Instituto Politécnico Nacional
Centro de Investigación en Computación

Probabilidad, procesos aleatorios e inferencia
Profesor. Dr. Jesús Alberto Martínez Castro

Apuntes del curso

Iván Rosales Martínez

Ciudad de México, a 4 de Diciembre del 2017.

Contenido

List of Figures

Capítulo 1

Tareas

1.1 Probabilidad

En cualquier experimento aleatorio siempre hay incertidumbre sobre si un suceso específico ocurrirá o no. Como medida de la oportunidad o probabilidad con la que podemos esperar que un suceso ocurra es conveniente asignar un número entre 0 y 1. Si estamos seguros de que el suceso ocurrirá decimos que su probabilidad es 1, pero si estamos seguros de que el suceso no ocurrirá decimos que su probabilidad es 0.[10]

1.1.1 Probabilidad Clásica

Probabilidad de un suceso es la razón entre el número de casos favorables y el número total de casos posibles, siempre que nada obligue a creer que alguno de estos casos debe tener lugar de preferencia a los demás, lo que hace que todos sean, para nosotros, igualmente posibles.

Laplace define la probabilidad de un suceso como el cociente entre el número de casos favorables, siempre que todos sean igualmente posibles.

La probabilidad clásica supone que el número de casos totales posibles sea finito, lo que implica otra limitación de su campo de aplicabilidad.

$$P(A) = \frac{\left|A\right|}{\left|S\right|} = \frac{\text{Numero de resultados favorables para A}}{\text{Numero total de resultados en S}} \tag{1.1}$$

1.1.2 Probabilidad Geometríca

La probabilidad de un evento representa la posibilidad de que ese evento ocurrirá. La probabilidad geométrica describe la posibilidad de que un punto esté en una parte de un segmento de línea o en una parte de una región.

Considere el segmento de línea $\overline{PQ}$, suponga que un punto X es escogido al azar.

Entonces, la probabilidad de que $\left|X\right|$ este en $\left|\overline{PQ}\right| = \frac{\text{Longitud de PR}}{Longitud de PQ}$.

Considere la región $\left|R\right|$ y la región $\left|S\right|$. Suponga que un punto P es escogido al azar.

Entonces, la probabilidad de que P este en $S = \frac{\text{Area de S}}{\text{Area de R}}$

1.1.3 Probabilidad Frecuentista

Si despúes de n repeticiones de un experimento, donde n es muy grande, un suceso ocurre h veces, entonces la probabilidad del suceso es $\frac{h}{n}$. Esto también se llama la probabilidad empírica del suceso.[10]

$$\lim_{n \to \infty} \left(\frac{h}{n}\right) = P(S) \tag{1.2}$$

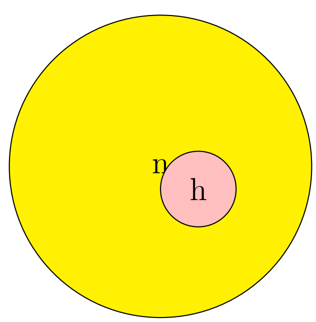


Figure 1.1: Sucesos h ocurridos en n repeticiones de un experimento.

La probabilidad de un suceso como el limite de la frecuencia relativa de aparición de un suceso, al repetirse indefinidamente el experimento aleatorio.

1.1.4 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 más confiable.

1.1.5 Probabilidad Axiomática

Dado un experimento con espacio muestral S y una familia de eventos A, tal que sus elementos cumplen con la leyes del álgebra de eventos se llama "Probabilidad Axiomática" ala función númerica P, cuyo dominio es A, y como rango el intervalo $[0, 1]$, siendo tal que los valores $P(E)$ para cualquier evento E en A, cumplen los tres axiomas de probabilidad.

1.1.6 Los axiomas de la probabilidad

Suponiendo que tenemos un espacio muestral S y S es discreto, entonces todos los subconjuntos corresponden a sucesos, y si S es continuo, solamente subconjuntos especiales (llamados medibles) corresponden a sucesos. A cada suceso A en la clase C de sucesos asociamos un número real $P(A)$, es decir P es una función de valor real definida en C. Así P se llama la función de probabilidad, y $P(A)$ la probabilidad del suceso A, si se satisfacen los axiomas siguientes:

- ***Axioma 1.***Para cada suceso A en la clase C

$$P(A) \geq 0$$

- ***Axioma 2.***Para el suceso cierto o seguro S de la clase C

$$P(S) = 1$$

- ***Axioma 3.***Para cualquier número de sucesos mutuamente excluyentes $A_1, B_2...$ en la clase C

$$P(A_1 \cup A_2 \cup ...) = P(A_1) + P(A_2) + ...$$

En particular, para solo dos sucesos mutuamente excluyentes A_1, A_2,

$$P(A_1 \cup A_2) = P(A_1) + P(A_2)$$

1.1.7 Espacios muestrales

El conjunto S de todos los resultados posibles de una experimento dado se llama el **Espacio muestral.** un resultado particular, esto es, un elemento de S, se llama un **punto muestral o muestra.** Un **evento** A es el conjunto de resultados o, en otras palabras, un subconjunto del espacio muestral S. El evento $|a|$ que consta de una muestra simple $a \in S$ se llama **evento elemental**. El conjunto vacio $\oslash$ y S de por sí son eventos; el $\oslash$ algunas veces se denomina el evento **imposible** (o imposibilidad), y S el evento **cierto o seguro**.[16]

Espacios muestrales finitos

Sea S un espacio mustral finito; digamos, $S = a_1, a_2, ..., a_n$. Un espacio finito de probabilidad se obtiene al asignar a cada punto $a_i \in S$ un número real p_i llamado **probabilidad** de a_i, que satisface las siguientes propiedades:

- cada p_i es no negativo, $p_i \geqslant 0$
- la suma de los p_i es uno, $p_1 + p_2 + ... + p_n = 1$

La probabilidad $P(A)$ de un evento A, se define como la suma de las probabilidades de los puntos A. Por conveniencia de símbolos escribimos $P(a_i)$ en lugar de $P(a_i)$.[10]

Espacios muestrales finitos equiprobables

Un espacio finito S de probabilidad, donde cada punto muestral tiene la misma probabilidad, se llamará **espacio equiprobable o uniforme**. En particular, si S contiene n puntos entonces la probabilidad de cada punto es $\frac{1}{n}$. Además, si un evento A contiene r puntos entonces su probabilidad es $(r)(\frac{1}{n}) = \frac{r}{n}$. En otras palabras;

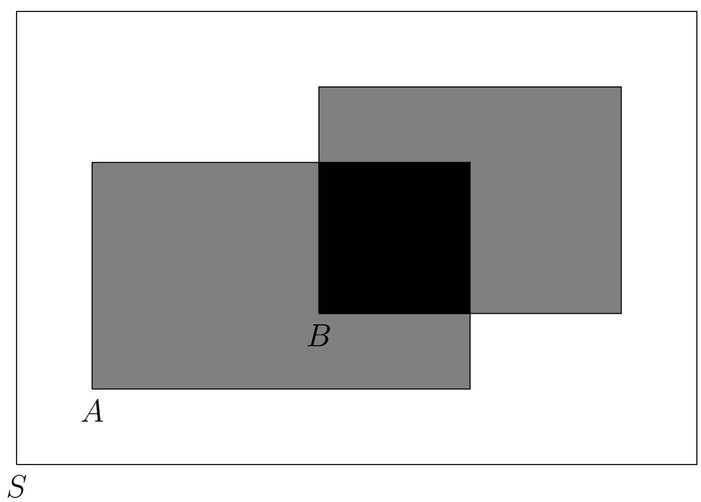
$$P(A) = \frac{\text{numero de maneras en que el evento A puede suceder}}{\text{numero de maneras en que el espacio muestral S puede suceder}} \tag{1.3}$$

La expresión "al azar" se usará solamente respecto a un espacio equiprobable; formalmente, la proposición "escoger un punto al azar de un conjunto S"significa que S es un espacio equiprobable, esto es, que cada punto muestral de S tiene la misma probabilidad.

Espacios muestrales infinitos

Si el espacio mustral S es infinito o contablemente infinito, entonces cada subconjunto de S es un evento. Por otra parte, si S es no contable, entonces algunos subconjuntos de S no pueden ser eventos. Sin embargo, en todos los casos los eventos forman una σ-álgebra ε de subconjuntos de S.[10]

Ejemplo: Sea un lápiz que cae de punto, en una caja rectangular y obsérvese el punto de fondo de la caja donde el lapiz toca primero. Aquí S está formado por todos los puntos de la superficie del fondo. Representemos estos puntos por el área rectangular dibujada a continuación. Sean A y B los eventos en que el lápiz cae en las respectivas áreas ilustradas en la figura. Este es un ejemplo de un espacio muestral que no es finito ni siquiera contablemente infinito, esto es que es no cantable.



Sea S un espacio muestral infinito contable: es decir $S = a_1, a_2, ...$ Como en el caso finito, obtenemos un espacio de probabilidad asignado a cada $a_i \in S$ un número real p_1, llamado su probabilidad, tal que:

- (i) $p_i \geq 0$ y
- (ii) $p_1 + p_2 + ... = \Sigma_{i=1}^{\infty} p_i = 1$

La probabilidad $P(A)$ de un evento A es entonces la suma de las probabilidades de sus puntos.

Los únicos espacios muestrales no contables S que consideraremos aquí son aquellos de medida gométrica finita $m(S)$ tales como la longitud, área o volumen, y en los cuales un punto se selecciona al azar. La probabilidad de un evento A, esto es, aquella en que el punto seleccionado pertenece a A, es entonces la relacioón de $m(A)$ a $m(S)$; o sea;

$$P(A) = \frac{\text{longitud de } A}{\text{longitud de } S}$$
$$P(A) = \frac{\text{area de } A}{\text{area de } S}$$
$$P(A) = \frac{\text{volumen de } A}{\text{volumen de } S}$$

Se dice que un espacio de probabilidad tal es **uniforme**.

1.1.8 Teoremas importantes sobre la probabilidad

De los axiomas anteriores podemos demostrar varios teoremas sobre probabilidad que son importantes.

- ***Teorema 1-14:*** Si $A_1 \subset A_2$ entonces $P(A_1) \leqq P(A_2)$ y $P(A_2 - A_1) = P(A_2) - P(A_1)$
- ***Teorema 1-15:*** Para cada suceso A

  $$0 \leqq P(A) \leqq 1$$

  es decir la probabilidad de un suceso está entre 0 y 1
- ***Teorema 1-16:*** $P(\oslash) = 0$ Es decir el suceso imposible tiene probabilidad cero.
- ***Teorema 1-17:***Si $A^{'}$ es el complemento de A entonces

  $$P(A^{'}) = 1 - P(A)$$

- ***Teorema 1-18:***Si $A = A_1 \cup A_2 \cup ... \cup A_n$ y $A_1, A_2, ..., A_n$son sucesos mutuamente excluyentes, entonces

  $$P(A) = P(A_1) + P(A_2) + ... + P(A_n)$$

  En particular si $A = S$, el espacio muestral, entonces

  $$P(A_1) + P(A_2) + ... + P(A_n) = 1$$

- ***Teorema 1-19:***Si A y B son dos suceso cuales quiera, entonces

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

  Generalizando, si A_1, A_2, A_3 son tres suceso cualesquiera, entonces

  $P(A_1 \cup A_2 \cup A_3) =$
  $= P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_2 \cap A_3) - P(A_3 \cap A_1) + P(A_1 \cap A_2 \cap A_3)$

- ***Teorema 1-20:***Para dos sucesos A y B

$$P(A) = P(A \cap B) + P(A \cap B^{'})$$

- ***Teorema 1-21:*** Si un suceso A debe resultar en uno de los sucesos mutuamente excluyentes $A_1, A_2, ..., A_n$ entonces

$$P(A) = P(A \cap A_1) + P(A \cap A_2) + ... + P(A \cap A_n)$$

1.2 Probabilidad condicional

Sean A y B dos sucesos tales que $P(A) > 0$.Se denota por $P(B|A)$ la probabilidad de B dado que A ha ocurrido. Puesto que se sabe que A ha ocurrido, se convierte en el nuevo espacio muestral remplazando el original S. De aqui llegamos a la definición:

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \tag{1.4}$$

$$P(A \cap B) \leqq P(A)P(B|A) \tag{1.5}$$

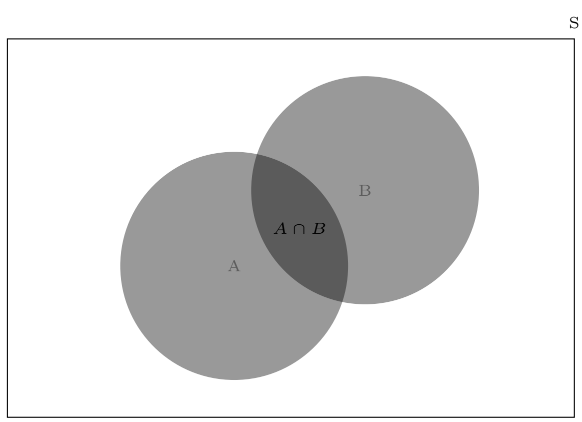


Figure 1.2: Intersección de los eventos A y B - $A \cap B$

Esto nos dice que la probabilidad de que tanto A y B ocurran es igual a la probabilidad de que A ocurra tantas veces la probabilidad de que B ocurra dado que A ha ocurrido. Se le llama a $P(B|A)$ la **probabilidad condicional** de B dada A, es decir la probabilidad de que B ocurra dado que A ha ocurrido.

En particular, si S es un espacio finito equiprobable y $|A|$ denota el número de elementos de un evento A, entonces:

$$P(A \cap B) = \frac{|A \cap B|}{|S|} \tag{1.6}$$

$$P(A) = \frac{|A|}{|S|} \tag{1.7}$$

y así

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{|A \cap B|}{|A|} \tag{1.8}$$

1.2.1 Teorema de la multiplicación para probabilidad condicional

- Para tres eventos cualesquiera A_1, A_2, A_3 tenemos:

$$P(A_1 A_2 A_3) = P(A_1)P(A_2|P(A_1))P(A_3|A_1 \cap A_2) \tag{1.9}$$

  En palabras, la probabilidad de que A_1 y A_2 y A_3 ocurran es igual a la probabilidad de que A_1 ocurra tantas veces la probabilidad de que A_2 ocurra dado que A_1 ha ocurrido tantas veces la probabilidad de que A_3 ocurra dado que A_1 y A_2 han ocurrido.El resultado se generaliza fácilmente a n sucesos.

- Si un suceso A debe resultar en uno de los sucesos mutuamente excluyentes $A_1, A_2, ..., A_n$, entonces:

$$P(A) = P(A_1)P(A|A_1) + P(A_2)P(A|A_2) + ... + P(A_n)P(A|A_n) \tag{1.10}$$

1.3 Teorema de Bayes

Supongamos que los eventos $A_1, A_2, ..., A_n$ forman una partición de un espacio muestral S, esto es, que los eventos A_i son mutuamente exclusivos y su unión es S. Ahora se B otro evento.[10]

Entonces:

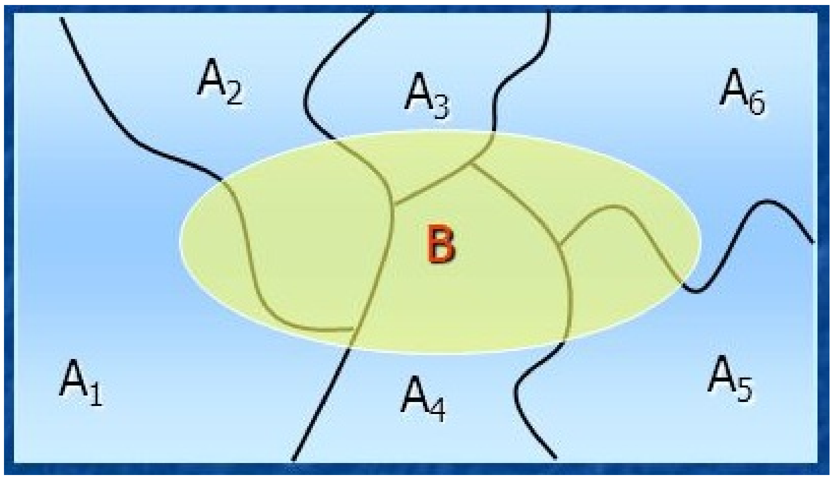


Figure 1.3: Ilustración: Teorema de Bayes

$$\begin{aligned} B = S \cap B &= (A_1 \cup A_2 \cup ... \cup A_n) \cap B && (1.11) \\ &= (A_1 \cap B) \cup (A_2 \cap B) \cup ... \cup (A_n \cap B) && (1.12) \end{aligned}$$

donde las $A_i \cap B$ son eventos mutuamente exclusivos. En consecuencia:

$$P(B) = P(A_1 \cap B) + P(A_2 \cap B) + \cdots + P(A_n \cap B) \quad (1.13)$$

Luego por el teorema de la multiplicación:

$$P(B) = P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + \cdots + P(A_n)P(B|A_n) \quad (1.14)$$

Por otra parte, para cualquier i, la probabilidad condicional de A_i dado B se define por:

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} \tag{1.15}$$

Sabiendo que $P(A_i \cap B) = P(A_i)P(B|A_i)$se puede remplazar $P(A_i \cap B)$, obteniendo así :

El Teorema de Bayes

Supongamos que los eventos $A_1, A_2, ..., A_n$ forman una partición de un espacio muestral S y que B es cualquier evento.[10]

Entonces para cualquier i:

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + \cdots + P(A_n)P(B|A_n)} \tag{1.16}$$

1.4 Técnicas de conteo

1.4.1 Principio fundamental del conteo

Si un evento puede realizar de n_1 maneras diferentes, y si, continuando el procedimiento un segundo evento puede realizarse de n_2 maneras diferentes, y si despúes de efectuado un tercer evento puede realizarse de n_3 maneras diferentes y así sucesivamente, entonces el número de maneras de que los eventos puedan realizarse en el orden indicado es el producto $n_1 n_2 n_3 ....$

1.4.2 Notación factorial

El producto de los enteros positivos desde 1 hasta n inclusive, se emplea con mucha frecuencia en matemáticas y se denota por el símbolo $n!$ (que se lee n factorial).

$$n! = (1)(2)(3)......(n-2)(n-1)(n) \tag{1.17}$$

Conviene también definir $0! = 1$.

1.5 Aproximación de Stirling

La fórmula de Stirling es una aproximación para factoriales grandes. Lleva el nombre en honor al matemático escocés del siglo XVIII James Stirling.

$$ln(n!) \approx nln(n) - n \tag{1.18}$$

Para n suficientemente grande, donde ln es el logaritmo natural.

La fórmula está dada por:

$$\lim_{n \to \infty} \frac{n!}{\sqrt{2\pi n}(\frac{n}{e})^n} = 1 \tag{1.19}$$

Que se reescribe frecuentemente como:

$$n! \approx \sqrt{2\pi n}(\frac{n}{e})^n \tag{1.20}$$

Formalmente:

$$n! = \sqrt{2\pi n}(\frac{n}{e})^n e^{\frac{1}{12n} - \frac{1}{360n^3} + \frac{1}{1260n^5} - \frac{1}{1680n^7} + \cdots} \tag{1.21}$$

Donde el último término del producto(la exponencial) tiende a 1 cuando n tiende a infinito.

La lista de los numeradores es: 1, -1, 1, -1, 1, -691, 1, -3617, 43867, -174611, ...

La lista de los denominadores es: 12, 360, 1260, 1680, 1188, 360360, 156, 122400, 244188, 125400, ...

Desarrollando este último término también se puede reescribir la fórmula como:

$$n! = \sqrt{2\pi n}(\frac{n}{e})^n (1 + \frac{1}{12n} - \frac{1}{288n^2} - \frac{139}{51840n^3} - \frac{571}{2488320n^4} + \cdots) \tag{1.22}$$

Acotando la fórmula:

$$\sqrt{2\pi n}(\frac{n}{e})^n e^{\frac{1}{12n+1}} < n! < \sqrt{2\pi n}(\frac{n}{e})^n e^{\frac{1}{12n}} \tag{1.23}$$

1.6 Permutaciones

Una ordenación de un conjunto de n objetos en un orden dado se llama una permutación de los objetos (tomados todos a la vez). una ordenación de un numero r de dichos objetos $r \leq n$, en un orden dado se llama una permutación r o una permutación de los n objetos, tomados r a la vez.

Ejemplo: Consideremos el conjunto de las letras $|a, b, c, d|$, entonces:

- (i) *bdca*, *dcba* y *acdb* son permutaciones de las cuatro letras tomadas todas a la vez.
- (ii) *bad*, *abd* y *cbd* son permutaciones de las cuatro letras tomadas 3 a la vez.
- (iii) *ad*, *cb*, *da* y *bd* son permutaciones de las cuatro letras tomadas 2 a la vez.

El número de permutaciones de n elementos tomados r a la vez se denota por:

$$P(n, r) \tag{1.24}$$

Para la deducción de la formula $P(n, r)$, el primer elemento de una permutación r de n objetos puede escogerse de n diferentes maneras, despúes el segundo elemento de la permutación puede escogerse de $n - 1$ maneras y sucesivamente el tercer elemento puede escogerse de $n-2$ maneras. Continuando en esta forma tenemos que el $r - esimo$ elemento de la permutación r puede escogerse de $n - (r - 1) = n - r + 1$ maneras, así:

$$ \tag{1.25}$$

$$P(n, r) = n(n-1)(n-2)......(n-r+1) = \frac{n(n-1)(n-2)......(n-r+1)(n-r)!}{(n-r)!} \tag{1.26}$$

$$P(n, r) = \frac{n!}{(n-r)!} \tag{1.27}$$

Para el caso especial de $r = n$, tenemos que:

$$P(n, n) = n(n-1)(n-2)......(3)(2)(1) = n! \tag{1.28}$$

Es decir: Las permutaciones de n objetos (tomados todos a la vez) son igual es a $n!$.

1.6.1 Permutaciones con repeticiones

Con frecuencia se desea saber el número de permutaciones de objetos, de los cuales algunos son iguales, como se indica a continuación. Se usa la formula general.

El número de permutaciones de n objetos de los cuales n_1 son iguales, n_2 son iguales , hasta n_r son iguales es:

$$\frac{n!}{n_1!n_2!n_3!....n_r!} \tag{1.29}$$

Comprobamos el teorema expuesto mediante el siguiente ejemplo: Deseamos formar todas las posibles palabras de 5 letras usando las letras empleadas en la palabra $DADDY$. Ahora se tiene $5! = 120$ permutaciones de los objetos D_1, A, D_2, D_3, Y, donde las tres D están marcadas. Observamos que las 6 permutaciones siguientes:

$$\begin{array}{c} D_1D_2D_3AY, \\ D_2D_1D_3AY, \\ D_3D_1D_2AY, \\ D_1D_3D_2AY, \\ D_2D_3D_1AY, \\ D_3D_2D_1AY \end{array}$$

Forman la misma palabra si se quitan la subíndices. Las 6 resultan del echo que hay $3! = (3)(2)(1) = 6$ maneras diferentes de colocar las 3 D en los 3 primeros lugares de la permutación. Esto es cierto para cada una de las otras posiciones posibles en donde las D aparezcan. Por consiguiente hay:

$$\frac{5!}{3!} = \frac{120}{60} = 20$$

Palabras diferentes de 5 letras, que pueden formarse tomando las letras de la palabra $DADDY$.

1.7 Pruebas ordenadas

Muchos problemas del ánalisis combinatorio y, en particular de probabilidad se relacionan con escoger una bola de una urna que contiene n bolas (o una carta de una baraja o una persona de una población). Cuando escogemos una bola de una urna r veces , definimos esta actividad como una prueba ordenada de tamaño r y se consideran dos casos:

- (i) **Pruebas con sustitución** En este caso cada bola se regresa a la urna antes de tomar la siguiente. Ahora puesto que hay n maneras diferentes para escoger cada bola, según el principio fundamental del conteo, hay:

$$((n)(n)(n).......(n)) \text{ r veces } = n^r$$

- (ii) **Pruebas sin sustitución** Aquí la bola no se devuelve a la urna antes de escoger la siguiente. Así no hay repeticiones en la prueba ordenada. O sea que, una prueba ordenada de tamaño r si sustitución es simplemente una permutación r de objetos de una urna. Por consiguiente hay:

$$P(n,r) = n(n-1)(n-2)......(n-r+1) = \frac{n(n-1)(n-2)......(n-r+1)(n-r)!}{(n-r)!}$$

Pruebas ordenadas de tamaño r sin sustitución tomadas de un grupo de n objetos.

1.8 Coeficientes del binomio

El símbolo $\binom{n}{r}$ se lee "nCr", donde n y r son números positivos con $r \leq n$, se define como a continuación:

$$\binom{n}{r} = \frac{n(n-1)(n-2)\cdots(n-r+1)}{1 \cdot 2 \cdot 3 \cdots (r-1)r} \tag{1.30}$$

Esta espresión es llamada coeficiente del binomio.

Nótese que $\binom{n}{r}$ tiene exactamente r factores tanto en el numerados como en el denominador, entonces:

$$\binom{n}{r} = \frac{n(n-1)(n-2)\cdots(n-r+1)}{1 \cdot 2 \cdot 3 \cdots (r-1)r} = \frac{n(n-1)(n-2)\cdots(n-r+1)(n-r)!}{1 \cdot 2 \cdot 3 \cdots (r-1)r(n-r)!} \tag{1.31}$$

(1.32)

$$\binom{n}{r} = \frac{n!}{r!(n-r)!} \tag{1.33}$$

Con esta formula y tomando en cuenta que $n - (n + r) = r$, se obtiene la seiguiente relación importante:

$$\binom{n}{n-r} = \binom{n}{r} \text{ o lo que es lo mismo, si} a + b = n \text{ , entonces } \binom{n}{a} = \binom{n}{b} \tag{1.34}$$

En base a que $0! = 1$ es importante definir:

$$\binom{n}{0} = \frac{n!}{0!(n)!} = 1$$

y, en particular

$$\binom{0}{0} = \frac{0!}{0!(0)!} = 1$$

1.9 Teorema del binomio

$$\tag{1.35}$$

$$(a + b)^n = \Sigma_{r=0}^{n} \binom{n}{r} a^{n-r} b^r \tag{1.36}$$

$$(a + b)^n = a^n + na^{n-1}b + \frac{n(n-1)}{1 \cdot 2} a^{n-2}b^2 + \cdots + nab^{n-1} + b^n \tag{1.37}$$

En el desarrollo de $(a + b)^n$ se debe de observar las siguientes propiedades:

1. Hay $(n + 1)$ términos.
2. La suma de los exponentes en a y en b en cada término es n
3. Los exponentes de a decrecen una unidad en cada término desde n hasta 0, los exponentes de b creecen similarmente de 0 a n.
4. El coeficiente de cualquier término es $\binom{n}{k}$ donde k es el exponente de a o de b.
5. Los coeficientes de términos equidistantes de los extremos son iguales.

1.10 Combinaciones

En una permutación nos interesa el orden de la distribucion de los objetos, así *abc* es una permutación diferente de *bca*. Sin embargo en muchos otros problemas nos interesa solamente en seleccionar o escoger objetos sin importar el orden. Dichas selecciones se llaman *Combinaciones*. Por ejemplo *abc* y *bca* son la misma combinación.

El número total de combinaciones de r objetos tomados de n, se denota por nCr , $C(n,r)$ o $\binom{n}{r}$, por lo que tenemos:

$$\tag{1.38}$$

$$\binom{n}{r} = nCr = \frac{n!}{r!(n-r)!} \tag{1.39}$$

$$\binom{n}{r} = \frac{n(n-1)(n-2)\cdots(n-r+1)}{r!} = \frac{nPr}{r!} \tag{1.40}$$

Por lo que se demostra fácilmente que :

$$\binom{n}{r} = \binom{n}{n-r} \text{ o } nCr = nCn - r \tag{1.41}$$

Ejemplo: ¿Cuál es el número total de maneras que pueden escogerse 3 cartas de un total de 8 diferentes?

$$C(8,3) = \binom{8}{3} = \frac{8!}{3!(8-3)!} = 56$$

1.11 Diagramas de árbol

Un diagrama de árbol es el dibujo que se usa para enumerar todos los resultados posibles de experimentos en donde cada experimento,puede suceder en un número finito de maneras. La construcción de diagramas de árbol se ilustra en el siguiente ejemplo.

