pulso optimiza

el latido de la optimización de procesos

Año 1 / Número 1 — Noviembre/Diciembre 2025

www.pulso.optimiza.org

EN ESTE NÚMERO:

Detrás de las puertas de Monty Hall
El concepto de entropía en simulación y comunicación
Cadenas de Markov: el lenguaje oculto detrás de muchos procesos
Desarrollo de aplicaciones de software para las asignaturas

Pulso de Optimiza

Revista de divulgación **publicada por el equipo docente** a cargo de las asignaturas Ingeniería de Procesos (Carrera de Ingeniería en Alimentos), Investigación Operativa (Carrera de Ingeniería Industrial) y Modelos, Simulación y Teoría de la Decisión (Licenciatura en Sistemas de Información – discontinuada). Departamento de Tecnología de la Universidad Nacional de Luján

Los contenidos de los artículos firmados reflejan la opinión de sus autores. Esta publicación no es una publicación oficial del Departamento de Tecnología ni de la Universidad Nacional de Luján.

Por todo lo no firmado: equipo docente de las asignaturas mencionadas arriba, Universidad Nacional de Luján. https://optimiza.org/1779-2/

https://www.pulso.optimiza.org

Año 1 Número 1 Noviembre / Diciembre 2025

Derechos reservados © 2025

Contenido del Número 1

pulso optimiza
el latido de la optimización de procesos

Detrás de las puertas de Monty Hall: intuición, probabilidad y una lectura bayesiana de la paradoja 4

El concepto de entropía en simulación y teoría de la comunicación 11

Cadenas de Márkov: el lenguaje probabilístico oculto detrás de muchos procesos. 22

Desarrollo de aplicaciones destinadas a los estudiantes de las asignaturas en reemplazo de software obsoleto 28

Desafíos 38

AÑO 1 – NÚMERO 1 – NOVIEMBRE/DICIEMBRE 2025

https://pulso.optimiza.org

Detrás de las puertas de Monty Hall: intuición, probabilidad y una lectura bayesiana de la paradoja

Gustavo Chijani

Ing. en Alimentos. Profesor responsable de Investigación Operativa y de Ingeniería de Procesos, Universidad Nacional de Luján, Depto. de Tecnología. Noviembre 2025

Resumen

El llamado *problema de Monty Hall* se ha convertido en uno de los ejemplos paradigmáticos donde la intuición humana entra en conflicto con el razonamiento probabilístico. A partir de un juego televisivo de la década de 1960, la formulación matemática del problema muestra que una estrategia de decisión simple —cambiar de puerta cuando el presentador revela información parcial— permite duplicar la probabilidad de éxito. En este trabajo se presenta una exposición estructurada del problema de Monty Hall: su origen histórico, una formulación rigurosa, la solución clásica por enumeración de casos y un desarrollo detallado de la solución mediante el teorema de Bayes. Se analizan además variantes del problema que muestran la sensibilidad de las probabilidades a las hipótesis sobre el comportamiento del presentador, así como evidencias empíricas sobre los sesgos cognitivos involucrados. Finalmente, se discuten analogías con situaciones reales de toma de decisiones bajo incertidumbre, típicas de la estadística aplicada, la investigación operativa y la gestión de riesgos.

1. Introducción

El problema de Monty Hall puede describirse en pocas líneas:

Un concursante se enfrenta a tres puertas cerradas. Detrás de una de ellas hay un premio (por ejemplo, un automóvil) y detrás de las otras dos hay premios no deseados (habitualmente caricaturizados como cabras). El concursante elige una puerta, pero no se abre todavía. El presentador, que conoce la ubicación del premio, abre entonces una de las otras dos puertas, revelando siempre una cabra. A continuación, ofrece al concursante la posibilidad de **cambiar** su elección inicial y pasar a la puerta restante.

La pregunta es: ¿le conviene al concursante cambiar de puerta, quedarse con su elección original, o es indiferente?

La intuición de muchas personas (incluyendo profesionales con formación matemática) es que, una vez abierta una puerta y descartada una cabra, quedan dos puertas y “la probabilidad es 50–50”. Sin embargo, el análisis formal muestra que, bajo ciertas hipótesis bien definidas, la probabilidad de ganar el premio si el concursante **cambia** de puerta es (2/3), mientras que si se queda con su elección original es solo (1/3).

Este aparente conflicto entre intuición y cálculo ha hecho del problema de Monty Hall un clásico de la divulgación y de la enseñanza de la probabilidad. Más allá de la anécdota, el problema constituye un laboratorio conceptual para discutir **probabilidad condicional**, **actualización bayesiana de creencias** y **sesgos cognitivos** en la toma de decisiones.

2. Origen histórico del problema

El nombre “Monty Hall” proviene del conductor del programa de televisión estadounidense *Let’s Make a Deal*, emitido inicialmente en la década de 1960. En el programa, Monty Hall ofrecía a los participantes elegir entre distintas puertas o cajas, algunas con premios valiosos y otras con premios de escaso valor. La mecánica real del programa era más rica y flexible que la versión académica: a veces se ofrecían canjes intermedios, sumas de dinero, o se modificaban las reglas según la dinámica del show.

La versión matemática del problema surge como una **idealización** del juego, fijando un protocolo simple y repetible. Bajo esta idealización, el problema fue formulado y analizado en la literatura estadística en la década de 1970, destacándose las cartas de Steve Selvin en *The American Statistician* (1975).

Antes de Monty Hall, ya existían paradojas probabilísticas emparentadas, como el problema de los tres prisioneros y ciertas variantes de las paradojas de Bertrand, donde la información parcial modifica de manera contraintuitiva las probabilidades condicionales.

La gran difusión popular del problema se produjo en 1990, cuando Marilyn vos Savant, (la persona con más IQ del mundo su IQ es 228) columnista de la revista *Parade*, publicó la respuesta correcta (recomendando cambiar de puerta) y recibió miles de cartas —incluyendo de académicos— que cuestionaban su razonamiento. La posterior confirmación matemática y empírica de su solución convirtió a Monty Hall en un caso emblemático de choque entre intuición y probabilidad.

3. Formulación rigurosa del problema

Muchas confusiones en torno al problema surgen de **no explicitar las hipótesis** sobre el comportamiento del presentador. En esta sección fijamos el modelo estándar que se adoptará en el resto del artículo.

Consideremos el siguiente protocolo:

1. Hay **tres puertas**: 1, 2 y 3.
2. Detrás de una de ellas se coloca un premio (auto), y detrás de las otras dos, cabras. La posición del premio se determina al azar, con probabilidad uniforme: $P(\text{auto en la puerta } i)=1/3 \quad i=1,2,3$
3. El concursante elige inicialmente una puerta, sin información previa. Denotaremos esta puerta como la **puerta elegida**.
4. El presentador (Monty):
    - conoce exactamente en qué puerta está el premio;
    - **siempre** abre una puerta que **no** es la elegida por el concursante;
    - **siempre** revela una cabra (nunca abre la puerta que contiene el auto);
    - **siempre** ofrece al concursante la posibilidad de cambiar a la otra puerta que queda cerrada.
5. El concursante debe decidir entre:
    - **permanecer** con su elección original, o
    - **cambiar** a la puerta restante.

La pregunta central es: bajo este protocolo, ¿qué estrategia maximiza la probabilidad de ganar el premio?

En estas condiciones, el problema está bien definido y puede abordarse con las herramientas estándar de la probabilidad.

4. Solución elemental: enumeración de casos

Una primera solución, accesible incluso sin formalismo bayesiano, consiste en enumerar los casos posibles. Sin pérdida de generalidad, supongamos que el concursante elige inicialmente la **puerta 1**. Dado que el premio se coloca al azar, existen tres casos igualmente probables:

1. El auto está en la puerta 1.
2. El auto está en la puerta 2.
3. El auto está en la puerta 3.

El comportamiento de Monty en cada caso es el siguiente:

- Si el auto está en la puerta 1 (caso 1), Monty puede abrir la 2 o la 3, ambas con cabra.
- Si el auto está en la puerta 2 (caso 2), Monty está obligado a abrir la puerta 3, porque es la única con cabra entre las no elegidas.
- Si el auto está en la puerta 3 (caso 3), Monty está obligado a abrir la puerta 2, por el mismo motivo.

Podemos resumirlo en la tabla:

Caso	Auto	Elección inicial	Puerta que abre Monty	Si permanece	Si cambia
1	1	1	2 o 3 (cabra)	Gana	Pierde
2	2	1	3 (cabra)	Pierde	Gana
3	3	1	2 (cabra)	Pierde	Gana

Dado que los tres casos son equiprobables (cada uno ocurre con probabilidad (1/3)):

- La estrategia de **permanecer** gana solo en el caso 1 → probabilidad de ganar = (1/3).
- La estrategia de **cambiar** gana en los casos 2 y 3 → probabilidad de ganar = (2/3).

Con esto ya se justifica que el concursante debería cambiar. Sin embargo, el verdadero valor didáctico y conceptual del problema aparece cuando se lo reformula y resuelve explícitamente mediante el **teorema de Bayes**, haciendo foco en la actualización de probabilidades ante nueva información.

5. Solución mediante el teorema de Bayes

5.1. Definición de eventos

Supondremos, nuevamente, que el concursante elige inicialmente la puerta 1. Definimos los eventos:

- *(C1): "el auto está en la puerta 1"*

- *(C2): “el auto está en la puerta 2”*
- *(C3): “el auto está en la puerta 3”*

Por simetría y aleatoriedad en la colocación del auto:
P(C1) = P(C2) = P(C3) = 1/3
Definimos además el evento:

(Hi) “Monty abre la puerta i y muestra una cabra.

- *(H3):* “Monty abre la puerta 3 y muestra una cabra”.

El objetivo es calcular las probabilidades a posteriori:
P(C1 | H3) y P(C2 | H3),
es decir, las probabilidades de que el auto esté en la puerta 1 o 2, dado que Monty abrió la puerta 3. La comparación entre estos valores determinará si conviene permanecer o cambiar.

5.2. Probabilidades condicionales del comportamiento de Monty

Bajo el protocolo definido en la Sección 3, las probabilidades *(P(H3 | Ci))* son:

- Si el auto está en la puerta 1 (C1), Monty debe abrir una de las puertas 2 o 3, ambas con cabra. Supondremos que elige entre ellas al azar, con probabilidad (1/2) cada una:
  P(H3 | C1) = 1/2

- Si el auto está en la puerta 2 (C2), Monty **no puede** abrir la 2 (tiene el auto) ni la 1 (elegida por el concursante). Se ve forzado a abrir la puerta 3:
  P(H3 | C2) = 1.

- Si el auto está en la puerta 3 (C3), Monty no puede abrir la 3, por contener el premio, por lo que debe abrir la puerta 2:[
  P(H3 | C3) = 0.

5.3. Cálculo de la probabilidad del evento observado

Aplicamos la ley de la probabilidad total para el evento *(H3):*
*P(H3) = P(H3 | C1)*P(C1) + P(H3 | C2)*P(C2) + P(H3 | C3)*P(C3*).
Sustituyendo los valores:
*P(H3) =1/2*1/3 + 1*1/3 + 0*1/3 =1/2*

5.4. Actualización bayesiana

El teorema de Bayes establece que:
P(Ci | H3) = (P(H3 | Ci) P(Ci)) / P(H3).*

Para (C1):
P(C1 | H3) = = (P(H3 | C1) P(C1)) / P(H3) = (1/2*1/3)/1/2 = 1/3*

Para (C2):
P(C2 | H3) = = (P(H3 | C2) P(C2)) / P(H3) = (1*1/3)/1/2 = 2/3*

Por lo tanto, **condicionado a que Monty abre la puerta 3 y muestra una cabra**, la probabilidad de que el auto esté:

- en la puerta 1 (la elección original) es (1/3),
- en la puerta 2 (la puerta disponible para el cambio) es (2/3).

