Horacio Bazán.

Tareas y Programas

Probabilidad de Procesos aleatorios de interferencia

Versión 1.0 – Diciembre 19, 2016

B16 - MCIC

Probabilidad de Procesos aleatorios de interferencia (http://http://148.204.64.83/).

Contents

1 Tareas

1.1 Señal WOW

La señal Wow es la denominación por la cual se conoce en círculos astronómicos a una captación de radio que constituiría el único mensaje recibido hasta la fecha que podría tener un origen extraterrestre y haber sido emitido por seres inteligentes.

El 15 de agosto de 1977 a las 23:16, el radiotelescopio Big Ear recibió una señal de radio de origen desconocido durante exactamente 72 segundos proveniente de la zona oriental de la constelación de Sagitario y alcanzando una intensidad 30 veces superior al ruido de fondo.

De acuerdo al protocolo utilizado, esta señal no fue grabada sino que fue registrada por la computadora del observatorio en una sección de papel continuo diseñada para tal efecto. Unos días después, el joven profesor de la Universidad Estatal de Ohio Jerry R. Ehman, que estaba trabajando como voluntario en el proyecto SETI revisando los registros de la computadora, descubrió la señal anómala más intensa que se hubiera detectado hasta entonces por un radiotelescopio. La señal fue conocida como Wow debido a la anotación que Jerry Ehman hizo en el papel continuo, denotando su sorpresa y emoción. La secuencia de dicha señal fue: **6EQUJ5.**

La computadora del radio-observatorio, una IBM 1130 equipada con 1 MB de disco duro y 32 KB de memoria RAM, se encargaba de convertir los datos recibidos directamente por el radio-telescopio a una serie de caracteres alfanuméricos. El software, diseñado por Bob Dixon y Jerry Ehman, era bastante sofisticado, ya que hacía continuos chequeos del funcionamiento del equipo y era capaz de ejecutar varios algoritmos de búsqueda simultáneamente, incluidos unos algoritmos de búsqueda capaces de aislar señales pulsantes o continuas.

Además, sirvió para solucionar la falta de espacio en los registros de impresora y el ahorro de tinta ya que se estaban rastreando cincuenta canales en la frecuencia del hidrógeno neutro (1420 MHz). Cada fila representaba los resultados de los datos recogidos durante aproximadamente doce segundos de búsqueda. Eran necesarios diez segundos para obtener las intensidades de todos los canales, y aproximadamente dos segundos para que la computadora procesara los datos recibidos. Las columnas representaban las intensidades para los cincuenta canales en rastreo, de 10kHz de ancho de banda cada uno, con el canal n.1 situado en el extremo izquierdo y el canal n.50 situado en el extremo derecho.

En la actualidad aún se investiga el origen de la señal. Las explicaciones de la señal van desde el mensaje de una civilización extraterrestre inteligente, hasta alguna interferencia cercana al radiotelescopio.

Para detectar con precisión la intensidad de una posible señal, la computadora basaba las mediciones tomando como referencia la medición anterior. Esto se hacía debido a que el ruido de fondo no es constante respecto al tiempo y necesitaban tener en todo momento una referencia actualizada del mismo para poder diferenciar lo que es señal de lo que es el ruido. Este proceso se llevaba a cabo en cinco pasos:

1. Se dividía en seis porciones la señal recibida en cada canal, de las cuales se separaban 1/6 del valor actual y 5/6 del valor anterior y eran separadas para eliminar el ruido de base.

2. El resto era dividido por la desviación estándar computada sobre 60 periodos (porciones de señal), 1/60 del valor actual más 59/60 del valor anterior.

3. El número calculado en el primer paso era dividido por el número calculado en el segundo. Esta operación daba la relación de ruido de la señal.

4. Después se tomaba la parte entera de esta relación de ruido de la señal.

5. Por último, el número entero era impreso con las siguientes modificaciones:

   (a) Si el valor era un 0, se representaba mediante un espacio en blanco;

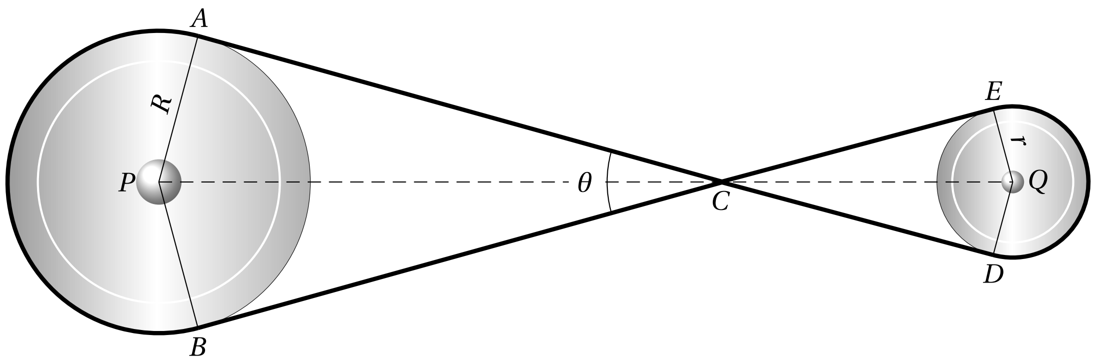
   (b) Los valores entre el 1 y el 9 eran impresos tal cual;

   (c) Los enteros del 10 al 35 eran representados con las letras mayúsculas que van de la A a la Z respectivamente.

Si alguna señal tenía una intensidad de 36 o superior, el programa simplemente empezaba de nuevo desde 0. Así, por ejemplo, el valor 39 sería convertido a 4 (39-35).

La secuencia "6EQUJ5" en el segundo canal del registro de la computadora representaba los siguientes valores de ruido de la señal:

- 6: los valores entre 6.0 y 6.999...
- E: los valores entre 14.0 y 14.999...
- Q: los valores entre 26.0 y 26.999...
- U: los valores entre 30.0 y 30.999...
- J: los valores entre 19.0 y 19.999...
- 5: los valores entre 5.0 y 5.999..



1.2 Efecto Peltier

En la naturaleza, los materiales están formados por moléculas compuestas por átomos enlazados entre sí. Según el tipo de enlace atómico y molecular, los electrones exteriores de cada átomo tienen mayor o menor posibilidad de moverse alrededor de los núcleos y electrones más internos. En los conductores, metales puros y aleaciones, los electrones exteriores menos ligados, pueden moverse en todo el material como si no pertenecieran a ningún átomo. Estos "electrones libres" tienen una distribución de energía que depende principalmente de la temperatura y del tipo de átomos que compone el metal.

Lo anterior es lo que da origen a los 3 "Efectos Termoeléctricos" **(Seebeck, Peltier y Thomson)**. El primer trabajo más directamente relacionado con ellos, fue el manuscrito que en marzo del año 1800 el físico italiano Alessandro Giuseppe Antonio Anastasio Volta (1745-1827) envió a la London Royal Society describiendo el desarrollo de la hoy conocida como Pila de Volta, formada por placas alternadas de metales diferentes en agua salada o ácida. Volta interpretó correctamente el trabajo de su amigo, el médico Galvani, y en su honor, lo denominó "Galvanismo".

Cuando se ponen en contacto 2 metales A y B diferentes, en la unión fluyen electrones en una dirección hasta equilibrar las fuerzas eléctricas debidas a la distribución inicial desigual. Esto hace que aparezca una diferencia de voltaje o "potencial de contacto" entre los metales (Efecto Volta), ya que uno quedó cargado negativamente por los electrones que recibió, y el otro cargado positivamente por la falta de los que perdió.

Volta descubrió este fenómeno en 1793 y estableció que de la siguiente serie de metales: (+) Rb K Na Al Zn Pb Sn Sb Bi Fe Cu Ag Au Pt (-), poniendo en contacto dos cualesquiera de ellos, el de la izquierda es el que se carga positivamente.

En 1815, el francés fabricante de relojes Jean Charles Athanase Peltier (1785-1845), a los 30 años decide dedicar su tiempo a la investigación. En particular, en 1820 Orsted descubre la interacción entre una corriente eléctrica y el magnetismo, Ampere demuestra y formula mateméticamente la interacción entre 2 corrientes, y Biot y Savart descubren que la intensidad del campo magnético producido por una corriente es inversamente proporcional a la distancia del conductor.

Después de la explosión de descubrimientos de ese año, en 1821 el físico alemán Thomas Johann Seebeck (1770-1831) descubre que al colocar a diferente temperatura las uniones de un lazo formado por dos metales distintos (cobre y bismuto), aparece una corriente eléctrica, que dependía de la diferencia de temperatura entre las uniones. Éste es el principio físico de los termopares utilizados en termometría **("Efecto Seebeck")**.

Dentro de la importante serie de descubrimientos de esos años, en 1834 Peltier descubre el fenómeno inverso al Seebeck, el "Efecto Peltier", por el cual, una corriente eléctrica que atraviesa las uniones de un lazo formado por dos metales diferentes, dependiendo del sentido de la corriente, genera calor en una unión y lo absorbe en la otra.

Entonces, concretamente, el principio físico del "efecto Peltier" es *que al conectar una fuente de corriente a un lazo formado por 2 conductores A y B, en una unión la corriente que va desde A hacia B es favorecida por el potencial de contacto, y en la otra, la corriente que va desde B hacia A debe vencer una barrera de energía debida al potencial de contacto opuesto.*

Por lo tanto, la corriente al atravesar las uniones, en una libera calor, y en la otra lo absorbe del medioambiente. Debido a que los metales tienen distribuciones electrónicas similares, los potenciales de contacto

son muy bajos (del orden de 100 mV) y el bombeo de calor mediante el efecto Peltier entre metales es muy pequeño. Esta es la razón por la que no se utilizó este fenómeno en refrigeradores hasta la segunda mitad del Siglo XX.

En efecto, con la invención del transistor en 1947 y el desarrollo de la Física del Estado Sólido, se fabricaron semiconductores donde las distribuciones electrónicas se controlan artificialmente agregando impurezas. Según el elemento de la impureza, el material se convierte en un semiconductor tipo-p donde los portadores de carga son positivos, o tipo-n con portadores negativos.

Deducción de $\Delta V = \Delta T + IR$ para un módulo Peltier

Aquí se demuestra que la caída de voltaje en un módulo Peltier no es simplemente $\Delta V = IR$, ya que existe además una contribución termoeléctrica: el potencial de contacto a ΔT , o voltaje Seebeck.

La potencia eléctrica total (en watt, W) entregada por la fuente que energiza al módulo Peltier, está dada por la caída de voltaje ΔV (en volt, V) y la corriente I (en ampere, A):

$$P_{total} = \Delta V I \tag{1.1}$$

Debido a la conservación de la energía, esta cantidad será igual a la potencia calórica (tasa de flujo de calor) total disipada por el módulo, una a través de la placa caliente H, y otra a través de la placa fría C. Es decir:

$$\Delta V I = P_H + P_C \tag{1.2}$$

donde $P_H > 0$ es la potencia calórica disipada desde H hacia el exterior, y donde $P_C < 0$ es la potencia calórica disipada desde C hacia el exterior. Esta última es negativa porque la placa absorbe calor del medio ambiente, en vez de entregárselo.

1.3 Conceptos

1.3.1 Probabilidad

La probabilidad es un método por el cual se obtiene la frecuencia de un acontecimiento determinado mediante la realización de experimentos aleatorios, de los que se conocen todos los resultados posibles, bajo condiciones suficientemente estables. La probabilidad es un evento o suceso que puede ser improbable, probable o seguro.

La teoría de probabilidades se ocupa de asignar un cierto número a cada posible resultado que pueda ocurrir en un experimento aleatorio, con el fin de cuantificar dichos resultados y saber si un suceso es más probable que otro.

1.3.2 Probabilidad Clásica

Es el número de resultados favorables a la presentación de un evento dividido entre el número total de resultados posibles. Asignación de probabilidad “a priori”, sin necesidad de realizar el experimento.

La probabilidad clásica o teórica se aplica cuando cada evento simple del espacio muestral tiene la misma probabilidad de ocurrir.

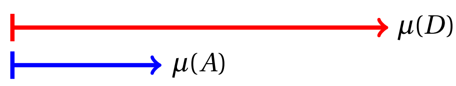
$$P(evento) = \frac{Resultados_{favorables}}{Eventos_{posibles}} \tag{1.3}$$

1.3.3 Probabilidad Geométrica

Los sucesos aleatorios que tienen lugar en un espacio muestral continuo pueden requerir imagenes geométricas por al menos dos razones : debido a la naturaleza del problema o por el de la naturaleza de la solución.

Algunos problemas, como pájaros en un alambre o el problema del palo roto en tres pedazos, por su naturaleza , surgen en un entorno geométrico. En general , podemos pensar en probabilidades geométricas como cantidades no negativas (que no excedan 1) siendo asignados a las subregiones de un tema de un dominio con ciertas reglas. Si la función μ es una expresión de esta asignación definida en un dominio D , entonces, por ejemplo , se requiere:

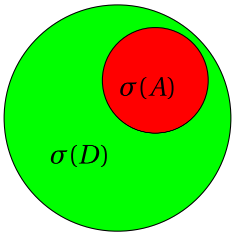
$$0 \leq \mu(A) \leq 1 \text{ donde } A \subset D, \text{ y } \mu(D) = 1 \tag{1.4}$$



Si definimos entonces: $\sigma(A)$ como el área del conjunto A y $\sigma(D)$ como el área de todo el Dominio, tendremos entonces:

$$\mu(A) = \frac{\sigma(A)}{\sigma(D)} \tag{1.5}$$

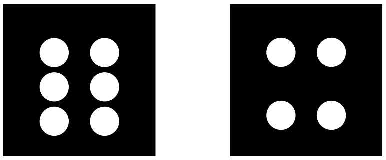
Al igual que implica el caso de las áreas de un circulo:



1.3.4 Probabilidad Frecuentista.

Se entiende por probabilidad frecuentista que por cuantas más veces se repita el experimento, al final las posibilidades de que ocurra cada uno de los sucesos será regular. Aunque cualquier comportamiento sea aleatorio, por proceso empírico llegaremos a una regularidad. Es cuando se lanza un dado y suponiendo cuantas veces cae el número que se seleccionó.

La estadística que estamos acostumbrados a utilizar es la estadística frecuentista, que es la que se desarrolla a partir de los conceptos de probabilidad y que se centra en el cálculo de probabilidades y los contrastes de hipótesis.



Por tanto, la forma de calcular la probabilidad es usar la frecuencia relativa, ya que si se trata de un experimento aleatorio en el cual se repite muchas veces, la frecuencia relativa se acercara mucho a la probabilidad del suceso $P(S)$.

1.3.5 Probabilidad Subjetiva

Es una forma de cuantificar por medio de factores de ponderación individuales, la probabilidad de que ocurra cierto evento, cuando no es posible de cuantificar de otra manera mas confiable.

$$P(E) = \frac{Número_{eventos}}{Total_{eventos}} \tag{1.6}$$

La probabilidad subjetiva se puede definir como la probabilidad asignada a un evento por parte de un individuo, basada en la evidencia que se tenga disponible. Esa evidencia puede presentarse en forma de frecuencia relativa de presentación de eventos pasados o puede tratarse simplemente de una creencia meditada.

1.3.6 Probabilidad Axiomática

Los axiomas de probabilidad son las condiciones mínimas que deben verificarse para que una función que definimos sobre unos sucesos determine consistentemente valores de probabilidad sobre dichos sucesos.

La probabilidad P de un suceso E, denotada por P(E), se define con respecto a un “universo” o espacio muestral Ω, conjunto de todos los posibles sucesos elementales, tal que P verifique los Axiomas de Kolmogoróv, enunciados por el matemático ruso de este nombre en 1933. En este sentido, el suceso E es, en términos matemáticos, un subconjunto de Ω

- Axiomas:

  1. $P(A) \geq 0$;
  2. $P(\Omega) = 1$
  3. P(A)= $\sum_{n=1}^{\infty} P(A_n)$

1.3.7 Posibilidad

Posibilidad es un término que proviene del lat'in *possibilitas*. El concepto hace referencia a la potencia, aptitud u ocasión para ser o existir algo, es decir, para que algo suceda.

Por ejemplo: un partido de fútbol tiene tres resultados posibles. Puede ganar el equipo local, ganar el equipo visitante o que se produzca un empate. No existe otra posibilidad. En otras palabras, no es posible que un partido finalice con un resultado que no sea alguno de los tres mencionados.

1.3.8 Proceso estocástico

En estadística, y específicamente en la teoría de la probabilidad, un proceso estocástico es un concepto matemático que sirve para tratar con magnitudes aleatorias que varían con el tiempo, o más exactamente para caracterizar una sucesión de variables aleatorias (estocásticas) que evolucionan en función de otra variable, generalmente el tiempo. Cada una de las variables aleatorias del proceso tiene su propia función de distribución de probabilidad y pueden o no, estar correlacionadas entre ellas.

Cada variable o conjunto de variables sometidas a influencias o efectos aleatorios constituye un proceso estocástico. Un proceso estocástico X_t puede entenderse como una familia uniparamétrica de variables aleatorias indexadas mediante el tiempo t. Los procesos estocásticos permiten tratar procesos dinámicos en los que hay cierta aleatoriedad.

1.3.9 Riesgo.

Riesgo es una medida de la magnitud de los daños frente a una situación peligrosa. El riesgo se mide asumiendo una determinada vulnerabilidad frente a cada tipo de peligro.

Más informalmente se habla de riesgo para hablar de la ocurrencia ante un potencial perjuicio o daño para las unidades, personas, organizaciones o entidades. Cuanto mayor es la vulnerabilidad mayor es el riesgo, pero cuanto más factible es el perjuicio o daño, mayor es el peligro. Por tanto, el riesgo se refiere sólo a la teórica “posibilidad de daño” bajo determinadas circunstancias, mientras que el peligro se refiere sólo a la teórica “probabilidad de daño” bajo esas circunstancias. Por ejemplo, desde el punto de vista del riesgo de daños a la integridad física de las personas, cuanto mayor es la velocidad de circulación de un vehículo en carretera mayor es el “riesgo de daño” para sus ocupantes, mientras que cuanto mayor es la imprudencia al conducir mayor es el “peligro de accidente” (y también es mayor el riesgo del daño consecuente).

1.3.10 Peligrosidad

La peligrosidad generalmente se refiere a la probabilidad de ocurrencia de una situación peligrosa. No debe confundirse peligrosidad con vulnerabilidad o riesgo que son conceptos relacionados pero diferentes del de peligrosidad.

Así se dice que un determinado conjunto A de situaciones o ubicaciones presentan mayor peligrosidad que otro conjunto de situaciones o ubicaciones B si se cumple la siguiente relación entre probabilidades condicionales que:

$$Prob(P \mid A) > Prob(P \mid B) \text{ ,donde } P \text{ es un evento potencialmente peligroso.} \tag{1.7}$$

1.3.11 Vulnerabilidad

El concepto de vulnerabilidad es diferente del de peligrosidad, la vulnerabilidad se definiría como la probabilidad de unos determinados daños D dado el hecho de que se produce un evento peligroso P en forma de probabilidad condicionada, la vulnerabilidad frente a daños de tipo D es:

$$V_{D|P} = Prob(D \mid P) \tag{1.8}$$

1.3.12 Amenaza

Las amenazas son aquellas que incluyen actos dirigidos, deliberados y sucesos no dirigidos, aleatorios o impredecibles (como puede ser un rayo).

Amenaza es la causa de riesgo que crea aptitud dañina sobre personas y bienes. En el ámbito económico las amenazas latentes (con posibilidad de existencia) es, por ejemplo, la causa origen de pérdida de dinero por baja de las cotizaciones de la bolsa, mientras que el riesgo de pérdida de las acciones es la posibilidad de daño monetario.

1.3.13 Función Factorial

El factorial de un entero positivo n, el factorial de n o “n factorial” se define en principio como el producto de todos los números enteros positivos desde 1 (es decir, los números naturales) hasta n. Por ejemplo:

Ejemplo 1.1

$n! = 1 \times 2 \times 3 \times ... \times (n-1) \times n$
$3! = 3 \times 2 \times 1 = 6$
$5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$

- $n! = 1 \times 2 \times 3 \times ... \times (n-1) \times n$
- $3! = 3 \times 2 \times 1 = 6$
- $5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$

1.3.14 Función Gamma

La función Gamma se define para todo número complejo z cuya parte real positiva de la siguiente forma:

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dz \tag{1.9}$$

Esta definición puede extenderse $\forall z \in \mathbb{C} - \mathbb{Z}^-$, siendo $\mathbb{Z}^-$ el conjunto de los números enteros negativos.

1.3.15 Experimentos Deterministas

Un experimento determinista es aquel que produce el mismo resultado cuando se le repite bajo las mismas condiciones. Ejemplos:

Ejemplo 1.2

1. Tirar piedras a al aire y todas caen.
2. Calentar objetos de metal y todos consiguen temperatura mayor a la temperatura ambiente.
3. Que salga el sol.
4. Que se haga de noche.
5. Un dado se tenga un resultado $n \mid 1 \leq n \leq 6$
6. De un mazo de 52 cartas de poker obtener alguna de la 4 diferentes clases de cartas (trebol, espada, diamante, corazón.)
7. En un jego de Fútbol soccer obtener alguno de los tres resultados (empate, victoría, derrota).
8. Al tener un bit que este sea 1 o 0.
9. Que al nacer un ser humano sea hombre o mujer.
10. Que al llegar en automovil hasta un semaforo se obtenga alguno de los tres colores (verde, amarillo, rojo).

1.3.16 Experimentos Aleatorios

Un experimento aleatorio es aquel que, cuando se le repite bajo las mismas condiciones, el resultado que se observa no siempre es el mismo y tampoco es predecible.

Ejemplo 1.3

1. Lanzar una moneda.
2. El número que se obtendra al lanzar un dado.
3. El resultado de un juego de fútbol.

4. El número de la lotería.
5. El caballo que ganará en una carrera.
6. La carta que se obtendra de un mazo de 52 cartas.
7. Personas que acabaran un curso.
8. El clima.
9. Sacar una bola de color especifico de un recipiente con varias bolas de colores.
10. Encotrar a alguien conocido en un evento.

1.4 Teorema del Binomio

El teorema del binomio esta dado por la ecuación:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k \tag{1.10}$$

Y el coeficiente Binomial es:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{1.11}$$

Ejemplo 1.4

Para n=10 entonces:

$$(x+y)^{10} = \sum_{k=0}^{n=10} \frac{10!}{k!(10-k)!} x^{10-k} y^k$$

$$(x+y)^{10} = \frac{10!}{0!(10-0)!} x^{10-0} y^0 + \frac{10!}{1!(10-1)!} x^{10-1} y^1 + \frac{10!}{2!(10-2)!} x^{10-2} y^2$$

$$+ \frac{10!}{3!(10-3)!} x^{10-3} y^3 + \frac{10!}{4!(10-4)!} x^{10-4} y^4 + \frac{10!}{5!(10-5)!} x^{10-5} y^5 + \frac{10!}{6!(10-6)!} x^{10-6} y^6$$

$$+ \frac{10!}{7!(10-7)!} x^{10-7} y^7 + \frac{10!}{8!(10-8)!} x^{10-8} y^8 + \frac{10!}{9!(10-9)!} x^{10-9} y^9 + \frac{10!}{10!(10-10)!} x^{10-10} y^{10}$$

Finalmente al sumar todos los términos tenemos:

$$(x+y)^{10} = x^{10}+10x^9y+45x^8y^2+120x^7y^3+210x^6y^4+255x^5y^5+210x^4y^6+120x^3y^7+45x^2y^8+10xy^9+y^{10}$$

2 Combinatoria

¿Qué diferencia hay entre Combianaciones y Permutaciones?

Normalmente usamos la palabra "combinación" descuidadamente, sin pensar en si el orden de las cosas es importante. En otras palabras:

“Mi ensalada de frutas es una combinación de manzanas, uvas y bananas”: no importa en qué orden pusimos las frutas, podría ser “bananas, uvas y manzanas” o “uvas, manzanas y bananas”, es la misma ensalada.

“La combinación de la cerradura es 472”: ahora sí importa el orden. “724” no funcionaría, ni “247”. Tiene que ser exactamente 4-7-2.

2.1 Permutaciones

Hay dos tipos de permutaciones:

1. Se permite repetir: como la cerradura de arriba, podría ser “333”.
2. Sin repetición: por ejemplo los tres primeros en una carrera. No puedes quedar primero y segundo a la vez.

Permutaciones con repetición

Son las más fáciles de calcular. Si tienes n cosas para elegir y eliges r de ellas, las permutaciones posibles son:

$$(n_1)(n_2)\ldots(n_r) = n^r \tag{2.1}$$

donde $\boldsymbol{n}$ es el número de cosas que puedes elegir, y eliges $\boldsymbol{r}$ de ellas, (Se puede repetir, el orden importa)

Permutaciones sin repetición

En este caso, se reduce el número de opciones en cada paso. Para describir este tipo de probabilidades usaremos la función factorial, las permutaciones posibles sin repeticion son:

$$\frac{n!}{(n-r)!} \tag{2.2}$$

donde ***n*** es el número de cosas que puedes elegir, y eliges ***r*** de ellas, (NO e puede repetir, el orden importa)

Ejemplo 2.1

Calcular el número de posibilidades que puede haber si se toman 3 numero de los cuales pueden ser 0 a 9.

Solución:
Sea ***n*** el conjunto de n"umeros que podemos tomar (0,1,...,0) y ***r*** el número de datos que se van a tomar:

$$(n_1)(n_2)(n_3) = (10)(10)(10) = 10^3 = 1000 \tag{2.3}$$

De lo anterior concluimos que tenemos 1000 posibilidades y que los n'umeros se pueden repetir.

Ejemplo 2.2

Calcular el número de posibilidades que puede haber si se toman 5 n'umeros de una bolsa de los cuales pueden ser 0 a 4.

Solución:
Sea ***n*** el conjunto de n"umeros que podemos tomar (0,1,2,3,4) y ***r*** el número de datos que se van a tomar:

$$(n_1)(n_2)(n_3)(n_4)(n_5) = (5)(5)(5)(5)(5) = 5^5 = 3125 \tag{2.4}$$

De lo anterior concluimos que tenemos 3125 posibilidades y que los n'umeros se pueden repetir.

Ejemplo 2.3

Calcular el número de posibilidades en que se pueden ordenar 15 bolas de billar.

Solución:

Sea $\boldsymbol{n}$ el conjunto de n"umeros que podemos tomar (1,2,...,15) y $\boldsymbol{r}$ el número de datos que se van a tomar, en este caso todos:

$$(n_1)(n_2)\dots(n_{15}) = (1)(2)\dots(15) = 15! = 1307674368000 \tag{2.5}$$

De lo anterior concluimos que tenemos 1307674368000 posibilidades de ordenar las 15 bolas de billar.

Ejemplo 2.4

Calcular el número de posibilidades en que se pueden ordenar 3 bolas de billar.

Solución:
Sea $\boldsymbol{n}$ el conjunto de n"umeros que podemos tomar (1,2,...,15) y $\boldsymbol{r}$ el número de datos que se van a tomar, en este caso 3:

$$\frac{n!}{(n-r)!} = \frac{15!}{(15-3)!} = \frac{15!}{12!} = 2730 \tag{2.6}$$

De lo anterior concluimos que tenemos 2730 posibilidades de ordenar las 3 bolas de billar de 15.

Ejemplo 2.5

De cuántas maneras se pueden dar primer y segundo premio entre 10 personas.

Solución:
Sea $\boldsymbol{n}$ el conjunto de personas que podemos que pueden obtener el primer o segundo lugar (1,2) y $\boldsymbol{r}$ el número de datos que se van a tomar, en este caso 2:

$$\frac{n!}{(n-r)!} = \frac{10!}{(10-2)!} = \frac{10!}{8!} = 90 \tag{2.7}$$

De lo anterior concluimos que tenemos 90 posibilidades de el primer y segundo lugar entre diez personas.

Ejemplo 2.6

De cu'antas formas distintas pueden sentarse ocho personas en una fila de butacas.

Solución:

Considerando que no se repiten los elementos. Una persona no se puede repetir.

$$(n_1)(n_2)\dots(n_8) = (1)(2)\dots(8) = 8! = 40320 \tag{2.8}$$

De lo anterior concluimos que tenemos 40320 posibilidades de ordenar las 8 personas en una fila de butacas.

Ejemplo 2.7

De cuántas formas pueden colocarse los 11 jugadores de un equipo de fútbol teniendo en cuenta que el portero no puede ocupar otra posición distinta que la portería.

Solución:
Considerando que el portero no puede ocupar otra posición. Disponemos de 10 jugadores que pueden ocupar 10 posiciones distintas.

$$(n_1)(n_2)\dots(n_{10}) = (1)(2)\dots(10) = 10! = 3628800 \tag{2.9}$$

De lo anterior concluimos que tenemos 3628800 posibilidades de ordenar el equipo de futbol.

Ejemplo 2.8

Una mesa presidencial está formada por ocho personas, de cuántas formas distintas se pueden sentar, si el presidente y el secretario siempre van juntos.

Solución:
Se forman dos grupos el primero de 2 personas y el segundo de 7 personas, en los dos se cumple que:

$$(P_2)(P_7) = (2)(7!) = 10080 \tag{2.10}$$

De lo anterior concluimos que tenemos 10080 posibilidades de ordenar la mesa.

Ejemplo 2.9

Cuatro libros distintos de matemáticas, seis diferentes de física y dos diferentes de química se colocan en un estante. De cuántas formas distintas es posible ordenarlos si: Los libros de cada asignatura deben estar todos juntos..

Solución:
Se forman dos grupos el primero de 2 personas y el segundo de 7 personas, en los dos se cumple que:

$$(P_4)(P_6)(P_2)(P_3) = (4!)(6!)(2!)(3!) = 207360 \tag{2.11}$$

De lo anterior concluimos que tenemos 207360 posibilidades de ordenar los libros y que las asignaturas queden juntas.

Ejemplo 2.10

Cuatro libros distintos de matemáticas, seis diferentes de física y dos diferentes de química se colocan en un estante. De cuántas formas distintas es posible ordenarlos si: Solamente los libros de matemáticas deben estar juntos.

Solución:
Se forman dos grupos el primero de 2 personas y el segundo de 7 personas, en los dos se cumple que:

$$(P_4)(P_9) = (4!)(9!) = 8709120 \tag{2.12}$$

De lo anterior concluimos que tenemos 8709120 posibilidades de ordenar los libros y que los de matemáticas queden juntos.

2.2 Combinaciones

Existen dos tipos de combinaciones (recuerda que ahora el orden no importa):

1. Se puede repetir: como monedas en tu bolsillo (5,5,5,10,10)
2. Sin repetición: como números de lotería (2,14,15,27,30,33).

Combinaciones sin repetición

Así que sólo tenemos que ajustar nuestra fórmula de permutaciones para reducir por las maneras de ordenar los objetos elegidos (porque no nos interesa ordenarlos):

$$\frac{n!}{(n-r)!} * \frac{1!}{r!} = \frac{n!}{r!(n-r)!} = \binom{n}{r} \tag{2.13}$$

donde ***n*** es el número de cosas que puedes elegir, y eliges ***r*** de ellas, (NO se puede repetir, el orden NO importa)

Demostracíon
El argumento presentado puede generalizarse de la siguiente forma. Si se tiene un conjunto con **n** elementos, de los cuales se van a escoger **k** de ellos, la elección (ordenada) puede hacerse de la siguiente forma:

$$n(n-1)(n-2)\dots(n-k+1) \tag{2.14}$$

Ya que en el primer paso se tienen **n** opciones, en el segundo se tienen **n-1**, en el tercero **n-2**, y así sucesivamente, terminando en el paso **k** que tendrá **n-k+1** opciones.

Ahora, hay que dividir el producto anterior entre el número de selecciones "equivalentes". Pero si se tiene **k** objetos, hay **k!** formas de permutarlos, es decir, **k!** formas de listarlos en distinto orden. Recordemos que **k!** se lee k-factorial y es igual a:

$$k! = (1)(2)(3)\dots(k) \tag{2.15}$$

Concluimos que el número de subconjuntos con **k** elementos, escogidos de un conjunto con **n** elementos es:

$$\binom{n}{r} = \frac{n(n-1)(n-2)\dots(n-k+1)}{(1)(2)(3)\dots(k-1)k} \tag{2.16}$$

Multiplicando el numerador y el denominador de la fracción por (1)(2)(3)... (n-k)

$$\binom{n}{r} = \frac{n(n-1)(n-2)\dots(n-k+1)}{(1)(2)(3)\dots(k-1)k} \tag{2.17}$$

Combinaciones con repetición

Sea A un conjunto con n elementos y m un natural menor o igual que n. Llamamos combinación con repetición de m elementos de A a todo subconjunto de m elementos de A en el que un elemento puede aparecer hasta m veces. En este caso sólo nos importa la naturaleza, no el orden y además podemos repetir elementos.

El número de combinaciones con repetición viene dado por:

$$\binom{n+r-1}{r} = \frac{(n+r-1)!}{r!(n-1)!} \tag{2.18}$$

donde ***n*** es el número de cosas que puedes elegir, y eliges ***r*** de ellas, (Se puede repetir, el orden NO importa)

Ejemplo 2.11

En una clase de 35 alumnos se quiere elegir un comit'e formado por tres alumnos. ¿Cuántos comit'es diferentes se pueden formar?.

Solución:
Considerando que:

- No entran todos los elementos.
- No importa el orden: Juan, Ana.
- No se repiten los elementos.

$$C_r^n = C_{35}^3 = 6545 \tag{2.19}$$

De lo anterior concluimos que tenemos 5445 posibilidades de formar los comités.

Ejemplo 2.12

¿De cuántas formas pueden mezclarse los siete colores del arco iris tomándolos de tres en tres?

Solución:
Considerando que:

- No entran todos los elementos.
- No importa el orden.
- No se repiten los elementos.

$$C_r^n = C_7^3 = 35 \tag{2.20}$$

De lo anterior concluimos que tenemos 35 posibilidades de combinar los colores.

Ejemplo 2.13

A una reunión asisten 10 personas y se intercambian saludos entre todos. ¿Cuántos saludos se han intercambiado?

Solución:
Considerando que:

- No entran todos los elementos.
- No importa el orden.
- No se repiten los elementos.

$$C_r^n = C_{10}^2 = 45 \tag{2.21}$$

De lo anterior concluimos que tenemos 45 saludos se han dado.

Ejemplo 2.14

En una bodega hay en un cinco tipos diferentes de botellas. De ¿cuántas formas se pueden elegir cuatro botellas?

Solución:
Considerando que:

- No entran todos los elementos. Sólo elije 4..
- No importa el orden. Da igual que elija 2 botellas de anís y 2 de ron, que 2 de ron y 2 de anís.
- Sí se repiten los elementos. Puede elegir más de una botella del mismo tipo.

$$C_r^n = C_5^4 = 45 = \frac{(5+4-1)!}{4!(5-1)!} = \frac{8!}{4!4!} = 70 \tag{2.22}$$

De lo anterior concluimos que tenemos 70 de elegir las botellas de vino.

Ejemplo 2.15

¿Cuántas apuestas de Lotería Primitiva de una columna han de rellenarse para asegurarse el acierto de los seis resultados, de 49?

Solución:
Considerando que:

- No entran todos los elementos.
- No importa el orden.
- No se repiten los elementos.

$$C_r^n = C_{49}^6 = 13983816 \tag{2.23}$$

De lo anterior concluimos que tenemos 13983816 apuestsa de loteria.

Ejemplo 2.16

¿Cuántas diagonales tiene un pentágono y cuántos triángulos se puede informar con sus vértices?

Solución: Vamos a determinar en primer lugar las rectas que se pueden trazar entre 2 vértices. Considerando que:

- No entran todos los elementos.
- No importa el orden.
- No se repiten los elementos.

$$C_r^n - 5 = C_5^2 - 5 = 5 \tag{2.24}$$

De lo anterior concluimos que tenemos 5 diagonales en un pentágono, se restaron los lados que determinan 5 rectas que no son diagonales.

Para determinar el número de triángulos dentro de un pentágono:

$$C_r^n = C_5^3 = 10 \tag{2.25}$$

De lo anterior concluimos que tenemos 10 triángulos en un pentágono.

Ejemplo 2.17

Un grupo, compuesto por cinco hombres y siete mujeres, forma un comité de 5 hombres y 3 mujeres. De cuántas formas puede formarse, si:

Considerando que:

1. Puede pertenecer a él cualquier hombre o mujer.

   Solución:

$$C_r^n = C_5^2 * C_7^3 = 10 * 35 = 350 \tag{2.26}$$

2. Una mujer determinada debe pertenecer al comité.

   Solución:

$$C_r^n = C_5^2 * C_6^2 = 10 * 15 = 150 \tag{2.27}$$

3. Dos hombres determinados no pueden estar en el ccomité.

   Solución:

$$C_r^n = C_3^2 * C_7^3 = 3 * 35 = 105 \tag{2.28}$$

Ejemplo **2.18**

Una persona tiene cinco monedas de distintos valores. ¿Cuántas sumas diferentes de dinero puede formar con las cinco monedas?

Solución:

$$C_r^n = C_5^1 + C_5^2 + C_5^3 + C_5^4 + C_5^5 = 31 \tag{2.29}$$

De lo anterior concluimos que existen 31 formas de formar las cantidades con 5 monedas.

Ejemplo **2.19**

Calcular el número de combinaciones de 10 elementos tomados de 4 en 4.

Solución:

$$C_r^n = C_{10}^4 = 210 \tag{2.30}$$

De lo anterior concluimos que existen 210 formas combinar estos elementos.

Ejemplo 2.20

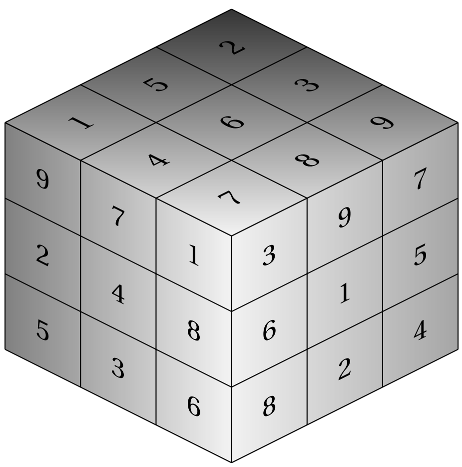
De cuántas formas se pueden ordenar tres bolas de billar de 16.

Solución:

$$C_r^n = C_{16}^6 = 560 \tag{2.31}$$

De lo anterior concluimos que existen 560 formas combinar estas bolas de billar.

3 Lecturas



3.1 Lectura: Principio del Palomar

El principio de palomar, es tambien conocido como principio de Dirichlet o principio de las cajas, en dicho principio se establece que si **n** palomas se distribuyen en **m** palomares, y si **n > m**, entonces al menos habrá un palomar con más de una paloma.

Una versión generalizada de este principio dice que, si n objetos discretos deben guardarse en **m** cajas, al menos una caja debe contener no menos de:

$$\left[\frac{n}{m}\right] \tag{3.1}$$

donde: **n** es el numero de objetos y **m** es el número de cajas

Este principio se utiliza para resolver cuestiones, tales como:

- En un conjunto de 32 personas al menos 2 celebran su cumplea'nos el mismo día del mes.
- Juan regresa de la lavanderia con 12 pares de calcetines, (cada par de distinto color) en una bolsa, al sacar cada calcetin de la bolsa aleatoriamente tendra que sacar cuando mucho trece calcetines para obtener el par.

- Vilma opera una computadora que tiene una unidad de cinta magnética para respaldar la información. Un día le dan una cinta que contiene 600,000 “palabras” de cuatro o menos letras minúsculas. En la cinta las palabras consecutivas se separan con un caracter en blanco. ¿Puede suceder que las 600,000 palabras sean distintas entre sí?.

3.2 Lectura: Envejecimineto - Oveja Dolly

La oveja Dolly (5 de julio de 1996-14 de febrero de 2003) fue el primer mamífero clonado a partir de una célula adulta. Sus creadores fueron los científicos del Instituto Roslin de Edimburgo (Escocia).

En 1999, un equipo de la firma de Biotecnología PPL Therapeutics estudió a fondo diversas estructuras celulares en la oveja y encontró que sus **telómeros**, necesarios para el crecimiento y la reproducción celular, eran ligeramente más cortos en Dolly que en otras ovejas de su edad. Ya desde entonces se especuló con la posibilidad de que Dolly fuera a envejecer más rápidamente de lo normal. Esta es la razón por la que su prematura muerte tiene una gran trascendencia.

Los telómeros (del griego telos, ”final” y meros, ”parte”) son los extremos de los cromosomas. Son regiones de ADN no codificante, altamente repetitivas, cuya función principal es la estabilidad estructural de los cromosomas en las células eucariotas, la división celular y el tiempo de vida de las estirpes celulares.

3.3 Lectura: Telómeros

Con una alta precisión, un grupo de científicos de la Universidad de Anglia del Este en EEUU ha desarrollado un método para calcular la velocidad a la que se envejece y que, por ende, permite pronosticar la esperanza de vida. Las primeras pruebas, llevadas a cabo con éxito en una bandada de aves silvestres, consisten en un simple análisis de sangre, el cual permite medir la duración de la vida de las regiones de ADN ubicadas en los extremos de los cromosomas, denominados telómeros. Estos últimos son los encargados de producir la división celular y determinar el tiempo de vida medio de una célula.

Los telómeros actúan como relojes internos del organismo y proporcionan una estimación precisa de la edad biológica de una persona, la cual es mucho más exacta que la edad real o cronológica a la hora de calcular el envejecimiento y los a’nos que restan de vida, siempre y cuando la muerte se produzca por causas naturales. ”La prueba te dice básicamente si los telómeros de una persona tienen una longitud normal para su grupo de edad, o si son más cortos o más largos de lo normal”, según ha explicado David Richardson, uno de los investigadores principales.Durante la próxima década estos análisis se llevarán a cabo de forma rutinaria. Los resultados del estudio han sido publicados en la revista científica **Molecular Ecology**.

Hoy sabemos que los telómeros son los principales responsables del envejecimiento, pues actúan como relojes internos del organismo: según se acortan, nosotros nos hacemos mayores. En los jóvenes, la longitud de los telómeros es de entre 8.000 y 10.000 nucleótidos (las moléculas que conforman el ADN y el ARN), y se van acortando cada vez que se producen nuevas divisiones celulares. Llegado a un punto el telómero es demasiado estrecho y la célula deja de dividirse o muere. Si los telómeros no se acortaran, o lográramos alargarlos, nuestro envejecimiento biológico se detendría.

En los a’nos treinta los científicos Hermann Joseph Muller y Barbara McClintock descubrieron los telómeros, unos peque’nos ”tapones” que se encuentran en los extremos de los cromosomas para garantizar su estabilidad estructural. El hallazgo les valió un premio Nobel, pero nunca podrían haber imaginado la importancia real de estas peque’nísimas regiones del ADN.