Ejemplo 1: Hallar el conjunto producto $A\text{x}B\text{x}C$ en donde $A = 1, 2$, $B = a, b, c$ y $C = 3, 4$, usamos el diagrama de árbol siguiente:

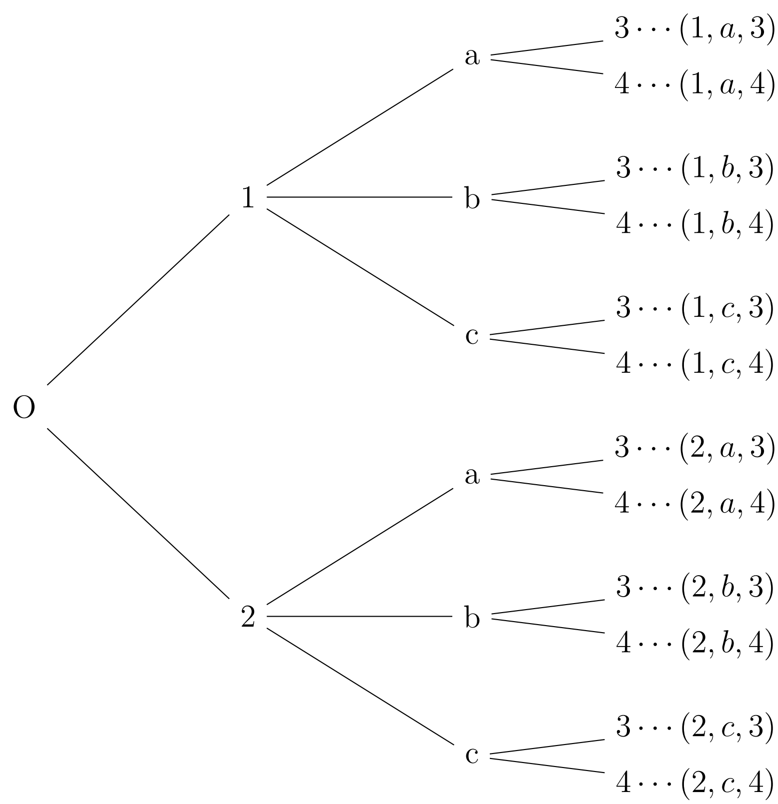


Figure 1.4: Diagrama de árbol del problema.

1.12 AND

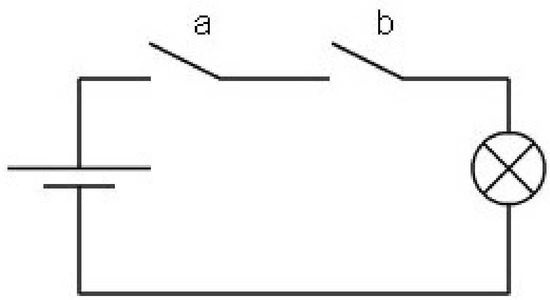


Figure 1.5: Representación de la compuerta lógica AND con circuito de interruptores

1.12.1 Principio Multiplicativo

Si se desea realizar una actividad que consta de r pasos, en donde el primer paso de la actividad a realizar puede ser llevado a cabo de N_1 maneras o formas, el segundo paso de N_2 maneras o formas y el r-ésimo paso de N_r maneras o formas, entonces esta actividad puede ser llevada a cabo de: $N_1 \times N_2 \times ...N_r$ maneras posibles.

El principio multiplicativo implica que cada uno de los pasos de la actividad deben ser llevados a efecto, uno tras otro.

Ejemplo:

De la ciudad A a la ciudad B hay 4 caminos diferentes y de la ciudad B a la ciudad C hay 3 caminos diferentes de cuantas maneras se podrá ir a de A a C?

Maneras posibles = Caminos de A a B $\times$ Caminos de B a C $= 4 \times 3 = 12$

1.13 OR

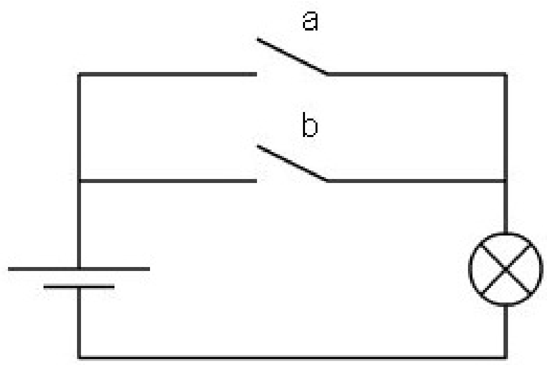


Figure 1.6: Representación de la compuerta lógica OR con circuito de interruptores

1.13.1 Principio Aditivo

Si se desea llevar a cabo una actividad la cual tiene formas alternativas para ser realizada, donde la primera de estas alternativas puede ser realizada de M maneras o formas, la segunda alternativa puede realizarse de N maneras o formas y la tercera de W maneras posibles, entonces esta actividad puede llevarse a cabo de $M + N + W$ ... maneras posibles.

Ejemplo:

Para pintar una casa de un solo color disponemos de 3 colores de pinturas brillantes y 4 colores de pinturas opacas. ¿De cuantas maneras podemos elegir la pintura para la casa?

La pintura brillante puede ser elegida de 3 maneras diferentes mientras que la opaca puede ser elegida de 4 maneras distintas por lo tanto:

Las maneras con las que podemos elegir pintar una casa son $= 3 + 4 = 7$.

1.14 Eigenvalores y Eigenvectores

Sea A una matriz de nxn, su ecuación de eigenvalor es:

$$A\vec{x} = \lambda\vec{x} \tag{1.42}$$

El número $\lambda \in C$ es un eigenvalor o valor propio.

Mientras que el valor x es el eigen vector o vector propio asociado a λ.

El espacio nulo de $(A - \lambda I)$ debe contener vectores diferentes de cero. Por lo tanto $(A - \lambda I)$ debe de ser singular.[9]

Sea A una matriz de nxn y sea λ un escalar. Entonces el polinomio característico de A se define como:

$$P(\lambda) = det(A - \lambda I) \tag{1.43}$$

Entonces λ_0 es una valor propio de A si y solo si:

$$P(\lambda_0) = det(A - \lambda_0 I) = 0 \tag{1.44}$$

1.14.1 Multiplicidad algebraica

Sea A una matriz de nxn con eigenvalores $\lambda_1, \lambda_2, ..., \lambda_m$ entonces su polinomio característico puede factorizarse como:

$$P(\lambda) = (-1)^n(\lambda - \lambda_1)^{r_1}(\lambda - \lambda_2)^{r_2}...(\lambda - \lambda_m)^{r_m} \tag{1.45}$$

Donde $r_1, ..., r_m$ son las multiplicidades algebraicas de los eigenvalores $\lambda_1, \lambda_2, ..., \lambda_m$ y cumplen que $r_1 + r_2 + ... + r_m = n$.[9]

1.14.2 Polinomio característico

$$det(\begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix}) = (a-\lambda)(d-\lambda) - cb = \tag{1.46}$$

$$= ad - d\lambda - a\lambda + \lambda^2 - cb = \lambda^2 - (a+d)\lambda + (ad - cb) =$$

$$= \lambda^2 - tr(A)\lambda + det(A)$$

$$P(\lambda) = (-1)^n\lambda^n + (-1)^{n-1}tr(A)\lambda^{n-1} + ... + det(A) \tag{1.47}$$

1.14.3 Método para encontrar los valores y vectores propios

1. Obtener el determinante de $(A - \lambda I)$, este determinante es un polinomio de grado n, y que comienza con λ^n.

2. Encontrar las raíces de este polinomio, las n raíces son los valores propios de A.

3. Para cada valor propio resolver la ecuación $(A - \lambda I)x = 0$. Debido a que el determinante es 0 existen soluciones ádemas de $x = 0$. Estas soluciones son los vectores propios.

1.14.4 Propiedades de los valores propios

Sea A una matriz de nxn. La suma de sus n valores propios es igual a la traza de la matriz.

$$tr(A) = \lambda_1, \lambda_2, ..., \lambda_n = a_{11} + a_{22} + ... + a_{nn} \tag{1.48}$$

El producto de n valores propios es igual al determinante de A.

$$det(A) = \lambda_1 \lambda_2 ... \lambda_n = \prod_{i=1}^{n} \lambda_i \tag{1.49}$$

1.15 Convenio de suma de Einstein

2 Índices repetidos $\Leftrightarrow$ Suma.

$$a_i b_i = \sum_{n=1}^{3} a_i b_i = a_1 b_1 + a_2 b_2 + a_3 b_3 \tag{1.50}$$

$$\overline{AB} = \left|\overline{A}\right|\left|\overline{B}\right| \cos(\theta) \tag{1.51}$$

Por convención :

$$= a_i b_i = a_1 b_1 + a_2 b_2 + a_3 b_3 \tag{1.52}$$

1.16 Tensor de Levi-Civita

$$\varepsilon_{ijk} = \begin{cases} +1 & \text{si} \quad P.pares \\ -1 & \text{si} \quad P.impares \\ 0 & \text{si} \quad Serepiten \end{cases} \tag{1.53}$$

Ejemplo:

$$123 \longleftrightarrow (-)321 \longleftrightarrow (-)(-)312 = 312$$

Para pasar de la secuencia 123 a la secuencia 312 se tiene que pasar por dos permutaciones las cuales generan dos cambios de signo.

$$(\overline{A} \times \overline{B}) = \begin{pmatrix} \widehat{i} & \widehat{j} & \widehat{k} \\ A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \end{pmatrix} = \widehat{i}(A_2 B_3 - A_3 B_2) - \widehat{j}(A_1 B_3 - A_3 B_1) + \widehat{k}(A_1 B_2 - A_2 B_1) \tag{1.54}$$

$$(\overline{A} \times \overline{B})_i = \varepsilon_{ijk} A_j B_k \tag{1.55}$$

Aplicando el convenio de suma:

$$(\overline{A} \times \overline{B})_i = \varepsilon_{i1k} A_1 B_k + \varepsilon_{i2k} A_2 B_k + \varepsilon_{i3k} A_3 B_k = \tag{1.56}$$

$$\begin{aligned} = {} & \varepsilon_{i11} A_1 B_1 + \varepsilon_{i12} A_1 B_2 + \varepsilon_{i13} A_1 B_3 \\ & + \varepsilon_{i21} A_2 B_1 + \varepsilon_{i22} A_2 B_2 + \varepsilon_{i23} A_2 B_3 \\ & + \varepsilon_{i31} A_3 B_1 + \varepsilon_{i32} A_3 B_2 + \varepsilon_{i33} A_3 B_3 \end{aligned}$$

En base a la definición del tensor de Levi-Civita las expresiones en rojo valen 0 por que los subíndices se repiten.

Evaluando $i = 1, 2, 3$ y tomando en cuenta la definición del tensor de Levi-Civita se obtienen las siguientes expresiones:

$$(\overline{A} \times \overline{B})_1 = \varepsilon_{123} A_2 B_3 + \varepsilon_{132} A_3 B_2 = A_2 B_3 - A_3 B_2 \tag{1.57}$$

$$(\overline{A} \times \overline{B})_2 = \varepsilon_{213} A_1 B_3 + \varepsilon_{231} A_3 B_1 = -A_1 B_3 + A_3 B_1 \tag{1.58}$$

$$(\overline{A} \times \overline{B})_3 = \varepsilon_{312} A_1 B_2 + \varepsilon_{321} A_2 B_1 = A_1 B_2 - A_2 B_1 \tag{1.59}$$

$$\begin{aligned} 123 &\to (-)213 \to (+)231 \\ 123 &\to (-)321 \to (+)312 \end{aligned}$$

$$X^{'\mu} = a^\mu_\nu X^\nu ; \nu, \mu = 0, 1, 2, 3 \tag{1.60}$$

$$= a^\mu_0 X^0 + a^\mu_1 X^1 + a^\mu_2 X^2 + a^\mu_3 X^3$$

$$X^{'0} = a^0_0 X^0 + a^0_1 X^1 + a^0_2 X^2 + a^0_3 X^3 \tag{1.61}$$

$$X^{'1} = a^1_0 X^0 + a^1_1 X^1 + a^1_2 X^2 + a^1_3 X^3 \tag{1.62}$$

$$X^{'2} = a^2_0 X^0 + a^2_1 X^1 + a^2_2 X^2 + a^2_3 X^3 \tag{1.63}$$

$$X^{'3} = a^3_0 X^0 + a^3_1 X^1 + a^3_2 X^2 + a^3_3 X^3 \tag{1.64}$$

$$\underbrace{\begin{pmatrix} X^{'0} \\ X^{'1} \\ X^{'2} \\ X^{'3} \end{pmatrix}}_{X'^\mu} = \underbrace{\begin{pmatrix} a^0_0 & a^0_1 & a^0_2 & a^0_3 \\ a^1_0 & a^1_1 & a^1_2 & a^1_3 \\ a^2_0 & a^2_1 & a^2_2 & a^2_3 \\ a^3_0 & a^3_1 & a^3_2 & a^3_3 \end{pmatrix}}_{a^\mu_{,\nu}} \underbrace{\begin{pmatrix} X^0 \\ X^1 \\ X^2 \\ X^3 \end{pmatrix}}_{X^\nu} \tag{1.65}$$

1.17 Población y muestra

Población (o universo) es cualquier colección finita o infinita de elementos o sujetos (población de pacientes que acuden al Hospital Central “Antonio María Pineda”, población de familias de una comunidad. Algunos autores establecen diferencias entre los términos universo y población, indicando con el primero un conjunto de personas, seres u objetos y con el segundo, un conjunto de números obtenidos midiendo o contando cierta característica de los mismos, de allí que un universo puede contener varias poblaciones. Por ejemplo, del universo de pacientes del Hospital Central Antonio María Pineda, una población estaría dada por el conjunto de las edades de esas personas, sus tallas, sus pesos, etc.

Se habla de que una población es finita cuando consta de un número limitado de elementos, ejemplo: todos los habitantes de una comunidad. Una población es infinita cuando no se pueden contabilizar todos sus elementos pues existen en número ilimitado, como por ejemplo, la población de insectos en el mundo. Una muestra es un subconjunto de la población, que se obtiene para averiguar las propiedades o características de esta última, por lo que interesa que sea un reflejo de la población, que sea representativa de ella, concepto al que volveremos más adelante.

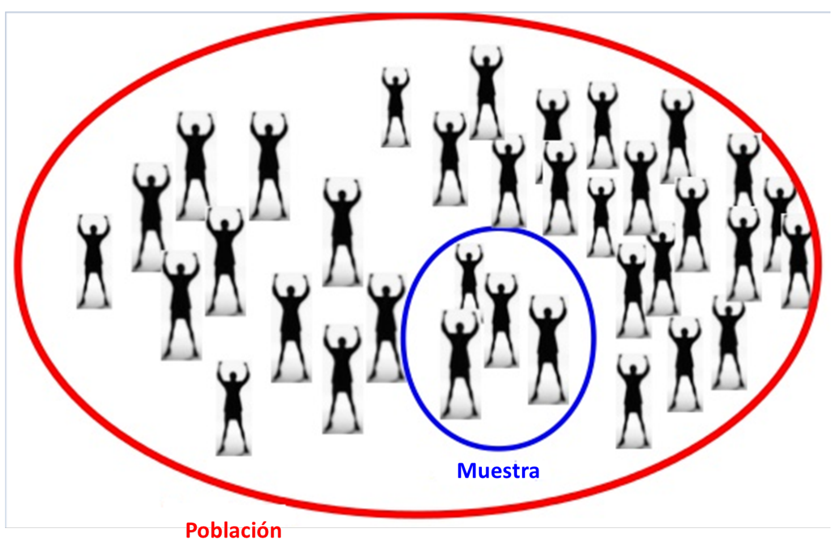


Figure 1.7: Representación: Población y Muestra

1.18 Histograma

En estadística, un histograma es una representación gráfica 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 caracterśtica.

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 aleatoria mente 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 in

nito 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 están separadas, esto es porque en el "x" ya no se representa un espectro continuo de valores, sino valores cuantitativos específi

cos como ocurre en un diagrama de barras cuando la característica que se representa es cualitativa o categórica. Su utilidad se hace más evidente cuando se cuenta con un gran número de datos cuantitativos y que se han agrupado en intervalos de clase.

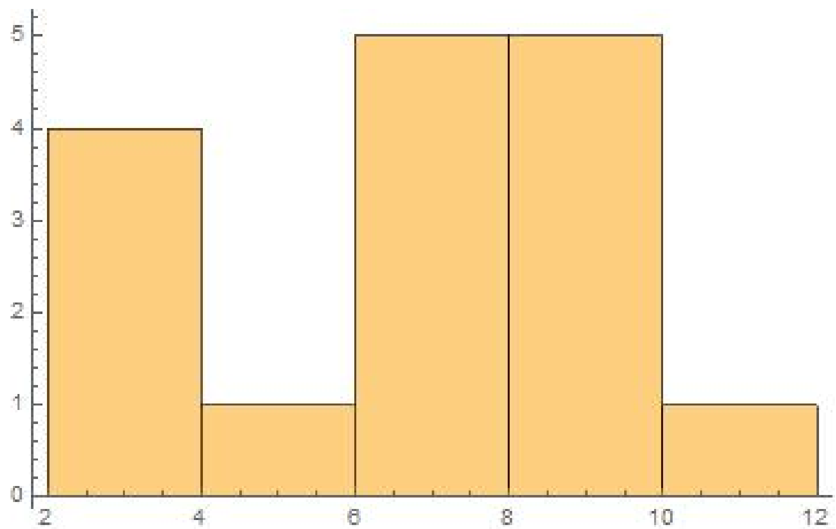


Figure 1.8: Ejemplo de histograma en Mathematica Wolfram

1.19 Medidas Descriptivas

Las medidas de tendencia central, de dispersión y de posición, son de relevante importancia en el momento de realizar estudios estadísticos. Las medidas de tendencia central son utilizadas para localizar el centro de un grupo de datos. La dispersión evalúa la separación o apartamiento de las medidas de los datos, respecto al centro.

Las medidas de posición ubican un elemento en un grupo de datos respecto a otro.

1.19.1 Medidas de tendencia central

Un promedio es un valor típico o representativo de un conjunto de datos. Como estos valores típicos tienden a encontrarse en el centro de los conjuntos de datos, ordenados de acuerdo con su magnitud, a los promedios se les conoce también como medidas de tendencia central. Se pueden definir varios tipos de promedios; los más usados son la media aritmetica, la mediana, la moda. Cada una de ellas tiene ventajas y desventajas de acuerdo con el tipo de datos y el propósito de su uso.[16]

Media Aritmética

La media aritmética, o brevemente la media, de un conjunto de N números $X_1, X_2, X_3, ..., X_N$ se denota así: $\overline{X}$ (que se lee "X barra") y está definida como:

$$\overline{X} = \frac{X_1, X_2, X_3, ..., X_N}{N} = \frac{\sum_{j=1}^{N} X_j}{N} \tag{1.66}$$

Si los números $X_1, X_2, X_3, ..., X_k$ se presentan $f_1, f_2, f_3, ..., f_k$ veces, respectivamente (es decir, se presentan con frecuencias $f_1, f_2, f_3, ..., f_k$), su media aritmética es:

$$\overline{X} = \frac{f_1X_1, f_2X_2, f_3X_3, ..., f_kX_k}{f_1 + f_2 + ... + f_k} = \frac{\sum_{j=1}^{k} f_jX_j}{\sum_{j=1}^{k} f_j} \tag{1.67}$$

Mediana

La mediana es el valor central que se localiza en una serie ordenada de datos. Para obtener la mediana de los números $x_1 = 13, x_2 = 15, x_3 = 9, x_4 = 6, x_5 = 4, x_6 = 12, x_7 = 11$, primero tenemos que ordenarlos:

4	6	9	11	12	13	15

Entonces la mediana es 11.

En caso de que el número de datos sea par, se tienen dos valores centrales y la mediana de estos dos valores es el punto medio entre ambas. La mediana es el centro geométrico de la distribución de los datos

Moda

La moda es el valor que con mayor frecuencia ocurre, e otras palabras, tiene la mayor probabilidad de ocurrir. En tal valor x, $f(x)$ es máxima. Algunas veces tenemos dos, tres o más valores que tienen relativamente grandes probabilidades de ocurrencia. En tales casos decimos que la distribución es bimodal, trimodal o multimodal respectivamente.

1.19.2 Varianza y desviación estándar

La varianza de X, mide el ”esparcimiento” o ”dispersion” de X.[10]

Sea X una variable aleatoria con la siguiente distibución:

x_1	x_2	...	x_n
$f(x_1)$	$f(x_2)$	...	$f(x_n)$

Entonces el valor esperado $E(X)$ y la *varianza* de X, denotada por $var(X)$, se definen como:

$$E(X) = x_1 f(x_1) + x_2 f(x_2) + ... + x_i f(x_i) = \sum_{i=1}^{\infty} x_i f(x_i) \tag{1.68}$$

$$var(X) = \sum_{i=1}^{n} (x_i - \mu)^2 f(x_i) = E((X-\mu)^2) \tag{1.69}$$

donde μ es la media de X. La desviación estándar de X, denotada por σ_X, es la raíz cuadrada (no-negativa) de $var(X)$:

$$\sigma_X = \sqrt{var(X)} \tag{1.70}$$

El teorema siguiente nos da una alternativa y algunas veces una fórmula más útil para calcular la varianza de la variable aleatoria X.

Teorema 5.4 [10]

Cuando las series pertinentes convergen absolutamente. Se puede demostrar que $var(X)$ existe si y sólo sí $\mu = E(X)$ y $E(X^2)$ existen ambos y en este caso la variaza se define.

$$var(X) = \sum_{i=1}^{n} x_i^2 f(x_i) - \mu^2 = E(X^2) - \mu^2 \tag{1.71}$$

1.19.3 Variables aleatorias continuas

Supóngase que X es una variable aleatoria cuyo conjunto imagen $X(S)$ es un conjunto continuo de números tales como un intervalo $|a \leq X \leq b|$ por lo que la probabilidad $P(a \leq X \leq b)$ está bien definida. Suponiendo que existe una función continua especial $f : R \longrightarrow R$ tal que $P(a \leq X \leq b)$ es el área bajo la curva de f entre $x = a$ y $x = b$ por lo tanto:

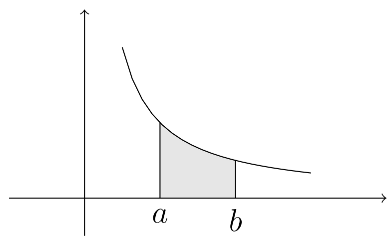


Figure 1.9: Área bajo la curva de la función f de a a b.

$$P(a \leq X \leq b) = \int_a^b f(x)dx \tag{1.72}$$

En este caso X es una variable aleatoria continua. La función f se llama función de distribución o de probabilidad continua (o función de densidad) de X, que satisface las condiciones:

$$f(x) \geq 0$$

$$\int_R f(x)dx = 1$$

Esto es f es no negativa y el área bajo la curva es 1.

El valor esperado $E(X)$ se define por:

$$E(X) = \int_R xf(x)dx \tag{1.73}$$

cuando existe. Las funciones de variables aleatorias se definen justamente como en el caso discreto y puede demostrarse que si $Y = \phi(X)$ entonces:

$$E(Y) = \int_R \phi(x) f(x) dx \tag{1.74}$$

cuando el miembro de la derecha existe. La varianza se define por:

$$var(X) = E((X - \mu)^2) = \int_R (x - \mu)^2 f(x) dx \tag{1.75}$$

Justamente como en el caso discreto, se puede demostrar que $var(X)$ existe si y sólo sí μ , $E(X)$ y $E(X)$ existen y por lo tanto:

$$var(X) = E(X^2) - \mu^2 = \int_R x^2 f(x) dx - \mu^2 \tag{1.76}$$

1.19.4 Sesgo

Con frecuencia una distribución no es simétrica con respecto a un máximo sino que tiene una "cola" más larga que la otra. Si la cola más larga se extiende a la derecha, se dice que la distribución está sesgada a la derecha, mientras que si la cola más larga se extiende a la izquierda, se dice que está sesgada a la izquierda. Las medidas que describen esta simetría se denominan los coeficientes de sesgo o sencillamente sesgo, que se determina por la siguiente expresión:

$$\alpha_3 = \frac{E[(X - \mu)^3]}{\sigma^3} = \frac{\mu_3}{\sigma^3} \tag{1.77}$$

Que es una cantidad adimensional. La medida α_3 es positiva o negativa si la distribución está sesgada a la derecha o a la izquierda respectivamente. Para una disitribución simétrica $\alpha_3 = 0$.

1.19.5 Curtosis

En algunos casos una distribución puede tener sus valores concentrados cerca de la media así que la distribución tiene un pico grande. En otros casos la distribución puede ser relativamente plana. Medidas del grado de apuntamiento de una distribución se llaman coeficientes de curtosis o simlemente curtosis. una medida empleada frecuentemente está dada por:

$$\alpha_4 = \frac{E[(X-\mu)^4]}{\sigma^4} = \frac{\mu_4}{\sigma^4} \tag{1.78}$$

Que es una cantidad adimensional. Comúnmente se compara con la curva normal que tiene un coeficiente curtosis igual a 3.

1.19.6 Coeficiente de Asimetría o Skewness

Skewness, determina el grado de asimetría que posee una distribución. Para el caso de funciones simétricas como la normal o la t-student, este coeficiente es cero, y analíticamente se representa por:

$$S_k = \frac{\sum(\frac{x_i - \overline{x}}{\sigma})^3}{n} \tag{1.79}$$

Donde n representa al tamaño muestral. Este indicador indica si la cola más larga de la distribución se encuentra desviada hacia la derecha, centrada o desviada hacia la izquierda de la distribución. Si la cola más larga se encuentra hacia la izquierda (derecha) de la distribución, el coeficiente de skewness será negativo (positivo) y se dirá que la distribución es sesgada a la izquierda (derecha).

