PROGRAMA DE DIRECCION GERENCIAL PDG' 2001

CURSO: MATEMATICA APLICADA A LAS DECISIONES

TEMA: "Elementos de la Teoría de Probabilidades. Problemas, Trabajos Prácticos y Casos"

Domingo Alberto Tarzia

Rosario (ARGENTINA), Abril 2001

Facultad de Ciencias Empresariales
UNIVERSIDAD AUSTRAL

Departamento de Matemática
Facultad de Ciencias Empresariales
Universidad Austral
Paraguay 1950, S2000FZF Rosario, ARGENTINA.
TEL.: (0341)-4814990 Int. 137 ; FAX: (0341)-4810505
E-Mail: tarzia@uaufce.edu.ar

PROGRAMA DE DIRECCIÓN GERENCIAL PDG'2001

CURSO: MATEMÁTICA APLICADA A LAS DECISIONES

TEMA: "ELEMENTOS DE LA TEORÍA DE PROBABILIDADES. PROBLEMAS, TRABAJOS PRÁCTICOS Y CASOS"

Domingo Alberto TARZIA

Rosario (ARGENTINA)

Abril 2001

RESUMEN.

Se presentan las definiciones y propiedades básicas de la teoría de probabilidades a los efectos de introducir el pensamiento matemático a través de las tablas de contingencia, diagramas de árboles y criterios para la toma de decisiones en problemas no determinísticos.

NOTA.

El presente texto ha servido de apoyo para desarrollar actividades en el tema "Elementos de la teoría de probabilidades. Problemas, trabajos prácticos y casos" como parte del curso "Matemática aplicada a las decisiones" del Programa de Dirección Gerencial PDG dictado en la Facultad de Ciencias Empresariales de la Universidad Austral (sede Rosario) desde el año 1998.

En el presente módulo se presentan:

- un resumen de las definiciones y propiedades básicas en la teoría de la probabilidad [Li, Me, Sp];
- la matemática de las distribuciones de probabilidades mas frecuentes (variables aleatorias discretas: binomial y Poisson; variables aleatorias continuas: normal y exponencial);
- actividades sobre probabilidad condicionada y tablas de contingencia (diagramas de árbol), probabilidad geométrica y probabilidad dinámica [CLGGLM];
- problemas, criterios determinísticos y no determinísticos para la toma de decisiones, trabajos prácticos y casos no clásicos sobre el tema [CLGGLM, Kr, Li, Me, Mi, SpBo, Sp, Th];
- un modelo de fidelidad de marca para marketing y sus consecuencias matemáticas [Ta1];
- las referencias básicas de la bibliografía utilizada.

INDICE

I. INTRODUCCIÓN A LA PROBABILIDAD.

Ejemplos de experimentos no determinísticos.

Se tratará de entender lo que se considera un experimento "aleatorio" o "no determínistico". (Más precisamente se darán ejemplos de fenómenos para los cuales los modelos no determínisticos son apropiados. Esta es una distinción que el lector deberá mantener presente. Así se refieren frecuentemente a experimentos no determínisticos o aleatorios, cuando de hecho se está hablando de un modelo no determínistico para un experimento.) No se pretende dar una definición precisa de diccionario para este concepto. En su lugar se darán numerosos ejemplos que la ilustran.

E_1: Se lanza un dado y se observa el número que aparece en la cara superior.
E_2: Se lanza una moneda cuatro veces y se cuenta el número total de caras obtenidas.
E_3: Se lanza una moneda cuatro veces y se observa la sucesión de caras y sellos obtenidos.
E_4: Se fabrica una lámpara eléctrica. Luego se prueba su duración conectándola en un portalámparas y se cuenta el tiempo transcurrido (en horas) hasta que se quema.
E_5: Se fabrican artículos hasta producir 10 no defectuosos. Se cuenta el número total de artículos manufacturados.
E_6: Se lanza un proyectil. Después de un tiempo determinado t, se anotan las tres componentes de la velocidad v_x, v_y, v_z.
E_7: Un termógrafo marca la temperatura continuamente en un período de 24 horas. En un sitio y en una fecha señalados, "leer" dicho termógrafo.

¿Qué tienen en común los experimentos anteriores? Los siguientes aspectos son importantes para la descripción de un *experimento aleatorio.*

(a) Es posible repetir cada experimento en forma indefinida sin cambiar esencialmente la condiciones.
(b) Aunque en general no se puede indicar cuál será un resultado *particular*, se puede describir el conjunto de *todos* los resultados *posibles* del experimento.
(c) A medida que el experimento se repite los resultados individuales parecen ocurrir en forma caprichosa. Sin embargo, como el experimento se repite un *gran* número de veces, aparece un patrón definido o regularidad. Esta regularidad hace posible la construcción de un modelo matemático preciso con el cual se analiza el experimento.

El espacio muestral.

Definición. Con cada experimento ϵ del tipo que se ha considerado, se define el *espacio muestral* como el conjunto de *todos* los resultados posibles de ϵ. Usualmente se designa este conjunto como S. (En nuestro contexto, S representa el conjunto universal descripto anteriormente).

Se considerará cada uno de los experimentos anteriores y se describirá el espacio muestral de cada uno. El espacio muestral S_i se referirá al experimento E_i.

- $S_1 = \{1, 2, 3, 4, 5, 6\}$
- $S_2 = \{0, 1, 2, 3, 4\}$

- $S_3 = \{$todas las sucesiones posibles de la forma a_1, a_2, a_3, a_4 donde $a_i = C$ ó S según si aparece cara o ceca en el i-ésimo lanzamiento$\}$.
- $S_4 = \{t \mid t \geq 0\}$.
- $S_5 = \{10, 11, 12, ...\}$.
- $S_6 = \left\{(v_x, v_y, v_z) \ / \ v_x, v_y, v_z \text{ son números reales}\right\}$.
- $S_7 =$ Este espacio muestral es el más importante de los que aquí se considera. Prácticamente se debe suponer que la temperatura en cierta localidad específica nunca puede subir o bajar con relación a ciertos valores, digamos M y m. Fuera de esta restricción, se debe admitir la posibilidad de que aparezca cualquier gráfica con determinadas carácterísticas. Es posible que ésta no tenga saltos (esto es, representará una función continua). Además, la gráfica tendrá ciertas características de suavidad que pueden resumirse en forma matemática al decir que la gráfica representa una función diferenciable. Así, finalmente se puede enunciar que el espacio muestral es

$$\left\{f \mid f \text{ es una función diferenciable, que satisface } m \leq f(t) \leq M, \text{ para todo tiempo } t\right\}.$$

Será importante analizar de nuevo el *número* de resultados en un espacio muestral. Surgen tres posibilidades: el espacio muestral puede ser finito, infinito numerable, o infinito no numerable. Refiriéndose a los ejemplos anteriores, notemos que S_1, S_2 y S_3 son finitos, S_5 es infinito numerable y S_4 , S_6 y S_7 son infinitos no numerables.

Eventos. Un evento A (respecto a un espacio muestral particular S asociado con un experimento ϵ) es un conjunto de resultados posibles. En la terminología de conjuntos, *un evento es un subconjunto del espacio muestral S*. En vista de lo expuesto previamente, esto significa que S es también un evento y también lo es el conjunto vacío $\emptyset$. Cualquier resultado individual también puede considerarse como un evento.

Los siguientes son ejemplos de eventos. Otra vez se referirá a los experimentos antes anotados: A_i se referirá a un evento asociado con el experimento E_i:

- A_1: Un número par ocurre; esto es, $A_1 = \{2, 4, 6\}$.
- A_2: $\{2\}$; es decir, ocurren dos caras.
- A_3: {CCCC, CCCS, CCSC, CSCC, SCCC}; es decir, salen más caras (C) que cecas (S).
- A_4: $\{t \mid t < 3\}$; es decir, la lámpara se quema en menos de tres horas.

Cuando el espacio muestral S es finito o infinito numerable, *todo* subconjunto se puede considerar como un evento. Sin embargo, si S es infinito no numerable, aparece una dificultad teórica. Resulta que no cualquier subconjunto concebible se puede considerar como un evento. Por razones que escapan al nivel de esta presentación, ciertos subconjuntos "no admisibles" deben ser excluidos. Por fortuna, tales conjuntos no admisibles en realidad no aparecen en las aplicaciones y, por tanto, no interesarán aquí. En lo que sigue se supondrá tácitamente que cada vez que se mencione un evento será de la clase que está permitido considerar. Se pueden usar ahora los diversos métodos para combinar conjuntos (es decir, eventos) y obtener los nuevos conjuntos (es decir, eventos) que se presentó con anterioridad:

(a) Si A y B son eventos, $A \cup B$ es el evento que ocurre si y sólo si A ó B (o ambos) ocurren;

(b) Si A y B son eventos, $A \cap B$ es el evento que ocurre si y sólo si A y B ocurren;

(c) Si A es un evento, A^c es el evento que ocurre si y sólo si A *no* ocurre;

(d) Si A_1, ..., A_n es cualquier colección finita de eventos, entonces $\bigcup_{i=1}^{n} A_i$ es el evento que ocurre si y sólo si *al menos uno* de los eventos A_i ocurre;

(e) Si A_1, ... , A_n es cualquier colección finita de eventos, entonces $\cap_{i=1}^{n} A_i$ es el evento que ocurre si y sólo si *todos* los eventos A_i ocurren simultáneamente.

(f) Si A_1, ..., A_n, ... es cualquier colección finita (numerable) de eventos, entonces $\bigcup_{i=1}^{\infty} A_i$ es el evento que ocurre si y sólo si *al menos uno* de los eventos A_i ocurre.

(g) Si A_1, ..., A_n, ... es cualquier colección infinita (numerable) de eventos, entonces $\bigcap_{i=1}^{\infty} A_i$ es el evento que ocurre si y sólo si *todos* los eventos A_i ocurren simultáneamente.

(h) Supóngase que S representa el espacio muestral asociado con un experimento ϵ y realizamos ϵ dos veces. Entonces S x S se puede utilizar para representar todos los resultados de esas dos repeticiones. Es decir, $(s_1, s_2) \in S \times S$ significa que s_1 resultó cuando se realizó ϵ la primera vez y s_2 cuando se realizó ϵ la segunda vez.

(i) Evidentemente, el ejemplo h se puede generalizar. Se considera n repeticiones de un experimento ϵ cuyo espacio muestral es S. Entonces $S \times S \times ... \times S = \{(s_1, s_2, ..., s_n) \mid s_i \in S, i = 1, ..., n\}$ representa el conjunto de todos los resultados posibles cuando ϵ se realiza n veces. En cierto sentido, S x S x ... x S es un espacio muestral en sí mismo, o sea el espacio muestral asociado con n repeticiones de ϵ.

Eventos mutuamente excluyentes.

Definición. Se dice que dos eventos, A y B son *mutuamente excluyentes* si no pueden ocurrir juntos. Se expresa esto escribiendo $A \cap B = \emptyset$; es decir, la intersección de A y B es el conjunto vacío.

Frecuencia Relativa. Se supone que se repite n veces el experimento ϵ, y sean A y B dos eventos asociados con ϵ. Sean n_A y n_B el número respectivo de veces que el evento A y el evento B ocurrieron en las n repeticiones.

Definición. Se llama *frecuencia relativa* del evento A en las n repeticiones de ϵ al número dado por $f_A = n_A/n$.

La frecuencia relativa f_A tiene las siguientes propiedades importantes, que son verificables fácilmente:

1) $0 \leq f_A \leq 1$.

2) $f_A = 1$ si y sólo si A ocurre cada vez en las n repeticiones.

3) $f_A = 0$ si y sólo si A no ocurre en las n repeticiones.

4) Si A y B son dos eventos que se excluyen mutuamente y si $f_{A \cup B}$ es la frecuencia relativa sociada al evento $A \cup B$, entonces $f_{A \cup B} = f_A + f_B$.

5) f_A, basada en las n repeticiones del experimento y considerada para una función de n, "converge" en cierto sentido probabilístico a P(A) (probabilidad del evento A) cuando $n \to +\infty$.

Probabilidad de un evento.

Definición. Sea ϵ un experimento y S un espacio muestral asociado con ϵ. Con cada evento A se asocia un número real, designado con P(A) y llamado *probabilidad de* A, el cual satisface las siguientes propiedades:

1) $0 \leq P(A) \leq 1$.

2) $P(S) = 1$.

3) Si A y B son dos eventos que se excluyen mutuamente, $P(A \cup B) = P(A) + P(B)$.

4) Si A_1, A_2, ..., A_n, ... son eventos que se excluyen mutuamente dos a dos, entonces

$$P(\cup_{i=1}^{\infty}) = P(A_1) + P(A_2) + \ldots + P(A_n) + \ldots = \sum_{n=1}^{\infty} P(A_i)$$

Teorema. (i) Si $\emptyset$ es el conjunto vacío, entonces $P(\emptyset) = 0$.

(ii) Si A^c es el evento complementario de A, entonces

$$P(A) = 1 - P(A^c)$$

(iii) Si A y B son eventos cualesquiera, entonces

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

(iv) Si A, B, y C son tres eventos cualesquiera, entonces

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C).$$

(v) Si $A \subseteq B$ entonces $P(A) \leq P(B)$.

II. EL ESPACIO MUESTRAL FINITO.

Se consideran sólo experimentos para los cuales el espacio muestral S consta de un número finito de elementos. La suposición que más comúnmente se hace para espacios muestrales finitos es que todos los resultados son igualmente probables. De ninguna manera esta suposición puede darse como un hecho; deberá justificarse con cuidado. En dicho caso, se puede definir la probabilidad de un evento de la siguiente manera:

$$P(A) = \frac{\text{número de maneras en que } \epsilon \text{ puede ocurrir favorable a A}}{\text{número de maneras en que } \epsilon \text{ puede ocurrir}}.$$

Es importante comprender que la expresión anterior de P(A) es sólo una consecuencia de la suposición de que todos los resultados son igualmente probables y sólo es aplicable cuando se satisface esta suposición. Sin duda no sirve como una definición general de probabilidad.

Métodos de enumeración.

A) Principio de multiplicación. Se supone que un procedimiento, designado como 1, puede hacerse de n_1 maneras. Se supone que un segundo procedimiento, designado como 2, se puede hacer de n_2 maneras. También se supone que cada una de las maneras de efectuar 1 puede ser seguida por cualquiera de las maneras de efectuar 2. Entonces el procedimiento que consta de 1 seguido por 2 se puede hacer de $n_1 n_2$ maneras.

Observación. Obviamente este principio puede extenderse a cualquier número de procedimientos. Si hay k procedimientos y el i-ésimo procedimiento se puede hacer de n_i maneras, $i = 1, 2, \ldots, k$ entonces el procedimiento que consiste en 1, seguido por 2 ,..., seguido por el procedimiento k puede hacerse de $n_1 n_2 ... n_1$ maneras.

B) Principio de adición. Se supone que un procedimiento, designado con 1, se puede hacer de n_1 maneras. Se supone que un segundo procedimiento, designado con 2, se puede hacer de n_2 maneras. Se supone además que *no* es posible que *ambos*, 1 y 2, se hagan juntos. Entonces, el número de maneras de como se puede hacer 1 ó 2 está dado por $n_1 + n_2$.

Observación. También este principio puede generalizarse como sigue: si hay k procedimientos y el i-ésimo procedimiento se puede hacer en n_i maneras, $i = 1, 2,..., k$, entonces el número de maneras como se puede hacer el procedimiento 1, ó el procedimiento 2 ó ... ó el procedimiento k está dado por $n_1 + n_2 + ... + n_k$, suponiendo que los procedimientos no se pueden realizar en forma conjunta.

C) Permutaciones.

a) Se supone que se tiene n objetos diferentes. ¿De cuántas maneras, se notará nPn, se pueden agrupar (permutar) estos objetos? En general, se considera el esquema siguiente. Agrupar los n objetos es equivalente a ponerlos, en algún orden específico, en una caja con n compartimientos. La primera casilla se puede llenar de cualquiera de las n maneras, la segunda de cualquiera de $(n-1)$ maneras, ..., y la última casilla de sólo una manera. Por tanto, aplicando el principio de multiplicación anterior, se ve que la caja se puede llenar de $n\,(n-1)(n-2)\,...\,1$ maneras.

