

Probabilidad, procesos aleatorios e inferencia

Itzael
Jiménez
Aranda



Índice

	Página
Capítulo 1	8
1.1 Experimentos probabilísticos y espacios muestrales	8
1.1.1 Experimentos deterministas	8
1.1.2 Experimentos aleatorios	8
1.1.3 Espacios muestrales	9
1.2 Tipos de probabilidad	9
1.2.1 Probabilidad clásica	9
1.2.2 Probabilidad geométrica	9
1.2.3 Probabilidad frecuentista	10
1.2.4 Probabilidad axiomática	10
1.2.5 Probabilidad subjetiva	10
1.2.6 Probabilidad lógica	10
1.3 Tipos de frecuencias	12
1.3.1 Frecuencia estadística	12
1.3.2 Frecuencia absoluta	12
1.3.3 Frecuencia relativa	12
1.4 Proceso estocástico	13
1.5 Riesgo	13
1.6 Peligrosidad	13
1.7 Amenaza	13
1.8 Vulnerabilidad	13
1.9 Expansión del teorema del binomio del 2 al 10	14
Capítulo 2	18
2.1 Ejemplos de permutaciones	18
2.2 Ejemplos de combinaciones	18
2.3 Triángulo de Pascal	19
2.3.1 Construcción del triángulo de Pascal	19
2.3.2 Otra manera de construir el triángulo de Pascal	21
2.3.3 Multiplos de 2 en el triángulo de Pascal	23
2.3.4 Multiplos de 5 en el triángulo de Pascal	24
2.3.5 Multiplos de 7 en el triángulo de Pascal	25
2.3.6 Multiplos de 11 en el triángulo de Pascal	26
2.4 Generalizando el problema de caminos posibles para el caso de un cuadrado	26
2.5 Generalizando el problema de caminos posibles para el caso de un rectángulo	29
2.6 Reseña de artículos	30
2.6.1 Reseña del artículo "Primates Count"	30
2.6.2 Reseña del artículo "Cicada-Generate Prime Numbers"	30
2.6.3 Reseña del artículo "Ant Odometer"	30
2.6.4 Reseña del artículo "Quipu"	31

Capítulo 3	31
3.1 Magnetic Cloaking	31
3.2 Compuertas lógicas con interruptores	31
3.3 Telómeros	33
3.4 Oveja Dolly	34
3.5 Histograma	37
3.6 Fórmula de Stirling	38
Capítulo 4	39
4.1 Paradojas de conjuntos	39
4.2 La paradoja del barbero	40
4.3 Paradoja y antinomia	41
4.4 Demostración de las leyes De Morgan	42
4.5 Demostración de uniones de conjuntos	44
4.5.1 Desarrollo de $P(A \cup B \cup C)$	44
4.5.2 Desarrollo de $P(A \cup B \cup C \cup D)$	50
4.6 Significado de $P(A B)$	56
Capítulo 5	57
5.1 Teorema de Bayes	57
5.2 Probabilidad bayesiana	58
5.3 Paradoja del falso positivo	59
5.4 Paradoja de Borel-Kolmogorov	61
5.5 Falacia del apostador	63
5.6 Simulación de un dado de 3 caras en ROOT	65
5.6.1 Histograma a 10 tiradas	65
5.6.2 Histograma a 100 tiradas	66
5.6.3 Histograma a 1,000 tiradas	67
5.6.4 Histograma a 10,000 tiradas	68
5.6.5 Código de ROOT	68
5.7 Derivadas e Integrales en Mathematica	69
Capítulo 6	71
6.1 Histograma en ROOT llenado con la distribución de Gauss	71
6.1.1 Histograma	71
6.1.2 Código de ROOT	71
6.2 Teorema del límite central	72
6.3 Explosión del Challenger	73
6.4 Densidad de probabilidad	75
6.5 Mínimos cuadrados	76
Capítulo 7	77
7.1 Ejercicios	77
7.1.1 Ejercicio 1	77
7.1.2 Ejercicio 2	83
7.1.3 Ejercicio 3	84
7.1.4 Ejercicio 4	85

7.1.5	Ejercicio 5	86
7.1.6	Ejercicio 6	87
7.1.7	Ejercicio 7	89
7.2	Grafos por criterio de divisibilidad	90
7.2.1	Grafo por criterio de divisibilidad por 3	92
7.2.2	Grafo por criterio de divisibilidad por 5	93
7.2.3	Grafo por criterio de divisibilidad por 7	93
7.2.4	Grafo por criterio de divisibilidad por 11	94
Capítulo 8	94
8.1	Máquina Enigma	94
8.2	El último teorema de Fermat y los Simpsons	98
8.3	Esperanza matemática	100
8.4	Símbolo de Levi-Civita	101
8.5	Convenio de suma de Einstein	103
8.6	Ventaja	104
Capítulo 9	106
9.1	Congruencia Zeller	106
9.1.1	Código en gcc	106
9.1.2	Código en Mathematica	109
9.2	Proyecto.- Lanzar 2 dados hasta obtener un doble 6	110
9.2.1	Abstract	110
9.2.2	Introducción.	110
9.2.3	Análisis Matemático.	110
9.2.4	Análisis Experimental.	111
9.2.5	Entender el problema.	113
9.2.6	Hacer un plan.	113
9.2.7	Poner el plan en marcha.	114
9.2.8	Experimento con dados reales.	114
9.2.9	Experimento con dados virtuales.	114
9.2.10	Simulación en Mathematica.	115
9.2.11	Simulación en ROOT.	118
9.2.12	Revisar y extender.	124
9.2.13	Conclusiones.	124
9.3	Examen propuesto	125
9.4	Examen resuelto	136
Capítulo 10	142
10.1	Curtosis	142
10.2	Skewness	144
10.3	Algoritmos para calcular π	146
10.3.1	Algoritmo de Gauss-Legendre	146
10.3.2	Algoritmo de Borwein	147
Capítulo 11	147
11.1	Resumen Capítulo 5 del libro Probabilidad Serie Schaum	147

11.1.1 Variables aleatorias	147
11.1.2 Distribución y Esperanza de una variable aleatoria finita	148
11.1.3 Varianza y Desviación estándar	150
11.1.4 Distribución conjunta	151
11.1.5 Variables aleatorias continuas	152
11.1.6 Esperanza, varianza y desviación estándar	153
11.2 Señal Wow	153
11.3 Reseña del artículo "Is Economics the Physical Science?"	158
Capítulo 12	160
12.1 Caminata aleatoria	160
12.2 Simulación caminata aleatoria	162
12.2.1 Código en ROOT	162
12.2.2 Histograma de posición final con paso a la derecha de 0.5 de probabilidad	164
12.2.3 Histograma de los pasos totales a la derecha con paso a la derecha de 0.5 de probabilidad	165
12.2.4 Ajuste de una Gaussiana de la posición final con paso a la derecha de 0.5 de probabilidad	166
12.2.5 Ajuste de una Gaussiana de los pasos totales a la derecha con paso a la derecha de 0.5 de probabilidad	167
12.2.6 Histograma de posición final con paso a la derecha de 0.6 de probabilidad	168
12.2.7 Histograma de los pasos totales a la derecha con paso a la derecha de 0.6 de probabilidad	169
12.2.8 Ajuste de una Gaussiana de la posición final con paso a la derecha de 0.6 de probabilidad	170
12.2.9 Ajuste de una Gaussiana de los pasos totales a la derecha con paso a la derecha de 0.6 de probabilidad	171
12.2.10 Histograma de posición final con paso a la derecha de 0.4 de probabilidad	172
12.2.11 Histograma de los pasos totales a la derecha con paso a la derecha de 0.4 de probabilidad	173
12.2.12 Ajuste de una Gaussiana de la posición final con paso a la derecha de 0.4 de probabilidad	174
12.2.13 Ajuste de una Gaussiana de los pasos totales a la derecha con paso a la derecha de 0.4 de probabilidad	175
12.3 Demostraciones caminata aleatoria	176
12.4 Simulación del movimiento Browniano	178
12.4.1 Código en R para simulación en 2D	178
12.4.2 Gráfica del movimiento Browniano en dos dimensiones	179
12.4.3 Código en R para simulación en 3D	179
12.4.4 Gráfica del movimiento Browniano en tres dimensiones	180
Capítulo 13	181
13.1 Relación de función Gamma en factorial	181
13.2 Función Delta de Dirac	181

Capítulo 14	184
14.1 Fork	184
14.2 NP (clase de complejidad)	184
14.3 Reseña capítulo 2 del hombre anumérico	185
Capítulo 15	191
15.1 Matrices semejantes	191
15.2 Problemas de eigen valores y vectores	193
15.2.1 Ejercicio 1	193
15.2.2 Ejercicio 2	194
15.2.3 Ejercicio 3	195
15.2.4 Ejercicio 4	198
15.2.5 Ejercicio 5	199
15.3 Matriz poco densa	200
15.4 Problemas de cadena de Markov	201
15.4.1 Ejercicio 1	201
15.4.2 Ejercicio 2	202
15.4.3 Ejercicio 3	202
15.4.4 Ejercicio 4	203
15.4.5 Ejercicio 5	205
Capítulo 16	206
16.1 Clasificaciones de Inteligencia Artificial	206
16.1.1 Inteligencia Artificial Débil	206
16.1.2 Inteligencia Artificial Fuerte	207
16.1.3 Inteligencia Artificial Distribuida	208
16.1.4 Ramas de la Inteligencia Artificial	209
Referencias	211

Lista de figuras

1	Dos niveles del triángulo de Pascal	19
2	Tres niveles del triángulo de Pascal	20
3	Cuatro niveles del triángulo de Pascal	20
4	Triángulo de Pascal	21
5	Caminos en el triángulo de Pascal	22
6	Multiplos de 2 en el triángulo de Pascal	23
7	Multiplos de 5 en el triángulo de Pascal	24
8	Multiplos de 7 en el triángulo de Pascal	25
9	Multiplos de 11 en el triángulo de Pascal	26
10	Caminos posibles para un cuadrado	26
11	Caminos posibles para un cuadrado	29
12	Compuerta AND	32
13	Compuerta OR	32
14	Telómero	33
15	Oveja Dolly	35

16	Histograma	38
17	Union de tres conjuntos	44
18	Union de 4 conjuntos	50
19	Paradoja de Borel -Kolmogorov	62
20	Falacia del apostador	64
21	Histograma de 10 lanzamientos	65
22	Histograma de 100 lanzamientos	66
23	Histograma de 1,000 lanzamientos	67
24	Histograma de 10,000 lanzamientos	68
25	Histograma en ROOT llenado con la distribución de Gauss	71
26	Challenger	74
27	Aproximación a una recta por medio de mínimos cuadrados	77
28	Representación del tornillo	78
29	Los 4 tornillos tienen longitudes superiores a L_1	79
30	Al menos 3 tornillos estén entre L_1 y L_2	80
31	Al menos 3 tornillos estén entre L_1 y L_2	81
32	Al menos 3 tornillos estén entre L_1 y L_2	82
33	A lo sumo tres tornillos con longitud menor a L_2	83
34	Los 4 tornillos tienen longitudes superiores a L_1	84
35	Ejemplo de un grafo por criterio de divisibilidad por 5	92
36	Grafo por criterio de divisibilidad por 3	92
37	Grafo por criterio de divisibilidad por 5	93
38	Grafo por criterio de divisibilidad por 7	93
39	Grafo por criterio de divisibilidad por 11	94
40	Máquina Enigma	95
41	Representación de los rotores de la máquina enigma	96
42	Esvástica Nazi	97
43	Primero capítulo de los Simpson donde aparece el teorema de Fermat	99
44	Segundo capítulo de los Simpson donde aparece el teorema de Fermat	100
45	Resultado de 10 lanzamientos.	116
46	Distribución normal de 10 lanzamientos.	116
47	Resultado de 100 lanzamientos.	116
48	Distribución normal de 100 lanzamientos.	117
49	Resultado de 1,000 lanzamientos.	117
50	Distribución normal de 1,000 lanzamientos.	117
51	Resultado de 10,000 lanzamientos.	118
52	Distribución normal de 10,000 lanzamientos.	118
53	Resultado de 10 lanzamientos.	120
54	Distribución normal de 10 lanzamientos.	121
55	Resultado de 100 lanzamientos.	121
56	Distribución normal de 100 lanzamientos.	122
57	Resultado de 1,000 lanzamientos.	122
58	Distribución normal de 1,000 lanzamientos.	123
59	Resultado de 1,000,000 lanzamientos.	123
60	Distribución normal de 1,000,000 lanzamientos.	124
61	Señal Wow	154

62	Posición final con paso a la derecha de 0.5 de probabilidad	164
63	Pasos totales a la derecha con paso a la derecha de 0.5 de probabilidad	165
64	Ajuste de una Gaussian de la posición final con paso a la derecha de 0.5 de probabilidad	166
65	Ajuste de una Gaussiana de los pasos totales a la derecha con paso a la derecha de 0.5 de probabilidad	167
66	Posición final con paso a la derecha de 0.6 de probabilidad	168
67	Pasos totales a la derecha con paso a la derecha de 0.6 de probabilidad	169
68	Ajuste de una Gaussiana de la posición final con paso a la derecha de 0.6 de probabilidad	170
69	Ajuste de una Gaussiana de los pasos totales a la derecha con paso a la derecha de 0.6 de probabilidad	171
70	Posición final con paso a la derecha de 0.4 de probabilidad	172
71	Pasos totales a la derecha con paso a la derecha de 0.4 de probabilidad	173
72	Ajuste de una Gaussiana de la posición final con paso a la derecha de 0.4 de probabilidad	174
73	Ajuste de una Gaussiana de los pasos totales a la derecha con paso a la derecha de 0.4 de probabilidad	175
74	Simulación de un movimiento Browniano en dos dimensiones	179
75	Simulación de un movimiento Browniano en tres dimensiones	180
76	Ramas de la Inteligencia Artificial	210

Capítulo 1

1.1 Experimentos probabilisticos y espacios muestrales

1.1.1 Experimentos deterministas

1. Salir al espacio sin un traje especial (no podremos respirar)
2. Aventarse de un avion sin paracaidas (moriaremos)
3. Poner agua al fuego (hervirá)
4. Sacar una fruta de una bolsa de manzanas
5. Tirar algun objeto hacia arriba (regresará hacia abajo)
6. Tirar un dado con caras iguales
7. Despues del Sábado es Domingo
8. Despues de la tarde es la noche
9. El tiempo
10. Probar algun dispositivo descompuesto (nunca funcionará)

1.1.2 Experimentos aleatorios

1. Lanzar un dado
2. Tomar una carta de un mazo
3. Ganar el melate
4. Que gane mi equipo favorito de deporte
5. Predecir el clima
6. Predecir el paso de los metioritos
7. Ganar una competencia
8. Sacar una pelota de una tombola
9. Aventar una moneda al aire
10. El precio del dolar en un año

1.1.3 Espacios muestrales

1. Las caras de un dado
2. Las cartas de una baraja
3. Los resultados de los partidos de football
4. Los números impresos en las pelotitas de la tombola
5. Los números posibles a seleccionar en el melate
6. Los distintos folios de los boletos de la lotería
7. Los caballos en una carrera
8. Los pilotos en una carrera
9. Las bolas de una urna
10. Las piezas fabricadas por cierta máquina

1.2 Tipos de probabilidad

1.2.1 Probabilidad clásica

Es el número de resultados favorables (A) de cierto evento dividido entre el número total de resultados posibles (Ω). Tambien es conocida como probabilidad a priori puesto que antes se sabe la probabilidad del resultado [14] [15].

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

1.2.2 Probabilidad geométrica

Describe la posibilidad de que un punto esté en cierta parte de un área, volumen, línea o alguna característica geométrica [6].

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

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

1.2.3 Probabilidad frecuentista

Por N número de veces que se repita un experimento, al final las posibilidades que ocurra cada uno de los sucesos será regular, tal como lanzar 3 dados al mismo tiempo, siempre será mayor la probabilidad de obtener una suma de 7 [6].

$$P(A) = \frac{n(A)}{n} \quad (4)$$

1.2.4 Probabilidad axiomática

Los axiomas de probabilidad son las condiciones mínimas que deben verificarse para que una función que definimos sobre unos sucesos determine consistentemente valores de probabilidad sobre dichos sucesos. La probabilidad P de un suceso E , denotada por $P(E)$, se define con respecto a un "universo" o espacio muestral Ω , conjunto de todos los posibles sucesos elementales, tal que P verifique los Axiomas de Kolmogoróv [6].

$$P(A) \geq 0 \quad (5)$$

$$P(\Omega) = 1 \quad (6)$$

$$P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k) \quad (7)$$

1.2.5 Probabilidad subjetiva

Se refiere a la probabilidad de ocurrencia de un suceso basado en la experiencia previa, la opinión personal o la intuición del individuo. Un ejemplo es preguntarle a un experto en cierto deporte cual es la probabilidad que gane un equipo sobre el otro, tomando en cuenta que conoce el historial de cada equipo, los jugadores, desempeño en el campo, etc. [?] [3].

1.2.6 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 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 [?] [3].

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 cumplimientos 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 enumorados 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 .

1.3 Tipos de frecuencias

1.3.1 Frecuencia estadística

Número de veces que sucede cierto evento en cierto experimento, por ejemplo el número de alumnos que saca 10 cada semestre en cierta materia.

1.3.2 Frecuencia absoluta

La frecuencia absoluta es el número de veces que aparece un determinado valor en un estudio estadístico.

Se representa por f_i .

La suma de las frecuencias absolutas es igual al número total de datos, que se representa por N .

$$\sum_{i=1}^{i=n} f_i = N \quad (8)$$

1.3.3 Frecuencia relativa

La frecuencia relativa es la división entre la frecuencia absoluta (f_i) de un determinado valor y el número total de datos (N).

La frecuencia relativa se puede expresar en tantos por ciento y se representa por n_i .

$$n_i = \frac{f_i}{N} \quad (9)$$

La suma de las frecuencias relativas es igual a 1.

Frecuencia acumulada

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

La frecuencia acumulada se representa por F_i .

1.4 Proceso estocástico

Colección o familia de variables aleatorias que cambian en función de otra variable, es decir con el tiempo. Por cada instante t tendremos una variable aleatoria distinta.

1.5 Riesgo

Corresponde a un valor relativo probable de pérdidas de toda índole en un sitio específico vulnerable a una amenaza particular. El riesgo es la probabilidad de un suceso, que usualmente se estima como un suceso negativo.

El riesgo resulta entonces del cruce probable en el espacio como en el tiempo de una amenaza de magnitud determinada y de un elemento relativamente vulnerable a ella. Por tanto, la magnitud del riesgo depende de la amenaza y del grado de vulnerabilidad [5] [12].

1.6 Peligrosidad

Probabilidad de daño en cierto evento.

1.7 Amenaza

Fenómeno con probabilidad de ocurrencia, magnitud e intensidad de daño.

1.8 Vulnerabilidad

Probabilidad de que un evento, según el grado de fragilidad de sus elementos, pueda sufrir daños a consecuencia de estar expuestos a una amenaza. La magnitud de estos daños estará asociada con el grado de vulnerabilidad [5] [12].

1.9 Expansión del teorema del binomio del 2 al 10

El teorema del binomio esta dado por la Eq(10) y el coeficiente binomial por Eq(11)

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

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

Desarrollar para $n = 2$ hasta $n = 10$

- Si $n=2$:

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

$k=0$:

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

El valor de $0!$ es 1:

$$x^2$$

Cuando $k=1$:

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

$$2xy$$

Cuando $k=2$:

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

$$y^2$$

Aplicamos la suma de los términos queda:

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

- Si n=3

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

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

Aplicamos la suma de los términos queda:

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

- Si n=4

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

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

Aplicamos la suma de los términos queda:

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

- Si n=5

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

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

Aplicamos la suma de los términos queda:

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

- Si n=6

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

$$(x+y)^6 = \frac{6!}{0!(6-0)!} x^{6-0} y^0 + \frac{6!}{1!(6-1)!} x^{6-1} y^1 + \frac{6!}{2!(6-2)!} x^{6-2} y^2 + \frac{6!}{3!(6-3)!} x^{6-3} y^3 + \\ \frac{6!}{4!(6-4)!} x^{6-4} y^4 + \frac{6!}{5!(6-5)!} x^{6-5} y^5 + \frac{6!}{6!(6-6)!} x^{6-6} y^6$$

Aplicamos la suma de los términos queda:

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

- Si n=7

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

$$(x+y)^7 = \frac{7!}{0!(7-0)!} x^{7-0} y^0 + \frac{7!}{1!(7-1)!} x^{7-1} y^1 + \frac{7!}{2!(7-2)!} x^{7-2} y^2 + \frac{7!}{3!(7-3)!} x^{7-3} y^3 + \\ \frac{7!}{4!(7-4)!} x^{7-4} y^4 + \frac{7!}{5!(7-5)!} x^{7-5} y^5 + \frac{7!}{6!(7-6)!} x^{7-6} y^6 + \frac{7!}{7!(7-7)!} x^{7-7} y^7$$

Aplicamos la suma de los términos queda:

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

- Si n=8

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

$$(x+y)^8 = \frac{8!}{0!(8-0)!} x^{8-0} y^0 + \frac{8!}{1!(8-1)!} x^{8-1} y^1 + \frac{8!}{2!(8-2)!} x^{8-2} y^2 + \frac{8!}{3!(8-3)!} x^{8-3} y^3 + ... \\ ... + \frac{8!}{4!(8-4)!} x^{8-4} y^4 + \frac{8!}{5!(8-5)!} x^{8-5} y^5 + \frac{8!}{6!(8-6)!} x^{8-6} y^6 + \frac{8!}{7!(8-7)!} x^{8-7} y^7 + ... \\ ... + \frac{8!}{8!(8-8)!} x^{8-8} y^8$$

Aplicamos la suma de los términos queda:

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

- Si n=9

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

$$\begin{aligned} (x+y)^9 &= \frac{9!}{0!(9-0)!} x^{9-0} y^0 + \frac{9!}{1!(9-1)!} x^{9-1} y^1 + \frac{9!}{2!(9-2)!} x^{9-2} y^2 + \frac{9!}{3!(9-3)!} x^{9-3} y^3 + \\ &\quad \frac{9!}{4!(9-4)!} x^{9-4} y^4 + \frac{9!}{5!(9-5)!} x^{9-5} y^5 + \frac{9!}{6!(9-6)!} x^{9-6} y^6 + \frac{9!}{7!(9-7)!} x^{9-7} y^7 + \\ &\quad \frac{9!}{8!(9-8)!} x^{9-8} y^8 + \frac{9!}{9!(9-9)!} x^{9-9} y^9 \end{aligned}$$

Aplicamos la suma de los términos queda:

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

- Si n=10

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

$$\begin{aligned} (x+y)^{10} &= \frac{10!}{0!(10-0)!} x^{10-0} y^0 + \frac{10!}{1!(10-1)!} x^{10-1} y^1 + \frac{10!}{2!(10-2)!} x^{10-2} y^2 + \frac{10!}{3!(10-3)!} x^{10-3} y^3 + \\ &\quad \frac{10!}{4!(10-4)!} x^{10-4} y^4 + \frac{10!}{5!(10-5)!} x^{10-5} y^5 + \frac{10!}{6!(10-6)!} x^{10-6} y^6 + \frac{10!}{7!(10-7)!} x^{10-7} y^7 + \\ &\quad \frac{10!}{8!(10-8)!} x^{10-8} y^8 + \frac{10!}{9!(10-9)!} x^{10-9} y^9 + \frac{10!}{10!(10-10)!} x^{10-10} y^{10} \end{aligned}$$

Aplicamos la suma de los términos queda:

$$\begin{aligned} (x+y)^{10} &= x^{10} + 10x^9y + 45x^8y^2 + 120x^7y^3 + 210x^6y^4 + 252x^5y^5 + 210x^4y^6 + 120x^3y^7 + \\ &\quad + 45x^2y^8 + 10xy^9 + y^{10} \end{aligned}$$

Capítulo 2

2.1 Ejemplos de permutaciones

1. La cantidad de opciones que tenemos para escoger en el menú de un restaurante de comida corrida
2. La cantidad de opciones de asignar un Presidente, Secretario y Tesorero de un equipo
3. Las diferentes maneras de tener un campeón y un subcampeón en un torneo de football
4. Las distintas opciones de asignar una clave a un candado
5. Las distintas opciones de seleccionar 4 números para el PIN de una tarjeta bancaria
6. Las distintas opciones de seleccionar una contraseña
7. Asignar la posición de salida de los autos de una carrera de formula 1
8. Las distintas opciones de iniciar las tareas del día
9. El trazo de una figura sin la necesidad de alzar el lápiz
10. Sacar canicas de una bolsa

2.2 Ejemplos de combinaciones

1. Las pelotas que van saliendo de una tombola
2. Las cartas que van saliendo en la lotería
3. Las cartas que se van recibiendo en una mano de poker
4. Las fichas que se van seleccionando en el inicio de un juego de domino
5. El conteo de las bolas de billar insertadas en las troneras
6. Las bolas de nieve en un cono
7. Las pinturas utilizadas para generar otro color
8. Los ingredientes para realizar una ensalada
9. El trazo de una figura sin importar si se alza el lápiz
10. Las cartas en la mesa de una mesa de Poker Texas

2.3 Triángulo de Pascal

En matemática, el triángulo de Pascal es una representación de los números combinatorios ordenados en forma triangular. Es llamado así en honor al matemático francés Blaise Pascal, quien introdujo esta notación en 1654, en su *Traité du triangle arithmétique*.¹ Si bien las propiedades y aplicaciones del triángulo fueron conocidas con anterioridad al tratado de Pascal por matemáticos indios, chinos o persas, fue Pascal quien desarrolló muchas de sus aplicaciones y el primero en organizar la información de manera conjunta.

2.3.1 Construcción del 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:

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}, \quad 0 \leq k \leq n \in \mathbb{N} \quad (12)$$

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.

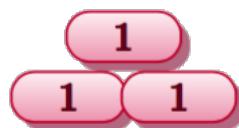


Figura 1: Dos niveles del triángulo de Pascal

Para $m = 2$ tenemos en primer lugar los dos números combinatorios de los extremos, que corresponden a $n = 0$ y $n = 2$, i.e. $\binom{2}{0} = 1$ y $\binom{2}{2} = 1$, que coloquemos debajo de los anteriores. Utilizando la Eq(12): $\binom{2}{1} = \binom{1}{0} + \binom{1}{1} = 1 + 1 = 2$. Si colocamos este número en el centro de la tercera fila observamos que su valor es la suma de los dos números que se encuentran sobre él:

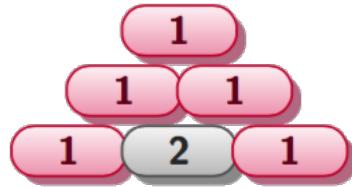


Figura 2: Tres niveles del triángulo de Pascal

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: $\binom{3}{1} = 1 + 2 = 3$, $\binom{3}{2} = 2 + 1 = 3$

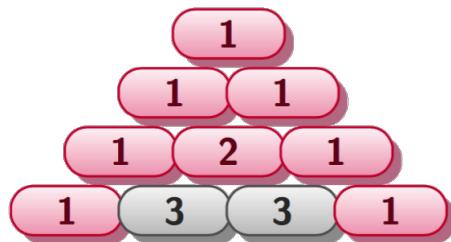


Figura 3: Cuatro niveles del triángulo de Pascal

Si continuamos este proceso inductivamente obtenemos el triángulo conocido como triángulo de Pascal.

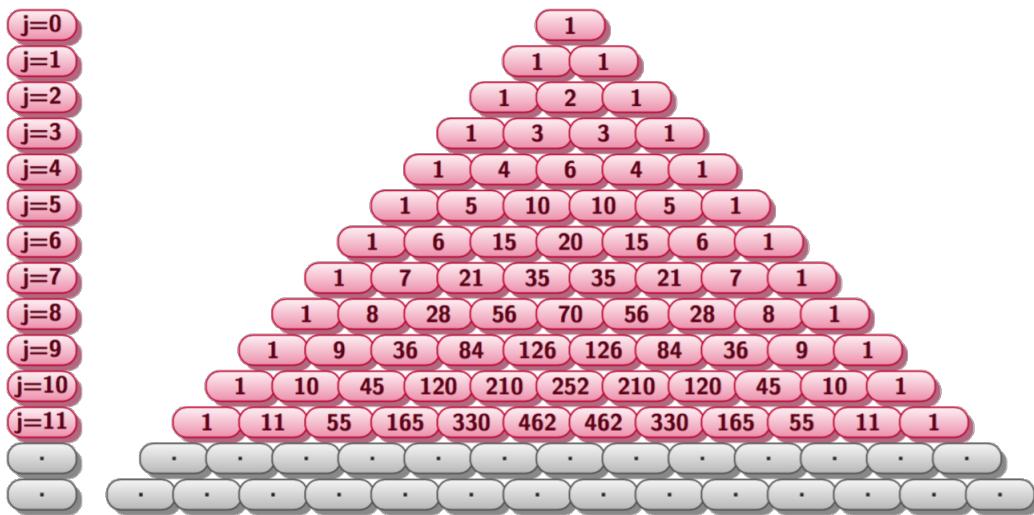


Figura 4: Triángulo de Pascal

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 $\binom{7}{4} = 35$

2.3.2 Otra manera de construir el triángulo de Pascal

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.