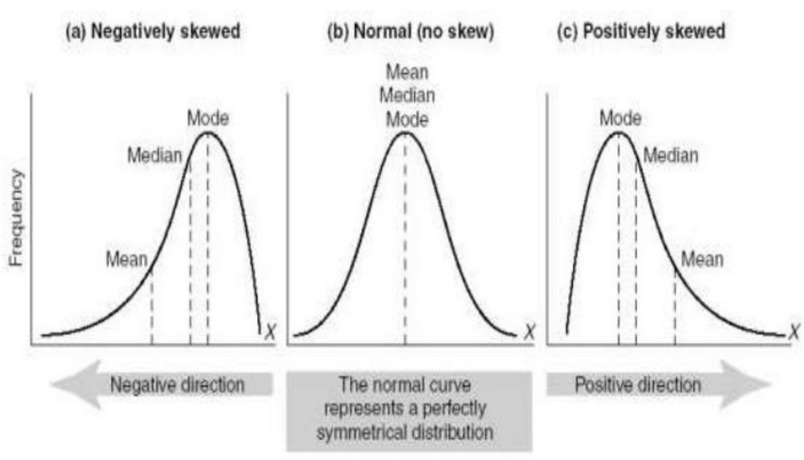


Figure 1.10: Coeficiente de asimetría: Skewness

1.20 Teorema del límite central

El teorema del limite central o teorema del límite indica que, en condiciones muy generales, si S_n es la suma de las variables aleatorias independientes y de varianza no nula pero finita, entonces la función S_n se aproxima bien a una distribución normal (también llamada distribución gaussiana). Así pues, el teorema asegura que esto ocurre cuando la suma de esas variables aleatorias e independientes es lo suficientemente grande.[11]

Sea $X_1, X_2, \cdots, X_n$ 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.[11]

$$\overline{X} = \frac{1}{n}\Sigma_{i=1}^{n} X_i \tag{1.80}$$

Tiene aproximadamente una distribución normal con $\mu_{\overline{X}} = \mu$ y $\sigma_{\overline{X}}^2 = \dfrac{\sigma^2}{n}$

1.20.1 Propiedades del teorema

- El teorema del límite central garantiza una distribución normal cuando n es suficientemente grande.
- Existen diferentes versiones del teorema, en función de las condiciones utilizadas para asegurar la convergencia. una de las más simples establece que es suficiente que las variables que se suman sean independientes, idénticamente distribuidas, con valor esperado y varianza finitas.[11]
- La aproximación entre las dos distribuciones es, en general, mayor en el centro de las mismas que en sus extremos o colas, motivo por el cual se prefiere el nombre "teorema del límite central".[11]
- Este teorema, perteneciente a la teoría de la probabilidad, encuentra aplicación en muchos campos relacionados, tales como la inferencia estadística o la teoría de la renovación.[11]

1.21 Caminata al azar

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.

$$pp...ppp...pp = p^{n_1}p^{n_2} \tag{1.81}$$

Siendo p y q las probabilidades de dar el paso a la izquierda o a la derecha, sabemos que para obtener la probabilidad de un evento por principio multiplicativo dice:

Si una acción puede realizarse de n_1 maneras diferentes y una segunda acción puede realizarse de n_2 maneras diferentes, entonces ambas acciones pueden realizarse secuencial mente de n_1n_2 maneras diferentes. Este principio multiplicativo se generaliza para cualquier número de acciones a realizar, esto es, si una primera acción se puede realizar de n_1 maneras diferentes, una segunda acción se puede realizar de n_2 maneras diferentes,..., y una r-ésima acción se puede realizar de n_r maneras diferentes, entonces las r acciones se pueden realizar de maneras diferentes.

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:

$$\frac{N!}{n_1!n_2!} \tag{1.82}$$

Entonces:

La probabilidad de dar n_1 pasos ala derecha y n_2 pasos ala izquierda en cualquier orden se obtendrá al multiplicar la probabilidad de obtener una secuencia individual por el número de secuencias.

$$W_N(n_1) = \frac{N!}{n_1!n_2!}p^{n_1}q^{n_2} \tag{1.83}$$

Dado que:

$$n_1 = \frac{1}{2}(N+m), n_2 = \frac{1}{2}(N-m) \tag{1.84}$$

Para obtener la probabilidad $P_N(m)$ de que la particula se encuentre en m despúes de N pasos.

$$P_N(m) = W_N(n_1) \tag{1.85}$$

Sustituimos n_1 y n_2 en $W_N(n_1)$ para dejar la expresión en términos de N y m:

$$W_N(n_1) = \frac{N!}{(\frac{1}{2}(N+m))!(\frac{1}{2}(N-m))!} p^{\frac{1}{2}(N+m)} q^{\frac{1}{2}(N-m)} \tag{1.86}$$

Ademas sabemos que $p+q=1$ entonces $q=1-p$, por lo tanto:

$$P_N(m) = W_N(n_1) = \frac{N!}{(\frac{1}{2}(N+m))!(\frac{1}{2}(N-m))!} p^{\frac{1}{2}(N+m)} (1-p)^{\frac{1}{2}(N-m)} \tag{1.87}$$

1.22 Demostraciones de las ecuaciones 23 a 26

1.22.1 Ecuación 23

Demuestre que $\Delta u = 0$

$$\Delta u = u - \overline{u} = 0 \tag{1.88}$$

$$\Delta u = \overline{u - \overline{u}} = \overline{u} - \overline{u} = 0 \tag{1.89}$$

$$\overline{u} = \frac{\Sigma u}{|u|} \tag{1.90}$$

$$\overline{\overline{u}} = \frac{\overline{u}}{1} \tag{1.91}$$

$$\tag{1.92}$$

$$\overline{u - \overline{u}} = \frac{\Sigma u - |u|\overline{u}}{|u|} = \frac{\Sigma u}{|u|} - \overline{u} = \overline{u} - \overline{u} = 0 \tag{1.93}$$

1.22.2 Ecuación 24

Demuestre:

$$\overline{(\Delta u)^2} = \Sigma_{i=0}^{M} P(u_i)(u_i - \overline{u})^2 \geq 0 \tag{1.94}$$

Se sabe que $P(u_i)$ es una manera de agregar peso a u_i y que por lo general es un valor positivo. También se sabe que para toda $u \in R$, $u^2 = (-u)^2 \geq 0$, por lo tanto queda demostrado.

1.22.3 Ecuación 25

Demuestre:

$$\overline{(u - \overline{u})^2} = \overline{u^2} - \overline{u}^2 \tag{1.95}$$

$$\tag{1.96}$$

$$\overline{(u - \overline{u})^2} = \frac{\Sigma(u - \overline{u})^2}{|u|} \tag{1.97}$$

$$= \frac{\Sigma u^2 - 2u\overline{u} + \overline{u}^2}{|u|} \tag{1.98}$$

$$= \overline{u^2} - 2\overline{u}^2 + \overline{u}^2 \tag{1.99}$$

$$= \overline{u^2} - \overline{u}^2 \tag{1.100}$$

1.22.4 Ecuación 26

Demostrar que $\overline{u^2} \geq \overline{u}^2$.

$$\tag{1.101}$$

$$\overline{u}^2 = (\frac{\Sigma u}{|u|})^2 = \frac{(\Sigma u)^2}{(|u|)^2} \tag{1.102}$$

$$\overline{u^2} = \frac{(\Sigma u)^2}{|u|} \tag{1.103}$$

Para el caso en el que solo hay un elemento:

$$\frac{\Sigma u^2}{|u|} = \frac{(\Sigma u)^2}{(|u|)^2} \tag{1.104}$$

De la ecuación :

$$\overline{(u-\overline{u})^2} = \overline{u^2} - \overline{u}^2 \tag{1.105}$$

Dado que el lado izquierdo de la ecuación siempre es positivo, se concluye que : $\overline{u^2} \geq \overline{u}^2$ Q.E.D.

1.23 Justificación de la ecuación 29

Justificar que :

$$\Sigma_{n_1=0}^{N} W_N(n_1) = 1 \tag{1.106}$$

Una vez que sustituimos:

$$W_N(n_1) = \frac{N!}{n_1!(N-n_1)!} p^{n_1} q^{N-n_1} \tag{1.107}$$

En:

$$\Sigma_{n_1=0}^{N} W_N(n_1) = 1 \tag{1.108}$$

Se obtiene:

$$\Sigma_{n_1=0}^{N} \frac{N!}{n_1!(N-n_1)!} p^{n_1} q^{N-n_1} = (p+q)^{\frac{1}{2}} \tag{1.109}$$

Por el teorema del binomio, dado que $q = 1-p$, $p+q = 1$, por lo tanto $(p+q)^{\frac{1}{2}} = 1$.

1.24 Explicación de la ecuación 31

Por definición:

$$\overline{n_1} = \Sigma_{n_1=0}^{N} W(n_1) n_1 = \Sigma_{n_1=0}^{N} \frac{N!}{n_1!(N-n_1)!} p^{n_1} q^{N-n_1} n_1 \tag{1.110}$$

Se puede apreciar que el término n_1 se puede obtener si:

$$n_1 p = p \frac{\partial}{\partial p} p^{n_1} \tag{1.111}$$

Ya que al aplicar el operador se obtiene:

$$p n_1 p^{n_1 - 1} = (n_1) p^{n_1} \tag{1.112}$$

Si sustituimos en la definición anterior en $\overline{n_1}$:

$$\Sigma_{n_1=0}^{N} \frac{N!}{n_1!(N-n_1)!} p \frac{\partial}{\partial p} p^{n_1} q^{N-n_1}$$

Como la suma está en terminos de n_1, se puede hacer que:

$$p \frac{\partial}{\partial p} \Sigma_{n_1=0}^{N} \frac{N!}{n_1!(N-n_1)!} p^{n_1} q^{N-n_1}$$

Esto esto es igual a:

$$p \frac{\partial}{\partial p} (p+q)^N$$

Al aplicar la derivada parcial, obtendremos:

$$p N (p+q)^{N-1} \tag{1.113}$$

1.25 Justificación de la ecuación 33

Justificar $\overline{n_2} = Nq$. De la expresión $pN(p+q)^{N-1}$, se sabe que $p+q=1$, por lo tanto $\overline{n_1} = Np$. Para obtener la ecuación 33:

$$q\frac{\partial}{\partial q}\Sigma_{n_2=0}^{N}\frac{N!}{n_2!(N-n_2)!}q^{n_2}p^{N-n_2}$$

De aquí:

$$q\frac{\partial}{\partial q}(p+q)^N$$

Al aplicar la derivada parcial:

$$\overline{n_2} = Nq \qquad (1.114)$$

1.26 Deducción de la ecuación 38

Encuentre $\overline{(n_1)^2}$.

Por definición:

$$\overline{(n_1)^2} = \Sigma_{n_1=0}^{N}\frac{N!}{n_1!(N-n_1)!}p^{n_1}q^{N-n_1}(n_1)^2 \qquad (1.115)$$

Partiendo de considerar verdadera la expresión:

$$(n_1)^2p^{n_1} = n_1(p\frac{\partial p}{)}p^{n_1} \qquad (1.116)$$

Entonces, obtenemos la expresión:

$$\overline{(n_1)^2} = (p\frac{\partial p}{)}^2 \Sigma_{n_1=0}^{N} \frac{N!}{n_1!(N-n_1)!} p^{n_1} q^{N-n_1} \tag{1.117}$$

De la cual

$$\overline{(n_1)^2} = (p\frac{\partial p}{)}(pN(p+q)^{N-1}) = p(pN(N-1)(p+q)^{N-2} + N(p+q)^{N-1})$$

Recordando que $p+q=1$,

$$\begin{aligned} \overline{(n_1)^2} &= p^2(N^2-N) + pN \\ &= Np(1+pN-p) = (Np)^2 - Npq \end{aligned}$$

1.27 Ecuaciones 39 y 40

El cálculo de la raíz cuadrática media (R.M.S.) $\Delta * n_1 \equiv [\overline{(\Delta n_1)^2}]^{\frac{1}{2}}$ es una medida linear sobre el rango en el cual n_1 está distribuido. Una buena medida de la anchura relativa de la distribución es:

$$\frac{\Delta * n_1}{\overline{n_1}} = \frac{\sqrt{Npq}}{Np} = \sqrt{\frac{q}{p}} \frac{1}{\sqrt{N}} \tag{1.118}$$

Para el caso en el que $p=q$.

$$\frac{\Delta * n_1}{\overline{n_1}} = \frac{1}{\sqrt{N}} \tag{1.119}$$

1.27.1 RMS- Valor cuadrático medio

El valor cuadrático medio o RMS es una medida estadística de la magnitud de una cantidad variable. Puede calcularse para una serie de valores discretos o para una función matemática de variable continua. El nombre deriva del hecho de que es la raíz cuadrada de la media aritmética de los cuadrados de los valores.

A veces la variable toma valores positivos y negativos, como ocurre, por ejemplo, en los errores de medida. En tal caso se puede estar interesado en obtener un promedio que no recoja los efectos del signo. Este problema se resuelve, mediante la denominada media cuadrática. Consiste en elevar al cuadrado todas las observaciones (así los signos negativos desaparecen), en obtener después su media aritmética y en extraer, finalmente, la raíz cuadrada de dicha media para volver a la unidad de medida original.

$$x_{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2} \tag{1.120}$$

1.28 Análisis de componentes principales (PCA. Principal Component Analysis)

El Análisis de Componentes Principales (PCA. Principal Component Analysis) es una técnica estadística de síntesis de la información, o reducción de la dimensión (número de variables). Es decir, ante un banco de datos con muchas variables, el objetivo será reducirlas a un menor número perdiendo la menor cantidad de información posible.[13]

El método de componentes principales tiene por objeto transformar un conjunto de variables, a las que se denomina originales, en un nuevo conjunto de variables denominadas componentes principales. Estas últimas se caracterizan por estar incorrelacionadas entre sí y, además, pueden ordenarse de acuerdo con la información que llevan incorporada. Como medida de la cantidad de información incorporada en una componente se utiliza su varianza.

Es decir, cuanto mayor sea su varianza mayor es la cantidad de información que lleva incorporada dicha componente. Por esta razón se selecciona como primera componente aquella que tenga mayor varianza, mientras que la última componente es la de menor varianza.

En general, la extracción de componentes principales se efectúa sobre variables normalizadas para evitar problemas derivados de la escala, aunque también se puede aplicar sobre variables expresadas en desviaciones respecto a la media.

Si p variables están normalizadas, la suma de las varianzas es p, ya que la varianza de una variable normalizada es por definición 1.

El nuevo conjunto de variables que se obtiene por el método de componentes principales es igual en número al de las variables originales. Es importante destacar que la suma de sus varianzas es igual a la suma de las varianzas de las variables originales. La diferencia entre ambos conjuntos de variables estriba en que las componentes principales se calculan de forma que estén incorrelacionadas entre sí. Cuando las variables originales están muy correlacionadas entre sí, la mayor parte de su variabilidad se puede explicar con muy pocas componentes. Si las variables originales estuvieran completamente incorrelacionadas entre sí, el análisis de componentes principales carecería por completo de interés, ya que en ese caso las componentes principales coincidirían con las variables originales. Las componentes principales se expresan como una combinación lineal de las variables originales. Desde el punto de vista de su aplicación, el método de componentes principales es considerado como un método de reducción, esto es, un método que permite reducir la dimensión del número de variables originales que se han considerado en el análisis.[13]

La reducción de muchas variables a pocas componentes puede simplificar la aplicación sobre estas últimas de otras técnicas multivariantes (regresión, clusters, etc.). El método de componentes principales puede considerarse como un método para la reducción de datos, tratar otros problemas como el de la rotación de factores, contrastes, etc. se hace dentro del análisis factorial que implica una mayor formalización. En este sentido, el método de componentes principales se inscribe dentro de la estadística descriptiva.

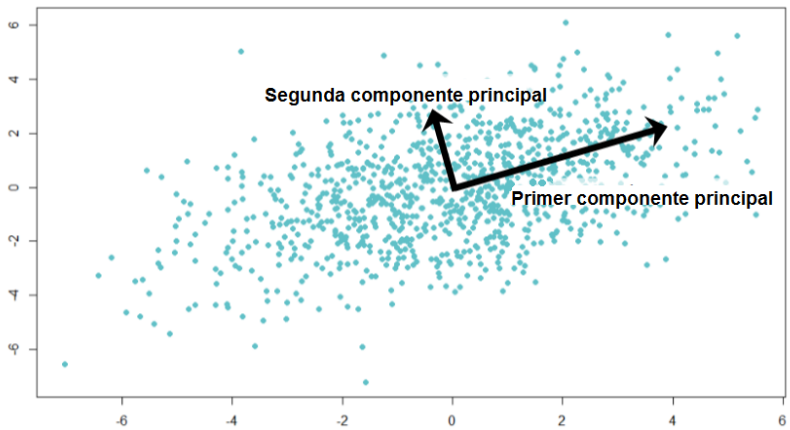


Figure 1.11: Componentes principales

1.28.1 Obtención de las componentes principales

En el análisis de componentes principales se dispone de una muestra de tamaño n acerca de p variables $X_1, X_2, ..., X_p$ (normalizadas o expresadas en desviaciones respecto a la media) inicialmente correlacionadas, para posteriormente obtener a partir de ellas un número $k \leq p$ de variables incorrelacionadas $Z_1, Z_2, ..., Z_k$ que sean combinación lineal de las variables iniciales y que expliquen la mayor parte de su variabilidad.

La **primera componente principal**, al igual que las restantes, se expresa como combinación lineal de las variables originales:

$$Z_{1i} = u_{11}X_{1i} + u_{12}X_{2i} + ... + u_{1p}X_{pi} \tag{1.121}$$

Para el conjunto de las n observaciones muestrales, la ecuación puede expresarse matricialmente:

$$\begin{pmatrix} Z_{11} \\ Z_{12} \\ ... \\ Z_{1n} \end{pmatrix} = \begin{pmatrix} X_{11} & X_{21} & ... & X_{p1} \\ X_{12} & X_{22} & ... & X_{p2} \\ ... & ... & ... & ... \\ X_{1n} & X_{2n} & ... & X_{pn} \end{pmatrix} \begin{pmatrix} u_{11} \\ u_{12} \\ ... \\ u_{1p} \end{pmatrix} \tag{1.122}$$

En notación abreviada:

$$Z_1 = Xu_1$$

Tanto si las X_j están tipificadas, como si están expresadas en desviaciones respecto de su media muestral, la media de Z_1 es cero, es decir, $E(Z_1) = E(Xu_1) = E(X)u_1 = 0$ La varianza de Z_1 será:

$$V(Z_1) = \frac{1}{n}\Sigma_{i=1}^{n} Z_{1i}^2 = \frac{1}{n} Z_1^{'} Z_1 = \frac{1}{n} u_1^{'} X^{'} X u_1 = u_1^{'} (\frac{1}{n} X^{'} X) u_1 = u_1^{'} V u_1 \tag{1.123}$$

Si las variables están expresadas en desviaciones respecto a la media, la expresión $\frac{1}{n}X^{'}X$ (matriz de inercia), es la matriz de covarianzas muestral a la que se denomina V (caso mas general) y para variables normalizadas $\frac{1}{n}X^{'}X$ es la matriz de correlaciones R.

La primera componente Z se obtiene de forma que su varianza sea máxima, sujeta a la 1 restricción de que la suma de los pesos u_{1j} al cuadrado sea igual a la unidad, es decir, la variable de los pesos o ponderaciones $(u_{11}, u_{12}, ..., u_{1p})^{'}$ se encuentra normalizada.

Se trata de hallar Z_1 maximizando $V(Z_1) = u_1^{'} V u_1$ con las restricción $\Sigma_{j=1}^{p} u_1^{'} u_1 = 1$ para afrontar el problema de maximización con restricciones se aplica el método de los multiplicadores de Lagrange, considerando la función Lagrangiana:

$$L = u_1^{'} V u_1 - \lambda (u_1^{'} u_1 - 1) \tag{1.124}$$

derivando respecto a u_1 e igualando a cero:

$$\frac{\partial L}{\partial u_1} = 2Vu_1 - 2\lambda u_1 = 0 \tag{1.125}$$

$$(V - \lambda I)u_1 = 0 \tag{1.126}$$

Se trata de un sistema homegéneo en u_1 que sólo tiene solución si el determinante de la matriz de los coeficientes es nulo: $\left|(V - \lambda I)u_1\right| = 0$. Lo cual es equivalente a decir que $\left|(V - \lambda I)u_1\right| = 0$ es un valor propio de la matriz V.

En general la ecuación $|(V - \lambda I)u_1| = 0$ tiene n raíces $\lambda_1, \lambda_2, ..., \lambda_n$, ordenadas de mayor a menor $\lambda_1 > \lambda_2 > ... > \lambda_n$.

Si en la ecuación $(V - \lambda I)u_1 = 0$ se multiplica a la derecha por $u_1^{'}$ se obtiene:

$$u_1^{'}(V - \lambda I)u_1 = 0$$

$$u_1^{'}Vu_1 = \lambda$$

$$V(Z_1) = u_1^{'}Vu_1 = \lambda$$

Por lo tanto para maximizar $V(Z_1)$, hayq ue tomar el mayor valor propio λ de la matriz V.

Tomando λ_1 como el mayor valor propio de V y tomando u_1 como su vector propio asociado normalizado $u_1^{'}u_1 = 1$, se define el vector de ponderaciones que se aplica a las variables iniciales para obtener la **primera componente principal**, componente que se expresa:

$$Z_1 = Xu_1 \tag{1.127}$$

La **segunda componente principal**, al igual que las restantes, se expresa como combinación lineal de las variables originales:

$$Z_{2i} = u_{21}X_{1i} + u_{22}X_{2i} + ... + u_{2p}X_{pi} \tag{1.128}$$

Para el conjunto de las n observaciones muestrales, la ecuación puede expresarse matricialmente:

$$\begin{pmatrix} Z_{21} \\ Z_{22} \\ ... \\ Z_{2n} \end{pmatrix} = \begin{pmatrix} X_{11} & X_{21} & ... & X_{p1} \\ X_{12} & X_{22} & ... & X_{p2} \\ ... & ... & ... & ... \\ X_{1n} & X_{2n} & ... & X_{pn} \end{pmatrix} \begin{pmatrix} u_{21} \\ u_{22} \\ ... \\ u_{2p} \end{pmatrix} \tag{1.129}$$

En notación abreviada:

$$Z_2 = Xu_2$$

Tanto si las X_j están tipificada, como si están expresadas en desviaciones respecto de su media muestral, la media de Z_2 es cero, es decir, $E(Z_2) = E(Xu_2) = E(X)u_2 = 0$.

La varianza de Z_2 será:

$$V(Z_2) = \frac{1}{n}\Sigma_{i=1}^{n} Z_{2i}^2 = \frac{1}{n}Z_2^{'}Z_2 = \frac{1}{n}u_2^{'}X^{'}Xu_2 = u_2^{'}(\frac{1}{n}X^{'}X)u_2 = u_2^{'}Vu_2 \tag{1.130}$$

La segunda componente Z_2 se obtiene de forma que su varianza sea máxima, sujeta a la restricción de que la suma de los pesos u_{2j} al cuadrado sea igual a la unidad, es decri la variable de los pesos o ponderaciones $(u_{21}, u_{22}, ..., u_{2p})^{'}$ se encuentra normalizada $(u_2^{'}u_2 = 1)$.

Por otra parte como Z_1 y Z_2 han de estar incorreladas se tienen que verificar:

$$0 = E(Z_2^{'}Z_1) = E(u_2^{'}X^{'}Xu_1) = u_2^{'}E(X^{'}X)u_1 = u_2^{'}Vu_1 \tag{1.131}$$

Se sabe que $Vu_1 = \lambda_1 u_1$(dado que u_1 es el vector propio de V asociado a su mayor valor propio λ_1. Multiplicando por $u_2^{'}$ se tiene:

$$u_2^{'}Vu_1 = u_2^{'}\lambda_1 u_1 = 0$$

$$u_2^{'}u_1 = 0$$

$$u_2 \perp u_1 \text{ Son ortogonales} \tag{1.132}$$

Hay que hallar Z_2 maximizando $V(Z_2) = u_2^{'}Vu_2$ con las restricciones $u_2^{'}u_2 = 1$ y $u_2^{'}Vu_1 = 0$ Aplicando el método de los multiplicadores de Lagrange, se considera la función:

$$L = u_2^{'}Vu_2 - 2\mu(u_2^{'}Vu_1) - \lambda(u_2^{'}u_2 - 1) \tag{1.133}$$

Derivando respecto a u_2 e igualando a cero:

$$\frac{\partial L}{\partial u_2} = 2Vu_2 - 2\mu V u_1 - 2\lambda u_2 = 0 \tag{1.134}$$

Dividiendo la expresión anterior por 2 y multiplicando por $u_1^{'}$ resulta:

$$u_1^{'} V u_2 - \mu u_1^{'} V u_1 - \lambda u_1^{'} u_2 = 0 \tag{1.135}$$

Siendo $Vu_1 = \lambda u_1 \Rightarrow V u_1^{'} = \lambda u_1^{'}$ y $V(Z_1) = u_1^{'} V u_1$, resulta:

$$\lambda u_1^{'} u_2 - \mu V(Z_1) - \lambda u_1^{'} u_2 = 0 \tag{1.136}$$

Como $u_1^{'} u_2 = 0$, $\mu V(Z_1) = 0$ y $\mu = 0$.

De donde:

$$\tag{1.137}$$

$$\frac{\partial L}{\partial u_2} = 2Vu_2 - 2\lambda u_2 = 0 \tag{1.138}$$

$$(V - \lambda I) u_2 = 0 \tag{1.139}$$

Se trata de un sistema homogéneo en u_2 que solo tiene solución si el determinante de la matriz de los coeficientes en nulo: $|(V - \lambda I)| = 0$.

En general la ecuación $|(V - \lambda I)| = 0$ tiene n raíces $\lambda_1, \lambda_2, ..., \lambda_n$, ordenadas de mayor a menor $\lambda_1 > \lambda_2 > ... > \lambda_n$.

Si la ecuación $(V - \lambda I) u_2 = 0$ se multiplica a la derecha por $u_2^{'}$ se tiene:

$$\tag{1.140}$$

$$u_2^{'} (V - \lambda I) u_2 = 0 \tag{1.141}$$

$$u_2^{'} V u_2 = \lambda \tag{1.142}$$

$$V(Z_2) = u_2^{'} V u_2 = \lambda \tag{1.143}$$

Por lo tanto, para maximizar $V(Z_2)$ se ha de tomar el segundo mayor valor propio λ de la matriz V (el primer mayor valor propio ya lo había tomado al obtener la primera componente principal).

Tomando λ_2 como el segundo mayor valor propio de V y tomando u_2 como su vector propio asociado normalizado $(u_2^{'}u_2) = 1$, ya se ha definido el vector de ponderaciones que se aplica a las variables iniciales para obtener la segunda componente principal, componente que vendrá definida como:

$$Z_2 = Xu_2 \tag{1.144}$$

Análogamente, la **componente principal h-ésima** se define como$Z_h = Xu_h$, donde u_h es el vector propio de V asociado a su h-ésimo mayor valor propio.

Suele denominarse también a u_h eje factorial h-ésimo.

1.29 Proceso Estocástico

Un proceso estocástico es una colección o familia de variables aleatorias X_t,con $t \in T$ ordenadas según el subíndice t que en general se suele identificar con el tiempo.[17]

Por tanto, para cada instante t tendremos una variable aleatoria distinta representada por X_t, con lo que un proceso estocástico puede interpretarse como una sucesión de variables aleatorias cuyas características pueden variar a lo largo del tiempo.

Por ejemplo, si observamos sólo unos pocos valores de t, tendríamos una imagen similar a la de la figura siguiente:

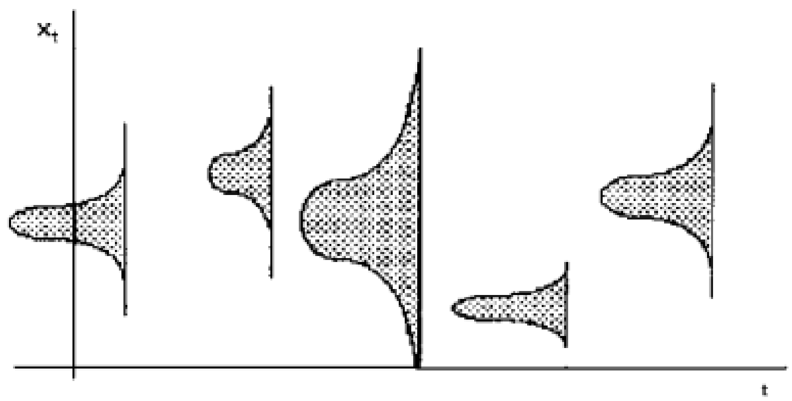


Figure 1.12: Proceso estocástico

en la que se representa para cada t la función de densidad correspondiente a X_t. Aunque en la figura se han representado unas funciones de densidad variables, un proceso estocástico no tiene por que presentar esas diferencias en la función de densidad a lo largo del tiempo.[17]

A los posibles valores que puede tomar la variable aleatoria se le denominaran estados, por lo que se puede tener un espacio de estados discreto y un espacio de estados continuo. Por otro lado, la variable tiempo puede ser de tipo discreto o de tipo continuo. En el caso del tiempo discreto se podría tomar como ejemplo que los cambios de estado ocurran cada día, cada mes, cada año, etc... En el caso del tiempo continuo, los cambios de estado se podrían realizar en cualquier instante.[1]

Por tanto, dependiendo de cómo sea el conjunto de subíndices T y el tipo de variable aleatoria dado por X_t se puede establecer la siguiente clasificación de los procesos estocásticos:

- Si el conjunto T es continuo, por ejemplo $\mathbb{R}^+$, diremos que X_t es un proceso estocástico de parámetro continuo.

