

Probabilidad, procesos aleatorios e inferencia

Andrea Viridiana
Soto Miranda



Índice

1	Tipos de experimentos y espacios muestrales	7
1.1	Ejemplos de experimentos deterministas	7
1.2	Ejemplos de experimentos aleatorios	7
1.3	Ejemplos de Espacios muestrales	7
2	Probabilidad	9
2.1	Probabilidad Clásica	9
2.2	Probabilidad geométrica	9
2.3	Probabilidad frecuentista	10
2.4	Probabilidad subjetiva	10
2.5	Probabilidad axiomática	10
2.5.1	Axiomas de la probabilidad	10
2.6	Probabilidad lógica	11
3	Proceso estocástico	11
4	Tipos de frecuencias	12
4.1	Frecuencia absoluta	12
4.2	Frecuencia relativa	12
4.3	Frecuencia acumulada	12
5	Conceptos	12
6	Teorema del binomio	13
6.1	Teorema del binomio $n=2$	13
6.2	Teorema del binomio $n=3$	13
6.3	Teorema del binomio $n=4$	14
6.4	Teorema del binomio $n=5$	14
6.5	Teorema del binomio $n=6$	14
6.6	Teorema del binomio $n=7$	15
6.7	Teorema del binomio $n=8$	15
6.8	Teorema del binomio $n=9$	15
6.9	Teorema del binomio $n=10$	15
7	Triángulo de Pascal	17
7.1	Triángulos de Pascal	20
8	El principio de la pichonera	22
9	Permutaciones, Combinaciones	23
9.1	Ejemplos de permutaciones	23
9.2	Ejemplos de combinaciones	23
9.3	Generalizando el problema de caminos posibles para el caso de un cuadrado	24
9.4	Generalizando el problema de caminos posibles para el caso de un rectángulo	27
9.5	Generalizando el problema de caminos posibles para el caso de un cubo	28
10	Reseñas de artículos	29
10.1	Primates Count	29
10.2	Cicada-Generated Prime Numbers	29
10.3	Ant Odometer	29
10.4	Quipu	29
11	Histograma	30
11.1	Ejemplos de histogramas	30
12	Compuertas NAND y XOR con electrónica	31

12.1	$\text{NAND} = (AB)'$	31
12.2	$\text{XOR} = AB' + A'B$	31
13	Fórmula de Stirling	32
14	Telómeros	33
15	Oveja Dolly	33
16	Antinomias y paradojas	34
16.1	Paradoja de Cantor: el conjunto de todos los conjuntos.	34
16.2	Paradoja de Russell.	34
16.3	Paradoja de Burali-Forti: Conjunto de todos los números ordinales.	34
16.4	Conjunto de todos los números cardinales.	34
16.5	Paradoja del falso positivo	34
16.6	Paradoja Borel-Kolmogorov	36
17	Conjuntos	38
17.1	Leyes de De Morgan	38
17.2	Desarrollar $P(A \cup B \cup C)$	40
17.3	Desarrollar $P(A \cup B \cup C \cup D)$	41
17.4	Significado de $P(A B)$	43
18	Falacia del apostador	44
19	Teorema de Bayes, probabilidad bayesiana	45
19.1	Teorema de Bayes	45
19.2	Probabilidad bayesiana	45
19.3	Problemas teorema de Bayes	47
20	Histograma dado de tres caras	50
21	Gráficas en ROOT	52
21.1	Gráfica de $\ln(x!)$	52
21.2	Gráfica de $x \ln(x) - x$	53
21.3	Gráfica de $\ln(x!) - (x \ln(x) - x)$	54
22	Integrales y derivadas con Mathematica	56
23	Teorema del límite central	58
24	Densidad de probabilidad	59
25	Mínimos cuadrados	61
26	Explosión del Challenger	62
27	Ejercicios 1.4 y 1.5	63
27.1	Ejercicio 1.4	63
27.2	Ejercicio 1.5	69
28	Grafos por criterios de divisibilidad	70
28.1	Grafo de divisibilidad de 3	71
28.2	Grafo de divisibilidad de 5	71
28.3	Grafo de divisibilidad de 7	72
28.4	Grafo de divisibilidad de 11	73
29	Símbolo de Levi-Civita	74
29.1	Calcular c_1 , c_2 y c_3 de $\mathbf{a} \times \mathbf{b}$	74

30 Convenio de suma de Einstein	76
31 Esperanza matemática	76
32 El último teorema de Fermat y los Simpsons	77
33 Máquina Enigma	79
34 Ventaja	81
35 Examen	82
36 Congruencia zeller	89
36.1 Congruencia Zeller en Mathematica	91
36.2 Congruencia Zeller en gcc	92
37 Simulación de dos dados	94
38 Introducción.	94
39 Análisis Matemático.	94
39.1 Lanzamiento de un dado.	94
39.2 Lanzamiento de dos dados.	94
40 Análisis Experimental.	95
40.1 Método Heurístico.	95
40.1.1 Primer principio: Entender el problema.	95
40.1.2 Segundo Principio: Hacer un plan.	95
40.1.3 Tercer Principio: Poner el plan en marcha.	96
40.1.4 Cuarto Principio: Revisar y extender.	96
40.2 Entender el problema.	96
40.3 Hacer un plan.	96
40.4 Poner el plan en marcha.	96
40.4.1 Experimento con dados reales.	96
40.4.2 Experimento con dados virtuales.	97
40.4.3 Simulación en Mathematica.	97
40.4.4 Simulación en ROOT.	101
40.5 Revisar y extender.	105
41 Conclusiones.	106
42 EXAMEN DE PROBABILIDAD, PROCESOS ALEATORIOS E INFERENCIA	107
43 Skewness y Curtosis	113
44 Algoritmos para calcular π	116
44.1 Algoritmo de Gauss-Legendre	116
44.2 Algoritmo de Borwein	116
44.3 Algoritmo de Chudnovsky	117
45 Señal WOW	118
46 Resumen Capitulo 5 del libro Probabilidad Serie Schaum	119
46.1 Variables aleatorias	119
46.1.1 Distribución y Esperanza de una variable aleatoria finita	119
46.1.2 Varianza y Desviación estándar	120
46.1.3 Distribución conjunta	121
46.2 Variables aleatorias continuas	122
46.2.1 Esperanza, varianza y desviación estándar	122

46.2.2 <u>Ejemplo de variable aleatoria continua</u>	123
47 Análisis ROC de falsos positivos	124
48 Random Walk	125
48.1 Aplicaciones	125
48.2 Histogramas en root	127
48.3 Demostraciones Random Walk	135
49 Función Gamma	137
50 Movimiento Browniano	138
50.1 Simulación Movimiento Browniano 2D en root	138
50.2 Simulación Movimiento Browniano 3D en root	140
51 Funciones de distribución de probabilidad	142
51.1 Función de distribución de Rademacher	144
52 Procesos NP	145
53 Matrices	146
53.1 Matrices semejantes	146
53.2 Eigen Valores y Eigen Vectores	147
53.3 Matrices cuadradas poco densas	158
54 Cadenas de Markov	159
54.1 Ejemplo de Cadenas de Markov	159
54.2 Ejemplo del hombre que viaja en carro y en metro	164
55 Teorema del punto fijo	165
56 El hombre anumérico Capítulo 2 Probabilidad y Coincidencia	166
57 Teorema de los cuatro colores	170
58 Inteligencia Artificial	171
58.1 Clasificación Inteligencia Artificial	171
59 Los experimentos de Mendel	172

Lista de Figuras

1	Probabilidad geométrica dada una longitud	9
2	Probabilidad geométrica dada un área	9
3	Probabilidad geométrica dada un área	10
4	Construcción de el triángulo de Pascal hasta $m=1$	17
5	Construcción de el triángulo de Pascal hasta $m=3$	17
6	Construcción de el triángulo de Pascal de $n=0$ hasta $n=12$	18
7	Construcción de el triángulo de Pascal, relación entre las filas	18
8	Triangulo de Pascal, pares e impares de diferente color	20
9	Triángulo de Pascal, multiples de 5 de diferente color	20
10	Triángulo de Pascal, multiples de 7 de diferente color	21
11	Triángulo de Pascal, multiples de 11 de diferente color	21
12	Principio de la pichonera	22
13	Caminos posibles para un cuadrado de 1×1	24
14	Caminos posibles en un cuadrado de 2×2	25
15	Caminos posibles en un rectangulo de 3×2	27
16	Caminos posibles para un cubo	28
17	Ejemplo de Histograma	30
18	Histograma del peso de un grupo de personas	31
19	Compuerta NAND con circuitos	31
20	Compuerta XOR con circuitos	31
21	Diagrama de Venn de 3 conjuntos	40
22	Diagrama de Venn de 4 conjuntos	41
23	Diagrama de árbol de 3 urnas con bolas de diferentes colores	48
24	Diagrama de árbol de "Alimentos Mr. Pollo"	49
25	Histograma dado de 3 caras con $n=100$	50
26	Histograma dado de 3 caras con $n=1000$	50
27	Histograma dado de 3 caras con $n=10000$	51
28	Grafica de $\ln(x!)$ en root	52
29	Grafica de $x\ln(x)-x$ en root	53
30	Grafica de $\ln(x!)-(x\ln(x)-x)$ en root	55
31	Teorema del limite central	59
32	Densidad de probabilidad	60
33	Mínimos cuadrados	61
34	Explosión del Challenger	62
35	Esquema de longitudes para el problema 1.4	63
36	Los 4 tornillos tienen longitudes superiores a L_1	64
37	Al menos 3 tornillos estén entre L_1 y L_2	65
38	Al menos 3 tornillos estén entre L_1 y L_2	66
39	Al menos 3 tornillos estén entre L_1 y L_2	67
40	A lo sumo tres tornillos con longitud menor a L_2	68
41	Diagrama de árbol de lanzamiento de una moneda y un dado	69
42	Grafo de divisibilidad de 3	71
43	Grafo de divisibilidad de 5	71
44	Grafo de divisibilidad de 7	72
45	Grafo de divisibilidad de 11	73
46	Primer capítulo de los Simpson donde aparece el teorema de Fermat	77
47	Segundo capítulo de los Simpson donde aparece el teorema de Fermat	78
48	Diagrama de árbol del ejercicio 5	84
49	Diagrama de árbol del ejercicio 8	87
50	Resultado de 10 lanzamientos.	98
51	Distribución normal de 10 lanzamientos.	98
52	Resultado de 100 lanzamientos.	99
53	Distribución normal de 100 lanzamientos.	99
54	Resultado de 1,000 lanzamientos.	99
55	Distribución normal de 1,000 lanzamientos.	100

56	Resultado de 10,000 lanzamientos.	100
57	Distribución normal de 10,000 lanzamientos.	100
58	Resultado de 10 lanzamientos.	102
59	Distribución normal de 10 lanzamientos.	102
60	Resultado de 100 lanzamientos.	103
61	Distribución normal de 100 lanzamientos.	103
62	Resultado de 1,000 lanzamientos.	104
63	Distribución normal de 1,000 lanzamientos.	104
64	Resultado de 1,000,000 lanzamientos.	105
65	Distribución normal de 1,000,000 lanzamientos.	105
66	Diagrama del árbol del ejercicio 2	108
67	Diagrama de árbol del ejercicio 6	110
68	Número de diagonales en un pentágono	112
69	Función de densidad de probabilidad y función de distribución de probabilidad	121
70	Función de distribución de probabilidad continua	122
71	Función de distribución de Ejemplo de variable aleatoria continua	123
72	Caminata al azar en dos dimensiones	126
73	Caminata al azar un una dimensión	126
74	Tabla de posición final, dependiendo los pasos dados.	126
75	Random Walk $p = 0.5$ y $q = 0.5$	129
76	Random Walk $p = 0.5$ y $q = 0.5$	130
77	Random Walk $p = 0.4$ y $q = 0.6$	131
78	Random Walk $p = 0.5$ y $q = 0.5$	132
79	Random Walk $p = 0.6$ y $q = 0.4$	133
80	Random Walk $p = 0.5$ y $q = 0.5$	134
81	Movimiento browniano en 2 dimensiones.	139
82	Movimiento browniano en 2 dimensiones.	139
83	Movimiento browniano en 3 dimensiones.	141
84	Gráfica de la Ecuación Característica de la Matriz A	151
85	Mapa coloreado con cuatro colores	170
86	Las siete variantes de Pisum sativum estudiadas por Mendel en sus experimentos	172
87	Experimentos de Mendel: 100% Semilla Amarilla	173
88	Experimentos de Mendel: 75% Semilla Amarilla, 25% Semilla Verde	173
89	Las siete variantes de Pisum sativum estudiadas por Mendel en sus experimentos	173

1 Tipos de experimentos y espacios muestrales

1.1 Ejemplos de experimentos deterministas

En estadística, un suceso determinista es un experimento o fenómeno que da lugar a un resultado cierto o seguro, es decir, cuando partiendo de unas mismas condiciones iniciales tenemos la certeza de lo que va a suceder. La relación causa-efecto se conoce en su totalidad.

- Medir el volumen de un gas cuando la presión y la temperatura son constantes produce teóricamente siempre el mismo resultado.
- Medir el ángulo de un rayo de luz reflejado en un espejo resulta siempre en el mismo resultado cuando el ángulo de incidencia es el mismo y el resto de las condiciones son constantes.
- Observar la temperatura a la que hierve el agua a una altitud dada.
- Observar el tiempo que tarda un objeto en caer al suelo cuando se le deja caer desde una altura dada.
- El año tiene 12 meses.
- La distancia recorrida en un tiempo t y una velocidad v .
- Obtener la corriente I de un circuito formado por una tensión eléctrica E y una resistencia R .
- Medir la resistencia R dado un resistor.
- El Domingo sigue después del día Lunes.
- Observar el tiempo de carga y descarga de un capacitor.

1.2 Ejemplos de experimentos aleatorios

En teoría de la probabilidad un experimento aleatorio es aquel que bajo el mismo conjunto aparente de condiciones iniciales, puede presentar resultados diferentes, es decir, no se puede predecir o reproducir el resultado exacto de cada experiencia particular.

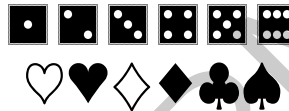
- Lanzar una moneda al aire, observar el resultado si el resultado es Águila o Sol.
- Lanzar un dado al aire y observar el número que cae, el resultado puede ser: 1,2,3,4,5,6.
- De una baraja de póker se extrae una carta, ¿Qué probabilidad hay de que la carta extraída sea de color negro?.
- En una caja se introducen 20 canicas de las cuales: 5 son azules, 7 verdes y las demás rojas. Si se extraen dos canicas ¿qué probabilidad hay de que salgan dos canicas rojas?.
- Registrar el número ganador en un juego de lotería.
- Registrar el consumo de electricidad de una casa-habitación en un día determinado.
- Observar el precio que tendrá el petróleo en un año.
- Registrar la precipitación pluvial anual en una zona geográfica determinada.
- Predecir el clima.
- Se fabrican artículos en una línea de producción y se cuenta el número de artículos defectuosos producidos en un periodo de 24 horas.



1.3 Ejemplos de Espacios muestrales

El espacio muestral de un experimento es el conjunto de todas las posibles salidas y se denota por la letra griega Ω (Omega mayúscula).

- Los lados de una moneda, $\Omega = \{\text{Águila}, \text{Sol}\}$.
- Las caras de un dado, $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- Número total de águila que sale al lanzar una moneda cuatro veces, $\Omega = \{0, 1, 2, 3, 4\}$.
- El número (N) de focos defectuosos en la producción total de 24 horas, $\Omega = \{0, 1, 2, \dots, N\}$.
- Las cartas de una baraja de poker.
- Todos los números de lotería que participaran cierto día.
- Las pelotitas que se encuentran en una tombola para participar en un sorteo.
- La colección de tarjetas de un juego de Bingo.



2 Probabilidad

La matemática es la lógica de la seguridad; probabilidad es la lógica de la incertidumbre. La probabilidad es extremadamente útil en una amplia variedad de campos, ya que proporciona herramientas para entender y explicar sucesos.

2.1 Probabilidad Clásica

La probabilidad clásica se define como el cociente entre el número de casos favorables ($\#A$) y el número total de casos ($\#\Omega$) dada por la Eq(1) [9].

$$P(A) = \frac{\#A}{\#\Omega} \quad (1)$$

2.2 Probabilidad geométrica

Es una extensión de la definición de probabilidad clásica en donde ahora la probabilidad de un evento A se calcula ya no a través de su cardinalidad sino mediante la determinación de su longitud, área, volumen o alguna característica geométrica según el experimento que se este realizando, y se representa por medio de la Eq(2),Eq(3) y Eq(4).

- Dada una longitud:

$$P(A) = \frac{\text{longitud}(A)}{\text{longitud}(\Omega)} \quad (2)$$

En la Fig (1) se muestra la probabilidad geométrica dada una longitud.

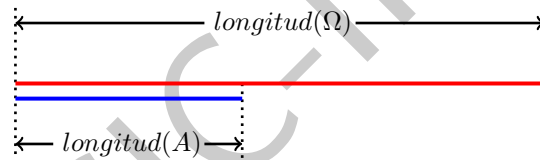


Fig. 1: Probabilidad geométrica dada una longitud

- Dada un área

$$P(A) = \frac{\text{area}(A)}{\text{area}(\Omega)} \quad (3)$$

En la Fig (2) se muestra la probabilidad geométrica dada un área.

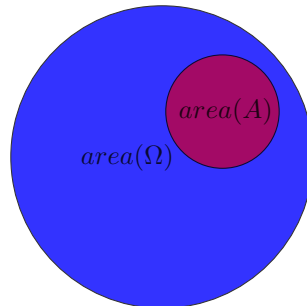


Fig. 2: Probabilidad geométrica dada un área

- Dado un volumen

$$P(A) = \frac{\text{volumen}(A)}{\text{volumen}(\Omega)} \quad (4)$$

En la Fig (3) se muestra la probabilidad geométrica dado un volumen.

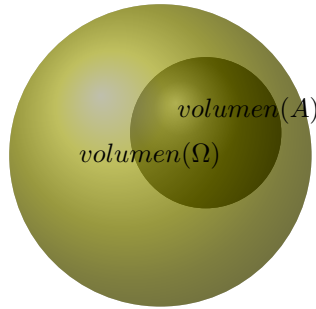


Fig. 3: Probabilidad geométrica dada un área

2.3 Probabilidad frecuentista

La probabilidad frecuentista en un experimento aleatorio se define como el cociente entre el número de veces que ocurre el evento A [$n(A)$] y el número total de repeticiones del experimento (n) dada por la Eq(5) [9].

$$P(A) = \frac{n(A)}{n} \quad (5)$$

2.4 Probabilidad subjetiva

Se refiere a la probabilidad que una persona asigna a un posible resultado de algún proceso y representa su propio criterio de la probabilidad de que se obtenga el resultado. Este juicio se basa en creencias e información sobre el proceso de cada persona. Otra persona, que puede tener diferentes creencias o información diferente, puede asignar una probabilidad diferente al mismo resultado.

2.5 Probabilidad axiomática

En la definición axiomática de la probabilidad no se establece la forma explícita de calcular las probabilidades sino únicamente se proponen las reglas que el cálculo de probabilidades debe satisfacer. Los siguientes tres postulados o axiomas fueron establecidos en 1933 por el matemático ruso Andrey Nikolaevich Kolmogorov.

2.5.1 Axiomas de la probabilidad

1. $-P(A) \geq 0$
2. $-P(\Omega) = 1$
3. $-P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k)$

2.6 Probabilidad lógica

La probabilidad es una propiedad objeto de los fenómenos o experimentos aleatorios puesto que admiten que es una característica intrínseca más de los sucesos alternativos en que pueden plasmarse estos fenómenos (como el peso, la longitud, la densidad, etc.).

Según los logicistas, la probabilidad como grado de creencia racional asociado a cada proposición o grupo de proposiciones, es una consecuencia que se desprende del diferente grado de implicación que pueda establecerse entre grupos de proposiciones o enunciados. La probabilidad está directamente relacionada con el cuerpo de evidencia existente, pudiendo variar con éste, de manera que los términos *<<cierto>>* y *<<probable>>* describen diversos grados de creencia racional en una proposición, aunque a posteriori la evidencia hará que las proposiciones sean simplemente verdaderas o falsas.

Para comprender mejor el sentido de la **probabilidad lógica** partamos, como ejemplo, de un silogismo donde las premisas iniciales sean:

$$A_1 \rightarrow \text{Todos los perros ladran}$$

$$A_2 \rightarrow \text{Pluto es un perro}$$

y la conclusión:

$$B \rightarrow \text{Luego Pluto ladra.}$$

Las premisas A_1 y A_2 llevan necesariamente, es decir, con absoluta certeza, a la conclusión B , por lo que el **grado de implicación** entre ellas será total. En este caso, los logicistas dicen que la probabilidad lógica es igual a la unidad, pues otorgan el campo de variación $[0;1]$ a su medida de probabilidad, asimilando $P=1$ a la situación de certeza.

Por el contrario, cuando la afirmación de las premisas iniciales conducen al no cumplimiento de la conclusión, ésta es imposible, y por tanto, la probabilidad asignada es cero.

Sin embargo, hay muchos otros casos, en nuestras construcciones lógicas, donde el grado de implicación no toma ninguno de estos dos valores extremos, sino que se debe situar en algún punto del continuo que existe entre cero y uno.

Consideremos un ejemplo de esta categoría, donde las premisas iniciales fueran:

$$A_1 \rightarrow \text{Siempre que llueve está nublado}$$

$$A_2 \rightarrow \text{Hoy está nublado}$$

y la conclusión:

$$B \rightarrow \text{Luego hoy lloverá.}$$

Observamos que ahora no existe certeza en la implicación de B respecto a los enunciados A_1 y A_2 , sino cierto grado de creencia racional que depende de la evidencia anterior para que, dadas estas premisas iniciales, se produzca la conclusión contenida en B .

3 Proceso estocástico

Un proceso estocástico es un modelo matemático de un experimento probabilístico que evoluciona en el tiempo y genera una secuencia de valores numéricos. Por ejemplo, un proceso estocástico se puede utilizar para modelar:

- La secuencia de los precios diarios de una población.
- La secuencia de las puntuaciones en un partido de fútbol.
- La secuencia de tiempos de fallo de una máquina;

Cada valor numérico en la secuencia se modela mediante una variable aleatoria, por lo que un proceso estocástico es simplemente una secuencia (finito o infinito) de variables aleatorias y no representa un cambio conceptual importante de nuestro marco básico.

4 Tipos de frecuencias

4.1 Frecuencia absoluta

La frecuencia absoluta es el número de veces que aparece un determinado valor en un estudio estadístico. Se representa por n_i . La suma de las frecuencias absolutas es igual al número total de datos, que se representa por N .

$$n_1 + n_2 + n_3 + \dots + n_n = N$$

4.2 Frecuencia relativa

La frecuencia relativa es el cociente entre la frecuencia absoluta de un determinado valor y el número total de datos. Esta representada por la Eq(6).

$$f_i = \frac{n_i}{N} \quad (6)$$

La suma de las frecuencias relativas es igual a 1.

4.3 Frecuencia acumulada

La frecuencia acumulada es la suma de las frecuencias absolutas de todos los valores inferiores o iguales al valor considerado. Se representa por F_i .

5 Conceptos

- Riesgo: Contingencia o proximidad de un daño.
- Peligrosidad: Riesgo o posibilidad de un daño o un mal.
- Vulnerabilidad: es la cualidad de vulnerable (que es susceptible de ser lastimado o herido ya sea física o moralmente).
- Amenaza: Anuncio de un mal o peligro.

6 Teorema del binomio

El teorema del binomio esta dado por la Eq(7) y el coeficiente binomial por Eq(8)

$$(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k \quad (7)$$

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \quad (8)$$

Si sustituimos Eq(8) en Eq(7) nos queda Eq(9)

$$(x + y)^n = \sum_{k=0}^n \frac{n!}{k!(n-k)!} x^{n-k} y^k \quad (9)$$

Desarrollar el teorema del binomio para n=2 hasta n=10

6.1 Teorema del binomio n=2

- Si n=2:

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

Cuando k=0:

$$\frac{2!}{0!(2-0)!} x^{2-0} y^0$$

El valor de 0! es 1 por lo que queda: x^2

Cuando k=1:

$$\frac{2!}{1!(2-1)!} x^{2-1} y^1$$

Haciendo las operaciones correspondientes queda: $2xy$

Y por último cuando k=2:

$$\frac{2!}{2!(2-2)!} x^{2-2} y^2$$

De este último término queda: y^2

Aplicando la suma en los términos obtenidos queda:

$$(x + y)^2 = x^2 + 2xy + y^2$$

6.2 Teorema del binomio n=3

- Si n=3

$$(x + y)^3 = \sum_{k=0}^3 \frac{3!}{k!(3-k)!} x^{3-k} y^k$$

$$(x + y)^3 = \frac{3!}{0!(3-0)!} x^{3-0} y^0 + \frac{3!}{1!(3-1)!} x^{3-1} y^1 + \frac{3!}{2!(3-2)!} x^{3-2} y^2 + \frac{3!}{3!(3-3)!} x^{3-3} y^3$$

$$(x + y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$$

6.3 Teorema del binomio n=4

- Si n=4

$$(x + y)^4 = \sum_{k=0}^4 \frac{4!}{k!(4-k)!} x^{4-k} y^k$$

$$(x + y)^4 = \frac{4!}{0!(4-0)!} x^{4-0} y^0 + \frac{4!}{1!(4-1)!} x^{4-1} y^1 + \frac{4!}{2!(4-2)!} x^{4-2} y^2 + \frac{4!}{3!(4-3)!} x^{4-3} y^3 + \frac{4!}{4!(4-4)!} x^{4-4} y^4$$

$$(x + y)^4 = x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4$$

6.4 Teorema del binomio n=5

- Si n=5

$$(x + y)^5 = \sum_{k=0}^5 \frac{5!}{k!(5-k)!} x^{5-k} y^k$$

$$(x + y)^5 = \frac{5!}{0!(5-0)!} x^{5-0} y^0 + \frac{5!}{1!(5-1)!} x^{5-1} y^1 + \frac{5!}{2!(5-2)!} x^{5-2} y^2 + \frac{5!}{3!(5-3)!} x^{5-3} y^3 + \frac{5!}{4!(5-4)!} x^{5-4} y^4 +$$

$$\frac{5!}{5!(5-5)!} x^{5-5} y^5$$

$$(x + y)^5 = x^5 + 5x^4y + 10x^3y^2 + 10x^2y^3 + 5xy^4 + y^5$$

6.5 Teorema del binomio n=6

- Si n=6

$$(x + y)^6 = \sum_{k=0}^6 \frac{6!}{k!(6-k)!} x^{6-k} y^k$$

$$(x + y)^6 = \frac{6!}{0!(6-0)!} x^{6-0} y^0 + \frac{6!}{1!(6-1)!} x^{6-1} y^1 + \frac{6!}{2!(6-2)!} x^{6-2} y^2 + \frac{6!}{3!(6-3)!} x^{6-3} y^3 + \frac{6!}{4!(6-4)!} x^{6-4} y^4 +$$

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

$$(x + y)^6 = x^6 + 6x^5y + 15x^4y^2 + 20x^3y^3 + 15x^2y^4 + 6xy^5 + y^6$$

6.6 Teorema del binomio n=7

- Si n=7

$$(x+y)^7 = \sum_{k=0}^7 \frac{7!}{k!(7-k)!} x^{7-k} y^k$$

$$(x+y)^7 = \frac{7!}{0!(7-0)!} x^{7-0} y^0 + \frac{7!}{1!(7-1)!} x^{7-1} y^1 + \frac{7!}{2!(7-2)!} x^{7-2} y^2 + \frac{7!}{3!(7-3)!} x^{7-3} y^3 + \frac{7!}{4!(7-4)!} x^{7-4} y^4 +$$

$$\frac{7!}{5!(7-5)!} x^{7-5} y^5 + \frac{7!}{6!(7-6)!} x^{7-6} y^6 + \frac{7!}{7!(7-7)!} x^{7-7} y^7$$

$$(x+y)^7 = x^7 + 7x^6y + 21x^5y^2 + 35x^4y^3 + 35x^3y^4 + 21x^2y^5 + 7xy^6 + y^7$$

6.7 Teorema del binomio n=8

- Si n=8

$$(x+y)^8 = \sum_{k=0}^8 \frac{8!}{k!(8-k)!} x^{8-k} y^k$$

$$(x+y)^8 = \frac{8!}{0!(8-0)!} x^{8-0} y^0 + \frac{8!}{1!(8-1)!} x^{8-1} y^1 + \frac{8!}{2!(8-2)!} x^{8-2} y^2 + \frac{8!}{3!(8-3)!} x^{8-3} y^3 + \frac{8!}{4!(8-4)!} x^{8-4} y^4 +$$

$$\frac{8!}{5!(8-5)!} x^{8-5} y^5 + \frac{8!}{6!(8-6)!} x^{8-6} y^6 + \frac{8!}{7!(8-7)!} x^{8-7} y^7 + \frac{8!}{8!(8-8)!} x^{8-8} y^8$$

$$(x+y)^8 = x^8 + 8x^7y + 28x^6y^2 + 56x^5y^3 + 70x^4y^4 + 56x^3y^5 + 28x^2y^6 + 8xy^7 + y^8$$

6.8 Teorema del binomio n=9

- Si n=9

$$(x+y)^9 = \sum_{k=0}^9 \frac{9!}{k!(9-k)!} x^{9-k} y^k$$

$$(x+y)^9 = \frac{9!}{0!(9-0)!} x^{9-0} y^0 + \frac{9!}{1!(9-1)!} x^{9-1} y^1 + \frac{9!}{2!(9-2)!} x^{9-2} y^2 + \frac{9!}{3!(9-3)!} x^{9-3} y^3 + \frac{9!}{4!(9-4)!} x^{9-4} y^4 +$$

$$\frac{9!}{5!(9-5)!} x^{9-5} y^5 + \frac{9!}{6!(9-6)!} x^{9-6} y^6 + \frac{9!}{7!(9-7)!} x^{9-7} y^7 + \frac{9!}{8!(9-8)!} x^{9-8} y^8 + \frac{9!}{9!(9-9)!} x^{9-9} y^9$$

$$(x+y)^9 = x^9 + 9x^8y + 36x^7y^2 + 84x^6y^3 + 126x^5y^4 + 126x^4y^5 + 84x^3y^6 + 36x^2y^7 + 9xy^8 + y^9$$

6.9 Teorema del binomio n=10

- Si n=10

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

$$(x+y)^{10} = \frac{10!}{0!(10-0)!} x^{10-0} y^0 + \frac{10!}{1!(10-1)!} x^{10-1} y^1 + \frac{10!}{2!(10-2)!} x^{10-2} y^2 + \frac{10!}{3!(10-3)!} x^{10-3} y^3 +$$

$$\frac{10!}{4!(10-4)!} x^{10-4} y^4 + \frac{10!}{5!(10-5)!} x^{10-5} y^5 + \frac{10!}{6!(10-6)!} x^{10-6} y^6 + \frac{10!}{7!(10-7)!} x^{10-7} y^7 +$$

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

$$(x+y)^{10} = x^{10} + 10x^9y + 45x^8y^2 + 120x^7y^3 + 210x^6y^4 + 252x^5y^5 + 210x^4y^6 + 120x^3y^7 +$$

$$45x^2y^8 + 10xy^9 + y^{10}$$

CIC-IPN

7 Triángulo de Pascal

El triángulo de Pascal se construye de la siguiente manera: se comienza en el número "1" centrado en la parte superior; después se escriben una serie de números en las casillas situadas en sentido diagonal descendente, a ambos lados, del siguiente modo: se suman las parejas de cifras situadas horizontalmente ($1 + 1$), y el resultado (2) se escribe debajo de dichas casillas; el proceso continúa escribiendo en las casillas inferiores la suma de las dos cifras situadas sobre ellas ($1 + 2 = 3$), etc. De manera general, esto se cumple así debido a la regla de Pascal, que indica que:

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}, 0 \leq k \leq n \quad (10)$$

Por lo tanto de esta manera podemos obtener todos los números combinatorios. Comenzamos con:

$$\binom{0}{0} = 1$$

El cual colocamos en el centro de la página. Los siguientes dos son:

$$\binom{1}{0} = 1$$

$$\binom{1}{1} = 1$$

que colocamos debajo, a ambos lados que habíamos colocado inicialmente, de modo que éste quede en el centro del espacio que separa los dos números nuevos como se muestra en la Fig(4).



Fig. 4: Construcción de el triángulo de Pascal hasta $m=1$

Veamos como se construye la fila que corresponde a $m = 3$. Ambos extremos valen 1: $\binom{3}{0} = \binom{3}{3} = 1$. El resto de los espacios los llenamos sumando en cada caso los dos valores que se encuentran por encima del espacio en cuestión: $\binom{3}{1} = 1 + 2 = 3$, $\binom{3}{2} = 2 + 1 = 3$ como se muestra en la Fig(5)

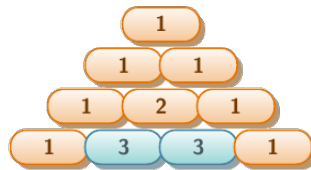


Fig. 5: Construcción de el triángulo de Pascal hasta $m=3$

Si continuamos este proceso inductivamente obtenemos el triángulo conocido como triángulo de Pascal y se va construyendo fila por fila como se muestra en la Fig(6).

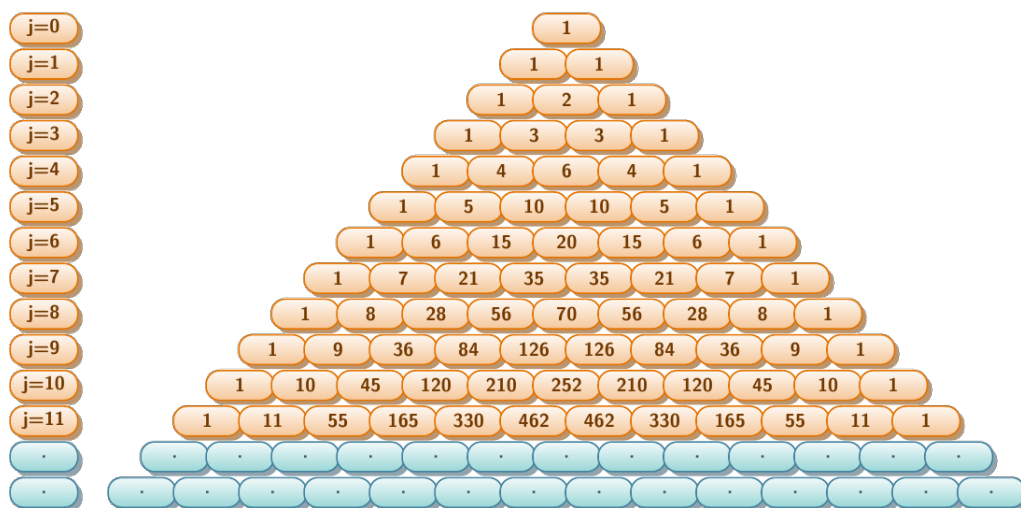


Fig. 6: Construcción de el triángulo de Pascal de $n=0$ hasta $n=12$

La fila j tiene $j + 1$ números, que corresponde a los números combinatorios $\binom{j}{i}$, para $0 \leq i \leq j$, es decir que cada fila comienza con el número combinatorio $\binom{j}{0}$. Observamos, en consecuencia, que el número que aparece en el lugar $i + 1$ de la fila j , es el número combinatorio $\binom{j}{i}$, por ejemplo, para hallar $\binom{7}{4}$ buscamos el lugar 5 de la fila 7 obtenemos $\binom{7}{4} = 35$

Otra manera de construir el triángulo es la siguiente. Cambiamos los números por puntos o nodos. Escribimos un 1 sobre el vértice superior, y luego, sobre cada nodo, el número de maneras que hay para llegar a este punto a partir del vértice superior, moviéndonos únicamente hacia abajo. El resultado es el triángulo de Pascal y se puede observar en la Fig(7).

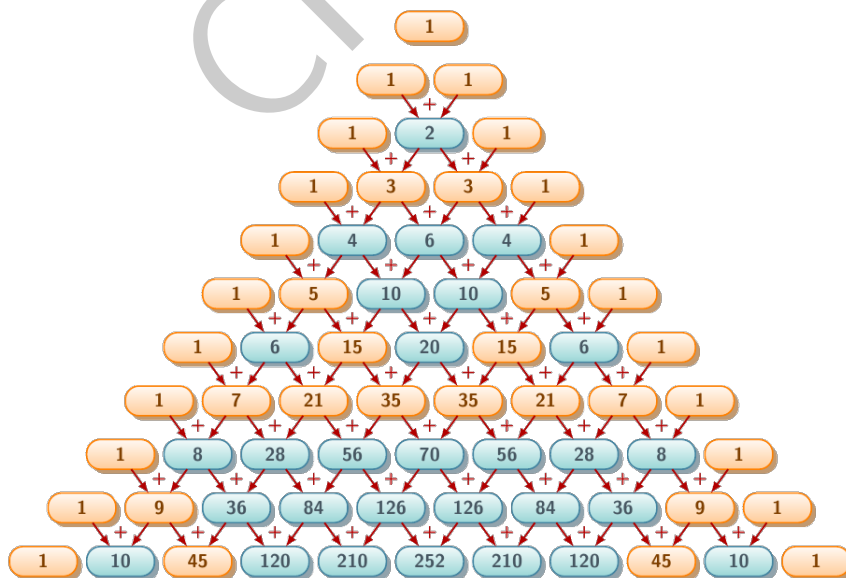


Fig. 7: Construcción de el triángulo de Pascal, relación entre las filas

Veamos una propiedad interesante del triángulo de Pascal. Si evaluamos la suma de los números en cada fila obtenemos 1, 2, 4, 8, 16, 32, 64, 128, etc. Parece natural la conclusión de que la suma de la n -ésima fila es 2^n . Esto es cierto y podemos probarlo por inducción. Sabemos que es cierto para las primeras filas. Para probar el paso inductivo observamos que cada número de la n -ésima fila es sumando para formar dos números de la siguiente fila: los que están por debajo de él, a ambos lados. Por lo tanto la suma de los números de la fila $n + 1$ es dos veces la suma de los números de la fila anterior. Esto completa el paso inductivo.

Si escribimos esta relación explícitamente obtenemos la siguiente identidad:

$$\binom{m}{0} = 1 + \binom{m}{1} = 1 + \dots + \binom{m}{m-1} + \binom{m}{m} = 2^m$$

También podemos observar que todas las cifras escritas en cada fila del triángulo, corresponden a los coeficientes del desarrollo binomial de la potencia de una suma:

$$(a + b)^0 = 1$$

$$(a + b)^1 = a + b$$

$$(a + b)^2 = a^2 + 2ab + b^2$$

$$(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$$

7.1 Triángulos de Pascal

- Triángulo de Pascal con los coeficientes pares en color azul y los impares en color naranja Fig(8):

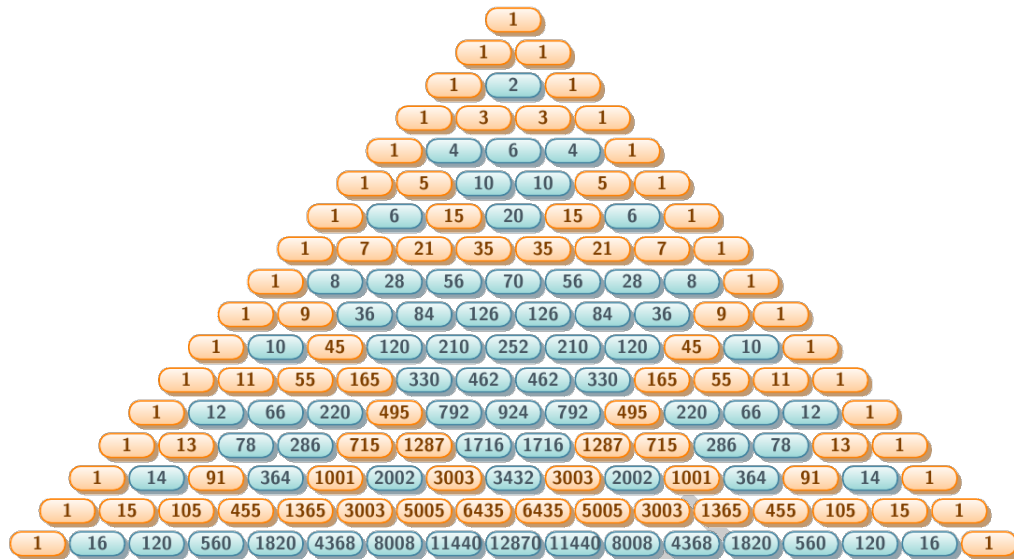


Fig. 8: Triángulo de Pascal, pares e impares de diferente color

- Triángulo de Pascal con los coeficientes múltiplos de 5 en color azul y los demás en color naranja Fig(9).

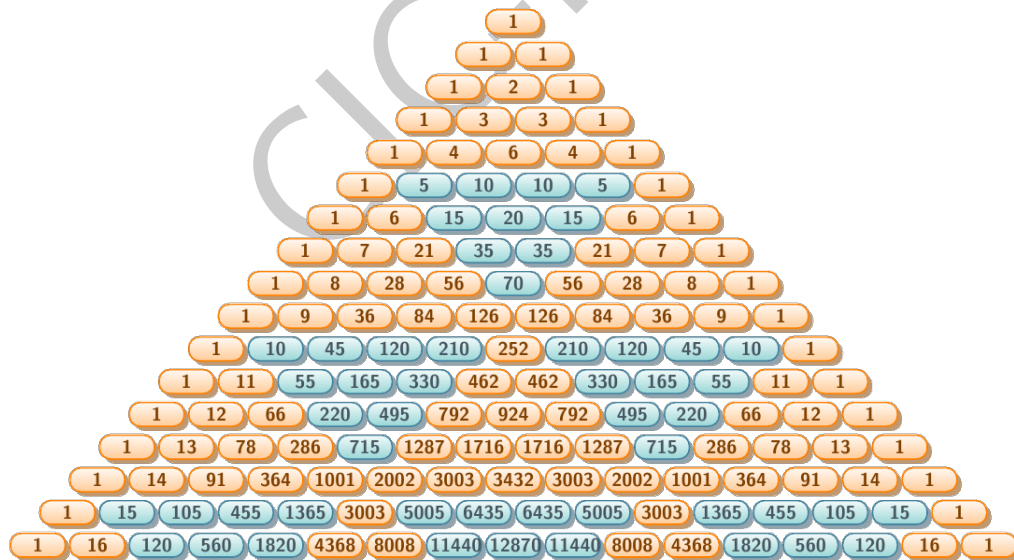


Fig. 9: Triángulo de Pascal, múltiplos de 5 de diferente color

- Triángulo de Pascal con los coeficientes múltiplos de 7 en color azul y los demás en color naranja Fig(10).

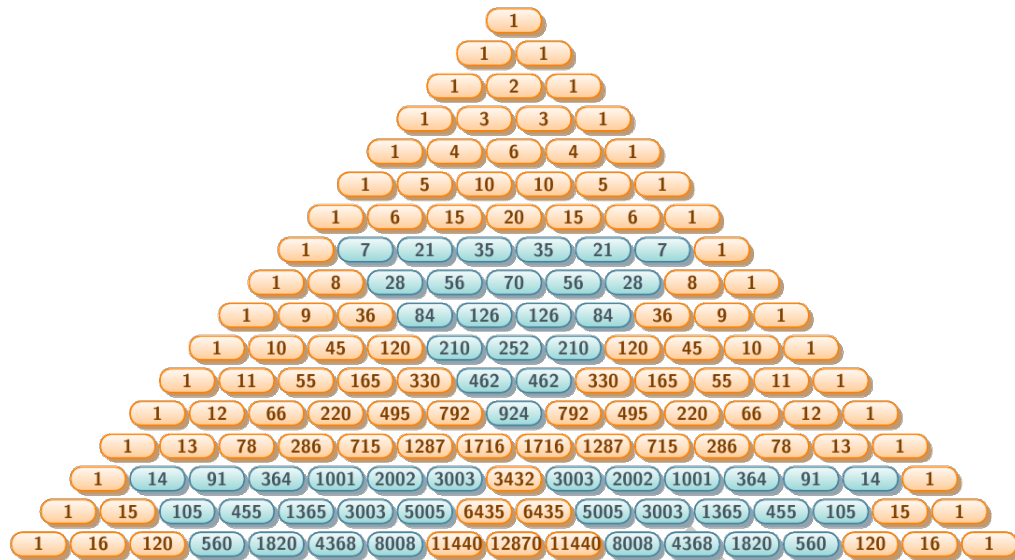


Fig. 10: Triángulo de Pascal, múltiplos de 7 de diferente color

- Triángulo de Pascal con los coeficientes múltiplos de 11 en color azul y los demás en color naranja Fig(11).

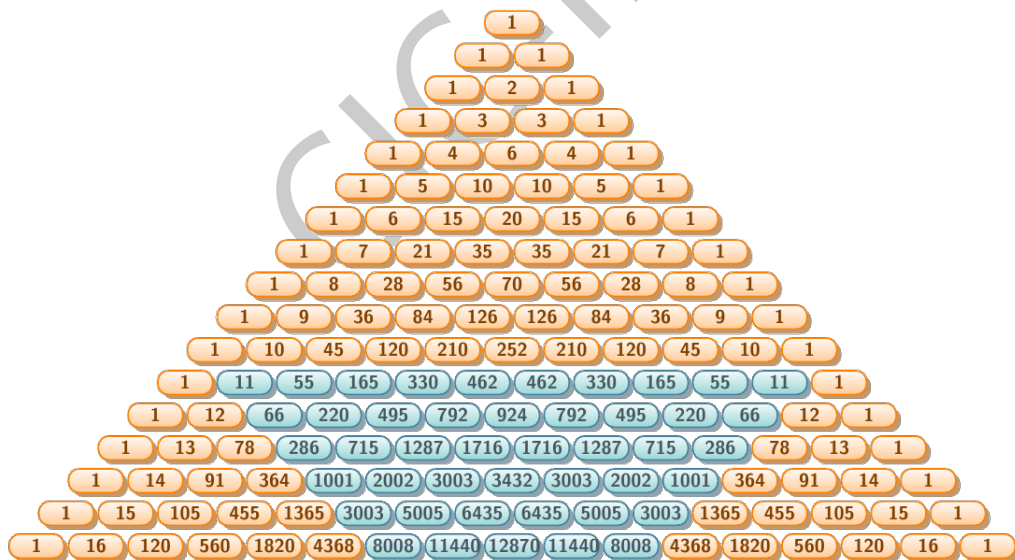


Fig. 11: Triángulo de Pascal, múltiplos de 11 de diferente color

8 El principio de la pichonera

Si k es un número entero positivo y $k+1$ o más objetos se colocan en k número de cajas, entonces hay al menos una caja que contiene dos o más de los objetos, como se representa en la Fig(12) [10].

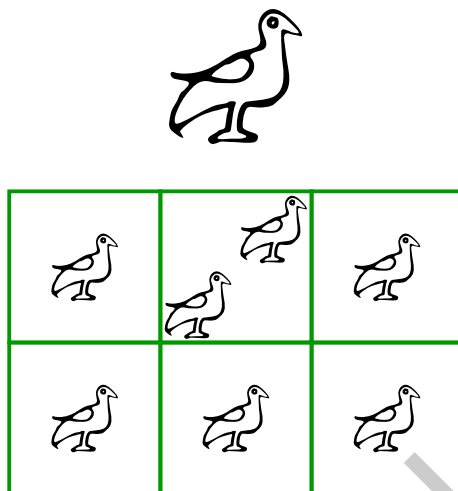


Fig. 12: Principio de la pichonera

Ejemplos:

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

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

- Los nombres de 10 personas son Alice, Bernard, Charles, mientras que sus apellidos son Lee, McDuff, 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.

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

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

$$27^4 + 27^3 + 27^2 + 27 = 551,880$$

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.

- 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.

9 Permutaciones, Combinaciones

9.1 Ejemplos de permutaciones

- Tres billetes de lotería se sacan de entre un total de 50. Si los billetes se han de distribuir a cada uno de tres empleados en el orden en que son sacados, el orden será importante. ¿Cuántos eventos simples están asociados con el experimento?

El número total de eventos es:

$$P(50, 3) = \frac{50!}{(50 - 3)!} = 117600$$

- Una máquina está compuesta de cinco partes que se pueden ensamblar en cualquier orden. Se ha de realizar una prueba para determinar el tiempo necesario para cada orden de ensamble. Si cada orden se ha de probar una vez, ¿cuántas pruebas deben efectuarse?

El número total de eventos es:

$$P(5, 5) = \frac{5!}{(5 - 5)!} = 120$$

- ¿Cuántas palabras distintas se pueden formar permutando las letras de la palabra auto?

$$P(4, 4) = \frac{4!}{(4 - 4)!} = 24$$

- Existen diez asistentes de profesor disponibles para calificar exámenes en un curso de cálculo en una gran universidad. El primer examen se compone de cuatro preguntas y el profesor desea seleccionar un asistente diferente para calificar cada pregunta (sólo un asistente por pregunta). ¿De cuántas maneras se pueden elegir los asistentes para calificar?

$$P(10, 4) = \frac{10!}{(10 - 4)!} = 5040$$

- ¿De cuántas maneras diferentes pueden ordenarse 10 bolas en una fila?

$$P(7, 7) = \frac{7!}{(7 - 7)!} = 13608$$

- Se requiere crear una contraseña que contenga 6 números del 1 al 10, ¿Cuántas posibles contraseñas se pueden formar, si los números no se repiten?

$$P(9, 6) = \frac{9!}{(9 - 6)!} = 93960$$

- Una persona tiene 4 pantalones, 3 camisas y 2 corbatas ¿De cuántas formas distintas puede vestirse esa persona sin repetir ningún día?
- Las formas en que se pueden ordenar los libros en un librero
- De 10 participantes de una carrera ¿Cuántas formas hay de entregar el premio?

9.2 Ejemplos de combinaciones

- 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 número de formas es:

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

- Cinco fabricantes producen cierto aparato electrónico, cuya calidad varía de un fabricante a otro. Si fuéramos a seleccionar tres fabricantes al azar, ¿cuál es la probabilidad de que la selección contenga exactamente dos de los tres mejores?

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

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

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

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

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

- El número de maneas en las cuales 3 cartas pueden escogerse de un total de 8 cartas diferentes es:

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

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

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

- Se tienen 2 billetes de \$500, 3 billetes de \$200 y 10 de \$50 si se eligen 4 billetes ¿Cuántas cantidades diferentes se tienen?
- En un juego de domino hay 28 fichas si cada jugador debe tener 7 ¿Cuántas combinaciones son posibles?
- Las maneras en que se puede ordenar la comida en un restaurante dependiendo el menú

9.3 Generalizando el problema de caminos posibles para el caso de un cuadrado

Se tiene un cuadrado como el que se muestra en la Fig(13) que se muestran a continuación, y se quiere llegar del punto A al punto B ¿cuántos posibles caminos hay para llegar?

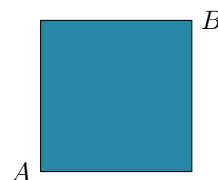
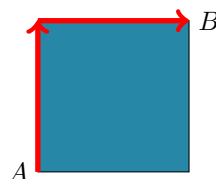


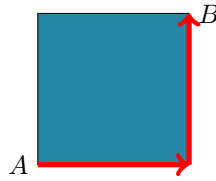
Fig. 13: Caminos posibles para un cuadrado de 1x1

Al recorrer el cuadrado se observan solo dos posibles caminos, los cuales se muestran a continuación, en las figuras:

- Primera opción: se recorre primero hacia arriba y luego hacia la derecha.