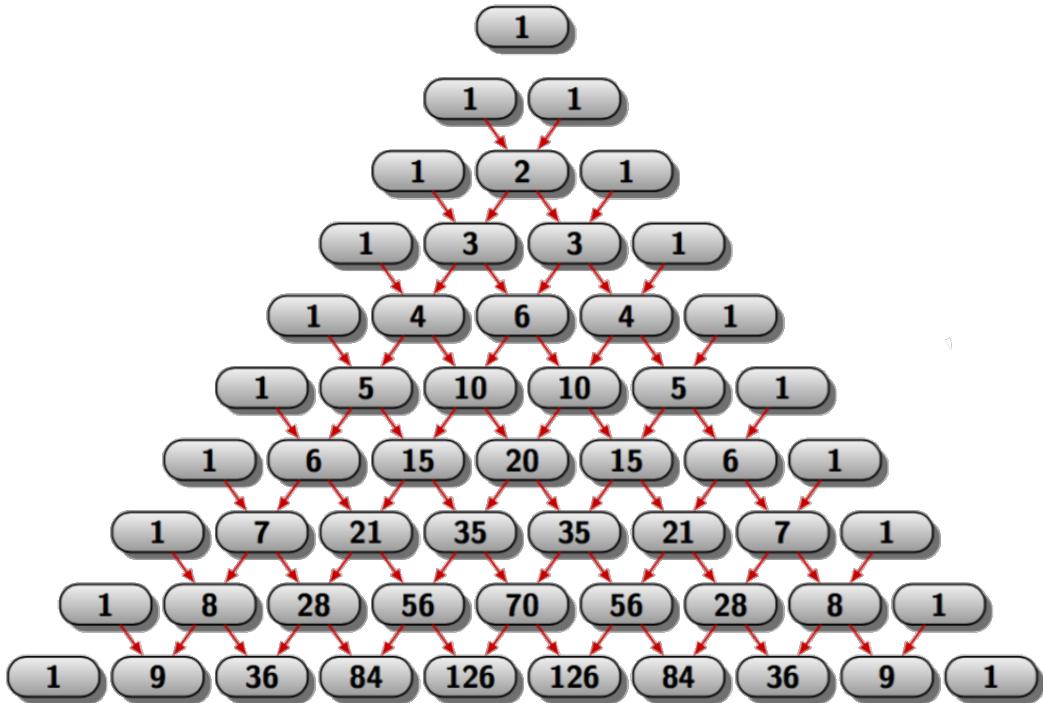


Figura 5: Caminos en el triángulo de Pascal

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:

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

Tambien 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:

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

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

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

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

2.3.3 Multiplos de 2 en el triángulo de Pascal

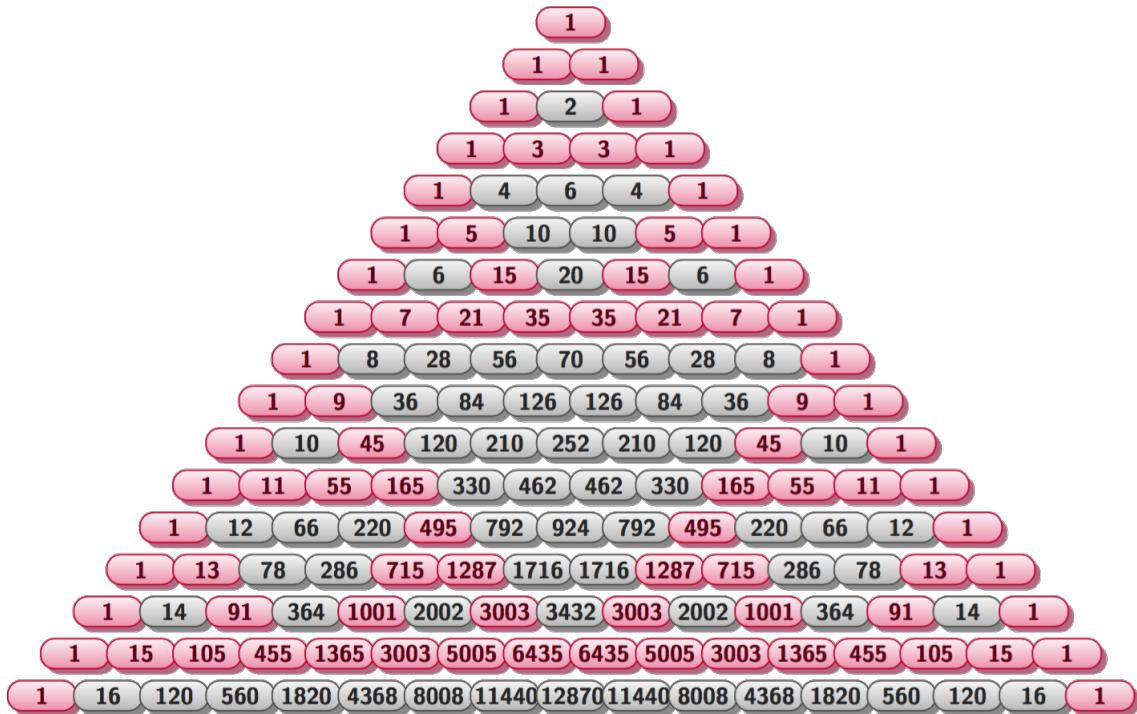


Figura 6: Multiplos de 2 en el triángulo de Pascal

2.3.4 Multiplos de 5 en el triángulo de Pascal

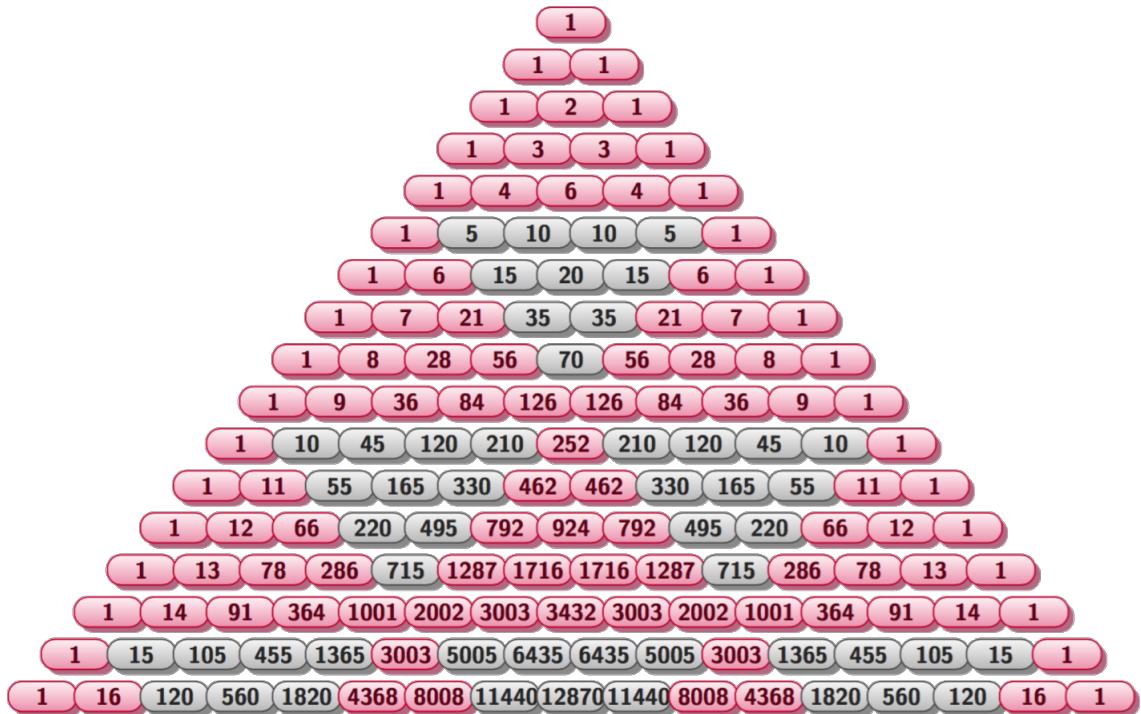


Figura 7: Multiplos de 5 en el triángulo de Pascal

2.3.5 Multiplos de 7 en el triángulo de Pascal

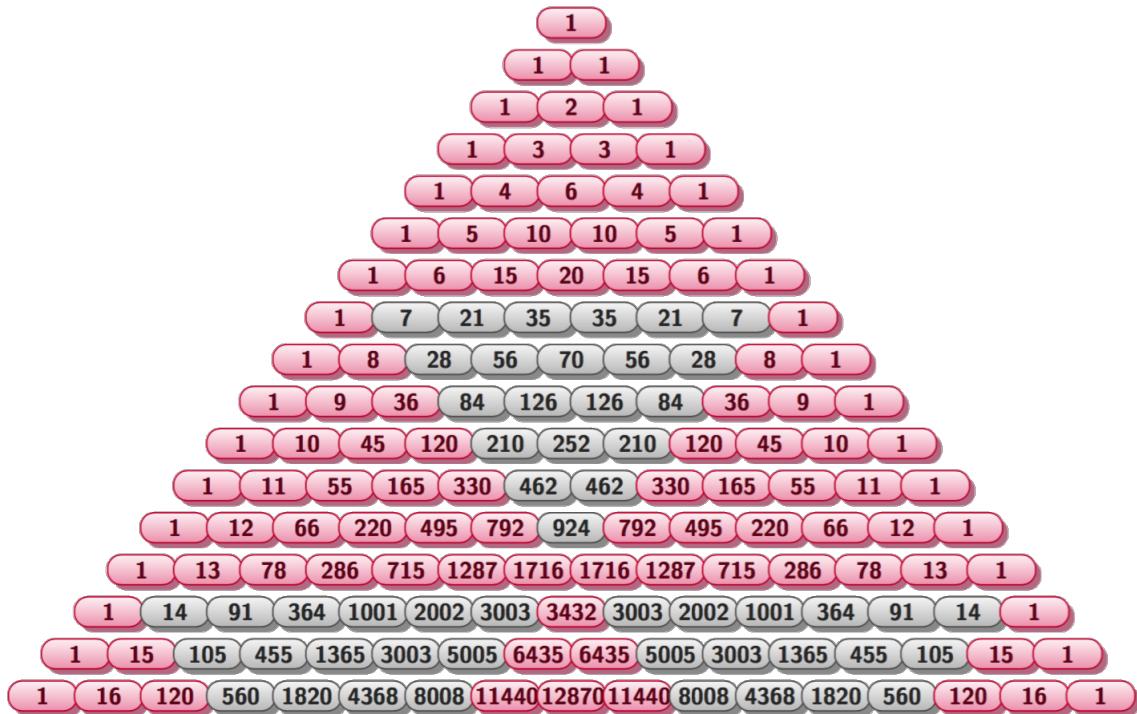


Figura 8: Multiplos de 7 en el triángulo de Pascal

2.3.6 Multiplos de 11 en el triángulo de Pascal

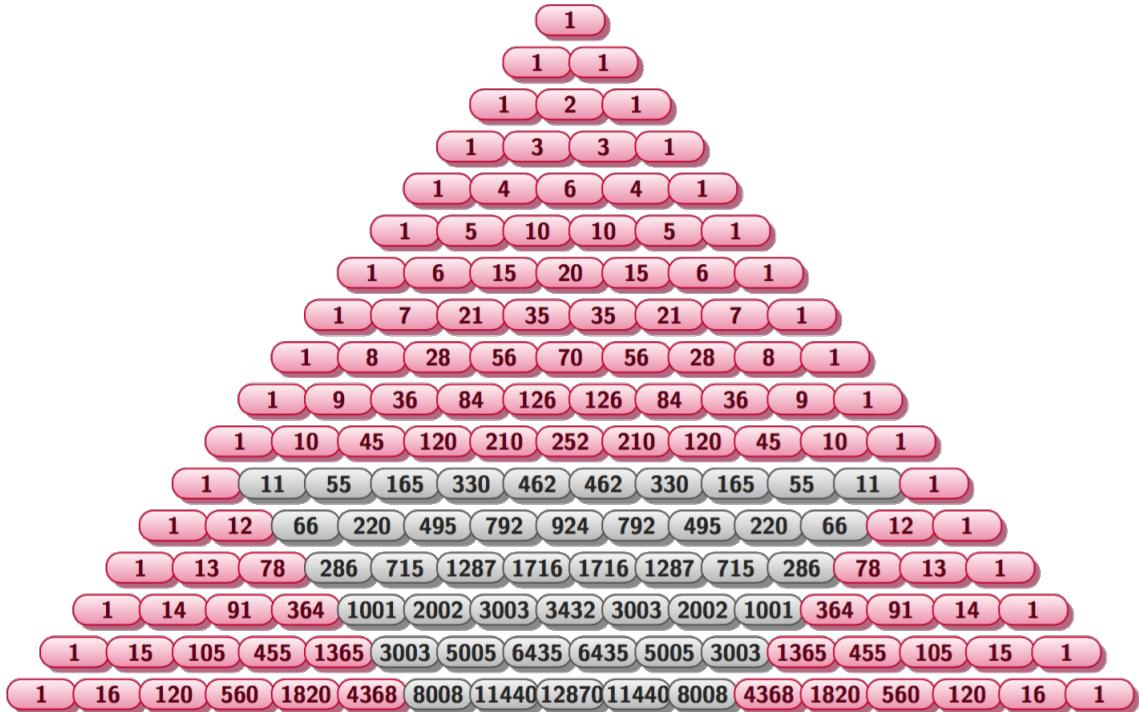


Figura 9: Multiplos de 11 en el triángulo de Pascal

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

Se tiene un cuadrado como el que se muestra en la figura, y se quiere llegar del punto A al punto B ¿cuántas posibles caminos hay para llegar?

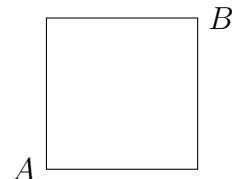
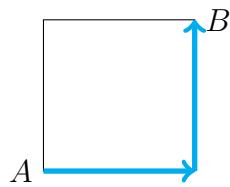


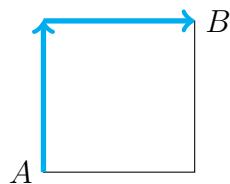
Figura 10: Caminos posibles para un cuadrado

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

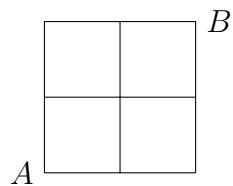
- Primera opción: se recorre primero a la derecha y después hacia arriba.



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

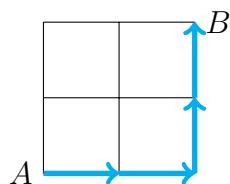


En una cuadrícula de 2×2 como la que se muestra a continuación ¿cuantos caminos posibles hay del punto A al punto B?

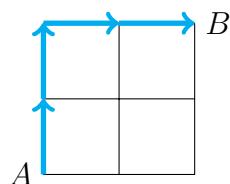


Empecemos a hacer los recorridos:

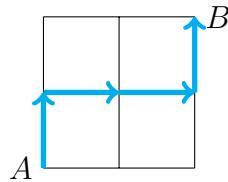
- Primer camino: se recorre dos hacia la derecha y dos hacia arriba.



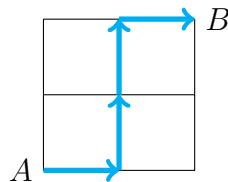
- Segundo camino: se recorre dos hacia arriba y dos hacia la derecha.



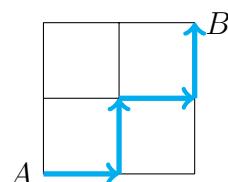
- Tercer camino: se recorre uno hacia arriba, dos hacia la derecha y por último uno hacia arriba.



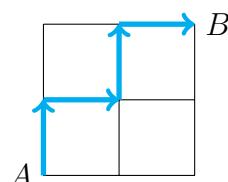
- Cuarto camino: se recorre uno hacia la derecha, dos hacia arriba y por último uno hacia la derecha.



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



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



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

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 permutaciones con repetición de 4 elementos y repeticiones 2 y 2.

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(13).

$$PR_n^{a,b,c,\dots} \frac{P_n}{a!b!c!\dots} \quad (13)$$

Para este ejemplo queda:

$$PR_4^{2,2} = \frac{P_4}{2!2!} = \frac{4!}{2!2!} = \frac{24}{4} = 6$$

Generalizando para un cuadrado de $n \times n$ se utiliza la Eq(13), donde:

- n es el número total de casillas que se recorren en cada camino,
- a son las casillas que se recorren hacia arriba y
- b son las casillas que se recorren hacia abajo.

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

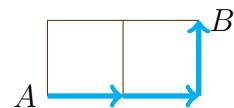
Dado un rectángulo de $m \times n$ dimensiones se quiere saber el número de caminos posibles que hay para llegar del punto A al punto B , en este caso se tiene un rectángulo de 2×1 .



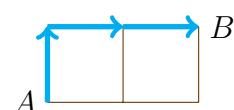
Figura 11: Caminos posibles para un cuadrado

Si hacemos todos los recorridos posibles quedaría de la siguiente forma:

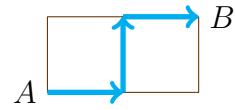
- Primer camino: se recorren dos hacia la derecha y uno hacia arriba.



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



- Tercer recorrido: se recorre uno hacia la derecha, uno hacia arriba y finalmente uno hacia la derecha.



Se observa que el número total de caminos es tres, si hacemos uso de la Eq(13), tomando en cuenta que se recorren siempre tres casillas y que de estas son dos hacia la derecha y una hacia arriba, queda:

$$PR_3^{2,1} = \frac{P_3}{2!1!} = \frac{3!}{2!1!} = \frac{6}{2} = 3$$

2.6 Reseña de artículos

2.6.1 Reseña del artículo ”Primates Count”

Tal vez el cerebro ya se de un animal o de algún humano es capaz de generar un contador cuando uno aprende a realizar un conteo superficial como lo conocemos, pero el cerebro genera un contador especial y no uno relacionado al contador que nosotros conocemos, y es por esto que es posible que sin darnos cuenta podemos contar de distintas maneras al igual los animales. Y seguramente otros sentidos estén interconectados con este contador puesto que un ejemplo sería la vista, pues el ejemplo del primate al llegar a 7 bananas en una pantalla ya no le es posible contarlas, nosotros tal vez podamos contar un poco más pero llegará el momento cuando ya no podamos contar ciertos objetos en la pantalla.

2.6.2 Reseña del artículo ”Cicada-Generate Prime Numbers”

Por lo visto la naturaleza está más relacionada con las matemáticas de lo que actualmente sabemos, puesto que como es posible que el insecto Cicada tenga un instinto de cuando debe de salir para disminuir las probabilidades de que un depredador la tome como su presa.

2.6.3 Reseña del artículo ”Ant Odometer”

Es posible que al momento de cortar el extremidades de las patas de la hormiga se dañe de alguna forma el contador que permite que regrese a su nido únicamente habiendo

recorrido grandes distancias y por ende este se reinicia o deja de funcionar por cierto tiempo. Y tal vez por esto cuando la hormiga al tener las protesis y al estar en su nido su contador está reiniciado y por eso como inicia en un nuevo conteo puede regresar sin problemas a su nido como en un inicio.

2.6.4 Reseña del artículo "Quipu"

Creo que esta es una forma de ver que los números como los conocemos sólo es es una representación gráfica que nuestro cerebro asocia al contador creado en él, creo que este contador puede recibir cualquier referencia gráfica para funcionar. Pues en los "Quipus", esta representación estaba hecha por nudos, colores, posiciones, direcciones, etc.

Capítulo 3

3.1 Magnetic Cloaking

En la última decada la investigación en invisibilidad se ha dearrollado basandose en las propiedades de los metamateriales, los cuales son materiales artificiales con propiedades electromagnéticas específicas que proceden de su estructura y no se su composición.

En un inicio se analizó como los campos electromagnéticos podían ser redireccionados utilizando estos materiales.

En la actualidad, se han desarrollado estructuras con los metamateriales, capaces de esconder objetos en dos y tres dimensiones, pero con dificultades para trabajar en altas frecuencias, como la de la luz visible.

3.2 Compuertas lógicas con interruptores

Se utilizarán botones como entradas lógicas y una resistencia como la carga que recibirá el resultado de la operación.

a) *AND*.

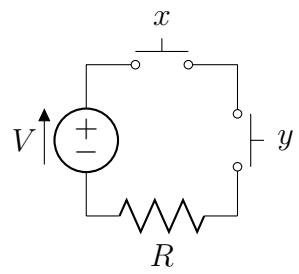


Figura 12: Compuerta AND

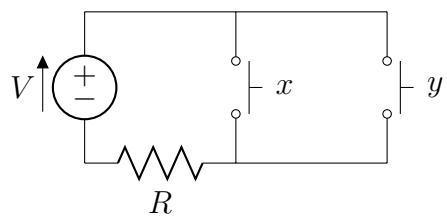
b) OR .

Figura 13: Compuerta OR

3.3 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 teómero, 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.

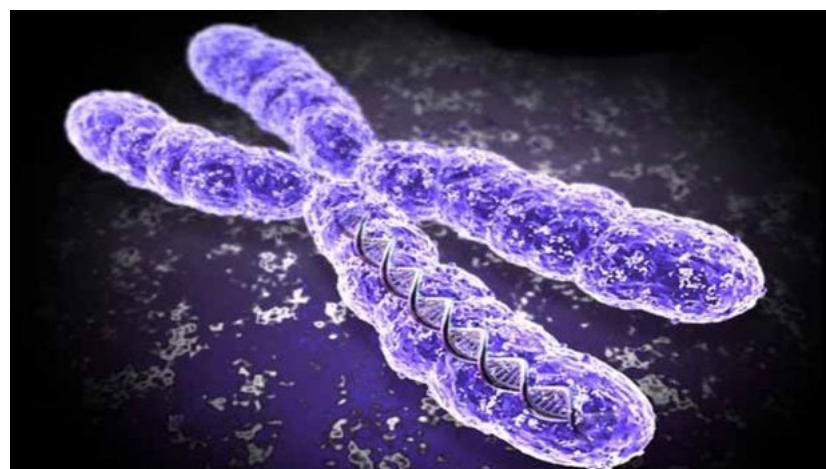


Figura 14: Telómero

3.4 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.

Por razones que se temían, la oveja Dolly tendrá una vida corta y envejecerá prematuramente. La causa esta en que ella se clono a partir de una célula que ya se había dividido varias veces cuando formaba parte de los tejidos de la madre, y el numero de veces que una célula puede dividirse esta inscrito en los cromosomas. Ello a su vez depende del largo de los telómeros. Recientemente los investigadores del Rosselin Institute, de Edimburgo, y de la compañía PPL Therapeutic, que clonaron a la Dolly, comunicaron que aun cuando ésta se encuentra en buenas condiciones de salud, posiblemente tenga una vida más corta, y envejezca prematuramente debido a que los telómeros de sus células son más cortos. El hallazgo no es inesperado. Por el contrario era un riesgo perfectamente posible de esperar, y por ello los investigadores tenían presente la necesidad de medir sus telómeros.

Para comprender lo que ello significa, es necesario recordar que los organismos multicelulares nacen con un determinado número de células, cuya vida de cada una es limitada, y por ello deben estar constantemente renovándose. Todos los tejidos, a excepción de las células cerebrales (salvo excepciones recientemente descritas) (Cae un dogma: el cerebro humano puede generar neuronas) tienen que continuar renovándose a lo largo de la vida. Para ello las células están en constante división, lo que produce células nuevas que reemplazan a las viejas que van muriendo. Es en este proceso donde parecen jugar un importante rol los telómeros. Estos son pequeños trozos de DNA que se encuentran en los extremos de los cromosomas y que no codifican proteínas. Lo que sucede es que en el mecanismo de replicación del DNA, que inicia la división celular, los cromosomas no pueden copiarse hasta el final, de modo que un pedacito de telómero que está en los extremos, es desechable y se pierde

con cada división celular. Es así como con cada división celular se va perdiendo un pedacito, hasta que llega un momento en que ya no hay telómero y entonces la célula ya no puede seguir dividiéndose. Es decir, los telómeros limitan el número de divisiones que una célula puede tener por lo que el órgano va perdiendo la posibilidad de renovarse a sí mismo. Ello está íntimamente relacionado con el proceso de envejecimiento celular, que de este modo estaría determinado genéticamente.



Figura 15: Oveja Dolly

Lo que ha sucedido con la oveja Dolly es que ella se clonó a partir de una célula de una oveja adulta, cuyas células ya se habían dividido un cierto número de veces, y por lo tanto sus telómeros se habían acortado en igual proporción. Al clonarse las células adultas de la madre retrocedió su reloj biológico volviendo a ser pluripotencial, pero en el proceso sus telómeros persistieron cortos. Lo probable es entonces que Dolly viva el tiempo que le queda de telómeros para que las células continúen dividiéndose. Pero como ésta lo recibió más corto, posiblemente sus células continúen con el restante número de divisiones, que son las mismas divisiones que le quedaron a las células de su madre.

Ahora cabe esperar que igual estudio de los telómeros se realice en los numerosos otros animales que han sido clonados siguiendo las técnicas que se usaron en la clonación de Dolly. Sin embargo, la hija de Dolly llamada Bonnie, que nació en la forma tradicional, tiene los telómeros de sus cromosomas normales, por lo que su expectativa de vida es la misma que una oveja normal.

La parte buena de la noticia es que esto va a desestimar las iniciativas para clonar un ser humano. Nadie va a querer clonarse si ya al nacer su expectativa de vida va a ser corta.

Lo que se sabe y no se sabe de los telómeros.

Un interesante y reciente trabajo realizado en ratas comienza a demostrarnos la importancia de los telomeros en la biología de este animal. Se observó que si este animal no puede mantener en forma sus telomeros, se torna prematuramente gris y tiene un alto riesgo de sufrir un cáncer.

Ronald DePinho y colaboradores del Dana Farber Cancer Institute y la Facultad de Medicina de la Universidad de Harvard, en Boston, modificaron genéticamente una rata a la que le suprimieron una enzima llamada "telomerasa". Esta enzima normalmente se encuentra en las células stem (Células Totipotenciales), que van a dar origen a los óvulos, los espermios y también a las células sanguíneas, todas las que necesitan dividirse muy rápido y continuamente. La telomerasa tiene una acción muy particular: ayuda a mantener la longitud de los telomeros, renovando su longitud. La mayor parte de las células humanas no tienen esta enzima, y por ello los telomeros se van gastando en la medida que se dividen las células. Ello no sucede en los óvulos, los espermios y en las células progenitoras sanguíneas humanas, que sí poseen esta enzima y pueden continuar dividiéndose.

Lo interesante es que los telomeros de las ratas son cinco veces más largos que los telomeros de las células humanas. Por ello parecía que en el caso de las ratas, los telomeros no están relacionados con el envejecimiento, como es el caso de los humanos. Por esta razón DePinho y sus colaboradores decidieron producir en la rata una modificación genética, privándolas de la enzima telomerasa, para así producir en ellas un envejecimiento como el que se observa en los humanos.

Los investigadores observaron que después de tres generaciones de ratas sin telomerasa, sus telomeros eran cortos y a consecuencia de ello se ponían grises en la medida que envejecían y en su piel aparecían manchas de vejez. También su estructura era débil y su piel muy delgada y sus heridas demoraban mucho en sanar (un 50 por ciento más de tiempo). Después de seis generaciones sin telomerasa, se acortaba también su vida, de modo que estas últimas vivían sólo dos tercios del tiempo, comparado con las ratas normales.

Pero lo otro que observaron fue que las ratas sin telomerasa tenían un riesgo seis veces mayor de desarrollar un cáncer, especialmente cánceres del sistema reproductivo, de la sangre y de la piel. Esto llamó la atención, porque se sabe que las células cancerosas son las únicas en que la telomerasa está presente y por ello se dividen sin fin hasta formar un tumor. DePinho piensa que el caos que se produce por deficiencia de telomerasa en la rata, durante el proceso de división celular se producen también fusiones y pérdidas de cromosomas, lo que podría gatillar el comienzo del cáncer.

Los autores reconocen que en las ratas sin telomerasa, en las que esperaban reproducir los signos que se observan en la vejez de los humanos, ésta aparentemente no se produjo. Así por ejemplo, las ratas no tuvieron cataratas, osteoporosis o problemas cardiovasculares, lo que es propio de la vejez de los humanos. Parecía entonces que el acortamiento de los telomeros no es suficiente para explicarse por si solo el

proceso del envejecimiento.

De todo esto puede deducirse otra conclusión interesante. Se sabe que las células tumorales son inmortales, y ellas lo son porque son capaces de regenerar continuamente sus telomeros. Ello porque las células tumorales son capaces de producir telomerasa en forma continua. Habría entonces una posibilidad cierta de inhibir la síntesis de telomerasa de las células tumorales, con lo que no podrían continuar dividiéndose indefinidamente. Podría ser ésta una forma de tratar un tumor canceroso.