- Si por el contrario T es discreto, por ejemplo $\mathbb{N}$, diremos que nos encontramos frente a un proceso estocástico de parámetro discreto.

- Si para cada instante t la variable aleatoria T es de tipo continuo, diremos que el proceso estocástico es de estado continuo.

- Si para cada instante t la variable aleatoria X_t es de tipo discreto, diremos que el proceso estocástico es de estado discreto.

1.30 Criterios de divisibilidad

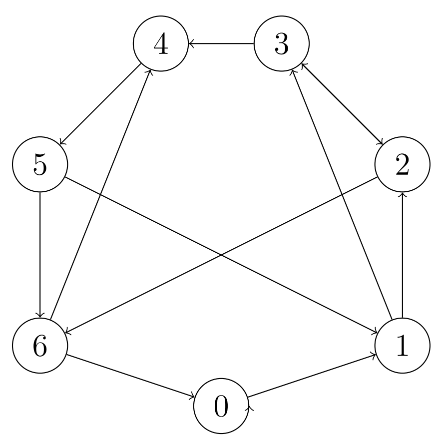
1.30.1 Divisibilidad entre 7

Existe una sencilla regla para comprobar si un número natural es divisible entre 7, que es la siguiente:

Separamos la cifra de las unidades del número inicial, la multiplicamos por 2 y se la restamos al resto del número (lo que quedó sin las unidades). Si obtenemos un múltiplo de 7 entonces el número inicial es múltiplo de 7, y si obtenemos un nḿero que no es múltiplo de 7 pues el inicial tampoco lo es. Si obtenemos un número demasiado grande y no sabemos si es múltiplo de 7 o no, repetimos el proceso anterior las veces necesarias hasta que lleguemos a un número del que sepamos si es o no múltiplo de 7.

El grafo de la divisibilidad entre 7

Es un grafo a partir del cual no sólo podemos saber si un número es divisible entre 7 de manera algo más rápida que con el algoritmo anterior (al menos bajo mi punto de vista) sino también el resto que deja dicha división en el caso de no serlo.



Para saber si un número natural es divisible entre 7 comenzamos en el cero, recorremos desde él tantas flechas negras como indique la primera cifra del número y después seguimos la flecha blanca que salga del punto al que hemos llegado. Tomamos la segunda cifra y hacemos lo mismo: desde el punto donde nos encontramos recorremos tantas flechas negras como indique la segunda cifra y después la flecha blanca que nos encontramos en el destino. Y así sucesivamente. Cuando lleguemos a la última cifra recorremos desde el punto donde nos encontremos tantas flechas negras como ella indique y el punto al que lleguemos nos dice el resto de dividir el número inicial entre 7.

Tomemos un número grande, digamos el 239058. Con el método anterior posiblemente tardaríamos un buen rato en comprobar si nuestro número es divisible entre 7 o no. Además no conoceríamos el resto de dicha división.

Criterio de divisibilidad del 11

Los divisores de 11 son: 1 y 11 pues es número primo.

Los primeros múltiplos de 11 son: 22, 33, 44, 55, 66, 77, 88, 99, 110, 121, 132, 143, 154, 165, 176, 187, 198, 209,..., 297, 308, 319, 330, 341, 352, 363 y muchos más.

Un número es divisible por 11 cuando la diferencia entre sus dígitos (se puede iniciar por la derecha o por la izquierda) que ocupan lugar par e impar es 0 o múltiplo de 11 .

Ejemplo 1 : Veamos si 57342 es divisible por 11. Los dígitos colocados en lugar par son: 7 y 4 y los de lugar impar 5, 3 y 2; Aplicando criterio: lugar par (7 + 4) – lugar impar (5 + 3 + 2) = 11 – 10 = 1 que no es 0 ni múltiplo de 11, luego 57342 no es divisible por 11.

Ejemplo 2: Probar que 101354 es divisible por 11. Dígitos que ocupan lugar impar (4 + 3 + 0) – dígitos par (5 + 1 +1) = 7 – 7 = 0, luego 101354 es divisible por 11.

Ejemplo 3: Determina el valor de X para que 679X251 sea divisible por 11. Fijamos el criterio: Cifras de lugar par (5 + X + 7) – cifras lugar impar (1 + 2 + 9 + 6) = 12 + X – 18 = X – 6. Para que sea divisible X – 6 = 0, =¿ X = 6, pues no puede ser 11, ya que entonces X =17, y ha de ser una sola cifra. Luego X = 6 y por tanto 6796251 es divisible por 11.

En otros temas se estudian las reglas del 2, 3, 4, 5, 10 6, del 7 del 8 del 9 del 13 15 del 17 del 19 23 29 y 43.

Criterio de divisibilidad del 3

Criterio de divisibilidad del 3: Para saber si un número es divisible entre 3, tenemos que comprobar que la suma de todos sus dígitos sea 3 o múltiplo de 3.

Por ejemplo: ¿Es 1098 divisible entre 3?

Sumamos todos los dígitos de 1098: 1 + 0 + 9 + 8 = 18 1 + 8 = 9
9 es un múltiplo de 3 por lo tanto 1098 es divisible por 3.

1.31 Congruencia Zeller

En teoría de calendarios nos enfrentamos de continuo con el problema de establecer expresiones analíticas de correspondencias entre números enteros que tienen la propiedad de ser muy cercanamente lineales. Por ejemplo, la correspondencia entre el año de un ciclo de cuatro años del calendario juliano y el número de días transcurridos es lineal, salvo la divergencia que se produce con los años bisiestos que hace aumentar en uno el número de días transcurridos. La conveniencia de establecer una relación analítica de una aplicación numérica como la anterior, es principalmente para su uso en programas informáticos, eludiendo de esta forma las tablas que se usaron en el pasado.[González]

Christian Zeller en una serie de trabajos publicados a final del siglo XIX hizo uso de este tipo de relaciones para el cálculo del día semanal de una fecha y para la obtención de la fecha de la Pascua. Tanto Zeller como otros despúes de él, obtuvieron estas relaciones por tanteo. Nosotros hemos dado forma al teorema, que hemos llamado de Zeller en honor a Christian Zeller y lo hemos usado sistemáticamente en cálculos calendaristas.[González]

$$h = (q + (2.6(m+1)) + K + (\frac{K}{4}) + (\frac{J}{4}) - 2J) mod 7 \qquad (1.145)$$

Donde:

- h = día de la semana (0 sábado, 1 domingo, 2 lunes,...)
- q número del día del mes
- m número del mes
- K año de la centuria
- J centuria

Capítulo 2

Prácticas

2.1 Triángulo de Sierpinski

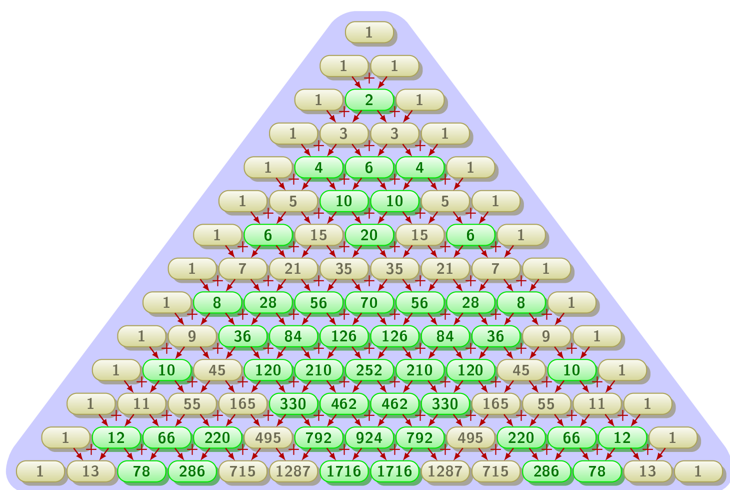


Figure 2.1: Triángulo de Sierpinski resaltando las posiciones múltiplos de 2

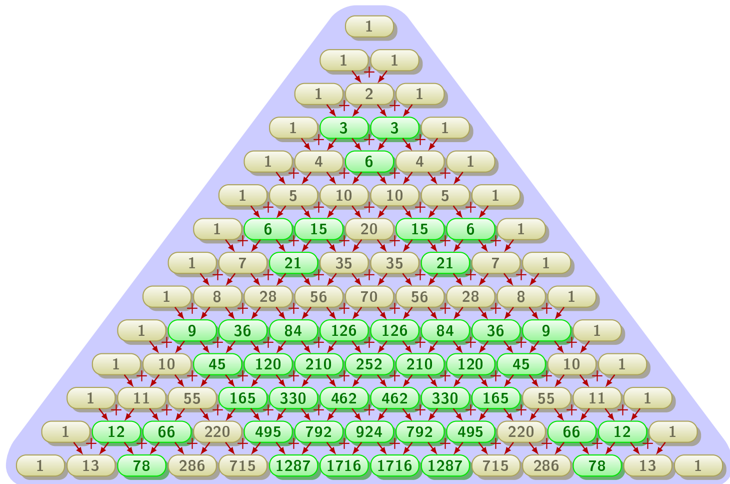


Figure 2.2: Triángulo de Sierpinski resaltando las posiciones múltiplos de 3

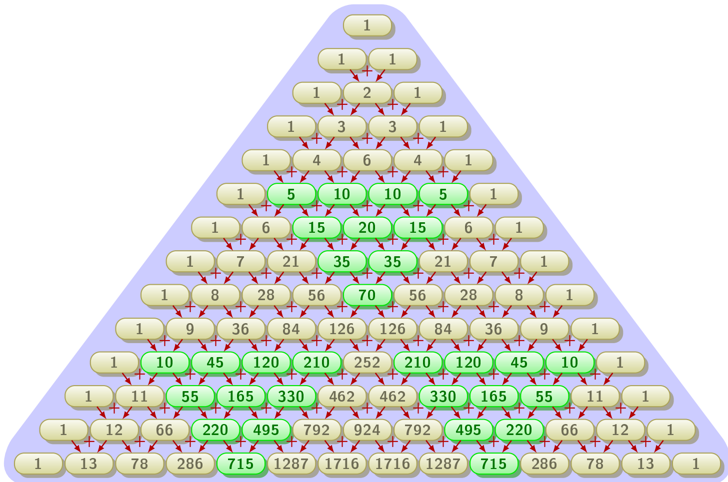


Figure 2.3: Triángulo de Sierpinski resaltando las posiciones múltiplos de 5

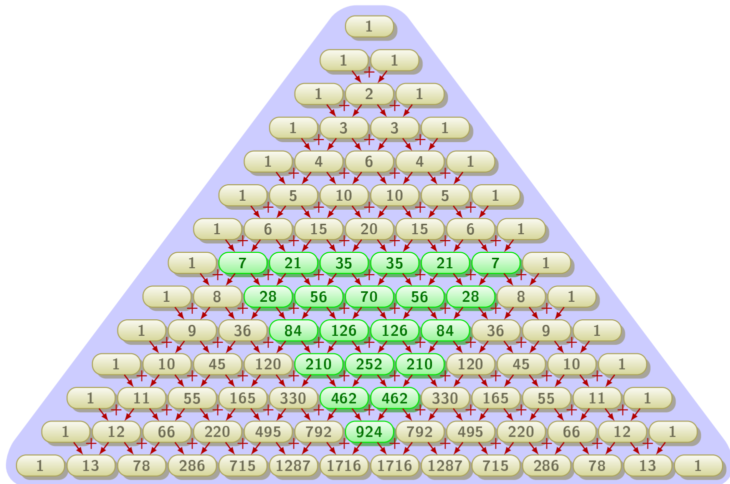


Figure 2.4: Triángulo de Sierpinski resaltando las posiciones múltiplos de 7

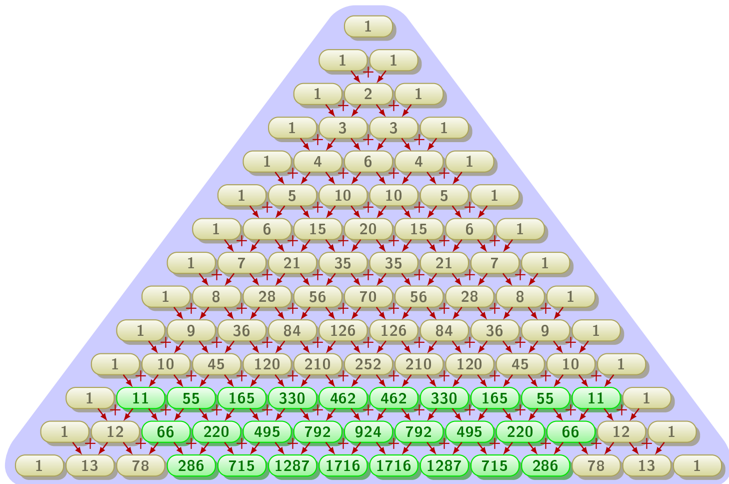


Figure 2.5: Triángulo de Sierpinski resaltando las posiciones múltiplos de 11

2.2 Graficación de lanzamientos aleatorios (Gráficas en ROOT)

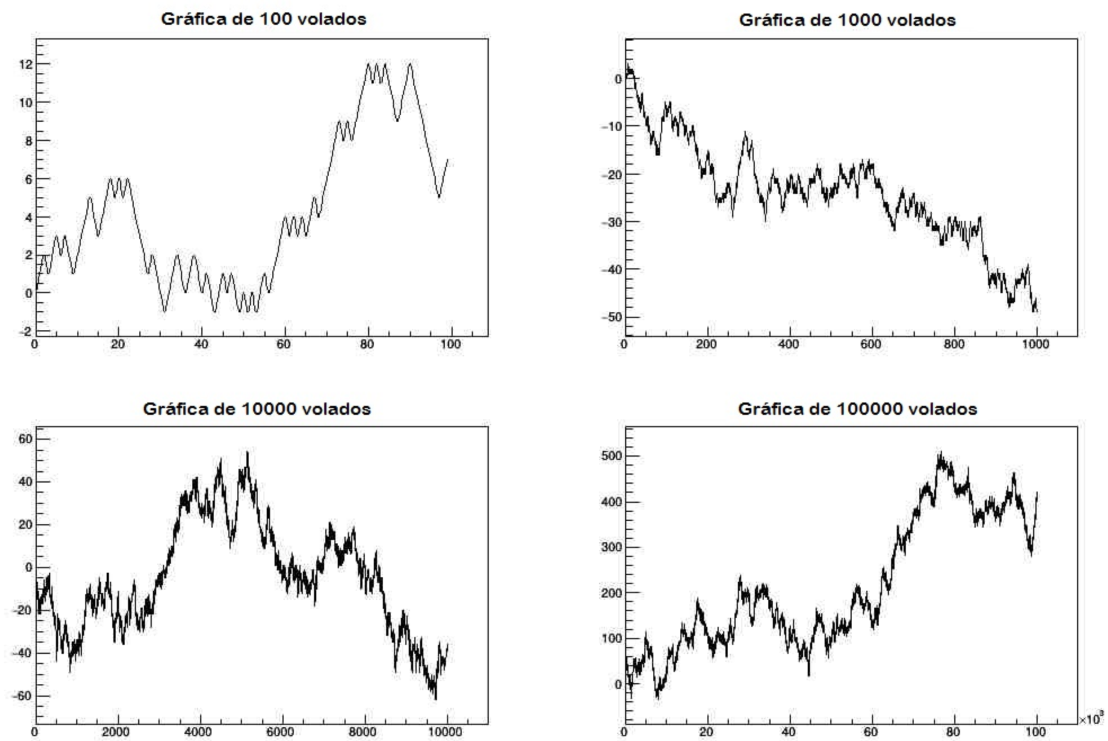


Figure 2.6: Gráfica de experimentos de volados aleatorios

2.3 Permutaciones

Programa en lenguaje C para encontrar las permutaciones posibles del conjunto de elementos, en este caso se muestra como conjunto de ejemplo, las letras "A,B,C,D".

```
#include <stdio.h>
#include <string.h>
#include <stdlib.h>
#include <iostream>

void permutaciones(char *conjunto, size_t card, size_t elem);

int main(void)
{
   char conjunto[] = "ABCD";

   size_t card = sizeof conjunto - 1;

   permutaciones(conjunto, card, 0);

   return 0;
}

void permutaciones(char *conjunto, size_t card, size_t elem)
{char temp;
   if (card > 1){
      int i;

      permutaciones(conjunto, card - 1, elem + 1);

      for (i = 1; i < card; i++){


         temp = conjunto[elem + i];
         memmove(conjunto + elem + 1, conjunto + elem, i);
         conjunto[elem] = temp;

         permutaciones(conjunto, card - 1, elem + 1);

         memmove(conjunto + elem, conjunto + elem + 1, i);
         conjunto[elem + i] = temp;
      }
   }else
      puts(conjunto);
}
```

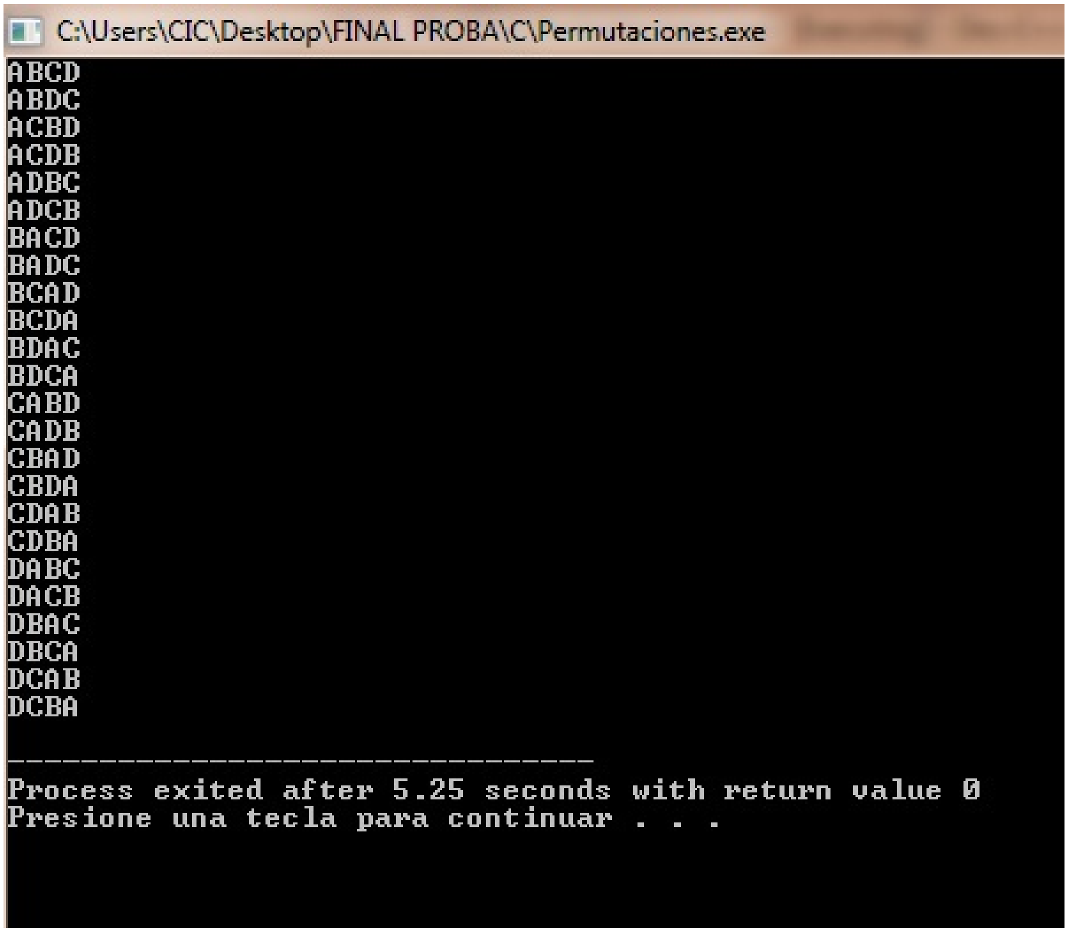


Figure 2.7: Permutaciones posibles del conjunto de elementos ABCD.

2.4 Caso general de probabilidad condicional en árbol

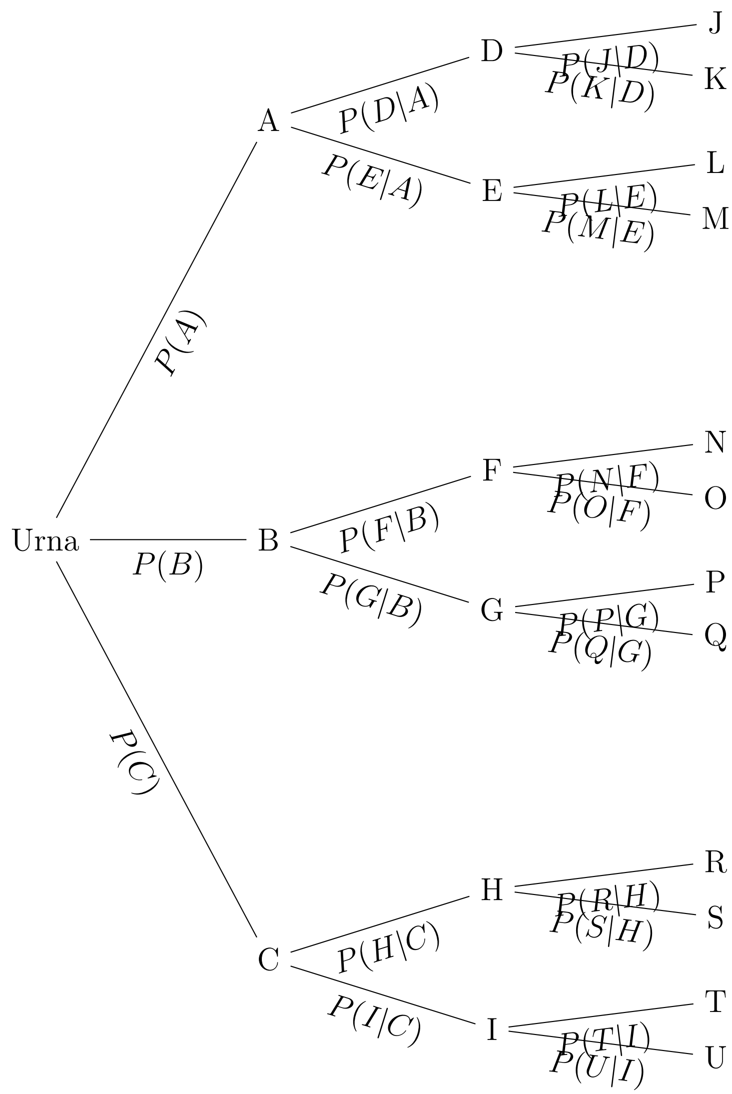


Figure 2.8: Diagrama de árbol que describe un problema de probabilidad condicional.

Probabilidad de ocurrencia de cada uno de los puntos del árbol.

$D : P(A \cap D) = P(A)P(D|A)$
$E : P(A \cap E) = P(A)P(E|A)$
$F : P(B \cap F) = P(B)P(F|B)$
$G : P(B \cap G) = P(B)P(G|B)$
$H : P(C \cap H) = P(C)P(H|C)$
$I : P(C \cap I) = P(C)P(I|C)$
$J : P(A \cap D \cap J) = P(A)P(D|A)P(J|A \cap D)$
$K : P(A \cap D \cap K) = P(A)P(D|A)P(K|A \cap D)$
$L : P(A \cap E \cap L) = P(A)P(E|A)P(L|A \cap E)$
$M : P(A \cap E \cap M) = P(A)P(E|A)P(M|A \cap E)$
$N : P(B \cap F \cap N) = P(B)P(F|B)P(N|B \cap F)$
$O : P(B \cap F \cap O) = P(B)P(F|B)P(O|B \cap F)$
$P : P(B \cap G \cap P) = P(B)P(G|B)P(P|B \cap G)$
$Q : P(B \cap G \cap Q) = P(B)P(G|B)P(Q|B \cap G)$
$R : P(C \cap H \cap R) = P(C)P(H|C)P(R|C \cap H)$
$S : P(C \cap H \cap S) = P(C)P(H|C)P(S|C \cap H)$
$T : P(C \cap I \cap T) = P(C)P(I|C)P(T|C \cap I)$
$U : P(C \cap I \cap U) = P(C)P(I|C)P(U|C \cap I)$

2.5 La conjetura de Gilbreath: Números Primos

```
In[9]:= With[{k = 10},
        con
          NestList[Abs[Subtract @@@ Partition[#, 2, 1]] &, Table[Prime[n], {n, k}], k - 1]
          lista de re··· val·· resta    particiona                tabla  número primo
        ] // TableForm
              forma de tabla
```

Out[9]//TableForm=

2	3	5	7	11	13	17	19	23	29
1	2	2	4	2	4	2	4	6	
1	0	2	2	2	2	2	2		
1	2	0	0	0	0	0			
1	2	0	0	0	0				
1	2	0	0	0					
1	2	0	0						
1	2	0							
1	2								
1									

Figure 2.9: Triángulo de números primos, desglose de la conjetura de Gilbreath

Capítulo 3

Cultura General

3.1 Los Simpson: Bart es un genio

Bart el genio. Ese fue el primer episodio genuino de Los Simpson, donde se estrenó la famosa y característica secuencia de inicio y donde debutó la famosa frase de Bart "Multiplícate por cero". Y lo más curioso de todo es que "Bart, el genio" contiene una gran dosis de matemáticas.

Maggie está contruyendo una torre con sus cubos alfabéticos que dice EMCSQU. Maggie pudo llegar a representar la famosa ecuación científica de Albert Einstein $E = mc^2$.

Esa línea ($rdrr$) no solo es la solución del problema, sino también la supuesta gracia del chiste. Tenemos dos cuestiones: por qué es tan divertido r dr r, y por qué es la respuesta al problema de matemáticas? La clase se ríe porque r dr r suena en inglés como har-dehar-har, una expresión que se ha usado para indicar una risa sarcástica como reacción a un chiste malo.

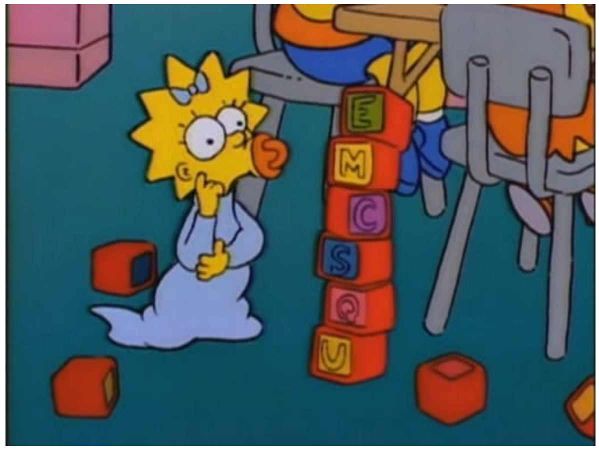


Figure 3.1: Imagen del segundo episodio de la primera temporada de la serie animada de televisión Los Simpson

Los cuadrados palindrómicos son sencillamente cuadrados que son iguales ya se lean desde un lado o desde el otro, como el 121(112), o 5221225(22852). Aquel niño de ocho años 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 698896(8362) 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.

Si nos basamos en Euclides para resolver este problema, podemos partir de la siguiente afirmación:

Supongamos que el número de primos es finito, y que todos esos números primos se han recopilado en una lista: $p_1, p_2, p_3, ..., p_n$.