- Segunda opción: se recorre primero a la derecha y después hacia arriba.



Si se presenta el caso de una cuadrícula de 2x2 como la que se muestra en la Fig(14), ¿Cuántos caminos posibles hay del punto A al punto B?

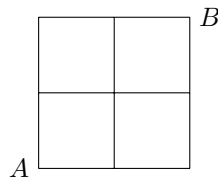
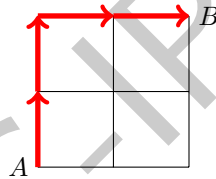


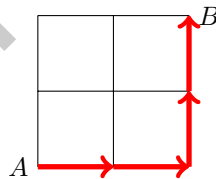
Fig. 14: Caminos posibles en un cuadrado de 2x2

Al hacer los diferentes recorridos se observa que hay 6 caminos posibles, sus secuencias se muestran a continuación, en las figuras:

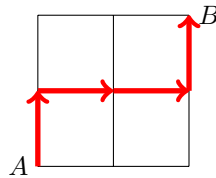
- Primer camino: se recorre dos hacia arriba y dos hacia la derecha.



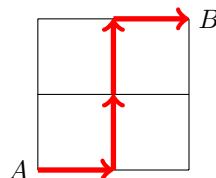
- Segundo camino: se recorre dos hacia la derecha y dos hacia arriba.



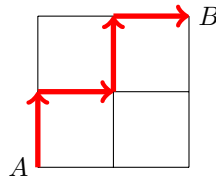
- Tercer camino: se recorre uno hacia arriba, dos hacia la derecha y por último uno hacia arriba.



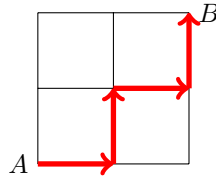
- Cuarto camino: se recorre uno hacia la derecha, dos hacia arriba y por último uno hacia la derecha.



- Quinto camino: se recorre uno hacia arriba, uno hacia la derecha, uno hacia arriba y por último uno hacia la derecha.



- Sexto camino: se recorre uno hacia la derecha, uno hacia arriba, uno hacia la derecha y por último uno hacia arriba.



¿Pero hay alguna forma de saber cuántos caminos hay sin tener que contarlos?

La respuesta es si, si se observa detenidamente en todos los recorridos que se hacen, siempre se recorren 2 casillas hacia la derecha y dos hacia arriba.

Los caminos del ejemplo anterior son Se puede representar con una combinación de n elementos donde se recorren k veces, se representa en la Eq(11)

$$C_{n,k} = \frac{n!}{k!(n-k)!} \quad (11)$$

Para este ejemplo queda:

$$C_{4,2} = \frac{4!}{2!(4-2)!} = \frac{24}{4} = 6$$

Generalizando para un cuadrado se utiliza la Eq(11), donde:

- n es el número total de casillas que se recorren en cada camino y
- k son las casillas que se recorren en cada dirección.

9.4 Generalizando el problema de caminos posibles para el caso de un rectángulo

Se tiene un rectángulo con $m \times n$ y se quiere saber ¿De cuántas formas distintas se puede recorrer del punto A al punto B?. En la Fig(9.4) se representa el caso de un rectángulo de 3×2 casillas como el que se muestra en la Fig(15).

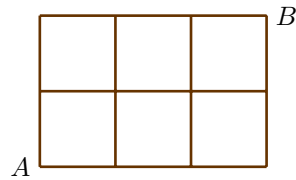


Fig. 15: Caminos posibles en un rectángulo de 3×2

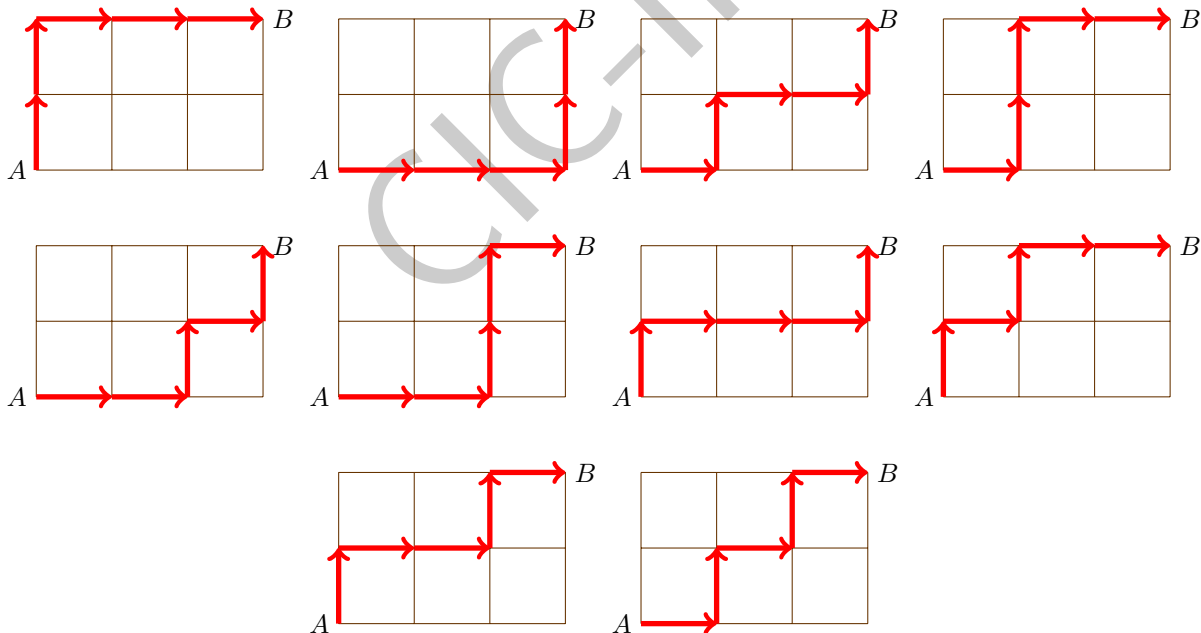
Se usa la ecuación de permutaciones con repetición de n elementos donde el primer elemento se repite a veces, el segundo b veces, el tercero c veces, ..., se muestra en la Eq(12).

$$PR_n^{a,b,c,\dots} = \frac{P_n}{a!b!c!\dots} \quad (12)$$

Sustituyendo los valores usados queda:

$$PR_5^{3,2} = \frac{P_5}{3!2!} = \frac{5!}{3!2!} = \frac{120}{12} = 10$$

Y el resultado se puede observar en las figuras que se muestran a continuación:



9.5 Generalizando el problema de caminos posibles para el caso de un cubo

Dado un cubo de $2 \times 2 \times 2$ casillas, se quiere saber la cantidad de recorridos posibles de un punto A a un punto B, retomando la Eq(12), y tomando en cuenta que al hacer los recorridos cualesquiera que sean se recorren 6 casillas a la vez (2 hacia la derecha, 2 hacia arriba y 2 hacia el fondo) queda así, una en la Fig(16) se muestra un cubo de $1 \times 1 \times 1$ en representación :

$$PR_6^{2,2,2} = \frac{P_6}{2!2!2!} = \frac{6!}{2!2!2!} = \frac{720}{8} = 90$$

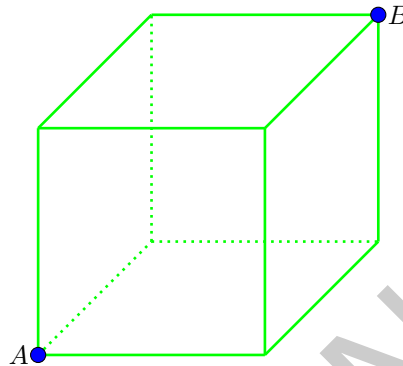


Fig. 16: Caminos posibles para un cubo

10 Reseñas de artículos

10.1 Primates Count

En este artículo se plantea una pregunta importante ¿Pueden contar los animales?, Se han hecho varios experimentos a lo largo del tiempo con varios animales como son: ardillas, ratas, chimpancés, algunos insectos y se observa que algunos de estos animales con ayuda de entrenamiento son capaces de distinguir números, otros son capaces de contar, pueden reconocer y reproducir una secuencias acústicas, sin embargo por su falta de habla y la escritura de símbolos las personas se niegan a aceptarlos como matemáticos.

El ser humano también ha necesitado de cierto entrenamiento para saber contar, por lo que se puede deducir que los animales saben contar a su manera y para las actividades que ellos necesitan en su vida cotidiana.

10.2 Cicada-Generated Prime Numbers

Las cigarras son animales que viven debajo de la tierra, se alimentan de el jugo de las raíces de las plantas, después emerge y mueren rápidamente, estas criaturas exhiben un comportamiento sorprendente: Su aparición se sincroniza con periodos de años en general, los números primos 13 y 17.

10.3 Ant Odometer

Las hormigas son insectos capaces de recorrer grandes distancias y encontrar de regreso el camino a casa, según experimentos se dedujo que las hormigas tienen un tipo de contador de pasos para lograr este objetivo. A una hormiga se le pusieron extensiones en algunas de sus patas, por lo que al salir en búsqueda de comida y tratar de regresar a casa no lo lograba, debido a que la diferencia en sus patas impedían que su sentido de orientación y el contar sus pasos fallara.

10.4 Quipu

Debido a que en la antigüedad no había tanto conocimiento del tiempo como el que hay ahora se recurrían a formas distintas para marcar el paso del tiempo, una de ellas es el caso de los quipus que son bancos de memoria hechos de cadenas y de nudos, para almacenar números y fueron originados por los antiguos Incas. Pueden almacenar mapas de construcción, patrones de danza etc., Todo dependía de la complejidad del quipu, cada tipo de nudo, posiciones, color y otros factores determinaban números diferentes que representaban objetos del mundo real.

11 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 característica. Así pues, podemos evidenciar comportamientos, observar el grado de homogeneidad, acuerdo o concisión entre los valores de todas las partes que componen la población o la muestra, o, en contraposición, poder observar el grado de variabilidad, y por ende, la dispersión de todos los valores que toman las partes, también es posible no evidenciar ninguna tendencia y obtener que cada miembro de la población toma por su lado y adquiere un valor de la característica aleatoriamente sin mostrar ninguna preferencia o tendencia, entre otras cosas. En el eje vertical se representan las frecuencias, es decir, la cantidad de población o la muestra, según sea el caso, que se ubica en un determinado valor o sub-rango de valores de la característica que toma la característica de interés, evidentemente, cuando este espectro de valores es infinito o muy grande el mismo es reducido a sólo una parte que muestre la tendencia o comportamiento de la población, en otras ocasiones este espectro es extendido para mostrar el alejamiento o ubicación de la población o la muestra analizada respecto de un valor de interés. En general se utilizan para relacionar variables cuantitativas continuas, pero también se lo suele usar para variables cuantitativas discretas, en cuyo caso es común llamarlo diagrama de frecuencias y sus barras están separadas, esto es porque en el "x" ya no se representa un espectro continuo de valores, sino valores cuantitativos específicos como ocurre en un diagrama de barras cuando la característica que se representa es cualitativa o categórica. Su utilidad se hace más evidente cuando se cuenta con un gran número de datos cuantitativos y que se han agrupado en intervalos de clase, en la Fig(17) se muestra un ejemplo de histograma.

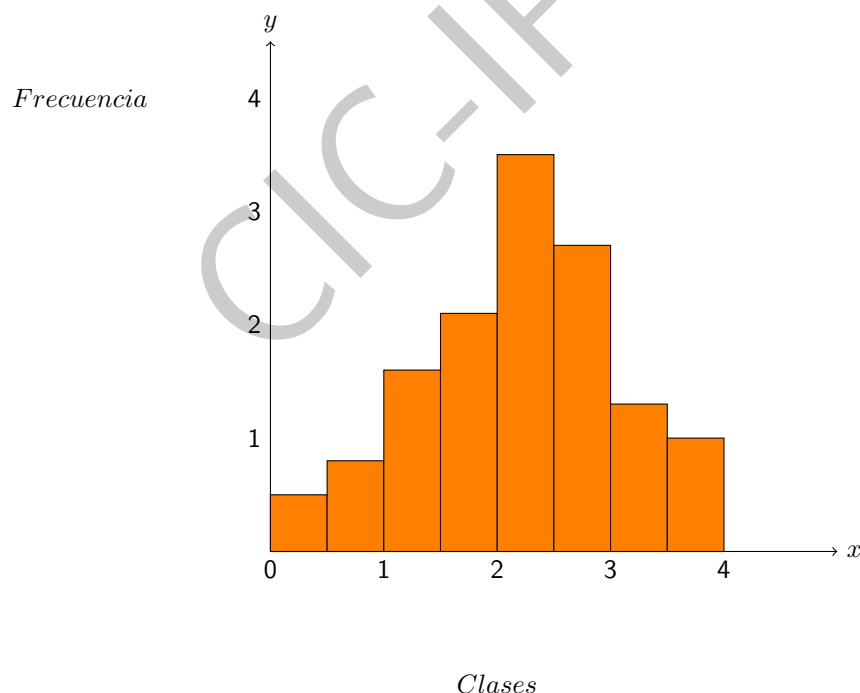


Fig. 17: Ejemplo de Histograma

11.1 Ejemplos de histogramas

En la Fig(18) se muestra un histograma, que se deriva de las muestras de peso tomadas a un grupo de personas.

- El peso de un grupo de personas

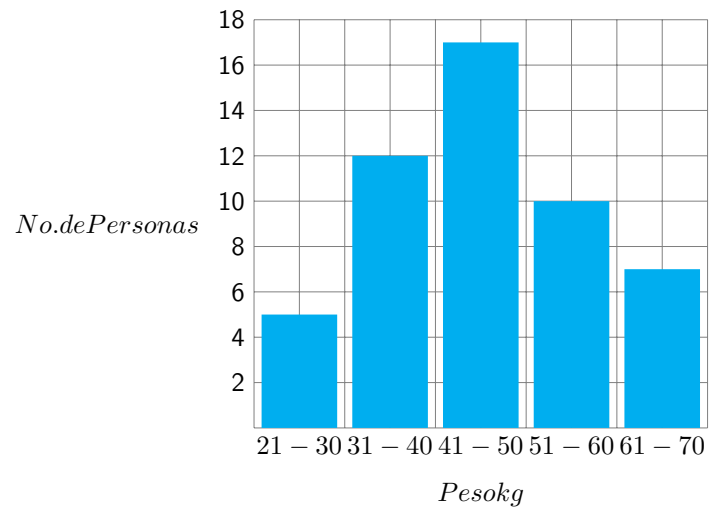


Fig. 18: Histograma del peso de un grupo de personas

12 Compuertas NAND y XOR con electrónica

12.1 NAND = $(AB)'$

En la Fig(19), se muestra la representación de la compuerta NAND mediante circuitos.

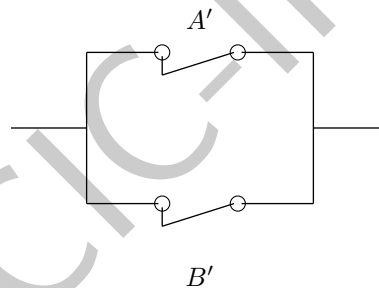


Fig. 19: Compuerta NAND con circuitos

12.2 XOR = $AB' + A'B$

En la Fig(20), se muestra la representación de la compuerta XOR mediante circuitos.

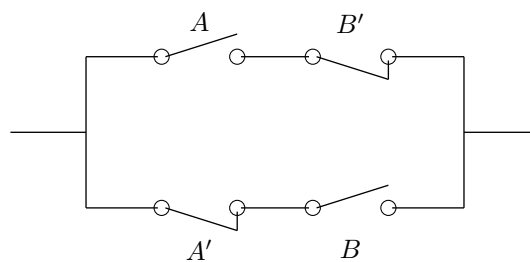


Fig. 20: Compuerta XOR con circuitos

13 Fórmula de Stirling

En matemáticas, la fórmula de Stirling es una aproximación para factoriales grandes. Lleva el nombre en honor al matemático escocés del siglo XVIII James Stirling. La aproximación se expresa como se muestra en la Eq(13).

$$\ln(n!) \approx n \ln(n) - n \quad (13)$$

Definición formal.

La fórmula de Stirling está dada por:

$$\lim_{n \rightarrow \infty} \frac{n!}{\sqrt{2\pi * n} \left(\frac{n}{e}\right)^n} = 1 \quad (14)$$

que se reescribe frecuentemente como:

$$n! \approx \sqrt{2\pi * n} \left(\frac{n}{e}\right)^n \quad (15)$$

más exactamente la fórmula es como sigue:

$$n! = \sqrt{2\pi * n} \left(\frac{n}{e}\right)^n e^{\frac{1}{12n} - \frac{1}{360n^3} + \frac{1}{1260n^5} - \frac{1}{1680n^7} + \dots} \quad (16)$$

donde el último término del producto tiene a 1 cuando n tiene a infinito.

14 Telómeros

Los telómeros son los extremos de los cromosomas. Son regiones de ADN no codificante, altamente repetitivas, cuya función principal es la estabilidad estructural de los cromosomas en las células eucariotas (células con un núcleo celular delimitado dentro de una doble capa lipídica, la envoltura nuclear, la cual es porosa y contiene su material hereditario, fundamentalmente su información genética), la división celular y el tiempo de vida de las estirpes celulares. Además están involucradas en enfermedades tan importantes como el cáncer.

Los organismos procariotas tienen cromosomas circulares que no poseen telómeros. Algunos procariotas poseen cromosomas lineales con secuencias teloméricas, cuya secuencia es diferente a la de eucariotas.

Los telómeros fueron descubiertos por Hermann Joseph Muller durante la década de los años 30 del siglo XX, que junto a Barbara McClintock recibieron el Premio Nobel. Desde entonces, se ha avanzado mucho en el conocimiento de los telómeros, gracias a las técnicas de genética molecular. Proponían que los telómeros, situados en los extremos de los cromosomas, tenían la función de prevenir que éstos se fusionaran al ponerse en contacto por sus extremos, lo que produciría consecuencias desastrosas para las células.

Los científicos Elizabeth H. Blackburn, Carol W. Greider y Jack W. Szostak son reconocidos con el Premio Nobel de medicina en 2009 por la descripción molecular de los telómeros, la demostración de su conservación evolutiva y el descubrimiento de la telomerasa, enzima central de la maquinaria celular para la síntesis del telómero, logrando un muy consistente modelo que explica el 'problema de la terminación de la replicación' (end-replication problem) y el mecanismo molecular de protección de los extremos cromosomales.

15 Oveja Dolly

La oveja Dolly, el primer mamífero clonado en el mundo, fue reprimida el 14 de febrero de este año. Ella sufría de un virus que le causó un tumor en el pulmón. Esto ha dado lugar a una nueva ronda de debate sobre la clonación, en particular sobre los problemas de envejecimiento de los animales clonados. Los científicos se han preocupado desde hace tiempo que los animales clonados podrían heredar su edad de su donante de la célula, siendo de este modo que nace viejo y morirá prematuramente. Los telómeros son las unidades que se encuentran al final de todos los cromosomas.

En el proceso de transferencia nuclear, una célula somática, con la longitud del telómero acortado, se transfiere a un ovocito enucleado y se activa para iniciar el desarrollo del embrión. Una pregunta que viene inmediatamente a la cabeza, ¿están los telómeros acortados de las células del donante restaurados en toda su longitud en los animales producidos por transferencia nuclear?. Era una preocupación de que si los telómeros acortados de hecho fueron heredados por los animales clonados, que estos animales pueden heredar el acortamiento de la vida del donante adulto en lugar de la de sus controles emparejados por edad producidos mediante reproducción tradicional. La muerte de Dolly, si es natural y prematura, podría ser el primer indicio de que estas preocupaciones eran válidas.



16 Antinomias y paradojas

Definición antinomia.

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

Definición de paradoja

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

Paradojas sobre teoría de conjuntos

16.1 Paradoja de Cantor: el conjunto de todos los conjuntos.

Sea C el conjunto de todos los conjuntos. Entonces todo subconjunto de C es así mismo un elemento de C ; luego, el conjunto potencia de C es un subconjunto de C ; pero esto implica que la cardinalidad del conjunto potencia es menor o igual a la cardinalidad de C . Pero entonces, según el teorema de Cantor, la cardinalidad de C debe ser menor a la cardinalidad del conjunto potencia. Así pues, el concepto de conjunto de todos los conjuntos lleva a una contradicción. George Cantor.

16.2 Paradoja de Russell.

Sea Z el conjunto de todos los conjuntos que no son elementos de sí mismos. Se pregunta ¿ Z es o no elemento de sí mismo? Si Z no pertenece a Z , entonces, por la definición de Z , Z pertenece a sí mismo. Pero si Z pertenece a Z , entonces por la definición de Z , Z no pertenece a sí mismo. En cualquiera de los dos casos hay contradicción.

Esta paradoja es análoga a la paradoja del barbero: En una aldea hay un barbero que afeita solamente a los hombres que no se afeitan ellos mismos. Se pregunta ¿Al barbero quién lo afeita?. Bertrand Russell.

16.3 Paradoja de Burali-Forti: Conjunto de todos los números ordinales.

Sea D el conjunto de todos los números ordinales. Por un teorema anterior, D es un conjunto bien ordenado; sea $A = \text{ord}(D)$. Considérese ahora $s(A)$ el conjunto de todos los números ordinales menores que A . Obsérvese que

Puesto que $S(A)$ consiste en todos los elementos de D que son anteriores a A , $S(A)$ es una sección inicial de D . Por un teorema previo, $A = \text{ord}(s(A))$; por tanto, $\text{ord}(s(a)) = A = \text{ord } D$

Por consiguiente D es isomorfo a una de sus secciones iniciales. Así pues el concepto de conjunto de todos los números ordinales lleva a una contradicción.

16.4 Conjunto de todos los números cardinales.

Sea A el conjunto de todos los números cardinales. Entonces, para cada cardinal a que pertenece a A , hay un conjunto A_a tal que a es igual a la cardinalidad de A_a .

Considérese el conjunto potencia de A . Nótese que el conjunto potencia de A es menor o igual a A , y en particular, la cardinalidad del conjunto potencia de A es menor o igual a la cardinalidad de A .

Pero por el teorema de Cantor, el concepto de conjunto de todos los números cardinales es contradictorio.

16.5 Paradoja del falso positivo

La paradoja del falso positivo es un resultado estadístico donde las pruebas con falsos positivos son más probables que resultados con positivos verdaderos, esto se produce cuando la población en general tiene una baja incidencia de una condición y la tasa de incidencia es inferior a la tasa de falsos positivos. La probabilidad de un resultado positivo de la

prueba se determina no sólo por la precisión de la prueba, sino también por las características de la población muestreada. Cuando la incidencia, la proporción de aquellos que tienen una condición dada, es más baja que la tasa de falso positivo de la prueba, incluso pruebas de que tienen una muy baja probabilidad de dar un falso positivo en un caso individual dará más falsos que verdaderos positivos en general. Por lo tanto, en una sociedad con muy pocas personas infectadas- menos proporcionalmente que la prueba da falsos positivos- será más que un resultado positivo para una enfermedad de forma incorrecta y no lo tiene que los que dan positivo con precisión y hacer. La paradoja ha sorprendido a muchos.

Es especialmente contrario a la intuición la interpretación de un resultado positivo en una prueba con una baja incidencia de la población después de haber tratado con resultados positivos extraídos de una población de alta incidencia. Si el ritmo de falsos positivos de la prueba es mayor que la proporción de la nueva población con la condición, a continuación, un administrador de la prueba cuya experiencia se ha elaborado a partir de las pruebas en una población de alta incidencia puede llegar a la conclusión de la experiencia de que un resultado positivo por lo general indica un tema positivo, cuando en realidad un falso positivo es mucho más probable que se hayan producido.

No ajustarse a la escasez de la condición en la nueva población, y la conclusión de que un resultado positivo de la prueba indica probablemente un tema positivo, a pesar de que la incidencia de la población está por debajo de la tasa de falsos positivos es una "tasa base de falacia".

Supongamos que un grupo de personas de las que el 1% sufre una cierta enfermedad, y el resto está bien:

$$P(enfermo) = 1\% = 0.01 \text{ y } P(sano) = 99\% = 0.99$$

Supongamos que aplicando una prueba a una persona que no tiene la enfermedad, hay una posibilidad del 1% de conseguir un falso positivo, esto es:

$$P(positivo | sano) = 1\% = 0.01 \text{ y } P(negativo | sano) = 99\% = 0.99$$

Finalmente, supongamos que aplicando la prueba a una persona que tiene la enfermedad, hay una posibilidad del 1% de un falso negativo, esto es:

$$P(negativo | enfermo) = 1\% = 0.01 \text{ y } P(positivo | enfermo) = 99\% = 0.99$$

Ahora, uno puede calcular lo siguiente:

La fracción de individuos en el grupo que están sanos y dan negativo:

$$P(sano \cap negativo) = P(sano) * P(negativo | sano) = 0.99 * 0.99 = 0.9801$$

La fracción de individuos en el grupo que están enfermos y dan positivo:

$$P(enfermo \cap positivo) = P(enfermo) * P(positivo | enfermo) = 0.01 * 0.99 = 0.0099$$

La fracción de individuos en el grupo que dan falso positivo:

$$P(sano \cap positivo) = P(sano) * P(positivo | sano) = 0.99 * 0.01 = 0.0099$$

La fracción de individuos en el grupo que dan falso negativo:

$$P(enfermo \cap negativo) = P(enfermo) * P(negativo | enfermo) = 0.01 * 0.01 = 0.0001$$

Además, la fracción de individuos en el grupo que dan positivo:

$$P(positivo) = P(sano \cap positivo) + P(enfermo \cap positivo) = 0.0099 + 0.0099 = 0.0198$$

Finalmente, la probabilidad de que un individuo realmente tenga la enfermedad, dado un resultado de la prueba positivo:

$$P(enfermo | positivo) = \frac{P(enfermo \cap positivo)}{P(positivo)} = \frac{0.0099}{0.0198} = 0.5$$

En este ejemplo, debería ser fácil ver la diferencia entre las probabilidades condicionadas $P(positivo | enfermo)$ (que es del 99%) y $P(enfermo | positivo)$ (que es del 50%): la primera es la probabilidad de que un individuo enfermo dé positivo en la prueba; la segunda es la probabilidad de que un individuo que da positivo en la prueba tenga realmente la enfermedad. Con los números escogidos aquí, este último resultado probablemente sería considerado inaceptable: la mitad de la gente que da positivo en realidad está sana.

16.6 Paradoja Borel-Kolmogorov

En teoría de probabilidad, la paradoja de Borel-Kolmogorov (algunas veces conocida como paradoja Borel's) es una paradoja relacionada a la probabilidad condicional con respecto a un evento de probabilidad cero (también conocida como conjunto nulo).

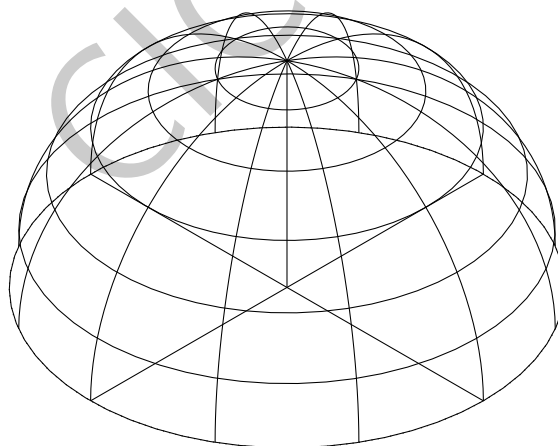
Supongamos que una variable aleatoria tiene una distribución uniforme en una esfera unitaria. ¿Cuál es la distribución condicional de un gran círculo? Porque de la simetría de una esfera, uno podría esperar que la distribución es uniforme e independiente de la selección de coordenadas. Sin embargo, dos análisis dan resultados contradictorios. Primero, notar que escogiendo un punto uniformemente en la esfera es equivalente a escoger la longitud λ uniformemente desde $[-\pi, \pi]$ y escogiendo la latitud ϕ desde $[-\frac{\pi}{2}, \frac{\pi}{2}]$ con densidad $\frac{1}{2} \cos \phi$. Entonces nosotros podemos observar dos diferentes grandes círculos.

Sí las coordenadas se eligen de modo que el gran círculo es un ecuador ($latitud \phi = 0$), la densidad condicional para una longitud α definida en el intervalo $[-\pi, \pi]$ es:

$$f(\lambda | \phi = 0) = \frac{1}{2\pi}$$

Sí el gran círculo es un meridiano con $\lambda = 0$ la densidad condicional para ϕ en el intervalo $[-\frac{\pi}{2}, \frac{\pi}{2}]$ es:

$$f(\phi | \lambda = 0) = \frac{1}{2} \cos \phi$$



¿Cuál es la probabilidad de que al escoger dos puntos al azar en la superficie de una esfera la distancia entre ellos sea menor que 10 min? Por la simetría de la esfera, podemos suponer que el primer punto es conocido. Así, una manera de responder a la pregunta es calculando la proporción de la superficie de la esfera que se halla a una distancia de 10 min del punto dado, esto es 2.1×10^{-6} .

Después de fijar el primer punto, también podemos suponer que conocemos el círculo máximo que conecta los dos puntos, ya que las posibilidades de que el segundo punto caiga en uno u otro círculo máximo en torno al primer punto son las mismas. Hay 360 grados, esto es, 2160 arcos de 10 min, en este círculo máximo. Solamente los puntos en los dos arcos vecinos al primer punto se hallan a una distancia de a lo más 10 min, de manera que la probabilidad que buscamos es de $2/2160$, o bien, 9.3×10^{-4} .

Borel dio la siguiente explicación sobre esta paradoja en su libro de texto de probabilidad publicado en 1909. Él afirmó que el primer método utilizado por Bertrand, en donde hace la suposición de que áreas iguales tienen probabilidades

iguales de contener al segundo punto, es correcto. Su segundo método, basado en el supuesto de que arcos iguales en el círculo máximo tienen probabilidades iguales de contener al segundo punto, es incorrecto. Escribiendo M y M' para los puntos que serán elegidos al azar en la esfera, Borel explicó el error de Bertrand de la siguiente manera:

...El error empieza cuando, después de fijar el punto M y el círculo máximo, uno supone que la probabilidad de que M' se halle en un arco dado del círculo máximo es proporcional a la longitud de ese arco. Si el arco no tiene grosor, entonces, hablando en el sentido estricto, habremos de asignar el valor cero a la probabilidad de que M y M' se hallen en el mismo círculo. Con el fin de evitar este factor de cero, lo cual hace imposible cualquier cálculo, debemos de considerar franjas delgadas de círculos máximos, los cuales pasan todos por M , donde las franjas son de tal manera que éstas se hacen cada vez más delgadas al acercarse a los polos -de los cuales M es uno de ellos- de modo que las franjas no se traslapen una sobre la otra.

De la solución a esta paradoja surge el siguiente principio: Es posible que necesitemos considerar un evento real observado de probabilidad distinta de cero como uno que es representado de una manera idealizada por un evento de probabilidad cero. La confusión en torno a las paradojas significó una fuente más de insatisfacción con la teoría clásica, basada en eventos equiprobables.

CIC-IPN

17 Conjuntos

17.1 Leyes de De Morgan

$$1. (\overline{A \cup B}) = \overline{A} \cap \overline{B}$$

$$2. (\overline{A \cap B}) = \overline{A} \cup \overline{B}$$

Demostración:

Sea x un elemento arbitrario del conjunto universal U .

$$1. (\overline{A \cup B}) \iff x \notin (A \cup B) \quad \{\text{Definición de complementario}\}$$

$$\iff \neg[x \in (A \cup B)] \quad \{\text{Negación}\}$$

$$\iff \neg[(x \in A) \vee (x \in B)] \quad \{\text{Definición de unión}\}$$

$$\iff \neg(x \in A) \wedge \neg(x \in B) \quad \{\text{De Morgan para } \vee\}$$

$$\iff (x \notin A) \wedge (x \notin B) \quad \{\text{Negación}\}$$

$$\iff (x \in \overline{A}) \wedge (x \in \overline{B}) \quad \{\text{Definición de complemento}\}$$

$$\iff x \in (\overline{A} \cap \overline{B}) \quad \{\text{Definición de intersección}\}$$

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

$$\forall x [x \in (\overline{A \cap B}) \iff x \in (\overline{A} \cap \overline{B})]$$

luego,

$$\overline{A \cup B} = \overline{A} \cap \overline{B}$$

Sea x un elemento arbitrario del conjunto universal U .

$$2. (\overline{A \cap B}) \iff x \notin (A \cap B) \quad \{\text{Definición de complementario}\}$$

$$\iff \neg[x \in (A \cap B)] \quad \{\text{Negación}\}$$

$$\iff \neg[(x \in A) \wedge (x \in B)] \quad \{\text{Definición de intersección}\}$$

$$\iff \neg(x \in A) \vee \neg(x \in B) \quad \{\text{De Morgan para } \wedge\}$$

$$\iff (x \notin A) \vee (x \notin B) \quad \{\text{Negación}\}$$

$$\iff (x \in \overline{A}) \vee (x \in \overline{B}) \quad \{\text{Definición de complemento}\}$$

$$\iff x \in (\overline{A} \cup \overline{B}) \quad \{\text{Definición de unión}\}$$

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

$$\forall x [x \in (\overline{A \cap B}) \iff x \in (\overline{A} \cup \overline{B})]$$

luego,

$$\overline{A \cap B} = \overline{A} \cup \overline{B}$$

CIC-IPN

17.2 Desarrollar $P(A \cup B \cup C)$

En la Fig(21) se observa el diagrama de Venn de 3 conjuntos.

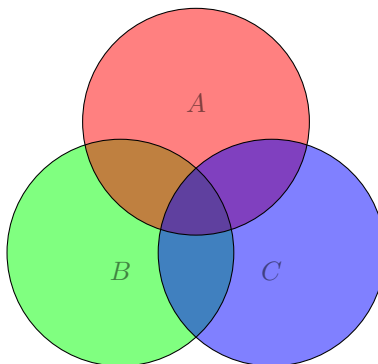


Fig. 21: Diagrama de Venn de 3 conjuntos

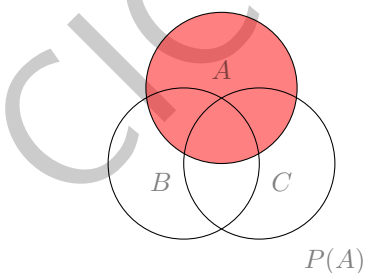
$$P(A \cup B \cup C) = P(A) + P(B) + P(C)$$

$$-P(A \cap B) - P(B \cap C) - P(C \cap A)$$

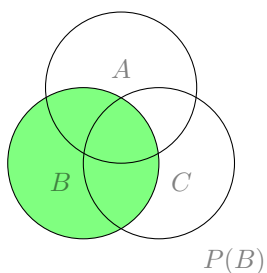
$$+P(A \cap B \cap C)$$

Debido a los términos tenemos:

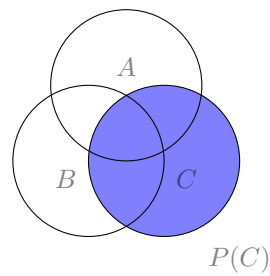
La probabilidad de A la podemos observar de la siguiente manera, en las figuras siguientes se va observando paso a paso el desarrollo de la expresión:



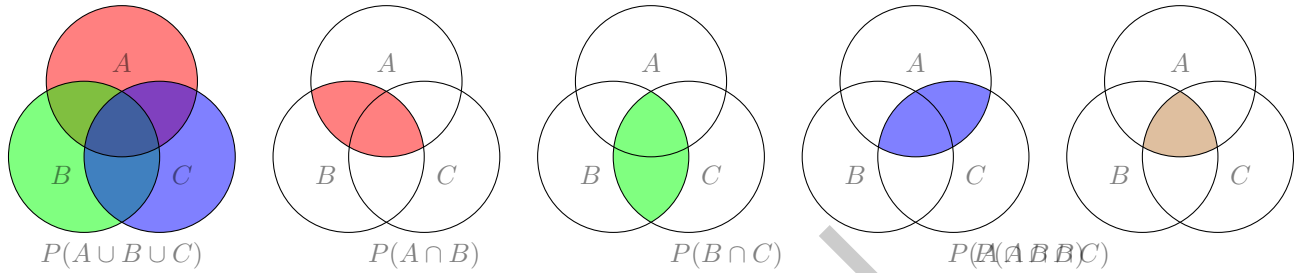
La probabilidad de B de la siguiente manera:



La probabilidad de C de la siguiente manera:



Al hacer la suma de probabilidades, podemos observar que se suman dos veces $P(A \cap B)$, $P(A \cap C)$ y $P(B \cap C)$, motivo por el cual se restan, pero debemos observar que también se elimina $P(A \cap B \cap C)$, para obtener la unión de los 3 conjuntos ahora debemos sumar $P(A \cap B \cap C)$ y con esto ya queda completo $P(A \cup B \cup C)$:



17.3 Desarrollar $P(A \cup B \cup C \cup D)$

En la Fig(22) se muestra el diagrama de Venn de 4 conjuntos.

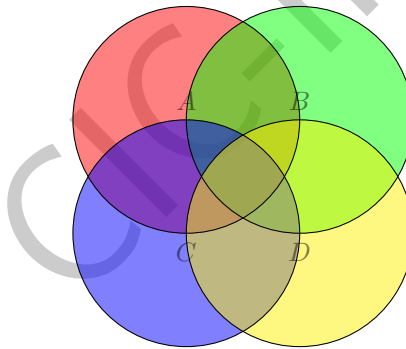


Fig. 22: Diagrama de Venn de 4 conjuntos

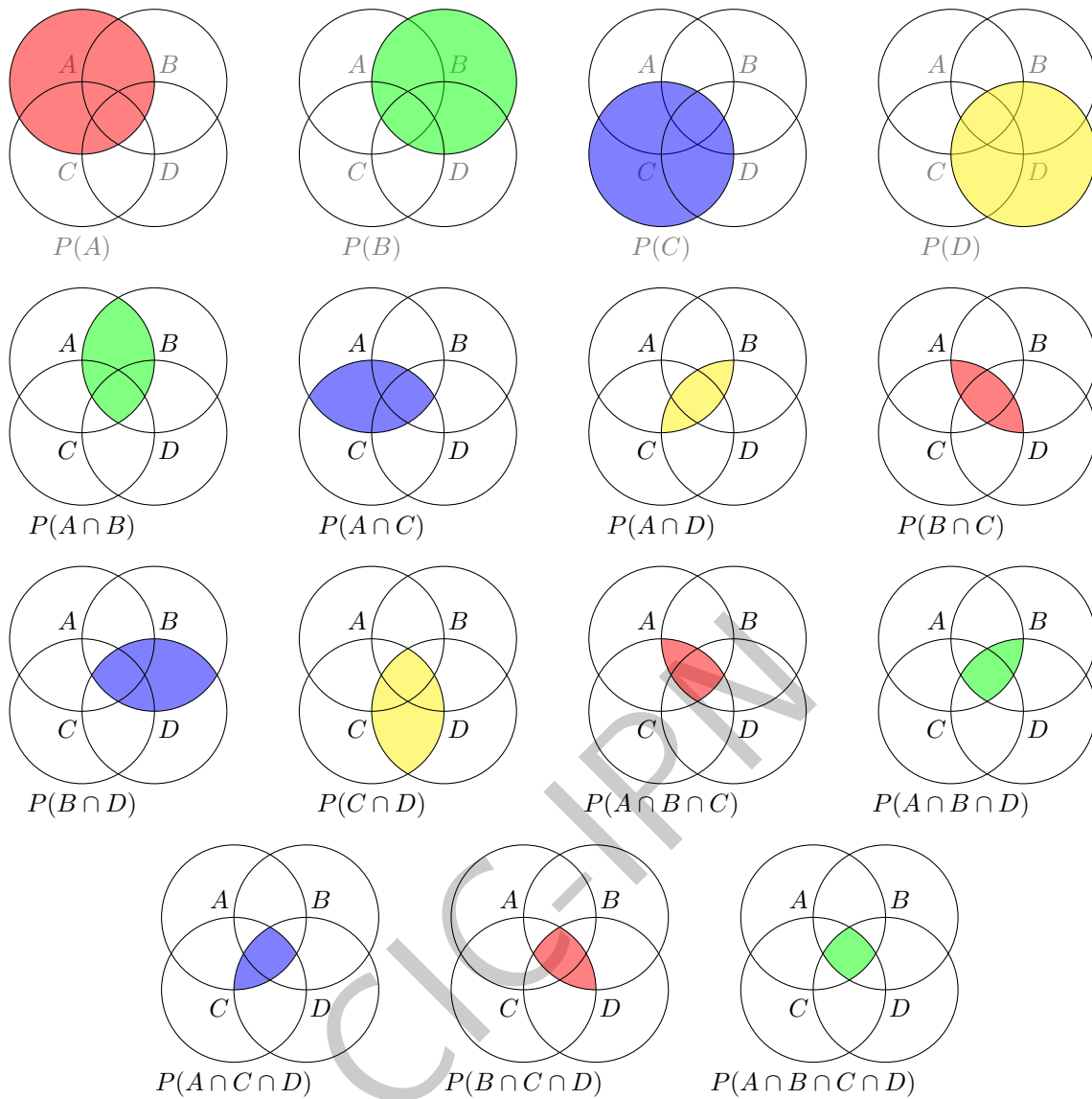
$$P(A \cup B \cup C \cup D) = P(A) + P(B) + P(C) + P(D)$$

$$-P(A \cap B) - P(A \cap C) - P(A \cap D) - P(B \cap C) - P(B \cap D) - P(C \cap D)$$

$$+P(A \cap B \cap C) + P(A \cap B \cap D) - P(A \cap C \cap D) + P(B \cap C \cap D)$$

$$-P(A \cap B \cap C \cap D)$$

De manera mas gráfica se puede observar como es que se va sumando las probabilidades y por consiguiente que es lo que se debe restar para que obtengamos solo $P(A \cup B \cup C \cup D)$:



La probabilidad de la unión de n sucesos queda como se muestra en a continuación.

$$P\left(\bigcup_{i=1}^n S_i\right) = \sum_{i=1}^n P(S_i) - \sum_{i,j=1;i < j}^n P(S_i \cap S_j) + \sum_{i,j,k=1;i < j < k}^n P(S_i \cap S_j \cap S_k) + \dots + (-1)^{n+1} P(S_1 \cap S_2 \cap \dots \cap S_n).$$

17.4 Significado de $P(A|B)$

Sean A y B dos eventos y supongamos que B tiene probabilidad estrictamente positiva. La probabilidad condicional del evento A dado el evento B se denota por el símbolo $P(A|B)$ y se define por la Eq(17) [7]

$$P(A | B) = \frac{P(A \cap B)}{P(B)} \quad (17)$$

El término $P(A|B)$ se lee "probabilidad de A dado B" y es claro a partir de la definición que es necesaria la condición $P(B)>0$ para que el cociente esté bien definido. Cuando $P(B)=0$ no existe una definición establecida para $P(A|B)$. En ocasiones se usa la expresión $P_B(A)$ para denotar a esta probabilidad. En la Eq(17), el evento B representa un evento que ha ocurrido y la probabilidad condicional $P(A | B)$ es la probabilidad de A modificada con la información adicional de que B ha ocurrido.

CIC-IPN

18 Falacia del apostador

La falacia del apostador, también conocida como la falacia de Monte Carlo, es la creencia errónea de que, si algo pasa más frecuente de lo normal durante cierto periodo, esto pasará menos frecuente en el futuro. Se cree erróneamente que los sucesos pasados afectan a los futuros en lo relativo a actividades aleatorias, como en muchos juegos de azar.

El uso de el termino de la falacia de Monte Carlo tiene origen en el más famosos ejemplo de este fenómeno, el cual ocurrió en el Casino Monte Carlo en 1913.

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.

Algunos ejemplos de falacia del apostador:

- Tirar una moneda y que caiga cierta cara puesto que ya vana vairas veces que sale la cara contraria
- Que salgan ciertas cartas en la mesa de Poker Texas puesto que van varios turnos que no han salido al juego
- La probabilidad de tener niño o niña puesto que se han tenido dos niños o niñas anteriormente (excluyendo influencias genéticas)
- Pensar que se tiene mayor probabilidad de ganar en la lotería jugando siempre el mismo número, esta la misma que jugando un número diferente cada vez: las probabilidades sólo dependen de los números en juego.



19 Teorema de Bayes, probabilidad bayesiana

19.1 Teorema de Bayes

El teorema de Bayes, en la teoría de la probabilidad, es una proposición planteada por el filósofo inglés Thomas Bayes en 1763,² que expresa la probabilidad condicional de un evento aleatorio A dado B en términos de la distribución de probabilidad condicional del evento B dado A y la distribución de probabilidad marginal de sólo A .

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

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

$$P(A_i | B) = \frac{P(B | A_i)P(A_i)}{P(B)} \quad (18)$$

donde:

$P(A_i)$ son las probabilidades a priori

$P(B | A_i)$ es la probabilidad de B en la hipótesis A_i

$P(A_i | B)$ son las probabilidades a posteriori

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

$$P(A_i | B) = \frac{P(B | A_i)P(A_i)}{\sum_{k=1}^n P(B | A_k)P(A_k)} \quad (19)$$

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

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

19.2 Probabilidad bayesiana

La probabilidad bayesiana es una de las diferentes interpretaciones del concepto de probabilidad. La interpretación bayesiana de la probabilidad puede ser vista como una extensión de la lógica proposicional que permite razonar con hipótesis, es decir, las proposiciones cuya verdad o falsedad son inciertas.

La probabilidad bayesiana pertenece a la categoría de las probabilidades probatorias; para evaluar la probabilidad de una hipótesis, la probabilista bayesiana especifica alguna probabilidad a priori, que se actualiza a continuación, a la luz de nuevos y relevantes datos (en pruebas). La interpretación bayesiana proporciona un conjunto estándar de los procedimientos y las fórmulas para realizar este cálculo.

En contraste con la interpretación de la probabilidad como la "frecuencia" o "propensión" de algún fenómeno, la probabilidad bayesiana es una cantidad que se asigna para el propósito de representar un estado de conocimiento, o un estado de creencia. En la vista bayesiana, una probabilidad se asigna a una hipótesis, mientras que bajo el punto de vista frecuentista, una hipótesis es típicamente probada sin ser asignada una probabilidad.

El término "bayesiano" se refiere al matemático del siglo 18 y teólogo Thomas Bayes, que proporcionó el primer tratamiento matemático de un problema no trivial de la inferencia bayesiana. El matemático Pierre-Simon Laplace fue pionero y popularizó lo que ahora se llama probabilidad bayesiana.

En términos generales, hay dos puntos de vista sobre la probabilidad bayesiana que interpretan el concepto de probabilidad de diferentes maneras. Según el punto de vista objetivista, las reglas de la estadística bayesiana pueden justificarse por exigencias de la racionalidad y la coherencia, y la interpretan como una extensión de la lógica. Según la visión subjetivista, cuantifica la probabilidad de una "opinión personal". Muchos métodos modernos de aprendizaje automático se basan en los principios bayesianos objetivistas.

CIC-IPN

19.3 Problemas teorema de Bayes

- En una fábrica hay dos máquinas. La máquina 1 realiza el 60% de la producción total y la máquina 2 el 40 %. De su producción total, la máquina 1 produce 3% de material defectuoso, la 2 el 5%. El asunto es que se ha encontrado un material defectuoso, cuál es la probabilidad de que este material defectuoso provenga de la máquina 2?

Solución: Sea M_1 el evento La máquina 1 produjo el material escogido, M_2 en evento La máquina 2 produjo el material escogido y finalmente sea D el evento El material escogido es defectuoso. El problema es encontrar $P(M_2|D)$ y observamos que la información que tenemos es $P(D|M_2)$. Por el teorema de Bayes tenemos entonces que

$$\begin{aligned} P(M_2|D) &= \frac{P(D|M_2)P(M_2)}{P(D|M_1)P(M_1) + P(D|M_2)P(M_2)} \\ &= \frac{\frac{5}{100} * \frac{40}{100}}{\frac{3}{100} * \frac{60}{100} + \frac{5}{100} * \frac{40}{100}} \\ &= \frac{10}{19} \end{aligned}$$

- En un laboratorio se descubrió una prueba para detectar cierta enfermedad y sobre la eficacia de dicha prueba se conoce lo siguiente: si se denota por E el evento de que un paciente tenga la enfermedad y por N el evento de que la prueba resulte negativa, entonces se sabe que $P(N^c|E) = 0.95$, $P(N|E^c) = 0.96$ y $P(E) = 0.01$. Con esta información uno podría pensar que la prueba es muy buena, sin embargo calcularemos las probabilidades $P(E|N)$ y $P(E|N^c)$ usando el teorema de Bayes.

$$\begin{aligned} P(E|N) &= \frac{P(N|E)P(E)}{P(N|E)P(E) + P(N|E^c)P(E^c)} \\ &= \frac{0.05 * 0.01}{0.05 * 0.01 + 0.96 * 0.99} \\ &= 0.000526 \end{aligned}$$

- Acabamos de recibir la noticia de que hemos dado positivo en una enfermedad. La prueba es 90% confiable. Queremos saber la probabilidad de que tenemos la enfermedad después nos enteramos de que el resultado de la prueba es positiva. Algunos lectores pueden sentir que esta probabilidad debe ser de aproximadamente 0,9. Sin embargo, este sentimiento ignora por completo la pequeña probabilidad de 0.0001 que tenía la enfermedad antes de tomar la prueba. Dejemos B_1 el evento de que usted tiene la enfermedad, y sea B_2 el evento de que no tiene la enfermedad. La eventos B_1 y B_2 forman una partición. Además, Sea A el caso de que la respuesta a la prueba es positiva. El evento es una información vamos a aprender que nos dice algo acerca de los elementos de separación. Entonces, por el teorema de Bayes.

$$\begin{aligned} P(B_1|A) &= \frac{P(A|B_1)P(B_1)}{P(A|B_1)P(B_1) + P(A|B_2)P(B_2)} \\ &= \frac{(0.9)(0.0001)}{(0.9)(0.0001) + (0.1)(0.9999)} = 0.00090 \end{aligned}$$

Por lo tanto, la probabilidad condicional de que tenga la enfermedad, dado el resultado de la prueba es de aproximadamente sólo 1 de cada 1000. Por supuesto, esta probabilidad condicional es aproximadamente 9 veces mayor que la probabilidad fue antes se pusieron a prueba, pero incluso la probabilidad condicional es bastante pequeña. Otra forma de explicar este resultado es el siguiente: Sólo una persona de cada 10.000 en realidad tiene la enfermedad, pero la prueba da una respuesta positiva por aproximadamente una persona de cada 10. Por lo tanto, el número de respuestas positivas es aproximadamente 1000 veces el número de personas que realmente tienen la enfermedad. En otras palabras, de cada 1000 personas para las que el ensayo ha dado una respuesta positiva, sólo una persona en realidad tiene la enfermedad. Este ejemplo ilustra no sólo el uso del teorema de Bayes, sino también la importancia de tomar en cuenta toda la información disponible en un problema.

- Tenemos tres urnas: A con 3 bolas rojas y 5 negras, B con 2 bolas rojas y 1 negra y C con 2 bolas rojas y 3 negras. Escogemos una urna al azar y extraemos una bola. Si la bola ha sido roja, ¿Cuál es la probabilidad de haber sido extraída de la urna A?