Hoy sabemos que los telómeros son los principales responsables del envejecimiento, pues actúan como relojes internos del organismo: según se acortan, nosotros nos hacemos mayores. En los jóvenes, la longitud de los telómeros es de entre 8.000 y 10.000 nucleótidos (las moléculas que conforman el ADN y el ARN), y se van acortando cada vez que se producen nuevas divisiones celulares. Llegado a un punto el telómero es demasiado estrecho y la célula deja de dividirse o muere.

Hemos encontrado una manera de alargar los telómeros humanos en al menos 1.000 nucleótidos, retrasando el reloj interno de estas células en lo equivalente a muchos a’nos de vida.

De un tiempo a esta parte, numerosos grupos de investigación han tratado de revertir este proceso pues, si los telómeros no se acortaran, o lográramos alargarlos, nuestro envejecimiento biológico se detendría.

En los últimos cincuenta a’nos el estudio de los telómeros ha producido hasta cinco premios Nobel, además de 16.000 artículos científicos. Pese a esto, aunque en diversas ocasiones se ha logrado alargar los telómeros o retrasar el acortamiento de estos, no se había desarrollado una técnica lo suficientemente segura como para ser probada en humanos.

Un equipo de la Universidad de Stanford ha descrito esta semana en la revista FASEB Journal una técnica que, aseguran, es mucho más eficaz y segura que las anteriores, lo que supondría un importante paso en el desarrollo de una terapia válida.

”Hemos encontrado una manera de alargar los telómeros humanos en al menos 1.000 nucleótidos, retrasando el reloj interno de estas células en lo equivalente a muchos a’nos de vida”, explica la doctora **Helen Blau**, profesora de microbiología e inmunología de la Universidad de Stanford y autora principal del estudio. En concreto, tal como especifica la nota de presentación del estudio, los científicos han logrado alargar los telómeros en 1.000 nucleótidos en sólo un par de días, lo que supone un rejuvenecimiento de más de una década.

Los investigadores han utilizado un ácido ribonucleico (RNA) modificado para transmitir instrucciones de los genes del ADN a las células encargadas de fabricar proteínas. El RNA utilizado en el experimento contenía una secuencia codificada de la telomerasa transcriptasa inversa (TER), el componente activo que genera de forma natural la enzima telomerasa, presente sólo en tejidos fetales y ciertas células madre. Esta enzima se asegura de que los telómeros de estas células permanezcan en plena forma en la siguiente generación, pero desaparece tras el nacimiento, lo que provoca que, a partir de entonces, los telómeros se acorten y empecemos a envejecer.

Al reaparecer el TER, los telómeros vuelven a crecer, pero esto puede suponer un problema si no se controla el proceso. Si las células tratadas empiezan a dividirse hasta el infinito pueden volverse muy peligrosas, pues en el proceso es bastante probable que se desarrolle un cáncer.

3.4 Lectura: Maquina Enigma

La criptografá, palabra que procede del griepo krypto (oculto) y graphos (escritura), es la disciplina cientḟica que se encarga del cifrado y descifrado de mensajes, es decir, enmascarar mensajes mediante un algoritmo de ofuscación que, además, debe permitir devolver el mensaje a su estado original.

Hubo un sistema de cifrado que fue usado por Alemania y que tuvo en jaque a los aliados, sobre todo, en el Atlántico Norte, donde los convoys de material procedente de Estados Unidos caían presa de los submarinos alemanes que se comunicaban entre sí utilizando el código que generaba uno de los inventos más fascinantes de esa época, la máquina Enigma.

La máquina Enigma fue inventada por un ingeniero alemán, Arthur Scherbius, un experto en electromecánica que, tras la Primera Guerra Mundial, quiso aplicar la tecnología existente para mejorar los sistemas de criptografía de los ejércitos. Su idea, patentada en febrero de 1918, consistía en aplicar el Cifrado de Vigenère o, dicho de otra forma, se aplicaba un algoritmo de sustitución de unas letras por otras. Como Scherbius no contaba con recursos para fabricarla, se asoció con Willie Korn que tenía una compa'nía llamada Enigma Chiffiermaschinen AG en Berlín. Ambos mejoraron el dise'no y en 1923 la presentaron en la Exhibición Postal Internacional de Berlín para el cifrado de secretos comerciales.

¿En qué consistía la máquina Enigma? La máquina Enigma era un dispositivo electromecánico, es decir, tenía una parte eléctrica y otra mecánica. El mecanismo consistía en una serie de teclas, con las letras del alfabeto, al igual que una máquina de escribir, que en realidad eran interruptores que accionaban los dispositivos eléctricos y hacían mover unos cilindros rotatorios. El funcionamiento, cara al usuario, era bastante sencillo. El operador tenía que teclear las letras de su mensaje y anotar las letras que devolvía la máquina (a través de un alfabeto que se iba iluminando). El código a usar se fijaba con las posiciones de los cilindros que constaban, cada uno, de 26 cables que se conectaban al teclado pero, con la particularidad, que el primer cilindro giraba un veintiseisavo de vuelta después de cada pulsación, de tal manera que la posición de las conexiones iba cambiando con cada entrada del teclado, obteniendo un cifrado polialfabético. Además, para dar mayor robustez, el segundo cilindro sólo daba un giro cuando el primero había completado 26 giros y el tercero cuando el segundo había dado sus correspondientes 26 y a'nadió la posibilidad de que los rodillos pudiesen ser intercambiados de posición, de manera que el número de posibilidades aumentase hasta tener 105.456 alfabetos.

Además, el sistema contaba con 6 cables de conexión que también permitían introducir modificaciones dado que podrían conectarse a 26 lugares (representando a las 16 letras del alfabeto de Enigma) lo que producía 100.391.791.500 maneras distintas de conectar los cables que unidos a los 105.456 alfabetos arrojaba 3.283.883.513.796.974.198.700.882.069.882.752.878.379.955.261.095.623. 685.444.055.315.226.006.433.616.627.409 posibilidades distintas de codificación. Durante la Segunda Guerra Mundial, Alemania contaba con una enorme ventaja porque el código de Enigma era, prácticamente, indescifrable.

La suma de estos factores obtuvo como resultado el descifrado de los mensajes de Enigma y, por tanto, la drástica disminución de las pérdidas Aliadas en el Atlántico Norte. Ante las sucesivas derrotas, los Alemanes evolucionaron Enigma y crearon una nueva máquina, la M4 pero fue vencida gracias a Colossus, un computador dise'nado para descifrar los códigos alemanes.

4 Demmostración combinatoria

4.1 Demostración: Permutación sin Reemplazo

El número de distintas alternativas en que **k** elementos pueden extraerse de manera consecutiva de una población se encuentra dada por:

$$P_s(k;n) = \frac{n!}{(n-k)!}, k \leq n \tag{4.1}$$

Aplicando el principio de la multiplicación con:

$$n_1 = n, n_2 = n-1, \ldots, n_k = n-(k-1) \tag{4.2}$$

Debido a que al efectuar cada extracción de la población disminuye en uno, luego:

$$P_s(k;n) = n(n-1)\ldots[n-(k-1)] = \frac{n!}{(n-K)!} \tag{4.3}$$

Si multiplicamos y dividimos por *(n-k)!*, se llega a la ultima igualdad.

4.2 Demostración: Permutación con Reemplazo

(Estadística de Maxwell - Boltzmann).

$$P_c(k;n) = n^k, \tag{4.4}$$

Aplique el principio de la multiplicación con:

$$n_1 = n_2 = \ldots = n_k = n \tag{4.5}$$

Dado que nunca se acaban los elementos de la población

4.3 Demostración: Combinación sin Reemplazo

(Estadística de Fermi-Dirac) el total de las maneras distintas en que pueden seleccionarse k elementos a la vez de una población de tama'no n es de:

$$C_s(k;n) = \binom{n}{k} = \frac{k!}{k!(n-k)!}, k \leq n \tag{4.6}$$

N Antes de la deducción vale la pena resaltar que una interpretación muy conveniente de las combinaciones sin reemplazo es la de reemplazar el número de subconjuntos de tama'no k que pueden formarse con n elementos.

Con k elementos de la población, '¿Cuántas permutaciones y combinaciones de tama'no k pueden hacerse? Para el caso de las permutaciones basta hacer $n = k$ y para las combinaciones con solo recurrir a la nota para darnos la pauta para seguir, entonces:

$$P_s(k;k) = k! \; y \; C_s(k;K) = 1 \tag{4.7}$$

Así que con los mimos elementos k se puede formar $k!$ ordenaciones de ellos, pero una sola combinación. Ahora si son n elementos de la población habrá $P_s(k;n)$ permutaciones, aunque de éstas hay $k!$ que constan de los mismos k elementos, bantando dividir el $k!$ para eliminar tal multiplicidad.

La fórmula cuya deducción acabamos de efectuar es una de las más relevantes, y existen algunos puntos lo suficientemente importantes como para recalcarlos.

$$\binom{n}{k} = \binom{n}{n-k} \tag{4.8}$$

1. La igualdad es tan evidente que un ligero vistazo a la fórmula es suficiente, aunque su importancia radica en su interpretación. Dónde $C_s(k;n)$ representa el número de subconjuntos de tama'no k de una población de n, luego por cada k elementos de la población automáticamente se dejan $(n-k)$ elementos ella. Así que habrá tantos subconjuntos de tama'no $(n-k)$ como de tama'no k.

2. Si en un lote de n articulos distinguimos entre dos clases distintas, digamos, unidades buenas y defectuosas, ¿de cuántas maneras en una muestra de tama'no k, pueden seleccionarse a j unidades buenas? Suponiendo que hay m artículos buenos en el lote y el resto son defectuosos la respuesta está dada por **la distrubución hipergeometríca**, la cual se deduce a partir del principio de multiplicación y es:

$$\binom{m}{j}\binom{n-m}{k-j} \tag{4.9}$$

Basta suponer que la primera etapa consiste en tomar los j artículos buenos y la segunda de tomar el resto, representado por $(k-j)$ de unidades malas.

4.4 Demostración: Combinación con Reemplazo

(Estadística de Bose-Einstein)

$$C_c(k;n) = \binom{n+k-1}{k} \tag{4.10}$$

N Suponga distinguibilidad entre los elementos de la población, usando demaostración por inducción tenemos lo siguiente:

$$\text{Si } k=1, \qquad C_c(1;n) = n \qquad y \qquad \binom{n}{1} = n \tag{4.11}$$

Primeramente verificamos la siguiente igualdad:

$$C_c(k+1;n) = \sum_{i=1}^{N} C_c(k,i) = (*) \tag{4.12}$$

Debido a la distinguibilidad de los elementos puede suponerse que a estos numerados del 1 al n y que cada muestra no ordenada aparece mostrándo a los elementos que la constituyen de manera creciente. Ahora la muestra podría constar de elementos que fuesen todos iguales, algunos desiguales o todos diferentes y de que cualquier miembro de la población puede serlo también de la muestra. Si dentro de los $(k+1)$ elementos se encuentra el 1 éste ocupa el primer sitio del arreglo, y los k miembros restantes se tomarán de entre n (el 1 inclusive, por haberse restituído) pudiento ésto efectuarse de $C_c(k;n)$ maneras distintas. Si la muestra no coincide al 1, el primer sitio estará ocupado por un número por lo menos igual a 2 y los k restantes números se extraen de entre $(n-1)$, pués se elimina al 1, habiéndo $C_c(k;n-1)$ alternativas de llevar esto a cabo. Repitiendo el proceso hasta el último caso, donde la muestra no contiene elemento alguno de los $(n-1)$ primeros números, se tiene a n como el mínimo número y los otros k se toman de una población de un elemento a saber 'n', seleccionándose estos de $C_c(k;1)$ formas distintas, y concluyéndose así la afirmación.

Además.

$$\binom{N}{j-1} + \binom{N}{j} = \binom{N+1}{j} \tag{4.13}$$

Lo cual no es difícil de probar, ya que si se fija un elemento de la población, por ejemplo x_0, $\binom{N+1}{j}$ representa a todos los subconjuntos de tama'no j de la población. Esta colección consta de subconjuntos que contienen a x_0 y de los que no lo contienen, habiendo $\binom{N}{j-1}$ miembro en la primera clase y $\binom{N}{J}$ en la segunda.

Si $N = k + i + 1 \qquad y \qquad j - 1 = k$, substituyendo en la ecuación anterior resulta:

$$\binom{i+k-1}{k} \qquad + \qquad \binom{i+k-1}{k+1} \qquad = \qquad \binom{i+k}{k+1} \tag{4.14}$$

O equivalente:

$$C_c(k,i) = \binom{i+k}{k+1} \qquad - \qquad \binom{i+k-1}{k+1} \tag{4.15}$$

La cual basta substituir en (*), aunado al hecho de que $\binom{k}{k+1} = 0$, para obtener el resultado desea.

4.5 Tensor de Levi-Civita

En matemáticas, y en particular en cálculo tensorial, se define el símbolo de Levi-Civita, también llamado el símbolo de permutación o tensor de Levi-Civita, como sigue:

$$\epsilon_{ijk} = \begin{cases} 1 & \text{- Permutaciones Pares} \\ 0 & \text{- Repite} \\ -1 & \text{- Permutaciones Impares} \end{cases} \tag{4.16}$$

Calculando los valores del tensor de Levi-Civita:

$$\epsilon_{123} = 1 = \begin{cases} 123 \\ 321 \\ 231 \end{cases} \tag{4.17}$$

$$\epsilon_{123} = -1 = \begin{cases} 321 \\ 132 \\ 213 \end{cases} \tag{4.18}$$

Calculando los valores del tensor de Levi-Civita:
Para cuando i = 1:

$$
\begin{aligned}
(\overline{A}X\overline{B})_i &= \epsilon_{ijk} A_j B_k && (4.19)\\
(\overline{A}X\overline{B})_1 &= \epsilon_{1jk} A_j B_k && (4.20)\\
(\overline{A}X\overline{B})_1 &= \epsilon_{1jk} A_j B_k && (4.21)\\
&= \epsilon_{11k} A_1 B_k + \epsilon_{12k} A_2 B_k + \epsilon_{13k} A_3 B_k && (4.22)\\
&= \epsilon_{111} A_1 B_1 + \epsilon_{112k} A_1 B_2 + \epsilon_{113} A_1 B_3 + && (4.23)\\
& \epsilon_{121} A_2 B_1 + \epsilon_{122k} A_2 B_2 + \epsilon_{123} A_2 B_3 + \epsilon_{131} A_3 B_1 + \epsilon_{132k} A_3 B_2 + \epsilon_{133_{A_3 B_3}} && (4.24)\\
&= \epsilon_{123} A_2 B_3 + \epsilon_{132k} A_3 B_2 && (4.25)
\end{aligned}
$$

Para cuando i = 2:

$$
\begin{aligned}
(\overline{A}X\overline{B})_i &= \epsilon_{ijk} A_j B_k && (4.26)\\
(\overline{A}X\overline{B})_2 &= \epsilon_{1jk} A_j B_k && (4.27)\\
(\overline{A}X\overline{B})_2 &= \epsilon_{1jk} A_j B_k && (4.28)\\
&= \epsilon_{21k} A_1 B_k + \epsilon_{22k} A_2 B_k + \epsilon_{23k} A_3 B_k && (4.29)\\
&= \epsilon_{211} A_1 B_1 + \epsilon_{212k} A_1 B_2 + \epsilon_{213} A_1 B_3 + && (4.30)\\
& \epsilon_{221} A_2 B_1 + \epsilon_{222k} A_2 B_2 + \epsilon_{223} A_2 B_3 + \epsilon_{231} A_3 B_1 + \epsilon_{232k} A_3 B_2 + \epsilon_{233_{A_3 B_3}} && (4.31)\\
&= = \epsilon_{213} A_1 B_3 + \epsilon_{231k} A_3 B_1 && (4.32)
\end{aligned}
$$

Para cuando i = 3:

$$(\overline{A}X\overline{B})_i = \epsilon_{ijk} A_j B_k \tag{4.33}$$

$$(\overline{A}X\overline{B})_3 = \epsilon_{1jk} A_j B_k \tag{4.34}$$

$$(\overline{A}X\overline{B})_3 = \epsilon_{1jk} A_j B_k \tag{4.35}$$

$$= \epsilon_{31k} A_1 B_k + \epsilon_{32k} A_2 B_k + \epsilon_{33k} A_3 B_k \tag{4.36}$$

$$= \epsilon_{311} A_1 B_1 + \epsilon_{312k} A_1 B_2 + \epsilon_{313} A_1 B_3 + \tag{4.37}$$

$$\epsilon_{321} A_2 B_1 + \epsilon_{322k} A_2 B_2 + \epsilon_{323} A_2 B_3 + \epsilon_{331} A_3 B_1 + \epsilon_{332k} A_3 B_2 + \epsilon_{333_{A_3 B_3}} \tag{4.38}$$

$$= \epsilon_{312} A_1 B_2 + \epsilon_{321k} A_2 B_1 \tag{4.39}$$

4.6 Permutaciones Circulares

Se utilizan cuando los elementos se han de ordenar ”en círculo”, (por ejemplo, los comensales en una mesa), de modo que el primer elemento que “se sitúe” en la muestra determina el principio y el final de muestra.

$$PC_n = P_{n-1} = (n-1)! \tag{4.40}$$

Ejemplo 4.1

Calcular las permutaciones cirulares de 7 elementos.

Solución:

$$PC_7 = (7-1)! = 720 \tag{4.41}$$

Ejemplo 4.2

¿De cuántas formas distintas pueden sentarse ocho personas alrededor de una mesa redonda?

Solución:

$$PC_8 = P_{8-1} = 5040 \tag{4.42}$$

5 Medidas de dispersión

Estudia la distribución de los valores de la serie, analizando si estos se encuentran más o menos concentrados, o más o menos dispersos. Existen diversas medidas de dispersión, entre las más utilizadas podemos destacar las siguientes:

5.1 Rango

Mide la amplitud de los valores de la muestra y se calcula por diferencia entre el valor más elevado y el valor más bajo.

5.2 Varianza

Mide la distancia existente entre los valores de la serie y la media. Se calcula como sumatorio de las diferencias al cuadrado entre cada valor y la media, multiplicadas por el número de veces que se ha repetido cada valor. El sumatorio obtenido se divide por el tama'no de la muestra.

$$S_x^2 = \frac{\sum (x_i - x_m)^2 (n_i)}{n} \tag{5.1}$$

La varianza siempre será mayor que cero. Mientras más se aproxima a cero, más concentrados están los valores de la serie alrededor de la media. Por el contrario, mientras mayor sea la varianza, más dispersos están.

5.3 Desviación típica

Se calcula como raíz cuadrada de la varianza.

5.4 Coeficiente de varización de Pearson

Se calcula como cociente entre la desviación típica y la media.

5.5 Wikipedia - Imagen

1. Imagen A, Dónde representa la forma que se tiene cuando de cuatro elementos se selecciona solo uno. Como se puede observar la imagen para realizar conteo, considerando que no se repiten y que el orden importa, entonces tenemos:

$$P\binom{n}{r} = \frac{n!}{(n-r)!} = 1 \tag{5.2}$$

2. Imagen B, Es este caso en partinular se tienen 4 elementos en 'n' de los cuales se toman los mismos 'n', por tal motivo significa que se van a considerar todas las permutaciones.

$$P\binom{n}{r} = \frac{n!}{(n-r)!} = 24 \tag{5.3}$$

3. Imagen C, Dónde representa la forma que se tiene cuando de cuatro elementos se selecciona solo dos. Como se puede observar la imagen para realizar conteo, considerando que no se repiten y que el orden importa, entonces tenemos:

$$P\binom{n}{r} = \frac{n!}{(n-r)!} = 12 \tag{5.4}$$

4. Imagen D, Dónde representa la forma que se tiene cuando de cuatro elementos se selecciona solo dos. Como se puede observar la imagen para realizar conteo, considerando que no se repiten y que el orden no importa, entonces tenemos:

$$P\binom{n}{r} = \frac{n!}{(n-r)!(r!)} = 6 \tag{5.5}$$

Ejemplo 5.1

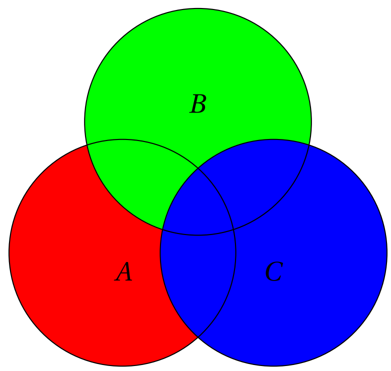
Cuantas permutaciones se pueden formar de un conjunto que tiene 5 elementos de los cuales se toman los mismo 5, considere que el orden si importa y que no se pueden repetir:
Aplicando la formula se tiene lo siquiente:

$$P\binom{n}{r} = \frac{n!}{(n-r)!} = 120 \tag{5.6}$$

Para tenerlo más claro vea la siguiente figura:

a,b,c,d,e	a,b,c,e,d	a,b,d,c,e	a,b,d,e,c	a,b,e,c,d
a,b,e,d,c	a,c,b,d,e	a,c,b,e,d	a,c,d,b,e	a,c,d,e,b
a,c,e,b,d	a,c,e,d,b	a,d,b,c,e	a,d,b,e,c	a,d,c,b,e
a,d,c,e,b	a,d,e,b,c	a,d,e,c,b	a,e,b,c,d	a,e,b,d,c
a,e,c,b,d	a,e,c,d,b	a,e,d,b,c	a,e,d,c,b	b,a,c,d,e
b,a,c,e,d	b,a,d,c,e	b,a,d,e,c	b,a,e,c,d	b,a,e,d,c
b,c,a,d,e	b,c,a,e,d	b,c,d,a,e	b,c,d,e,a	b,c,e,a,d
b,c,e,d,a	b,d,a,c,e	b,d,a,e,c	b,d,c,a,e	b,d,c,e,a
b,d,e,a,c	b,d,e,c,a	b,e,a,c,d	b,e,a,d,c	b,e,c,a,d
b,e,c,d,a	b,e,d,a,c	b,e,d,c,a	c,a,b,d,e	c,a,b,e,d
c,a,d,b,e	c,a,d,e,b	c,a,e,b,d	c,a,e,d,b	c,b,a,d,e
c,b,a,e,d	c,b,d,a,e	c,b,d,e,a	c,b,e,a,d	c,b,e,d,a
c,d,a,b,e	c,d,a,e,b	c,d,b,a,e	c,d,b,e,a	c,d,e,a,b
c,d,e,b,a	c,e,a,b,d	c,e,a,d,b	c,e,b,a,d	c,e,b,d,a
c,e,d,a,b	c,e,d,b,a	d,a,b,c,e	d,a,b,e,c	d,a,c,b,e
d,a,c,e,b	d,a,e,b,c	d,a,e,c,b	d,b,a,c,e	d,b,a,e,c
d,b,c,a,e	d,b,c,e,a	d,b,e,a,c	d,b,e,c,a	d,c,a,b,e
d,c,a,e,b	d,c,b,a,e	d,c,b,e,a	d,c,e,a,b	d,c,e,b,a
d,e,a,b,c	d,e,a,c,b	d,e,b,a,c	d,e,b,c,a	d,e,c,a,b
d,e,c,b,a	e,a,b,c,d	e,a,b,d,c	e,a,c,b,d	e,a,c,d,b
e,a,d,b,c	e,a,d,c,b	e,b,a,c,d	e,b,a,d,c	e,b,c,a,d
e,b,c,d,a	e,b,d,a,c	e,b,d,c,a	e,c,a,b,d	e,c,a,d,b
e,c,b,a,d	e,c,b,d,a	e,c,d,a,b	e,c,d,b,a	e,d,a,b,c
e,d,a,c,b	e,d,b,a,c	e,d,b,c,a	e,d,c,a,b	e,d,c,b,a

6 Teoría de Conjuntos



6.1 Primer Ley Distributiva

La ley de distributividad establece que la suma y el producto evaluado permanece el mismo valor incluso cuando el valor de los elementos este alterado

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \tag{6.1}$$

Demostración:
Proponemos $x \in A \cup (B \cap C)$. si $x \in A \cup (B \cap C)$ entonces x es incluso en A ó en (B y C).

$$\text{Proponemos } x \in A \cup (B \cap C)\text{.si } x \in A \cup (B \cap C) \tag{6.2}$$
$$\text{entonces } x \text{es incluso en } A \text{ ó en } (B y C). \tag{6.3}$$
$$x \in A \text{ ó } x \in (B y C) \tag{6.4}$$
$$x \in A \text{ ó } \{x \in B \text{ y } x \in C\} \tag{6.5}$$
$$\{x \in \text{ ó } x \in B\} \text{ y } \{x \in A \text{ y } x \in C\} \tag{6.6}$$
$$x \in (A \text{ ó } B) \text{ y } x \in (A \text{ ó } C) \tag{6.7}$$
$$x \in (A \cup B) \cap x \in (A \cap C) \tag{6.8}$$
$$x \in (A \cup B) \cap (A \cup C) \tag{6.9}$$
$$x \in A \cup (B \cap C) \Rightarrow x \in (A \cup B) \cap (A \cup C) \tag{6.10}$$
$$\tag{6.11}$$

Por lo tanto

$$A \cup (B \cap C) \subset (A \cup B) \cap (A \cup C) \tag{6.12}$$

$$\text{Proponemos } x \in (A \cup B) \cap (A \cup C)\text{.si } x \in (A \cup B) \cap (A \cup C) \tag{6.13}$$
$$\text{entonces } x \text{está en } (A \text{ ó } B) \text{ y } x \text{está en } (A ó C). \tag{6.14}$$
$$x \in (A ó B) \text{ y } x \in (A \text{ ó } C) \tag{6.15}$$
$$\{x \in A \text{ ó } x \in B\} \text{ y } \{x \in A \text{ ó } x \in C\} \tag{6.16}$$
$$x \in A \text{ ó } \{x \in B \text{ y } x \in C\} \tag{6.17}$$
$$x \in A \text{ ó } \{x \in (B \text{ y } C)\} \tag{6.18}$$
$$x \in A \cup \{x \in (B \cap C)\} \tag{6.19}$$
$$x \in A \cup (B \cap C) \tag{6.20}$$
$$x \in (A \cup B) \cap (A \cup C) \Rightarrow x \in A \cup (B \cap C) \tag{6.21}$$
$$\tag{6.22}$$

Por lo tanto

$$(A \cup B) \cap (A \cup C) \subset A \cup (B \cap C) \tag{6.23}$$

De la ecuación 3 y 4:

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \tag{6.24}$$

6.2 Segunda Ley Distributiva

La ley de distributividad establece que la suma y el producto evaluado permanece el mismo valor incluso cuando el valor de los elementos este alterado.

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \tag{6.25}$$

Demostración:

$$\text{Proponemos} x \in A \cap (B \cup C).\text{si } x \in A \cap (B \cup C). \tag{6.26}$$
$$\text{entonces} x \in A \text{ y } x \in (B \text{ ó } C). \tag{6.27}$$
$$x \in A \text{ y } x \in (B \text{ ó } C) \tag{6.28}$$
$$x \in A \text{ y } \{x \in B \text{ y } x \in C\} \tag{6.29}$$
$$\{x \in A \text{ y } x \in B\} \text{ ó } \{x \in A \text{ y } x \in C\} \tag{6.30}$$
$$x \in (A \text{ ó } B) yx \in (A \text{ ó } C) \tag{6.31}$$
$$x \in (A \cap B) \cup x \in (A \cap C) \tag{6.32}$$
$$x \in (A \cap B) \cup (A \cap C) \tag{6.33}$$
$$x \in A \cap (B \cup C) \Rightarrow x \in (A \cap B) \cup (A \cap C) \tag{6.34}$$
$$\tag{6.35}$$

Por lo tanto:

$$(A \cap B) \cup (A \cap C) \subset A \cap (B \cup C) \tag{6.36}$$
$$\text{Proponemos } x \in (A \cap B) \cup (A \cap C).\text{si } x \in (A \cap B) \cup (A \cap C). \tag{6.37}$$
$$entonces x \in (A \cap B) \text{ ó } x \in (A \cap C). \tag{6.38}$$
$$x \in (A \text{ y } B) \text{ ó } x \in (A \text{ y } C) \tag{6.39}$$
$$\{x \in A y x \in B\} \text{ ó } \{x \in A \text{ y } x \in C\} \tag{6.40}$$
$$x \in A \text{ y } \{x \in B \text{ ó } x \in C\} \tag{6.41}$$
$$x \in A \text{ y } \{x \in (B \text{ ó } C)\} \tag{6.42}$$
$$x \in A \cap \{x \in (B \cup C)\} \tag{6.43}$$
$$x \in A \cup (B \cap C) \tag{6.44}$$
$$x \in (A \cap B) \cup (A \cap C) \Rightarrow x \in A \cap (B \cup C) \tag{6.45}$$
$$\tag{6.46}$$

Por lo tanto:

$$(A \cap B) \cup (A \cap C) \subset A \cap (B \cup C) \tag{6.47}$$

De la ecuación 7 y 8:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \tag{6.48}$$

6.3 Leyes de DeMorgan

Dados dos conjuntos A y B en un universal ϑ , se verifica:

$$(A \cup B)' = A' \cap B' \tag{6.49}$$

$$(A \cap B)' = A' \cup B' \tag{6.50}$$

Demostración:

$$(A \cup B)' = A' \cap B' \tag{6.51}$$

En efecto, sea x un elemento arbitrario del conjunto universal ϑ . Entonces,

$$\begin{aligned}
x \in (A \cup B)' &\leftrightarrow x \notin (A \cup B) && \text{Definición de complementario} && (6.52)\\
&\leftrightarrow \neg[x \in (A \cup B)] && \text{Negación} && (6.53)\\
&\leftrightarrow \neg[(x \in A) \vee (x \in B)] && \text{Definición de unión} && (6.54)\\
&\leftrightarrow \neg(x \in A) \wedge \neg(x \in B) && \text{De Morgan} && (6.55)\\
&\leftrightarrow (x \notin A) \wedge (x \notin B) && \text{Negación} && (6.56)\\
&\leftrightarrow (x \in A') \wedge (x \in B') && \text{Definición de complementario} && (6.57)\\
&\leftrightarrow x \in (A' \cap B') && \text{Definición de intersección} && (6.58)\\
& && && (6.59)
\end{aligned}$$

y al ser x un elemento arbitrario de ϑ, se sigue que:

$$\forall x[x \in (A \cup B)' \leftrightarrow x \in (A' \cap B')] \tag{6.60}$$

Por lo tanto tenemos que:

$$(A \cup B)' = A' \cap B' \tag{6.61}$$

Con esto se prueba análogamente la ley de DeMorgan.

6.4 Principio De La Dualidad

Traduce conceptos, teoremas o estructuras matemáticas en otros conceptos, teoremas o estructuras, a menudo por medio de una operación de involución, esto es: Si la dualidad de A es B, entonces de forma análoga, dualidad de B es A, cumpliéndose una relacion uno a uno. Como a veces la involución tiene puntos fijos, la dualidad de A es a veces A (ella misma). Por ejemplo, el Teorema de Desargues en la geometría proyectiva es Dual a ella misma en este sentido.

El concepto de dualidad es amplia mente usado en diversos representaciones matemáticas, tales como lógica, lógica boleana, teoría de conjuntos.

La dualidad en el caso exclusivo de teoría de conjuntos, se entiende de la siguiente manera: grupos bajo la operación de de unión, intersección y complemento, satisfacen varias leyes (identidades) y cada una de estas leyes se pueden representar de forma dual, es decir, al construir la dualidad de una expresión, es necesario intercambiar operaciones:

- Unión por Intersección $\cup \to \cap$
- Intersección por Unión $\cap \to \cup$
- Conjunto universo por vacío $U \to \oslash$
- Conjunto vacío por universo $\oslash \to U$

Ahora si aplicamos este principio a una identidad, se tiene.

$$(U \cap A) \cup (B \cap A) = A \tag{6.62}$$

Aplicando la dualidad:

$$(\oslash \cup A) \cap (B \cup A) = A \tag{6.63}$$

Ahora bien, como se enuncia anteriormente, la dualidad en un principio que se puede aplicar en una gran variedad de áreas del conocimiento, a continuación se muestra la dualidad en lógica boleana. Se realiza lo homólogo a la teoría de conjuntos, se reemplazan las operaciones de suma por multiplicación, los 1 por 0 y viceversa.

Sea:

$$x(y+0) \text{ y } \bar{x}\cdot 1 + (\bar{y}+z) \tag{6.64}$$

Aplicando el principio de la dualidad de las ecuaciones correspondientes.

$$x + (y \cdot 1) \tag{6.65}$$

$$(\bar{x}+1)\cdot(\bar{y}\cdot z) \tag{6.66}$$

7 Histogramas

En estadística, un histograma es una representación grf́ica de una variable en forma de barras, donde la superficie de cada barra es proporcional a la frecuencia de los valores representados, ya sea en forma diferencial o acumulada. Sirven para obtener una "primera vista" general, o panorama, de la distribución de la población, o la muestra, respecto a una característica, cuantitativa y continua, de la misma y que es de interés para el observador (como la longitud o la masa). De esta manera ofrece una visión en grupo permitiendo observar una preferencia, o tendencia, por parte de la muestra o población por ubicarse hacia una determinada región de valores dentro del espectro de valores posibles (sean infinitos o no) que pueda adquirir la característica. Así pues, podemos evidenciar comportamientos, observar el grado de homogeneidad, acuerdo o concisión entre los valores de todas las partes que componen la población o la muestra, o, en contraposición, poder observar el grado de variabilidad, y por ende, la dispersión de todos los valores que toman las partes, también es posible no evidenciar ninguna tendencia y obtener que cada miembro de la población toma por su lado y adquiere un valor de la característica aleatoriamente sin mostrar ninguna preferencia o tendencia, entre otras cosas.

En el eje vertical se representan las frecuencias, es decir, la cantidad de población o la muestra, según sea el caso, que se ubica en un determinado valor o sub-rango de valores de la característica que toma la característica de interés, evidentemente, cuando este espectro de valores es infinito o muy grande el mismo es reducido a sólo una parte que muestre la tendencia o comportamiento de la población, en otras ocasiones este espectro es extendido para mostrar el alejamiento o ubicación de la población o la muestra analizada respecto de un valor de interés.

En general se utilizan para relacionar variables cuantitativas continuas, pero también se lo suele usar para variables cuantitativas discretas, en cuyo caso es común llamarlo diagrama de frecuencias y sus barras estń separadas, esto es porque en el "x" ya no se representa un espectro continuo de valores, sino valores cuantitativos específicos como ocurre en un diagrama de barras cuando la característica que se representa es cualitativa o categórica. Su utilidad se hace mś evidente cuando se cuenta con un gran número de datos cuantitativos y que se han agrupado en intervalos de clase.

Ejemplos de su uso es cuando se representan franjas de edades o altura de la muestra, y, por comodidad, sus valores se agrupan en clases, es decir, valores continuos. En los casos en los que los datos son cualitativos (no numéricos), como sexto grado de acuerdo o nivel de estudios, es preferible un diagrama de sectores.

Los histogramas son mś frecuentes en ciencias sociales, humanas y económicas que en ciencias naturales y exactas. Y permite la comparación de los resultados de un proceso.

7.1 Tipos de Histogramas

- **Diagramas de barras simples**.Representa la frecuencia simple (absoluta o relativa) mediante la altura de la barra la cual es proporcional a la frecuencia simple de la categoría que representa.

- **Diagramas de barras compuesta**.Se usa para representar la información de una tabla de doble entrada o sea a partir de dos variables, las cuales se representan así; la altura de la barra representa la frecuencia simple de las modalidades o categorías de la variable y esta altura es proporcional a la frecuencia simple de cada modalidad.

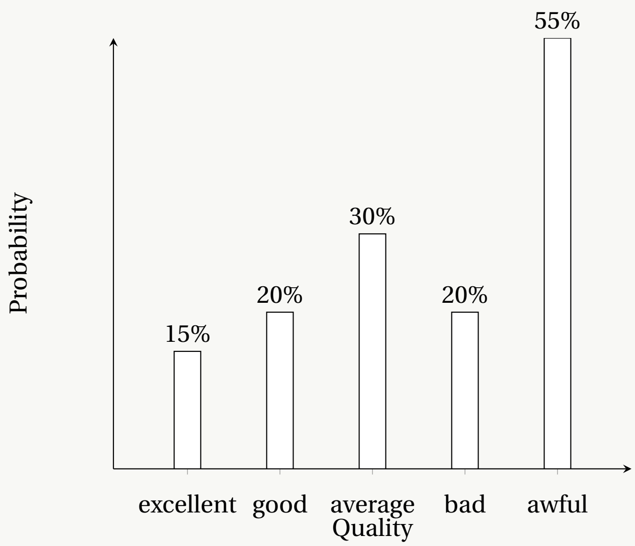
- **Diagramas de barras agrupadas**. Se usa para representar la información de una tabla de doble entrada o sea a partir de dos variables, el cual es representado mediante un conjunto de barras como se clasifican respecto a las diferentes modalidades.

- **Polígono de frecuencias**. Es un gráfico de líneas que de las frecuencias absolutas de los valores de una distribución en el cual la altura del punto asociado a un valor de las variables es proporcional a la frecuencia de dicho valor.

- **Ojiva porcentual**. Es un gráfico acumulativo, el cual es muy útil cuando se quiere representar el rango porcentual de cada valor en una distribución de frecuencias.

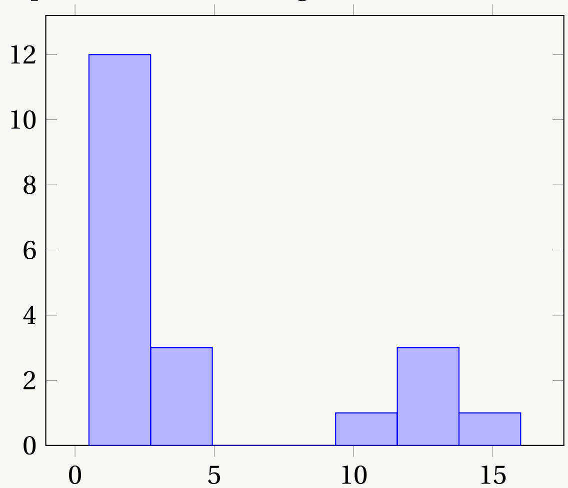
Ejemplo 7.1

Histograma con representación en barras.



Ejemplo 7.2

Representación de Histograma



7.2 Aproximación de Stirling

$$n! \approx \sqrt{2\pi n}(\frac{n}{e}) \tag{7.1}$$

De acuerdo a la función Gamma de Euler podemos describir el factorial de cualquier número como:

$$\Gamma(n+1) = \int_{x=0}^{\infty} x^n \cdot e^{-x} dx \tag{7.2}$$

Donde $n \geq 0$, cuando n=10 e integrar obtenemos:

$$f(x) = x^n \cdot e^{-x} \tag{7.3}$$

Esta ecuación puede ser representada por la integral Gaussiana teniendo el máximo punto cuando n=0:

$$f(x) = \frac{10}{e}^{10*} \cdot e^{\frac{-x^2}{20}} \tag{7.4}$$

Para poder transformar la función Gamma a la integral de Euler tenemo que definimos las variables de la integral como:
$u = x + n$ y $du = dx$ y sustituyendo:
$I = \int_{u=-n}^{\infty} (u+n)^n \cdot e^{u+n} du$ factorizando $(u+n)^n$ la integral es:

$I = \int_{u=-n}^{\infty} (n)^n \cdot (1+\frac{u}{n})^n \cdot e^{-u \cdot n} du$, tomando las constantes fuera de nuestra integral tenemos:

$I = (n)^n \cdot \int_{u=-n}^{\infty} (1+\frac{u}{n})^n \cdot e^{-u \cdot n} du$ (1)
Resolviendo la integral y usando Series de Maclaurin para expander el logaritmo:

$$ln(1+\frac{u}{n})^n = n \cdot ln(1+\frac{u}{n}) \quad (7.5)$$
$$\Longrightarrow \quad ln(1+\frac{u}{n} = \frac{u}{n} - \frac{u^2}{2n^2}) \quad (7.6)$$

Finalmente sustituyendo 2 tenemos:

$$n \cdot ln(1+\frac{u}{n}) \approx u - \frac{u^2}{2n} \quad (7.7)$$

Sustituyendo en 1 tenemos: $I_2 = (\frac{n}{e})^n \cdot \int_{u=-n}^{n} e^{\frac{u^2}{2n}} du$, notemos que regresamos a la integral Gaussiana:

$$\int_{\infty}^{\infty} e^{\frac{u^2}{2n}} du = \sqrt{2\pi n} \quad (7.8)$$

Por lo tanto decimos que n tiende al infinito y el factorial de n tiende al valor de abajo que es la aproximación de Stirling:

$$n! \sim (\frac{n}{2})^n \cdot \sqrt{2\pi n} \quad (7.9)$$

7.3 Número de Erdös

Este es un número que pretende cuantificar el grado de separacion entre las personas que han publicado un artículo de investigación, en este caso particular el mateático Paul Erdös.

¿Cómo se calcula?

Esto es muy sencillo, se considera que Erdös tiene un número 0.

Si se ha colaborado con el, directamente en la publicacion de un artículo se tiene número de Erdös 1.

Si se ha colaborado con alguien que colabóro directamente con Erdös se tiene número 2 y así sucesivamente.

Esta métrica pretende darnos una medida de cuan separados estamos de una persona famosa, esto es en un máximo de seis personas puede uno contactar por ejemplo a Barack Obama.

Se asegura que Paul Erdös es el segundo más prolífico matemático de todos los tiempos, siendo superado solamente por Leonhard Euler, el gran matemático del siglo XVIII. La producción de Erdös es más o menos de 1,500 artículos publicados, y muchos están aún por publicarse después su muerte.

Erdŏ utilizaba café, pastillas de cafeína y Benzedrina para trabajar 19 horas al día, 7 días a la semana.

Para él, ”el café era una sustancia que los matemáticos convertían en teoremas”. Cuando sus amigos le aconsejaban bajar el ritmo y descansar, siempre respondía lo mismo: ”Habrá mucho tiempo para descansar en la tumba”. Erdös viajaba constantemente y vivía enfocado totalmente hacia la matemática evitando la compa’nía social, el sexo y las grandes comidas.

Erdös seguía publicando un artículo a la semana incluso a los setenta a’nos. Erdös, sin duda, tenía el mayor número de coautores (alrededor de 500) entre los matemáticos de todos las especialidades.