Podemos explorar las consecuencias de esa afirmación multiplicando todos los primos de la lista y añadiendo 1, con lo cual crearemos un nuevo número $N = p_1p_2p_3...p_n + 1$. Este nuevo número N es un número primo o un número no primo, pero en cualquier caso, contradice la afrimación de Euclides.

- Si N es un número primo, entonces no está en la lista original. Por lo tanto, la afirmación de que hay una lista completa es falsa.
- Si N no es un número primo, entonces debe tener divisores. Esos divisores tienen que ser primos, por que los primos de la lista original dejarán un residuo de 1 cuando se dividan por N. Por lo tanto, de nuevo la afirmación de que hay una lista completa es falsa. En resumen, ¿la afirmación de Euclides es falsa?, su lista finita no contiene todos los números primos. Además, cualquier intento de reparar esa afirmación añadiendo más números primos a la lista está condenado al fracaso, por que todo el argumento se puede repetir, demostrando que la lista con números primos añadidos sigue sin estar completa. Este argumento prueba que cualquier

lista de números primos estará siempre incompleta, significa que debe de haber un número infinito de primos.

3.2 Hormigas cuentakilometros

Hormigas del desierto del Sahara cuentan con podómetros "incluídos" que les permiten calcular distancias exactas y usan la dirección del Sol para orientarse. Experimentos muestran que al agregarles extensiones a sus patas, al intentar regresar al hormiguero se pasaban de largo. Mientras que si se les recortaban, no alcanzaban a llegar. Esto sugiere que el tamaño de sus extremidades es de suma importancia.

Figure 3.2: Hormigas del desierto del Sahara

3.3 Animales que cuentan

Hay animales como los pericos y ardillas que pueden ser entrenados para contar. Otros animales muestran comportamientos en los que se muestra un cierto tipo de capacidad para contar. Hay quienes han entrenado a primates para distinguir números. Tal es el caso de Tesutsuro Matsuzawa quien entreno a un chimpancé para que reconociera números del 1 al 6, o Michael Beran quien entreno a 2 chimpancés para que hicieran match de un numero con una serie de puntos. Despúes de 3 años, estos pudieron hacerlo pero con un mayor margen de error.

Figure 3.3: Experimento chimpancés

3.4 Números primos generados por cigarras

Algunas cigarras muestran un comportamiento muy peculiar, en la que salen a la superficie para reproducirse en periodos de tiempo de 13 o 17 años. Posiblemente debido a una evolución para evitar ser devorados por sus depredadores con periodos de vida más cortos.

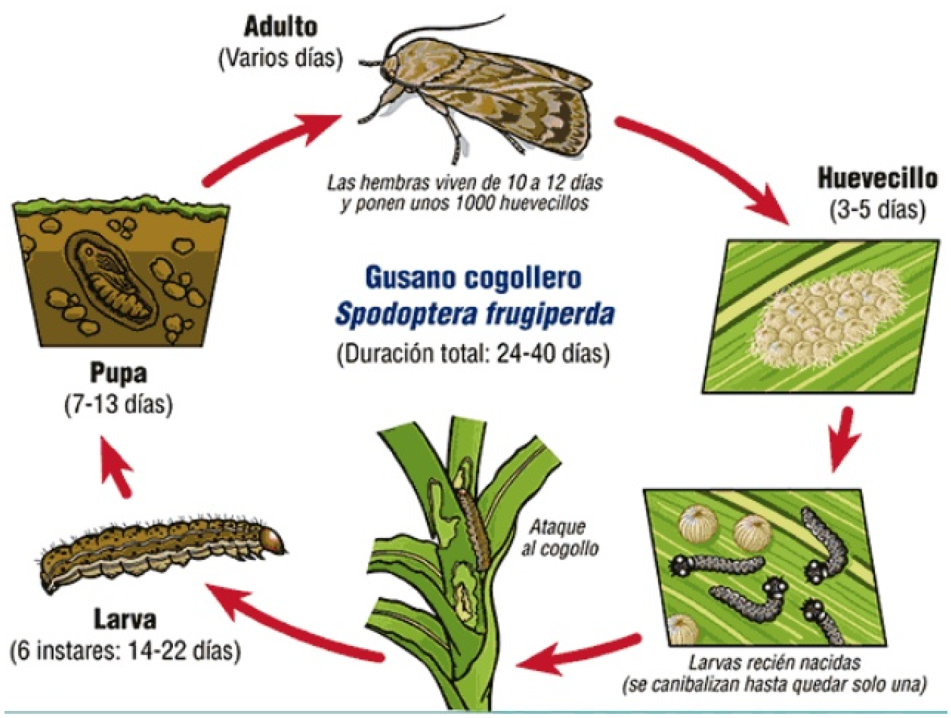


Figure 3.4: Ciclo de vida de la cigarra

3.5 Hueso de Ishango

Huesos de babuino Ishango encontrados en la República democrática del Congo con marcas a lo largo del mismo han causado especulación con respecto al hecho de si quienes los utilizaban los usaban para cosas más allá de un simple conteo. Incluso se bromea con la frase "La menstruación creo las matemáticas" al haber quien hace la hipótesis de que las marcas de los huesos era para ayudar a las mujeres a llevar un control de su ciclo menstrual.

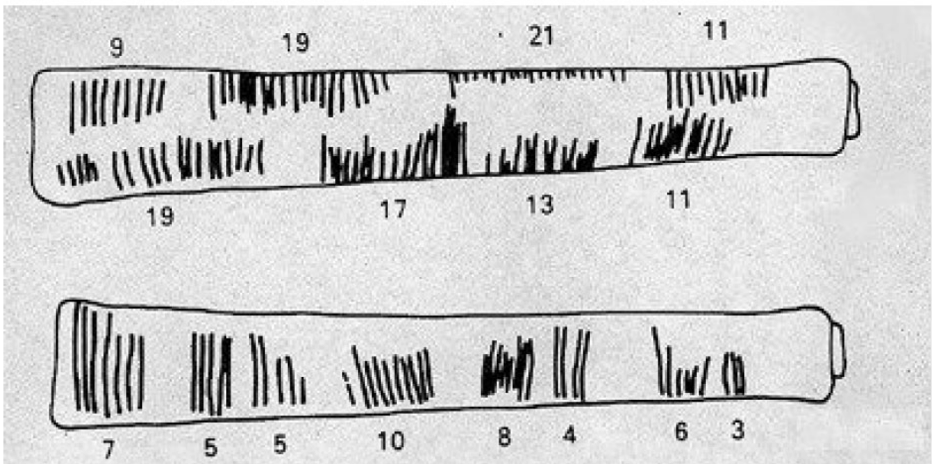


Figure 3.5: Hueso de Ishango

3.6 La maquina Enigma

Enigma era el nombre de una máquina que disponía de un mecanismo de cifrado rotatorio, que permitía usarla tanto para cifrar como para descifrar mensajes. Varios de sus modelos fueron muy utilizados en Europa desde inicios de los años 1920.

Su fama se debe a haber sido adoptada por las fuerzas militares de Alemania desde 1930. Su facilidad de manejo y supuesta inviolabilidad fueron las principales razones para su amplio uso. Su sistema de cifrado fue finalmente descubierto y la lectura de la información que contenían los mensajes supuesta mente protegidos es considerado, a veces, como la causa de haber podido concluir la Segunda Guerra Mundial al menos dos años antes de lo que hubiera acaecido sin su descifrado.

La máquina equivalente británica, Typex, y varias estadounidenses, como la SIGABA (o M-135-C en el ejército), eran similares a Enigma. La primera máquina moderna de cifrado rotatorio, de Edward Hebern, era considerablemente menos segura, hecho constatado por William F. Friedman cuando fue ofrecida al gobierno de Estados Unidos.

Figure 3.6: Ilustración:Maquina electromecánica de cifrado rotativo. Maquina Enigma

La máquina Enigma fue un dispositivo electromecánico, lo que significa que usaba una combinación de partes mecánicas y eléctricas. El mecanismo estaba constituido fundamentalmente por un teclado similar al de las máquinas de escribir cuyas teclas eran interruptores eléctricos, un engranaje mecénico y un panel de luces con las letras del alfabeto.

La parte eléctrica consistía en una batería que encendía una lámpara de una serie de ellas, que representan cada una de las diferentes letras del alfabeto.

El corazón de la máquina Enigma era mecánico y constaba de varios rotores conectados entre sí. Cada rotor es un disco circular plano con 26 contactos eléctricos en cada cara, uno por cada letra del alfabeto. Cada contacto de una cara está conectado o cableado a un contacto diferente de la cara contraria. Por ejemplo, en un rotor en particular, el contacto número 1 de una cara puede estar conectado con el contacto número 14 en la otra cara y el contacto número 5 de una cara con el número 22 de la otra. Cada uno de los cinco rotores proporcionados con la máquina Enigma estaba cableado de una forma diferente y los rotores utilizados por el ejército alemán poseían un cableado distinto al de los modelos comerciales.

Dentro de la máquina había, en la mayoría de las versiones, tres ranuras para alojar los rotores. Cada uno de los rotores se encajaba en la ranura correspondiente de forma que sus contactos de salida se conectaban con los contactos de entrada del rotor siguiente. El tercer y último rotor se conectaba, en la mayoría de los casos, a un reflector que conectaba el contacto de salida del tercer rotor con otro contacto del mismo rotor para realizar el mismo proceso pero en sentido contrario y por una ruta diferente. La existencia del reflector diferencia a la máquina Enigma de otras máquinas de cifrado de la época basadas en rotores. Este elemento, que no se incluía en las primeras versiones de la máquina, posibilitaba que la clave utilizada para el cifrado se pudiera emplear en el descifrado del mensaje. Se pueden observar en la parte superior de la imagen los tres rotores con sus correspondientes protuberancias dentadas que permitían girarlos a mano, colocándolos en una posición determinada.

Cuando se pulsaba una tecla en el teclado, por ejemplo la correspondiente a la letra A, la corriente eléctrica procedente de la batería se dirigía hasta el contacto correspondiente a la letra A del primer rotor. La corriente atravesaba el cableado interno del primer rotor y se situaba, por ejemplo, en el contacto correspondiente a la letra J en el lado contrario. Supongamos que este contacto del primer rotor estaba alineado con el contacto correspondiente a la letra X del segundo rotor. La corriente llegaba al segundo rotor y seguía su camino a través del segundo y tercer rotor, el reflector y de nuevo a través de los tres rotores en el camino de vuelta. Al final del trayecto, la salida del primer rotor se conectaba a la lámpara correspondiente a una letra, distinta de la A, en el panel de luces. El mensaje de cifrado se obtenía por tanto sustituyendo las letras del texto original por las proporcionadas por la máquina.

Cada vez que se introducía una letra del mensaje original, pulsando la tecla correspondiente en el teclado, la posición de los rotores variaba. Debido a esta variación, a dos letras idénticas en el mensaje original, por ejemplo AA, les correspondían dos letras diferentes en el mensaje cifrado, por ejemplo QL. En la mayoría de las versiones de la máquina, el primer rotor avanzaba una posición con cada letra. Cuando se habían introducido 26 letras y por tanto el primer rotor había completado una vuelta completa, se avanzaba en una muesca la posición del segundo rotor, y cuando éste terminaba su vuelta, se variaba la posición del tercer rotor. El número de pasos que provocaba el avance de cada uno de los rotores, era un parámetro con figurable por el operario.

Debido a que el cableado de cada rotor era diferente, la secuencia exacta de los alfabetos de sustitución variaba en función de qué rotores estaban instalados en las ranuras (cada máquina disponía de cinco), su orden de instalación y la posición inicial de cada uno. A estos datos se les conocía con el nombre de configuración inicial, y eran distribuidos, mensualmente al principio y con mayor frecuencia a medida que avanzaba la guerra, en libros a los usuarios de las máquinas.

El funcionamiento de las versiones más comunes de la máquina Enigma era simétrico en el sentido de que el proceso de descifrado era análogo al proceso de cifrado. Para obtener el mensaje original sólo había que introducir las letras del mensaje cifrado en la máquina, y ésta devolvía una a una las letras del mensaje original, siempre y cuando la configuración inicial de la máquina fuera idéntica a la utilizada al cifrar la información. La Máquina Enigma tiene dos características fundamentales:

1. En el texto únicamente pueden aparecer las 26 letras del alfabeto anglosajón, A-Z. No se permiten espacios en blanco (las palabras siguen unas a otras sin separación entre ellas), signos de puntuación ni números. Si se quiere incluir alguno de esos elementos en el texto se debe hacer usando una palabra o acrónimo que lo represente (PUNTO o CERO por ejemplo). A efectos prácticos esto significa que podemos (y debemos) representar un mensaje como una serie de números en el rango 0..25 (0-A, 1-B, 2-C, . . . 25-Z).

2. La Máquina Enigma es simétrica: Cuando se cifra un mensaje con una determinada clave, si el texto cifrado se vuelve a cifrar con la máquina usando la misma clave se obtiene el mensaje original. Es decir, la clave de cifrado y descifrado es la misma.

Las partes principales de la Máquina Enigma son:

3.7 Los experimentos de Mendel

Mendel descubrió los principios fundamentales de la genética gracias a sencillos y a la vez ingeniosos experimentos realizados con, variedades de plantas de arvejas de la especie Pisum sativum. El había estudiado estas plantas desde mucho antes de iniciar sus experimentos sobre la herencia y había logrado identificar varios tipos de plantas (variedades) que diferían entre sí en una o más características.

Seleccionó esta especie de arvejas porque presenta características que la hacen idónea para realizar este tipo de experimentos, como la presencia de caracteres claramente distinguibles (textura y color en la semilla), la capacidad de originar un gran número de descendientes en poco tiempo, y la de auto fecundarse (dionicas), característica que le permitió obtener individuos ”puros” para un rasgo determinado. Por ejemplo, mientras ciertas variedades presentaban solo flores blancas, otras tenían solo flores púrpura y cada una de las variedades generaba siempre individuos con las mismas características.

Mendel fue hábil y riguroso al realizar sus diseños experimentales, evitando que influyeran otras variables que pudieran perjudicar sus resultados. Se planteó analizar la herencia de siete características que, de acuerdo con sus observaciones previas, diferían de manera clara entre las variables. Protegió a los descendientes, para que en el periódo de floreción no fueran contaminados por otros tipos de polen y manipuló los híbridos, para que no tuvieran perturbaciones en su fertilidad. Además, llevó un control para que los caracteres se expresaran de la misma proporción, así como de las condiciones experimentales y utilizó una misma letra del alfabeto para determinar cada característica (Por ejemplo: $A = variantedominantes$ y $a = varianterecesiva$).

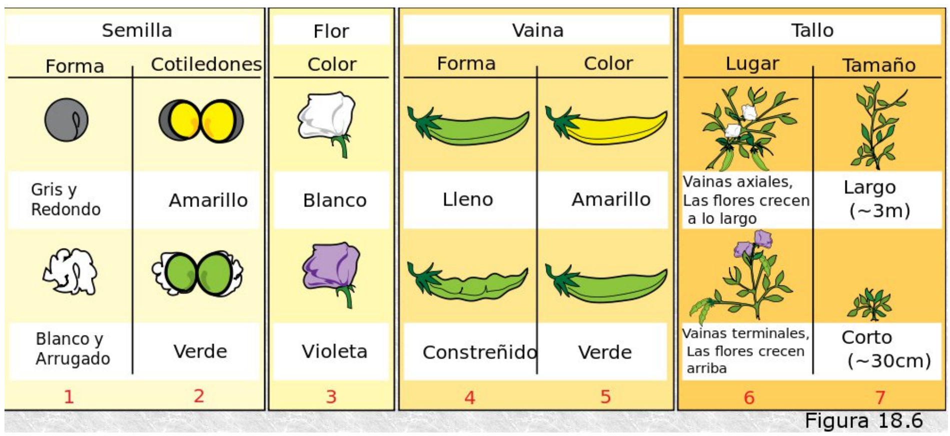


Figure 3.7: Las siete variantes de Pisum sativum estudiadas por Mendel en sus experimentos

3.7.1 Monohibridismo

una vez seleccionados los caracteres sobre los que se fijaría su atención. Pero en vez de analizar las siete al mismo tiempo, optó por estudiar la herencia utilizando una cada vez, es decir, realizo experimentos entre plantas que diferían en una característica y luego analizó a la progenie que era híbrida para dicho carácter. Es por esta razón que este tipo de cruzamientos se denomina monohibridismo, vale decir, producción de híbridos entre variedades que difieren en un solo carácter.

Los resultados de los experimentos de monohibridismo revelaron que en la descendencia se observaba solo una de las características estudiadas. Actualmente se les denomina fenotipos dominantes a los que se expresan y recesivos a los que no lo hacen. Por ejemplo, el color amarillo es dominante sobre el verde, por lo tanto, todas las semillas obtenidas serán amarillas.

Mendel no se limitó al análisis de la descendencia obtenida del cruce entre las mismas variedades, sino que planteó un nuevo experimento: cruzar individuos híbridos (obtenidos del cruce de las especies "puras") y analizar su descendencia. Para organizar estos cruzamientos, designó con símbolos a los individuos: la generación con los que inició los cruzamientos experimentales serían P (parentales), la primera generación de descendientes F1 (fila 1) y la segunda generación, F2.

El análisis de la F2 reveló que aquellos fenotipos que no se observaban en la F1 volvían a presentarse en la F2, obtenida del cruzamiento entre dos individuos de la F1 (híbridos). Al cuantificar, de dio cuenta de que los fenotipos dominantes eran siempre tres veces más frecuentes que los recesivos a la F1.

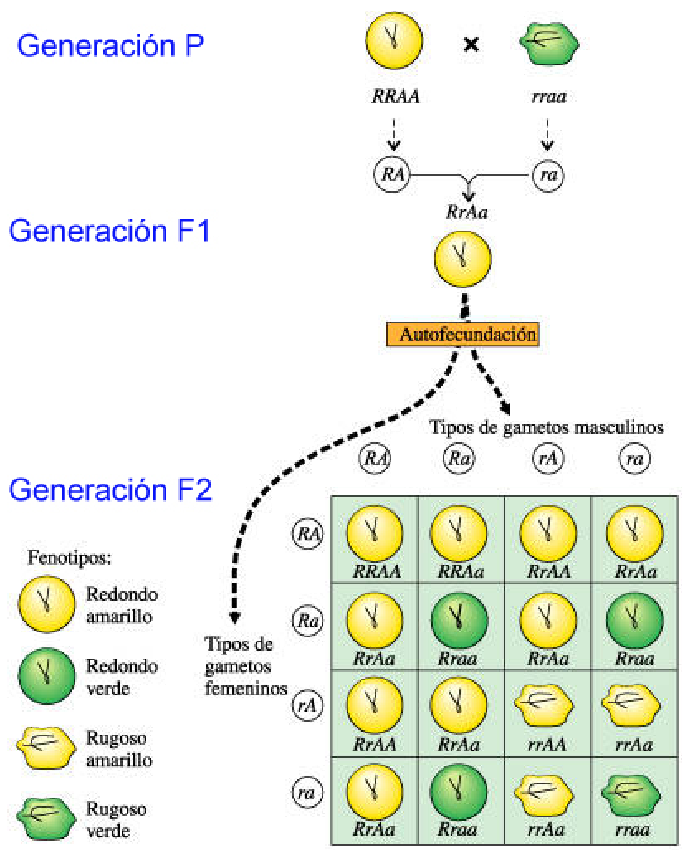


Figure 3.8: Experimentos de Mendel

En resumen los parentales eran individuos puros que presentaban semillas de color amarillo y verde. La F1 solamente produjo arvejas de semillas amarillas. Al realizar el cruce de F1 (híbridas), que presentaban semillas amarillas, se obtuvo en la F2 mayor cantidad de individuos de semillas amarillas, pero parte de ellas eran verde. El color verde había desaparecido de F1, por lo tanto, Mendel infirió que los individuos F2 llevaban "escondida" esta característica.

A partir de los resultados de los cruzamientos monohíbridos, Mendel infirió que los factores de la herencia se encuentran de a pares en los individuos, y que se separan al azar durante la formación de los gametos. Esta conclusión fue derivada de las siguientes observaciones:

- Todos los individuos de la generaciíon F1 presentaban una característica igual a la de uno de los padres, a la cual llamó dominante. El resultado anterior podrá explicarse si en la F1 cada individuo presenta un factor hereditario de cada padre. Como uno de ellos es dominante y el otro recesivo, siempre se expresará el dominante.

- En la F2, alrededor del 75 de los individuos surgán con fenotipo dominante y el 25, recesivo; equivalente a la proporción 3:1. Como estos individuos provenían de padres que presentaban un factor dominante y uno recesivo,entonces originaban gametos de dos tipos, unos con el factor dominante y otros con el recesivo, en igual proporción. Al cruzarse dos individuos de este tipo, se originaban las siguientes combinaciones posibles.

A partir de estos experimentos desarrollo varias leyes:

1. Primera ley o principio de la uniformidad: Cuando se cruzan dos individuos de raza pura, los híbridos resultantes son todos iguales. El cruce de dos individuos homocigotos, uno dominante (AA) y el otro recesivo (aa), origina sólo individuos heterocigotos (Aa).

2. Segunda ley o principio de la segregación: Ciertos individuos son capaces de transmitir un carácter aunque en ellos no se manifieste. El cruce de dos individuos de la F1 (Aa) daría origen a una segunda generación

   lial en la cual reaparece el fenotipo a, a pesar de que todos los individuos de la F1 eran de fenotipo A. Esto hace presumir a Mendel que el carácter a no había desaparecido, sino que sólo había sido opaca do por el carácter A pero que, al reproducirse un individuo, cada carácter se segrega por separado.

3. Tercera ley o principio de la combinación independiente: Mendel trabajó este cruce en guisantes, en los cuales las características que el observaba (color de la semilla y rugosidad de su superficie) se encontraban en cromosomas separados. De esta manera, observó que los caracteres se transmitán independientemente unos de otros. Esta ley, sin embargo, deja de cumplirse cuando existe vinculación (dos genes están en locus muy cercanos y no se separan en la meiosis).

3.8 ¿Cómo puedo calcular en que se gasta la memoria en los dispositivos de almacenamiento masivo?

Los usuarios pueden definir objetos con nombre llamados archivos y que están constituidos por una secuencia de bits, bytes, líneas o registros. Se referencian mediante su nombre. El sistema operativo suministra una serie de operaciones especiales (llamadas al sistema) para la manipulación de los archivos:[2]

- Como unidad: Apertura, cierre, creación, borrado, copia, ...
- Accediendo al contenido: Lectura y escritura, actualización, inserción, ...

3.8.1 Gestión del espacio en disco

Dos estrategias posibles para el almacenamiento de un archivo de n bytes en un dispositivo de memoria secundaria:

- Asignación contigua: Al archivo se le asignan n bytes consecutivos de espacio de disco. Si el archivo crece (hay que tener en cuenta que los archivos son estructuras de datos con una alta volatilidad) probablemente tendrá que ser movido en el disco (análogo a mover segmentos en memoria central, con la diferencia de que esta última operación es más rápida).
- Asignación no contigua: El archivo se divide en m bloques de tamaño fijo que se almacenan en disco en bloques no necesariamente contiguos.

La segunda estrategia es la más usual, pero nos encontramos con un primer problema en la gestión del espacio de disco: El establecimiento del tamaño del bloque, es decir, de la unidad mínima de asignación del espacio de disco. Para elegir el tamaño del bloque, se van a tener en cuenta los siguientes datos:

- Organización fśica del disco: Dada la forma en que se organizan los discos, el sector, la pista y el cilindro son candidatos obvios para la unidad de asignación.
- Aprovechamiento del espacio de disco: Si se elige una unidad de asignación grande (cilindro) y el tamaño medio de los ficheros es pequeño, se desaprovecha gran cantidad de espacio de disco por fragmentación interna.
- Ritmo de transferencia de datos: Si se elige una unidad de asignación pequeña, cada archivo estará dividido en muchos bloques. Como la lectura de cada bloque requiere un tiempo de búsqueda, un tiempo de retraso rotacional y un tiempo de transferencia, la lectura de un archivo (dividido en muchos bloques) será lenta.

Los dos últimos criterios están en contraposición, por lo que la elección del tamaño del bloque deberá ser un compromiso entre los mismos.

En la siguiente gráfica, se muestra, por un lado, el ritmo de transferencia de datos frente a T (tamaño del bloque), observándose que este es mayor a medida que aumenta T. También se ha representado en la misma gráfica el porcentaje de utilización del espacio de disco, suponiendo un tamaño de medio de archivo de 1 KB. Se observa que, a medida que aumenta el tamaño del bloque, el porcentaje de utilización del espacio de disco disminuye (mayor fragmentación interna).[de Memoria]

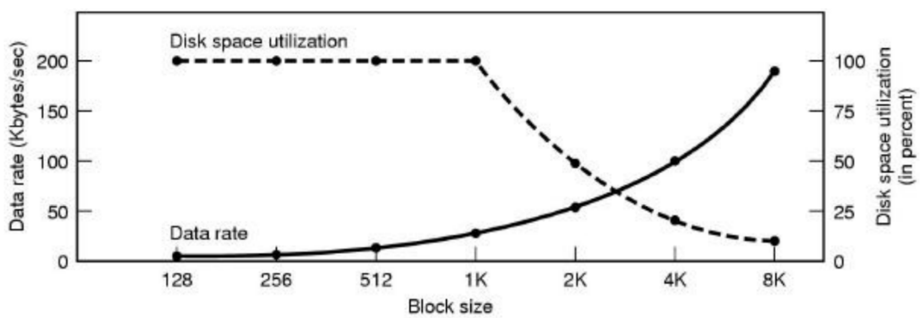


Figure 3.9: Ritmo de transferencia de datos

Surge pues un conflicto inherente entre la eficiencia del tiempo y del espacio (una buena utilización del espacio de disco implica ritmos de transferencia de datos bajos y viceversa). Compromiso normal: Se elige el tamaño de bloque de 512 B, 1 KB o 2 KB. Si se elige un tamaño de bloque de 1 KB sobre un disco con un tamaño de sector de 512 Bytes, el sistema de archivos leerá o escribirá dos sectores consecutivos y los tratará como una unidad simple e indivisible (hay que tener en cuenta que el sector es la unidad más pequeña que se puede leer o escribir en el disco). [de Memoria]

3.9 Paradoja

Una paradoja (del latín paradoxa, plural de paradoxon, "lo contrario a la opinión común") o antilogía es una idea extraña opuesta a lo que se considera verdadero a la opinión general.

Una declaración esencialmente contradictoria basada en un razonamiento válido de suposiciones lógicas.

3.10 Antinomia

Conflicto o contradicción entre dos leyes, principios racionales, ideas o actitudes.

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.

3.11 Diferencia entre paradoja y antinomia

Ejemplos de paradoja:

- Paradoja del hotel infinito: un hotel de infinitas habitaciones puede aceptar más huéspedes, incluso si está lleno.
- 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."

Ejemplos de Antinomia:

- Ignore este mensaje
- No tolero la intolerancia
- Prohibido prohibir
- Nunca digo la verdad
- Se que no se nada

3.12 Paradoja del Barbero

Propuesta por Bertrand Russell, dice: El único barbero de la ciudad dice que afeitará a todos aquellos que no se afeiten a sí mismos.

Pregunta: ¿Quien afeitará al barbero? Si no se afeita a sí mismo sera una de las personas de la ciudad que no se afeitan a sí mismas, con lo cual debería de afeitarse, siendo por tanto una de las personas que se afeitan a sí mismas, no debiendo por tanto afeitarse.

Figure 3.10: Barbería

¿Existe un conjunto de todos los conjuntos que no se contienen a sí mismos?.

Llamemos M al "conjunto de todos los conjuntos que no se contienen a sí mismos como miembros". Es decir:

$$M = \{x : x \notin x\}$$

Según la teoría de conjuntos de Cantor, la expresión anterior se puede representar por:

$$\forall x$$
$$x \in M \Leftrightarrow x \notin x$$

Es decir "Cada conjunto es elemento de M si y sólo si no es elemento de sí mismo". Ahora, en vista de que M es un conjunto, se puede substituir x por M en la segunda expresión, de donde se obtiene:

$$M \in M$$
$$M \notin M$$

Lo cual es absurdo.