3.5 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 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.

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

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

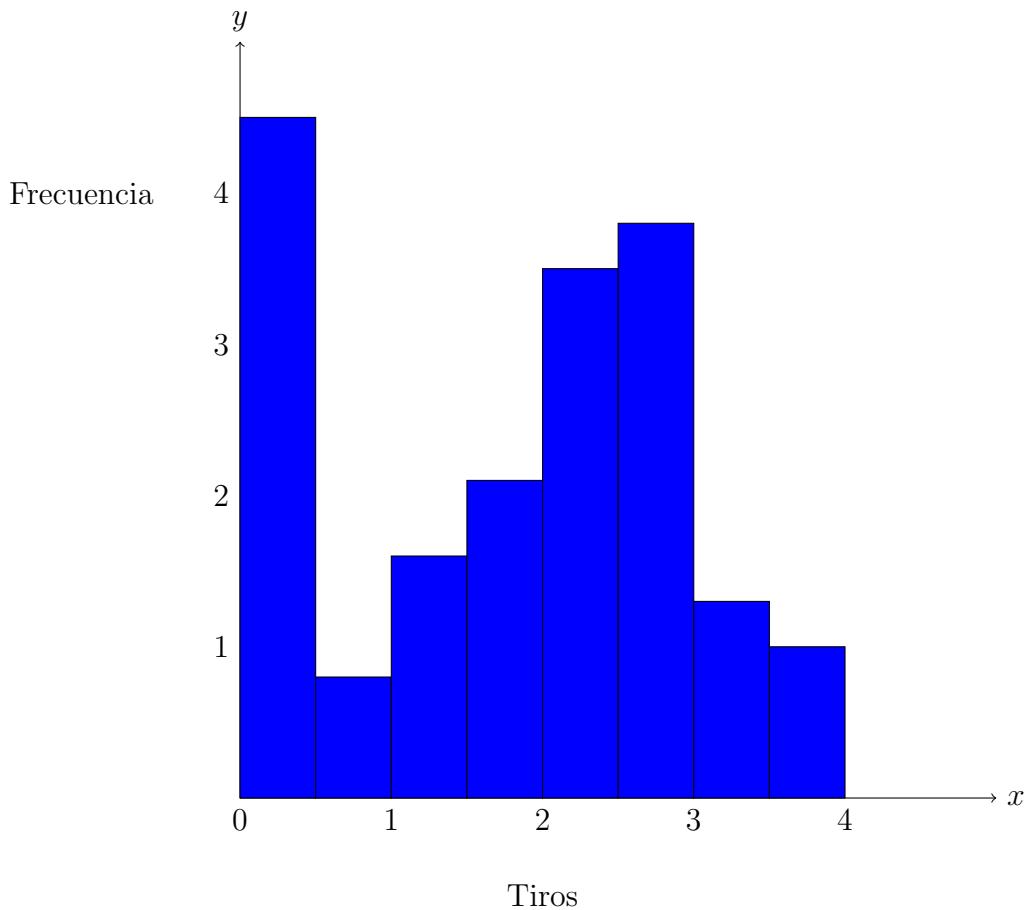


Figura 16: Histograma

3.6 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(14).

$$\ln(n!) \approx n \ln(n) - n \quad (14)$$

Definición formal.

La fórmula de Stirling está dada por:

$$\lim_{n \rightarrow \infty} \frac{n!}{\sqrt{2\pi n} \left(\frac{n}{e}\right)^n} = 1 \quad (15)$$

que se reescribe frecuentemente como:

$$n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \quad (16)$$

más exactamente la fórmula es como sigue:

$$n! = \sqrt{2\pi n} \left(\frac{n}{e}\right)^n e^{\frac{1}{12n} - \frac{1}{360n^3} + \frac{1}{1260n^5} - \frac{1}{1680n^7} + \dots} \quad (17)$$

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

Capítulo 4

4.1 Paradojas de conjuntos

La definición de conjunto que la mayoría de la gente conoce es: un conjunto es una colección de objetos. De esto tenemos los clásicos ejemplos: el conjunto de las bancas del salón, el conjunto de los perros, el conjunto de los gatos, el conjunto de las manzanas, y demás ilustrativos ejemplos que no incluyo por falta de tiempo, pero en resumen, cualesquiera objetos que se nos ocurran podemos meterlos en una bolsa y la bolsa con los objetos es un conjunto. En particular, un conjunto puede contener otros conjuntos, por ejemplo:

$$A = 1, 2, 3, 1, 2, 3, \dots$$

Podemos entonces preguntarnos si es posible que un conjunto se contenga a si mismo, i.e. si A es un conjunto, ¿es posible que A pertenezca a A ? Formalmente eso si se puede escribir como: $A := A$; o en palabras, A es el conjunto que contiene a A . ¿Puedes imaginarte un conjunto que se contenga a sí mismo?, ¿que propiedades se te ocurren que debe tener?

A finales del siglo XVIII Georg Cantor, Gottlob Frege y otros matemáticos realizaron un gran y meticuloso trabajo en la teoría de conjuntos con la esperanza de formular bases firmes tanto para la matemática como para la filosofía. Con este fin, Frege comenzó un tratado de varios volúmenes titulado Grundgesetze der Arithmetik (Fundamentos de la Aritmética o Las Bases de la Aritmética). En 1902 el segundo volumen de dicho tratado estaba en imprenta cuando Frege recibió una carta del filósofo y matemático Bertrand Russell. En dicha carta, Russell ponía el libro en alto, pero agregaba Hay un único punto en donde he encontrado una dificultad. Este único punto resultó ser la paradoja que estremeció los fundamentos de la lógica y le dió un giro a la lógica de conjuntos intuitiva de Cantor hacia una teoría axiomática.

La paradoja de Russell suele llamarse también la paradoja del Barbero de Sevilla (o simplemente paradoja del Barbero), esto se debe a que un manera famosa de enunciar esta paradoja es la siguiente: cada hombre de Sevilla es afeitado por el barbero de Sevilla si y sólo si no se afeita a sí mismo. Entonces, ¿el barbero de Sevilla se afeita a sí mismo?. Ahora, dado un conjunto A es natural pensar en los subconjuntos de A, y luego, pensar en el conjunto de los subconjuntos de A, es una manera de obtener un nuevo conjunto a partir de uno ya dado [2].

4.2 La paradoja del barbero

La paradoja de Russell ha sido expresada en varios términos más cotidianos, el más conocido es la paradoja del barbero que se puede enunciar de la siguiente manera:

En un lejano poblado de un antiguo emirato había un barbero llamado As-Samet diestro en afeitar cabezas y barbas, maestro en escamondar pies y en poner sanguijuelas. Un día el emir se dio cuenta de la falta de barberos en el emirato, y ordenó que los barberos sólo afeitaran a aquellas personas que no pudieran hacerlo por sí mismas. Cierta noche el emir llamó a As-Samet para que lo afeitara y él le contó sus angustias:

-En mi pueblo soy el único barbero. No puedo afeitar al barbero de mi pueblo, ¡que soy yo!, ya que si lo hago, entonces puedo afeitarme por mí mismo, por lo tanto ¡no debería afeitarme! Pero, si por el contrario no me afeito, entonces soy una persona que no se puede afeitar.

El emir pensó que sus pensamientos eran tan profundos, que lo premió con la mano de la más virtuosa de sus hijas. Así, el barbero As-Samet vivió para siempre feliz y barbón.

En lógica de primer orden, la paradoja del barbero se puede expresar como:

$$\forall x \quad afeita(x, \text{barbero}) \Leftrightarrow \neg afeita(x, x)$$

Donde $afeita(x, y)$ significa x es afeitado por y . Lo anterior se leería como "Cada persona es afeitada por el barbero si y sólo si no se afeita a sí misma". Es importante notar la semejanza entre las ecuaciones. Al substituir x por barbero se obtiene:

$$afeita(\text{barbero}, \text{barbero}) \Leftrightarrow \neg afeita(\text{barbero}, \text{barbero})$$

Es decir que el barbero se afeita a sí mismo si y sólo si no se afeita a sí mismo, lo cual es una contradicción.

Pero Russell duda sobre esta formulación, él mismo comenta En una ocasión me fue sugerida una formulación que no era válida; a saber, la cuestión de si el barbero se afeita o no a sí mismo. Ustedes pueden definir al barbero como alguien que afeita a todos aquellos, y sólo aquellos, que no se afeitan a sí mismos. La pregunta ahora es: ¿se afeita el barbero a sí mismo?

Explicación de la paradoja

Los conjuntos son reuniones de cosas, por ejemplo de coches, libros, personas, etc. y en este sentido los llamaremos conjuntos normales.

La característica principal de un conjunto normal es que no se contiene a sí mismo. Pero también existen conjuntos de conjuntos, como 2^M , que es el conjunto de subconjuntos de M .

Un conjunto de conjuntos es normal salvo si podemos hacer que se contenga a sí mismo. Esto último no es difícil si tenemos el conjunto de todas las cosas que NO son libros y como un conjunto no es un libro, el conjunto de todas las cosas que NO son libros formará parte del conjunto de todas las cosas que NO son libros. Estos conjuntos que se contienen a sí mismos se llaman conjuntos singulares.

Está claro que un conjunto dado o bien es normal o bien es singular, no hay término medio, o se contiene a sí mismo o no se contiene. Ahora tomemos el conjunto C como el conjunto de todos los conjuntos normales. ¿Qué clase de conjunto es C ? ¿Normal o Singular?

Si es normal, estará dentro del conjunto de conjuntos normales, que es C , luego ya no puede ser normal. Si es singular, no puede estar dentro del conjunto de conjuntos normales, luego no puede estar en C , pero si no puede estar en C entonces no es singular [11].

4.3 Paradoja y antinomia

Una paradoja (del latín *paradoxa*, plural de *paradoxon*, lo contrario a la opinión común) o antilogía es una idea extraña opuesta a lo que se considera verdadero a la opinión general. También se considera paradoja a una proposición en apariencia falsa o que infringe el sentido común, pero no conlleva una contradicción lógica, en contraposición a un sofisma que solo aparenta ser un razonamiento verdadero. En retórica, es una figura de pensamiento que consiste en emplear expresiones o frases que implican contradicción. Un ejemplo de paradoja es la "Paradoja de Jevons", más conocida como efecto rebote. La paradoja es estímulo para la reflexión.

A menudo los filósofos se sirven de las paradojas para revelar la complejidad de la realidad. La paradoja también permite demostrar las limitaciones de las herramientas

tas de la mente humana. Así, la identificación de paradojas basadas en conceptos que a simple vista parecen simples y razonables ha impulsado importantes avances en la ciencia, la filosofía y las matemáticas.

Antinomia (del griego, contra, contradicción en la ley) es un término empleado en la lógica y la epistemología que, en sentido laxo, significa paradoja o contradicción irresoluble.

Immanuel Kant sostuvo que cuando la razón rebasa la experiencia posible a menudo cae en varias antinomias; es decir, perspectivas igualmente racionales pero contradictorias. En el vocabulario kantiano, razón no remite a una facultad que cumpla la función de establecer las verdades racionales pues, según su planteamiento, esta rebasa la experiencia posible y se vuelve trascendente. Por ejemplo, Kant pensaba que se podía llegar, a partir de la suposición de que el mundo tiene un comienzo en el tiempo, a la conclusión de que no lo tenía, y viceversa. El estudio de tales fenómenos forma parte del programa crítico de Kant para determinar los límites de la ciencia y de la investigación filosófica.

De hecho, las antinomias no tienen en cuenta las limitaciones de alcance del razonamiento lógico, como a menudo se cree. Esto se debe a que la conclusión de que hay una limitación se deriva (supuestamente) de una antinomia por razonamiento lógico; por lo tanto, toda limitación de la validez del razonamiento lógico impone una limitación a la conclusión de que el razonamiento lógico tiene una limitación (éste es un argumento por autorreferencia).

En resumen, en cuanto a la validez del razonamiento lógico en su totalidad, las antinomias se aíslan solas; son como discontinuidades dispersas dentro del campo de la lógica, incapaces de poner cosa alguna en duda, salvo a sí mismas.

En la lógica, la existencia de dos enunciados contradictorios sobre un objeto, con la fundamentación lógica convincente por igual es una antinomia.

4.4 Demostración de las leyes De Morgan

$$1. (\overline{A \cup B}) = \overline{A} \cap \overline{B} \quad (18)$$

$$2. (\overline{A \cap B}) = \overline{A} \cup \overline{B} \quad (19)$$

Demostración [8]:

Sea x un elemento arbitrario del conjunto universal U .

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

$$\begin{aligned}
 &\iff \neg[x \in (A \cup B)] \quad \{\text{Negación}\} \\
 &\iff \neg[(x \in A) \vee (x \in B)] \quad \{\text{Definición de unión}\} \\
 &\iff \neg(x \in A) \wedge \neg(x \in B) \quad \{\text{De Morgan para } \vee\} \\
 &\iff (x \notin A) \wedge (x \notin B) \quad \{\text{Negación}\} \\
 &\iff (x \in \overline{A}) \wedge (x \in \overline{B}) \quad \{\text{Definición de complemento}\} \\
 &\iff x \in (\overline{A} \cap \overline{B}) \quad \{\text{Definición de intersección}\}
 \end{aligned}$$

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

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

luego,

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

Sea x un elemento arbitrario del conjunto universal U .

$$\begin{aligned}
 2. (\overline{A \cap B}) &\iff x \notin (A \cap B) \quad \{\text{Definición de complementario}\} \\
 &\iff \neg[x \in (A \cap B)] \quad \{\text{Negación}\} \\
 &\iff \neg[(x \in A) \wedge (x \in B)] \quad \{\text{Definición de intersección}\} \\
 &\iff \neg(x \in A) \vee \neg(x \in B) \quad \{\text{De Morgan para } \wedge\} \\
 &\iff (x \notin A) \vee (x \notin B) \quad \{\text{Negación}\} \\
 &\iff (x \in \overline{A}) \vee (x \in \overline{B}) \quad \{\text{Definición de complemento}\} \\
 &\iff x \in (\overline{A} \cup \overline{B}) \quad \{\text{Definición de unión}\}
 \end{aligned}$$

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

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

luego,

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

4.5 Demostración de uniones de conjuntos

4.5.1 Desarrollo de $P(A \cup B \cup C)$

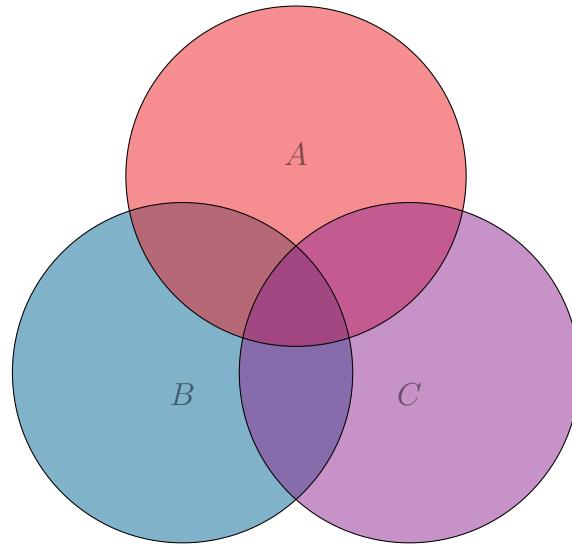
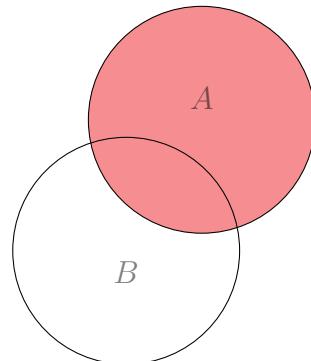
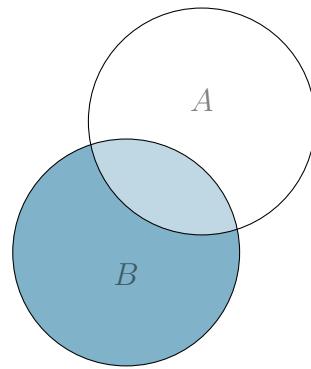


Figura 17: Union de tres conjuntos

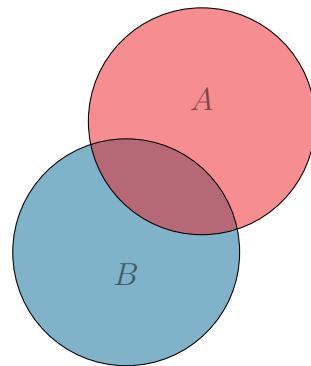
Iniciamos con $P(A \cup B)$. La probabilidad de A la podemos observar de la siguiente manera:



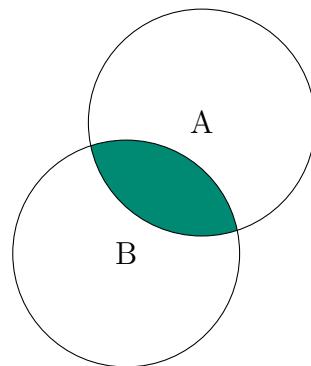
La probabilidad de B de la siguiente manera:



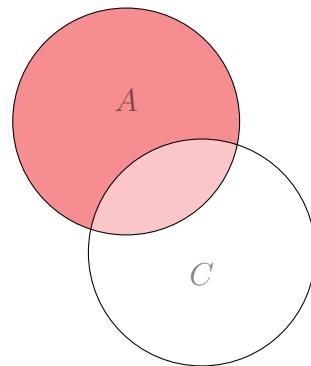
Podemos observar que teniendo las dos probabilidades simultáneas tenemos dos veces $P(A \cap B)$.



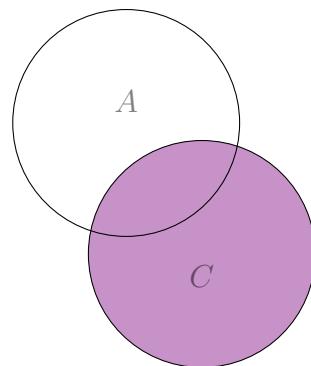
Por lo tanto debemos de restar una vez el término $P(A \cap B)$.



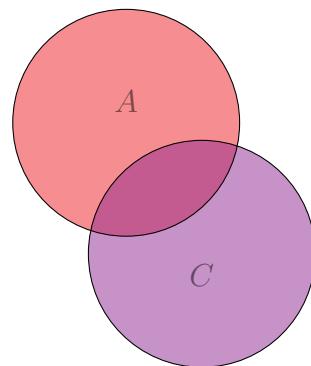
Continuamos con $P(A \cup C)$. La probabilidad de A la podemos observar de la siguiente manera:



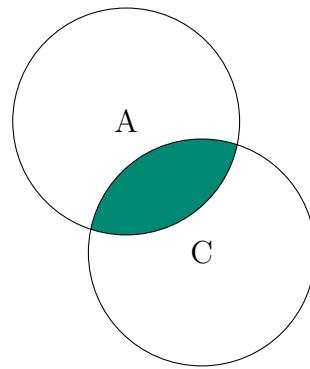
La probabilidad de C de la siguiente manera:



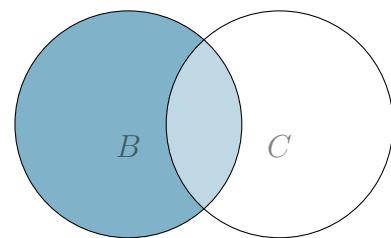
Podemos observar que teniendo las dos probabilidades simultáneas tenemos dos veces $P(A \cap C)$.



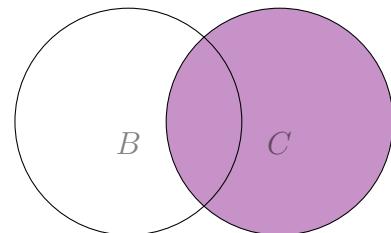
Por lo tanto debemos restar una vez el término $P(A \cap C)$.



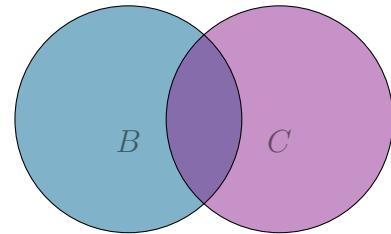
Continuamos con $P(B \cup C)$. La probabilidad de B la podemos observar de la siguiente manera:



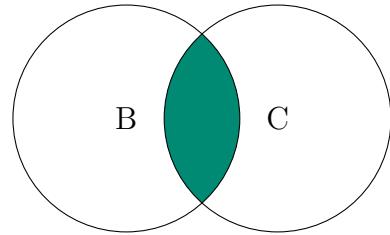
La probabilidad de C de la siguiente manera:



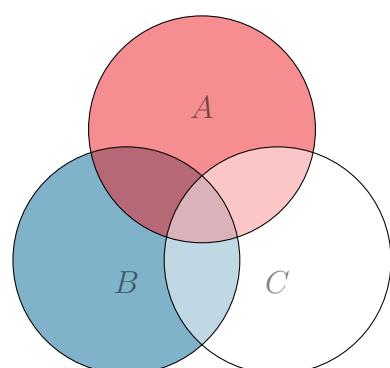
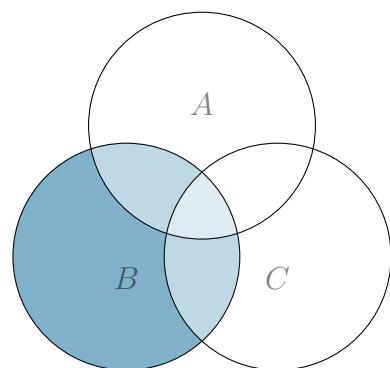
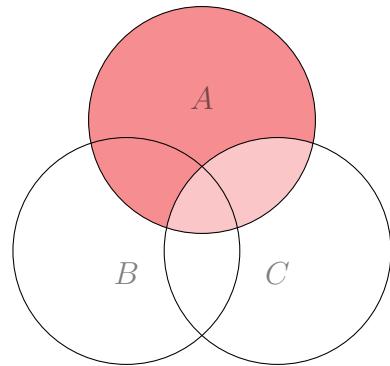
Podemos observar que teniendo las dos probabilidades simultáneas tenemos dos veces $P(B \cap C)$.

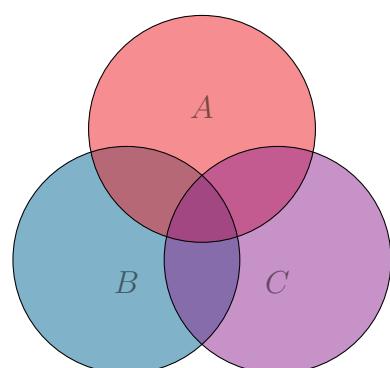
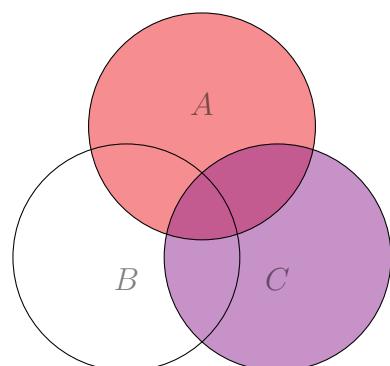
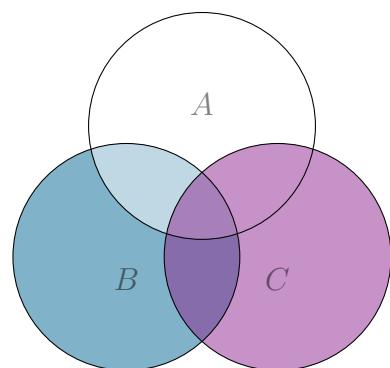
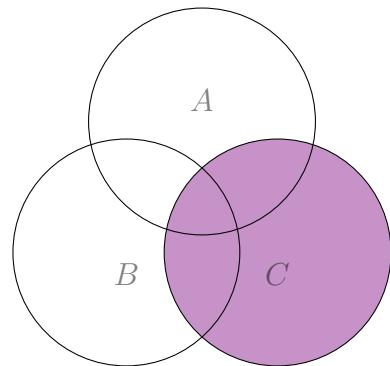


Por lo tanto debemos restar una vez el término $P(B \cap C)$.



Observemos los pasos anteriores con los tres conjuntos al mismo tiempo.





Al restar los terminos $P(A \cap B)$, $P(B \cap C)$ y $P(C \cap A)$ estamos restando tambien el termino $P(A \cap B \cap C)$ por lo tanto debemos de sumarlo.

Por lo que la ecuación queda:

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) - P(A \cap B) \\ &\quad - P(B \cap C) - P(C \cap A) + P(A \cap B \cap C) \end{aligned} \tag{20}$$

4.5.2 Desarrollo de $P(A \cup B \cup C \cup D)$

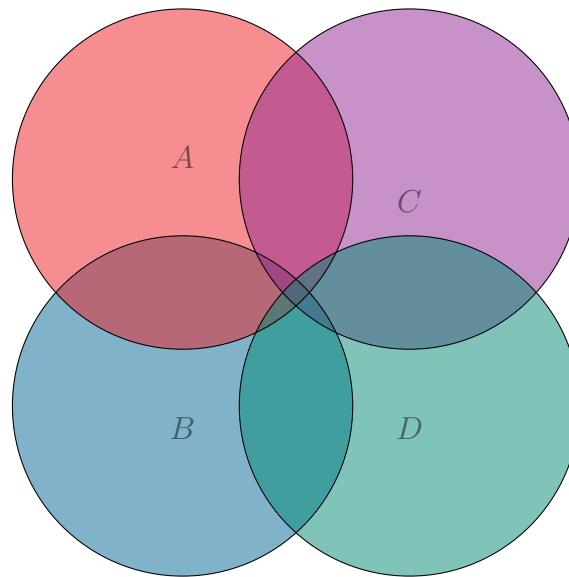
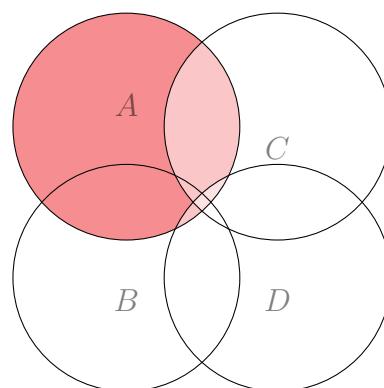


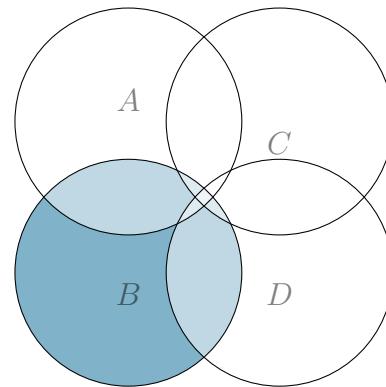
Figura 18: Unión de 4 conjuntos

Observemos las probabilidades de cada conjunto por separado.

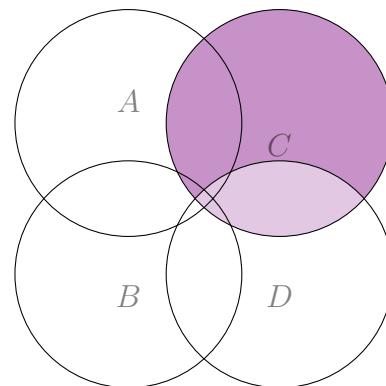
$$P(A)$$



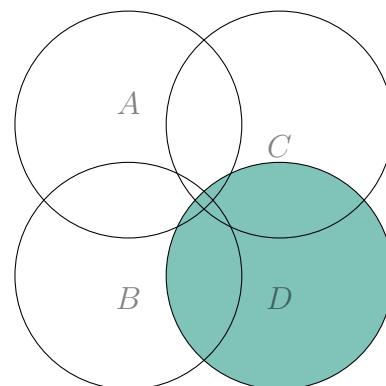
$$P(B)$$



$P(C)$

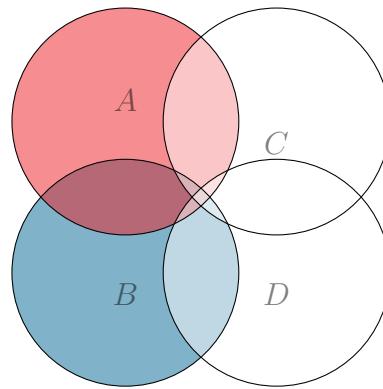


$P(D)$



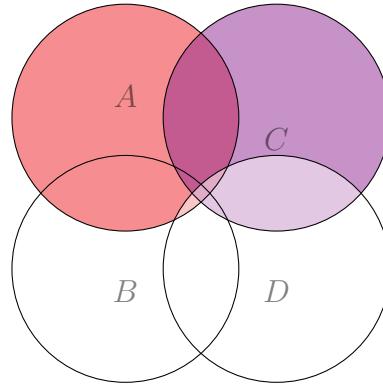
Ahora observemos las uniones de las probabilidades.