7.4 Número divisible por 11

Se dice que un número es divisible entre 11, cuando la diferencia de la suma de las cifras de posición par y la suma de las cifras de posición impar es igual a cero u once:

Ejemplo 7.3

4,158 = (4+5) - (1+8) = 0 ,Entonces 4,158 **es** divisible entre 11 Por tanto $\frac{4{,}158}{11} = 378$

Ejemplo 7.4

28,765 = (2+7+5) - (8-6) = 0 ,Entonces 28,765 **es** divisible entre 11

Por tanto $\frac{28{,}765}{11} = 2,615$

Ejemplo 7.5

979 = (9+9) - (7) = 11 ,Entonces 979 **es** divisible entre 11

Por tanto $\frac{979}{11} = 89$

Ejemplo 7.6

5,147 = (5+4) - (1+7) = 1 ,Entonces 5147 **no** es divisible entre 11

Por tanto $\frac{5{,}147}{11} = 467.909091$

7.5 Cuantificadores

En lógica formal, los cuantificadores son expresiones que indican la cantidad de veces que un predicado o propiedad P se satisface dentro de una determinada clase (por ejemplo, pertenencia, equivalencia u orden). Existen muchos tipos de cuantificadores, entre los más utilizados están:

- Cuantificador universal: $\forall x, y, \ldots$

- Cuantificador existencial: $\exists x, y, \ldots$
- Cuantificador existencial único: $\exists! x, y, \ldots$
- Negación del cuantificador existencial: $\nexists x, y, \ldots$

7.6 Fullereno

Un fullereno (también, fulereno) es una molécula compuesta por carbono que puede adoptar una forma geométrica que recuerda a una esfera, un elipsoide, un tubo (llamado nanotubo) o un anillo. Los fullerenos son similares al grafito, compuesto de hojas de anillos hexagonales enlazadas, pero conteniendo anillos pentagonales y a veces heptagonales, lo que impide que la hoja sea plana. Los fullerenos son la tercera forma molecular estable conocida de carbono, tras el grafito y el diamante.

Los fullerenos fueron descubiertos en 1985 por Harold Kroto, Robert Curl y Richard Smalley, lo que les valió la concesión del Premio Nobel de Química en 1996.

El primer fullereno descubierto fue el C_{60}, que consta de 12 pentágonos y 20 hexágonos. Cada pico corresponde a un átomo de carbono y cada lado a un enlace covalente. Tiene una estructura idéntica a la cúpula geodésica o un balón de fútbol. Por esta razón, se le llama ”buckminsterfullereno” (en homenaje al arquitecto Buckminster Fuller quien dise’nó la cúpula geodésica) o ”futboleno”. Los fullerenos esféricos reciben a menudo el nombre de buckyesferas y los cilíndricos el de buckytubos o nanotubos.

Destacan tanto por su versatilidad para la síntesis de nuevos compuestos. Su naturaleza y forma se han hecho ampliamente conocidas en la ciencia y en la cultura en general, por sus características físicas, químicas, matemáticas y estéticas.

7.7 Los Simpson y las matemáticas

7.7.1 Capítulo I

Reiss empezó con El ahorcamiento inesperado y otros entretenimientos matemáticos, y luego se gastó todo el dinero de su paga en otros libros de pasatiempos de Gardner. A la edad de ocho a’nos escribió a Gardner explicándole que era seguidor suyo, y haciendo una observación muy aguda sobre los cuadrados palindrómicos: que tendían a tener un número de dígitos impar. Los cuadrados palindrómicos son sencillamente cuadrados que son iguales ya se lean desde un lado o desde el otro, como el 121 (11^2), o 5,221,225 (2285^2).

Aquel ni’no de ocho a’nos tenía toda la razón, porque hay treinta y cinco números de este tipo en menos de cien mil millones, y solo uno de ellos, el 698,896 (836^2) tiene un número de dígitos par. Reiss admitió ante mí de mala gana que aquella carta a Gardner contenía también una pregunta. Le preguntaba si había una cantidad finita o infinita de números primos. Ahora recuerda aquella pregunta con cierta vergüenza: ”todavía veo aquella carta perfectamente, y era una pregunta estúpida e ingenua”.

7.8 Teorema del Límite Central

La distribución de la media muestral de una población normal es una distribución normal con la misma media poblacional y con desviación típica el error estándar. Este hecho nos permite calcular probabilidades cuando tenemos una muestra de una variable con distribución normal y desviación típica conocida. Cuando no conocemos la desviación típica de la variable, también podemos hacer cálculos con la distribución t de Student.

En esta sesión veremos cómo debemos proceder cuando no sabemos si la variable de interés sigue una distribución normal o no, o cuando sabemos seguroque su distribución no es normal.

Cuando la muestra es lo bastante grande, la solución nos viene dada por uno de los resultados fundamentales de la estadística: el teorema del límite central. Lo introduciremos con un caso particular: el estudio de la binomial.

Teorema 7.1 (Límite central)

Sea X_1 X_1, X_2 X_2, ..., X_n X_n un conjunto de variables aleatorias, independientes e idénticamente distribuidas de una distribución con media μ y varianza $\sigma^2 \neq 0$. Entonces, si n es suficientemente grande, la variable aleatoria

8 Exámen I - Técnicas de conteo

8.1 Exámen a propuesto

Probabilidad, procesos aleatorios e interferencia

Autor: Horacio Rodríguez Bazán.
Estatus: Creación
Nombre: ________________________________

Resuelva los siguientes problemas, relacionados a técnicas de conteo.

1. Dadas ls siguientes ecuaciones, explique la relación que tiene asi como el significado de sus terminos.

$$\frac{n!}{r!(n-r)!} \quad (1) \tag{8.1}$$

$$\frac{(r+n-1)!}{r!(n-1)!} \quad (2) \tag{8.2}$$

$$\frac{n!}{(n-r)!} \quad (3) \tag{8.3}$$

Solución: Las siguientes formulas son básicas en técnicas de conteo, las cuales nos permiten contar elementos de una colección de objetos:

(a) La primer ecuación esta relacionada a combinaciones con las siguientes restricciones:
   - No importa el orden
   - No permite repeticiones

(b) La segunda ecuación esta relaciona de la misma manera que la anterior a combinaciones, con una diferencia:
   - No importa el orden
   - Sí permite repeticiones

(c) La tercer ecuación es diferente a las anterior, ya que esta ecuación esta relacionada con permutaciones en dónde los elementos a permutar tienen las siguientes restucciones:

- Sí importa el orden
- No permite repeticiones

2. En un restaurante nos dan tres bolas de helados a escoger entre: vainilla, plátano, fresa, chocolate y melocotón. ¿Cuántas combinaciones podremos hacer?

Solución: Sea k = 3 bolas de helado y n = 5 sabores de helado, entonces tenemos la siguiente formula sobre combinaciones:

$$C\binom{n}{k} = \frac{(k+n-1)!}{k!(n-1)!} \tag{8.4}$$

$$C\binom{5}{3} = \frac{(3+5-1)!}{3!(5-1)!} \tag{8.5}$$

$$= \frac{7!}{3!*4!} \tag{8.6}$$

$$= \frac{7*6*5*\cancel{4!}}{3*2*1*\cancel{4!}} \tag{8.7}$$

$$= \frac{210}{6} = \mathbf{35} \tag{8.8}$$

3. En un catering se quieren repartir los 20 canapés de 4 en 4. ¿Cuántas combinaciones de canapés podremos formar?

Solución: Sea k = 3 bolas de helado y n = 5 sabores de helado, entonces tenemos la siguiente formula sobre combinaciones:

$$C\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{8.9}$$

$$C\binom{20}{4} = \frac{20!}{4!(20-4)!} \tag{8.10}$$

$$= \frac{20!}{4!*16!} \tag{8.11}$$

$$= \frac{20*19*18*17*\cancel{16!}}{4*3*2*1*\cancel{16!}} \tag{8.12}$$

$$= \frac{116280}{24} = \mathbf{4845} \tag{8.13}$$

4. ¿De cuántas maneras se pueden dar los primeros tres lugares entre 10 personas?

   Solución: Sea n el conjunto de personas que podemos que pueden obtener el primero,segundo y tercer lugar (1,2,3) y r el número de datos que se van a tomar, en este caso 3:

$$P\binom{n}{k} = \frac{n!}{(n-k)!} \tag{8.14}$$

$$P\binom{10}{3} = \frac{10!}{(10-3)!} \tag{8.15}$$

$$= \frac{10!}{7!} \tag{8.16}$$

$$= \frac{10*9*8*\cancel{7!}}{\cancel{7!}} \tag{8.17}$$

$$= \frac{720}{1} = \mathbf{720} \tag{8.18}$$

5. Una mesa presidencial está formada por ocho personas, ¿De cuántas formas distintas se pueden sentar, si el presidente y el secretario siempre van juntos?

   Solución: Se forman dos grupos el primero de 2 personas y el segundo de 7 personas, en los dos se cumple que:

$$Pn = n! \tag{8.19}$$

$$P2 = 2! = 2 \tag{8.20}$$

$$Pk = n! \tag{8.21}$$

$$P7 = 7! = 5040 \tag{8.22}$$

$$PnPk = P2*P7 \tag{8.23}$$

$$= 2!*7! \tag{8.24}$$

$$= 2*5040 = \mathbf{1080} \tag{8.25}$$

6. Cuatro libros distintos de matemáticas, seis diferentes de física y dos diferentes de química se colocan en un estante. ¿De cuántas formas distintas es posible ordenarlos si: Los libros de cada asignatura deben estar todos juntos?

   Solución: Se tienen que calcular 3 permutaciones direntes de la siguiente manera:

   - Una para los 4 libros de matemáticas

- Una para los 6 libros de física
- Una para los 2 libros de química

$$
\begin{aligned}
Pm &= n! && (8.26)\\
P4 &= 4! = 24 && (8.27)\\
Pf &= n! && (8.28)\\
P6 &= 6! = 720 && (8.29)\\
Pq &= n! && (8.30)\\
P2 &= 2! = 720 && (8.31)\\
& && (8.32)\\
PmPfPq &= 4! * 6! * 2! * 3! && (8.33)\\
&= 24 * 720 * 2 * 6 && (8.34)\\
&= \mathbf{207360} && (8.35)
\end{aligned}
$$

7. Cuatro libros distintos de matemáticas, seis diferentes de física y dos diferentes de química se colocan en un estante. ¿De cuántas formas distintas es posible ordenarlos si: Solamente los libros de matemáticas deben estar juntos?

   Solución: Se tienen que calcular 3 permutaciones direntes de la siguiente manera:

   - Una para los 4 libros de matemáticas, son los unicos que nos interesan que esten juntos.

$$
\begin{aligned}
Pm &= n! && (8.36)\\
P4 &= 4! = 24 && (8.37)\\
& && (8.38)\\
PmPfPq &= 4! * 9! && (8.39)\\
&= 24 * 362880 && (8.40)\\
&= \mathbf{8709120} && (8.41)
\end{aligned}
$$

8. ¿De cuántas formas distintas pueden sentarse nueve personas alrededor de una mesa redonda?

   Solución: Se tiene un problema de permutaciones circulaes que se ataca de la siguiente manera:

$$PC_n = P_{n-1} = (n-1)! \quad (8.42)$$
$$PC_9 = (9-1)! \quad (8.43)$$
$$= 8! \quad (8.44)$$
$$= \mathbf{40320} \quad (8.45)$$

9. ¿De cuántos modos distintos podemos ubicar las cifras del 1 al 7 en una figura que tiene un centro y seis posiciónes alrededor?

   Solución: En este poblema es necesario considerar la forma dada 'flor' donde se tiene una posición al centro y seis alrededor, es decir, tenemos un circulo con seis posiciones, de tal manera que el número de formas en que podemos ubicar estas cifras es:

$$PC_n = P_{n-1} = (n-1)! \quad (8.46)$$
$$PC_6 = 7 * (6-1)! \quad (8.47)$$
$$= 7 * 5! \quad (8.48)$$
$$= 7 * 120! \quad (8.49)$$
$$= \mathbf{840} \quad (8.50)$$

10. En una clase de 30 alumnos, 20 juegan fútbol y el resto baloncesto.¿De cuántas maneras se pueden seleccionar 3 alumnos de entre los que juegan al fútbol y 2 de entre los que juegan al baloncesto?

    Solución: Se tienen dos grupos A de 20 alumnos que juega fútbol y un grupo B de 10 alumnos que juega baloncesto, de tal manera que la el numero de formas que se pueden eligir es el producto de las combianaciones.

    Calculanado las formas en que se eligen dos personas que juegan fútbol

$$C\binom{n}{k} = \frac{n!}{k!(n-k)!} \quad (8.51)$$
$$C\binom{20}{3} = \frac{20!}{3!(20-3)!} \quad (8.52)$$
$$= \frac{20!}{3! * 17!} \quad (8.53)$$
$$= \frac{20 * 19 * 18 * \cancel{17!}}{3 * 2 * 1 * \cancel{17!}} \quad (8.54)$$
$$= \frac{6840}{6} = \mathbf{1140} \quad (8.55)$$

Calculanado las formas en que se eligen dos personas que juegan baloncesto

$$C\binom{n}{k} = \frac{n!}{k!(n-k)!} \quad (8.56)$$

$$C\binom{10}{2} = \frac{10!}{2!(10-2)!} \quad (8.57)$$

$$= \frac{10!}{2! * 8!} \quad (8.58)$$

$$= \frac{10 * 9 * \cancel{8!}}{2 * 1 * \cancel{8!}} \quad (8.59)$$

$$= \frac{90}{2} = \mathbf{45} \quad (8.60)$$

De tal forma que el total de combianaciones es el producto de las combinaciones anteriores:

$$C_f * C_b = 1140 * 25 \quad (8.61)$$

$$= \mathbf{28500} \quad (8.62)$$

11. ¿Cuántos números de 5 cifras son divisibles por 5?

    Solución: Para que un número sea divisible por cinco debe terminar en ”0” ó ”5” de tal manera que tenemos lo siguiente.

$$C = N_1 * N_2 * N_3 * N_4 * N_5 \quad (8.63)$$

$$C = 9 * 10 * 10 * 10 * 2 \quad (8.64)$$

$$= \frac{20!}{3! * 17!} \quad (8.65)$$

$$= \mathbf{18000} \quad (8.66)$$

12. Se tiene un cierto número de objetos, y se sabe que si se toman 7 objetos distintos se pueden hacen tantas combinaciones como si se tomaran 5. ¿Cuánto son los objetos?

    Solución: Para que un número sea divisible por cinco debe terminar en ”0” ó ”5” de tal manera que tenemos lo siguiente.

$$
\begin{aligned}
C\binom{n}{k} &= C\binom{n}{k} && (8.67)\\
C\binom{n}{7} &= C\binom{n}{7} && (8.68)\\
\frac{n!}{k!(n-k)!} &= \frac{n!}{k!(n-k)!} && (8.69)\\
\frac{n!}{7!(n-7)!} &= \frac{n!}{5!(n-5)!} && (8.70)\\
\frac{n!}{7!(n-7)!}*\frac{1}{n!} &= \frac{n!}{5!(n-5)!}*\frac{1}{n!} && (8.71)\\
\frac{1}{7!(n-7)!} &= \frac{1}{5!(n-5)!} && (8.72)\\
1*7!(n-7)! &= 1*5!(n-5)! && (8.73)\\
7!(n-7)! &= 5!(n-5)! && (8.74)\\
\frac{7!}{5!}(n-7)! &= (n-5)! && (8.75)\\
\frac{7*6*\cancel{5!}}{\cancel{5!}}(n-7)! &= (n-5)! && (8.76)\\
42(n-7)! &= (n-5)! && (8.77)\\
42(n-7)(n-8)(n-9)\ldots &= (n-5)(n-6)(n-7)(n-8)(n-9)\ldots && (8.78)\\
42 &= \frac{(n-5)(n-6)\cancel{(n-7)}\cancel{(n-8)}\cancel{(n-9)}\ldots}{\cancel{(n-7)}\cancel{(n-8)}\cancel{(n-9)}\ldots} && (8.79)\\
42 &= (n-5)(n-6) && (8.80)\\
42 &= (n^2-6n-5n+36 && (8.81)\\
0 &= (n^2-11n-12 && (8.82)\\
0 &= (n-12)(n+1) && (8.83)\\
n &= \mathbf{12} && (8.84)
\end{aligned}
$$

13. Dado el siguiente conjunto de letras (A,B,C,D,E), construya el árbol que represente las combinaciones de 5 tomando 3, dónde no se permiten repeticiones y el orden no importa.

    Solución: Primero calculamos el número de combianaciones posibles, de tal manera que.

$$C\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{8.85}$$

$$C\binom{5}{3} = \frac{5!}{3!(5-3)!} \tag{8.86}$$

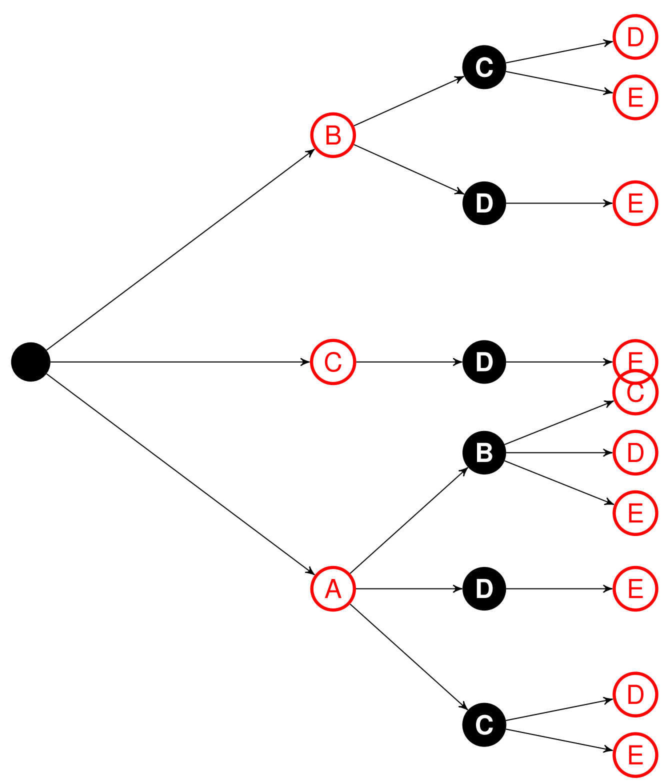
$$= \frac{5!}{3! * 2!} \tag{8.87}$$

$$= \frac{5 * 4 * \cancel{3!}}{2! * \cancel{3!}} \tag{8.88}$$

$$= \frac{20}{2} = \mathbf{10} \tag{8.89}$$

$$a,b,c \mid a,b,d \mid a,b,e \mid a,c,d \mid a,c,e \mid a,d,e \mid b,c,d \mid b,c,e \mid b,d,e \mid c,d,e \tag{8.90}$$

De tal forma que el diagrama de árbol se puede representar de la siguiente manera:



14. Obtener el número de triángulos que se pueden formar en un polígono de 20 lados, considere que un triangulo esta formado por tres vértices.

Solución: Para que un triángulo sea formado requiere de 3 puntos, de tal manera que.

$$
\begin{aligned}
C\binom{n}{k} &= \frac{n!}{k!(n-k)!} && (8.91)\\
C\binom{20}{3} &= \frac{20!}{3!(20-3)!} && (8.92)\\
&= \frac{20!}{3! * 17!} && (8.93)\\
&= \frac{20 * 19 * 18 * \cancel{17!}}{3 * 2 * 1 * \cancel{17!}} && (8.94)\\
&= \frac{6840}{6} = \mathbf{1140} && (8.95)
\end{aligned}
$$

15. ¿Cuántas iniciales de cuatro letras distintas se pueden formar usando los alfabetos del idioma Inglés (26 letras) de forma que la última de las cuatro palabras siempre es una consonante ?

    Solución: Para la solución de este problema consideree que hay cinco vocale, de tal forma que la cuarta letra solo puede tener 21 posibilidades.

$$
\begin{aligned}
C &= L_1 * L_2 * L_3 * L_4 && (8.96)\\
&= 26 * 26 * 26 * 21 && (8.97)\\
&= 26^3 * 21 && (8.98)\\
&= 17576 * 21 && (8.99)\\
&= \mathbf{369096} && (8.100)
\end{aligned}
$$

16. ¿Cuál es el número total de formas en que Juan puede distribuir 9 regalos distintas entre sus 8 novias distintas tales que cada una de ellas obtiene al menos un regalo ?

    Solución: Como cada novia debe recibir un regalo. El número de maneras distintas 8 regalos pueden ser seleccionados es : $C\binom{n}{k} = C\binom{9}{8} = 9$ formas. El número de maneras en que cada novia obtiene uno de estos 8 regalos es 8!. Número total de formas que 8 regalos pueden ser distribuidos es de $9 * 8!$. Ahora el último regalo se puede dar a cualquiera de las 8 novias , por tanto, el número total de formas de la distribución es:

$$C = C\binom{n}{k} * 8! * 8 \tag{8.101}$$
$$= C\binom{9}{8} * 8! * 8 \tag{8.102}$$
$$= 9 * 8 * 8! \tag{8.103}$$
$$= \mathbf{72*8}! \tag{8.104}$$

17. ¿Cuántas veces aparece el número 7 entre los números 1 y 1000?

    Solución: Considere que en el número 007 aparece una vez, en 017 dos veces y en 777 tres veces, de tal forma que:

    Cuando solo esta en una posición:

$$C = d_1 * d_2 * d_3 \tag{8.105}$$
$$= 1 * 9 * 9 \tag{8.106}$$
$$= 81 * 3 \tag{8.107}$$
$$= \mathbf{243}! \tag{8.108}$$

    Cuando solo esta en dos posiciones:

$$C = d_1 * d_2 * d_3 \tag{8.109}$$
$$= 1 * 1 * 9 \tag{8.110}$$
$$= 9 * 3 \tag{8.111}$$
$$= \mathbf{27}! \tag{8.112}$$

    Cuando solo esta en tres posiciones:

$$C = d_1 * d_2 * d_3 \tag{8.113}$$
$$= 1 * 1 * 1 \tag{8.114}$$
$$= 1 * 3 \tag{8.115}$$
$$= \mathbf{3} \tag{8.116}$$

    Entonces el total es $T = N_1 + N_2 + N_3 = 243 + 2 * 27 + 3 = 243 + 54 + 3 = \mathbf{300}$

18. ¿De cuántas formas se pueden sentar 15 personas en dos mesas redondas con capacidades de 8 y 7 personas?

    Solución: Calculemos las permutaciones de cada una de las mesas

$$
\begin{aligned}
P_{m7} &= P_{n-1} = (n-1)! && (8.117)\\
&= (7-1)! && (8.118)\\
&= \mathbf{6}! && (8.119)
\end{aligned}
$$

$$
\begin{aligned}
P_{m8} &= P_{n-1} = (n-1)! && (8.120)\\
&= (8-1)! && (8.121)\\
&= \mathbf{7}! && (8.122)
\end{aligned}
$$

    Entonces, el resultado es considerando que se llena la mesa de 8 personas el resultado seria $P\binom{15}{8} * 7! * 6!$

19. ¿Cuántos números enteros positivos de 5 dígitos existe la suma de cuyos dígitos son impares ?

    Solución: Considere los números de 5 cifras en le primera poisción no puede haber cero y que el últumo digito define si es par o impar por lo tanto al final solo podemos tener cinco posibles números (1,3,5,7,9).

$$
\begin{aligned}
C &= D_1 * D_2 * D_3 * D_4 * D_5 && (8.123)\\
&= 9 * 10 * 10 * 10 * 5 && (8.124)\\
&= 9 * 10^3 * 5 && (8.125)\\
&= 45 * 1000 && (8.126)\\
&= \mathbf{45000} && (8.127)
\end{aligned}
$$

20. Hay 2 hermanos entre un greupo de 20 personas, ¿De cuántas formas puede ser ordenado el grupo de tal forma que haya exactamente una persona entre los hermanos?

    Solución: Consideremos a los hermanos como un persona y que estos pueden estar en un lado o en otro, de tal forma que tenemos 19 personas que podemos organizar en un circulo.

$$
\begin{aligned}
P_C &= P_{n-1} = (n-1)! && (8.128)\\
&= 2! * (19-1)! && (8.129)\\
&= \mathbf{2 * 18!} && (8.130)
\end{aligned}
$$

8.2 Exámen a resolver

Probabilidad, procesos aleatorios e interferencia

Autor: Francisco Javier Vazquez Vazquez.
Estatus: Resolución
Nombre: Horacio Rodríguez Bazán___________

1. Desarrolle $(a+b)^6$ utilizando el teorema del binomio
   Solución: Usando la ecuación del teorema del binomio, tenemos:

$$(a+b)^n = \sum_{k=0}^{n} \tbinom{n}{k} a^{n-k} b^k \tag{8.131}$$

$$(x+y)^6 = \binom{6}{0} x^{6-0} y^0 + \binom{6}{1} x^{6-1} y^1 + \binom{6}{2} x^{6-2} y^2 + \binom{6}{3} x^{6-3} y^3 + \tag{8.132}$$

$$+ \binom{6}{4} x^{6-4} y^4 + \binom{6}{5} x^{6-5} y^5 + \binom{6}{6} x^{6-6} y^6 \tag{8.133}$$

$$= \binom{6}{0} x^6 + \binom{6}{1} x^5 y^1 + \binom{6}{2} x^4 y^2 + \binom{6}{3} x^3 y^3 + \binom{6}{4} x^2 y^4 + \binom{6}{5} x^1 y^5 + \binom{6}{6} y^6 \tag{8.134}$$

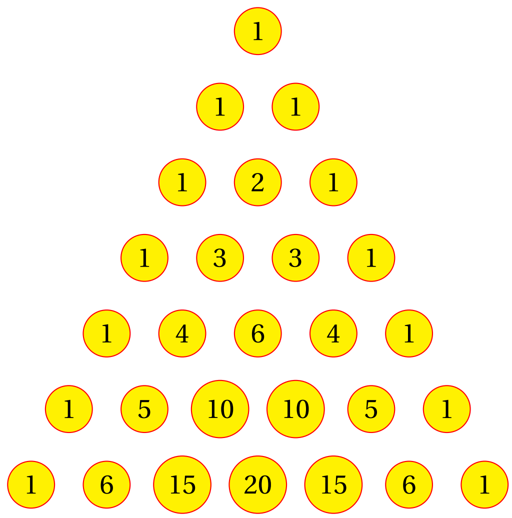
$$= 1x^6 + 6x^5 y^1 + 15x^4 y^2 + 20x^3 y^3 + 15x^2 y^4 + 6x^1 y^5 + 1y^6 \tag{8.135}$$

$$(x+y)^6 = x^6 + 6x^5 y + 15x^4 y^2 + 20x^3 y^3 + 15x^2 y^4 + 6xy^5 + y^6 \tag{8.136}$$

2. Dados los siguiente coeficientes, hacer el triángulo de Pascal:

   1 6 15 20 15 6 1

   Solución: Tomado como base el problema anterior y haciendo uso del teorema del binomio, podemos representar los siguientes coeficientes binomiales realizando las iteraciones correspondientes: De tal forma que el triángulo de Pascal se puede representar de la siguiente manera.



3. Un estudiante debe responder 8 preguntas de un exámen de 10 preguntas, determine lo siguiente:

   (a) ¿De cuantas formas puede elegir las 8 preguntas?

   (b) ¿De cuantas formas puede elegir las 8 preguntas si debe responder obligatoriamente las primeras 3 preguntas?

   Solución:

   (a) Las preguntas pueden ser seleccionadas de un conjunto de 10 preguntas en donde se deben elegir 8 y el orden es irrelevante; utilizando:

$$C\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{8.137}$$

$$C\binom{10}{8} = \frac{10!}{8!(10-8)!} \tag{8.138}$$

$$= \frac{10!}{8! * 2!} \tag{8.139}$$

$$= \frac{10 * 9 * \cancel{8!}}{2 * 1 * \cancel{8!}} \tag{8.140}$$

$$= \frac{90}{2} = \mathbf{45} \tag{8.141}$$

   (b) Si el estudiante debe responder obligatoriamente las primeras 3 preguntas, solo podra elegir las siguientes 5 preguntas de un conjunto de 7 preguntas, generando todas las posibles combinaciones:

$$C\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{8.142}$$

$$C\binom{7}{5} = \frac{7!}{5!(7-5)!} \tag{8.143}$$

$$= \frac{7!}{5!*2!} \tag{8.144}$$

$$= \frac{7*6*\cancel{5!}}{2*1*\cancel{5!}} \tag{8.145}$$

$$= \frac{42}{2} = \mathbf{21} \tag{8.146}$$

4. Redusca la siguiente expresión a su forma mas simple en terminos de n
   $\frac{(n+2)!}{n!}$
   Solución:

$$= \frac{(n+2)!}{n!} \tag{8.147}$$

$$= \frac{(n+2)*(n+1)*(n)*(n-1)\ldots*3*2*1}{(n)(n-1)\ldots*3*2*1} \tag{8.148}$$

$$= \frac{(n+2)*(n+1)*\cancel{(n)*(n-1)}\ldots*\cancel{3*2*1}}{\cancel{(n)(n-1)}\ldots*\cancel{3*2*1}} \tag{8.149}$$

$$= (n+2)(n+1) \tag{8.150}$$

$$= n^2+n+2n+2 \tag{8.151}$$

$$= n^2+3n+2 \tag{8.152}$$

5. Desarrolle y simplifique la expansion de $(x^2-2y)^6$

   Solución: Usando la ecuación del teorema del binomio, tenemos:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k \tag{8.153}$$

$$(x^2-2y)^6 = \binom{6}{0}(x^2)^{6-0}(-2y)^0 + \binom{6}{1}(x^2)^{6-1}(-2y)^1 + \binom{6}{2}(x^2)^{6-2}(-2y)^2 + \tag{8.154}$$

$$+ \binom{6}{3}(x^2)^{6-3}(-2y)^3 + \binom{6}{4}(x^2)^{6-4}(-2y)^4 + \binom{6}{5}(x^2)^{6-5}(-2y)^5 + \tag{8.155}$$

$$+ \binom{6}{6}(x^2)^{6-6}(-2y)^6 \tag{8.156}$$

$$= \binom{6}{0}x^{12} + \binom{6}{1}x^{10}(-2y) + \binom{6}{2}x^8(-2y)^2 + \binom{6}{3}x^6(-2y)^3 + \binom{6}{4}x^4(-2y)^4 + \tag{8.157}$$

$$+ \binom{6}{5}x^2(-2y)^5 + \binom{6}{6}(-2y)^6 \tag{8.158}$$

$$= x^{12} + 6x^{10}(-2y) + 15x^8(-2y)^2 + 20x^6(-2y)^3 + 15x^4(-2y)^4 + 6x^2(-2y)^5 + (-2y)^6 \tag{8.159}$$

$$= x^{12} - 12x^{10}y + 60x^8y^2 - 160x^6y^3 + 240x^4y^4 - 192x^2y^5 + 64y^6 \tag{8.160}$$

6. (a) ¿Cuántos numeros de 3 digitos pueden ser formados por el conjunto de números A=0,1.......9 si no pueden existir numeros repetidos? y
   (b) ¿Cuántos números pueden formarse si el último dígito debe ser el numero 0 y no pueden existir numeros repetidos?

Solución:

(a) En primer inciso se requieren formar números de 3 dígitos empleando todos los elementos del conjunto sin repetir números, de esta forma podemos generar las permutaciones de el número total de elementos tomando solamente 3 de ellos a la vez:

$$P\binom{n}{k} = \frac{n!}{(n-k)!} \tag{8.161}$$

$$P\binom{10}{3} = \frac{10!}{(10-3)!} \tag{8.162}$$

$$= \frac{10!}{7!} \tag{8.163}$$

$$= \frac{10*9*8*\cancel{7!}}{\cancel{7!}} \tag{8.164}$$

$$= \mathbf{720} \tag{8.165}$$

(b) En la segunda pregunta solo podemos elegir 9 elementos de 10 y nuestras permutaciones consistirán unicamente en números de 2 dígitos pues el tercero ya estara definido, de esta forma aplicamos:

$$P\binom{n}{k} = \frac{n!}{(n-k)!} \tag{8.166}$$

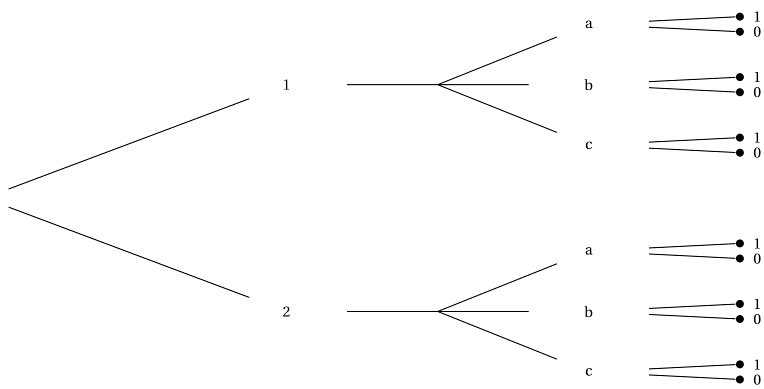
$$P\binom{9}{2} = \frac{9!}{(9-2)!} \tag{8.167}$$

$$= \frac{9!}{7!} \tag{8.168}$$

$$= \frac{9*8*\cancel{7!}}{\cancel{7!}} \tag{8.169}$$

$$= \mathbf{72} \tag{8.170}$$

7. Dados los iguientes conjuntos A=1,2, B=a,b,c y C=0,1, represente mediante un diagrama de arbol el producto de AxBxC
Solución: El árbol puede ser representado de la siguiente manera:



8. Determine mediante una expresión cuantos subconjuntos pueden existir dentro de un conjunto X que contiene N elementos
Solución: El número de subconjuntos de X conjunto con $r \leq n$ esta dado por $\binom{n}{r}$, esto se puede representar de las siguiente expresión:

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \ldots + \binom{n}{n-1} + \binom{n}{n} \tag{8.171}$$

(8.172)

9. Demuestre que $\binom{17}{6} = \binom{16}{5} + \binom{16}{6}$

Solución: Tomando como punto de partida $\binom{16}{5}+\binom{16}{6}$ y usando la formula de combinaciones se procede a demostrar:

$$\binom{17}{6} = \frac{n!}{r!(n-r)!}+\frac{n!}{r!(n-r)!} \tag{8.173}$$

$$= \frac{16!}{5!(16-5)!}+\frac{16!}{6!(16-6)!} \tag{8.174}$$

$$= \frac{16!}{5!(11)!}+\frac{16!}{6!(10)!} \tag{8.175}$$

$$= \frac{16!}{5!(11)!}*\frac{6}{6}+\frac{16!}{6!(10)!}*\frac{11}{11}, \text{ multiplicamos por uno ambos terminos} \tag{8.176}$$

$$= \frac{6*16!}{6*5!(11)!}+\frac{11*16!}{6!(10)!*11} \tag{8.177}$$

$$= \frac{6*16!}{6!*11!}+\frac{11*16!}{6!*11!} \tag{8.178}$$

$$= \frac{6*16!+11*16!}{6!*11!} \tag{8.179}$$

$$= \frac{16!(6+11)!}{6!*11!} \tag{8.180}$$

$$= \frac{16!(17)}{6!*11!} \tag{8.181}$$

$$= \frac{17!}{6!*11!} \tag{8.182}$$

10. Un departamento de policias consiste en 6 oficiales, el departamento de policias debe tener 3 oficiales patrullando las callles, 2 oficiales trabajando tiempo completo dentro de la estación y 1 oficial como reserva en la estación, ¿cuantas divisiones de los 10 oficiales en 3 grupos pueden ser posibles?
Solución:

Para resolver este problema utilizamos el teorema de coeficientes multinomial, que reprsenta el numero de posibles combinaciones de n objetos divididas en r distintos grupos :

$$\binom{n}{n_1, n_2, ..., n_r} = \frac{n!}{n_1! * n_2!, ..., n_r!} \tag{8.183}$$

$$= \frac{6!}{1! * 2! * 3!} \tag{8.184}$$

$$= \frac{720}{12} \tag{8.185}$$

$$= 60 \tag{8.186}$$

11. Escribe la formula para la aproximación de un numero factorial demasiado grande
    Solución: Usamos aproximacion de Stirling
    $n! = \sqrt{2\pi n n^n e^{-n}}$

12. Encuentra el número mínimo de estudiantes en una clase para asegurar que tres de ellos nacieron en el mismo mes.
    Solución: En este problema se hace uso del principio de la pichonera en el cual nuestra variable $n = 12$ representa las pichoneras, el problema menciona que pueden existir tres de ellos nacidos en el mismo mes, esto significa $k + 1 = 3$, y $k = 2$, usando el principio de la pichonera tenemos que el mínimo de estudiantes es igual a $kn + 1 = 25$

13. Una clase esta conformada por 8 estudiantes . Encuentra el numero de forma de:
    Seleccionar un comite con 4 estudiantes.

    Solución:

$$C\binom{n}{r} = \frac{n!}{n!(n-r)!} \tag{8.187}$$

$$= \frac{8!}{4!(8-4)!} \tag{8.188}$$

$$= \frac{8!}{4! * 4!} \tag{8.189}$$

$$= \frac{8 * 7 * 6 * 5 * \cancel{4!}}{4! * \cancel{4!}} \tag{8.190}$$

$$= \frac{1680}{24} \tag{8.191}$$

$$= \mathbf{70} \tag{8.192}$$

14. Desarrolla los coeficientes multinomiales de: $(x_1 + x_2 + x_3)^2$
    Solución:

$$
\begin{aligned}
&= \tbinom{2}{2,0,0} x_1^2 x_2^0 x_3^0 + \tbinom{2}{0,2,0} x_1^0 x_2^2 x_3^0 + \tbinom{2}{0,0,2} x_1^0 x_2^0 x_3^2 + && (8.193)\\
&\quad \tbinom{2}{1,1,0} x_1^1 x_2^1 x_3^0 + \tbinom{2}{1,0,1} x_1^1 x_2^0 x_3^1 + \tbinom{2}{0,1,1} x_1^0 x_2^1 x_3^1 && (8.194)\\
&= x_1^2 + x_2^2 + x_3^2 + 2x_1x_2 + 2x_1x_3 + 2x_2x_3 && (8.195)
\end{aligned}
$$

15. Se requiere asignar habitaciones a un grupo de n estudiantes dentro del piso de un edificio con n habitaciones hubicadas de forma lineal, considerando que solo un estudiante debe ser asignado a una habitación y dados dos estudiantes A y B, de cuantas formas pueden asignarse las habitaciones de modo que la habitación del estudiante A este junto a la habitación del estudiante B.

    Solución: Existen $n-1$ formas de elegir un par de habitaciones juntas y 2 formas de asignar a los estudiantes A y B de modo que las habitaciones esten juntas en cada par de habitaciones de modo que existen $2(n-1)$ formas de acomodar A y B, los otros $n-2$ estudiantes pueden ser asignados de $n-2$ formas, entonces el número total de formas es igual a :
    $2(n-1)(n-2)! = 2(n-1)!$

16. Cuantos números par de 2 dígitos pueden formarse? Considere el conjunto de numeros A=0,1,2,...9 en cada dígito
    Solución: Dado E el evento de elegir un digito para la pocision de las unidades y sea F el evento de elegir un número para las decenas, tenemos que el total de números par es igual a:
    Utilizando el principio multiplicativo $(9)(5) = 45$ numeros.

17. Demuestre que : $k\binom{n}{k} = n\binom{n-1}{k-1}$

    Solución:

$$
\begin{aligned}
k\binom{n}{k} &= \frac{kn!}{k!(n-k)!} && (8.196)\\
&= \frac{n(n-1)!}{(k-1)![(n-1)-(k-1)]!} && (8.197)\\
&= n\binom{n-1}{k-1} && (8.198)
\end{aligned}
$$

18. Demuestre que: $k\binom{n}{k} = (n-k+1)\binom{n}{k-1}$

    Solución:

$$k\binom{n}{k} = \frac{kn!}{k!(n-k)!} \tag{8.199}$$

$$= \frac{(n-k+1)n!}{(k-1)![(n-k+1)]!} \tag{8.200}$$

$$= (n-k+1)\binom{n}{k-1} \tag{8.201}$$

19. Demuestre que : $\binom{n+1}{r} = \binom{n}{r-1} + \binom{n}{r}$

    Solución: Sabemos que:

$$\binom{n}{r-1} + \binom{n}{r} = \frac{n!}{(r-1)!(n-r+1)!} + \frac{n!}{r!(n-r)!} \tag{8.202}$$

$$= \frac{n!}{(r-1)!(n-r+1)!} * \frac{r}{r} + \frac{n!}{r!(n-r)!} * \frac{n-r+1}{n-r+1} \tag{8.203}$$

$$= \frac{rn!}{r(r-1)!(n-r+1)!} + \frac{(n-r+1)n!}{r!(n-r+1)(n-r)!} \tag{8.204}$$

$$= \frac{rn!}{r(r-1)!(n-r+1)!} + \frac{(n-r+1)n!}{r!(n-r+1)!} \tag{8.205}$$

$$= \frac{rn! + (n-r+1)n!}{r!(n-r+1)!} \tag{8.206}$$

$$= \frac{[r+(n-r+1)]n!}{r!(n-r+1)!} \tag{8.207}$$

$$= \frac{(n+1)n!}{r!(n-r+1)!} \tag{8.208}$$

$$= \frac{(n+1)!}{r!(n-r+1)!} \tag{8.209}$$

$$= \binom{n+1}{r} \tag{8.210}$$

20. Simplifique : $\dfrac{n!}{(n-1)!}$

    Solución: $\frac{n!}{(n-1)!} = \frac{n(n-1)(n-2)...3\cdot 2\cdot 1}{(n-1)(n-2)...3\cdot 2\cdot 1} = n$

$$\frac{n!}{(n-1)!} = \frac{n(n-1)(n-2),\ldots,3*2*1}{(n-1)(n-2),\ldots,3*2*1} \tag{8.211}$$

$$= \frac{n\cancel{(n-1)}\cancel{(n-2)},\ldots,\cancel{3}*\cancel{2}*\cancel{1}}{\cancel{(n-1)}\cancel{(n-2)},\ldots,\cancel{3}*\cancel{2}*\cancel{1}} \tag{8.212}$$

$$= n \tag{8.213}$$

8.3 Exámen a evaluar

Probabilidad, procesos aleatorios e interferencia

Autor: Sandra Lizeth Sánchez Gonzáles
Estatus: Revisión
Nombre: Victor Manuel Rangel Fajardo____________

1. Cuántas permutaciones distintas pueden formarse con todas las letras de cada una de las siguientes palabras: i)Sufre, ii)examen, iii)abandonado.