3.13 Falacia del apostador

La falacia del jugador o falacia de Montecarlo es una falacia lógica por la que se cree erróneamente que los sucesos pasados afectan a los futuros en lo relativo a actividades aleatorias, como en muchos juegos de azar. Puede comprender las siguientes ideas equivocadas:

Un suceso aleatorio tiene más probabilidad de ocurrir porque no ha ocurrido durante cierto período. Un suceso aleatorio tiene menos probabilidad de ocurrir porque ha ocurrido durante cierto período Un suceso aleatorio tiene más probabilidad de ocurrir si no ocurrió recientemente Un suceso aleatorio tiene menos probabilidad de ocurrir si ocurrió recientemente Las anteriores son ideas equivocadas que surgen cotidianamente en razonamientos sobre probabilidades, muchos de los cuales se han estudiado con gran profundidad. Mucha gente pierde dinero apostando debido a su creencia errónea en esta falacia.[Casanova Leal]

Sencillamente, las probabilidades de que algo suceda la próxima vez no están necesariamente relacionadas con lo que ya sucedió, especialmente en muchos juegos de azar. Esto suele resumirse en la frase "Los dados (o moneda) no tiene memoria", pues su naturaleza es la misma, independiente del número de tiros y resultados previos.[Casanova Leal]

Figure 3.11: Ruleta Casino

3.14 La paradoja del cumpleaños

El problema del cumpleaños, también llamado paradoja del cumpleaños, establece que de un conjunto de 23 personas, hay una probabilidad del $50,7\%$ de que al menos dos personas de ellas cumplan años el mismo día. Para 57 o más personas la probabilidad es mayor del 99%.

Figure 3.12: Problema del cumpleaños

En sentido estricto esto no es una paradoja ya que no es una contradicción lógica; no es una paradoja pero es una verdad matemática que contradice la intuición común. Mucha gente piensa que la probabilidad es mucho más baja, y que hacen falta muchas más personas para que se alcance la probabilidad del 50%. Si una habitación tuviera 367 personas, por el Principio del pichonera (Principio de Dirchlet) sabemos que habría al menos dos personas cumpliendo años en la misma fecha, ya que un año normal tiene 365 días, y uno bisiesto tiene 366.

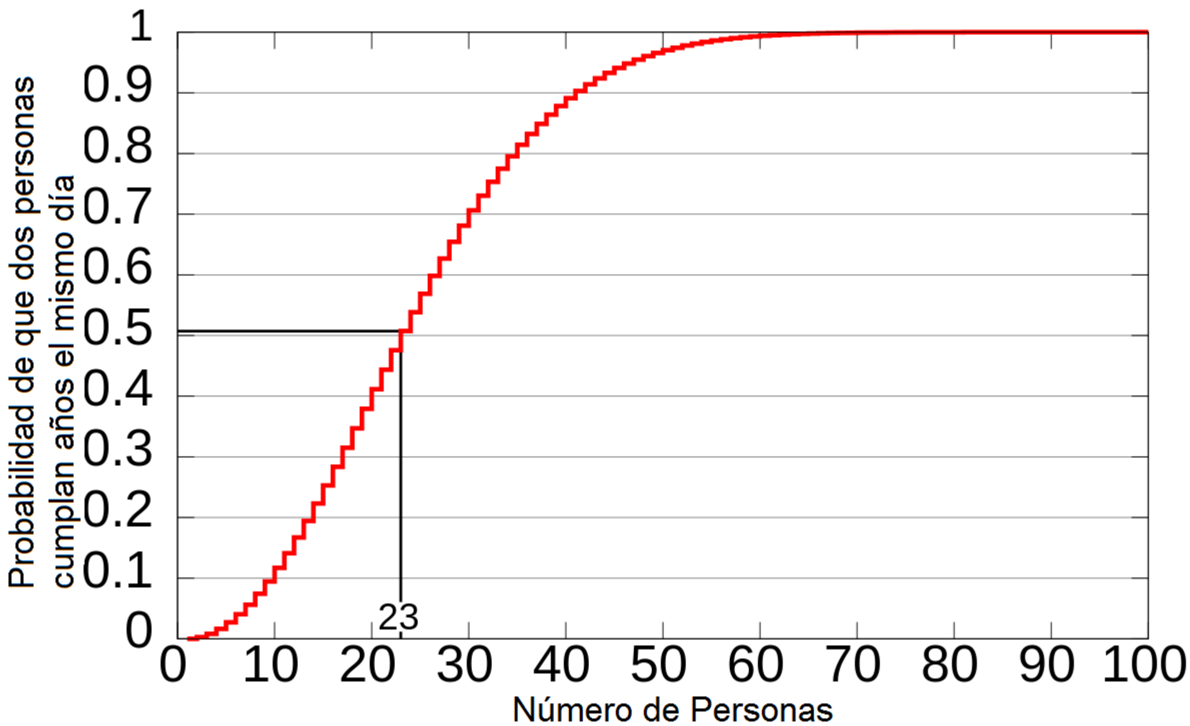


Figure 3.13: Gráfica que muestra el comportamiento del fenómeno probabilista en base al número de personas

Calcular esta probabilidad es el problema del cumpleaños. La teoría fue descrita en American Mathematical Monthly en 1938 en la teoría de Estimación del total de población de peces en un lago de Zoe Emily Schnabel, bajo el nombre de captura-recaptura estadística.

La clave para entender la paradoja del cumpleaños es pensar que hay muchas probabilidades de encontrar parejas que cumplan años el mismo dá. Específicamente, entre 23 personas, hay $\frac{(23)(22)}{2} = 253$ pares, cada uno de ellos un candidato potencial para cumplir la paradoja.

Hay que entender que si una persona entrase en una habitación con 22 personas, la probabilidad de que cualquiera cumpla años el mismo día que quien entra, no es del 50%, es mucho más baja. Esto es debido a que ahora sólo hay 22 pares posibles. El problema real de la paradoja del cumpleaños consiste en preguntar si el cumpleaños de cualquiera de las 23 personas coincide con el cumpleaños de alguna de las otras personas.

Calculemos la probabilidad de que, en una habitación con n personas, al menos dos cumplan años el mismo día, desechando los años bisiestos y las personas gemelas, y asumiendo que existen 365 cumpleaños que tienen la misma probabilidad. El truco es calcular primero la probabilidad de que n cumpleaños sean diferentes. La probabilidad de que ninguna persona cumpla años el mismo día viene dada por:

$$p = (\frac{365}{365})(\frac{364}{365})(\frac{363}{365})...(\frac{365-n+1}{365}) \tag{3.1}$$

Porque la segunda persona no puede tener el mismo cumpleaños que el primero $\frac{364}{365}$, la tercera persona no puede tener el mismo cumpleaños que las dos primeras $\frac{363}{365}$, etc.

Usando notación factorial, puede ser escrita como:

$$p = \begin{cases} \frac{365!}{365^n(365-n)!} & \text{si} \quad 1 \leqslant n \leqslant 365 \\ \\ 0 & \text{si} \quad 365 < n \end{cases} \tag{3.2}$$

Ahora, $(1-p)$ es la probabilidad de que al menos dos personas tengan el mismo día de cumpleaños. Para $n = 23$ se obtiene una probabilidad de alrededor de 0,507.

$$1 - p = \begin{cases} 1 - \frac{365!}{365^n(365-n)!} & \text{si} \quad 1 \leqslant n \leqslant 365 \\ \\ 1 & \text{si} \quad 365 < n \end{cases} \tag{3.3}$$

En contraste, la probabilidad que cualquiera en una habitación de n personas (sin contarse a si mismo) tengan el mismo día de cumpleaños que usted está dada por:

$$1 - (\frac{364}{365})^n \tag{3.4}$$

La solución se puede generalizar para incluir a los nacidos un 29 de febrero, naturalmente de un año bisiesto. Es una solución, puede haber otras, la ventaja de ésta es que es exacta y sencilla. Se usa el algoritmo anterior (con 365, haya personas nacidas en años bisiestos o no) con los siguientes cambios:

Sean nb las personas presentes que cumplen años el 29 de febrero.

- Si $nb = 0$; Aplicar Algoritmo. Termina
- Si $nb = 1$; $n = n - 1$; Aplicar Algoritmo. Termina
- Si $nb > 1$; hay al menos 2 personas con la misma fecha de cumpleaños. Termina

3.14.1 Ejemplo:

En una fiesta se encuentran 23 personas bebiendo y bailando. ¿Cuál es la probabilidad de que haya al menos 2 personas con la misma fecha de cumpleaños? ¿Cuantás personas se necesitan para que haya un 70% de probabilidad de que al menos 2 personas cumplan años el mismo día?
De la solución al problema del cumpleaños se sabe que la probabilidad de que n personas cumplan años el mismo día es:

$$P(n\_Personas) = 1 - \left(\frac{365}{365}\frac{364}{365}\cdots\frac{365-n+1}{365}\right)$$
$$= 1 - \frac{365!}{365^n(365-n)!}$$

De tal manera que la probabilidad para 23 personas es:

$$P(23) = 1 - \frac{365!}{365^{23}(342)!} = 0.492702766$$
$$\approx 1 - e^{-\frac{23^2}{2*365}} = 0.515509538$$

3.15 Problema de Monty Hall

El juego está inspirado en el concurso televisivo Let's Make a Deal (Hagamos un trato), emitido entre 1963 y 1986 en la televisión americana y su nombre proviene del presentador del concurso, Monty Hall.

El concurso geneó bastante polémica en relación a posibles soluciones del problema matemático latente y muestra las intuiciones incorrectas en relación a la probabilidad condicional. La formulación más conocida de dicho problema (Bohl, Liberatore y Nydick, 1995) se reproduce a continuación:

Supongamos que estás en un concurso, y se te ofrece escoger entre tres puertas: detás de una de ellas hay un coche, y detrás de las otras cabras. Escoges una puerta, digamos la No 1, y el presentador, que sabe lo que hay detrás de las puertas, abre otra, digamos la No 3, que contiene una cabra. Entonces te pregunta: ¿No prefieres escoger la No 2? ¿Es mejor para ti cambiar tu elección?.

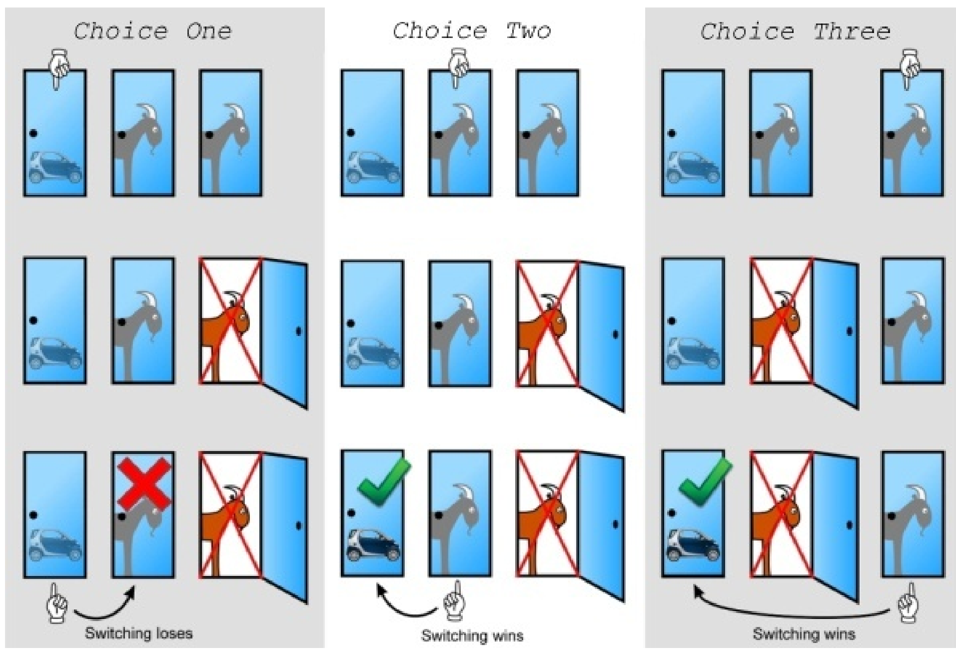


Figure 3.14: Elecciones de la puerta ganadora

Consideramos, en primer lugar, el experimento "puerta que tiene el premio" (cada puerta tiene probabilidad 1/3). A continuación, consideramos la puerta que se elige (1/3 cada puerta). Estos dos primeros experimentos son independientes. El tercer experimento es la puerta que abre el locutor, que es dependiente de los anteriores. [4]

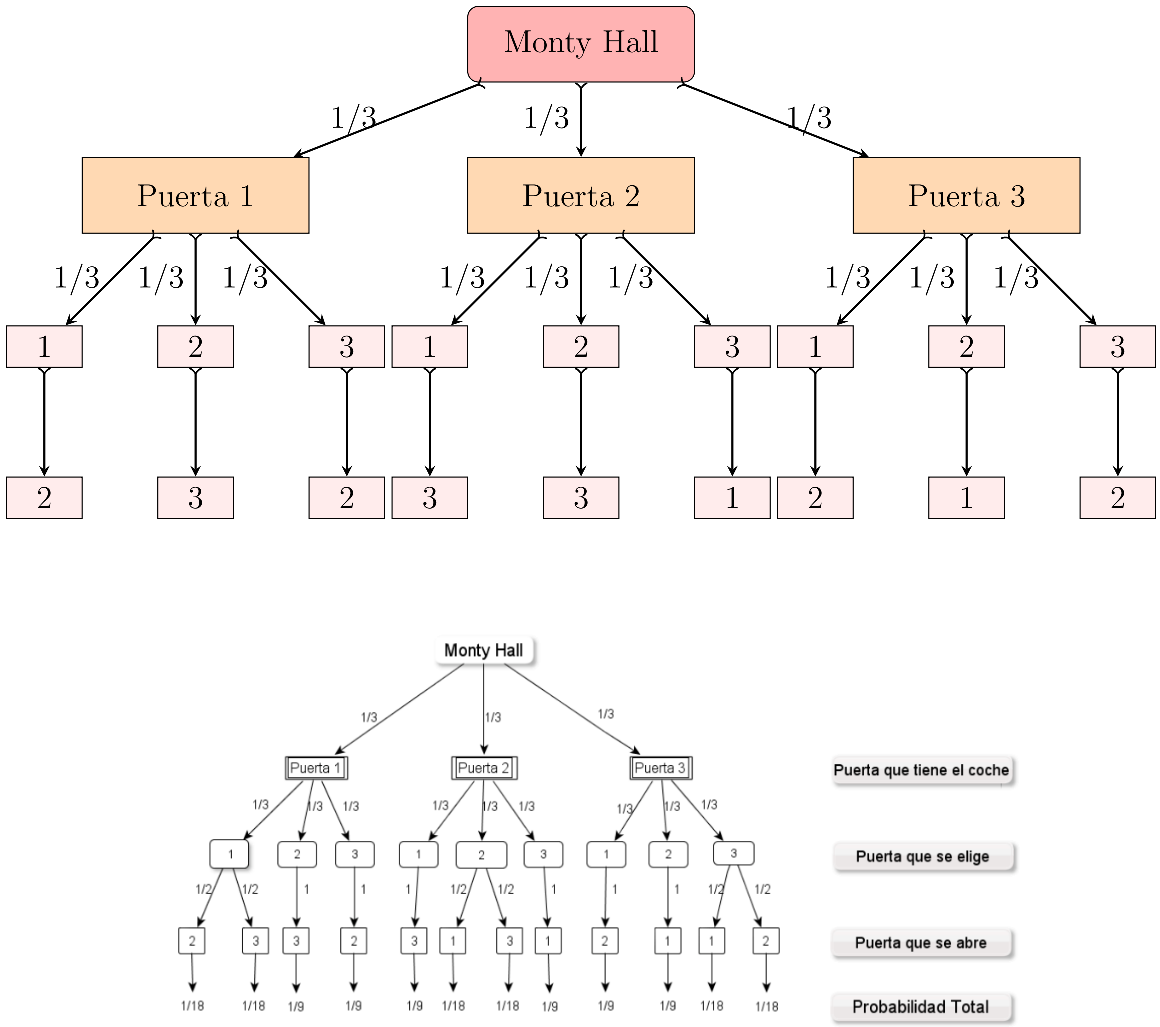


Figure 3.15: Ilustración del juego

Inicialmente al elegir una puerta, las posibilidades de ganar son de 1/3. Cada uno de estos sucesos (si decidimos cambiar de puerta cuando el presentador nos da la opción) tiene probabilidad 1/9.

Suponemos que cambiamos. Si estando el coche en la puerta 1, elegimos la puerta 1, el presentador te abre bien la puerta 2 o la 3, cada una de ellas con probabilidad 1/18, en total 1/9. Si cambiamos de puerta perdemos con probabilidad 1/9.

Como hay tres puertas, la probabilidad total de ganar cambiando sería 2/3 y la de perder cambiando sería 1/3.

3.16 Ejemplos de Tensores

3.16.1 Tensor de Curvatura

En geometría diferencial, el tensor de curvatura de Riemann, o simplemente tensor de curvatura o tensor de Riemann, supone una generalización del concepto de curvatura de Gauss, definido para superficies, a variedades de dimensiones arbitrarias. Representa una medida de la separación de la métrica de la variedad respecto de la métrica euclídea.

Si M es una variedad riemanniana, el tensor de curvatura vendrá definido a partir de la conexión de Levi-Civita. El tensor métrico g podrá utilizarse para subir o bajar índices del tensor de curvatura. En particular, la versión completamente covariante del tensor es:

$$R_{\rho\sigma\mu\nu} = g_{\rho\alpha} R^{\alpha}_{\sigma\mu\nu} \tag{3.5}$$

3.17 Movimiento de Rotación

Rotación es el movimiento de cambio de orientación de un cuerpo o un sistema de referencia de forma que una línea (llamada eje de rotación) o un punto permanece fijo.

La rotación de un cuerpo se representa mediante un operador que afecta a un conjunto de puntos o vectores. El movimiento rotatorio se representa mediante el vector velocidad angular, que es un vector de carácter deslizante y situado sobre el eje de rotación. Cuando el eje pasa por el centro de masa o de gravedad se dice que el cuerpo gira sobre sí mismo.[14]

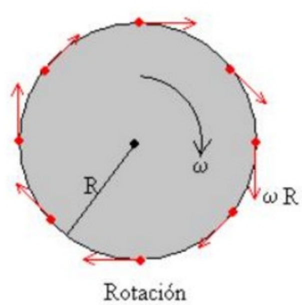


Figure 3.16: Movimiento de rotación

3.18 Movimiento de Traslación

El movimiento de traslación es aquel giro que realizan los cuerpos alrededor de un centro provocados por la gravedad, al hacer la traslación los objetos lo hacen bastante separados del centro a través del cual están girando, lo que denota que el radio de separación que existe entre dicho cuerpo y el centro es bastante amplio comparado con sus propias dimensiones.[14]

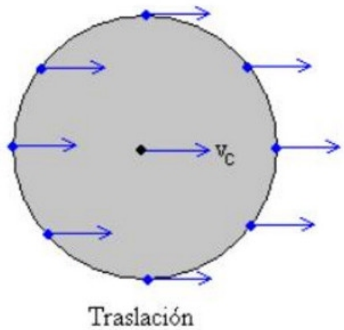


Figure 3.17: Movimiento de traslación

3.19 Movimiento de Nutación

Nutación (del latín "nutare", cabecear u oscilar) es un movimiento ligero irregular en el eje de rotación de objetos simétricos que giran sobre su eje. Ejemplos comunes son los giroscopios, los trompos y los planetas. Una nutación pura es el movimiento del eje de rotación que mantiene el primer ángulo de Euler (precesión) constante.

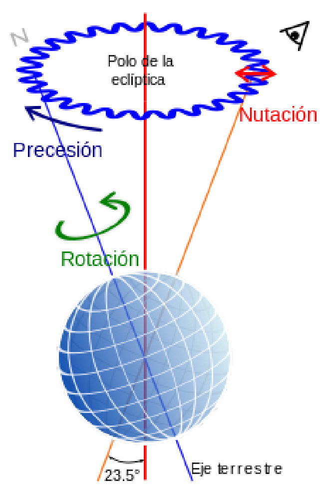


Figure 3.18: Movimiento de nutación

3.20 Movimiento de Precesión

La precesión o movimiento de precesión nutación es el movimiento asociado con el cambio de dirección en el espacio, que experimenta el eje instantáneo de rotación de un cuerpo.

Un ejemplo de precesión lo tenemos en el movimiento que realiza una peonza o trompo en rotación. Cuando su eje de rotación no es vertical, la peonza posee un movimiento de cabeceo similar al de precesión.

Más exactamente una precesión pura es aquel movimiento del eje de rotación que mantiene su segundo ángulo de Euler (nutación) constante. Este movimiento de nutación también se da en el eje de la Tierra.

Hay dos tipos de precesión: la precesión debida a los momentos externos, y la precesión sin momentos de fuerzas externos.

3.21 Reseña del libro "The Math Book, Clifford A. Pickover"

3.21.1 Stirling Formula

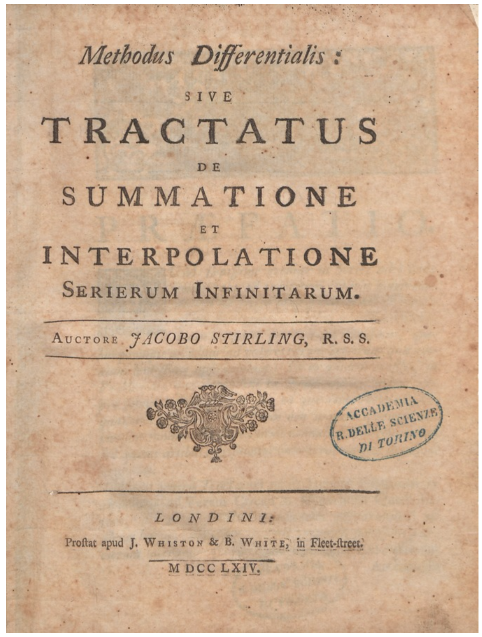
En una época donde los factoriales se encontraban por todos lados. La notación factorial la introdujo el matematico francés Christian Kramp en 1808, lo cual ha sido muy útil para problemas de combinatoria, cuando se calcula el factorial de un número muy grande es conveniente utilizar métodos de aproximación como la formúla de Striling:

$$n! \approx \sqrt{2\pi}e^{-n}n^{n-\frac{1}{2}}; e = 2.71828; \pi = 3.14159$$

Donde para valores grandes de n:

$$ln(n!) \approx nln(n) - n; n! \approx n^n e^{-n}$$

Striling presento la aproximación de $n!$ en su más famoso trabajo Methodus Differentialis iniciando su carrera en medio de conflictos políticos y religiosos.



Methodus Differentialis:
SIVE
TRACTATUS
DE
SUMMATIONE
ET
INTERPOLATIONE
SERIERUM INFINITARUM.

AUCTORE *JACOBO STIRLING*, R. S. S.

LONDINI:
Proſtat apud J. WHISTON & B. WHITE, in Fleet-ſtreet.
M DCC LXIV.

Figure 3.19: Methodus Differentialis

La formula marco la transformación de las matemáticas del siglo XVII al siglo XVIII, desde los principios del cálculo de Newton y 90 años despúes la formula de Stirling, por lo que las matemáticas dejaron de ser na actividad amateur y se convirtió en un trabajo profesional.

3.21.2 Bayes Theorem

El teorema de Bayes fue introducido por el matemático británico Thomas Bayes para el calculo de la probablidad condicional, la cual se refiere a la probabilidad que tiene un evento A de ocurrir dada la ocurrencia de un evento B $P(A|B) = [P(B|A) \times P(A)]/P(B)$.

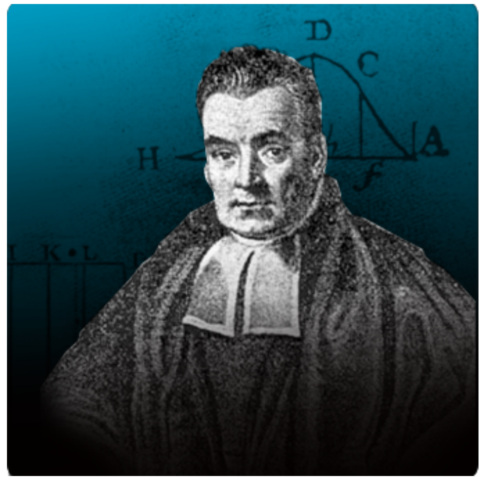


Figure 3.20: Retrato del matemático Thomas Bayes

Supongamos que en una primer caja tienes 10 pelotas de golf y 30 pelotas de billar, en una segunda caje tienes 20 de cada una, seleccionas una caja aleatoriamente y tomas una bola y te sale una pelota de billar, entonces ¿Qué probabilidad hayq de que hayas escogido la primer caja dado que tienes una pelota de billar en la mano? El evento A corresponde a escoger la primer caja y el evento B a sacar una pelota de billar, la probabilidad de seleccionar una caja es de 0.5 y la probabilidad de sacar una pelota de billa r de la primer caja es de 0.75, la probabilidad de sacar una de billar de la segunda caja es de 0.5 por lo que $0.75 \times 0.5 + 0.5 \times 0.5 = 0.625$ Se puede calcular esta probabilidad gracias al teorema de bayes como $P(A|B) = 0.6$.

3.21.3 Pigeonhole Principle

El principio fue introducido en 1834 por el matemático alemán Johann Dirichlet haciendo referencia a ”El principio de las cajoneras” sin embargo se uso formalmente en una seria revista de matemáticas en 1940.

El principio consiste en que si se tienen m pichoneras y n pichones, se puede asegurar que al menos una pichonera alberga a mas de un pichón si $n > m$.

El principio ha sido generalizado a aplicaciones probabilistas que dicen que si se sitúan n pichones aleatoriamente en m pichoneras con una probabilidad uniforme de $\frac{1}{m}$, entonces por lo menos un casillero tendrá más de paloma con probabilidad $1 - \frac{m!}{[(m-n)!m^n]}$.

Figure 3.21: Ilustración: pichones por celda

3.21.4 Principia Mathematica

Dos filósofos y matemáticos británicos, Bertrand Russell y Alfed North Whitehead, colaboraron por 8 años con su trabajo *Principa Mathematica* para demostrar que las matemáticas pueden ser declaradas usando conceptos de lógica como clases y relaciones entre miembros de una clase.

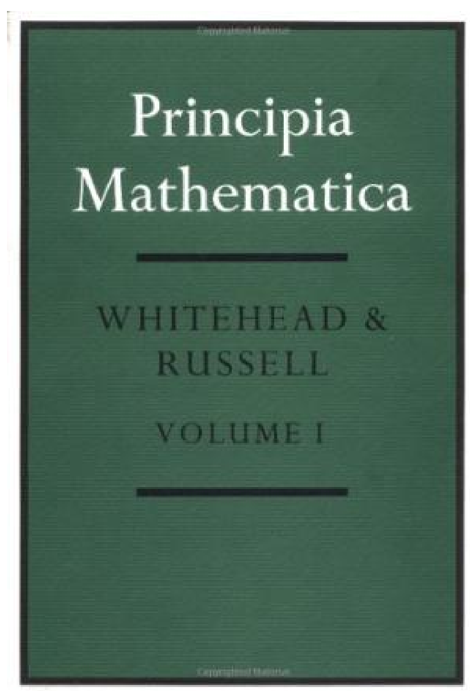


Figure 3.22: Portada: Principia Mathematica

El libro es incluido en los libros de no ficción mas importantes del siglo XX por la biblioteca moderna.