Así el número de permutaciones de n objetos diferentes está dado por la expresión:

$$nPn = n!$$

b) De nuevo se considera n objetos diferentes. Esta vez se desea escoger r de esos objetos, $0 \leq r \leq n$, y se permuta el r elegido. Se indica el número de manera de hacerlo con nPr. Se recurre otra vez al esquema anterior de llenar una caja que tiene n compartimientos; ahora se detiene el proceso después que se ha llenado el compartimiento r-ésimo. Así, el primer compartimiento puede llenarse de n maneras, el segundo de $(n-1)$ maneras, ..., y el r-ésimo compartimiento de $n-(r-1)$ maneras. De este modo se puede realizar el procedimiento completo, usando de nuevo el principio de multiplicación, y se obtiene

$$n\,(n-1)(n-2)\,....\,(n-r+1)$$

maneras. Usando la notación factorial, se puede escribir

$$nPr = \frac{n!}{(n-r)!}.$$

Esta expresión también se conoce como *arreglo o variación.*

D) Combinaciones. Se considera nuevamente n objetos diferentes. Esta vez se está interesado en contar el número de maneras como se puede escoger r de esos n objetos sin considerar el orden.

Para obtener el resultado general se recuerda la fórmula derivada anteriormente: el número de maneras de elegir r objetos entre n y permutar los r elegidos es igual a $n!/(n-r)!$ Sea C el número de maneras de elegir r entre n, sin considerar el orden. (Esto es, el número buscado es C.) Se observa que una vez que se han escogido los r artículos, hay $r!$ maneras de permutarlos. Por tanto, aplicando una vez más el principio de multiplicación, junto con el resultado anterior, se obtiene la expresión:

$$C\ r! = \frac{n!}{(n-r)!}.$$

Así, el número de maneras de elegir r entre n objetos diferentes, sin considerar el orden, está dado por

$$C = \frac{n!}{r!\ (n-r)!} = \binom{n}{r}$$

que se conoce como el número combinatorio de n tomado de a r.

Se generaliza el problema anterior. Supóngase que se tiene N artículos. Si se elige n de esos al azar, sin sustitución, hay $\binom{N}{n}$ muestras posibles diferentes, todas las cuales tienen la misma probabilidad de ser escogidas. Si los N artículos están formados de r_1 A's y r_2 B's (con $r_1 + r_2 = N$) entonces la probabilidad de que los n artículos elegidos contengan exactamente s_1 A's y $(n - s_1)$ B's está dada por

$$\frac{\binom{r_1}{s_1}\binom{r_2}{n-s_1}}{\binom{N}{r}}.$$

La anterior se llama *probabilidad hipergeométrica.*

Observación: Es muy importante especificar, cuando se habla de escoger artículos al azar, si se lo hace con o sin sustitución. En una descripción más realista se propondrá esta última. Por ejemplo, cuando se inspecciona un número de artículos manufacturados con el propósito de descubrir cuántos defectuosos podría haber, en general, no se pretende inspeccionar el mismo artículo dos veces. Previamente se ha observado que el número de maneras de escoger r objetos entre n, sin considerar el orden está dado por $\binom{n}{r}$. El número de maneras de escoger r artículos entre n, con sustitución, está dado por n^r. Aquí se está interesado en el orden en que se escogieron los artículos.

E) Permutaciones cuando no todos los objetos son diferentes. En todos los métodos de enumeración presentados se ha supuesto que todos los objetos considerados eran diferentes (esto es, distinguibles). Sin embargo, no siempre este es el caso. Se supone, entonces, que se tiene n objetos tales que hay n_1 de una clase, n_2 de una segunda clase,..., n_k de una k-ésima clase, donde $n_1 + n_2 + \ldots. + n_k = n$. Entonces, el número de permutaciones de estos n objetos está dada por

$$\frac{n!}{n_1!\ n_2!\ \ldots.\ n_k!}$$

Nótese que si todos los objetos son diferentes, se tiene $n_i = 1$, $i = 1, 2, \ldots, k$, y, por tanto, la fórmula anterior se reduce a $n!$, el resultado obtenido en forma previa.

III. PROBABILIDAD CONDICIONAL E INDEPENDENCIA.

Probabilidad condicional. Sean A y B dos eventos asociados con un experimento ϵ. Se indica con $P(B/A)$ la *probabilidad condicional del evento B, dado que A ha ocurrido.* Cada vez que se calcula $P(B/A)$, esencialmente se está calculando $P(B)$ respecto al espacio muestral reducido A, en vez de espacio muestral original S. Cuando se calcula $P(B)$ se pregunta qué tan probable es que se esté en B, sabiendo que se debe estar en S, y cuando se evalua $P(B/A)$ se pregunta qué tan probable es que se esté en B, sabiendo que se debe estar en A (es decir, el espacio muestral se ha reducido de S a A.).

Definición. Se define

$$P(B/A) = \frac{P(A \cap B)}{P(A)}$$

siempre que sea $P(A) > 0$.

Definición. Se dice que los eventos $B_1, B_2, ..., B_k$ representan una partición del espacio muestral S si:

a) $B_i \cap B_j = \emptyset$, $\forall\, i \neq j$;

b) $\cup_{i=1}^{k} B_i = S$

c) $P(B_i) > 0$, $\forall$ i=1, ..., k.

En otras palabras , cuando se efectúa el exprimento ϵ, ocurre uno y sólo uno de los eventos B_i. Sea A algún evento respecto a S y sea $B_1, B_2, ..., B_k$ una partición de S, con lo cual se puede escribir:

$$A = \cup_{i=1}^{k} (A \cap B_i) = (A \cap B_1) \cup (A \cap B_2) \cup .... \cup (A \cap B_k)$$

y por ende, se tiene:

$$P(A) = \sum_{i=1}^{k} P(A \cap B_i) = P(A \cap B_1) + P(A \cap B_2) + ... + P(A \cap B_k)$$

Sin embargo, cada término $P(A \cap B_i)$ se puede expresar como $P(A/B_i)\,P(B_i)$ y, por lo tanto, se obtiene el llamado *teorema de la probabilidad total*:

$$P(A) = \sum_{i=1}^{k} P(A/B_i)\, P(B_i) = P(A/B_1)\, P(B_1) + P(A/B_2)\, P(B_2) + ... + P(A/B_k)\, P(B_k).$$

Observación. Este resultado representa una relación muy útil, ya que cuando se busca $P(A)$ frecuentemente puede ser difícil calcularlo de manera directa. Sin embargo, con la información adicional de que B_i ha ocurrido, se puede calcular $P(A/B_i)$ y por ende se puede entonces usar la fórmula anterior.

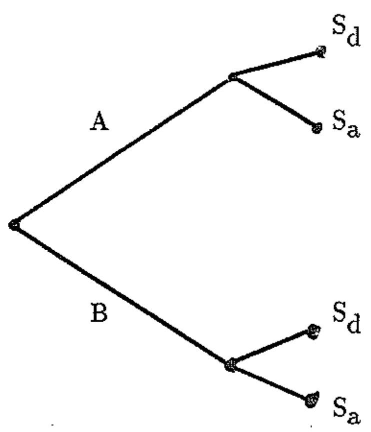
Por otro lado, cuando se necesite calcular $P(B_i/A)$ se podrá calcular esta probabilidad como una conscuencia de lo anterior. Sean $B_1, ..., B_k$ una partición del espacio muestral S y A un evento asociado con S. Aplicando la definición de probabilidad condicional, se puede escribir:

$$P(B_i/A) = \frac{P(A/B_i)\ P(B_i)}{\sum\limits_{i=1}^{k} P(A/B_i)\ P(B_i)}, \quad i = 1, \ldots\ldots, k\ .$$

Este resultado se conoce como *Teorema de Bayes.* También se le llama fórmula para la probabilidad de las "causas". Puesto que las B_i son una partición del espacio muestral, uno y sólo uno de los eventos B_i ocurre. (Esto es, uno de los eventos B_i debe ocurrir y solamente uno.) Por lo tanto, la fórmula anterior da la probabilidad de un B_i particular (esto es, una "causa"), dado que el evento A ha ocurrido. Para aplicar este teorema se debe conocer los valores de las $P(B_i)$, con $i = 1, ..., k$. Muy a menudo esos valores no son conocidos y esto limita la aplicabilidad del resultado.

La siguiente ilustración del Teorema de Bayes dará la oportunidad de introducir la idea de *diagrama de árbol*, método muy útil para analizar ciertos problemas. Supóngase que varias cajas de caramelos son de dos tipos A y B. El tipo A contiene 70% de caramelos dulces y 30% de caramelos ácidos, mientras que en el tipo B dichos porcentajes están invertidos. Más aún, supóngase que el 60% de todas las cajas de caramelos son del tipo A, mientras que el resto son del tipo B.

Ahora se está ante el siguiente problema de decisión. Usted recibe una caja de dulces de tipo desconocido. Se le permite sacar una muestra de caramelo (una situación ciertamente no real, pero que permite presentar las ideas importantes sin mucha complicación y con esta información debe decir si cree que se le ha sido ofrecido una caja del tipo A o del tipo B). El siguiente " diagrama de árbol" (llamado así por las diversas trayectorias o ramas que aparecen) ayudará a analizar el problema (S_d y S_a indican la elección de un caramelo dulce o ácido respectivamente).



Se pueden realizar algunos cálculos simples:

$$P(A) = 0{,}60\ ; \qquad P(B) = 0{,}40\ ; \qquad P(S_d/A) = 0{,}70\ ;$$

$$P(S_a/A) = 0{,}30\ ; \qquad P(S_d/B) = 0{,}30\ ; \qquad P(S_a/B) = 0{,}70\ .$$

Lo que en realidad se desea conocer son las siguientes probabilidades

$$P(A/S_d), \qquad P(A/S_a), \qquad P(B/S_d) \quad \text{y} \quad P(B/S_a).$$

Esto es, suponiendo que realmente se escogió un caramelo dulce, ¿qué decisión estaríamos más inclinados a hacer?

Para ello se deben comparar $P(A/S_d)$ y $P(B/S_d)$. Utilizando la fórmula de Bayes se tiene:

$$P(A/S_d) = \frac{P(S_d/A)\ P(A)}{P(S_d/A)\ P(A)\ +\ P(S_d/B)\ P(B)} = (0,7)(0,6)\ /\ [(0,7)(0,6) + (0,3)(0,4)] = 7/9.$$

Un cálculo similar da $P(B/S_d) = 2/9$. De este modo, con base en la evidencia que se tiene (es decir, la obtención de un caramelo dulce) es 3,50 veces más probable que se trate de una caja del tipo A que del tipo B. Por lo tanto, se decidiría, posiblemente, que el caramelo se obtuvo de una caja tipo A. (Por supuesto, se podría estar equivocado. Lo interesante del análisis anterior es que se elige la alternativa que parece más probable con base en los pocos datos que se tiene).

En términos del diagrama de árbol, lo que realmente se necesitaba (y se hizo) en los cálculos precedentes fue un análisis "hacia atrás". Esto es, dado lo que se observa, en este caso S_d, ¿qué tan probable era escoger el tipo A?

Eventos independientes. Dos eventos A y B que no pueden ocurrir de manera simultánea verifican que $A \cap B = \emptyset$. Tales eventos se designaron mutuamente excluyentes. Si A y B son mutuamente excluyentes entonces $P(A/B)=0$, porque la ocurrencia de B impide la ocurrencia de A. Por otra parte, para el caso particular $A \subseteq B$ se tendrá que $P(B/A)=1$.

En algunos de los casos anteriores, sabiendo que B ocurrió, se tiene una información precisa sobre la probabilidad de la ocurrencia de A. Sin embargo, hay muchos casos en los cuales se sabe que si un evento B ocurre, no tiene influencia alguna en la ocurrencia o no ocurrencia de otro evento A.

Definición. Se dice que A y B son dos *eventos independientes* si y sólo si

$$P(A \cap B) = P(A)\ P(B)\ .$$

Observación. Esta definición es esencialmente equivalente a la que antes se sugirió, es decir, que A y B son independientes si $P(B/A)=P(B)$ y $P(A/B) = P(A)$. Esta última forma es un poco más intuitiva, porque afirma precisamente lo que se ha estado tratando de decir antes: A y B son independientes si el conocimiento de la ocurrencia de A no influye de modo alguno en la probabilidad de ocurrencia de B.

Definición. Se dice que los tres eventos, A, B y C, son mutuamente independientes si y sólo si todas las condiciones siguientes se satisfacen:

$$P(A \cap B) = P(A)\ P(B)\ , \qquad P(A \cap C) = P(A)\ P(C)\ ,$$
$$P(B \cap C) = P(B)\ P(C)\ , \qquad P(A \cap B \cap C) = P(A)\ P(B)\ P(C)\ .$$

Finalmente se generaliza esta noción a n eventos a través de la siguiente definición.

Definición. Los n eventos $A_1, A_2, ..., A_n$ son mutuamente independientes si y sólo si se tiene

$$P(A_{i_1} \cap A_{i_2} \cap ..... \cap A_{i_k}) = P(A_{i_1})\ P(A_{i_2})\ ....\ P(A_{i_k})\ ,\ \forall\ k=2, 3, ..., n\ .$$

IV. VARIABLES ALEATORIAS UNIDIMENSIONALES.

Definición. Sea ϵ un experimento y S el espacio muestral asociado con él. Una función X que asigna a cada uno de los elementos $s \in S$, un número real X (s), se le llama *variable aleatoria.*

Definición. Sea X una variable aleatoria. Si el número de valores posibles de X (es decir su recorrido) es finito o infinito numerable, se llama a X una *variable aleatoria discreta.* Esto es, se pueden anotar los valores posibles de X como $x_1, x_2, ..., x_n, ...$. En el caso finito, la lista termina y en el caso infinito numerable, la lista continúa indefinidamente.

Definición. Sea X una variable aleatoria discreta, es decir que el recorrido de X consta, a lo más, de un número de valores, $x_1, x_2, ..., x_n, ...$ que es infinito numerable. Con cada resultado posible x_i se asocia un número $p(x_i) = P(X=x_i)$, llamado *probabilidad de* x_i. Los números $p(x_i)$, i=1,2, ... deben satisfacer las condiciones siguientes:

a) $p(x_i) \geq 0 \; , \forall \, i \; ,$

b) $\sum_{i=1}^{\infty} p(x_i) = 1$

La función p que antes se definió, se llama *función de probabilidad (o función de probabilidad puntual) de la variable aleatoria X.* La colección de pares $(x_i \, , p(x_i))$, i=1, 2, ... , algunas veces se llama distribución de probabilidad de X.

Definición. Se considera un experimento ϵ y sea A un evento asociado con ϵ. Se supone que $P(A) = p$, y por lo tanto, $P(A^c) = 1 - p$. Se considera n repeticiones independientes de ϵ. Por lo tanto, el espacio muestral consiste en todas las sucesiones posibles $\{a_1, a_2, ..., a_n\}$, donde cada a_i es A ó A^c, según A ó A^c ocurra en la i-ésima repetición de ϵ (hay 2^n de tales sucesiones). Aún más, se supone que $P(A) = p$ es el mismo para todas las repeticiones. Se define la variable aleatoria X como sigue: X es igual al número de veces que ocurrió el evento A. Se llama a X una variable aleatoria binomial con los parámetros n y p. Sus valores posibles obviamente son el 0, 1, 2, ..., n. (Decimos en forma equivalente que X tiene una *distribución binomial*). Las repeticiones individuales de E se llamarán ensayos de Bernoulli.

Definición. Se dice que X es una *variable aleatoria continua*, si existe una función f, llamada *función de densidad de probabilidad* (fdp) de X, que satisface las siguientes condiciones:

a) $f(x) \; \geq \; 0 \; , \forall \, x,$

b) $\int_{-\infty}^{+\infty} f(x) \, dx = 1$

c) Se tiene que $P(a \leq X \leq b) = \int_a^b f(x) \, dx \, , \forall \, a, \, b \in \mathbb{R}$, tal que $a < b$.

Definición. Sea X una variable aleatoria, discreta o continua. Se llama *función de distribución acumulativa de la variable aleatoria X* (abreviada fda) a la función dada por la expresión

$$F(x) = P(X \leq x) .$$

Teorema. a) Si X es una variable aleatoria discreta, entonces F viene dada por

$$F(x) = \sum_i p(x_i) ,$$

donde la suma se toma sobre todos los índices i que satisfacen $x_i \leq x$.

b) Si X es una variable aleatoria continua con función de probabilidad f, entonces F viene dada por

$$F(x) = \int_{-\infty}^{x} f(s)\, ds .$$

Más aún:

i) La función F no es decreciente. Esto es, si $x_1 \leq x_2$ entonces se tiene que $F(x_1) \leq F(x_2)$.

ii) Se tienen los siguientes límites:

$$F(-\infty) = 0 , \qquad F(\infty) = 1 .$$

c) Sea F la función de distribución acumulativa de una variable aleatoria continua con función de distribución f. Entonces, se tiene que f es la derivada de F, es decir:

$$f(x) = \frac{dF}{dx}(x) ,$$

para todo x en el cual F es diferenciable.

d) Sea X una variable aleatoria discreta con valores posibles $x_1, x_2, \ldots$, y se supone que es posible rotular dichos valores de modo que $x_1 < x_2 < \ldots$. Sea F la función de distribución acumulativa de X. Entonces,

$$p(x_i) = P(X = x_i) = F(x_i) - F(x_{i-1}) .$$

Observación. Una advertencia sobre la terminología puede ser de utilidad. Esta terminología, aunque no es nuy uniforme, ha llegado a estandarizarse. Cuando se habla de la distribución de probabilidades de una variable aleatoria X se indica su fdp f si X es continua, o su función de probabilidad puntual p definida para $x_1, x_2, \ldots$ si X es discreta. Cuando se habla de la función de distribución acumulativa, o algunas veces sólo de la función de distribución, siempre se refiere a F, donde $F(x) = P(X \leq x)$.