   Solución:

   (a) Sea $n = 5$, entonces $5!$;
   (b) Sea $n = 6$, con 2 elementos repetidos, entonces $\frac{6!}{2!}$;
   (c) Sea $n = 10$, con $3, 2, 2, 2$ elementos repetidos, entonces $\frac{10!}{3!2!2!2!}$

2. Simplificar $\dfrac{(n+2)!}{n!}$

   Solución:

$$\frac{(n+2)(n+1)\cancel{(n!)}}{\cancel{(n!)}} = (n+2)(n+1) = n^2 + 3n + 2 \tag{8.214}$$

3. i) En cuantas maneras diferentes pueden ser colocadas 7 personas en 7 sillas, ii) cuantas si son colocadas en una mesa circular.

   Solución:

   Sea $n = 7$
   i)Para sillas de forma lineal 7!
   ii)Para sillas en forma circular $(n-1)! = 6!$

4. Encuentre el valor de n si $P(n, 4) = 42P(n, 2)$

   Solución:

$$\begin{aligned} n(n-1)(n-2)(n-3) &= 42n(n-1) & (8.215)\\ (n-2)(n-3) &= 42 & (8.216)\\ & n^2-5n-36 & (8.217)\\ n_1 = 9 \text{ y } n_2 = -4 & & (8.218) \end{aligned}$$

Como n debe de ser positiva, entonces $n = 9$

5. Demuestra el siguiente teorema:

$$\binom{k+1}{r} = \binom{k}{r-1} + \binom{k}{r}$$

Solución:

$$\begin{aligned} \binom{k}{r-1} &= \frac{k!}{r!(k-(r-1))!} = \frac{k!}{\frac{r!}{r}(k-r+1)!} & (8.219)\\ \binom{k}{r} &= \frac{k!}{r!(k-r)!} = \frac{k!}{r!\frac{(k-r+1)!}{k-r+1}} & (8.220)\\ \binom{k+1}{r} &= \frac{k!}{\frac{r!}{r}(k-r+1)!} + \frac{k!}{r!\frac{(k-r+1)!}{k-r+1}} & (8.221)\\ \binom{k+1}{r} &= \frac{rk!}{r!(k-r+1)!} + \frac{k!(k-r+1)}{r!(k-r+1)!} & (8.222)\\ \binom{k+1}{r} &= \frac{k!(r+(k-r+1))}{r!(k-r+1)!} = \frac{k!(k-r+1)}{r!(k-r+1)!} & (8.223)\\ \binom{k+1}{r} &= \frac{(k+1)!}{r!(k-r+1)!} & (8.224) \end{aligned}$$

6. En cuantas maneras puede formarse un comité que consta de 3 hombres y 2 mujeres y estos son escogidos de 7 hombres y 5 mujeres.

Solución:

- 3 hombres de 7 $\therefore \binom{7}{3}$
- 2 mujeres de 5 $\therefore \binom{5}{2}$

Usando el principio multiplicativo, hay:

$$\frac{7!}{3!4!} \cdot \frac{5!}{2!3!} = 350 maneras \tag{8.225}$$

7. i) Encuentra el numero de las distintas permutaciones que pueden ser formadas de la palabra "AVANZA"
ii) Cuantas de estas permutaciones empiezan y terminan con "A"
iii) Cuantas tienen 3 "A's"juntas
$iiii$) Cuantas empiezan con "A"y terminan con "Z"

Solución:

i)Sea $n = 6$, con 3 elementos repetidos, entonces $\frac{6!}{3!} = 120$
ii)Si empieza y termina con A, entonces $n = 4 \therefore 4! = 24$
iii)Se consideran las A's con un solo elementos, entonces $n = 4 \therefore 4! = 24$
iv)Sea $n = 4$ quitando conociendo el primer y ultimo elementos, además hay 2 elementos repetidos, entonces $\frac{4!}{2!} = 12$

8. Cuantas soluciones distinta tiene la siguiente ecuación:

$$x_1 + x_2 + x_3 + x_4 = 100 \text{tal que} \tag{8.226}$$

$$x_1 \in \{1, 2, 3, \ldots\}; x_2 \in \{2, 3, 4, \ldots\}; x_3, x_4 \in \{0, 1, 2, 3, \ldots\} \tag{8.227}$$

Solución:
se dice:

$$\binom{n+k-1}{k} = \binom{n+k-1}{n-1} \tag{8.228}$$

Nos adecuamos las retricciones para coincidir con la formula general
proponemos:

$$y_1 = x_1 - 1 \text{ entonces} \in \{0, 1, 2, 3, \ldots\}\ y_2 = x_2 - 2 \text{ entonces} \in \{0, 1, 2, 3, \ldots\}$$

El problema queda

$$y_1 + 1 + y_2 + 2 + x_3 + x_4 = 100 \text{ donde } y_1; y_2; x_3; x_4 \in \{0, 1, 2, 3, \ldots\} \tag{8.229}$$

$$y_1 + y_2 + x_3 + x_4 = 97 \text{ donde } y_1; y_2; x_3; x_4 \in \{0, 1, 2, 3, \ldots\} \tag{8.230}$$

$$\binom{4+97-1}{3} = \binom{100}{3} \tag{8.231}$$

9. Cuantos mensajes telegráficos diferentes se pueden enviar utilizando exactamente 4 puntos y 5 rayas

   Solución:
   Sea $n = 9$, además hay 4 y 5 elementos repetidos (puntos y rayas), entonces:

$$\frac{9!}{4!5!} = 126 \text{ mensajes} \tag{8.232}$$

10. Si en un salón hay 14 estudiantes ¿Cuantos partidos diferentes de voleibol se podrían realizar si los equipos son de 6 jugadoras ?

    Solución:
    Se forman dos equipos, tal que:

    - Primer equipo: $\binom{14}{6}$
    - Segundo equipo: $\binom{8}{6}$

    Como una sexteta puede pertenecer a ambas combinaciones, se divide entre la permutación de los 2 equipos.

$$\frac{\frac{14!}{6!(14-6)!} \cdot \frac{8!}{6!(8-6)!}}{2!} = 42042 \text{ partidos} \tag{8.233}$$

11. Encuentra el termino que contiene y^6 en la expancion de $(3xy^2 - z^2)^7$

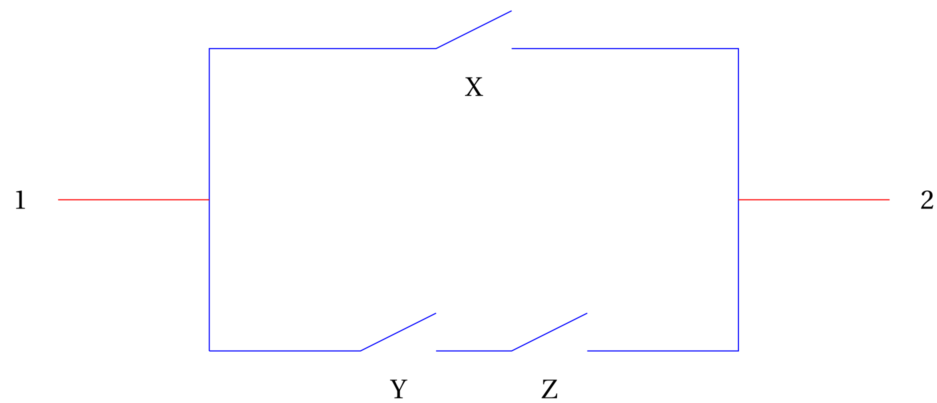
    Solución:
    Sea sabe que el teorema del binomio es:

$$(a+b)^n = \sum_{r=0}^{n} \binom{n}{r} a^{n-r} b^r \tag{8.234}$$

y si queremos el término y^6, entonces $n - r = 3$, si se sabe que $n = 7$, entonces $r = 4$

$$\binom{7}{4} (3xy^2)^{7-4} (z^2)^4 = 945x^3y^6z^8 \tag{8.235}$$

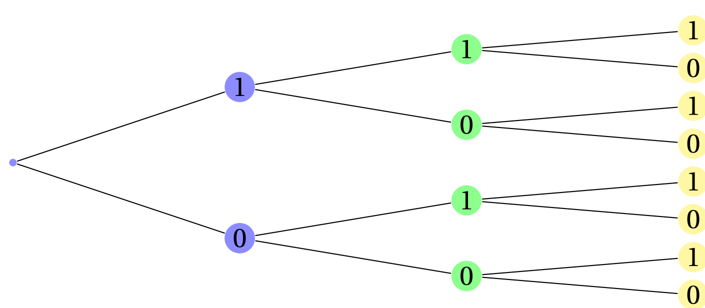
12. En el circuito mostrado en la siguiente figura,la corriente fluye de la terminal 1 a la terminal 2, siempre que el interruptor X esté cerrado, o que los interruptores Y y Z, ambos estén cerrados.



El experimento consiste en observar el funcionamiento de un interruptor,que puede presentar uno de dos estados: 0, abierto o 1,cerrado, generando el espacio muestral $S_1 = \{0, 1\}$.
Observando el funcionamiento de los tres interruptores, simultáneamente. Construya el diagrama de árbol asociado a tal experimento.

Solución:

El funcionamiento de cada interruptor es independiente del funcionamiento de los otros dos, por lo que cada interruptor puede presentar uno de dos estados: 0, abierto o 1, cerrado. Entonces, el diagrama de árbol correspondiente es:



Donde los estados de activación son:

X	Y	Z
1	0	0
1	1	0
1	0	1
1	1	1
0	1	1

13. La selección mexicana está integrada por 25 jugadores en total, de los cuales tres son porteros, siete defensas, diez medios y cinco delanteros. ¿De cuántas maneras puede el entrenador integrar un equipo de once jugadores, si cualquiera de ellos puede ocupar cualquier posición?

    Solución:

$$\binom{25}{11} = \frac{25!}{11!(25-11)!} = 4457400 \tag{8.236}$$

14. Del problema anterior ¿De cuántas maneras puede integrar el entrenador en equipo que tenga un portero,cuatro defensas, cuatro medios y dos delanteros?

    Solución:

    - Para el portero es: $\binom{3}{1}$
    - Para los defensas es: $\binom{7}{4}$
    - Para los medios es: $\binom{10}{4}$
    - Para los delanteros es: $\binom{5}{2}$

$$C\binom{3}{1} * C\binom{7}{4} * C\binom{10}{4} * C\binom{5}{2} = 3 * 840 * 210 * 10 = 5,292,000 \text{ Equipos} \tag{8.237}$$

15. Cuatro libros distintos de matemáticas, seis diferentes de física y dos diferentes de química se colocan en un estante. De cuántas formas distintas es posible ordenarlos si, los libros de cada asignatura deben estar todos juntos.

    Solución:

$$4! \cdot 6! \cdot 2! \cdot 3! = 207360 \tag{8.238}$$

16. ¿Cuántas diagonales tiene un pentágono y cuántos triángulos se puede formar con sus vértices?, si se sabe que:

    - No entran todos los elementos.
    - No importa el orden.
    - No se repiten los elementos.

    Solución: Vamos a determinar en primer lugar las rectas que se pueden trazar entre 2 vértices y suponiendo que entre dos vértices ya hay una línea, entonces:

    hay $C\binom{5}{2}$, a las que tenemos que restar los lados que determinan 5 rectas que no son diagonales.

$$C\binom{5}{2} - 5 = \frac{5!}{2!3!} - 5 = 5 \text{ Diagonales} \tag{8.239}$$

$$C\binom{5}{3} - 5 = \frac{5!}{3!2!} = 10 \text{ Triangulos} \tag{8.240}$$

17. Demuestre:

$$\binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \binom{n}{3} + \ldots \pm \binom{n}{n} = 0 \tag{8.241}$$

Solución:
expandir : $(1-1)^n$

$$(a+b)^n = \sum_{r=0}^{n} \binom{n}{r} a^{n-r} b^r \tag{8.242}$$

$$(1+(-1))^n = \binom{n}{0} 1^{n-0}(-1)^0 + \binom{n}{1} 1^{n-1}(-1)^1 + \ldots + \binom{n}{n} 1^{n-n}(-1)^n = 0 \tag{8.243}$$

$$(1+(-1))^n = \binom{n}{0} \cancel{1^{n-0}} - \binom{n}{1} \cancel{1^{n-1}} + \ldots + \binom{n}{n} 1^{n-n}(-1)^n = 0 \tag{8.244}$$

18. Con las cifras 1, 2 y 3, ¿cuántos números de cinco cifras pueden formarse? ¿Cuántos son pares?

- Sí entran todos los elementos
- Sí importa el orden.
- Sí se repiten los elementos.

Solución:

$$3 \times 3 \times 3 \times 3 \times 3 = 3^5 = 243 \tag{8.245}$$

son pares :

$$3 \times 3 \times 3 \times 3 \times 3 = 3^4 = 81 \text{son pares} \tag{8.246}$$

19. ¿Cuántos números de cinco cifras distintas se pueden formar con las cifras impares? ¿Cuántos de ellos son mayores de 70.000?

    - Sí entran todos los elementos.
    - Sí importa el orden.
    - No se repiten los elementos.

    Solución:
    hay 5 números impares que se pueden ordenar en 5 lugares, esto es:

    $$5 \times 4 \times 3 \times 2 \times 1 = 5! = 120 \tag{8.247}$$

    Ahora considerando solo aquellos mayores de 70,000, entonces el primer posición será ocupada por el 7 o el 9 (2!), los números restantes serán acomodados en las 4 posiciones faltantes (4!), entonces hay:

    $$2! \cdot 4! = 48 \text{números mayores a } 70{,}000 \tag{8.248}$$

20. Calcular

    $$i)\binom{7}{5,3,6} \tag{8.249}$$

    $$ii)\binom{12}{10,3,5} \tag{8.250}$$

    Solución:

    $$i)\frac{14!}{5! \cdot 3! \cdot 6!} \tag{8.251}$$

    $$ii)\text{no se puede ya que excede el número máximo de elementos } (10+3+5 \neq 12) \tag{8.252}$$

9 Probabilidad condicional

9.1 Teorema de Bayes

El teorema de Bayes, en la teoría de la probabilidad, es una proposición planteada por el filósofo inglés Thomas Bayes (1702-1761) en 1763, que expresa la probabilidad condicional de un evento aleatorio A dado B en términos de la distribución de probabilidad condicional del evento B dado A y la distribución de probabilidad marginal de sólo A.

En términos más generales y menos matemáticos, el teorema de Bayes es de enorme relevancia puesto que vincula la probabilidad de A dado B con la probabilidad de B dado A. Es decir, que sabiendo la probabilidad de tener un dolor de cabeza dado que se tiene gripe, se podría saber (si se tiene algún dato más), la probabilidad de tener gripe si se tiene un dolor de cabeza. Muestra este sencillo ejemplo la alta relevancia del teorema en cuestión para la ciencia en todas sus ramas, puesto que tiene vinculación íntima con la comprensión de la probabilidad de aspectos causales dados los efectos observados.

Sea $\{A_1, A_2, \ldots, A_i, \ldots, A_n\}$ un conjunto de sucesos mutuamente excluyentes y exhaustivos, y tales que la probabilidad de cada uno de ellos es distinta de cero (0). Sea B un suceso cualquiera del que se conocen las probabilidades condicionales $P(B|A_i)$. Entonces, la probabilidad $P(A_i|B)$ viene dada por la expresión:

$$P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{P(B)} \tag{9.1}$$

donde:

- $P(A_i)$ son las probabilidades a priori,
- $P(B|A_i)$ es la probabilidad de B en la hipótesis A_i,
- $P(A_i|B)$ on las probabilidades a posteriori.

9.2 Fórmula Bayes

Con base en la definición de Probabilidad condicionada, obtenemos la Fórmula de Bayes, también conocida como la Regla de Bayes:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{k=1}^{N} P(B|A_k)P(A_k)} \tag{9.2}$$

Esta fórmula nos permite calcular la probabilidad condicional $P(A_i|B)$ de cualquiera de los eventos $P(A_i)$, dado B.

9.3 Aplicaciones

El teorema de Bayes es válido en todas las aplicaciones de la teoría de la probabilidad. Sin embargo, hay una controversia sobre el tipo de probabilidades que emplea. En esencia, los seguidores de la estadística tradicional sólo admiten probabilidades basadas en experimentos repetibles y que tengan una confirmación empírica mientras que los llamados estadísticos bayesianos permiten probabilidades subjetivas. El teorema puede servir entonces para indicar cómo debemos modificar nuestras probabilidades subjetivas cuando recibimos información adicional de un experimento. La estadística bayesiana está demostrando su utilidad en ciertas estimaciones basadas en el conocimiento subjetivo a priori y el hecho de permitir revisar esas estimaciones en función de la evidencia empírica es lo que está abriendo nuevas formas de hacer conocimiento. Una aplicación de esto son los clasificadores bayesianos que son frecuentemente usados en implementaciones de filtros de correo basura o spam, que se adaptan con el uso. Otra aplicación se encuentra en la fusión de datos, combinando información expresada en términos de densidad de probabilidad proveniente de distintos sensores.

Como observación, se tiene $\sum\limits_{k=1}^{N} P(A_i|B) = 1$ y su demostración resulta trivial.

Como aplicaciones puntuales:

1. El diagnóstico de cáncer.
2. Evaluación de probabilidades durante el desarrollo de un juego de bridge por Dan F. Waugh y Frederick V. Waugh.
3. Probabilidades a priori y a posteriori.
4. Un uso controvertido en la Ley de sucesión de Laplace.
5. En el testeo de hipótesis en Ciencia Política cuando se usa metodología process tracing.

Ejemplo 9.1

En la sala de pediatría de un hospital, el 60% de los pacientes son niñas. De los niños el 35% son menores de 24 meses. El 20% de las niñas tienen menos de 24 meses. Un pediatra que ingresa a la sala selecciona un infante al azar.

- Determine el valor de la probabilidad de que sea menor de 24 meses.
- Si el infante resulta ser menor de 24 meses. Determine la probabilidad que sea una niña.

Solución: Se definen los sucesos:

- Suceso H: seleccionar una niña.
- Suceso V: seleccionar un niño.
- Suceso M: infante menor de 24 meses.

En los ejercicios de probabilidad total y teorema de bayes, es importante identificar los sucesos que forman la población y cuál es la característica que tienen en común dichos sucesos. Estos serán los sucesos condicionados.

1. En este caso, la población es de los infantes. Y la característica en común es que sean menores de 24 meses. Por lo tanto, la probabilidad de seleccionar un infante menor de 24 meses es un ejemplo de probabilidad total. Su probabilidad será:

$$P(M) = P(H) * P(\frac{M}{H}) + P(V) * P(\frac{M}{V}) \tag{9.3}$$
$$= (0.6)(0.2) + (0.4)(0.35) \tag{9.4}$$
$$= 0.26 \tag{9.5}$$
$$= 26\% \tag{9.6}$$

2. Para identificar cuando en un ejercicio se hace referencia al teorema de bayes, hay que partir de reconocer esta es una probabilidad condicionada y que la característica común de los sucesos condicionantes ya ha ocurrido. Entonces, la probabilidad de que sea niña una infante menor de 24 meses será:

$$P(\frac{H}{M}) = \frac{P(H) * P(\frac{M}{H})}{P(H) * P(\frac{M}{H}) + P(V) * P(\frac{M}{V})} \tag{9.7}$$
$$= \frac{(0.6)(0.2)}{(0.6)(0.2) + (0.4)(0.35)} \tag{9.8}$$
$$= \frac{0.12}{0.26} \tag{9.9}$$
$$= 0.46 \tag{9.10}$$
$$= 46\% \tag{9.11}$$

Ejemplo 9.2

Un médico cirujano se especializa en cirugías estéticas. Entre sus pacientes, el 20% se realizan correcciones faciales, un 35% implantes mamarios y el restante en otras cirugías correctivas. Se sabe además, que son de genero masculino el 25% de los que se realizan correcciones faciales, 15% implantes mamarios y 40% otras cirugías correctivas. Si se selecciona un paciente al azar, determine:

- Determine la probabilidad de que sea de género masculino
- Si resulta que es de género masculino, determine la probabilidad que se haya realizado una cirugía de implantes mamarios.

Solución: Se definen los sucesos:

- Suceso F: pacientes que se realizan cirugías faciales
- Suceso M: pacientes que se realizan implantes mamarios
- Suceso O: pacientes que se realizan otras cirugías correctivas
- Suceso H: pacientes de género masculino

1. La probabilidad de que sea de género masculino se refiere a un problema de probabilidad total, ya que es el suceso condicionado y las cirugías los condicionantes. Dicho valor será:

$$
\begin{aligned}
P(H) &= P(F)*P(\frac{H}{F})+P(M)*P(\frac{H}{M}+P(O)*P(\frac{O}{H}) && (9.12)\\
&= (0.2)(0.25)+(0.35)(0.15)+(0.45)(0.40) && (9.13)\\
&= 0.28 && (9.14)\\
&= 28\% && (9.15)\\
& && (9.16)
\end{aligned}
$$

2. Como el suceso condicionado ha ocurrido entonces se aplica el teorema de bayes, luego, el valor de la probabilidad será:

$$
\begin{aligned}
P(\frac{M}{H}) &= \frac{P(M)*P(\frac{H}{M})}{P(F)*P(\frac{H}{F})+P(M)*P(\frac{H}{M}+P(O)*P(\frac{O}{H})} && (9.17)\\
&= \frac{(0.35)(0.15)}{(0.2)(0.25)+(0.35)(0.15)+(0.45)(0.40)} && (9.18)\\
&= \frac{0.0525}{0.2825} && (9.19)\\
&= 0.19 && (9.20)\\
&= 19\% && (9.21)\\
& && (9.22)
\end{aligned}
$$

Ejemplo 9.3

Un Doctor dispone de tres equipos electrónicos para realizar ecosonogramas. El uso que le da a cada equipo es de 25% al primero, 35% el segundo en y 40% el tercero. Se sabe que los aparatos tienen probabilidades de error de 1%, 2% y 3% respectivamente. Un paciente busca el resultado de una ecografía y observa que tiene un error. Determine la probabilidad de que se ha usado el primer aparato.
Solución: Se definen los sucesos:

- Suceso P: seleccionar el primer aparato
- Suceso S: seleccionar el segundo aparato
- Suceso T: seleccionar el tercer aparato
- Suceso E: seleccionar un resultado con error

Se puede observar que la pregunta es sobre determinar la probabilidad de que un examen errado sea del primer aparato, es decir, ya ha ocurrido el error. Por lo tanto, debemos recurrir al teorema de bayes. Claro está, que es necesario de igual forma obtener la probabilidad de que los aparatos produzcan un resultado erróneo, por lo tanto:

$$P(\frac{P}{E}) = \frac{P(P)*P(\frac{E}{P})}{P(P)*P(\frac{E}{P})+P(S)*P(\frac{E}{S}+P(T)*P(\frac{E}{T})} \tag{9.23}$$

$$= \frac{(0.25)(0.01)}{(0.25)(0.01)+(0.35)(0.02)+(0.4)(0.03)} \tag{9.24}$$

$$= \frac{0.0025}{0.0215} \tag{9.25}$$

$$= 0.12 \tag{9.26}$$

$$= 12\% \tag{9.27}$$

10 Paradojas

Una paradoja (del latín paradoxa, plural de paradoxon, ”lo contrario a la opinión común”, y este del griego [parádoxa], plural de [parádoxon]) o antilogía es una idea extraña opuesta a lo que se considera verdadero a la opinión general. También se considera paradoja a una proposición en apariencia falsa o que infringe el sentido común, pero no conlleva una contradicción lógica, en contraposición a un sofisma que solo aparenta ser un razonamiento verdadero. En retórica, es una figura de pensamiento que consiste en emplear expresiones o frases que implican contradicción. Un ejemplo de paradoja es la “Paradoja de Jevons”, más conocida como efecto rebote. La paradoja es estímulo para la reflexión. A menudo los filósofos se sirven de las paradojas para revelar la complejidad de la realidad. La paradoja también permite demostrar las limitaciones de las herramientas de la mente humana. Así, la identificación de paradojas basadas en conceptos que a simple vista parecen simples y razonables ha impulsado importantes avances en la ciencia, la filosofía y las matemáticas.

Según su veracidad y las condiciones que las forman:

Paradojas verídicas Son resultados que aparentan tal vez ser absurdos a pesar de ser demostrable su veracidad. A esta categoría pertenecen la mayor parte de las paradojas matemáticas.

- Paradoja del cumpleaños: ¿Cuál es la probabilidad de que dos personas en una reunión cumplan años el mismo día?
- Paradoja de Galileo: A pesar de que no todos los números son cuadrados perfectos, no hay más números que cuadrados perfectos.
- Paradoja del hotel infinito: Un hotel de infinitas habitaciones puede aceptar más huéspedes, incluso si está lleno.
- Paradoja de la banda esférica: No es una paradoja en sentido estricto, pero choca con nuestro sentido común debido a que tiene una solución que parece imposible.

Antinomias Son paradojas que alcanzan un resultado que se autocontradice, aplicando correctamente modos aceptados de razonamiento. Muestran fallos en un modo de razón, axioma o definición previamente aceptados. Por ejemplo, la Paradoja de Grelling-Nelson señala problemas genuinos en nuestro modo de entender las ideas de verdad y descripción. Muchos de ellos son casos específicos, o adaptaciones, de la Paradoja de Russell.

- Paradoja de Russell: ¿Existe un conjunto de todos los conjuntos que no se contienen a sí mismos?
- Paradoja de Curry: ”Si no me equivoco, el mundo se acabará en diez días”.
- Paradoja del mentiroso:4 ”Esta oración es falsa”.
- Paradoja de Grelling-Nelson: ¿Es la palabra ”heterológico”, que significa ”que no describe a sí mismo”, heterológica?
- Paradoja de Berry: ”El menor entero positivo que no se puede definir con menos de quince palabras”.
- Paradoja de la suerte: Es de mala suerte ser supersticioso.
- Paradoja de los números interesantes: Todo número entero presenta alguna propiedad interesante específica, y por tanto el conjunto de los números no-interesantes es vacío

Antinomias de definición Estas paradojas se basan en definiciones ambiguas, sin las cuales no alcanzan una contradicción. Este tipo de paradojas constituye un recurso literario, en cuyo empleo se ha destacado el escritor inglés G. K. Chesterton, a quién se llamó el ”príncipe de las paradojas”. Sirviéndose de los múltiples sentidos de las palabras, buscaba marcar contrastes que llamaran la atención sobre alguna cuestión comúnmente poco considerada. Estas paradojas, como en su libro ”Las paradojas de Mr. Pond” (1936), se resuelven en el transcurso de los relatos al clarificar un sentido o añadir alguna información clave.

- Paradoja sorites: ¿En qué momento un montón deja de serlo cuando se quitan granos de arena?
- Paradoja de Teseo: Cuando se han reemplazado todas las partes de un barco, ¿sigue siendo el mismo barco?
- Paradoja de Boixnet: Pienso, luego existo, mas cuando no pienso, ¿no existo?
- Ejemplos de Paradoja en Chesterton: ”Era un extranjero muy deseable, y a pesar de eso, no lo deportaron”. ”Una vez conocí a dos hombres que estaban tan completamente de acuerdo que, lógicamente, uno mató al otro”.

Paradojas condicionales Sólo son paradójicas si se hacen ciertas suposiciones. Algunas de ellas muestran que esas suposiciones son falsas o incompletas. El huevo o la gallina: Esta trae a colación el antiguo dilema sobre quien fue primero, ¿el huevo o la gallina?

- Paradoja de Newcomb: Cómo jugar contra un oponente omnisciente.
- Paradoja de San Petersburgo: La gente solo arriesgará una pequeña cantidad para obtener una recompensa de valor infinito.
- Paradoja del viaje en el tiempo: ¿Qué pasaría si viajas en el tiempo y matas a tu abuelo antes de que conozca a tu abuela?
- Paradoja de la serpiente: Si una serpiente se empieza a comer su cola, acaba comiéndose absolutamente todo su cuerpo, ¿dónde estaría la serpiente, si está dentro de su estómago que, a su vez, está dentro de ella?

- Paradoja de Pinocho : ¿Qué pasaría si Pinocho dijera: ”Ahora mi nariz crecerá”?. Crecería porque estaría mintiendo, a su vez al crecer su nariz se expresa la validez de la frase propiamente dicha anteriormente como resultado de esto estaría diciendo la verdad ya que pasó lo que predijo, entonces al ser verdad lo que dijo no tendría que haberle crecido la nariz.

Según el área del conocimiento al que pertenecen Todas las paradojas se consideran relacionadas con la lógica, que antiguamente se consideraba parte de la filosofía, pero que ahora se ha formalizado y se ha incluido como una parte importante de la matemática. A pesar de ello, muchas paradojas han ayudado a entender y a avanzar en algunas áreas concretas del conocimiento.
Paradojas en Matemática / Lógica

- Paradoja de Banach-Tarski
- Paradoja de Frege

Paradojas sobre la probabilidad y la estadística

- Paradoja del cumpleaños: ¿Cuál es la probabilidad de que dos personas en una reunión cumplan años el mismo día?
- Paradoja de Simpson: Al agregar datos, podemos encontrar relaciones engañosas.
- Paradoja de Arrow: No puedes tener todas las ventajas de un sistema de votación ideal al mismo tiempo. Problema de Monty Hall: Y tras la puerta número dos... (¿Por qué la probabilidad no es intuitiva?)
- Paradoja de San Petersburgo: Cómo no merece la pena arriesgar mucho para ganar un premio infinito. Fenómeno Will Rogers: Sobre el concepto matemático de la media, trata sobre la media o mediana de dos conjuntos cuando uno de sus valores es intercambiado entre ellos, dando lugar a un resultado aparentemente paradójico.
- Paradoja de los dos sobres: Uno de los sobres contiene el doble de dinero que el otro. Sin importar cuál de los dos sobres esté en mi poder, las probabilidades siempre indican que es favorable cambiarlo por el sobre restante.

Paradojas sobre lógica A pesar de que todas las paradojas se consideran relacionadas con la lógica, hay algunas que afectan directamente a su bases y postulados tradicionales. Las paradojas más importantes relacionadas directamente con el área de la lógica son las antinomias, como la paradoja de Russell, que muestran la inconsistencia de las matemáticas tradicionales. A pesar de ello, existen paradojas que no se autocontradicen y que han ayudado a avanzar en conceptos como demostración y verdad.

- Paradoja del actual rey de Francia: ¿Es cierta una afirmación sobre algo que no existe?
- Paradoja del cuervo (o cuervos de Hempel): Una manzana roja incrementa la probabilidad de que todos los cuervos sean negros.
- Regresión infinita del presupuesto: ”Todo nombre que designa un objeto puede convertirse a su vez en objeto de un nuevo nombre que designe su sentido”.

Paradojas sobre el infinito El concepto matemático de infinito, al ser contrario a la intuición, ha generado muchas paradojas desde que fue formulado. Es importante resaltar que estos casos muestran una paradoja pero no en el sentido de una contradicción lógica, sino en el sentido de que muestran un resultado contrario a la intuición, pero demostrablemente cierto.

- Paradoja de Galileo: A pesar de que no todos los números son números cuadrados, no hay más números que números cuadrados.
- Paradoja del hotel infinito: Un hotel de infinitas habitaciones puede aceptar más huéspedes, incluso si está lleno.
- Conjunto de Cantor: Cómo quitar elementos de un conjunto y que siga teniendo el mismo tamaño.
- Cuerno de Gabriel (o Trompeta de Torricelli): ¿Cómo puede ser necesaria una superficie infinita para contener un volumen finito?
- Paradojas de Zenón. Mediante el concepto de división al infinito, Zenón trató de demostrar que el movimiento no puede existir, confirmando así la filosofía de su maestro, Parménides. Las más conocidas son la ”dicotomía” y la paradoja de ”Aquiles y la tortuga”.

Paradojas geométricas

- Ilusiones ópticas
- La serie de Fibonacci
- Disposición de hojas en un tallo
- División áurea
- Espiral logarítmica
- ¿Interior o exterior?
- Problema de los puentes de Königsberg
- Botella de Klein
- Banda de Möbius
- Problema de los cuatro colores

Paradojas en Física Richard Feynman en sus Lectures on Physics, aclara que en la Física realmente no existen las paradojas, sino que en las paradojas físicas hay siempre una mala interpretación de alguno o ambos razonamientos que componen la paradoja. Esto no es necesariamente válido en otras disciplinas donde las paradojas reales pueden existir.

- Paradoja de Bell: Plantea un problema clásico de relatividad especial.

- Paradoja de Olbers: ¿Por qué, si hay infinitas estrellas, el cielo es negro? Olberts calculó que la luminosidad del cielo correspondería a una temperatura del orden de los 5.500 °C, que, de hecho, no se observa. Actualmente se sabe que la luminosidad calculada por Olberts no llega a ser tal por el importante corrimiento al rojo de las fuentes de luz más alejadas, hecho que la teoría más aceptada atribuye al alejamiento de las galaxias o expansión del universo. Además se oponen la edad finita del universo, sus cambios notables durante su historia y que la cantidad de galaxias no es infinita. La paradoja proviene de un tiempo en el que no se conocían las galaxias y tendía a creerse que el universo era infinito y estático, por lo que también era plausible que hubiera infinitas estrellas.6
- Paradoja de Maxwell o Demonio de Maxwell: Una aparente paradoja clásica de la termodinámica.
- Paradoja de los gemelos: Cuando uno de los hermanos regresa de un viaje a velocidades cercanas a las de la luz descubre que es mucho más joven que su hermano.
- Paradoja de Einstein-Podolsky-Rosen: Una paradoja sobre la naturaleza de la mecánica cuántica propuesta por estos tres físicos.
- Paradoja de Fermi: Si el Universo estuviera poblado por civilizaciones avanzadas tecnológicamente, ¿dónde están?
- El experimento de Young. Una paradoja cuántica en su versión electrón a electrón. En el experimento de Young se pueden hacer pasar electrones por una doble rendija uno a uno de manera corpuscular, como si fueran partículas, obteniéndose sin embargo una figura de interferencias.
- Paradoja de Schrödinger: La paradoja por excelencia de la mecánica cuántica.
- Paradoja de D'Alembert: Relacionada con la resistencia de los cuerpos ante fluidos viscosos y no viscosos, en Mecánica de Fluidos.

Paradojas en Economía

- Paradoja de Abilene: Un grupo de personas frecuentemente toman decisiones contra sus propios intereses.
- Paradoja del ahorro: Si todo el mundo trata de ahorrar durante una recesión, la demanda agregada caerá y los ahorros totales de la población serán más bajos, esta paradoja es similar a la paradoja de Kalecki.
- Paradoja de Allais: En cierto tipo de apuestas, aun cuando la gente prefiere la certeza a la incertidumbre, si se plantea de manera diferente el problema, preferirán la incertidumbre que antes rechazaban.
- Paradoja de Bertrand: Dos jugadores que alcanzan el mismo equilibrio de Nash se encuentran cada uno sin ningún beneficio.
- Paradoja del pájaro en el arbusto: ¿Por qué las personas evitan el riesgo?
- Paradoja del valor (o paradoja del diamante y el agua): ¿Por qué es más barata el agua que los diamantes, siendo que los humanos necesitan agua, y no diamantes, para sobrevivir?

- Paradoja de Edgeworth: Con restricciones de capacidad, no puede haber ningún equilibrio.
- Paradoja de Ellsberg: En cierto tipo de apuestas, aun cuando sean lógicamente equivalentes las personas apostar por algo que contra algo, es decir, obtienen mayor utilidad apostando a favor.
- Paradoja de Gibson: ¿Por qué están los tipos de interés y los precios positivamente correlacionados?
- Paradoja de Giffen: ¿Puede ser que los pobres coman más pan aunque suba su precio?
- Paradoja de Jevons: Un incremento en la eficiencia conlleva un mayor incremento en la demanda.
- Paradoja de Kalecki de los costes: Un descenso generalizado de los salarios (reducción de costes) y precios fijos lejos en lugar de aumentar los beneficios reducen las ventas por una caída de la demanda agregada.
- Paradoja de Leontief: Algunos países exportan bienes intensivos en trabajo e importan bienes intensivos en capital, en contradicción con la teoría de Heckscher-Ohlin.
- Paradoja de Parrondo: Es posible jugar en dos juegos que ocasionan pérdidas alternativamente para acabar ganando.
- Paradoja de San Petersburgo: Cómo no merece la pena arriesgar mucho para ganar un premio infinito
- Paradoja del votante: Cuantas más personas participen en una elección por votación, menor será el beneficio de ir a votar, al ser cada votante menos decisivo.

10.1 Paradoja del cuervo

La paradoja del cuervo es una paradoja propuesta por el filósofo alemán Carl Hempel en la década de 1940 para ilustrar un problema donde la lógica inductiva desafía a la intuición. Esta paradoja se conoce también como paradoja de la negación o paradoja de Hempel. Cuando durante miles de años la gente ha observado hechos que se acomodan bien en el marco de una teoría como la ley de la gravedad, tendemos a creer que dicha teoría tiene una alta probabilidad de ser cierta y nuestra confianza en ella aumenta con cada nueva observación de acuerdo con ella. Este tipo de razonamiento puede sintetizarse en el principio de inducción:

- Si se observa un caso particular X consistente con la teoría T, entonces la probabilidad de que T sea cierta aumenta.

Hempel da un ejemplo del principio de inducción. Propone como teoría ”Todos los cuervos son negros”. Si ahora examinamos a un millón de cuervos, y observamos que todos son negros, nuestra creencia en la teoría ”todos los cuervos son negros” crecerá ligeramente con cada observación. En este caso, el principio de inducción parece razonable.

Ahora bien, la afirmación ”todos los cuervos son negros” es equivalente en lógica a la afirmación ”todas las cosas no-negras son no-cuervos”. Por lo tanto, observar una manzana roja proporciona evidencia empírica para sostener esta segunda afirmación. Una manzana roja es una cosa no-negra, y cuando la examinamos, vemos que es un no-cuervo. Así que, por el principio de inducción, el observar una manzana roja debería incrementar nuestra confianza en la creencia de que todos los cuervos son negros.

Hay filósofos que han ofrecido varias soluciones a este desafío a la intuición. El lógico estadounidense Nelson Goodman ha sugerido añadir restricciones a nuestro propio razonamiento, como no considerar nunca que un caso válido ”Todos los P son Q” sí valida también ”Ningún P es Q”.

Otros filósofos han cuestionado el ”principio de equivalencia”. A lo mejor, la manzana roja debe aumentar nuestra creencia en la teoría ”todas las cosas no-negras son no-cuervos” sin aumentar nuestra creencia en la teoría ”todos los cuervos son negros”. Esta sugerencia también ha sido cuestionada, sin embargo, con el argumento de que no puedes tener distinto nivel de creencia en dos afirmaciones si sabes que ambas son o ciertas o falsas al mismo tiempo. Goodman, y más tarde, Quine, usaron el término predicado proyectable para describir las expresiones, como cuervo y negro, que sí permiten el uso de generalizaciones inductivas. Los predicados no proyectables son aquellos como no-negro y no-cuervo, que aparentemente no lo permiten. (Ver también verjo, otro predicado no proyectable inventado por Goodman.) Quine sugirió que es una cuestión empírica cuáles, si alguno, de los predicados son proyectables, y observa que en un universo de infinitos objetos, el complemento de un predicado proyectable debe ser siempre no proyectable. Esto tendría la consecuencia de que, a pesar de que ”todos los cuervos son negros” y ”todas las cosas no-negras son no-cuervos” deben ser validados al mismo tiempo, ambos derivan su apoyo de cuervos negros, y no de no-cuervos no-negros.

Algunos filósofos han defendido que es nuestra intuición la que falla. Observar una manzana roja realmente incrementa la probabilidad de que todos los cuervos sean negros. Después de todo, si alguien te diese todas las cosas no-negras del universo, y pudieses ver que no hay ningún cuervo entre ellas, podrías concluir entonces que todos los cuervos son negros. El ejemplo solo desafía a la intuición porque el conjunto de cosas no-negras es con diferencia más grande que el conjunto de cuervos. Así, observar otra cosa no-negra que no sea un cuervo debería cambiar muy poco nuestra creencia en la teoría si lo comparamos con la observación de otro cuervo que sí sea negro.

10.2 Paradoja de Galileo

La paradoja de Galileo es una demostración de una de las sorprendentes propiedades de los conjuntos infinitos. El carácter paradójico se da por poner en entredicho el principio de que el todo es mayor que sus partes.

En su último trabajo científico, Dos nuevas ciencias, Galileo Galilei hizo dos afirmaciones aparentemente contradictorias acerca de los números enteros positivos. Primero, algunos números tienen la propiedad de ser un cuadrado perfecto (esto es, el cuadrado de un entero, desde ahora llamado simplemente cuadrado), mientras que otros no la tienen. Por ello, el conjunto de todos los números, incluyendo tanto a los cuadrados como a los no cuadrados, tiene que ser mayor que el conjunto de los cuadrados. Sin embargo, por cada cuadrado hay exactamente un número que es su raíz cuadrada, y por cada número hay exactamente un cuadrado. Por lo tanto, no puede haber más de un tipo que de otro. Este es uno de los primeros usos, aunque no el primero, de demostración a través de una función biyectiva.

En sus célebres “Diálogos” Galileo llegó a la conclusión de que los conceptos de menor, igual y mayor sólo se aplicaban a conjuntos finitos, y no tenían sentido aplicados a conjuntos infinitos. En el siglo XIX, Cantor, usando los mismos métodos, demostró que a pesar de que el resultado de Galileo era correcto si se aplicaba a los números enteros, o incluso a los racionales, la conclusión general no era cierta: algunos conjuntos infinitos son mayores que otros, en el sentido de que no se pueden relacionar en una correspondencia biunívoca. No obstante, es notable que Galileo haya demostrado que el número de puntos en un segmento es el mismo que en un segmento algo mayor, pero, por cierto, no llegó a la demostración de Cantor sobre la existencia de varios infinitos ni a su concepto de número transfinito. Galileo tenía en mente en ese momento otros asuntos: estaba anotando las contradicciones en las paradojas de Zenón para abrir camino a su teoría matemática del movimiento.

10.3 Paradoja El huevo o la gallina

¿El huevo o la gallina? es un dilema que proviene de la expresión "¿Qué fue primero, el huevo o la gallina?", ya que las gallinas ponen huevos y de ellos provienen los pollos. Esta expresión llena de ambigüedad condujo a los filósofos antiguos a cuestionar cómo se originó la vida y el Universo.

En el habla popular, referirse a la cuestión "el huevo o la gallina" hace hincapié a la inutilidad de preguntarse quién fue primero, pues esta sentencia es una falacia del tipo "círculo vicioso". Es en esta forma de ver el problema en que yace la naturaleza fundamental de la cuestión, ya que la respuesta literal es un tanto obvia. El dilema, entonces, procura elevarse a cuestiones metafísicas, con un planteamiento metafórico. Para entender mejor dicha representación metafórica, la pregunta se puede reformular de la siguiente manera: "¿Qué vino primero: X que no puede venir sin Y, o Y que no puede venir sin X?"