Llamamos R a "sacar una bola roja" y

N a "sacar una bola negra".

En la Fig(23) se muestra el diagrama de árbol de las probabilidades de ocurrencia de los sucesos R o N para cada una de las tres urnas.

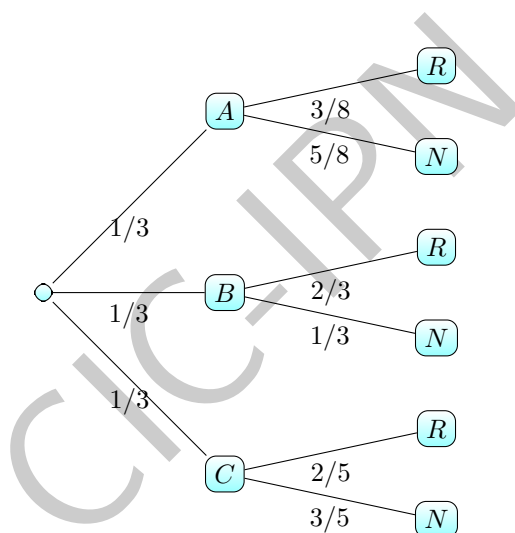


Fig. 23: Diagrama de árbol de 3 urnas con bolas de diferentes colores

$$\begin{aligned}
 P(A|R) &= \frac{P(A)P(R|A)}{P(A)P(R|A) + P(B)P(R|B) + P(C)P(R|C)} \\
 &= \frac{\frac{1}{3} \cdot \frac{3}{8}}{\frac{1}{3} \cdot \frac{3}{8} + \frac{1}{3} \cdot \frac{2}{3} + \frac{1}{3} \cdot \frac{2}{5}} = \frac{45}{173} = 0.260
 \end{aligned}$$

- En la empresa Alimentos Mr Pollo el 20% de los empleados son ingenieros y otro 20% son economistas. El 75% de los ingenieros ocupan un cargo directivo y el 50% de los economistas también, mientras que los demás trabajadores (no ingenieros y no economistas) solamente el 20% ocupa un puesto directivo. ¿Cuál es la probabilidad de que un empleado directivo elegido al azar sea ingeniero?

Para este problema se genera el diagrama de árbol que se muestre en la Fig(24).

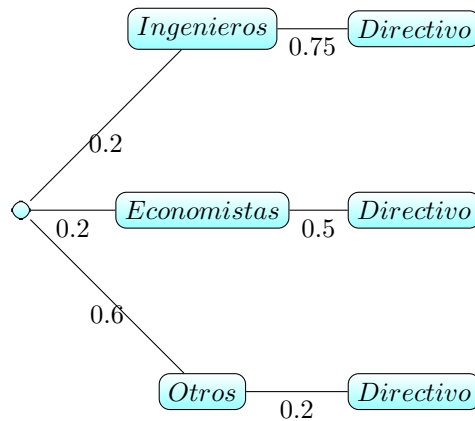


Fig. 24: Diagrama de árbol de "Alimentos Mr. Pollo"

$$P(\text{Ingeniero}|\text{Directivo}) = \frac{0.2 * 0.75}{0.2 * 0.75 + 0.2 * 0.5 + 0.6 * 0.2}$$

20 Histograma dado de tres caras

Hacer la simulación de un dado de tres caras y realizar un histograma que represente los resultados obtenidos, para: 100, 1000 y 10000 veces

- 100 veces

```

TRandom a;
TH1I hdado1("Datos","Simulacion dado 3 caras",10,0,4)
int aux1=0;
for(int i = 0; i < 100; i++)aux1 = a.Integer(3)+1; hdado1.Fill(aux1);
hdado1.Draw()

```

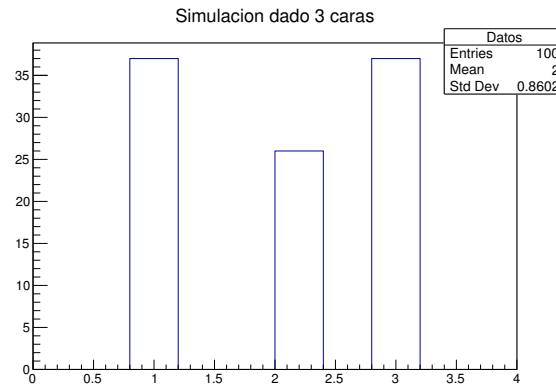


Fig. 25: Histograma dado de 3 caras con n=100

- 1000 veces

```

TRandom a;
TH1I hdado1("Datos","Simulacion dado 3 caras",10,0,4)
int aux1=0;
for(int i = 0; i < 1000; i++)aux1 = a.Integer(3)+1; hdado1.Fill(aux1);
hdado1.Draw()

```

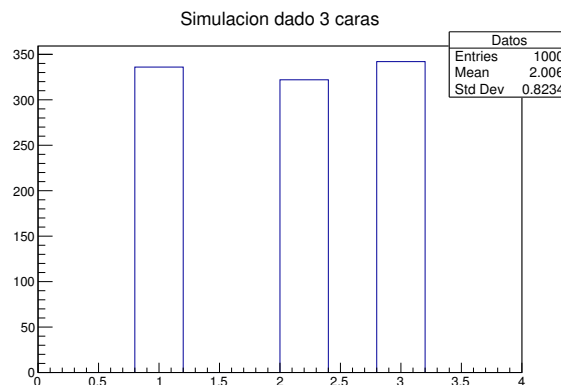


Fig. 26: Histograma dado de 3 caras con n=1000

- 10000 veces

```

TRandom a;
TH1I hdado1("Datos","Simulacion dado 3 caras",10,0,4)
int aux1=0;

```

```
for(int i = 0; i < 10000; i++)aux1 = a.Integer(3)+1; hdado1.Fill(aux1);  
hdado1.Draw()
```

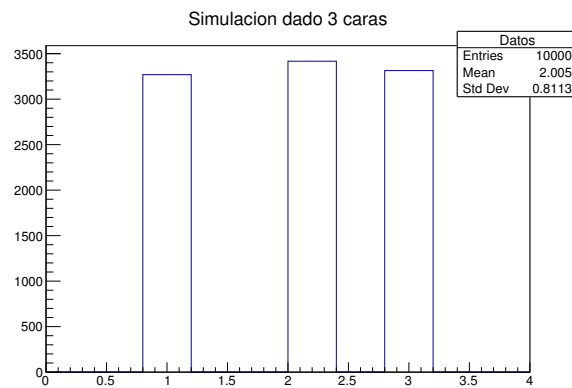


Fig. 27: Histograma dado de 3 caras con $n=10000$

21 Gráficas en ROOT

21.1 Gráfica de $\ln(x!)$

```
//Andrea Viridiana Soto Miranda
//Probabilidad , procesos aleatorios e inferencia
//CIC-IPN
void Grafica_Lnx(){
    TCanvas *c1 = new TCanvas ("c1"," Grafica de la funcion ln(x!)" ,200,10,700,500);
    c1->SetFillColor(32);
    c1->SetGrid();
    const Int_t n=170;
    Double_t x[n],y[n];
    Double_t fact;
    fact =1;
    for(Int_t i=1; i<n ; i++) {
        x[i] = i ;
        fact=fact*i;
        y[i]=log( fact);
        printf("i_%i_%f_%f\n",i,x[i],y[i]);
    }
    TGraph *gr = new TGraph (n,x,y);
    gr->SetLineColor(3);
    gr->SetLineWidth (2);
    gr->SetMarkerColor(4);
    gr->SetMarkerStyle(21);
    gr->SetTitle(" Grafica ln(x!)" );
    gr->GetXaxis()->SetTitle("X");
    gr->GetYaxis()->SetTitle("ln(x!)" );
    gr->Draw("ACP");
    // TCanvas ::Update()draws the frame , after which one can change it
    c1->Update();
    c1->GetFrame()->SetFillColor(21);
    c1->GetFrame()->SetBorderSize(12);
    c1->Modified();
}
```

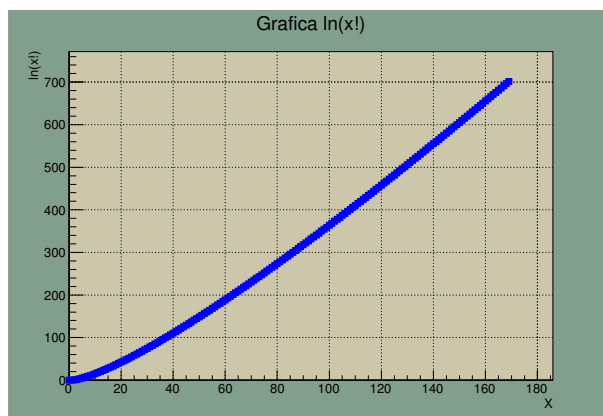


Fig. 28: Gráfica de $\ln(x!)$ en root

21.2 Gráfica de $x\ln(x)-x$

```
//Andrea Viridiana Soto Miranda
//Probabilidad , procesos aleatorios e inferencia
//CIC-IPN
void Grafica_xlnx_x(){
    TCanvas *c1 = new TCanvas ("c1"," Grafica de la funcion  $x\ln(x)-x$ ",200,10,700,500);
    c1->SetFillColor(32);
    c1->SetGrid();
    const Int_t n=170;
    Double_t x[n],y[n];
    Double_t fact;
    fact =1;
    for(Int_t i=1; i<n ; i++) {
        x[i] = i ;
        y[i]=(x[i]*log(x[i]))-x[i];
        printf("i %i %f %f\n",i,x[i],y[i]);
    }
    TGraph *gr = new TGraph (n,x,y);
    gr->SetLineColor(3);
    gr->SetLineWidth (2);
    gr->SetMarkerColor(4);
    gr->SetMarkerStyle(21);
    gr->SetTitle(" Grafica  $x\ln(x)-x$ ");
    gr->GetXaxis()->SetTitle("X");
    gr->GetYaxis()->SetTitle(" $x\ln(x)-x$ ");
    gr->Draw("ACP");
    // TCanvas :: Update() draws the frame , after which one can change it
    c1->Update();
    c1->GetFrame()->SetFillColor(21);
    c1->GetFrame()->SetBorderSize(12);
    c1->Modified();
}
```

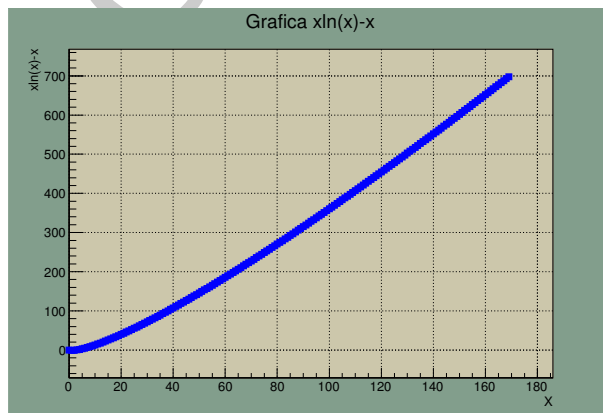


Fig. 29: Gráfica de $x\ln(x)-x$ en root

21.3 Gráfica de $\ln(x!)-(x\ln(x)-x)$

```
//Andrea Viridiana Soto Miranda
//Probabilidad , procesos aleatorios e inferencia
//CIC-IPN
void Grafica_diferencia(){
    TCanvas *c1 = new TCanvas ("c1"," Grafica de la funcion  $\ln(x!)-(x\ln(x)-x)$ ",200,10,700,500);
    c1->SetFillColor(32);
    c1->SetGrid();
    const Int_t n=170;
    Double_t x[n],y[n];
    Double_t fact;
    fact =1;
    for(Int_t i=1; i<n ; i++) {
        x[i] = i ;
        fact=TMath::Factorial(x[i]);
        y[i]=log(fact)-(x[i]*log(x[i])-x[i]);
        printf("i %i %f %f\n",i,x[i],y[i]);
    }
    TGraph *gr = new TGraph (n,x,y);
    gr->SetLineColor(3);
    gr->SetLineWidth (2);
    gr->SetMarkerColor(4);
    gr->SetMarkerStyle(21);
    gr->SetTitle(" Grafica  $\ln(x!)-(x\ln(x)-x)$ ");
    gr->GetXaxis()->SetTitle("X");
    gr->GetYaxis()->SetTitle("  $\ln(x!)-(x\ln(x)-x)$ ");
    gr->Draw("ACP");
    // TCanvas ::Update()draws the frame , after which one can change it
    c1->Update();
    c1->GetFrame()->SetFillColor(21);
    c1->GetFrame()->SetBorderSize(12);
    c1->Modified();
}
Double_t TMath::Factorial(Int_t n)
{
    // Compute factorial(n).

    if (n <= 0) return 1.;
    Double_t x = 1;
    Int_t b = 0;
    do {
        b++;
        x *= b;
    } while (b != n);
    return x;
}
```

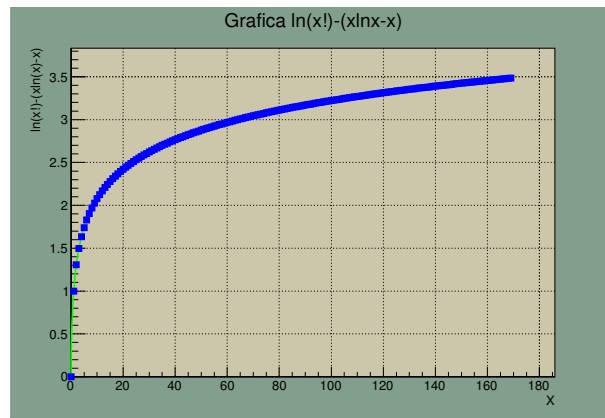


Fig. 30: Grafica de $\ln(x!) - (x \ln x - x)$ en root

22 Integrales y derivadas con Mathematica

- `Integrate[Sin[3x]Cos[x + 1], x]`

$$\frac{1}{8}(-2\cos[1 - 2x] - \cos[1 + 4x])$$

- `Integrate[1/(x + 1), x]`

$$\text{Log}[1 + x]$$

- `Integrate[Exp[-(x/2)], x]`

$$-2e^{-x/2}$$

- `Integrate[x^2Exp[4x + 1], x]`

$$\frac{1}{32}e^{1+4x}(1 - 4x + 8x^2)$$

- `Integrate[Tan[x]Cos[2x], x]`

$$-\frac{1}{2}\cos[2x] + \text{Log}[\cos[x]]$$

- `D[Tan[x]Sin[2x], x]`

$$\sec[x]^2\sin[2x] + 2\cos[2x]\tan[x]$$

- `D[1/(Sqrt[x^2 + 5x + 6]), x]`

$$-\frac{5+2x}{2(6+5x+x^2)^{3/2}}$$

- `D[xExp[x + 8], x]`

$$e^{8+x} + e^{8+x}x$$

- $D[1/(\sin[x])^2, x]$

$$-2\cot[x]\csc[x]^2$$

- $D[\operatorname{ArcSinh}[x], x]$

$$\frac{1}{\sqrt{1+x^2}}$$

CIC-IPN

23 Teorema del límite central

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

Sea $\mathcal{N}(\mu, \sigma^2)$ la función de densidad de la distribución normal definida como:

$$f_{\mu, \sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

con una media μ y una varianza σ^2 . El caso en el que su función de densidad sea $\mathcal{N}(0, 1)$ a la distribución se le conoce como normal estándar.

Se define S_n la suma de n variables aleatorias, independientes, idénticamente distribuidas, y con una media μ y varianza σ^2 finitas ($\sigma^2 \neq 0$):

$$S_n = X_1 + \dots + X_n$$

de manera que, la media de S_n es $n \cdot \mu$ y la varianza $n \cdot \sigma^2$, dado que son variables aleatorias independientes. Con tal de hacer más fácil la comprensión del teorema y su posterior uso, se hace una estandarización de S_n como:

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

para que la media de la nueva variable sea igual a 0 y la desviación estándar sea igual a 1. Así, las variables Z_n convergerán en distribución a la distribución normal estándar $N(0,1)$, cuando n tienda a infinito. Como consecuencia, si $\phi(z)$ es la función de distribución de $N(0,1)$, para cada número real z :

$$\lim_{n \rightarrow \infty} P_r(Z_n \leq z) = \phi(z)$$

donde $P_r()$ indica probabilidad y \lim se refiere a límite matemático.

De manera formal, normalizada y compacta el enunciado del teorema es:

Sea X_1, X_2, \dots, X_n un conjunto de variables aleatorias, independientes e idénticamente distribuidas con media μ y varianza $0 < \sigma^2 < \infty$. Sea:

$$S_n = X_1 + \dots + X_n$$

entonces:

$$\lim_{n \rightarrow \infty} P_r\left(\frac{S_n - n\mu}{\sigma\sqrt{n}}\right) = \phi(z)$$

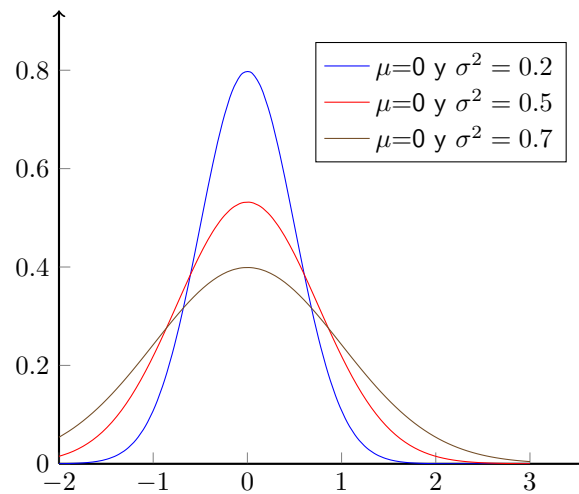


Fig. 31: Teorema del limite central

24 Densidad de probabilidad

En la teoría de la probabilidad, la función de densidad de probabilidad, función de densidad, o, simplemente, densidad de una variable aleatoria continua describe la probabilidad relativa según la cual dicha variable aleatoria tomará determinado valor. La probabilidad de que la variable aleatoria caiga en una región específica del espacio de posibilidades estará dada por la integral de la densidad de esta variable entre uno y otro límite de dicha región. La función de densidad de probabilidad (FDP o PDF en inglés) es no-negativa a lo largo de todo su dominio y su integral sobre todo el espacio es de valor unitario.

Una función de densidad de probabilidad caracteriza el comportamiento probable de una población en tanto especifica la posibilidad relativa de que una variable aleatoria continua X tome un valor cercano a x .

Una variable aleatoria X tiene densidad f , siendo f una función no-negativa integrable de Lebesgue, si:

$$P[a \leq X \leq b] = \int_a^b f(x)dx$$

Por lo tanto, si F es la función de distribución acumulativa de X , entonces:

$$F(x) = \int_{-\infty}^x f(u)du$$

y (si f es continua en x)

$$f(x) = \frac{d}{dx} F(x)$$

Intuitivamente, puede considerarse $f(x)dx$ como la probabilidad de X de caer en el intervalo infinitesimal $[x, x + dx]$.

Se define como el cociente entre la probabilidad de X de tomar un valor en el intervalo $[x, x + dx]$ y dx , siendo dx un infinitésimo. La mayoría de las funciones de densidad de probabilidad requieren uno o más parámetros para especificarlas totalmente. Recíprocamente respecto de la definición ya desarrollada, pueden hacerse las siguientes consideraciones. La probabilidad de que una variable aleatoria continua X quede ubicada entre los valores a y b está dada por el desenvolvimiento en el intervalo de la FDP; de los valores comprendidos en el rango entre a y b .

$$P[a \leq X \leq b] = \int_a^b f(x)dx$$

La FDP es la derivada (cuando existe) de la función de distribución:

$$f(x) = \frac{d}{dx} F(x)$$

Así, si F es la función de distribución acumulativa de X , entonces:

$$F(x) = \int_{-\infty}^x f(u) du$$

y (si f es continua en x)

$$f(x) = \frac{d}{dx} F(x)$$

En la Fig(32), se muestra un ejemplo de densidad de probabilidad, en este caso es la campana de Gauss con diferentes parametros.

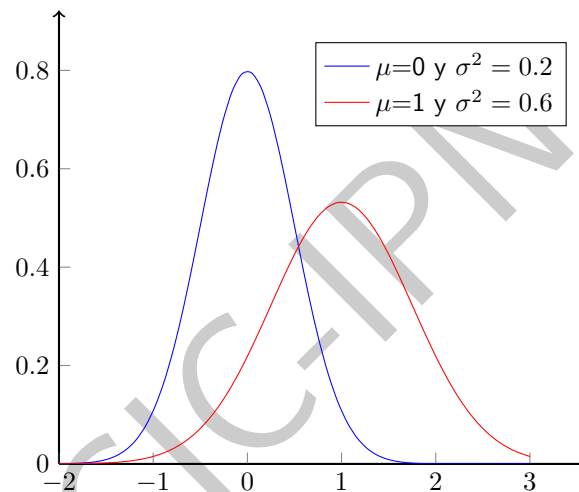


Fig. 32: Densidad de probabilidad

25 Mínimos cuadrados

Mínimos cuadrados es una técnica de análisis numérico enmarcada dentro de la optimización matemática, en la que, dados un conjunto de pares ordenados variable independiente, variable dependiente y una familia de funciones, se intenta encontrar la función continua, dentro de dicha familia, que mejor se aproxime a los datos (un "mejor ajuste"), de acuerdo con el criterio de mínimo error cuadrático.

En su forma más simple, intenta minimizar la suma de cuadrados de las diferencias en las ordenadas (llamadas residuos) entre los puntos generados por la función elegida y los correspondientes valores en los datos. Específicamente, se llama mínimos cuadrados promedio (LMS) cuando el número de datos medidos es 1 y se usa el método de descenso por gradiente para minimizar el residuo cuadrado. Se puede demostrar que LMS minimiza el residuo cuadrado esperado, con el mínimo de operaciones (por iteración), pero requiere un gran número de iteraciones para converger.

Desde un punto de vista estadístico, un requisito implícito para que funcione el método de mínimos cuadrados es que los errores de cada medida estén distribuidos de forma aleatoria. El teorema de Gauss-Markov prueba que los estimadores mínimos cuadráticos carecen de sesgo y que el muestreo de datos no tiene que ajustarse, por ejemplo, a una distribución normal. También es importante que los datos a procesar estén bien escogidos, para que permitan visibilidad en las variables que han de ser resueltas (para dar más peso a un dato en particular, véase mínimos cuadrados ponderados).

La técnica de mínimos cuadrados se usa comúnmente en el ajuste de curvas. Muchos otros problemas de optimización pueden expresarse también en forma de mínimos cuadrados, minimizando la energía o maximizando la entropía, en la Fig(33) se muestra una gráfica de aproximación por mínimos cuadrados.

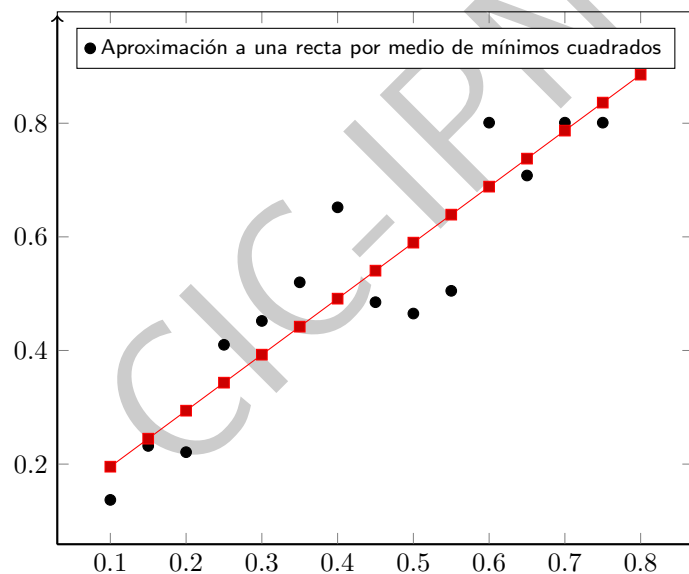


Fig. 33: Mínimos cuadrados

26 Explosión del Challenger

La temperatura para la cual se habían hecho las consideraciones para el lanzamiento del challenger no eran las mismas que se tenían en el día del lanzamiento, este se estuvo retrasando el despegue por algunos detalles que se habían presentado, en total 7 veces al final deciden lanzarlo el Martes 28 de Enero de 1986 a las 9:38 A. M. a pesar de las advertencias de los fabricantes e ingenieros a cargo.

Después de un último retraso de 2 horas debido al hielo formado en varias partes de la Base de Lanzamiento.

Se ajustó a hora de lanzamiento para las 11:38 A.M.

73 segundos antes del desastre. El cohete propulsor se enciende. Dentro del cohete derecho, los anillos se endurecieron debido a las temperaturas bajo cero, no pueden mantener selladas las uniones de campo inferiores, los anillos comienzan a quemarse.

Mientras que los gases super calientes se escapan, pequeños trozos de escoria de aluminio provenientes del combustible para cohetes se acumulan y bloquean el agujero, evitando una catástrofe en la plataforma de lanzamiento.

15 segundos antes del desastre el challenger entra en una corriente en chorro muy estrecha que se mueve rápidamente y el transbordador se sacude violentamente. Entonces la escoria de aluminio se suelta, casi de inmediato una llama aparece en el cohete derecho.

8 segundos antes del desastre. Con la intensidad de un soplete la llama penetra el tanque externo y el hidrógeno líquido comienza a derramarse.

1 segundo antes del desastre el acoplamiento entre el cohete y el tanque se desprende y toda la sección inferior se separa. El infierno empuja el compartimiento del hidrógeno líquido hacia arriba, contra el contenedor lleno de oxígeno, justo cuando la nariz del cohete impacta contra la parte superior del tanque externo.

Se queman casi dos millones de litros de combustible en forma instantanea y el transbordador estalla.



Fig. 34: Explosión del Challenger

27 Ejercicios 1.4 y 1.5

27.1 Ejercicio 1.4

Una máquina de fabricar tornillos se desajusta de forma que produce tornillos de longitud indeterminada. Se sabe que la probabilidad de que la longitud de un tornillo sea inferior a L_1 es p_1 , y de que sea superior a L_2 es p_2 , donde $L_2 > L_1$ [8]

Se extraen cuatro tornillos al azar, determínese:

- Probabilidad de que los cuatro tengan longitud superior a L_1 .
- Probabilidad de que tres de ellos tengan longitudes entre L_1 y L_2 .
- Probabilidad de que, a lo sumo, tres sean de longitud inferior a L_2 .
- Probabilidad de que, como máximo, haya dos tornillos con longitud comprendida entre L_1 y L_2 .

En la Fig(35), se muestra el esquema de longitudes empleado en el problema 1.4.

SOLUCIÓN

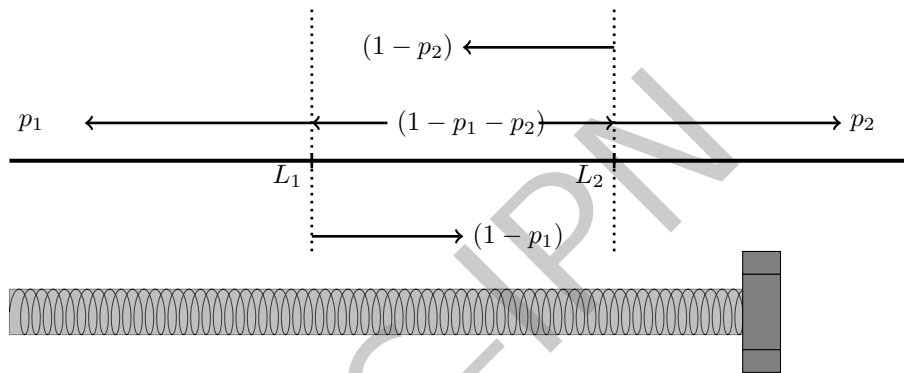


Fig. 35: Esquema de longitudes para el problema 1.4

Para el inciso a) se genera el diagrama de árbol mostrado en la Fig(36).

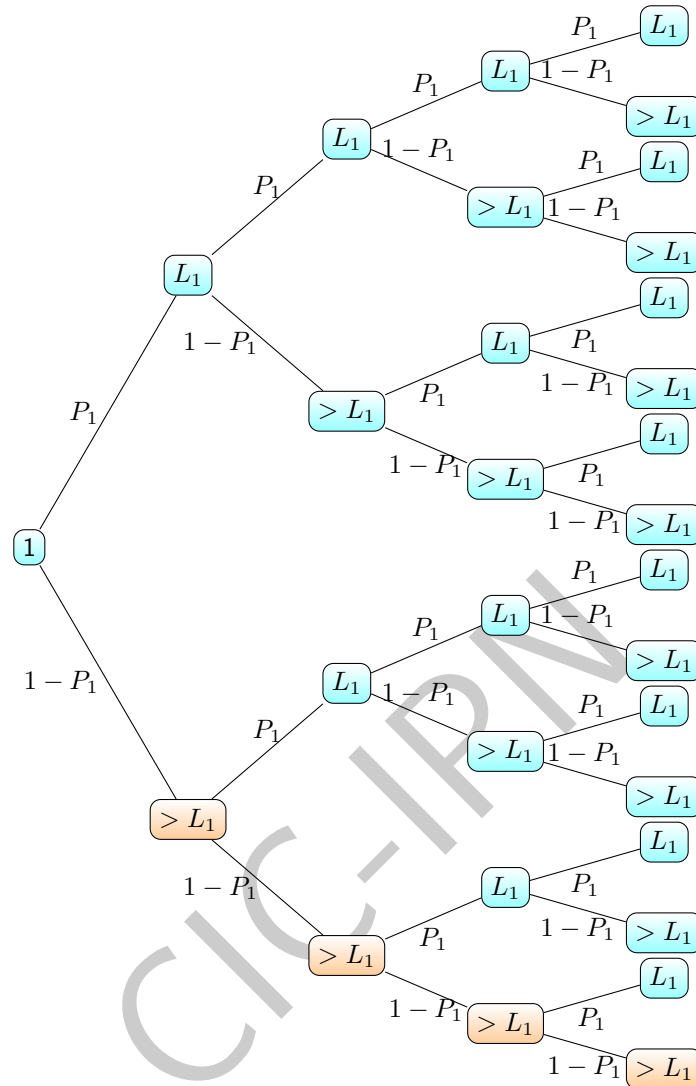


Fig. 36: Los 4 tornillos tienen longitudes superiores a L_1

Los casos en los que se cumple la condición de que los cuatro tengan longitud mayor a L_1 , están resaltadas de verde, el resultado es multiplicar las probabilidades que corresponden al caso, por lo que queda:

$$P(4\text{torn}.supL_1) = (1 - p_1)(1 - p_1)(1 - p_1)(1 - p_1) = (1 - p_1)^4$$

En el caso del inciso b), designamos L_3 como la Longitud entre L_1 y L_2 con P_3 que es $1 - P_1 - P_2$ y se genera el diagrama de árbol que se muestra en la Fig(37,38,39):

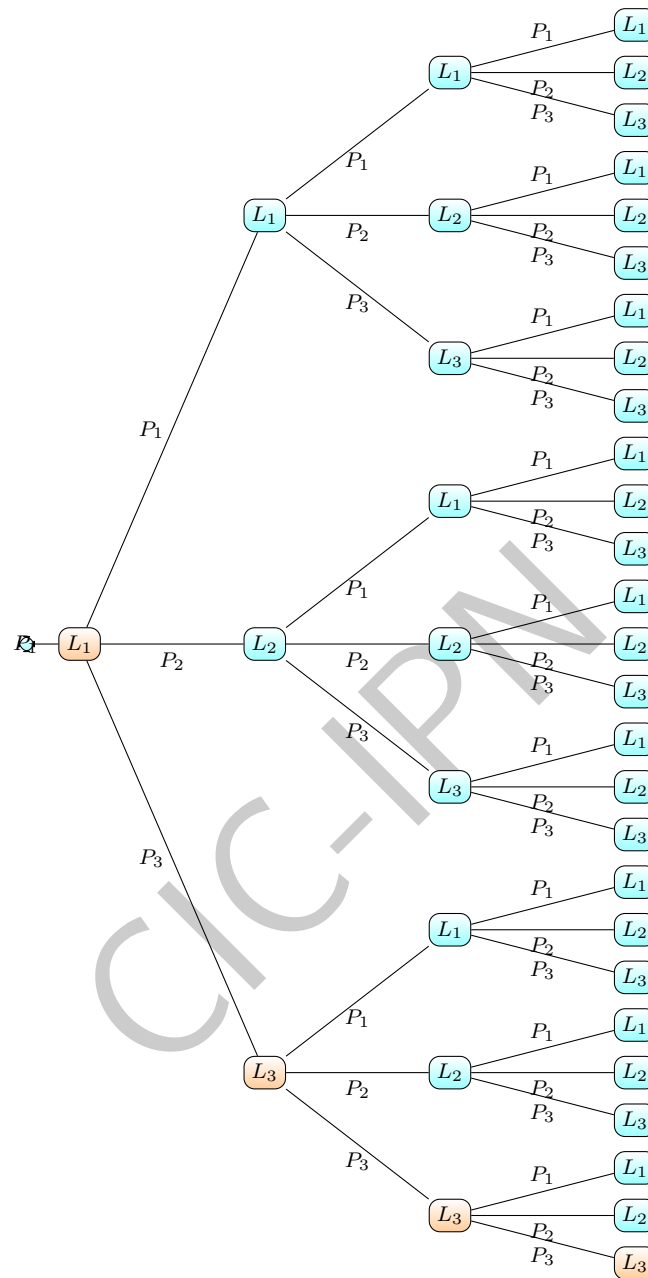


Fig. 37: Al menos 3 tornillos esten entre L_1 y L_2

En esta primera rama solo se obtiene $P_1P_3P_3P_3$ como se muestra en la Fig(37).

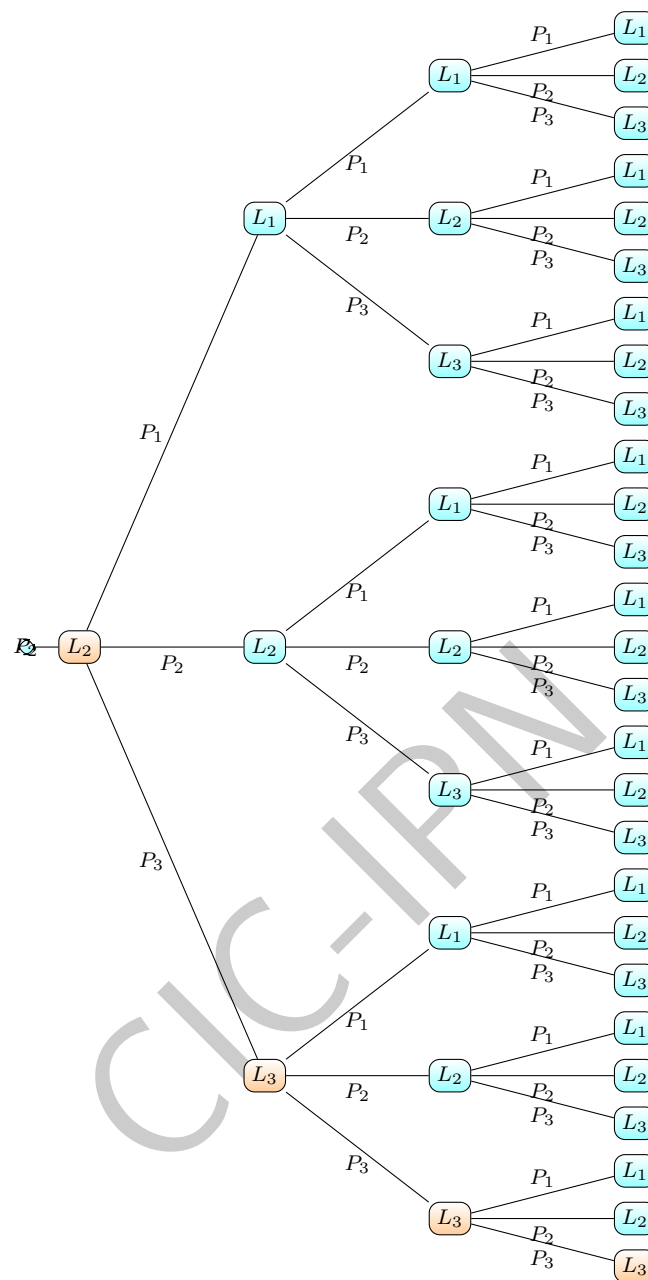
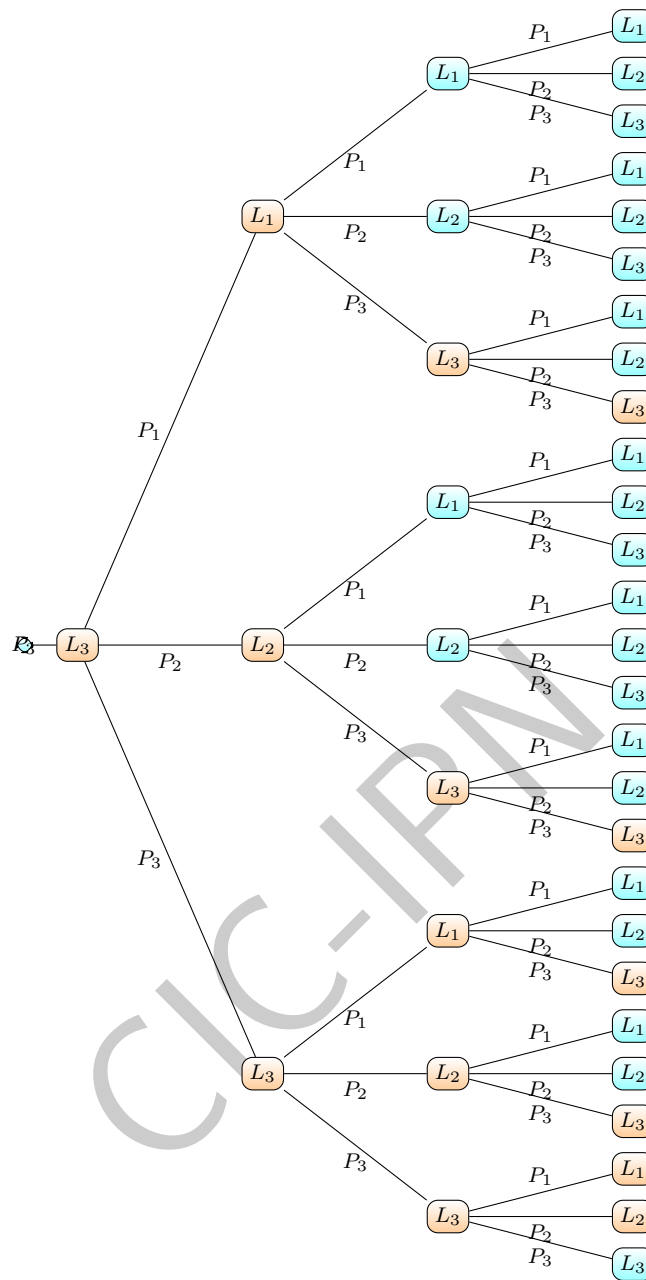


Fig. 38: Al menos 3 tornillos esten entre L_1 y L_2

En esta segunda rama solo se obtiene $P_2P_3P_3P_3$ como se muestra en la Fig(38).

Fig. 39: Al menos 3 tornillos esten entre L_1 y L_2

En a última rama solo se obtiene $P_3P_1P_3P_3 + P_3P_2P_3P_3 + P_3P_3P_1P_3 + P_3P_3P_2P_3 + P_3P_3P_3P_1 + P_3P_3P_3P_2$ como se muestra en la Fig(39)

Si agrupamos los resultados de las tres ramas queda: $P_1P_3P_3P_3 + P_2P_3P_3P_3 + P_3P_1P_3P_3 + P_3P_2P_3P_3 + P_3P_3P_1P_3 + P_3P_3P_2P_3 + P_3P_3P_3P_1 + P_3P_3P_3P_2$

Observamos que hay cuatros formas diferentes en que caiga P_1 y P_3 y de la misma forma P_2 y P_3 por lo que se agrupan los términos de la siguiente manera:

$$P(3\text{tornillos entre } L_1 \text{ y } L_2) = 4P_1P_3^2 + 4P_2P_3^2$$

Sustituyendo P_3 queda $P(3\text{tornillos entre } L_1 \text{ y } L_2) = 4P_1(1 - P_1 - P_2)^3 + 4P_2(1 - P_1 - P_2)^3$

Para el inciso c) se genera el diagrama de árbol que se muestra en la Fig(40).

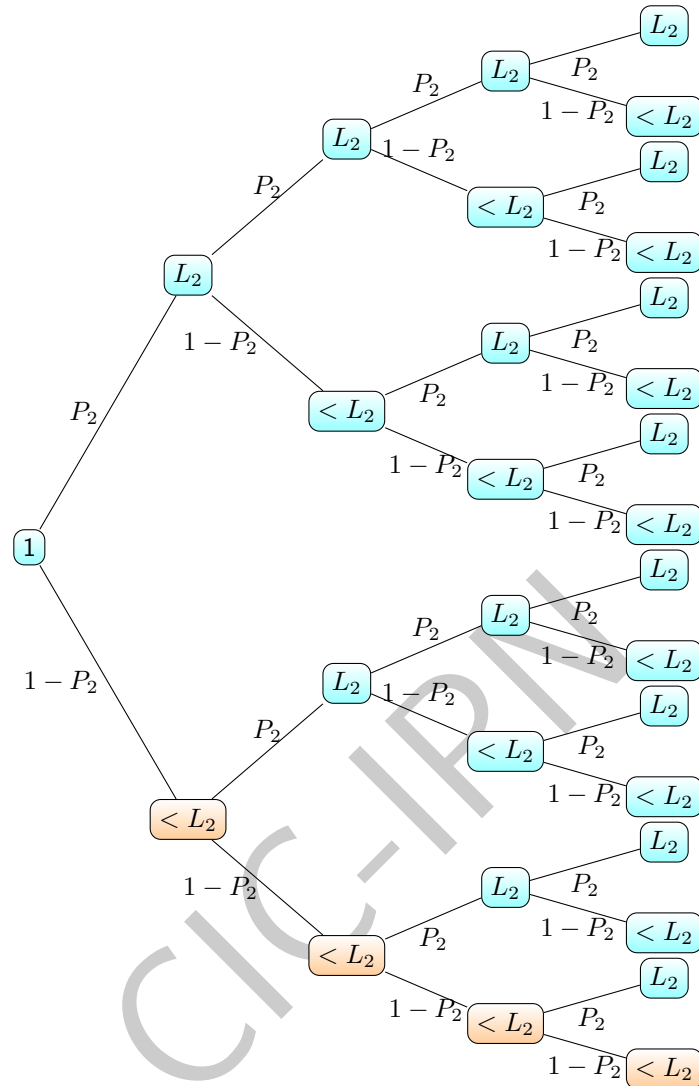


Fig. 40: A lo sumo tres tornillos con longitud menor a L_2

Considerando el suceso complementario decimos que $P(\text{máx}3\text{tor} < L_2) = 1 - P(4\text{tor} < L_2) = 1 - (1 - P_2)(1 - P_2)(1 - P_2)(1 - P_2) = 1 - (1 - P_2)^4$

27.2 Ejercicio 1.5

Un experimento consiste en lanzar una moneda y a continuación un dado, cuatro veces consecutivas.

¿Cuál es la probabilidad de que se obtenga cara antes de que salga un 6 en el dado si la probabilidad de obtener cara en la moneda es $2/3$? diagrama de árbol mostrado en la Fig(41), el cual por motivos de espacio se limita a poner los casos en los cuales se cumple la condición los cuales están en color naranja [8].

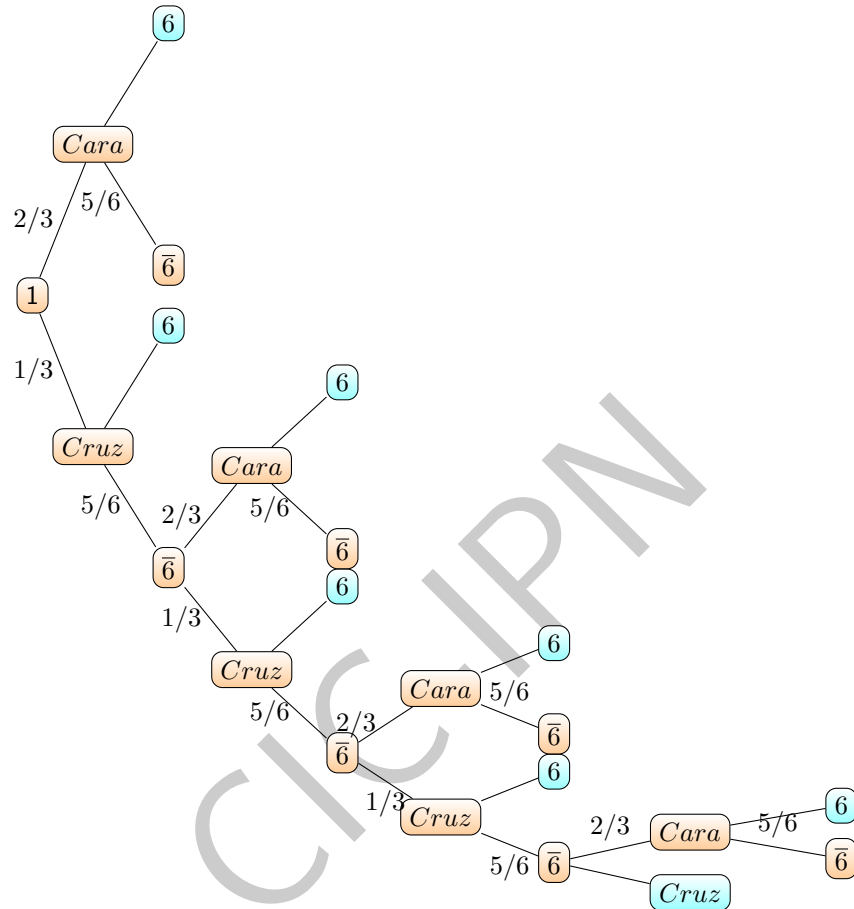


Fig. 41: Diagrama de árbol de lanzamiento de una moneda y un dado

La solución es : $P(S) = P(en1_{er}experimento) + P(en2_{do}experimento) + P(en3_{er}experimento) + P(en4_{to}experimento)$

$$P(S) = \frac{2}{3} \frac{5}{6} + \left(\frac{1}{3} \frac{5}{6}\right) \left(\frac{2}{3} \frac{5}{6}\right) + \left(\frac{1}{3} \frac{5}{6}\right)^2 \left(\frac{2}{3} \frac{5}{6}\right) + \left(\frac{1}{3} \frac{5}{6}\right)^3 \left(\frac{2}{3} \frac{5}{6}\right)$$

$$= \frac{10}{18} \left[1 + \frac{5}{18} + \frac{5}{18}^2 + \frac{5}{18}^3 \right] = 0.765$$

28 Grafos por criterios de divisibilidad

Que un número sea divisible entre otro quiere decir que, en un lenguaje sencillo, al dividir (división euclídea) el primero entre el segundo se obtiene de resto cero, es decir, que la división es exacta.

Si tenemos, por ejemplo, una pizza de 8 porciones y somos 3 comensales, se trata de ver si tocamos a un número entero de porciones cada persona (que 8 sea divisible entre 3) o si, por el contrario, sobra alguna o algunas de las porciones y hay que partirla o partirlas en trozos más pequeños para que todos comamos lo mismo y no quede nada (que 8 no sea divisible entre 3).

Que un número sea divisible entre otro es equivalente a decir que dicho número es múltiplo del otro. Por ejemplo, 6 es divisible entre 2 (al dividir 6 entre 2 se obtiene de cociente 3 y de resto 0), y 6 es múltiplo de 2 (6 es igual a 2 por 3).

Ver si un número es divisible entre otro cuando los números son pequeños es relativamente sencillo. Sin embargo, cuando tenemos números más grandes resulta algo más complicado.

Para facilitar esta labor surgen los criterios de divisibilidad. Digamos que son unas reglas que empleamos para saber si un número es divisible entre otro sin necesidad de realizar la división.

En la escuela se suelen enseñar, normalmente, los criterios de divisibilidad entre 2, 3, 5, 6 (derivado de los criterios de divisibilidad entre 2 y 3), 9, 10 y 11.

Por ejemplo, un número es divisible entre 2 si el número termina en 0 ó una cifra par (2, 4, 6, 8). O, un número es divisible entre 5 si la última cifra es 0 ó 5.

Pero hay otros criterios de divisibilidad que no suelen verse, quizás porque ya no son tan sencillos, aunque realmente algunos sí que lo son.

Supongamos que queremos hacer un grafo para n divisibilidad, entonces tenemos que hacer n nodos y marcarlos desde el 0, 1, ..., $n - 1$.

Ahora agregaremos flechas negras a cada nodo i hacía el nodo $i + 1$, y también desde el nodo $n - 1$ hacía el nodo 0.

Ahora agregaremos flechas azules a cada nodo i hacía el nodo $\text{mod}(\frac{i*10}{n})$. Esto es, desde i hacía el resto que da cuando dividimos $\frac{i*10}{n}$. Por ejemplo, si $i = 2$ y $n = 7$, se deberá de poner agregar una flecha azul desde el nodo 2 hacía el nodo 6 porque el resto despues de dividir $\frac{2*10}{7}$ es 6.

Ya que tenemos nuestro grafo ahora veremos como tenemos que recorrerlo para saber si el número a analizar es divisible entre cierto número (dado por el grafo) o no.

Siempre iniciamos en el nodo 0 del grafo, ahora de nuestro número a analizar tomamos su primer cifra y recorremos ese número de nodos por medio de las flechas negras, al estar en el ultimo nodo ahora avanzamos una vez por la flecha azul y vamos hacía el nodo al que nos dirige, ahora tomamos la siguiente cifra de nuestro número a analizar, y hacemos lo mismo, recorremos ese número de nodos por medio de las flechas negras, al llegar al ultimo nodo avanzamos al siguiente nodo por medio de la flecha azul y vamos hacía el nodo al que nos dirige y asi sucesivamente. Si al terminar nuestro recorrido nos quedamos en el nodo 0 esto quiere decir que el número a analizar si es divisible por el número del grafo.

Por ejemplo si construimos el grafo para el número divisible 5, queda como el de la Fig(43):

Nuestro número a analizar será 255, tomamos la primer cifra, que es el 2, por lo tanto recorremos dos nodos por medio de las flechas negras, entonces terminamos en el nodo 2, ahora avanzamos un nodo por medio de la flecha azul, ahora quedamos en el nodo 0, tomamos ahora la siguiente cifra que es 5, de nuevo recorremos 5 nodos por las flechas negras y quedamos en el nodo 0, recorremos un nodo por medio de la flecha azul, y quedamos en el mismo nodo, el 0. Tomamos la ultima cifra y vemos que terminamos en el nodo 0 de nuevo, por lo tanto 255 si es divisible con 5.

28.1 Grafo de divisibilidad de 3

En la Fig(42) se muestra el grafo de divisibilidad de 3.

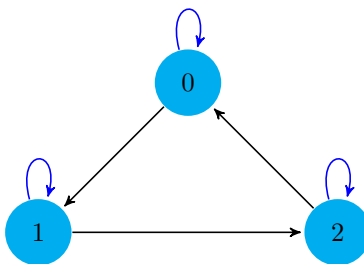


Fig. 42: Grafo de divisibilidad de 3

28.2 Grafo de divisibilidad de 5

En la Fig(43) se muestra el grafo de divisibilidad de 5.