Estos valores son las **probabilidades posteriores** resultantes de la actualización bayesiana al observar la acción de Monty. La conclusión es inmediata: el concursante maximiza su probabilidad de ganar si **cambia** de puerta.

6. Variantes del problema y sensibilidad a las hipótesis

El resultado anterior depende críticamente de las hipótesis sobre el comportamiento del presentador. Modificar estas hipótesis conduce a otros problemas, que pueden modelarse también de manera bayesiana:

- Si Monty **no siempre abre una puerta**, sino solo a veces, la ausencia de apertura se convierte en información adicional que afecta las probabilidades.
- Si Monty **elige al azar** una puerta a abrir, sin conocer la ubicación del premio, la probabilidad de que al abrir una cabra queden dos puertas "simétricas" puede cambiar el análisis.
- Si el concursante sospecha que Monty tiene una **estrategia sesgada** (por ejemplo, abrir siempre la puerta 3 cuando puede), las probabilidades condicionales *(P(H3 | Ci))* cambian, y con ello la actualización bayesiana.

Estas variantes muestran que el problema de Monty Hall puede entenderse como un ejemplo simple de **decisión bajo información asimétrica**: el presentador tiene más información que el concursante, y el concursante debe inferir, a partir del comportamiento observable de Monty, cómo actualizar sus creencias.

7. Evidencia empírica y sesgos cognitivos

Diversos estudios en psicología del razonamiento probabilístico han mostrado que la mayoría de las personas:

- tiende a considerar que, tras abrirse una puerta con cabra, las dos puertas restantes están "en igualdad de condiciones";
- subestima el rol de la **información condicional** incorporada por la acción del presentador;
- muestra resistencia a aceptar que "cambiar" es sistemáticamente mejor que "quedarse".

En algunos experimentos, la proporción de participantes que elige cambiar es muy baja; incluso después de explicarles la solución correcta, muchos siguen percibiendo el problema como "50–50".

Curiosamente, en experimentos con animales (por ejemplo, palomas entrenadas con un esquema análogo), estos tienden a aprender la estrategia óptima de cambiar de "puerta" con mayor rapidez

que los humanos, simplemente por refuerzo repetido. Esto ha sido citado como ejemplo de cómo la intuición humana puede quedar atrapada en representaciones pobres de la estructura probabilística, mientras que un mecanismo de aprendizaje puramente empírico "descubre" la estrategia óptima.

El problema de Monty Hall se ha convertido así en un caso de estudio en:

- heurísticas y sesgos,
- dificultades con la probabilidad condicional,
- y enseñanza de estadística y probabilidad en distintos niveles educativos.

8. Aplicaciones y analogías en toma de decisiones

Aunque el escenario de las "tres puertas y las cabras" es puramente lúdico, la estructura del problema ofrece analogías con situaciones reales donde se debe decidir si **revisar una decisión inicial** a la luz de información nueva.

Algunos ejemplos conceptuales:

- **Diagnóstico médico**: ante un test con ciertas tasas de falsos positivos y falsos negativos, la aparición de nueva información (un segundo estudio, un resultado de laboratorio adicional) obliga a recalcular las probabilidades posteriores de enfermedad, de modo análogo a la actualización bayesiana que se ve en Monty Hall.
- **Gestión de riesgos**: en contextos de riesgo operativo o financiero, la aparición de señales parciales (por ejemplo, indicadores de estrés en un sistema) puede interpretarse como "puertas abiertas" que obligan a revisar la probabilidad asignada a escenarios de falla.
- **Investigación operativa y analítica de decisiones**: la distinción entre un **prior** (probabilidad inicial), la información observada y el **posterior** (probabilidad actualizada) está en la base de modelos bayesianos de decisión, líneas de ensamblado con inspección, sistemas de mantenimiento preventivo, etc.

En todos estos casos, el mensaje central que ilustra Monty Hall es que la información nueva rara vez debe tratarse como un adorno: requiere un **recálculo sistemático** de las probabilidades involucradas, y el teorema de Bayes proporciona el marco formal para hacerlo.

9. Conclusiones

El problema de Monty Hall combina tres ingredientes que lo han convertido en un clásico:

1. Una **formulación extremadamente simple**, comprensible en segundos.
2. Una **respuesta correcta contraintuitiva**, que contradice la percepción espontánea de la mayoría de las personas.
3. Una **estructura matemática rica**, que permite discutir probabilidad condicional, actualización bayesiana, información asimétrica y sesgos cognitivos.

En este artículo se ha presentado:

- el contexto histórico del problema;
- la formulación rigurosa del modelo estándar;

- la solución clásica por enumeración de casos;
- un desarrollo explícito basado en el teorema de Bayes;
- variantes que muestran la sensibilidad a las hipótesis sobre el comportamiento del presentador;
- y algunas conexiones con problemas reales de toma de decisiones bajo incertidumbre.

Más allá de su carácter lúdico, Monty Hall constituye una herramienta didáctica y conceptual valiosa en cursos de probabilidad, estadística, investigación operativa y analítica de decisiones. Sirve para enfatizar que la intuición probabilística puede ser engañosa, y que el uso correcto de la información disponible exige un marco formal de razonamiento, del cual el teorema de Bayes es una pieza central.

Sin embargo, la gran reflexión que ilustra el verdadero valor de lo que nos enseña la estadística es:

¿Qué pasa si seguimos el juego, cambiamos de puerta como mejor decisión y cuando abrimos y se resuelve el juego EL AUTO ESTA EN PUERTA INICIAL? A la espera de vuestros comentarios.

Referencias

- Selvin, S. (1975). A problem in probability (letter to the editor). *The American Statistician*, 29(1), 67.
- Gill, R. D. (2011). The Monty Hall Problem is not a probability puzzle — it's a challenge in mathematical modelling. *Statistica Neerlandica*, 65(1), 58–71.
- vos Savant, M. (1990). Ask Marilyn. *Parade Magazine*.
- Piattelli-Palmarini, M. (1991). *Inevitable Illusions: How Mistakes of Reason Rule Our Minds*. Wiley.
- Gardner, M. (1959). Mathematical Games: Problems involving questions of probability and ambiguity. *Scientific American*.

El concepto de entropía en simulación y teoría de la comunicación

Alejandro Roberti

Ing. en Alimentos. Profesor consulto del Departamento de Tecnología de la Universidad Nacional de Luján. Ex responsable de las asignaturas Ingeniería de Procesos, Investigación Operativa y Modelos, Simulación y Teoría de la Decisión

En nuestra publicación ***Optimiza 3: Modelos, Simulación y Teoría de la Decisión*** hemos referenciado el concepto "entropía" en el Capítulo 1, al mencionar "la tendencia al desgaste de los sistemas"; en el Capítulo 4, al mencionar los requisitos que debe cumplir un generador de números aleatorios, donde mencionamos el *blanqueamiento* de Von Neuman o *destilación de entropía*, como método que ofrece "seguridad de Shanon" y también en el Capítulo 1, destinado a Teoría de la Información, donde hablamos nuevamente de Claude Shannon, aunque no directamente de entropía, pero sí de su teoría matemática de la información que tiene uno de sus pilares en ella.

Debemos señalar que la entropía es un concepto muy particular que se utiliza en el campo de la física, específicamente en termodinámica, en la parametrización de los contenidos de energía de un sistema, y que tiene la particularidad de que no es una magnitud ni una ley física. Lo que la hace especial, además, es que no se cumple estrictamente, ni aun en sus propios ámbitos. Esto se debe a que se trata de un concepto estadístico referido a los comportamientos macroscópicos de la materia o de la energía en base a los estados microscópicos (microestados) de ese estado (macroestado). Esto implica que solo se puedan abordar esos comportamientos en términos estadísticos.

Si bien el párrafo anterior puede ser un galimatías y aún cuestionable, y pidiendo permiso a los termodinámicos y a los físicos, vamos a tratar de explicar algo de lo que concierne a nuestros temas desde un punto de vista muy simplificado con el objetivo meramente informativo y sin pretender presentarlo como un tratado sobre la entropía.

Debemos recordar que hemos oído conceptos muy arraigados. como por ejemplo:

"*la entropía de los sistemas tiende naturalmente a aumentar* ***a medida que transcurre el tiempo***"

y que solemos recurrir a casos de la naturaleza para demostrar que esto es cierto: por ejemplo, sabemos que una copa de vidrio (estado cuyos componentes están ordenados) que se rompe (estado cuyos componentes están desordenados) nunca se recompone espontáneamente y vuelve a ser una copa aunque pase mucho tiempo. Llamamos a esto "proceso irreversible" y es una de las bases del segundo principio de la termodinámica cuando se aplica a fenómenos que implican el uso, el aprovechamiento o la simple disipación de la energía.

Cuando nos hemos referimos a las propiedades de los **sistemas** y, entre ellas mencionamos a la entropía con la siguiente definición:

> "**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."

En esta definición estamos relacionando claramente varios conceptos fundamentales: *información, energía, orden, aleatoriedad y tiempo*.

Por experiencia sabemos que hay comportamientos de la materia que son irreversibles, es decir que no tienen una simetría "natural" a menos que se use energía externa. El ejemplo clásico es el de un recipiente lleno de agua, separada en dos cavidades por una compuerta: el agua de la izquierda tiene disuelta cierta cantidad de tinta y la derecha está pura. Así tenemos agua azul a la izquierda y clara a la derecha. Si sacamos o abrimos la compuerta, vemos que, al cabo de un tiempo, el agua se mezcla (el agua con colorante se difunde en todo el recipiente) y tenemos el recipiente completo con agua coloreada. Por más que dejemos pasar tiempo, lo contrario no ocurre: no queda el lado izquierdo con agua azul y el derecho con agua clara o viceversa, aunque pongamos nuevamente el tabique o aunque esperemos mucho tiempo. (Fig. 1)

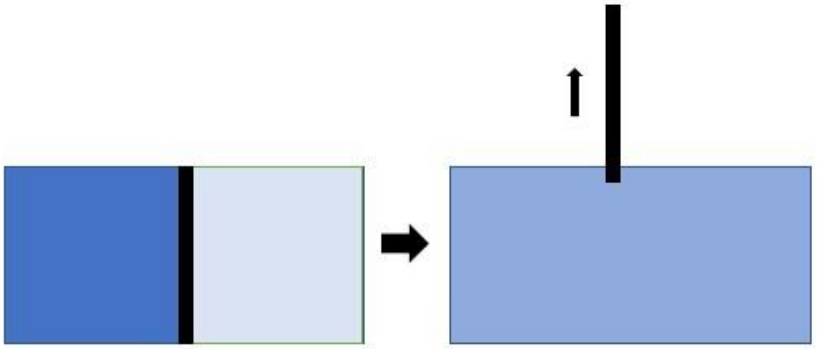


Fig. 1. A la izquierda un recipiente con agua separado en dos sectores por una compuerta. Uno de ellos tiene un colorante y el otro no. A la derecha, al abrir la compuerta ambos líquidos se mezclan y el color se difunde.

Ahora vamos a suponer que en el lado "azul", con la compuerta puesta en su lugar, solo tenemos 4 moléculas de colorante, y que las tenemos perfectamente identificadas: son las moléculas "A", "B", "C" y "D". (Fig. 2)

Acá podemos notar algunas cosas interesantes. La primera es que tenemos toda la información posible sobre el estado del sistema "agua con color" sabemos dónde están todas y cada una de las moléculas del colorante (en el lado izquierdo) y donde NO están (en el lado derecho). Además sabemos que, en estas condiciones, no hay otra posibilidad.