$P(A \cup B)$



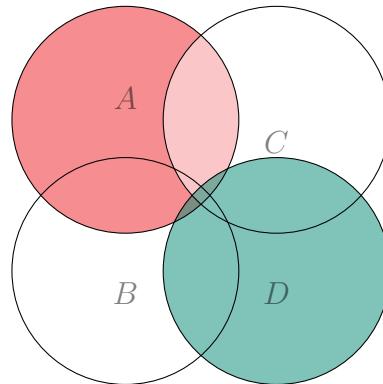
Podemos observar que tenemos dos veces la $P(A \cap B)$ por lo que es necesario restarla una vez.

$$P(A \cup C)$$



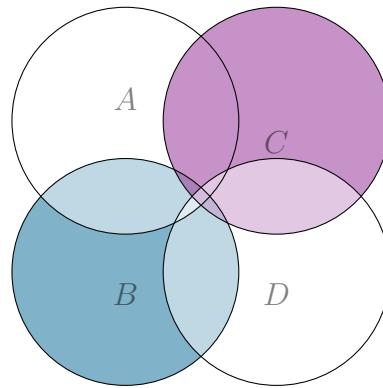
Igual que en el caso anterior tenemos dos veces la $P(A \cap C)$ por lo que es necesario restarla una vez.

$$P(A \cup D)$$



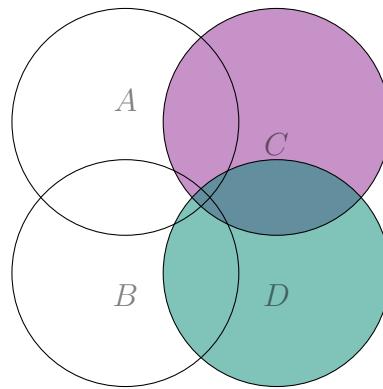
Igual que en el caso anterior tenemos dos veces la $P(A \cap D)$ por lo que es necesario restarla una vez.

$$P(C \cup B)$$



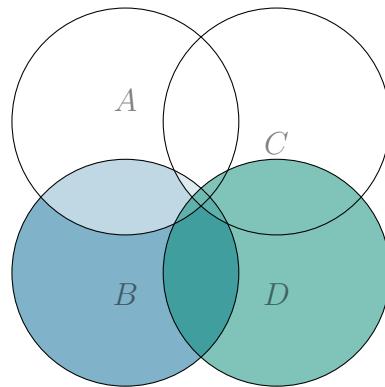
Igual que en el caso anterior tenemos dos veces la $P(C \cap B)$ por lo que es necesario restarla una vez.

$$P(C \cup D)$$



Igual que en el caso anterior tenemos dos veces la $P(C \cap D)$ por lo que es necesario restarla una vez.

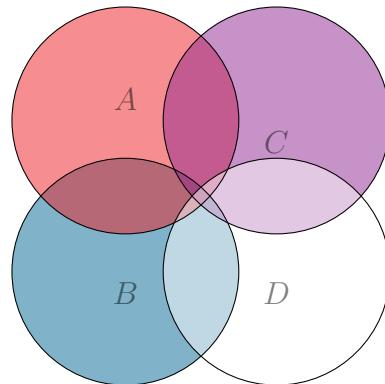
$$P(B \cup D)$$



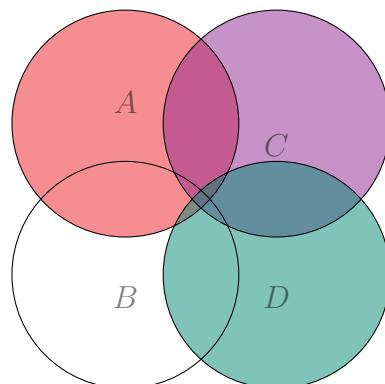
Igual que en el caso anterior tenemos dos veces la $P(B \cap D)$ por lo que es necesario restarla una vez.

Ahora observemos la union de 3 conjuntos. Llegando a este caso ya tenemos consideradas todas las restas de las intersecciones entre pares de conjuntos. Igual que en el caso de $P(A \cup B \cup C)$ al eliminar las intersecciones de cada triple de conjutnos estamos eliminado la interseccion de estos 3 conjuntos por la tanto hay que sumar este termino.

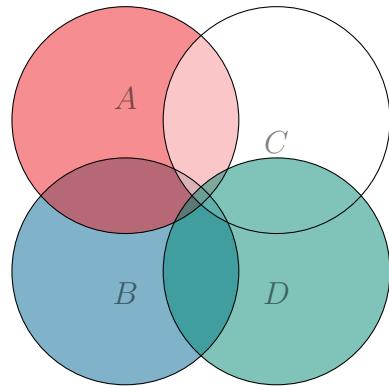
$$P(A \cup B \cup C)$$



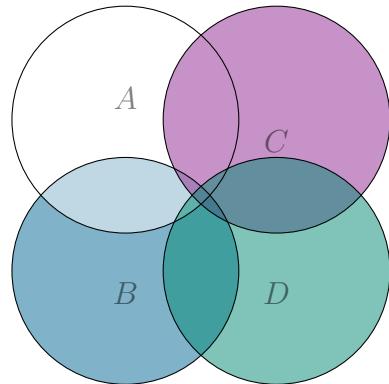
$$P(A \cup C \cup D)$$



$$P(A \cup B \cup D)$$

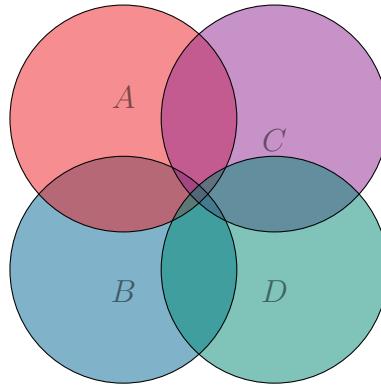


$$P(C \cup B \cup D)$$



Por ultimo al tener considerado todo lo anterior podemos observar que estamos considerando dos veces los terminos $P(A \cap B \cap C)$, $P(A \cap B \cap D)$, $P(A \cap C \cap D)$ y $P(B \cap C \cap D)$. Por lo tanto debemos de restar una vez cada termino.

Al quitar los terminos anteriores de igual forma estamos eliminando el termino $P(A \cap B \cap C \cap D)$ por lo que es necesario sumarlo. En la siguiente figura podemos observar como $P(A \cap B \cap C \cap D)$ forma parte de las intersecciones de cada tercia de conjuntos.



Por lo que la ecuación queda:

$$\begin{aligned}
 P(A \cup B \cup C \cup D) = & P(A) + P(B) + P(C) + P(D) \\
 -P(A \cap B) - P(A \cap C) - P(A \cap D) - P(B \cap C) - P(B \cap D) - P(C \cap D) \\
 +P(A \cap B \cap C) + P(A \cap B \cap D) + P(A \cap C \cap D) + P(B \cap C \cap D) \\
 -P(A \cap B \cap C \cap D)
 \end{aligned} \tag{21}$$

4.6 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(22):

$$P(A | B) = \frac{P(A \cap B)}{P(B)} \tag{22}$$

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(22), 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.

Capítulo 5

5.1 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, 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.

$$P(A_i | B) = \frac{P(B | A_i)P(A_i)}{P(B)} \quad (23)$$

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:

$$P(A_i | B) = \frac{P(B | A_i)P(A_i)}{\sum_{k=1}^n P(B | A_k)P(A_k)} \quad (24)$$

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.

dad. 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 [3].

5.2 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 frequentista, 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.

Una de las principales críticas al enfoque bayesiano es que dos personas cuyas opiniones iniciales difieren podrán llegar a conclusiones distintas a pesar de observar exactamente los mismos datos. Si bien esto es cierto en un sentido estricto, conforme se va acumulando la evidencia, las opiniones de esas dos personas deberán converger. En este sentido, la objetividad a la que aspira la ciencia podría replantearse de manera más realista como un consenso de subjetividades.[11]

5.3 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:

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

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:

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

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

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

Ahora, uno puede calcular lo siguiente:

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

$$P(\text{sano} \cap \text{negativo}) = P(\text{sano}) \times P(\text{negativo} | \text{sano}) = 0.99 \times 0.99 = 0.9801$$

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

$$P(\text{enfermo} \cap \text{positivo}) = P(\text{enfermo}) \times P(\text{positivo} | \text{enfermo}) = 0.01 \times 0.99 = 0.0099$$

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

$$P(\text{sano} \cap \text{positivo}) = P(\text{sano}) \times P(\text{positivo} | \text{sano}) = 0.99 \times 0.01 = 0.0099$$

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

$$P(\text{enfermo} \cap \text{negativo}) = P(\text{enfermo}) \times P(\text{negativo} | \text{enfermo}) = 0.01 \times 0.01 = 0.0001$$

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

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

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

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

En este ejemplo, debería ser fácil ver la diferencia entre las probabilidades condicionadas $P(\text{positivo} | \text{enfermo})$ (que es del 99%) y $P(\text{enfermo} | \text{positivo})$ (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.

5.4 Paradoja de 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 (tambien 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 ($\text{latitud} \phi = 0$), la densidad condicional para una longitud α definida en el intervalo $[-\pi, \pi]$ es:

$$f(\lambda | \phi = 0) = \frac{1}{2\pi} \quad (25)$$

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:

$$f(\phi | \lambda = 0) = \frac{1}{2} \cos \phi \quad (26)$$

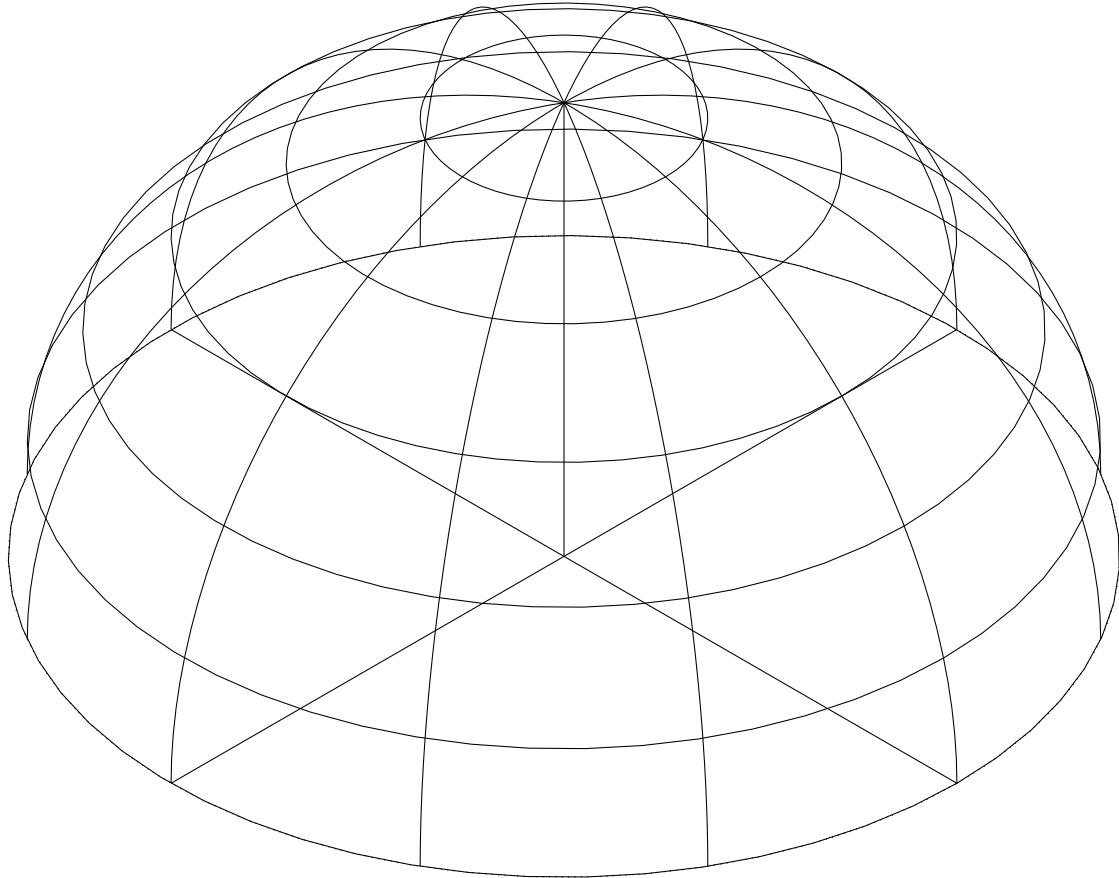


Figura 19: Paradoja de Borel -Kolmogorov

¿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 10min 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 [3].

5.5 Falacia del apostador

La falacia del apostador, tambien conocida como la falacia de Monte Carlo, es la creencia erronea 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 famoso ejemplo de este fenómeno, el cual ocurrio 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.

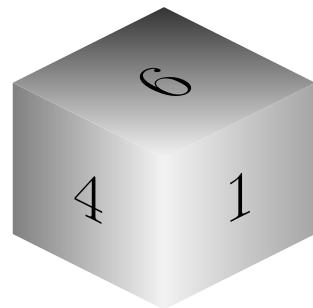


Figura 20: Falacia del apostador

5.6 Simulación de un dado de 3 caras en ROOT

5.6.1 Histograma a 10 tiradas

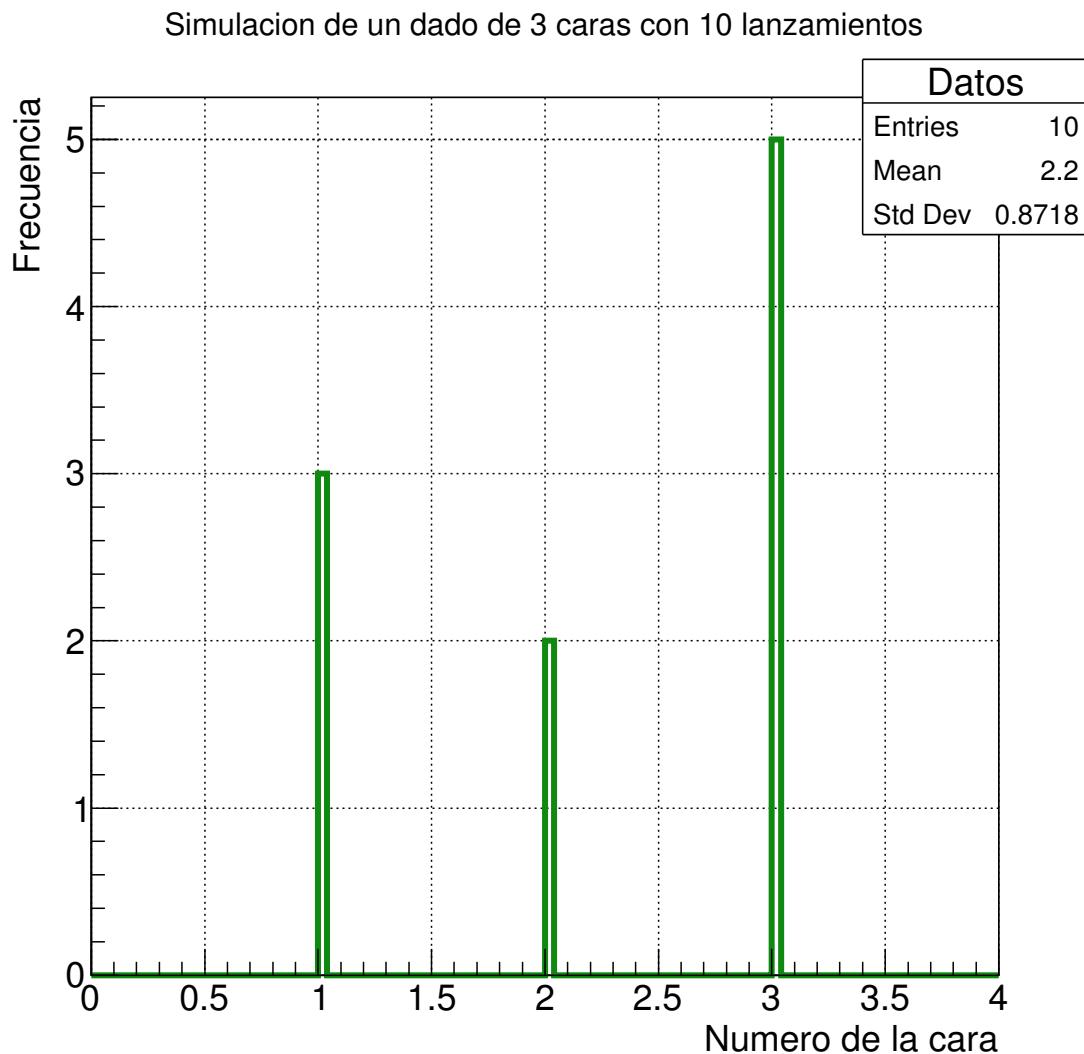


Figura 21: Histograma de 10 lanzamientos

5.6.2 Histograma a 100 tiradas

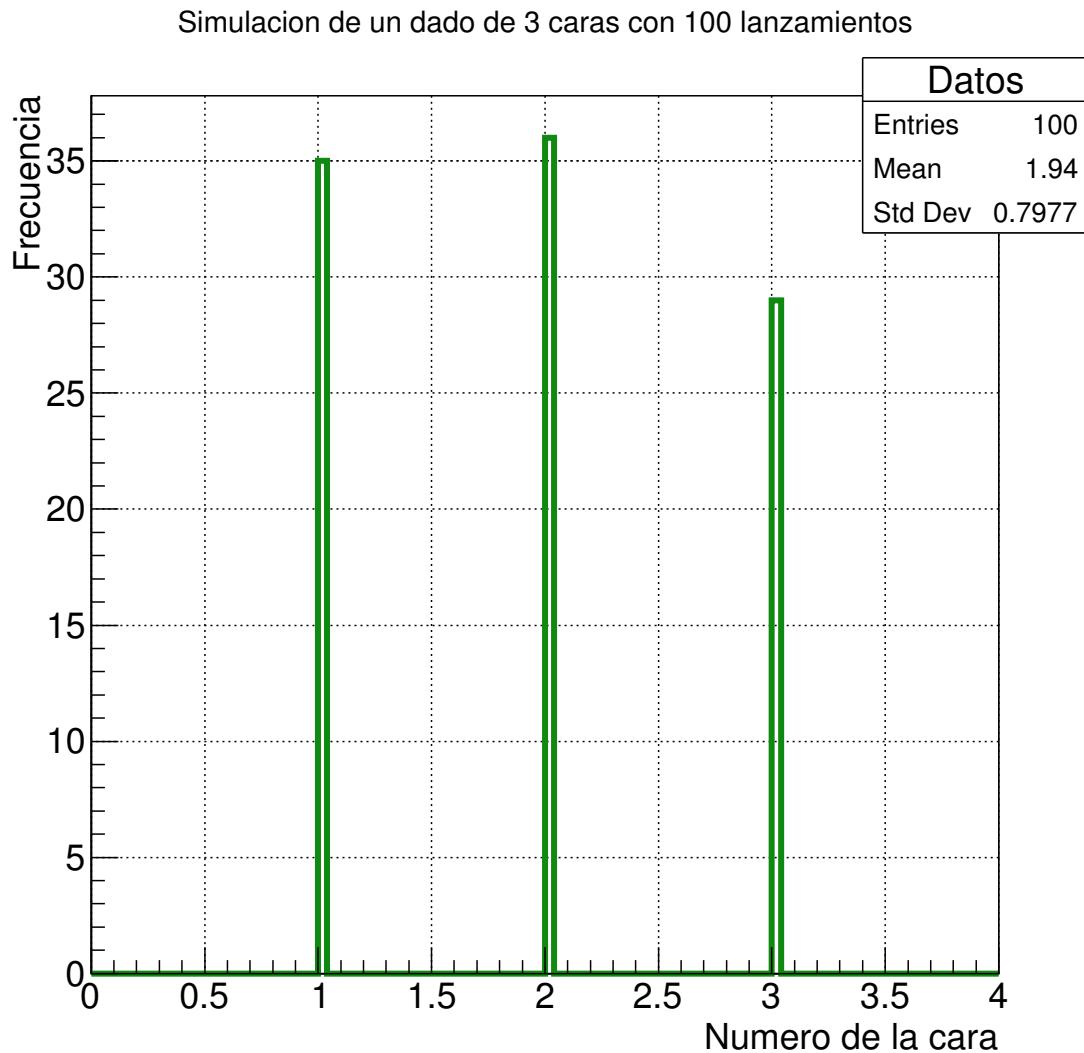


Figura 22: Histograma de 100 lanzamientos

5.6.3 Histograma a 1,000 tiradas

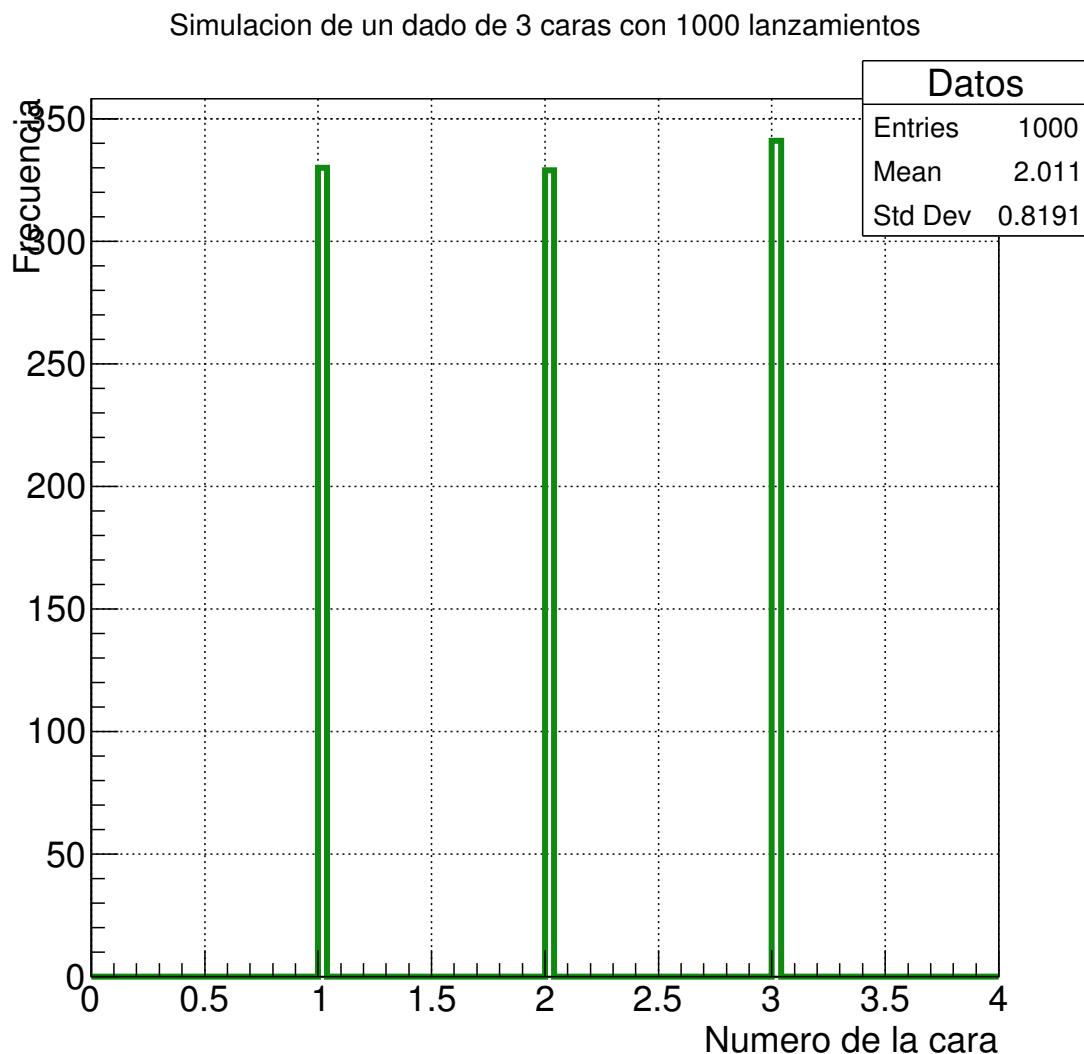


Figura 23: Histograma de 1,000 lanzamientos

5.6.4 Histograma a 10,000 tiradas

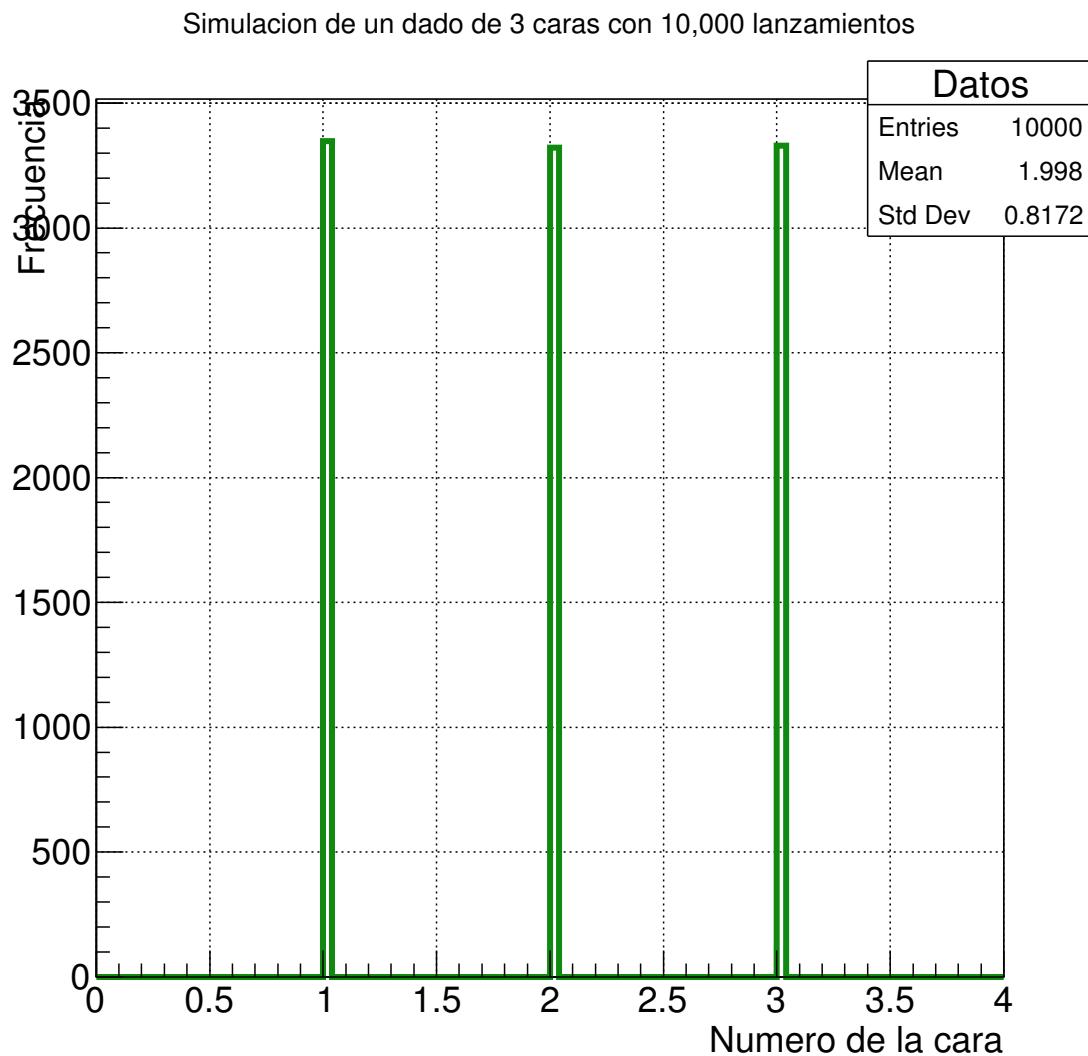


Figura 24: Histograma de 10,000 lanzamientos

5.6.5 Código de ROOT

```
//Autor: Itzael Jimenez Aranda
//Instituto: CIC - IPN
//Materia: Probabilidad, procesos aleatorios e inferencia
void simulacion_dado(){
    TRandom3 *r = new TRandom3(0);
    TCanvas *c1 = new TCanvas(
        "Simulacion_dado_3_caras",
        "Simulacion_de_un_dado_de_3_caras",
        ,200
    )
}
```

```

        ,10
        ,700
        ,700);
c1->SetGrid();

TH1I* hdado = new TH1I("Datos"
                      ,"Simulacion con 10 lanzamientos"
                      ,100
                      ,0
                      ,4);
hdado->SetLineColor(209);
hdado->SetLineWidth(4);
hdado->GetXaxis()->SetTitle("Número de la cara");
hdado->GetYaxis()->SetTitle("Frecuencia");

int aux = 0;
for(int i = 0; i < 10; i++){
    aux = (r->Integer(3))+1;
    hdado->Fill(aux);
}
hdado->Draw();
}