V. CARACTERÍSTICAS DE LAS VARIABLES ALEATORIAS.

Valor esperado o esperanza matemática.

Definición. Sea X una variable aleatoria discreta con valores posibles $x_1, \ldots, x_n, \ldots$, y sea $p(x_i) = P(X = x_i)$, $i = 1, 2, \ldots, n, \ldots$. Entonces el *valor esperado* de X (o *esperanza matemática* de X) que se denota con E(X), se define como

$$E(X) = \sum_{i=1}^{\infty} x_i \, p(x_i)$$

si la serie $\sum_{i=1}^{\infty} x_i \, p(x_i)$ converge absolutamente, es decir si $\sum_{i=1}^{\infty} |x_i| \, p(x_i) < \infty$. Este número también se designa como *valor promedio* de X.

Definición. Sea X una variable aleatoria continua con función de probabilidad f. El *valor esperado* de X se define como

$$E(X) = \int_{-\infty}^{+\infty} x\, f(x)\, dx$$

Nuevamente puede suceder que esta integral (impropia) no converja. Por lo tanto, se dice que E(X) existe si y sólo si la siguiente integral

$$\int_{-\infty}^{+\infty} |x|\, f(x)\, dx$$

es finita.

Varianza.

Definición. Sea X una variable aleatoria. Se define la *varianza* de X, que se denota con Var(X), V(X) ó σ_X^2 como sigue:

$$\sigma_X^2 = V(X) = E\Big([X - E(X)]^2\Big).$$

La raíz cuadrada positiva de V(X) se llama *desviación estándar* de X y se designa con σ_X.

Teorema. Sea X una variable aleatoria. La varianza de X puede expresarse como

$$\sigma_X^2 = V(X) = E(X^2) - [E(X)]^2 .$$

Desigualdad de Chebyshev. Sea X una variable aleatoria con $E(X) = \mu$ y sea c un número real cualquiera. Entonces, si $E(X-c)^2$ es finita y $\in$ es cualquier número positivo, se tiene

$$P[|X - c| \geq \in] \leq \frac{1}{\in^2} E(X-c)^2.$$

Las formas siguientes, equivalentes a la desigualdad anterior, son inmediatas:

(a) Al considerar el evento complementario se obtiene:

$$P[|X - c| < \in] \geq 1 - \frac{1}{\in^2} E(X-c)^2.$$

(b) Al elegir $c = \mu$ se obtiene

$$P[|X - \mu| < \in] \geq \frac{\mathrm{Var}(X)}{\in^2}.$$

(c) Al elegir $c = \mu$ y $\in = k\sigma$, donde $\sigma^2 = \mathrm{Var}(X) > 0$, se obtiene

$$P[|X - c| \geq k\sigma] \geq k^{-2}.$$

Esta última forma indica especialmente cómo la varianza mide el "grado de la concentración" de la probabilidad próxima a $E(X) = \mu$.

VI. MATEMÁTICA DE LAS DISTRIBUCIONES DE PROBABILIDAD

A. Distribuciones discretas:

1) Distribución binomial:

Se define :

$$P(r) \equiv \binom{n}{r} p^r q^{n-r}$$

como la probabilidad de que sucedan r éxitos en n pruebas donde :

$r \in \{0,1,2,\ldots,n\}$: variable aleatoria que representa el número de aciertos ;
$n \in \mathbb{N}$: tamaño de la muestra o número de lanzamientos ;
$p \in (0,1)$: probabilidad de un acierto ;
$q=1-p \in (0,1)$: probabilidad de un fracaso ;
parámetros : n , p ; $\binom{n}{r} = \frac{n!}{r!\,(n-r)!}$: número combinatorio .

Verifique que se tienen las siguientes relaciones :

(i) P determina una distribución de probabilidad pues :

$$\sum_{r=0}^{n} P(r)=1 \quad \text{(Utilice el binomio de Newton)} .$$

(ii) La media aritmética, esperanza o momento de orden 1 está dada por :

$$E \equiv \sum_{r=0}^{n} r\, P(r) = np .$$

(iii) El momento de orden 2 está dado por :

$$\sum_{r=0}^{n} r^2\, P(r)=np(q+np) .$$

(iv) La varianza está dada por :

$$\sum_{r=0}^{n} (r-E)^2\, P(r) = \sum_{r=0}^{n} r^2\, P(r) - E^2 = npq$$

y la desviación estándar por $\sqrt{npq}$.

(v) Si se considera que $np=m \in \mathbb{R}^+$ (fijo) se obtiene el siguiente límite :

$$\lim_{\substack{n \to +\infty \\ np=m}} P(r) = \frac{m^r e^{-m}}{r!} \quad \text{(Ver (2))} .$$

2) Distribución de Poisson:

Se define

$$P(x) \equiv \frac{e^{-m}\, m^x}{x!}$$

como la probabilidad de que sucedan x eventos por unidad de medida donde :

$x \in \mathbb{N} \cup \{0\}$: variable aleatoria que representa el número de ocurrencias por unidad de medida;
$m \in \mathbb{N}$: número promedio de ocurrencias por unidad de medida ;
parámetro : m .

Verifique que se tienen las siguientes relaciones :

(i) P determina una distribución de probabilidad pues :

$$\sum_{k=0}^{+\infty} P(k) = 1 \qquad \text{(Utilice la serie de potencias } e^t = \sum_{n=0}^{+\infty} \frac{t^n}{n!} \text{ con radio de convergencia } +\infty \text{).}$$

(ii) La media aritmética, esperanza o momento de orden 1 está dada por :

$$E \equiv \sum_{k=0}^{+\infty} k\, P(k) = m\,.$$

(iii) El momento de orden 2 está dado por :

$$\sum_{k=0}^{+\infty} k^2\, P(k) = m^2 + m\,.$$

(iv) La varianza está dada por :

$$\sum_{k=0}^{+\infty} (k-E)^2\, P(k) = \sum_{k=0}^{+\infty} k^2\, P(k) - E^2 = m$$

y la desviación estándar por $\sqrt{m}$.

B. Distribuciones continuas.

3) Distribución normal:

La función de densidad de probabilidad está dada por :

$$f(x) \equiv \frac{1}{\sigma\sqrt{2\pi}} \; e^{-\dfrac{(x-\mu)^2}{2\sigma^2}} \qquad (\text{Normal } N(\mu,\sigma))$$

donde

| x ∈ ℝ : variable aleatoria continua ;
| parámetros : $\mu \in \mathbb{R}$, $\sigma > 0$.

Verifique que se tienen las siguientes relaciones :

(i) f genera una distribución de probabilidad pues :

$$\int_{-\infty}^{+\infty} f(x)\, dx = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-t^2} dt = 1 \qquad (\text{Utilice la sustitución } t = \frac{x-\mu}{\sigma\sqrt{2}}) .$$

(ii) La media aritmética, esperanza o momento de primer orden está dada por

$$E \equiv \int_{-\infty}^{+\infty} x f(x)\, dx = \mu .$$

(iii) El momento de orden 2 está dado por :

$$\int_{-\infty}^{+\infty} x^2 f(x)\, dx = \mu^2 + \sigma^2 .$$

(iv) La varianza está dada por

$$\int_{-\infty}^{+\infty} (x - E)^2 f(x)\, dx = \int_{-\infty}^{+\infty} x^2 f(x)\, dx - E^2 = \sigma^2 ,$$

y la desviación estándar por σ.

(v) El cambio de variable $z = \frac{x-\mu}{\sigma}$ transforma la función f= f(x) (Normal $N(\mu,\sigma)$) en g=g(z) (Normal $N(0,1)$).

(vi) La gráfica de la función f=f(x) es una campana (campana de Gauss) que tiene las siguientes características (pasar de la variable x a la variable z ó t cuando sea conveniente) :

(a) Moda : $Mo \equiv \underset{x \in \mathbb{R}}{\text{Máx}} \; f(x) = f(\mu) = \frac{1}{\sigma\sqrt{2\pi}} \quad (f'(\mu) = 0 \, , \, f''(\mu) < 0)$;

(b) Mediana : $Md \equiv f(\mu) = \frac{1}{\sigma\sqrt{2\pi}}$ pues

$$\int_{-\infty}^{\mu} f(x)\,dx = \int_{\mu}^{+\infty} f(x)\,dx = \frac{1}{2}$$

(c) f es cóncava en $(-\sigma,\sigma)$, es convexa en $(-\infty,-\sigma)\cup(\sigma,+\infty)$ y tiene puntos de inflexión en $x=\mu\pm\sigma$ $(f''(\mu\pm\sigma)=0)$.

(d) Se tienen las siguientes probabilidades de que la variable aleatoria pertenezca a ciertos intervalos importantes :

$$P(|x|\leq\sigma)=P(|z|\leq 1)=\frac{1}{\sqrt{2\pi}}\int_{-1}^{1} e^{-\frac{z^2}{2}}\,dz = 0{,}6827\ ;$$

$$P(|x|\leq 1{,}645\,\sigma)=P(|z|\leq 1{,}645)=\frac{1}{\sqrt{2\pi}}\int_{-1{,}645}^{1{,}645} e^{-\frac{z^2}{2}}\,dz = 0{,}90\ ;$$

$$P(|x|\leq 1{,}96\,\sigma)=P(|z|\leq 1{,}96)=\frac{1}{\sqrt{2\pi}}\int_{-1{,}96}^{1{,}96} e^{-\frac{z^2}{2}}\,dz = 0{,}95\ ;$$

$$P(|x|\leq 2\sigma)=P(|z|\leq 2)=\frac{1}{\sqrt{2\pi}}\int_{-2}^{2} e^{-\frac{z^2}{2}}\,dz = 0{,}9545\ ;$$

$$P(|x|\leq 2{,}58\,\sigma)=P(|z|\leq 2{,}58)=\frac{1}{\sqrt{2\pi}}\int_{-2{,}58}^{2{,}58} e^{-\frac{z^2}{2}}\,dz = 0{,}99\ ;$$

$$P(|x|\leq 3\sigma)=P(|z|\leq 3)=\frac{1}{\sqrt{2\pi}}\int_{-3}^{3} e^{-\frac{z^2}{2}}\,dz = 0{,}9973\ .$$

e) Calcule el primer cuartil Q_1, el tercer cuartil Q_3, y la desviación cuartílica.

4) Distribución exponencial:

La función de densidad de probabilidad está dada por :

$$f(t) \equiv \lambda\, e^{-\lambda t} \quad ,$$

donde :

$t > 0$: variable aleatoria que representa el tiempo entre llegadas sucesivas ;
$\lambda > 0$: es la tasa promedio de llegadas (similar al parámetro m de Poisson) ;
$\frac{1}{\lambda}$: es el tiempo promedio entre llegadas
parámetro : λ .

Se puede extender la definición de la función f a todo el conjunto de números reales $\mathbb{R}$ de la siguiente manera :

$$f(t) \equiv \begin{cases} \lambda\, e^{-\lambda t} & \text{si} \quad t \geq 0 \,, \\ 0 & \text{si} \quad t < 0. \end{cases}$$

Verifique que se tienen las siguientes relaciones :

(i) f genera una distribución de probabilidad pues :

$$\int_0^{+\infty} f(t)\, dt = \int_0^{+\infty} e^{-u}\, du = 1$$

(ii) La media aritmética, esperanza o momento de orden 1 está dada por :

$$E \equiv \int_0^{+\infty} t\, f(t)\, dt = \frac{1}{\lambda} \int_0^{+\infty} u\, e^{-u}\, du = \frac{1}{\lambda} \qquad \left(\text{utilice } \int_0^{+\infty} u\, e^{-u}\, du = 1\right).$$

(iii) El momento de orden 2 está dado por

$$\int_0^{+\infty} t^2 f(t)\, dt = \frac{1}{\lambda^2} \int_0^{+\infty} u^2 e^{-u}\, du = \frac{2}{\lambda^2} \qquad \left(\text{utilice } \int_0^{+\infty} u^2 e^{-u}\, du = 2\right).$$

(iv) La varianza está dada por

$$\int_0^{+\infty} (t - E)^2 f(t)\, dt = \int_0^{+\infty} t^2 f(t)\, dt - E^2 = \frac{1}{\lambda^2}$$

y la desviación estándar por $\frac{1}{\lambda}$.

(v) Las probabilidades acumuladas "más que" están dadas por:

$$P(t>a) \equiv \int_{a}^{+\infty} f(t)\,dt = e^{-\lambda a} \quad , \quad \forall\, a \geq 0 .$$

(vi) Calcule el primer, segundo y tercer cuartil, y la desviación cuartílica.

(vi) La distribución de probabilidad exponencial es del tipo J invertida (grafique la función $f=f(t)$) .

(vii) La distribución de probabilidad exponencial no tiene memoria pues se tiene la siguiente relación para la probabilidad condicional :

$$P(t>b/t>a) \equiv \frac{P(t>b)}{P(t>a)} = P(t>(b-a)).$$

VII. TABLAS DE CONTINGENCIA Y DIAGRAMA DE ARBOL

A continuación se plantean y se analizan diversas actividades sobre:

- Probabilidad condicionada y tablas de contingencia ([CLGGLM], actividad 13, pp. 121 – 129);
- Probabilidad geométrica ([CLGGLM], actividad 16, pp. 137 – 145);
- Probabilidad dinámica ([CLGGLM], actividad 18, pp. 149 – 151);

con el objetivo de realizar diagramas de árboles en problemas no determinísticos (análogo al método de bifurcación en problemas de lógica-matemática [Ta2]) y poder tomar adecuadamente una decisión. Se resalta la resolución de problemas que tienen diagramas de árboles con una o varias ramas infinitas

Actividad Nº 13: Probabilidad condicionada y tablas de contingencia

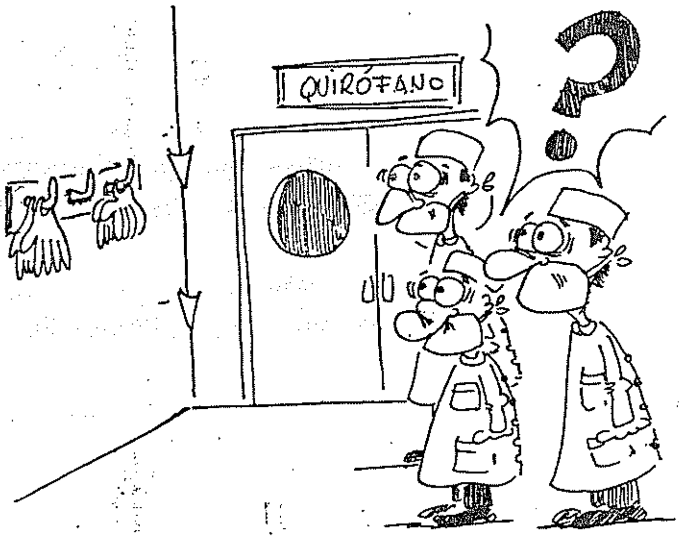
Material

No es necesario.

Tablas de contingencia

Comenzamos por dos problemas en los que se introduce un modo de presentar los datos de gran interés para el cálculo de probabilidades:

Problema 1. Diagnósticos



Una revista médica publica que para diagnosticar las lesiones de hígado existen dos procedimientos: "el histológico" y el "gráfico". El segundo procedimiento no es tan preciso como el primero, pero conlleva menor riesgo.

Para verificar la bondad del procedimiento gráfico se estudiaron con él 1.160 lesiones de hígado, comprobándose más tarde si el diagnóstico fue correcto o no.