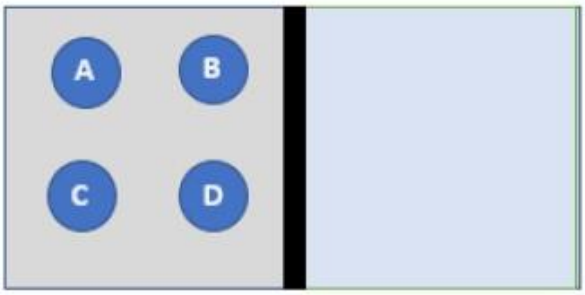


Fig. 2. Antes de abrir la compuerta el sector coloreado tiene todas las moléculas del colorante. Si queremos describir el recipiente tenemos la información completa de su estado.

Ahora vamos a abrir la compuerta y vemos que tenemos cuatro posibles estados:

1) Que las moléculas permanezcan todas a la izquierda
2) Que solamente una de ellas pase a la derecha y las otras tres queden a la izquierda
3) Que la mitad de ellas pasen a un lado y la otra mitad no pase
4) Que tres pasen a la derecha y una permanezca en la izquierda
5) Que las cuatro pasen a la derecha.

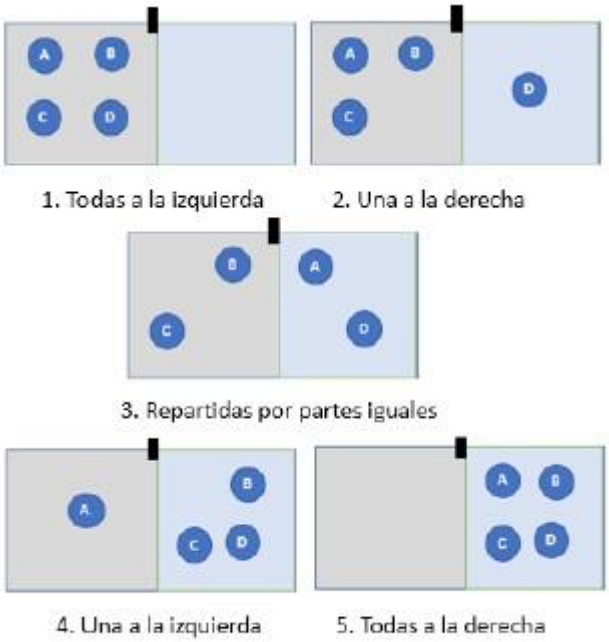


Fig. 3 Todos los macroestados posibles

Podemos decir que cada uno de estos estados es un "estado no específico" o macroestado. Tenemos, entonces, cinco macroestados para el caso de un recipiente que originalmente tenía dos cavidades y cuatro moléculas.

En el primer estado, tenemos información completa sobre la situación de cada molécula. Cada una de las cuatro está en el mismo sector. No hay otra posibilidad. Esto puede ocurrir entonces en un solo y único caso (o microestado) sobre los cinco posibles.

No es así el segundo caso, una a la derecha y tres a la izquierda, porque acá tendremos estas posibilidades:

Moléculas en cada lado	
Izquierda	Derecha
A – B – C	D
A – B – D	C
A – C – D	B
B – C – D	A

Como vemos para el macroestado "1 a la derecha, tres a la izquierda" hay cuatro estados particulares posibles (microestados) y tenemos que admitir que debemos buscar información extra para saber en cuál de esos microstados estamos. Hemos perdido información respecto al caso anterior. También sabemos que cada "microestado" tiene la misma probabilidad.

Para el tercer caso, “2 a la derecha, 2 a la izquierda”, tenemos seis posibles escenarios:

Moléculas en cada lado	
Izquierda	Derecha
A – B	C – D
A – D	C – B
A – C	B – D
B – C	A – D
B – D	A – C
C – D	A – B

Como vemos ahora hay seis microestados posibles con la misma probabilidad y tenemos que averiguar, conseguir información, sobre que moléculas están a la derecha, ya que puede haber una sola combinación de seis posibles. Hemos perdido aún más información.

Obviamente si analizamos lo que ocurriría con tres moléculas pasando a la derecha tendríamos el mismo razonamiento que en caso 2. Y si estuvieran todas a la derecha sería lo mismo que el punto de partida.

Podríamos hacer un gráfico de estados posibles (microestados) para cada uno de los estados inespecíficos. Los microestados posibles están en el eje de la “y” y los macroestados en el de las “x”

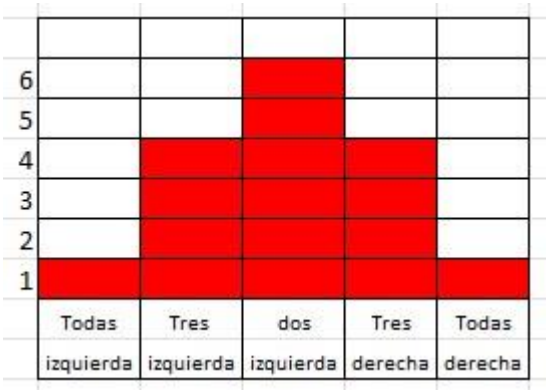


Gráfico 1.

Nos encontramos con que hay 16 microestados posibles y que solo nos encontraremos en el macroestado “***todas a la izquierda***” una vez cada 16 estados que se den (1/16) y 1/16 (0,065) veces que estén “***todas a la derecha***”. Esos dos serían casos de “máximo orden posible”, aunque son posibles, tenemos pocas chances de que ocurran.

Los casos con un desorden un poco mayor podrán darse con algunas posibilidades más, porque van a aparecer 4 veces entre 16 posibles, que es 4/16 = 0,25, pero el caso de máximo desorden lo vamos a encontrar 6 veces entre 16 posibles (6/16 = 0,375).

Si hacemos la suma de estos números, tenemos

Todos a la izquierda + una a la derecha + dos a la derecha + tres a la derecha. + todos a la derecha

$$(1/16) + (4/16) + (6/16) + (4/16) + (1/16) = (1 + 4 + 6 + 4 + 1) / 16 = 16/16$$

Esta suma nos dice que forzosamente debemos estar en uno de estos 16 casos. Además podemos decir que estamos calculando las probabilidades de encontrarnos en un estado determinado y que solamente nos encontraremos en uno de ellos porque al ser la suma igual a uno, las probabilidades son excluyentes.

La pregunta es, ¿El sistema evoluciona por alguna ley de la naturaleza al máximo desorden? La respuesta es **No.** Ocurre algo mucho más simple: es más probable encontrar al sistema en un estado donde existan muchas más combinaciones de microestados que en cualquier otro que tenga menos combinaciones. Simplemente, al haber más microestados en un macroestado, es más probable caer en uno que en el que en otro.

Si además pensamos que el colorante no tiene cuatro, sino miles de millones de moléculas, veremos que las probabilidades de encontrar un sistema ordenado son absolutamente despreciables porque se demuestra que, a medida que aumentamos el número de componentes la distribución de estados posibles centrada en el caso más frecuente (Gráfico 1) se hace cada vez más estrecha[1].

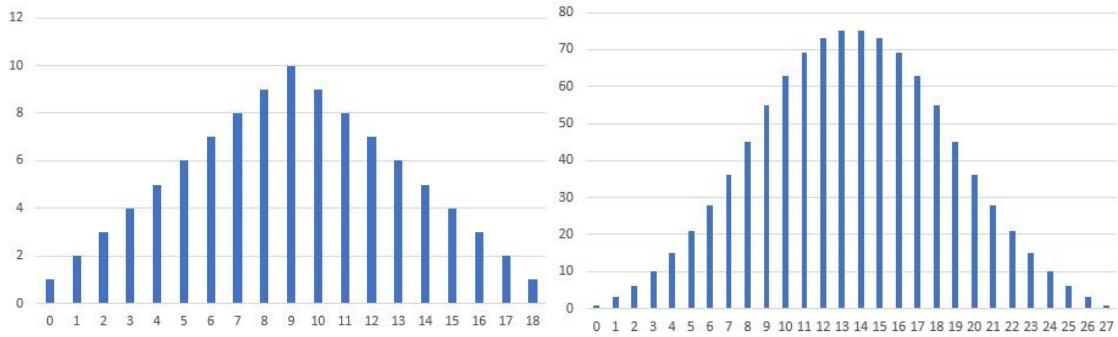


Gráfico 2 (izq.) frecuencia de aparición de sumas con 2 bolas — **Gráfico 3**. (Der.) con 3 bolas

Lo mismo ocurrió con la información: a medida que hay posibilidades de encontrar las moléculas distribuidas en cualquier orden, vamos perdiendo información sobre el lugar donde encontraremos una molécula en particular.

En el caso 1 y 5, el sistema tiene el *máximo orden posible* y solamente tenemos 1 grado de libertad en 16 opciones para describir exactamente el sistema y saber en qué recinto está cada una de las moléculas. *Menos ordenado*, casos 2 y 4, ahora tenemos cuatro combinaciones diferentes para cada recinto de posibles situaciones de cada molécula. *Totalmente desordenado*, ahora tenemos 6 posibles combinaciones de estados posibles para cada una de las moléculas, hemos perdido aún más información.

[1] Un ejemplo claro que se usa habitualmente es con dados: si se usa un dado la probabilidad de encontrar cualquier cara es la misma para las 6 caras. Si se usan dos dados, la probabilidad de encontrar la suma de caras vemos un mínimo de frecuencias para la suma 2 o la suma 12 que solo se pueden conseguir si ambos dados tienen un as o ambos un 6. La máxima frecuencia aparece con la suma 7, que se puede conseguir con 6 combinaciones de ambos dados: 1-6, 2-5, 3-4, 4-3, 5-2 y 6-1. Es claro que si usamos 1000 dados, nadie esperaría encontrar el número 1000 o el 6000, ya que eso solo se podría lograr si los 1000 dados cayeran simultáneamente en 1 o en 6, aunque la probabilidad existe, en términos prácticos es casi cero. [(1/6) elevado a la 1000]. Si en vez de dados sumamos los números de dos bolas numeradas del 0 al 9, tendremos una distribución como la del gráfico 2. Y de tres bolas también del 0 al 9, como la del gráfico 3. Ahí comprobamos como se estrecha la curva de frecuencias.

La **PRIMERA CONCLUSIÓN** de estas líneas es que la entropía no es una propiedad intrínseca de la materia o de la física o de la termodinámica o de la información. Es un principio estadístico.[2]

La **SEGUNDA CONCLUSIÓN** es que podemos definir la entropía exclusivamente como un fenómeno estadístico si usamos la expresión siguiente:

$$S = k \ln X,$$

Siendo *S* entropía, *k* la constante de *Boltzman*, y *X* la cantidad de microestados posibles equiprobables. Debemos prestar atención a este término: equiprobables. Por ejemplo *una molécula a la derecha/tres a la izquierda* es un estado válido si la probabilidad de que la molécula sea la "A" es la misma que la probabilidad de que sea la "B", la "C" o la "D".

Por ejemplo, si tenemos que calcular la entropía de una moneda (bit), que tiene solo dos estados, cara o ceca, nos quedaría, $\boldsymbol{S = k \ln 2}$. Podemos asimilar *X* a la información disponible.

De nuevo, en la moneda, esa información es cara (0) y ceca (1). En el caso de nuestras 4

moléculas tendremos $\boldsymbol{S = k \ln 2^4}$, lo que da que S es proporcional a 16… ya que 2 es el número de estados posibles (derecha o izquierda) y 4 es el número de componentes totales.

La **TERCERA CONCLUSIÓN** es que la entropía está relacionada con la pérdida de información. A medida que aumenta la entropía, hemos visto, aumenta la pérdida de información. Aumenta la incertidumbre. Aumenta la *aleatoriedad*.

Basado en estos conceptos (que acá hemos simplificado al máximo posible) es que se trabaja en base a entropía de sistemas para diseñar generadores pseudoaleatorios criptográficamente seguros.