```

5.7 Derivadas e Integrales en Mathematica

```

F[x] := x^3 + 10 x^2 + 5 x + 2
F[y] := 10 y^9/y
F[z] := Log[z]*Sin[2 z]
F[n] := Cos[n]*Sin[n]^3
F[k] := Cos[k]*Sin[k]^3*Exp[k]
D[F[x], x]
D[F[y], y]
D[F[z], z]
D[F[n], n]
D[F[k], k]
Integrate[F[x], x]
Integrate[F[y], y]
Integrate[F[z], z]
Integrate[F[n], n]
Integrate[F[k], k]

5 + 20 x + 3 x^2

80 y^7

```

$$2 \cos[2z] \log[z] + \sin[2z]/$$

$$3 \cos[n]^2 \sin[n]^2 - \sin[n]^4$$

$$3 E^k \cos[k]^2 \sin[k]^2 + E^k \cos[k] \sin[k]^3 - E^k \sin[k]^4$$

$$2x + (5x^2)/2 + (10x^3)/3 + x^4/4$$

$$(10y^9)/9$$

$$1/2 \cosIntegral[2z] - 1/2 \cos[2z] \log[z]$$

$$\sin[n]^4/4$$

$$1/680 E^k (-68 \cos[2k] + 20\cos[4k] + 34\sin[2k] - 5\sin[4k])$$

Capítulo 6

6.1 Histograma en ROOT llenado con la distribución de Gauss

6.1.1 Histograma

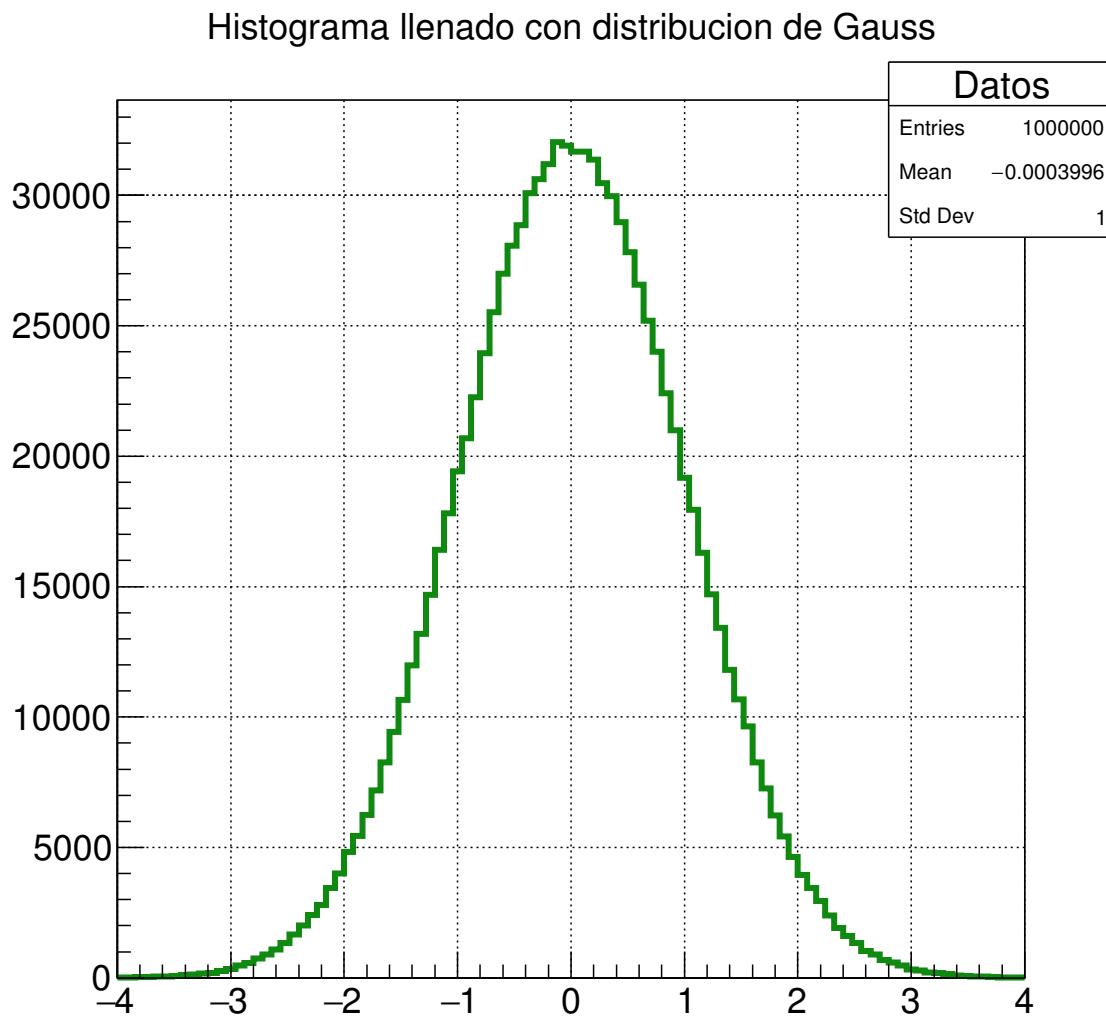


Figura 25: Histograma en ROOT llenado con la distribución de Gauss

6.1.2 Código de ROOT

```
//Autor: Itzael Jimenez Aranda  
//Instituto : CIC - IPN
```

```
//Materia: Probabilidad , procesos aleatorios e inferencia
void simulacion_gauss(){
    TCanvas *c1 = new TCanvas("simulacion_gauss"
                            , "Distribucion_de_Gauss"
                            , 200
                            , 10
                            , 700
                            , 700);
    c1->SetGrid();

    TH1I* hgauss;
    hgauss = new TH1I("Datos"
                       , "Histograma llenado con distribucion de Gauss"
                       , 100
                       , -4
                       , 4);
    hgauss->SetLineColor(209);
    hgauss->SetLineWidth(4);
    hgauss->FillRandom("gaus", 1000000);
    hgauss->Draw();
}
```

6.2 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:

$$f_{\mu, \sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (27)$$

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 dis-

tribuidas, y con una media μ y varianza σ^2 finitas ($\sigma^2 \neq 0$):

$$S_n = X_1 + \dots + X_n \quad (28)$$

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:

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} \quad (29)$$

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 :

$$\lim_{n \rightarrow \infty} P_r(Z_n \leq z) = \phi(z) \quad (30)$$

donde $Pr()$ 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:

$$S_n = X_1 + \dots + X_n \quad (31)$$

entonces:

$$\lim_{n \rightarrow \infty} P_r\left(\frac{S_n - n\mu}{\sigma\sqrt{n}}\right) = \phi(z) \quad (32)$$

6.3 Explosión del Challenger

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 instantánea y el transbordador estalla.



Figura 26: Challenger

6.4 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:

$$P[a \leq X \leq b] = \int_a^b f(x)dx \quad (33)$$

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

$$F(x) = \int_{-\infty}^x f(u)du \quad (34)$$

y (si f es continua en x)

$$f(x) = \frac{d}{dx} F(x) \quad (35)$$

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 .

$$P[a \leq X \leq b] = \int_a^b f(x)dx \quad (36)$$

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

$$f(x) = \frac{d}{dx}F(x) \quad (37)$$

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

$$F(x) = \int_{-\infty}^x f(u)du \quad (38)$$

y (si f es continua en x)

$$f(x) = \frac{d}{dx}F(x) \quad (39)$$

6.5 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-Márkov 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.

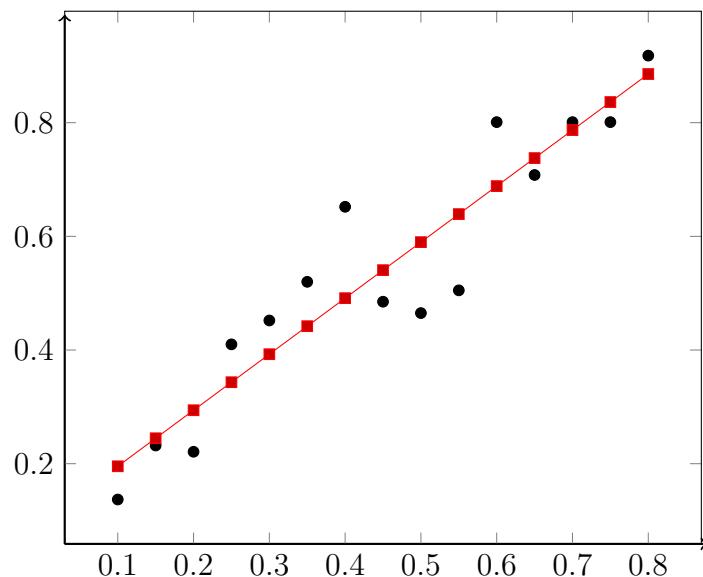


Figura 27: Aproximación a una recta por medio de mínimos cuadrados

Capítulo 7

7.1 Ejercicios

7.1.1 Ejercicio 1

Una máquina que fabrica tornillos se desajusta de forma que produce tornillos de longitud indeterminada. Se sabe que la probabilidad de que un tornillo tenga una longitud menor a L_1 es p_1 y superior a L_2 es p_2 , donde $L_2 > L_1$.

Se extraen cuatro tornillos al azar, calcular:

- Probabilidad de que los cuatro tornillos tengan longitud superior a L_1
- Probabilidad de que tres de ellos tengan longitudes entre L_1 y L_2

- c) Probabilidad de que, a lo sumo, tres sean de longitud inferior a L_2

Las posibles probabilidades asignables a las distintas longitudes se resumen en la siguiente imagen:

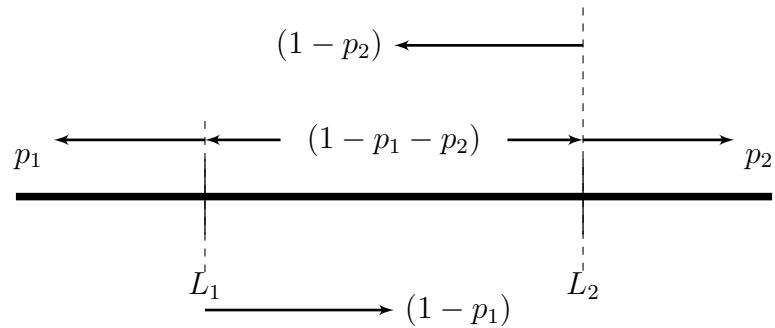


Figura 28: Representación del tornillo

Para el inciso a) se genera el siguiente árbol.

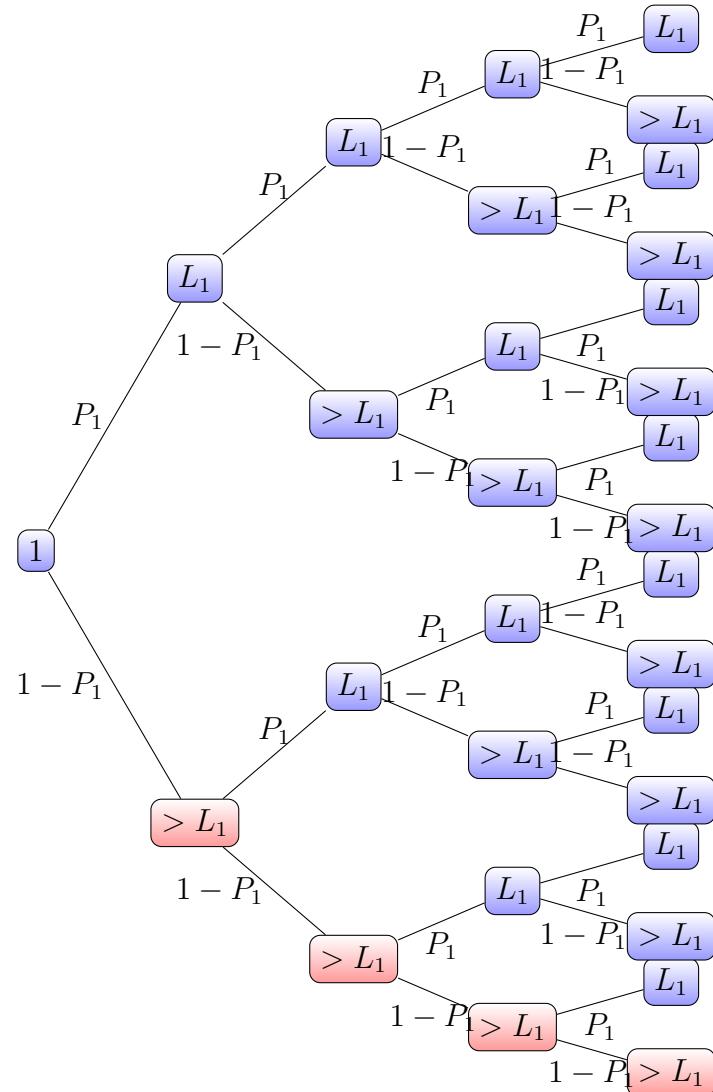


Figura 29: Los 4 tornillos tienen longitudes superiores a L_1

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

$$P(4 \text{torn.sup} L_1) = (1 - p_1)(1 - p_1)(1 - p_1)(1 - p_1) = (1 - p_1)^4$$

En el caso del inciso b), designamos L_3 como la Longitud entre L_1 y L_2 con P_3 que es $1 - P_1 - P_2$: En esta primera rama solo se obtiene $P_1P_3P_3P_3$

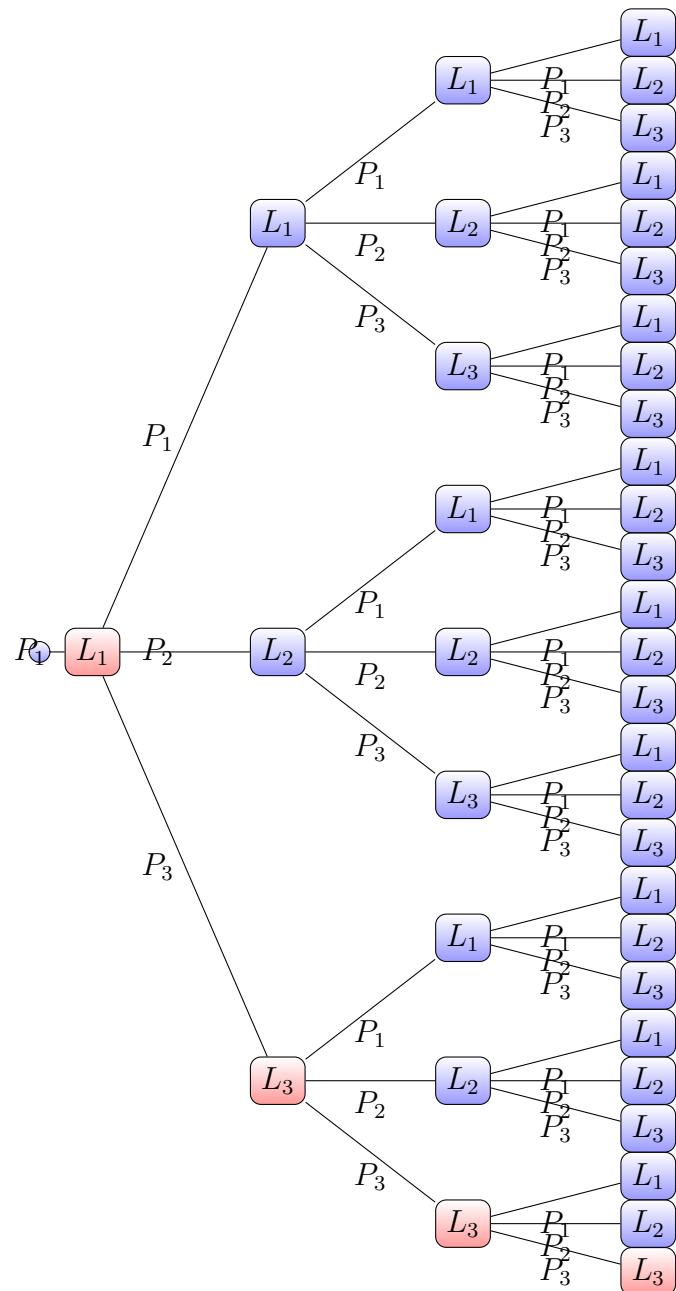
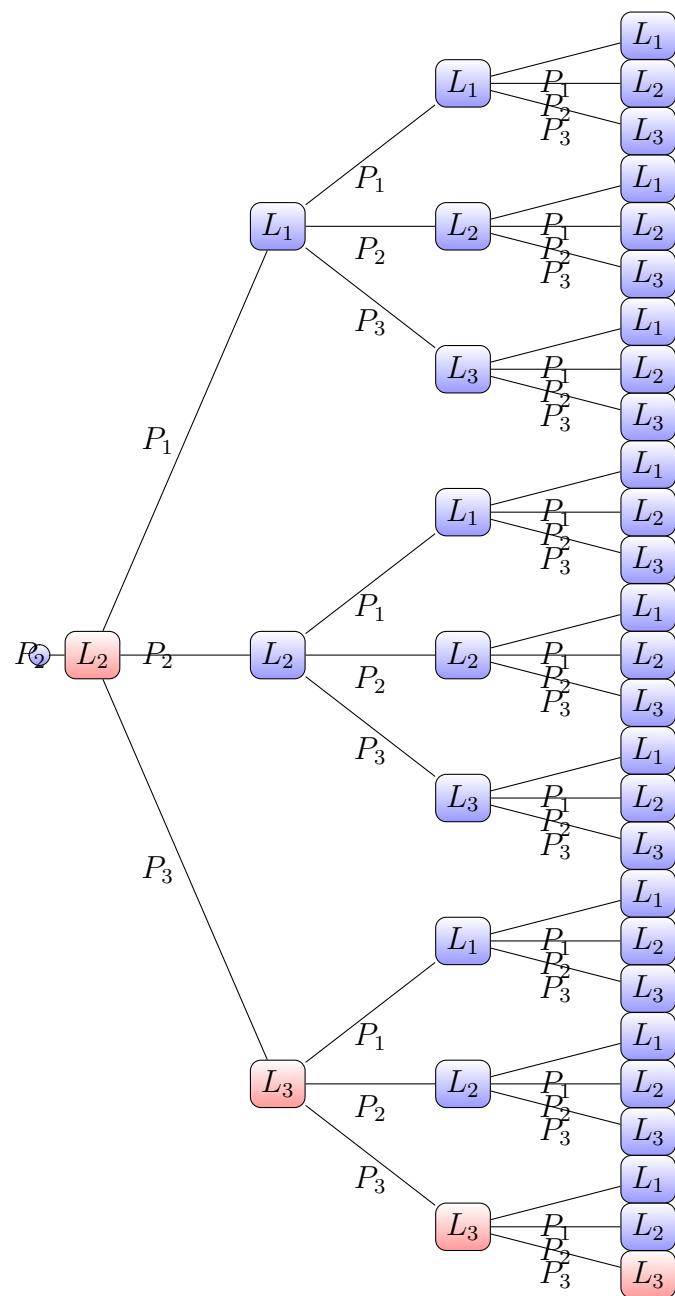


Figura 30: Al menos 3 tornillos estén entre L_1 y L_2

Figura 31: Al menos 3 tornillos esten entre L_1 y L_2

En esta segunda rama solo se obtiene $P_2P_3P_3P_3$

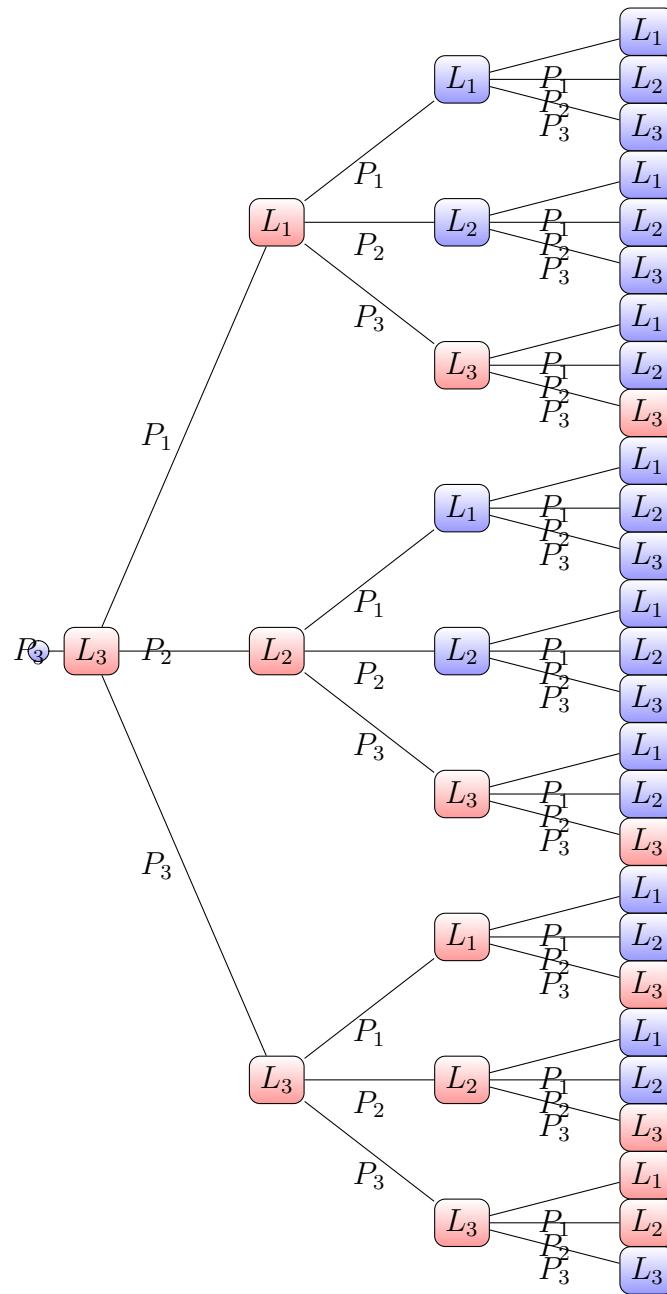


Figura 32: Al menos 3 tornillos estén entre \$L_1\$ y \$L_2\$

En la última rama solo se obtiene: $P_3P_1P_3P_3 + P_3P_2P_3P_3 + P_3P_3P_1P_3 + P_3P_3P_2P_3 + P_3P_3P_3P_1 + P_3P_3P_3P_2$

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 cuatro 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:

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

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

Para el inciso c) se genera el siguiente árbol.

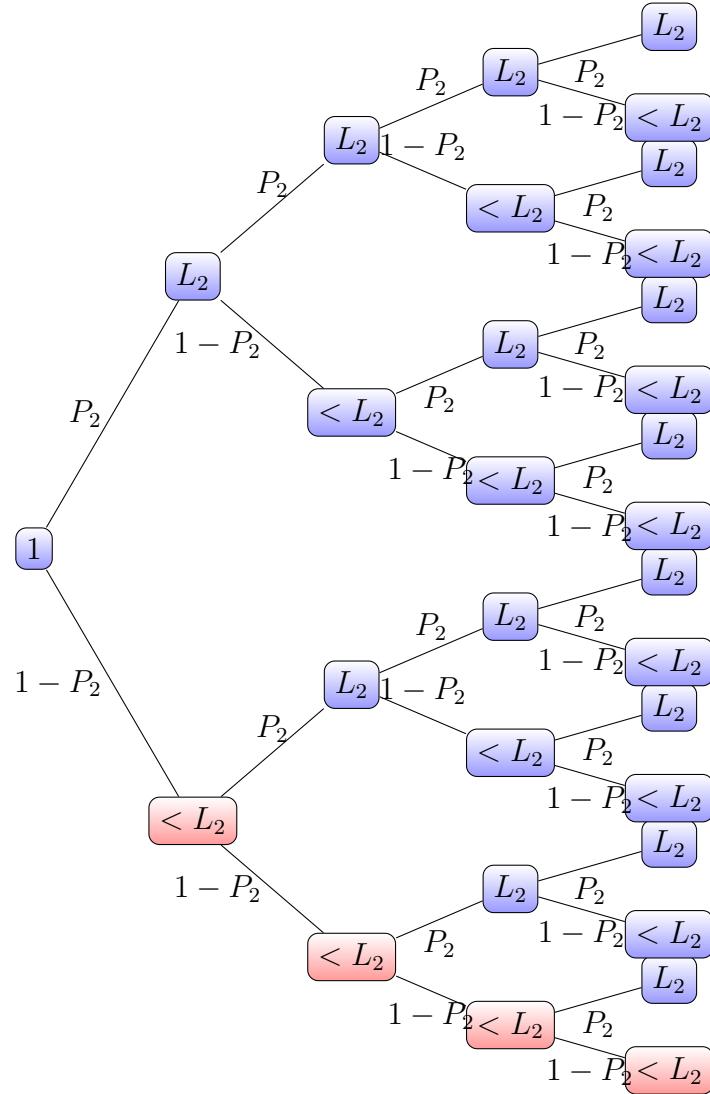


Figura 33: A lo sumo tres tornillos con longitud menor a L_2

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

7.1.2 Ejercicio 2

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 $2/3$?

Para este ejercicio se genera el siguiente árbol, el cual por motivos de espacio se limita a poner los casos en los cuales se cumple la condición los cuales estan en color naranja.

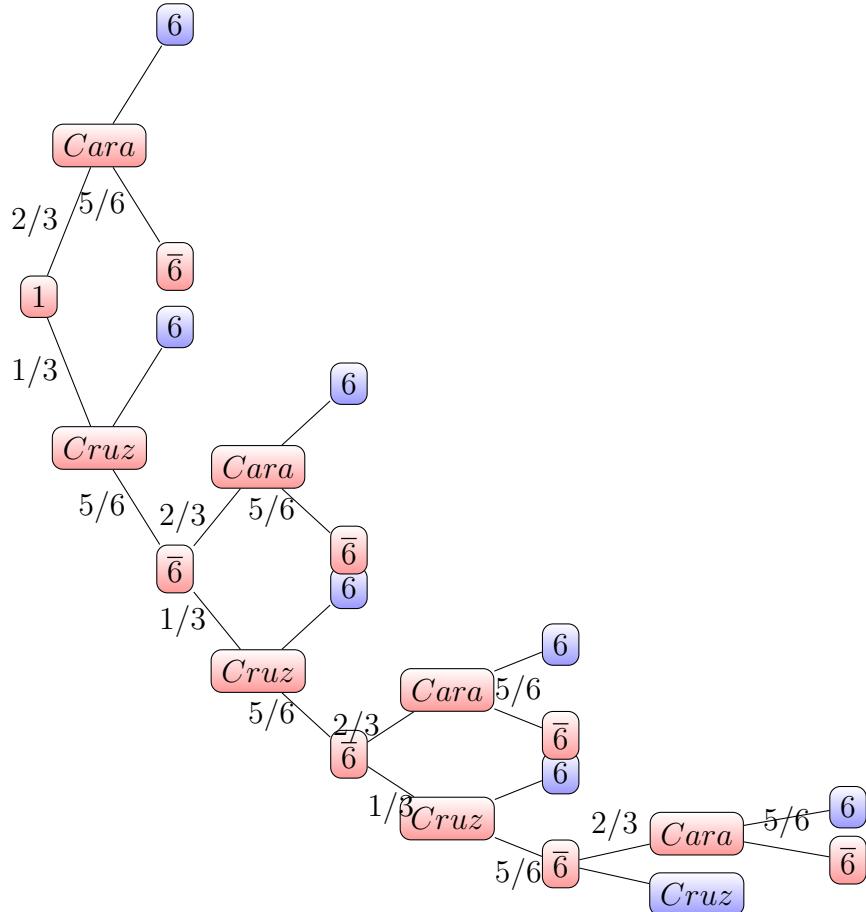


Figura 34: Los 4 tornillos tienen longitudes superiores a L_1

La solución es : $P(S) = P(\text{en 1er experimento}) + P(\text{en 2do experimento}) + P(\text{en 3er experimento}) + P(\text{en 4to experimento})$

$$\begin{aligned}
 P(S) &= \frac{2}{3} \cdot \frac{5}{6} + \left(\frac{1}{3} \cdot \frac{5}{6}\right) \left(\frac{2}{3} \cdot \frac{5}{6}\right) + \left(\frac{1}{3} \cdot \frac{5}{6}\right)^2 \left(\frac{2}{3} \cdot \frac{5}{6}\right) + \left(\frac{1}{3} \cdot \frac{5}{6}\right)^3 \left(\frac{2}{3} \cdot \frac{5}{6}\right) \\
 &= \frac{10}{18} \left[1 + \frac{5}{18} + \frac{5^2}{18} + \frac{5^3}{18}\right] = 0.765
 \end{aligned}$$

7.1.3 Ejercicio 3

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.

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

Solución:

Se definen los sucesos, H : seleccionar una niña, suceso V : seleccionar un niño y suceso M : infante menor de 24 meses.

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

- 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:

$$\begin{aligned} P(M) &= P(H) \cdot P(M | H) + P(V) \cdot P(M | V) \\ &= 0.6 \times 0.2 + 0.4 \times 0.35 \\ &= 0.26 \end{aligned}$$

- 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:

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

7.1.4 Ejercicio 4

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

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

Solución:

Se definen los sucesos F : pacientes que se realizan cirugías faciales, M : pacientes que se realizan implantes mamarios, O : pacientes que se realizan otras cirugías correctivas, H : pacientes de género masculino.