Los datos obtenidos fueron distintos según que la lesión fuera maligna o benigna. He aquí la tabla:

TABLA 1	Diagnostico correcto	Diagnostico incorrecto	TOTALES
Lesiones malignas	418	38	456
Lesiones benignas	608	96	704
TOTALES	1.026	134	1.160

a) Interpreta el significado de cada uno de los números de la tabla.

b) Para juzgar la precisión del procedimiento es importante conocer la proporción de cada caso respecto al total de 1.160 lesiones. Expresa la tabla anterior en frecuencias relativas y en porcentajes; es decir, completa las dos tablas de la página siguiente:

TABLA 2	C	I	TOTAL
M	0'36		
B			
TOTAL			

TABLA 3	C	I	TOTAL
M	36		
B			
TOTAL			

¿Cómo son los sucesos C e I entre sí? ¿Y los M y B?

c) Se va a diagnosticar a un paciente por el procedimiento gráfico. ¿Qué probabilidades asignarás a los siguientes sucesos aleatorios

C y M; C y B; I y M; I y B; C, I?

d) ¿Qué probabilidad asignas al suceso "lesión maligna diagnosticada correctamente"? ¿Y al suceso "lesión benigna diagnosticada correctamente"? ¿Son independientes los sucesos C y M?

Para responder a estas preguntas recuerda que según vimos en la actividad número 10, si dos sucesos son independientes se verificaba que:

$$P(S1 \text{ y } S2) = P(S1) \cdot P(S2)$$

Mientras que si son dependientes se verifica que:

$$P(S1 \text{ y } S2) = P(S1) \cdot P(S2/S1)$$

de donde se tiene:

$$P(S2/S1) = P(S1 \text{ y } S2) / P(S1)$$

Problema 2. Accidentes de tráfico

Con objeto de que disminuya el número de accidentes de circulación es preciso tomar diversos tipos de medidas: campañas de información en prensa, radio y televisión; mejora de la seguridad pasiva de los vehículos; mejora del trazado y señalización de las vías (calles y carreteras); vigilancia por parte de los agentes de tráfico, etc.

Para ello es necesario disponer de información, disponer de datos. El Instituto Nacional de Estadística (INE), en su Anuario de 1975 publicó para el año 1973 los siguientes datos:

TABLA 4	En carretera	En zona urbana	TOTALES
Con victimas	34.092	32.295	66.387
Solo daños materiales	11.712	20.791	32.503
TOTALES	45.804	53.086	98.890

a) Completa la siguiente tabla de porcentajes, obtenida a partir de la anterior:

TABLA 5	A	no A	TOTALES
B	34		67
no B			33
TOTALES	46	54	100

b) Como sabes que:

$$P(B/A) = \frac{P(A \text{ y } B)}{P(A)}$$

calcula en este problema P(B/A) y P(B/no A) y decide si A y B son independientes. ¿Es P(B/A) = P(B)?.

Las tablas 1, 2, 3 del problema del diagnóstico, y las 4 y 5 del de los accidentes de tráfico reciben el nombre de **tablas de contingencia**, pues en ellas figuran todas las posibilidades, o contingencia, de los sucesos compuestos que son intersección de otros sucesos, esto es:

A y $B = A \cap B$; A y no $B = A \cap \overline{B}$; no A y $B = \overline{A} \cap B$; no A y no $B = \overline{A} \cap \overline{B}$

El esquema de dichas tablas es el siguiente:

TABLA 6	A	$\overline{A}$	Totales
B	$P(A \cap B)$	$P(\overline{A} \cap B)$	$P(B)$
$\overline{B}$	$P(A \cap \overline{B})$	$P(\overline{A} \cap \overline{B})$	$P(\overline{B})$
Totales	$P(A)$	$P(\overline{A})$	1

y pueden figurar en ellas, en lugar de probabilidades, frecuencias relativas o porcentajes.

Estas tablas además de darnos probabilidades de forma directa nos permiten el cálculo de probabilidades de sucesos condicionados; por ejemplo P(B/A) y P(B/no A).

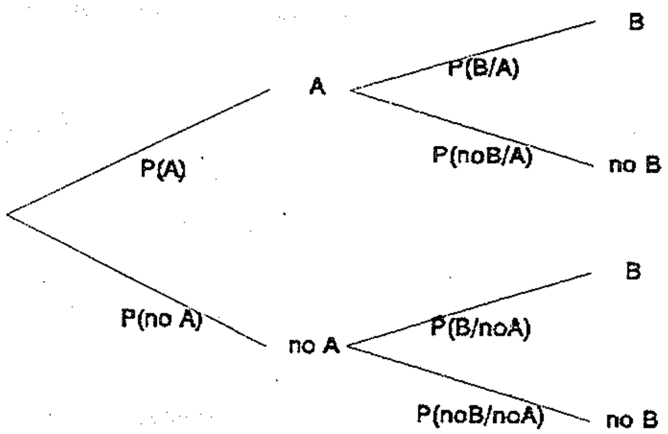
Existen tablas de contingencia más complicadas. Si consideramos que los accidentes pueden ser leves, graves o mortales; en carretera o en zona urbana, el diagrama de contingencia será:

TABLA 7	L	G	M	TOTALES
C				
U				
TOTALES				1

Lo fundamental de estas tablas es que los sucesos L, G y M sean incompatibles dos a dos y su unión sea el espacio muestral, y que los sucesos C y U también lo sean y compongan el espacio muestral.

Relación entre los diagramas de contingencia y los de árbol

La tabla 6 se puede poner en forma de diagrama de árbol de la forma siguiente:



Multiplicando las probabilidades de las distintas ramas que van hasta un punto dado, por ejemplo B, se obtiene la probabilidad de la intersección de los sucesos que se encuentran por el camino:

$$P(A \text{ y } B) = P(A) \cdot P(B/A)$$

$$P(A \text{ y no } B) = P(A) \cdot P(\text{no } B/A)$$

etc.

Los diagramas de contingencia y los de árbol están íntimamente relacionados: dado uno de ellos podemos construir el otro. Esto es muy importante, ya que unas veces los datos del problema permiten construir rápidamente uno de ellos y a partir de él podemos construir el otro diagrama, que nos dará la respuesta.

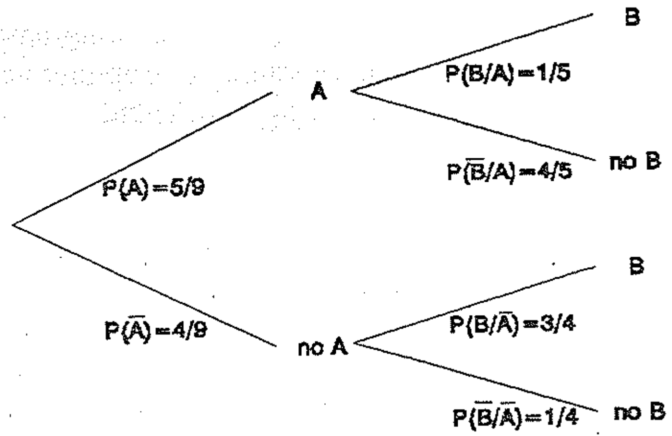
Teniendo en cuenta que:

$$P(B/A) = P(A \text{ y } B) / P(A) \qquad [1]$$

y dado el diagrama de contingencia:

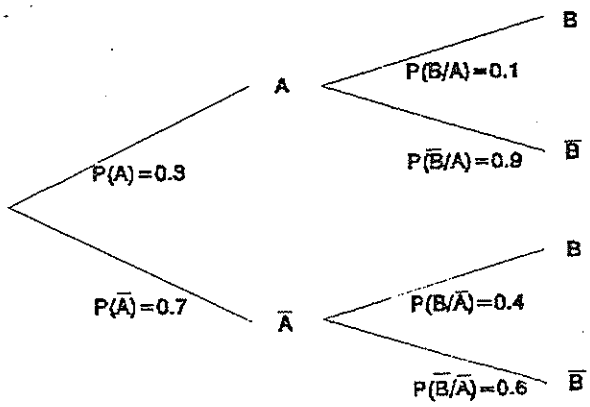
	A	no A	Totales
B	P(A y B)=1/9	P(no A y B)=3/9	P(B)=4/9
no B	P(A y no B)=4/9	P(no A y no B)=1/9	P(no B)=5/9
Totales	P(A)=5/9	P(no A)=4/9	1

puedes construir el de árbol:



Sin más que hacer en cada caso el cociente indicado en la expresión [1].

Recíprocamente, dado el diagrama de árbol, puedes construir el diagrama de contingencia:



Obtenemos, multiplicando los números de una misma rama, las probabilidades P(A y B), P(A y no B), P(no A y B), P(no A y no B), es decir, el diagrama de contingencia.

	A	$\bar{A}$	**Totales**
B	$P(A \cap B) =$ 0´03	$P(\bar{A} \cap B) =$ 0´28	$P(B) =$ 0´31
$\bar{B}$	$P(A \cap \bar{B}) =$ 0´27	$P(\bar{A} \cap \bar{B}) =$ 0´42	$P(\bar{B}) =$ 0´69
Totales	$P(A) =$ 0´30	$P(\bar{A}) =$ 0´70	**1**

Problemas a resolver

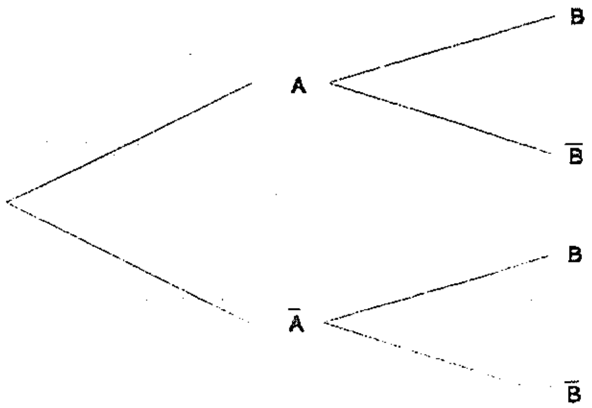
Ahora que ya conocemos los diagramas de contingencia y las relaciones con los diagramas de árbol te vamos a proponer varios problemas que resolverás aplicando lo que has aprendido en esta actividad.

1º

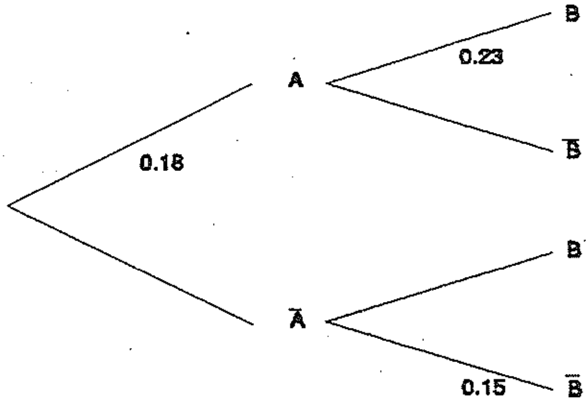
a) Dado el siguiente diagrama de contingencia

	A	**no A**	**Totales**
B	0,41	0,10	0,51
no B	0,30	0,19	0,49
Totales	0,71	0,29	1

construye el diagrama en árbol.



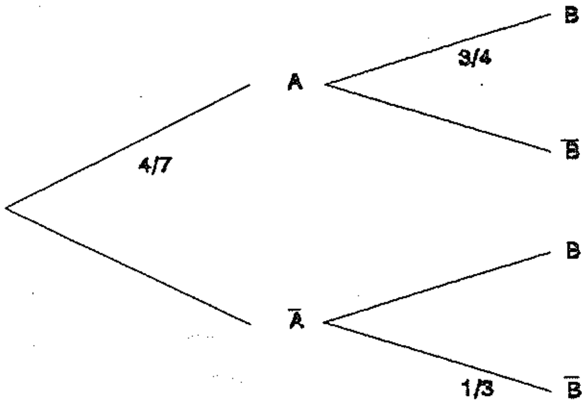
b) Dado el siguiente diagrama en árbol



construye el diagrama de contingencia:

	A	no A	Totales
B			
no B			
Totales			1

2º Completa el siguiente diagrama de probabilidades:



Construye con él una tabla de contingencia y a partir de ella un nuevo diagrama de árbol poniendo el suceso B en primer lugar.

3º Disponemos de dos cajas, M y N, la caja M contiene 7 bolas negras y 3 blancas, y la N 5 negras y 3 blancas. Se saca una bola al azar de una de las dos cajas, también al azar, y resulta ser blanca. Calcular la probabilidad de que proceda de la caja M.

4º Se han observado 50 enfermos de piel tratados con un nuevo antibiótico y otros 70 enfermos no tratados. Anotamos las curaciones al cabo de 2 semanas, los resultados han sido:

	Tratados	No tratados
Curados	40	20
No Curados	10	50

Si se emplean estos datos para asignar probabilidades,

a) ¿Qué probabilidad existe de que un enfermo curado haya sido tratado?
b) ¿Qué probabilidad existe de que un enfermo curado no haya sido tratado?

5º Una determinada empresa se dedica a la fabricación de cubiertas para las ruedas de los automóviles. Por los controles de calidad que esta empresa lleva a cabo se puede concluir diciendo que el 5% de las cubiertas producidas son defectuosas. En el laboratorio de control de calidad hay instalado un dispositivo que detecta el 90% de las cubiertas defectuosas, pero también califica como defectuosas el 2% de las correctas. El empresario está interesado en conocer las siguientes probabilidades:

* Probabilidad de que sea correcta una cubierta calificada como defectuosa por el dispositivo.
* Probabilidad de que sea defectuosa una cubierta calificada por el dispositivo como correcta.

¿Puedes ayudarle tú?

Nota: Consideremos los sucesos:

A = "Ser correcta"
B = "Ser calificada como correcta"
no A = "Ser defectuosa"
no B = "Ser calificada defectuosa"

El problema pregunta: P(A/no B) y P(no A/B). Para resolver el problema puedes construir primero el diagrama de árbol, el diagrama de contingencia y finalmente responder a las probabilidades pedidas a través del diagrama de contingencia.

$$P(A/\text{no } B) = P\,(A \text{ y no } B)\;/\;P(\text{no } B)$$

$$P(\text{no } A/B) = P\,(B \text{ y no } A)\;/\;P(B)$$

6º En un municipio hay tres partidos políticos: Progresista, Liberal y Moderado. Se efectúa un referéndum para decidir si un cierto día se declara fiesta local. He aquí los resultados en %, en función del partido al que votó cada ciudadano en las últimas elecciones:

		ULTIMAS ELECCIONES			
		Pr	Lib	Mod	Abs
REFERÉNDUM	SI	15 %	25 %	12 %	8 %
	NO	25 %	5 %	8 %	2 %

a) ¿Qué porcentaje votó a cada partido en las últimas elecciones?

b) ¿Qué probabilidad hay de que una persona tomada al azar haya votado Sí en el referéndum, P(Sí)?

c) Calcular las siguientes probabilidades:

P(Pr/Sí) P(Sí/Pr) P(Lib/Sí) P(Sí/Lib)
P(Mod/Sí) P(Sí/Mod) P(Abs/Sí) P(Sí/Abs)

d) Los sucesos "votar moderado en la última votación" y "votar Sí en el referéndum" son dependientes o independientes?

Contenidos implícitos en esta actividad

- Tablas de contingencia.
- Probabilidad condicionada.
- Relación entre los diagramas de contingencia y diagramas de árbol.

Actividad Nº 16: Probabilidad geométrica

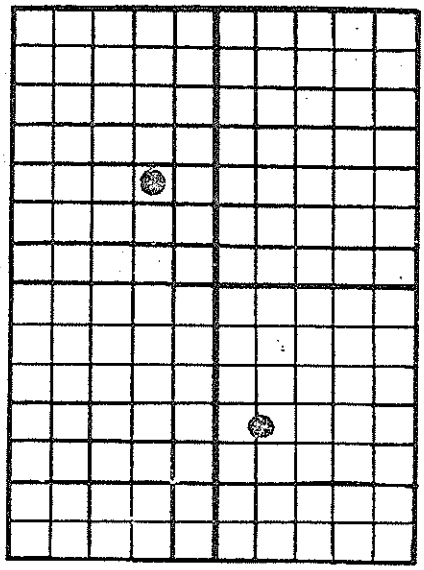
Reglas del juego

Este juego consiste en experimentar al fenómeno aleatorio "lanzar una moneda de cien pesetas sobre una cuadrícula de 60 x 60 mm". Cada alumno realizará el experimento cien veces.

Materiales necesarios

Cada alumno necesita:

- Una moneda de 100 pesetas.
- Una cuadrícula de 60 x 60 mm, que obtendrás uniendo cuatro hojas DIN A4, como muestra la figura, y después haciendo el cuadriculado:
- Unos folios para ir anotando los resultados, conclusiones o conjeturas.



Desarrollo del juego

1º- ¿Puedes formular alguna conjetura sobre qué es lo que ocurrirá al lanzar una moneda de cien pesetas sobre una cuadrícula de 60 x 60 mm?.