El concepto es obtener secuencias de bits mediante un generador binario que soporte la llamada "prueba del siguiente bit": *Dados los primeros k bits de una secuencia aleatoria, no debería haber ningún algoritmo que pueda predecir el bit k+1 con una probabilidad de éxito superior a 0,5.*

Veamos esto en términos de entropía. Como se trata de bits binarios, podemos usar

$$\boldsymbol{S = k \log_2 X}.$$

Si graficamos en las x la probabilidad de cada uno de los estados posibles de ocurrencia de un bit aleatorio (0 o 1, o cara y ceca) y en ordenadas la medida de entropía, llamando 1 a la entropía de la máxima probabilidad de ocurrencia del bit, tenemos que para una probabilidad de 0,5 habrá una entropía máxima de 1. ($\log_2 2 = 1$)

[2] A pesar de esta afirmación debemos recordar que las interacciones entre las partículas reales existen y contribuyen a que aumente la entropía termodinámica. Este ejemplo es simplificador.

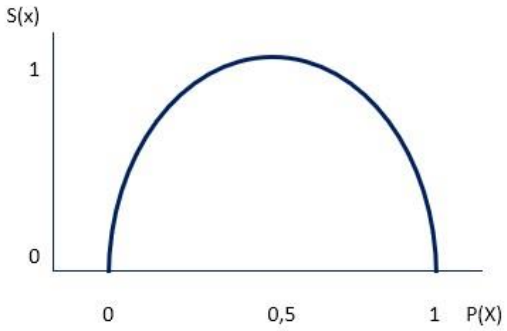


Fig. 4 – Máxima entropía de 1 bit

von Neumann probó que un algoritmo simple puede eliminar cualquier sesgo de un flujo de bits (a este proceso se lo conoce como destilación de entropía o blanqueamiento) si se aplica en conjunto con un procedimiento de generación de una serie aleatoria.

Veamos como ejemplo una secuencia de 1 bit, (digamos que genera una secuencia en que cada bit es *k*.) Le agregamos como parte de la destilación un bit "de blanqueo" que genera la secuencia *k+1,* vamos a obtener dos bits con cuatro posibles combinaciones: 00 – 01 – 10 – 11. El blanqueamiento de von Neuman nos entregaría al final un solo bit proveniente de desechar de estos cuatro los casos 00 y 11 porque presupone que en ellos el segundo bit fue "anunciado" por el primero y nos dejaría solamente con las combinaciones 01 y 10, los que se presentarían, como dijimos, como un solo bit: para 01 obtendríamos 0 y para 10 un 1.

Hemos obtenido un bit de entropía máxima usando un bit para blanquearlo.

Bit k	Bit k+1	Bit de salida
0	0	—
0	1	0
1	0	1
1	1	—

Si la entropía es función de la probabilidad de que X sea igual a 1, cuando la probabilidad P(X=1) sea 0,5 entonces todos los resultados posibles son equiprobables. Máxima entropía, máxima impredecibilidad. Es la prueba de la moneda. Si un proceso tiene la misma impredecibilidad que tirar una moneda, entonces es verdaderamente aleatorio.

De esta manera, la mención que hemos hecho a la teoría de comunicación en el capítulo 1, cuando la describimos como una aplicación de la TGS, encierra un grado de complejidad ya que la propuesta de Claude Shannon tiene como punto principal la entropía. En este caso se trata de una magnitud física (información) que es una secuencia de caracteres cuyo nivel de información se puede medir con la entropía y el mensaje en términos de cadena cuya información se mide con una secuencia de bits.

El principio es el siguiente: supongamos el caso de un mensaje. La probabilidad de acceder al contenido de cada mensaje no es igual para todos los mensajes. Sabiendo esto el emisor va a "gastar" más bits en codificar los mensajes menos probables y menos en los más probables. Por ejemplo si el mensaje es "***hay que medir 10 centímetros***", es muy probable que si se emite solamente "***medir 10 c***" el receptor "entienda" (no pierda información) que la c se refiere a "centímetros" y no note la ausencia de "***hay que"*** ya que aporta mucho menos información que

el resto. Dicho de otra manera no parece que sea necesario mandar el mensaje completo. El contexto referido a medida (en infinitivo) y la unidad que comienza con ***c*** son suficientes. Por el contrario, si el mensaje original fuera "***hay que adquirir 10 computadoras***" y nos limitamos a mandar el mensaje "***adquirir 10 c"*** seguramente necesitamos más información, porque si bien es cierto que ***hay que*** sigue siendo de baja importancia, en este contexto el carácter "***c***" perdió información porque puede ser interpretado como "*computadora*", "*carcaza*", "*conector*", "*cable*", etc. Debemos aumentar la longitud del mensaje porque la información alta que había en el carácter "***c***" del primer mensaje se perdió en el segundo caso.

Por esto se usa para todos los mensajes emitidos por una fuente un promedio ponderado de la longitud del código que se calcula en función de las probabilidades de ocurrencia de dicha longitud. Este promedio es la entropía de la fuente y sirve para cosas tan cotidianas como, por ejemplo, comprimir un archivo para enviarlo por correo electrónico. Se demuestra que la entropía de una fuente depende de la probabilidad de cada carácter emitido. (por ejemplo, en español, la "a" es más frecuente que la "w" en cualquier conjunto, por tanto "w" transmite más información que "a").

En contexto de un mensaje, no cambia mucho la información si se transmite "***iremos al cine***" como "***iremos cine***". En este caso, la partícula omitida ("***al***") posee muy poca información. En cambio si hemos transmitido solamente "***iremos al***" hemos perdido mucha información al omitir la partícula "***cine***"

Es común que para medir la información se ejemplifique con la cantidad de mensajes que serían necesarios para responder la supuesta pregunta que originó estos mensajes. Muchas veces la respuesta es binaria (SI – NO) y con 1 solo bit alcanza para responder una pregunta simple (¿estás en viaje?). Pero no así para otras preguntas (¿En qué fecha te queda bien reunirnos en Rosario?) Esta definición llevó a Shannon a imaginar que, sin importar lo compleja que sea la pregunta, la misma pregunta se podía desglosar en una cantidad de preguntas simples que nos llevaría a la respuesta.

Veamos un ejemplo: Queremos saber quién produjo un informe de entre los miembros de un equipo formado por dos hombres y dos mujeres, cuyos nombres son ***Juan, Pedro, Ana y Belén***. Como la pregunta ¿Quién escribió el informe? No nos sirve porque no hay respuestas por SI o por NO, vamos a averiguarlo mediante preguntas de respuestas binaria:

1) El que escribió el informe, es hombre?, La respuesta por **SI** o por **NO** elimina al 50% de los candidatos.

2) La segunda pregunta, en función de la respuesta a la primera sería: Su nombre comienza con J? (si era hombre) Su nombre comienza con A? (si era mujer). Con esta segunda respuesta tenemos el nombre preciso. Hemos necesitado solamente 2 bits para reconstruir un mensaje con 4 posibles respuestas.

De esta manera vemos que dividiendo la cantidad de opciones en grupos con algún criterio podemos elegir cualquier cantidad de opciones y dividirlas sucesivamente por la mitad hasta llegar a una única alternativa. Para saber cuántas veces vamos a dividir una cantidad por 2 para llegar a la unidad usamos nuevamente el logaritmo en base 2 del número de preguntas.

$$H = log_2 x$$

En este caso el log_2 de 4 es 2. Si el caso fuera saber mi PIN de acceso a Window, que es un número de 5 dígitos, tendría que responder log_2 de 100.000 = 16,6 preguntas o 16 bits para acceder al número correcto. Quizá en este punto se entiende un poco mejor la destilación de entropía de von Neumann que mencionamos un poco más arriba.

Sin embargo este ejemplo supone que la mitad es masculina y la mitad es femenina. Pero, no siempre esto es posible, por ejemplo, ¿Qué ocurre si sabemos de antemano que solo puede haber un hombre y dos mujeres?

Digamos que son *Juan, Ana, Belén y Beatriz*. Ahora necesitaríamos hacer estas preguntas, que se responden por SI o por NO:

1) Es Mujer? (si la respuesta es NO ya conocemos el mensaje y hemos usado un solo bit)
2) Si la respuesta es SI: ¿su nombre comienza con "B"? (si la respuesta es NO ya sabemos que es Ana y hemos gastado 2 bits)
3) Si la respuesta es SI: ¿Su nombre finaliza con "z"? y obtenemos la respuesta gastando 3 bits

Hemos necesitado 1 bit en un camino, 2 en el segundo y 3 en el tercero.

En estos casos la primera respuesta condiciona la(s) siguiente(s) pregunta(s) y cambia el estado en antes de la pregunta y después.

Si en lugar de usar el log_2 de la cantidad x de estados, ahora probamos con la expresión

$$H = log_2 \frac{opciones\ antes\ de\ contestar}{opciones\ remanentes\ después\ de\ contestar}$$

Tendremos para el primer ejemplo, (cuando había 2 hombres y 2 mujeres): SI (es hombre)

$$H = log_2 \frac{4}{2} = 1$$

Segunda pregunta, tanto como si la primera respuesta fue SI como si fue NO)

$$H = log_2 \frac{2}{1} = 1$$

Como vemos, el número de bits es igual al que encontramos anteriormente con el otro método.

En el segundo ejemplo (un solo hombre) las preguntas no dividen por la mitad. La primera es para el saber si es hombre o no:

$$H = log_2 3 = 0{,}41 \text{ para "no"}$$

Quiere decir que ahora esta pregunta tiene menos información que antes, ya que "vale" 0,4 bits (o sea, esta sola vale casi como media pregunta de las de antes o como dos preguntas si fue contestada como si), para saber, con una sola que se trata de un hombre. Así que resulta práctico ordenar las preguntas según la importancia que tienen en el mensaje. Sabemos que 1 de cada cuatro personas es Hombre, pero que 3 de cada 4 es mujer. Tendremos 2 opciones antes de preguntar y solo 1 después de hacerlo.

$$H = log_2 \frac{3}{2} = 0{,}58$$

Que significa que esta respuesta también tiene poca información (excepto que la respuesta sea SI), ya que seguimos dudando entre personas restantes, lo que nos lleva a hacer una tercera pregunta, que, al ser la última, siempre valdrá un bit:

$$H = log_2 \frac{2}{1} = 1$$

Fuimos pasando de 4 personas a 3, de 3 a 2 y de 2 a 1.

Esta forma de ver el problema se relaciona con la probabilidad de que sea un hombre (1/4) o que sea una mujer (3/4) y aún en ese caso que su nombre comience con "B" (2/3) o con "A". Por este motivo, como los números son los mismos pero inversos, la expresión de Shannon también la podemos escribir así pensando en hacer una media ponderada de respuesta que necesito según las preguntas:

$$H = log_2 \frac{1}{p}$$

¿Cómo se expresa esto como información? Si supiéramos la respuesta (*p = 1*) no necesitamos información. Si tenemos un 50% de probabilidad de saber la respuesta, (*p = 0,5*) tenemos 1 bit de información, y si, en cambio, la probabilidad de conocer la respuesta tiende a cero, más bits de información necesitamos y obtenemos de ella.

Finalmente, como estamos usando probabilidades que pueden ser diferentes en cada uno de los pasos, lo que haremos es usar una media ponderada para calcular la cantidad final de entropía del mensaje:

$$I = \sum_i p_i log_2 \left(\frac{1}{p_i}\right)$$

En el ejemplo de 1 hombre y 3 mujeres, tenemos una probabilidad de 0,25 para hombre y de 0,75 para mujer, por lo tanto, necesitaremos 0,8 bits promedio ponderado para llegar a la

respuesta.

En la Figura 5 comprobamos que para llegar a Juan necesitamos 1 bit, para Ana 2 bits y para el resto (Beatriz o Belén), 3 bits.