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

$$\begin{aligned} P(H) &= P(F) \cdot P(H | F) + P(M) \cdot P(H | M) + P(O) \cdot P(O | H) \\ &= 0.2 \times 0.25 + 0.35 \times 0.15 + 0.45 \times 0.40 \\ &= 0.28 \end{aligned}$$

- b) Como el suceso condicionado ha ocurrido entonces se aplica el teorema de bayes, luego, el valor de la probabilidad será:

$$\begin{aligned} P(M | H) &= \frac{P(M) \cdot P(H | M)}{P(F) \cdot P(H | F) + P(M) \cdot P(H | M) + P(O) \cdot P(O | H)} \\ &= \frac{0.35 \times 0.15}{0.2 \times 0.25 + 0.35 \times 0.15 + 0.45 \times 0.40} \\ &= 0.19 \end{aligned}$$

7.1.5 Ejercicio 5

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

Solución:

Se definen los sucesos P : seleccionar el primer aparato, S : seleccionar el segundo aparato T : seleccionar el tercer aparato E : seleccionar un resultado con error.

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

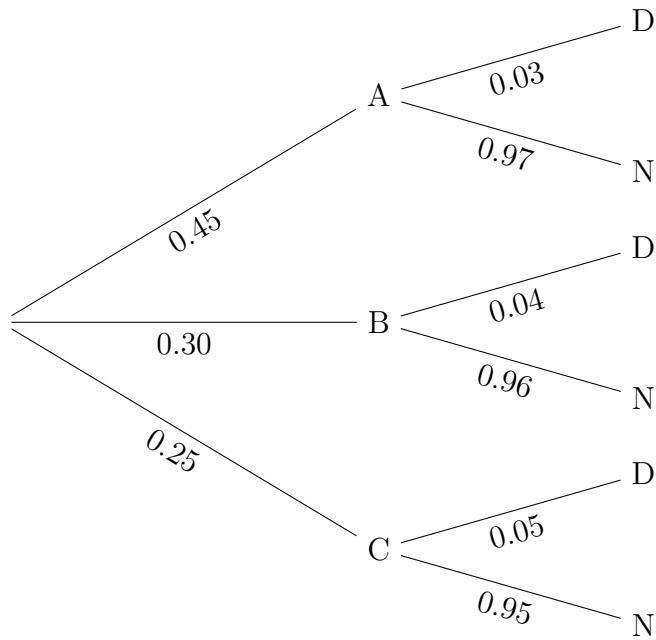
$$\begin{aligned} P(P | E) &= \frac{P(P) \cdot P(E | P)}{P(P) \cdot P(E | P) + P(S) \cdot P(E | S) + P(T) \cdot P(E | T)} \\ &= \frac{0.25 \times 0.01}{0.25 \times 0.01 + 0.35 \times 0.02 + 0.4 \times 0.03} \\ &= 0.116 \end{aligned}$$

7.1.6 Ejercicio 6

Tres máquinas, A , B y C , producen el 45%, 30% y 25%, respectivamente, del total de las piezas producidas en una fábrica. Los porcentajes de producción defectuosa de estas máquinas son del 3%, 4% y 5%.

- a) Seleccionamos una pieza al azar; calcula la probabilidad de que sea defectuosa.
- b) Tomamos, al azar, una pieza y resulta ser defectuosa; calcula la probabilidad de haber sido producida por la máquina B.
- c) ¿Qué máquina tiene la mayor probabilidad de haber producido la citada pieza defectuosa?

Solución: Sea D = "la pieza es defectuosa" y N = "la pieza no es defectuosa". Podemos representar el problema en un diagrama de árbol



- a) Para calcular la probabilidad de que la pieza elegida sea defectuosa, $P(D)$, tomemos en cuenta todos los casos cuando la pieza es defectuosa de todas las máquinas:

$$\begin{aligned}
 P(D) &= P(A) \cdot P(D/A) + P(B) \cdot P(D/B) + P(C) \cdot P(D/C) \\
 &= 0.45 \cdot 0.03 + 0.30 \cdot 0.04 + 0.25 \cdot 0.05 \\
 &= 0.038
 \end{aligned}$$

- b) En este inciso nos están pidiendo que la máquina haya sido B dado que la pieza fue defectuosa ($P(B|D)$), por lo tanto utilizamos el teorema de Bayes:

$$\begin{aligned}
 P(B|D) &= \frac{P(B) \cdot P(D|B)}{P(A) \cdot P(D|A) + P(B) \cdot P(D|B) + P(C) \cdot P(D|C)} \\
 &= \frac{0.30 \cdot 0.04}{0.45 \cdot 0.03 + 0.30 \cdot 0.04 + 0.25 \cdot 0.05} \\
 &= \frac{12}{38} \\
 &= 0.316
 \end{aligned}$$

- c) Calculamos $P(A|D)$ y $P(C|D)$, comparándolas con el valor de $P(B|D)$ ya calculado. Aplicando el teorema de Bayes, obtenemos:

$$\begin{aligned}
 P(A|D) &= \frac{P(A) \cdot P(D|A)}{P(A) \cdot P(D|A) + P(B) \cdot P(D|B) + P(C) \cdot P(D|C)} \\
 &= \frac{0.45 \cdot 0.03}{0.45 \cdot 0.03 + 0.30 \cdot 0.04 + 0.25 \cdot 0.05} \\
 &= \frac{135}{380} \\
 &= 0.355
 \end{aligned}$$

$$\begin{aligned}
 P(C|D) &= \frac{P(C) \cdot P(D|C)}{P(A) \cdot P(D|A) + P(B) \cdot P(D|B) + P(C) \cdot P(D|C)} \\
 &= \frac{0.25 \cdot 0.05}{0.45 \cdot 0.03 + 0.30 \cdot 0.04 + 0.25 \cdot 0.05} \\
 &= \frac{125}{380} \\
 &= 0.329
 \end{aligned}$$

La máquina con mayor probabilidad de haber producido la pieza defectuosa es A .

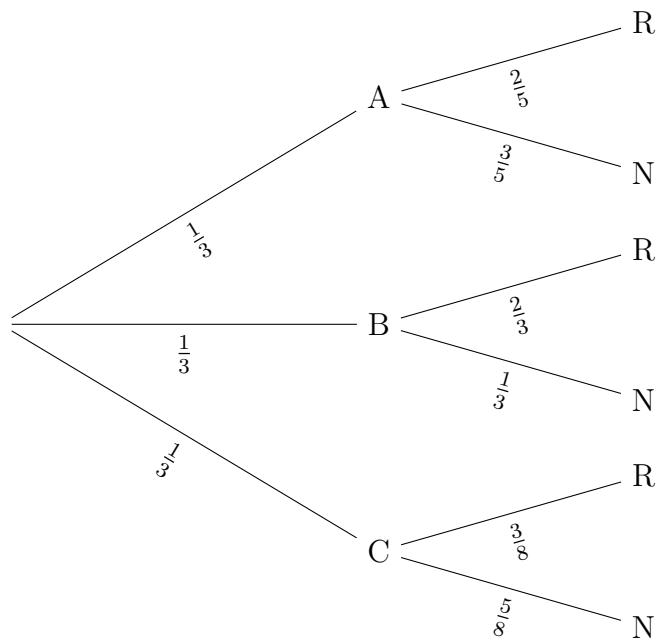
7.1.7 Ejercicio 7

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 ?

Solución:

Llamamos R = "sacar bola roja" y N = "sacar bola negra". Dibujemos un diagrama de árbol para ver las distintas probabilidades de ocurrencia de los sucesos R o N para cada una de las tres urnas, por supuesto la probabilidad de seleccionar una urna es de $\frac{1}{3}$.

La probabilidad pedida es, haber sacado la bola de la urna dado que la bola es roja ($P(A|R)$). Utilizando el teorema de Bayes, tenemos:



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

7.2 Grafos por criterio 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 hacia el nodo $i + 1$, y también desde el nodo $n - 1$ hacia el nodo 0.

Ahora agregaremos flechas azules a cada nodo i hacia el nodo $\text{mod}(\frac{i*10}{n})$. Esto es, desde i hacia 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 hacia 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 hacia 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 hacia el nodo al que nos dirige y así 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 [1].

Por ejemplo si construimos el grafo para el número divisible 5, queda como el de la figúra siguiente:

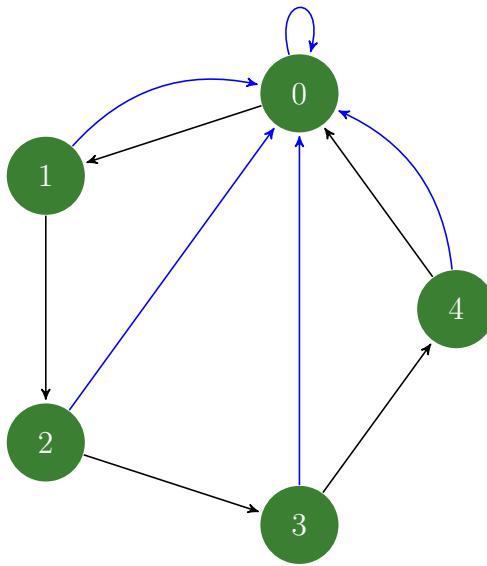


Figura 35: Ejemplo de un grafo por criterio de divisibilidad por 5

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.

7.2.1 Grafo por criterio de divisibilidad por 3

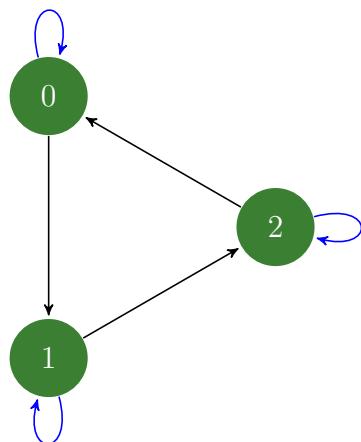


Figura 36: Grafo por criterio de divisibilidad por 3

7.2.2 Grafo por criterio de divisibilidad por 5

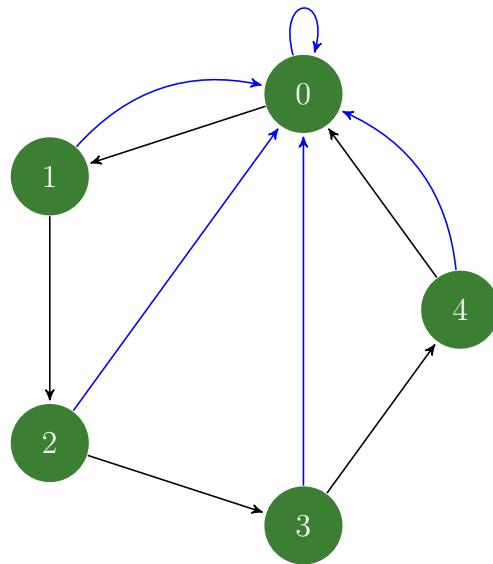


Figura 37: Grafo por criterio de divisibilidad por 5

7.2.3 Grafo por criterio de divisibilidad por 7

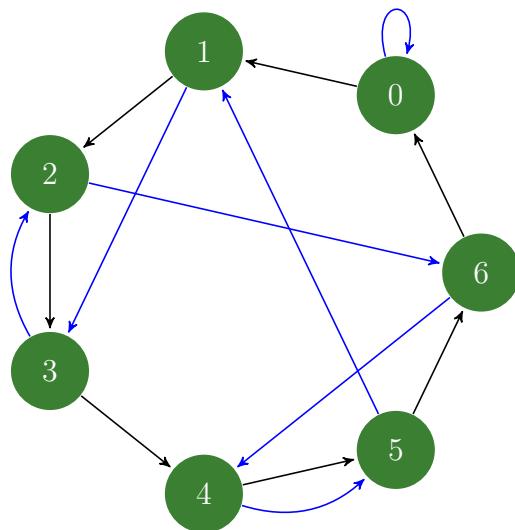


Figura 38: Grafo por criterio de divisibilidad por 7

7.2.4 Grafo por criterio de divisibilidad por 11

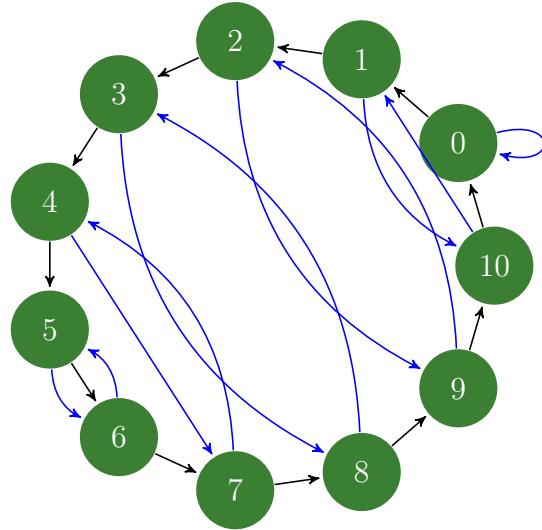


Figura 39: Grafo por criterio de divisibilidad por 11

Capítulo 8

8.1 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 convoyes 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 Vigenére 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.



Figura 40: Máquina Enigma

¿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:

$$\begin{aligned} & 3.283.883.513.796.974.198.700.882.069.882.752.878.379.955.261.095. \\ & 623.685.444.055.315.226.006.433.616.627.409.666.933.182.371.154. \\ & \quad 802.769.920.000.000.000 \end{aligned}$$

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.

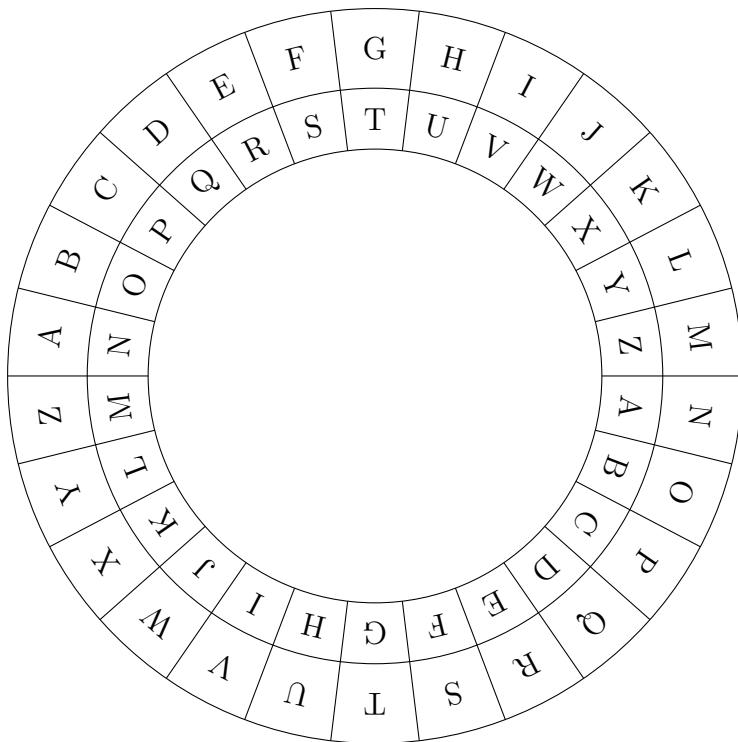


Figura 41: Representación de los rotores de la máquina enigma

¿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.

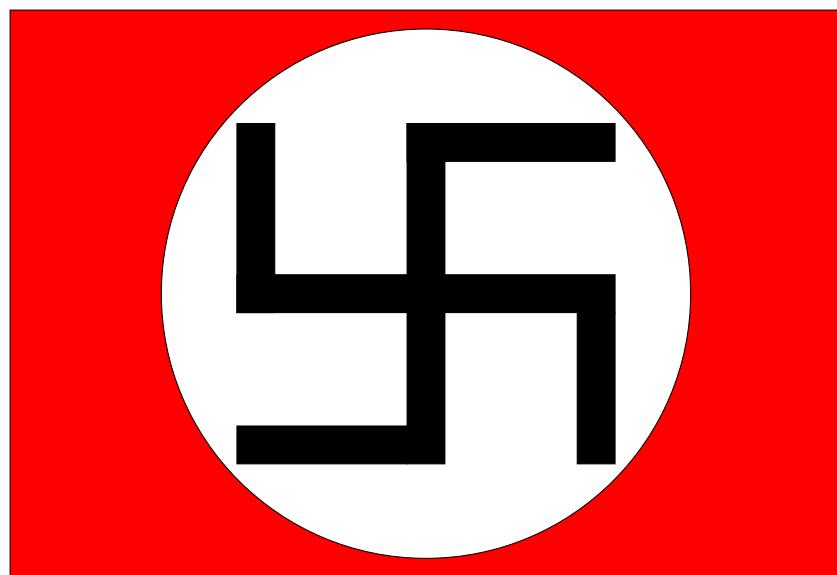


Figura 42: Esvástica Nazi

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.

8.2 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:

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:

$$1782^{12} = 1025397835622633634807550462948226174976$$

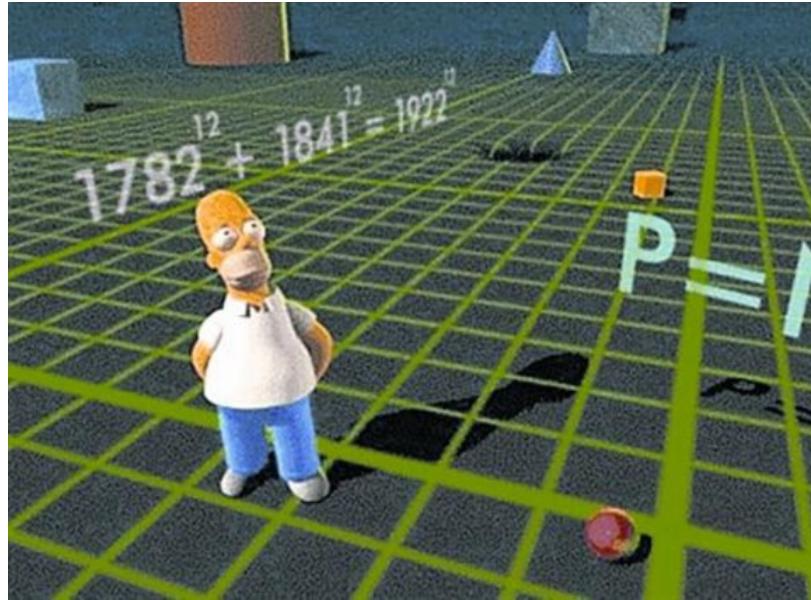


Figura 43: Primero capítulo de los Simpson donde aparece el teorema de Fermat

$$1841^{12} = 1515812422991955541481119495194202351681$$

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

$$1922^{12} = 2541210259314801410819278649643651567616$$

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 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 Simpson:

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

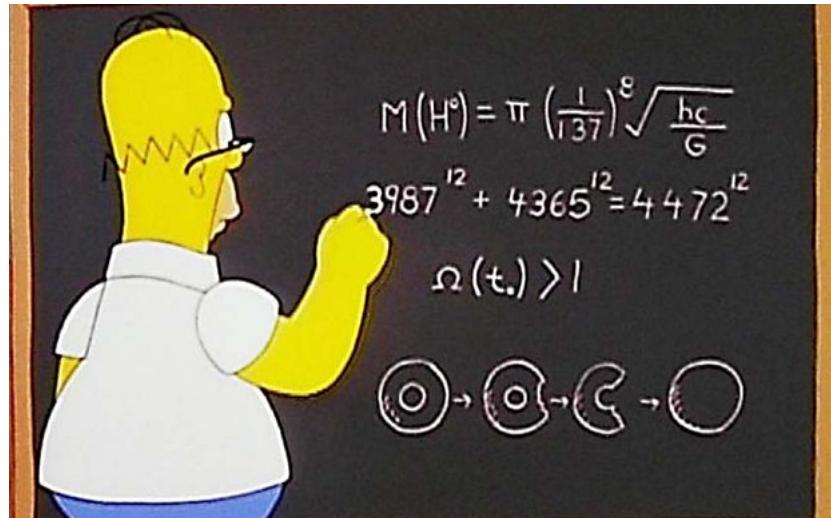


Figura 44: Segundo capítulo de los Simpson donde aparece el teorema de Fermat

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.

8.3 Esperanza matemática

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

$$\mathbb{E}(X) = x_1 p_1 + x_2 p_2 + \cdots + x_i p_i = \sum_{i=1}^n x_i p_i \quad (40)$$

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.

Por ejemplo, el valor esperado cuando tiramos un dado equilibrado de 6 caras es 3,5. Podemos hacer el cálculo

$$\begin{aligned}\mathbb{E}(X) &= 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} \\ &= \frac{1+2+3+4+5+6}{6} = 3.5\end{aligned}$$

y cabe destacar que 3.5 no es un valor posible al rodar el dado. En este caso, en el que todos los sucesos son de igual probabilidad, la esperanza es igual a la media aritmética.

Una aplicación común de la esperanza matemática es en las apuestas o los juegos de azar. Por ejemplo, la ruleta americana tiene 38 casillas equiprobables. La ganancia para acertar una apuesta a un solo número paga de 35 a 1 (es decir, cobramos 35 veces lo que hemos apostado y recuperamos la apuesta, así que recibimos 36 veces lo que hemos apostado). Por tanto, considerando los 38 posibles resultados, la esperanza matemática del beneficio para apostar a un solo número es:

$$(-1 \times \frac{37}{38}) + (35 \times \frac{1}{38})$$

que es -0,0526 aproximadamente. Por lo tanto uno esperaría, en media, perder unos 5 céntimos por cada dólar que apuesta, y el valor esperado para apostar 1 dólar son 0.9474 euros. En el mundo de las apuestas, un juego donde el beneficio esperado es cero (no ganamos ni perdemos) se llama un "juego justo".

Nota: El primer paréntesis es la "esperanza" de perder tu apuesta de \$1, por eso es negativo el valor. El segundo paréntesis es la esperanza matemática de ganar los \$35. La esperanza matemática del beneficio es el valor esperado a ganar menos el valor esperado a perder.

8.4 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:

$$\mathbf{a} \times \mathbf{b} = \mathbf{c}, c_i = \sum_{j,k=1}^3 \epsilon_{ijk} a_j b_k \quad (41)$$

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

- Calculando c_1

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

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

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

$$c_1 = \epsilon_{111} a_1^0 b_1 + \epsilon_{112} a_1^0 b_2 + \epsilon_{113} a_1^0 b_3 + \epsilon_{121} a_2^0 b_1 + \epsilon_{122} a_2^0 b_2 + \epsilon_{123} a_2^1 b_3 + \epsilon_{131} a_3^0 b_1 + \epsilon_{132} a_3^1 b_2 + \epsilon_{133} a_3^0 b_3$$

$$c_1 = a_2 b_3 - a_3 b_2$$

- 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} :

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

$$c_1 = \epsilon_{211} a_1^0 b_1 + \epsilon_{212} a_1^0 b_2 + \epsilon_{213} a_1^- b_3^1 + \epsilon_{221} a_2^0 b_1 + \epsilon_{222} a_2^0 b_2 + \\ \epsilon_{223} a_2^0 b_3 + \epsilon_{231} a_3^1 b_1 + \epsilon_{232} a_3^0 b_2 + \epsilon_{233} a_3^0 b_3$$

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

- 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} :

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

$$c_1 = \epsilon_{311} a_1^0 b_1 + \epsilon_{312} a_1^1 b_2 + \epsilon_{313} a_1^0 b_3 + \epsilon_{321} a_2^- b_1^1 + \epsilon_{322} a_2^0 b_2 + \epsilon_{323} a_2^0 b_3 + \\ \epsilon_{331} a_3^1 b_1 + \epsilon_{332} a_3^0 b_2 + \epsilon_{333} a_3^0 b_3$$

$$c_2 = a_1 b_2 - a_2 b_1$$

8.5 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:

$$\mathbf{u} = u_1x_1 + u_2x_2 + u_3x_3 + \dots + u_nx_n \quad (42)$$

esta puede expresarse convecinalmente como la sumatoria:

$$\mathbf{u} = \sum_{i=1}^n u_i x_i \quad (43)$$

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.

$$\mathbf{u} = u_i x_i \quad (44)$$

8.6 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.

Capítulo 9

9.1 Congruencia Zeller

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

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

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.

9.1.1 Código en gcc

```
//Autor: Itzael Jimenez Aranda
//Instituto: CIC-IPN
//Materia: Probabilidad, procesos aleatorios e inferencia
#include<stdio.h>

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

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

```

    }
else{
    m = mes;
    a = anno;
}
d = dia;
printf("Dia: %i", d);
printf("Mes: %i", m);
printf("Ano: %i", a);
//Calculamos K y J
K = a % 100;
J = a / 100;
printf("K: %i", K);
printf("J: %i", J);

//Aplicamos la formula
h = d+((m+1)*26)/10)+K+(K/4)+(J/4)-(2*J);
printf("\n%i\n", h);
h = h%7;
printf("\n%i\n", h);
//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{
        do{
            printf("Ingrese el dia: ");
            scanf("%i", &dia);
            printf("Ingrese el mes: ");
            scanf("%i", &mes);
            if(dia < 1)
                printf("Error, dia 0 no existe\n");
            if(mes < 1 || mes > 12)
                printf("Error, rango de mes: 1-12\n");
            if(mes > 0 && mes < 13){
                if(mes != 2){
                    if(dia > meses[mes]){
                        printf("Error, en el mes seleccionado,");
                        printf("no existe tal dia.\n");
                    }
                }
            }
        }
    }
}

```

```

        aux = 0;
    }
    else{
        aux = 1;
    }
}
else{
    if (dia > 29){
        printf("Error , _en _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 _bisiesto \n" );
        aux = 0;
    }
    if (ano%4 == 0 && ano%100 == 0 && ano%400 != 0){
        printf(" El _ano _seleccionado _no _es _bisiesto \n" );
        aux = 0;
    }
}
}while (aux == 0);
int z = zeller(dia , mes , ano);
char *dia_sem;
switch (z)
{
    case 1:
        dia_sem = " Domingo" ;
        break;
    case 2:
        dia_sem = " Lunes" ;
        break;
    case 3:
        dia_sem = " Martes" ;
        break;
    case 4:
        dia_sem = " Miercoles" ;
}

```

```

        break;
    case 5:
        dia_sem = "Jueves";
        break;
    case 6:
        dia_sem = "Viernes";
        break;
    case 0:
        dia_sem = "Sabado";
        break;
    }
}

printf ("\nLa_fecha_corresponde_con_un_%s\n", dia_sem);

return 0;
}

```

9.1.2 Código en Mathematica

```

dia = 29;
mes = 2;
anio = 2008;
If[Mod[anio, 4] == 0 && dia == 29 && mes == 2,
  If[Mod[anio, 100]==0, If[Mod[anio,400]==0, bis=0, bis=1],
    bis = 0], bis = 0];

If[bis == 0,
  If[mes == 1, {m = 13; a = anio - 1;},
    If[mes==2,{m=14;a=anio-1;},{m=mes;a=anio;}]]
  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 = "anno_no_bisiesto"];
  StringForm["El ``de ``de ``fue ``", dia, mes, anio, d]

```

9.2 Proyecto.- Lanzar 2 dados hasta obtener un doble 6

9.2.1 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.

9.2.2 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 [4].

9.2.3 Análisis Matemático.

Lanzamiento de un dado

Para el caso de un dado tenemos la Eq.(45).

$$P(as) = 1 - \frac{5^n}{6} \quad (45)$$

Donde n es el número de lanzamientos realizados. Para $n = 4$ tenemos la Eq.(46).

$$P(as) = 1 - \frac{5^4}{6} = \frac{671}{1296} \approx 0.517 \quad (46)$$

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

Lanzamiento de dos dados.

Para el caso de dos dados tenemos la Eq.(47).

$$P(doble\ as) = 1 - \frac{35^n}{36} \quad (47)$$

Donde n es el número de lanzamientos realizados. Para $n = 24$ tenemos la Eq.(48).

$$P(doble\ as) = 1 - \frac{35^{24}}{36} \approx 0.491 \quad (48)$$

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.(49).

$$P(doble\ as) = 1 - \frac{35^{25}}{36} \approx 0.505 \quad (49)$$

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.

9.2.4 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.

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.

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.

Segundo Principio: Hacer un plan.

Mientras se resuelven más 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 ordenada de los pasos a seguir.
- Eliminar diferentes posibilidades.
- Considerar casos especiales.

- Usar razonamiento directo.
- Usar una fórmula.

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.

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 [9].

9.2.5 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.

9.2.6 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.

9.2.7 Poner el plan en marcha.

9.2.8 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								
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.

9.2.9 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								
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.

9.2.10 Simulación en Mathematica.

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

Código en Mathematica

```
k = 0;
limite = 10;
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]
Export["simulacion_doble_6_1.eps", Histogram[c, limite]]
d = StandardDeviation[c]
e = Mean[c]
Plot[PDF[NormalDistribution[e, d], c], {c, -100, 200}]
Export["simulacion_doble_6_2.eps",
  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.(45,46).