A pesar de que *Principia* logró proveer de varias derivaciones de teoremas importantes de matemáticas fue criticado por algunos supuestos del libro, como el axioma del infinito que parecía ser mas una suposición empírica que lógica.

En *Principia*,los autores demuestran que 1 + 1 = 2. La editoral del libro, había decidido que la publicación *Principa* daría como resultado una pérdida estimada de 600 libras. Sólo después de que los autores acordaron dar algo de dinero a la editorial fue el libro publicado.

3.21.5 Infinite Monkey Theorem

El teorema del mono infinito afirma que un mono pulsando teclas al azar sobre un teclado durante un periodo de tiempo de 10 horas casi seguramente podrá escribir finalmente cualquier texto dado.

Figure 3.23: Infinite Monkey Theorem

En este contexto, el término casi seguramente es un término matemático con un sentido preciso y el ”mono” no es en realidad un mono, sino que se trata de una metáfora de la creación de una secuencia aleatoria de letras al infinito.

La idea original fue planteada por Émile Borel, dijo que si un ejercito de monos mecanografiaran diez horas al día era extremadamente improbable que pudiesen producir algo que fuese igual a lo contenido en los libros de las bibliotecas más ricas del mundo y aun así, en comparación, sería aún más inverosímil que las leyes de la estadística fuesen violadas. Para Borel, el propósito de la metáfora de los monos era ilustrar la magnitud de un acontecimiento extraordinariamente improbable.

3.21.6 Differential Analyzer

Las ecuaciones diferenciales juegan un papel importante en la física, la ingeniería, la química, economía y otras disciplinas de números.

En 1927 el ingeniero americano $VannevarBush$ y sus colegas desarrollaron un analizador diferencial (DA), una computadora analógica integrando metódos de ecuaciones diferenciales con muchas variables independientes.

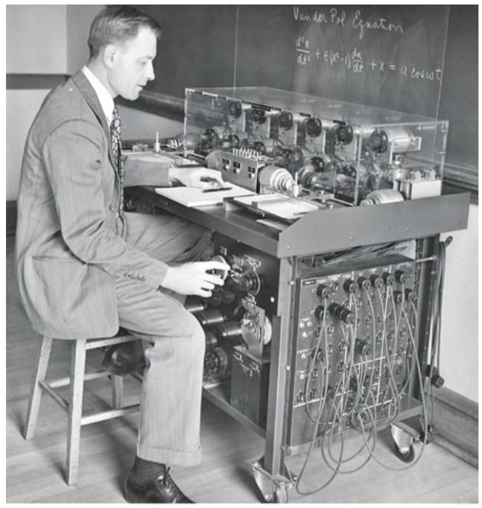


Figure 3.24: Vannevar Bush with Differential Analyzer

La versión anterior de este tipo de dispositivos tenía sus raíces en el analizador de armónicos de Lord Kelvin en los EE.UU., los investigadores que trabajan en la base de la fuerza aérea wright-Patterson y más escuela de ingeniería eléctrica en la universidad construir los dispositivos DA en parte para la creación de mesas de tiro de artillería, antes de la invención de ENIAC (integración numérica electrónica y computador).

3.22 CUDA

3.22.1 ¿Qué es CUDA?

CUDA es una arquitectura de cálculo paralelo de NVIDIA que aprovecha la gran potencia de la GPU (unidad de procesamiento gráfico) para proporcionar un incremento extraordinario del rendimiento del sistema.[12]

Gracias a millones de GPUs CUDA vendidas hasta la fecha, miles de desarrolladores, científicos e investigadores están encontrando innumerables aplicaciones prácticas para esta tecnología en campos como el procesamiento de vídeo e imágenes, la biología y la química computacional, la simulación de la dinámica de fluidos, la reconstrucción de imágenes de TC, el análisis sísmico o el trazado de rayos, entre otras.

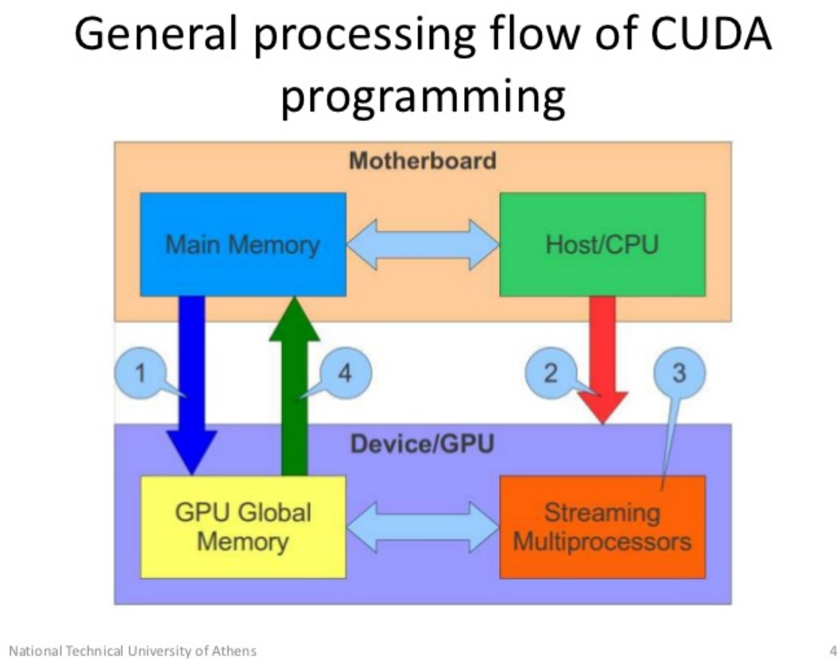


Figure 3.25: Procesamiento general de CUDA

3.22.2 Procesamiento paralelo con CUDA

Los sistemas informáticos están pasando de realizar el "procesamiento central" en la CPU a realizar "coprocesamiento" repartido entre la CPU y la GPU. Para posibilitar este nuevo paradigma computacional, NVIDIA ha inventado la arquitectura de cálculo paralelo CUDA, que ahora se incluye en las GPUs GeForce, ION Quadro y Tesla GPUs, lo cual representa una base instalada considerable para los desarrolladores de aplicaciones.

En el mercado de consumo, prácticamente todas las aplicaciones de vídeo se han acelerado, o pronto se acelerarán, a través de CUDA, como demuestran diferentes productos de Elemental Technologies, MotionDSP y LoiLo, Inc.

CUDA ha sido recibida con entusiasmo por la comunidad científica. Por ejemplo, se está utilizando para acelerar AMBER, un simulador de dinámica molecular empleado por más de 60.000 investigadores del ámbito académico y farmacéutico de todo el mundo para acelerar el descubrimiento de nuevos medicamentos.

En el mercado financiero, Numerix y CompatibL introdujeron soporte de CUDA para una nueva aplicación de cálculo de riesgo de contraparte y, como resultado, se ha multiplicado por 18 la velocidad de la aplicación. Cerca de 400 instituciones financieras utilizan Numerix en la actualidad.Un buen indicador de la excelente acogida de CUDA es la rápida adopción de la GPU Tesla para aplicaciones de GPU Computing. En la actualidad existen más de 700 clusters de GPUs instalados en compañías Fortune 500 de todo el mundo, lo que incluye empresas como Schlumberger y Chevron en el sector energético o BNP Pariba en el sector bancario.

Por otra parte, la reciente llegada de los últimos sistemas operativos de Microsoft y Apple (Windows 8 y Snow Leopard) está convirtiendo el GPU Computing en una tecnología de uso masivo. En estos nuevos sistemas, la GPU no actúa únicamente como procesador gráfico, sino como procesador paralelo de propósito general accesible para cualquier aplicación. [12]

3.23 La 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), Ian Wilmut y Keith Campbell.

Dolly fue en realidad una oveja resultado de una combinación nuclear desde una célula donante diferenciada a un óvulo no fecundado y anucleado (sin núcleo). La célula de la que venía Dolly era una ya diferenciada o especializada, procedente de un tejido concreto, la glándula mamaria, de un animal adulto (una oveja Finn Dorset de seis años), lo cual suponía una novedad. Hasta ese momento se creía que sólo se podían obtener clones de una célula embrionaria, es decir, no especializada. Cinco meses después nacía Dolly, que fue el único cordero resultante de 277 fusiones de óvulos anucleados con núcleos de células mamarias.[15]

Dolly vivió siempre en el Instituto Roslin. Allí fue cruzada con un macho Welsh Mountain para producir seis crías en total. De su primer parto nace Bonnie, en abril de 1998. Al año siguiente, Dolly produce mellizos: Sally y Rosie, y en el siguiente parto trillizos: Lucy, Darcy y Cotton. En el otoño de 2001, a los cinco años, Dolly desarrolla artritis comenzando a caminar dolorosamente, siendo tratada exitosamente con pastillas antiinflamatorias.

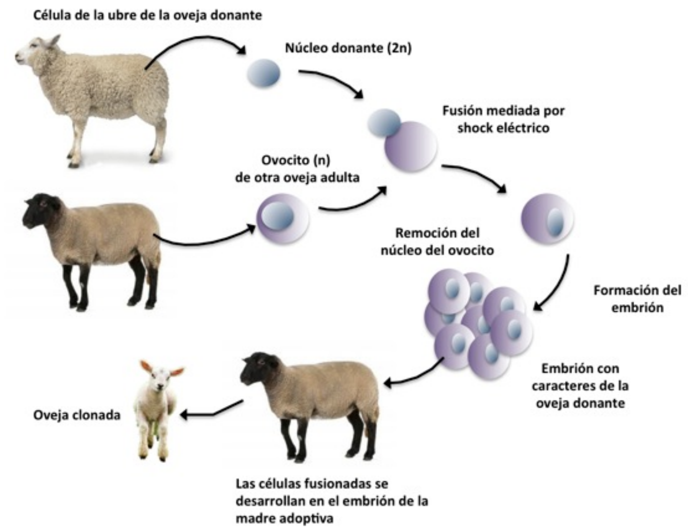


Figure 3.26: Proceso de clonación

Por razones que se temían, la oveja Dolly tendrá una vida corta y envejecerá prematuramente. La causa esta en que ella se clono a partir de una célula que ya se había dividido varias veces cuando formaba parte de los tejidos de la madre, y el numero de veces que una célula puede dividirse esta inscrito en los cromosomas. Ello a su vez depende del largo de los telomeros.

Figure 3.27: Ilustración: cromosomas

Recientemente los investigadores del Rosselin Institute, de Edimburgo, y de la compañía PPL Therapeutic, que clonaron a Dolly, comunicaron que aun cuando ésta se encuentra en buenas condiciones de salud, posiblemente tenga una vida más corta, y envejezca prematuramente debido a que los telómeros de sus células son más cortos. El hallazgo no es inesperado. Por el contrario era un riesgo perfectamente posible de esperar, y por ello los investigadores tenían presente la necesidad de medir sus telomeros (Nature, vol 339, pág. 316, año 1999).

3.24 Telómeros

Los estudios pioneros de Hermann Müller (premio Nobel en 1945) trabajando con la mosca del vinagre, Drosophila melanogaster, y de Barbara McClintock (también premio Nobel en 1983) estudiando el maíz, Zea mays, habían concluído que los extremos de los cromosomas poseen unas estructuras especiales denominadas telómeros (del griego telos: final y meros: parte). Estas estructuras representarían una protección fundamental para que los cromosomas no resulten fusionados por la maquinaria de reparación de la célula ya que permiten distinguir estos extremos de cualquier otra rotura en el ADN celular que debe ser reparada y vuelta a fusionar.

Los telómeros son los extremos de los cromosomas. Una analogía común es imaginar que son como las fundas de plástico del extremo de los cordones de los zapatos, que impiden que estos se deshilachen. Los telómeros están formados por repeticiones en tándem de una secuencia de ADN y proteínas asociadas.[5]

El problema celular se produce cuando el acortamiento de los telómeros es tal que durante el proceso replicativo estos no puedan preservar el ADN de los cromosomas, produciendo la inconsistencia del material genético.[5]

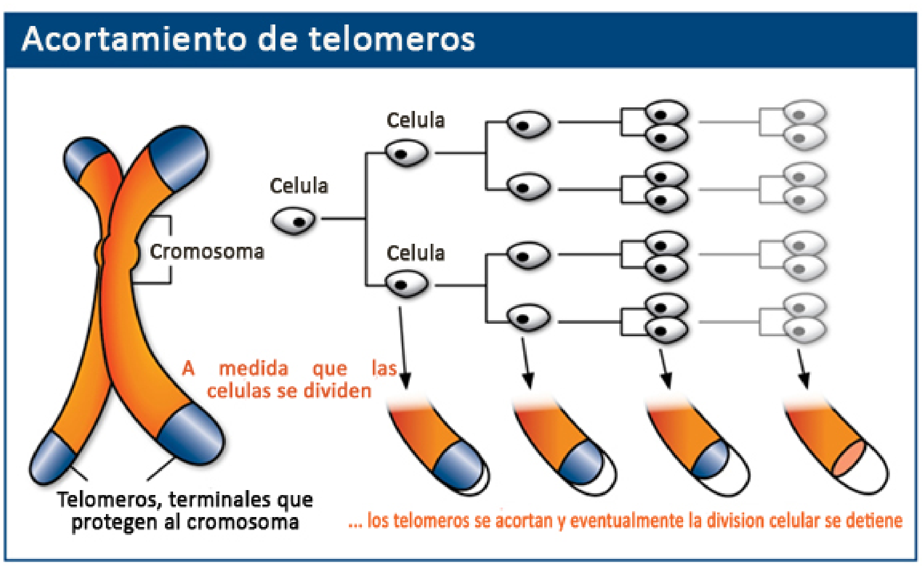


Figure 3.28: Ilustración de telómeros acortados

3.25 Desastre del Challenger

Muchos recuerdan haber visto en directo como explotaba la nave a los 73 segundos de vuelo, terminando inmediatamente con las vidas de los siete astronautas que formaban su tripulación. También recuerdan horrorizados como en la investigación subsiguiente se descubrió que la NASA había utilizado un diseño erróneo de los cohetes de combustible sólido para satisfacer las demandas de los políticos, que también habrían insistido en que el lanzamiento siguiera adelante aunque las condiciones no eran seguras, y como al final se consolaron diciendo que era el precio inevitable del progreso.[18]

Figure 3.29: Ilustración de la explosión: Challenger

1. Casi nadie vio el accidente en directo. La CNN en efecto retransmitió todo en directo, aunque poca gente veía todavía televisión por cable entonces, pero a esas alturas los despegues de los transbordadores eran casi, con lo que apenas despertaban interés en los medios y la mayoría de las emisoras ya no estaban emitiendo en directo cuando se produjo el accidente. Eso sí, todas volvieron a pinchar la señal de la NASA rápidamente después de éste, pero lo que la gente recuerda como imágenes en directo en realidad es una grabación. Quienes sí lo vieron en directo fueron muchos niños, ya que en ese vuelo iba Christa McAuliffe, la primera profesora en el espacio, y la NASA había preparado las cosas para que los colegios pudieran seguir el despegue vía satélite.

2. El Challenger explotó. No en el sentido tradicional de la palabra. Lo que sucedió es que el hidrógeno y oxígeno líquido que escaparon del tanque principal se incendiaron formando una enorme bola de fuego a algo más de 15 kilómetros de altura, pero no hubo ninguna onda de choque ni detonación; de hecho los cohetes de combustible sólido siguieron volando sin haber sido dañados por ninguna explosión. Lo que destrozó el transbordador fueron las fuerzas aerodinámicas cuando al quedar suelto se giró de lado contra el sentido de la marcha.

3. La tripulación murió al instante. Casi con toda certeza no, y esta es sin duda la parte más angustiosa del accidente, pues aunque el transbordador fue reducido a piezas la cabina se conservó prácticamente intacta y siguió subiendo por inercia hasta alcanzar casi 22 kilómetros de altura para luego iniciar el descenso hasta el agua, a dónde llegó 2 minutos y 45 segundos después del accidente, y todo parece indicar que la tripulación seguía viva en ese momento, aunque no está claro si estaban conscientes o no. Lo que sí está claro es que el impacto contra el agua provocó una des aceleración de unas 200 veces la fuerza de la gravedad, lo que destruyó la estructura y todo lo que había en su interior.

4. Los cohetes de combustible sólido eran inseguros a causa de injerencias políticas. Estos cohetes tenían fallos que podían ser mejorados y se estaba trabajando continuamente en ello, pero no eran espacial mente peligrosos si se respetaban los parámetros de seguridad, ni el diseño era consecuencia de interferencias de los políticos en el proceso. Ciertamente lo que acabó con el transbordador fue una llama que se escapó de una de las juntas situadas entre cada una de las cuatro secciones que forman estos cohetes, pero hacerlos de una sola pieza también representa riesgos, y las juntas habían funcionado adecuadamente siempre que fueron utilizadas dentro de sus parámetros.

5. El sellado entre secciones de los cohetes era más débil de lo debido a causa de ciertas leyes de protección medioambiental. Se dice que la NASA se vio obligada por ley a dejar de usar un sellador muy efectivo pero con un alto contenido de amianto para usar uno con menos amianto y que resultaba ser más débil, pero este cambio de sellador no tuvo nada que ver con el accidente y de hecho ya se había producido antes de cualquier prohibición al respecto. Cualquiera de las versiones del sellador tenía sus problemas, pero lo que acabó con el Challenger fue una anillo de goma que formaba parte de una de las juntas entre secciones del cohete y que debido a las bajas temperaturas reinantes la noche anterior al despegue había perdido su flexibilidad y se rompió bajo la presión de los gases durante el lanzamiento.Es como si coges un chicle, normalmente flexible, y lo metes en el congelador: se volverá rígido y muy fácil de romper, y esto es exactamente lo que le pasó a aquella junta.

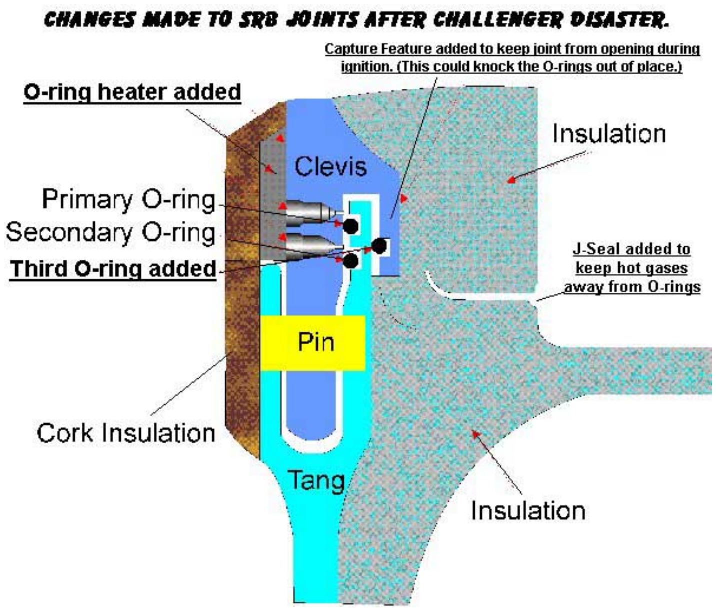


Figure 3.30: Ilustración: Junta que ocacionó el desastre del Challenger

6. Presiones políticas forzaron el lanzamiento. Había presiones dentro de la propia NASA para hacer ya el lanzamiento, pues ya se había retrasado en varias ocasiones, y el principal motivo parecen haber sido dos sondas espaciales que el Challenger tenía que poner en órbita y cuya ventana de lanzamiento terminaba en cuatro meses. El rumor de que se insistió en el lanzamiento desde la propia Casa Blanca para que el presidente Reagan pudiera comunicarse con la tripulación durante el debate del estado de la unión no tiene ningún fundamento, pues les tocaba dormir durante este debate y además no había previsto ningún tipo de comunicación con Tierra en esos momentos el programa de la misión.

7. El accidente fue el precio inevitable a pagar a cambio del progreso. Esto no son más que racionalizaciones para salvar el culo por parte de aquellos responsables de una gestión incompetente de todo el proceso. El desastre no tenía que haber ocurrido, y los responsables de la NASA se equivocaron al decidir lanzar a pesar de las protestas e inquietudes de los ingenieros que se dieron cuenta de que los anillos de goma podían fallar.

Capítulo 4

Problemas y ejercicios

4.1 Problemas del principio de la pichonera (Principio de Dirichlet)

1. Los nombres de 10 personas son Alice, Bernard, Charles, mientras que sus apellidos son Lee, Mac, Montana, entonces al menos 2 personas tienen el mismo nombre y apellido.

   Hay 9 nombres y apellidos diferentes que seleccionar, pero son diez personas en total. Si consideramos a las 10 personas como las palomas y a los nombres y apellidos como los nidos por el principio de Dirichlet al menos dos personas tienen el mismo nombre y apellido.

2. Laura 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?.

   A partir de las reglas del reglas del producto y de la suma, el número total de palabras distintas posibles, de cuatro o menos letras es:

   $$274 + 273 + 272 + 27 = 551,880$$

   Estas 551,880 palabras si las consideramos como los nidos y las 600,000 palabras de la cinta como las palomas, de acuerdo al principio de Dirichlet al menos una palabra se repite en la cinta.

3. Cualquier subconjunto de tamaño seis del conjunto $S = 1, 2, 3, 4, 5, 6, 7, 8, 9$ debe contener al menos dos elementos cuya suma es 10.

Aquí los números 1, 2, 3, 4, 5, 6, 7, 8, 9 son las palomas, y los nidos son los subconjuntos 1, 9, 2, 8, 3, 7, 4, 6, 5. Cuando las palomas van a sus respectivos nidos, deben ocupar al menos uno de los subconjuntos cuyos miembros suman 10.

4. Juan regresa de la lavandería con 12 pares de calcetines, (cada par de distinto color) en una bolsa, al sacar cada calcetín de la bolsa aleatoriamente tendrá que sacar cuando mucho trece calcetines para obtener el par.

5. En un conjunto de 32 personas que cumplen años el mismo mes al menos 2 celebran su cumpleaños el mismo dá del mes.

   Si consideramos a las personas como palomas y a los dás del mes como los nidos y aplicamos el principio de Dirichlet, al menos dos o más personas cumpliran años el mismo día del mes.

4.2 10 Problemas de Combinaciones

1. Una tarjeta de circuito impreso se puede comprar de entre cinco proveedores. ¿En cuántas formas se pueden escoger tres proveedores de entre los cinco? Como es sólo importante cuáles tres se han escogido, no el orden de selección, el nḿero de formas es:

$$C(5,3) = \frac{5!}{3!(5-3)!} = 10$$

2. Encontrar el número de combinaciones al sacar dos de las cuatro letras A, B, C and D.

$$C(4,2) = \frac{4!}{2!(4-2)!} = 6$$

3. ¿Cuántos equipos distintos de 5 personas pueden escogerse de un grupo de 8 personas?

$$C(8,5) = \frac{8!}{5!(8-5)!} = 56$$

4. ¿Cuál es el número total de maneras que pueden escogerse 3 cartas de un total de 8 diferentes?

$$C(8,3) = \frac{8!}{3!(8-3)!} = 56$$

5. Si un niño escoge 4 juguetes de una caja que contiene 12 ¿Cuántos juguetes distintos puede escoger?

$$C(12,4) = \frac{12!}{4!(12-4)!} = 495$$

6. Un comité de 4 personas va a ser seleccionado de un grupo de 3 estudiantes de 4to. Año, 4 de 3ro. Y 5 de 2do. Si dos estudiantes de tercero no son elegibles. ¿De cuántas maneras puede seleccionarse el comité con 2 estudiantes de 2do., 1 de 3ro. Y 1 de 4to?

$$C(5,2)C(2,1)C(3,1) = (\frac{5!}{2!(5-2)!})(\frac{2!}{1!(2-1)!})(\frac{3!}{1!(3-1)!}) = 60$$

7. En una prueba de atletismo en la que participan 8 atletas se pueden clasificar sólo 3 para la final. ¿Cuántos grupos distintos de finalistas se pueden formar?

$$C(8,3) = \frac{8!}{3!(8-3)!} = 56$$

8. En una clase de 35 alumnos se quiere elegir un comité formado por 3 alumnos. ¿Cuántos comités diferentes se pueden formar?

$$C(35,3) = \frac{35!}{3!(35-3)!} = 6545$$

9. Una pizzería ofrece diez ingredientes adicionales para su pizza ¿De cuántas maneras un cliente puede seleccionar cuatro ingredientes adicionales para su pizza?

$$C(10,4) = \frac{10!}{4!(10-4)!} = 120$$

10. Susana es una de 7 oficinistas de una empresa pequeña. Se seleccionarán a 3 de estos trabajadores para formar parte de un comité ¿De cuántas maneras diferentes se puede seleccionar a tres de estas personas para formar parte del comité?

$$C(7,3) = \frac{7!}{3!(7-3)!} = 35$$

4.3 Problemas de Probabilidad de Canicas con y sin devolución

1. una caja contiene 10 canicas, 3 son rojas. Escogemos dos canicas al azar. Encuentra la probabilidad de que ninguna de ellas sea roja, (a) con remplazo y (b) sin remplazo.

   Consideremos los sucesos: A: Primera canica no roja. B: Segunda canica no roja.

   $$P(A) = \frac{7}{10}$$

   (a). Si le muestreo es con remplazo, la situación para la segunda elección es idéntica que para la primera y $P(B) = \frac{7}{10}$, los sucesos son independientes y la respuesta es:

   $$P(A \cap B) = P(A) \cdot P(B) = (0.7)(0.7) = 0.49$$

   (b). Si es sin remplazo hemos de tener en cuenta que una vez extraída la primer canica, quedan solo 9 y 3 deben ser rojas, así: $P(B|A) = \frac{6}{9} = \frac{2}{3}$, por lo que en este caso la respuesta es:

   $$P(A \cap B) = P(A)P(B|A) = (\frac{7}{10})(\frac{2}{3}) \approx 0.47$$

2. Tres máquinas A, B, C producen $60\%, 30\%$y 10%, respectivamente, el total de productos de una fábrica. Los porcentajes de falla de cada máquina son: 2%, 3% y 4%. Un producto es seleccionado al azar y este es defectuoso, encuentre la probabilidad de que el item haya sido producido por la máquina C.[10]

Solución:

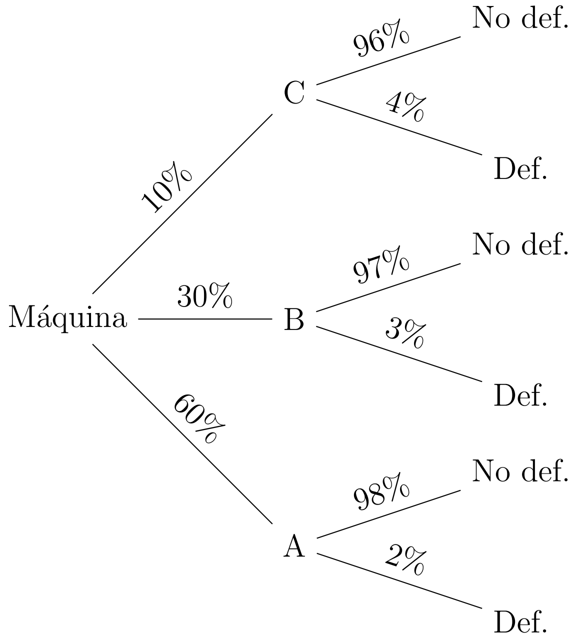


Figure 4.1: Diagrama de árbol del problema.

$$P(C|X) = \frac{P(C)P(X|C)}{P(A)P(X|A) + P(B)P(X|B) + P(C)P(X|C)}$$
$$= \frac{10 \times 4}{(60 \times 2) + (30 \times 3) + (10 \times 4)} = \frac{4}{25}$$

X= Defectuoso.

3. Una urna contiene 7 canicas rojas y 3 canicas blancas. Se sacan 3 canicas de la urna una tras otra. Hallar la probabilidad p de que las 2 primeras sean rojas y la tercera blanca.