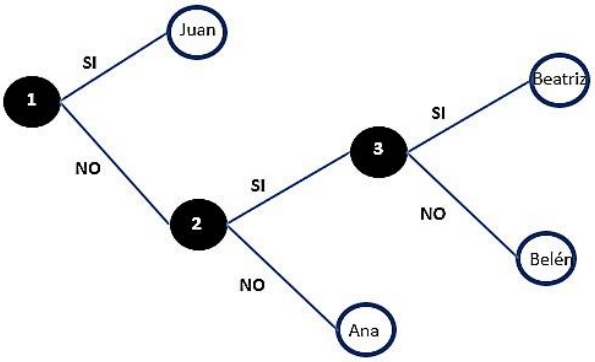


Fig. 5 – Entropía condicionada por la respuesta

La probabilidad de que sea Juan es de ¼, podemos escribir

$$Juan = p_1 log_2\left(\frac{1}{p_1}\right) = 0{,}25 \times log_2\left(\frac{1}{0{,}25}\right) = 0{,}5$$

Nos queda un remanente de 3 personas. El factor ponderado de que sea Ana será

$$Ana = p_2 log_2\left(\frac{1}{p_2}\right) = 0{,}333 \times log_2\left(\frac{1}{0{,}33}\right) = 0{,}53$$

Ahora solo nos quedan 2 personas, cada una de ellas tiene la misma probabilidad y su factor de ponderación será

$$Belen/Beatriz = p_3 log_2\left(\frac{1}{p_3}\right) = 0{,}5 \times log_2\left(\frac{1}{0{,}5}\right) = 0{,}5$$

Entonces, la cantidad media de bits necesarios será

$$I = p_1 log_2\left(\frac{1}{p_1}\right) + p_2 log_2\left(\frac{1}{p_2}\right) + p_3 log_2\left(\frac{1}{p_3}\right) = 0{,}5 + 0.53 + 0{,}5 = 1{,}53$$

Que significa que llegamos a la pregunta final con menos de 2 bits.

Obviamente, este es un ejemplo pueril, aunque sigue siendo válido es situaciones más complejas.

Referencias:

John von Neumann, 1963, “Various techniques for use in connection with random digits” . The Collected Works of John von Neumann. Pergamon Press, pp. 768-770 ISBN 0- 8-009566-6

McKay, David, 2003. “Information Theory, Inference and Learning Algorithms”, Cambridge University Press.

Cadenas de Márkov: el lenguaje probabilístico oculto detrás de muchos procesos.

Introducción

Más de cien años después de que Andrei Markov formulara su idea sobre “el futuro que depende solo del presente”, sus modelos siguen escondidos detrás de muchas herramientas modernas de optimización, gestión de riesgos e inteligencia artificial. En este artículo proponemos una mirada conceptual y aplicada a los modelos markovianos, pensados como un lenguaje natural para describir procesos bajo incertidumbre en donde en un principio desconocemos la probabilidad de ocurrencia de los eventos

1. Por qué hablar hoy de modelos markovianos

En un cajero automático, en una línea de producción, en un *call center* o en la evolución de la cartera de créditos, hay algo en común: los sistemas **cambian de estado con el tiempo** y lo hacen con un fuerte componente de incertidumbre.

En la práctica, solemos preocuparnos por preguntas como:

- ¿Cuánto tiempo va a esperar un cliente?
- ¿Con qué frecuencia se detiene una máquina?
- ¿Cuál es la probabilidad de que un cliente entre en mora?
- ¿Qué fracción del tiempo mi proceso está “sano” y qué fracción está “en problema”?

Los **modelos markovianos** ofrecen un marco matemático preciso, pero sorprendentemente intuitivo para abordar estas preguntas. Nos permiten representar el sistema como un conjunto de **estados** y un patrón de **transiciones probabilísticas** entre ellos, y a partir de ahí analizar comportamientos de corto y largo plazo, comparar políticas y tomar decisiones de optimización.

2. Un poco de historia: de Markov a la optimización moderna

El punto de partida está en la obra de **Andrei Andreevich Markov (1856–1922)**, matemático ruso que, a comienzos del siglo XX, se propuso estudiar secuencias de variables aleatorias que no son independientes, pero en las que el futuro inmediato solo depende del estado presente.

Markov introdujo lo que hoy llamamos **cadenas de Markov** analizando, entre otros problemas, la sucesión de letras en textos literarios. La idea central puede resumirse así:

Para predecir el próximo estado, alcanza con saber dónde estoy ahora. No necesito toda la historia.

Décadas más tarde, **Andrei Kolmogorov** aportó los fundamentos formales de la probabilidad moderna y los **procesos estocásticos**, dando un marco riguroso a los procesos con propiedad de Markov, tanto en tiempo discreto como en tiempo continuo.

A partir de allí, los modelos markovianos se integraron a varios campos que hoy son familiares para los lectores de PULSO OPTIMIZA:

- **Teoría de colas** y dimensionamiento de recursos.
- Investigación de operaciones y programación dinámica.
- Fiabilidad y mantenimiento de sistemas complejos.
- Ingeniería financiera y gestión de riesgos.
- Y, más recientemente, **machine learning** e **inteligencia artificial**, a través de procesos de decisión de Markov y cadenas de Markov ocultas.

Lo interesante es que muchas de las herramientas modernas que hablamos con naturalidad (simuladores, algoritmos de control, modelos de *scoring*, métodos de aprendizaje por refuerzo) tienen, en su corazón, una estructura marcadamente markoviana.

3. El diccionario básico markoviano

Para poder trabajar con estos modelos con comodidad, conviene fijar algunos conceptos clave.

3.1 Proceso estocástico

Un **proceso estocástico** es una familia de variables aleatorias que describen cómo evoluciona en el tiempo una cierta magnitud de interés. Ejemplo clásico: el número de clientes en un sistema de colas minuto a minuto.

3.2 Espacio de estados

El **espacio de estados** es el conjunto de situaciones posibles del sistema. No estamos modelando cada detalle microscópico, sino las categorías que realmente importan para decidir.

Algunos ejemplos:

- Cola en un cajero: 0, 1, 2, 3, … clientes en el sistema.
- Estado de una máquina: operativa, degradada, en falla, en reparación.
- Estado de un cliente: normal, atrasado, en mora, castigado.

La elección de estados es ya una decisión de modelización: demasiados detalles vuelven el modelo inmanejable; muy pocos, lo vuelven irrelevante.

3.3 Propiedad de Markov

Un proceso tiene **propiedad de Markov** si, en términos formales,

$$\mathbb{P}(X_{n+1} = j \mid X_n = i, X_{n-1}, \dots, X_0) = \mathbb{P}(X_{n+1} = j \mid X_n = i)$$

Es decir, la probabilidad del próximo estado depende **solo** del estado actual, no del recorrido completo que llevó hasta él.

Intuitivamente: “el pasado se resume en el presente”.

3.4 Cadenas de Markov en tiempo discreto

Cuando el tiempo avanza en pasos (t = 0, 1, 2, …) y en cada paso el sistema puede cambiar de estado, hablamos de una **cadena de Markov en tiempo discreto**.

La información esencial se concentra en la **matriz de transición** P, donde cada elemento p_{ij}indica la probabilidad de pasar del estado $i$al estado $j$en un paso de tiempo:

$$P = [p_{ij}], p_{ij} = \mathbb{P}(X_{n+1} = j \mid X_n = i)$$

Cada fila de la matriz suma 1, porque desde un estado dado el sistema necesariamente pasará a alguno de los estados (incluido eventualmente el mismo).

3.5 Estados transientes y recurrentes

Dentro de una cadena de Markov, algunos estados son "de paso" y otros son "de residencia".

- Un estado es **recurrente** si, una vez visitado, la probabilidad de volver a él en algún momento futuro es 1.
- Es **transiente** si hay una probabilidad positiva de que, una vez que lo dejamos, no volvamos nunca.

Esta clasificación es fundamental para saber si ciertos escenarios son excepcionales o forman parte estructural del funcionamiento del sistema.

3.6 Distribución estacionaria

Bajo condiciones razonables (cadena irreducible, aperiódica), las potencias sucesivas de la matriz de transición convergen y la distribución de estados se acerca a una **distribución estacionaria** π, que cumple:

$$\pi = \pi P$$

Esta distribución describe, en el largo plazo, **qué fracción del tiempo** pasará el sistema en cada estado.

Desde el punto de vista de la optimización, esto es oro: nos permite cuantificar, por ejemplo:

- Porcentaje de tiempo con cola larga.
- Porcentaje de tiempo con máquina parada.
- Probabilidad de encontrar a un cliente típico en estado de mora.

3.7 Procesos markovianos en tiempo continuo

Cuando las transiciones no ocurren en pasos discretos, sino en un tiempo continuo (fallas, reparaciones, llegadas de clientes, etc.), hablamos de **procesos de Markov en tiempo continuo**. (a veces conocidos como procesos de nacimiento-muerte)

En lugar de una matriz de probabilidades por período, trabajamos con una **matriz de tasas de transición** Q, donde cada elemento q_{ij} refleja la "velocidad" a la que el sistema pasa del estado i al estado j.

Estos modelos son centrales en fiabilidad, mantenimiento e ingeniería de sistemas.

4. De la realidad al modelo: cómo se construye una cadena de Markov útil

Desde la perspectiva de optimización de procesos, la pregunta práctica es: ¿Cómo paso de un sistema real, con sus datos y complejidades, a un modelo markoviano que me ayude a decidir?

Un camino típico es:

1. **Definir los estados relevantes** Se identifican las situaciones que realmente afectan el desempeño o la decisión. Ejemplo: para una línea de envasado puede bastar con tres estados: *operativa*, *en ajuste* y *detenida por falla*.
2. Elegir el tipo de tiempo
    - ¿Tiene sentido un modelo en **tiempo discreto** (por turno, por día, por mes)?
    - ¿O las transiciones ocurren naturalmente en **tiempo continuo** (fallas, llegadas, reparaciones)?
3. **Estimar probabilidades o tasas de transición** A partir de datos históricos, registros de mantenimiento, tiempos de atención, etc., se construyen:
    - Probabilidades de pasar de un estado a otro en un período, o
    - Tasas de transición por unidad de tiempo.
4. **Construir la matriz de transición (o de tasas)** Esta matriz sintetiza el comportamiento dinámico del sistema. Una vez obtenida:
    - Se pueden simular escenarios.
    - Se calculan probabilidades a corto plazo.
    - Se estudia el comportamiento estacionario.
5. **Derivar indicadores de gestión** A partir del modelo se obtienen métricas como:
    - Probabilidad de cola mayor que cierto umbral.
    - Porcentaje de tiempo con máquina parada.
    - Tiempo esperado hasta la falla.
    - Probabilidad de que un cliente llegue a estado “castigado”.
6. **Evaluar alternativas de decisión** Modificando parámetros (cantidad de recursos, políticas de mantenimiento, reglas de crédito, etc.) se compara el impacto sobre estos indicadores y se elige la alternativa que optimiza el desempeño.

5. Relevancia actual: ¿dónde los modelos markovianos encuentran su lugar de trabajo?

5.1 Industria 4.0 y monitoreo de procesos

En entornos conectados, con sensores y sistemas de supervisión en tiempo real, es natural describir el estado de una planta, de una línea o de un recurso mediante un número acotado de estados (normal, alerta, crítico, fuera de servicio). Las transiciones entre estos estados, alimentadas por datos, encajan perfectamente en marcos markovianos.

5.2 Gestión de riesgos y finanzas

La evolución de un cliente en una cartera de créditos (al día, atrasado, en mora) o la transición de un activo entre distintas calificaciones de riesgo puede modelarse con cadenas de Markov. Estas sirven para proyectar la calidad futura de la cartera y alimentar decisiones de provisiones, límites y estrategias de cobranza.

5.3 Fiabilidad y mantenimiento

Equipos complejos se modelan como sistemas que saltan entre estados de operación, degradación, falla y reparación. Con un modelo markoviano se pueden diseñar políticas de mantenimiento preventivo o predictivo, evaluar contratos de servicio y cuantificar el costo esperado de paradas.