Valor promedio: 38.4

Desviación estándar: 36.936

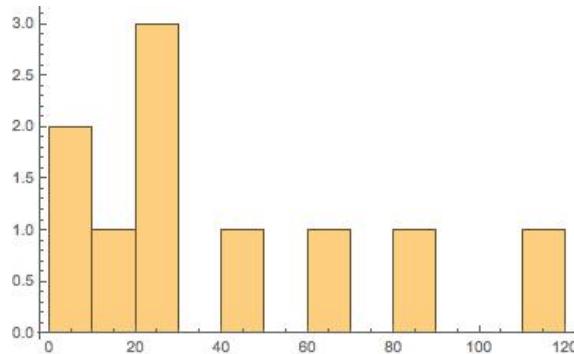


Figura 45: Resultado de 10 lanzamientos.

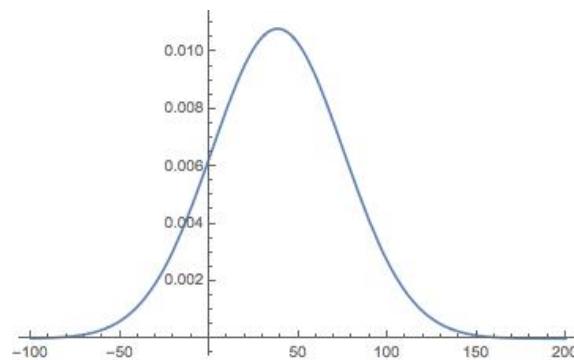


Figura 46: Distribución normal de 10 lanzamientos.

- Para 100 repeticiones del experimento, fig.(47,48).

Valor promedio: 35.46

Desviación estándar: 30.6566

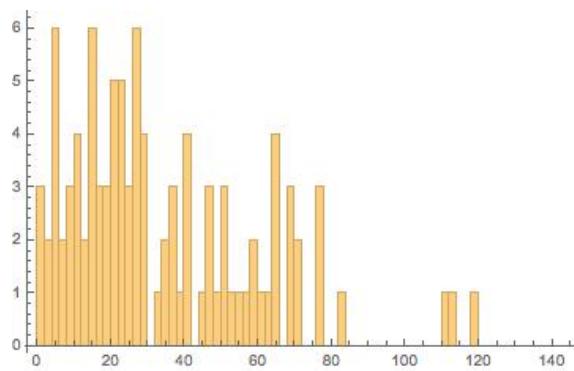


Figura 47: Resultado de 100 lanzamientos.

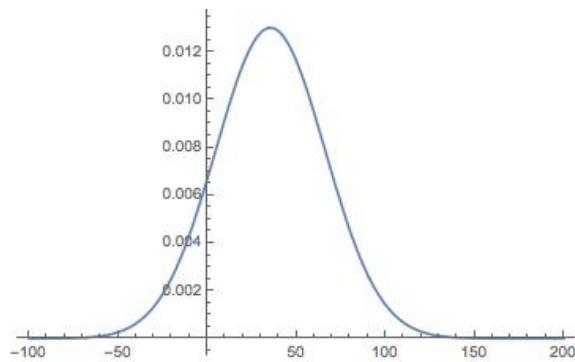


Figura 48: Distribución normal de 100 lanzamientos.

- Para 1,000 repeticiones del experimento, fig.(49,50).
Valor promedio: 34.627
Desviación estándar: 35.1388

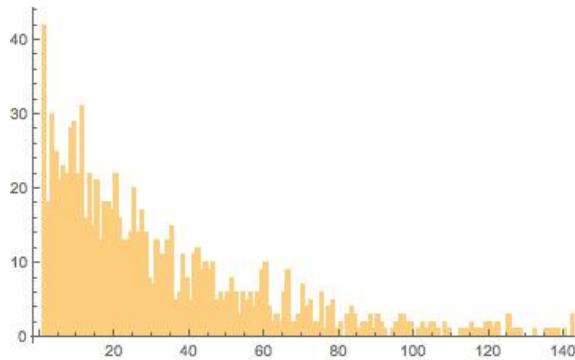


Figura 49: Resultado de 1,000 lanzamientos.

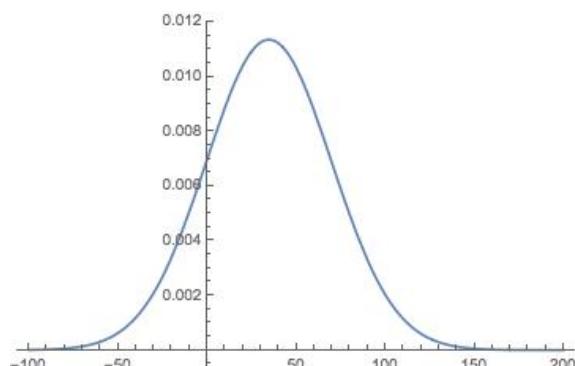


Figura 50: Distribución normal de 1,000 lanzamientos.

- Para 10,000 repeticiones del experimento, fig.(51,52).
Valor promedio: 35.7256
Desviación estándar: 34.5559

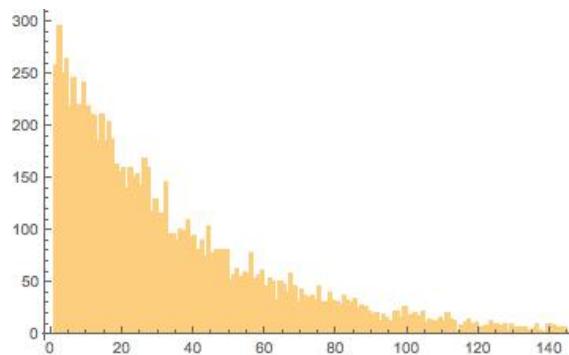


Figura 51: Resultado de 10,000 lanzamientos.

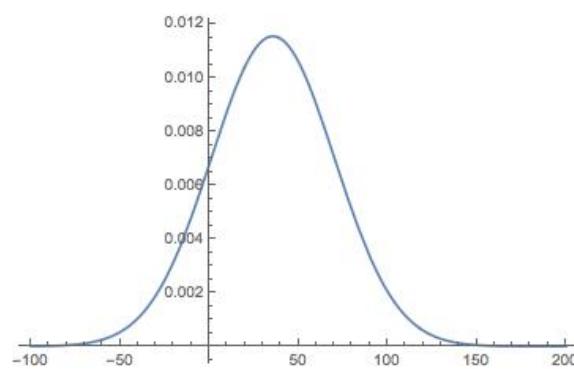


Figura 52: Distribución normal de 10,000 lanzamientos.

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.

9.2.11 Simulación en ROOT.

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

Código en ROOT

```
//Autor: Itzael Jimenez Aranda
//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
                            , 20)
```

```

        ,10
        ,700
        ,700);
c1->SetGrid();

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

TH1I* hdado= new TH1I("Datos"
                      , "Lanzamientos para obtener un doble 6"
                      ,100
                      ,0
                      ,50);
hdado->SetLineColor(209);
hdado->SetLineWidth(4);
hdado->GetXaxis()->SetTitle("Lanzamientos");
hdado->GetYaxis()->SetTitle("Frecuencia");

int dado1 = 0;
int dado2 = 0;
int conteo = 0;
int repeticiones = 10000000;
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(209);
dis->SetLineWidth(4);
TString s;
s = "Distribucion_de_los_lanzamientos";
s += "Lanzamientos;Cara_dado";
dis->SetTitle(s);
dis->Draw();
}

```

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

- Para 10 repeticiones del experimento, fig.(53, 54).

Valor promedio: 17.67

Desviación estándar: 9.695

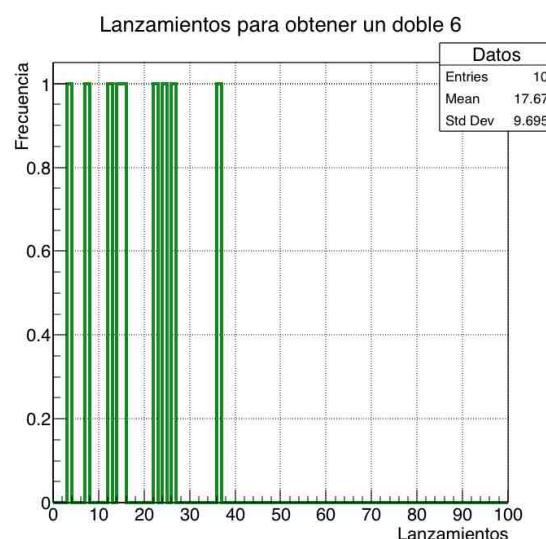


Figura 53: Resultado de 10 lanzamientos.

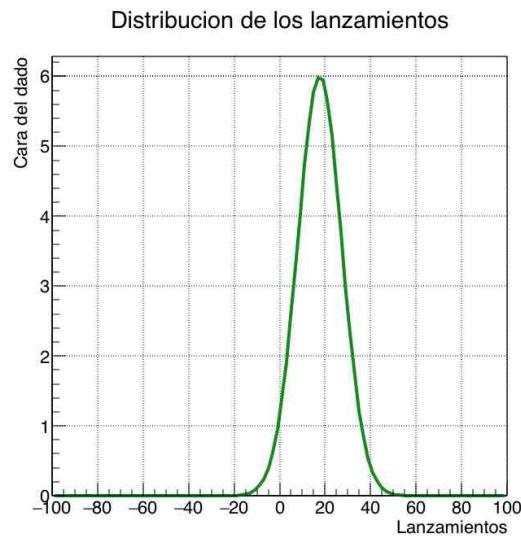


Figura 54: Distribución normal de 10 lanzamientos.

- Para 100 repeticiones del experimento, fig.(55, 56).

Valor promedio: 25.96

Desviación estándar: 23.395

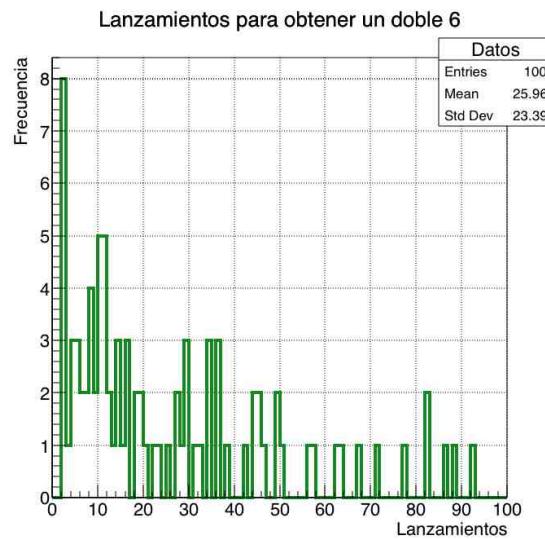


Figura 55: Resultado de 100 lanzamientos.

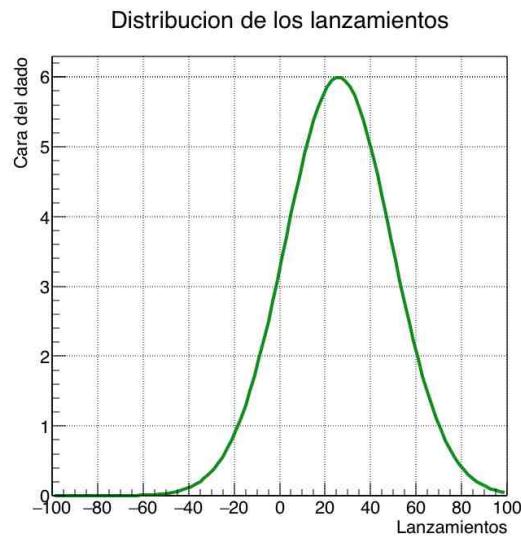


Figura 56: Distribución normal de 100 lanzamientos.

- Para 1,000 repeticiones del experimento, fig.(57, 58).

Valor promedio: 29.96

Desviación estándar: 24.89

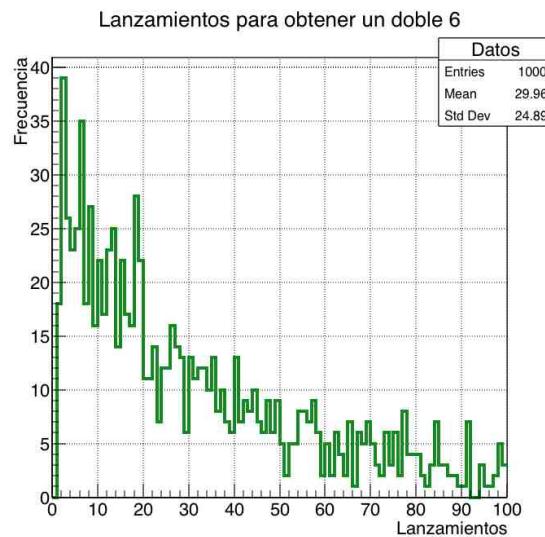


Figura 57: Resultado de 1,000 lanzamientos.

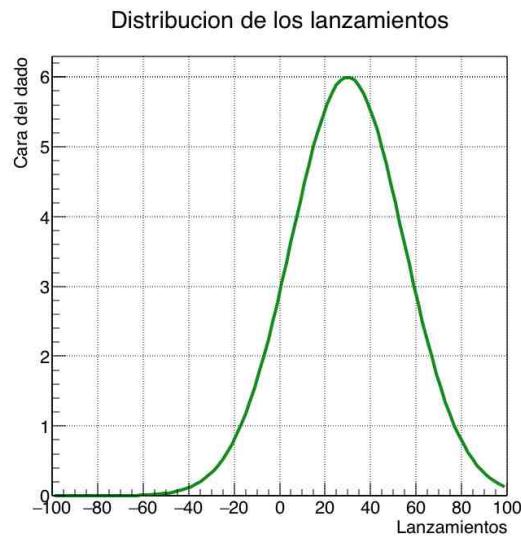


Figura 58: Distribución normal de 1,000 lanzamientos.

- Para 1,000,000 repeticiones del experimento, fig.(59, 60).

Valor promedio: 29.5

Desviación estándar: 24

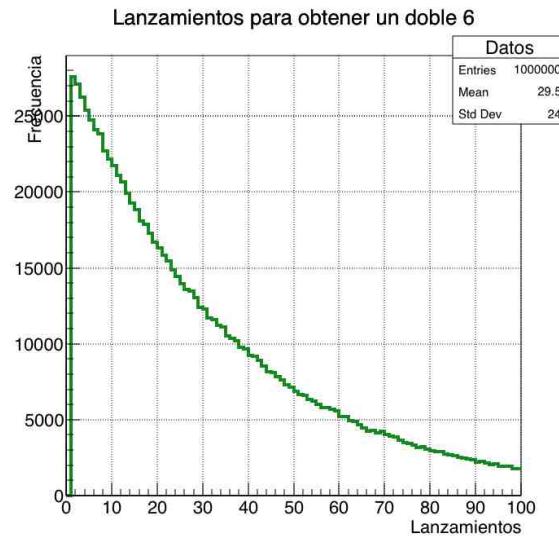


Figura 59: Resultado de 1,000,000 lanzamientos.

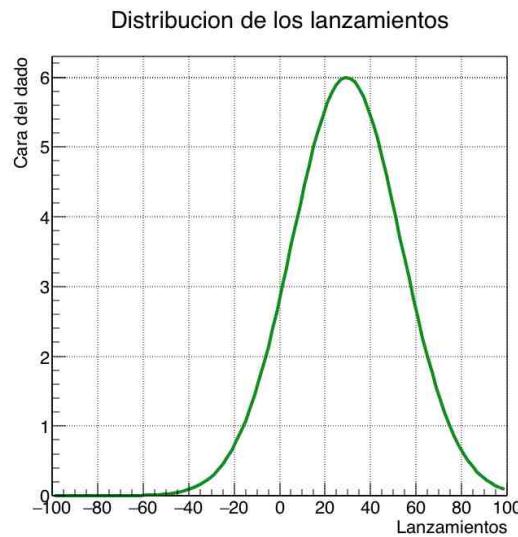


Figura 60: Distribución normal de 1,000,000 lanzamientos.

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.

9.2.12 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.

9.2.13 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 más grandes, para lograr obtener un resultado más aproximado al que se espera de manera teórica.

9.3 Examen propuesto

Ejercicio 1.- Tome en cuenta un experimento donde se lanzan 2 dados, determine:

- a) La cantidad de eventos del espacio muestral
- b) La cantidad de casos cuando la suma al lanzar los dos dados sea un múltiplo de 2
- c) La probabilidad de que la suma sea múltiplo de 3
- d) La suma más probable
- e) La probabilidad de la suma más probable
- f) La sumas menos probables
- g) La probabilidad de las sumas menos probable

Solución:

+	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

La tabla anterior muestra los posibles resultados de la suma al lanzar los dos dados, con esta podemos apoyarnos para resolver los incisos del problema.

a) Tenemos 36 eventos en nuestro espacio muestral, puesto que se tienen 6 posibilidades al tirar el primer dado y luego 6 más del segundo dado por cada posibilidad del primer dado. Generalizando: No importa la cantidad de dados podemos obtener la cantidad de eventos de nuestro espacio muestral con: 6^n , tomando en cuenta que los dados son de 6 caras.

$$\Omega = 36$$

b) Los casos cuando la suma es un múltiplo de 2 podemos observarlos de color rojo en la tabla.

$$A = 18$$

c) Para obtener la probabilidad de que la suma sea un múltiplo de 3 tenemos que considerar los casos favorables y los casos totales y con esto se obtiene la probabilidad. Los casos favorables podemos contarlos en la tabla, que serían los casos donde la suma da 3,6,9 y 12.

$$P(B) = \frac{n(B)}{n(\Omega)} = \frac{12}{36} = \frac{1}{3}$$

d) La suma más probable la encontramos en el número 7 puesto que podemos observar en la tabla que es el número que aparece más veces, esto se debe a que hay más combinaciones en los resultados de los dados para que nos de la suma de 7.

$$\text{Suma más probable} = 7$$

e) Al igual que el inciso c consideraremos los casos favorables y los casos totales. En este caso el número de casos favorables son 6.

$$P(\text{Suma más probable}) = \frac{n(\text{Suma más probable})}{n(\Omega)} = \frac{6}{36} = \frac{1}{6}$$

f) La suma menos probable la encontramos en los números 2 y 12 porque sólo hay una manera en cada caso en de que los dados nos puedan dar esas sumas.

$$\text{Suma menos probable} = 2$$

g) Al igual que el inciso c consideraremos los casos favorables y los casos totales. En este caso el número de casos favorables son 2.

$$P(\text{Suma menos probable}) = \frac{n(\text{Suma menos probable})}{n(\Omega)} = \frac{2}{36} = \frac{1}{18}$$

Ejercicio 2.- Una urna contiene 6 bolas azules y 4 blancas. Se extraen 2 bolas sucesivamente.

- a) Calcular la probabilidad de que la primera sea blanca y la segunda azul habiendo reposición
- b) Calcular la probabilidad de que la primera sea blanca y la segunda azul no habiendo reposición

Solución:

a) Si hay reposición entonces cuando tomamos la primer bola se tiene un número de eventos totales, pero como esta bola se vuelve a regresar a la urna la segunda bola tendrá el mismo número de eventos totales. En este caso tenemos sucesos independientes.

$$P(\text{blanca} \cap \text{azul}) = P(\text{blanca})P(\text{azul})$$

$$P(\text{blanca}) = \frac{n(\text{blanca})}{n(\Omega)} = \frac{4}{10}$$

$$P(\text{azul}) = \frac{n(\text{azul})}{n(\Omega)} = \frac{6}{10}$$

$$P(\text{blanca} \cap \text{azul}) = P(\text{blanca})P(\text{azul}) = \frac{4}{10} \cdot \frac{6}{10} = \frac{24}{100} = \frac{6}{25}$$

a) Si no hay reposición entonces cuando tomamos la primer bola se tiene un número de eventos totales, pero cuando se toma la segunda bola esta no tendrá el mismo número de eventos totales, puesto que el espacio muestral tendrá una bola menos. En este caso tenemos sucesos dependientes.

$$P(\text{blanca} \cap \text{azul}) = P(\text{blanca})P(\text{azul}|\text{blanca})$$

$$P(\text{blanca}) = \frac{n(\text{blanca})}{n(\Omega)} = \frac{4}{10}$$

$$P(\text{azul}|\text{blanca}) = \frac{n(\text{azul})}{n(\Omega(\text{actual}))} = \frac{6}{9}$$

$$P(\text{blanca} \cap \text{azul}) = P(\text{blanca})P(\text{azul}|\text{blanca}) = \frac{4}{10} \cdot \frac{6}{9} = \frac{24}{90} = \frac{4}{15}$$

Ejercicio 3.- A partir de 5 matemáticos y 7 físicos hay que constituir una comisión de 2 matemáticos y 3 físicos. ¿De cuántas formas podrá hacerse?, si:

- a) Todos son elegibles
- b) Un físico en particular ha de estar en esa comisión
- c) Dos matemáticos concretos no pueden estar juntos

Solución:

- a) Puesto que todos son elegibles y no importa el orden en este caso entonces son combinaciones.

Primero obtenemos las combinaciones del grupo de matemáticos, se tienen 5 matemáticos y se necesitan dos:

$$C_{5,2} = 10$$

Ahora las combinaciones del grupo de físicos, se tiene 7 físicos y se necesitan 3

$$C_{7,3} = 35$$

Por lo tanto se pueden formar $10 \times 35 = 350$ comisiones posibles.

- b) Si se fija a alguno de los físicos se tendría en ese grupo nada más 6 físicos y la selección de dos de ellos. El grupo de matemáticos permanece igual que el inciso anterior. Por lo tanto:

$$C_{6,2} = 15$$

Por lo tanto se pueden formar $10 \times 15 = 150$ comisiones posibles.

- c) Como necesitamos seleccionar a dos matemáticos y hay un par de ellos que no pueden estar juntos, o sea tenemos sólo un grupo que no es posible, puesto que

cualquiera de estos dos matemáticos por separado puede estar con todos los demás sin problema. Entonces obtenemos el número de grupos pares que se pueden tener de los 5 matemáticos quitandole el grupo que no es posible formar.

$$C_{5,2} = 10 - 1 = 9$$

Las combinaciones del grupo de físicos permanece igual (35 grupos de 3 físicos), porlo que el total de comisiones que pueden formarse es $9 \times 35 = 315$

Ejercicio 4.- Sean A_1 , A_2 y A_3 sucesos tales que $A_1 \cup A_2 \cup A_3 = \Omega$ y $A_1 \cap A_2 = A_1 \cap A_3 = A_2 \cap A_3$. Sabiendo que $P(A_1) = \frac{1}{4}$ y $P(A_2) = \frac{1}{2}$ hallar $P(A_3)$.

Solución:

Sabemos que $P(\Omega) = 1$, así desarrollando esta expresión y sustituyendo los datos que nos da el problema, se tiene que:

$$P(\Omega) = 1 = P(A_1 \cup A_2 \cup A_3)$$

Donde:

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) &= P(A_1) + P(A_2) + P(A_3) \\ &\quad - P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) \\ &\quad + P(A_1 \cap A_2 \cap A_3) \end{aligned}$$

Sabemos que $A_1 \cap A_2 = A_1 \cap A_3 = A_2 \cap A_3$ por lo tanto tenemos:

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) &= P(A_1) + P(A_2) + P(A_3) \\ &\quad - 3P(A_1 \cap A_2) + P(A_1 \cap A_2 \cap A_3) \end{aligned}$$

$P(A_1 \cap A_2 \cap A_3)$ equivale a:

$$P(A_1 \cap (A_2 \cap A_3)) = P(A_1 \cap (A_1 \cap A_2)) = P(A_1 \cap A_2)$$

Sustituyendo:

$$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - 3P(A_1 \cap A_2) + P(A_1 \cap A_2)$$

$$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - 2P(A_1 \cap A_2)$$

Por ultimo sustituyendo valores dados por el problema:

$$P(A_1 \cup A_2 \cup A_3) = \frac{1}{4} + \frac{1}{2} + P(A_3) - 2P(A_1 \cap A_2)$$

$$\begin{aligned}
 P(\Omega) &= \frac{1}{4} + \frac{1}{2} + P(A_3) - 2P(A_1 \cap A_2) \\
 1 &= \frac{1}{4} + \frac{1}{2} + P(A_3) - 2P(A_1 \cap A_2) \\
 P(A_3) &= -\frac{1}{4} - \frac{1}{2} + 1 + 2P(A_1 \cap A_2) \\
 P(A_3) &= 2P(A_1 \cap A_2) + \frac{1}{4}
 \end{aligned}$$

Ejercicio 5.- Siendo A_1 , A_2 y B tres sucesos tales que $A_1 \cap A_2 = \emptyset$, demostrar que $P(A_1 \cup A_2|B) = P(A_1|B) + P(A_2|B)$.

Solución:

Dado que sabemos que $P(S_1|S_2) = \frac{P(S_1 \cap S_2)}{P(S_2)}$ continuamos a aplicar esta regla a la ecuación propuesta.

$$\begin{aligned}
 P(A_1 \cup A_2|B) &= \frac{P(A_1 \cup (A_2 \cap B))}{P(B)} \\
 &= \frac{P((A_1 \cup A_2) \cap B)}{P(B)} \\
 &= \frac{P((A_1 \cap B) \cup (A_2 \cap B))}{P(B)}
 \end{aligned}$$

Sabemos que $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, por lo tanto:

$$\frac{P((A_1 \cap B) \cup (A_2 \cap B))}{P(B)} = \frac{P(A_1 \cap B) + P(A_2 \cap B) - P(A_1 \cap A_2 \cap B)}{P(B)}$$

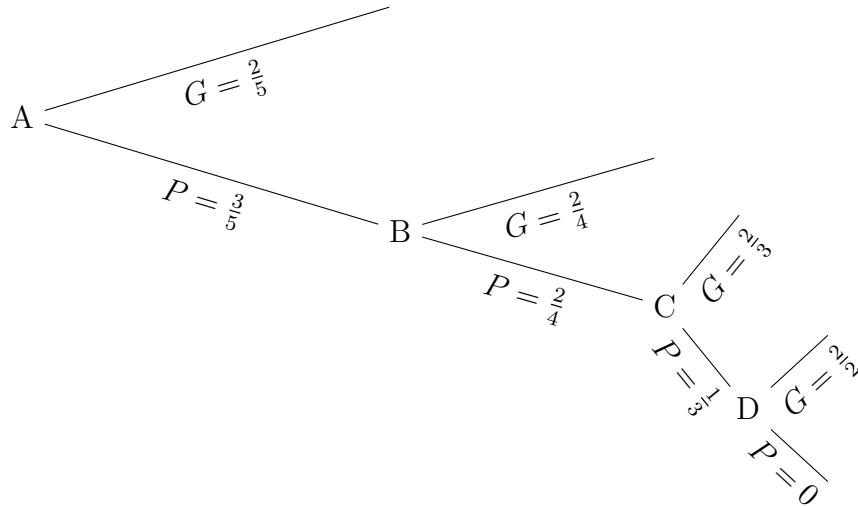
Donde $P(A_1 \cap A_2) = \emptyset$, sustituyendo:

$$\begin{aligned}
 \frac{P(A_1 \cap B) + P(A_2 \cap B) - P(A_1 \cap A_2 \cap B)}{P(B)} &= \frac{P(A_1 \cap B) + P(A_2 \cap B) - P(\emptyset \cap B)}{P(B)} \\
 &= \frac{P(A_1 \cap B) + P(A_2 \cap B)}{P(B)} \\
 &= \frac{P(A_1 \cap B)}{P(B)} + \frac{P(A_2 \cap B)}{P(B)} \\
 &= P(A_1|B) + P(A_2|B)
 \end{aligned}$$

Ejercicio 6.- Una bolsa contiene dos bolas blancas y tres bolas negras. Cada una de cuatro personas, A , B , C y D , en ese orden, saca una bola y no la repone. El primero que saque una bola blanca gana. Determinar las probabilidades de ganar de A , B , C y D .

Solución:

Podemos generar un árbol para poder observar con mayor detenimiento el problema.