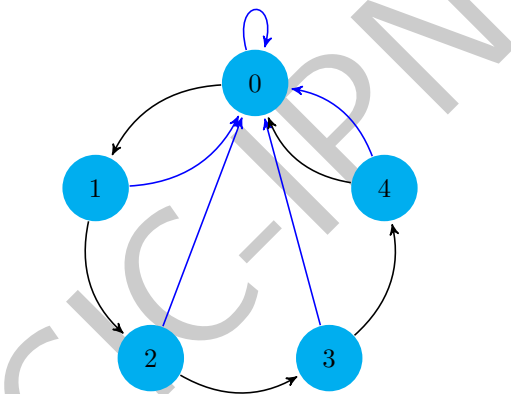


Fig. 43: Grafo de divisibilidad de 5

28.3 Grafo de divisibilidad de 7

En la Fig(44) se muestra el grafo de divisibilidad de 7.

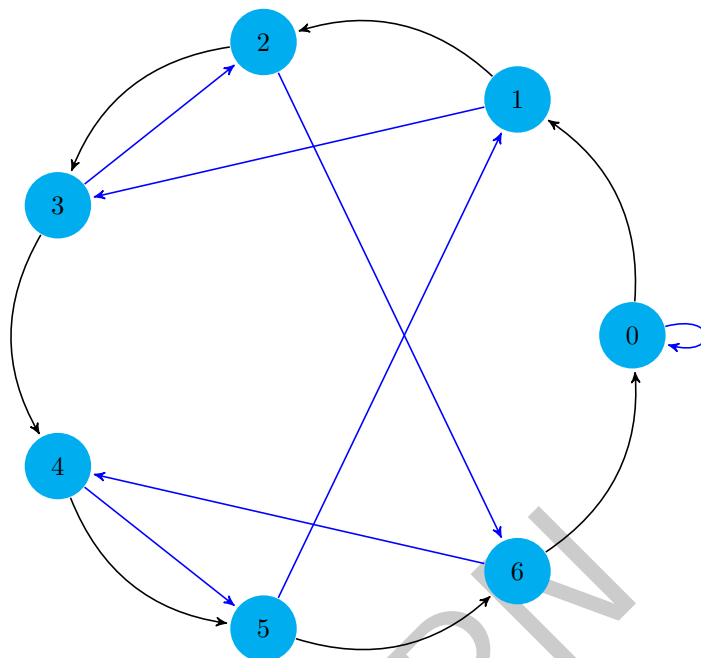


Fig. 44: Grafo de divisibilidad de 7

28.4 Grafo de divisibilidad de 11

En la Fig(45) se muestra el grafo de divisibilidad de 11.

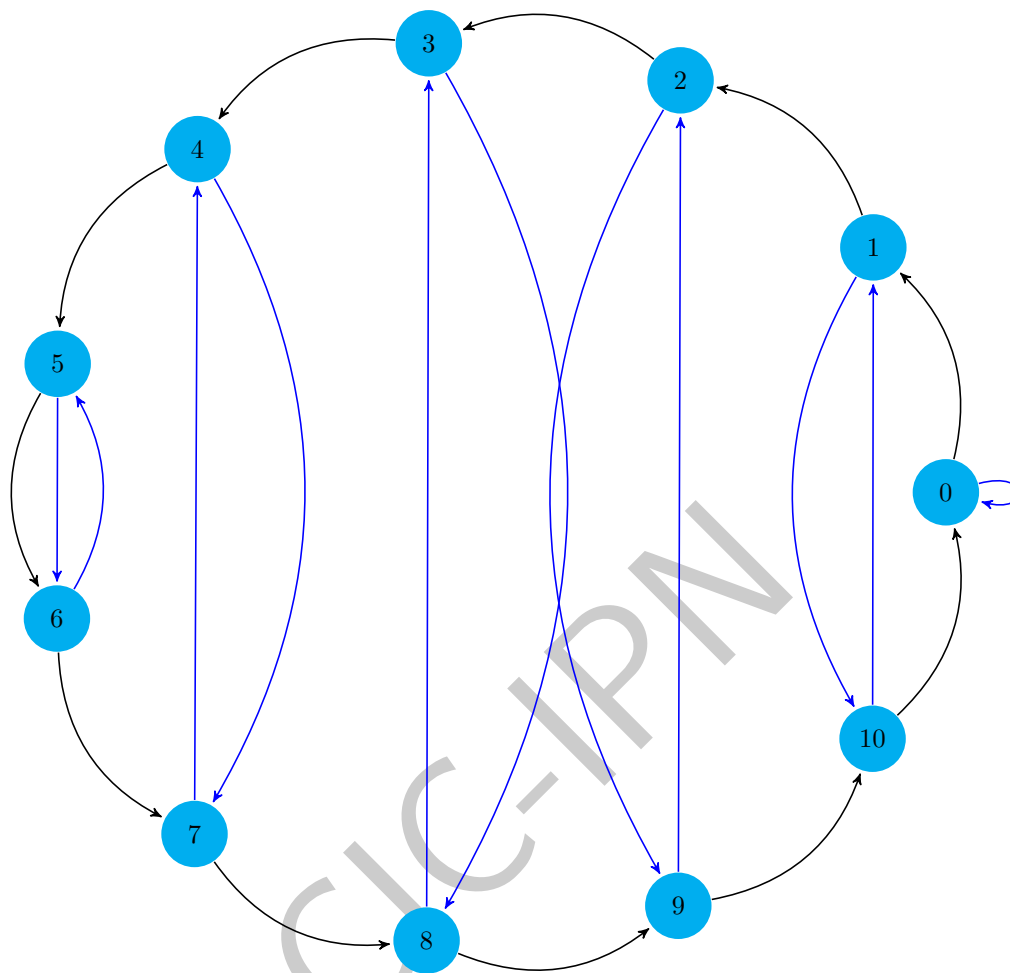


Fig. 45: Grafo de divisibilidad de 11

29 Símbolo de Levi-Civita

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

$$\epsilon_{ijk} = \begin{cases} +1 & \text{si } (i, j, k) \text{ es } (1, 2, 3), (2, 3, 1) \text{ o } (3, 1, 2) \\ -1 & \text{si } (i, j, k) \text{ es } (3, 2, 1), (1, 3, 2) \text{ o } (2, 1, 3) \\ 0 & \text{de otro modo } i = j \text{ o } j = k \text{ o } k = i \end{cases}$$

nombrado así por Tullio Levi-Civita. Se utiliza en muchas áreas de las matemáticas y en física. Por ejemplo, en álgebra lineal, el producto vectorial de dos vectores se puede escribir como:

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \sum_{i=1}^3 \left(\sum_{j,k=1}^3 \epsilon_{ijk} a_j b_k \right) \mathbf{e}_i$$

o más simplemente:

$$\mathbf{a} \times \mathbf{b} = \mathbf{c}, c_i = \sum_{j,k=1}^3 \epsilon_{ijk} a_j b_k$$

29.1 Calcular c_1 , c_2 y c_3 de $\mathbf{a} \times \mathbf{b}$

- Calculando c_1

$$\begin{aligned} c_1 &= \sum_{j,k=1}^3 \epsilon_{1jk} a_j b_k = \epsilon_{111} a_1 b_1 + \epsilon_{112} a_1 b_2 + \epsilon_{113} a_1 b_3 + \\ &\quad \epsilon_{121} a_2 b_1 + \epsilon_{122} a_2 b_2 + \epsilon_{123} a_2 b_3 + \\ &\quad \epsilon_{131} a_3 b_1 + \epsilon_{132} a_3 b_2 + \epsilon_{133} a_3 b_3 \end{aligned}$$

Se tomamos en cuenta el valor de ϵ_{ijk} :

$$\epsilon_{ijk} = \begin{cases} +1 & \text{si } (i, j, k) \text{ es } (1, 2, 3), (2, 3, 1) \text{ o } (3, 1, 2) \\ -1 & \text{si } (i, j, k) \text{ es } (3, 2, 1), (1, 3, 2) \text{ o } (2, 1, 3) \\ 0 & \text{de otro modo } i = j \text{ o } j = k \text{ o } k = i \end{cases}$$

$$\begin{aligned} c_1 &= \cancel{\epsilon_{111} a_1 b_1} + \cancel{\epsilon_{112} a_1 b_2} + \cancel{\epsilon_{113} a_1 b_3} + \\ &\quad \epsilon_{121} a_2 b_1 + \cancel{\epsilon_{122} a_2 b_2} + \epsilon_{123} a_2 b_3 + \\ &\quad \epsilon_{131} a_3 b_1 + \cancel{\epsilon_{132} a_3 b_2} + \cancel{\epsilon_{133} a_3 b_3} \end{aligned}$$

$$c_1 = a_2 b_3 - a_3 b_2$$

- Calculando c_2

$$c_2 = \sum_{j,k=1}^3 \epsilon_{2jk} a_j b_k = \epsilon_{211} a_1 b_1 + \epsilon_{212} a_1 b_2 + \epsilon_{213} a_1 b_3 +$$

$$\epsilon_{221} a_2 b_1 + \epsilon_{222} a_2 b_2 + \epsilon_{223} a_2 b_3 +$$

$$\epsilon_{231} a_3 b_1 + \epsilon_{232} a_3 b_2 + \epsilon_{233} a_3 b_3$$

Se tomamos en cuenta el valor de ϵ_{ijk} :

$$\epsilon_{ijk} = \begin{cases} +1 & \text{si } (i, j, k) \text{ es } (1, 2, 3), (2, 3, 1) \text{ o } (3, 1, 2) \\ -1 & \text{si } (i, j, k) \text{ es } (3, 2, 1), (1, 3, 2) \text{ o } (2, 1, 3) \\ 0 & \text{de otro modo } i = j \text{ o } j = k \text{ o } k = i \end{cases}$$

$$c_1 = \cancel{\epsilon_{211} a_1 b_1} + \cancel{\epsilon_{212} a_1 b_2} + \cancel{\epsilon_{213} a_1 b_3} +$$

$$\cancel{\epsilon_{221} a_2 b_1} + \cancel{\epsilon_{222} a_2 b_2} + \cancel{\epsilon_{223} a_2 b_3} +$$

$$\epsilon_{231} a_3 b_1 + \epsilon_{232} a_3 b_2 + \epsilon_{233} a_3 b_3$$

$$c_2 = -a_1 b_3 + a_3 b_1$$

- Calculando c_3

$$c_3 = \sum_{j,k=1}^3 \epsilon_{3jk} a_j b_k = \epsilon_{311} a_1 b_1 + \epsilon_{312} a_1 b_2 + \epsilon_{313} a_1 b_3 +$$

$$\epsilon_{321} a_2 b_1 + \epsilon_{322} a_2 b_2 + \epsilon_{323} a_2 b_3 +$$

$$\epsilon_{331} a_3 b_1 + \epsilon_{332} a_3 b_2 + \epsilon_{333} a_3 b_3$$

Se tomamos en cuenta el valor de ϵ_{ijk} :

$$\epsilon_{ijk} = \begin{cases} +1 & \text{si } (i, j, k) \text{ es } (1, 2, 3), (2, 3, 1) \text{ o } (3, 1, 2) \\ -1 & \text{si } (i, j, k) \text{ es } (3, 2, 1), (1, 3, 2) \text{ o } (2, 1, 3) \\ 0 & \text{de otro modo } i = j \text{ o } j = k \text{ o } k = i \end{cases}$$

$$c_3 = \cancel{\epsilon_{311} a_1 b_1} + \cancel{\epsilon_{312} a_1 b_2} + \cancel{\epsilon_{313} a_1 b_3} +$$

$$\epsilon_{321} a_2 b_1 + \epsilon_{322} a_2 b_2 + \epsilon_{323} a_2 b_3 +$$

$$\epsilon_{331} a_3 b_1 + \epsilon_{332} a_3 b_2 + \epsilon_{333} a_3 b_3$$

$$c_3 = a_1 b_2 - a_2 b_1$$

30 Convenio de suma de Einstein

Se denomina convenio de suma de Einstein, notación de Einstein o notación indexada a la convención utilizada para abreviar la escritura de sumatorios, en el que se suprime el símbolo de sumatorio (representado con la letra griega sigma - \sum). El convenio fue introducido por Albert Einstein en 1916. Se aplica en matemáticas en especial a los cálculos realizados en álgebra lineal destinados a la física. El convenio se aplica sólo a sumatorios sobre índices repetidos. El convenio se usa especialmente con tensores donde es muy frecuente la operación de suma sobre índices repetidos y sería muy fatigoso escribir explícitamente los signos de sumatorios.

Dada una expresión lineal en \mathbb{R}^n en la que se escriben todos sus términos de forma explícita:

$$u = u_1x_1 + u_2x_2 + u_3x_3 + \dots + u_nx_n$$

esta puede expresarse concisamente como la sumatoria:

$$u = \sum_{i=1}^n u_i x_i$$

La notación de Einstein obtiene una expresión aún más condensada eliminando el signo de sumatorio y entendiendo que en la expresión resultante un índice indica suma sobre todos los posibles valores del mismo.

$$u = u_i x_i$$

31 Esperanza matemática

En estadística la esperanza matemática (también llamada esperanza, valor esperado, media poblacional o media) de una variable aleatoria X , es el número $\mathbb{E}[X]$ que formaliza la idea de valor medio de un fenómeno aleatorio.

Cuando la variable aleatoria es discreta, la esperanza es igual a la suma de la probabilidad de cada posible suceso aleatorio multiplicado por el valor de dicho suceso. Por lo tanto, representa la cantidad media que se "espera" como resultado de un experimento aleatorio cuando la probabilidad de cada suceso se mantiene constante y el experimento se repite un elevado número de veces. Cabe decir que el valor que toma la esperanza matemática en algunos casos puede no ser "esperado" en el sentido más general de la palabra - el valor de la esperanza puede ser improbable o incluso imposible.

Para una variable aleatoria discreta con valores posibles x_1, x_2, \dots, x_n y sus probabilidades representadas por la función de probabilidad $p(x_i)$ la esperanza se calcula como ejemplo:

$$\begin{aligned}\mathbb{E}[X] &= x_1 p(X = x_1) + \dots + x_n p(X = x_n) \\ &= \sum_{i=1}^n x_i p(x_i)\end{aligned}$$

Para una variable aleatoria absolutamente continua, la esperanza se calcula mediante la integral de todos los valores y la función de densidad $f(x)$:

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x) dx$$

La definición general de esperanza se basa, como toda la teoría de la probabilidad, en el marco de la teoría de la medida y se define como la siguiente integral:

$$\mathbb{E}[X] = \int_{-\Omega} X dP$$

La esperanza también se puede simbolizar con $\mu = \mathbb{E}[X]$

32 El último teorema de Fermat y los Simpsons

Empecemos con el teorema de Fermat; es uno de los teoremas más famosos en la historia de la matemática. Utilizando la notación moderna, se puede enunciar de la siguiente manera:

Si n es un número entero mayor que 2, entonces no existen números enteros positivos x , y y z , tales que se cumpla la igualdad:

$$x^n + y^n = z^n$$

En los simpsons en 1995 se emite un capítulo, en el que, entre otras cosas, Homero salta a otra dimensión y, a través de una puerta, aparece en el mundo humano. Mientras Homer está en ese mundo entre lo animado y lo humano aparece en imagen lo siguiente:

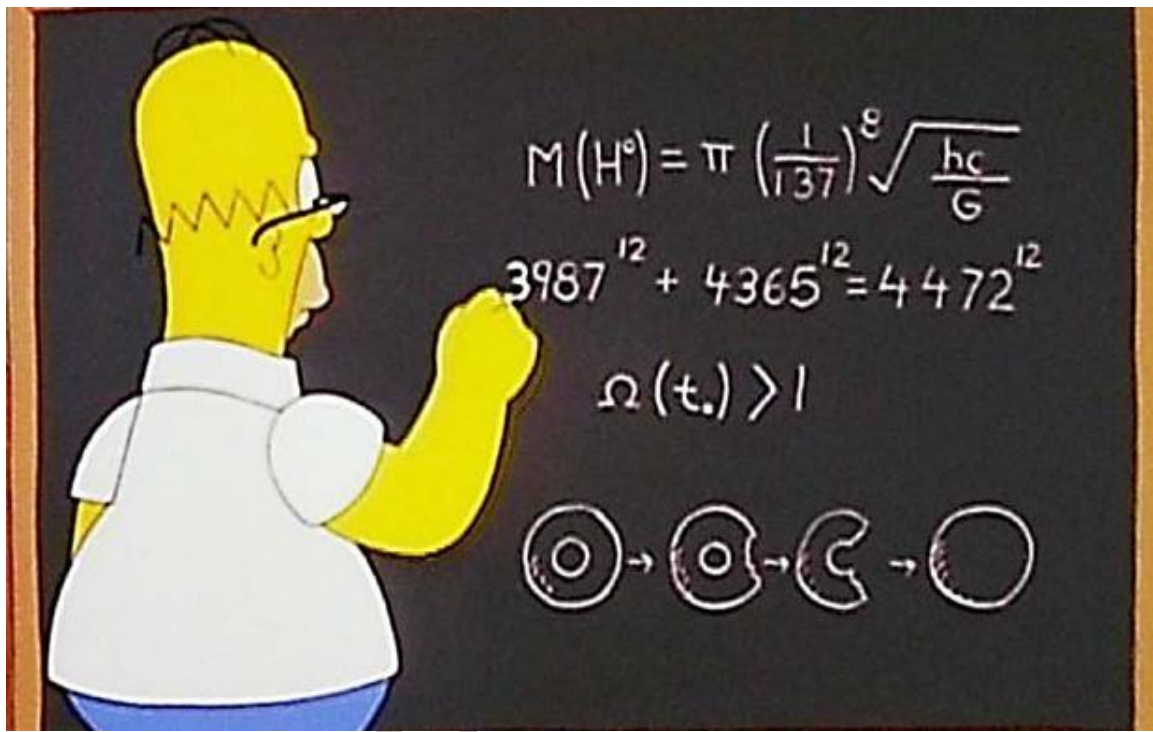


Fig. 46: Primer capítulo de los Simpson donde aparece el teorema de Fermat

Homero en primer plano y una igualdad a su espalda:

$$1782^{12} + 1841^{12} = 1922^{12}$$

Si la igualdad la verificamos en algún dispositivo de baja capacidad como una calculadora, esto, evidentemente, demuestra que la igualdad es cierta. Pues no, esa igualdad no es cierta. Y no es cierta por lo siguiente: el último teorema de Fermat fue demostrado en ese mismo año, 1995. Incógnita resuelta, la igualdad no es cierta y su inclusión en ese capítulo es simplemente una broma de los creadores de la serie. Pero, por qué en una calculadora simple sí se cumple?. Pues muy sencillo. Vamos a ver concretamente los resultados de cada una de las operaciones orrojados por una dispositivo y software más potente, por ejemplo Mathematica:

$$1782^{12} = 1025397835622633634807550462948226174976$$

$$1841^{12} = 1515812422991955541481119495194202351681$$

$$1782^{12} + 1841^{12} = 2541210258614589176288669958142428526657$$

$$1922^{12} = 2541210259314801410819278649643651567616$$

Como podemos ver la suma de las dos primeras potencias y la tercera se parecen mucho. De hecho coinciden en las 9 primeras cifras, y si redondeamos los dos números a 10 cifras son iguales. Esa es la clave, David X. Cohen, uno de los guionistas y productores de Futurama y Los Simpson's había escrito un programa que buscaba combinaciones de x , y , z y n que parecían no cumplir el último teorema de Fermat en una calculadora.

Pero Cohen no se rindió. Y tiempo después, en 1998 apareció la siguiente imagen en otro capítulo de los simpsons:

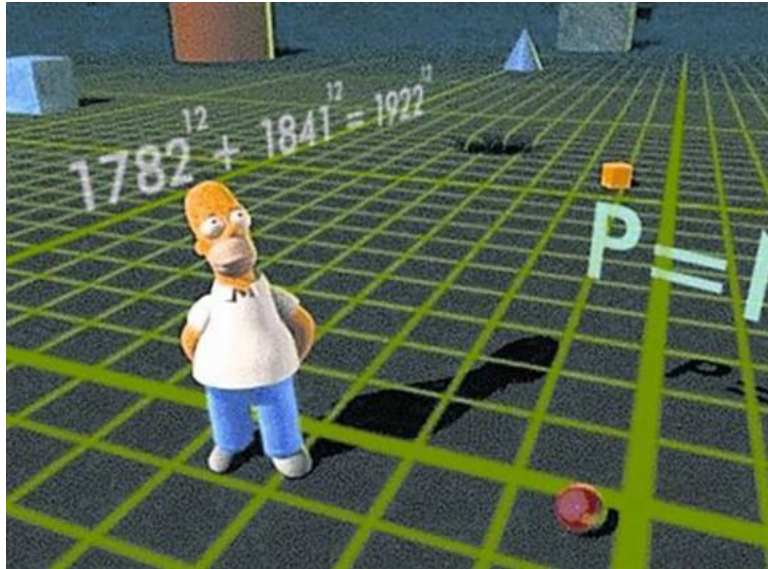


Fig. 47: Segundo capítulo de los Simpson donde aparece el teorema de Fermat

Donde:

$$3987^{12} = 16134474609751291283496491970515151715346481$$

$$4365^{12} = 47842181739947321332739738982639336181640625$$

$$3987^{12} + 4365^{12} = 63976656349698612616236230953154487896987106$$

$$4472^{12} = 63976656348486725806862358322168575784124416$$

Nuevamente si redondeamos a 10 cifras ambos números obtenemos el mismo resultado. Y eso es lo que hace una calculadora de bajos recursos: redondea hasta donde, digamos, puede trabajar. Es el problema de las calculadoras: cuando trabajamos con un cierto número de cifras nos salimos del rango máximo de trabajo de la misma y eso produce un error que la máquina soluciona redondeando. De todas formas, aun conociendo ese error, es complicadísimo encontrar un ejemplo como los dos que encontró Cohen. Y es que ser Licenciado en Física por la Universidad de Harvard además de tener un Máster en Ciencias Computacionales por la Universidad de Berkeley tenía que servir de algo.

33 Máquina Enigma

La criptografía, palabra que procede del griego *krypto* (oculto) y *graphos* (escritura), es la disciplina científica que se encarga del cifrado y descifrado de mensajes, es decir, enmascarar mensajes mediante un algoritmo de ofuscación que, además, debe permitir devolver el mensaje a su estado original. La criptografía, históricamente, ha estado vinculada a las campañas militares y los secretos de los gobiernos, de hecho, se dice que la primera técnica criptográfica se usó en el siglo V a.C en la antigua Grecia e incluso, posteriormente, por Julio César en sus campañas militares. Pero si hay un período donde la criptografía tomó una vital importancia fue en la Segunda Guerra Mundial, en la que el cifrado y el descifrado de códigos se convirtió en otro frente más de lucha.

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

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

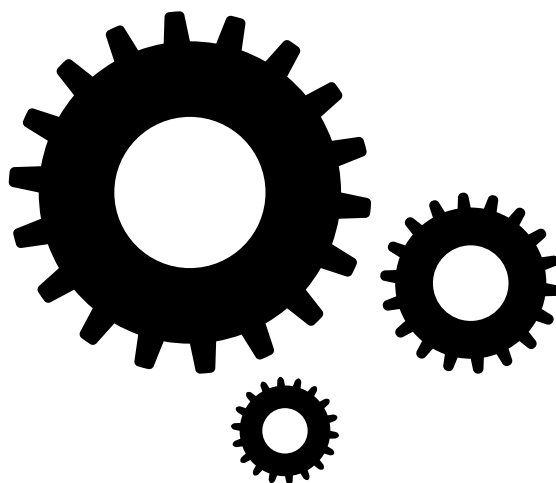
¿En qué consistía la máquina Enigma?

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

Además, el sistema contaba con 6 cables de conexión que también permitían introducir modificaciones dado que podrían conectarse a 26 lugares (representando a las 16 letras del alfabeto de Enigma) lo que producía 100.391.791.500 maneras distintas de conectar los cables que unidos a los 105.456 alfabetos arrojaba 3.283.883.513.796 .974.198.700.
882.069.882.752.878.379.955.261.095.623.685.444.055.315.226.006.433.616.627.409.666.933.182.371.154.
802.769.920.000.000.000 posibilidades distintas de codificación.

En 1933, Alemania nacionalizó la compañía Enigma Chiffiermaschinen AG y pasó a equipar a todo el ejército Alemán que utilizó estas máquinas de cifrado, a las que le añadieron un cuarto cilindro para complicar aún descifrar de los mensajes. Uno de los primeros casos conocidos de uso de Enigma fue durante la Guerra Civil Española donde fueron utilizadas por el Ejército Español, que contaba con el apoyo de la Alemania nazi.

Durante la Segunda Guerra Mundial, Alemania contaba con una enorme ventaja porque el código de Enigma era, prácticamente, indescifrable; además, el ejército alemán cambiaba cada día el código a usar, de tal forma que los Aliados contaban con un único día para descifrarlo porque, al otro día, se volvía a cambiar (algo que enviaban codificando la clave del día siguiente durante la transmisión codificada del día). El cifrado de comunicaciones alemanas mantuvo en jaque a los aliados durante gran parte de la Guerra puesto que, en todos los frentes, se usaba Enigma para codificar las comunicaciones y, además, cada ejército usaba códigos distintos.



Y cómo se logró vencer a Enigma? La máquina fue vencida gracias a varios factores:

- Estas databan de 1923 y eran modelos comerciales que se distribuyeron por todo el mundo. Si bien es cierto que las máquinas se fueron complicando en sus versiones militares, el principio de funcionamiento ya era conocido.
- La codificación de un mensaje en Enigma obligaba a los operadores a introducir 3 letras, dos veces, al iniciar el mensaje, una especie de bandera. La Luftwaffe no modificaba esta secuencia y, por tanto, era un patrón que siempre se repetía y que fue aprovechado por Marian Rejewski que fue capaz de descifrar el código gracias a unas máquinas electromecánicas denominadas bombas criptológicas que eran unas máquinas Enigma de procesamiento en paralelo y que buscaban las codificaciones posibles. Los criptógrafos polacos trabajaron junto a los británicos en Bletchley Park para descifrar los códigos alemanes (en este equipo se encontraba Alan Turing).
- El 9 de mayo de 1941 la Royal Navy capturó el submarino alemán U-110 (a cuya tripulación del fallaron las cargas explosivas para hundirlo) y pudo hacerse con una máquina Enigma y con el preciado libro de claves. Esta captura se mantuvo en secreto y se hizo creer a la opinión pública que el submarino había sido hundido, para que las claves no fuesen cambiadas.

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

La historia de Enigma es fascinante así como su funcionamiento (que podéis probar en este simulador). Creo que es una maravilla tecnológica y, en el fondo, de un funcionamiento extremadamente avanzado para la época en la que se diseñó. Tanto es así que gracias a la criptografía durante la Segunda Guerra Mundial nacieron los primeros computadores que, precisamente, se destinaron a descifrar códigos.

Hoy en día Enigma sigue siendo un objeto de estudio y uno de los elementos que más expectación genera en cualquier exposición sobre la Segunda Guerra Mundial.

34 Ventaja

A la hora de poder determinar el origen etimológico del término ventaja tenemos que dejar patente que este se encuentra en un vocablo francés: *avantage*, que, a su vez, emana del adverbio *avant*, que puede traducirse como antes y que deriva del vocablo latino *abante*.

Una ventaja es una superioridad o una mejoría de algo o alguien respecto de otra cosa o persona. Puede definirse como una condición favorable que algo o alguien tiene. Por ejemplo: El equipo local tiene ventaja ya que su rival cuenta con cuatro jugadores lesionados, La ventaja de esta heladera es que consume un 15% menos de energía en comparación a los modelos similares de otras marcas, Me molesta que hayas querido sacar ventaja de la situación. El concepto de ventaja es habitual en el deporte o en los juegos con tanteador. La ventaja es una situación de supremacía que tiene un jugador o un equipo hasta que puede concretar el triunfo. Si un equipo de fútbol gana 3 a 2 a su rival, puede decirse que tiene una ventaja de un gol.

De igual modo, dentro del ámbito deportivo, suele emplearse el concepto de ventaja para referirse a los puntos que un equipo saca a otro en la tabla de clasificación de una competición determinada. Así, por ejemplo, podríamos expresar la siguiente frase: El F.C. Barcelona lleva ocho puntos de ventaja sobre el Real Madrid.

Todo ello sin olvidar que existe el término jugador de ventaja. Este se utiliza para referirse a aquella persona que participa en la partida de un juego concreto o en una competición y que realiza todo tipo de trampas para poder proclamarse ganador. En concreto, se considera que aquel no sólo utiliza la cautela y la astucia sino también diversos engaños para conseguir su propósito. La noción también se utiliza en la economía y las finanzas. La ventaja comparativa es aquella que disfruta un país sobre otro para la elaboración de un producto, ya que está en condiciones de producir a menor costo. Un país como Argentina tiene ventajas comparativas para la producción agropecuaria ya que el entorno natural facilita dichas labores. La ventaja competitiva, por otra parte, es una ventaja que tiene una empresa frente a su competencia. Dicha ventaja puede ser el valor de marca, una patente tecnológica, la capacidad de sus recursos humanos o una protección estatal, por ejemplo. Es vital tener claro que para que cualquier ventaja sea considerada como competitiva y pueda ser eficaz a la par que eficiente debe calificarse o identificarse por las siguientes señas de identidad: se puede aplicar a varias situaciones diferentes que tengan lugar en el seno del mercado, debe estar totalmente ajustada a las normativas legales existentes, tiene que ser posible mantenerla y además debe marcar una considerable diferencia con respecto a la que ofrezca la competencia. Concretamente entre los tipos de ventajas competitivas que puede tener una empresa determinada sobre otra están las cualidades superiores de sus productos o la buena reputación que aquella tenga. Cabe destacar que la idea de aprovechar o utilizar las ventajas es aceptada en el marco de una competencia sana. En cambio, si alguien intenta sacar ventaja de una confusión o un momento complejo, la actitud será condenada a nivel social.

El término ventaja se utiliza para designar la diferencia a favor que una persona ostenta respecto de otra, es decir, no podrá existir la ventaja nunca sino se encuentra esta misma contrapuesta y en estrecha relación con una situación similar que experimenta otra persona.

La ventaja se puede obtener en diversos órdenes y ámbitos de la vida, por ejemplo en el campo laboral la ventaja puede conseguirse luego de haber observado un buen desempeño en alguna tarea que nos encomendaron a mí y a otro compañero que se desempeña en mi misma área y que claro, requería una resolución inmediata y yo logré sobresalir en la resolución, en este caso obtuve una ventaja respecto de mi colega y además conseguí un ascenso, esto también podría leerse como una ventaja.

Antes les decía que una ventaja es algo que se puede observar en varios órdenes de la vida y si me tengo que poner a pensar en el ámbito en el cual esta se encuentra muy presente a todo momento es en el ambiente del deporte y se la entiende como el margen que un jugador o equipo le deja a su adversario. Es muy común que allí y en oportunidad de algún partido de fútbol o de tenis, especialmente en este último caso, escuchemos a los comentaristas decir que tal o cual jugador tienen ventaja sobre el saque respecto de su adversario o en el caso del fútbol también, cuando un equipo se impone en resultado y en juego a otro se suele decir que el equipo A le sacó una contundente ventaja al equipo B.

Pero también, si nos mudamos al mundo del marketing y de la dirección estratégica nos encontramos con que la ventaja ocupa un lugar preponderante por ser el bien tan ansiado que toda empresa intentará sacarle a su competencia, ya que al igual que en el deporte, la ventaja, es el margen de diferencia a favor que siempre una empresa intentará sacarle a su más directo competidor.

En tanto, para lograr esta ventaja que el marketing llama precisamente ventaja competitiva será necesario que esta observe las siguientes características: única, imposible de imitar, posible de mantener en el tiempo.

Circunstancia o condición a favor. Lo que da superioridad en cualquier cosa. Circunstancia a favor. Provecho, beneficio.

35 Examen

Ejercicio 1.- En una clase de 10 alumnos van a distribuirse 3 premios. Averiguar de cuántos modos puede hacerse si:

- a) los premios son diferentes
- b) los premios son iguales

Solución:

- a) Si los premios son diferentes importa el orden con que se reparten los regalos por lo tanto

$$P_{10,3} = \frac{10!}{(10-3)!} = \frac{10 \cdot 9 \cdot 8 \cdot 7!}{7!} = 720$$

maneras de distribuir los premios si son diferentes

- b) Si los premios son iguales

$$C_{10,3} = \frac{10!}{3!(10-3)!} = \frac{10 \cdot 9 \cdot 8 \cdot 7!}{3! \cdot 7!} = \frac{10 \cdot 9 \cdot 8}{3 \cdot 2 \cdot 1} = 120$$

Ejercicio 2.- ¿De cuántas maneras pueden colocarse en línea nueve bolas de las que 4 son blancas, 3 amarillas y 2 azules?

Solución:

Disponemos de 9 bolas, m=9 Formamos grupos de 9 bolas, n=9

Observaciones:

- Influye el orden de los elementos
- Las bolas del mismo color son iguales

Utilizamos la fórmula de permutaciones con repetición de 9 elementos (m=n) tomados de 4 en 4 (a), de 3 en 3 (b) y de 2 en 2 (c)

$$P_n^{a,b,c} = \frac{n!}{a!b!c!}$$

$$P_9^{4,3,2} = \frac{9!}{4!3!2!} = \frac{9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4!}{4! \cdot 3! \cdot 2! \cdot 1 \cdot 2} = \frac{2520}{2} = 1260$$

Ejercicio 3.- ¿Cuántos números de 4 dígitos se pueden formar con las cifras 0,1,...,9

- a) permitiendo repeticiones;
- b) sin repeticiones;
- c) si el último dígito ha de ser 0 y no se permiten repeticiones?

Observaciones

- Para que un número sea de 4 dígitos su primer dígito debe ser distinto de cero

Solución

- a) Por lo tanto hay nueve posibilidades para el primer dígito y 10 para los siguientes 3

Por lo que queda $9 \cdot 10^3 = 9000$ números posibles

- b) En este caso hay 9 posibilidades para el primero, 9 para el segundo, 8 para el tercero y 7 para el cuarto no hay repeticiones $9 \cdot 9 \cdot 8 \cdot 7 = 4536$

- c) Fijamos el último dígito y no hay repeticiones $9 \cdot 8 \cdot 7 \cdot 1 = 504$ números

Ejercicio 4.- En una fábrica hay dos máquinas. La máquina 1 realiza el 60% de la producción total y la máquina 2 el 40 %. De su producción total, la máquina 1 produce 3% de material defectuoso, la 2 el 5%. El asunto es que se ha encontrado un material defectuoso, cuál es la probabilidad de que este material defectuoso provenga de la máquina 2?

Solución: Sea M_1 el evento La máquina 1 produjo el material escogido, M_2 en evento La máquina 2 produjo el material escogido y finalmente sea D el evento El material escogido es defectuoso. El problema es encontrar $P(M_2|D)$ y observamos que la información que tenemos es $P(D|M_2)$. Por el teorema de Bayes tenemos entonces que

$$\begin{aligned}
 P(M_2|D) &= \frac{P(D|M_2)P(M_2)}{P(D|M_1)P(M_1) + P(D|M_2)P(M_2)} \\
 &= \frac{\frac{5}{100} \cdot \frac{40}{100}}{\frac{3}{100} \cdot \frac{60}{100} + \frac{5}{100} \cdot \frac{40}{100}} \\
 &= \frac{10}{19}
 \end{aligned}$$

Ejercicio 5.- Una caja contiene 8 bolas rojas, 3 blancas y 9 azules. Si se sacan 3 bolas al azar, determina la probabilidad de que:

- las 3 sean rojas;
- las 3 sean blancas;
- 2 sean rojas y 1 blanca;
- al menos 1 sea blanca;
- sea 1 de cada color;
- salgan en el orden rojo, blanca y azul.

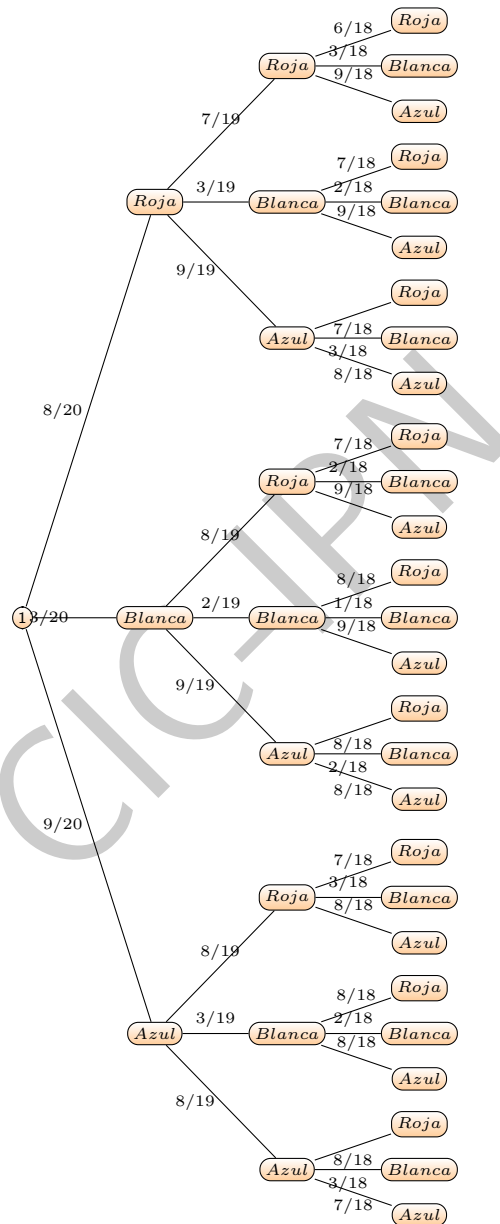


Fig. 48: Diagrama de árbol del ejercicio 5

Solución:

Hay 20 bolas de diferentes colores en total en la caja:

a) La probabilidad de que al extraer la primera bola esta sea roja es: $\frac{8}{20}$, para la segunda bola es $\frac{7}{19}$ debido a que ya se saco una bola roja y solo quedan 19 en la caja y para la tercera es $\frac{6}{18}$.

Por lo tanto

$$P(3 \text{ rojas}) = \left(\frac{8}{20}\right) \left(\frac{7}{19}\right) \left(\frac{6}{18}\right) = \frac{336}{6840} = \frac{14}{285}$$

b) De acuerdo al inciso anterior

$$P(3 \text{ blancas}) = \left(\frac{3}{20}\right) \left(\frac{2}{19}\right) \left(\frac{1}{18}\right) = \frac{6}{6840} = \frac{1}{1140}$$

c) Hay tres formas diferentes de que salgan 2 bolas rojas y 1 blanca por lo que:

$$P(2 \text{ rojas } 1 \text{ blanca}) = 3 \left(\frac{8}{20}\right) \left(\frac{7}{19}\right) \left(\frac{3}{18}\right) = \frac{504}{6840} = \frac{7}{95}$$

d) Considerando que se puede obtener 1 blanca, 2 blancas o 3 blancas queda:

$$\begin{aligned} P(\text{al menos } 1 \text{ blanca}) &= 6P(1 \text{ roja } 1 \text{ blanca } 1 \text{ azul}) + 3P(2 \text{ rojas } 1 \text{ blanca}) + 3P(2 \text{ azules } 1 \text{ blanca}) \\ &\quad + 3P(2 \text{ blancas } 1 \text{ roja}) + 3P(2 \text{ blancas } 1 \text{ azul}) + P(3 \text{ blancas}) \\ &= 6 \left(\frac{8}{20}\right) \left(\frac{3}{19}\right) \left(\frac{9}{18}\right) + 3 \left(\frac{8}{20}\right) \left(\frac{7}{19}\right) \left(\frac{3}{18}\right) \\ &\quad + 3 \left(\frac{9}{20}\right) \left(\frac{8}{19}\right) \left(\frac{3}{18}\right) + 3 \left(\frac{3}{20}\right) \left(\frac{2}{19}\right) \left(\frac{8}{18}\right) \\ &\quad + 3 \left(\frac{3}{20}\right) \left(\frac{2}{19}\right) \left(\frac{9}{18}\right) + \left(\frac{3}{20}\right) \left(\frac{2}{19}\right) \left(\frac{1}{18}\right) \\ &= \frac{2760}{6840} = \frac{23}{57} \end{aligned}$$

e) Sea una de cada color, existen 6 formas diferentes y queda:

$$6P(1 \text{ roja } 1 \text{ blanca } 1 \text{ azul}) = 6 \left(\frac{8}{20}\right) \left(\frac{3}{19}\right) \left(\frac{9}{18}\right) = 6 \frac{216}{6840} = \frac{18}{95}$$

f) Salgan en el orden roja, blanca y azul, solo hay un resultado posible ya que el orden si importa.

$$P(1 \text{ roja } 1 \text{ blanca } 1 \text{ azul}) = \left(\frac{8}{20}\right) \left(\frac{3}{19}\right) \left(\frac{9}{18}\right) = \frac{216}{6840} = \frac{3}{95}$$

Ejercicio 6.- Se extrae al azar una carta de una baraja española. Halla las siguientes probabilidades:

- a) Que sea un rey o un as;
- b) Que sea una copa o una figura;
- c) Que no sea figura.

La baraja española contiene 40 cartas repartidas en 4 palos (oros, bastos, espadas y copas).

- a) La probabilidad de que sea un rey o un as

$P(\text{sea un Rey o un As})$

$$(R \cup A) = P(R) + P(A) = \frac{4}{40} + \frac{4}{40} = \frac{8}{40}$$

- b) $P(\text{sea Copa o Figura})$

$$P(C \cup F) = P(C) + P(F) - P(C \cap F) = \frac{10}{40} + \frac{12}{40} - \frac{3}{40} = \frac{19}{40}$$

- c) $P(\text{no sea Figura})$

$$P(\overline{F}) = 1 - P(F) = 1 - \frac{12}{40} = \frac{28}{40}$$

Ejercicio 7.- Sean A y B los sucesos tales que:

$$P(A) = 0.4$$

$$P(\overline{A} \cap B) = 0.4$$

$$P(A \cap B) = 0.1$$

Calcula $P(A \cup B)$ y $P(B)$.

Solución

$$P(B) = P(\overline{A} \cap B) + P(A \cap B) = 0.4 + 0.1 = 0.5$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.4 + 0.5 - 0.1 = 0.8$$

Ejercicio 8.- En la empresa Alimentos Mr Pollo el 20% de los empleados son ingenieros y otro 20% son economistas. El 75% de los ingenieros ocupan un cargo directivo y el 50% de los economistas también, mientras que los demás trabajadores (no ingenieros y no economistas) solamente el 20% ocupa un puesto directivo. ¿Cuál es la probabilidad de que un empleado directivo elegido al azar sea ingeniero?

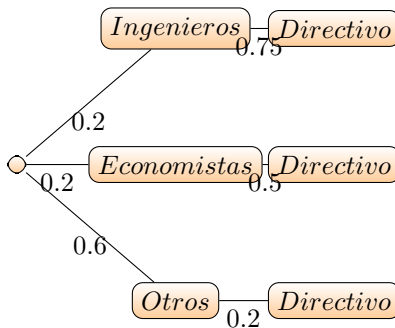


Fig. 49: Diagrama de árbol del ejercicio 8

$$P(\text{Ingeniero}|\text{Directivo}) = \frac{0.2 * 0.75}{0.2 * 0.75 + 0.2 * 0.5 + 0.6 * 0.2}$$

Ejercicio 9.- En la sala de pediatría de un hospital, el 60% de los pacientes son niñas. De los niños el 35% son menores de 24 meses. El 20% de las niñas tienen menos de 24 meses. Un pediatra que ingresa a la sala selecciona un infante al azar.

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

Solución

Se definen los sucesos:

Suceso H: seleccionar una niña.

Suceso V: seleccionar un niño.

Suceso M: infante menor de 24 meses.

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

$$P(M) = P(H) \cdot P(M|H) + P(V) \cdot P(M|V) = 0.6 \cdot 0.2 + 0.4 \cdot 0.35 = 0.26$$

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

$$\begin{aligned} P(H|M) &= \frac{P(H) \cdot P(M|H)}{P(H) \cdot P(M|H) + P(V) \cdot P(M|V)} \\ &= \frac{0.6 \cdot 0.2}{0.6 \cdot 0.2 + 0.4 \cdot 0.35} = \frac{0.12}{0.260} = 0.46 \end{aligned}$$

Ejercicio 10.- Un jugador de Rol, en una partida de Dungeons and Dragons, para salvarse de un conjuro de Raistlin, necesita sacar un 18 en el lanzamiento de los dados. El Dungeon Master le ofrece lanzar tres dados de seis caras o uno de diez junto con uno de ocho.

- ¿En cual de estas dos alternativas es más probable obtener un 18 y salvarse del conjuro? Explique su respuesta
- ¿Sería la respuesta la misma si hubiese que sacar 17 o más para evitar el conjuro?

Solución

a) Para sacar 18 con 3 dados de seis caras tiene que ocurrir el suceso (6,6,6) que tiene una probabilidad $\frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{216}$. Para obtener 18 con un dado de diez caras y otro de ocho tiene que ocurrir el suceso (10, 8) cuya probabilidad es $\frac{1}{10} \cdot \frac{1}{8} = \frac{1}{80}$. Obviamente esta probabilidad es mayor que la anterior.

b) Para obtener 17 o más con los tres dados tiene que ocurrir el suceso:

$(5, 6, 6) \cup (6, 5, 6) \cup (6, 6, 5) \cup (6, 6, 6)$ cuya probabilidad es: $\frac{4}{216}$

Para conseguir el mismo resultado con los dos dados tiene que ocurrir: $(10, 7) \cup (9, 8) \cup (10, 8)$ que tiene probabilidad de $\frac{3}{80}$ que sería también mayor que con los tres dados.

36 Congruencia zeller

La congruencia de Zeller es un algoritmo ideado por Julius Christian Johannes Zeller para calcular el día de la semana de cualquier fecha del calendario.

Para el calendario juliano es:

$$h = \left(q + \left[\frac{(m+1)26}{10} \right] + K + \left[\frac{K}{4} \right] + 5 - J \right) \bmod 7,$$

donde:

h es el día de la semana (0 = sábado, 1 = domingo, 2 = lunes,...),

q es el día del mes,

m es el mes,

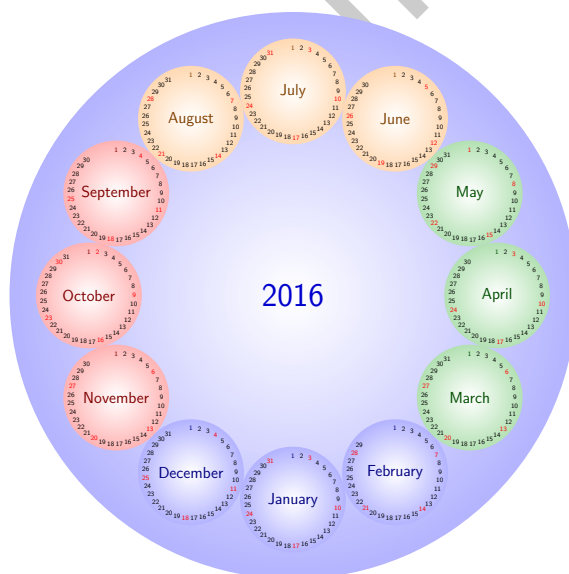
J es la centuria (es realidad año / 100) y

K el año de la centuria (año mod 100).

Enero y febrero se cuentan como meses 13 y 14 del año anterior. Observe, que el 2 de enero del 2013, es m=13; año=2012

Es oportuno recordar que la función mod es el residuo que queda de la división de dos números.

En las implementaciones informáticas en las que el módulo de un número negativo es negativo, la manera más sencilla de obtener un resultado entre 0 y 6 es reemplazar $-2J$ por $+5J$ y $-J$ por $+6J$.



Análisis

Estas fórmulas se basan en la observación de que el día de la semana progresa de una manera predecible basada en cada subparte de esa fecha. Cada término de la fórmula se usa para calcular el desplazamiento necesario para obtener el día correcto de la semana.

Por tanto, para el calendario gregoriano, las diversas partes de esta fórmula pueden entenderse así:

q representa la progresión del día de la semana basada en el día del mes, dado que cada día sucesivo resulta en un desplazamiento adicional de 1 en el día de la semana.

K representa la progresión del día de la semana basada en el año. Suponiendo que cada año tiene 365 días, la misma fecha de cada año sucesivo será desplazada por un valor de **$365 \bmod 7 = 1$** .

Como hay 366 días en cada año bisiesto, esto se debe tener en cuenta añadiendo un día adicional al valor de desplazamiento del día de la semana. Esto se logra añadiendo $\left[\frac{K}{4}\right]$ al desplazamiento. Este término se calcula como un resultado entero. Cualquier resto que pueda haber es descartado.

Usando una lógica similar, se puede calcular la progresión del día de la semana para cada centuria observando que hay 36524 días en una centuria normal, y 36525 en cada centuria divisible por 400. Dado que **$36525 \bmod 7 = 6$** y **$36524 \bmod 7 = 5$** , el término: $\left[\frac{J}{4}\right] - 2J$ refleja esto (de nuevo usando división entera y descartando cualquier resto fraccional). Para evitar los números negativos, este término se puede reemplazar por $5J + \left[\frac{J}{4}\right]$ con un resultado equivalente.

El término $\left[\frac{(m+1)26}{10}\right]$ se puede explicar de la siguiente manera. Zeller observó que, al iniciar cada año el 1 de marzo, el día de la semana de cada mes sucesivo progresaba multiplicando el mes por un valor constante y descartando el resto fraccional.

La función global, **$\bmod 7$** , normaliza el resultado para que se encuentre en el intervalo de 0 a 6, lo que da el índice del día de la semana correcto para la fecha analizada.

La razón por la que la fórmula difiere para el calendario juliano es que este calendario no tiene una regla aparte para las centurias bisiestas y está desplazado con respecto al calendario gregoriano un número fijo de días. Ambas diferencias se pueden tener en cuenta reemplazando el término $\left[\frac{J}{4}\right] - 2J$ por el término $5 - J$, o $5 + 6J$ para evitar números negativos.

Dado que el calendario gregoriano fue adoptado en diferentes momentos en diferentes partes del mundo, la ubicación de un evento es significativa a la hora de determinar el día de la semana correcto de una fecha que tuvo lugar durante este periodo de transición.