Para resolver dicha paradoja se propusieron soluciones en diferentes areas tales:

Filosofía

Diversos filósofos reflexionaron ante el dilema:

- Aristóteles (384 a. C.-322 a. C.) afirmó que lo primero en existir fue la gallina.
- Plutarco (46-126) hizo una comparación del dilema con la creación del mundo.
- Macrobio (395-423 AD), destacó la trascendencia de la cuestión.
- Stephen Hawking y Christopher Langan, concluyeron que fue primero el huevo que la gallina. Langan presentó un detallado estudio del problema en 2001.
- David Papineau, filósofo académico del King's College de Londres, opina que primero fue el huevo. Papineau comentó que un huevo es un huevo de gallina si en su interior lleva un pollo al igual que si un canguro pusiese un huevo, y de él saliese un avestruz, el huevo sería de avestruz y no de canguro.

Teología

Las teorías creacionistas dan dos respuestas opuestas a la pregunta, tanto fácilmente resolver el dilema, porque el creador es la creación de las gallinas adultas, tanto poner los huevos, y luego encontrar una solución para la eclosión, se evita el problema de la regresión infinita.

Quienes sólo admiten la interpretación de la Biblia de forma literal, desde el punto de vista del creacionismo, indican que las aves fueron creadas al igual que el resto del universo. El relato bíblico de la creación dijo que Dios creó a las aves, pero no menciona a los huevos.

Sin embargo, la evolución teísta establece que los pollos fueron creados por Dios a partir de huevos de gallinas. Así que Dios creó a las gallinas mediante evolución y pudo haberlas creado a partir de huevos. En escritos hindúes, los pájaros fueron creados por Dios mediante seres superhumanos según se dice en el Puranas y Dharmashastra.

Evolución
Como las especies cambian con el paso del tiempo durante el proceso evolutivo, la primera gallina moderna antecesora de las gallinas domésticas no puede ser clasificada como tal. El ADN sólo puede ser modificado antes del nacimiento, por lo cual la madre de la primera gallina (la cual no era una gallina tal como la conocemos), sufrió una mutación en la gestación mediante la cual el embrión que llevaba dentro cambió y se convirtió en la primera gallina, propiamente dicha. A partir de ese momento apareció la primera gallina, tal como lo entendemos ahora. Por tanto el huevo sería primero que la gallina.

Pero esta deducción sería sólo aplicable a la pregunta específica de qué fue antes ”¿La gallina o el huevo de gallina?” (entendiendo huevo de gallina como el primero en contener una gallina), ya que si nos referimos al huevo, en general, dado que en la pregunta no se específica que el huevo sea de gallina, el huevo sería primero, ya que lógicamente el animal antecesor de la primera gallina también era ovíparo, y el hecho de que animales anteriores como los dinosaurios también eran ovíparos.

Tiempo cíclico

En el Budismo existe la creencia sobre la ”Rueda del Tiempo” el cual ve al tiempo como cíclico y con repetición de eras, tal como otras culturas de Mesoamérica (aztecas, mayas) y los indios nativos de Norte América creen. Esta idea da una respuesta distinta a la pregunta de ”quién es primero” cuando está combinado con el concepto del Eterno Retorno el cual es bien conocido en Occidente gracias a Nietzsche. Este concepto asume que el tiempo es eternamente repetitivo, por consiguiente, no existe ”primero” en la eternidad. Nada ha sido creado, siempre ha existido. La respuesta es entonces ”ninguno es primero”. El tiempo cíclico no permite que exista un ”primero”.

10.4 Paradoja de Berry

La paradoja de Berry es la aparente contradicción que deriva de frases como ésta:

”*El menor entero positivo que no se puede definir con menos de quince palabras*”

El siguiente argumento parece probar que esta frase define un único entero positivo N. El número de frases que se pueden formar con menos de quince palabras es finito. Algunas de estas frases pueden describir un entero positivo específico, por ejemplo ”mil trescientos veintisiete”, ”el primer número primo mayor que cien millones” o ”dos elevado a trece”. Sin embargo, otras de las frases describen cosas que no son enteros, por ejemplo ”William Shakespeare” o ”Torre Eiffel”. En cualquier caso, el conjunto A de enteros que se pueden definir con menos de quince palabras es finito. Puesto que A es finito, no puede contener todos los enteros positivos, de modo que tiene que haber un número entero positivo N que sea el menor de todos los números enteros positivos que no están contenidos en A.

Entonces, como hemos dicho, el número N es el menor entero positivo que no se puede definir con menos de quince palabras. Pero entonces resulta que la frase ”El menor entero positivo que no se puede definir con menos de quince palabras” definirá a este número N. Y, sin embargo, dicha frase tiene sólo catorce palabras.

Esto es claramente paradójico, y parece sugerir que ”que no se puede definir con menos de quince palabras” no está bien definido. Sin embargo, es posible construir una expresión análoga con lenguaje matemático formal, como ha hecho Gregory Chaitin. A pesar de que la expresión análoga en lenguaje formal no lleva a una contradicción lógica, sí tiene ciertos resultados imposibles, incluyendo un teorema de incompletitud similar al teorema de la incompletitud de Gödel.

La paradoja de Berry fue propuesta por Bertrand Russell (Russell, 1906). Russell a su vez, la atribuyó a G. G. Berry, bibliotecario en jefe de la biblioteca Bodleiana de la Universidad de Oxford (cf. Russell and Whitehead 1910), que había sugerido la idea de estudiar la paradoja asociada a la expresión ”el primer número ordinal que no se puede definir”.

Resolución de la paradoja:

Se suele aceptar que la paradoja de Berry y otras paradojas similares (como la paradoja de Richard) provienen de la interpretación de conjuntos de expresiones que se autorreferencian. De acuerdo con (Russell y Whitehead, 1910) estas paradojas ”encarnan falacias de círculo vicioso”. Resolver una de estas paradojas significa localizar exactamente dónde comienza el error en el uso del lenguaje y restringirlo para evitarlas.

Algunas expresiones de este tipo no presentan la paradoja:

”*El menor entero positivo que no se puede definir con menos de dos palabras*”

que bajo cualquier uso razonable del idioma español describe al 31, ya que ”Treinta y uno” son tres palabras y cualquier definición indirecta de ese número (como ”el número de días en enero”, o incluso ”El menor entero positivo que no se puede definir con menos de dos palabras”) tienen necesariamente dos o más palabras.

10.5 Paradoja de los dos sobres

El problema de los dos sobres, también conocido como paradoja del cambio, es un acertijo o paradoja lógico, filosófico y de probabilidad y matemáticas recreativas, de especial interés en teoría de la decisión y la interpretación bayesiana de la probabilidad.

Una posible formulación del problema es la siguiente:

”Digamos que te dan dos sobres indistinguibles, cada uno de los cuales contiene una suma positiva de dinero. Un sobre contiene dos veces más que el otro. Puedes escoger un sobre y quedarte con cualquier cantidad que contenga. Escoges un sobre al azar, pero antes de abrirlo se te ofrece la posibilidad de tomar el otro sobre en su lugar.”

Es posible argumentar que será ventajoso para ti cambiar los sobres, demostrando que el valor esperado de un cambio de sobres es mayor que la cantidad que tiene en tu sobre. Esto conduce al absurdo de que es beneficioso cambiar sobres indefinidamente.

Las razones para cambiar de sobres:

1. Sea A la cantidad que contiene mi sobre.
2. La probabilidad de que A sea el sobre con menos dinero es 1/2, y la de que sea el sobre con la cantidad mayor es también 1/2.
3. El otro sobre puede contener o bien 2A o bien A/2.
4. Si A es la cantidad menor entonces el otro sobre contiene 2A.
5. Si A es la cantidad mayor entonces el otro sobre contiene A/2.
6. Por lo tanto el otro sobre contiene 2A con probabilidad 1/2 y A/2 con probabilidad 1/2.
7. Por lo tanto el valor esperado de dinero en el otro sobre es

$$\frac{1}{2}2A+\frac{1}{2}\frac{A}{2}=\frac{5}{4}A \tag{10.1}$$

8. Este valor es mayor que A, por lo tanto ganó cambiando los sobres.
9. Después del cambio, puedo llamar B al contenido de mi sobre y razonar exactamente igual que antes.
10. Concluyo que lo más racional es cambiar de sobres una y otra vez.
11. Para ser racional, deberé entonces cambiar de sobres indefinidamente.
12. Dado que parece más racional limitarse a abrir un sobre, tengo una contradicción.

Explicación de la paradoja:

A pesar de que en primera instancia parecería que el otro sobre es mejor siempre, se comete un error en la paradoja de los dos sobres al comparar A con un valor esperado ya que A es el valor efectivo del sobre y no el valor esperado de este, de hecho el valor esperado de los dos sobres es:

$$\frac{5}{4}A \tag{10.2}$$

11 Probabilidad

11.1 Teorema de Ramsey

En combinatoria el teorema de Ramsey establece que en cualquier esquema de color aplicado a un grafo de un grafo completo suficientemente grande, se hallarán subgrafos completos monocromáticos. Para dos colores, el teorema enuncia que para cualquier par de enteros positivos (r,s), existe al menos un entero positivo R(r,s) como aquel para cualquier grafo completo en R(r,s) vértices, cuyas aristas (o ramas) están coloreados de rojo o azul, existe un subgrafo completo de r vértices que es totalmente azul, o un subgrafo completo de s vértices que es totalmente rojo. R(r,s) representa un entero que depende conjuntamente de r y s. Se entiende que representa al entero más pequeño para el que el teorema aplica. A ese número se le llama número de Ramsey.

El teorema de Ramsey es fundacional en combinatoria. La primera versión de estos resultados fueron probados por F. P. Ramsey. Esto inició la teoría combinatoria, ahora llamada teoría de Ramsey, que busca regularidad en medio del desorden: condiciones generales para la existencia de subestructuras con propiedades regulares.

Una extensión de este teorema se aplica a cualquier número finito de colores, en lugar de solamente dos. Más precisamente, el teorema enuncia que por cualquier número dado de colores ”c”, y cualquier entero n1,...,nc, existe un número, R(n1, ..., nc), que si las aristas de un grafo completo de orden R(n1, ...,nc) se colorea con c colores diferentes, entonces para algún i entre 1 y c, debe contener un subgrafo completo de orden ni cuyas aristas son de color i. El caso especial de arriba donde c = 2 (y n1 = r y n2 = s).

Otra generalización se obtiene al considerar grafos que no sean completos. Son conocidos todos los valores de R(G1,G2) si G1 y G2 tienen a lo más 5 vértices salvo cuando G1 ó G2 es el grafo completo de 5 vértices y el otro es o bien el grafo completo de 5 vértices o bien el grafo completo de 5 vértices menos una arista.

11.2 Probabilidades Subjetivas

Existen muchos eventos de interés cuyas probabilidades de ocurrencia no las podemos calcular de acuerdo con los métodos axiomático, clásico y de frecuencia relativa (empírica). sino que se basan en el ”grado de creencia” acerca de que tenga o no lugar un determinado hecho como, por ejemplo:

- Exista vida en algún planeta distante,

- En los póximos diez años se descubra algún remedio contra el cáncer.
- Determinada persona se vaya a destacar en la universidad,
- Ó una persona se enferme,
- Una determinada máquina se dañe.

Sin embargo, poca gente se muestra renuente a concederles probabilidades & los eventos anteriores. Inclusive, con mucha frecuencia oímos decir que hay un 20% de posibilidades de que llueva mañana, que el Junior gane, etc. Aquella probabilidad que nos permite asignarles probabilidades a eventos tales como éstos se denomina probabilidad subjetiva.

Teorema 11.1 (PROBABILIDAD SUBJETIVA)

La PROBABILIDAD SUBJETIVA o PERSONAL se puede definir como la probabilidad que expresa un grado de creencia individual sobre la posibilidad de que un evento ocurra. Al método de asignar estas probabilldades se le conoce como MÉTODO SUBJETIVO.

La asignación de la probabilidad subjetiva a un evento dado no depende del tratamiento matemático ni de la noción de experimentos repetibles.

La magnitud de la probabilidad que una persona asigna subjet¡vamente a un evento depende del grado de crédito que le dé a la ocurrencia del evento. Esa es la razón por la que es posible asignarles probabilidades & eventos que sólo se presentan una vez, como por ejemplo, el evento de ganar una determinada competencia atlética. A diferencia del método de probabilidad de frecuencia relativa, la probabilidad subjetiva no depende de la posibilidad de repetición de un experimento.

Ejemplo 11.1

1. Si afirmamos que la probabilidad de que salga cara al lanzar una moneda es 1/2, lo que tenemos en mente es que la moneda no parece estar trucada y que resultará igualmente probable que salga cara o cruz. Al enunciar esta probabilidad sub jetiva, no estamos pensando necesariamente en términos de la experimentación repetida, sino que estamos interesado por un único lanzamiento de la moneda. Nuestra evaluación de la probabilidad subjetiva implica que consideraríamos justa una apuesta que consistiese en pagar 5.000 pesos si saliera cruz y recibir 5.000 pesos si saliera cara. Si fueramos a recibir más de 5.000 pesos si del lanzamiento resultase una cara, consideraría?os favorable la apuesta.
2. De modo similar, si creemos que la probabilidad de que un caballo gane una determinada carrera es 0,4, estamos dando nuestra opinión personal de que existe una posibilidad de 40 entre 100 de que gane. Dada esta creencia, considerariamos justa una apuesta en la que perdiésemos dos dólares si el caballo no ganase y tres dólares en caso contrario.

Debemos insistir en que las probabilidades subjetivas son personales; no se requiere que diferentes individuos consideren que el mismo evento debe tener lugar con las mismas probabilidades, como se explica a través de las situaciones del siguiente ejemplo:

Ejemplo 11.2

1. En el ejemplo del lanzamiento de una moneda, la mayoría de la gente llegaria a la conclusión de que la probabilidad apropiada para el resultado ”cara” es 1/2. Sin embargo, un individuo con nnís infonnación sobre la moneda en em3stión podría creer otra cosa.

2. En el ejemplo de las carreras de caballos, es probable que dos apostadores cuenten con diferentes probabilidades subjetivas. Por ejemplo, pueden no tener la misma información, e incluso aunque la tuvieran, podrían interpretarla de distinta forma.

3. Es obvio que los inversionistas individuales no cuentan con las mismas opiniones sobre el probable futuro comportamiento de la bolsa. Sus probabilidades subjetivas deben ser vistas como dependientes del conocimiento que tienen y su manera de interpretarlo.

Ya hemos explicado que, en el caso de apuestas, como carreras de caballos y pronósticos deportivos, a menudo se determina la probabilidad de ocurrencia de un evento usando probabilidad subjetiva y se establece comúnmente en términos de oportunidades.

11.3 Modelo Ising

El modelo de Ising es un modelo físico propuesto para estudiar el comportamiento de materiales ferromagnéticos. Se trata de un modelo paradigmático de la Mecánica Estadística, en parte porque fue uno de los primeros en aparecer, pero sobre todo porque es de los pocos modelos útiles (no sólo pedagógicamente) que tiene solución analítica exacta (esto es, sin cálculos aproximados). Esto lo hace muy útil para ensayar nuevos tipos de aproximaciones y luego comparar con el resultado real.

El modelo de Ising fue inventado por el físico Wilhelm Lenz (1920), que lo concibió como un problema para su alumno Ernst Ising para demostrar que el sistema presentaba una transición de fase. Ising (1925) demostró que en una dimensión no existía tal transición de fase, resolviéndolo en su tesis de 1924, aunque le provocó una profunda desmoralización e hizo que renunciara a la física estadística. El modelo bidimensional de Ising de retícula cuadrada es mucho más difícil, y solamente se le dio una descripción analítica mucho más tarde, por Lars Onsager (1944), que demostró que la física estadística era capaz de describir transiciones de fase (pues como se verá, éste modelo presenta una) lo que terminó de consolidar definitivamente la mecánica estadística. Por lo general, se resuelve mediante un método de transferencia de matriz, aunque existen diferentes enfoques, más relacionados con la teoría cuántica de campos.

11.4 Teorema de los cuatro colores

En teoría de grafos, el teorema de los cuatro colores (o teorema de la minimalidad cromática) es un teorema sobre la coloración de grafos que establece lo siguiente:

Teorema 11.2 (Cuatro colores)

Dado cualquier mapa geográfico con regiones continuas, este puede ser coloreado con cuatro colores diferentes, de forma que no queden regiones adyacentes con el mismo color.

Asumiendo que las regiones adyacentes comparten no solo un punto, sino todo un segmento de borde (frontera) en común.

Tres colores son suficientes para mapas simples, pero en algunos casos es necesario un cuarto color adicional, esto es, cuando una región a colorear queda encerrada por un número impar de regiones que se tocan formando un ciclo. El teorema de los cinco colores, cuya demostración es corta y elemental, establece que cinco colores son suficientes para colorear un mapa y fue probado en el siglo XIX por Heawood. Una serie de pruebas falsas y falsos contraejemplos han aparecido desde el primer enunciado del teorema de los cuatro colores en 1852.

El problema del mapa de cuatro colores fue planteado, por primera vez, por el estudiante Francis Guthrie en 1852, lo que fue comunicado a Augustus de Morgan. La conjetura se hizo famosa con la declaración de Arthur Cayley, en 1878, en el sentido de que la había abordado. Fue resuelto, a mediados de 1970, por Kenneth Appel y Wolfgang Haken.

11.4.1 Formulación precisa del teorema

En primer lugar, todas las esquinas y puntos en común que pertenecen a tres o más países, deben ser ignoradas. Sin esta restricción, los mapas extraños (utilizando las regiones del área finita pero perímetro infinito) pueden requerir más de cuatro colores.

En segundo lugar, para el propósito del teorema cada "país" tiene que ser una región simplemente conexa o continua. En el mundo real, esto no es cierto (por ejemplo, Alaska como parte de los Estados Unidos, Nakhchivan como parte de Azerbaiyán, y Kaliningrado como parte de Rusia no son regiones continuas). Debido a que el territorio de un país en particular debe ser del mismo color, si se permitiesen "países" no continuos, cuatro colores podrían no ser suficientes. Por ejemplo, considérese un mapa simplificado:

En este mapa, las dos regiones A pertenecen a un mismo país, y por lo tanto, deben ser del mismo color. En consecuencia, este mapa requiere cinco colores, puesto que las dos regiones A son contiguas con las otras cuatro regiones, y cada una de estas regiones son contiguas entre sí. Si hay tres regiones A, entonces se necesitan seis o más colores; se pueden construir mapas que requieren un número arbitrariamente elevado de colores. Un escenario similar también se puede dar si el color azul se reserva para el agua.

Una versión más simple del teorema utiliza la teoría de grafos. El conjunto de las regiones de un mapa se puede representar de manera más abstracta como un grafo simple no dirigido asociando un vértice para cada región y una arista para cada par de regiones que comparten un segmento de borde. Esta representación del mapa con vértices y aristas es un grafo dual y el problema de colorear países se cambia por la coloración del grafo. Este grafo es plano, o sea, que se puede dibujar en el plano sin cruce de aristas mediante la colocación de cada vértice en un lugar elegido arbitrariamente dentro de la región a la que corresponde. Con la terminología de la teoría de grafos, el teorema de cuatro colores establece que:

Teorema 11.3 (Cuatro colores)

Si G, es un grafo plano, entonces $\chi(G) \leq 4$

Es decir, los vértices de cada grafo plano pueden ser coloreados con un máximo de cuatro colores de modo que no existan dos vértices adyacentes con el mismo color. $\chi(G)$ corresponde al número cromático.

11.5 Problema de los puentes de Königsberg

El problema de los puentes de Königsberg, también llamado más específicamente problema de los siete puentes de Königsberg, es un célebre problema matemático, resuelto por Leonhard Euler en 1736 y cuya resolución dio origen a la teoría de grafos. Su nombre se debe a Königsberg, la ciudad de Prusia Oriental y luego de Alemania que desde 1945 se convertiría en la ciudad rusa de Kaliningrado.

Esta ciudad es atravesada por el río Pregel, en ruso ”Pregolya”, el cual se bifurca para rodear con sus brazos a la isla Kneiphof, dividiendo el terreno en cuatro regiones distintas, las que entonces estaban unidas mediante siete puentes llamados Puente del herrero, Puente conector, Puente verde, Puente del mercado, Puente de madera, Puente alto y Puente de la miel. El problema fue formulado en el siglo XVIII y consistía en encontrar un recorrido para cruzar a pie toda la ciudad, pasando sólo una vez por cada uno de los puentes, y regresando al mismo punto de inicio.

11.5.1 Contextualización del problema

Leonhard Euler llegó a Prusia en 1741, a la edad de 34 años, donde vivió hasta 1766 para luego regresar a San Petersburgo. Durante esos años trabajó en la Academia Prusiana de las Ciencias, donde desarrolló una prolífica carrera como investigador. Euler fue contemporáneo de varios otros famosos matemáticos y pensadores procedentes de aquella ciudad, tales como Immanuel Kant, Johann Georg Hamann y Christian Goldbach, por lo que Königsberg fue en ese tiempo un importante epicentro científico.

Es en este ambiente y por estos años en que surge la formulación del problema de los puentes de Königsberg, propagándose a modo de juego y de trivia matemática entre los intelectuales de la época.

11.5.2 Análisis y solución del problema

El problema, formulado originalmente de manera informal, consistía en responder a la siguiente pregunta: Dado el mapa de Königsberg, con el río Pregel dividiendo el plano en cuatro regiones distintas, que están unidas a través de los siete puentes, ¿es posible dar un paseo comenzando desde cualquiera de estas regiones, pasando por todos los puentes, recorriendo sólo una vez cada uno, y regresando al mismo punto de partida?

La respuesta es negativa, es decir, no existe una ruta con estas características. El problema puede resolverse aplicando un método de fuerza bruta, lo que implica probar todos los posibles recorridos existentes. Sin embargo, Euler en 1736 en su publicación ”Solutio problematis ad geometriam situs pertinentis” demuestra una solución generalizada del problema, que puede aplicarse a cualquier territorio en que ciertos accesos estén restringidos a ciertas conexiones, tales como los puentes de Königsberg.

Para dicha demostración, Euler recurre a una abstracción del mapa, enfocándose exclusivamente en las regiones terrestres y las conexiones entre ellas. Cada puente lo representó mediante una línea que unía a dos puntos, cada uno de los cuales representaba una región diferente. Así el problema se reduce a decidir si existe o no un camino que comience por uno de los puntos azules, transite por todas las líneas una única vez, y regrese al mismo punto de partida.

11.6 Articulos

11.6.1 How can Bayes theorem assign a probability to the existence of God?

How can Bayes theorem assign a probability to the existence of God?

A patient goes to see a doctor. The doctor performs a test with 99 percent reliability that is, 99 percent of people who are sick test positive and 99 percent of the healthy people test negative. The doctor knows that only 1 percent of the people in the country are sick. Now the question is: if the patient tests positive, what are the chances the patient is sick?

The intuitive answer is 99 percent, but the correct answer is 50 percent, and Bayes's theorem gives us the relationship between what we know and what we want to know in this problem. What we are given what we know is $p(+ \mid s)$, which a mathematician would read as ”the probability of testing positive given that you are sick”; what we want to know is $p(s|+)$, or ”the probability of being sick given that you tested positive.” The theorem itself reads $p(s|+) = p(+ \mid s)p(s)/p(+)$, although what Reverend Bayes, who lived from 1702 to 1761, actually said was something simpler. Bayes stated the defining relationship expressing the probability you test positive AND are sick as the product of the likelihood that you test positive GIVEN that you are sick and the ”prior” probability that you are sick (that is, the probability the patient is sick, prior to specifying a particular patient and administering the test).

Rather than relying on Bayes's math to help us with this, let us consider another illustration. Imagine that the above story takes place in a small country, with 10,000 people. From the prior $p(s) = 0.01$, we know that 1 percent, or 100 people, are sick, and 9,900 are healthy. If we administer the test to everyone, the most probable result is that 99 of the 100 sick people test positive. Since the test has a 1 percent error rate, however, it is also probable that 99 of the healthy people test positive. Now if the doctor sends everyone who tests positive to the national hospital, there will be an equal number of healthy and sick patients. If you meet one, even though you are armed with the information that the patient tested positive, there is only a 50 percent chance this person is sick.

Now imagine the doctor moves to another country, performing the same test, with the same likelihood ($p(+ \mid s)$) and, for that matter, the same success rate for healthy people, which we might call $p(- \mid h)$, ”the probability of scoring negative given that one is healthy.” In this country, however, we suppose that only one in every 200 people is sick. If a new patient tests positive, it is actually more probable that the patient is healthy than sick. The doctor needs to update the prior.

The importance of accurate data in quantitative modeling is central to the subject raised in the question: using Bayes's theorem to calculate the probability of the existence of God. Scientific discussion of religion is a popular topic at present, with three new books arguing against theism and one, University of Oxford professor Richard Dawkins's book The God Delusion, arguing specifically against the use of Bayes's theorem for assigning a probability to God's existence. (A Google news search for ”Dawkins” turns up 1,890 news items at the time of this writing.) Arguments employing Bayes's theorem calculate the probability of God given our experiences in the world (the existence of evil, religious experiences, etc.) and assign numbers to the likelihood of these facts given existence or nonexistence of God, as well as to the prior belief of God's existence the probability we would assign to the existence of God if we had no data from our experiences. Dawkins's argument is not with the veracity of Bayes's theorem itself, whose proof is direct and unassailable, but rather with the lack of data to put into this formula by those employing it to argue for the existence of God. The equation is perfectly accurate, but the numbers inserted are, to quote Dawkins, ”not measured quantities but & personal judgments, turned into numbers for the sake of the exercise.”

Note that although this is receiving much attention now, quantifying one's judgments for use in Bayesian calculations of the existence of God is not new. Richard Swinburne, for example, a philosopher of science turned philosopher of religion (and Dawkins's colleague at Oxford), estimated the probability of God's existence to be more than 50 percent in 1979 and, in 2003, calculated the probability of the resurrection [presumably of both Jesus and his followers] to be "something like 97 percent." (Swinburne assigns God a prior probability of 50 percent since there are only two choices: God exists or does not. Dawkins, on the other hand, believes "there's an infinite number of things that some people at one time or another have believed in ... God, Flying Spaghetti Monster, fairies, or whatever," which would correspondingly lower each outcome's prior probability.) In reviewing the history of Bayes's theorem and theology, one might wonder what Reverend Bayes had to say about this, and whether Bayes introduced his theorem as part of a similar argument for the existence of God. But the good reverend said nothing on the subject, and his theorem was introduced posthumously as part of his solution to predicting the probability of an event given specific conditions. In fact, while there is plenty of material on lotteries and hyperbolic logarithms, there is no mention of God in Bayes's "Essay towards Solving a Problem in the Doctrine of Chances," presented after his death to the Royal Society of London in 1763.

One primary scientific value of Bayes's theorem today is in comparing models to data and selecting the best model given those data. For example, imagine two mathematical models, A and B, from which one can calculate the likelihood of any data given the model ($p(D \mid A)$ and $p(D \mid B)$). For example, model A might be one in which spacetime is 11-dimensional, and model B one in which spacetime is 26-dimensional. Once I have performed a quantitative measurement and obtained some data D, one needs to calculate the relative probability of the two models: $p(A \mid D)/p(B \mid D)$. Note that just as in relating $p(+ \mid s)$ to $p(s \mid +)$, I can equate this relative probability to $p(D \mid A)p(A)/p(D \mid B)p(B)$. To some, this relationship is the source of deep joy; to others, maddening frustration.

The source of this frustration is the unknown priors, $p(A)$ and $p(B)$. What does it mean to have prior belief about the probability of a mathematical model? Answering this question opens up a bitter internecine can of worms between "the Bayesians" and "the frequentists," a mathematical gang war which is better not entered into here. To oversimplify, "Bayesian probability" is an interpretation of probability as the degree of belief in a hypothesis; "frequentist probability is an interpretation of probability as the frequency of a particular outcome in a large number of experimental trials. In the case of our original doctor, estimating the prior can mean the difference between more-than-likely and less-than-likely prognosis. In the case of model selection, particularly when two disputants have strong prior beliefs that are diametrically opposed (belief versus nonbelief), Bayes's theorem can lead to more conflict than clarity.

More generally, Bayes's theorem is used in any calculation in which a "marginal" probability is calculated (e.g., $p(+)$, the probability of testing positive in the example) from likelihoods (e.g., $p(+ \mid s)$ and $p(+ \mid h)$, the probability of testing positive given being sick or healthy) and prior probabilities ($p(s)$ and $p(h)$): $p(+) = p(+ \mid s)p(s) + p(+ \mid h)p(h)$. Such a calculation is so general that almost every application of probability or statistics must invoke Bayes's theorem at some point. In that sense Bayes's theorem is at the heart of everything from genetics to Google, from health insurance to hedge funds. It is a central relationship for thinking concretely about uncertainty, and given quantitative data, which is sadly not always a given for using mathematics as a tool for thinking clearly about the world.

11.7 Skewness - Asimetría estadística

Las medidas de asimetría son indicadores que permiten establecer el grado de simetría (o asimetría) que presenta una distribución de probabilidad de una variable aleatoria sin tener que hacer su representación gráfica. Como eje de simetría consideramos una recta paralela al eje de ordenadas que pasa por la media de la distribución. Si una distribución es simétrica, existe el mismo número de valores a la derecha que a la izquierda de la media, por tanto, el mismo número de desviaciones con signo positivo que con signo negativo. Decimos que hay asimetría positiva (o a la derecha) si la ”cola” a la derecha de la media es más larga que la de la izquierda, es decir, si hay valores más separados de la media a la derecha. Diremos que hay asimetría negativa (o a la izquierda) si la ”cola” a la izquierda de la media es más larga que la de la derecha, es decir, si hay valores más separados de la media a la izquierda.

11.7.1 Medidas de Asimetría

Coeficiente de asimetría de Fisher

En teoría de la probabilidad y estadística, la medida de asimetría más utilizada parte del uso del tercer momento estándar. La razón de esto es que nos interesa mantener el signo de las desviaciones con respecto a la media, para obtener si son mayores las que ocurren a la derecha de la media que las de la izquierda. Sin embargo, no es buena idea tomar el momento estándar con respecto a la media de orden 1. Debido a que una simple suma de todas las desviaciones siempre es cero. En efecto, si por ejemplo, los datos están agrupados en k clases, se tiene que:

$$\sum_{i=1}^{k} f_i(x_i-\mu)=\sum_{i=1}^{k} f_i x_i-\mu\sum_{i=1}^{k} f_i=\mu-\mu=0 \tag{11.1}$$

en donde x_i representa la marca de la clase i-ésima y f_i denota la frecuencia relativa de dicha clase. Por ello, lo más sencillo es tomar las desviaciones al cubo.

El coeficiente de asimetría de Fisher, representado por γ_1, se define como:

$$\gamma_1=\frac{\mu_3}{\sigma^3} \tag{11.2}$$

donde μ_3 es el tercer momento en torno a la media y σ es la desviación estándar.

- Si $\gamma_1>0$, la distribución es asimétrica positiva o a la izquierda.
- Si $\gamma_1<0$, la distribución es asimétrica negativa o a la derecha.
- Si la distribución es simétrica, entonces sabemos que $\gamma_1=0$. El recíproco no es cierto: es un error común asegurar que si $\gamma_1=0$ entonces la distribución es simétrica (lo cual es falso).

Coeficiente de asimetría de Pearson

Sólo se puede utilizar en distribuciones uniformes, unimodales y moderadamente asimétricas. Se basa en que en distribuciones simétricas la media de la distribución es igual a la moda.

$$A_p=\frac{\mu-moda}{\sigma} \tag{11.3}$$

donde μ es el momento central de orden 1, que corresponde a la media aritmética de la variable X.

Si la distribución es simétrica, $\mu=moda$ y $A_p=0$. Si la distribución es asimétrica positiva la media se sitúa por encima de la moda y, por tanto, $A_p>0$.

Coeficiente de asimetría de Bowley

Está basado en la posición de los cuartiles y la mediana, y utiliza la siguiente expresión:

$$A_B = \frac{Q_{3/4} + Q_{1/4} - 2Me}{Q_{3/4} - Q_{1/4}} \tag{11.4}$$

En una distribución simétrica el tercer cuartil estará a la misma distancia de la mediana que el primer cuartil. Por tanto $A_B = 0$.

Si la distribución es positiva o a la derecha, $A_B > 0$.

11.8 Caminata al azar

El problema de la caminata al azar, nos dice que una persona situada en un poste puede dar pasos a la izquierda y a la derecha, estos pasos pueden ser combinados, es decir, que no todos tienen que ser en un mismo sentido. Se quiere determinar la posición final del home despues de haber dado cierto número de pasos:

$$n_1 = \text{número de pasos a la derecha} \tag{11.5}$$

$$n_2 = \text{número de pasos a la izquierda} \tag{11.6}$$

Se requiere calcular la probabilidad $PN(m)$ de encontrar a la partícula (despúes de N pasos) en la posición.

Definamos la distancia que se recorre la persona, es decir, número de pasos por la longitud del paso, esto se puede representar como:

$$x = ml \tag{11.7}$$

Resulta claro que: de la ecuación anterior conocemos que si multiplicamos el número de pasos por su longitud nos dara la posición final de tal suerte que lo podemos acotar de la siguiente manera:

$$-N \leq m \leq N \tag{11.8}$$

Tambien, podemos expresar la ecuación anterior en terminos de los pasos a la derecha e izquierda como la suma de ambos:

$$N = n_1 + n_2 \tag{11.9}$$

Por construcción se tiene que:

$$m = n_1 - n_2 \tag{11.10}$$

Lo cual permite escribir:

$$\begin{aligned} m &= n_1 - n_2 && (11.11) \\ &= n_1 - (Nn_1) && (11.12) \\ &= 2n_1 - N && (11.13) \end{aligned}$$

Empleando técnicas de conteo se puede obtener la probabilidad de obtener cualquier secuencia de n_1 pasos a la derecha y n_2 pasos a la izquierda al multiplicar sus respectivas probabilidades, i.e.:

$$PPPP\dots PPP\dots QQQQ\dots QQQ = P^{n_1}Q^{n_2} \tag{11.14}$$

$$\tag{11.15}$$

Por otro lado, el número de formas en que se pueden dar n pasos en total considerando n_1 pasos a la derecha y n_2 pasos a la izquierda es:

$$P\binom{n}{r} = \frac{N!}{n_1!n_2!} \tag{11.16}$$

11.9 Curva ROC

En la Teoría de detección de señales una curva ROC (acrónimo de Receiver Operating Characteristic, o Característica Operativa del Receptor) es una representación gráfica de la sensibilidad frente a la especificidad para un sistema clasificador binario según se varía el umbral de discriminación. Otra interpretación de este gráfico es la representación de la razón o ratio de verdaderos positivos (VPR = Razón de Verdaderos Positivos) frente a la razón o ratio de falsos positivos (FPR = Razón de Falsos Positivos) también según se varía el umbral de discriminación (valor a partir del cual decidimos que un caso es un positivo). ROC también puede significar Relative Operating Characteristic (Característica Operativa Relativa) porque es una comparación de dos características operativas (VPR y FPR) según cambiamos el umbral para la decisión. No se suele utilizar ROC aislado, debemos decir ”curva ROC” o ”análisis ROC”. Sobre la historia del acrónimo ROC consultar Swets (1996).

El análisis de la curva ROC, o simplemente análisis ROC, proporciona herramientas para seleccionar los modelos posiblemente óptimos y descartar modelos subóptimos independientemente. El análisis ROC se relaciona de forma directa y natural con el análisis de coste/beneficio en toma de decisiones diagnósticas.

La curva ROC se desarrolló por ingenieros eléctricos para medir la eficacia en la detección de objetos enemigos en campos de batalla mediante pantallas de radar, a partir de lo cual se desarrolla la Teoría de Detección de Señales (TDS). El análisis ROC se aplicó posteriormente en medicina, radiología, psicología y otras áreas durante varias décadas. Sólo recientemente ha encontrado aplicación en áreas como aprendizaje automático (o machine learning en inglés), y minería de datos (data mining en inglés).

Consideremos un problema de predicción de clases binario, en la que los resultados se etiquetan positivos (p) o negativos (n). Hay cuatro posibles resultados a partir de un clasificador binario como el propuesto. Si el resultado de una exploración es p y el valor dado es también p, entonces se conoce como un Verdadero Positivo (VP); sin embargo si el valor real es n entonces se conoce como un Falso Positivo (FP). De igual modo, tenemos un Verdadero Negativo (VN) cuando tanto la exploración como el valor dado son n, y un Falso Negativo (FN) cuando el resultado de la predicción es n pero el valor real es p. Un ejemplo aproximado de un problema real es el siguiente: consideremos una prueba diagnóstica que persiga determinar si una persona tiene una cierta enfermedad. Un falso positivo en este caso ocurre cuando la prueba predice que el resultado es positivo, cuando la persona no tiene realmente la enfermedad. Un falso negativo, por el contrario, ocurre cuando el resultado de la prueba es negativo, sugiriendo que no tiene la enfermedad cuando realmente sí la tiene.

Definamos un experimento a partir de P instancias positivas y N negativas. Los cuatro posibles resultados se pueden formular en una Tabla de contingencia (o Matriz de confusión) 2x2 como sigue:

Valor en la realidad

	p	n	
p'	TP	FP	P'
n'	FN	FT	N'
	P	N	

11.10 Ejemplo - Probabilidad Cáncer

Ejemplo 11.3

Del ejemplo visto en clase que menciona: Se realizó un estudio de cáncer con una fiabilidad del 98%, es decir, que sí la persona está enferma de cáncer la prueba será fiable en 98%, por lo que tendrá un margen de error del 2%, y aplica de la misma manera sí la persona no tiene cancer:

Considérense los siguientes cuatros resultados en terminos de porcentaje dada la descipción del problema: Entonces tenemos la siguiente matriz de confusión.

Estudios de Cancer

	p	n	
p'	TP=98%	FP=2%	P'
n'	FN=2%	FT=98%	N'
	P	N	

11.11 Errores de tipo I y de tipo II

En un estudio de investigación, el error de tipo I también denominado error de tipo alfa (?)1 o falso positivo, es el error que se comete cuando el investigador no acepta la hipótesis nula H_0 siendo esta verdadera en la población. Es equivalente a encontrar un resultado falso positivo, porque el investigador llega a la conclusión de que existe una diferencia entre las hipótesis cuando en realidad no existe. Se relaciona con el nivel de significancia estadística.

La hipótesis de la que se parte H_0 aquí es el supuesto de que la situación experimental presentaría un "estado normal". Si no se advierte este "estado normal", aunque en realidad existe, se trata de un error estadístico tipo I. Algunos ejemplos para el error tipo I serían:

- Se considera que el paciente está enfermo, a pesar de que en realidad está sano; hipótesis nula: El paciente está sano.
- Se declara culpable al acusado, a pesar de que en realidad es inocente; hipótesis nula: El acusado es inocente.
- No se permite el ingreso de una persona, a pesar de que tiene derecho a ingresar; hipótesis nula: La persona tiene derecho a ingresar.

En un estudio de investigación, el error de tipo II, también llamado error de tipo beta (β) (β es la probabilidad de que exista este error) o falso negativo, se comete cuando el investigador no rechaza la hipótesis nula siendo esta falsa en la población. Es equivalente a la probabilidad de un resultado falso negativo, ya que el investigador llega a la conclusión de que ha sido incapaz de encontrar una diferencia que existe en la realidad.

Se acepta en un estudio que el valor del error beta esté entre el 5 y el 20%.

Contrariamente al error tipo I, en la mayoría de los casos no es posible calcular la probabilidad del error tipo II. La razón de esto se encuentra en la manera en que se formulan las hipótesis en una prueba estadística. Mientras que la hipótesis nula representa siempre una afirmación enérgica (como por ejemplo H_0 : "Promedio $\mu = 0$") la hipótesis alternativa, debido a que engloba todas las otras posibilidades, es generalmente de naturaleza global (por ejemplo H_1 : "Promedio $\mu \neq 0$"). El gráfico de la derecha ilustra la probabilidad del error tipo II (rojo) en dependencia del promedio μ desconocido.

El poder o potencia del estudio representa la probabilidad de observar en la muestra una determinada diferencia o efecto, si existe en la población. Es el complementario del error de tipo II (1- β).

Representación de los valores posibles de la probabilidad de un error tipo II (rojo) en el ejemplo de un test de significancia estadística para el parámetro μ. El error tipo II depende del parámetro μ . Mientras más cerca se encuentre este del valor supuesto bajo la hipótesis nula, mayor es la probabilidad de ocurrencia del error tipo II. Debido a que el verdadero valor de μ es desconocido al hacer la presunción de la hipótesis alternativa, la probabilidad del error tipo II, en contraste con el error tipo I (azul), no se puede calcular.

11.12 Distribución binomial

En estadística, la distribución binomial es una distribución de probabilidad discreta que cuenta el número de éxitos en una secuencia de n ensayos de Bernoulli independientes entre sí, con una probabilidad fija p de ocurrencia del éxito entre los ensayos. Un experimento de Bernoulli se caracteriza por ser dicotómico, esto es, sólo son posibles dos resultados. A uno de estos se denomina éxito y tiene una probabilidad de ocurrencia p y al otro, fracaso, con una probabilidad q = 1 - p. En la distribución binomial el anterior experimento se repite n veces, de forma independiente, y se trata de calcular la probabilidad de un determinado número de éxitos. Para n = 1, la binomial se convierte, de hecho, en una distribución de Bernoulli.

Para representar que una variable aleatoria X sigue una distribución binomial de parámetros n y p, se escribe:

$$X \sim B(n, p) \tag{11.17}$$

Las siguientes situaciones son ejemplos de experimentos que pueden modelizarse por esta distribución:

- Se lanza un dado diez veces y se cuenta el número X de tres obtenidos:
  entonces $X \sim B(10, 1/6)$
- Se lanza una moneda dos veces y se cuenta el número X de caras obtenidas:
  entonces $X \sim B(2, 1/2)$

11.13 Distribución acumulada (CDF)

En la teoría de la probabilidad y en estadística, una función de distribución acumulada (CDF) describe la probabilidad de que una variable aleatoria real X sujeta a cierta ley de distribución de probabilidad se sitúe en la zona de valores menores o iguales a x.

Intuitivamente, asumiendo la función f como la ley de distribución de probabilidad, la CDF sería la función con la recta real como dominio, con imagen del área hasta aquí de la función f, siendo aquí el valor x para la variable aleatoria real X.

La CDF asocia a cada valor x, la probabilidad del evento: “la variable X toma valores menores o iguales a x”.

Las Funciones de Distribución Acumulativa se emplean también para especificar la distribución de variables aleatorias multivariantes.

Es convención usar una F mayúscula para una CDF, en contraste con la f minúscula usada para una Función de Densidad de Probabilidad y/o para una Función de Probabilidad.