2º- Sobre una cuadrícula lanzas la moneda cien veces y anotas el número de veces que la moneda toca alguna de las líneas de la cuadrícula.

Llamamos al suceso "la moneda toca alguna de las líneas de la cuadrícula", suceso S. Con los datos de la experiencia, rellena la siguiente tabla:

	f(S)	$f_r(S)$
SUCESO "S"		
100 Lanzamientos		

3º- Recogemos los datos obtenidos por todos los alumnos de la clase y los colocamos en la siguiente tabla:

	$f(S)$ Frecuencias absolutas de los Alumnos	$f_r(S)$ Frecuencia relativa total
S		
	100 X N lanzamientos	

$$f_r(S) = \frac{\text{suma de las frecuencias absolutas}}{100 \times N} \; ; \; N = \text{número de Alumnos}$$

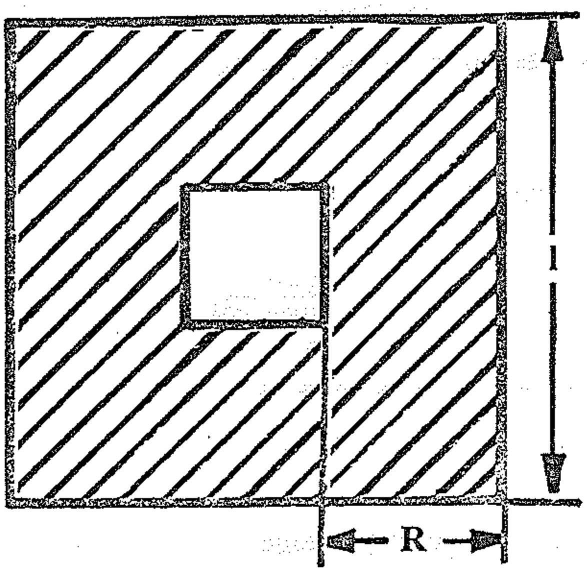
A la vista de los resultados obtenidos en la última tabla, ¿qué probabilidad asignarás al suceso S? Razónalo.

4º Recuerda la ley de Laplace, que utilizamos para el cálculo de probabilidades cuando los sucesos elementales son equiprobables:

$$P(S) = \frac{\text{casos favorables}}{\text{casos posibles}}$$

Con esta ley también calcularemos **probabilidades geométricas**, como vas a ver a continuación.

Si tomamos un cuadrado de la cuadrícula, la moneda pisará alguna línea de la cuadrícula si su centro cae dentro del área rayada:



l = 60 mm.
R = radio de la moneda.
ϕ = diámetro de la moneda = 24,5 mm.

Según la ley de Laplace la probabilidad de que al lanzar la moneda pise alguna línea de la cuadrícula será:

$$P(S) = \frac{\text{área de la zona rayada}}{\text{área del cuadrado}} = \frac{l^2-(l-2R)^2}{l^2} =$$

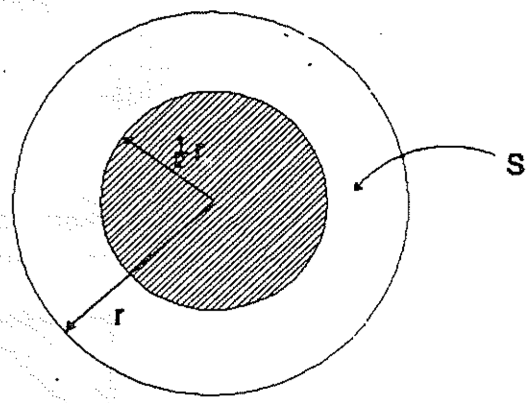
$$= \frac{l^2}{l^2} - \frac{(l-2R)^2}{l^2} = 1-(\frac{l-2R}{l})^2 = 1-(1-\frac{\phi}{l})^2$$

$$P(S) = 1-(1-\frac{24'5}{60})^2 = 0'65$$

5º- Compara el resultado anterior con el que se obtuvo, de forma experimental, en el punto 3 y comenta la primera ley de los grandes números.

6º- Con el fin de que veas cómo se hacen los cálculos de probabilidades geométricas te resolveremos varios problemas.

Problema 1: En el interior de un círculo se selecciona un punto al azar. Hallar la probabilidad de que el punto quede más cercano al centro que a la circunferencia.
Denotamos por S el conjunto de los puntos interiores al círculo de radio r y denotamos por A el conjunto de puntos interiores al círculo concéntrico de radio 1/2 de r. Así A está formado por aquellos puntos de D que están más cercanos a su centro que a su circunferencia.



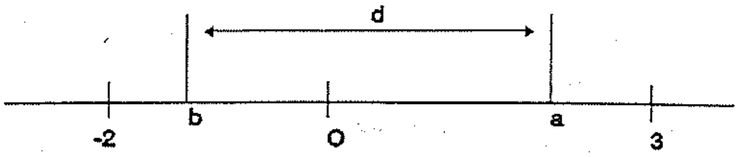
La probabilidad pedida vale:

$$P(A) = \frac{\text{área de } A}{\text{área de } S} = \frac{\pi(\frac{1}{2}r)^2}{\pi r^2} = \frac{1}{4}$$

Problema 2: Sobre una línea recta se seleccionan al azar los puntos a y b tales que

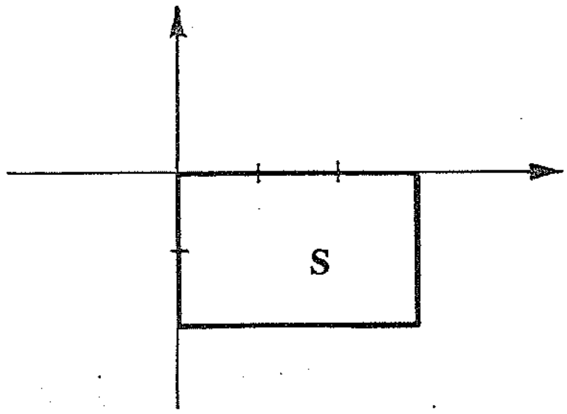
$$-2 \le b \le 0 \;\; y \;\; 0 \le a \le 3$$

como se muestra en la figura siguiente:



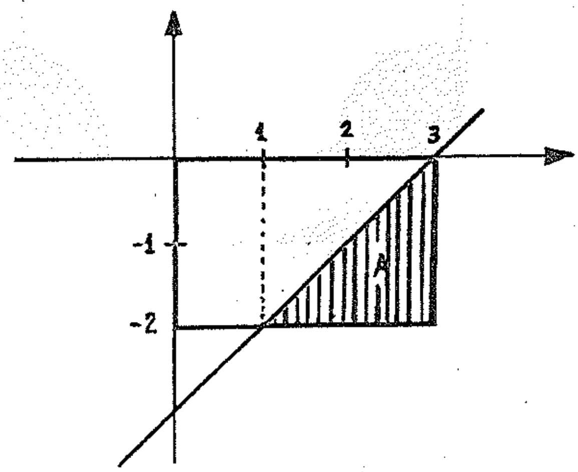
Hallar la probabilidad de que la distancia d entre a y b sea mayor que 3.

Los puntos (a,b) posibles son los que se muestran en el diagrama siguiente:



e incluidos en el rectángulo.

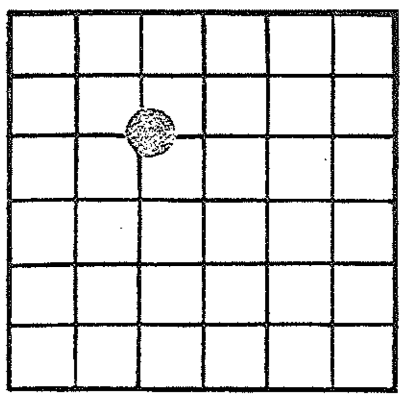
El conjunto de puntos favorables (a,b) serán aquellos para los cuales d = a - b $\geq$ 3 y son los puntos de S que caen por debajo de la línea x - y = 3 y forman por tanto la superficie sombreada.



En consecuencia:

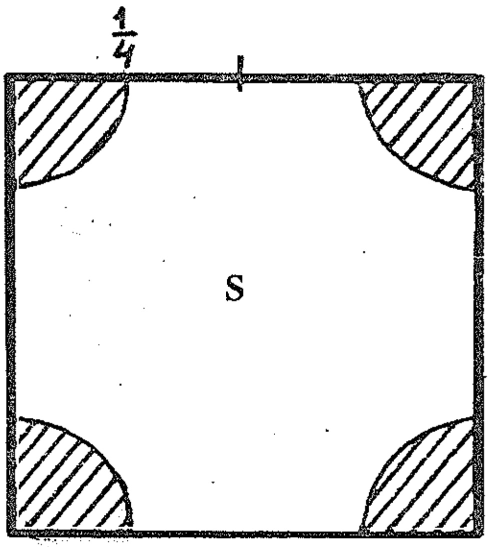
$$P(A) = \frac{\textit{área de } A}{\textit{área de } S} = \frac{2}{6} = \frac{1}{3}$$

Problema 3: Se tiene una cuadrícula de 1 x 1 cm. Se lanza un disco sobre ella de 1/2 cm de diámetro al azar. Hallar la probabilidad de que el disco cubra un punto de intersección de la cuadrícula.



Tomamos un cuadrado de la cuadrícula y sea S el conjunto de los puntos interiores a este cuadrado.

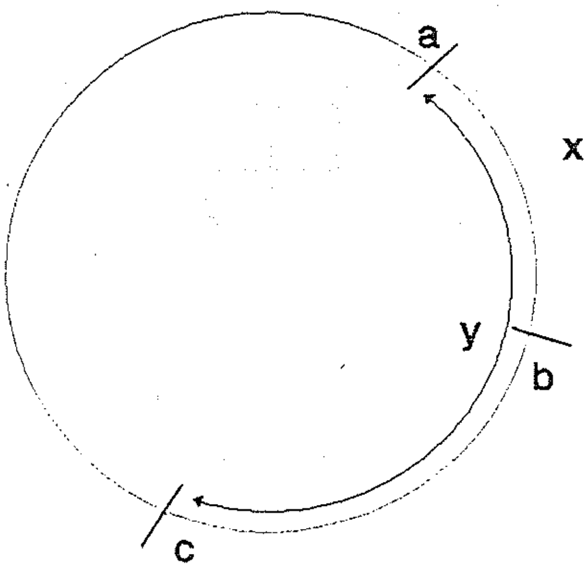
Denotamos por A el conjunto de puntos de S de distancia a las esquinas 1/4 cm. Así si el centro del disco cae en S cubrirá un punto de intersección de la cuadrícula sí y sólo sí su centro cae en un punto de A.



A= sombreada

Según esto:

$$P(A) = \frac{\text{área de } A}{\text{área de } S} = \frac{\pi \left(\frac{1}{4}\right)^2}{1^2} = \frac{\pi}{16}$$



Problema 4: Tres puntos a, b, c de una circunferencia se escogen al azar. Hallar la probabilidad de que los puntos caigan sobre el mismo semicírculo.

Supongamos que la longitud de la circunferencia sea 2s. Denotamos por x la longitud del arco ab en el sentido del movimiento de las agujas del reloj y denotamos por y la longitud del arco ac en el mismo sentido. Se debe cumplir:

$$0 < x < 2s$$

$$0 < y < 2s$$

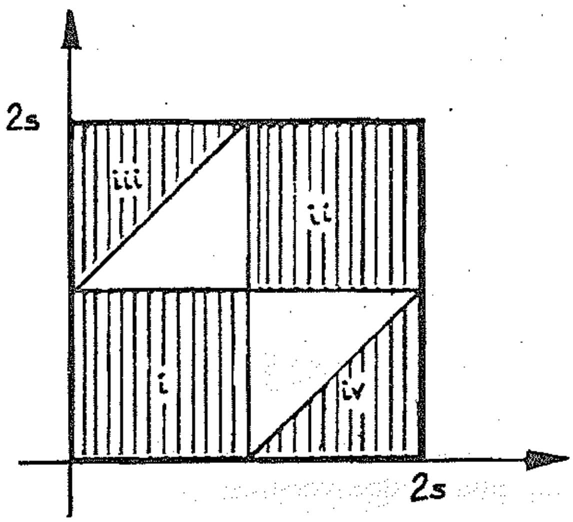
Sea S el conjunto de los puntos de R^2 para los cuales se cumplen las condiciones anteriores: Sea A el subconjunto de S para los cuales se cumplen las condiciones siguientes:

i) $x,y < s$

ii) $x,y > s$

iii) $x < s,\ y - x > s$

iv) $y < s,\ x - y > s$



Entonces A consta de aquellos puntos para los cuales se cumple que a, b, c caen sobre el semicírculo. Así

$$P(A) = \frac{\text{área de } A}{\text{área de } S} = \frac{3s^2}{4s^2} = \frac{3}{4}$$

Problema 5: Dado un segmento cualquiera, hallar la probabilidad de obtener, por trisección, los tres lados de un triángulo.

Entendemos por trisección la elección de dos puntos al azar del interior del segmento. Es evidente que no supone restricción alguna identificar al segmento con el intervalo [0,1]. La elección de dos puntos en el segmento es entonces la elección de dos números tales que

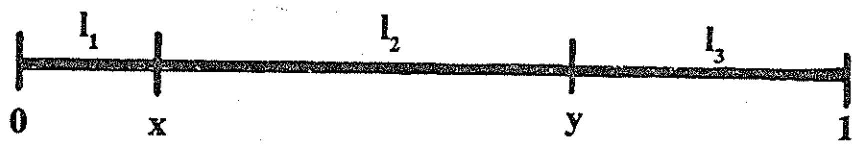
$x \in (0,1)$

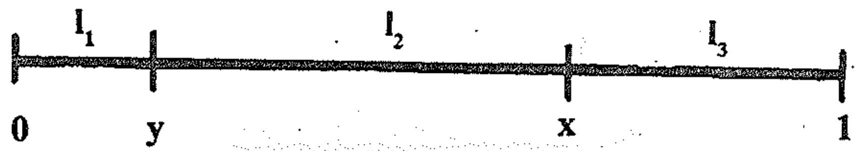
$y \in (0,1)$

$x \neq y.$

La condición necesaria y suficiente para que con tres segmentos

se pueda formar un triángulo es que la longitud de cada uno de los segmentos sea menor que la suma de los otros dos. Refiriendo estas condiciones a los números x e y obtenemos:

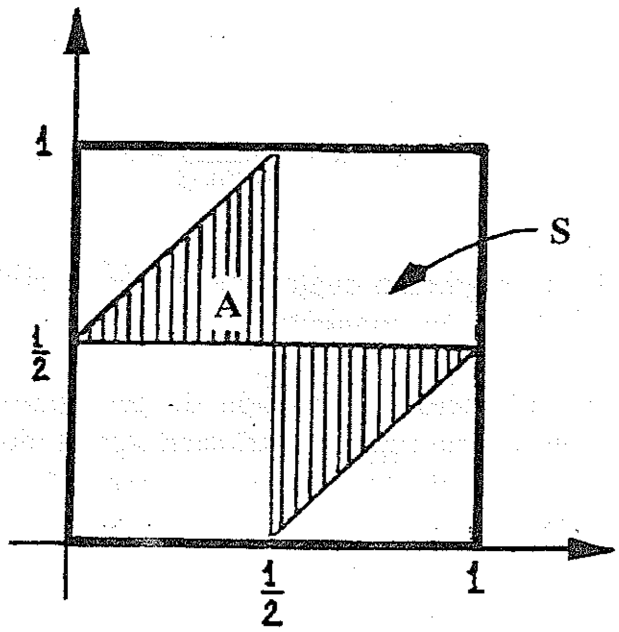




$$x<y;\ 0<x<\frac{1}{2};\ y-x<\frac{1}{2};\ \frac{1}{2}<y<1$$

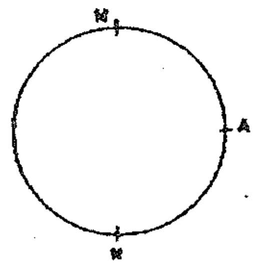
$$y<x;\ 0<y<\frac{1}{2};\ x-y<\frac{1}{2};\ \frac{1}{2}<x<1$$

Interpretando geométricamente estas condiciones, identificando (x,y) con un punto del plano, resulta que los puntos que verifican la condición pertenecen a la parte rayada de la figura:



La probabilidad es por tanto:

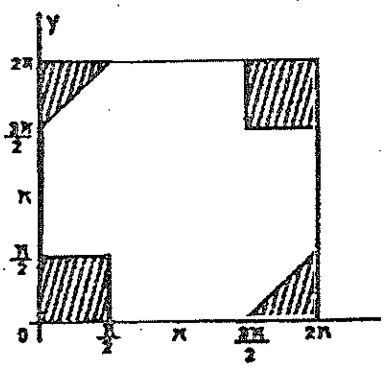
$$P(A) = \frac{\acute{a}rea\ de\ A}{\acute{a}rea\ de\ S} = \frac{\frac{1}{8}+\frac{1}{8}}{1} = \frac{1}{4}$$



Problema 6: En una circunferencia se escogen al azar tres puntos A, B y C.