36.1 Congruencia Zeller en Mathematica

El siguiente código realiza el algoritmo de congruencia Zeller en Mathematica, la fecha se introduce mediante las variables día, mes y año, una vez definidas estas variables se ejecuta el algoritmo.

```
(*Andrea Viridiana Soto Miranda*)
(*Congruencia Zeller en Mathematica*)
\
(*Probabilidad , procesos aleatorios e inferencia*)
(*CIC-IPN*)

dia = 28;
mes = 2;
año = 2010;
If[ dia == 29 && mes == 2,
  If[ Mod[ año , 4] == 0,
    If[ Mod[ año , 100] == 0,
      If[ Mod[ año , 400] == 0, bis = 0, bis = 1], bis = 0], bis = 1],
  bis = 0]
(* If [Mod[ año,4]==0&&dia==29&&mes==2, If [Mod[ año,100]==0, If [Mod[ año ,\
400]==0, bis=0, bis=1], bis=0], bis=1];*)

If[ bis == 0,
  If[ mes == 1, {m = 13; a = año - 1;},
    If[ mes == 2, {m = 14; a = año - 1;}, {m = mes; a = año;}] ]
J = Quotient[a, 100];
K = Mod[a, 100];
h = Mod[( dia + Quotient[(m + 1)*26, 10] + K + Quotient[K, 4] +
  Quotient[J, 4] - 2*J), 7], h = 7];

If[h == 0, d = sabado];
If[h == 1, d = domingo];
If[h == 2, d = lunes];
If[h == 3, d = martes];
If[h == 4, d = miercoles];
If[h == 5, d = jueves];
If[h == 6, d = viernes];
If[h == 7, d = "año no bisiesto"];
StringForm["El '' de '' de '' fue ''", dia, mes, año, d]
```

Algunos resultados obtenidos son:

- día:9; mes=6; año:1991
El 9 de 6 de 1991 fue domingo
- día:29; mes=2; año:2000
El 29 de 2 de 2000 fue martes
- día:29; mes=2; año:2001
El 29 de 2 de 2000 fue año no bisiesto

36.2 Congruencia Zeller en gcc

```
//Autor: Soto Miranda Andrea Viridiana
//Congruencia Zeller en gcc
//Instituto: CIC-IPN
//Materia: Probabilidad, procesos aleatorios e inferencia
#include<stdio.h>

int zeller(int dia, int mes, int anno){

    int h, K, J;
    //Tomamos en cuenta el caso de
    //Enero y Febrero
    if(mes <= 2){
        mes = mes + 12;
        anno = anno - 1;
    }
    else{
        mes = mes - 2;
    }

    //Calculamos K y J
    K = anno % 100;
    J = anno / 100;

    //Aplicamos la formula
    h = (700 + ((26 * mes - 2) / 10) + dia);
    h = (h + K + (K/4) + ((J/4) + 5*J)) % 7;
    //Regresamos el resultado
    return (h);
}

int main(void){
    //Dias de los meses
    int meses[13] = {0, 31, 28, 31, 30, 31, 30, 31,
                    , 31, 30, 31, 30, 31};

    int dia, mes, ano, aux;
    aux = 0;
    printf("Congruencia_zeller_en_gcc\n");
    do{
        printf("Ingrese el dia: ");
        scanf("%i", &dia);
        printf("Ingrese el mes: ");
        scanf("%i", &mes);
        if(dia < 1)
            printf("Error, el dia no existe\n");
        if(mes < 1 || mes > 12)
            printf("Error, el rango de mes: 1-12\n");
        if(mes > 0 && mes < 13){
            if(mes != 2){
                if(dia > meses[mes]){
                    printf("Error, el mes seleccionado, ");
                    printf("no existe tal dia.\n");
                    aux = 0;
                }
                else{
                    aux = 1;
                }
            }
            else{
                if(dia > 29){
                    printf("Error, el mes seleccionado, ");
                    printf("no existe tal dia.\n");
                }
                else{
                    aux = 1;
                }
            }
        }
    }
    while (dia < 1 || mes < 1 || mes > 12 || aux == 0);
}
```

```
printf(" Ingrese el ano:");
scanf("%i", &ano);

if(mes == 2 && dia == 29){
    if(ano%4 != 0){
        printf(" El ano seleccionado no es bisiestro\n");
        aux = 0;
    }
    if(ano%4 == 0 && ano%100 == 0 && ano%400 != 0){
        printf(" El ano seleccionado no es bisiestro\n");
        aux = 0;
    }
}
}while(aux == 0);

int z = zeller(dia, mes, ano);
char *dia_sem;
switch (z)
{
    case 0:
        dia_sem = "Domingo";
        break;
    case 1:
        dia_sem = "Lunes";
        break;
    case 2:
        dia_sem = "Martes";
        break;
    case 3:
        dia_sem = "Miercoles";
        break;
    case 4:
        dia_sem = "Jueves";
        break;
    case 5:
        dia_sem = "Viernes";
        break;
    case 6:
        dia_sem = "Sabado";
        break;
}

printf (" \nLa fecha corresponde con un %s\n", dia_sem);

return 0;
}
```

37 Simulación de dos dados

Abstract

En este trabajo se realiza un análisis del problema del caballero de Méré para la obtención de un as doble al lanzar dos dados.

El problema consiste en determinar la cantidad de lanzamientos necesarios para lograr que la probabilidad de obtener un as doble (o cualquier figura o número que se muestre en los dados) sea de por lo menos 0.5. Esto es para lograr que al realizar una apuesta sobre este tipo de juego se tenga una probabilidad ventajosa.

Para realizar el análisis se propuso una serie de experimentos y simulaciones, y sus resultados se compararon con un resultado que se obtiene de manera teórica.

Después de realizar las simulaciones, se observa que los valores obtenidos son bastante cercanos a los valores que se esperan de manera teórica, sin embargo no son totalmente exactos, probablemente debido a los generadores de números aleatorios utilizados.

38 Introducción.

En el siglo 17 un famoso apostador, el caballero de Méré, recurrió a la ayuda de Blaise Pascal para obtener una explicación de porqué perdía constantemente.

En esta época, en Francia, se solía apostar por el evento de obtener al menos un as en cuatro tiros de un dado. Una variación de este juego era obtener al menos un doble as al tirar 2 dados 24 veces.

Según el razonamiento del caballero de Méré, obtener dos ases con dos dados es $1/6$ veces menos probable que obtener un as con un solo tiro (lo cual es correcto). Para compensar, el caballero de Méré pensó que, como para obtener un as con un dado se necesitaban cuatro tiradas para tener una probabilidad mayor al 50%, con dos dados el número de tiradas debía de ser multiplicado por seis (haciendo 24 tiradas) para tener un 50% de probabilidad de obtener un as doble. Sin embargo, el perdía constantemente. [1].

39 Análisis Matemático.

39.1 Lanzamiento de un dado.

Para el caso de un dado tenemos la Eq.(20).

$$P(as) = 1 - \frac{5^n}{6} \quad (20)$$

Donde n es el número de lanzamientos realizados. Para $n = 4$ tenemos la Eq.(21).

$$P(as) = 1 - \frac{5^4}{6} = \frac{671}{1296} \approx 0.517 \quad (21)$$

La cual es mayor a 0.5 y por lo tanto genera una apuesta ventajosa.

39.2 Lanzamiento de dos dados.

Para el caso de dos dados tenemos la Eq.(22).

$$P(doble\ as) = 1 - \frac{35^n}{36} \quad (22)$$

Donde n es el número de lanzamientos realizados. Para $n = 24$ tenemos la Eq.(23).

$$P(doble\ as) = 1 - \frac{35^{24}}{36} \approx 0.491 \quad (23)$$

La cual es menor a 0.5 y por lo tanto es una apuesta con desventaja.

Sin embargo, si se hacen 25 lanzamientos tenemos la Eq.(24).

$$P(\text{doble as}) = 1 - \frac{35^{25}}{36} \approx 0.505 \quad (24)$$

La cual es mayor 0.5.

Por lo tanto, observamos que, según los cálculos realizados, con 25 lanzamientos se logra obtener una probabilidad ventajosa al realizar una apuesta en favor de que saldrá por lo menos un as doble.

40 Análisis Experimental.

Para complementar el resultado obtenido analíticamente, realizamos un experimento con el cual se espera confirmar los datos para el lanzamiento de dos dados.

Este experimento se realizará siguiendo el métodos heurístico, el cual se describe a continuación.

40.1 Método Heurístico.

El matemático George Pólya en su libro *How to Solve It* plantea el método heurístico para resolver problemas.

Este consta de cuatro pasos fundamentales:

1. Entender el problema.
2. Hacer un plan para resolver el problema.
3. Poner el plan en marcha.
4. Revisar el trabajo realizado y analizar como podría mejorarse.

40.1.1 Primer principio: Entender el problema.

Para analizar este paso podemos hacernos las siguientes preguntas:

- Qué se necesita encontrar o demostrar?
- Puedes expresar el problema en tus propias palabras?
- Tienes toda la información requerida para resolver el problema?

Estos son algunos ejemplos de preguntas que podemos analizar para comprender el problema, aunque en general cualquier método que nos ayude a comprender mejor el problema es aceptado.

40.1.2 Segundo Principio: Hacer un plan.

Mientras se resuelven mas problemas parecidos, la habilidad para escoger una estrategia adecuada va incrementando.

Algunas de las estrategias que se pueden seguir son las siguientes:

- Hacer una lista ordenado de los pasos a seguir.
- Eliminar diferentes posibilidades.
- Considerar casos especiales.
- Usar razonamiento directo.
- Usar una fórmula.

40.1.3 Tercer Principio: Poner el plan en marcha.

Para que este paso se realice de la mejor manera, es necesario haber realizado un plan adecuado. Si se tiene el plan adecuado, lo solamente se necesita paciencia y cuidado al aplicar la solución planeada.

Si lo que se planeo no funciona, es necesario desecharlo y pensar en una mejor solución.

40.1.4 Cuarto Principio: Revisar y extender.

Tomarse el tiempo para revisar lo que se hizo, lo que funcionó y lo que no funcionó, puede traer muchas ganancias.

Esto nos puede ayudar a mejorar la formas en que resolvemos futuros problemas, si están relacionados con el problema original.

[5]

40.2 Entender el problema.

Siguiendo el método descrito, comenzamos con una breve descripción del problema.

Necesitamos comprobar de manera experimental que los resultados obtenidos analíticamente para el problema del caballero de Méré son válidos.

Esto es, comprobar que para tener una probabilidad de por lo menos 50% de obtener un doble as en el lanzamiento de dos dados, es necesario realizar 25 lanzamientos.

40.3 Hacer un plan.

Para analizar el problema se realizarán los siguientes pasos (se utilizarán dados con números del 1 al 6):

1. Realizar físicamente lanzamientos con dos dados hasta obtener un seis doble. El número de lanzamientos requeridos será registrado y posteriormente se repetirá el experimento 50 veces. Se obtendrá un promedio del número de lanzamientos requeridos para obtener el evento de interés.
2. Se repetirá el primer paso pero en esta ocasión se utilizará una aplicación móvil para simular en lanzamiento de los dados.
3. Se hará una simulación en computadora con un gran número de repeticiones del experimento para observar el comportamiento con esta cantidad de datos.

40.4 Poner el plan en marcha.

40.4.1 Experimento con dados reales.

Se realizan repeticiones del experimento tirando dos dados hasta obtener un seis doble y registrando el número de tiradas que fue requerido en cada caso.

En la tabla (1) se registran los valores obtenidos en el experimento. La columna Exp. representa las iteraciones del experimento y la columna n representa el número de tiradas requeridas en esa iteración para obtener el seis doble.

Exp.	n		Exp.	n		Exp.	n		Exp.	n		Exp.	n
1	16		11	9		21	45		31	60		41	54
2	18		12	14		22	1		32	36		42	42
3	22		13	5		23	18		33	13		43	7
4	25		14	62		24	20		34	1		44	86
5	27		15	52		25	20		35	53		45	40
6	80		16	2		26	66		36	36		46	17
7	113		17	26		27	1		37	68		47	22
8	29		18	56		28	1		38	86		48	40
9	1		19	99		29	36		39	6		49	1
10	12		20	55		30	15		40	51		50	6

Table 1: Resultados de experimento con dados reales.

De estos datos obtenemos que se requirió de un promedio de **33.42** tiradas para obtener un seis doble.

40.4.2 Experimento con dados virtuales.

Se realizan repeticiones del experimento, utilizando una aplicación que simula el lanzamiento de dos dados, hasta obtener un seis doble y registrando el número de lanzamientos requeridos en cada caso.

En la tabla (2) se registran los valores obtenidos en el experimento. La columna Exp. representa las iteraciones del experimento y la columna n representa el número de tiradas requeridas en esa iteración para obtener el seis doble.

Exp.	n		Exp.	n		Exp.	n		Exp.	n		Exp.	n
1	74		11	63		21	6		31	1		41	69
2	7		12	46		22	72		32	29		42	8
3	26		13	121		23	10		33	12		43	3
4	80		14	18		24	34		34	14		44	8
5	32		15	12		25	18		35	23		45	16
6	47		16	56		26	24		36	10		46	38
7	10		17	15		27	72		37	17		47	40
8	50		18	12		28	16		38	24		48	19
9	8		19	38		29	33		39	17		49	65
10	5		20	10		30	26		40	18		50	43

Table 2: Resultados de experimento con dados virtuales.

De estos datos obtenemos que se requirió de un promedio de **30.3** tiradas para obtener un seis doble.

40.4.3 Simulación en Mathematica.

Se utilizó el siguiente programa en Mathematica para obtener los resultados para un gran número de iteraciones del experimento.

```
(* Simulación de dos dados en mathematica *)
\
(* Probabilidad, procesos aleatorios e inferencia *)
(* CIC-IPN *)
k = 0;
limite = 100;
c = Table[0, {limite}];
While[k < limite, i = 1;
  While[i <= 150, a = RandomInteger[{1, 6}];
    (* Print[a]; *) b = RandomInteger[{1, 6}];
    (* Print[b]; *) If[a == 6 && b == 6, Break[], i++] (* Print[
i]; *) k++;
```

```

c[[k]] = i
Histogram[c, limite]
d = StandardDeviation[c];
e = Mean[c];
Plot[PDF[NormalDistribution[e, d], c], {c, -100, 200}]

```

Variando la variable *limite* para diferentes números de iteraciones, se obtienen los siguientes datos:

- Para 10 repeticiones del experimento, fig.(50,51).

Valor promedio: 38.4

Desviación estándar: 36.936

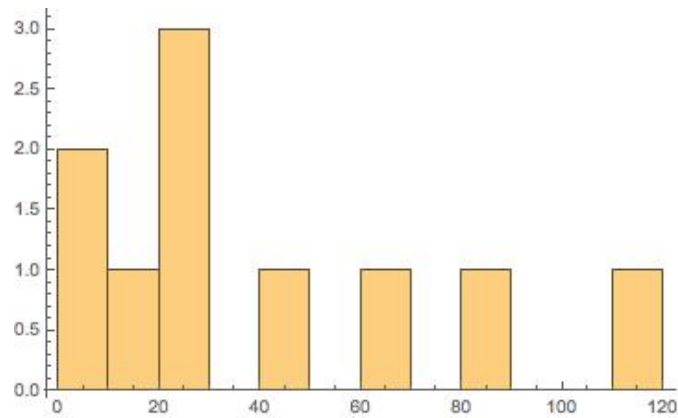


Fig. 50: Resultado de 10 lanzamientos.

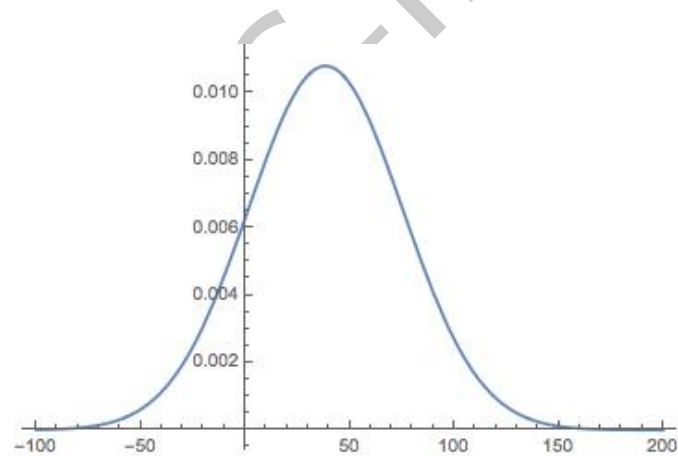


Fig. 51: Distribución normal de 10 lanzamientos.

- Para 100 repeticiones del experimento, fig.(52,53).

Valor promedio: 35.46

Desviación estándar: 30.6566

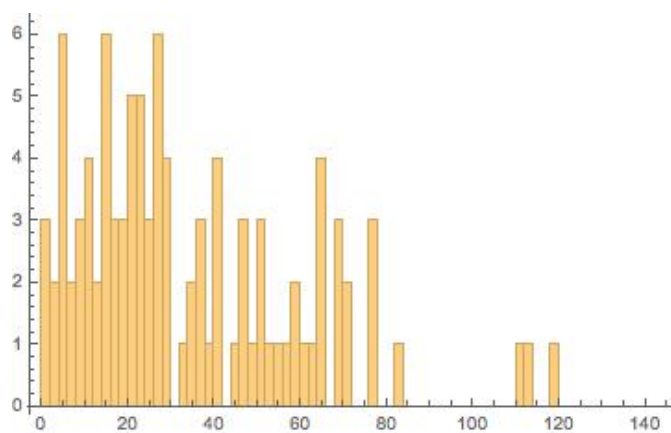


Fig. 52: Resultado de 100 lanzamientos.

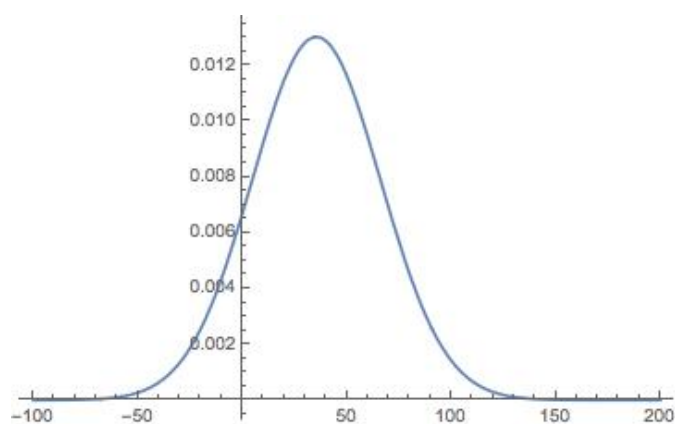


Fig. 53: Distribución normal de 100 lanzamientos.

- Para 1,000 repeticiones del experimento, fig.(54,55).

Valor promedio: 34.627

Desviación estándar: 35.1388

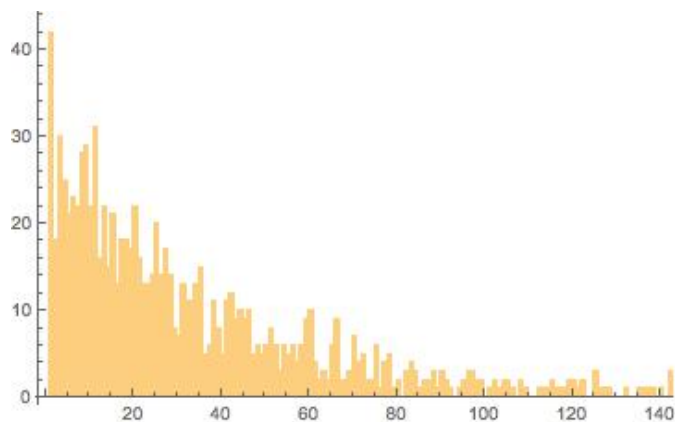


Fig. 54: Resultado de 1,000 lanzamientos.

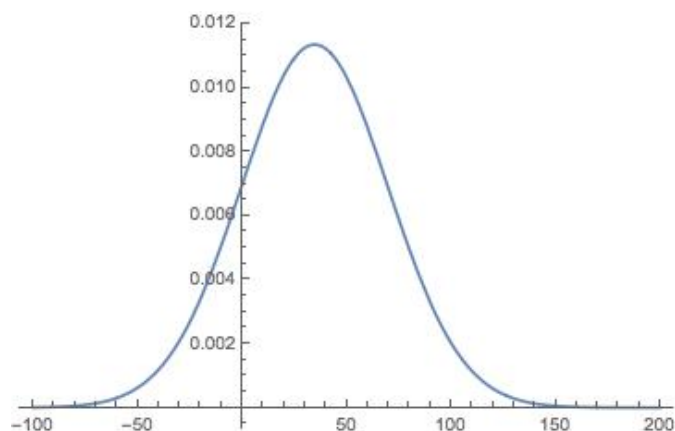


Fig. 55: Distribución normal de 1,000 lanzamientos.

- Para 10,000 repeticiones del experimento, fig.(56,57).

Valor promedio: 35.7256

Desviación estándar: 34.5559

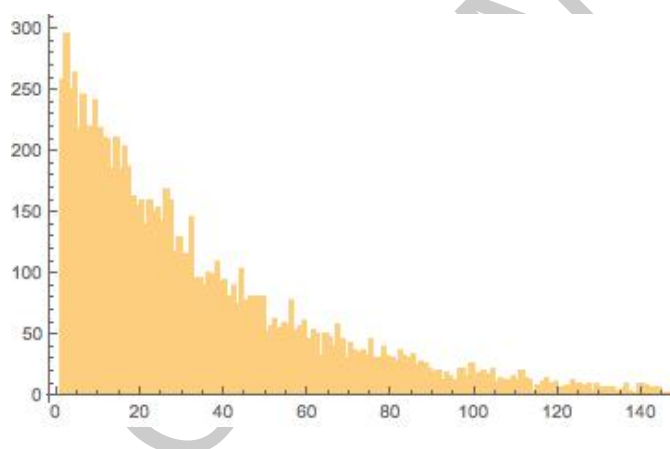


Fig. 56: Resultado de 10,000 lanzamientos.

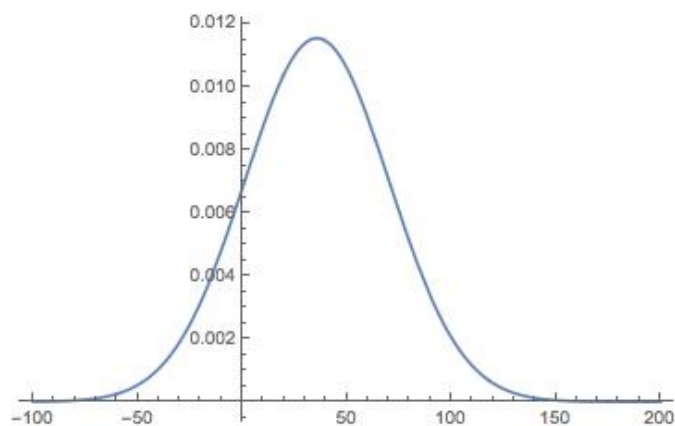


Fig. 57: Distribución normal de 10,000 lanzamientos.

En los resultados obtenidos podemos observar que el promedio de tiradas requeridas para obtener el doble seis se encontró

cerca de los **35** lanzamientos en cada uno de los experimentos.

40.4.4 Simulación en ROOT.

Se utilizó el siguiente programa en ROOT para simular el experimento con un número diferente de iteraciones.

```
//Autor: Soto Miranda Andrea Viridiana
//Simulación de dos dados
//Instituto: CIC-IPN
//Materia: Probabilidad, procesos aleatorios e inferencia
void simulacion_doble_6(){
    TRandom3 *r = new TRandom3(0);
    TCanvas *c1 = new TCanvas("lanzamientos_doble_6"
                              , "Lanzamientos"
                              , 20
                              , 10
                              , 700
                              , 700);

    c1->SetGrid();

    TCanvas *c2 = new TCanvas("distribucion_doble_6"
                              , "Distribucion"
                              , 800
                              , 10
                              , 700
                              , 700);

    c2->SetGrid();

    TH1* hdado = new TH1("Datos"
                        , "Lanzamientos para obtener un doble_6"
                        , 100
                        , 0
                        , 50);

    hdado->SetLineColor(225);
    hdado->SetLineWidth(4);
    hdado->GetXaxis()->SetTitle("Lanzamientos");
    hdado->GetYaxis()->SetTitle("Frecuencia");

    int dado1 = 0;
    int dado2 = 0;
    int conteo = 0;
    int repeticiones = 1000;
    for(int i = 0; i < repeticiones; i++){
        while((dado1 != 6) || (dado2 != 6)){
            dado1 = (r->Integer(6))+1;
            dado2 = (r->Integer(6))+1;
            conteo++;
        }
        hdado->Fill(conteo);
        conteo = 0;
        dado1 = 0;
        dado2 = 0;
    }
    c1->cd();
    hdado->Draw();
    double mean, dev;
    mean = hdado->GetMean();
    dev = hdado->GetStdDev();
    cout << mean << endl;
    cout << dev << endl;
    c2->cd();
    TF1 *dis = new TF1("Gauss", "gaus", -100, 100);
    dis->SetParameters(6, mean, dev);
    dis->SetLineColor(225);
    dis->SetLineWidth(4);
    TString s;
    s = "Distribucion de los lanzamientos;";
    s += "Lanzamientos; Cara del dado";
    dis->SetTitle(s);
    dis->Draw();
}
```

}

Cambiando el valor de la variable *repeticiones* se obtienen los siguientes resultados:

- Para 10 repeticiones del experimento, fig.(58, 59).

Valor promedio: 17.67

Desviación estándar: 9.695

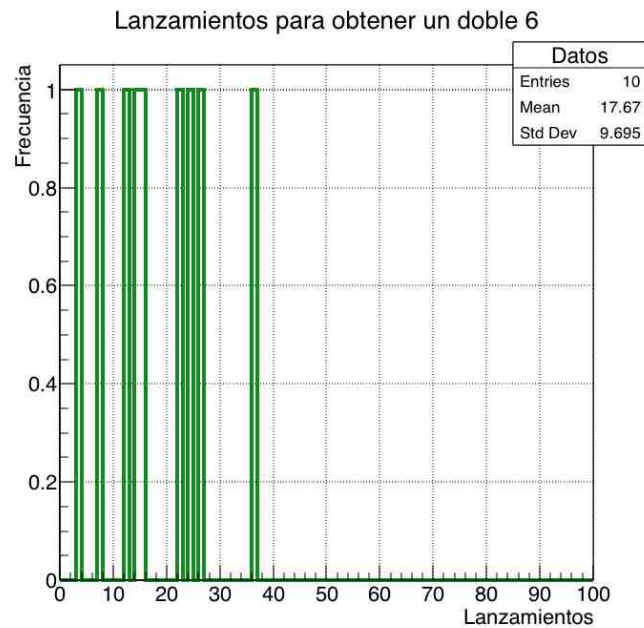


Fig. 58: Resultado de 10 lanzamientos.

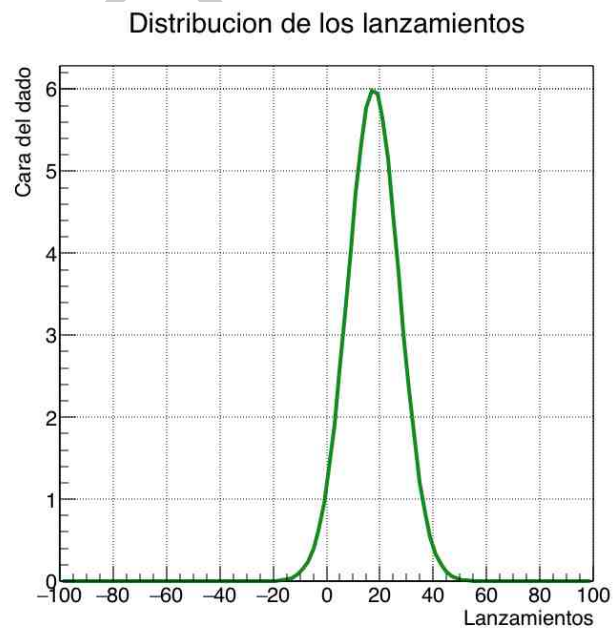


Fig. 59: Distribución normal de 10 lanzamientos.

- Para 100 repeticiones del experimento, fig.(60, 61).

Valor promedio: 25.96

Desviación estándar: 23.395

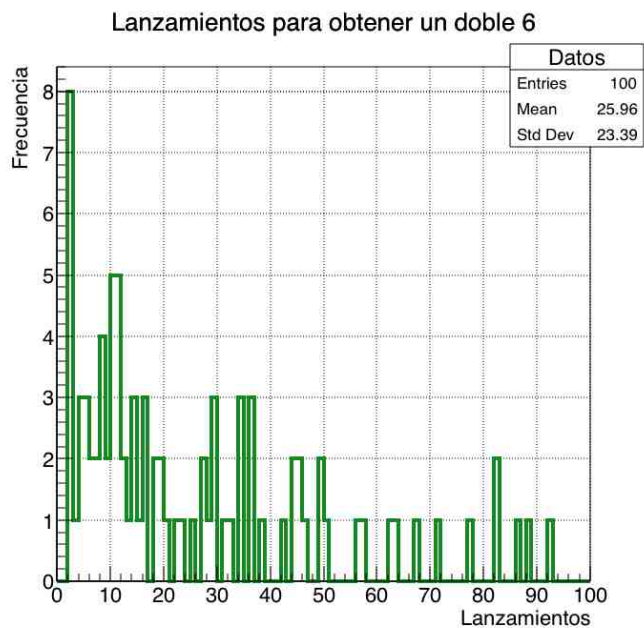


Fig. 60: Resultado de 100 lanzamientos.

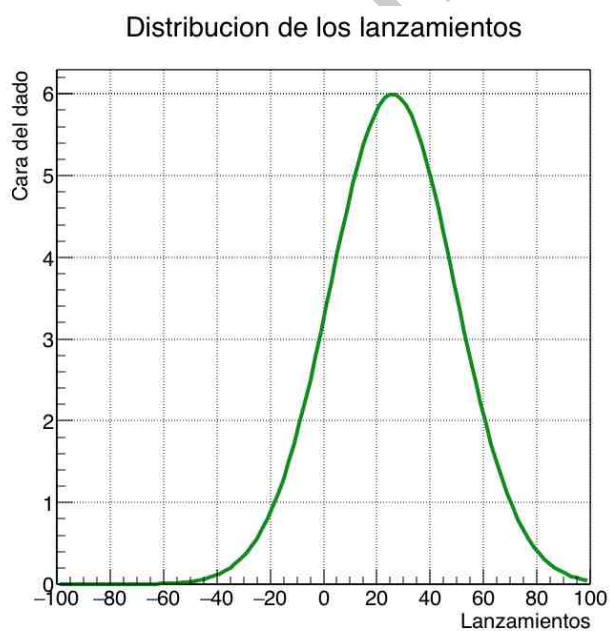


Fig. 61: Distribución normal de 100 lanzamientos.

- Para 1,000 repeticiones del experimento, fig.(62, 63).

Valor promedio: 29.96

Desviación estándar: 24.89

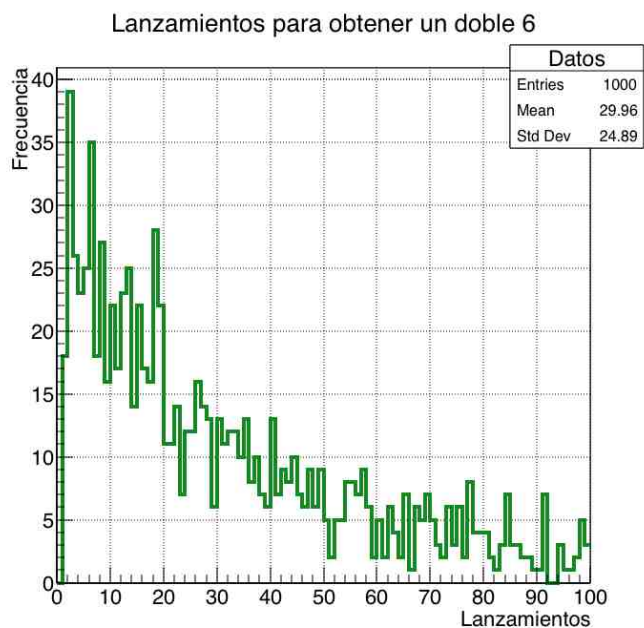


Fig. 62: Resultado de 1,000 lanzamientos.

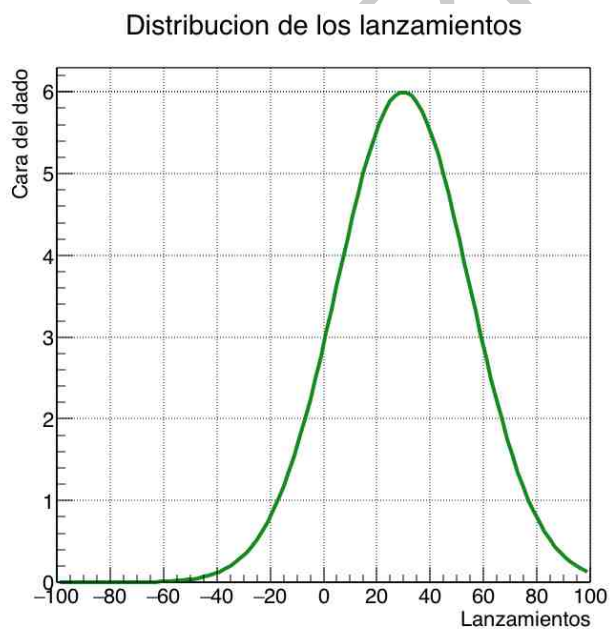


Fig. 63: Distribución normal de 1,000 lanzamientos.

- Para 1,000,000 repeticiones del experimento, fig.(64, 65).

Valor promedio: 29.5

Desviación estándar: 24

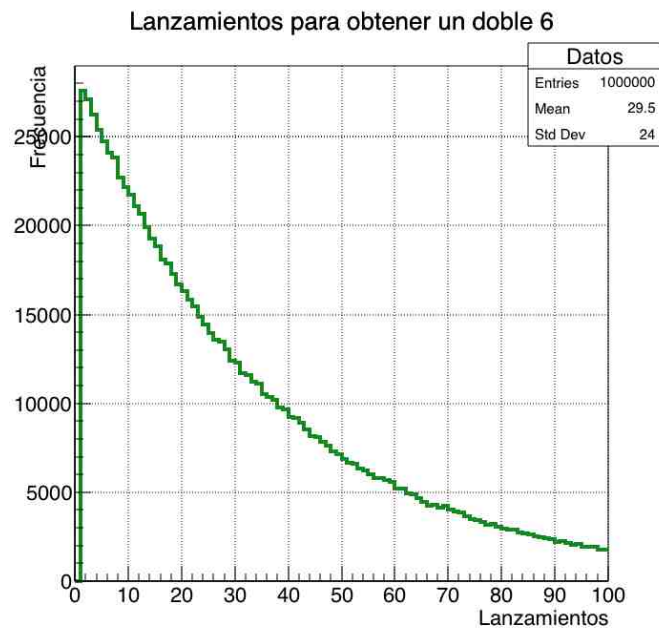


Fig. 64: Resultado de 1,000,000 lanzamientos.

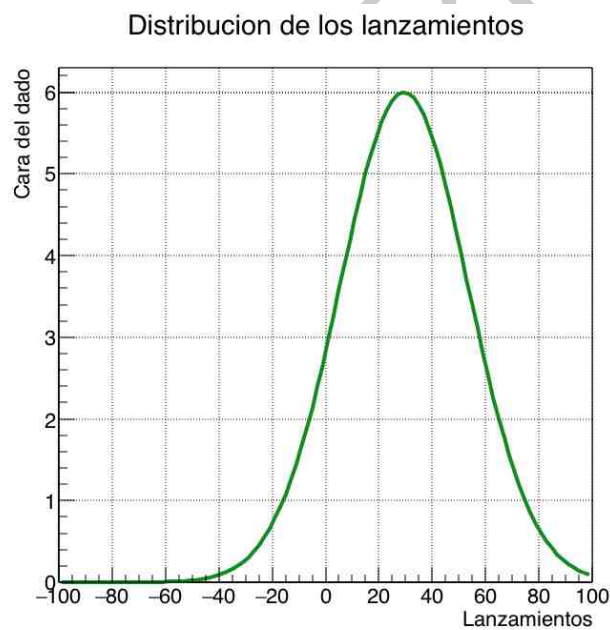


Fig. 65: Distribución normal de 1,000,000 lanzamientos.

En los resultados obtenidos podemos observar que el promedio de tiradas requeridas para obtener el doble seis se encontró cerca de los **29** lanzamientos en cada uno de los experimentos.

40.5 Revisar y extender.

Analizando los resultados obtenidos, nos damos cuenta que el promedio de lanzamientos para los cuales se obtiene por lo menos un doble seis ronda los valores entre 17 y 38 lanzamiento, el cual es un margen muy amplio de error.

Analizando solamente los resultados de los experimentos con mayor número de repeticiones observamos que el valor medio varía de 29 hasta 35 lanzamiento, el cual es un margen mas aceptable y considerablemente cercano al valor esperado de 25 lanzamientos.

Para obtener un resultado mas próximo al valor teórico, se podrían buscar otras formas de generar números aleatorios que nos permitan reproducir el experimento con mayor exactitud.

41 Conclusiones.

Al realizar este experimento se compararon valores teóricos con valores prácticos obtenidos de manera experimental y a través de simulaciones.

Con los valores obtenidos de manera experimental no se obtiene un valor demasiado exacto, ya que debido a la naturaleza del experimento se requiere de mucho tiempo para obtener una muestra lo suficientemente grande como para que el resultado obtenido con ella pueda ser comparado con los resultados teóricos.

Por otro lado, con los valores obtenidos a través de simulaciones se obtiene un resultado bastante estable, pero aún no es lo suficientemente cercano al valor conseguido de manera teórica. Esto posiblemente se debe a los generadores de números aleatorios utilizados, que por su naturaleza computacional, no son totalmente aleatorios.

Para mejorar el experimento, se podrían analizar otros métodos de generación de números aleatorios, y al mismo tiempo, trabajar con muestras experimentales mas grandes, para lograr obtener un resultado mas aproximado al que se espera de manera teórica.

42 EXAMEN DE PROBABILIDAD, PROCESOS ALEATORIOS E INFERENCIA

Elaborado por: Aldo Nicolás Alejandro Miranda Aguilar

Nombre: Andrea Viridiana Soto Miranda

1.- Un estudiante responde al azar a 4 preguntas de verdadero o falso.

- Escriba el espacio muestral.
- Escriba el suceso responder "falso" a una sola pregunta.
- Escriba el suceso responder "verdadero" al menos a 3 preguntas.
- Escriba la unión de estos dos sucesos, la intersección y la diferencia entre el 2 y el 1.
- La colección formada por estos 5 sucesos, más el suceso seguro y el suceso imposible Constituyen un σ -álgebra?

SOLUCIÓN

El espacio muestral es el conjunto de todos los sucesos elementales. Los sucesos elementales son cada uno de los resultados posibles del experimento aleatorio, indescomponibles en otros más simples. Como el experimento consiste en responder al azar a cuatro preguntas, cada uno de los posibles patrones de respuesta constituirá un suceso elemental.

$$E = \{(V, V, V, V), (V, V, V, F), (V, V, F, V), (V, F, V, V), \\ (F, V, V, V), (V, V, F, F), (V, F, V, F), (V, F, F, V), \\ (F, V, V, F), (F, V, F, V), (F, F, V, V), (V, F, F, F), \\ (F, V, F, F), (F, F, V, F), (F, F, F, V), (F, F, F, F)\}$$

b) El Suceso responder falso a una sola pregunta será el subconjunto del espacio muestral formado por todos los sucesos elementales en que solo hay una respuesta falso, lo llamaremos A y será:

$$A = \{(V, V, V, F) \cup (V, V, F, V) \cup (V, F, V, V) \cup (F, V, V, V)\}$$

c) El suceso responder verdadero al menos a 3 preguntas, lo llamaremos B y será:

$$B = (V, V, V, F) \cup (V, V, F, V) \cup (V, F, V, V) \cup (F, V, V, V) \cup (V, V, V, V)$$

d) Observando los sucesos elementales que los componen se deducen inmediatamente los siguientes resultados:

$$A \cup B = B; \\ A \cap B = A; \\ B - A = \{(V, V, V, V)\}$$

e) La colección formada por el suceso A, el B, la unión de ambos, su intersección, y su diferencia, más el suceso seguro y el suceso imposible, no constituye un σ -álgebra. Para demostrarlo basta comprobar que se incumple una de las dos condiciones. Por ejemplo, el suceso A incumple la segunda porque su contrario no pertenece a la colección

2.- La prevalencia de la diabetes es del 4%. La glucemia basal diagnóstica correctamente el 95% de los diabéticos, pero da un 2% de falsos positivos. Diagnosticada una persona Cuál es la probabilidad de que realmente sea diabética?

SOLUCIÓN

Sea D el suceso de tener diabetes, \bar{D} el suceso de no tenerla y $Gl+$ el suceso de dar positivo en la prueba de la glucemia basal. Los datos del problema nos dicen que:

$$\begin{aligned}P(D) &= 0.04; \\P(\bar{D}) &= 0.96; \\P(Gl+ | D) &= 0.95; \\P(Gl+ | \bar{D}) &= 0.02\end{aligned}$$

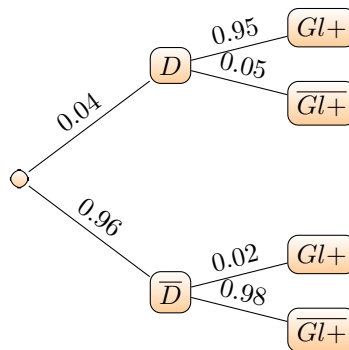


Fig. 66: Diagrama del árbol del ejercicio 2

Entonces el teorema de Bayes, escrito en los términos de este problema nos dice que:

$$P(D|Gl+) = \frac{P(Gl+ | D)P(D)}{P(Gl+ | D)P(D) + P(Gl+ | \bar{D})P(\bar{D})}$$

sustituyendo:

$$P(D|Gl+) = \frac{0.95(0.04)}{0.95(0.04) + 0.02(0.96)} = \frac{0.038}{0.038 + 0.0192} = 0.664$$

3.- En un sistema de comunicación, los mensajes se codifican en base 3, es decir, con los tres símbolos 0, 1 y 2. La probabilidad de emitir cualquiera de los tres símbolos es la misma y se ha observado que el 0 y el 2 nunca se confunden, es decir, la probabilidad de recibir 0 si se ha emitido 2 (respectivamente, recibir 2 si se ha emitido 0) es cero. Además, sea cual sea el símbolo que se emite, la probabilidad de recibir ese mismo símbolo es 0,9 y finalmente se sabe que cuando se emite 1, la probabilidad de recibir 0 y la probabilidad de recibir 2, es la misma. Calcular:

- La probabilidad de recibir 1
- La probabilidad de que el símbolo emitido haya sido 0 si se recibió 0

SOLUCIÓN

a) Se definen los sucesos:

E_0 = "se emite el símbolo 0",

E_1 = "se emite el símbolo 1",

E_2 = "se emite el símbolo 2",

R_0 = "se recibe el símbolo 0",

R_1 = "se recibe el símbolo 1" y

R_2 = "se recibe el símbolo 2"

Con esta notación, se tiene por los datos del enunciado que

$$\begin{aligned}
 P(E_0) &= P(E_1) = P(E_2) = \frac{1}{3} \\
 P(R_0|E_2) &= P(R_2|E_0) = 0 \\
 P(R_0|E_0) &= P(R_1|E_1) = P(R_2|E_2) = 0.9 \\
 P(R_0|E_1) &= P(R_2|E_1)
 \end{aligned}$$

Por tanto, la probabilidad de recibir 1, usando el teorema de la probabilidad total, es

$$\begin{aligned}
 P(R_1) &= P(R_1 \cap (E_0 \cup E_1 \cup E_2)) \\
 &= P(R_1|E_0)P(E_0) + P(R_1|E_1)P(E_1) + P(R_1|E_2)P(E_2) \\
 &= (1 - P(R_0|E_0) - P(R_2|E_0)) \cdot \frac{1}{3} + 0.9 \cdot \frac{1}{3} + (1 - P(R_0|E_2) - P(R_2|E_2)) \cdot \frac{1}{3} \\
 &= (1 - 0.9 - 0) \cdot \frac{1}{3} + 0.9 \cdot \frac{1}{3} + (1 - 0 - 0.9) \cdot \frac{1}{3} = 0.3667
 \end{aligned}$$

b) Para resolver este apartado se utiliza el teorema de Bayes, de forma que:

$$\begin{aligned}
 P(R_0|E_0) &= \frac{P(R_0|E_0)P(E_0)}{P(R_0|E_0)P(E_0) + P(R_0|E_1)P(E_1) + P(R_0|E_2)P(E_2)} \\
 &= \frac{0.9 \cdot \frac{1}{3}}{0.9 \cdot \frac{1}{3} + P(R_0|E_1) \cdot \frac{1}{3} + 0 \cdot \frac{1}{3}}
 \end{aligned}$$

Dado que $P(R_0|E_1) = P(R_2|E_1) = x$, se tiene que al ser $P(R_1|E_1) = 0.9$, entonces $x + 0.9 + x = 1$ y por tanto $x = 0.05$. Por último

$$P(E_0|R_0) = \frac{0.9 \cdot \frac{1}{3}}{0.9 \cdot \frac{1}{3} + 0.05 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3}} = 0.9474$$

4.- Decir si es verdadera o falsa la siguiente afirmación. En caso de que sea verdadera, demostrarlo y en caso de que sea falsa dar un contraejemplo:

"Si $B \subset A$ con $0 < P(A) < 1$, entonces B y A no pueden ser independientes"

$$\begin{aligned}
 &\text{Si } B \subset A \\
 &\Rightarrow A \cap B = B \\
 &\Rightarrow P(A \cap B) = P(B)
 \end{aligned}$$

Por lo tanto, para que $P(A \cap B) = P(A) \cdot P(B)$, tendría que ser necesario que $P(A) = 1$.

Por lo que podemos decir que la afirmación es **verdadera**.

5.- De cuántas formas diferentes se pueden cubrir los puestos de presidente, vicepresidente y tesorero de un club de fútbol sabiendo que hay 12 posibles candidatos?

NO entran todos los elementos

Sí importa el orden

NO se permiten repeticiones

Por lo tanto

$$\# \text{ de formas} = \text{Permutaciones}(12, 3) = \frac{12!}{(12-3)!} = 1320$$

6.- Una persona cruza dos semáforos para ir al trabajo. La probabilidad de que cada uno de ellos esté rojo es de 0.4; de que esté amarillo, 0.2, y de que esté verde, 0.4. Elabore el árbol de probabilidad y calcule la probabilidad de que uno esté en verde y el otro en rojo.

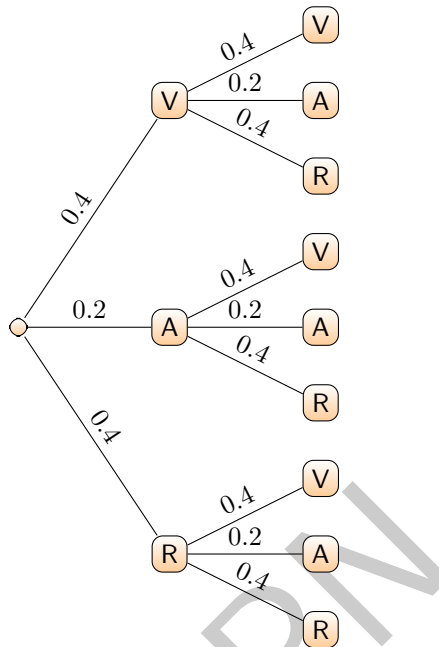


Fig. 67: Diagrama de árbol del ejercicio 6

Se aplica la regla de la suma o de la probabilidad total.

$$P(RV) + P(VR) = 0.4 \cdot 0.4 + 0.4 \cdot 0.4 = 0.32$$

7.- Un estudiante para aprobar un examen que consta de 10 preguntas, debe contestar 7 de ellas. De cuántas maneras puede hacer la selección para aprobar el examen?

NO entran todos los elementos

NO importa el orden

NO se permiten repeticiones

Por lo tanto el resultado lo obtenemos con las combinaciones de 7 en 10

$$\binom{10}{7} = \frac{10!}{7!(10-7)!} = 120$$

8.- Demuestre que $\binom{n}{r}$ es simétrico para $0 \leq r \leq n$

Para que $\binom{n}{r}$ sea simétrico se debe cumplir que:

$$\binom{n}{r} = \binom{n}{n-r}$$

Entonces construimos con $r = (n - r)$

$$\begin{aligned}
 \binom{n}{n-r} &= \frac{n!}{(n-r)!(n-(n-r))!} \\
 &= \frac{n!}{(n-r)!(n-n+r)!} \\
 &= \frac{n!}{r!(n-r)!} \\
 &= \binom{n}{r}
 \end{aligned}$$

9.- El portero titular de un equipo de fútbol para 8 de cada 10 penaltis, mientras que el suplente solo para 5. el portero suplente juega, por termino medio, 15 minutos en cada partido (90 minutos). Si se lanza un penalti y no se para cuál es la probabilidad de que estuviera jugando el portero titular?

Se consideran los sucesos:

P = el portero para un penalti

T = juega el portero titular

S = juega el portero suplente ($S = T^c$)

Con probabilidades:

$$\begin{aligned}
 P(S) &= \frac{15}{90} = \frac{1}{6} \\
 P(T) &= 1 - P(S) = 1 - \frac{1}{6} = \frac{5}{6} \\
 P(P|T) &= \frac{8}{10} = \frac{4}{5} \\
 P(P^c|T) &= 1 - \frac{4}{5} = \frac{1}{5} \\
 P(P|S) &= \frac{5}{10} = \frac{1}{2} \\
 P(P^c|S) &= 1 - \frac{1}{2} = \frac{1}{2}
 \end{aligned}$$

para calcular $P(T|P^c)$ se aplica el teorema de Bayes, con los sucesos T y S como sistema completo de sucesos:

$$P(T|P^c) = \frac{P(P^c|T)P(T)}{P(P^c|T)P(T) + P(P^c|S)P(S)} = \frac{\frac{1}{5} \cdot \frac{5}{6}}{\frac{1}{5} \cdot \frac{5}{6} + \frac{1}{2} \cdot \frac{1}{6}} = \frac{10}{15} \approx 0.66667$$

10.- Cuántas diagonales se pueden trazar dentro de un pentágono regular?

Una diagonal la podemos definir como la recta que va desde un vértice a a otro vértice b dentro del pentágono, tal que b no sea vecino de a y $a \neq b$, por lo tanto, podemos decir que para cada vértice del polígono podemos escoger otros dos vértices con los cuales formar una diagonal y entonces podemos decir que el número de vertices esta dado por

$$\binom{5}{2}$$

Sin embargo, hay que tomar en cuenta que la diagonal ab es la misma que la diagonal ba , por lo que tenemos que dejar de contar estas repeticiones dividiendo entre 2, entonces

$$\# \text{ de diagonales} = \frac{1}{2} \cdot \binom{5}{2} = \frac{1}{2} \cdot 10 = 5$$

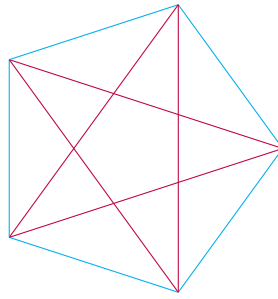


Fig. 68: Número de diagonales en un pentágono

CIC-IPN

43 Skewness y Curtosis

Curtosis En teoría de la probabilidad y estadística, la curtosis es una medida de la forma. Así, las medidas de curtosis tratan de estudiar la proporción de la varianza que se explica por la combinación de datos extremos respecto a la media en contraposición con datos poco alejados de la misma. Una mayor curtosis implica una mayor concentración de datos muy cerca de la media de la distribución coexistiendo al mismo tiempo con una relativamente elevada frecuencia de datos muy alejados de la misma. Esto explica una forma de la distribución de frecuencias con colas muy elevadas y con un centro muy apuntado.

Un coeficiente de apuntamiento o de curtosis es el basado en el cuarto momento con respecto a la media y se define como:

$$\beta_2 = \frac{\mu_4}{\sigma^4} \quad (25)$$

donde μ_4 es el 4 momento centrado o con respecto a la media y σ es la desviación estándar.

