganancia total   : "& acumula - 40
```

(Donde XY es un valor genérico que significa, por ejemplo celda A1, en ese caso XY = A1).

Ejemplo de salida:

```
Ingresos totales : 126

Ganancia total   : 86
```

Con este método existe la posibilidad de construir una tabla de ganancias suponiendo que disponemos de 0, 1, 2, 3 o 4 copias:

```
DIM diaria(4)
DIM acumula(4)
CLS
RANDOMIZE
FOR i = 1 TO 30
a = RND
IF a < .15 THEN a = 0: GOTO final
IF a < .4 THEN a = 1
IF a < .85 THEN a = 2
IF a < .95 THEN a = 3
IF a < 1 THEN a = 4
final:
IF a = 0 THEN diaria(1) = 0: diaria(2) = 0: diaria(3) = 0: diaria(4) = 0
```

```
IF a = 1 THEN diaria(1) = 3: diaria(2) = 3: diaria(3) = 3: diaria(4) = 3
IF a = 2 THEN diaria(1) = 3: diaria(2) = 6: diaria(3) = 6: diaria(4) = 6
IF a = 3 THEN diaria(1) = 3: diaria(2) = 6: diaria(3) = 9: diaria(4) = 9
IF a = 4 THEN diaria(1) = 3: diaria(2) = 6: diaria(3) = 9: diaria(4) = 12
       FOR w=1 to 4
       Acumula(w) = acumula(w) + diaria(w)
       NEXT w
NEXT i
Range("XY").Value =  "        *** VIDEO OPERATIVA ***"
Range("XY").Value = "     Ingresos                    Ganancia neta "
Range("XY").Value =  " ------------------------------------------------"
Range("XY").Value = "  0 copia :"; 0
FOR w=1 to 4
Range("Xw").Value = w;" copias :"; acumula(w);" menos ";25*w;" más ";5*w;" :";
acumula(w) - 25*w + 5*w
NEXT w
```

La ventaja es que así podemos simular varias veces o en períodos de más días.

Como ejercitación busquemos introducir cambios en el código para obtener resultados partiendo de otras hipótesis, por ejemplo: cambio de las probabilidades según el mes del año, según el día de la semana, según una determinada probabilidad de aparición de títulos impactantes, vacaciones, etc.

Modelo de colas M/M/1[72]

La teoría de Sistema Líneas de Espera, desarrollada en el Libro 4 muestra que para obtener indicadores de rendimiento aceptables en muy importante disponer de herramientas de simulación adecuadas.

Así cualquier programa dedicado a la solución de este tipo de problemas, como el primitivo WinQSB, por ejemplo, generalmente presenta dos formas de obtener indicadores para ese tema. (Análisis de Colas y Simulación de Colas, en el caso de WinQSB) o, como es lo que ocurre con Rockwell Arena, directamente empleando simulación. En *www.optimiza.org* se puede encontrar una hoja de Excel que reproduce las prestaciones de WinQSB en teoría de colas.

Buscando abordar el tema de la manera más amena posible, desarrollaremos un sistema muy simple de un servidor con llegadas y salidas de clientes que siguen distribuciones exponenciales.

Caso

El Banco IRCS ha instalado un sistema de cinco cajeros automáticos en un centro comercial de 5 plantas que funciona 24 horas al día y desea comprender el funcionamiento del conjunto, para lo cual analiza uno de los cajeros en las horas pico. Se estima entonces que a ese cajero arriba un promedio de sesenta clientes por hora, $\lambda = 60$ y tiene una capacidad de atención promedio de $\mu = 66$ clientes por hora. Como $\mu > \lambda$ haremos un análisis de estado estacionario.

Las llegadas (clientes que solicitan servicio) a un sistema cliente—servidor pueden ser determinísticas o estocásticas. En este caso se trata de procesos estocásticos, lo que originan el símbolo M tanto para las llegadas como las salidas si usamos la notación M/M

[72] En este libro, en el Capítulo 4 (Introducción a la simulación/Diferencias entre métodos numéricos, métodos analíticos y simulación/Simulación) presentamos un modelo M/M/1 en hoja Excel más simple que el que vemos en esta sección.

Proceso de llegada probabilístico: el **tiempo** entre llegadas sucesivas es variable. En los casos de colas se utiliza, generalmente, la distribución exponencial. Así, la función de densidad para las llegadas de clientes, que se denomina λ, (número promedio de llegadas por unidad de tiempo) está dada por

$$f(t) = \left(\frac{1}{\lambda}\right) e^{-\lambda T}$$

Para un tiempo T determinado usamos la función de densidad para calcular la probabilidad de que el siguiente cliente llegue dentro de esas *T* unidades de tiempo a partir de la llegada del anterior:

$$P(tiempo\ entre\ llegadas \leq T) = 1 - e^{-\lambda T}$$

En este caso, los clientes llegan al banco a una tasa promedio diaria de 60 clientes cada hora, λ = 60 clientes/h. De esta manera, si en este instante ingresara un cliente, nos podríamos preguntar ¿Cuál es la probabilidad de que llegue un cliente nuevo en los próximos 10 minutos?

Como establecimos un tiempo entre llegadas, T, de 10 minutos (T=1/6h), entonces podremos calcular de esta manera la probabilidad de que ocurra una llegada, ya que conocemos la tasa de llegadas media λ = 60 clientes/h

$$P\left(tiempo\ entre\ llegadas \leq \frac{1}{6}\right) = 1 - e^{-60\times\frac{1}{6}}$$

Esta probabilidad también la podríamos obtener usando hoja de cálculo, empleando, en EXCEL la siguiente función:

=DISTR.EXP.N(t; λ; Acumulado)

que, en este caso será:

DISTR.EXP.N(0,16667;60;VERDADERO)

=DISTR.EXP(0,16667;60;VERDADERO) – (versiones antiguas)

Otra manera de abordar esta cuestión es refiriéndose al número de clientes que llegan. Por ejemplo, la probabilidad de que dos clientes lleguen dentro de los diez minutos siguientes, en este caso, al ser una variable discreta, la distribución de probabilidad para el número de llegadas será una distribución de *Poisson*:

$$P(que\_en\_tiempo\_T\_lleguen\_k\_clientes) = \frac{e^{-\lambda T}(\lambda.T)^k}{k!}$$

Siguiendo el caso anterior, con λ = 60 clientes/h y T = 1/6 h. La probabilidad de que lleguen dos clientes (k = 2) en los próximos 10 minutos será:

$$P(que\_en\_10\_\min utos\_lleguen\_2\_clientes) = \frac{e^{-60x1/6}(60/6)^2}{2x1} = 0{,}00227$$

También podemos obtener esta probabilidad empleando hoja de cálculo, con la siguiente función:

=POISSON.DIST(K;X;Acumulado),
=POISSON(K;X;Acumulado) (versiones antiguas)

donde en K colocamos el número de eventos k = 2, y en X el producto λT = 60/6 = 10. y en Acumulado ponemos FALSO, ya que no queremos la probabilidad acumulada.

Si buscáramos la probabilidad acumulada la expresión que utilizaríamos sería:

$$P(tiempo\_entre\_llegadasT \leq k) = \sum_{0}^{k} \frac{e^{-\lambda T}(\lambda.T)^k}{k!}$$

y en el caso de usar Excel, incluiríamos el valor lógico Verdadero como parámetro de Acumulado en la función anterior.

Resolución analítica.

Los indicadores de rendimiento que obtenemos con la solución analítica son:

a. Factor de tráfico, o intensidad de tráfico,

$$\tau = \frac{\lambda}{\mu} = \frac{60}{66} = 0{,}9091$$

Esta expresión representa el número de clientes que llegan por cada cliente que es atendido, entonces el factor de tráfico es, en realidad, la **probabilidad de que esté ocupado el servidor.**

b. Probabilidad de que no haya clientes en el sistema (sistema o servidor no ocupado) *(P_0)*

$$P_0 = 1 - \tau = 1 - 0{,}9091 = 0{,}0909$$

(el 9% de las veces el cliente encuentra el cajero vacío, que es lo mismo que afirmar que el 91% de las veces está ocupado)

c. Largo promedio de la fila

$$L_q = \frac{\tau^2}{1-\tau} = \frac{0{,}9091^2}{1-0{,}9091} = 9{,}0909$$

El cajero tendrá un promedio de nueve clientes haciendo cola, sin contar el que está siendo atendido.

d. Tiempo promedio de espera en la cola

$$W_q = \frac{L_q}{\lambda} = \frac{9{,}0909}{60} = 0{,}1515$$

cada cliente que llega debe esperar aproximadamente 9 minutos antes de poder pasar al cajero.

e. Tiempo de espera en el sistema

$$W = W_q + \frac{1}{\mu} = 0{,}1515 + \frac{1}{66} = 0{,}1667$$

cada cliente invierte 10 minutos en el sistema: hacer la cola y ser atendido

f. Promedio de clientes en el sistema: los 9 que hacen cola más el que está siendo atendido

$$L = \lambda W = 60 \times 0{,}1667 = 10$$

g. Probabilidad de que un cliente tenga que esperar

$$P_w = 1 - P_0 = \tau = 0{,}9091.$$

h. Probabilidad de que haya *n* clientes cuando un cliente llegue al sistema

$$P_n = \tau^n P_0$$

Con esta expresión podemos construir la siguiente tabla de probabilidades

n	P_n
0	0.0909
1	0.0826
2	0.0751
3	0.0683
…	…

La probabilidad de que no haya más de tres clientes en el sistema cuando llega un determinado cliente es

$$\Sigma P_i = P_0 + P_1 + P_2 + P_3 = 0{,}0909 + 0{,}0826 + 0{,}0751 + 0{,}0683 = 0{,}3169$$

Podemos agregar a la tabla anterior una columna de acumulados de *Pn*:

n	P_n	*Pn acu*
0	0.0909	0,0909
1	0.0826	0,1735
2	0.0751	0,2486
3	0.0683	0,3169
4	0,0621	0,3790
...	...	...

i. Utilización del cajero

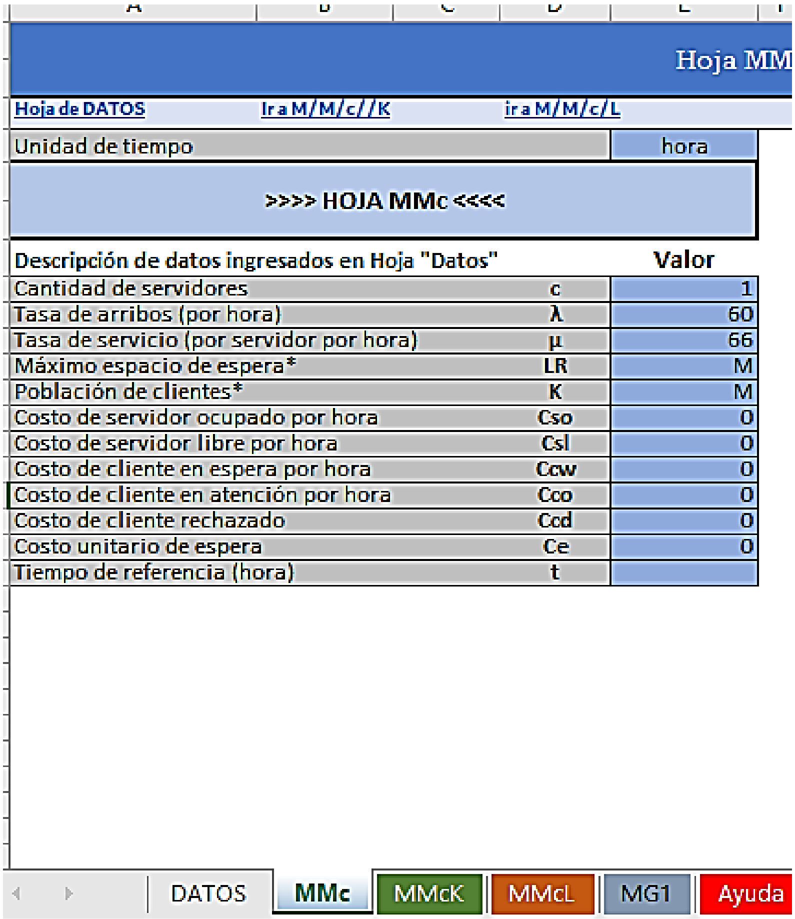
$$U = \tau = 0{,}90901$$

El 91% del tiempo el cajero está en uso.

Resolución con "Análisis de Colas" desarrollado para Excel (2020)[73]

Comenzamos seleccionando la hoja correspondiente al modelo M/M/1 porque tenemos llegadas de clientes con distribución exponencial, lo mismo que para la salida y estamos analizando un solo cajero. La hoja a usar es, entonces, la que aparece como "MMc".

Completamos la ventana de la izquierda con los datos del caso, según vemos en la figura siguiente.


Hoja MM

Hoja de DATOS | Ir a M/M/c//K | ir a M/M/c/L

Unidad de tiempo		hora

>>>> HOJA MMc <<<<

Descripción de datos ingresados en Hoja "Datos"		Valor
Cantidad de servidores	c	1
Tasa de arribos (por hora)	λ	60
Tasa de servicio (por servidor por hora)	µ	66
Máximo espacio de espera*	LR	M
Población de clientes*	K	M
Costo de servidor ocupado por hora	Cso	0
Costo de servidor libre por hora	Csl	0
Costo de cliente en espera por hora	Ccw	0
Costo de cliente en atención por hora	Cco	0
Costo de cliente rechazado	Ccd	0
Costo unitario de espera	Ce	0
Tiempo de referencia (hora)	t	

DATOS | MMc | MMcK | MMcL | MG1 | Ayuda

Los parámetros completados con cero indican que no se conocen o no los tenemos en cuenta en este caso. Los que están completados con sendas letras M indican que no hay límites. En este ejemplo no tenemos límites para capacidad de la cola y población de clientes.

[73] "Análisis de Colas" es un Libro de Excel publicado en 2020 que está disponible en optimiza.org y permite resolver, en forma analítica modelos M/M/c, M/M/c/K, M/M/c/L y M/G/1.