5.4 Sistemas de colas y servicios

Call centers, cajeros automáticos, centros de salud, sistemas logísticos urbanos… en todos ellos el número de clientes en sistema y los tiempos de espera pueden tratarse mediante modelos de colas de naturaleza markoviana. Esto permite dimensionar recursos para alcanzar niveles de servicio deseados con costos razonables.

5.5 Inteligencia artificial y aprendizaje por refuerzo

En procesos de decisión de Markov (MDP) se combinan:

- Estados (la situación actual del sistema).
- Acciones (las decisiones posibles en cada estado).
- Recompensas (beneficios o costos asociados a cada transición).

Los MDP son la base teórica de muchos algoritmos de **aprendizaje por refuerzo**, donde un agente "aprende" políticas óptimas a partir de la experiencia. Así, los modelos markovianos conectan directamente la tradición de la investigación de operaciones con la IA contemporánea.

6. Algunas aplicaciones ilustrativas

6.1 Cola en un cajero automático

- **Estados:** 0, 1, 2, … personas en el sistema.
- **Interés:** probabilidad de que la cola supere cierta longitud, tiempo medio de espera, impacto de agregar un segundo cajero.
- Con un modelo de colas markoviano (por ejemplo, un M/M/1 o M/M/c) se pueden comparar configuraciones sin necesidad de experimentar "en vivo" con los clientes.

6.2 Línea de producción con paradas

- **Estados:** operativa, parada por falla, en reparación, en mantenimiento preventivo.
- **Interés:** porcentaje de tiempo disponible, frecuencia de paradas, costo esperado de indisponibilidad.
- Un modelo markoviano permite evaluar cómo cambia la disponibilidad si alteramos la estrategia de mantenimiento o la calidad de los repuestos.

6.3 Inventarios y quiebres de *stock*

- **Estados:** *stock* alto, medio, bajo, quiebre.
- **Interés:** probabilidad de quiebre de *stock*, tiempo medio entre quiebres, costo combinado de inventario y falta.
- Con una cadena de Markov ajustada a la demanda y reposición, es posible diseñar políticas de pedido más robustas.

6.4 Evolución de una cartera de créditos

- **Estados:** al día, 30 días de atraso, 60 días de atraso, castigado.
- **Interés:** probabilidad de default, distribución esperada de la cartera en el futuro, impacto de distintas políticas de seguimiento.
- La matriz de transición, estimada con datos históricos, se convierte en una herramienta de proyección y gestión de riesgo.

7. Mirando hacia adelante: enseñanzas de los modelos markovianos

Más allá de su contenido técnico, los modelos markovianos dejan varias enseñanzas valiosas para quienes trabajan en optimización y gestión:

1. **Pensar en estados y transiciones** Obligan a abstraer la realidad en términos de situaciones clave y cambios entre ellas. Esa forma de pensar simplifica la complejidad sin caer en modelos ingenuos.
2. **Aceptar el azar, pero estructurarlo** La incertidumbre deja de ser "ruido incomprensible" y se convierte en un conjunto de probabilidades explícitas, que se pueden medir, discutir y mejorar.
3. **Equilibrar corto y largo plazo** La distinción entre dinámica transitoria y comportamiento estacionario ayuda a evitar decisiones que sólo alivian el presente a costa de degradar el futuro del sistema.
4. **Aprovechar los datos de manera inteligente** En un entorno con cada vez más información (sensores, registros, logs), los modelos markovianos son una forma natural de convertir datos históricos en modelos dinámicos útiles para decidir.
5. **Tender puentes entre disciplinas** Desde la teoría de colas hasta la inteligencia artificial, pasando por mantenimiento y finanzas, los modelos markovianos funcionan como un lenguaje común que conecta áreas tradicionalmente separadas.

8. Conclusión

Los modelos markovianos nacieron como una curiosidad matemática a principios del siglo XX y hoy constituyen una herramienta silenciosa pero central en la optimización de procesos, la gestión de riesgos y el diseño de sistemas inteligentes.

Para quienes trabajan en mejora continua y toma de decisiones, aprender a mirar sus sistemas como **cadenas de estados y transiciones probabilísticas** no es solamente un ejercicio teórico: es una forma poderosa de ganar claridad en entornos complejos y de fundamentar decisiones que, inevitablemente, se toman bajo incertidumbre.

Desarrollo de aplicaciones destinadas a los estudiantes de las asignaturas en reemplazo de software obsoleto

El presente artículo está basado en el Proyecto de Desarrollo de la Asignatura presentado en el Departamento de Tecnología (UNLu) con la meta de reemplazar todas las funciones del viejo programa WinQSB utilizado anteriormente y que está en estado de obsolescencia técnica. Para mayor practicidad se decidió implementarlo en la modalidad de plantillas con macros de MS 365 Excel

Objetivo

Como se trata de un proceso dinámico y de desarrollo constante sirve para profundizar los conocimientos disciplinares de los integrantes del equipo de docentes y a la vez para generar herramientas de formación tanto para futuros docentes investigadores como para estudiantes de grado.

Paralelamente permite cubrir la falta de software especializado en este tipo de problemas mediante el desarrollo de una aplicación simple. Es esperable que, subsidiariamente, se logre un mayor acercamiento de los estudiantes al empleo de herramientas informáticas de uso extendido, entre ellas, las hojas de cálculo y la programación.

Fundamentos:

El equipo docente a cargo de las disciplinas ha detectado, con el paso de los años, la desaparición, por obsolescencia tecnológica, de la disponibilidad de software capaz de brindar soluciones rápidas tanto al estudiante como al profesional que antes era cubierto perfectamente por software avanzado de alto costo (por ejemplo LINDO/LINGO) orientado solo a aspectos restringidos (programación lineal, transporte, etc.) o por software de aplicación más amplia, que hace varias décadas que fue discontinuado y con alto grado de obsolescencia pero que abarcaba la mayor parte de los temas (WinQSB).

Sin embargo, en algunos temas, como el de la toma de decisión en ambientes concretos como los que en este artículo se ejemplifican, directamente no existen herramientas informáticas de aplicación específica.

Desde hace tiempo, en las asignaturas, se viene trabajando con los estudiantes en la propuesta de creación de algoritmos y soluciones informáticas capaces de resolver los casos vacantes.

La utilización de hojas de cálculo que se propone a los estudiantes durante el desarrollo de la actividad docente muestra, año a año, una curiosa resistencia inicial a su empleo por parte de ellos que, a poco se va transmutando en entusiasmo por usarla, lo que — por otra parte — es fomentado por el equipo docente.

De allí surgió la idea de comenzar a proponer el uso de estas planillas en casos en que el software existente no da respuestas o da respuestas muy pobres, tales como planificación de inventarios,

rentabilidad esperada en condiciones de incertidumbre o planificación de producción en condiciones de incertidumbre, como es el caso de ejemplo que sigue.

Método de trabajo.

Partiendo de un caso hipotético (un producto perecedero) desarrollamos el modelo algebraicamente y lo comparamos y testeamos con un modelo matricial generado en hoja de cálculo, que, de superar la prueba, será usada como herramientas de práctica para el estudiante y los profesionales.

El modelo construido proviene de optimizar la decisión aplicando el criterio de Savage en la **Teoría de Toma de Decisiones bajo condiciones de riesgo o de incertidumbre**, conocido como "*criterio del mínimo arrepentimiento*" o "*mínimo lamento*". Se considera que éste es uno de los criterios que en un escenario incierto puede generar una decisión óptima generalmente aceptada como la más eficiente.

Desarrollo

Se buscará determinar el lote óptimo de producción de un producto perecedero, bajo condiciones de incertidumbre. Esa condición se aplica a la variable aleatoria DEMANDA, suponiendo desconocimiento de la naturaleza de la misma ni posibilidades de establecer esa naturaleza en forma subjetiva.

Se trata de un producto perecedero que tiene un lapso de vida útil, a partir del cual se considera cualitativamente no apto. Para desarrollar el modelo se dispone de la siguiente información determinística:

v: Precio de venta unitario del producto **[$/un]**
c: Costo total unitario del producto **[$/un]**
$\mathbf{v_r}$: Valor residual unitario del producto **[$/un],** referido a las unidades sobrantes
$\mathbf{c_s}$: Costo de escasez de las unidades faltantes en **[$/un]**.

Definiciones:

X: Demanda del producto. Son los valores discretos, $\mathbf{X_j}$, que puede tomar esa demanda

Donde $X_j \geq 0$; $X_j \in Z$

Los valores de demanda del mercado, $\mathbf{X_j}$, pertenecen al intervalo [d,D],

siendo **d**: la demanda mínima del producto y

D: la demanda máxima del producto

Además

$v > c > v_r$

$v > c > c_s$

P. Producción en unidades

B. Beneficio en $

Con las anteriores descripciones, se construye una herramienta algebraica que permita resolver el problema planteado, mediante la utilización del criterio MINIMAX de Savage

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

aplicable a lanzamientos de nuevos productos, para los cuales no se tiene información estadística previa, ni una correlación satisfactoria con vinculación a otro producto similar en el mercado.

La producción, en unidades, **P,** no tiene restricciones para cubrir cualquier valor establecido de la demanda X. Buscamos un resultado operativo, que será el cálculo del **Beneficio** (**B**) y para lograrlo estableceremos dos casos.

Caso en que demanda es menor que lo que se produce

$X \leq P$, donde resulta que $B = v.X - c.P + v_r (P - X)$

que puede expresarse como

$$X \leq P \rightarrow B = (v - v_r) X + (v_r - c) P$$

Caso en que la demanda supera lo que se produce

$X \geq P$, donde resulta que $B = v P - c P - c_s (X - P)$

$$X \geq P \rightarrow B = (v + c_s - c) P - c_s X$$

Puede definirse como "*mejor acto*", cuando $X = P$

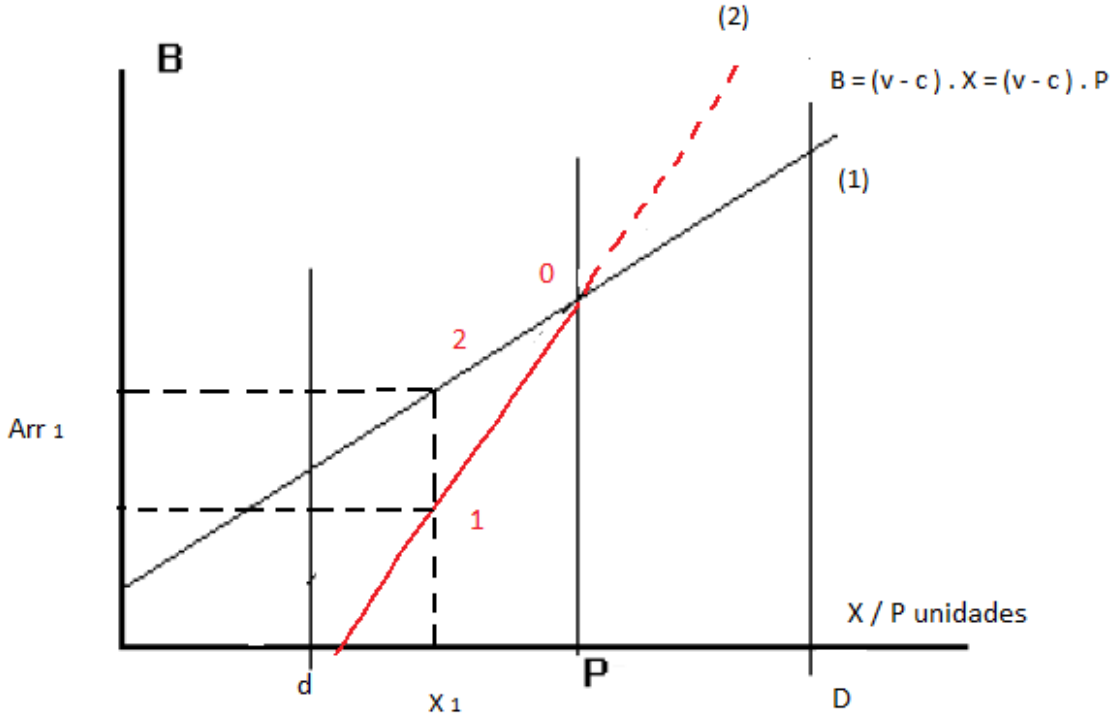
$$B = (v - c) X = (v - c) P$$

Esta definición se hace necesaria para la construcción de la función de ARREPENTIMIENTO tal como la definiera Savage.