Calcular la probabilidad de que los tres estén situados en un mismo arco de 90º.

Supuesto elegido el punto A tomémoslo como origen de medida de longitudes de arco en una circunferencia que podemos suponer de radio 1.



El punto B se puede escoger de forma que la longitud x del arco AB esté comprendida en

$$[0, \frac{\pi}{2}] \cup [\frac{3\pi}{2}, 2\pi]$$

Elegido el punto B, la medida del arco AC verificará:

a) Si $0 \le x \le \frac{\pi}{2}$ *entonces* $0 \le y \le \frac{\pi}{2}$ *o bien* $x - \frac{\pi}{2} + 2\pi \le 2\pi$

b) Si $\frac{3\pi}{2} \le x \le 2\pi$ *se verificará* $\frac{3\pi}{2} \le y \le 2\pi$ *o bien* $0 \le y \le x + \frac{\pi}{2} - 2\pi$

En consecuencia (x,y) ha de ser un punto de la parte rayada de la figura, cuyo área es

$$2\left(\frac{\pi^2}{4} + \frac{\frac{\pi}{2} \times \frac{\pi}{2}}{2}\right) = \frac{3\pi^2}{4}$$

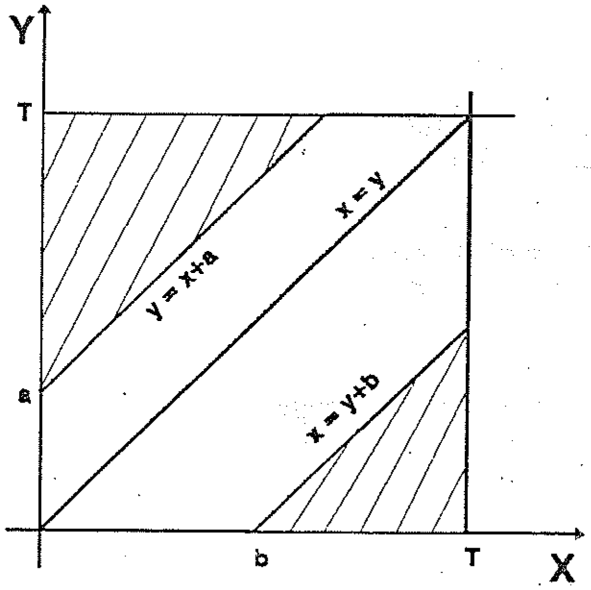
La probabilidad buscada es

$$\frac{\frac{3\pi^2}{4}}{4\pi^2} = \frac{3}{16}$$

Problema 7: Un tren llega a una estación en un instante al azar del intervalo (0,T), parando **a** minutos. A la misma estación acude un autobús al azar e independientemente en el mismo intervalo y parando **b** minutos, **a** y **b** < T.

1) Calcular la probabilidad de que el tren llegue antes que el autobús.
2) Determinar la probabilidad de que se encuentren.
3) Suponiendo que se encuentren en la estación, determinar la probabilidad de que el tren llegue antes que el autobús.

Interpretaremos el problema geométricamente.



Sean x el instante de llegada del tren e y el de llegada del autobús.

1) Es la probabilidad del suceso $x < y$:

$$P=\frac{\frac{1}{2}T^2}{T^2}=\frac{1}{2}$$

2) Es la probabilidad de la suma de los sucesos:

$$x < y < x+a$$
$$y < x < y+b;$$

luego

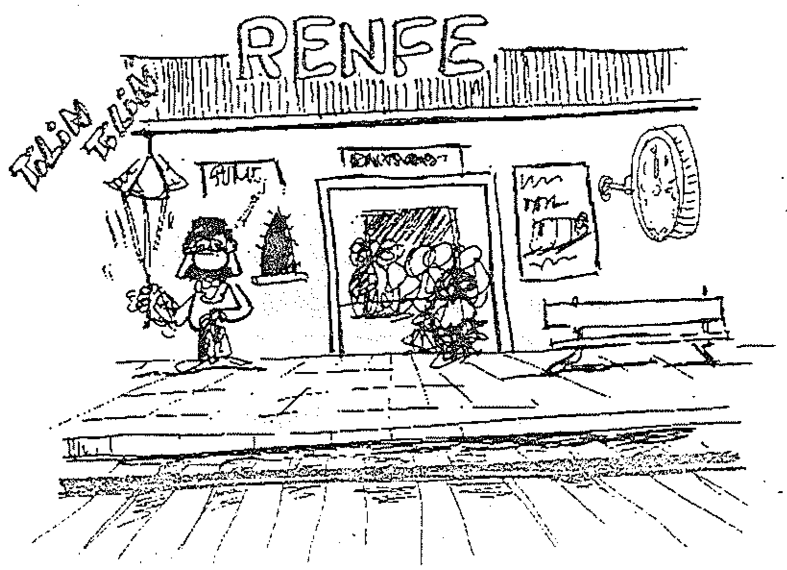
$$P=\frac{\frac{1}{2}T^2-\frac{1}{2}(T-a)^2}{T^2}+\frac{\frac{1}{2}T^2-\frac{1}{2}(T-b)^2}{T^2}=\frac{(a+b)T-\frac{1}{2}(a^2+b^2)}{T^2}$$

3) Es

$$P=\frac{\frac{1}{2}T^2-\frac{1}{2}(T-a)^2}{(a+b)T-\frac{1}{2}(a^2+b^2)}=\frac{aT-\frac{1}{2}a^2}{(a+b)T-\frac{1}{2}(a^2+b^2)}$$

Contenidos implícitos de esta actividad

- Probabilidades geométricas.



Actividad Nº 18
Probabilidad dinámica

Reglas del juego

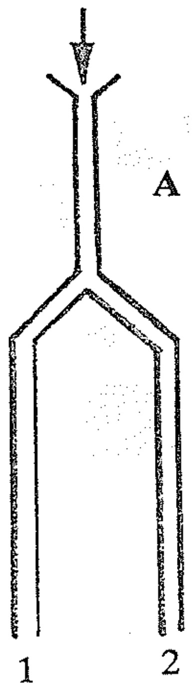
Utilizando fichas, simular situaciones de problemas de probabilidad.

Materiales necesarios

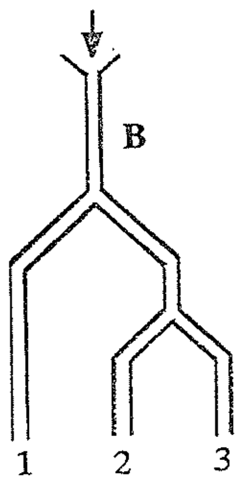
Un gran número de fichas.

Etapas del desarrollo del juego

Embudos



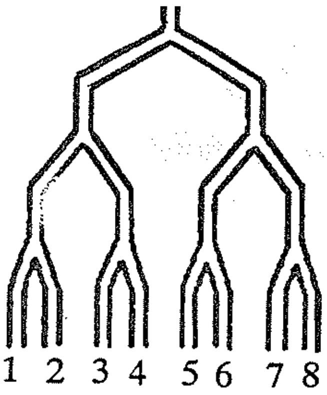
Si dejamos caer un gran número de bolitas, perdigones, fichas, etc, en los embudos A y B que aparecen dibujados observaremos que ocurre lo siguiente:



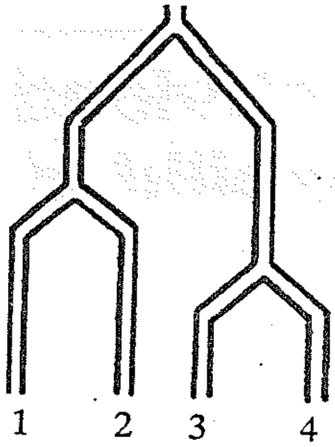
Las posibilidades que tiene un objeto de ir por un camino u otro es igual en cada bifurcación; como en cada bifurcación aparecen dos ramas, la probabilidad de que un objeto pase por cada una de ellas es 1/2 en cada cruce; expresado en porcentaje, el 50%.

Así, si lanzamos 100 objetos por el embudo A, esperamos que aproximadamente sean 50 los que salen por el canal 1 y otros 50 por el canal 2. Si los 100 lanzamientos los realizamos en el embudo B, esperamos que aproximadamente 50 caigan por el canal 1 y 25 por los canales 2 y 3.

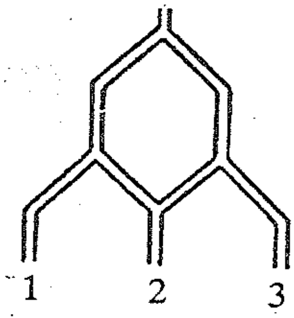
Supón que ahora dejamos caer un montón de perdigones en cada uno de los embudos C, D y F. En estos casos, ¿qué proporción de perdigones esperas que salga por cada canal?



Embudo C

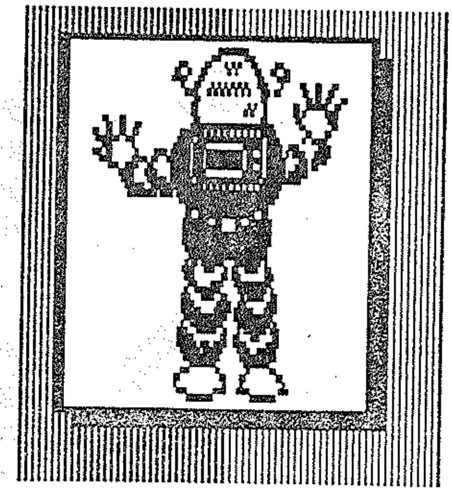


Embudo D



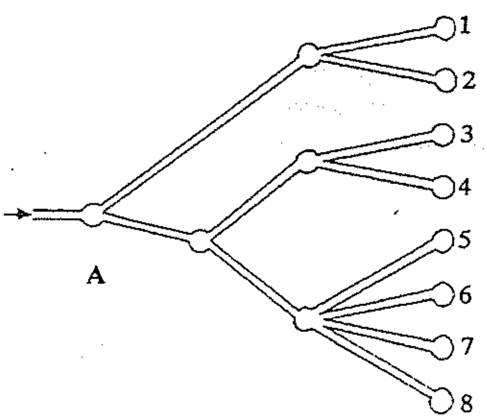
Embudo E

Laberintos

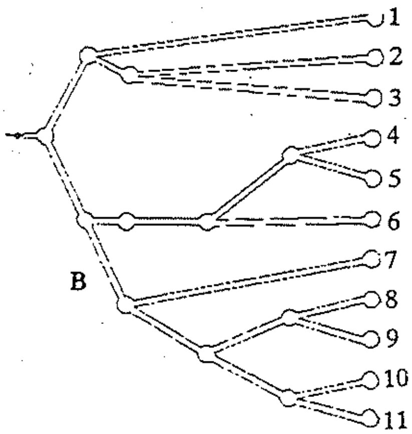


Se pone un robot en un laberinto y empieza a explorarlo. En cada bifurcación es igual de probable que el robot continúe por un camino que por otro (excepto que no puede retroceder por el mismo camino por el que ha llegado). Hay trampas al final de los caminos. ¿En cuál de las trampas es más probable que acabe el robot, o son todas igualmente probables?

Imagina que repetimos la experiencia muchas veces ¿En qué proporción de ocasiones caerá en cada trampa? Resuelve, para cada uno de los laberintos dibujados.

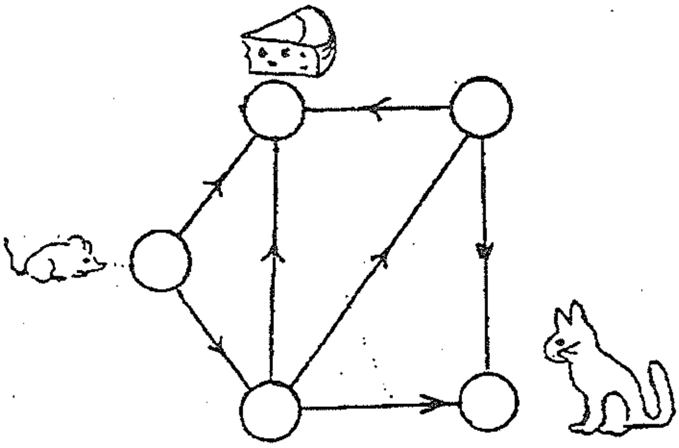


A



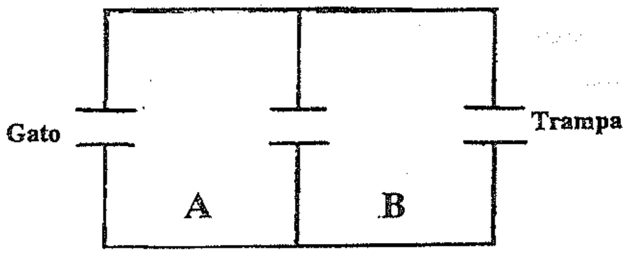
B

Ratones



A).- Se introducen doce ratones por la entrada del laberinto de la figura, que tiene dos salidas, una en la que hay un gato y otra en la que se encuentra un queso. Si el ratón llega a donde está el queso, se lo come y sale libre, pero si va a parar a la salida donde se encuentra el gato, es comido por éste.

¿Cuántos ratones se salvarán? ¿Cuántos serán comidos por el gato? ¿Y si hay 24 ó 36 ratones? ¿Cuál es la probabilidad de que el ratón se coma el queso? ¿Cuál la de que sea comido por el gato?



B).- Un ratón se mueve entre dos habitaciones A y B. Si sale de A es atrapado por el gato, mientras que si sale de B cae en una trampa, como se ve en el diagrama. Inicialmente parte de la habitación A y su movimiento se realiza de la siguiente forma: de A a B con probabilidad 3/4, de B a A con probabilidad 7/8. Hallar la probabilidad de que: a) lo coja el gato, b) lo coja la trampa.

C).- Un ratón puede estar en 6 posiciones (como se indica en la figura) de forma que si está en la posición i, luego se mueve a las posiciones i+1 o i-1 con probabilidades 2/3 y 1/3 respectivamente, pero si llega a las posiciones 1 o 6 queda atrapado en sendas trampas. En este paseo aleatorio, si el ratón parte de la posición 3 determinar las probabilidades de ser atrapado en las trampas 1 y 6.

Estrategias

D).- Un jugador necesita cinco millones de pesetas para pagar una deuda, pero solo tiene un millón y su Banco no está dispuesto a prestarle; decide intentar conseguir la cantidad total jugando a cara o cruz con una estrategia audaz: en cada jugada se apuesta una cantidad de dinero tal que si gana llegue de la forma mas rápida posible a su objetivo. Estudiar la estrategia a seguir por el jugador y calcular la probabilidad que tiene de conseguirlo.

VIII. PROBLEMAS CON PROBABILIDADES

1) Ir a dos fábricas. Un ingeniero trabaja en dos fábricas A y B. Los ómnibus que lo llevan a las fábricas parten del mismo lugar. A las horas exactas sale el que lo lleva a la fábrica A, y a las horas y cuarto el que lo lleva a la fábrica B. Si sale de su casa sin preocuparse de la hora y toma el primer ómnibus que llega, ¿Cuál es la probabilidad de que llegue a cada una de las fábricas?

2) Nueva sociedad. Un grupo financiero estudia el lanzamiento de una nueva sociedad con una inversión de $ 500.000. De acuerdo a sus expertos, tres casos solamente pueden ocurrir:

- La sociedad se desarrolla rápidamente obteniendo el 30 % del mercado. Se tiene una ganancia de $ 800.000 con una probabilidad del 10 %;
- La sociedad comienza normalmente obteniendo el 10 % del mercado. Se tiene una ganancia de $ 200.000 con una probabilidad del 70 %;
- La sociedad cierra después de una tentativa infructuosa. Los expertos estiman al 20 % la probabilidad de esta eventualidad.

¿Cuál es la ganancia esperada ?

Observación. La ganancia esperada no corresponde a la ganancia de ninguna de las tres posibilidades. Ella indica solamente la ganancia media que se obtendría con la hipótesis irreal que se podría ejecutar la inversión un número grande de veces. Esta medida permite tomar en cuenta la ganancia de cada posibilidad futura teniendo en cuenta como peso la probabilidad de dicha realización. En otras palabras, la ganancia esperada es la media aritmética (esperanza) de las ganancias de cada posibilidad, ponderadas por su probabilidad respectiva.