La función distribución puede obtenerse a partir de la función de probabilidad respectiva. La CDF en el caso de una variable aleatoria X discreta, puede establecerse como:

$$F(x) = \sum_{x_i \leq x} f(x_i) \tag{11.18}$$

Para una variable aleatoria X continua, CDF surge como:

$$F(x) = \int_{-\infty}^{x} f(t)\,dt \tag{11.19}$$

Debe observarse que una definición del tipo ”menor o igual”, ≤ podría sustituirse por estrictamente ”menor” ’<’. Esto produciría una función diferente, pero cualquiera de las funciones F puede deducirse a partir de la otra f.

También se podría cambiar por una determinada por mayor (>) en lugar de ”menor” ’<’ y deducir las propiedades de esta nueva función.

Sólo es preciso ajustar las formulaciones y definiciones a lo pretendido en cada caso.
En países de lengua inglesa, una convención es usar una desigualdad de este tipo $\leq$ en lugar de una desigualdad estricta (<).

12 Multiplicadores de Lagrange

En los problemas de optimización, el método de los multiplicadores de Lagrange, llamados así en honor a Joseph Louis Lagrange, es un procedimiento para encontrar los máximos y mínimos de funciones de múltiples variables sujetas a restricciones. Este método reduce el problema restringido con n variables a uno sin restricciones de n + k variables, donde k es igual al número de restricciones, y cuyas ecuaciones pueden ser resueltas más fácilmente. Estas nuevas variables escalares desconocidas, una para cada restricción, son llamadas multiplicadores de Lagrange. El método dice que los puntos donde la función tiene un extremo condicionado con k restricciones, están entre los puntos estacionarios de una nueva función sin restricciones construida como una combinación lineal de la función y las funciones implicadas en las restricciones, cuyos coeficientes son los multiplicadores.

La demostración usa derivadas parciales y la regla de la cadena para funciones de varias variables. Se trata de extraer una función implícita de las restricciones, y encontrar las condiciones para que las derivadas parciales con respecto a las variables independientes de la función sean iguales a cero.

12.1 Introducción

Consideremos un caso bidimensional. Supongamos que tenemos la función, f (x, y), y queremos maximizarla, estando sujeta a la condición:

$$g(x,y) = c \tag{12.1}$$

donde c es una constante. Podemos visualizar las curvas de nivel de f dadas por:

$$f(x,y) = d_n \tag{12.2}$$

para varios valores de dn, y el contorno de g dado por g(x, y) = c. Supongamos que hablamos de la curva de nivel donde g = c. Entonces, en general, las curvas de nivel de f y g serán distintas, y la curva g = c por lo general intersectará y cruzará muchos contornos de f. En general, moviéndose a través de la línea g=c podemos incrementar o disminuir el valor de f. Sólo cuando g=c (el contorno que estamos siguiendo) toca tangencialmente (no corta) una curva de nivel de f, no se incrementa o disminuye el valor de f. Esto ocurre en el extremo local restringido y en los puntos de inflexión restringidos de f.

Un ejemplo familiar puede ser obtenido de los mapas climatológicos, con sus curvas de nivel de presión y temperatura (isóbaras e isotermas respectivamente): el extremo restringido ocurrirá donde los mapas superpuestos muestren curvas que se tocan.

Geométricamente traducimos la condición de tangencia diciendo que los gradientes de f y g son vectores paralelos en el máximo. Introduciendo un nuevo escalar, λ, resolvemos:

$$\nabla[f(x,y) - \lambda(g(x,y) - c)] = 0 \tag{12.3}$$

para $\lambda \neq 0$.

Una vez determinados los valores de λ, volvemos al número original de variables y así continuamos encontrando el extremo de la nueva ecuación no restringida.

$$F(x,y) = f(x,y) - \lambda(g(x,y) - c) \tag{12.4}$$

de forma tradicional. Eso es, $F(x,y) = f(x,y)$ para todo (x, y) satisfaciendo la condición porque $g(x,y) - c$ es igual a cero en la restricción, pero los ceros de $\nabla F(x,y)$ están todos en $g(x,y) = c$.

12.2 El método de los multiplicadores de Lagrange

Sea $f(x)$ una función definida en un conjunto abierto n-dimensional $x \in R$. Se definen s restricciones $gk(x) = 0, k = 1, \ldots, s$, y se observa (si las restricciones son satisfechas) que:

$$h(\mathbf{x}, \lambda) = f - \sum_{k=1}^{s} \lambda_k g_k \tag{12.5}$$

Se procede a buscar un extremo para h,

$$\frac{\partial h}{\partial x_i} = 0 \tag{12.6}$$

lo que es equivalente a:

$$\frac{\partial f}{\partial x_i} = \sum_{k}^{s} \lambda_k \frac{\partial g_k}{\partial x_i} \tag{12.7}$$

12.3 Demostración

Comenzemos con el caso de una restricción.

Sea una superficie M contenida en R^n definida por $g(x) = 0$ y sea $f(x)$ la función a obtener su punto crítico. Si $p \in M$ un punto crítico entonces se ha de cumplir:

$$\nabla f \cdot v = 0 \tag{12.8}$$

para todo v vector tangente a M en p (es decir, sea cual sea la dirección en la que nos desplacemos en M, el incremento de f a primer orden es nulo) La anterior condición significa que ∇f es perpendicular al tangente a M en p y dado que dim M=n-1 existe un único vector perpendicular linealmente independiente que viene dado por ∇g, de modo que se tiene:

$$\nabla f = \lambda \nabla g \tag{12.9}$$

para algún número λ

En el caso de que M esté definida por varias restricciones $g_1, \ldots, g_k$ el conjunto de vectores perpendiculares al tangente a M en p viene generado por $\nabla g_1, \nabla g_2, ... \nabla g_k$ de modo que al ser ∇f perpendicular al vector tangente a M en p este ha de ser de la forma:

$$\nabla f = \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2 + ... + \lambda_k \nabla g_k \tag{12.10}$$

para unos ciertos números $\lambda_1, ..., \lambda_k$

Los multiplicadores desconocidos λk se determinan a partir de las ecuaciones con las restricciones y conjuntamente se obtiene un extremo para h que al mismo tiempo satisface las restricciones (i.e. gk=0), lo que implica que f ha sido optimizada.

El método de multiplicadores de Lagrange es generalizado por las condiciones de Karush-Kuhn-Tucker.

12.4 Clasificadores

El término clasificador se utiliza en referencia al algoritmo utilizado para asignar un elemento entrante no etiquetado en una categoría concreta conocida. Dicho algoritmo, permite pues, ordenar o disponer por clases elementos entrantes, a partir de cierta información característica de éstos.

Una manera de implementar un clasificador es seleccionar un conjunto de ejemplos etiquetados y tratar de definir una regla que permita asignar una etiqueta a cualquier otro dato de entrada.

En ocasiones, el término clasificador también es utilizado para referirse a la función matemática que implementa el algoritmo de clasificación.

Para la implementación de un clasificador es necesario tener en cuenta una serie de características concretas. La selección de éstas, sin embargo, no es una tarea sencilla.

La adición de muchos parámetros irrelevantes hace más difícil la clasificación para todos los métodos. Además, a medida que vamos añadiendo más información se van incrementando las dimensiones del espacio, hecho que supone que la optimización sea cada vez más difícil.

12.4.1 Tipos de clasificadores

- **Clasificador Bayesiano**

  Es un clasificador basado en el aprendizaje que calcula la probabilidad de cada hipótesis de los datos y realiza predicciones sobre éstas. Estimar la probabilidad es complejo, pero se simplifica si se considera que los atributos son independientes dada la hipótesis: P(E1, E2, ...EN | H) = P(E1 | H) P(E2 | H) ... P(EN | H) Por lo que la probabilidad de la hipótesis dada la evidencia puede estimarse como: P(H | E1, E2, ...EN) = [P(H) P(E1 | H) P(E2 | H) ... P(EN | H)] / P(E)

- **Clasificador Parzen**

  Se trata de un clasificador basado en la estimación no paramétrica, que a diferencia de la paramétrica, donde se obtiene la función de densidad de probabilidad estimando los parámetros desconocidos de un modelo conocido, no se conoce el modelo. Esta técnica consiste básicamente en variaciones de la aproximación del histograma de una función de densidad de probabilidad desconocida. Este algoritmo se utiliza en la la clasificación de imágenes o para clasificar y aprender simultáneamente. El problema de este clasificador es que tiene un tiempo de ejecución elevado.

- **Clasificador Backpropagation**

  En algunos problemas el uso de modelos simples de densidad paramétrica o el uso de modelos de histogramas no dan los resultados deseados. En estos casos se buscan modelos de densidad más sofisticados. Las redes neuronales son una técnica de aproximación paramétrica útil para construir modelos de densidad. El modelo de red neuronal habitual que utiliza este algoritmo consiste en una red con una capa de entrada con tantos nodos como entradas tengan, una capa oculta con un número de nodos variable que dependerá de las características del problema, y una capa de salida con tantos nodos como posibles salidas tenga.

- **Clasificador con PCA**

  El clasificador en cuestión utiliza el análisis de componentes principales (también conocido como ACP o PCA) con el objetivo de tratar de reducir la dimensión de un grupo de datos y hallar las causas de su variabilidad para poder ordenarlas según su importancia. Máquinas de vectores de soporte “Support vector machine” Una máquina de vectores de soporte utiliza una técnica que aprende de dos clases distintas de entrada. Como un clasificador de una sola clase, la descripción dada por los datos de los vectores de soporte elabora una frontera de decisión alrededor de los datos de aprendizaje. Los datos son mapeados donde se busca la máxima separación entre clases. La función de frontera tiene la misión de separar los datos en clases distintas, cada una formando un agrupamiento.

12.4.2 Tipos de aprendizaje

- **Aprendizaje supervisado**

  El aprendizaje supervisado es una técnica de aprendizaje artificial que elabora una función matemática (hipótesis) a partir de datos de entrenamiento previamente etiquetados. Donde el usuario con unos datos de entrenamiento en una máquina puede deducir entre un conjunto de datos de entrada a que clase pertenecen los datos de salida.

- **Aprendizaje no supervisado**

Los algoritmos en aprendizaje no supervisado no disponen de un conjunto de entrenamiento que permita conocer las etiquetas de los datos, así pues, se hace necesario el uso de técnicas de agrupamiento que intentan construir estas etiquetas.

Este sistema de agrupamiento (o clustering) tiene como finalidad catalogar los objetos en conjuntos tales que los que estén en el mismo sean muy semejantes entre sí, mientras que el grado de semejanza entre grupos diferentes sea bajo.

Aun así, uno de los problemas que presenta este método es la toma de decisiones a la hora de escoger un patrón entre todos los proporcionados.

- **Aprendizaje semi-supervisado**

  El aprendizaje semi-supervisado surge como consecuencia de la dificultad que se tiene para obtener los datos etiquetados requeridos por los métodos no supervisados, ya que éstos deben ser etiquetados por un experto de forma manual convirtiéndose en un trabajo pesado.

  Este aprendizaje es una combinación del aprendizaje supervisado y no supervisado. Puesto que asignar etiquetas o clases de datos puede ser muy costoso, el aprendizaje semi-supervisado recurre a la opción de usar a la vez una colección limitada de datos etiquetados y un conjunto más extenso de datos no etiquetados, mejorando así la construcción de los modelos.

  En este método, se ha de tener en cuenta que no siempre los datos no etiquetados son de ayuda al proceso de aprendizaje. Por lo general se asume que los datos no etiquetados siguen la misma distribución que los etiquetados para que el uso de datos sin etiquetar sea útil.

- **Aprendizaje por refuerzo**

  El aprendizaje por refuerzo no trata de aprender a partir de un conjunto de ejemplos sino a través de la experiencia.

  En este tipo de aprendizaje, el agente dispone de un conjunto de opciones disponibles para llegar a resolver una determinada tarea. Para ello, debe escoger en primer lugar, la opción que considere más oportuna para lograr su objetivo. Una vez hecho, el agente en cuestión recibe información de retroalimentación del entorno sobre su desempeño. Posteriormente, dicha información pueda ser utilizada para modificar su comportamiento. El aprendizaje por refuerzo conlleva pues, una fuerte carga de ensayo y error.

12.5 Ejercicios - Probabilidad condicional

Resuelva los siguientes ejercicios relacionados a probabilidad condicional.

Ejemplo **12.1**

Servicios de telecomunicaciones. Suponga que es de interés conocer la probabilidad de que un usuario de Telex, que tiene señal de televisión satelital, tenga servicio ilimitado de larga distancia de cobertura nacional. El espacio muestral se reduce automáticamente, se condiciona a la ocurrencia de un evento, se restringe a los usuarios que tienen televisión satelital. Ahora bien, 1 de cada 5 usuarios tiene señal de televisión y 3 de cada 25 tiene tanto señal de televisión como larga distancia nacional, por lo que, por cada 25 usuarios, habrá 5 que tengan señal de televisión y 3 de ellos también tendrán servicio de larga distancia nacional; es decir, por cada 5 usuarios con televisión satelital, habrá 3 con servicio de larga distancia nacional. Entonces, la probabilidad de que un usuario tenga servicio de larga distancia nacional, dado que tiene señal de televisión satelital, es $3/5$. Calcule las siguientes probabilidades de que un usuario de Telex:

1. Tenga señal de televisión satelital, dado que tiene Internet de banda ancha.
2. Tenga Internet de banda ancha, dado que tiene señal de televisión satelital.
3. No tenga Internet, dado que tiene señal de televisión satelital.
4. Tenga señal de televisión satelital, dado que no tiene servicio de larga distancia.
5. Tenga servicio de larga distancia, dado que tiene señal de televisión.
6. Tenga larga distancia o señal de televisión, dado que tiene Internet.

Solución:

1. Tenga señal de televisión satelital, dado que tiene Internet de banda ancha:
$$P(T|I) = \frac{P(I \cap T)}{P(I))} = \frac{0.10}{0.30} = \frac{1}{3} \tag{12.11}$$
2. Tenga Internet de banda ancha, dado que tiene señal de televisión satelital:
$$P(I|T) = \frac{P(I \cap T)}{P(T))} = \frac{0.10}{0.20} = \frac{1}{2} \tag{12.12}$$
3. No tenga Internet, dado que tiene señal de televisión satelital:
$$P(\overline{I}|T) = \frac{P(\overline{I} \cap T)}{P(T))} = \frac{0.10}{0.20} = \frac{1}{2} \tag{12.13}$$
4. Tenga señal de televisión satelital, dado que no tiene servicio de larga distancia:
$$P(T|\overline{L}) = \frac{P(T \cap \overline{L})}{P(\overline{L}))} = \frac{0.00}{0.30} = 0 \tag{12.14}$$
5. Tenga servicio de larga distancia, dado que tiene señal de televisión:
$$P(L|T) = \frac{P(L \cap T)}{P(T))} = \frac{0.20}{0.20} = 1 \tag{12.15}$$

6. Tenga larga distancia o señal de televisión, dado que tiene Internet:

$$
\begin{aligned}
P(L \cup UI) &= P(L|I) + P(T|I) - P(L \cap U|I) && (12.16)\\
&= \frac{P(L \cap I)}{P(I)} + \frac{P(T \cap I)}{P(I)} - \frac{P(L \cap T \cap I)}{P(I)} && (12.17)\\
&= \frac{0.23}{0.30} + \frac{0.10}{0.30} - \frac{0.10}{0.30} && (12.18)\\
&= \frac{0.23}{0.30} && (12.19)
\end{aligned}
$$

Ejemplo 12.2

Dado. Considere el experimento consistente en lanzar un dado y observar la cara que queda hacia arriba. Sean los eventos: A = cae par, B = cae 2 o 4 y C = cae 1 o 2; las probabilidades correspondientes son: P(A) = 1/2, P(B) = 1/3 y P(C) = 1/3. Calcularemos las siguientes probabilidades condicionales, utilizando la interpretación clásica o a priori:
Solución:

$$
\begin{aligned}
P(A|B) &= \frac{N(A \cap B)}{N(B)} = \frac{N(\text{cae 2 o 4})}{N(\text{cae 2 o 4})} = 1 && (12.20)\\
P(B|A) &= \frac{N(A \cap B)}{N(A)} = \frac{N(\text{cae 2 o 4})}{N(\text{cae par})} = \frac{2}{3} && (12.21)\\
P(A|C) &= \frac{N(A \cap C)}{N(C)} = \frac{N(\text{cae 2})}{N(\text{cae 1 o 2})} = \frac{1}{2} && (12.22)\\
P(C|A) &= \frac{N(A \cap C)}{N(A)} = \frac{N(\text{cae 2})}{N(\text{cae par})} = \frac{1}{3} && (12.23)
\end{aligned}
$$

Ejemplo 12.3

Considere una urna que contiene 6 bolas rojas, 4 blancas y 5 azules, de la que se extraen sucesivamente de la urna tres bolas, con reemplazo. Sean los eventos:

- A = {sale bola azul}
- B = {sale bola blanca}
- R = {sale bola roja}

Calcule las siguientes probabilidades:

1. Que salgan en el orden roja, azul, blanca.
2. Que salgan en el orden azul, roja azul.
3. Que salgan tres bolas blancas.

4. Que salgan una roja, una azul y una blanca, sin importar el orden.

5. Que salgan dos azules y una roja.

Solución:

1. Que salgan en el orden roja, azul, blanca:

$$P(R \cap A \cap B) = P(R) * P(A|R) * P(B|R \cap A) \quad (12.24)$$

$$= \frac{6}{15} * \frac{5}{15} * \frac{4}{15} = \frac{120}{3375} = \frac{8}{225} \quad (12.25)$$

2. Que salgan en el orden azul, roja azul:

$$P(A \cap R \cap A) = P(A) * P(R|A) * P(A|A \cap R) \quad (12.26)$$

$$= \frac{5}{15} * \frac{6}{15} * \frac{5}{15} = \frac{150}{3375} = \frac{2}{45} \quad (12.27)$$

3. Que salgan tres bolas blancas:

$$P(B \cap B \cap B) = P(B) * P(B|B) * P(B|B \cap B) \quad (12.28)$$

$$= \frac{4}{15} * \frac{4}{15} * \frac{4}{15} = \frac{64}{3375} \quad (12.29)$$

4. Que salgan una roja, una azul y una blanca, sin importar el orden:

$$P(\text{1 roja, 1 azul y 1 blanca}) = 6 * P(R \cap A \cap B) \quad (12.30)$$

$$= 6\left(\frac{8}{225}\right) = \frac{16}{75} \quad (12.31)$$

5. Que salgan dos azules y una roja:

$$P(\text{2 azules y 1 roja}) = 3\left(\frac{2}{45}\right) = \frac{2}{15} \quad (12.32)$$

Ejemplo 12.4

Tetraedro. Si en lugar de lanzar un dado, se lanza un tetraedro,cuyas caras están numeradas del 1 al 4, definimos los eventos: A = {1, 2}, B = {1, 3}, C = {1, 4}, las probabilidades correspondientes son: P(A) = 1/2, P(B) = 1/2, P(C) = 1/2. Ahora, utilizando el criterio de Laplace, calculamos las siguientes probabilidades conjuntas:

Solución:

$$P(A \cap B) = \frac{N(A \cap B)}{N} = \frac{1}{4} \quad (12.33)$$

$$P(A \cap C) = \frac{N(A \cap C)}{N} = \frac{1}{4} \quad (12.34)$$

$$P(B \cap C) = \frac{N(B \cap C)}{N} = \frac{1}{4} \quad (12.35)$$

$$P(A \cap B \cap C) = \frac{N(A \cap B \cap C)}{N} = \frac{1}{4} \quad (12.36)$$

Vemos que los eventos A, B y C son independientes por pares, pues:

$$P(A \cap B) = P(A)P(B) = \frac{1}{2} * \frac{1}{2} = \frac{1}{4} \quad (12.37)$$

$$P(A \cap C) = P(A)P(C) = \frac{1}{2} * \frac{1}{2} = \frac{1}{4} \quad (12.38)$$

$$P(B \cap C) = P(B)P(C) = \frac{1}{2} * \frac{1}{2} = \frac{1}{4} \quad (12.39)$$

$$P(A \cap B \cap C) = \frac{N(A \cap B \cap C)}{N} = \frac{1}{4} \quad (12.40)$$

Sin embargo, A, B y C no son mutuamente independientes, ya que:

$$P(A \cap B \cap C) = P(A)P(B)P(C) = \frac{1}{2} * \frac{1}{2} * \frac{1}{2} = \frac{1}{8} \quad (12.41)$$

Ejemplo 12.5

Defectuosos. Tres máquinas A, B y C producen, respectivamente, 50%, 30% y 20% del número total de artículos de una fábrica; los porcentajes de desperfectos de producción de estas máquinas son, respectivamente, 4%, 2% y 3%. Existe interés en determinar cuál es el porcentaje de artículos defectuosos en toda la fábrica.
Solución:
Intuitivamente, tal porcentaje se puede obtener sumando los productos de los porcentajes de producción de cada máquina por sus correspondientes porcentajes de desperfectos:

$$(0.50)(0.04) + (0.30)(0.02) + (0.20)(0.03) = 0.02 + 0.006 + 0.006 = 0.032 \quad (12.42)$$

$$(12.43)$$

El porcentaje total de artículos defectuosos es de 3.2%, formalmente:
Sean los eventos:

A = {artćulo producido por la máquina A}
B = {artćulo producido por la máquina B}
C = {artćulo producido por la máquina C}
D = {artćulo defectuoso}

Con las siguientes probabilidades:

$$P(A) = 0.50 \quad , \quad P(D \mid A) = 0.04 \tag{12.44}$$

$$P(B) = 0.30 \quad , \quad P(D \mid B) = 0.02 \tag{12.45}$$

$$P(B) = 0.20 \quad , \quad P(D \mid C) = 0.03 \tag{12.46}$$

Entonces:

$$\begin{aligned} P(D) &= P(A)P(D \mid A) + P(B)P(D \mid B) + P(C)P(D \mid C) && (12.47)\\ &= (0.50)(0.04) + (0.30)(0.02) + (0.20)(0.03) && (12.48)\\ &= 0.02 + 0.006 + 0.006 && (12.49)\\ &= 0.032 && (12.50) \end{aligned}$$

12.6 Eigenvalores y Eigenvectores

En álgebra lineal, los vectores propios, autovectores o eigenvectores de un operador lineal son los vectores no nulos que, cuando son transformados por el operador, dan lugar a un múltiplo escalar de sí mismos, con lo que no cambian su dirección. Este escalar λ recibe el nombre valor propio, autovalor, valor característico o eigenvalor. A menudo, una transformación queda completamente determinada por sus vectores propios y valores propios. Un espacio propio, autoespacio, eigenespacio o subespacio fundamental asociado al valor propio λ es el conjunto de vectores propios con un valor propio común.

La palabra alemana eigen, que se traduce en español como propio, se usó por primera vez en este contexto por David Hilbert en 1904 (aunque Helmholtz la usó previamente con un significado parecido). Eigen se ha traducido también como inherente, característico o el prefijo auto-, donde se aprecia el énfasis en la importancia de los valores propios para definir la naturaleza única de una determinada transformación lineal. Las denominaciones vector y valor característicos también se utilizan habitualmente.

12.6.1 Ejercicios - Eigenvalores y Eigenvectores

Resuelva los siguientes ejercicios relacionados a Eigenvalores y Eigenvectores.

Ejemplo 12.6

Encuentre los valores propios de la matriz $A = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}$

Solución:
Por definición sabemos que: Si A es una matriz cuadrada, entonces un escalar λ es un valor propio de A si satisface la ecuacion: $Det(\lambda I - A) = 0$

$$\lambda I - A = \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 3 & 2 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} \lambda - 3 & -2 \\ 1 & \lambda \end{bmatrix} \tag{12.51}$$

$$Det(\lambda I - A) \quad = \quad \lambda^2 - 3\lambda + 2 = 0 \tag{12.52}$$

Resolviendo la ecuación tenemos: $\lambda = 1$ y $\lambda = 2$ son los valores propios de la matriz A.

Ejemplo 12.7

Encuentre los valores y vectores propios de la matriz $A = \begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix}$

Solución:
Por definición sabemos que: Si A es una matriz cuadrada, entonces un escalar λ es un valor propio de A si satisface la ecuacion: $Det(\lambda I - A) = 0$

$$\lambda I - A = \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix} = \begin{bmatrix} \lambda - 2 & 12 \\ -1 & \lambda + 5 \end{bmatrix} \tag{12.53}$$

$$Det(\lambda I - A) \quad = \quad (\lambda + 1)(\lambda + 2) = 0 \tag{12.54}$$

Resolviendo la ecuación tenemos: $\lambda = -1$ y $\lambda = -2$ son los valores propios de la matriz A.
Sea $X = \begin{bmatrix} X1 \\ X2 \end{bmatrix}$ un vector propio de A, X es un vector propio A correspondiente a λ si y solo si X es una solución no trivial de $(\lambda I - A)X = 0$, es decir, solución no trivial para:

$$\begin{bmatrix} \lambda - 2 & 12 \\ -1 & \lambda + 5 \end{bmatrix} \begin{bmatrix} X1 \\ X2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{12.55}$$

Si $\lambda = -1$ la ecuación se transforma en:

$$\begin{bmatrix} -3 & 12 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} X1 \\ X2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{12.56}$$

$$\begin{cases} -3X_1 + 12X_2 = 0 & X_1 = 4X_2 \\ -X_1 + 4X_2 = 0 & X_2 = t \end{cases}$$

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} 4t \\ t \end{bmatrix} = t \begin{bmatrix} 4 \\ 1 \end{bmatrix} \text{Son los vectores propios de A correspondientes a } \lambda = -1 \quad (12.57)$$

Si $\lambda = -2$ la ecuación se transforma en:

$$\begin{bmatrix} -4 & 12 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} X1 \\ X2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (12.58)$$

$$\begin{cases} -4X_1 + 12X_2 = 0 & X_1 = 3X_2 \\ -X_1 + 3X_2 = 0 & X_2 = t \end{cases}$$

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} 3t \\ t \end{bmatrix} = t \begin{bmatrix} 3 \\ 1 \end{bmatrix} \text{Son los vectores propios de A correspondientes a } \lambda = -2 \quad (12.59)$$

12.7 Polinomio de Bezier

Se denomina curvas de Bézier a un sistema que se desarrolló hacia los años 1960 para el trazado de dibujos técnicos, en el diseño aeronáutico y en el de automóviles. Su denominación es en honor a Pierre Bézier, quien ideó un método de descripción matemática de las curvas que se comenzó a utilizar con éxito en los programas de CAD.

Las curvas de Bézier fueron publicadas por primera vez en 1962 por el ingeniero francés Pierre Bézier, que las usó posteriormente con profusión en el diseño de las diferentes partes de los cuerpos de un automóvil, en sus años de trabajo en la Renault. Las curvas fueron desarrolladas por Paul de Casteljau usando el algoritmo que lleva su nombre. Se trata de un método numéricamente estable para evaluar las curvas de Bézier.

Posteriormente, los inventores del PostScript, lenguaje que permitió el desarrollo de sistemas de impresión de alta calidad desde el ordenador, introdujeron en ese código el método de Bézier para la generación del código de las curvas y los trazados. El lenguaje PostScript sigue empleándose ampliamente y se ha convertido en un estándar de calidad universal; por ello, los programas de diseño vectorial como Adobe Illustrator, el extinto Macromedia FreeHand y Corel Draw, tres de los programas más importantes de dibujo vectorial y otros como Inkscape, denominan ”bézier” a algunas de sus herramientas de dibujo, y se habla de ”trazados bézier”, ”pluma bézier”, ”lápiz bézier”, etc. Su facilidad de uso la ha estandarizado en el diseño gráfico, extendiéndose también a programas de animación vectorial, como Adobe Flash, y retoque fotográfico (bitmap), como Photoshop y Gimp, donde se usa para crear formas cerradas o selecciones.

La idea de definir geométricamente las formas no es demasiado compleja: un punto del plano puede definirse por coordenadas. Por ejemplo, un punto A tiene unas coordenadas (x1, y1) y a un punto B le corresponde (x2,y2). Para trazar una recta entre ambos basta con conocer su posición.

Si en lugar de unir dos puntos con una recta se unen con una curva, surgen los elementos esenciales de una curva Bézier; los puntos se denominan ”puntos de anclaje” o ”nodos”. La forma de la curva se define por unos puntos invisibles en el dibujo, denominados ”puntos de control”, ”manejadores” o ”manecillas”.

La curva de Bézier de grado n puede ser generalizada de la siguiente manera. Dados los puntos $P_0, P_1, ..., P_n$, la curva de Bézier es del tipo:

$$\mathbf{B}(t) = \sum_{i=0}^{n} \binom{n}{i} \mathbf{P}_i (1-t)^{n-i} t^i = \mathbf{P}_0 (1-t)^n + \binom{n}{1} \mathbf{P}_1 (1-t)^{n-1} t + \cdots + \mathbf{P}_n t^n \text{ , } t \in [0,1]. \tag{12.60}$$

Ejemplo 12.8

Por ejemplo, una curva de orden cinco ($n = 5$) quedaría como:

$$\mathbf{B}(t) = \mathbf{P}_0 (1-t)^5 + 5\mathbf{P}_1 t(1-t)^4 + 10\mathbf{P}_2 t^2 (1-t)^3 + 10\mathbf{P}_3 t^3 (1-t)^2 + 5\mathbf{P}_4 t^4 (1-t) + \tag{12.61}$$

$$\mathbf{P}_5 t^5 \text{ , } t \in [0,1]. \tag{12.62}$$

Esta ecuación puede ser expresada de manera recursiva como sigue: sea la expresión $\mathbf{B}_{\mathbf{P}_0\mathbf{P}_1\ldots\mathbf{P}_n}$ que denota la curva de Bézier determinada por los puntos $P_0, P_1, ..., P_n$. Entonces:

$$\mathbf{B}(t) = \mathbf{B}_{\mathbf{P}_0\mathbf{P}_1\ldots\mathbf{P}_n}(t) = (1-t)\mathbf{B}_{\mathbf{P}_0\mathbf{P}_1\ldots\mathbf{P}_{n-1}}(t) + t\mathbf{B}_{\mathbf{P}_1\mathbf{P}_2\ldots\mathbf{P}_n}(t) \tag{12.63}$$

En otras palabras, el grado n de la curva de Bézier es una interpolación entre los dos grados $n-1$ de las curvas de Bézier.

TerminologíaTerminología
Existe una terminología asociada exclusivamente para este tipo de curvas. Se tiene:

$$\mathbf{B}(t) = \sum_{i=0}^{n} \mathbf{P}_i \mathbf{b}_{i,n}(t), \quad t \in [0,1] \tag{12.64}$$

donde las polinomiales

$$\mathbf{b}_{i,n}(t) = \binom{n}{i} t^i (1-t)^{n-i}, \quad i = 0, \ldots n \tag{12.65}$$

son conocidas como polinomios de Bernstein de grado n, definidos por $0^0 = 1$.
Los puntos P_i son llamados puntos de control de las curvas de Bézier. El polígono formado por la conexión de los puntos de Bézier con rectas, comenzando por P_0 y terminando en P_n, se denomina polígono de Bézier (o polígono de control). La envolvente convexa del polígono de Bézier contiene las curvas de Bézier.

12.8 Cadena de Márkov

En la teoría de la probabilidad, se conoce como cadena de Márkov o modelo de Márkov a un tipo especial de proceso estocástico discreto en el que la probabilidad de que ocurra un evento depende solamente del evento inmediatamente anterior. Esta característica de falta de memoria recibe el nombre de propiedad de Markov.

Recibe su nombre del matemático ruso Andréi Márkov (1856-1922), que lo introdujo en 1907.

12.8.1 Definición formal

En matemática se define como un proceso estocástico discreto que cumple con la propiedad de Márkov, es decir, si se conoce la historia del sistema hasta su instante actual, su estado presente resume toda la información relevante para describir en probabilidad su estado futuro.

Una cadena de Márkov es una secuencia $X_1, X_2, X_3, \ldots$ de variables aleatorias. El dominio de estas variables es llamado espacio estado; el valor de Xn es el estado del proceso en el tiempo n. Si la distribución de probabilidad condicional de X_{n+1} en estados pasados es una función de X_n por sí sola, entonces:

$$P(X_{n+1} = x_{n+1} | X_n = x_n, X_{n-1} = x_{n-1}, \ldots, X_2 = x_2, X_1 = x_1) = P(X_{n+1} = x_{n+1} | X_n = x_n). \tag{12.66}$$

Donde xi es el estado del proceso en el instante i. La identidad mostrada es la propiedad de Márkov.

12.8.2 Ejemplos - Cadena de Márkov

Resuelva los siquientes problemas usando cadenas de Márkov

Ejemplo 12.9

El departamento de estudios de mercado de una fábrica estima que el 20% de la gente que compra un producto un mes, no lo comprará el mes siguiente. Además, el 30% de quienes no lo compren un mes lo adquirirá al mes siguiente. En una población de 1000 individuos, 100 compraron el producto el primer mes.

1. ¿Cuántos lo comprarán al mes próximo?
2. ¿Y dentro de dos meses?

Solución:

1. Con esa información construimos la matriz 2x2. $P^{(0)}$ representa la situación:

$$P^{(0)} = \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix} \tag{12.67}$$

$$(C,N) = \begin{bmatrix} 100 & 900 \end{bmatrix} \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix} = (350, 650) \tag{12.68}$$

El primer mes comprarán C=350 y no comprarán N=650

2. Para el segundo caso tenemos lo siguiente:

$$P^{(2)} = \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix} \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix} = \begin{bmatrix} 0.7 & 0.3 \\ 0.45 & 0.55 \end{bmatrix} \tag{12.69}$$

$$(C,N) = \begin{bmatrix} 100 & 900 \end{bmatrix} \begin{bmatrix} 0.7 & 0.3 \\ 0.45 & 0.55 \end{bmatrix} = (475, 525) \tag{12.70}$$

El segundo mes compraán C=475 y no comprarán N= 525

Ejemplo 12.10

En una población de 10,000 habitantes, 5000 no fuman, 2500 fuman uno o menos de un paquete diario y 2500 fuman más de un paquete diario. En un mes hay un 5% de probabilidad de que un no fumador comience a fumar un paquete diario, o menos, y un 2% de que un no fumador pase a fumar más de un paquete diario. Para los que fuman un paquete, o menos, hay un 10% de probabilidad de que dejen el tabaco, y un 10% de que pasen a fumar más de un paquete diario. Entre los que fuman más de un paquete, hay un 5% de probabilidad de que dejen el tabaco y un 10% de que pasen a fumar un paquete, o menos. ¿Cuántos individuos habrá de cada clase el próximo mes?

Solución:

$$P^{(1)} = \begin{array}{c|ccc} & 0 & 1 & 2 \\ \hline 0 & 0.93 & 0.05 & 0.02 \\ 1 & 0.10 & 0.80 & 0.10 \\ 2 & 0.05 & 0.10 & 0.85 \end{array} \tag{12.71}$$

NF= No fuman
FC= fuman uno o menos de un paquete diarios
FCC= fuman más de un paquete diario.

$$(NF, FCFCC) \quad = \quad \begin{bmatrix} 5000 & 2500 & 2500 \end{bmatrix} \begin{bmatrix} 0.93 & 0.05 & 0.02 \\ 0.10 & 0.80 & 0.10 \\ 0.05 & 0.10 & 0.85 \end{bmatrix} = (5025, 2500, 2475) \qquad (12.72)$$

Después de un mes habrán NF=5025, FC=2500, FCC=2475

Ejemplo 12.11

Una urna contiene dos bolas sin pintar. Se selecciona una bola al azar y se lanza una moneda. Si la bola elegida no está pintada y la moneda produce cara, pintamos la bola de rojo; si la moneda produce cruz, la pintamos de negro. Si la bola ya está pintada, entonces cambiamos el color de la bola de rojo a negro o de negro a rojo, independientemente de si la moneda produce cara o cruz.

1. Modele el problema como una cadena de Markov y encuentre la matriz de probabilidades de transición.

Solución:
Identificando estados: Vamos a utilizar vectores con el siguiente formato: [S R N] donde S es el número de bolas sin pintar, R el número de bolas rojas y N el número de bolas negras que hay en la urna.

- E0: [0 1 1] Una bola roja y una negra.
- E1: [0 2 0] Dos bolas rojas.
- E2: [0 0 2] Dos bolas negras.
- E3: [2 0 0] Inicialmente, cuando las dos bolas están sin pintar.
- E4: [1 1 0] Una bola pintada de rojo.
- E5: [1 0 1] Una bola pintada de negro.

Probabilidades de transición:

Como el estado de la urna después del siguiente lanzamiento de la moneda depende solo del pasado del proceso hasta el estado de la urna después del lanzamiento actual, se trata de una cadena de Markov. Como las reglas no varían a través del tiempo, tenemos una cadena estacionaria de Markov. La matriz de transición es la siguiente:

Cálculos de las probabilidades de transición si el estado actual es (1 1 0):

Evento	Probabilidad	Nuevo evento
Sacar cara en el lanzamiento y escoger una bola sin pintar	1/4	(0 2 0)
Escoger bola roja	1/2	(1 0 1)
Sacar cruz en el lanzamiento y escoger una bola sin pintar	1/4	(0 1 1)

Para ver cómo se forma la matriz de transición, determinamos la fila (1 1 0) Si el estado actual (1 1 0), entonces debe suceder uno de los eventos que aparecen en la tabla anterior. Así, el siguiente estado será (1 0 1) con la probabilidad 1/2; (0 2 0) con probabilidad 1/4 y (0 1 1) con probabilidad 1/4. En la siguiente imagen se da la representación gráfica de esta matriz de transición.

$$= \begin{bmatrix} 0 & 1/2 & 1/2 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 1/4 & 1/4 & 0 & 0 & 0 & 1/2 \\ 1/4 & 0 & 1/4 & 0 & 1/2 & 0 \end{bmatrix} \tag{12.73}$$

Ejemplo 12.12

En un pueblo, al 90% de los días soleados le siguen días soleados, y al 80% de los días nublados le siguen días nublados. Con esta información modelar el clima del pueblo como una cadena de Markov.

Solución:

Se trata de una cadena de Markov con dos estados Soleado, Nublado que para abreviar representaremos por S, N, siendo la matriz de probabilidades de transición

$$P = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \tag{12.74}$$

Ejemplo 12.13

Suponga que toda la industria de refresco produce dos colas: Coca Cola y Pepsi Cola. Cuando una persona ha comprado Coca Cola hay una probabilidad de 90% de que siga comprándola la vez siguiente. Si una persona compró Pepsi, hay 80% de que repita la vez siguiente. Se pide:

1. Si una persona actualmente es comprador de Pepsi. ¿Cuál es la probabilidad de que compre Coca Cola pasadas dos compras a partir de hoy?

2. Si en la actualidad una persona es comprador de Coca Cola. ¿Cuál es la probabilidad de que compre Coca Cola pasadas tres compras a partir de ahora?

3. Supongamos que el 60% de toda la gente toma hoy Coca Cola y el 40% Pepsi. A tres compras

a partir de ahora, ¿Qué fracción de los compradores estará tomando Coca Cola?.

Solución:
La situación se puede modelar como una cadena de Markov con dos estados CocaCola, PepsiCola= C, P. La matriz de transición para el orden C,P, es:

$$P^{(0)} = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \tag{12.75}$$

1. Se pide la probabilidad de transición en dos pasos, es decir que se pide el valor en fila 2, columna 1 para la matriz P^2 , obteniéndose que este es: $0.2.0.9 + 0.8.0.2 = 0.34$

$$P^{(2)} = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} = \begin{bmatrix} 0.83 & 0.17 \\ 0.34 & 0.66 \end{bmatrix} \tag{12.76}$$

2. Al igual que en el apartado anterior se pide el valor de probabilidad de transición en fila 1 y columna 1 para la matriz P^3. La matriz es:

$$P^{(2)} = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix} \begin{bmatrix} 0.83 & 0.17 \\ 0.34 & 0.66 \end{bmatrix} = \begin{bmatrix} 0.781 & 0.219 \\ 0.438 & 0.562 \end{bmatrix} \tag{12.77}$$

3. El vector de probabilidad inicial es (0.6, 0.4), por tanto la probabilidad de consumir ambos estados a partir de tres etapas es:

$$(0.6, 0.4)P^{(3)} = \begin{bmatrix} 0.6 & 0.4 \end{bmatrix} \begin{bmatrix} 0.781 & 0.219 \\ 0.438 & 0.562 \end{bmatrix} = (0.6438, 0.3562) \tag{12.78}$$

12.9 ArgMax

En la matemática, los argumentos de los máximos (abreviado arg max o argmax) son los puntos del dominio de alguna función en la que se maximizan los valores de la función. En contraste con los máximos globales, refiriéndose a las salidas más grandes de una función , Arg max se refiere a las entradas o argumentos en los que las salidas de función son lo más grandes posible.

12.10 Maximum a posteriori estimation - MAP

En estadística bayesiana, una estimación máxima a posteriori (MAP) es un modo de distribución posterior. El MAP puede utilizarse para obtener una estimación puntual de una cantidad no observada sobre la base de datos empíricos. Está estrechamente relacionado con el método de estimación de máxima verosimilitud (ML) de Fisher, pero emplea un objetivo de optimización aumentada que incorpora una distribución previa sobre la cantidad que se desea estimar. Por tanto, la estimación del MAP puede considerarse como una regularización de la estimación del ML.

Supongamos que queremos estimar un parḿetro poblacional no observado θ sobre la base de las observaciones x. Sea f la distribución de muestreo de x, de modo que $f(x \mid \theta)$ es la probabilidad de x cuando el parámetro de población subyacente es θ. Entonces la función:

$$\theta \mapsto f(x \mid \theta) \tag{12.79}$$

Se conoce como la funció de verosimilitud y la estimación:

$$\hat{\theta}_{\mathrm{ML}}(x) = \underset{\theta}{\arg\max}\, f(x \mid \theta) \tag{12.80}$$

Es la estimación de máxima verosimilitud de θ.
Supongamos ahora que existe una distribución anterior g sobre θ. Esto nos permite tratar θ como una variable aleatoria como en las estadísticas bayesianas. Podemos calcular la distribución posterior de θ usando el teorema de Bayes:

$$\theta \mapsto f(\theta \mid x) = \frac{f(x \mid \theta)\, g(\theta)}{\displaystyle\int_{\vartheta \in \Theta} f(x \mid \vartheta)\, g(\vartheta)\, d\vartheta} \tag{12.81}$$

Donde g es la función de densidad de θ, Θ es el dominio de g.