F	G	H	I	J	K	L

MMc – ANALISIS DE COLAS - OPTIMIZA - VER. 20.1.0

ir a M/G/1 ayuda

análisis de sensibilidad

Hoja

MODELO INGRESADO>>> **M/M/1/** **MMc**

Tabla de resultados

Tasa de arribos (clientes/hora)	λ	60,0000	cl/hora
Tasa de servicio (clientes/hora)	µ	66,0000	cl/hora
Uso del sistema	U	0,9091	90,9091%
Media de clientes en el sistema	L	10,0000	
Media de clientes en la cola	Lq	9,0909	
Tiempo medio cliente en sistema	W	0,1667	hora
Tiempo medio cliente en cola	Wq	0,1515	hora
Probabilidad de servidor(es) libre(s)	Po	0,0909	9,0909%
Probabilidad de sistema ocupado	Pb o Pw	0,9091	90,9091%
Media de clientes rechazados por hora			

En el centro de la hoja, parte superior aparecerán los indicadores de rendimiento del sistema analizado. Luego de las tasas de arribos y salidas (que son datos) tendremos el grado de utilización del sistema, el largo promedio del sistema y de la cola y los demás indicadores.

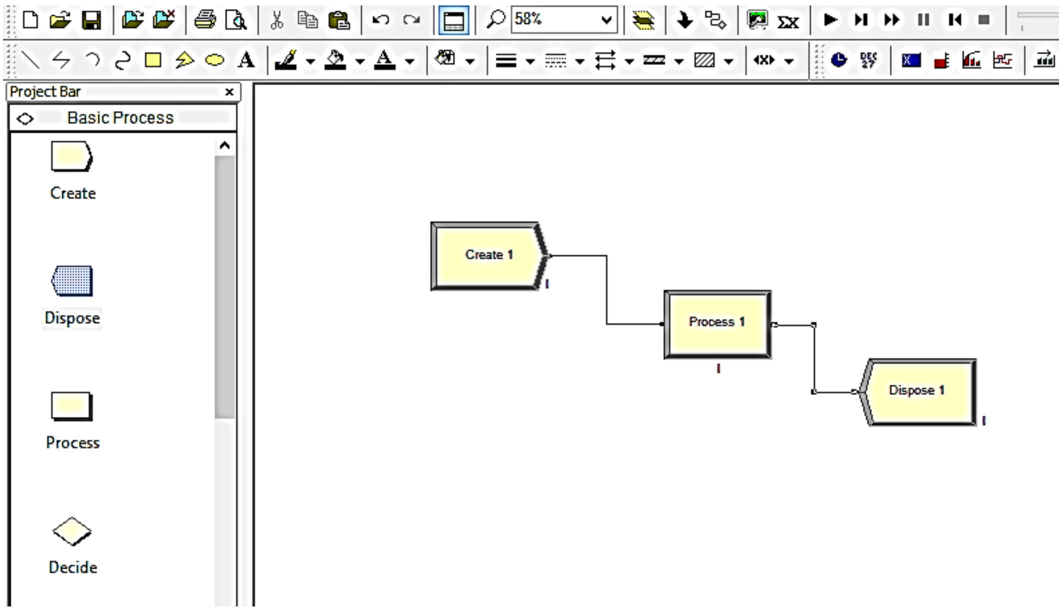
A la derecha de la hoja disponemos de una tabla de probabilidades para *n* clientes y su correspondiente probabilidad acumulada.

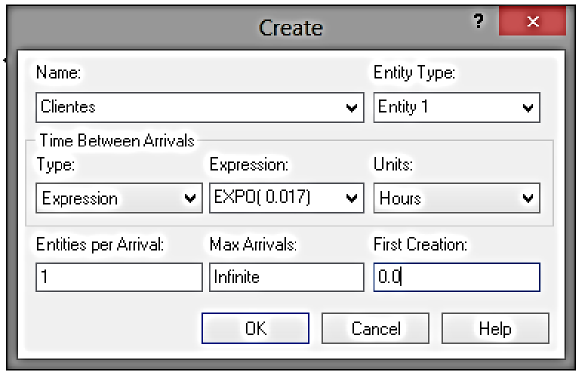
documentación

Resumen de probabilidades

n	P(n)	P(n) ACU
0	0,0909	0,0909
1	0,0826	0,1736
2	0,0751	0,2487
3	0,0683	0,3170
4	0,0621	0,3791
5	0,0564	0,4355
6	0,0513	0,4868
7	0,0467	0,5335
8	0,0424	0,5759
9	0,0386	0,6145
10	0,0350	0,6495
11	0,0319	0,6814
12	0,0290	0,7103
13	0,0263	0,7367
14	0,0239	0,7606
15	0,0218	0,7824
16	0,0198	0,8022
17	0,0180	0,8201
18	0,0164	0,8365
19	0,0149	0,8514
20	0,0135	0,8649

En la hoja de diseño de ARENA se crean tres entidades: Una con "*Create*", la siguiente con "*Process*" y la última con "*Dispose*". Hacer esto es muy sencillo pues a la izquierda de la pantalla de Arena hay una ventana vertical llamada "*Basic Process*" donde simplemente seleccionamos cada uno de los procesos y los arrastramos a la ventana de diseño





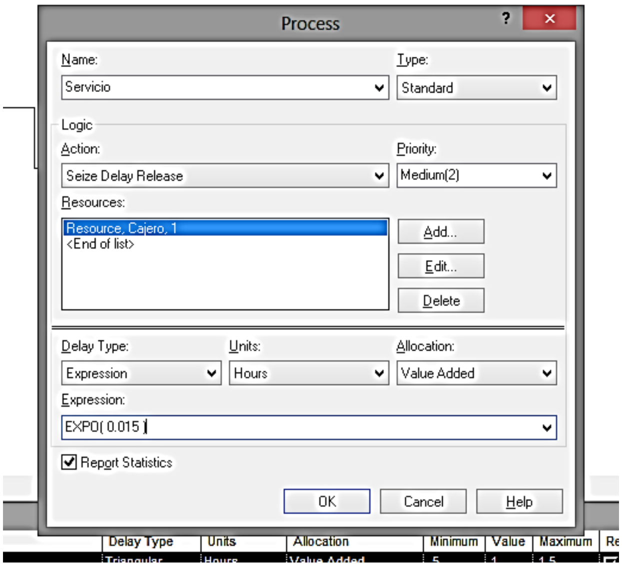
Una vez que las tres entidades están en la ventana, hacemos clic en cada una de ellas y así accedemos a la pantalla de configuración:

Completamos la entidad "*Create*", de la siguiente manera:

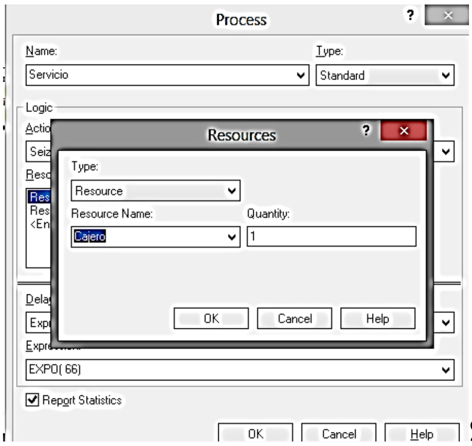
El **Nombre** será "Clientes", dejando señalado que es una entrada, en este caso es la única.

En "**Tiempo entre arribos**" hay tres listas desplegables en las que seleccionamos "*Expression*" en **tipo**, y buscamos en **expresión** "*EXPO(Mean)*", colocando entre paréntesis el valor del tiempo entre clientes que arriban. (1/60 = 0,017) La unidad de tiempo — horas — la fijamos en **unidades**.

Para configurar la segunda entidad, "*Process*", haremos algo parecido. Colocaremos



la etiqueta "Servicio". En la lista que tiene el nombre **Action** seleccionamos la acción "*Seize Delay Release*" (liberar al cliente con retardo), y nos va a aparecer una ventana inferior con una lista de recursos disponibles y tres botones (**Agregar**, **Editar** y **Borrar**).



Si pulsamos en "**Agregar**", aparece abajo una nueva ventana (**Recursos**) en la que escribimos el nombre del recurso, en este caso escribimos "Cajero". Damos OK y esta ventana se cierra, volvemos así a la original.

Podemos ya completar la parte inferior con el tipo de retraso, dicho de otra manera, acá tenemos que definir de qué manera demoraremos al cliente en esta estación:

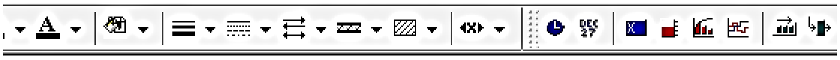
Para este caso elegiremos en la lista "*Data Type*" la opción "*Expression*", y debajo de ella, en la lista "*Expression*" elegimos el tipo EXPO() con una media de 66, por lo tando escribimos entre los paréntesis de EXPO() el valor correspondiente a

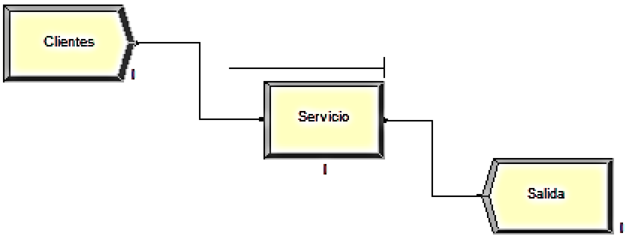
$$\frac{1}{\lambda} = \frac{1}{66} = 0{,}15$$

Solo nos resta pulsar en Aceptar para cerrar esta ventana.

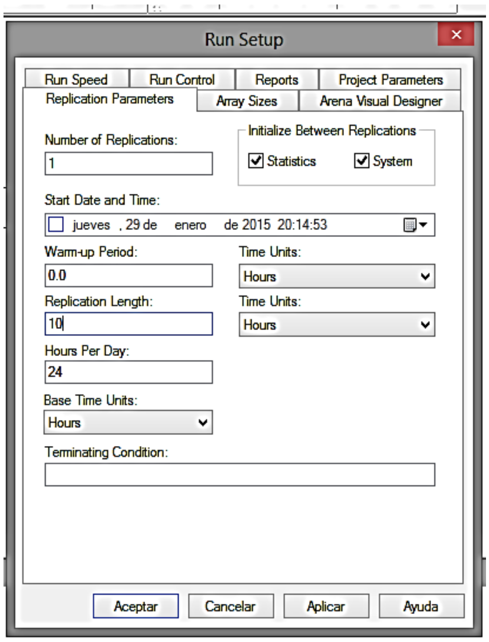
Por último seleccionamos la tercera entidad "*Disposse*", que es, simplemente, donde van a parar los clientes luego del servicio, donde lo único que hay que hacer es colocar el nombre, en este caso "Salida" porque acá no hay ningún proceso ni demora.

El aspecto del lienzo principal es como se ve en la figura, apareciendo encima de "Servicio" una línea donde veremos la cola en tiempo de ejecución-





Luego vamos al menú “*Run*”, y elegimos “*Setup*” y allí buscamos la pestaña “*Replication parameters*”, donde pondremos los parámetros que deseemos para correr la simulación. En este caso seleccionamos hacer una sola replicación que simule 10 hs de funcionamiento, debemos tener en cuenta que este parámetro, por defecto, está en infinito. Si lo dejamos en infinito, entonces no se detendrá la simulación nunca hasta que, en forma manual, decidamos interrumpir el proceso. Debemos prestar mucha atención a que las unidades de tiempo sean coherentes.



El siguiente paso es, en el mismo menú “*Run*” seleccionar “*Go*” o pulsar F5. La velocidad de simulación es controlable en la pestaña “*Run Speed*”.

Una vez finalizado el proceso, se tiene acceso a las estadísticas e informes, en varias páginas, que pueden ser convertidas a un archivo CSV apto para hoja de cálculo:

Unnamed Project
Entity
Time
Other
Queue
Resource

Replications: 1 Time Units : Hours

Entity

Time

VA Time	Average	Half Width	Minimum Value	Maximum Value
Entity 1	0.01513149	0,001036016	0.00002126	0.0992

NVA Time	Average	Half Width	Minimum Value	Maximum Value
Entity 1	0.00	0,000000000	0.00	0.00

Wait Time	Average	Half Width	Minimum Value	Maximum Value
Entity 1	0.1564	(Correlated)	0.00	0.3924

Transfer Time	Average	Half Width	Minimum Value	Maximum Value
Entity 1	0.00	0,000000000	0.00	0.00

Other Time	Average	Half Width	Minimum Value	Maximum Value
Entity 1	0.00	0,000000000	0.00	0.00

Total Time	Average	Half Width	Minimum Value	Maximum Value
Entity 1	0.1715	(Correlated)	0.00008989	0.4051

Other

Number In	Value
Entity 1	599.00

Number Out	Value
Entity 1	594.00

WIP	Average	Half Width	Minimum Value	Maximum Value
Entity 1	10.1922	(Correlated)	0.00	27.0000

Model Filename: Model3 Page 2 of 3

Unnamed Project

11:42:46 **Category Overview** enero 30, 2015

Unnamed Project

Replications: 1 Time Units: Hours

Queue

Time

Waiting Time	Average	Half Width	Minimum Value	Maximum Value
Servicio.Queue	0.1561	(Correlated)	0.00	0.3924

Other

Number Waiting	Average	Half Width	Minimum Value	Maximum Value
Servicio.Queue	9.2920	(Correlated)	0.00	26.0000

Resource

Usage

Unnamed Project

Instantaneous Utilization

	Average	Half Width	Minimum Value	Maximum Value
Cajero	0.9002	(Insufficient)	0.00	1.0000

Number Busy

	Average	Half Width	Minimum Value	Maximum Value
Cajero	0.9002	(Insufficient)	0.00	1.0000

Number Scheduled

	Average	Half Width	Minimum Value	Maximum Value
Cajero	1.0000	(Insufficient)	1.0000	1.0000

Scheduled Utilization

	Value
Cajero	0.9002

Total Number Seized

	Value
Cajero	595.00

Model Filename: C:\Users\Alejandro\SkyDrive\Documentos\Model3 Page 3 of 3

Simulación con Excel, desarrollo propio[74].

Para poder comparar los resultados entre Arena u otros simuladores como el mencionado WinQSB, hemos desarrollado uno propio en hoja de cálculo Excel empleando VBA, que describimos con más detalle en el Capítulo 6, "estudio de caso" y cuyo código y detalles de implementación están en el parágrafo siguiente (Simulación con Excel, desarrollo propio)

La siguiente es una ventana obtenida corriendo el simulador con un tiempo máximo simulado de 10 horas:

Colas M/M/1

DATOS DE ENTRADA	
SERVIDORES:	1
UNIDAD DE TIEMPO	Hora
TASA DE SERVICIO POR Hora (μ)	66
TASA DE LLEGADAS POR Hora (λ)	60
	1000
Tiempo máximo simulado(*)	10
*Ingresar en formato decimal (Ej.: 1:30 es 1,5)	
VOY A SIMULAR POR N° MAXIMO DE	TIEMPO

COLAS Versión 2.0.0
Simulador de un sistema M/M/1
Capacidad Máxima: 1000 clientes
Optimiza
Julio 2021

Start!!

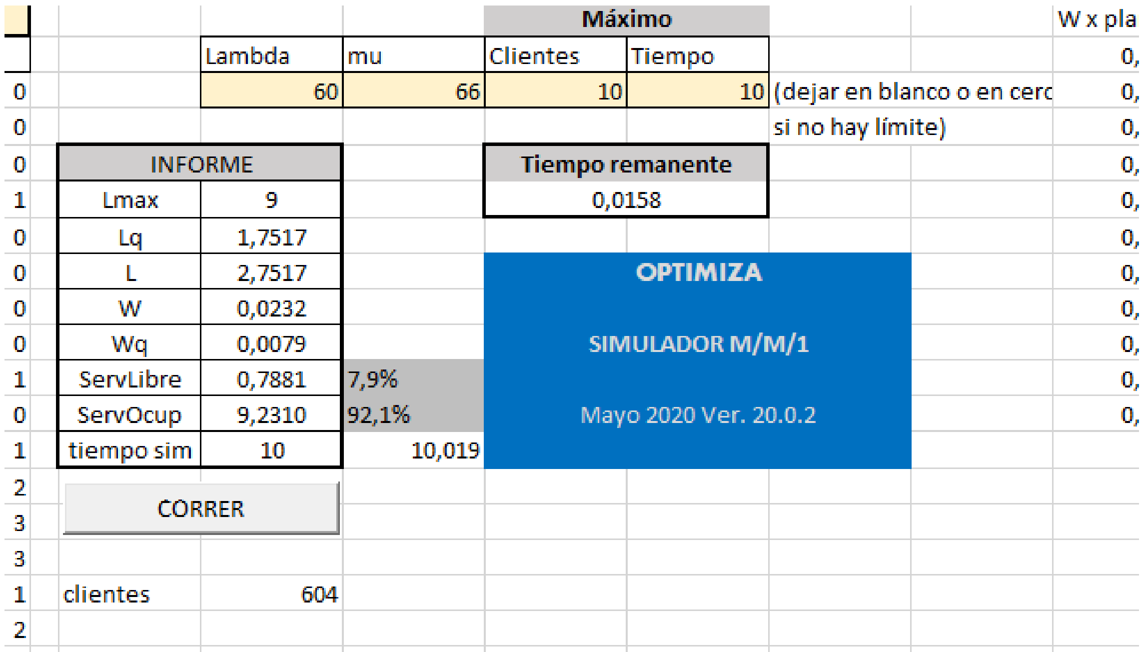
AYUDA: 1) Entradas | 2) Salidas | 3) Numérico

SIMULACION		
# Clientes	579	
Tiempo total	10,0149	Hora
Lq medio	4,3730	
L medio	5,3730	
W	0,0406	Hora
Wq	0,0319	Hora
Serv Libre [%]	13,1600	%
Serv Ocupado (τ)(*) - Pw	86,8400	%
Po	13,1600	%
Lq máxima	18	
L al momento de corte	14	
Tiempo adicional(**)	0,16028513	
*Equivale a factor de tráfico		
** hasta que se atiende el ultimo cliente en la cola		

NUMERICO		
Serv Ocupado (τ)*	0,9091	90,91%
Po	0,0909	9,09%
Lq	9,0909	

Cliente#	H Llegada	T esper
1	0,00669	0,00
2	0,03537	0,00
3	0,04790	0,03
4	0,04832	0,03
5	0,05137	0,04
6	0,12057	0,00
7	0,14071	0,00
8	0,18224	0,00
9	0,18654	0,01
10	0,21490	0,00
11	0,24275	0,00
12	0,25145	0,00
13	0,25263	0,01
14	0,25944	0,02
15	0,33306	0,00
16	0,35008	0,02
17	0,35031	0,05
18	0,36004	0,06
19	0,36780	0,06

[74] Para conocer más sobre este simulador, consultar Capítulo 6, "Estudio de caso"



Comparación de resultados

Observando los tres métodos empleados y tomando solamente alguno de los indicadores de rendimiento se pueden hacer evaluaciones sobre la calidad de los datos obtenidos:

Indicador	Excel analítico	WinQSB simulado	ARENA simulado	Excel simulado
Uso Gral. del sistema	90,9091%	37,61%	90,02	86,84
Largo del sistema L	10	0,56	10	5,3730
Largo Cola Lq	9,09	0,087	9,29	4,3730
Tiempo del sistema W	0,1667 h	0,662 h	0,1715 h	0,0406
Probabilidad Po	9,0909%		1	13,36
Probabilidad Pb	90,90%		90,02%	86,84
Tpo simulado	n/c	10 h	10 h	10 h
Clientes			595	579

Como se ve, en WinQSB simulado hay diferencias. Ambos fueron corridos en simulaciones de 10 horas, pero si la de WinQSB se alarga, por ejemplo, a 500 horas, los valores convergen.

Modelo de simulación de colas M/M/1 en hoja de cálculo.

Presentamos este sencillo modelo de gestión de colas basado en un reloj sincrónico con el cual hemos obtenido resultados similares al simulador de WinQSB. El objetivo de este modelo es poner a prueba la capacidad de desarrollar paradigmas de simulación empleando herramientas de fácil disponibilidad y sin emplear programación especializada ni software dedicado.

El código completo y los comentarios los dividimos en partes para más claridad.

1 Parte: preparación y definición de variables