3) Oscar y sus dos tías. Oscar tiene dos tías a quienes visita cada viernes por la noche. Cada una de ellas vive sobre la ruta de la línea de ómnibus B, una en dirección sur y la otra en dirección norte. Cada viernes, alrededor de las 7 de la tarde, Oscar camina hacia la parada y toma el primer ómnibus que pasa, en una dirección o en la opuesta. En ambas direcciones la frecuencia es de 10 minutos. Al cabo de algún tiempo, una de las tías llama a Oscar reclamando que sólo había ido una vez, mientra que ha visitado 9 veces a su hermana. ¿Cómo se explica esto?

4) Gracia de un condenado a muerte. En el lejano reino de Juegolandia, a los condenados a muerte se les concedía la gracia de que su vida dependiera de que sacaran una bola blanca de una bolsa que contenía 50 bolas blancas y 50 negras. Pero en cierta ocasión, un reo pidió la gracia de que se le dejara distribuir las bolas de otro modo antes de hacer el sorteo. Tras algunas discusiones, se le concedió la gracia y preparó dos bolsas: en una colocó una sola bola blanca; en otra bolsa colocó 49 blancas y 50 negras. ¿Cuál resultó de este modo la probabilidad de sacar blanca?

5) Azar en el círculo. En el interior de un círculo se selecciona un punto al azar. Halle la probabilidad de que el punto quede más cercano al centro que a la circunferencia.

6) Azar de un disco en una cuadrícula. Se tiene una hoja cuadriculada con cuadrados de 1cm x 1cm. Se lanza sobre ella, al azar, un disco de 1/4 cm de radio. Halle la probabilidad de que el disco cubra un

punto de intersección de la cuadrícula.

7) Un jugador arriesgado. Un jugador necesita cinco millones para pagar una deuda, pero solo tiene un millón y su Banco no está dispuesto a prestarle; decide intentar conseguir la cantidad total jugando a cara o cruz con una estrategia audaz: en cada jugada se apuesta una cantidad de dinero tal que si gana llegue de la forma más rápida posible a su objetivo. Estudie la estrategia a seguir por el jugador y calcule la probabilidad que tiene de conseguirlo.

8) Azar en una recta. Sobre una línea recta se seleccionan al azar los puntos a y b tales que:

$$-2 \le b \le 0 \qquad \text{y} \qquad 0 \le a \le 3 \;.$$

Halle la probabilidad de que la distancia entre a y b sea mayor que 3.

Definición.– Dados dos números positivos a y b, se definen la media aritmética MA y la media geométrica MG de la siguiente manera :

$$MA = \frac{a+b}{2} \qquad ; \qquad MG = \sqrt{ab} \;.$$

9) (i) Demuestre que :

(a) $MG \le MA$; (b) $MG = MA \Leftrightarrow a = b$; (c) $x\,(1-x) \le \frac{1}{4}$, $\forall\; x \in [0,1]$.

(ii) ¿Cuál es el rectángulo de área máxima que tiene por perímetro p?

(iii) La población de la Argentina está dada en la siguiente tabla:

año	población (en millones de habitantes)
1960	20
1970	23
1980	28
1990	33

(a) Calcule las medias aritmética y geométrica de la población de los años 1960 y 1980. Compare los valores hallados con los que nos brinda la tabla para el año 1970.

(b) Idem (a) para los años 1970 y 1990.

(c) Estime el número de habitantes que tendría la Argentina en el año 2000, utilizando solamente los datos de la tabla.

Observación.– Los resultados obtenidos con la MA y la MG no son muy diferentes. En general, para problemas referentes al crecimiento o decrecimiento de poblaciones la MG da resultados más aproximados a la realidad. Naturalmente que éstos son valores estimados, es decir, aproximados, pues pueden ocurrir fenómenos (catástrofes, gran inmigración o emigración) que los haga variar.

IX. TRABAJOS PRACTICOS Y CASOS

TRABAJO PRACTICO 1: Estrategias con urnas y bolas.

Una urna contiene α bolas blancas y β bolas negras; se sabe que $\alpha + \beta \geq 3$. Los jugadores A y B juegan un juego con la urna y las bolas. Se consideran dos estrategias:

(I) El jugador A toma al azar una bola de la urna. Si es blanca entonces él gana, en caso contrario pierde;

(II) El jugador A toma al azar una bola y la tira afuera de la urna sin mirarla. El jugador B toma entonces una bola negra. Luego, A toma otra bola de la urna. Si esta segunda bola es blanca entonces A gana. En caso contrario, pierde.

(i) Calcule la probabilidad de ganar que tiene el jugador A en las estrategias I y II.

(ii) ¿ Cuál de las dos estrategias es preferible para el primer jugador A?

TRABAJO PRACTICO 2: El tren y el ómnibus. Un tren llega a una estación en un instante del intervalo (0,T) al azar, parando α minutos. A la misma estación acude un ómnibus al azar, independientemente del tren, en el mismo intervalo de tiempo (0, T) parando β minutos (con $\alpha < T$ y $\beta < T$).

(i) Calcule la probabilidad de que el tren llegue antes que el ómnibus;

(ii) Determine la probabilidad de que se encuentren;

(iii) Suponiendo que se encuentren en la estación, determine la probabilidad de que el tren llegue antes que el ómnibus;

(iv) Analice los resultados anteriores dando diferentes valores a α y a β.

TRABAJO PRACTICO 3: El duelo triangular. Tres políticos A, B y C deciden resolver sus diferencias mediante un duelo a pistola bajo las siguientes reglas. Luego de sortear quien dispara en primer, segundo y tercer lugar, los tres se ubican en cada uno de los vértices de un triángulo equilátero. Se conviene que cada uno disparará un tiro por turno, continuando en el orden sorteado hasta que dos de ellos queden fuera de combate. Cada uno puede disparar en la dirección que desee. Además, conocen que las precisiones de tiro de A, B y C son de 100%, 80% y 50% respectivamente.

Suponiendo que cada uno adopta la estrategia más favorable y que nadie muere a causa de una bala perdida que no había sido disparada contra él,

(i) ¿Quién tiene más chances de sobrevivir?

(ii) ¿Cuáles son las probabilidades de cada uno? Demuestre que se tiene que:

$$P_A = \frac{3}{10} \ ; \qquad P_B = \frac{8}{45} \ ; \qquad P_C = \frac{47}{90} \, .$$

(iii) Suponiendo que C no dispara al aire, calcule las probabilidades de cada uno de sobrevivir.

(iv) Si la precisión de C es del x % ¿Cuál es el mínimo valor de x que garantice la victoria a C?

Observación: dado que A y B son los mejores tiradores, ambos tratarán de eliminarsese el uno contra el otro, luego, la mejor estrategia para C es disparar al aire hasta uno A ó B caiga. De este modo, C tiene el primer disparo contra el sobreviviente con una chance del 50 % de vencer.

TRABAJO PRACTICO 4. La compañía X está tratando de decidir sobre la compra de una máquina nueva, la cual se utilizará exclusivamente en la fabricación de cierto producto. Actualmente existen dos máquinas que pueden ser satisfactorias para el fin perseguido.

Si se compra la máquina A, se invertirán \$ 10.000 y se ahorrará \$ 1 por unidad, en relación con el proceso de producción que se utiliza en la actualidad. Si se compra la máquina B, se invertirán \$60.000 y se ahorrarán \$3 por unidad producida. Ambas máquinas tienen una vida útil de 5 años. Las condiciones futuras del mercado son algo inciertas y se han resumido en las siguientes estimaciones sobre la probabilidad correspondiente a un volumen total de ventas dado, para los próximos 5 años:

Ventas Totales (en unidades)		*Probabilidad*
10000	....	0,10
20000	....	0,30
30000	....	0,40
40000	....	0,20

Sin tomar en cuenta el problema de la actualización financiera de los ingresos futuros, ¿cuál es la máquina que debería comprar la empresa X? ¿Cuáles son los ahorros esperados correspondientes a cada una de esas acciones alternativas?

CASO 1: Producción óptima de productos por semana. Se supone que la demanda D, por semana, de cierto producto es una variable aleatoria con determinada distribución de probabilidades, $P(D = n) = p(n)$, $n = 0, 1, 2, \ldots$. Supóngase que el costo para el proveedor es \$ C_1 por artículo, mientras que él lo vende a \$ C_2. Cualquier artículo que no se venda al término de la semana, debe almacenarse con un costo de \$ C_3 por artículo. Si el fabricante decide producir N artículos al comienzo de la semana, entonces:

(i) ¿Cuál es su utilidad esperada por semana?

(ii) ¿Para qué valor de N es máxima la ganancia esperada? Para realizar este estudio se analizará el siguiente caso: Se supone además que para la demanda aleatoria D es apropiada la siguiente distribución de probabilidades:

$$P(D = n) = \frac{1}{5}\,,\ n = 1, 2, 3, 4, 5.$$

(a) Entonces calcule, en este caso, la expresión de la utilidad esperada por semana.

(b) Calcule el N óptimo para el caso concreto de $C_1 = 3$, $C_2 = 9$, $C_3 = 1$.

(c) Pruebe, en general, que el valor N óptimo se obtiene en el caso en que $N < 5$ y que los parámetros C_1, C_2 y C_3 satisfagan la siguiente desigualdad:

$$C_2 < 10\,C_1 + 9\,C_3 .$$

Más aún, en dicho caso el N óptimo se obtiene como la parte entera o la parte entera más uno del valor real x_o dado por la expresión:

$$x_o = \frac{1}{2} + \frac{5\,(C_2 - C_1)}{C_2 + C_3}.$$

CASO 2: Postres preparados. Un fabricante de postres planifica el lanzamiento de un nuevo tipo de plan. Le ofrecen dos posibilidades:

- venderlo a un precio elevado para hacer un producto de prestigio;
- venderlo a un precio accesible (barato) para realizar una venta masiva y obtener una importante ganancia por cantidad.

Su competidor principal tiene también un proyecto de este tipo, pero al presente no se conoce todavía cuál será su decisión. Tres posibilidades del futuro pueden ser consideradas, y sus consecuencias están descriptas en el cuadro 1 siguiente:

Unidad: \$ 10.000	precio bajo (\$)	precio elevado (\$)	probabilidad del evento
No competencia (NC)	20	-2	0,50
Competencia con buena reacción del mercado (CBR)	5	25	0,20
Competencia con débil reacción del mercado (CDR)	-10	10	0,30

(i) Represente el árbol de decisión correspondiente al problema del fabricante;

(ii) ¿Cuál es la mejor decisión a tomar, y cuánto es su esperanza de ganancia (o ganancia esperada)?;

(iii) ¿Qué permite conocer con certeza, cuál de las tres posibilidades del futuro ocurrirá? En otras palabras, cuál es el monto a pagar por la encuesta que nos permita pasar de información imperfecta (dada por el cuadro anterior) a información perfecta.

(iv) Un asesor externo, de quién se conoce la seguridad de sus estudios, propone sus servicios para prever la reacción de la competencia y del mercado; las previsiones del experto están dadas en el siguiente cuadro 2:

RR: Reacción real

RP: Reacción prevista

RR \ RP	NC	CBR	CDR
NC	0,80	0,10	0,10
CBR	0,10	0,70	0,20
CDR	0,10	0,30	0,60

Observación. Cuando hay competencia y se tiene una débil reación del mercado (caso CDR), el experto prevé:

NC en el 10 % de los casos (probabilidad 0,10)

CBR en el 30 % de los casos (probabilidad 0,30)

CDR en el 60 % de los casos (probabilidad 0,60).

Si el análisis del experto cuesta \$ 15.000 ¿Conviene contratarlo?

CASO 3: Curso de idioma. Una empresa lanza un curso de idioma por CD (discos compactos). La demanda puede ser de 1000, 2000, 3000, 4000 ó 5000 unidades. La creación del curso costó \$ 100.000. El juego de CD cuesta \$ 200 y se vende a \$ 400. Como todos los CD a producir deben ser realizados en el mismo momento, la sociedad debe decidir el número de cursos a producir. Por razones de prestigio, los cursos que no se pudieron vender, no pueden venderse en liquidación; pero los CD se borran y se liquidan a \$ 100 el juego.

I) Formulación del problema. El problema puede ser representado por una matriz, en la cual cada fila corresponde a una decisión de la empresa, y cada columna corresponde a una posibilidad futura de demanda. De acuerdo a los datos anteriores coloque en la intersección de cada fila y cada columna el beneficio (o pérdida) correspondiente (en unidades de \$ 1.000 para facilitar las operaciones).

P \ D	1000	2000	3000	4000	5000
1000					
2000					
3000					
4000					
5000					

D: Demanda de CD ; P: Producción de CD

II) A continuación se estudiarán varios criterios que permitirán elegir la solución adecuada:

(a) El criterio maximin. Su interés principal, además de su simplicidad, reside en su óptica conservadora: se elige la estrategia para la cual la ganancia mínima es la más grande posible. Este criterio tiene por objetivo principal dar una garantía: "se va a ganar al menos tanto", o "se va a perder a lo sumo tanto".

Sea $g(p_i, d_j)$ la ganancia obtenida para un nivel de producción de p_i unidades (i=1,...,5) y una demanda de d_j unidades (j = 1, ..., 5). Entonces

$$\underset{j}{\mathrm{Min}}\ g(p_i, d_j)$$

será la ganancia más pequeña posible para la producción p_i dada (i = 1, ..., 5).

La estrategia maximin es la que maximiza estos mínimos, es decir halla el nivel de producción p_i de manera tal que se tenga

$$\underset{i}{\mathrm{Max}}\ \underset{j}{\mathrm{Min}}\ g(p_i, d_j) = \underset{j}{\mathrm{Min}}\ g(p_i, d_j)\ .$$

Calcule la estrategia del criterio maximin, es decir ¿cuál es el nivel de producción p_i (i = 1, ..., 5) a elegir y la ganancia a obtener?

(b) El criterio maximax. Este criterio es, por el contrario, extremadamente optimista. Se elige la estrategia que aporta la mayor ganancia, en el mejor de los casos posibles. La estrategia p_i a elegir es la dada por

$$\operatorname*{Max}_{i} \operatorname*{Max}_{j} g(p_i, d_j) = \operatorname*{Max}_{j} g(p_i, d_j) .$$

Calcule la estrategia del criterio maximax.

Observación: Este criterio es de un uso poco frecuente.

(c) Criterio de Laplace. Se utiliza cuando la probabilidad de cada posibilidad futura parece ser la misma (constante) o que se ignoran sus probalidades. Este criterio consiste en elegir la estrategia que maximiza la ganancia esperada, dando la misma probabilidad a cada evento futuro. La estrategia p_i (i=1, ..., 5) a elegir es la dada por

$$G(p_i) = \operatorname*{Max}_{k} G(p_k)$$

donde $G(p_i)$ es la ganancia esperada de la estrategia p_i (con n posibilidades futuras con probabilidad $1/n$ cada una; en nuestro caso se tiene que $n = 5$), es decir:

$$G(p_i) = \sum_{j=1}^{n} \frac{1}{n} g(p_i, d_j) .$$

Calcule la estrategia del criterio de Laplace.

(d) Criterio del minimax (conocido también como del lamento o costo de oportunidad mínimo, es decir que sea mínimo lo que se deja de ganar).

Se supone que se conoce a priori la demanda. Se busca hallar la estrategia que permita obtener la mejor ganancia para esta demanda. En la columna correspondiente a esta demanda se reemplaza cada ganancia por la diferencia entre dicha ganancia y la ganancia correspondiente a la producción optimal para la demanda. Se obtiene de esta manera una tabla identificando, para la demanda dada, el lamento o lo que se deja de ganar, correspondiente a una producción no óptima para esta demanda. Esta modificación de la tabla se realiza para cada valor posible de la demanda. Se elige la estrategia que minimiza el lamento máximo.

Calcule la estrategia del criterio minimax.

Observación: Ningún criterio es en sí mismo, mejor que otro. Todo depende de la situación en la que se encuentre la empresa. El sólo criterio que se puede, más o menos, eliminar es el maximax.

III) Arboles de decisión probabilísticos.

La empresa que lo edita tiene 6 posibilidades de acción:

- no lanzar el producto (nunca debe olvidarse esta posibilidad nula);
- producir 1.000 juegos de CD;
- producir 2.000 juegos de CD;

- producir 3.000 juegos de CD;
- producir 4.000 juegos de CD;
- producir 5.000 juegos de CD.

Se distinguirán 5 estados posibles de demanda: 1000, 2000, 3000, 4000, 5000 juegos de CD.