No obstante, está más extendida la siguiente definición del coeficiente de curtosis:

$$g_2 = \frac{\mu_4}{\sigma^4} - 3 \quad (26)$$

donde al final se ha sustraído 3 (que es la curtosis de la Normal) con objeto de generar un coeficiente que valga 0 para la Normal y tome a ésta como referencia de apuntamiento:

Tomando, pues, la distribución normal como referencia, una distribución puede ser:

- más apuntada y con colas menos anchas que la normal leptocúrtica.
- menos apuntada y con colas más anchas que la normal - platicúrtica.
- la distribución normal es mesocúrtica.

En la distribución normal se verifica que $\mu_4 = 3\sigma^4$, donde μ_4 es el momento de orden 4 respecto a la media y σ la desviación típica.

Así tendremos que:

- Si la distribución es leptocúrtica $\beta_2 > 3$ y $g_2 > 0$
- Si la distribución es platicúrtica $\beta_2 < 3$ y $g_2 < 0$
- Si la distribución es mesocúrtica $\beta_2 = 3$ y $g_2 = 0$

Otra forma de medir la curtosis se obtiene examinando la fórmula de la curtosis de la suma de variables aleatorias. Si Y es la suma de n variables aleatorias estadísticamente independientes, todas con igual distribución X , entonces $Kurt[Y] = \frac{Kurt[X]}{n}$, complicándose la fórmula si la curtosis se hubiese definido como $\frac{\mu_4}{\sigma^4}$.

Skewness En probabilidad y estadística, skewness es una medida de asimetría del valor real de una variable aleatoria de una distribución de probabilidad acerca de su media. El valor de skewness puede ser positivo o negativo, o incluso indefinido.

La interpretación cualitativa de la falta de simetría es complicado. Para una distribución unimoda, una falta de simetría negativa indica que la cola en el lado izquierdo de la función de densidad de probabilidad es más larga o más gorda que el lado derecho - esto no distingue esas formas.

A la inversa, falta de simetría positiva indica que la cola en el lado derecho es más larga o más gorda que del lado izquierdo. En algunos casos cuando la cola es larga pero la otra cola es gorda, skewness no obedece una regla simple. Por ejemplo, un valor cero indica que la cola en ambos lados de la media está balanceada, el cual es el caso para una distribución simétrica, pero es también cierto para una distribución asimétrica donde las asimetrías no son pares, tal como una cola larga pero delgada, y otra corta pero gorda. Además, en distribuciones multimodas y distribuciones discretas, skewness es también difícil de interpretar. Importantemente, el skewness no determina la relación de la media y mediana.

Las medidas de asimetría son indicadores que permiten establecer el grado de simetría (o asimetría) que presenta una distribución de probabilidad de una variable aleatoria sin tener que hacer su representación gráfica.

Como eje de simetría consideramos una recta paralela al eje de ordenadas que pasa por la media de la distribución. Si una distribución es simétrica, existe el mismo número de valores a la derecha que a la izquierda de la media, por tanto, el mismo número de desviaciones con signo positivo que con signo negativo. Decimos que hay asimetría positiva (o a la derecha) si la "cola" a la derecha de la media es más larga que la de la izquierda, es decir, si hay valores más separados de la media a la derecha. Diremos que hay asimetría negativa (o a la izquierda) si la "cola" a la izquierda de la media es más larga que la de la derecha, es decir, si hay valores más separados de la media a la izquierda.

Coeficiente de asimetría de Fisher

En teoría de la probabilidad y estadística, la medida de asimetría más utilizada parte del uso del tercer momento estándar. La razón de esto es que nos interesa mantener el signo de las desviaciones con respecto a la media, para obtener si son mayores las que ocurren a la derecha de la media que las de la izquierda. Sin embargo, no es buena idea tomar el momento estándar con respecto a la media de orden 1. Debido a que una simple suma de todas las desviaciones siempre es cero. En efecto, si por ejemplo, los datos están agrupados en k clases, se tiene que:

$$\sum_{i=1}^k f_i(x_i - \mu) = \sum_{i=1}^k f_i x_i - \mu \sum_{i=1}^k f_i = \mu - \mu = 0 \quad (27)$$

en donde x_i representa la marca de la clase i -ésima y f_i denota la frecuencia relativa de dicha clase. Por ello, lo más sencillo es tomar las desviaciones al cubo.

El coeficiente de asimetría de Fisher, representado por γ_1 , se define como:

$$\gamma_1 = \frac{\mu_3}{\sigma^3} \quad (28)$$

donde μ_3 es el tercer momento en torno a la media y σ es la desviación estándar.

Si $\gamma_1 > 0$, la distribución es asimétrica positiva o a la izquierda.

Si $\gamma_1 < 0$, la distribución es asimétrica negativa o a la derecha.

Si la distribución es simétrica, entonces sabemos que $\gamma_1 = 0$. El recíproco no es cierto: es un error común asegurar que si $\gamma_1 = 0$ entonces la distribución es simétrica (lo cual es falso).

Coeficiente de asimetría de Pearson

Sólo se puede utilizar en distribuciones uniformes, unimodales y moderadamente asimétricas. Se basa en que en distribuciones simétricas la media de la distribución es igual a la moda.

$$A_p = \frac{\mu - moda}{\sigma} \quad (29)$$

donde μ es el momento central de orden 1, que corresponde a la media aritmética de la variable X .

Si la distribución es simétrica, $\mu = moda$ y $A_p = 0$. Si la distribución es asimétrica positiva la media se sitúa por encima de la moda y, por tanto, $A_p > 0$.

Coeficiente de asimetría de Bowley

Está basado en la posición de los cuartiles y la mediana, y utiliza la siguiente expresión:

$$A_B = \frac{Q_{\frac{3}{4}} + Q_{\frac{1}{4}} - 2M_e}{Q_{\frac{3}{4}} - Q_{\frac{1}{4}}} \quad (30)$$

En una distribución simétrica el tercer cuartil estará a la misma distancia de la mediana que el primer cuartil. Por tanto $A_B = 0$.

Si la distribución es positiva o a la derecha, $A_B > 0$.

La asimetría resulta útil en muchos campos. Muchos modelos simplistas asumen una distribución normal, esto es, simétrica en torno a la media. La distribución normal tiene una asimetría cero. Pero en realidad, los valores no son nunca

perfectamente simétricos y la asimetría de la distribución proporciona una idea sobre si las desviaciones de la media son positivas o negativas. Una asimetría positiva implica que hay más valores distintos a la derecha de la media.

Las medidas de asimetría, sobre todo el coeficiente de asimetría de Fisher, junto con las medidas de apuntamiento o curtosis se utilizan para contrastar si se puede aceptar que una distribución estadística sigue la distribución normal. Esto es necesario para realizar numerosos contrastes estadísticos en la teoría de inferencia estadística.

CIC-IPN

44 Algoritmos para calcular π

En toda la historia se han descubierto/creado muchas maneras para calcular aproximaciones del número Pi.

44.1 Algoritmo de Gauss-Legendre

Partimos de los siguientes datos iniciales:

$$a_0 = 1 \quad b_0 = \frac{1}{\sqrt{2}} \quad t_0 = \frac{1}{4} \quad p_0 = 1$$

A partir de ellos realizamos la siguientes operaciones:

$$a_{n+1} = \frac{a_n + b_n}{2} \quad (31)$$

$$b_{n+1} = \frac{2}{\sqrt{a_n b_n}} \quad (32)$$

$$t_{n+1} = t_n - p_n(a_n - a_{n+1})^2 \quad (33)$$

$$p_{n+1} = 2p_n \quad (34)$$

Entonces Pi se aproxima de la siguiente forma:

$$\pi \approx \frac{(a_n + b_n)^2}{4t_n} \quad (35)$$

El método tiene convergencia de segundo orden. Es decir, en cada iteración duplicamos el número de dígitos exactos obtenidos en la iteración anterior.

44.2 Algoritmo de Borwein

Partimos de los siguientes datos iniciales:

$$a_0 = 6 - 4\sqrt{2}$$

$$y_0 = \sqrt{2} - 1$$

Con ellos operamos de la siguiente forma:

$$y_{k+1} = \frac{1 - (1 - y_k^4)^{\frac{1}{4}}}{1 + (1 - y_k^4)^{\frac{1}{4}}} \quad (36)$$

$$a_{k+1} = a_k(1 + y_{k+1})^4 - 2^{2k+3}y_{k+1}(1 + y_{k+1} + y_{k+1}^2) \quad (37)$$

Entonces se tiene lo siguiente:

$$\lim_{k \rightarrow \infty} (a_k) = \frac{1}{\pi} \quad (38)$$

La convergencia de este método es cuártica. Es decir, en cada iteración se consiguen el cuádruple de dígitos exactos que en la iteración anterior. Existen variaciones de este método que consiguen en cada iteración muchos más dígitos exactos.

44.3 Algoritmo de Chudnovsky

El algoritmo de Chudnovsky es un algoritmo creado por **David Volfovich Chudnovsky** y **Gregory Volfovich Chudnovsky**, mediante el cual podemos obtener muy buenas aproximaciones del número Pi. Se basa en la siguiente expresión relacionada con el número Pi que encontró Ramanujan:

$$\frac{1}{\pi} = \frac{2\sqrt{2}}{9801} \sum_{k=0}^{\infty} \frac{(4k)!(1103 + 26390k)}{(k!)^4 396^{4k}}$$

La expresión del algoritmo de Chudnovsky es la siguiente:

$$\frac{1}{\pi} = \frac{2\sqrt{2}}{9801} \sum_{k=0}^{\infty} \frac{(-1)^k (6k)!(13591409 + 545140134k)}{(3k)!(k!)^3 640320^{3k+3/2}}$$

y con ella obtenemos 14 decimales exactos más de Pi con cada término de la misma.

CIC-IPN

45 Señal WOW

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

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

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

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

Todos los intentos posteriores de obtener una señal de la misma dirección no han encontrado nada inusual.

Posibles explicaciones

El código de letras "6EQUJ5" posee un significado específico.

Las intensidades recibidas del espacio se codifican de la siguiente manera:

0 = Intensidad 0 "1"..."9" = Intensidad 1,...Intensidad 9 "a", "b",...= Intensidad 11, Intensidad 12 y así sucesivamente.

El valor 'U' (intensidad entre 30.0 y 30.999) fue la mayor jamás registrada por el telescopio.²

El telescopio Big-Ear estaba fijo y empleaba la rotación de la Tierra para escanear el cielo. Por la velocidad de rotación de la Tierra y la ventana de observación del telescopio, éste sólo podía observar un punto cualquiera durante 72 segundos. Por eso, se da por hecho que una señal extraterrestre se registraría durante exactamente 72 segundos, y que la grabación de la intensidad de esa señal mostraría un pico gradual durante los primeros 36 segundos (hasta que la señal llegara al centro de la ventana de observación del telescopio) para luego mostrar un descenso gradual. Por eso, tanto la duración de la seal WOW!, 72 segundos, como su forma, corresponderían a un origen extraterrestre. Hay tres hipótesis que intentan explicar el fenómeno:

La señal proviene de una emisión de radio de un satélite artificial que atravesó esa órbita en ese instante. La emisión de radio fue producto de un acontecimiento astronómico de enorme potencia. La señal tiene su origen en una civilización avanzada tecnológicamente con un potente transmisor.

Mucho se ha especulado sobre esta última, alegando que, de querer comunicarse con otra civilización, se debería usar la frecuencia del hidrógeno.

Jerry R. Ehman, el descubridor de la seal, siempre ha expresado sus dudas de que la señal pueda tener un origen extraterrestre. Según sus palabras, deberíamos haberla visto de nuevo cuando la buscamos más de cincuenta veces; algo me sugiere que se trató de una señal con origen terrestre que simplemente fue reflejada por algún pedazo de basura espacial (We should have seen it again when we looked for it 50 times. Something suggests it was an Earth-sourced signal that simply got reflected off a piece of space debris).

En enero de 2016, se planteó una explicación que involucra a dos cometas, 266P/Christensen y P/2008 Y2 (Gibbs), que en la época no habían sido descubiertos (de hecho, fueron descubiertos durante la década de 2010) y se encontraban en aquella zona del espacio cuando se registró la señal. Según Antonio Paris, astrónomo del St. Petersburg College de Florida, se sabe que ambos cuerpos celestes liberan grandes nubes de hidrógeno cuando entran en nuestro Sistema Solar y reciben el impacto de la radiación solar. Esas estelas de hidrógeno hubieran bastado, según el astrónomo, por sí solas para generar una lectura anómala y muy intensa en las mediciones.

46 Resumen Capitulo 5 del libro Probabilidad Serie Schaum

46.1 Variables aleatorias

Una variable aleatoria X de un espacio muestral S es una función de S en el conjunto \mathbb{R} de los números reales tal que la imagen inversa de cada intervalo de \mathbb{R} es un evento (o suceso) de S .

Hacemos énfasis en que si S es un espacio discreto en el cual cada subconjunto es un suceso, entonces cada función de valores reales de S es una variable aleatoria. Por otra parte, se puede comprobar que si S es no contable, entonces ciertas funciones de valores reales de S no son variables aleatorias.

Si X y Y son variables aleatorias del mismo espacio muestral S , entonces $X + Y$, $X + k$, kX y XY (donde k es un número real) son funciones de S definidas por:

$$(X + Y)(s) = X(s) + Y(s)$$

$$kX(s) = kX(s)$$

$$(X + k)(s) = X(s) + k$$

$$(XY)(s) = X(s)Y(s)$$

para todo $s \in S$. Se puede comprobar que estas variables también son aleatorias.

Usamos la notación abreviada $P(X = a)$ y $P(a \leq X \leq b)$ para la probabilidad de los sucesos " X toma el valor a " y " X toma valores en el intervalo $[a, b]$." Esto es,

$$P(X = a) = P(\{s \in S : X(s) = a\})$$

$$\text{y } P(a \leq X \leq b) = P(\{s \in S : a \leq X(s) \leq b\})$$

46.1.1 Distribución y Esperanza de una variable aleatoria finita

Sea X una variable aleatoria de un espacio muestral S con el conjunto de imagen finito; a saber, $X(S) = \{x_1, x_2, \dots, x_n\}$. Convertimos $X(S)$ en un espacio de probabilidad definiendo la probabilidad de x_i como $P(X = x_i)$ que escribimos $f(x_i)$. Esta función f de $X(S)$, o sea, definida como $f(x_i) = P(X = x_i)$, se llama la *función de distribución o probabilidad de X* y se expresa generalmente en la forma de tabla:

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

La distribución f satisface las condiciones

$$(i) f(x_i) \geq 0 \text{ y } (ii) \sum_{i=1}^n f(x_i) = 1$$

La media o esperanza (o: valor esperado) de X , denotada por $E(X)$ o μ_X , o simplemente E o μ se define como

$$E(X) = x_1 f(x_1) + x_2 f(x_2) + \dots + x_n f(x_n) = \sum_{i=1}^n x_i f(x_i)$$

Esto es, $E(X)$ es el promedio ponderado de los valores posibles de X , cada valor por su probabilidad.

Ejemplo 5.1: Se lanza un par de dados corrientes. Obtenemos el espacio finito equiprobable S que consta de las 36 parejas ordenadas de números entre 1 y 6:

$$S = \{(1, 1), (1, 2), \dots, (6, 6)\}$$

Sea X que hace corresponder a cada punto (a, b) de S el máximo de sus números, o sea, $X(a, b) = \max(a, b)$. Entonces X es una variable aleatoria cuyo conjunto imagen es

$$X(S) = \{1, 2, 3, 4, 5, 6\}$$

Calculamos la distribución f de X :

$$f(1) = P(X = 1) = P(\{(1, 1)\}) = \frac{1}{36}$$

$$f(2) = P(X = 2) = P(\{(2, 1), (2, 2), (1, 1)\}) = \frac{3}{36}$$

$$f(3) = P(X = 3) = P(\{(3, 1), (3, 2), (3, 3), (2, 2), (1, 3)\}) = \frac{5}{36}$$

$$f(4) = P(X = 4) = P(\{(4, 1), (4, 2), (4, 3), (4, 4), (3, 4), (2, 4), (1, 4)\}) = \frac{7}{36}$$

$$f(5) = P(X = 5) = P(\{(5, 1), (5, 2), (5, 3), (5, 4), (5, 5), (4, 5), (3, 5), (2, 5), (1, 5)\}) = \frac{9}{36}$$

$$f(6) = P(X = 6) = P(\{(6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6), (5, 6), (4, 6), (3, 6), (2, 6), (1, 6)\}) = \frac{11}{36}$$

Esta información se pone en la tabla de la siguiente manera:

x_i	1	2	3	4	5	6
$f(x_i)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

Calculamos la media de X :

$$E(X) = \sum x_i f(x_i) = 1 \cdot \frac{1}{36} + 2 \cdot \frac{3}{36} + 3 \cdot \frac{5}{36} + 4 \cdot \frac{7}{36} + 5 \cdot \frac{9}{36} + 6 \cdot \frac{11}{36} = \frac{161}{36} = 4.47$$

Ahora sea Y que hace corresponder a cada punto (a, b) de S la suma de sus números, o sea, $Y(a, b) = a + b$. Entonces Y es también una variable aleatoria de S con conjunto imagen

$$Y(S) = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

A continuación la distribución de Y :

y_i	2	3	4	5	6	7	8	9	10	11	12
$g(y_i)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Obtenemos, por ejemplo, $g(4) = \frac{3}{36}$ del hecho de que, $(1, 3), (2, 2)$ y $(3, 1)$ son aquellos puntos de S para los que la suma de componentes es 4; por tanto

$$g(4) = P(Y = 4) = P(\{(1, 3), (2, 2), (3, 1)\}) = \frac{3}{36}$$

La media de Y se calcula como sigue:

$$E(Y) = \sum y_i g(y_i) = 2 \cdot \frac{1}{36} + 3 \cdot \frac{2}{36} + \cdots + 12 \cdot \frac{1}{36} = 7$$

Los siguientes diagramas describen gráficamente las distribuciones anteriores:

46.1.2 Varianza y Desviación estándar

La media de una variable aleatoria X mide, en cierto sentido, el valor "promedio" de X . La varianza de X mide el "esparcimiento" o "dispersión" de X , se denota por $\text{var}(X)$ y se define como

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

La desviación estándar de X , denotada por σ_x , es la raíz cuadrada (no negativa) de $\text{var}(X)$

$$\sigma_x = \sqrt{\text{var}(X)}$$

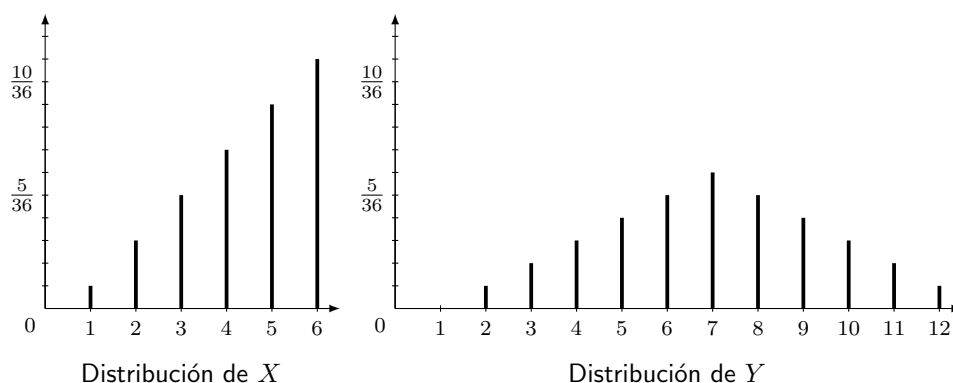


Fig. 69: Función de densidad de probabilidad y función de distribución de probabilidad

46.1.3 Distribución conjunta

Sean X y Y variables aleatorias de un espacio muestral S con los respectivos conjuntos imagen

$$X(S) = \{x_1, x_2, \dots, x_n\} \text{ y } Y(S) = \{y_1, y_2, \dots, y_m\}$$

Formamos el conjunto producto

$$X(S) \times Y(S) = \{(x_1, y_1), (x_1, y_2), \dots, (x_n, y_m)\}$$

en un espacio de probabilidad definiendo la *probabilidad* de la pareja ordenada (x_i, y_i) como $P(X = x_i, Y = y_i)$ que escribimos $h(x_i, y_i)$. Esta función h de $X(S) \times Y(S)$, esto es, definida por $h(x_i, y_i) = P(X = x_i, Y = y_i)$, se llama *distribución conjunta o función de probabilidad conjunta* de X y Y y se da de forma de tabla por lo general:

X \ Y	Y				
	y_1	y_2	...	y_m	Suma
x_1	$h(x_1, y_1)$	$h(x_1, y_2)$...	$h(x_1, y_m)$	$f(x_1)$
x_2	$h(x_2, y_1)$	$h(x_2, y_2)$...	$h(x_2, y_m)$	$f(x_2)$
...
x_n	$h(x_n, y_1)$	$h(x_n, y_2)$...	$h(x_n, y_m)$	$f(x_n)$
Suma	$g(y_1)$	$g(y_2)$...	$g(y_m)$	

Las funciones f y g anteriores se definen por

$$f(x_i) = \sum_{j=1}^m h(x_i, y_j) \text{ y } g(y_j) = \sum_{i=1}^n h(x_i, y_j)$$

o sea, $f(x_i)$ es la suma de los elementos de la fila i -ésima y $g(y_j)$ es la suma de los elementos de la columna j -ésima; son llamados *distribuciones marginales* y son, de hecho, las distribuciones (individuales) de X y Y respectivamente. La distribución conjunta de h satisface las condiciones

$$(i) h(x_i, y_j) \geq 0$$

$$\text{y } (ii) \sum_{i=1}^n \sum_{j=1}^m h(x_i, y_j)$$

Ahora si X y Y son variables aleatorias con la distribución conjunta anterior (y las respectivas medias μ_X y μ_Y), entonces la *covarianza* de X y Y denotada por $\text{cov}(X, Y)$, se define por

$$\text{cov}(X, Y) = \sum_{i,j} (x_i - \mu_X)(y_j - \mu_Y)h(x_i, y_j) = E[(X - \mu_X)(Y - \mu_Y)]$$

La correlación de X y Y , denotada por $\rho(X, Y)$, se define por

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}$$

La correlación no es dimensionada y tiene las siguientes propiedades:

$$(i) \rho(X, Y) = \rho(Y, X)$$

$$(ii) -1 \leq \rho \leq 1$$

$$(iii) \rho(X, X) = 1, \rho(X, -X) = -1$$

$$(iv) \rho(aX + b, cY + d) = \rho(X, Y), \text{ si } a, c \neq 0$$

46.2 Variables aleatorias continuas

Sea X una variable aleatoria continua. Decimos que la función integrable y no negativa $f(x) : \mathbb{R} \rightarrow \mathbb{R}$ es la función de densidad de probabilidad si para cualquier intervalo $[a, b]$ de \mathbb{R} se cumple la igualdad

$$P(a \leq X \leq b) = \int_a^b f(x) dx$$

Como podemos observar es el área sombreada.

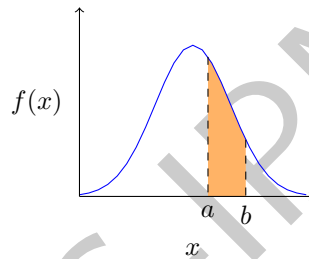


Fig. 70: Función de distribución de probabilidad continua

Por lo tanto decimos que X es una variable continua. a función f se llama función de distribución o de probabilidad continua (o función de densidad) de X que satisface las condiciones [6].

$$(i) f(x) \geq 0$$

$$(ii) \int_{\mathbb{R}} f(x) dx = 1$$

Esto es, f no es negativa y el área total bajo su curva es 1.

46.2.1 Esperanza, varianza y desviación estándar

La esperanza matemática se define por

$$E(X) = \int_{\mathbb{R}} xf(x) dx$$

Para obtener la varianza se obtiene primero la esperanza matemática de los valores elevados al cuadrado.

$$E(X^2) = \int_{\mathbb{R}} x^2 f(x) dx$$

Y con esto podemos obtener la varianza que se define por

$$\text{var}(X) = E(X^2) - \mu^2$$

Y por último se define la desviación estándar como

$$\sigma_x = \sqrt{\text{var}(X)}$$

46.2.2 Ejemplo de variable aleatoria continua

Sea X una variable aleatoria continua con la distribución siguiente:

$$f(x) = \begin{cases} \frac{1}{2}x & \text{si } 0 \leq x \leq 2 \\ 0 & \text{en otro caso} \end{cases}$$

Entonces obtener $P(1 \leq x \leq 1.5)$, $E(X)$, $var(X)$ y σ_x

Solución Se construye la gráfica del problema como se muestra en la figura

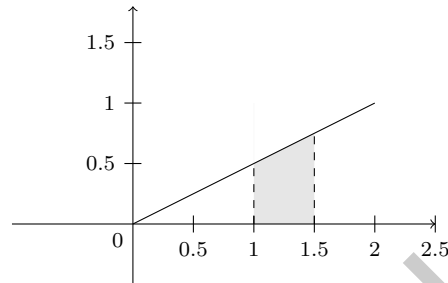


Gráfico de f

Fig. 71: Función de distribución de Ejemplo de variable aleatoria continua

Primero calculamos

$$P(1 \leq x \leq 1.5) = \int_1^{1.5} \frac{1}{2}x = \left[\frac{1}{4} \cdot x^2 \right]_1^{1.5} = \frac{1}{4} (1.5^2 - 1^2) = \frac{5}{16}$$

Obteniendo esperanza matemática, esperanza matemática con X^2 , varianza y desviación estándar

$$E(X) = \int_{\mathbb{R}} x f(x) dx = \int_0^2 \frac{1}{2} x^2 dx = \left[\frac{x^3}{6} \right]_0^2 = \frac{4}{3}$$

$$E(X^2) = \int_{\mathbb{R}} x^2 f(x) dx = \int_0^2 \frac{1}{2} x^3 dx = \left[\frac{x^4}{8} \right]_0^2 = 2$$

$$var(X) = E(X^2) - \mu^2 = 2 - \frac{16}{9} = \frac{2}{9}$$

$$\sigma_x = \sqrt{var(X)} = \sqrt{\frac{2}{9}} = \frac{1}{3}\sqrt{2}$$

47 Análisis ROC de falsos positivos

En la Teoría de detección de señales una curva ROC (acrónimo de Receiver Operating Characteristic) es una representación gráfica de la sensibilidad frente a (1 - especificidad) para un sistema clasificador binario según se varía el umbral de discriminación. Otra interpretación de este gráfico es la representación de la razón o ratio de verdaderos positivos (VPR = Razón de Verdaderos Positivos) frente a la razón o ratio de falsos positivos (FPR = Razón de Falsos Positivos) también según se varía el umbral de discriminación (valor a partir del cual decidimos que un caso es un positivo). ROC también puede significar Relative Operating Characteristic (Característica Operativa Relativa) porque es una comparación de dos características operativas (VPR y FPR) según cambiamos el umbral para la decisión.

El análisis de la curva ROC, o simplemente análisis ROC, proporciona herramientas para seleccionar los modelos posiblemente óptimos y descartar modelos subóptimos independientemente de (y antes de especificar) el coste de la distribución de las dos clases sobre las que se decide. La curva ROC es también independiente de la distribución de las clases en la población (en diagnóstico, la prevalencia de una enfermedad en la población). El análisis ROC se relaciona de forma directa y natural con el análisis de coste/beneficio en toma de decisiones diagnósticas.

La curva ROC se desarrolló por ingenieros eléctricos para medir la eficacia en la detección de objetos enemigos en campos de batalla mediante pantallas de radar, a partir de lo cual se desarrolla la Teoría de Detección de Seales (TDS). El análisis ROC se aplicó posteriormente en medicina, radiología, psicología y otras áreas durante varias décadas. Sólo recientemente ha encontrado aplicación en áreas como aprendizaje automático (o machine learning en inglés), y minería de datos (data mining en inglés).

Un modelo de clasificación (clasificador o Clasificadores (matemático) o Diagnóstico) es una función que permite decidir cuáles de un conjunto de instancias están relacionadas o no por pertenecer a un mismo tipo o clase. El resultado del clasificador o del diagnóstico puede ser un número real (valor continuo), en cuyo caso el límite del clasificador entre cada clase debe determinarse por un valor umbral (por ejemplo para determinar si una persona tiene hipertensión basándonos en una medida de presión arterial), o puede ser un resultado discreto que indica directamente una de las clases.

Consideremos un problema de predicción de clases binario, en la que los resultados se etiquetan positivos (p) o negativos (n). Hay cuatro posibles resultados a partir de un clasificador binario como el propuesto. Si el resultado de una exploración es p y el valor dado es también p, entonces se conoce como un Verdadero Positivo (VP); sin embargo si el valor real es n entonces se conoce como un Falso Positivo (FP). De igual modo, tenemos un Verdadero Negativo (VN) cuando tanto la exploración como el valor dado son n, y un Falso Negativo (FN) cuando el resultado de la predicción es n pero el valor real es p. Un ejemplo aproximado de un problema real es el siguiente: consideremos una prueba diagnóstica que persiga determinar si una persona tiene una cierta enfermedad. Un falso positivo en este caso ocurre cuando la prueba predice que el resultado es positivo, cuando la persona no tiene realmente la enfermedad. Un falso negativo, por el contrario, ocurre cuando el resultado de la prueba es negativo, sugiriendo que no tiene la enfermedad cuando realmente sí la tiene.

Definamos un experimento a partir de P instancias positivas y N negativas. Los cuatro posibles resultados se pueden formular en una Tabla de contingencia (o Matriz de confusión) 2x2 como sigue:

		Valor en la realidad		total
		p	n	
Predicción outcome	p'	Verdaderos Positivos	Falsos Positivos	P'
	n'	Falsos Negativos	Verdaderos Negativos	N'
total		P	N	

48 Random Walk

La caminata aleatoria, abreviado en inglés como RW (Random Walks), es una formalización matemática de la trayectoria que resulta de hacer sucesivos pasos aleatorios. Por ejemplo, la ruta trazada por una molécula mientras viaja por un líquido o un gas, el camino que sigue un animal en su búsqueda de comida, el precio de una acción fluctuante y la situación financiera de un jugador pueden tratarse como una caminata aleatoria. El término caminata aleatoria fue introducido por Karl Pearson en 1905.¹ Los resultados del estudio de las caminatas aleatorias han sido aplicados a muchos campos como la computación, la física, la química, la ecología, la biología, la psicología o la economía. En particular en este último campo la teoría del paseo aleatorio de Burton G. Malkiel en su obra *A Random Walk Down Wall Street* (que se traduce al español como *Un Paseo Aleatorio Por Wall Street*) se fundamenta en la hipótesis de los mercados eficientes, desarrollado en tres formas o hipótesis. En física el modelo ha servido, por ejemplo, para modelar el camino seguido por una molécula que viaja a través de un líquido o un gas (movimiento browniano). En ecología, se emplea para modelar los movimientos de un animal de pastoreo, etc. Varios tipos diferentes de caminos aleatorios son de interés. A menudo, los caminos aleatorios se suponen que son cadenas de Markov o proceso de Markov, pero otros caminos más complicados también son de interés. Algunos caminos aleatorios se dan en grafos finitos, otros en la recta, en el plano, o en dimensiones mayores, mientras algunos caminos aleatorios se dan en grupos.

En su forma más general, las caminatas aleatorias son cualquier proceso aleatorio donde la posición de una partícula en cierto instante depende solo de su posición en algún instante previo y alguna variable aleatoria que determina su subsecuente dirección y la longitud de paso. Los caminos aleatorios también varían con respecto al tiempo. Casos específicos o límites de estos incluyen la caminata de un borracho, el vuelo de Lévy y el movimiento browniano. Los paseos aleatorios están relacionados con los modelos de difusión y son un tema fundamental en la discusión de los procesos de Markov. Varias propiedades de los paseos aleatorios incluyen distribuciones dispersas, tiempos de primer cruce y rutas de encuentro.

48.1 Aplicaciones

- En genética de poblaciones, el camino aleatorio describe las propiedades estadísticas de la deriva genética.
- En física, los caminos aleatorios son utilizados como modelos simplificados del movimiento browniano y difusión tales como el movimiento aleatorio de las moléculas en líquidos y gases. Además, los caminos aleatorios y algunos de los caminos que interactúan consigo mismos juegan un papel en la teoría cuántica de campos.
- En biología matemática, los caminos aleatorios son utilizados para describir los movimientos individuales de los animales, para apoyar empíricamente los procesos de biodifusión, y en ocasiones para desarrollar la dinámica de poblaciones.
- En otros campos de las matemáticas, el camino aleatorio se utiliza para calcular las soluciones de la ecuación de Laplace, para estimar la media armónica, y para varias construcciones en el análisis y la combinatoria.
- En informática, los caminos aleatorios son utilizados para estimar el tamaño de la Web. En la World Wide Web conference-2006, Bar-Yossef et al. publicó sus descubrimientos y algoritmos para lo mismo.
- En el procesamiento de imágenes, los caminos aleatorios son utilizados para determinar las etiquetas (es decir, "objeto" o "fondo") para asociarlas con cada píxel. Este algoritmo se suele denominar como algoritmo de segmentación del camino aleatorio.
- Publicidad de los productos en redes sociales.
- Redes inalámbricas.
- Motores de búsqueda.

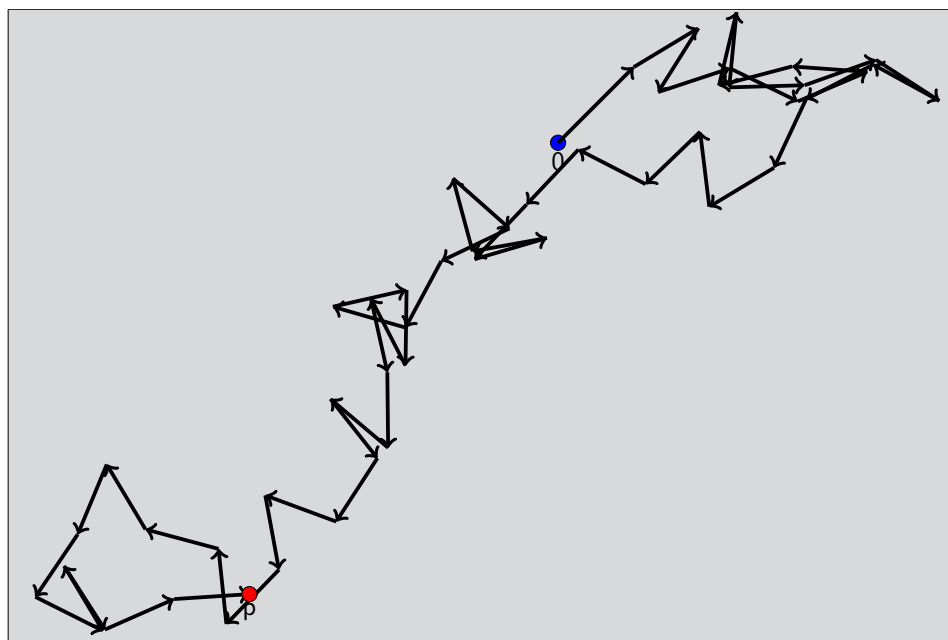


Fig. 72: Caminata al azar en dos dimensiones

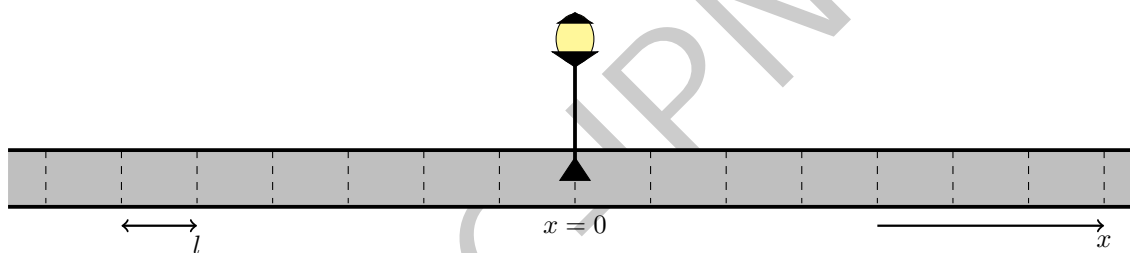


Fig. 73: Caminata al azar un una dimensión

Caminata al azar.

Caminata al azar en una dimensión.

n_1 = número de pasos a la derecha.

n_2 = número de pasos a la izquierda.

				n_1	n_2	m
→	→	→	}	3	0	3
→	→	←	}	2	1	1
→	←	→				
←	→	→				
→	←	←	}	1	2	-1
←	→	←				
←	←	→				
←	←	←	}	0	3	-3

Fig. 74: Tabla de posición final, dependiendo los pasos dados.

48.2 Histogramas en root

Código para generar histogramas "Random Walk"

```
//Autor: Soto Miranda Andrea Viridiana
//Random Walk
//Instituto: CIC-IPN
//Materia: Probabilidad, procesos aleatorios e inferencia
void random_walk(){
    TRandom3 *r = new TRandom3(0);
    TCanvas *c1 = new TCanvas("Random_Walk",
                              "posici\'on final",
                              ,20
                              ,10
                              ,700
                              ,700);

    c1->SetGrid();

    TCanvas *c2 = new TCanvas("Random_Walk2",
                              "Pasos a la derecha",
                              ,800
                              ,10
                              ,700
                              ,700);

    c2->SetGrid();

    TH1F* h_posicion= new TH1F("Datos",
                              "Posici#acute{o}n Final",
                              ,100
                              ,-20
                              ,20);

    h_posicion->SetLineColor(225);
    h_posicion->SetLineWidth(4);
    h_posicion->GetXaxis()->SetTitle("Posici#acute{o}n Final");
    h_posicion->GetYaxis()->SetTitle("m/N");
    TH1F* h_derecha= new TH1F("Datos",
                              "Pasos a la derecha",
                              ,100
                              ,0
                              ,20);

    h_derecha->SetLineColor(225);
    h_derecha->SetLineWidth(4);
    h_derecha->GetXaxis()->SetTitle("Pasos a la derecha");
    h_derecha->GetYaxis()->SetTitle("n1/N");

    int pasos=0;
    int m=0;
    int repeticiones = 100000;
```



```
int n1=0;
int n2=0;
int prob_derecha=5;
int paso=0;
for(int i = 0; i < repeticiones; i++){
    while(pasos<20){
        paso = (r->Integer(10))+1;
        if(paso<=prob_derecha){
            m--;
            pasos++;
        }
        else if(paso>prob_derecha){
            m++;
            pasos++;
            n1++;
        }
    }
    h_posicion->Fill(m);
    h_derecha->Fill(n1);
    m = 0;
    pasos = 0;
    n1 = 0;
    n2=0;
}
c1->cd();
h_posicion->Scale(1/h_posicion->Integral());
h_posicion->Draw();

c2->cd();
h_derecha->Scale(1/h_derecha->Integral());
h_derecha->Draw();
}
```

- Probabilidad $p = 0.5$ y $q = 0.5$

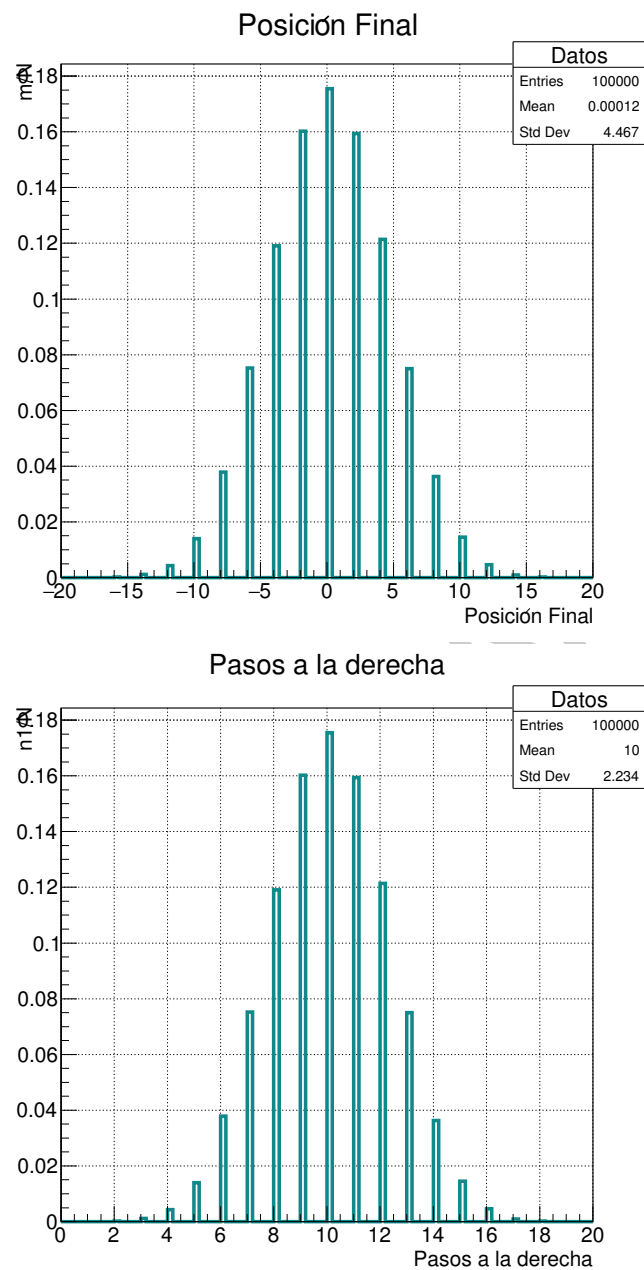
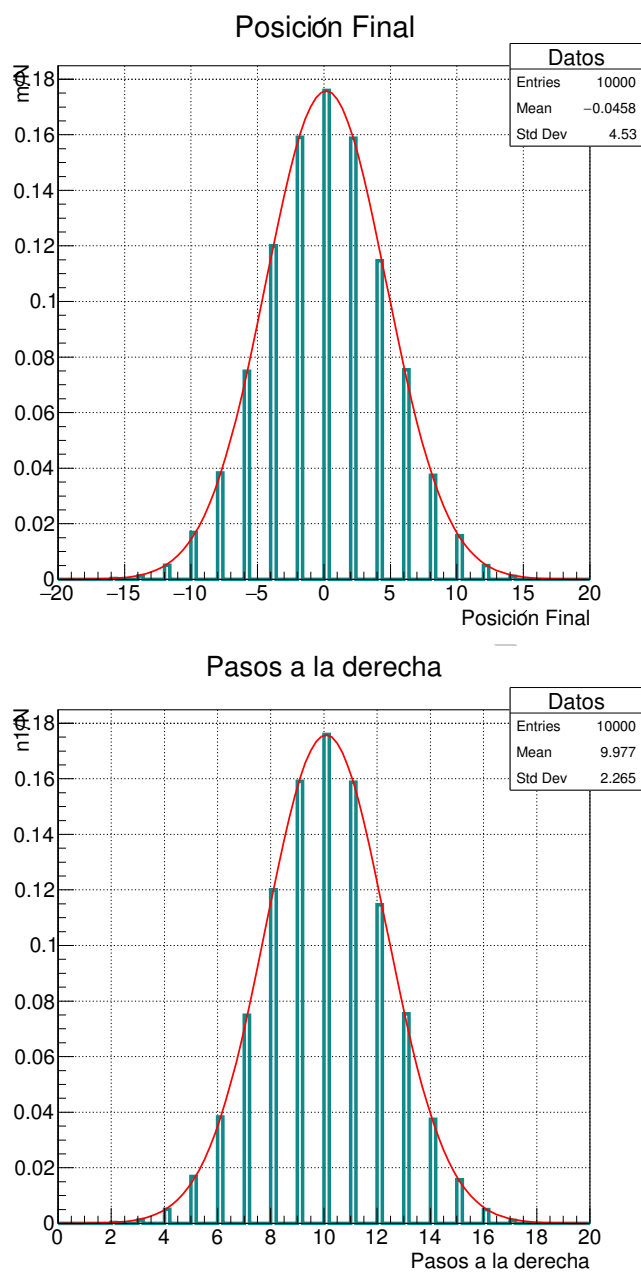


Fig. 75: Random Walk $p = 0.5$ y $q = 0.5$

Fig. 76: Random Walk $p = 0.5$ y $q = 0.5$

- Probabilidad $p = 0.4$ y $q = 0.6$

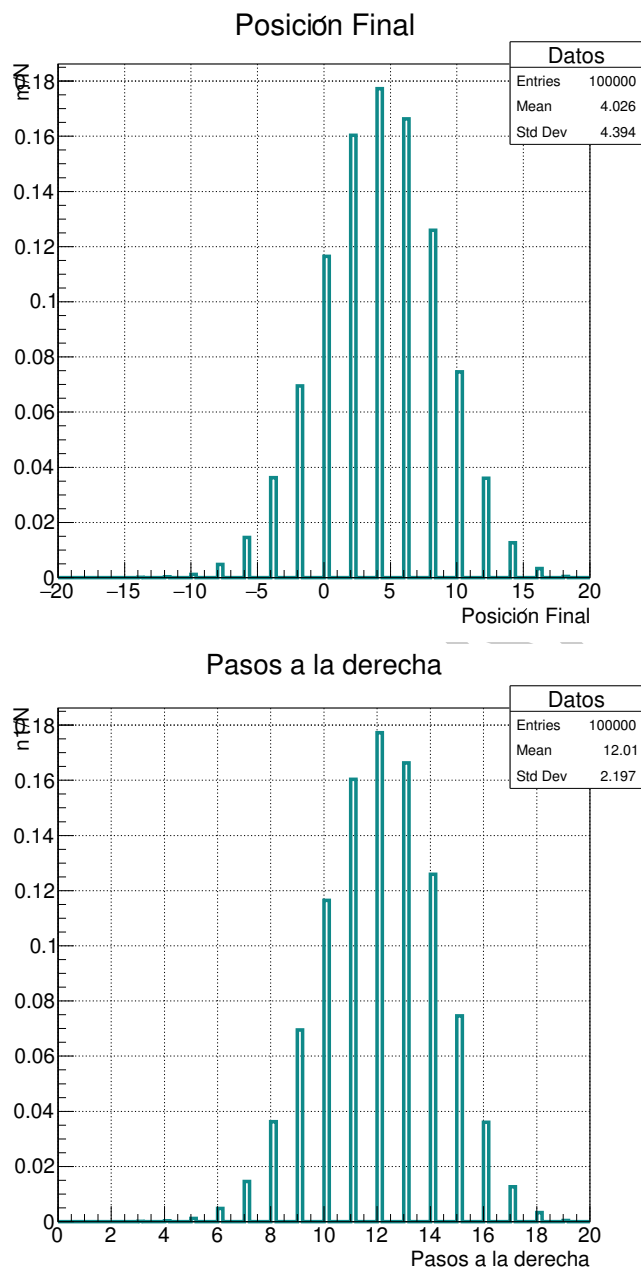
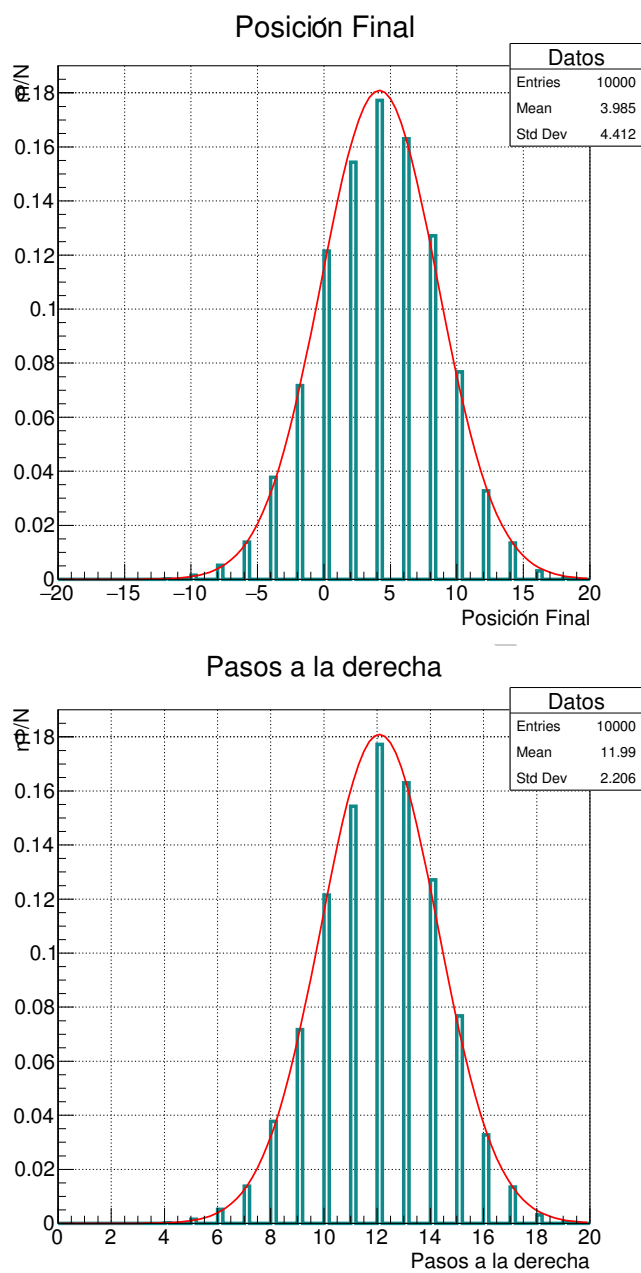


Fig. 77: Random Walk $p = 0.4$ y $q = 0.6$

Fig. 78: Random Walk $p = 0.5$ y $q = 0.5$

- Probabilidad $p = 0.6$ y $q = 0.4$

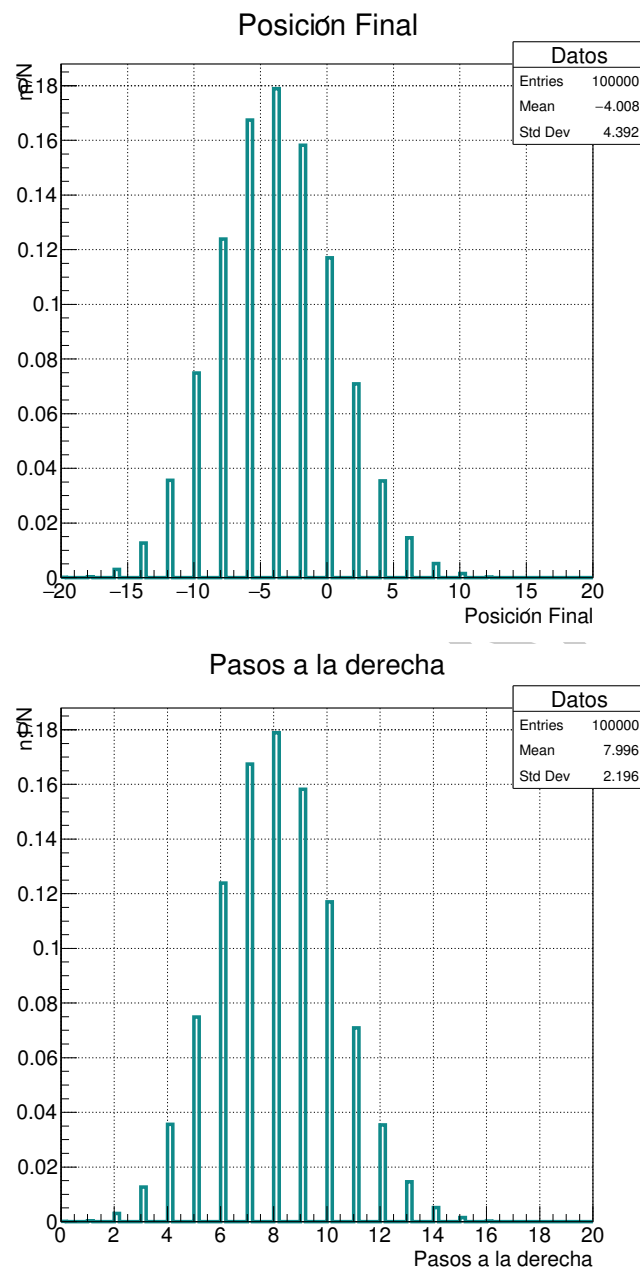
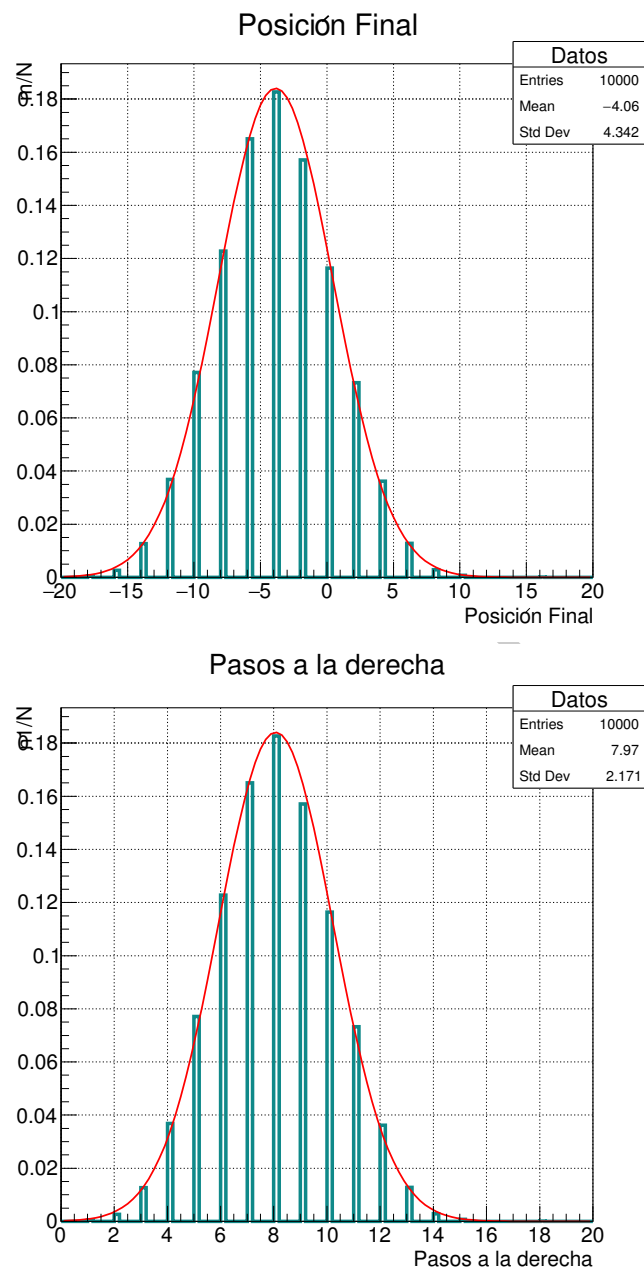


Fig. 79: Random Walk $p = 0.6$ y $q = 0.4$

Fig. 80: Random Walk $p = 0.5$ y $q = 0.5$

48.3 Demostraciones Random Walk

- Ecuación 23

$$\begin{aligned}\overline{\Delta u} &= \overline{u - \bar{u}} = \bar{u} - \bar{u} \\ &= \bar{u} - \bar{u} \\ &= \sum_{i=1}^M P(u_i) \cdot u_i - \bar{u} \\ &= \bar{u} - \bar{u}\end{aligned}$$

- Ecuación 24

$$\overline{(\Delta u)^2} = \sum_{i=1}^M P(u_i)(u_i - \bar{u})^2 \geq 0$$

Debido a que $P(u_i)$ no puede ser negativa y al elevar $(u_i - \bar{u})^2$ aunque $u_i - \bar{u}$ sea negativo, al elevarlo al cuadrado el valor se vuelve positivo y puede ser igual o mayor a cero.