```
Sub MM1()
'Limpieza de la tabla de cola en la hoja
   Range("A3:I3000").ClearContents
'Definición de variables
    Dim atendido As Integer      'contador de clientes atendidos
    Dim cl As Integer            'contador de clientes que llegan
    Dim i As Integer             'contador de renglones de la hoja
    Dim L As Integer             'Largo del sistema
    Dim lambda As Integer        'tasa de llegadas
    Dim mu As Integer            'tasa de salidas
    Dim W As Double              'tiempo en el sistema
    Dim Wq As Double             'tiempo en la cola
    Dim tLL(1 To 1000) As Double 'instante que se produce la llegada
                                      del cliente cl
    Dim tESP(1 To 1000) As Double 'tiempo de espera del cliente cl
    Dim tSER(1 To 1000) As Double 'tiempo de servicio del cliente cl
    Dim tDES(1 To 1000) As Double 'tiempo de servidor desocupado esperando
                                      al cliente cl
    Dim tSAL(1 To 1000) As Double 'tiempo de servicio del cliente cl
    lambda = Range("L3").Value   'Carga del valor de lambda que está en la hoja
    mu = Range("M3").Value       'idem, de mu
    'definimos nucli como variable para dar fin al proceso. Puede ser número
       de clientes (Celda N3) o tiempo máximo (Celda O3)
    If Range("O3").Value = 0 Then
    nucli = Range("N3").Value
    Else
    nucli = Range("O3").Value
    End If
'Ponemos en marcha el algoritmo de aleatorización
    Randomize
```

2 Parte: El evento "llegada del primer cliente" provoca una actualización de reloj y de estado de sistema