Se supone que se ha realizado un estudio del mercado a fin de predecir mejor la demanda. La encuesta ha producido los resultados siguientes:

Demanda	Probabilidad
1.000	0,10
2.000	0,30
3.000	0,40
4.000	0,15
5.000	0,05

(i) Represente el problema bajo la forma de un árbol de decisión y distinga las 26 situaciones finales posibles.
(ii) Calcule, para cada decisión de producción, la ganancia esperada:

Producción	Ganancia Esperada
1.000	
2.000	
3.000	
4.000	
5.000	

(iii) Reduzca el árbol reemplazando las ramas relativas a las diversas posibilidades de demanda por la ganancia esperada que le corresponde. Por lo tanto, el problema fue llevado a una elección entre las 6 acciones posibles, cada una de las cuales tiene una ganancia esperada, con lo cual se puede eligir la alternativa cuya ganancia esperada sea la más elevada. ¿Cuál es dicha alternativa?

(iv) Valor de la información.

En general, es muy interesante poder determinar cuánto más se podría obtener de ganancia teniendo una mejor información sobre el futuro; de allí la necesidad de información adicional.

Se supone que se puede hacer una cierta publicidad del curso de idioma y que se lo puede vender por subscripción, con lo cual se podrá producir conociendo la demanda por anticipado. Por lo tanto, ¿cuánto se podría invertir en publicidad y en organización para que esta estrategia sea rentable?

Se supone, para simplificar, que esto no modifica la demanda, pero permite solamente predecirla. Bajo subscripción, se producirá exactamente la demanda.

Calcule la ganancia esperada si se produce de acuerdo a la demanda obtenida (es el cálculo de la ganancia esperada con información perfecta).

Observación. Por ende, considerando sólo la tabla de probabilidades para cada demanda se pudo calcular anteriormente la ganancia esperada máxima (dada por \$ 350.000 para una producción de 3.000 CD). Por otro lado, con información perfecta, se pudo obtener una ganancia esperada de \$ 450.000. Por lo tanto, la información extra para conocer la demanda exacta vale \$ 100.000, es decir, es el monto máximo que se podrá pagar para el conjunto de gastos relativos a la venta por subscripción.

X. UN MODELO SIMPLE DE FIDELIDAD DE MARCA Y SUS CONSECUENCIAS

A continuación se presenta un modelo de fidelidad o lealtad de los clientes hacia una determinada marca para un dado producto para el cual existen sólo dos marcas en competencia. Se obtienen resultados de interés en marketing y se dan algunos consejos que deben tenerse en cuenta por las empresas que comercializan dichas marcas a los efectos de no perder participación en el mercado. Se sigue el análisis efectuado en [Ta1].

Un modelo simple de fidelidad de marca y sus consecuencias

por
Domingo Alberto Tarzia
Investigador Principal del CONICET
Director Depto. Matemática
U.A. - F.C.E.

"La importancia del presente modelo radica en el hecho de conocer cómo varía el mercado período a período y como se estabiliza..."

Los analistas de mercado se interesan en la preferencia de un cliente hacia una determinada marca, y en el efecto que tiene esa fidelidad o lealtad en la participación de cada marca en el mercado de un determinado producto.

En este trabajo se presenta un modelo simple de lealtad a la marca que corresponde a un producto dado P para el cual existen sólo dos marcas «M_1» y «M_2». Se realizan las siguientes suposiciones (una probabilidad del 85% se expresará matemáticamente a través del número real 0,85%):

(S1) Un cliente que compra la marca «M_1» en un período determinado t tiene una probabilidad a (con $0<a<1$) de comprar nuevamente «M_1» y una probabilidad 1 - a de comprar la marca «M_2» (es decir, de cambiar de marca) en el período siguiente t+1;

(S2) Un cliente que compra la marca «M_2» en un período determinado t tiene una probabilidad b (con $0<b<1$) de comprar nuevamente «M_2» y una probabilidad 1 - b de comprar la marca «M_1» (es decir, de cambiar de marca) en el período siguiente t+1;

(S3) Las probabilidades dadas anteriormente en (S1) y en (S2) no varían de período en período;

(S4) El comportamiento del comprador de una marca depende sólo de la compra inmediata anterior y es estadísticamente independiente de las otras compras anteriores;

(S5) En un determinado momento t (condición inicial del presente estudio) la marca «M_1» tiene una porción α (con $0 \leq \alpha \leq 1$) del mercado del producto P (en el sentido de las probabilidades) y por ende la marca «M_2» tiene la porción restante $1 - \alpha$ del mercado del producto P.

Observación 1: De las hipótesis (S1) y (S2) se desprende que los clientes de la marca «M_2» son más leales que los de la marca «M_1» si y sólo si el porcentaje de los que compran la marca «M_2» en el período t y vuelven a comprar la marca «M_2» en el siguiente período t + 1 es mayor que el porcentaje de los que compran la marca «M_1» en el período t y vuelven a comprar la marca «M_1» en el siguiente período t + 1, lo cual se expresa matemáticamente por la siguiente desigualdad: $b > a$.

Esta propiedad es también válida recíprocamente, es decir, que los clientes de la marca «M_1» son más leales que los de la marca «M_2» si y sólo si $a > b$.

Las dos hipótesis (S1) y (S2) se pueden resumir en la siguiente tabla:

Marca comprada en el período t	Marca comprada en el período t + 1	
	Marca «M_1»	Marca «M_2»
Marca «M_1»	a	1 - a
Marca «M_2»	1 - b	b

Luego de presentar el modelo y sus hipótesis, será de mucho interés que se puedan responder las siguientes preguntas:

(1) ¿Qué porcentaje del mercado del producto P tendrán la marca «M_1» y la marca «M_2» en los períodos de tiempo t + 1, t + 2, t + 3, etc.?

(2) ¿El mercado del producto P tiende a

estabilizarse asintóticamente (es decir, cuándo se deja evolucionar el sistema)?

(3) Si la respuesta a (2) es afirmativa, ¿cómo será la convergencia: lenta o rápida?

(4) ¿Las respuestas a las preguntas anteriores (1) a (3) dependerán (y de qué forma) o serán independientes de la posición inicial del mercado del producto P?

Observación 2: Si el mercado del producto P se estabiliza en la posición de equilibrio se tendrá que el número de clientes que dejan la marca «M1» por la marca «M2» quedará balanceado con aquellos que cambian de la marca «M2» hacia la marca «M1».

Observación 3: La respuesta a las preguntas anteriores es de suma importancia para las empresas que producen las dos marcas «M_1» y «M_2» del producto P, a pesar de que en el modelo se supone que las empresas no reaccionan ante cambios del mercado (lo cual obviamente, no es cierto en la realidad). Pero, el presente modelo le brindará a las empresas la necesidad de hacer algo ante un cambio de lealtad a su marca, pues si no reaccionan (no realizan nada en el tiempo) sabrán lo que les sucederá, es decir el porcentaje del mercado que perderán!.

Se obtiene el siguiente resultado:

Teorema: Si x_n representa el porcentaje del mercado que tiene la marca «M_1» al instante $t_n = t+n$ (por lo tanto, $y_n=1-x_n$ representa el porcentaje del mercado que tiene la marca «M_2» al instante t_n) entonces se tienen los siguientes resultados:

R_1) Se tiene la siguiente relación que da la variación de x_n+1 (porcentaje del mercado que tiene la marca «M_1» al instante t_n+1) en función de x_n:

(1) $x_{n+1}=1-b+(a+b-1)\,x_n$, $\forall n \geq 1$,

con la condición inicial $x_1 = a$;

R2) El porcentaje x_{n+1} de la marca «M_1» del mercado del producto P al instante t_{n+1} viene dado en función de los coeficientes α, a y b de la siguiente manera:

(2)
$$x_{n+1}=(1-b)\,\frac{1-(a+b-1)^n}{2-(a+b)}+(a+b-1)^n\,\alpha\ ;$$

R3) El porcentaje x_n de la marca «M_1» del mercado del producto P al instante t_n converge al número $x \in (0,1)$ cuando $n\to+\infty$, el cual no depende del coeficiente α y viene dado en función de los coeficientes a y b de la siguiente manera:

(3)
$$x = \lim_{n\to+\infty} x_n = \frac{1-b}{2-a-b} = \frac{1}{1+\frac{1-a}{1-b}}\,.$$

Demostración: La prueba de estos resultados matemáticos (que no se explicitarán aquí) surgen de la utilización de la teoría de árboles de probabilidades, serie geométrica y su suma, principio de inducción matemática, y, sucesión y su límite.

Observación 4: La relación (2) da la expresión de la sucesión $(x_n)_{n\,\varepsilon N}$ en función de los parámetros a, b y α. Si se tiene el caso particular a+b-1=0 (b=1-a), entonces se deduce que:

(4)
$$x_{n+1}=\frac{1-b}{2-(a+b)}=a\ ,\quad \forall n \geq 1\ ,$$

que no depende del coeficiente α y del número de períodos n. Este hecho es lógico pues al ser a+b=1 el mercado está dividido en dos porciones que no varían: una porción del mercado con porcentaje «a» que compra la marca «M_1» y la otra porción del mercado con porcentaje «1-a» que compra la marca «M_2».

Observación 5: Si se tiene el caso límite en que a=b=1 (es decir, los clientes de ambas marcas son leales al 100%) entonces se deduce que:

(5)
$$x_{n+1}=x_n\ ,\ y_{n+1}=y_n\ ,\forall n \geq 1\ ,$$

que indica que el mercado no varía y permanece constante en la condición inicial, es decir $x_n = \alpha$, $y_n=1-\alpha, \forall n \geq 1$.

Observación 6: Si se tiene el caso particular a=b (ambas marcas tienen la misma fidelidad) entonces se deduce que x=0.50, es decir que las dos marcas «M_1» y «M_2» compartirán el mercado en partes iguales (el 50% para cada una de las dos marcas).

Observación 7: Si se supone que la sucesión $(x_n)_{n\,\varepsilon N}$ tiene límite x, entonces éste puede calcularse de una manera más simple a través de la relación (1) pues al pasar al límite cuando $n\to+\infty$ se tiene que

(6)
$$x=1-b+(a+b-1)\,x$$

de donde surge (3), pero tiene el inconveniente de no poder dar una expresión de x_n en función de n y de los otros parámetros del sistema: a, b y α.

Consecuencias del modelo:

1) El mercado siempre se estabiliza y la convergencia es, en general, rápida (será tanto más rápida como más pequeño sea el número |a+b-1|);

2) El resultado final, al estabilizarse el mercado, no depende del coeficiente α que indica cómo estaba compuesto inicialmente el mercado del producto P con las dos marcas «M_1» y «M_2»;

3) Si a > b (la fidelidad de la marca «M_1» es mayor que la de «M_2») entonces el mercado evolucionará al límite x, dado por la fórmula (3), siendo x > 0.50;

4) Si a=b=1 (ambas marcas tienen el 100% de fidelidad) entonces el mercado no evoluciona y permanece siempre constante con $x_n=\alpha$, $y_n=1-\alpha, \forall n \geq 1$;

5) Si a=b (ambas marcas tiene igual fidelidad) entonces el mercado evolucionará al límte x=0.50 que indica que cada marca tendrá el 50% del mercado.

6) Si a+b=1 (b=1-a) entonces el mercado sólo evoluciona en el primer período al pasar de $x_1=\alpha$ a $x_2=a$ y luego

permanece siempre constante con $x_{n+1}=a$, $y_{n+1}=1-a=b, \forall\, n\geq 1$;

Observación 8: De las consecuencias 2 y 3 anteriores se deduce que la marca que tenga mayor fidelidad obtendrá más del 50% del mercado independientemente de lo que poseía inicialmente.

Ejemplo: Para tener un ejemplo de cómo varía el porcentaje x_n de la marca «M_1» del mercado del producto P (y por ende, el porcentaje $y_n=1-x_n$ de la marca «M_2» del mercado del producto P) se considerará el siguiente caso:

α = 0,50 (el mercado está dividido en partes iguales entre las dos marcas)
a = 0,80 (un 80% de los clientes de la marca «M_1» vuelven a comprar «M_1»)
b = 0,50 (un 50% de los clientes de la marca «M_2» vuelven a comprar «M_2»).

Entonces, aplicando la fórmula (2), para distintos períodos de tiempo, se deduce que:

$x_1=0{,}50$, $x_2=0{,}65$,
$x_3=0{,}695$, $x_4=0{,}7085$,
$x_5=0{,}7126$, $x_6=0{,}7138$,.......
$x=\frac{5}{7} \simeq 0{,}7143$

(la marca «M_1» poseerá el 71,43% del mercado), con lo cual a partir del tercer período (n=3) el mercado se encuentra cerca del valor de estabilización con un error menor al 1%, al ser $x-x_4 \simeq 0{,}0058$.

Al final del trabajo, se presenta una tabla con numerosos casos y en cada uno de ellos se explicitan: $\alpha=x_1$, a, b, x_2, x_3, x_4, x_5, x_6, x_7, x_8 y $x=\lim_{n\to+\infty} x_n$

El ejemplo anterior se presenta como el primer caso.

Conclusión final:

La importancia del presente modelo radica en el hecho de conocer cómo varía el mercado período a período y como se estabiliza (si no se efectúa ningún cambio o ninguna reacción por parte de las empresas que producen las marcas «M_1» y «M_2»). Un estudio análogo para n marcas daría conclusiones más realistas, pero que no cambian la filosofía de lo realizado.

Observaciones y consejos finales:

1) El liderazgo en un mercado puede durar poco si ingresa al mercado un nuevo competidor con mayor fidelidad de marca;

2) Debemos preguntarnos: ¿Cuál es la rotación de nuestra cartera de clientes?

3) Tal vez debamos preocuparnos menos por aumentar nuestra participación en el mercado y más por aumentar la fidelidad de nuestros clientes actuales. S. XXI

x1= alpha	a	b	x2	x3	x4	x5	x6	x7	x8	x=lim xn
0.5	0.8	0.5	0.65	0.695	0.7085	0.71255	0.713765	0.71413	0.714239	0.714286
0.5	0.2	0.5	0.35	0.395	0.3815	0.38555	0.384335	0.3847	0.38459	0.384615
0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
0.2	0.8	0.5	0.56	0.668	0.7004	0.71012	0.713036	0.713911	0.714173	0.714286
0.1	0.8	0.5	0.53	0.659	0.6977	0.70931	0.712793	0.713838	0.714151	0.714286
0.5	0.5	0.8	0.35	0.305	0.2915	0.28745	0.286235	0.285871	0.285761	0.285714
0.5	0.1	0.9	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
0.5	0.1	0.5	0.3	0.38	0.348	0.3608	0.35568	0.357728	0.356909	0.357143
0.4	0.9	0.9	0.42	0.436	0.4488	0.45904	0.467232	0.473786	0.479028	0.5
0.3	0.9	0.9	0.34	0.372	0.3976	0.41808	0.434464	0.447571	0.458057	0.5
0.2	0.9	0.9	0.26	0.308	0.3464	0.37712	0.401696	0.421357	0.437085	0.5
0.01	0.9	0.9	0.108	0.1864	0.24912	0.299296	0.339437	0.371549	0.39724	0.5

X. BIBLIOGRAFÍA

- [CLGGLM] M.C. DE LA CRUZ LOPEZ – C. GONZALEZ GARCIA – J. LLORENTE MEDRANO, "Actividades sobre el azar y la probabilidad", Narcea – MEC, Madrid (1993).
- [Kr] S.G. KRANTZ, "Techniques of problem solving", American Math. Soc., Providence (1997).
- [Li] S. LIPSCHUTZ, "Probabilidad", McGraw Hill, México (1991).
- [Me] P.L. MEYER, 'Probabilidad y aplicaciones estadísticas", Addison-Wesley Iberoamericana, Wilmington (1992).
- [Mi] R. MIATELLO, "Actividades investigativas en la enseñanza", Revista de Educación Matemática, 12 (1997), 19 – 33.
- [Sp] M.R. SPIEGEL, "Probabilidad y estadística", McGraw Hill, México (1991).
- [SpBo] W.A. SPURR – C.H. BONINI, "Toma de decisiones en administración mediante métodos estadísticos", Editorial Limusa, Mexico (1980).
- [Ta1] D.A. TARZIA, "Un simple modelo de fidelidad de marca y sus consecuencias", Revista Siglo XXI, 3 No. 4 (1994), 23 – 25.
- [Ta2] D.A. TARZIA, "Cómo pensar, entender, razonar, demostrar y crear en Matemática", MAT-Serie B, Rosario, # 1 (2000).
- [Th] H. THIRIEZ, "Initiation au calcul économique", Dunod, Paris (1987).