- Ecuación 25

$$\begin{aligned}\overline{(u - \bar{u})^2} &= \overline{u^2 - 2u\bar{u} + \bar{u}^2} \\ &= \sum_{i=1}^M P(u_i)(u_i - \bar{u})^2 \\ &= \sum_{i=1}^M P(u_i)(u_i^2 - 2u_i\bar{u} + \bar{u}^2) \\ &= \sum_{i=1}^M P(u_i) \cdot u_i^2 - 2\bar{u} \sum_{i=1}^M P(u_i) \cdot u_i + \bar{u}^2 \\ &= \sum_{i=1}^M P(u_i) \cdot u_i^2 - 2\bar{u}\bar{u} + \bar{u}^2 \\ &= \overline{u^2} - \bar{u}^2\end{aligned}$$

- Ecuación 26

$$\begin{aligned}\overline{u^2} &\geq \bar{u}^2 \\ \sum_{i=1}^M P(u_i)(u_i - \bar{u})^2 &\geq \left[\sum_{i=1}^M P(u_i)(u_i - \bar{u}) \right]^2\end{aligned}$$

- Ecuación 31

$$\begin{aligned}n_1 p^{n_1} &= p \frac{\partial}{\partial p} (p^{n_1}) \\ &= p \cdot (n_1 p^{n_1-1}) \\ &= n_1 p^{n_1-1+1} \\ &= n_1 p^{n_1}\end{aligned}$$

- Ecuación 32

$$\overline{n_1} = Np.$$

$$\sum_{i=1}^N n_1 p$$

- Ecuación 33

$$\overline{n_2} = Nq.$$

$$\sum_{i=1}^N n_2 q$$

- Ecuación 34

$$\overline{n_1} + \overline{n_2} = Np + Nq = N(p + q) = N$$

Debido a que $p+q=1$ solo queda N

- Ecuación 35

$$\overline{m} = \overline{n_1 - n_2} = \overline{n_1} - \overline{n_2} = N(p - q)$$

Si $p = q$ se tiene que $\overline{m} = 0$.

- Ecuación 38

$$\begin{aligned} n_1^2 p^{n_1} &= n_1 p \frac{\partial}{\partial p} (p^{n_1}) = \left(p \frac{\partial}{\partial p} \right)^2 (p^{n_1}) \\ &= p^2 \frac{\partial^2}{\partial p^2} (p^{n_1}) \\ &= p^2 \frac{\partial}{\partial p} (n_1 p^{n_1-1}) \\ &= n_1 \cdot p^2 \frac{\partial}{\partial p} (p^{n_1-1}) \\ &= n_1 \cdot p^2 ((n_1 - 1) p^{n_1-2}) \\ &= n_1 (n_1 - 1) (p^{n_1-2+2}) \\ &= n_1 (n_1 - 1) (p^{n_1}) \end{aligned}$$

49 Función Gamma

En matemáticas, la función gamma (denotada como $\Gamma(z)$), es una aplicación que extiende el concepto de factorial a los números complejos. La notación fue propuesta por Adrien-Marie Legendre. Si la parte real del número complejo z es positiva, entonces la integral

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

converge absolutamente; esta integral puede ser extendida a todo el plano complejo, excepto a los enteros negativos y al cero. Si n es un entero positivo, entonces

$$\Gamma(n) = (n-1)!$$

lo que nos muestra la relación de esta función con el factorial. De hecho, la función gamma extiende el concepto de factorial a cualquier valor complejo de z . La función gamma aparece en varias funciones de distribución de probabilidad, por lo que es bastante usada tanto en probabilidad y estadística como en combinatoria.

Si la parte real del número complejo z es positiva $\text{Re}(z) > 0$, entonces la integral

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

converge absolutamente. Usando la integración por partes, se obtiene la siguiente propiedad:

$$\Gamma(z+1) = z\Gamma(z)$$

Esta ecuación funcional generaliza la relación $n! = n(n-1)!$ del factorial.

Se puede evaluar $\Gamma(1)$ analíticamente:

$$\Gamma(1) = \int_0^{\infty} e^{-t} dt = \lim_{k \rightarrow \infty} -e^{-t} \Big|_0^k = -0 - (-1) = 1.$$

Combinando estos dos resultados se deduce que el factorial es un caso particular de la función gamma:

$$\Gamma(n+1) = n\Gamma(n) = \dots = n!\Gamma(1) = n!$$

Aplicando la fórmula a unos pocos valores:

$$\begin{aligned}\Gamma(2) &= \Gamma(1+1) = 1\Gamma(1) = 1! = 1 \\ \Gamma(3) &= \Gamma(2+1) = 2\Gamma(2) = 2 \cdot 1! = 2! = 2 \\ \Gamma(4) &= \Gamma(3+1) = 3\Gamma(3) = 3 \cdot 2! = 3! = 6 \\ \Gamma(n+1) &= n\Gamma(n) = n \cdot (n-1)! = n!\end{aligned}$$

50 Movimiento Browniano

50.1 Simulación Movimiento Browniano 2D en root

El código que se muestra a continuación simula el movimiento browniano en 2 dimensiones

```
//Autor: Soto Miranda Andrea Viridiana
//Movimiento Browniano
//Instituto: CIC-IPN
//Materia: Probabilidad, procesos aleatorios e inferencia

#include "TGraph.h"

void movimiento_browniano();

void movimiento_browniano()
{
    TCanvas *c1 = new TCanvas("Movimiento_Browniano"
                              , "Movimiento_Browniano"
                              , 20
                              , 10
                              , 700
                              , 700);

    Double_t x=0;
    Double_t y=0;
    Double_t angulo=0;
    TRandom3 *r = new TRandom3(0);
    Int_t dim = 10000;

    TGraph *g = new TGraph();
    g->SetTitle("Movimiento_Browniano");
    for (Int_t i = 0; i < dim; i++)
    {
        angulo = (r->Integer(360))+1;
        g->SetPoint(i,x,y);
        x=x+cos(angulo);
        y=y+sin(angulo);
    }
    g->SetLineWidth(1);
    g->SetLineColor(225);
    g->GetXaxis()->SetTitle("Desplazamiento_en_x");
    g->GetYaxis()->SetTitle("Desplazamiento_en_y");
    g->Draw("AL");
}
```

En la Fig(81) y Fig(82) se muestra la trayectoria que resulta de esta simulación.

Movimiento Browniano

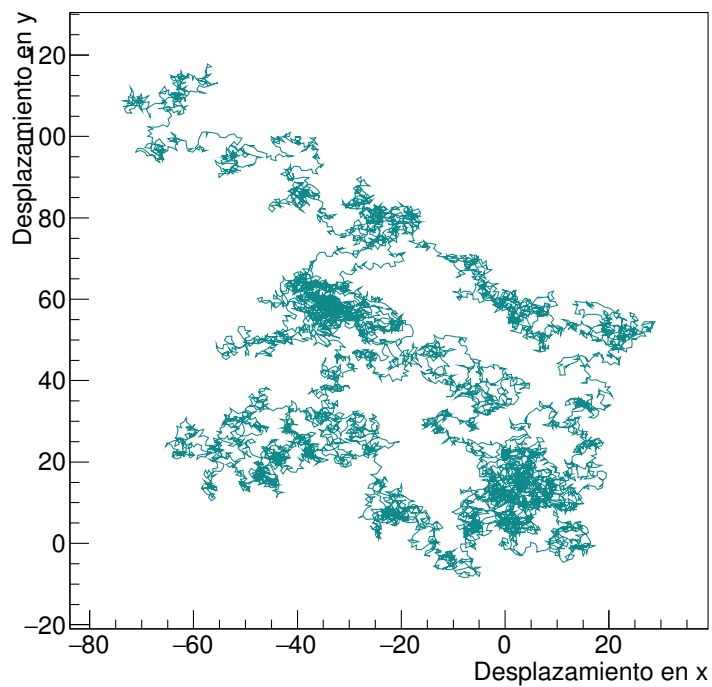


Fig. 81: Movimiento browniano en 2 dimensiones.

Movimiento Browniano

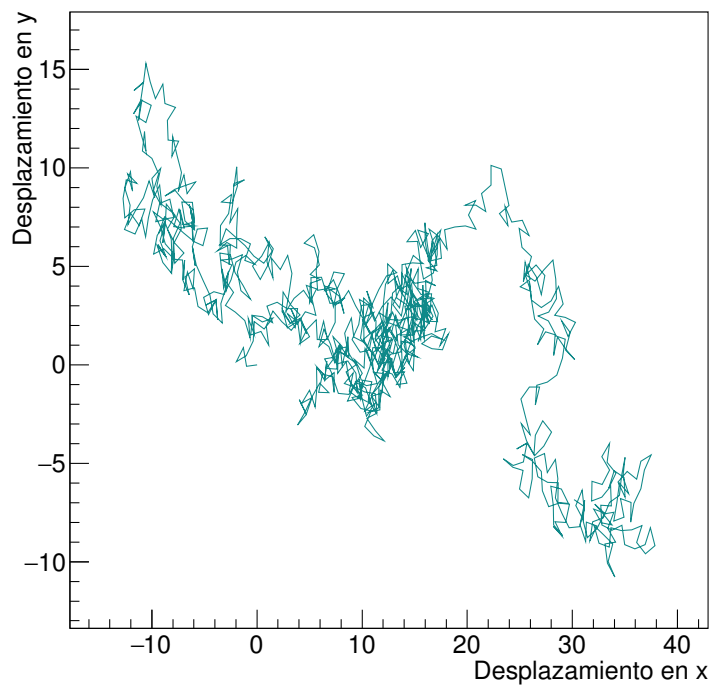


Fig. 82: Movimiento browniano en 2 dimensiones.

50.2 Simulación Movimiento Browniano 3D en root

El código que se muestra a continuación simula el movimiento browniano en 3 dimensiones

```
//Autor: Soto Miranda Andrea Viridiana
//Movimiento Browniano
//Instituto: CIC-IPN
//Materia: Probabilidad, procesos aleatorios e inferencia

#include "TGraph.h"
#include "TMath.h"
void movimiento_browniano3D();

void movimiento_browniano3D()
{
    TCanvas *c1 = new TCanvas("Movimiento_Browniano", "Movimiento_Browniano", 20, 10, 700, 700);

    Double_t x=0;
    Double_t y=0;
    Double_t z=0;
    Double_t theta=0;
    Double_t phi=0;
    Int_t dim = 10000;
    TRandom3 *r = new TRandom3(0);
    TGraph2D *g = new TGraph2D();
    g->SetTitle("Movimiento_Browniano");
    for (Int_t i = 0; i < dim; i++)
    {
        theta = (r->Integer(360))+1;
        phi = (r->Integer(360))+1;
        g->SetPoint(i, x, y, z);
        x=x+(cos(theta)*sin(phi));
        y=y+(sin(theta)*sin(phi));
        z=z+cos(phi);
    }
    g->SetLineWidth(1);
    g->SetLineColor(225);
    g->Draw("LINE");
}
```

En la Fig(83) se muestra la trayectoria que resulta de esta simulación.

Movimiento Browniano

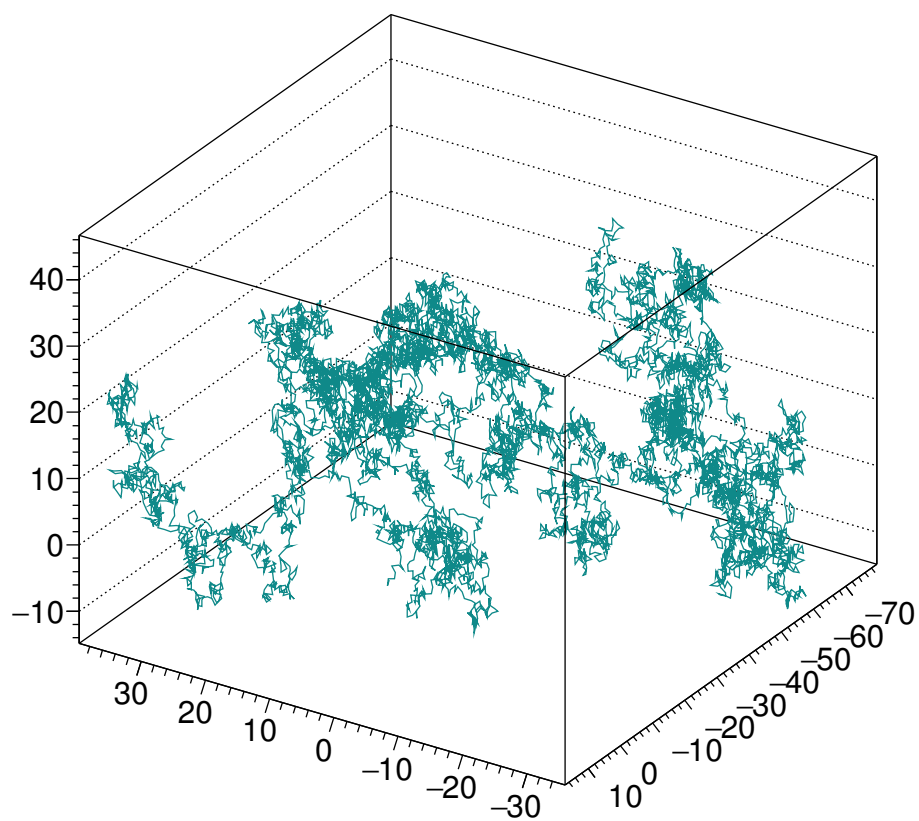


Fig. 83: Movimiento browniano en 3 dimensiones.

51 Funciones de distribución de probabilidad

	Función de distribución	Función de distribución acumulativa	Media (μ)	Varianza (σ^2)	Desviación estándar (σ)
Bernoulli $X \sim Be(p)$	$P(X = k) = p^k(1-p)^{1-k}$ $k = 0, 1$	$F(x) = \begin{cases} 0 & x < 0 \\ 1-p & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$	$\mu = p$	$\sigma^2 = p(1-p)$	$\sigma = \sqrt{p(1-p)}$
Binomial $b(X; n, p)$ $X \sim B(n, p)$ $X \sim Bi(n, p)$	$b(k; n, p) = \binom{n}{k} p^k q^{n-k}$ $\binom{n}{k} = \frac{n!}{k!(n-k)!}$	$F(x; p, n) = \sum_{k=0}^n \binom{n}{k} p^k q^{n-k}$	$\mu = np$	$\sigma^2 = npq$	$\sigma = \sqrt{npq}$
Uniforme	$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{de otra forma} \end{cases}$	$F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x < b \\ 1 & x \geq b \end{cases}$	$\mu = \frac{1}{2}(a+b)$	$\sigma^2 = \frac{1}{12}(b-a)^2$	$\sigma = \frac{b-a}{\sqrt{12}}$
Exponencial	$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$	$F(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$	$\mu = \frac{1}{\lambda}$	$\sigma^2 = \frac{1}{\lambda^2}$	$\sigma = \frac{1}{\lambda}$
Normal $N(\mu, \sigma^2)$	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$F(x) = \int_{-\infty}^x \frac{e^{-x^2/2}}{\sqrt{2\pi}}$	$\mu = \sum_{k=1}^n x_k P(x_k)$	$\sigma^2 = \sum_{k=1}^n (x_k - \mu)^2 P(x_k)$	$\sigma = \sqrt{\sum_{k=1}^n (x_k - \mu)^2 P(x_k)}$
Poisson	$f(x) = P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$ $x = 0, 1, 2, \dots$	$F(x; \lambda) = \sum_{x=0}^n \frac{e^{-\lambda} \lambda^x}{x!}$	$\mu = \lambda$	$\sigma^2 = \lambda$	$\sigma = \sqrt{\lambda}$
Hipergeométrica	$P(X = x) = \frac{\binom{b}{x} \binom{r}{n-x}}{\binom{b+r}{n}}$		$\mu = \frac{nb}{b+r}$	$\sigma^2 = \frac{nbr(b+r-n)}{(b+r)^2(b+r-1)}$	$\sigma = \sqrt{\frac{nbr(b+r-n)}{(b+r)^2(b+r-1)}}$
Cauchy	$f(x) = \frac{a}{\pi(x^2+a^2)}$ $a > 0, -\infty < x < \infty$	$F(x) = 0.5 + \frac{\arctan(x/a)}{\pi}$	No definido	No definido	No definido
Gamma	$f(x) = \begin{cases} \frac{x^{\alpha-1} e^{-x/\beta}}{\beta^\alpha \Gamma(\alpha)} & x > 0 \\ 0 & x \leq 0 \end{cases}$		$\mu = \alpha\beta$	$\sigma^2 = \alpha\beta^2$	$\sigma = \sqrt{\alpha}\beta$
Beta	$f(x) = \begin{cases} \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)} & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$		$\mu = \frac{\alpha}{\alpha+\beta}$	$\sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	$\sigma = \sqrt{\alpha}\beta$

	Gráfica de función de distribución	Gráfica de distribución acumulativa
Bernoulli		
Binomial		
Uniforme		
Exponencial		

51.1 Función de distribución de Rademacher

En teoría de la probabilidad y estadística, la distribución Rademacher (que lleva el nombre de Hans Rademacher) es una distribución discreta de probabilidad que una variable aleatoria X tiene una probabilidad del 50% de ser $+1$ o -1 .

Una serie de Rademacher distribuye las variables pueden considerarse como un simple camino aleatorio simétrico, donde el tamaño del paso es 1.

Formulación matemática La función de masa de probabilidad de esta distribución es

$$f(k) = \begin{cases} \frac{1}{2} & \text{if } k = -1 \\ \frac{1}{2} & \text{if } k = +1 \\ 0 & \text{otherwise.} \end{cases}$$

Y la función de distribución de probabilidad esta dada por:

$$F(k) = \begin{cases} 0, & k < -1 \\ \frac{1}{2}, & -1 \leq k < 1 \\ 1, & k \geq 1 \end{cases}$$

Puede también ser escrito como una función de densidad de probabilidad, en términos de la función delta de Dirac, como:

$$f(k) = \frac{1}{2}(\delta(k-1) + \delta(k+1)).$$

Aplicaciones

La distribución Rademacher se ha utilizado en bootstrapping.

La distribución Rademacher se puede utilizar para demostrar que se distribuye normalmente y no correlacionado no implica independiente.

52 Procesos NP

Los recursos comúnmente estudiados en complejidad computacional son:

- El tiempo: mediante una aproximación al número de pasos de ejecución que un algoritmo emplea para resolver un problema.
- El espacio: mediante una aproximación a la cantidad de memoria utilizada para resolver el problema.

Los problemas se clasifican en conjuntos o clases de complejidad (L, NL, P, PCompleto, NP, NP-Completo, NP Duro...). Nosotros nos vamos a centrar en las clases P y NP.

La mayoría de los algoritmos pueden ser resueltos en un tiempo Polinomial, hablando de entradas de tamaño n , en el peor de los casos el tiempo de ejecución es $O(n^k)$ para alguna constante k . Pero no todos los problemas pueden ser resueltos en tiempo polinomial, hay problemas como el famoso "Problema de parada" de Turing, el cual no puede ser resuelto por ninguna computadora, no importa el tiempo que este se ejecute. También hay problemas que pueden ser resueltos, pero no en un tiempo de ejecución $O(n^k)$ para alguna constante k .

En general, pensamos en los problemas que se pueden resolver por medio de algoritmos en tiempo polinómico como manejable, o fácil, y los problemas que requieren tiempo superpolynomial como intratable, o difíciles.

La clase NP

La clase NP está compuesta por los problemas que tienen un certificado sucinto (también llamado testigo polinómico) para todas las instancias cuya respuesta es un Sí. La única forma de que tengan un tiempo polinomial es realizando una etapa aleatoria, incluyendo el azar de alguna manera para elegir una posible solución, y entonces en etapas posteriores comprueba si esa solución es correcta.

En otras palabras, dada una solución para una cierta instancia, es posible comprobar que es válida en tiempo $O(n^k)$. En el caso de SAT (Problema de satisfacibilidad booleana), dado una asignación de valores de verdad, se puede comprobar fácilmente si la fórmula es cierta o no. Una NMT puede "adivinar" la solución en $O(n)$ y verificarla en tiempo polinómico.

Complejidad de NP

Para analizar la pregunta $P = NP$, resulta muy útil el concepto de complejidad NP. De manera informal, los problemas de complejidad NP son los problemas más "difíciles" en NP en el sentido de que ellos son los que son más probable no se encuentren en P. Problemas NP-difíciles son aquellos para los cuales cualquier problema en NP puede ser reducido en tiempo polinómico. Los problemas de complejidad NP son aquellos problemas NP-difícil que se encuentran en NP. Por ejemplo, la versión de problema de decisión del problema del vendedor viajero es completamente NP. Así ningún caso de ningún problema en NP puede ser transformado mecánicamente en una parte del problema del vendedor viajero, en tiempo polinómico. Por lo tanto, si el problema del vendedor viajero estuviera contenido en P, entonces $P = NP$. El problema del vendedor viajero es uno de muchos problemas NP-completos. Si cualquier problema NP-completo se encuentra contenido en P, entonces se verificaría que $P = NP$. Desafortunadamente, se ha demostrado que muchos problemas importantes son NP-completos y no se conoce la existencia de ningún algoritmo rápido para ellos.

La definición anterior de NP permite considerar de manera natural una clase de problemas complementarias. La co-NP está compuesta por los problemas que tienen un contraejemplo sucinto para todas las instancias cuya respuesta es NO.

53 Matrices

53.1 Matrices semejantes

Dos matrices cuadradas A y B se dice que son semejantes si existe una matriz cuadrada regular P tal que:

$$B = P \cdot A \cdot P^{-1}$$

Propiedades Las matrices semejantes comparten las siguientes propiedades:

(1) La semejanza es una relación de equivalencia.

- Reflexiva. Está claro que toda matriz es semejante a sí misma.
- Simétrica. Si existe P invertible tal que $A = PBP^{-1}$ entonces existe Q invertible tal que $B = QAQ^{-1}$ (basta tomar $Q = P^{-1}$)

$$A \sim B \Rightarrow B \sim A$$

$$B = PAP^{-1} \Rightarrow B = P^{-1}B(P^{-1})^{-1}$$

- Transitiva. Si existe P, Q invertibles tales que $A = PBP^{-1}$ y $B = QCQ^{-1}$ entonces existe R invertible tal que $A = RCR^{-1}$ (basta tomar $T=PQ$). Por lo tanto,

$$\left. \begin{matrix} A \sim B \\ B \sim C \end{matrix} \right\} \Rightarrow A \sim C$$

(2) Dos matrices semejantes son equivalentes.

(3) Dos matrices semejantes tienen la misma dimensión.

(4) Dos matrices semejantes tienen el mismo determinante.

(5) Dos matrices semejantes tienen el mismo rango.

Ejemplo Comprobar que las matrices A y B son semejantes considerando la matriz de paso Q.

$$A = \begin{bmatrix} 6 & 1 & -1 \\ 0 & 5 & 1 \\ 2 & -1 & 1 \end{bmatrix}$$

$$B = \begin{bmatrix} 4 & 0 & 2 \\ 0 & 4 & 2 \\ 0 & 2 & 4 \end{bmatrix}$$

$$Q = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & -1 \end{bmatrix}$$

Para poder comprobar que las matrices A y B son semejantes tomando como matriz de paso Q se necesita verificar que $A = Q^{-1}BQ$. Para ellos se plantea:

$$\begin{bmatrix} 6 & 1 & -1 \\ 0 & 5 & 1 \\ 2 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 & 1/2 \\ -1/2 & 1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 \end{bmatrix} \cdot \begin{bmatrix} 4 & 0 & 2 \\ 0 & 4 & 2 \\ 0 & 2 & 4 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & -1 \end{bmatrix}$$

$$\begin{bmatrix} 6 & 1 & -1 \\ 0 & 5 & 1 \\ 2 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 4 \\ -2 & 3 & 2 \\ 2 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & -1 \end{bmatrix}$$

$$\begin{bmatrix} 6 & 1 & -1 \\ 0 & 5 & 1 \\ 2 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 1 & -1 \\ 0 & 5 & 1 \\ 2 & -1 & 1 \end{bmatrix}$$

De esta manera se comprueba que A y B son semejantes.

53.2 Eigen Valores y Eigen Vectores

En álgebra lineal, los vectores propios, autovectores o eigenvectores de un operador lineal son los vectores no nulos que, cuando son transformados por el operador, dan lugar a un múltiplo escalar de sí mismos, con lo que no cambian su dirección. Este escalar λ recibe el nombre valor propio, autovalor, valor característico o eigenvalor. A menudo, una transformación queda completamente determinada por sus vectores propios y valores propios. Un espacio propio, autoespacio, eigenespacio o subespacio fundamental asociado al valor propio λ es el conjunto de vectores propios con un valor propio común.

El polinomio $p(\lambda) = \det(\mathbf{A} - \lambda \mathbf{I})$ es llamado el polinomio característico de \mathbf{A} . Las raíces $p(\lambda) = 0$ del polinomio son eigenvalores de \mathbf{A} .

Un vector $\mathbf{v} \neq 0$ que satisface $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ es un eigenvector de \mathbf{A} . Al par ordenado (λ, \mathbf{v}) se le llama eigenpar.

Ejemplos de eigenvalores y eigen vectores

(1) Sea una matriz

$$A = \begin{pmatrix} 4 & 1 \\ -6 & -3 \end{pmatrix}$$

Determinar sus valores y vectores propios (eigenvalores y eigenvectores).

La matriz identidad I multiplicada por λ para matrices 2×2 queda

$$\lambda I = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$$

La diferencia

$$(\lambda I - A) = \begin{pmatrix} \lambda - 4 & -1 \\ 6 & \lambda + 3 \end{pmatrix}$$

Cuyo determinante se puede escribir

$$\begin{aligned} \det(\lambda I - A) &= (\lambda - 4)(\lambda + 3) + 6 \\ &= \lambda^2 - \lambda - 12 + 6 = \lambda^2 - \lambda - 6 \end{aligned}$$

que es el polinomio característico. Mediante factorización o por la fórmula general obtenemos sus raíces, o valores característicos,

$$\lambda_1 = 3, \quad \lambda_2 = -2$$

Para encontrar los vectores propios (eigenvectores) de la matriz A se utiliza $(\lambda I - A)\vec{v} = 0$ donde \vec{v} es el vector propio:

$$(\lambda I - A)\vec{v} = \begin{pmatrix} \lambda - 4 & -1 \\ 6 & \lambda + 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Substituyendo $\lambda_1 = 3$ se tiene

$$\begin{pmatrix} 3 - 4 & -1 \\ 6 & 3 + 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ 6 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

que representa el siguiente sistema homogéneo de ecuaciones lineales:

$$-x - y = 0 \quad 6x + 6y = 0$$

cuya solución evidentemente es

$$x = -y$$

y y es variable libre. De modo que el vector propio que corresponde a λ_1 es:

$$\vec{v}_1 = \begin{pmatrix} -y \\ y \end{pmatrix} = y \begin{pmatrix} -1 \\ 1 \end{pmatrix} =$$

Si $y = 1$

$$\vec{v}_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix} =$$

Verificando que $A\vec{v}_1 = \lambda_1 \vec{v}_1$

$$\begin{pmatrix} 4 & 1 \\ -6 & -3 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = 3 \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} -4 + 3 \\ 6 - 3 \end{pmatrix} = \begin{pmatrix} -3 \\ 3 \end{pmatrix}$$

$$\begin{pmatrix} -3 \\ 3 \end{pmatrix} = \begin{pmatrix} -3 \\ 3 \end{pmatrix}$$

El segundo vector propio se obtiene, de la misma manera, substituyendo $\lambda_2 = -2$, y queda

$$\begin{pmatrix} -2 - 4 & -1 \\ 6 & -2 + 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -6 & -1 \\ 6 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

de donde el vector \vec{v}_2 es

$$\vec{v}_2 = \begin{pmatrix} -\frac{1}{6}y \\ y \end{pmatrix} = y \begin{pmatrix} -\frac{1}{6} \\ 1 \end{pmatrix}$$

Si $y = 6$

$$\vec{v}_2 = \begin{pmatrix} -1 \\ 6 \end{pmatrix}$$

Verificando que $A\vec{v}_2 = \lambda_2 \vec{v}_2$

$$\begin{pmatrix} 4 & 1 \\ -6 & -3 \end{pmatrix} \begin{pmatrix} -1 \\ 6 \end{pmatrix} = -2 \cdot \begin{pmatrix} -1 \\ 6 \end{pmatrix}$$

$$\begin{pmatrix} -4+6 \\ 6-18 \end{pmatrix} = \begin{pmatrix} 2 \\ -12 \end{pmatrix}$$

$$\begin{pmatrix} 2 \\ -12 \end{pmatrix} = \begin{pmatrix} 2 \\ -12 \end{pmatrix}$$

(2) Sea una matriz

$$A = \begin{pmatrix} 2 & -2 & 3 \\ 10 & -4 & 5 \\ 5 & -4 & 6 \end{pmatrix}$$

Se busca que el determinante de la matriz $(\lambda I - A)$ sea cero:

$$\begin{aligned} (\lambda I - A) &= \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 2 & -2 & 3 \\ 10 & -4 & 5 \\ 5 & -4 & 6 \end{pmatrix} \\ &= \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} - \begin{pmatrix} 2 & -2 & 3 \\ 10 & -4 & 5 \\ 5 & -4 & 6 \end{pmatrix} \\ &= \begin{pmatrix} \lambda-2 & 2 & -3 \\ -10 & \lambda+4 & -5 \\ -5 & 4 & \lambda-6 \end{pmatrix} \end{aligned}$$

El determinante de la matriz igualado a cero, está dado por

$$\begin{aligned} \det(\lambda I - A) &= \begin{vmatrix} \lambda-2 & 2 & -3 \\ -10 & \lambda+4 & -5 \\ -5 & 4 & \lambda-6 \end{vmatrix} = \\ &= \lambda^3 - 4\lambda^2 + 5\lambda - 2 = 0 \end{aligned}$$

Cuyas raíces son

$$\lambda_1 = 1$$

$$\lambda_2 = 1$$

$$\lambda_3 = 2$$

que son los valores propios o eigenvalores de la matriz A.

Para encontrar los vectores propios o eigenvalores de la matriz A se utiliza $(\lambda I - A)\vec{v} = 0$ donde \vec{v} es el vector propio:

$$(\lambda I - A)\vec{v} = \begin{pmatrix} \lambda-2 & 2 & -3 \\ -10 & \lambda+4 & -5 \\ -5 & 4 & \lambda-6 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Se sustituyen los valores propios $\lambda_1 = \lambda_2 = 1$, y se resuelve el sistema homogéneo de ecuaciones lineales resultante

$$(\lambda I - A)\vec{v} = \begin{pmatrix} -1 & 2 & -3 \\ -10 & 5 & -5 \\ -5 & 4 & -5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$-x + 2y - 3z = 0$$

$$-10x + 5y - 5z = 0$$

$$-5x + 4y - 5z = 0$$

cuyas soluciones son:

$$\frac{1}{5}y$$

$$\frac{3}{5}y$$

El vector propio, que por ahora se denominará \vec{E}_1 se puede expresar como

$$\vec{E}_1 = \begin{pmatrix} \frac{1}{5}y \\ y \\ \frac{3}{5}y \end{pmatrix} = \frac{1}{5}y \begin{pmatrix} 1 \\ 5 \\ 3 \end{pmatrix}$$

Si $y = 5$

$$\vec{E}_1 = \begin{pmatrix} 1 \\ 5 \\ 3 \end{pmatrix}$$

Para el valor propio $\lambda_3 = 2$

$$\begin{pmatrix} 0 & 2 & -3 \\ -10 & 6 & -5 \\ -5 & 4 & -4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

se obtiene el vector propio

$$\vec{E}_2 = \begin{pmatrix} \frac{4}{15}y \\ y \\ \frac{10}{15}y \end{pmatrix} = \frac{1}{15}y \begin{pmatrix} 4 \\ 15 \\ 10 \end{pmatrix}$$

Si $y = 15$

$$\vec{E}_2 = \begin{pmatrix} 4 \\ 15 \\ 10 \end{pmatrix}$$

(3) Considere la matriz dada por:

$$A = \begin{pmatrix} 8 & 2 & 4 & 1 \\ 2 & 6 & -2 & 1 \\ 4 & 3 & -8 & 0 \\ -2 & 3 & 4 & 5 \end{pmatrix}$$

La ecuación característica de la matriz **A** esta dada por

$$0 = |A - \lambda I_4| = p(\lambda) = -1636 + 756\lambda - 49\lambda^2 - 11\lambda^3 + \lambda^4$$

La siguiente figura muestra la gráfica de la ecuación característica, es evidente, de la figura que existen 4 raíces de la ecuación característica y, por lo tanto, los eigenvalores de la matriz **A** están dadas por

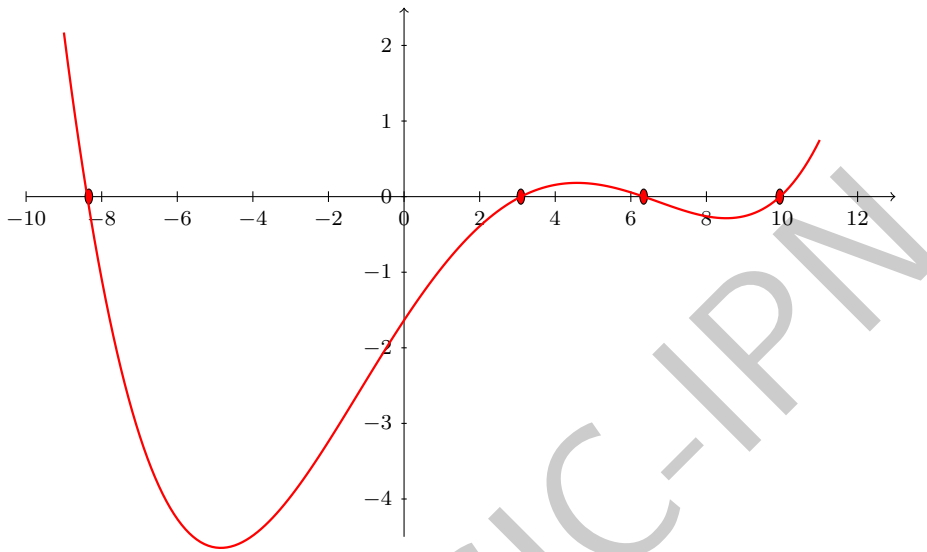


Fig. 84: Gráfica de la Ecuación Característica de la Matriz A

$$\lambda_1 = -8.3824, \lambda_2 = 3.09465, \lambda_3 = 6.3394, \lambda_4 = 9.9483$$

Para la determinación de los eigenvectores asociados a $\lambda_1 = -8.3824$, se tienen que realizar los siguientes cálculos.

- Determinar la matriz **A** - $\lambda_1 I_4$, dada por

$$A - \lambda_1 I_4 = \begin{pmatrix} 16.3824 & 2 & 4 & 1 \\ 2 & 14.3824 & -2 & 1 \\ 4 & 3 & 0.3824 & 0 \\ -2 & 3 & 4 & 13.3824 \end{pmatrix}$$

ida de la matriz

- Obtener la forma escalonada reducida de la matriz **A** - $\lambda_1 I_4$, dada por

$$\begin{pmatrix} 16.3824 & 2.0 & 4.0 & 1.0 \\ 0 & 14.1382 & -2.4883 & 0.8779 \\ 0 & 0 & 5.0592 & 13.3030 \\ 0 & 0 & 0 & 0.0000000009 \end{pmatrix}$$

- Despreciando el término (4,4), el eigenvector asociado al eigenvalor λ_1 está dado por la solución al sistema

$$\begin{pmatrix} 16.3824 & 2.0 & 4.0 & 1.0 \\ 0 & 14.1382 & -2.4883 & 0.8779 \\ 0 & 0 & 5.0592 & 13.3030 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

La solución está dada por

$$\vec{b}_{\lambda_1} = \begin{pmatrix} -0.245318 \\ 0.199614 \\ 1 \\ -0.380310 \end{pmatrix}$$

Es importante señalar que cualquier múltiplo escalar de \vec{b}_{λ_1} es igualmente un eigenvector de la matriz **A** asociado al eigenvalor λ_1 .

Procediendo de igual manera con los demás eigenvalores, se tiene que

- 1.- Para $\lambda_2 = 3.09465$, se tiene que

$$\vec{b}_{\lambda_2} = \begin{pmatrix} 0.064788 \\ -0.457957 \\ -0.100473 \\ 1 \end{pmatrix}$$

- 2.- Para $\lambda_3 = 6.3394$, se tiene que

$$\vec{b}_{\lambda_3} = \begin{pmatrix} -0.58024 \\ 0.215377 \\ -0.116801 \\ 1 \end{pmatrix}$$

- 3.- Para $\lambda_4 = 9.94834$, se tiene que

$$\vec{b}_{\lambda_4} = \begin{pmatrix} 1 \\ 0.37564 \\ 0.28564 \\ 0.05446 \end{pmatrix}$$

- (4) Encontrar los eigenvectores de la matriz:

$$A = \begin{pmatrix} 1 & 4 \\ 3 & 5 \end{pmatrix}$$

Para encontrar los eigenvalores

$$(\lambda I - A) = \begin{pmatrix} \lambda - 1 & -4 \\ -3 & \lambda - 5 \end{pmatrix}$$

$$\det(\lambda I - A)(\lambda - 1)(\lambda - 5) - 12 = \lambda^2 - 6\lambda - 7$$

Se calculan las raíces que en este caso son:

$$\lambda_1 = -1, \lambda_2 = 7$$

Debido a que se trata de una matriz de 2×2 , el eigenvector propuesto es

$$\vec{v}_1 = \begin{pmatrix} a \\ b \end{pmatrix}$$

Para el primer eigenvalor $\lambda_1 = -1$:

$$\begin{pmatrix} 1 & 4 \\ 3 & 5 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = -1 \begin{pmatrix} a \\ b \end{pmatrix}$$

Multiplicando:

$$\begin{pmatrix} a + 4b \\ 3a + 5b \end{pmatrix} = \begin{pmatrix} -a \\ -b \end{pmatrix}$$

De esta forma:

$$\begin{pmatrix} a + 4b \\ 3a + 5b \end{pmatrix} - \begin{pmatrix} -a \\ -b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 2a + 4b \\ 3a + 6b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

y nos da las siguientes ecuaciones:

$$2a + 4b = 0 \quad 3a + 6b = 0$$

que nos da: $a = -2b$, para ambas ecuaciones. Entonces el eigenvector está formado de la siguiente forma:

$$\vec{v}_1 = \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -2b \\ b \end{pmatrix} = b \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

En forma general el valor de b puede ser cualquier valor real. Así entonces un eigenvector correspondiente a $\lambda_1 = -1$ es:

$$\vec{v}_1 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

Para el segundo eigenvalor $\lambda_2 = 7$

$$\begin{pmatrix} 1 & 4 \\ 3 & 5 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = 7 \begin{pmatrix} a \\ b \end{pmatrix}$$

Multiplicando:

$$\begin{pmatrix} a + 4b \\ 3a + 5b \end{pmatrix} = \begin{pmatrix} 7a \\ 7b \end{pmatrix}$$

De esta forma:

$$\begin{pmatrix} a + 4b \\ 3a + 5b \end{pmatrix} - \begin{pmatrix} 7a \\ 7b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} -6a + 4b \\ 3a - 2b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

y nos da las siguientes ecuaciones:

$$-6a + 4b = 0 \quad 3a - 2b = 0$$

que nos da: $a = \frac{2}{3}b$, para ambas ecuaciones. Entonces el eigenvector está formado de la siguiente forma:

$$\vec{v}_2 = \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \frac{2}{3}b \\ b \end{pmatrix} = b \begin{pmatrix} \frac{2}{3} \\ 1 \end{pmatrix}$$

En forma general el valor de b puede ser cualquier valor real. Así entonces un eigenvector correspondiente a $\lambda_2 = 7$ es:

$$\vec{v}_2 = \begin{pmatrix} \frac{2}{3} \\ 1 \end{pmatrix}$$

Otro vector que pertenece a la familia es cuando $b=3$

$$\vec{v}_2 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

Cualquiera de estos dos eigenvectores es correspondiente a $\lambda_2 = 7$

(4) Encontrar los eigenvectores de la siguiente matriz

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \\ 4 & -4 & 5 \end{pmatrix}$$

Dados los eigenvalores $\lambda_1 = 1$, $\lambda_2 = 2$ y $\lambda_3 = 3$. Cada uno de estos eigenvalores tiene un conjunto de eigenvectores correspondientes. Como se trata de una matriz de 3×3 , el eigenvector propuesto tiene que ser de dimensiones 3×1 :

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

Para el primer eigenvalor $\lambda_1 = 1$:

$$\begin{pmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \\ 4 & -4 & 5 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 1 \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

Multiplicando:

$$\begin{pmatrix} a + 2b - c \\ a + c \\ 4a - 4b + 5c \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

De esta forma

$$\begin{pmatrix} a + 2b - c \\ a + c \\ 4a - 4b + 5c \end{pmatrix} - \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 2b - c \\ a - b + c \\ 4a - 4b + 4c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

y nos da las siguientes ecuaciones:

$$2b - c = 0 \quad a - b + c = 0 \quad 4a - 4b + 4c = 0$$

De las cuales la segunda y la tercera son equivalentes. La primera nos da: $c = 2b$, sustituyendo en la segunda ecuación, tenemos:

$$a - b + c = a - b + 2b = a + b = 0$$

Entonces:

$$a = -b$$

Entonces el eigenvector está formado de la siguiente forma:

$$\vec{v}_1 = \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} -b \\ b \\ 2b \end{pmatrix} = b \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}$$

Para el primer eigenvalor $\lambda_2 = 2$:

$$\begin{pmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \\ 4 & -4 & 5 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 2 \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

Multiplicando:

$$\begin{pmatrix} a + 2b - c \\ a + c \\ 4a - 4b + 5c \end{pmatrix} = \begin{pmatrix} 2a \\ 2b \\ 2c \end{pmatrix}$$

De esta forma

$$\begin{pmatrix} a + 2b - c \\ a + c \\ 4a - 4b + 5c \end{pmatrix} - \begin{pmatrix} 2a \\ 2b \\ 2c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} -a + 2b - c \\ a - 2b + c \\ 4a - 4b + 3c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

y nos da las siguientes ecuaciones:

$$-a + 2b - c = 0 \quad a - 2b + c = 0 \quad 4a - 4b + 3c = 0$$

De las cuales la primera y la segunda son equivalentes. Multiplicando la primera por cuatro y sumándola con la tercera nos da:

$$-4a + 8b - 4c = 0 \quad 4a - 4b + 3c = 0 \quad 4b - c = 0$$

Así

$$c = 4b$$

Sustituyendo en la segunda ecuación, tenemos:

$$a - 2b + c = a - 2b + 4b = a + 2b = 0$$

Entonces:

$$a = -2b$$

Entonces el eigenvector está formado de la siguiente forma:

$$\vec{v}_2 = \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} -2b \\ b \\ 4b \end{pmatrix} = b \begin{pmatrix} -2 \\ 1 \\ 4 \end{pmatrix}$$

En forma general el valor de b puede ser cualquier valor real. Así entonces un eigenvector correspondiente a $\lambda_2 = 2$ es:

$$\vec{v}_2 = \begin{pmatrix} -2 \\ 1 \\ 4 \end{pmatrix}$$

Para el primer eigenvalor $\lambda_3 = 3$:

$$\begin{pmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \\ 4 & -4 & 5 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 3 \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

Multiplicando:

$$\begin{pmatrix} a + 2b - c \\ a + c \\ 4a - 4b + 5c \end{pmatrix} = \begin{pmatrix} 3a \\ 3b \\ 3c \end{pmatrix}$$

De esta forma

$$\begin{pmatrix} a + 2b - c \\ a + c \\ 4a - 4b + 5c \end{pmatrix} - \begin{pmatrix} 3a \\ 3b \\ 3c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} -2a + 2b - c \\ a - 3b + c \\ 4a - 4b + 2c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

y nos da las siguientes ecuaciones:

$$-2a + 2b - c = 0 \quad a - 3b + c = 0 \quad 4a - 4b + 2c = 0$$

De las cuales la primera y la segunda son equivalentes. Multiplicando la segunda por -4 y sumándola con la tercera nos da:

$$-4a + 12b - 4c = 0 \quad 4a - 4b + 3c = 0 \quad 8b - 2c = 0$$

Así

$$c = 4b$$

Sustituyendo en la segunda ecuación, tenemos:

$$a - 3b + c = a - 3b + 4b = a + b = 0$$

Entonces:

$$a = -b$$

Entonces el eigenvector está formado de la siguiente forma:

$$\vec{v}_3 = \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} -b \\ b \\ 4b \end{pmatrix} = b \begin{pmatrix} -1 \\ 1 \\ 4 \end{pmatrix}$$

En forma general el valor de b puede ser cualquier valor real. Así entonces un eigenvector correspondiente a $\lambda_3 = 3$ es:

$$\vec{v}_3 = \begin{pmatrix} -1 \\ 1 \\ 4 \end{pmatrix}$$

53.3 Matrices cuadradas poco densas

Una matriz poco densa es aquella que está formada por elementos que en su mayoría son ceros, como se muestra en la siguiente matriz.

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 5 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 6 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 7 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 8 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Este tipo de matrices son matrices cuadradas que se dividen en los siguientes tipos:

- Matriz triangular superior

En este tipo de matriz los elementos iguales a cero se encuentran debajo de la diagonal. Ejemplo:

$$M = \begin{pmatrix} 5 & 8 & 2 & 7 \\ 4 & -3 & -6 & 0 \\ 9 & -5 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

- Matriz triangular inferior

En este tipo de matrices los elementos iguales a cero se encuentran por encima de la diagonal principal. Ejemplo:

$$M = \begin{pmatrix} 5 & 0 & 0 & 0 \\ 8 & 2 & 0 & 0 \\ 6 & 1 & -4 & 0 \\ 3 & 7 & 6 & 9 \end{pmatrix}$$

- Matriz tridiagonal

En este tipo de matrices los elementos diferentes de cero se encuentran en la diagonal principal o en las diagonales por debajo o encima de ella. Ejemplo:

$$M = \begin{pmatrix} 5 & 8 & 0 & 0 \\ 4 & 5 & 2 & 0 \\ 0 & 9 & -5 & 2 \\ 0 & 0 & 9 & -5 \end{pmatrix}$$

54 Cadenas de Markov

Las cadenas de Markov a un tipo especial de proceso estocástico discreto en el que la probabilidad de que ocurra un evento depende solamente del evento inmediatamente anterior. Esta característica de falta de memoria recibe el nombre de propiedad de Markov.

En matemática se define como un proceso estocástico discreto que cumple con la propiedad de Markov, es decir, si se conoce la historia del sistema hasta su instante actual, su estado presente resume toda la información relevante para describir en probabilidad su estado futuro.