```
    i = 3             'contador de renglón de hoja. Comienza en el 3ero.
    cl = 1            'el número de cliente es el 1
    L = 1             'el largo del sistema será 1
    atendido = 1      'el contador de clientes atendidos será 1
    tLL(cl) = -((Log(1 - Rnd)) / lambda).      'averiguamos cual es el tiempo
                      de llegada del cliente 1
    tESP(cl) = 0      'como se trata del primer cliente, no hay tiempo de espera
    tSER(cl) = -((Log(1 - Rnd)) / mu) ' averiguamos cual es el tiempo de
                      atención del cliente 1
    tSAL(cl) = tLL(cl) + tESP(cl) + tSER(cl) ' calculamos la hora en que sale
                      el cliente 1
    tDES(cl) = tLL(cl) 'El tiempo en que el servidor está libre va desde 0 hasta
                      que llega el primer cliente
    W = tSAL(cl) - tLL(cl) 'el tiempo en que el cliente está en el sistema
    Wq = 0            ' el cliente no tuvo que esperar
    'impresión del primer renglón (i=1) de clientes
    Cells(i, 1) = i
    Cells(i, 2) = cl
    Cells(i, 3) = tLL(cl)
```

```
    Cells(i, 4) = tESP(cl)
    Cells(i, 5) = tSER(cl)
    Cells(i, 6) = tDES(cl)
    Cells(i, 7) = tSAL(cl)
    Cells(i, 8) = L
    Cells(i, 9) = L - 1
    Cells(i, 19) = W
    Cells(i, 20) = Wq

    cldesp = cl      ‘registramos cual es el cliente que acaba de salir
                     de la cola
    uldesp = cl      ‘registramos cual es el último cliente despachado
```

3 Parte: Trabajamos con todos los demás clientes

‘registramos un posible evento: llegada del siguiente cliente

```
    cl = cl + 1     ‘el numero de clientes se incrementa en 1
    tLL(cl) = tLL(cl - 1) + (-((Log(1 - Rnd)) / lambda)) ‘averiguamos cual es la
hora de llegada del cliente subsiguiente

    tESP(cl) = Application.Max(tSAL(atendido), tLL(cl)) - tLL(cl)’ a la hora
de llegada o a la hora de salida del que está siendo atendido le restamos la
hora de llegada del nuevo. Ese será el tiempo de espera del que llegóa
    tSER(cl) = (-(Log(1 - Rnd)) / mu)
    tSAL(cl) = tLL(cl) + tESP(cl) + tSER(cl)
    L = L + 1
    W = tSAL(cl) - tLL(cl) + W
    Wq = tESP(cl) + Wq 'tSAL(cl) - tLL(cl) - tSER(cl) + Wq


nidito:
        'se averigua si ese evento es anterior o posterior a la salida del
cliente previo
    If Application.Max((tSAL(atendido)), tLL(cl)) = tLL(cl) Then
            'en caso que la salida del previo fuera anterior se registra el
evento
            'y se convierte en previo al siguiente de la cola
            L = L - 1
            If L < 0 Then L = 0
            atendido = atendido + 1
            If atendido > cl Then atendido = cl
            GoTo nidito
    End If
    'como es válida la llegada simulada, se avanza al siguiente renglón
    tDES(cl) = tLL(cl) - tSAL(cl - 1)
    If tDES(cl) < 0 Then tDES(cl) = 0
    i = i + 1
    'se imprime el renglón i
    Cells(i, 1) = i
    Cells(i, 2) = cl
    Cells(i, 3) = tLL(cl)
    Cells(i, 4) = tESP(cl)
    Cells(i, 5) = tSER(cl)
    Cells(i, 6) = tDES(cl)
    Cells(i, 7) = tSAL(cl)
    Cells(i, 8) = L
    Cells(i, 19) = tSAL(cl) - tLL(cl)
    Cells(i, 20) = tSAL(cl) - tLL(cl) - tSER(cl)

    If L - 1 < 0 Then Cells(i, 9) = 0 Else Cells(i, 9) = L - 1
    ' si no superamos el máximo de clientes volvemos a generar una llegada
    If Range("O3").Value = "" Then

        Range("N15").Value = Range("L13").Value
```

```
        If cl < nucli Then GoTo nido
    Else
        Range("N15").Value = Range("L13").Value
        If tLL(cl) < nucli Then GoTo nido
    End If
    Range("L17").Value = cl
    Application.ScreenUpdating = True
    Stop

    ' si terminamos calculamos W y Wq
    Range("L9").Value = W / cl

    Range("L10").Value = Wq / cl
    Range("N6").Value = tSAL(cl) - tLL(atendido)
  'calculo de tiempo desocupado del servidor
a = 0
For i = 1 To nucli1
a = tDES(i) + a
Next i
Range("L18").Value = a