Desarrollando los dos casos, se encuentra:

Primer caso

Considerando una producción P cualquiera sea y una demanda X_1, tal que $X_1 < P$, y graficando $B = f(X/P)$, resulta



La recta (1), representa la recta de los mejores actos, donde se cumple la condición X = P, es decir, se produjo P y resultó una demanda X tal que X = P, o, dicho de otra manera, la demanda en justamente lo producido.

La recta (2) representa la situación en donde habiendo producido P se da una demanda menor, X_1 tal que $X_1 < P$, donde

$$B = (v - v_r)\, X + (v_r - c)\, P \qquad \text{Ecuación 1}$$

en este caso, el punto 1 representa el Beneficio de que, habiendo producido P, se da la demanda X_1 ($X_1 < P$), y el punto 2, representa el punto máximo que se hubiera dado si $P \rightarrow X_1$ y se cumple que

$$X_1 = P.$$

De acuerdo con Savage, se define la función de arrepentimiento como

$$Arr = \max\,(b_{ij} - b_{ij})$$

Es decir, el máximo beneficio posible (punto 2) menos lo que realmente sucedió (punto 1). Entonces, para este análisis resulta que

$$Arr_1 = (v - c)\, X_1 - [(v - v_r)\, X_1 + P\,(v_r - c)]$$

Reordenando

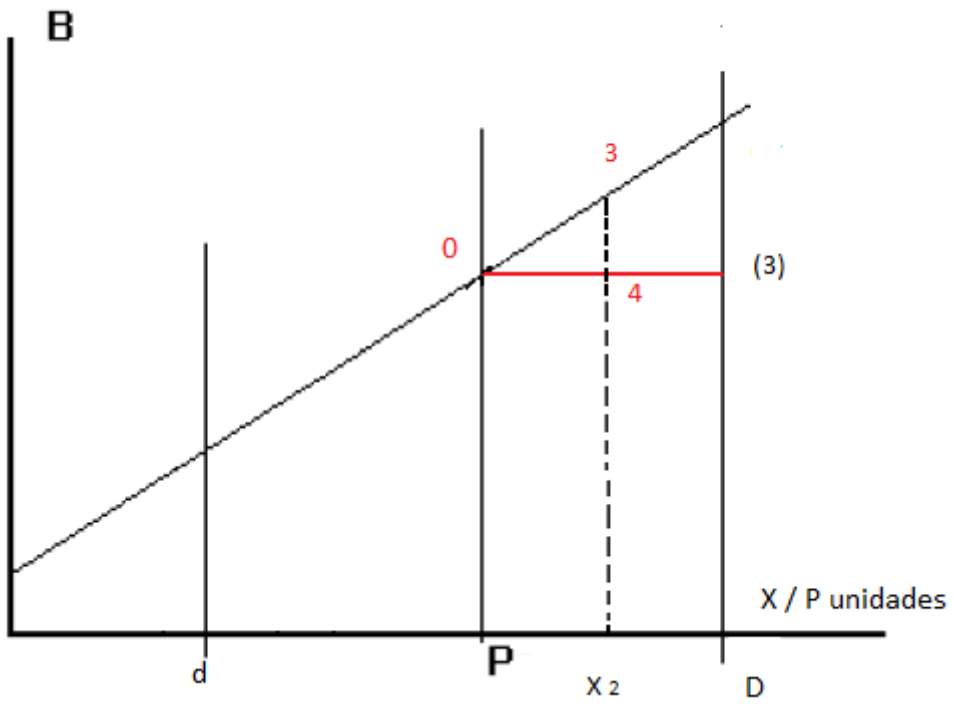
$$Arr_1 = (v_r - c)\, X_1 + P\,(v_r - c)$$

Se observa que cuando $X_1 \rightarrow d$, el Arr_1 es máximo, luego

$$Arr_{1max} = (v - c)\, d - [(v - v_r)\, d + P\,(v_r - c\,)]$$

Primer caso

Si $X \geq P$, mayor demanda que lo que se produjo, resulta el siguiente análisis
Si se supone $X_2 > P$ y se grafica como en el caso anterior



Donde el punto 3, es

$$B = (v - c)\, X_2$$

Y el punto 4, es

$$B = (v - c + c_s)\, P - c_s\, X_2 \qquad \text{Ecuación 2}$$

Aplicando Savage y observando que el arrepentimiento es máximo cuando $X_2 \rightarrow D$, se puede escribir

$$Arr_{2\ max} = (v - c)\, D - [(v - c + c_s)\, P - c_s\, D]$$

El modelo plantea que es posible arrepentirse por defecto (Arr_2) o por exceso (Arr_1). La solución óptima aparece en el punto que iguala ambos arrepentimientos, dado que el arrepentimiento relativo es cero.

$$Arr_1\ max = Arr_2\ max$$

luego

$$(v - c)\, d - [(v - v_r)\, d + P\,(v_r - c)] = (v - c)\, D - [(v - c + c_s)\, P - c_s\, D]$$

Reagrupando resulta

$$(v - c)\, d - (v - v_r)\, d - (v - c)\, D + c_s\, D = (v_r - c)\, P - (v - c + c_s)\, P$$

Donde

$$P = \frac{[(v - c) - (v - vr)]\,.\,d + [cs - (v - c)].\,D}{(vr - c) - (v - c + cs)}$$

Finalmente, P óptimo con criterio de Savage, resulta

$$P = \frac{(vr - c).\,d + (cs - v + c).\,D}{vr - v + cs}$$

(Ecuación 3)

Aplicación práctica

Seguimos con el mismo caso de un producto perecedero que tiene un período de vida útil y fecha de vencimiento. Como es nuevo en el mercado, hay antecedentes o cualquier otra información acerca de la demanda que tendrá, tampoco es posible establecer un análisis subjetivo acerca de la naturaleza aleatoria de la demanda.

El precio de venta unitario (v) es 150 $/un y el costo total unitario (c) es de 70 $/un. El producto NO vendido se destina a un mercado secundario con un valor residual (v_r) de 20 $/un. También se decide penalizar la demanda insatisfecha con un costo de escasez (c_s) de 10 $/un.

Dada la naturaleza del problema se pide determinar el **lote optimo de producción** que satisface la decisión optima de acuerdo con el criterio de Savage, bajo condiciones de incertidumbre, y con la variable *demanda del producto* que es aleatoria discreta, y está comprendida entre un mínimo de 100 un y máximo de 800 unidades para ese mercado específico.

Entonces

v = 150 $/un	c = 70 $/un
v_r = 20 $/un	c_s = 10 $/un
d = 100 un	D = 800 un

$$P = \frac{(vr - c) \cdot d + (cs - v + c) \cdot D}{vr - v - cs}$$

$$P = \frac{(20 - 70) \times 100 + (10 - 150 + 70) \times 800}{20 - 150 - 10}$$

$$P = \frac{-61000}{-140} = 435\ un$$

Lote óptimo de producción = 435 Unidades

Esta es una herramienta que permite rápidamente determinar en un escenario de incertidumbre un valor óptimo según el criterio de Savage.

Uso de hoja de cálculo

La manera de incentivar al estudiante para utilizar esta herramienta que resultó apta consiste en realizar estos procedimientos:

a) Utilizar una hoja de cálculo (en este caso Excel) para implementar una matriz de arrepentimientos, necesariamente limitada en tamaño, la que llamaremos "matriz física"
b) Calcular el nivel de producción P sin limitaciones de tamaño empleando la Ecuación 3
c) Generar en un libro (en este caso de Excel) una macro en *Visual Basic for Aplications* (VBA) que sea capaz de resolver matrices cuadradas de hasta 10000 filas.

Los resultados de ambos métodos pueden ser contrastados con la propuesta algebraica presentada más arriba.

Construcción de una matriz física

En este caso se procede a construir una matriz de ejemplo de 5 x 5 siendo la superior de

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1												demanda Xj			
2		V>C>VR — V>C>CS							Producc Pi	1	2	3	4	5	
3		Valor de venta	V	20		Borrador			1	5	-4	-13	-22	-31	
4		Valor residual	VR	10					2	0	10	1	-8	-17	
5		Costo	C	15					3	-5	5	15	6	-3	
6		Costo escasez	CS	9		Calcular			4	-10	0	10	20	11	
7		Demanda minima	d --> Xj	1					5	-15	-5	5	15	25	
8		Demanda Máxima	D --> Pi	5											
9															
10									MaxB(j)==>	5	10	15	20	25	
11															
12	MAXIMOS BENEFICIOS POR FILA		P(i)max	25	i=	5	(Producc)		1	0	14	28	42	56	
13									2	5	0	14	28	42	
14									3	10	5	0	14	28	
15	MINIMO ARREPENTIMIENTO		Arr(i)min	15	i=	4	(Producc)		4	15	10	5	0	14	
16	(por software)								5	20	15	10	5	0	
17	CALCULO TEORICO CON CS					-1	(Producc)								
18						44,625									
19	CALCULO TEORICO SIN CS					2	(Producc)								
20									4,00						
21	CALCULO DIRECTO MATRICIAL					4									
22										0	14	28	42	56	
23										5	0	14	28	42	
24										10	5	0	14	28	
25										15	10	5	0	14	
26										20	15	10	5	0	
27															

ellas la matriz básica de beneficios calculada a partir de una diagonal que representa los valores X < P que estarán en la diagonal y en las celdas por debajo de ella (fig 2) y los valores X > P en las celdas por encima de la diagonal (fig 2). Estos valores son calculados según la ecuación 1, para el primer caso

$$B = (v - v_r)\, X + (v_r - c)\, P$$

y con la ecuación 2, en el segundo

$$B = (v - c + c_s)\, P - c_s\, X_2$$

=(V-VR)*J2+(VR-COSTO)*I3

I	J	K	L	M	N
	demanda Xj				
Producc Pi	1	2	3	4	5
1	5	-4	-13	-22	-31
2	0	10	1	-8	-17
3	-5	5	15	6	-3
4	-10	0	10	20	11
5	-15	-5	5	15	25

Figura 1

=(V-COSTO)*I3-CS*(K$2-$I$3)

I	J	K	L	M	N	O
	demanda Xj					
ducc Pi	1	2	3	4	5	
1	5	-4	-13	-22	-31	
2	0	10	1	-8	-17	
3	-5	5	15	6	-3	
4	-10	0	10	20	11	
5	-15	-5	5	15	25	

Figura 2

Lo que sigue, es un vector MaxB(i,j) que se utiliza para mostrar los máximos beneficios de la matriz anterior calculando el máximo en cada demanda, según se ve en la figura 3.

Luego, (Figura 4) se construye la matriz de arrepentimientos, restando en cada celda al máximo beneficio correspondiente a esa demanda el beneficio real dado por la producción realizada.