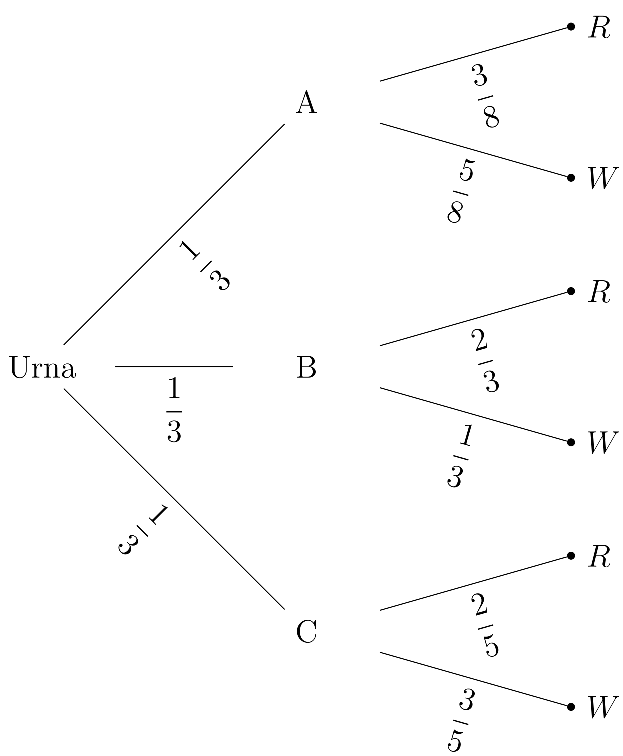
   La probabilidad de que la primera canica sea roja es de $\frac{7}{10}$ puesto que hay 7 rojas entre las 10 canicas. Si la primera canica es roja, entonces la probabilidad de que la segunda canica sea roja es $\frac{6}{9}$ puesto que quedan 6 rojas entre las 9 canicas restantes. Si las dos primeras son rojas, entonces la probabilidad de que la tercera sea blanca es $\frac{3}{8}$ puesto que quedan 3 blancas entre las 8 canicas restantes en la urna. Entonces por el teorema de la multiplicación.

$$p = (\frac{7}{10})(\frac{6}{9})(\frac{3}{8}) = \frac{7}{40}$$

4. Se nos dan tres urnas de la siguiente manera:
Una urna A contiene 3 canicas rojas y 5 blancas.
Una urna B contiene 2 canicas rojas y 1 blanca.
Una urna C contiene 2 canicas rojas y 3 blancas.

Se selecciona una urna al azar y se saca una bola de la urna. Si la canica es roja, ¿cuál es la probabilidad de que proceda de la urna A?

Se construye el diagrama de árbol de la siguiente manera:



Buscamos la probabilidad de que se seleccione A, dado que la canica es roja; esto es, $P(A|R)$. Con el fin de hallar $P(A|R)$ es necesario calcular primero $P(A \cap R)$ y $P(R)$.

La probabilidad de que se seleccione A y se saque una bola roja es

$$(\frac{1}{3})(\frac{3}{8}) = \frac{1}{8}$$

esto es,

$$P(A \cap R) = \frac{1}{8}$$

Puesto que hay tres trayectorias que conducen a bola roja

$P(R) = (\frac{1}{3})(\frac{3}{8}) + (\frac{1}{3})(\frac{2}{3}) + (\frac{1}{3})(\frac{2}{5}) = \frac{173}{360}$

Entonces:

$$P(A|R) = \frac{P(A \cap R)}{P(R)} = \frac{\frac{1}{8}}{\frac{173}{360}} = \frac{45}{173}$$

Alternadamente, por el teorema de Bayes.

$$P(A|R) = \frac{P(A)P(R|A)}{P(A)P(R|A) + P(B)P(R|B) + P(C)P(R|C)}$$

$$P(A|R) = \frac{(\frac{1}{3})(\frac{3}{8})}{(\frac{1}{3})(\frac{3}{8}) + (\frac{1}{3})(\frac{2}{3}) + (\frac{1}{3})(\frac{2}{5})} = \frac{45}{173}$$

5. Una moneda cargada de tal manera que la probabilidad de que caiga sol sea de 60% se lanza 5 veces. ¿Cuál es la probabilidad de que caiga algún águila?.

   Solución:

   La probabilidad de que ninguna águila caiga en 5 lanzamientos es:

$$\begin{aligned} P(NoAguilas) &= \left(\frac{60}{100}\right)^5 \\ P(AlgunAguila) &= 1 - P(NoAguilas) \\ &= 1 - \left(\frac{60}{100}\right)^5 \end{aligned}$$

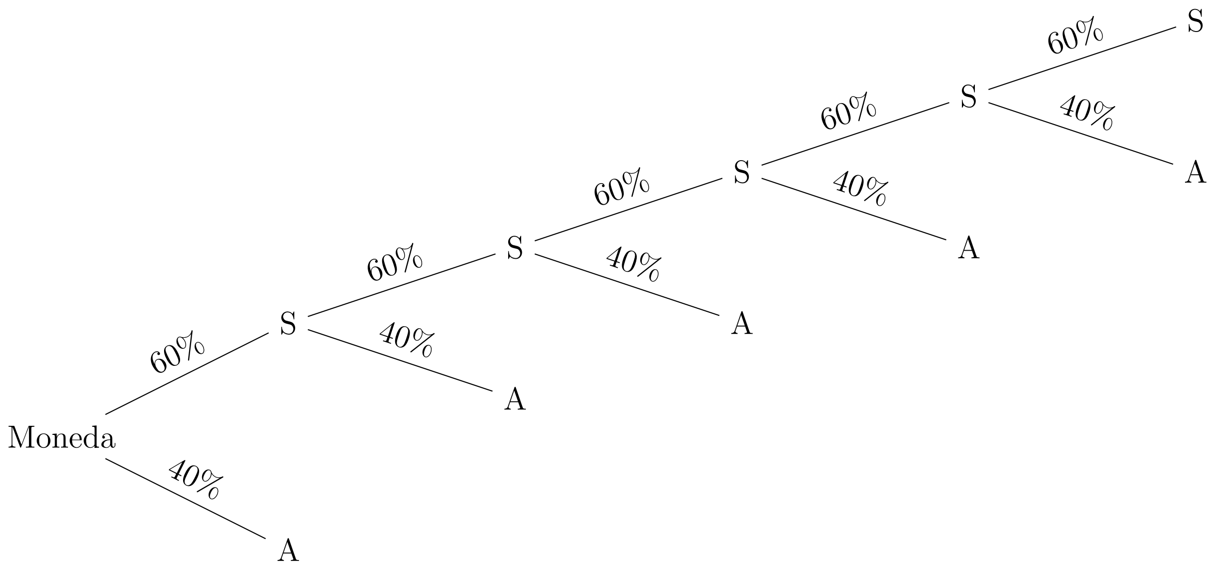
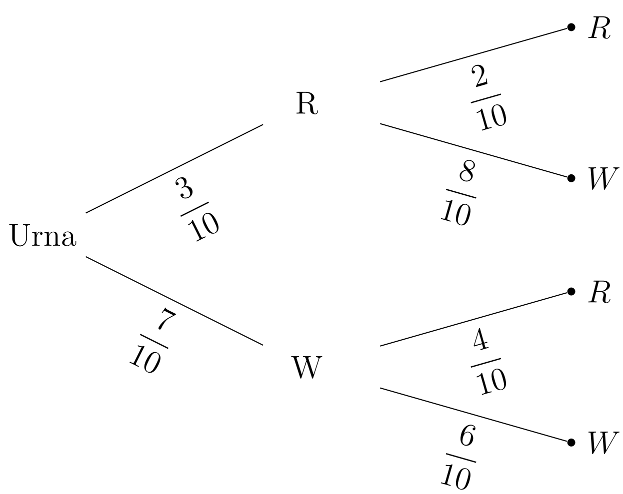


Figure 4.2: Diagrama de árbol simplificado.

6. Una urna tiene 3 canicas rojas y 7 blancas. Se saca una canica de la urna y se remplaza por una de otro color. Se saca de la urna una segunda canica.
(i). Hallar la probabilidad p de que la segunda canica sea roja.
(ii). Si ambas canicas son del mismo color, ¿cuál es la probabilidad p de que las dos sean blancas?

Diagrama de árbol:



(i). Dos trayectorias del diagrama de árbol conducen a la canica roja:

$$p = (\frac{3}{10})(\frac{2}{10}) + (\frac{7}{10})(\frac{4}{10}) = \frac{17}{50}.$$

(ii). La probabilidad de que ambas canicas fueran blancas es:

$$(\frac{7}{10})(\frac{6}{10}) = \frac{21}{50}$$

La probabilidad de que ambas canicas fueran del mismo color, esto es, la probabilidad del espacio muestral reducido, es:

$$(\frac{3}{10})(\frac{2}{10}) + (\frac{7}{10})(\frac{6}{10}) = \frac{12}{25}$$

por lo tanto la probabilidad condicional es:

$$p = \frac{\frac{21}{50}}{\frac{12}{25}} = \frac{7}{8}$$

7. Problema 1.44 (Murray R. Spiegel) Una caja contiene 8 bolas rojas, 3 blancas y 9 azules. Si se extraen 3 bolas aleatoriamente sin remplazamiento, determinar la probabilidad de que (a) las 3 bolas sena rojas, (b) las 3 bolas sean blancas, (c) 2 sean rojas y 1 blanca, (d) al menos 1 sea blanca, (e) se extraiga una de cada color, (f) las bolas sean extraídas en el orden rojo, blanco, azul.

- (a) Método 1. Sean R_1, R_2, R_3 los sucesos, "bola roja en la primera extracción", "bola roja en la segunda extracción", "bola roja en la tercer extracción", respectivamente. Así $R_1 \cap R_2 \cap R_3$ representa el suceso "las 3 bolas extraídas son rojas". De esta manera tenemos:

$$P(R_1 \cap R_2 \cap R_3) = P(R_1)P(R_2|R_1)P(R_3|R_1 \cap R_2) = (\frac{8}{20})(\frac{7}{19})(\frac{6}{18}) = \frac{14}{285}$$

- (a) Método 2.

  $$\text{Probabilidad pedida} = \frac{\text{numero de grupos de 3 bolas entre 8 rojas}}{\text{numero de grupos de 3 bolas entre 20}} = \frac{C(8,3)}{C(20,3)} = \frac{14}{285}$$

- (b). Empleando el segundo método indicado en la parte (a).

$$P(3 \text{ bolas blancas}) = \frac{C(3,3)}{C(20,3)} = \frac{1}{1140}$$

- (c). P(2 bolas rojas y 1 blanca)=

$$= \frac{\text{(grupos de 2 entre 8 bolas rojas)(grupos de 1 entre 3 bolas blancas)}}{\text{numero de grupos de 3 bolas entre 20}} =$$

$$= \frac{(C(8,2))(C(3,1))}{C(20,3)} = \frac{7}{95}$$

- (d). $P(\text{ninguna blanca}) = \dfrac{C(17,3)}{C(20,3)} = \dfrac{34}{57}$. Entonces:

$$P(\text{almenos una blanca}) = 1 - \frac{34}{57} = \frac{23}{57}$$

- (e). $P(\text{1 de cada color}) = \dfrac{(C(8,1))(C(3,1))(C(9,1))}{C(20,3)} = \dfrac{18}{95}$

- (f). $P(\text{extraer las bolas en orden rojo, blanco, azul}) = \dfrac{1}{3!}P(\text{1 de cada color})$

$= \dfrac{1}{6}(\dfrac{18}{95}) = \dfrac{3}{95}$, usando (e)

4.4 Propuesta de Examen

1. Dados los conjuntos $U = \{a, b, c, d, e, f, g\}$, $A = \{a, b, c, d, e\}$, $B = \{a, c, e, g\}$ y $C = \{b, e, f, g\}$. Hallar:

   - $A \cup C$
   - $B \cap A$
   - $C \diagdown B$
   - $B^c \cup C$
   - $C^c \cap A$
   - $(A \diagdown C)^c$
   - $(A \diagdown B^c)^c$
   - $(A \cap A^c)^c$

   Solución:

   - $A \cup C = \{a, b, c, d, e, f, g\} = U$
   - $B \cap A = \{a, c, e\}$
   - $C \diagdown B = \{b, f\}$
   - $B^c \cup C = \{b, d, e, f, g\}$
   - $C^c \cap A = \{a, c, d\}$
   - $(A \diagdown C)^c = \{b, e, f, g\}$
   - $(A \diagdown B^c)^c = \{b, d, f, g\}$
   - $(A \cap A^c)^c = U$

2. Demostrar que si $A \subset B$ y que $B \subset C$ entonces $A \subset C$.

   Solución: Sea x cualquier elemento de A, es decir $x \in A$. Entonces ya que $A \subset B$, es decir que cada elemento de A esta en B, se tiene $x \in B$. Puesto que también $B \subset C$, tenemos $x \in C$. Así cada elemento de A es un elemento de C y por lo tanto $A \subset C$.

3. Hallar el conjunto potencia $P(S)$ del conjunto $S = |1, 2, 3|$.

   Solución:

   EL conjunto potencia $P(S)$ del conjunto S es la clase de todos los subconjuntos de S; éstos son:$\{1, 3\}, \{2, 3\}, \{1, 2\}, \{1\}, \{2\}, \{3\}$ y el conjunto vacío ϕ. Por tanto:

$$P(S) = \{S, \{1,3\}, \{2,3\}, \{1,2\}, \{1\}, \{2\}, \{3\}, \phi\}$$

4. Simplifique la expresión $\frac{(n+2)!}{n!}$.

$$\frac{(n+2)!}{n!} = \frac{(n+2)(n+1)n(n-1)(n-2)...(3)(2)(1)}{n(n-1)(n-2)...(3)(2)(1)} = (n+2)(n+1) = n^2 + 3n + 2$$

5. ¿Cu'al es el número de permutaciones diferentes de las 11 letras de la palabra MISSISSIPPI?

   Solución: puesto que la palabra esta compuesta de una letra M, cuatro I, cuatro S y dos P, entonces:

$$\frac{11!}{(1!)(4!)(4!)(2!)} = 34,650 permutaciones$$

6. Se quieren sentar 5 hombres y 4 mujeres en una fila de manera que las mujeres ocupen los sitios pares. ¿De cuántas formas pueden sentarse?

   Solución: Los hombres pueden sentarse de $P(5,5)$ formas, y las mujeres de $P(4,4)$. Cada ordenación de los hombres puede asociarse a cada ordenación de las mujeres así que:

$$P(5,5)P(4,4) = 5!4! = (120)(24) = 2880$$

7. ¿Cuántas permutaciones distintas pueden formarse con todas las letras de cada una de las siguientes palabras: (i)tema (ii)campana (iii)estadísticas

   - $4! = 24$, puesto que hay cuatro letras distintas.
   - $\frac{7!}{3!} = 840$ puesto que hay 7 letras de las cuales 3 son a.
   - $\frac{12!}{3!2!2!2!}$, puesto que hay 12 letras de las cuales 3 son s, 2 son t, 2 son i y 2 son a.

8. ¿De cuántas maneras puede escogerse un comité, compuesto de 3 hombres y 2 mujeres, de un grupo de 7 hombre y 5 mujeres?

Solución: De los 7 hombres se pueden escoger 3 de $C(7,3)$ maneras, y de las 5 mujeres se pueden escoger 2 de $C(5,2)$ maneras.

Por consiguiente el comité puede escogerse de $(C(7,3))(C(5,2)) = \frac{(7)(6)(5)}{(1)(2)(3)}.\frac{(5)(4)}{(1)(2)} =$ 350 maneras.

9. Hallar el número de subconjuntos de un conjunto X que contiene n elementos.

   Solución: Hay dos posibilidades para cada elemento de X : o pertenece al subconjunto o no pertenece, por consiguiente hay:

$$\underbrace{(2)(2)(2)...(2)}_{\text{n veces}} = 2^n$$

   maneras de formar un subconjunto de X, o sea, hay 2^n subconjuntos diferentes de X.

10. Laura 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?.

    A partir de las reglas del reglas del producto y de la suma, el número total de palabras distintas posibles, de cuatro o menos letras es:

$$274 + 273 + 272 + 27 = 551,880$$

    Estas 551,880 palabras si las consideramos como los nidos y las 600,000 palabras de la cinta como las palomas, de acuerdo al principio de Dirichlet al menos una palabra se repite en la cinta, por lo tanto las 600,000 palabras no pueden ser distintas entre sí.

11. ¿De cuántas maneras puede un profesor escoger uno o más estudiantes de 6 elegibles?

    Solución: De acuerdo al problema, hay $2^6 = 64$ subconjuntos de conjuntos de 6 estudiantes. Sin embargo, el conjunto vacio debe de ser excluido puesto que se escogen uno o mas estudiantes. En consecuencia hay $2^6 - 1 = 64 - 1 = 63$ maneras de escoger los estudiantes.

12. En una clase hay 12 alumnos ¿de cuántas maneras los 12 alumnos pueden presentar 3 pruebas diferentes si a cada prueba le corresponden 4 alumnos?

    Solución: Hay $C(12,4)$ maneras de escoger 4 alumnos que tomen la primer prueba, posteriormente hay $C(8,4)$ maneras de escoger 4 estudiantes que tomen la segunda prueba. El resto de estudiantes toma la tercer prueba. O sea que por todas, hay $(C(12,4))(C(8,4)) = (495)(70) = 34,650$ maneras para que los estudiantes presenten las pruebas.

13. Una tarjeta de circuito impreso se puede comprar de entre cinco proveedores. ¿En cuántas formas se pueden escoger tres proveedores de entre los cinco?

    Solución: Como es sólo importante cuáles tres se han escogido, no el orden de selección, el nḿero de formas es:

$$C(5,3) = \frac{5!}{3!(5-3)!} = 10$$

14. ¿Cuántas placas para automóvil pueden hacerse si cada placa consta de dos letras diferentes seguidas de tres dígitos diferentes?

$$(26)(25)(10)(9)(8) = 468,000 \text{ placas diferentes}$$

15. Una pizzería ofrece diez ingredientes adicionales para su pizza ¿De cuántas maneras un cliente puede seleccionar cuatro ingredientes adicionales para su pizza?

$$C(10,4) = \frac{10!}{4!(10-4)!} = 120$$

16. En una prueba de atletismo en la que participan 8 atletas se pueden clasificar sólo 3 para la final. ¿Cuántos grupos distintos de finalistas se pueden formar?

$$C(8,3) = \frac{8!}{3!(8-3)!} = 56$$

17. Un comité de 4 personas va a ser seleccionado de un grupo de 3 estudiantes de 4to. Año, 4 de 3ro. Y 5 de 2do. Si dos estudiantes de tercero no son elegibles. ¿De cuántas maneras puede seleccionarse el comité con 2 estudiantes de 2do., 1 de 3ro. Y 1 de 4to?

$$C(5,2)C(2,1)C(3,1) = \frac{5!}{2!(5-2)!} x \frac{2!}{1!(2-1)!} x \frac{3!}{1!(3-1)!} = 60$$

18. una caja contiene 10 canicas, 3 son rojas. Escogemos dos canicas al azar. Encuentra la probabilidad de que ninguna de ellas sea roja, (a) con remplazo y (b) sin remplazo.

Consideremos los sucesos: A: Primera canica no roja. B: Segunda canica no roja.

$$P(A) = \frac{7}{10}$$

(a). Si le muestreo es con remplazo, la situación para la segunda elección es idéntica que para la primera y $P(B) = \frac{7}{10}$, los sucesos son independientes y la respuesta es:

$$P(A \cap B) = P(A) \cdot P(B) = (0.7)(0.7) = 0.49$$

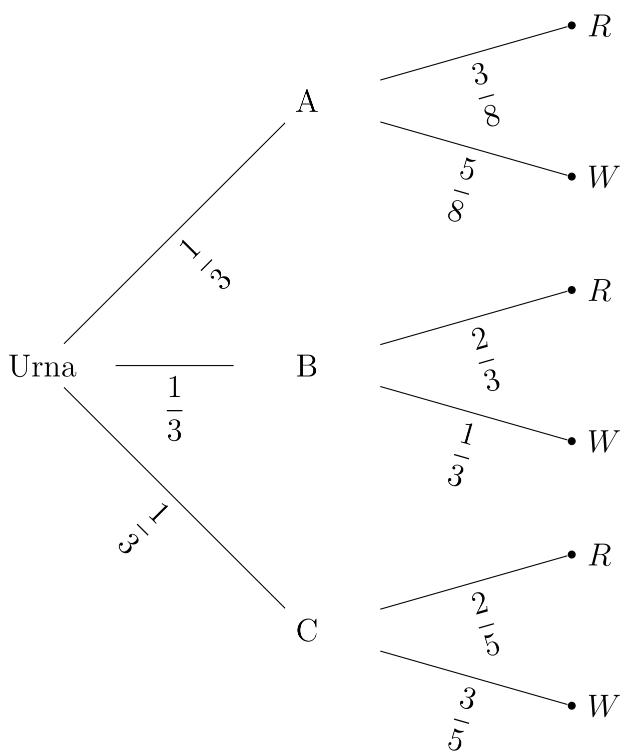
(b). Si es sin remplazo hemos de tener en cuenta que una vez extraída la primer canica, quedan solo 9 y 3 deben ser rojas, así: $P(B|A) = \frac{6}{9} = \frac{2}{3}$, por lo que en este caso la respuesta es:

$$P(A \cap B) = P(A)P(B|A) = (\frac{7}{10})(\frac{2}{3}) \approx 0.47$$

19. Se nos dan tres urnas de la siguiente manera:
Una urna A contiene 3 canicas rojas y 5 blancas.
Una urna B contiene 2 canicas rojas y 1 blanca.
Una urna C contiene 2 canicas rojas y 3 blancas.

Se selecciona una urna al azar y se saca una bola de la urna. Si la canica es roja, ¿cuál es la probabilidad de que proceda de la urna A?

Se construye el diagrama de árbol de la siguiente manera:



Buscamos la probabilidad de que se seleccione A, dado que la canica es roja; esto es, $P(A|R)$. Con el fin de hallar $P(A|R)$ es necesario calcular primero $P(A \cap R)$ y $P(R)$.

La probabilidad de que se seleccione A y se saque una bola roja es

$$(\frac{1}{3})(\frac{3}{8}) = \frac{1}{8}$$

esto es,

$$P(A \cap R) = \frac{1}{8}$$

Puesto que hay tres trayectorias que conducen a bola roja

$$P(R) = (\frac{1}{3})(\frac{3}{8}) + (\frac{1}{3})(\frac{2}{3}) + (\frac{1}{3})(\frac{2}{5}) = \frac{173}{360}$$

Entonces:

$$P(A|R) = \frac{P(A \cap R)}{P(R)} = \frac{\frac{1}{8}}{\frac{173}{360}} = \frac{45}{173}$$

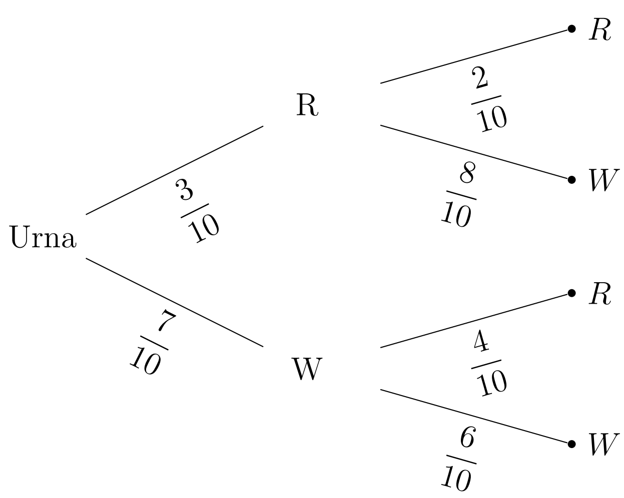
Alternadamente, por el teorema de Bayes.

$$P(A|R) = \frac{P(A)P(R|A)}{P(A)P(R|A) + P(B)P(R|B) + P(C)P(R|C)}$$

$$P(A|R) = \frac{(\frac{1}{3})(\frac{3}{8})}{(\frac{1}{3})(\frac{3}{8}) + (\frac{1}{3})(\frac{2}{3}) + (\frac{1}{3})(\frac{2}{5})} = \frac{45}{173}$$

20. Una urna tiene 3 canicas rojas y 7 blancas. Se saca una canica de la urna y se remplaza por una de otro color. Se saca de la urna una segunda canica.
(i). Hallar la probabilidad p de que la segunda canica sea roja.
(ii). Si ambas canicas son del mismo color, ¿cuál es la probabilidad p de que las dos sean blancas?

Diagrama de árbol:



(i). Dos trayectorias del diagrama de árbol conducen a la canica roja:

$$p = (\frac{3}{10})(\frac{2}{10}) + (\frac{7}{10})(\frac{4}{10}) = \frac{17}{50}.$$

(ii). La probabilidad de que ambas canicas fueran blancas es:

$$(\frac{7}{10})(\frac{6}{10}) = \frac{21}{50}$$

La probabilidad de que ambas canicas fueran del mismo color, esto es, la probabilidad del espacio muestral reducido, es:

$$(\frac{3}{10})(\frac{2}{10}) + (\frac{7}{10})(\frac{6}{10}) = \frac{12}{25}$$

por lo tanto la probabilidad condicional es:

$$p = \frac{\frac{21}{50}}{\frac{12}{25}} = \frac{7}{8}$$

Bibliografía

[1] Aguilera, A. (1993). *Metodos de Aproximacion de Estimadores en el ACP de un Proceso Estocastico*. PhD thesis, Tesis doctoral, Universidad de Granda.

[2] Alcalde Lancharro, E., García López, M., and Eduardo Alcalde Lancharro, M. G. L. (1994). *Informática básica*.

[Casanova Leal] Casanova Leal, H. Falacia de división en el comportamiento social del apostador en las loterías de cupones. *Fermentum*, 25(73).

[4] Contreras, J. M., Batanero, C., and Fernández, J. (2010). Problema de monty hall. un análisis semiótico. In *XIII Simposio de la SEIEM*.

[5] Cottliar, A. S. and Slavutsky, I. R. (2001). Telómeros y actividad de telomerasa: su participación en el envejecimiento y el desarrollo neoplásico. *Medicina (Buenos Aires)*, 61:335–42.

[de Memoria] de Memoria, I. G. Paginación en memoria virtual por: Edgar a. mendieta i. gestión de memoria.

[González] González, W. S. Teorema de congruencia de zeller: Aplicaciones.

[8] Hernández, A. C. (2007). Criterios de divisibilidad. *UNIVERSIDAD lA*, page 181.

[9] Johnson, D. E. et al. (2000). *Métodos multivariados aplicados al análisis de datos*. Number 519.5 J6.

[10] Lipschutz, S. S. J. J. C. et al. (2000). *Introducción a la probabilidad y estadística*. Number 519.5 L56.

[11] Mason, R. D., Lind, D. A., Marchal, W. G., and Lozano, M. C. H. (1998). *Estadística para administración y economía*. Number 658.00212 M376E 1998. Alfaomega México DF.

[12] Nvidia, C. (2007). Compute unified device architecture programming guide.

[13] Pla, L. E. (1986). *Análisis multivariado: método de componentes principales*. Number 519.535 P696. OEA, Washington, DC (EUA). Secretaría General. Programa Regional de Desarrollo Científico y Tecnológico.

[14] Severino, Y. and Daniel, J. (2012). La esfera celeste. B.S. thesis.

[15] Silver, L. (2000). Vuelta al edén más allá de la clonación en un mundo feliz. *Persona y Bioética*, (11-12).

[16] Spiegel, M. R., Losada Villasante, A., et al. (1970). *Estadística*.

[17] Winston, W. L. and Goldberg, J. B. (2005). Investigación de operaciones: aplicaciones y algoritmos.

[18] Youngson, R. (2003). *¡ Fiasco! aprendiendo de los errores de la ciencia*. Ediciones Robinbook.