End Sub
```

Modelo de inventarios

Vamos a trabajar con un almacén de determinado producto que tiene un precio de venta al público ***r*** por unidad. La llegada de clientes al almacén es un proceso de Poisson cuyo parámetro es λ y la cantidad demandada por cada cliente es ***G.*** La política del encargado del almacén es mantener un inventario basado en un criterio de reposición cuando el nivel es bajo, lo que se hace según una revisión periódica.

Para simular todo el sistema, primero vamos a establecer los parámetros siguientes:

- llamaremos ***x*** al nivel del inventario en cualquier instante. Es una variable de estado del sistema.
- ***p*** será el nivel de inventario mínimo
- ***P*** será el nivel de inventario "optimo" máximo
- El costo unitario será ***c(y)*** sobre el producto solicitado y depende de una función de costo sobre las unidades pedidas, que a su vez las llamamos ***y***.
- ***y*** es una variable de estado del sistema. Si vale cero es que no hay pedidos pendientes
- ***L*** es el tiempo que demora el pedido en llegar al almacén, se abona al final de ***L***
- ***h*** es el costo de almacén
- En caso de que un cliente llegue al almacén y haga un pedido mayor a las unidades que quedan en el depósito, se les venden las que quedan y se pierde la venta del remanente
- Cuando el sistema se simula se llama ***t*** al tiempo transcurrido en la simulación

En resumen, las variables de simulación serán:

var_nivel() para guardar los instantes en los que el nivel de inventario pasa a ser cero y que se usará después para saber cuánto tiempo el almacén estuvo vacío

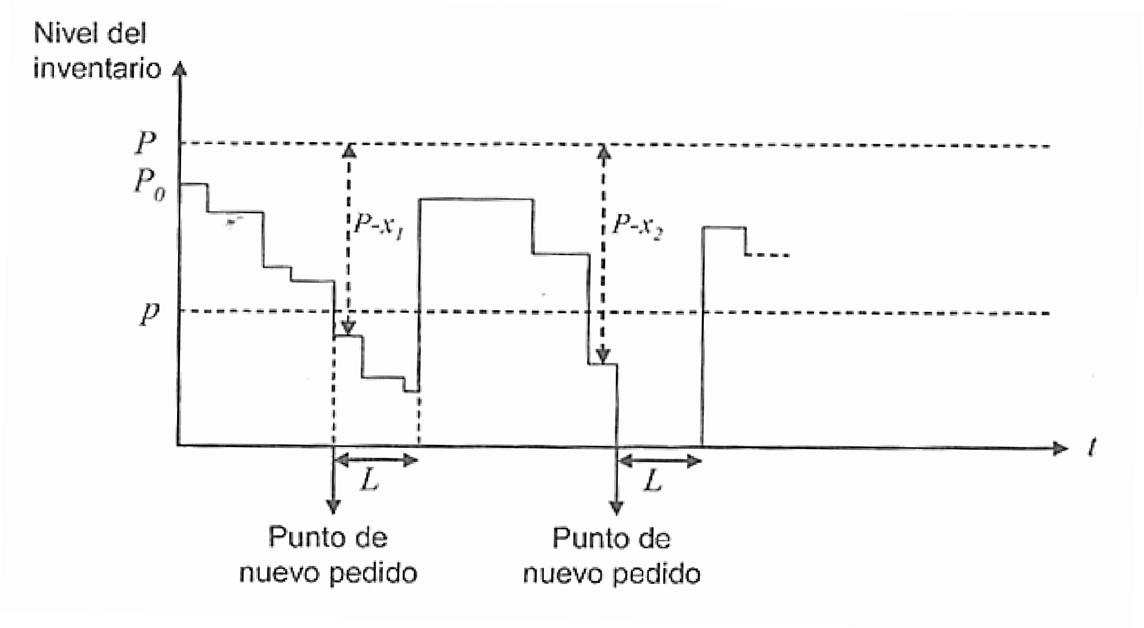
tc instante de llegada de cliente

tp instante de llegada de pedidos al proveedor

- SUCESOS
  - ***TSuc.tc*** llegada de un nuevo cliente al sistema
  - ***TSuc.tp*** llegada de un pedido
- CONTADORES:
  - ***C*** costo total por pedidos hasta el tiempo ***t***
  - ***H*** costo total por almacén hasta el tiempo ***t***
  - ***R*** beneficios por ventas hasta el tiempo ***t***
  - ***Nc+*** número de clientes servidos satisfactoriamente hasta el tiempo ***t***
  - ***Nc-*** número de clientes a los que no se ha podido satisfacer completamente hasta el instante ***t***
  - ***t0*** tiempo en el que inventario está en cero
- VARIABLES DE SALIDA:
  - ***beneficio*** beneficio esperado
  - ***%cliente_satisf*** clientes (%) cuya demanda se satisfizo completamente
  - ***%invent_0*** tiempo (%) en que el inventario estuvo en cero

CRITERIO DE FUNCIONAMIENTO:

- Si ***x*** < ***p*** Y no hay pedidos pendientes:
  Pedir y = P – x unidades



Algoritmo

El primer suceso que ocurre es la llegada de un cliente al sistema suponiendo que el nivel del inventario inicial es mayor que el mínimo (ver figura): P0 > p. Para lograr esto se llama a la rutina ***Llegada_cliente***.

Luego se entra en un bucle donde se van llamando a las rutinas que provocan los sucesos en el orden en que ocurren. Se finaliza calculando los indicadores del sistema utilizando los contadores

1. - Preparación de variables
x = P0 y = Nc- = Nc+ = C = R = H = 0 TSuc.tc = TSuc.tp = M t = tsuc = t0 = var_nivel(0) = 0
2. – Generación del número aleatorio Z ➔ Exp(λ)
Si Z > T

```
    Beneficio = %client_satisf = %invent_0 = 0
        FIN
Caso contrario: Llegada_cliente(Z)
```

3. – Sucesos siguientes

```
MIENTRAS SEA: [Tsuc.tp <> M] O [TScuc.tc <> M]
    SI TSuc.tc < TSuc.tp
            tsuc = TSuc.tc
            TSuc.tc = M
            Llegada_cliente(tsuc)
    SI TSuc.tp < TSuc.tc
            tsuc = TSuc.tp
            TSuc.tp = M
            Llegada_pedido(tsuc)
```

4. – Cálculo de las medias de comportamiento del sistema

```
beneficio = R – C – H
%invent_0 = t0 / t
%client_satisf = NC+ / (Nc+ + Nc-)
FIN
```

Subrutinas

En ***Llegada_cliente()*** lo primero que debe realizarse es actualizar los costos de almacén hasta ese momento dependiendo del nivel de inventario ***x*** desde el suceso anterior (guardado en ***t***) hasta este instante.

Luego se actualiza el tiempo de simulación, se genera la demanda del cliente que acaba de llegar y se comprueba si se puede satisfacer.

Si el inventario alcanza, se realiza la venta completa, actualizando la variable, disminuyendo el nivel de inventario en la cantidad demandada y se contabiliza como cliente satisfecho.

En el caso que no alcance el inventario, se vende el resto del depósito, se pone el inventario en 0, se cuenta como cliente insatisfecho y se guarda en var_***nivel0*** el instante para poder contar el tiempo en que el inventario estuvo en cero.

Finalmente, se haya satisfecho totalmente o no el pedido, quedan dos pasos: uno, se verifica si el nivel del inventario en ese momento en más chico que ***p***. Si es así y además no hay pedidos pendientes se solicita la cantidad ***P - x*** y se guarda en la lista de sucesos el instante en que llegará el nuevo pedido (tiempo actual más ***L***). Segundo, generar el tiempo que transcurre hasta la llegada del nuevo cliente y verificar si este acontecimiento cae o no dentro del plazo máximo de simulación ***T***.

```
Llegada_clientes()
H = H + (tsuc – t) * h * x
t = tsuc

Generar demanda →G
SI demanda <= x
        R = R + demanda * r
        x = x – demanda
        Nc+ = Nc+ + 1
SI NO
        R = R + (x*r)
```

```
        x = 0
        Nc- = Nc- + 1
        var_nivel0 = t

SI x <= p AND y = 0
        y = P – x
        TSuc.tp = t + L

Generar Y → Exp(λ)
SI t + Y < T
        TSuc.tc = t + Y
VOLVER
```

La segunda subrutina, ***Llegada_pedido()*** se refiere a la llegada de un pedido al sistema en el momento contado desde el tiempo en que fue pedido más el tiempo guía. Se deben actualizar los costos de almacén hasta ese momento, actualizar el tiempo, actualizar el pago al proveedor, dependiendo del número de unidades que llegaron ***y***, se aumenta el nivel del inventario, se pone en cero las pendientes de entrega y se calcula el tiempo en que el nivel de inventario fue cero si estaba en ese valor restando el tiempo actual y el guardado en ***var_nivel0***

```
Llegada_pedido()
H = H + (tsuc – t) * h * x
t = tsuc
C = C + c(y)
x = x + y
y = 0