El método de estimación máxima a posteriori estima entonces θ como el modo de la distribución posterior de esta variable aleatoria:

$$\begin{aligned}
\hat{\theta}_{\mathrm{MAP}}(x) &= \underset{\theta}{\arg\max}\, f(\theta \mid x) && (12.82)\\
&= \underset{\theta}{\arg\max}\, \frac{f(x \mid \theta)\, g(\theta)}{\displaystyle\int_{\vartheta} f(x \mid \vartheta)\, g(\vartheta)\, d\vartheta} && (12.83)\\
&= \underset{\theta}{\arg\max}\, f(x \mid \theta)\, g(\theta). && (12.84)
\end{aligned}$$

El denominador de la distribución posterior (denominada probabilidad marginal) no depende de θ y por lo tanto no juega ningún papel en la optimización. Obsérvese que la estimación MAP de θ coincide con la estimación ML cuando el g es uniforme (es decir, una función constante).

Cuando la función de pérdida es de la forma:

$$L(\theta, a) = \begin{cases} 0, & \text{if } |a - \theta| < c \\ 1, & \text{otherwise} \end{cases} \tag{12.85}$$

Como c pasa a 0, el estimador de Bayes se aproxima al estimador MAP, siempre que la distribución de θ sea unimodal. Pero generalmente un estimador MAP no es un estimador de Bayes a menos que θ sea discreto.

13 Exámen II - Probabilidad condicional

Probabilidad, procesos aleatorios e interferencia

Autor: Horacio Rodríguez Bazán.
Estatus: Creación
Nombre: ________________________________

Resuelva los siguientes problemas, relacionados a probabilidad condicional.

1. Una familia tiene dos niños. Encuentra la probabilidad que los dos niños sean mujeres, dado que al menos uno de los dos niños nació en invierno. Considera que las cuatro estaciones son igualmente probables, que el género es independiente de la estación y que los dos géneros son igualmente probables.

   Solución:

   Sea:

   (a) A = dos mujeres

   (b) B = almenos una mujer en invierno

   (c) $A \cap B$ = dos mujeres y almenos una mujer en invierno

   por la de finición de probabilidad condicional:

   $$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \tag{13.1}$$

   La probabilidad de que el niño sea mujer es $\frac{1}{2}$ y la probabilidad de que el niño nazca en invierno es $\frac{1}{4}$, por lo que la probabilidad de que un niño sea mujer y que nazca en invierno es: $\frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$, ahora la probabilidad de que el otro niño sea mujer es: $\frac{1}{2}$ y que nazca en cualquier estación $\frac{1}{4}$. Entonces la probabilidad de que sean dos mujeres y la primera haya nacido en invierno es: $\frac{1}{8} \cdot \frac{1}{8} = \frac{1}{64}$.

Utilizando la analogía anterior se pueden obtener las probabilidades, se puede hacer un diagrama de árbol de cuatro niveles donde el primer nivel consta de 4 elementos ($P = 1/4$) que representan las estaciones del año, el segundo nivel (de 8 elementos, dos por cada estación) representa al género (hombre o mujer) con $P = 1/2$, hasta este nivel representa solo un niño, ahora para representar al segundo niño, en el tercer nivel se representan las estaciones del año($P = 1/4$) y el último nivel representa el género.

Ahora para obtener la $P(B)$ se toma en cuenta solo las ramas donde al menos un niño de género mujer nació en invierno lo que genera las siguientes expresiones:

$$P(B) = 7(\frac{1}{64}) + \frac{1}{8} = \frac{7}{64} + \frac{8}{64} = \frac{15}{64} \tag{13.2}$$

Ahora se recorre el árbol buscando esta vez solo las ramas donde nacierón dos mujeres y al menos una mejer nació en invierno, generando la siguiente expresión:

$P(A \cap B) = \frac{3}{64} \cdot \frac{4}{64} = \frac{7}{64}$

entonces:

$$P(A \mid B) = \frac{7/64}{15/64} = \frac{7}{15} \tag{13.3}$$

2. Se cuentan con dos moneda, una moneda normal, y otra moneda preparada que cae cara con una probabilidad de $3/4$. Se escoge una moneda al azar y se lanzas 3 veces. Esta cae cara las 3 veces. Con esta información ¿cuál es la probabilidad de que la moneda escogida sea la moneda normal?

   Solución:

   Sea:

   (a) A= el evento de que la moneda escogida cae cara 3 veces
   (b) B= el evento de escoger la moneda normal
   (c) C= el evento de escoger la moneda preparada

   Entonces, el evento de elegir una moneda al azar es:

$$P(B) = P(C) = 1/2 \tag{13.4}$$

   Si consideramos de obtener cara con la moneda normal es $1/2$, y la probabilidad condicional de que caiga tres veces cara dado que fue la moneda normal:

$$P(A \mid B) = (1/2) \cdot (1/2)^3 \tag{13.5}$$

donde el primer término representa la probabilidad de escoger la moneda normal, y el segundo la probabilidad de que al lanzar 3 veces la moneda, obtengamos 3 caras.

Ahora bien para obtener la probabilidad de que salgan 3 caras en el evento de lanzar la moneda ($P(A)$), se hace uso del teorema de la probabilidad total, entonces:

$$P(A) = P(A \mid B)P(B) + P(A \mid C)P(C) \tag{13.6}$$

Se sabe que la probabilidad de que la moneda preparada caiga cara es $3/4$, por lo que:

$$P(A) = (1/2)^3 \cdot (1/2) + (3/4)^3 \cdot (1/2) \tag{13.7}$$

Usando el teorema de Bayes, entonces:

$$\begin{aligned}
P(B \mid A) &= \frac{P(A \mid B)P(B)}{P(A)} && (13.8)\\
&= \frac{P(A \mid B)P(B)}{P(A \mid B)P(B) + P(A \mid C)P(C)} && (13.9)\\
&= \frac{(1/2)^3 \cdot 1/2}{(1/2)^3 \cdot 1/2 + (3/4)^3 \cdot 1/2} && (13.10)\\
&\approx 0.23 && (13.11)
\end{aligned}$$

3. Una amiba vive en un estanque. Pasado un minuto, la amiba morirá, se dividirá en dos o permanecerá igual, los tres eventos son igualmente probables, y en los minutos consecuentes las amibas resultantes tendrán el mismo comportamiento, de forma independiente. ¿Cuál es la probabilidad que la población de amibas eventualmente muera?

   Solución:

   Sea:

   (a) M= el evento de que todas las amibas mueran.

   (b) A_i= el evento de que la amiba inicial se divida en i amibas, considerando el caso mas sencillo, para el primer minutos se pueden tener las siguientes amibas $i = 0, 1, 2$

Entonces, para el primer minuto hay 3 opciones:

$$P(M) = P(M \mid A_0) \cdot \frac{1}{3} + P(M \mid A_1) \cdot \frac{1}{3} + P(M \mid A_2) \cdot \frac{1}{3} \tag{13.12}$$

La $P(M \mid A_0) = 1$ ya que mure la amiba pasado el primer minuto,
La $P(M \mid A_1) = P(M)$ ya que la amiba no cambia, (no muere ni se divide), por lo que regresamos al caso inicial.

La $P(M \mid A_2) = P(M)^2$, se generan dos amibas y estas cuentan con las mismas características que la amiba inicial; entonces:

$$P(M) = \frac{1}{3} + \frac{P(M)}{3} + \frac{P(M)^2}{3} \tag{13.13}$$

$$\frac{1}{3}P(M)^2 - \frac{2}{3}P(M) + \frac{1}{3} = 0 \tag{13.14}$$

Resolviendo la eciación:

$$P(M) = 1 \tag{13.15}$$

Por lo que la población de amibas morirá con el paso del tiempo.

4. Es posible tener eventos A, B, C tal que $P(A \mid C) < P(B \mid C)$ y $P(A \mid C^c) < P(B \mid C^c)$, entonces ¿$P(A) > P(B)$?. Muestra que esto es posible.

   Solución:

   No es posible, usando la ley total de la probabilidad se tiene:

$$P(A) = P(A \mid C)P(C) + P(A \mid C^c) < P(B \mid C)P(C) + P(B \mid C^c) = P(B) \tag{13.16}$$

$$P(A) < P(B) \tag{13.17}$$

5. Sean tres urnas con las siguientes composiciones de bolas blancas y negras:

$$U_1 : (3\, blancas\, y\, 2\, negras)$$
$$U_2 : (4\, blancas\, y\, 2\, negras)$$
$$U_2 : (1\, blancas\, y\, 4\, negras)$$

Calcúlese:

a) Probabilidad de extraer bola blanca.

b) Probabilidad de que una bola negra que se ha extraido procesa de la segunda urna.

Solución

a) Suponemos inicialmente, puesto que nada indica lo contrario, que las tres urnas son equiprobables:

$$P(U_1) = P(U_2) = P(U_3) = \frac{1}{3} \tag{13.18}$$

Por el teorema de la probabilidad total:

$P(blanca) = P(blanca|U_1)P(U_1) + P(blanca|U_2)P(U_2) + P(blanca|U_3)P(U_3)$

$= \frac{3}{5}\frac{1}{3} + \frac{4}{6}\frac{1}{3} + \frac{1}{5}\frac{1}{3} = \frac{22}{45}$

b) Por el teorema de Bayes:

$$P(U_2|negra) = \frac{P(negra|U_2)P(U_2)}{P(negra)} \tag{13.19}$$

donde $P(negra)$ se puede determinar, como antes en el caso de la bola blanca, por el teorema de la probabilidad total y, por tanto.

$$P(U_2|negra) = \frac{P(negra|U_2)P(U_2)}{P(negra|U_1)P(U_1) + P(negra|U_2)P(U_2) + P(negra|U_3)P(U_3)} \tag{13.20}$$

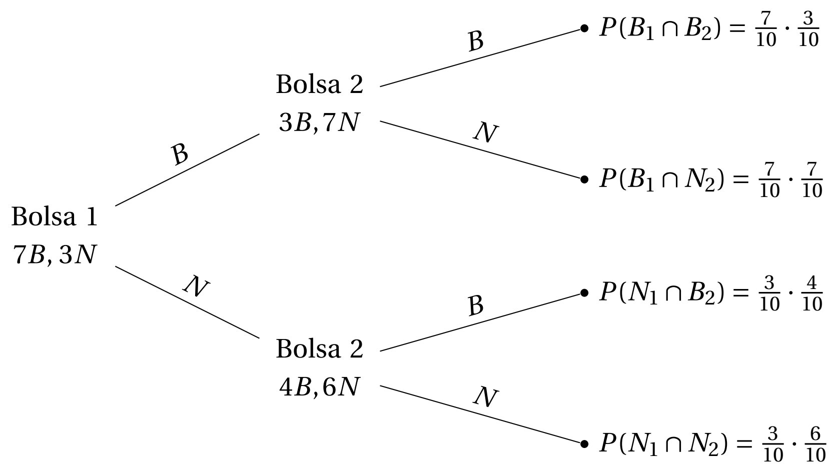
$$P(U_2|negra) = \frac{\frac{2}{6}\frac{1}{3}}{\frac{2}{5}\frac{1}{3} + \frac{2}{6}\frac{1}{3} + \frac{4}{5}\frac{1}{3}} = \frac{5}{23} \tag{13.21}$$

6. Tenemos dos bolsas, la primera con 10 bolas, 7 blancas y 3 negras, y la segunda con 9 bolas, 3 blancas y 6 negras.

   Se extrae al azar una bola de la primera bolsa y se pasa a la segunda. De esta bolsa, tambien al azar, se saca una bola, calcúlese la probabilidad que esta bola sea blanca.

   Solución:

   En el diagrama siguiente tenemos las distintas maneras de legar a la extraccion de la segunda bola:



En el diagrama se aprecia que hay dos caminos para obtener bola blanca en la segunda extracción. caminos que hay qye tomar en cuenta a la hora de calcular la probabilidad, pues al final del experimento se puede llegar por uno u por otro, sin saber de antemano cual tendra lugar.

El diagrama en términos de sucesos, adopta la forma anterior, siendo B_1 la extracciøón de bola blanca en la i-ésima extracciøón, y análogamente para N_1 como bola negra.

EL suceso de bola blanca es la segunda extracción sin establecer condición alguna (camino seguido), (B_2), es igual a

$$(B_2) = (B_1 \cap B_2) \cup (N_1 \cap B_2) \tag{13.22}$$

Se puede llegar a B_2 a través de dos sucesos disjuntos, los del segundo miembro de laigualdas, siendo su probabilidad

$$P(B_1 \cap B_2) = P(B_1)P(B_2|B_1) \tag{13.23}$$

$$P(N_1 \cap B_2) = P(N_1)P(B_2|N_1) \tag{13.24}$$

por lo cual la probabilidad de extraer bola blanca, $P(B_2)$, es:

$$P(B_2) = P(B_1)P(B_2|B_1) + P(N_1 \cap B_2) = P(N_1)P(B_2|N_1) \tag{13.25}$$

$$P(B_2) = \frac{7}{10}\frac{4}{10} + \frac{3}{10}\frac{3}{10} = \frac{37}{100} \tag{13.26}$$

7. En el centro de Ecatepec, un ladrón es perseguido por un coche de policia y al llegar a un determinado cruce en San Agustín se encuentra tres posibles calles por las que huir A, B y C, de tal manera que las dos ultimas son tan estrechas que por ellas no cabe el coche de la policia, si bien el ladrón va tan nervioso que no es consciente de ello.

   Si huye por la calle A le atrapan seguro puesto que el final de la misma está cortado por otra patrulla de la policia. Si huye por la calle C escapa seguro puesto que no esta vigilada. Si huye por la calle B se encuentra con que está cortada y que se bifurca en dos callejuelas: la BA, que conduce a la calle A y la BC que lleva a la calle C.

   ¿Cuál es la probabilidad de que el ladrón sea atrapado?

   Solución:

$$\begin{aligned} P(atraparle) = {} & \text{P(A) P(atraparle | A)} + \\ & \text{P(B)}\ \text{P(atraparle | B)} + \\ & \text{P(C) P(atraparle | C)} \end{aligned} \tag{13.27}$$

$$\begin{aligned} \text{P(atraparle | B)} = {} & \text{P(BA)P(atraparle | BA)} + \\ & \text{P(BC)P(atraparle | BC)} \end{aligned} \tag{13.28}$$

$$\begin{aligned} P(atraparle) = {} & \text{P(A) P(atraparle | A)} + \\ & \text{P(B)}\ \text{(P(BA) P(atraparle | BA)} + \\ & \text{P(BC) P(atraparle | BC))} + \\ & \text{P(C) P(atraparle | C)} = \frac{1}{3}1 + \frac{1}{3}(\frac{1}{2}1 + \frac{1}{2}0) + \frac{1}{3}0 = \frac{1}{2} \end{aligned} \tag{13.29}$$

8. Supongamos que haya un análisis para detectar el cáncer con una fiabilidad del 98 por ciento; es decir, si uno tiene cáncer el análisis dará positivo el 98 por ciento de las veces, y si no lo tiene, dará negativo el 98 por ciento de las veces. Supongamos además que el 0,5 por ciento de la población, una de cada doscientas personas, padece verdaderamente cáncer. Imaginemos que uno se ha sometido al análisis y que su mádico le informa con tono pesimista que ha dado positivo. ¿Hasta quá punto ha de deprimirse esa persona? Lo sorprendente del caso es que dicho paciente ha de mantenerse prudentemente optimista. El por quá de este optimismo lo encontraremos al determinar la probabilidad condicional de que uno tenga un cáncer sabiendo que el análisis ha dado positivo. Supongamos que se hacen 10.000 pruebas de cáncer. ¿Cuántas de ellas darán positivo?

   Solución:

   En promedio, 50 de estas 10.000 personas (el 0,5 por ciento de 10.000) tendrán cáncer, y como el 98 por ciento de ellas darán positivo, tendremos 49 análisis positivos. Por otra parte, el 2 por ciento de las 9.950 personas restantes, que no padecen cáncer, también darán positivo, con un total de 199 análisis positivos $(0,02)(9.950) = 199)$. Así, del total de 248 positivos $(199+49 = 248)$, la mayoría (199) son falsos positivos, y la probabilidad condicional de padecer el cáncer sabiendo que se ha dado positivo es sólo $\frac{49}{248}$

9. Dos jugadores A y B compiten en un test que incluye una serie de preguntas. Para cada una de las preguntas, las probabilidades de que A y B den la respuesta correcta son α y β respectivamente, para todas las preguntas. Se consideran todas las preguntas eventos independientes. El juego termina cuando un jugador responde una pregunta correctamente. Defina la probabilidad de que

   el jugador A gane si:

   a) El jugador A responda la primera pregunta.
   b) El jugador B responda la primera pregunta:

   Solución: Primer definimos las condiciones:

   (a) -Ω consiste en el espacio de muestreo que son todas las posibles secuencias infinitas de respuestas
   (b) Evento A determina que el jugador A responda la primera pregunta correctamente
   (c) Evento F el juego termina despues de la primera pregunta
   (d) Evento W el jugador A gana
   (e) Evento A^c determina que el jugador B responda la primera pregunta correctamente

   Se quiere obtener $P(W|A)$ y $P(W|A^c)$ Usando el teorema de la probabilidad total y la particion

   dada por $\{F, F^c\}$

$$P(W|A) = P(W|A\cap F)P(F|A) + P(W|A\cap F^c)P(F^c|A) \tag{13.30}$$

   Ahora tenemos: P(F|A)=P[A responda correctamente la primera pregunta]=$alpha$

$$P(F^c|A) = 1 - alpha, \tag{13.31}$$
$$P(W|A\cap F) = 1 \tag{13.32}$$

   pero $P(W|A\cup F^c) = P(W|A^c)$ entonces tenemos que:

$$P(W|A) = (1 \times \alpha) + (P(W|A^c) \times (1-\alpha)) = \alpha + P(W|A^c)(1-\alpha) \tag{13.33}$$

de forma similar:

$$P(W|A^c) = P(W|A^c \cap F)P(F|A^c) + P(W|A^c \cap F^c)P(F^c|A^c) \tag{13.34}$$

entonces definimos:

$P(F|A^c)$=P[B responda la primer pregunta correctamente]=β $P(F^c|A) = 1-\beta$

pero $P(W|A^c \cap F) = 0$, finalmente $P(W|A^c \cap F^c) = P(W|A)$, entonces tenemos:

$$P(W|A^c) = (0 \times \beta) + (P(W|A) \times (1-\beta)) = P(W|A)(1-\beta) \tag{13.35}$$

Resolviendo la euación 1 y 3 tenemos que:

$$P(W|A) = \frac{\alpha}{1-(1-\alpha)(1-\beta)} \tag{13.36}$$

$$P(W|A^c) = \frac{(1-\beta)\alpha}{1-(1-\alpha)(1-\beta)} \tag{13.37}$$

10. La información de un dispositivo es transmitida como una secuencia de numeros binarios(bits), sin embargo el ruido en el canal interrumpe la señal, en el que un digito transmitido como 0 es recibido como 1 con una probabilidad $1-\alpha$, con una interrución similar cuando el digito 1 es transmitido, se ha observado que en un gran numero de señales transmitidas, los 0s y 1s son transmitidos en el radio 3:4.

    Dado que la secuencia 101 es recibida, cual es la distribucion de probabilidad sobre las señales transmitidas?, asume que la trasmisión y recepción son procesos independientes.

    Solución:

(a) El espacio de muestreo Ω consiste en todas las posibles secuencias de longitud 3 que es:
$\{000, 001, 010, 011, 100, 101, 110, 111\}$
A conjunto correspondiente a eventos de señales
$\{S_0, S_1, S_2, S_3, S_4, S_5, S_6, S_7\}$ y eventos de recepcion
$\{R_0, R_1, R_2, R_3, R_4, R_5, R_6, R_7\}$

Hemos observado el evento R_5 que 101 fue recibido, deseamos obtener $P(S_i)|R_5$ para i=0,1,...7. Sin mebargo en la información dada solo tenemos $P(R_5|S_i)$, primero usando el teorema de la probabilidad total calculamos $P(R_5)$

$$P(R_5) = \sum_{i=0}^{7} P(R_5|S_i)P(S_i) \tag{13.38}$$

Si consideramos $P(R_5|S_0)$, si 000 es transmitido, la probabilidad de que 101 sea recibido es $(1-\alpha) \times \alpha \times (1-\alpha) = \alpha(1-\alpha)^2$, completando la evaluación tenemos:

$$P(R_5|S_0) = \alpha(1-\alpha)^2 \tag{13.39}$$
$$P(R_5|S_1) = \alpha^1(1-\alpha) \tag{13.40}$$
$$P(R_5|S_2) = (1-\alpha)^3 \tag{13.41}$$
$$P(R_5|S_3) = \alpha(1-\alpha)^2 \tag{13.42}$$
$$P(R_5|S_4) = \alpha^2(1-\alpha) \tag{13.43}$$
$$P(R_5|S_5) = \alpha^3 \tag{13.44}$$
$$P(R_5|S_6) = \alpha(1-\alpha)^2 \tag{13.45}$$
$$P(R_5|S_7) = \alpha^2(1-\alpha) \tag{13.46}$$

La información de los digitos transmitidos es la probabilidad de transmitir 1 que es $4/7$, entonces tenemos:

$$P(S_0) = (\frac{3}{7}^3) \tag{13.47}$$

$$P(S_1) = (\frac{4}{7})(\frac{3}{7}^2) \tag{13.48}$$

$$P(S_2) = (\frac{4}{7})(\frac{3}{7}^2) \tag{13.49}$$

$$P(S_3) = (\frac{4}{7}^2)(\frac{3}{7}) \tag{13.50}$$

$$P(S_4) = (\frac{4}{7}(\frac{3}{7}^2) \tag{13.51}$$

$$P(S_5) = (\frac{4}{7}^2)(\frac{3}{7}) \tag{13.52}$$

$$P(S_6) = (\frac{4}{7}^2)(\frac{3}{7}) \tag{13.53}$$

$$P(S_7) = (\frac{4}{7}^3) \tag{13.54}$$

De esta forma tenemos:

$$P(R_5) = \frac{48\alpha^3 + 136\alpha^2(1-\alpha) + 123\alpha(1-\alpha)^2 + 36(1-\alpha)^3}{343} \tag{13.55}$$

Finalmente usando el teorema de bayes tenemos que la probabillidad de distribución sobre: $\{S_0, S_1, S_2, S_3, S_4, S_5, S_6, S_7\}$

$$P(S_i|R_5) = \frac{P(R_5|S_i)P(S_i)}{P(R_5)} \tag{13.56}$$

La probabilidad de recepción de S_5 es:

$$P(S_5|R_5) = \frac{P(R_5|S_5)P(S_5)}{P(R_5)} = \frac{48\alpha^3}{48\alpha^3 + 136\alpha^2(1-\alpha) + 123\alpha(1-\alpha)^2 + 36(1-\alpha)^3} \tag{13.57}$$

11. En una una ciudad hay 51% son hombres y 49% mujeres, un adulto es seleccionado aleatoriamente para una encuesta, encuentra las siguientes dos probabilidades:
a)La encuesta es acerca de las personas que fuman, 9.5% de los hombres fuman mientras que el 1.7% de las mujeres también fuman, encuentra la probabilidad de que una persona entrevistada al azar sea fumador y hombre.

Solución: Definimos:

(a) M=hombres
(b) M^c=mujeres
(c) C=fumadores

(d) C^c=no fumadores

Basados en la información definimos:

(a) $P(M) = 0.51$

(b) $P(M^c) = 0.49$

(c) $P(C|M = 0.095)$ probabilidad de que un hombre tomado al azar sea fumador

(d) $P(C|M^c) = 0.017$ probabilidad de que una mujer seleccionada al azar sea fumadora

Aplicando teorema de Bayes queremos obtener la probabilidad de que la persona seleccionada al azar sea fumador.

$$\frac{P(M)P(C|M)}{(P(M)P(C|M) + (P(M^c)P(C|M^c)} \tag{13.58}$$

$$\frac{0.51 \times 0.095}{(0.51 \times 0.095) + (0.49 \times 0.017)} = 0.85329341 \tag{13.59}$$

12. Tres máquinas denominadas A, B y C, producen un 43%, 26% y 31% de la producción total de una empresa respectivamente, se ha detectado que un 8%, 2% y 1.6% del producto manufacturado por estas máquinas es defectuoso, a. Se selecciona un producto al azar y se encuentra que es defectuoso, ¿cual es la probabilidad de que el producto haya sido fabricado en la máquina B?

Solución:

Se definen los eventos:

(a) D = evento de que el producto seleccionado sea defectuoso

(b) A = evento de que el producto sea fabricado en la máquina A

(c) B = evento de que el producto sea fabricado por la máquina B

(d) C = evento de que el producto sea fabricado por la máquina C

$$P(B|D) = \frac{p(B)p(C|D)}{p(D)} \tag{13.60}$$

$$= \frac{p(B)p(D|B)}{p(A)p(D|A) + p(B)p(D|B) + p(C)p(D|C)} \tag{13.61}$$

$$P(B|D) = \frac{(0.26 * 0.02)}{(0.43 * 0.08 + 0.26 * 0.02 + 0.31 * 0.016)} \tag{13.62}$$

$$= \frac{0.0052}{0.04456} \tag{13.63}$$

$$= 0.116697 \tag{13.64}$$

13. A un paciente de un hospital se le aplica una prueba para la detección de una enfermedad, dicha enfermedad solo afecta al 1% de la población. El resultado es positivo indicando que el paciente tiene la enfermedad. Siendo D el evento que el paciente tenga la enfermedad y T de que el la prueba resulte positiva.

    Suponiendo que la prueba tiene una precisión del 95%, existen diferentes medidas para calcular la precisión, pero en este este problema se considera $P(T|D)$ Y $P(T^c|D^c)$. La probabilidad $P(T|D)$ es conocida como la sensibilidad o verdaderos positivos y $P(T^c|D^c)$ son los falsos negativos.

    Encuentra la probabilidad de que el paciente tenga la enfermedad dada la evidencia proporcionada por el test.

    Solución: Aplicando el teorema de bayes y el teorema de probabilidad completa:

$$P(D|T) = \frac{P(T|D)P(D)}{P(T)} \tag{13.65}$$

$$= \frac{P(T|D)P(D)}{P(D|T)P(D) + P(T|D^c)P(D^c)} \tag{13.66}$$

$$= \frac{0.95 \times 0.01}{0.95 \times 0.01 + 0.05 \times 0.99} \tag{13.67}$$

$$= 0.16 \tag{13.68}$$

14. Servicios de telecomunicaciones. Suponga que es de interés conocer la probabilidad de que un usuario de Telex, que tiene señal de televisión satelital, tenga servicio ilimitado de larga distancia de cobertura nacional. El espacio muestral se reduce automáticamente, se condiciona a la ocurrencia de un evento, se restringe a los usuarios que tienen televisión satelital. Ahora bien, 1 de cada 5 usuarios tiene señal de televisión y 3 de cada 25 tiene tanto señal de televisión como larga distancia nacional, por lo que, por cada 25 usuarios, habrá 5 que tengan señal de televisión y 3 de ellos también tendrán servicio de larga distancia nacional; es decir, por cada 5 usuarios con televisión satelital, habrá 3 con servicio de larga distancia nacional. Entonces, la probabilidad de que un usuario tenga servicio de larga distancia nacional, dado que tiene señal de televisión satelital, es $3/5$. Calcule las siguientes probabilidades de que un usuario de Telex:

    (a) Tenga señal de televisión satelital, dado que tiene Internet de banda ancha.
    (b) Tenga Internet de banda ancha, dado que tiene señal de televisión satelital.
    (c) No tenga Internet, dado que tiene señal de televisión satelital.
    (d) Tenga señal de televisión satelital, dado que no tiene servicio de larga distancia.
    (e) Tenga servicio de larga distancia, dado que tiene señal de televisión.

(f) Tenga larga distancia o señal de televisión, dado que tiene Internet.

Solución:

(a) Tenga señal de televisión satelital, dado que tiene Internet de banda ancha:

$$P(T|I) = \frac{P(I \cap T)}{P(I))} = \frac{0.10}{0.30} = \frac{1}{3} \tag{13.69}$$

(b) Tenga Internet de banda ancha, dado que tiene señal de televisión satelital:

$$P(I|T) = \frac{P(I \cap T)}{P(T))} = \frac{0.10}{0.20} = \frac{1}{2} \tag{13.70}$$

(c) No tenga Internet, dado que tiene señal de televisión satelital:

$$P(\bar{I}|T) = \frac{P(\bar{I} \cap T)}{P(T))} = \frac{0.10}{0.20} = \frac{1}{2} \tag{13.71}$$

(d) Tenga señal de televisión satelital, dado que no tiene servicio de larga distancia:

$$P(T|\bar{L}) = \frac{P(T \cap \bar{L})}{P(\bar{L}))} = \frac{0.00}{0.30} = 0 \tag{13.72}$$

(e) Tenga servicio de larga distancia, dado que tiene señal de televisión:

$$P(L|T) = \frac{P(L \cap T)}{P(T))} = \frac{0.20}{0.20} = 1 \tag{13.73}$$

(f) Tenga larga distancia o señal de televisión, dado que tiene Internet:

$$P(L \cup UI) = P(L|I) + P(T|I) - P(L \cap U|I) \tag{13.74}$$

$$= \frac{P(L \cap I)}{P(I)} + \frac{P(T \cap I)}{P(I)} - \frac{P(L \cap T \cap I)}{P(I)} \tag{13.75}$$

$$= \frac{0.23}{0.30} + \frac{0.10}{0.30} - \frac{0.10}{0.30} \tag{13.76}$$

$$= \frac{0.23}{0.30} \tag{13.77}$$

15. Dado. Considere el experimento consistente en lanzar un dado y observar la cara que queda hacia arriba. Sean los eventos: A = cae par, B = cae 2 o 4 y C = cae 1 o 2; las probabilidades correspondientes son: P(A) = 1/2, P(B) = 1/3 y P(C) = 1/3. Calcularemos las siguientes probabilidades condicionales, utilizando la interpretación clásica o a priori:

Solución:

$$P(A|B) = \frac{N(A \cap B)}{N(B)} = \frac{N(\text{cae 2 o 4})}{N(\text{cae 2 o 4})} = 1 \tag{13.78}$$

$$P(B|A) = \frac{N(A \cap B)}{N(A)} = \frac{N(\text{cae 2 o 4})}{N(\text{cae par})} = \frac{2}{3} \tag{13.79}$$

$$P(A|C) = \frac{N(A \cap C)}{N(C)} = \frac{N(\text{cae 2})}{N(\text{cae 1 o 2})} = \frac{1}{2} \tag{13.80}$$

$$P(C|A) = \frac{N(A \cap C)}{N(A)} = \frac{N(\text{cae 2})}{N(\text{cae par})} = \frac{1}{3} \tag{13.81}$$

16. Considere una urna que contiene 6 bolas rojas, 4 blancas y 5 azules, de la que se extraen sucesivamente de la urna tres bolas, con reemplazo. Sean los eventos:

   - A = {sale bola azul}
   - B = {sale bola blanca}
   - R = {sale bola roja}

   Calcule las siguientes probabilidades:

   (a) Que salgan en el orden roja, azul, blanca.
   (b) Que salgan en el orden azul, roja azul.
   (c) Que salgan tres bolas blancas.
   (d) Que salgan una roja, una azul y una blanca, sin importar el orden.
   (e) Que salgan dos azules y una roja.

   Solución:

   (a) Que salgan en el orden roja, azul, blanca:

$$P(R \cap A \cap B) = P(R) * P(A|R) * P(B|R \cap A) \tag{13.82}$$
$$= \frac{6}{15} * \frac{5}{15} * \frac{4}{15} = \frac{120}{3375} = \frac{8}{225} \tag{13.83}$$

   (b) Que salgan en el orden azul, roja azul:

$$P(A \cap R \cap A) = P(A) * P(R|A) * P(A|A \cap R) \tag{13.84}$$
$$= \frac{5}{15} * \frac{6}{15} * \frac{5}{15} = \frac{150}{3375} = \frac{2}{45} \tag{13.85}$$

   (c) Que salgan tres bolas blancas:

$$P(B \cap B \cap B) = P(B) * P(B|B) * P(B|B \cap B) \tag{13.86}$$
$$= \frac{4}{15} * \frac{4}{15} * \frac{4}{15} = \frac{64}{3375} \tag{13.87}$$

   (d) Que salgan una roja, una azul y una blanca, sin importar el orden:

$$P(\text{1 roja, 1 azul y 1 blanca}) = 6 * P(R \cap A \cap B) \tag{13.88}$$
$$= 6\left(\frac{8}{225}\right) = \frac{16}{75} \tag{13.89}$$

   (e) Que salgan dos azules y una roja:

$$P(\text{2 azules y 1 roja}) = 3\left(\frac{2}{45}\right) = \frac{2}{15} \tag{13.90}$$

17. Tetraedro. Si en lugar de lanzar un dado, se lanza un tetraedro,cuyas caras están numeradas del 1 al 4, defi nimos los eventos: A = {1, 2}, B = {1, 3}, C = {1, 4}, las probabilidades correspondientes son: P(A) = 1/2, P(B) = 1/2, P(C) = 1/2. Ahora, utilizando el criterio de Laplace, calculamos las siguientes probabilidades conjuntas:

Solución:

$$P(A\cap B) = \frac{N(A\cap B)}{N} = \frac{1}{4} \quad (13.91)$$

$$P(A\cap C) = \frac{N(A\cap C)}{N} = \frac{1}{4} \quad (13.92)$$

$$P(B\cap C) = \frac{N(B\cap C)}{N} = \frac{1}{4} \quad (13.93)$$

$$P(A\cap B\cap C) = \frac{N(A\cap B\cap C)}{N} = \frac{1}{4} \quad (13.94)$$

Vemos que los eventos A, B y C son independientes por pares, pues:

$$P(A\cap B) = P(A)P(B) = \frac{1}{2}*\frac{1}{2} = \frac{1}{4} \quad (13.95)$$

$$P(A\cap C) = P(A)P(C) = \frac{1}{2}*\frac{1}{2} = \frac{1}{4} \quad (13.96)$$

$$P(B\cap C) = P(B)P(C) = \frac{1}{2}*\frac{1}{2} = \frac{1}{4} \quad (13.97)$$

$$P(A\cap B\cap C) = \frac{N(A\cap B\cap C)}{N} = \frac{1}{4} \quad (13.98)$$

Sin embargo, A, B y C no son mutuamente independientes, ya que:

$$P(A\cap B\cap C) = P(A)P(B)P(C) = \frac{1}{2}*\frac{1}{2}*\frac{1}{2} = \frac{1}{8} \quad (13.99)$$

18. A y B juegan 12 partidas de ajedrez. A gana seis, B gana cuatro y en dos quedan en tablas. Otro día acuerdan jugar un torneo de 3 partidas. Hallar la probabilidad de que:

    (a) A gane las tres;

    (b) En dos partidas queden en tablas;

    (c) A y B ganen alternadamente;

    (d) B gane al menos una partida.

    Solución:

    Espacio muestral:

    $$\Omega = \{\omega = (\omega_1, \omega_2, \omega_3) : \omega_i \in \{A, B, X\}, i = 1, 2, 3\} \quad (13.100)$$

    Donde: cada $\omega_i (i = 1, 2, 3)$ representa el resultado de la i-ásima partida del torneo resultado:

    - A si gana el jugador A
    - B si gana el jugador B
    - X si quedan tablas

los resultados de las partidas son independientes.
Tambián sabemos:

$$P(A) = \frac{1}{2}\ ,\ P(B) = \frac{1}{3},\ P(X) = \frac{1}{6}$$

(a) Sea M el suceso "A gana las tres partidas". Luego, $M = (A, A, A)$ y la probabilidad es $P(M) = (\frac{1}{2})^3 = \frac{1}{8}$.

(b) Definimos el suceso N = "En dos partidas del torneo A y B quedan en tablas". Se tiene entonces que

$$N = N1 \cup N2 \cup N3,$$

Siendo

$$N_i = \{\omega = (\omega_1, \omega_2, \omega_3) : \omega_i \in \{A, B\}; \omega_j = X, \text{para todo; } j = 1, 2, 3,\ j \neq i\},\ i = 1, 2, 3 \quad (13.101)$$

Nótese que los sucesos elementales que son permutaciones de un mismo conjunto de resultados individuales, tienen la misma probabilidad de ocurrir. Por lo tanto:

$$\begin{aligned} P(N) &= 3 \cdot P(A, X, X) + 3 \cdot P(B, X, X) \\ &= 3 \cdot \frac{1}{2} \cdot \left(\frac{1}{6}\right)^2 + 3 \cdot \frac{1}{3} \cdot \left(\frac{1}{6}\right)^2 = \frac{5}{72} \end{aligned}$$

(c) Sea el suceso S = "A y B ganan alternadamente". Así, como cada torneo consta de 3 partidas, se obtiene que:

$$P(S) = P(A, B, A) + P(B, A, B) = \frac{1}{3} \cdot \left(\frac{1}{6}\right)^2 + 3 + \frac{1}{2}\left(\frac{1}{3}\right)^2 = \frac{5}{36}$$

(d) Representemos por T el suceso "B gana al menos una partida". La probabilidad deseada es:

$$\begin{aligned} &P(T) = 1 - P(T^c) = 1 - P((A, A, A)) - C_{3,1} \cdot \\ &P((A, X, X)) - C_{3,1} \cdot P(A, A, X) - P(X, X, X) \end{aligned} \quad (13.102)$$

$$= 1 - \left(\frac{1}{8} + 3 \cdot \frac{1}{72} + 3 \cdot \frac{1}{24} + \frac{1}{216}\right) = \frac{19}{27} \quad (13.103)$$

19. Se dispone de tres cajas idánticas y cada caja contiene dos cajones. Una caja contiene una moneda de oro en cada cajón, otra contiene una moneda de plata en cada cajón, y la tercera una moneda de oro en un cajón, y una moneda de plata en el otro cajón.

(a) Se selecciona aleatoriamente una caja, ¿cuál es la probabilidad de que la caja seleccionada contenga monedas de diferentes metales?

(b) Se selecciona aleatoriamente una caja, y al abrir aleatoriamente un cajón, nos encontramos con una moneda de oro. ¿Cuál es la probabilidad de que el otro cajón contenga una moneda de plata?

Solución:

Numeremos las cajas: la caja 1 es aquella caja con 2 monedas de oro, la 2 es la que tiene 2 monedas de plata y, finalmente, las monedas de diferente metal (una de oro y otra de plata) se encuentran

en la caja 3.
Sea entonces el espacio muestral

$$\Omega = \{\omega = (\omega_1, \omega_2) : \omega_1 \in \{1,2,3\}; \omega_2 \in \{O, P\}\} \tag{13.104}$$

donde ω_1 representa la caja escogida y ω_2 el metal de la moneda que contiene un cajón de esa caja escogido de forma aleatoria. $P(\{(1,P)\}) = 0$, pero $P(\{(1,O)\}) = \frac{1}{3}$

(a) Sea el suceso A_i = "Se selecciona la caja$i$",$i = 1,2,3$. En nuestro caso, $i = 3$, por lo que el sucesoA_3 es equivalente a "La caja seleccionada tiene monedas de diferentes metales". Sean también los sucesosB_o = "La moneda en el cajón escogido de forma aleatoria en la caja seleccionada es de oro"y B_p el correspondiente a una moneda de plata. Se tiene entonces que:

$$P(A_3) = P(\{(3,O)\}) + P(\{(3,P)\}) \tag{13.105}$$

$$= P(B_o|A_3) \cdot P(A_3) + P(B_p|A_3) \cdot P(A_3) = \tag{13.106}$$

$$2 \cdot \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{3} \tag{13.107}$$

(b) Si uno de los cajones contiene una moneda de oro y el otro una de plata, la caja en la que se encuentran es la B. Por lo tanto, la probabilidad pedida, es:

$$P(A_3|B_o) = \frac{P(B_o|A_3) \cdot P(A_3)}{P(B_o)} = \tag{13.108}$$

$$\frac{P(B_o|A_3) \cdot P(A_3)}{\sum_{i=1}^{3} P(B_o|A_i) \cdot P(A_i)} = \frac{\frac{1}{2} \cdot \frac{1}{3}}{1 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3}} = \frac{1}{3} \tag{13.109}$$

20. Un Doctor dispone de tres equipos electrónicos para realizar ecosonogramas. El uso que le da a cada equipo es de 25% al primero, 35% el segundo en y 40% el tercero. Se sabe que los aparatos tienen probabilidades de error de 1%, 2% y 3% respectivamente. Un paciente busca el resultado de una ecografía y observa que tiene un error. Determine la probabilidad de que se ha usado el primer aparato.