=MAX(J3:J7)

I	J	K	L	M	N	O
	demanda Xj					
Producc Pi	1	2	3	4	5	
1	5	-4	-13	-22	-31	
2	0	10	1	-8	-17	
3	-5	5	15	6	-3	
4	-10	0	10	20	11	
5	-15	-5	5	15	25	
MaxB(j)==>	(J3:J7)	10	15	20	25	

Figura 3

J13 =J10-J4

	H	I	J	K	L	M	N	O
1			demanda Xj					
2		Producc Pi	1	2	3	4	5	
3		1	5	-4	-13	-22	-31	
4		2	0	10	1	-8	-17	
5		3	-5	5	15	6	-3	
6		4	-10	0	10	20	11	
7		5	-15	-5	5	15	25	
8								
9								
10		MaxB(j)==>	5	10	15	20	25	
11								
12		1	0	14	28	42	56	
13		2	5	0	14	28	42	
14		3	10	5	0	14	28	
15		4	15	10	5	0	14	
16		5	20	15	10	5	0	
17								

Figura 4

Los últimos pasos consistirán en construir un vector columna para señalar los máximos arrepentimientos en cada nivel de producción elegido, mediante la función

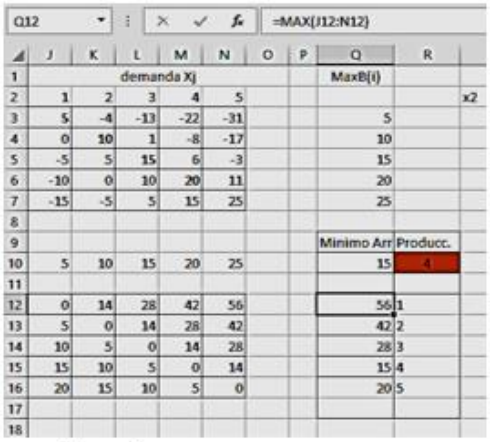
=MAX(fila-de-matriz-de-arrepentimiento)

consignando el valor de cada P a la derecha del vector. (Figura 5)

Con ello se podrá calcular el mínimo de esos máximos y la producción a la que corresponde, mediante las funciones

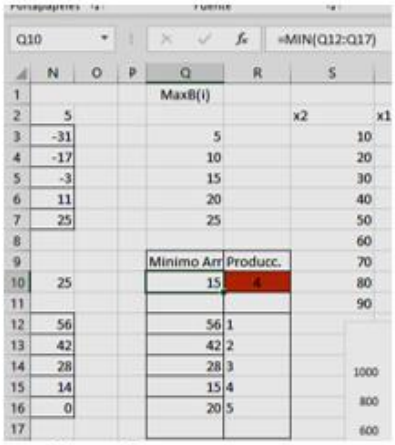
=MIN(vector-de-máximos-arrepentimientos)

=MIN(Q12:Q17) (Figura 6)



	J	K	L	M	N	O	P	Q	R	
1			demanda Xj					MaxB(i)		
2	1	2	3	4	5					x2
3	5	-4	-13	-22	-31			5		
4	0	10	1	-8	-17			10		
5	-5	5	15	6	-3			15		
6	-10	0	10	20	11			20		
7	-15	-5	5	15	25			25		
8										
9								Minimo Arr	Producc.	
10	5	10	15	20	25			15	4	
11										
12	0	14	28	42	56			56	1	
13	5	0	14	28	42			42	2	
14	10	5	0	14	28			28	3	
15	15	10	5	0	14			15	4	
16	20	15	10	5	0			20	5	
17										
18										

Figura 5

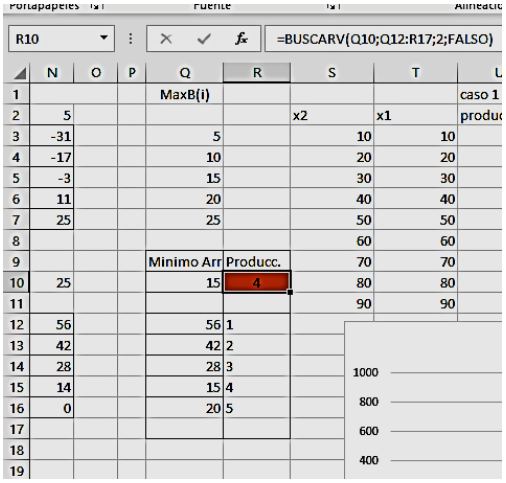


	N	O	P	Q	R	S
1				MaxB(i)		
2	5					x2 x1
3	-31			5		10
4	-17			10		20
5	-3			15		30
6	11			20		40
7	25			25		50
8						60
9				Minimo Arr	Producc.	70
10	25			15	4	80
11						90
12	56			56	1	
13	42			42	2	
14	28			28	3	1000
15	14			15	4	800
16	0			20	5	600
17						

Figura 6

=BUSCARV(valor-mínimo-hallado [en] vector-de-máximos-arrepentimientos-y-columna-a-la-derecha [y escribir] la segunda columna)

=BUSCARV(Q10;Q12:R17;2;FALSO) (Figura 7)

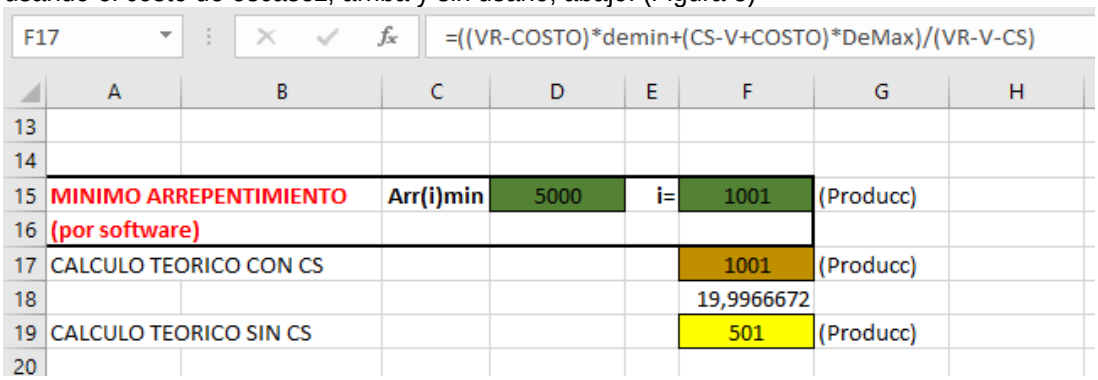


	N	O	P	Q	R	S	T	U
1				MaxB(i)				caso 1
2	5					x2	x1	produc
3	-31			5		10	10	
4	-17			10		20	20	
5	-3			15		30	30	
6	11			20		40	40	
7	25			25		50	50	
8						60	60	
9				Minimo Arr	Producc.	70	70	
10	25			15	4	80	80	
11						90	90	
12	56			56	1			
13	42			42	2			
14	28			28	3			
15	14			15	4			
16	0			20	5			
17								
18								
19								

Figura 7

Cálculo teórico utilizando la ecuación 3

Para implementar este parte simplemente habilitamos un sector de la hoja con este cálculo, usando el costo de escasez, arriba y sin usarlo, abajo. (Figura 8)



F17 =((VR-COSTO)*demin+(CS-V+COSTO)*DeMax)/(VR-V-CS)

	A	B	C	D	E	F	G	H
13								
14								
15	MINIMO ARREPENTIMIENTO		Arr(i)min	5000	i=	1001	(Producc)	
16	(por software)							
17	CALCULO TEORICO CON CS					1001	(Producc)	
18						19,9966672		
19	CALCULO TEORICO SIN CS					501	(Producc)	
20								

Figura 8

La razón de este doble cálculo se explica en las Conclusiones, más adelante.

Generación de un programa (macro) en VBA capaz de resolver matrices de hasta 10000 columnas.

El último paso será construir un programa en VBA que permita construir matrices cuadradas de hasta 10000 columnas. (se puede solicitar el código en optimiza.org)

Como complemento se agregaron, en la misma hoja de cálculo utilizada para resolver la matriz anterior dos "botones" destinados a correr el programa y a borrar los datos ingresados, respectivamente. El detalle de la parte de hoja destinada a la macro es el que se ve en la figura siguiente

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1										demanda Xj				
2		V>C>VR — V>C>CS							Producc Pi	1	2	3	4	5
3		Valor de venta	V	20		Borrador			1	5	-4	-13	-22	-31
4		Valor residual	VR	10					2	0	10	1	-8	-17
5		Costo	C	15					3	-5	5	15	6	-3
6		Costo escasez	CS	9		Calcular			4	-10	0	10	20	11
7		Demanda minima	d --> Xj	1					5	-15	-5	5	15	25
8		Demanda Máxima	D --> Pi	5										
9														
10									MaxB(j)==>	5	10	15	20	25
11														
12	MAXIMOS BENEFICIOS POR FILA		P(i)max	25	i=	5	(Producc)		1	0	14	28	42	56
13									2	5	0	14	28	42
14									3	10	5	0	14	28
15	MINIMO ARREPENTIMIENTO		Arr(i)min	15	i=	4	(Producc)		4	15	10	5	0	14
16	(por software)								5	20	15	10	5	0
17	CALCULO TEORICO CON CS					-1	(Producc)							
18						44,625								
19	CALCULO TEORICO SIN CS					2	(Producc)							
20									4,00					
21	CALCULO DIRECTO MATRICIAL					4								
22										0	14	28	42	56
23										5	0	14	28	42
24										10	5	0	14	28
25										15	10	5	0	14
26										20	15	10	5	0

Como se ve, hay un sector de carga de datos (tabla superior izquierda) y renglones con resultados y comparaciones con el método matricial (si es posible por el tamaño de la matriz). La matriz inferior derecha (en verde) está construida con la macro y sirve para comparar el funcionamiento con la matriz calculada en la primera parte de este artículo..

Resultados y conclusiones

Los resultados del trabajo son satisfactorios. Se dispone de una herramienta capaz de resolver con seguridad casos en los cuales el software disponible no logra resultados. Se han encontrado algunas fallas menores en el cálculo de casos con costo de escasez, donde hay divergencia de resultados probablemente originadas en un cálculo algebraico redundante y que está siendo analizado.

Este desarrollo servirá para:

- incorporar una más a las herramientas ya construidas por el equipo en hoja de cálculo (inventarios multimodal, pronóstico de rentabilidad bajo incertidumbre, gestión de proyectos PERT-CPM, y otras).

- afianzar la enseñanza avanzada de manejo de hojas de cálculo y programación en los estudiantes de ingeniería como herramienta profesional de aplicación directa
- posibilidades didácticas mayores al poder mostrar en exposiciones teóricas como se resuelven los casos complejos planteados.

Para quienes estén interesados ofrecemos, además del código en lenguaje VBA de la macro utilizada, el libro utilizado. Se puede solicitar utilizando los medios de contacto que figuran en www.optimiza.org

RECREO

En esta sección queremos presentarte, a manera de recreo, algunos desafíos matemáticos simples para pensar, aprender y divertirnos un rato.

Desafíos

Desafío 1 – ¿Demostración brillante o trampa oculta?

En una charla entre colegas aparece la siguiente "demostración" sorprendente:

Supongamos que a y b son números reales no nulos tales que $a = b$.

1. $a = b$
2. $a^2 = ab$
3. $a^2 - b^2 = ab - b^2$
4. $(a - b)(a + b) = b(a - b)$
5. Simplificando el factor común $(a - b)$ en ambos lados:

$$a + b = b$$

6. Como $a = b$, entonces:

$$b + b = b \Rightarrow 2b = b \Rightarrow 2 = 1.$$

Conclusión: ¡hemos demostrado que $2 = 1$!

¿En qué paso está el error? ¿Por qué la "demostración" no es válida?

Esperamos comentarios.

Desafío 2 – El test de calidad engañoso

Una planta fabrica piezas, y se sabe que aproximadamente el 1 % de las piezas producidas es defectuosa.

Se dispone de una prueba de control de calidad que tiene las siguientes características:

- Si la pieza es defectuosa, el test la detecta como defectuosa el 98 % de las veces. (Sensibilidad = 98 %)
- Si la pieza es buena, el test la indica incorrectamente como defectuosa el 5 % de las veces.
  (Falsos positivos = 5 %, es decir, especificidad = 95 %)

Se toma una pieza al azar de la producción y el test indica: “defectuosa”.

1. ¿Cuál es la probabilidad de que la pieza esté realmente defectuosa dado que el test salió positivo?
2. ¿Por qué esa probabilidad puede ser mucho menor de lo que la intuición sugiere?