SI var_nivel0 < t
        t0 = t0 + (t – var_nivel0)
        var_nivel0 = M

VOLVER
```

CAPITULO 8

Lenguajes de Simulación

A pesar de que para construir simuladores se puede utilizar cualquier lenguaje de programación con razonable capacidad matemática, disponemos de una variedad de software específico para simulación de sistemas discretos.

Cuando nos referimos a lenguajes específicos para simulación nos encontramos que ellos son comparables con C+, Pascal, etc. pero tienen capacidades y herramientas especiales que facilitan la tarea de simular.

Algunas herramientas que encontramos incluidas en ellos son:

- Mecanismo de reloj
- Métodos para secuencias los sucesos que ocurren
- Generador de números y variables aleatorias
- Facilidades para definir y maquetar las entidades y sus estados
- Recolección y análisis estadístico del uso de recursos y entidades
- Herramientas para validar el modelo

La desventaja de estos paquetes suele ser: precio, velocidad baja de ejecución y poca disponibilidad para diferentes SO.

Podemos clasificarlos en función de que estrategia utilizan para el reloj del sistema:

1.- **Por programación de sucesos**. El programador detalla las acciones asociadas a la ocurrencia de cada suceso. El lenguaje automatiza el proceso de muestreo a partir de las distribuciones, cronología, recuperación de sucesos y recopilación estadística. (GASP, SLAM, SIMAN, SIMSCRIPT)

2.- **Orientados a procesos**. Usan bloques que se pueden unir y forman una red que describe el flujo de movimiento u operaciones. Se basan en un enfoque de caja negra sobre procesos de entrada y salida. Son más fáciles de usar y simples, aunque menos flexibles. (GPSS, SIMULA, SIMNET II, SIMAN)

3.- **Combinados**. MODSIM, SIMSCRIPT II.5, SIMAN

4.- **Simuladores**, o paquetes de sencillo manejo aplicado a situaciones específicas: LANNET II.5 desarrollado para redes locales a partir de SIMSCRIPT; SIMFACTORY y AUTO MOD para sistemas de fabricación.

Se han desarrollado simuladores de propósito general. Dos de ellos, muy utilizados, permiten incluir componentes discretos y continuos simultáneamente, EXTENDED SUITE y ARENA (http://www.arenasimulation.com). Permiten describir gráficamente el modelo y facilitan la implementación mediante la importación de librerías y bloques, conectados para ver el flujo de entidades y uso de caja de diálogo, además de capacidades estadísticas, gráficas y animaciones.

Por otra parte es bueno recordar que no es imprescindible el uso de software específico para construir simuladores (en este libro para mostrar los ejemplos sencillos que utilizamos no los hemos empleado). Así cualquier lenguaje desde FORTRAM en adelante se puede usar. Simplemente es más complejo usarlos. Incluso, si se es consciente de las limitaciones inherentes, con una hoja Excel se podrán solucionar las mayoría de los problemas que abordemos.

ANEXOS

Anexo I

Fragmento del capítulo “CICLOS” de la novela de Neal Stephenson “1. El código Enigma” que forma parte de la trilogía “CRIPTONOMICON”. Publicado por NOVA en 2005 de la edición original de 1999[75].

(...)

—Perdóname. —Alan frena de pronto y baja de la bicicleta. Levanta la rueda trasera del pavimento, la hace girar con la mano libre, luego se agacha y tira de la cadena. Contempla el mecanismo con toda atención, interrumpida por algunos estornudos.

La cadena de la bicicleta de Turing tiene un eslabón débil. La rueda trasera tiene un radio doblado. Cuando el eslabón y el radio entren en contacto, la cadena se romperá y caerá sobre la carretera. No sucede a cada vuelta; en caso contrario la bicicleta sería completamente inútil. Sólo sucede cuando el eslabón y la rueda se encuentran en cierta posición relativa.

Basándose en suposiciones razonables respecto a la velocidad que el doctor Turing puede mantener, un ciclista enérgico (digamos 25 km/h) y el radio de la rueda trasera de la bicicleta (un tercio de metro), si el eslabón débil golpease contra el radio doblado a cada vuelta, la cadena se caería cada tercio de segundo.

De hecho, la cadena no cae a menos que el radio doblado y el eslabón débil coincidan. Ahora, supongamos que describimos la posición de la rueda trasera usando la θ habitual. Por simplificar, digamos que cuando la rueda empieza en la posición donde el radio doblado es capaz de golpear el eslabón débil (aunque sólo si el eslabón débil está ahí para ser golpeado) entonces $\theta = 0$. Si usas grados como unidades, durante una revolución completa de la rueda θ llegará hasta los 359 grados antes de volver a 0, en cuyo punto el radio doblado volverá a estar en posición de golpear la cadena. Y ahora supongamos que describes la posición de la cadena con la variable C de la siguiente forma muy simple: asignas un número a cada eslabón de la cadena. El eslabón débil tiene el número 0, el siguiente el 1, y a continuación, hasta l=1 donde l es el número total de eslabones de la cadena. Una vez más, para simplificar, digamos que cuando la cadena se encuentra en la posición donde el eslabón débil es capaz de golpear el radio doblado (aunque sólo si el radio doblado está ahí para ser golpeado) entonces C=0.

Entonces, para intentar descubrir cuándo caerá la cadena de la bicicleta del doctor Turing, todo lo que precisamos saber sobre la bicicleta está contenido en los valores de θ y C. Ese par de números define el estado de la bicicleta. La bicicleta tiene muchos estados posibles y puede haber muchos valores diferentes de (θ, C) pero sólo uno de esos estados, el (0, 0), es el que hará que la bicicleta caiga.

---

[75] Las notas de pie de página en este anexo, así como en el resto del volumen, son notas del autor de este texto. La nota con asterisco (*) casi al final del anexo, es original de la novela.

Supongamos que empezamos en ese estado, es decir, con ($\theta = 0$, C = 0), pero la cadena no ha caído porque el doctor Turing (conociendo muy bien el estado de su bicicleta en un momento dado) se ha detenido en medio de la carretera (casi provocando una colisión con su amigo y colega Lawrence Pritchard Waterhouse, porque la máscara antigás le bloquea la visión periférica). El doctor Turing ha tirado de la cadena hacia un lado mientras la adelanta ligeramente, evitando así que golpee el radio doblado. Ahora vuelve a subirse a la bicicleta y sigue pedaleando. La circunferencia de la rueda trasera es de unos dos metros, así que cuando se ha trasladado unos dos metros sobre la carretera, la rueda ha dado una vuelta completa y ha alcanzado de nuevo la posición $\theta = 0$, siendo ésa la posición, recuerden, en la que el radio doblado está en posición para golpear el eslabón débil.

¿Qué hay de la cadena? Su posición, definida por C, comienza en 0 y llega a 1 cuando el siguiente eslabón se traslada a la posición fatal, luego 2 y así sucesivamente. La cadena debe moverse en sincronía con los dientes del engranaje en el centro de la rueda trasera, y ese engranaje tiene *n* dientes, por lo que después de una revolución completa de la rueda trasera, de nuevo $\theta =0$, C=n. Después de una segunda vuelta completa de la rueda trasera, de nuevo $\theta =0$ pero ahora C=2n. En la siguiente C=3n y así sucesivamente. Pero hay que recordar que la cadena no es infinita sino un bucle con sólo *l* posiciones; en C=*l* vuelve a C=0 y repite el ciclo. Por lo que al calcular el valor de C es necesario realizar aritmética modular, es decir, si la cadena tiene un centenar de eslabones (*l*=100) y el número total de eslabones que han sido desplazados es 135, entonces el valor de C no es 135 sino 35. Cuando tienes un número superior o igual a *l*, restas repetidamente *l* hasta que obtienes un número menor que *l*. Los matemáticos escriben esa operación como mod *l*. Por tanto, los valores sucesivos de C, cada vez que la rueda trasera da una vuelta hasta $\theta =0$, son:

$$C_i = n \bmod l,\ 2n \bmod l,\ 3n \bmod l, ...,\ in \bmod l$$

donde i = (1, 2, 3,... ∞)

más o menos, dependiendo de cuánto tiempo quiera Turing seguir pedaleando en su bicicleta. Después de un rato, a Waterhouse ya le parece infinitamente largo.

La cadena de la bicicleta de Turing se caerá cuando la bicicleta alcance el estado *{* $\theta =0$, C=0) y visto lo escrito anteriormente, eso sucederá cuando i (que no es más que un contador que indica cuantas vueltas ha dado la rueda trasera) alcanza algún valor hipotético tal que in mod /=0, o, para explicarlo claramente, sucederá si hay algún múltiplo de n (como, oh[76], 2n, 3n, 395n o 109.948.368.443n) que resulte también ser un múltiplo de *l*. En realidad, puede haber muchos de esos llamados múltiplos comunes, pero desde un punto de vista práctico el único que importa es el primero —el mínimo común múltiplo, o MCM— porque ése será el que se alcance primero y el que hará caer la cadena.

Si, digamos, el engranaje tiene veinte dientes (n=20) y la cadena tiene cien eslabones (*1*=100), entonces después de un giro de la rueda tenemos C=20, después de dos C=40, luego 60, luego 80 y finalmente 100.

Pero como tomamos el módulo aritmético, ese valor debe cambiarse por 0. Por tanto, después de cinco vueltas de la rueda trasera, hemos llegado al estado ($\theta =0$, C=0) y la cadena de Turing caerá. Cinco revoluciones de la rueda trasera sólo le harán avanzar diez metros, y por tanto, con esos valores de *l* y n la bicicleta es prácticamente inútil. Claro está, todo eso es cierto si Turing es tan estúpido como para empezar a pedalear con la bicicleta en el estado-que-hace-caer-la-cadena. Si, en el momento de empezar a pedalear, se encuentra en su lugar en el estado ($\theta =0$,

[76] El autor de Optimiza buscó dos ediciones en español del libro y en ambas encontró la expresión “oh”. No sabemos qué significa.

C=l), entonces los valores subsiguientes serán C=21, 41, 61, 81, 1, 21... y así sucesivamente; la cadena nunca se caerá. Pero se trata de un caso degenerado, donde «degenerado» tiene el significado matemático de «enojosamente aburrido». En teoría, siempre que Turing ponga su bicicleta en el estado correcto antes de aparcarla fuera del edificio, nadie podrá robársela; la cadena se caerá apenas después de haber avanzado diez metros.

Pero si la cadena de Turing tiene ciento y un eslabones (*l*=101) y después de cinco revoluciones tenemos C=100, y después de seis tenemos C=19, luego

C= 39, 59, 79, 99,18, 38, 58, 78, 98,17, 37, 57, *77,* 97,16, 36, 56, 76, 96,15, 35, 55, 75, 95,14, 34, 54, 74, 94,13, 33, 53, 73, 93,12, 32, 52, 72, 92,11, 31, 51, 71, 91,10, 30, 50, 70, 90, 9,29, 49, 69, 89, 8,28,48, 68, 88, 7,27,47, 67, 87, 6, 26, 46, 66, 86, 5,25, 45, 65, 85,4,24,44, 64, 84, 3,23, 43, 63, 83,2,22,42, 62, 82,1, 21,41, 61, 81, 0