Solución:

Se definen los sucesos:

- Suceso P: seleccionar el primer aparato
- Suceso S: seleccionar el segundo aparato
- Suceso T: seleccionar el tercer aparato
- Suceso E: seleccionar un resultado con error

Se puede observar que la pregunta es sobre determinar la probabilidad de que un examen errado sea del primer aparato, es decir, ya ha ocurrido el error. Por lo tanto, debemos recurrir al teorema de bayes. Claro está, que es necesario de igual forma obtener la probabilidad de que los aparatos produzcan un resultado erróneo, por lo tanto:

$$P(P|E) = \frac{P(P)\cdot P(E|P)}{P(P)\cdot P(E|P)+P(S)\cdot P(E|S)+P(T)\cdot P(E|T)} \quad (13.110)$$

$$P(P|E) = \frac{0.25\cdot 0.01}{0.25\cdot 0.001+0.35\cdot 0.02+0.4\cdot 0.03} = 0.116 \cong 12 = 12\% \quad (13.111)$$

21. Se sabe que el 65% de los accidentes de tráfico que se producen durante la noche de los sábados se deben a la ingesta excesiva de alcohol, el 25% se deben a la imprudencia del conductor (sobrio) y el resto a otras causas, (fallo mecánico...etc.). En estos accidentes, el resultado es nefasto el 30% de las veces en el primer caso, el 20% en el segundo y el 5% en el tercero.
Calcular la probabilidad de que uno de estos accidentes tenga resultado nefasto.
Solución:
A_1 al suceso "tener un accidente por circular con una ingesta excesiva de alcohol" A_2 al suceso "tener un accidente por imprudencia del conductor"y A_3 al suceso "tener un accidente por otras causas". Estos sucesos son incompatibles dos a dos y su unión es el espacio muestral, por lo que se verifican las hipòtesis del teorema de la probabilidad total. Sea N el suceso "tener resultado nefasto"

$$\begin{aligned} P(N) &= P(A_1)\cdot P(\frac{N}{A_1})+P(A_2)\cdot P(\frac{N}{A_2})+P(A_3\cdot P(\frac{N}{A_3}) && (13.112)\\ &= 0.65\cdot 0.3+0.25\cdot 0.2+0.1\cdot 0.05 = 0.25 && (13.113)\end{aligned}$$

22. Se tienen tres monedas cargadas, donde se conoce que la primera tiene una probabilidad de 0.3 de obtenerse cara, la segunda una probabilidad de 0.4 de ocurrir sello y la tercera una probabilidad de 0.4 de salir cara. Un jugador escoge al azar una de las monedas y la lanza dos veces. ¿Cuál es la probabilidad de obtener dos caras?
Solución:

    (a) Experimento aleatorio 1: Escoger al azar de una moneda
        i. Propósito: Determinar cuĺ de las monedas fue escogida
    (b) Experimento aleatorio 2: Lanzamiento por primera vez de la moneda
        i. Propósito: Determinar lo ocurrido en la parte superior de la moneda
    (c) Experimento aleatorio 3: Lanzamiento por segunda vez de la moneda
        i. Propósito: Determinar lo ocurrido en la parte superior de la moneda

Definición de los eventos:

(a) M_i : se seleccionó la moneda i, $i = 1,2,3$

(b) C_j : ocurrió cara en el primer lanzamiento j, $j = 1,2$

Por el teorema de la probabilidad total tenemos:

$$P(C_1 \cap C_2) = P(M_1)P(C_1 \cap C_2)/M_1 + P(M_2)P(C_1 \cap C_2)/M_2 + P(M_3)P(C_1 \cap C_2)/M_3 \quad (13.114)$$

$$= (1/3) \times (0.3)^2 + (1/3) \times (0.6)^2 + (1/3) \times (0.4^2) \approx 0.203 \quad (13.115)$$

23. De los eventos A, B, C y D se tiene la siguiente información:

- A y B son independientes, B y C son independientes también A y C son independientes
- De A, B y C pueden ocurrir a lo sumo 2 de ellos
- $A \cup B \cup C$ y D son mutuamente excluyentes
- A, B y C ocurren cada uno con una probabilidad de r
- El evento D tiene una probabilidad igual a $1/5$

Encuentre el valor de r para el cual la probabilidad de que ocurra al menos uno de los cuatro eventos anteriores sea mx́ima.
Solución:

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) \cdots \quad (13.116)$$

$$\cdots - P(B \cap C) - P(A \cap B \cap C) \quad (13.117)$$

$$= r + r + r - r^2 - r^2 - r^2 - 0 = 3r - 3r^2 \quad (13.118)$$

de lo cual se sigue

$$P(A \cup B \cup C \cup D) = P(A \cup B \cup C) + P(D) = 3r - 3r^2 + \frac{1}{5} \quad (13.119)$$

De esto calcular la probabilidad máxima es sencillo:

$$P(A \cup B \cup C \cup D) = f(r) = 3r - 3r^2 + \frac{1}{5} \quad (13.120)$$

$$\Longrightarrow f'(r) = 3 - 6r = 0 \Longrightarrow r = \frac{1}{2} \quad (13.121)$$

$$\Longrightarrow f''(r) = -6 \Longrightarrow f''(\frac{1}{2}) = -6 < 0 \quad (13.122)$$

por lo tanto para $r = \frac{1}{2}$ la probabilidad de que ocurra al menos uno de los eventos es máxima.

24. La probabilidad de que una persona que se detiene en una gasolinera solicite revisión de neumt́icos es 0.12, la probabilidad de que pida revisión de aceite es 0.29 y la probabilidad de que pida ambas cosas es 0.07. ¿Cuál es la probabilidad de que no solicite la revisión de neumt́icos ni de aceite?
Solución:
Sea "A" el evento de pedir revisión de aceite y "N" el evento de pedir revisión de neumt́icos. Como ambos eventos son compatibles, tenemos:

$$P(N \cup A) = P(N) + P(A) - P(N \cap A) = 0.12 + 0.29 - 0.07 = 0.34 \quad (13.123)$$

La probabilidad de no pedir revisión de neumíticos ni de aceite, es el complemento de pedir cualquiera de estos servicios o ambos. Por lo que tenemos:

$$P(N \cup A)^c = 1 - P(N \cup A) = 1 - 0.34 = 0.66 \tag{13.124}$$

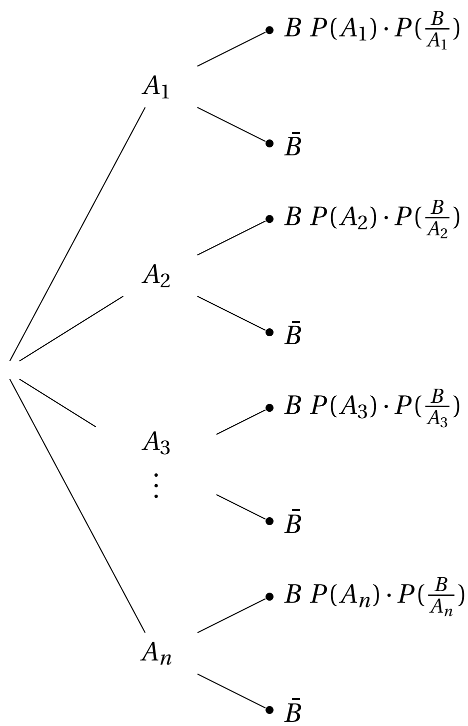
25. El capataz de un grupo de 20 obreros, pide la opinión de dos de ellos seleccionados aleatoriamente sobre las nuevas disposiciones de seguridad en la construcción. Si 12 estń a favor y 8 están en contra, ¿cuál es la probabilidad de que los dos obreros elegidos por el capataz estén en contra?
Solución:
Llamemos "R_c" el evento de que el primer obrero seleccionado esté en contra y "S_c" el evento de que el segundo obrero seleccionado esté en contra. Se trata de eventos condicionales, puesto que lo que realmente se pide es hallar la probabilidad de que el segundo obrero esté en contra, previo cumplimiento de que el primer seleccionado estuvo en contra, es decir, se pide $P(Rc \cap Sc)$. Aplicando el principio de la multiplicación en la probabilidad condicional y teniendo en cuenta que el evento condición es que el primer obrero estuvo en contra, tenemos:

$$P(R_c \cap S_c) = P(R_c) * P(S_c|R_c) = 8/20 * 7/19 = 56/380 \tag{13.125}$$

26. Se sabe por experiencia que de cada 100 clientes que entran a un supermercado, 18 pagan con tarjeta débito. Si se seleccionan 12 clientes aleatoriamente uno tras otro, ¿cuál es la probabilidad de que al menos uno pagará con tarjeta débito?
Solución:
Si de cada 100 clientes 18 pagan con tarjeta débito, entonces la probabilidad de pagar con tarjeta débito es 0.18 y la probabilidad de no pagar con éste documento serél complemento, es decir 0.82. Los ensayos entre cliente y cliente son independientes.
De los 12 clientes, pagarń con tarjeta débito: $0, 1, 2, 3, \ldots$ o 12 clientes y éstos 13 resultados constituyen el espacio muestral, cuya probabilidad es igual a uno, según el segundo axioma de probabilidad. Como se pide calcular la probabilidad de que mínimo uno de los doce pague con tarjeta débito, ya que la probabilidad del espacio muestral es "1", le debemos restar la probabilidad de que "0" (cero) paguen con tarjeta débito. La probabilidad pedida será entonces:

$$P(\text{al menos uno}) = 1 - \underbrace{0.82 \times 0.82 \times \cdots \times 0.82}_{\text{Doce veces}} = 0.9076 \tag{13.126}$$

27. Demuestre que se puede obtener $P(B) = \sum_{i=1}^{n} P(B \mid A_i) \cdot P(A_i)$ la probabilidad total del suceso B(la probabilidad de B) en el siguiente árbol



Solución:

Primero se toman las ramas que nos llevan a B y por definición se sabe que $P(A\cap B) = P(A)\cdot P(B \mid A) = P(B)\cdot P(A \mid B)$ (con probabilidades positivas), por tanto se tiene:

$$A_1 \rightarrow P(A_1\cap B) = P(A_1)\cdot P(B \mid A_1) \tag{13.127}$$
$$A_2 \rightarrow P(A_2\cap B) = P(A_2)\cdot P(B \mid A_2) \tag{13.128}$$
$$A_3 \rightarrow P(A_3\cap B) = P(A_3)\cdot P(B \mid A_3) \tag{13.129}$$
$$\vdots \qquad\qquad \vdots \tag{13.130}$$
$$A_n \rightarrow P(A_n\cap B) = P(A_n)\cdot P(B \mid A_n) \tag{13.131}$$

Ahora sumando las ecuaciones previas y dado que se requiere la probabilidad de B se tiene:

$$P(B) = P(A_1\cap B) + P(A_2\cap B) + P(A_3\cap B) + \ldots + P(A_n\cap B) \tag{13.132}$$

Por último y reduciendo la sumatoria de la ecuación previa, se tiene:

$$P(B) = \sum_{i=1}^{n} P(B \mid A_i)\cdot P(A_i) \tag{13.133}$$

28. Se tienen 3 lotes de lápices dados por A,B,C; y su siguiente margen de defecto esta dado respectivamente por: $\frac{1}{10}$, $\frac{5}{100}$ y $\frac{15}{100}$ La cantidad de cada lote ésta dada por $\frac{36}{100}$, $\frac{60}{100}$ y $\frac{4}{100}$ respectivamente. ¿Cuál es la probabilidad de que al tomar un lápiz de cualquiera de los lotes, éste se encuentre defectuoso?

Solución:

Se definen los sucesos:
Suceso A: seleccionar el lote A
Suceso B: seleccionar el lote B
Suceso C: seleccionar el lote C
Suceso D: seleccionar un resultado con defecto

Se sabe que:

$$P(D) = \sum_{i=1}^{n} P(D \mid A_i) \cdot P(A_i) \tag{13.134}$$

Por tanto se tiene:

$$P(D) = P(D \mid A_1) \cdot P(A_1) + P(D \mid A_2) \cdot P(A_2) + P(D \mid A_3) \cdot P(A_3) \tag{13.135}$$

Ahora representando los sucesos se tiene:

$$P(D) = P(D \mid A) \cdot P(A) + P(D \mid B) \cdot P(B) + P(D \mid C) \cdot P(C) \tag{13.136}$$

Sustituyendo valores se obtiene:

$$P(D) = \frac{1}{10} \cdot \frac{36}{100} + \frac{5}{100} \cdot \frac{60}{100} + \frac{15}{100} \cdot \frac{4}{100} \tag{13.137}$$

$$P(D) = \frac{36}{1000} + \frac{300}{10000} + \frac{60}{10000} = \frac{36}{1000} + \frac{36}{1000} = \frac{72}{1000} = 0.072 \tag{13.138}$$

29. Demuestre que:

$$P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{\sum_{j=1}^{n} P(B \mid A_j) \cdot P(A_j)} \tag{13.139}$$

Solución:

Por definición se sabe que $P(A \cap B) = P(A) \cdot P(B \mid A) = P(B) \cdot P(A \mid B)$ por tanto:

Utilizando sólamente a $P(A_i \cap B) = P(B) \cdot P(A_i \mid B)$ se despeja $P(A_i \mid B)$ y se obtiene:

$$P(A_i \mid B) = \frac{P(A_i \cap B)}{P(B)} \tag{13.140}$$

Ahora para la parte del númerador de la ecuación 15 se tiene que:

$$P(A_i \cap B) = P(A_i) \cdot P(B \mid A_i) \tag{13.141}$$

Para la parte del denominador de la ecuación 15 se tiene que:

$$P(B) = P(A_1 \cap B) + P(A_2 \cap B) + P(A_3 \cap B) + \ldots + P(A_n \cap B) \tag{13.142}$$

$$P(B) = \sum_{j=1}^{n} P(B \mid A_j) \cdot P(A_j) \tag{13.143}$$

Por último uniendo el resultado de las ecuaciones anteriores se concluye que:

$$P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{\sum_{j=1}^{n} P(B \mid A_j) \cdot P(A_j)} \tag{13.144}$$

30. Un paciente tiene una prueba para una enfermedad llamada condicionitis, una condición médica que afecta al 1% de la población. La prueba resultante es positiva, la prueba muestra que el paciente tiene la enfermedad, Dejemos que D sea el evento de que el paciente tenga la enfermedad y T el evento de que la prueba sea positiva.

    Suponiendo que la prueba tiene 95% de presición; hay diferentes mediciones de la presición de una prueba, pero en este problema se asume que $P(T|D) = 0.95$ y que $P(T^c|D^c) = 0.95$. La cantidad $P(T|D)$ es conocida como la sensibilidad o tasa de verdadero positivo de la prueba y $P(T^c|D^c)$ es conocida como la especifidad o tasa de verdadero negativo.

    Encontrar la probabilidad condicional de que el paciente tiene condicionitis, dada la evidencia probada por el resultado de la prueba

    Solución:

    Aplicando el teorema de Bayes y la ley de la probabilidad total se tiene:

$$P(D|T) = \frac{P(T|D) \cdot P(D)}{P(T)} \tag{13.145}$$

$$P(D|T) = \frac{P(T|D) \cdot P(D)}{P(T|D) \cdot P(D) + P(T|D^c) \cdot P(D^c)} \tag{13.146}$$

$$P(D|T) = \frac{0.95 \cdot 0.01}{0.95 \cdot 0.01 + 0.05 \cdot 0.99} \tag{13.147}$$

$$P(D|T) \approx 0.16 \tag{13.148}$$

Existe un 16% de oportunidad de que el paciente tenga condicionitis, dado que su prueba es positiva

14 Maximum Likelihood

14.1 Maximum Likelihood

En estadística, la estimación por máxima verosimilitud (conocida también como EMV y, en ocasiones, MLE por sus siglas en inglés) es un método habitual para ajustar un modelo y estimar sus parámetros.

Supóngase que se tiene una muestra $x1, x2, \ldots, xn$ de n observaciones independientes e idénticamente distribuidas extraídas de una función de distribución desconocida con función de densidad (o función de probabilidad) $f_0(\cdot)$. Se sabe, sin embargo, que f_0 pertenece a una familia de distribuciones $f(\cdot \mid \theta)$, llamada modelo paramétrico, de manera que f0 corresponde a $\theta = \theta_0$, que es el verdadero valor del parámetro. Se desea encontrar el valor $\hat{\theta}$ (o estimador) que esté lo más próximo posible al verdadero valor θ_0.
Tanto x_i como θ pueden ser vectores.

La idea de este método es el de encontrar primero la función de densidad conjunta de todas las observaciones, que bajo condiciones de independencia, es:

$$f(x_1, x_2, \ldots, x_n \mid \theta) = f(x_1|\theta) \cdot f(x_2|\theta) \cdots f(x_n|\theta) \tag{14.1}$$

Observando esta función bajo un ángulo ligeramente distinto, se puede suponer que los valores observados $x_1, x_2, \ldots, x_n$ son fijos mientras que θ puede variar libremente. Esta es la función de verosimilitud:

$$\mathcal{L}(\theta \mid x_1, \ldots, x_n) = \prod_{i=1}^{n} f(x_i|\theta). \tag{14.2}$$

En la práctica, se suele utilizar el logaritmo de esta función:

$$\hat{\ell}(\theta \mid x_1, \ldots, x_n) = \ln \mathcal{L} = \sum_{i=1}^{n} \ln f(x_i|\theta). \tag{14.3}$$

El método de la máxima verosimilitud estima θ_0 buscando el valor de θ que maximiza $\hat{\ell}(\theta|x)$. Este es el llamado estimador de máxima verosimilitud (MLE) de θ_0:

$$\hat{\theta}_{\text{mle}} = \underset{\theta\in\Theta}{\arg\max}\, \hat{\ell}(\theta \mid x_1, \ldots, x_n). \tag{14.4}$$

En ocasiones este estimador es una función explícita de los datos observados $x_1, \ldots, x_n$, pero muchas veces hay que recurrir a optimizaciones numéricas. También puede ocurrir que el máximo no sea único o no exista.

En la exposición anterior se ha asumido la independencia de las observaciones, pero no es un requisito necesario: basta con poder construir la función de probabilidad conjunta de los datos para poder aplicar el método. Un contexto en el que esto es habitual es el del análisis de series temporales.

Ejemplo 14.1

Distribución uniforme discreta:
Supóngase que n bolas numeradas de 1 a n se colocan en una urna y que una de ellas se extrae al azar. Si se desconoce n, su EMV es el número m que aparece en la bola extraída: la función de verosimilitud es 0 para $n < m$ y $\frac{1}{n}$ para $n \geq m$; que alcanza su máximo cuando $n = m$. La esperanza matemática de $\hat{n}$, es $\frac{(n+1)}{2}$. Como consecuencia, el EMV de n infravalorará el verdadero valor de n por $\frac{(n?1)}{2}$.

Distribución discreta con parámetros discretos
Supóngase que se lanza una moneda sesgada al aire 80 veces. La muestra resultante puede ser algo así como $x_1 = H, x_2 = T, ..., x_{80} = T$, y se cuenta el número de caras, "H". La probabilidad de que salga cara es p y la de que salga cruz, $1 - p$ (de modo que p es el parámetro θ). Supóngase que se obtienen 49 caras y 31 cruces. Imagínese que la moneda se extrajo de una caja que contenía tres de ellas y que éstas tienen probabilidades p iguales a 1/3, 1/2 y 2/3 aunque no se sabe cuál de ellas es cuál.

A partir de los datos obtenidos del experimento se puede obtener saber cuál es la moneda con la máxima verosimilitud. Usando la función de probabilidad de la distribución binomial con una muestra de tamaño 80, número de éxitos igual a 49 y distintos valores de p, la función de verosimilitud toma tres valores siguientes:

$$\Pr(\mathrm{H} = 49 \mid p = 1/3) = \binom{80}{49}(1/3)^{49}(1-1/3)^{31} \approx 0.000,$$

$$\Pr(\mathrm{H} = 49 \mid p = 1/2) = \binom{80}{49}(1/2)^{49}(1-1/2)^{31} \approx 0.012,$$

$$\Pr(\mathrm{H} = 49 \mid p = 2/3) = \binom{80}{49}(2/3)^{49}(1-2/3)^{31} \approx 0.054.$$

La verosimilitud es máxima cuando p = 2/3 y éste es, por lo tanto, el EMV de p.

14.2 PCA vs ICA

PCA: En estadística, el análisis de componentes principales (en español ACP, en inglés, PCA) es una técnica utilizada para reducir la dimensionalidad de un conjunto de datos.

Técnicamente, el ACP busca la proyección según la cual los datos queden mejor representados en términos de mínimos cuadrados. Esta convierte un conjunto de observaciones de variables posiblemente correlacionadas en un conjunto de valores de variables sin correlación lineal llamadas componentes principales.

El ACP se emplea sobre todo en análisis exploratorio de datos y para construir modelos predictivos. El ACP comporta el cálculo de la descomposición en autovalores de la matriz de covarianza, normalmente tras centrar los datos en la media de cada atributo.

Debe diferenciarse del análisis factorial con el que tiene similaridades formales y en el cual puede ser utilizado como un método de aproximación para la extracción de factores.

El ACP construye una transformación lineal que escoge un nuevo sistema de coordenadas para el conjunto original de datos en el cual la varianza de mayor tamaño del conjunto de datos es capturada en el primer eje (llamado el Primer Componente Principal), la segunda varianza más grande es el segundo eje, y así sucesivamente. Para construir esta transformación lineal debe construirse primero la matriz de covarianza o matriz de coeficientes de correlación. Debido a la simetría de esta matriz existe una base completa de vectores propios de la misma. La transformación que lleva de las antiguas coordenadas a las coordenadas de la nueva base es precisamente la transformación lineal necesaria para reducir la dimensionalidad de datos. Además las coordenadas en la nueva base dan la composición en factores subyacentes de los datos iniciales.

Una de las ventajas del ACP para reducir la dimensionalidad de un grupo de datos, es que retiene aquellas características del conjunto de datos que contribuyen más a su varianza, manteniendo un orden de bajo nivel de los componentes principales e ignorando los de alto nivel. El objetivo es que esos componentes de bajo orden a veces contienen el aspecto "más importante" de esa información.

ICA: El Análisis de Componentes Independientes (ACI) (en inglés ICA) es un método computacional que sirve para separar una señal multivariante en subcomponentes aditivos suponiendo que la señal de origen tiene una independencia estadística y es no-Gausiana. éste es un caso especial de separación ciega de las señales.

El análisis de componentes independientes (ACI) es una generalización del análisis de componentes principales (ACP), en ambos casos se practica una transformación lineal de los datos originales, aunque la diferencia básica es que el ACI no requiere que las variables originales tengan una distribución gausiana.

ICA se encuentra muy relacionado al problema de la separación ciega de fuentes (inglés Blind Signal Separation, BSS) el cual consiste en determinar a través de un arreglo de transductores las señales de las fuentes originales que intervienen en una mezcla, sin información alguna de las señales originales ni de las ponderaciones de la mezcla. éste es un problema clásico de procesamiento de señales.

Asumir la independencia es correcto en la mayoría de los casos; por tanto la separación ciega por ACI de una señal mezclada da muy buenos resultados. Un problema clásico (enunciado por primera vez en 1985 por Christian Jutten y Jeanny Hérault) consiste en un uso sencillo de ACI es el problema de la fiesta, donde la señal que consiste en una mezcla de voces de gente hablando en la misma habitación es separada. Normalmente el problema se simplifica asumiendo que no hay retrasos ni ecos. Una cosa importante a considerar es que si hay N fuentes (voces) en la habitación, al menos hacen falta N observaciones (micrófonos) para obtener la señal original.

ICA tiene un enfoque estadístico y parte de la hipótesis de que las señales originales (fuentes) son estadísticamente independientes y los procedimientos que lo siguen se basan en propiedades de las fuentes y de las observaciones (mezclas), caracterizadas por sus distribuciones de probabilidad. Se pretende que si las señales originales son estadísticamente independientes, las señales recuperadas también deben de serlo. Por lo tanto, los métodos relacionados con ICA consideran como condición primordial la independencia estadística. Buscando dentro de espacios multidimensionales las componentes principales (en la dirección de máxima varianza) de las variables a determinar. El problema cosiste en encontrar un transformación que encuentre la inversa de la matriz con la cual se mezcló.

15 Programas

15.1 Cuadro mágico

```
#include <stdio.h>
#include <math.h>
#include<graphics.h>
int
main (void)
{
int negui;
int result;
int bere;
int newx;

int icontador=0;
int x=0;
int y=0;
int vx=0;
int vy=0;

  system("clear");
  printf ("\nWelcome to my magic square\n");
  printf ("\tPlease, enter the number of the square... ");
  scanf ("%d", &negui);
  result=negui%2;
if ( result == 0 ) {
  printf ("The square is an even number, so...\n");
  printf ("\t Please retry and enter an odd number\n");
 }
else
 {
   if (negui >2 )
    {
    printf ("\tThanks you have chose a correct square...\n");
```

```
 bere=pow(negui,3);
 bere=(bere+negui)/2;
 printf ("\nThis square magic the sum will be = %d\n",bere);
     int Matriz[negui][negui];
// Variable a utilizar como contador de números
int icontador=0;
// Número de números que hay que contar y limites de la matriz
int limite=negui*negui;
int limitex = negui - 1;
int limitey = negui - 1;

// Inicializar los valores de la matriz a 0
for( x=0 ; x<negui ; x++)
{
   for( y=0 ; y<negui ; y++ )
    {
        Matriz[x][y] = 0;
    }
}
// Establecer la posición x e y inicial de la matriz donde almacenar
// el primer numero siendo en la primera fila posición central
x = negui / 2;
y = 0;
for(icontador=1;icontador<=limite;icontador++)
{
   Matriz[y][x] = icontador;
   vx = x;
   vy = y;
   x++;
   if( x > limitex ) x = 0;
   y--;
   if( y < 0 ) y = limitey;
   if (Matriz[y][x]>0)
   {
      x = vx ;
      y = vy +1 ;
   }
}

// Visualizar la matriz
printf("\nMagic Square \n\n\n");
for( x=0 ; x<negui ; x++)
{
    for( y=0 ; y<negui ; y++ )
     {
```

```
            printf("|%3d|",Matriz[x][y]);
        }
        printf("\n");
    }

    printf ("\n\n");
        int gd = DETECT,gm,left=50,top=50;
        initgraph(&gd,&gm,NULL);
int y;
     for( y=0 ; y<=negui ; y++ )
      {
          newx=negui*35;
          line(left, top, left + newx, top);
          line(top, left, top, left + newx);
          top=top+35;
      }
        delay(5000);
        closegraph();
    }
   else
    {
     printf ("\nThe square magic cannot be one, please retry\n\n");
    }
 }
    return 0;
}
```

15.2 Fractal Sierpinski

```
#include <stdio.h>
#include <math.h>
#include<graphics.h>

int ramdon(void);

int main (void)
{
int negui;
int puntos;
int randx;
int randy;
int randside;
int AX=10;
int AY=10;
int BX=10;
int BY=470;
```

```
int CX=610;
int CY=470;

  system("clear");
  srand( time( NULL ) );
  printf ("\nWelcome to Sierpinski fractal\n");
  printf ("\tPlease, enter the number of points... ");
  scanf ("%d", &negui);

  int gd = DETECT,gm;
  initgraph(&gd,&gm,NULL);
  int maxx=getmaxx();
  int maxy=getmaxy();

  putpixel(AX, AY, 3);
  putpixel(BX, BY, 3);
  putpixel(CX, CY, 3);

  randx = rand() % maxx;
  randy = rand() % maxy;

  for (puntos=0;puntos<negui;puntos++){
    delay(1);
    randside = rand() %9;

    if (randside==0 || randside==1 || randside==2){
       randx=(randx+AX)/2;
       randy=(randy+AY)/2;
       putpixel(randx,randy,2);

    }
    if (randside==3 || randside==4 || randside==5){
       randx=(randx+BX)/2;
       randy=(randy+BY)/2;
       putpixel(randx,randy,2);
    }
    if (randside==6 || randside==7 || randside==8){
       randx=(randx+CX)/2;
       randy=(randy+CY)/2;
       putpixel(randx,randy,2);
    }
  }
delay(5000);
closegraph();
```

```
 return 0;
}
```

15.3 Poligonos

```
#include <stdio.h>
#include <math.h>
#include<graphics.h>

#define PI 3.141592653589
#define CIRCLE          1
#define PENDULUM        2
#define POLYGON         3
#define POLYGON_ROTATING 4
#define POLYGON_ONE_POINT 5
#define SALIR           7

int main (void)
{
int negui;
int j,i,g;
int x,y;
int xc, yc;
double r,a;
int opcion;
system("clear");
int gd = DETECT,gm;
initgraph(&gd,&gm,NULL);
int maxx=getmaxx();
int maxy=getmaxy();
      while (1){
      printf ("\n Welcome to POLYGONO program :\n ");
      printf ("\n Presiona 1.- Circle running...");
      printf ("\n        2.- Pendulum");
      printf ("\n        3.- Poliygon");
      printf ("\n        4.- Poliygon Rotating ");
      printf ("\n        5.- Poliygon Rotating in one point ");
      printf ("\n        4.- SALIR -> ");
      scanf ("%d",&opcion);

      switch (opcion){
            case CIRCLE :
            {
            for (i=0;i<maxx + 40;i++){
                  cleardevice();
                  setcolor(0);
```

```
            outtextxy(100, 25, "Running circle...");
            setcolor(3);
            line(0, maxy/2 + 40, maxx, maxy/2 + 40);
            setcolor(2);
            circle(0+i, maxy/2,40);
            delay(7);
      }
      getch();
      cleardevice();
      break;
      }
      case PENDULUM :
      {
      for (i=0;i<=200;i++){
            cleardevice();
            setcolor(0);
            outtextxy(100, 15, "Pendulum circle...");
            setcolor(3);
            circle(maxx/2, 40,5);
            circle(maxx/2 +i +1, maxy/2 - i,20);
            delay(10);
      }
      getch();
      cleardevice();
      break;
      }
      case POLYGON :
      {
      printf ("\n\nType the number of sides: ");
      scanf ("%d",&negui);
      cleardevice();
      int maxx=getmaxx();
      int maxy=getmaxy();
      float px,py,px1,py1,ang;
      float edge=200;
      float dist_edge;
      float x0,y0;
      float gir=0;

      //Dibujando coordenadas X,Y
      setcolor(13); //7
      line(0, maxy/2, maxx, maxy/2);
      line(maxx/2, 0, maxx/2, maxy);
      setcolor(5);
      circle(maxx/2, maxy/2,200);
```

```
dist_edge=2*M_PI/negui;
x0=maxx/2;
y0=maxy/2;
if (negui%2==0)
{
px1=edge*(cos((M_PI/2)+(dist_edge/2)+gir));
       if(px1<0)
              {px1=(maxx/2)-(px1*(-1));}
       else
              {px1=maxx/2+px1;}
py1=edge*(sin((M_PI/2)+(dist_edge/2)+gir));
       if(py1<0)
              {py1=(maxy/2)+(py1*(-1));}
       else
              {py1=y0-py1;}
px=px1;
py=py1;
ang=(M_PI/2)+dist_edge+(dist_edge/2)+gir;
}
else{
px1=edge*(cos((M_PI/2)+gir));
       if(px1<0)
              {px1=(maxx/2)-(px1*(-1));}
       else
              {px1=maxx/2+px1;}
py1=edge*(sin((M_PI/2)+gir));
       if(py1<0)
              {py1=(maxy/2)+(py1*(-1));}
       else
              {py1=y0-py1;}
px=px1;
py=py1;
ang=(M_PI/2)+dist_edge+gir;
}
setcolor(2);
for (i=0;i<negui;i++)
{
       px1=edge*(cos(ang));
       if(px1<0)
              {px1=(maxx/2)-(px1*(-1));}
       else
              {px1=maxx/2+px1;}
       py1=edge*(sin(ang));
       if(py1<0)
              {py1=(maxy/2)+(py1*(-1));}
```

```
            else
                    {py1=y0-py1;}
    line(px,py,px1,py1);
    px=px1;
    py=py1;
    ang=ang+dist_edge;
    }
    getch();
    cleardevice();
    break;
    }
    case POLYGON_ROTATING :
    {
    printf ("\n\n Type the number of sides: ");
    scanf ("%d",&negui);
    cleardevice();
    int maxx=getmaxx();
    int maxy=getmaxy();
    float px,py,px1,py1,ang;
    float edge=200;
    float dist_edge;
    float x0,y0;
    float gir=0;
    //Dibujando coordenadas X,Y
    setcolor(13); //7
    line(0, maxy/2, maxx, maxy/2);
    line(maxx/2, 0, maxx/2, maxy);
    setcolor(5); //10
    circle(maxx/2, maxy/2,200);
    dist_edge=2*M_PI/negui;
    x0=maxx/2;
    y0=maxy/2;
    for(g=0;gir<100; g++)
    {
    delay(100);
    cleardevice();
    setcolor(13);
    line(0, maxy/2, maxx, maxy/2);
    line(maxx/2, 0, maxx/2, maxy);
    setcolor(5); //10
    circle(maxx/2, maxy/2,200);

            if (negui%2==0)
            {
            px1=edge*(cos((M_PI/2)+(dist_edge/2)+gir));
```

```
        if(px1<0)
                {px1=(maxx/2)-(px1*(-1));}
        else
                {px1=maxx/2+px1;}
py1=edge*(sin((M_PI/2)+(dist_edge/2)+gir));
        if(py1<0)
                {py1=(maxy/2)+(py1*(-1));}
        else
                {py1=y0-py1;}
px=px1;
py=py1;
ang=(M_PI/2)+dist_edge+(dist_edge/2)+gir;
}
else{
px1=edge*(cos((M_PI/2)+gir));
        if(px1<0)
                {px1=(maxx/2)-(px1*(-1));}
        else
                {px1=maxx/2+px1;}
py1=edge*(sin((M_PI/2)+gir));
        if(py1<0)
                {py1=(maxy/2)+(py1*(-1));}
        else
                {py1=y0-py1;}
px=px1;
py=py1;
ang=(M_PI/2)+dist_edge+gir;
}
setcolor(2);
for (i=0;i<negui;i++)
{
        px1=edge*(cos(ang));
        if(px1<0)
                {px1=(maxx/2)-(px1*(-1));}
        else
                {px1=maxx/2+px1;}
        py1=edge*(sin(ang));
        if(py1<0)
                {py1=(maxy/2)+(py1*(-1));}
        else
                {py1=y0-py1;}
line(px,py,px1,py1);
px=px1;
py=py1;
ang=ang+dist_edge;
```

```
                }
        gir=g*1;
        }
        getch();
        cleardevice();
        break;
        }
        case POLYGON_ONE_POINT :
        {
        printf ("\n\n Type the number of sides: ");
        scanf ("%d",&negui);
        cleardevice();
        int maxx=getmaxx();
        int maxy=getmaxy();
        float px,py,px1,py1,pxtmp,pytmp,ang,inc;
        float edge=80;
        float dist_edge;
        float x0,y0;
        float gir=0;
        dist_edge=2*M_PI/negui;
        x0=maxx/2;
        y0=maxy/2;
        delay(100);
                if (negui%2==0)
                {
                px1=edge*(cos((M_PI/2)+(dist_edge/2)+gir));
                        if(px1<0)
                                {px1=(maxx/2)-(px1*(-1));}
                        else
                                {px1=maxx/2+px1;}
                py1=edge*(sin((M_PI/2)+(dist_edge/2)+gir));
                        if(py1<0)
                                {py1=(maxy/2)+(py1*(-1));}
                        else
                                {py1=y0-py1;}
                px=px1;
                py=py1;
                ang=(M_PI/2)+dist_edge+(dist_edge/2)+gir;
                }
                else{
                px1=edge*(cos((M_PI/2)+gir));
                        if(px1<0)
                                {px1=(maxx/2)-(px1*(-1));}
                        else
                                {px1=maxx/2+px1;}
```

```
py1=edge*(sin((M_PI/2)+gir));
        if(py1<0)
                {py1=(maxy/2)+(py1*(-1));}
        else
                {py1=y0-py1;}
px=px1;
py=py1;
ang=(M_PI/2)+dist_edge+gir;
}
pxtmp=px1;
pytmp=py1;
setcolor(2);
float radius=sqrt ((x0 - px1) * (x0 - px1) + (y0 - py1) * (y0 - p
for (i=0;i<negui;i++)
{
        px1=edge*(cos(ang));
        if(px1<0)
                {px1=(maxx/2)-(px1*(-1));}
        else
                {px1=maxx/2+px1;}
        py1=edge*(sin(ang));
        if(py1<0)
                {py1=(maxy/2)+(py1*(-1));}
        else
                {py1=y0-py1;}
line(px,py,px1,py1);
px=px1;
py=py1;
ang=ang+dist_edge;
}
setcolor(10);
px1=pxtmp;
py1=pytmp;
x0=(maxx/2) - (radius/2);
y0=(maxy/2) - (radius/2);
ang=(M_PI/2)+dist_edge;
for (i=0;i<negui;i++)
{
        px1=edge*(cos(ang));
        if(px1<0)
                {px1=(maxx/2)-(px1*(-1));}
        else
                {px1=maxx/2+px1;}
        py1=edge*(sin(ang));
        if(py1<0)
```

```
                              {py1=(maxy/2)+(py1*(-1));}
                        else
                              {py1=y0-py1;}
                  line(px,py,px1,py1);
                  px=px1;
                  py=py1;
                  ang=ang+dist_edge;
                  }
            getch();
            cleardevice();
            break;
            }
            case SALIR :
            {
            cleardevice();
            printf ("Adios");
            getch();
            exit(0);
            break;
            }
            default :
             printf ("\n\nElije una Opcion Corecta");
      }
      }
getch();
closegraph();
return 0;
}
```

15.4 Bayes Probabilidad

```
#include <stdio.h>
//#include <math.h>
#include<graphics.h>


int ramdon(void);

int main (void)
{
int one_red;
int one_black;
int one_bag=0;
int two_red;
int two_black;
int two_bag=0;
```

```
int i;
int num_execution=0;
int total_probability=0;
float percent_total=0;
int rand_one=0;
int rand_two=0;
float tmp_rand_two=0;

  srand( time( NULL ) );

  system("clear");
  printf ("\nWelcome to Bayes probability\n");
  printf ("\t This program will calculate the probability to get two red balls\n");
  printf ("\t getting one red from the bag one and the other one from the bag two.\
  printf ("Bag one...\n");
  printf ("\tPlease, enter the number of red balls... ");
  scanf ("%d", &one_red);
  printf ("\tPlease, enter the number of black balls... ");
  scanf ("%d", &one_black);

  printf ("\nBag two...\n");
  printf ("\tPlease, enter the number of red balls... ");
  scanf ("%d", &two_red);
  printf ("\tPlease, enter the number of black balls... ");
  scanf ("%d", &two_black);

  printf ("\n How many time do you want to repit the experiment?...");
  scanf ("%d", &num_execution);
  printf ("\nBag one - Balls red %d Balls black %d \n", one_red, one_black);
  printf ("Bag two - Balls red %d Balls black %d\n\n", two_red, two_black);

  one_bag= one_red + one_black;
  two_bag= two_red + two_black;



  // Ramdon de la primera bolsa
  for (i=0;i<num_execution;i++){
    rand_one=0;
    rand_two=0;
    rand_one = rand() % one_bag;
    rand_one = rand_one +1;

    if (rand_one <= one_red){
     rand_one =2;
```

```
   two_red = two_red +1;
   printf ("ID %d - Ball in bag one is: %d \n",i+1, rand_one);
  }

  if (rand_one > one_red){
   rand_one =1;
   two_black=two_black +1;
   printf ("ID %d - Ball in bag one is: %d \n",i+1, rand_one);
  }

 // Ramdon de la segunda bolsa
  rand_two = rand() % two_bag;
  tmp_rand_two=rand_two;
  rand_two = rand_two +1;

  if (rand_two <= two_red){
   rand_two =2;
   printf ("ID %d - Ball in bag two is: %d \n\n",i+1, rand_one);
  }

  if (rand_two > two_red){
   rand_two =1;
   printf ("ID %d - Ball in bag two is: %d \n\n",i+1, rand_two);
  }


  if (rand_one==2 && rand_two==2){
    total_probability ++;
  }

 printf ("\n El valor de rand2 es %f \n", tmp_rand_two/10);

 }

percent_total=(total_probability * 100)/num_execution;

printf("\n The total of matches is:... %d of %d, which is:.. %.2f %% \n\n",total_pr

return 0;
}
```

15.5 Juego de la vida

```
#include <stdio.h>
#include <math.h>
#include<graphics.h>
```

```
int ramdon(void);

int main (void)
{
int negui;
int i;
int j;
int x=20;
int y=20;
int datos[x][y];
int on_off;
int rand_x;
int rand_y;
int sum=0;

  system("clear");
  srand( time( NULL ) );
  printf ("\nWelcome to The Game of the Life\n");
  printf ("\tPlease, enter the number of interactions... ");
  scanf ("%d", &negui);

  int gd = DETECT,gm;
  initgraph(&gd,&gm,NULL);
  int maxx=getmaxx();
  int maxy=getmaxy();

//This for will set all the matrix with 0
for (i=0;i<x;i++){
  for(j=0;j<y;j++){
    datos[i][j]=0;
  }
}

//This for the first eight seeds
for (i=0;i<10;i++){
 rand_x=rand() % x;
 rand_y=rand() % y;
 datos[rand_x][rand_y]=1;
}

//This will generate the random seeds
while (negui!=0){
  rand_x=rand() % x;
  rand_y=rand() % y;
```

```
  sum=sum+datos[rand_x-1][rand_y-1]+datos[rand_x][rand_y-1]+datos[rand_x+1][rand_y
  sum=sum+datos[rand_x-1][rand_y]+datos[rand_x+1][rand_y];
  sum=sum+datos[rand_x-1][rand_y+1]+datos[rand_x][rand_y+1]+datos[rand_x+1][rand_y+
  on_off = rand() % 2;
  if (sum==3 || sum==2){
    datos[rand_x][rand_y]=1;
    setcolor(2);
    bar(20*rand_y, 20*rand_x, 20*(rand_y+1), 20*(rand_x+1));
  }

  if (sum==1 || sum > 3){
    datos[rand_x][rand_y]=0;
    setcolor(0);
    bar(20*rand_y, 20*rand_x, 20*(rand_y+1), 20*(rand_x+1));
  }
  negui--;
  sum=0;
  delay(1);
}


delay(400);
closegraph();

  return 0;
}
```

15.6 Fractar de Koch

```
#/usr/bin/env python3
# -*- coding: utf-8 -*-
from math import sin, cos, pi, atan2
from pylab import *

def copoVonKoch(lado, n):
    x_vertice1 = 0
    y_vertice1 = 0

    x_vertice2 = lado * cos(2 * pi / 3)
    y_vertice2 = lado * sin(2 * pi / 3)

    x_vertice3 = lado * cos(pi / 3)
    y_vertice3 = lado * sin(pi / 3)

    curvaVonKoch(x_vertice1, y_vertice1, x_vertice2, y_vertice2, n)
    curvaVonKoch(x_vertice2, y_vertice2, x_vertice3, y_vertice3, n)
```

```
   curvaVonKoch(x_vertice3, y_vertice3, x_vertice1, y_vertice1, n)

def curvaVonKoch(xi, yi, xf, yf, n):
   if n == 0:
      plot([xi, xf], [yi, yf], lw=1.0, color='b')
   elif n > 0:
      x1 = xi + (xf - xi) / 3.0
      y1 = yi + (yf - yi) / 3.0

      x3 = xf - (xf - xi) / 3.0
      y3 = yf - (yf - yi) / 3.0

      radio = hypot(x3 - x1, y3 - y1)
      alpha = atan2((y3 - y1), (x3 - x1))
      alpha += pi / 3.0
      x2 = x1 + radio * cos(alpha)
      y2 = y1 + radio * sin(alpha)

      curvaVonKoch(xi, yi, x1, y1, n - 1)
      curvaVonKoch(x1, y1, x2, y2, n - 1)
      curvaVonKoch(x2, y2, x3, y3, n - 1)
      curvaVonKoch(x3, y3, xf, yf, n - 1)

def dibuja(lado, n):
   axes().set_xlim(0, lado)
   axes().set_ylim(-2, lado / 2.0)
   axes().set_aspect(1.0)
   title('Curva De Von Koch')
   xlim(0, lado)
   curvaVonKoch(0, 0, lado, 0, n)
   show()
   axes().set_xlim(-lado, lado)
   axes().set_ylim(-2, lado + 0.5 * lado)
   axes().set_aspect(1.0)
   title('Copo De Nieve De Koch')
   xlim(-lado, lado)
   copoVonKoch(lado, n)
   show()

def principal():
   print('Fractal de Koch.\n')
   print('-------------------------------------')
   n=7
   lado=5
   dibuja(lado, n)
```

```
principal()
```

Bibliography

[1] W. Gautschi. *Numerical Analysis. An Introduction.* Birkhäuser, 1997.

[2] P. Henrici.*Essentials of Numerical Analysis.* Wiley, New York, 1982.

[3] Santaló. Luis. A. *Probabilidad e inferencia estadística.* Edwin S.,1970.

[4] Pliego. Francisco. Javier. Martín. and. Pérez. Luis. Ruiz-Maya. Fundamentos de probabilidad Editorial Paraninfo., 2006.