Una cadena de Markov es una secuencia X_1, X_2, X_3, \dots de variables aleatorias. El dominio de estas variables es llamado espacio estado; el valor de X_n es el estado del proceso en el tiempo n . Si la distribución de probabilidad condicional de X_{n+1} en estados pasados es una función de X_n por sí sola, entonces:

$$P(X_{n+1} = x_{n+1} | X_n = x_n, X_{n-1} = x_{n-1}, \dots, X_1 = x_1) = P(X_{n+1} = x_{n+1} | X_n = x_n)$$

Las cadenas de Markov deben cumplir las siguientes propiedades:

- (i) Cada resultado pertenece a un conjunto finito de resultados a_1, a_2, \dots, a_n llamado espacio de estados del sistema; si el resultado de la n –ésima prueba es a_i , entonces decimos que el sistema está en estado a_i en la vez n o en el paso n –ésimo.
- (ii) El resultado de una prueba depende a lo sumo del resultado de la prueba inmediatamente precedente y no de cualquier otro resultado previo; con cada par de estados (a_i, a_j) se establece la probabilidad p_{ij} de que a_j suceda inmediatamente después de que suceda a_i .

A un proceso estocástico tal, se llama *cadena de Markov* (finita). Los números p_{ij} , llamados probabilidades de transición, pueden ordenarse en una matriz

$$\begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{pmatrix}$$

llamada matriz de transición.

Así, para cada estado a_i corresponde la i –ésima fila $(p_{i1}, p_{i2}, \dots, p_{im})$ de la matriz de transición P ; si el sistema está en estado a_i entonces este vector fila representa las probabilidades de todos los resultados posibles de la prueba siguiente y, por tanto, es un vector de probabilidad.

El elemento p_{ij} en la matriz de transición P de una cadena de Markov es la probabilidad de que el sistema cambie del estado a_i al estado a_j en un paso: $a_i \rightarrow a_j$.

Ahora supóngase que después de un tiempo arbitrario, la probabilidad de que el sistema esté en estado a_i es p_i ; denotamos estas probabilidades por el vector de probabilidad $p = (p_1, p_2, \dots, p_m)$ que se denomina *distribución de probabilidad* del sistema para tal tiempo. En particular, denotaremos por

$$p^{(0)} = (p_1^{(0)}, p_2^{(0)}, \dots, p_m^{(0)})$$

la *distribución de probabilidad* de paso n –ésimo, o sea la distribución después de los primeros n pasos.

54.1 Ejemplo de Cadenas de Markov

- 1.- Los hábitos de estudio de un estudiante son como sigue. Si estudia una noche, está 70% seguro de no estudiar la noche siguiente. Por otra parte, si no estudia una noche, está 60% seguro de no estudiar tampoco la noche siguiente. A la larga, ¿con qué frecuencia estudia?

Los estados del sistema son S (de estudiar) y T (no estudiar). La matriz de transición es:

$$P = \begin{matrix} & \begin{matrix} S & T \end{matrix} \\ \begin{matrix} S \\ T \end{matrix} & \begin{pmatrix} 0.3 & 0.7 \\ 0.4 & 0.6 \end{pmatrix} \end{matrix}$$

para averiguar qué sucede a la larga, tenemos que hallar el vector de probabilidad fijo único t de P .

buscamos un vector de probabilidad de la forma $t = (x, 1 - x)$ tal que $tA = t$

$$(x \quad 1 - x) \begin{pmatrix} 0.3 & 0.7 \\ 0.4 & 0.6 \end{pmatrix} = (x \quad 1 - x)$$

Multiplicando queda:

$$0.3x + 0.4 - 0.4x = x, \quad 0.7x + 0.6 - 0.6x = 1 - x$$

Resolvemos las ecuaciones y queda:

$$x = \frac{4}{11}$$

por lo tanto el vector de probabilidad queda $t = (4/11, 7/11)$, por lo que el estudiante estudia 4/11 de las veces.

- 2.- Un psicólogo hace los supuestos siguientes que conciernen al comportamiento de ratas sujetas a régimen especial de alimentación. Para una prueba particular, 80% de las ratas que giraron para la derecha e, el experimento previo hicieron lo mismo en esta prueba, y 60% de aquellas que giraron para la izquierda en el experimento previo, giraron hacia la derecha en esta prueba. Si 50% van a la derecha en la primera prueba, ¿ qué se podría predecir para, (i) la segunda prueba?, (ii) ¿la tercera prueba?, (iii) ¿la milésima prueba?

Los estados del sistema son R(derecha) y L(izquierda). La matriz de transición es:

$$P = \begin{matrix} & \begin{matrix} R & L \end{matrix} \\ \begin{matrix} R \\ L \end{matrix} & \begin{pmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{pmatrix} \end{matrix}$$

La distribución de probabilidad para la primera prueba es $p = (0.5, 0.5)$. Para calcular la distribución de probabilidad para el paso siguiente, esto es, la segunda prueba, multiplicamos p por la matriz de transición P :

$$(0.5, 0.5) \begin{pmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{pmatrix} = (0.7, 0.3)$$

Por consiguiente, en la segunda prueba se espera que el 70% de las ratas vayan a la derecha y el 30% a la izquierda. Para calcular la distribución de probabilidad para la tercera prueba, multiplicamos la de la segunda prueba por P :

$$(0.7, 0.3) \begin{pmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{pmatrix} = (0.74, 0.26)$$

Así en la tercera prueba se espera que el 74% de las ratas se dirijan a la derecha y el 26% se dirija hacia la izquierda.

Suponemos que la distribución de probabilidad para la milésima prueba es esencialmente la distribución de probabilidad estacionaria de la cadena de Markov, esto es, el vector de probabilidad fijo único t de la matriz de transición P .

Calculando el vector de probabilidad fijo único t de la matriz de transición P

$$(x \quad 1 - x) \begin{pmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{pmatrix} = (x \quad 1 - x)$$

Multiplicando queda:

$$0.8x + 0.6 - 0.6x = x, \quad 0.2x + 0.4 - 0.4 = 1 - x$$

Resolvemos las ecuaciones para encontrar el valor de x :

$$x = \frac{3}{4}$$

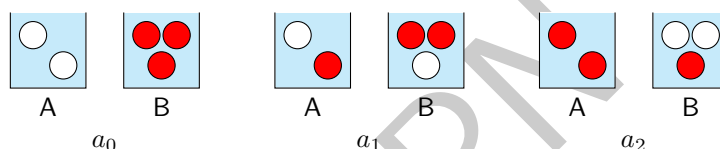
Y el vector queda

$$t = \left(\frac{3}{4}, 1 - \frac{3}{4}\right) = \left(\frac{3}{4}, \frac{1}{4}\right)$$

Por tanto se espera que la prueba mil 75% de las ratas giren a la derecha y el 25% restante hacia la izquierda.

- 3.- Hay 2 bolas blancas en una urna A y 3 en una urna B . A cada paso del proceso se selecciona una bola de cada urna y las dos bolas escogidas se intercambian. Sea el estado a_i del sistema el número i de bolas rojas de la urna A . (i) Hallas la matriz de transición P . (ii) ¿Cuál es la probabilidad de que haya 2 bolas rojas en la urna A después de 3 pasos? (iii) A la larga, ¿cuál es la probabilidad de que haya 2 bolas rojas en la urna A ?

(i) Hay tres estados, a_0, a_1, a_2 y se describen en los siguientes diagramas:



Si el sistema está en el estado a_0 , entonces tiene que escogerse una bola blanca de la urna A y una roja de la urna B , así que el sistema tiene que pasar al estado a_1 . En consecuencia, la primera fila de la matriz de transición es $(0, 1, 0)$.

Supongamos que el sistema está en un estado a_1 . Puede pasar al estado a_0 si y sólo si se escoge una bola roja de la urna A y una blanca de la urna B ; la probabilidad de que suceda esto es $\frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$. Así, $p_{10} = \frac{1}{6}$. El sistema puede pasar del estado a_1 al a_2 si y sólo si se selecciona una bola blanca de la urna A y una roja de la urna B ; la probabilidad de que suceda esto es $\frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}$. Así $p_{12} = \frac{1}{3}$. En consecuencia, la probabilidad de que el sistema permanezca en el estado a_1 es $p_{11} = 1 - \frac{1}{6} - \frac{1}{3} = \frac{1}{2}$. Así, la segunda fila de la matriz de transición es $(\frac{1}{6}, \frac{1}{2}, \frac{1}{3})$. (Obsérvese que p_{11} también se puede obtener del hecho de que el sistema permanezca en estado a_1 si se saca o una bola blanca de cada urna, con probabilidad $\frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$, o una bola roja de cada urna con probabilidad $\frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}$; así, $p_{11} = \frac{1}{6} + \frac{1}{3} = \frac{1}{2}$.)

Ahora supongamos que el sistema está en el estado a_2 . Tiene que sacarse una bola roja de la urna A . Si se escoge una bola roja de la urna B , con probabilidad $\frac{1}{3}$, entonces el sistema permanece en estado a_2 y se escoge una bola blanca de la urna B , con probabilidad de $\frac{2}{3}$, entonces el sistema pasa al estado a_1 . Noté que el sistema nunca puede pasar del estado a_2 al estado a_0 . Así, la fila tercera de la matriz de transición es $(0, \frac{2}{3}, \frac{1}{3})$. Esto es,

$$P = \begin{matrix} & \begin{matrix} a_0 & a_1 & a_2 \end{matrix} \\ \begin{matrix} a_0 \\ a_1 \\ a_2 \end{matrix} & \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ 0 & \frac{2}{3} & \frac{1}{3} \end{pmatrix} \end{matrix}$$

(ii) El sistema empieza en estado a_0 , esto es, $p^{(0)} = (1, 0, 0)$. Así;

$$p^{(1)} = p^{(0)}P = (0, 1, 0), \quad p^{(2)} = p^{(1)}P = \left(\frac{1}{6}, \frac{1}{2}, \frac{1}{3}\right), \quad p^{(3)} = p^{(2)}P = \left(\frac{1}{12}, \frac{23}{36}, \frac{5}{18}\right)$$

En consecuencia, la probabilidad de que haya 2 bolas rojas en la urna A después de 3 pasos es $\frac{5}{18}$.

(iii) Buscamos el vector de probabilidad fijo único t de la matriz de transición P . Primero hallamos un vector fijo $u = (x, y, z)$:

$$(x, y, z) \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ 0 & \frac{2}{3} & \frac{1}{3} \end{pmatrix} = (x, y, z) \quad \text{o} \quad \begin{cases} \frac{1}{6}y = x \\ x + \frac{1}{2}y + \frac{2}{3}z = y \\ \frac{1}{3}y + \frac{1}{3}z = z \end{cases}$$

Hacemos, $s = 1$. Luego por la primera ecuación $y = 6$ y por la tercera ecuación $z = 3$. Por tanto, $u = (1, 6, 3)$. Multipliquemos u por $\frac{1}{(1+6+3)} = \frac{1}{10}$ para obtener, el vector de probabilidad fijo único $t = (0.1, 0.6, 0.3)$ requerido. Así a la larga, 30% de las veces habrá 2 bolas rojas en la urna A.

4.- Considérense lanzamientos repetidos de un dado corriente. Sea X_n el máximo de los números que resulten en las n primeras pruebas.

(i) Hallar la matriz de transición P de la cadena de Markov. ¿La matriz es regular?

(ii) Hallar $p^{(1)}$, la distribución de probabilidad después del primer lanzamiento.

(iii) Hallar $p^{(2)}$ y $p^{(3)}$.

(i) El espacio de estados de la cadena de Markov es $\{1, 2, 3, 4, 5, 6\}$. La matriz de transición es:

$$P = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \begin{pmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ 0 & \frac{2}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ 0 & 0 & \frac{3}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ 0 & 0 & 0 & \frac{4}{6} & \frac{1}{6} & \frac{1}{6} \\ 0 & 0 & 0 & 0 & \frac{5}{6} & \frac{1}{6} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}$$

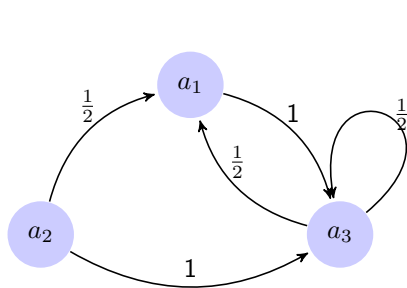
Obtenemos, por ejemplo, la tercera fila de la matriz como sigue. Supóngase que el sistema está en el estado 3, o sea, el máximo de los números que resultan en las n primeras pruebas es 3. Entonces, el sistema permanece en estado 3 si un 1, 2 ó 3 sucede en la prueba $(n+1)$; por tanto $p_{33} = \frac{3}{6}$. Por otra parte, el sistema pasa al estado 4, 5 ó 6, respectivamente, si un 4, 5 ó 6 sucede en la prueba $(n+1)$; por tanto $p_{34} = p_{35} = p_{36} = \frac{1}{6}$. El sistema nunca puede pasar al estado 1 ó 2 puesto que salió un 3 en una de las pruebas; de aquí $p_{31} = p_{32} = 0$. Así, la tercera fila de la matriz de transición es $(0, 0, \frac{3}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6})$. Las otras filas se obtienen similarmente.

La matriz no es regular puesto que el estado 6 es absorbente, o sea, hay un 1 en la diagonal principal, fila 6.

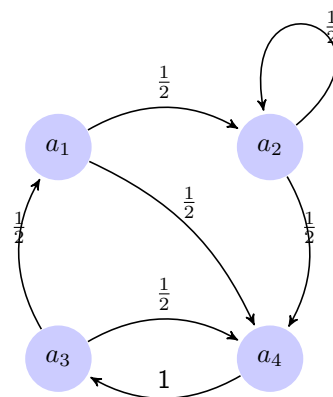
(ii) En el primer lanzamiento del dado el estado del sistema X_1 es el número que sale; por tanto, $p^{(1)} = (\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6})$.

(iii) $p^{(2)} = p^{(1)}P = (\frac{1}{36}, \frac{3}{36}, \frac{5}{36}, \frac{7}{36}, \frac{9}{36}, \frac{11}{36})$. $p^{(3)} = p^{(2)}P = (\frac{1}{216}, \frac{7}{216}, \frac{19}{216}, \frac{37}{216}, \frac{61}{216}, \frac{91}{216})$.

5.- Las probabilidades de transición de una cadena de Markov pueden representarse por un diagrama, llamado *diagrama de transición*, donde una probabilidad positiva p_{ij} es señalada por una flecha del estado a_i al estado a_j . Hallar la matriz de transición de cada uno de los diagramas de transición siguientes:



(i)



(ii)

- (i) Obsérvese primero que el espacio de estados es a_1, a_2, a_3 y, por tanto, la matriz de transición es de la forma

$$P = \begin{matrix} & \begin{matrix} a_1 & a_2 & a_3 \end{matrix} \\ \begin{matrix} a_1 \\ a_2 \\ a_3 \end{matrix} & \begin{pmatrix} & & \\ & & \\ & & \end{pmatrix} \end{matrix}$$

La fila i –ésima de la matriz se obtiene hallando aquellas flechas que parten de a_i en el diagrama: el número anexo a la flecha de a_i a a_j es la componente j –ésima de la fila i –ésima. Entonces la matriz de transición es

$$P = \begin{matrix} & \begin{matrix} a_1 & a_2 & a_3 \end{matrix} \\ \begin{matrix} a_1 \\ a_2 \\ a_3 \end{matrix} & \begin{pmatrix} 0 & 0 & 1 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} \end{matrix}$$

- (ii) El espacio de estados es a_1, a_2, a_3, a_4 . La matriz de transición es

$$P = \begin{matrix} & \begin{matrix} a_1 & a_2 & a_3 & a_4 \end{matrix} \\ \begin{matrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{matrix} & \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{matrix}$$

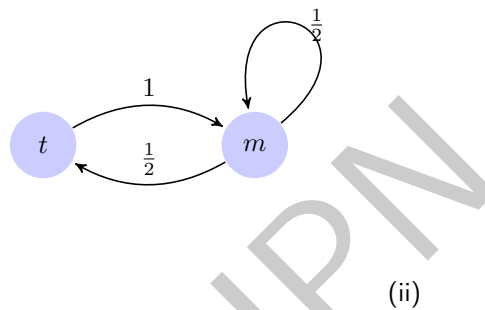
54.2 Ejemplo del hombre que viaja en carro y en metro

Un hombre o maneja su carro o toma el tren para ir a trabajar cada día. Supóngase que nunca toma el tren dos días seguidos; pero si maneja para trabajar, entonces al día siguiente es tan posible que maneje de nuevo como que tome el tren.

El espacio de estados del sistema es $t(\text{tren})$, $m(\text{manejar})$. Este proceso estocástico es una cadena de Markov puesto que los resultados de un día dependen únicamente de lo que sucedió el día anterior. La matriz de transición de la cadena de Markov es:

$$P = \begin{matrix} & \begin{matrix} t & m \end{matrix} \\ \begin{matrix} t \\ m \end{matrix} & \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \end{matrix}$$

Cuyo diagrama de transiciones es

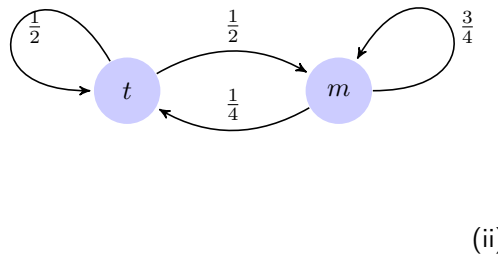


La primera fila de la matriz corresponde al hecho de que nunca toma el tren dos días seguidos y por tanto es seguro que manejará al día siguiente de usar el tren. La segunda fila de la matriz corresponde al hecho de que al día siguiente de manejar, manejará o tomará el tren con igual probabilidad.

Encontrar la probabilidad de que el sistema cambie del estado t al estado m en dos pasos

$$P^2 = \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix}$$

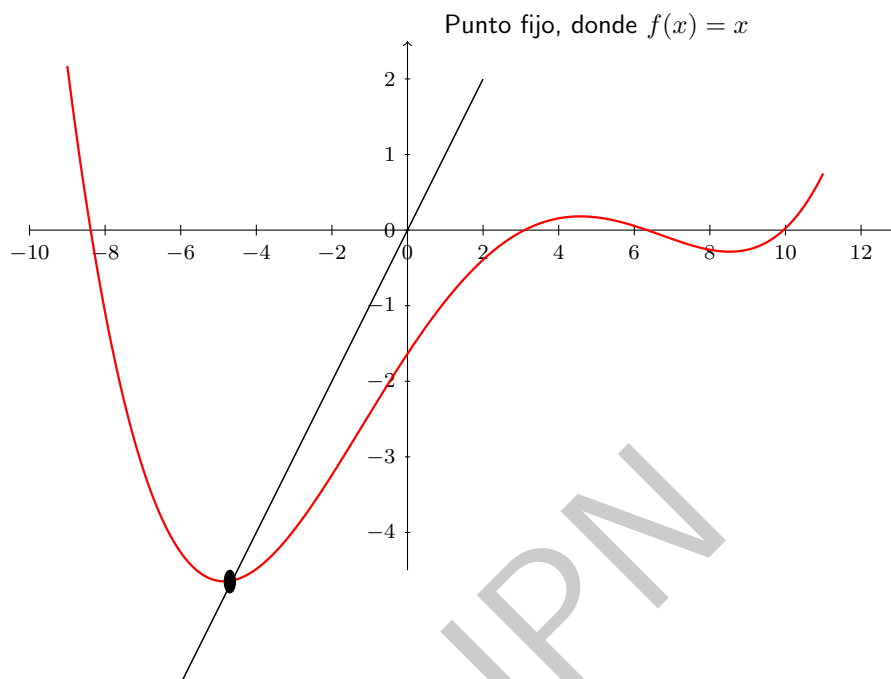
Cuyo diagrama de transiciones es



Por lo tanto la probabilidad de que el diagrama cambien del estado t al m en dos pasos es $\frac{1}{2}$

55 Teorema del punto fijo

En matemáticas, un teorema del punto fijo es un teorema que especifica condiciones bajo las cuales se puede afirmar que una función f sobre un dominio dado (con rango en el mismo dominio) tiene, al menos, un punto fijo; es decir, que existe un punto x en dicho dominio para el cual: $f(x) = x$.



Hay diferentes teoremas del punto fijo. Por ejemplo:

- **El teorema del punto fijo de Banach**

En análisis matemático el teorema del punto fijo de Banach (también llamado teorema de la aplicación contractiva) es una de las herramientas más importantes para demostrar la existencia de soluciones de numerosos problemas matemáticos. El teorema garantiza la existencia y unicidad de puntos fijos de ciertas funciones definidas sobre espacios métricos y proporciona un método para encontrarlos.

Sea (X, d) un espacio métrico completo y f una aplicación. Se dice que f es contractiva si existe una constante K con $0 < K < 1$ tal que $d(f(x), f(y)) \leq Kd(x, y)$ para cualesquiera $x, y \in X$. Un punto fijo x_0 de f es un punto de X tal que $f(x_0) = x_0$. Entonces el teorema del punto fijo de Banach dice:

Sea (X, d) un espacio métrico completo y sea $f : X \rightarrow X$ una aplicación contractiva en X . Entonces existe un único punto fijo de f .

Además, el teorema establece que para todo punto x de X la sucesión $x, f(x), f(f(x)), \dots$ converge a dicho punto fijo.

56 El hombre anumérico
















































































































































































Capítulo 2

Probabilidad y Coincidencia

1. Algunos cumpleaños y un cumpleaños determinado

Sigmund Freud señaló en cierta ocasión que las coincidencias no existen. Carl Jung habló de los misterios de la sincronización. Y en general la gente habla de por aquí e ironías por allá. Tanto si las llamamos coincidencias, sincronizaciones o ironías, resulta que son mucho más frecuentes que lo que la gente cree.

El año tiene 366 días (incluimos el 29 de febrero), tendríamos que reunir 367 personas para estar seguros de que por lo menos dos personas del grupo han nacido el mismo día. Con esto se hace referencia a el principio de la pichonera.

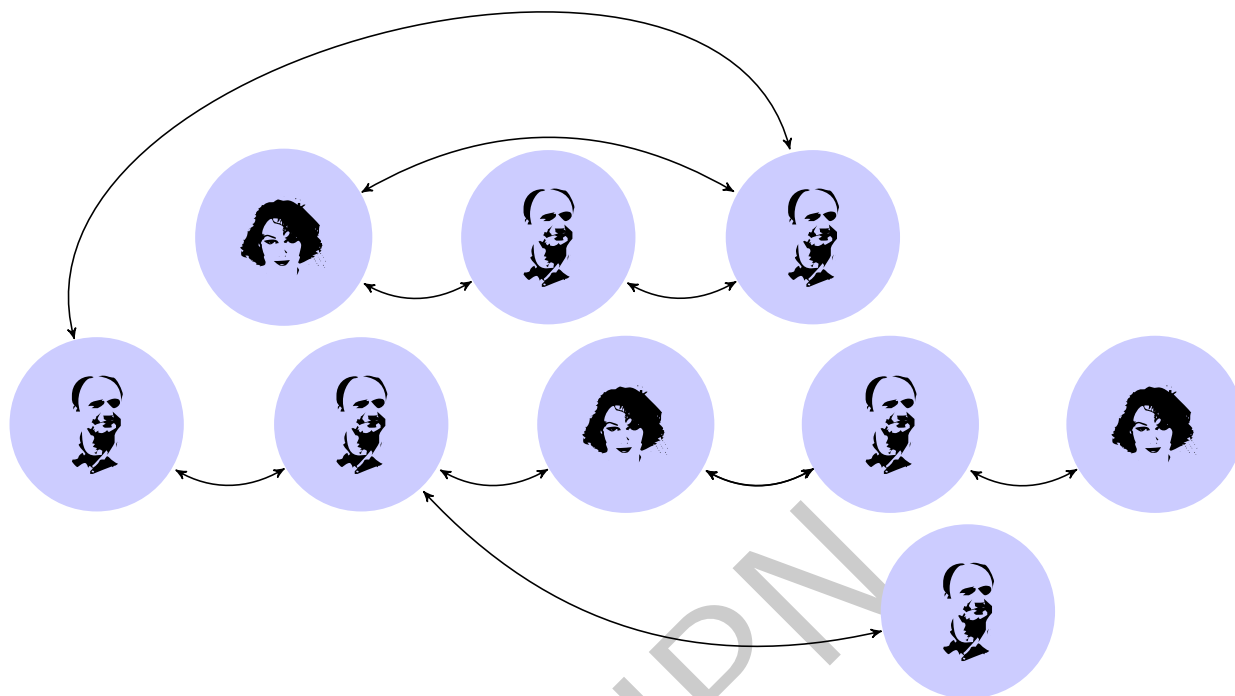
						
						
						
						
						
						
						
						
						
						
						
						
						
						
						
						
						
						
						
						
						
						
						
						
						

Si extendemos la tabla anterior, hasta cumplir los 366 días podemos observar en la tabla anterior, que habrá al menos un día en que dos personas cumplan años el mismo día.

Se analizan otros casos como el de ¿qué pasa si nos contentamos con tener una certeza de sólo el 50%?, Contemplando que 183 es la mitad de 366, a primera vista diríamos que esa es la cantidad necesaria, pero no es así. Haciendo el análisis se encuentra que solo son necesarias veintitrés personas, dos o más de ellas han nacido el mismo día.

2. Encuentros fortuitos

Dos extraños, procedentes de puntos opuestos de los Estados Unidos, se sientan juntos en un viaje de negocios a Milwaukee y descubren que la mujer de uno de ellos estuvo en un campo de tenis que dirigía un conocido del otro. Esta clase de coincidencias es sorprendentemente corriente. Si suponemos que cada uno de los aproximadamente 200 millones de adultos que viven en los Estados Unidos conoce a unas 1.500 personas, las cuales están razonablemente dispersas por todo el país, entonces la probabilidad de que cada dos tengan un conocido en común es del uno por ciento, y la de que estén unidos por una cadena con dos intermediarios es mayor que el noventa y nueve por ciento.



Podemos entonces estar prácticamente seguros, si aceptamos estas suposiciones, de que dos personas escogidas al azar, como los extraños del viaje de negocios, estarán unidos por una cadena de dos intermediarios como mucho. Que durante su conversación pasen lista de las 1.500 personas que conoce cada uno (así como de los conocidos de estas), y así sean conscientes de la relación y de los dos intermediarios, es ya un asunto más dudoso.

Las suposiciones en que basamos la deducción anterior se pueden relajar un tanto. Quizás el adulto medio conozca menos de 1.500 personas o, lo que es más probable, la mayoría de la gente que conoce vive cerca y no está dispersa por todo el país. Incluso en este caso, menos favorable, es inesperadamente alta la probabilidad de que dos personas escogidas al azar estén unidas por una cadena de como mucho dos intermediarios.

El psicólogo Stanley Milgrim emprendió un enfoque más empírico del problema de los encuentros fortuitos. Tomó un grupo de personas escogidas al azar, dio un documento a cada miembro del grupo y le asignó un individuo destinatario al que tenía que transmitir el documento. Las instrucciones eran que cada persona tenía que mandar el documento a aquel de sus conocidos que más probablemente conociera al destinatario, instruyéndole para que hiciera lo mismo, hasta que el documento llegara a su destino. Milgrim encontró que el número de intermediarios iba de dos a diez, siendo cinco el número más frecuente. Aunque menos espectacular que el argumento probabilístico anterior, el resultado de Milgrim es más impresionante. Aporta bastante a la explicación de cómo las informaciones confidenciales, los rumores y los chistes corren tan rápidamente entre cierta población.

Si el destinatario es un personaje conocido, el número de intermediarios es aún menor, sobre todo si uno está relacionado con uno o dos personajes célebres. ¿Cuántos intermediarios hay entre tú y el presidente Reagan? Pongamos que sean N . Entonces el número de intermediarios entre tú y el secretario general Gorbachov es menor o igual que $(N + 1)$, pues Reagan y Gorbachov se conocen. ¿Cuántos intermediarios hay entre tú y Elvis Presley? Aquí tampoco pueden ser más de $(N + 2)$, pues Reagan conoce a Nixon y este conoció a Presley. La mayoría de las personas se sorprenden al darse cuenta de lo corta que es la cadena que les une a cualquier personaje célebre.

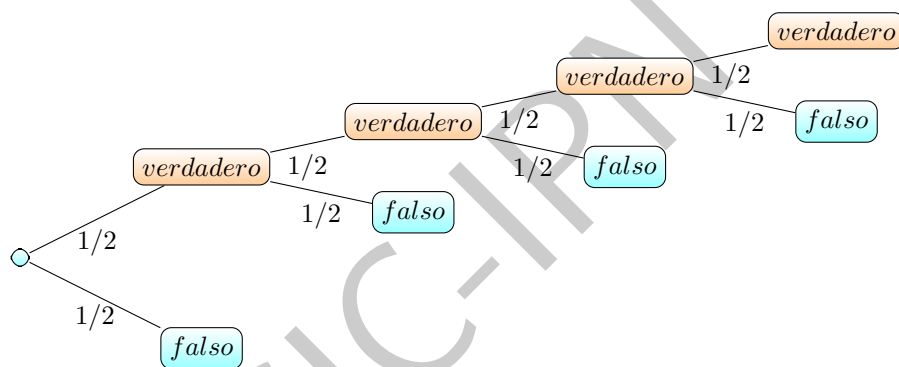
3. Un timo bursátil

Los asesores de bolsa están en todas partes y es muy probable encontrar alguno que diga cualquier cosa que uno esté dispuesto a oír. Normalmente son enérgicos, parecen muy expertos y hablan una extraña jerga de opciones de compra y de venta, cupones de cero y cosas por el estilo. A la luz de mi humilde experiencia, la mayoría no tiene mucha idea de lo que está hablando, pero cabe esperar que algunos sí.

Si durante seis semanas seguidas recibieras por correo las predicciones de un asesor de bolsa acerca de cierto índice del mercado de valores y las seis fueran acertadas, ¿estarías dispuesto a pagar por recibir la séptima predicción? Supón que estás realmente interesado en hacer una inversión y también que te han planteado la pregunta antes de la crisis del 19 de octubre de 1987. Si estuvieras dispuesto a pagar por esa predicción (y si no, también), piensa en el siguiente timo.

Uno que se hace pasar por asesor financiero imprime un logotipo en papel de lujo y envía 32.000 cartas a otros tantos inversores potenciales en un cierto valor de la bolsa. Las cartas hablan del elaborado sistema informático de su compañía, de su experiencia financiera y de sus contactos. En 16.000 de las cartas predice que las acciones subirán y, en las otras 16.000, que bajarán. Tanto si suben las acciones como si bajan, envía una segunda carta pero sólo a las 16.000 personas que recibieron la <<predicción>> correcta. En 8.000 de ellas, se predice un alza para la semana siguiente, y en las 8.000 restantes, una caída. Ocurra lo que ocurra, 8.000 personas habrán recibido ya dos predicciones acertadas. Manda una tercera tanda de cartas, ahora sólo a estas 8.000 personas, con una nueva predicción de la evolución del valor para la semana siguiente: 4.000 predicen un alza y 4.000 una caída. Pase lo que pase, 4.000 personas habrán recibido tres predicciones acertadas seguidas.

Sigue así unas cuantas veces más, hasta que 500 personas han recibido seis predicciones correctas seguidas. En la siguiente carta se les recuerda esto y se les dice que para seguir recibiendo una información tan valiosa por séptima vez habrán de aportar 500 dólares. Si todos pagan, nuestro asesor les saca 250.000 dólares. Si se hace esto a sabiendas y con intención de defraudar, es un timo ilegal. Y sin embargo, se acepta si lo hacen involuntariamente unos editores serios pero ignorantes de boletines informativos sobre la bolsa, los curanderos o los televangelistas. El puro azar siempre deja lugar a una cantidad suficiente de aciertos que permiten justificar casi cualquier cosa a alguien predispuesto a creer.



El árbol anterior muestra como es que va avanzando el árbol de probabilidad, dependiendo las cartas que va enviando, lo que quiere decir que seguira creciendo de la misma forma hasta llegar a 500 cartas a las personas donde ha acertado en el resultado.

4. Valores esperados: de los análisis de sangre al juego del chuck-a-luck

Aunque lo más llamativo sean los valores extremos y las coincidencias, lo que suele proporcionar más información son los valores medios o los valores <<esperados>>. El valor esperado de una cantidad es la media de los valores que toma, pesados según sus probabilidades respectivas. Por ejemplo, si $1/4$ de las veces la cantidad vale 2, $1/3$ vale 6, otro $1/3$ de las veces vale 15 y el $1/12$ restante vale 54, el valor esperado de dicha magnitud es 12.

5. Eligiendo cónyuge

Hay dos maneras de enfocar el amor: con el corazón y con la cabeza. Por separado, ninguno de los dos da buenos resultados, pero juntos... tampoco funcionan demasiado bien. Sin embargo, si se emplean ambos a la vez, quizá las probabilidades de éxito sean mayores. Es muy posible que, al recordar amores pasados, alguien que enfoque sus romances con el corazón se lamente de las oportunidades perdidas y que piense que nunca jamás volverá a amar así.

6. Las coincidencias y la ley

En 1964 una mujer rubia peinada con una cola de caballo robó el bolso a otra mujer en Los ángeles. La ladrona huyó a pie, pero posteriormente alguien la reconoció cuando montaba en un coche amarillo conducido por un negro con barba y bigote.

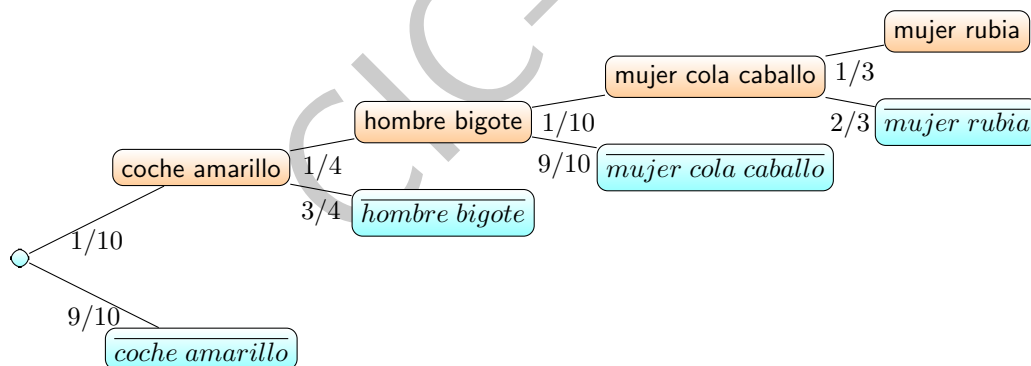
Las investigaciones de la policía acabaron por encontrar a una mujer rubia con cola de caballo que regularmente frecuentaba la compañía de un negro de barba y bigote que tenía un coche amarillo. No había ninguna prueba fehaciente

que relacionara a la pareja con el delito, ni testigos que pudieran identificar a ninguno de los dos. Se estaba de acuerdo, no obstante, en los hechos citados.

El fiscal basó sus conclusiones en que, como la probabilidad de que tal pareja existiera era tan baja, la investigación de la policía tenía que haber dado con los verdaderos culpables. Asignó las siguientes probabilidades a las características en cuestión: coche amarillo: $1/10$; hombre con bigote: $1/4$; mujer con cola de caballo: $1/10$; mujer rubia: $1/3$; hombre negro con barba: $1/10$; pareja interracial en un coche: $1/1.000$. El fiscal arguyó que como estas características eran independientes, la probabilidad de que todas ellas concurrieran en una pareja elegida al azar había de ser: $1/10 \cdot 1/4 \cdot 1/10 \cdot 1/3 \cdot 1/10 \cdot 1/1.000 = 1/12.000.000$, un número tan pequeño que la pareja había de ser culpable. El jurado les condenó.

Los condenados recurrieron ante el Tribunal Supremo de California, que anuló la sentencia sobre la base de otro razonamiento probabilístico. El abogado defensor de la pareja arguyó que $1/12.000.000$ no era la probabilidad que había que considerar. En una ciudad de las dimensiones de Los Ángeles, con unos 2.000.000 de parejas, no era tan improbable, sostenía, que hubiera más de una que reuniera todas las características mencionadas, dado que ya había por lo menos una pareja: la condenada. Basándose en la distribución binomial de probabilidad y en el $1/12.000.000$, se puede calcular dicha probabilidad, que resulta ser de aproximadamente el 8 por ciento, que, aunque pequeña, permite un margen de duda razonable. El Tribunal Supremo de California aceptó la argumentación del abogado y revocó la sentencia anterior.

Independientemente de las dudas que uno pueda tener con respecto a cómo se obtuvo la cifra de 12.000.000, la rareza por sí misma no prueba nada. Cuando le dan a uno una mano de bridge de trece cartas, la probabilidad de que le den precisamente esa mano concreta es menor que una seiscientos mil millonésima. Y a pesar de ello, será absurdo que, después de recoger las trece cartas, esa persona las examine detenidamente, calcule que la probabilidad de tener precisamente esas trece cartas es menor que una seiscientos mil millonésima y concluya que no puede ser que le hayan dado precisamente esa mano porque es muy improbable que esto ocurra. En determinados contextos, la improbabilidad es algo que no sorprende. Cada mano de bridge es muy improbable. También lo son las manos de póker y los billetes de lotería. En el caso de la pareja californiana, la improbabilidad es más significativa. Sin embargo, el razonamiento correcto es el de su abogado defensor. Y a propósito, si las 3.838.380 maneras de escoger seis números de entre cuarenta son todas igualmente probables ¿cómo es que la mayoría de la gente prefiere un billete de lotería con la combinación 2 13 17 20 29 36 a otro con la combinación 1 2 3 4 5 6? Esta es, me parece, una pregunta bastante interesante.



Las demás características también pueden ser representadas por medio del árbol, considerando las probabilidades de cada evento [4].

57 Teorema de los cuatro colores

El teorema de los cuatro colores consiste básicamente, en que cualquier mapa puede ser coloreado solamente con cuatro colores distintos de tal manera que dos regiones adyacentes (es decir, que tienen una frontera en común y no sólo un punto) no tengan el mismo color.

Aunque parece un problema no matemático, sin embargo lo es y su demostración no es nada sencilla, ha costado mucho esfuerzo y 125 años el conseguirlo.

El asunto de los cuatro colores se convirtió formalmente en un problema matemático recién en 1850. El responsable de otorgar esa categoría al problema que enfrentaban los cartógrafos fue un estudiante inglés llamado Francis Guthrie, quien intuyó que el mecanismo empleado podía ser demostrado. Dado que el problema superaba su formación matemática, lo comentó con su hermano Frederick, quien había sido alumno del prestigioso matemático inglés Augustus De Morgan. Nacido en la India en 1806, De Morgan fue el primer presidente de la Sociedad de Matemáticas de Londres y tutor de la genial Ada Lovelace. Fue además el autor de las leyes fundamentales del álgebra de la lógica que llevan su nombre (La negación de la conjunción es equivalente a la disyunción de las negaciones y La negación de la disyunción es equivalente a la conjunción de las negaciones), por lo que a priori parecía la persona indicada para demostrar este teorema, que fue inicialmente conocido con el nombre de problema de Guthrie. Sin embargo, no pudo hacerlo.

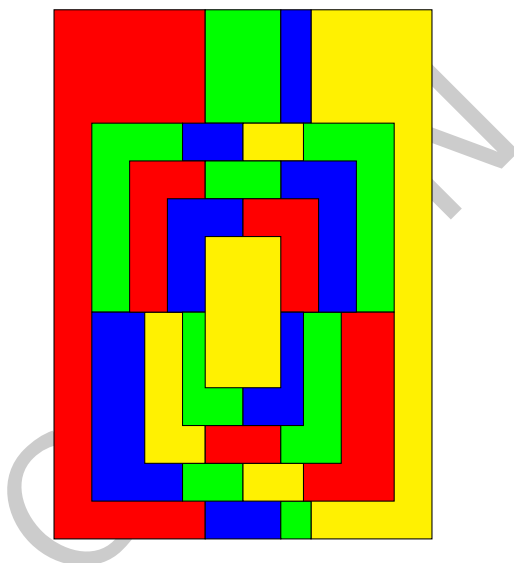


Fig. 85: Mapa coloreado con cuatro colores

Lejos de olvidarse del problema que su alumno le había planteado, De Morgan consideró que el problema del coloreado de mapas era lo suficientemente interesante como para enviarle una carta a su colega Sir William Hamilton, quien era famoso por postular la estructura de los números cuaterniónicos. La leyenda cuenta que un día de 1843 Hamilton paseaba por el puente de Brongham, que cruza el canal Real de Dublín y, de repente, en un momento de inspiración, comprendió la estructura de los cuaterniones. De inmediato grabó sobre una piedra utilizando su navaja la idea. Los cuaterniones resultarían fundamentales para edificar la física relativista y cuántica, y para demostrar el teorema propuesto por Lagrange según el cual cualquier número entero puede escribirse como la suma de 4 cuadrados perfectos.

58 Inteligencia Artificial

En el ámbito de las ciencias de la computación se denomina como inteligencia artificial a la facultad de razonamiento que ostenta un agente que no está vivo, tal es el caso de un robot, por citar uno de los ejemplos más populares, y que le fue conferida gracias al diseño y desarrollo de diversos procesos gestados por los seres humanos. Cabe destacarse que además del poder de razonar, estos dispositivos son capaces de desarrollar muchas conductas y actividades especialmente humanas como puede ser resolver un problema dado, practicar un deporte, entre otros.

Ahora bien, es importante destacar que nos podremos encontrar con diversos tipos de inteligencia artificial, como ser: sistemas que piensan como los seres humanos (se ocupan de imitar el pensamiento humano, siendo un ejemplo las redes neuronales artificiales, que justamente imitan el funcionamiento del sistema nervioso), sistemas que actúan como los humanos (son aquellos que imitan el comportamiento del hombre, el ejemplo más claro es el del robot), sistemas que piensan de modo racional (se caracterizan por imitar el pensamiento lógico de los humanos, es decir, en una cuestión concreta razonan como expertos. Se los emplea especialmente a la hora de buscar mejoras en la calidad y en la rapidez de un servicio) y sistemas que actúan de manera racional (imitan de manera racional el comportamiento de una persona, siendo capaces de percibir el entorno y actuar en consecuencia).

58.1 Clasificación Inteligencia Artificial

Inteligencia Artificial Débil IA débil (también conocido como IA estrecha) se define como la inteligencia artificial racional que se centra típicamente en una tarea estrecha. La inteligencia de la IA débil es limitada.

Siri es un buen ejemplo de la inteligencia estrecha. Siri opera dentro de un rango limitado previamente definido, no hay ninguna inteligencia genuina, sin conciencia, sin vida, a pesar de ser un ejemplo sofisticado de IA débil. En Forbes (2011), Ted Greenwald escribió: "El matrimonio iPhone/Siri representa la llegada de la IA híbrida, combinando técnicas de la IA estrecha, además de un acceso masivo a datos en la nube".⁴ El investigador de IA, Ben Goertzel, indicó en su blog en el 2010 que Siri era "muy estrecho y frágil" evidenciada por resultados molestos si se le hacían preguntas fuera de los límites de la aplicación.

Algunos comentaristas piensan que la IA débil podría ser peligrosa. En 2013 George Dvorsky declaró vía io9: "La IA estrecha podría noquear nuestra red eléctrica, dañar plantas de energía nucleares, causar un colapso económico a escala global, desviar los vehículos autónomos y a los robots..."

Inteligencia Artificial Fuerte

La Inteligencia Artificial Fuerte es aquella inteligencia artificial que iguala o excede la inteligencia humana promedio la inteligencia de una máquina que exitosamente puede realizar cualquier tarea intelectual de cualquier ser humano. Es un objetivo importante para la investigación sobre IA y un tópico interesante para la ciencia ficción.

La IA Fuerte (IAF) igualmente es referida como Inteligencia General Artificial o AGI por sus siglas en inglés, esta es la habilidad de ejecutar acciones generales inteligentes. La ciencia ficción asocia a la IAF con cualidades humanas como la conciencia, la sensibilidad, la sapiencia y el autoconocimiento.

Hasta el momento, la Inteligencia Artificial Fuerte se mantiene como una aspiración o mejor dicho, es hipotética; esto a pesar de los grandes avances en el campo y el mejoramiento de complejos algoritmos matemáticos. Algo importante a destacar es que el surgimiento y aplicación de este tipo de inteligencia artificial supondría un nuevo problema o mejor dicho, incógnitas morales sobre la posición que deberíamos tomar como seres humanos ante un ser que por su nueva inteligencia podría no considerarse solamente como una máquina más, otra incógnita es si este podría alcanzar un grado donde llegue a tener conciencia.

59 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*. él 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 caracteres. 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 autofecundarse (dioicas), 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 variedades. Protegió a los descendientes, para que en el período de floración 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 en la misma proporción, así como de las condiciones experimentales y utilizó una misma letra del alfabeto para determinar cada característica (por ejemplo, A=variante dominante y a=variante recesiva).















Características	Variantes	
Textura de la semilla	Lisa 	Rugosa 
Color de la semilla	Amarilla 	Verde 
Color de los pétalos de la flor	Púrpura 	Blanco 
Aspecto general de la vaina	Comprimida 	Hinchada 
Color de la vaina inmadura	Amarillo 	Verde 
Posición de la flor	Axial 	Terminal 
Longitud del tallo	Largo 	Corto 

Fig. 86: Las siete variantes de *Pisum sativum* estudiadas por Mendel en sus experimentos

Monohibridismo

Una vez seleccionados los caracteres sobre los que fijaría su atención, Mendel se dispuso a analizar cómo se transmitían estas características a la descendencia. Pero en vez de analizar las siete al mismo tiempo, optó por estudiar la herencia utilizando una cada vez, es decir, realizó 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 (filial 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, se dio cuenta de que los fenotipos dominantes eran siempre tres veces más frecuentes que los recesivos en la F1.

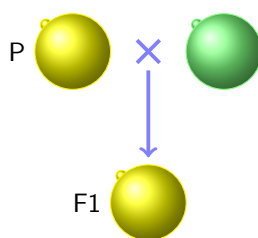


Fig. 87: Experimentos de Mendel: 100% Semilla Amarilla

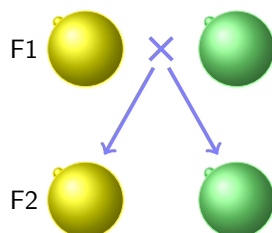


Fig. 88: Experimentos de Mendel: 75% Semilla Amarilla, 25% Semilla Verde

En síntesis, 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.

Rasgo	Cruzamiento	F1 (porcentaje)	F2 (N° individuos)	Proporción (dominantes:recesivos)
Superficie de la semilla		100% semilla lisa	5.474 semillas lisas 1.850 semillas rugosas	3:1
Color de la semilla		100% semilla amarilla	6.022 semillas amarillas 2.001 semillas verdes	3:1
Color de la flor		100% flor púrpura	705 flor púrpura 224 flor blanca	3:1
Forma de la vaina		100% vaina comprimida	882 vaina comprimida 299 vaina hinchada	3:1
Color de la vaina		100% vaina verde	428 vaina verde 152 vaina amarilla	3:1
Posición de la flor		100% flor axial (a los lados)	651 flor axial 207 flor terminal	3:1
Tamaño de la planta		100% planta alta	787 planta alta 277 planta baja	3:1

Fig. 89: Las siete variantes de *Pisum sativum* estudiadas por Mendel en sus experimentos

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 generación F1 presentaban una característica igual a la de uno de los padres, a la cual llamó dominante. El resultado anterior podría 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 surgían 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.

Factor aportado por la madre	Factor aportado por el padre	Característica resultante
Dominante (AA o Aa)	Dominante (AA o Aa)	Dominante (AA o Aa)
Dominante (AA o Aa)	Recesivo (aa)	Dominante (AA o Aa)
Recesivo (aa)	Dominante (AA o Aa)	Dominante (AA o Aa)
Recesivo (aa)	Recesivo (aa)	Recesivo (aa)

Actualmente se utiliza el término gen para referirse a los "factores de la herencia" descritos por Mendel. Se les denomina alelos a las variantes de un gen, es decir, dominantes o recesivos. A las combinaciones alélicas se les reconoce como genotipos, que pueden ser de dos tipos: los homocigotos, cuyos genes solamente expresan un mismo carácter; si son dominantes, se representan con dos letras mayúsculas, por ejemplo: M (semilla amarilla) y si son recesivos, con dos letras minúsculas: aa (semilla verde). Los heterocigotos, como en la F1, presentan en cada cromosoma un alelo distinto, es decir, que puede expresar características diferentes; se representa como Aa. La letra mayúscula indica la característica dominante, que será expresada en el fenotipo.

Variantes	Genotipo	Ejemplos de fenotipo
AA	Homocigoto dominante (ambos alelos dominantes)	Semillas amarillas
Aa	Heterocigoto (un alelo dominante y uno recesivo)	Semillas amarillas
aa	Homocigoto recesivo (ambos alelos recesivos)	Semillas verdes
Recesivo (aa)	Recesivo (aa)	Recesivo (aa)

De esta manera, los resultados obtenidos por Mendel, hace muchísimo tiempo atrás, se han complementado con los conocimientos actuales, postulándose los principios mendelianos o de la herencia. El primero de ellos es el de segregación al azar de los alelos, y sostiene que "los genes se encuentran de a pares en los individuos, y se separan en iguales proporciones durante la formación de los gametos".

Con el objeto de facilitar la comprensión del fenómeno de segregación de los alelos y de los genotipos esperados en la descendencia, se usan tablas de doble entrada para señalar los genotipos de los padres, los gametos producidos por los mismos y los genotipos de los hijos resultantes al fecundarse dichos gametos. Además, conociendo los genotipos es posible saber los fenotipos de cada uno de los individuos. Esta representación se conoce como tablero de Punnet y explica la segregación de alelos en la formación de gametos en padres heterocigotos y el resultado esperado para la descendencia, que puede representarse como proporción o porcentaje [3].

A partir de estos experimentos desarrollo varias leyes :

- 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).
- 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) dará origen a una segunda generación filial 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 opacado por el carácter A pero que, al reproducirse un individuo, cada carácter se segrega por separado.
- Tercera ley o principio de la combinación independiente: Mendel trabajó este cruce en guisantes, en los cuales las características que él observaba (color de la semilla y rugosidad de su superficie) se encontraban en cromosomas separados. De esta manera, observó que los caracteres se transmitían 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) [2].

References

- [1] A. Bogomolny. Chevalier de mere's problem from interactive mathematics miscellany and puzzles. <http://www.cut-the-knot.org/Probability/ChevalierDeMere.shtml>. Accessed 27 March 2016.
- [2] Pedro Rodríguez Encinas. Experimentos y leyes de mendel. <https://biologiaygeologia4eso.wordpress.com/2011/06/09/experimentos-y-leyes-de-mendel/>.
- [3] Biología Escolar. Los experimentos de mendel. <http://www.biologiaescolar.com/2014/07/experimentos-de-mendel.html>.
- [4] José Ramón Franco. El hombre anumérico. *Números. Revista de Didáctica de las Matemáticas*, 70:135–136, 2009.
- [5] G.Polya. *How to Solve It*. Princeton University Press, 1957.
- [6] Seymour Ferro Duque Lipschutz et al. *Probabilidad*. Number 519.2 L5P7. 1991.
- [7] Departamento de Matemáticas Facultad de Ciencias UNAM Luis Rincón. Introducción a la probabilidad. *Números. Revista de Didáctica de las Matemáticas*, 2013.
- [8] FJ Martín Pliego, José María Montero Lorenzo, and Luis Ruiz Maya Pérez. *Problemas de probabilidad*. Editorial Paraninfo, 2006.
- [9] Francisco Javier Martín Pliego and Luis Ruiz-Maya Pérez. *Fundamentos de probabilidad*. Editorial Paraninfo, 2006.
- [10] Kenneth H Rosen and Kamala Krithivasan. *Discrete mathematics and its applications*, volume 6. McGraw-Hill New York, 1999.