Así que no será hasta la revolución 101 de la rueda trasera que la bicicleta vuelva al estado *(8=0,* C=0) cuando cae la cadena. Durante ese centenar más uno de vueltas, la bicicleta de Turing ha recorrido un quinto de kilómetro, que no está mal. Así que la bicicleta se puede usar.

Sin embargo, al contrario que en el caso degenerado, no es posible situar la bicicleta en un estado tal que la cadena nunca caiga. Tal cosa puede demostrarse repasando la lista anterior de valores de C y comprobando que todo posible valor de C, todo posible valor entre 0 y 100, está en la lista. Eso significa que no importa en qué valor esté C cuando Turing empieza a pedalear, tarde o temprano llegará al C=0 fatal y la cadena caerá. Por tanto, Turing puede dejar la bicicleta en cualquier sitio con la confianza de que, si la roban, no recorrerá más de un quinto de kilómetro sin que la cadena se caiga.

La diferencia entre el caso degenerado y el caso no degenerado está relacionada con las propiedades de los números implicados. La combinación de (n=20, *l*=100) tiene propiedades radicalmente diferentes con respecto a (n=20, *l*=101). La diferencia principal es que 20 y 101 son «primos relativos», lo que significa que no tienen factores comunes.

Eso significa que su MCM es un número grande —de hecho, es igual a *1* x n = 20 x 1001 = 2020. Mientras que el MCM de 20 y 100 es sólo 100. La bicicleta *1*=101 tiene un periodo largo —pasa por muchos estados diferentes antes de volver al principio—, mientras que la bicicleta *l*=100 tiene un periodo de unos pocos estados.

Supongamos que la bicicleta de Turing fuese una máquina de cifrado que actuase por sustitución alfabética, lo que es lo mismo que decir que reemplazaría cada una de las 26 letras del alfabeto por alguna otra letra. Una A en el texto original se podría convertir en una T en el texto cifrado, B podría transformarse en F, C podría convertirse en M, y así hasta llegar a la Z. Por sí mismo, sería un código absurdamente fácil de romper; cosa de niños. Pero supongamos que el esquema de sustitución cambiase de una letra a la siguiente. Es decir, supongamos que la primera letra del texto original fuese cifrada usando cierto alfabeto de sustitución, la segunda letra del texto original fuese cifrada usando un alfabeto de sustitución completamente diferente, y la tercera con otro diferente, y así sucesivamente. Eso se conoce como un cifrado polialfabético.

Supongamos que la bicicleta de Turing fuese capaz de generar un alfabeto diferente para cada uno de sus diferentes estados. Por tanto, el estado *(θ =0,* C=0) correspondería, digamos, a este alfabeto de sustitución:

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
Q G U W B I Y T F K V N D O H E P X L Z R C A S J M

pero el estado (θ =180, C=15) correspondería a este otro, diferente:

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
B O R I X V G Y P F J M T C Q N H A Z U K L D S E W

Dos letras no serían cifradas usando el mismo alfabeto de sustitución, es decir, hasta que la bicicleta no llegase de nuevo al estado inicial (θ *=0,* C=0) y empezase a repetir el ciclo. Eso significa que se trata de un sistema polialfabético periódico. Ahora bien, si la máquina tuviese un periodo corto, se repetiría con frecuencia, y por tanto sería útil, como sistema de cifrado, sólo contra los niños. Cuanto más largo sea el periodo (cuanto mayor sea su primitividad relativa) con menos frecuencia vuelve al mismo alfabeto de sustitución, y más seguro es.

La Enigma de tres rotores es ese tipo de sistema (es decir, polialfabético periódico). Sus rotores, como el sistema de la bicicleta de Turing, contienen ciclos dentro de ciclos. Su periodo es 17.576, lo que significa que el alfabeto de sustitución que cifra la primera letra del mensaje no volverá a emplearse hasta que se llegue a la letra 17.577. Pero con Tiburón[77], los alemanes han añadido un cuarto rotor, elevando el periodo hasta 456.976. Los rotores se sitúan en una posición inicial diferente elegida al azar al comienzo de cada mensaje. Como los mensajes alemanes nunca llegan a los 450.000 caracteres, la Enigma nunca usa dos veces el mismo alfabeto de sustitución en un mismo mensaje, razón por la que los alemanes la consideran un buen sistema.

Un grupo de aviones de transporte pasan por encima de sus cabezas, muy probablemente en dirección al aeródromo de Bedford. Los aviones producen un zumbido diatónico curiosamente musical, como una gaita tocando dos tonos simultáneamente. Eso recuerda a Lawrence otro fenómeno más relacionado con la rueda de la bicicleta y la máquina Enigma.

—¿Sabes por qué los aviones suenan así? —pregunta.
—No, ahora que lo pienso. —Turing vuelve a quitarse la máscara antigás. Tiene la boca algo abierta y mueve los ojos de un lado a otro. Lawrence lo ha pillado por sorpresa.
—Me di cuenta en Pearl. Los motores de los aviones son rotatorios[78] —dice Lawrence—. Por tanto, deben tener un número impar de cilindros.
—¿Por tanto?
—Si tuviesen un número par, los cilindros estarían directamente en oposición, a ciento ochenta grados, y no funcionarían mecánicamente.
—¿Porqué no?
—Lo he olvidado. Pero no funcionaría.

Alan arquea las cejas. Claramente no está convencido.
—Es algo relativo a los cigüeñales —aventura Waterhouse, poniéndose algo a la defensiva.
—No estoy seguro de estar de acuerdo —dice Alan.
—Vamos a estipularlo... considéralo una condición de contorno —dice Waterhouse. Pero sospecha que Alan ya está concentrado, diseñando mentalmente un motor rotatorio de avión con un número par de cilindros.
—En todo caso, si los miras, todos tienen un número impar de cilindros —sigue diciendo Lawrence—. Por lo que el sonido de la expulsión se combina con el sonido de la hélice para producir ese sonido de dos tonos.

Alan vuelve a subir a la bicicleta y pedalea por el bosque sin hablar.

---

[77] Nombre dado por la Marina de Alemania durante la Segunda Guerra Mundial a su modificación de la máquina “Enigma” (N del autor de Optimiza)

[78] En la época en que se ambientó esta parte de la novela, ya no quedaban muchos aviones con motor rotativo.

En realidad, no han estado hablando sino más bien mencionando ciertas ideas y dejando que el otro desarrolle las implicaciones. Es una forma extremadamente eficaz de comunicarse; elimina los elementos redundantes de los que se quejaba Alan en el caso de FDR[79] y Churchill.

Waterhouse está pensando en ciclos dentro de ciclos. Ya ha decidido que la sociedad humana es uno de esos supuestos de ciclos dentro de ciclos* y ahora intenta decidir si es como la bicicleta de Turing (funciona bien durante un rato, y de pronto la cadena se cae; de ahí la ocasional guerra mundial) o como la máquina Enigma (se mueve incomprensiblemente durante un tiempo, y luego de pronto los rotores se alinean como en un tragaperras y todo queda claro en una especie de epifanía global o, si se prefiere, Apocalipsis) o como un motor rotatorio de avión (gira, gira y gira; no sucede nada especial, simplemente produce mucho ruido).

* No tiene datos reales para sostenerlo, pero le parece una idea genial.

[79] FDR son las iniciales usuales en esa época para referirse a Franklin Delano Roosevelt, presidente de EE.UU. durante la mayor parte de la Segunda Guerra Mundial (N del autor de Optimiza)

Anexo II

La máquina "Enigma"

Si bien la fama que tiene esta máquina se debe a su uso por parte del ejército y la marina alemanas durante la Segunda Guerra Mundial, en realidad fue desarrollada por empresas privadas para fines comerciales en la década de 1920.

El ejército alemán la comenzó a utilizar poco antes del inicio de la guerra con algunas mejoras y, posteriormente, la marina le agregó algunas modificaciones importantes en materia de seguridad.

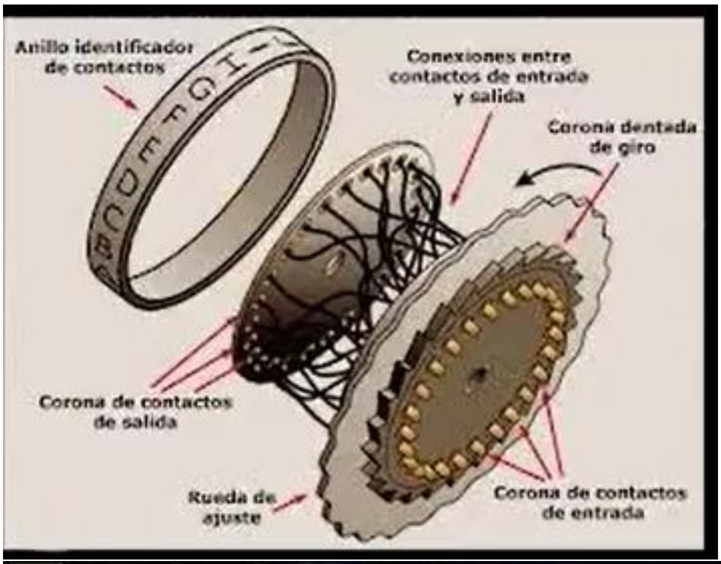
De todas maneras, y hasta el fin de la guerra, los alemanes creían disponer de un equipo de cifrado similar en prestaciones a la "libreta de un solo uso" pero automatizado y seguro, sin sospechar que los mensajes que usaban ese canal estaban siendo descifrados por la inteligencia británica a partir de poco tiempo de iniciada la guerra. Esto genera dos historias, una sobre cómo funciona la máquina y la otra como fue roto el código.

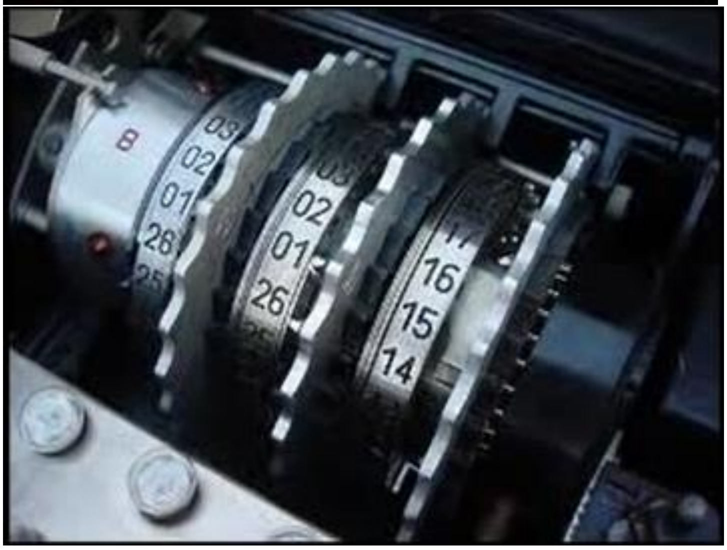
Principio de funcionamiento

La máquina Enigma, como fue utilizada en la guerra, consistía en una caja de dimensiones similares a cualquiera que contuviera una máquina de escribir de aquella época. Efectivamente el aspecto no es muy diferente, aunque tiene dos "teclados". Uno convencional, con teclas reales, y el otro similar, pero en lugar de teclas eran las mismas letras con una lamparita atrás. Encima hay lugar para tres o cuatro rotores moleteados (tres en la versión ejército, cuatro en la versión marina)

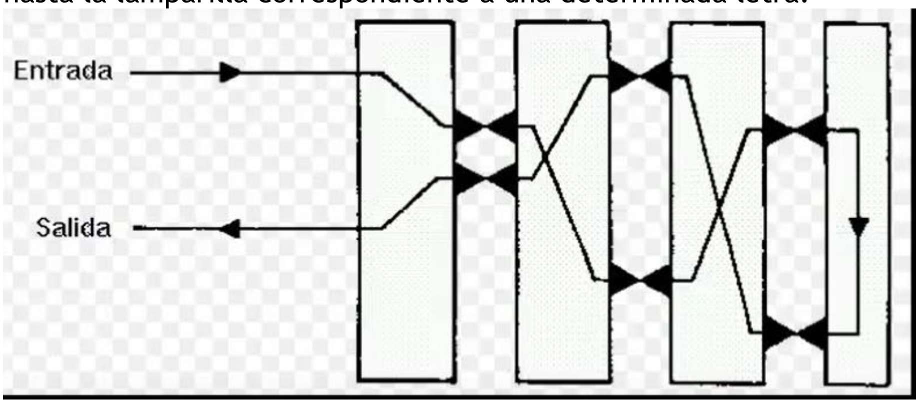
Los discos, como muestra la figura eran conexiones entre contactos de una cara a la otra, y según el tipo de máquina, había tres o cuatro combinaciones para usar de los cinco disponibles, llamadas "juego". En concreto, cada máquina, supongamos del ejército, se entregaba con cinco discos. Cada letra, en todos los discos corresponde a un número. En la máquina se instalaban tres de esos cinco. Eso da sesenta juegos o variaciones de combinación de discos.

$$V_{m,n} = m \times m - 1 \times m - 2 \times \ldots \times m - n$$
$$V_{5,3} = 5 \times 4 \times 3 = 60$$





Una vez instalados los tres discos la máquina estaba configurada de manera tal que si el operador pulsaba una tecla, cerraba un circuito eléctrico a través de los discos y hasta la lamparilla correspondiente a una determinada letra.



Digamos que la máquina está con el disco 1 en la posición 1 (la "A" arriba) que equivale que la entrada sea por el contacto 1[80], salida en 10 (entra en contacto 1 y sale en contacto 10), el disco 2 también en la posición 1, entra por el 10 (que está

[80] Si está con la "B" arriba la entrada es por el contacto 2 y así sucesivamente

enfrente del 10 del disco 1) y sale por el 17 y en el disco 3, entra en 17, que está enfrente del anterior y sale por el 5. Ahora ocurre algo ingenioso. Una vez que se estableció este camino, el cableado obliga a hacer un camino “de retorno” a través de los tres discos, llamémosle “rebote”.

Así, mediante ese rebote entra al mismo disco pero “por atrás” en otro número y pasa al disco 2 de nuevo y luego al otro, cambiando cada vez de contacto como a la ida. Supongamos que sale por el 21 (que corresponde a la letra “P”). Así sea cual fuere la letra que tocó el operador se enciende la “P”. Este camino de vuelta tiene un objetivo: si el operador pulsó la tecla, por ejemplo “T”, y la pulsa miles de veces, jamás va a aparecer la letra “T”. Es físicamente imposible.

Pero esto no es todo. Como ya dijimos, el análisis anterior solo sirve si hay una “posición” inicial para cada rotor. Lo que dijimos en el ejemplo es que si ponemos los tres rotores en la posición “1” aparece la “P” si apretamos una tecla dada, por ejemplo la “J”.

Entonces hay tres fuentes de aleatoriedad:
1) La posición de los discos tiene 26 posibilidades cada uno,
2) la combinación interna de cables que es diferente en cada rotor
3) los discos elegidos que son 3 de 5 posibles.

Veamos, las combinaciones de cableado posible en cada disco es 26, así que en tres discos será 26 x 26 x 26 = 17.576. En términos criptoanalíticos, ese es el número de alfabetos de sustitución disponibles. Pero, además, no era lo mismo poner los tres rotores en un orden dado que en otro. Por ejemplo, usar los rotores 1-3-5 no daba la misma combinación que usar los rotores 3-5-1 o 3-1-5 o 5-3-1 o 5-1-3.

Si a ese valor lo multiplicamos entonces por 60 variaciones de combinación de discos tenemos un total de 1.054.560 posibilidades.

Dicho de otra manera, si el operador pulsó, digamos, la tecla “J” y se encendió la “P”, la probabilidad de predecir matemáticamente cual se iba a encender sería de uno en un millón (1×10^{-6})

Pero, todavía quedan dos aspectos más:

El primero, las máquinas militares tenían un cableado cruzado adicional muy fácil de cambiar. En la parte frontal hay un cuadro de clavijas que permite puentear letras entre sí, digamos, la “P” con la “S”. De esta manera, en nuestro ejemplo, en realidad no se encendió la “P”, sino que se encendió la “S”. Este truco que permite puentear 20 letras, nos da 10 pares de letras con 2 combinaciones, o sea 2^{10} posibilidades y un número total de combinaciones que es

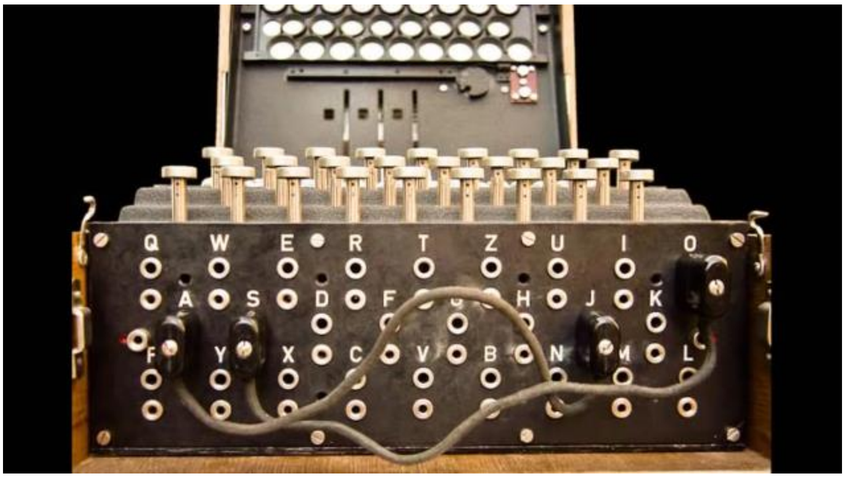
$$\frac{26!}{6! \times 10! \times 2^{10}} = 151 \times 10^{12}$$

Estas son las combinaciones posibles. Esto se debe multiplicar por el numero anterior y daría algo así como $1{,}51 \times 10^{20}$, aproximadamente.

Pero si bien es una cifra enorme y aterradora, no es muy difícil de vulnerar por un buen equipo criptoanalista. El secreto mayor de la máquina está en que, al apretar la cada tecla hay un sistema de criquet que hace que el primer disco se mueva una posición, y después de 26 teclas, el segundo avanza una posición también y al cabo de 676 teclas el tercero avanza también. Esto quiere decir que, por ejemplo, la letra que originó esa “P”, si es pulsada repetidamente va a originar una letra diferente cada vez.

Así, las salidas solamente volverán a repetirse después de 26^3 letras pulsadas. Como 26^3 es 17.576 para el ejército y 26^4 es 456.976 para la marina, un criptógrafo debería esperar una repetición solamente en un mensaje de 17.577 o de 456.977 letras de largo, mensaje que los alemanes jamás mandarían, ya que usualmente eran muchísimo más breves[81]. Como ejemplo, la letra que en nuestro ejemplo originó la “S”, solo va a volver a originarla después de pulsarla 17.576 veces.

A la medianoche de cada día se cambiaba la configuración de todas las máquinas. Todos los receptores tenían un libro de códigos con una configuración inicial para cada máquina para cada mes y día. Esa configuración indicaba que discos se iban a usar ese día, en qué orden y con que cableado. Era obvio que ese libro podía ser capturado, y por lo tanto no era seguro, así que el emisor del mensaje, supongamos en Berlín a las 5 de la mañana, emitía una combinación de código a usar ese día, pero en el código que está en el libro del día. Por ejemplo, hoy corresponde usar la combinación RST, que indica como se colocan los rotores y otras cosas. En el libro de códigos figura el día de hoy como TVB. El operador lanza el mensaje codificado con TVB, así: WQKGHY.



En el otro extremo se recibe WQKGHY, se usa TVB y se obtiene RSTRST, porque los alemanes al principio lo mandaban repetido. El operador sabe que debe usar entonces la combinación RST y de ahí en adelante descifra el resto del mensaje con la nueva configuración. Curiosamente esa repetición de letras ayudó a resolver el código.

Este es un ejemplo de libro de código real del día 7 de julio de 1941:
Rotores II, IV y V
Posición inicial de cada rotor 02, 21 y 12 (significa letra B U y L)

[81] En realidad, y por norma, los mensajes se emitían en longitudes máximas de 256 caracteres cada uno. A esto se lo llamaba “Parte”. Un mensaje largo podía tener varias partes. Siempre comenzaba con texto sin codificar (Fecha, destinatario y remitente, por ejemplo) y el resto en grupos de 4 letras. Un mensaje nunca tenía espacios y había códigos especiales de grupos de letras para indicar algunas puntuaciones o paréntesis. el primer y segundo grupo de 4 letras llevaba en código el código del resto del mensaje.

Puentes entre letras: AV BS CG DL FU HZ IN KM OW RX

El receptor configuraba la máquina de esta manera. Recibía el primer mensaje y leía el código de la configuración real que debía usar para seguir leyendo el resto del mensaje.

En síntesis. El mensaje tenía una primera parte sin cifrar con la fecha y las partes que tenía el mensaje (cada mensaje solo podía tener 256 caracteres, por eso se mandaba en partes de ese tamaño) luego venían tres letras codificadas repetidas (6 caracteres) que daban la posición inicial de los tres rotores de ese día. Poniendo los rotores en esas tres letras se leía lo que restaba del mensaje.

Los aliados descifran Enigma

A pesar de todo esto, los aliados lograron descifrar Enigma dando lugar a una verdadera guerra oculta sin armas de fuego, solamente matemáticas y estadísticas dentro de ambientes de espionaje.

La primera máquina fue capturada por ejército polaco, en cuanto comenzó la guerra y sus matemáticos lograron averiguar todo lo que era capaz de hacer y determinar todos sus parámetros operativos. Los franceses, por otra parte también avanzaron en ello. Esa máquina, cuando Polonia y luego Francia cayeron, fue a un laboratorio en la campiña inglesa (Bletchey Park) junto con todo el importante trabajo de los matemáticos polacos y franceses. Para ese entonces los alemanes habían dejado de usar las tres letras repetidas, conscientes que eran una violación de seguridad, y habían agregado los rotores adicionales hasta 5.

Pero un principio fundamental del criptoanálisis es que conocer el código y los parámetros no es lo mismo que descifrar un mensaje.

Los ingleses, junto a matemáticos polacos y norteamericanos, liderados por Alan Turing y aportes de Friedman se abocaron a tratar de entender como se transmitía el código de cada día (posición inicial de los discos y que discos usaban) y que combinación concreta correspondía al día. Lo primero que se hizo es analizar mensajes por medios tradicionales (repeticiones, patrones de frecuencia de letras, omisiones, etc.)

Luego se comenzó una búsqueda por fuerza bruta, mediante una máquina (bautizada “bomba” por los polacos) que se conserva actualmente en el lugar donde se trabajaba, que hoy es un museo. Esa máquina fue perfeccionada por Alan Turing y era una computadora analógica.

Como el código duraba solo hasta las 12 de la noche de cada día, la máquina debía descifrarlo antes de esa hora, cosa que no se logró nunca y, de haberlo hecho al día siguiente ya no tendría valor y debía volverse al principio.

Finalmente se descubrió que algunos mensajes comenzaban siempre con algún texto idéntico entre las varias decenas de se emitían todos los días y que se referían al código del día. Esto ocurría por operadores que usaban siempre las mismas letras, como por ejemplo sus iniciales para poder sincronizar los códigos a usar. Por análisis criptográfico ese código se rompió con esos mensajes y cuando un operador alemán que probaba el sistema transmitió la misma letra muchas veces (T), el análisis descubrió que en un texto moderadamente largo no aparecía nunca la T, y sabiendo que la característica principal de la Enigma es que jamás podía aparecer la tecla pulsada (por eso del camino inverso del cableado entre rotores), pudieron deducir que combinaciones de letras podía asumir la T. Fue el principio del fin.

A partir de ese momento, los aliados con la ayuda de “bomba” dispusieron de la lectura de gran cantidad de los mensajes enigma. La parte estadística aparece cuando tuvieron que reaccionar a los mensajes de manera tal que cada acción estuviera dentro de las probabilidades de ser tomadas sin conocer los mensajes. Por ejemplo, sabían que los alemanes habían decidido atacar el punto x tal día a tal hora. Se preguntaban, ¿Que probabilidades hay de que los defensores del punto x reforzaran la posición sin saber nada del ataque?, o, ¿se ha capturado últimamente a algún soldado alemán que pudiera tener esa información? O se montaba una red falsa de espionaje y se hacía correr el rumor de que un espía de esa red había avisado que iban a atacar x tal día. En oportunidades, grupos de “partisanos” en territorio ocupado emitían por radio en claro esa información minutos antes de la Hora H del Día D del ataque a x. Se defendía la posición solo si era probable defenderla sin levantar sospechas de violación de Enigma.

De todas maneras, la guerra terminó en 1945 y recién en 1960 se supo de que los aliados habían roto Enigma y personajes tan famosos como Alan Turing fallecieron sin que nunca nadie le reconociera su aporte a la tarea de ganar la guerra. El secreto total sobre el equipo participante en Bletchey Park y los pormenores se conocieron recién en la década de 1990 a 2000. Setenta años después de ocurrido. La máquina, inclusive, siguió utilizándose hasta mediados del siglo XX.

Recomendamos ver la película “*The imitation game*” (Código Enigma o El Juego de la imitación, en español) que, si bien no es directamente sobre Enigma, trata aspectos importantes de esa época y de la odisea matemática realizada para romper el código, así como el trágico destino de uno de los matemáticos más brillantes de la historia, a quien le debemos la computación digital y los principios de inteligencia artificial, y que fue destruido por una sociedad profundamente hipócrita y homofóbica.