Al inicio A tiene la probabilidad de ganar (tomar una bola blanca) de $\frac{2}{5}$, puesto que el total de bolas es de 5 y blancas hay 2 en ese momento.

$$A = \frac{2}{5}$$

La probabilidad de que B gane está relacionada con que A pierda, si observamos en el árbol la probabilidad de que A pierda es de $\frac{3}{5}$, además debemos de tomar en cuenta que la bola que tomó A ya no está en la bolsa, y por ende el total de bolas se reduce a 1, y también como A perdió la bola que no está es la negra, por lo tanto:

$$B = P(A_{Pierda}) \cdot P(B_{Gane})$$

$$B = \frac{3}{5} \cdot \frac{2}{4} = \frac{3}{10}$$

De igual forma la probabilidad de que C gane está relacionada con que A y B pierdan y tomando en cuenta que ya hay 2 bolas menos en la bolsa y que son negras. Por lo tanto:

$$C = P(A_{Pierda}) \cdot P(B_{Pierda}) \cdot P(C_{Gane})$$

$$C = \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{2}{3} = \frac{1}{5}$$

Por ultimo si A , B y C sacaron una bola negra, entonces ya no hay más bolas negras en la bolsa y terminan quedando sólo las dos blancas. Entonces para D :

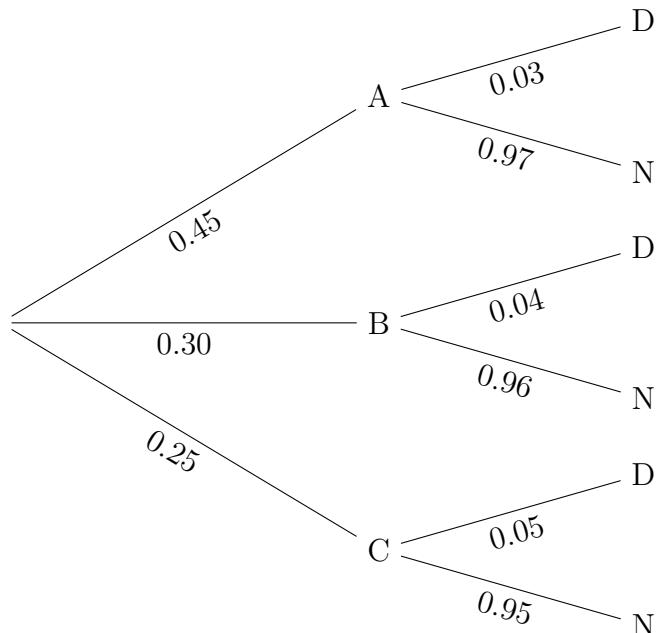
$$C = P(A_{Pierda}) \cdot P(B_{Pierda}) \cdot P(C_{Pierda}) \cdot P(D_{Gane})$$

$$C = \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{1}{3} \cdot \frac{2}{2} = \frac{1}{10}$$

Ejercicio 7.- Tres máquinas, A , B y C , producen el 45%, 30% y 25%, respectivamente, del total de las piezas producidas en una fábrica. Los porcentajes de producción defectuosa de estas máquinas son del 3%, 4% y 5%.

- Seleccionamos una pieza al azar; calcula la probabilidad de que sea defectuosa.
- Tomamos, al azar, una pieza y resulta ser defectuosa; calcula la probabilidad de haber sido producida por la máquina B .
- ¿Qué máquina tiene la mayor probabilidad de haber producido la citada pieza defectuosa?

Solución: Sea D = "la pieza es defectuosa" y N = "la pieza no es defectuosa". Podemos representar el problema en un diagrama de árbol



- Para calcular la probabilidad de que la pieza elegida sea defectuosa, $P(D)$, tomemos en cuenta todos los casos cuando la pieza es defectuosa de todas las máquinas:

$$\begin{aligned}
 P(D) &= P(A) \cdot P(D|A) + P(B) \cdot P(D|B) + P(C) \cdot P(D|C) \\
 &= 0.45 \cdot 0.03 + 0.30 \cdot 0.04 + 0.25 \cdot 0.05 \\
 &= 0.038
 \end{aligned}$$

- En este inciso nos están pidiendo que la máquina haya sido B dado que la pieza fue defectuosa ($P(B|D)$), por lo tanto utilizamos el teorema de Bayes:

$$\begin{aligned}
 P(B|D) &= \frac{P(B) \cdot P(D|B)}{P(A) \cdot P(D|A) + P(B) \cdot P(D|B) + P(C) \cdot P(D|C)} \\
 &= \frac{0.30 \cdot 0.04}{0.45 \cdot 0.03 + 0.30 \cdot 0.04 + 0.25 \cdot 0.05} \\
 &= \frac{12}{38} \\
 &= 0.316
 \end{aligned}$$

c) Calculamos $P(A|D)$ y $P(C|D)$, comparándolas con el valor de $P(B|D)$ ya calculado. Aplicando el teorema de Bayes, obtenemos:

$$\begin{aligned}
 P(A|D) &= \frac{P(A) \cdot P(D|A)}{P(A) \cdot P(D|A) + P(B) \cdot P(D|B) + P(C) \cdot P(D|C)} \\
 &= \frac{0.45 \cdot 0.03}{0.45 \cdot 0.03 + 0.30 \cdot 0.04 + 0.25 \cdot 0.05} \\
 &= \frac{135}{380} \\
 &= 0.355
 \end{aligned}$$

$$\begin{aligned}
 P(C|D) &= \frac{P(C) \cdot P(D|C)}{P(A) \cdot P(D|A) + P(B) \cdot P(D|B) + P(C) \cdot P(D|C)} \\
 &= \frac{0.25 \cdot 0.05}{0.45 \cdot 0.03 + 0.30 \cdot 0.04 + 0.25 \cdot 0.05} \\
 &= \frac{125}{380} \\
 &= 0.329
 \end{aligned}$$

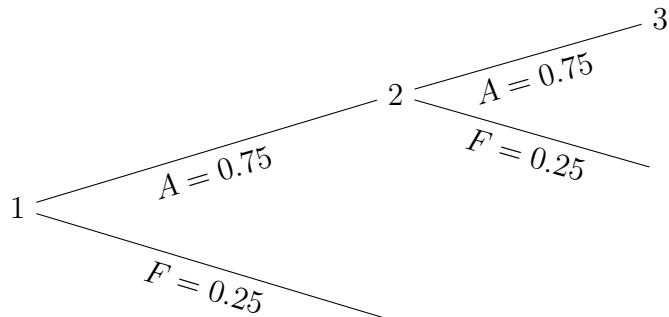
La máquina con mayor probabilidad de haber producido la pieza defectuosa es A .

Ejercicio 8.- Un jugador de baloncesto suele acertar el 75% de sus tiros desde el punto de lanzamiento de personales. Si acierta el primer tiro, puede tirar de nuevo. Calcular la probabilidad de:

- a) Haga dos puntos.
- b) Haga un punto.
- c) No haga puntos.

Solución:

a) Podemos hacer un diagrama de árbol $A = \text{Acierto}$ y $F = \text{Fallo}$, tomando en cuenta que $P(A) = 0.75$ y $P(F) = 1 - P(A) = 1 - 0.75 = 0.25$.



Por lo que:

$$P(\text{dos puntos}) = P(A) \cdot P(A) = 0.75 \cdot 0.75 = 0.56$$

b) La probabilidad de acierto de un punto está dado por que el jugador acertó el primer tiro pero falla el segundo.

$$P(\text{un punto}) = P(A) \cdot P(F) = 0.75 \cdot 0.25 = 0.19$$

c) Finalmente la probabilidad de que no haga ningún punto viene dado por la probabilidad de fallo.

$$P(\text{sin puntos}) = 1 - P(A) = 1 - 0.75 = 0.25$$

Ejercicio 9.- Cada muestra de aire tiene 10% de posibilidades de contener una molécula rara particular. Suponga que las muestras son independientes con respecto a la presencia de la molécula rara. Encuentre la probabilidad de que en las siguientes 18 muestras, exactamente 2 contengan la molécula rara.

Solución:

Sea $X = \text{número de muestras de aire que contienen la molécula rara en las siguientes 18 muestras analizadas}$. Entonces X es una variable aleatoria binomial con $p = 0.1$, $q = 0.9$, $n = 18$ y $k = 2$. Donde:

- p es la probabilidad de éxito, en este caso el éxito es que se contenga la molécula rara que es de 10%

- q es la probabilidad de fracaso, en este caso el fracaso es que no se contenga la molécula rara
- n es el número de muestras
- k es el número que tomaremos de las muestras para hacer el experimento

Recordando la fórmula de la distribución binomial:

$$f(x) = \binom{n}{k} p^k q^{n-k}$$

Sustituyendo valores:

$$P(X = 2) = \binom{18}{2} (0.1)^2 \cdot (0.9)^{16}$$

Donde $\binom{18}{2}$ es:

$$\binom{18}{2} = \frac{18!}{2!(16)!} = \frac{18 \cdot 17}{2} = 153$$

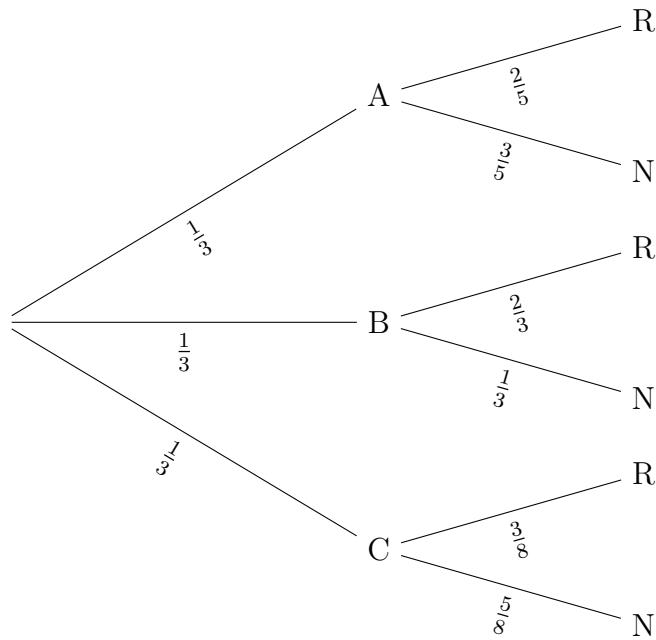
Por lo tanto:

$$P(X = 2) = 153 \cdot (0.1)^2 \cdot (0.9)^{16} = 0.284$$

Ejercicio 10.- 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 ?

Solución:

Llamamos R = "sacar bola roja" y N = "sacar bola negra". Dibujemos un diagrama de árbol para ver las distintas probabilidades de ocurrencia de los sucesos R o N para cada una de las tres urnas, por supuesto la probabilidad de seleccionar una urna es de $\frac{1}{3}$.



La probabilidad pedida es, haber sacado la bola de la urna dado que la bola es roja ($P(A|R)$). Utilizando el teorema de Bayes, tenemos:

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

9.4 Examen resuelto

Instrucciones: Resuelva lo siguiente:

1. Probar que si $P(A^c) = \alpha$ y $P(B^c) = \beta$, entonces $P(A \cap B) \geq 1 - \alpha - \beta$.

Solución:

Sea

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (50)$$

entonces despejando $P(A \cap B)$

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

por los axiomas de Kolmogorov se sabe que

$$0 \leq P(A \cup B) \leq 1$$

de ahí que

$$P(A) + P(B) - P(A \cup B) \geq P(A) + P(B) - 1$$

de Eq(1)

$$P(A \cap B) \geq P(A) + P(B) - 1$$

dado que $P(A) = 1 - \alpha$, $P(B) = 1 - \beta$, entonces

$$\begin{aligned} P(A \cap B) &\geq 1 - \alpha + 1 - \beta - 1 \\ &\geq 1 - \alpha - \beta \end{aligned}$$

2. Dado $P(A) = 0.5$ y $P(A \cup B) = 0.6$, encontrar $P(B)$ si

- (a) A y B son mutuamente excluyentes
- (b) A y B son independientes
- (c) $P(A|B) = 0.4$

Solución:

- (a) A y B son mutuamente excluyentes

Si A y B son mutuamente excluyentes, se debe cumplir que

$$P(A \cup B) = P(A) + P(B) \quad (51)$$

$$P(B) = P(A \cup B) - P(A) \quad (52)$$

entonces

$$\begin{aligned} P(B) &= 0.6 - 0.5 \\ &= 0.1 \end{aligned}$$

- (b) A y B son independientes Si A y B son independientes se debe cumplir que

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= P(A) + P(B) - P(A)P(B) \\ &= P(A) + P(B)[1 - P(A)] \end{aligned}$$

despejando $P(B)$ se tiene

$$\begin{aligned} P(B) &= \frac{P(A \cup B) - P(A)}{1 - P(A)} \\ &= \frac{0.6 - 0.5}{1 - 0.5} = 0.2 \end{aligned}$$

(c) $P(A|B) = 0.4$ Si $P(A|B) = 0.4$ se debe cumplir

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (53)$$

$$P(A \cap B) = P(B)P(A|B) \quad (54)$$

dado que

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (55)$$

Sustituyendo (9) en (10)

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(B)P(A|B) \\ &= P(A) + P(B)[1 - P(A|B)] \\ P(B) &= \frac{P(A \cup B) - P(A)}{1 - P(A|B)} \\ &= \frac{0.6 - 0.5}{1 - 0.4} \\ &= 1/6 \end{aligned}$$

3. Cinco por ciento de la gente tiene alta presión en la sangre. De la gente con alta presión en la sangre, 75% toman alcohol, mientras sólo 50% de la gente sin alta presión toman alcohol. Se desea saber que porcentaje de la gente que toma alcohol tiene alta presión en la sangre?.

Solución:

Sean los eventos

A: evento que denota la gente con alta presión en la sangre.
 B: evento que denota la gente que toma alcohol.

De la información dada se sabe que:

$$\begin{aligned} P(A) &= 0.05 \\ P(A^c) &= 0.95 \\ P(B|A) &= 0.75 \\ P(B|A^c) &= 0.5 \end{aligned}$$

Se pide hallar $P(A|B)$, empleando la regla de Bayes

$$\begin{aligned} P(A|B) &= \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)} \\ &= \frac{(0.75)(0.05)}{(0.75)(0.05) + (0.5)(0.95)} \\ &= 0.0732 \end{aligned}$$

4. Si $P(A) = 1/3$ y $P(B^c) = 1/4$, pueden A y B ser disjuntos? Explique.

Si son disjuntos, entonces

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

sustituyendo $P(A) = 1/3$ y $P(B) = 1 - 1/4 = 3/4$, se tiene que

$$\begin{aligned} P(A \cup B) &= \frac{1}{3} + \left(1 - \frac{1}{4}\right) \\ &= \frac{13}{12} \end{aligned}$$

por los axiomas de Kolmogorov se tiene que $0 \leq P(A \cup B) \leq 1$ y dado que $P(A \cup B) > 1$, se concluye que A y B no pueden ser disjuntos.

5. Si A y B son independientes y $P(A) = P(B) = 1/2$. Cuál es el valor de $P((A \cap B^c) \cup (A^c \cap B))$?

Solución:

Empleando la regla para la unión de dos eventos y dado que A y B son independientes entonces

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

y

$$\begin{aligned} P((A \cap B^c) \cup (A^c \cap B)) &= P(A \cap B^c) + P(A^c \cap B) - P((A \cap B^c) \cap (A^c \cap B)) \\ &= P(A \cap B^c) + P(A^c \cap B) - P(A \cap B^c)P(A^c \cap B) \\ &= P(A)P(B^c) + P(A^c)P(B) - P(A \cap B^c)P(A^c \cap B) \\ &= P(A)P(B^c) + P(A^c)P(B) - P(A)P(B^c)P(A^c)P(B) \end{aligned}$$

dado que $P(A) = P(B) = 1/2$ entonces $P(A^c) = P(B^c) = 1/2$, sustituyendo en la expresión anterior se tiene

$$\begin{aligned} P((A \cap B^c) \cup (A^c \cap B)) &= P(A)P(B^c) + P(A^c)P(B) - P(A)P(B^c)P(A^c)P(B) \\ &= \frac{1}{2} \frac{1}{2} + \frac{1}{2} \frac{1}{2} - \frac{1}{2} \frac{1}{2} \frac{1}{2} \frac{1}{2} \\ &= \frac{1}{4} + \frac{1}{4} - \frac{1}{16} \\ &= \frac{7}{16} \end{aligned}$$

6. Suponga que B_1 , B_2 y B_3 son mutuamente excluyentes. Si $P(B_j) = 1/3$ y $P(A|B_j) = j/6$ para $j = 1, 2, 3$, determine cuál es el valor de $P(A)$?

Solución:

Por el teorema de probabilidad total

$$\begin{aligned}
 P(A) &= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3)P(B_3) \\
 &= \sum_{j=1}^3 P(A|B_j)P(B_j) \\
 &= \sum_{j=1}^3 \frac{j}{6} \frac{1}{3} \\
 &= \frac{1}{18} \sum_{j=1}^3 j \\
 &= \frac{1}{18} (1 + 2 + 3) \\
 &= \frac{1}{3}
 \end{aligned}$$

7. Demuestre que $\binom{n+1}{k+1} = \frac{n+1}{k+1} \binom{n}{k}$

Demostración:

$$\begin{aligned}
 \binom{n+1}{k+1} &= \frac{(n+1)!}{(k+1)!(n-k)!} \\
 &= \frac{n!(n+1)}{(k+1)!(n-k)!} \\
 &= \frac{n!(n+1)}{k!(k+1)(n-k)!} \\
 &= \frac{n!}{k!(n-k)!} \frac{n+1}{k+1} \\
 &= \binom{n}{k} \frac{n+1}{k+1} \\
 &= \frac{n+1}{k+1} \binom{n}{k}
 \end{aligned}$$

8. Dos jugadores A y B alternadamente e independientemente lanzan una moneda, el primer jugador en obtener una cara gana, supóngase que el jugador A lanza primero.

- (a) Si la moneda es honesta ¿Cuál es la probabilidad de que A gane?.
- (b) Supóngase que $P(\text{cara}) = p$ no necesariamente $1/2$. ¿Cuál es la probabilidad de que A gane?
- (c) Demuestre que para cualquier p , $0 < p < 1$, $P(\text{A gana}) > 1/2$

Solución:

- (a) Si la moneda es honesta ¿Cuál es la probabilidad de que A gane?. Sea la probabilidad de que A gane

$$\begin{aligned} P(A \text{ gana}) &= \sum_{i=1}^{\infty} P(\text{A gana en el } n\text{-ésimo lanzamiento}) \\ &= \frac{1}{2} + \left(\frac{1}{2}\right)^2 \frac{1}{2} + \left(\frac{1}{2}\right)^4 \frac{1}{2} + \dots \end{aligned}$$

note que la suma corresponde a una progresión geométrica, donde $a = 1/2$ y $r = 1/4$, por lo que

$$S = \frac{a(1 - r^n)}{1 - r}$$

cuando $n \rightarrow \infty$

$$\begin{aligned} S &= \frac{a}{1 - r} \\ &= \frac{1/2}{1 - 1/4} \\ &= \frac{2}{3} \end{aligned}$$

Por lo que $P(A \text{ gana}) = 2/3$.

- (b) Supóngase que $P(\text{cara}) = p$ no necesariamente $1/2$. ¿Cuál es la probabilidad de que A gane?

El procedimiento es el mismo que el caso anterior solo generalizamos para cualquier $P(\text{cara})$.

$$\begin{aligned} P(A \text{ gana}) &= \sum_{k=1}^{\infty} P(\text{A gana en el } k\text{-ésimo lanzamiento}) \\ &= p + (1 - p)^2 p + (1 - p)^4 p \dots \\ &= \sum_{k=1}^{\infty} (1 - p)^{2k} p \end{aligned}$$

note que la suma corresponde a una progresión geométrica, donde $a = p$ y $r = (1 - p)^2$, por lo que

$$S = \frac{a(1 - r^n)}{1 - r}$$

cuando $n \rightarrow \infty$

$$\begin{aligned} S &= \frac{a}{1 - r} \\ &= \frac{p}{1 - (1 - p)^2} \\ &= \frac{1}{2 - p} \text{ donde } 0 < p \leq 2 \end{aligned}$$

Por lo que $P(A \text{ gana}) = \frac{1}{2-p}$.

- (c) Demuestre que para cualquier p , $0 < p < 1$, $P(A \text{ gana}) > 1/2$

Se pide

$$P(A \text{ gana}) > \frac{1}{2}$$

Sustituyendo el resultado del inciso anterior

$$\begin{aligned} \frac{1}{2-p} &> \frac{1}{2} \\ 2 &> 2-p \\ p &> 0 \end{aligned}$$

considerando $\lim_{p \rightarrow 0} = 1/2$ y $\lim_{p \rightarrow 1} = 1$, y como se tiene que $0 < p < 1$, entonces necesariamente $P(A \text{ gana}) > 1/2$.

9. Hallar n si $P(n, 2)=72$.

Solución:

$$P_2^n = \frac{n!}{(n-2)!} = n(n-1) = n^2 - n = 72$$

Puesto que n debe ser positivo, la única respuesta es 9.

10. Muestre el valor de $P(A \cup B \cup C)$

Solución:

$$\begin{aligned} P(A \cup (B \cup C)) &= P(A) + P(B \cup C) - P(A \cap (B \cup C)) \\ &= P(A) + P(B) + P(C) - P(B \cap C) - P((A \cup B) \cap (A \cup C)) \\ &= P(A) + P(B) + P(C) - P(B \cap C) - P(A \cap B) - P(A \cap C) \\ &\quad + P(A \cup B \cap A \cap C) \\ &= P(A) + P(B) + P(C) - P(B \cap C) - P(A \cap B) - P(A \cap C) \\ &\quad + P(A \cup B \cap C) \end{aligned}$$

Capítulo 10

10.1 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:

$$\beta_2 = \frac{m\mu_4}{\sigma^4} \quad (56)$$

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:

$$g_2 = \frac{m\mu_4}{\sigma^4} - 3 \quad (57)$$

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 $\frac{\mu_4}{\sigma^4}$.

10.2 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:

$$\sum_{i=1}^k f_i(x_i - \mu) = \sum_{i=1}^k f_i x_i - \mu \sum_{i=1}^k f_i = \mu - \mu = 0 \quad (58)$$

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:

$$\gamma_1 = \frac{\mu_3}{\sigma^3} \quad (59)$$

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.

$$A_p = \frac{\mu - \text{moda}}{\sigma} \quad (60)$$

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

Si la distribución es simétrica, $\mu = \text{moda}$ 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 ex-

presión:

$$A_B = \frac{Q_{\frac{3}{4}} + Q_{\frac{1}{4}} - 2M_e}{Q_{\frac{3}{4}} - Q_{\frac{1}{4}}} \quad (61)$$

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.

10.3 Algoritmos para calcular π

En toda la historia se han descubierto/creado muchas maneras para calcular aproximaciones del número Pi.

10.3.1 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:

$$a_{n+1} = \frac{a_n + b_n}{2} \quad (62)$$

$$b_{n+1} = \frac{2}{\sqrt{a_n b_n}} \quad (63)$$

$$t_{n+1} = t_n - p_n(a_n - a_{n+1})^2 \quad (64)$$

$$p_{n+1} = 2p_n \quad (65)$$

Entonces π se aproxima de la siguiente forma:

$$\pi \approx \frac{(a_n + b_n)^2}{4t_n} \quad (66)$$

El método tiene convergencia de segundo orden. Es decir, en cada iteración duplicamos el número de dígitos exactos obtenidos en la iteración anterior.

10.3.2 Algoritmo de Borwein

Partimos de los siguientes datos iniciales:

$$\begin{aligned} a_0 &= 6 - 4\sqrt{2} \\ y_0 &= \sqrt{2} - 1 \end{aligned}$$

Con ellos operamos de la siguiente forma:

$$y_{k+1} = \frac{1 - (1 - y_k^4)^{\frac{1}{4}}}{1 + (1 - y_k^4)^{\frac{1}{4}}} \quad (67)$$

$$a_{k+1} = a_k(1 + y_{k+1})^4 - 2^{2k+3}y_{k+1}(1 + y_{k+1} + y_{k+1}^2) \quad (68)$$

Entonces se tiene lo siguiente:

$$\lim_{k \rightarrow \infty} (a_k) = \frac{1}{\pi} \quad (69)$$

Capítulo 11

11.1 Resumen Capítulo 5 del libro Probabilidad Serie Schaum

11.1.1 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 [10].

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:

$$\begin{aligned}(X + Y)(s) &= X(s) + Y(s) \\ kX(s) &= kX(s) \\ (X + k)(s) &= X(s) + k \\ (XY)(s) &= X(s)Y(s)\end{aligned}$$

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,

$$\begin{aligned}P(X = a) &= P(\{s \in S : X(s) = a\}) \\ \text{y } P(a \leq X \leq b) &= P(\{s \in S : a \leq X(s) \leq b\})\end{aligned}$$

11.1.2 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 $f(x_i) = P(S = x_i)$, 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

$$E(X) = x_1 f(x_1) + x_2 f(x_2) + \cdots + x_n f(x_n) = \sum_{i=1}^n x_i f(x_i)$$

Esto es, $E(X)$ es el promediado 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 conuento 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
$g(x_i)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

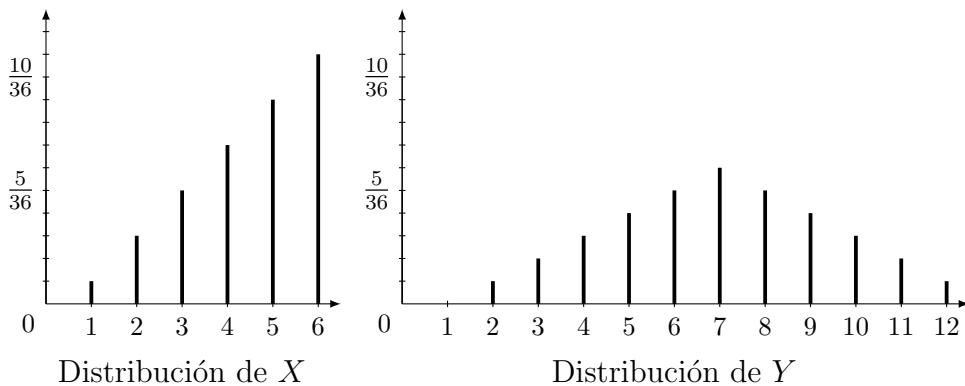
Obtenemos, por ejemplo, $g(4) = \frac{3}{36}$ del hecho de que, $(1,3), (2,2)$ y $(3,1)$ son aquellos puntos de S para los que la suma de componentes es 4; por tanto

$$g(4) = P(Y = 4) = P((1,3), (2,2), (3,1)) = \frac{3}{36}$$

La media de Y se calcula como sigue:

$$E(Y) = \sum y_i g(y_i) = 2 \cdot \frac{1}{36} + 3 \cdot \frac{2}{36} + \dots + 12 \cdot \frac{1}{36} = 7$$

Los siguientes diagramas describen gráficamente las distribuciones anteriores:



11.1.3 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 $\text{var}(X)$

$$\sigma_x = \sqrt{\text{var}(X)}$$

11.1.4 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_1	y_2	\dots	y_m	Suma
x_1		$h(x_1, y_1)$	$h(x_1, y_2)$	\dots	$h(x_1, y_m)$	$f(x_1)$
x_2		$h(x_2, y_1)$	$h(x_2, y_2)$	\dots	$h(x_2, y_m)$	$f(x_2)$
\dots		\dots	\dots	\dots	\dots	\dots
x_n		$h(x_n, y_1)$	$h(x_n, y_2)$	\dots	$h(x_n, y_m)$	$f(x_n)$
Suma		$g(y_1)$	$g(y_2)$	\dots	$g(y_m)$	

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

$$(i) h(x_i, y_j) \geq 0$$

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

$$\text{cov}(X, Y) = \sum_{i,j} (x_i - \mu_X)(y_i - \mu_Y)h(x_i, y_j) = E[(X - \mu_X)(Y - \mu_Y)]$$

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 dimensinada y tiene las siguientes propiedades:

$$(i) = \rho(X, Y) = \rho(Y, X)$$

$$(ii) -1 \leq \rho \leq 1$$

$$(iii) \rho(X, X) = 1, \rho(X, -X) = -1$$

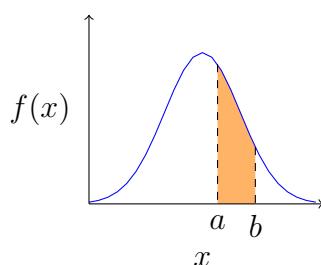
$$(iv) \rho(aX + b, cY + d) = \rho(X, Y), \text{ si } a, c \neq 0$$

11.1.5 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.



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

$$\begin{aligned} (i) f(x) &\geq 0 \\ (ii) \int_{\mathbb{R}} f(x)dx &= 1 \end{aligned}$$

Esto es, f no es negativa y el área total bajo su curva es 1.

11.1.6 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.

$$E(X^2) = \int_{\mathbb{R}} = x^2f(x)dx$$

Y con esto podemos obtener la varianza que se define por

$$var(X) = E(X^2) - \mu^2$$

Y por último se define la desviación estándar como

$$\sigma_x = \sqrt{var(X)}$$

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.

11.2 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.

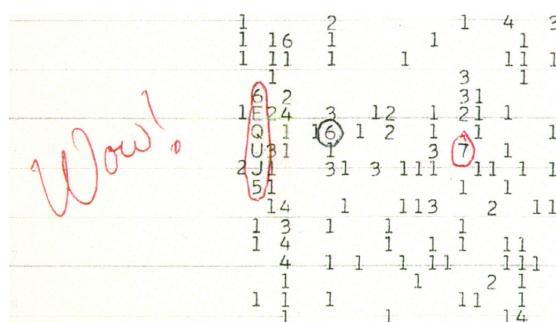


Figura 61: Señal Wow

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.

Detección de la señal

La computadora del radio-observatorio, una IBM 1130 equipada con 1 MB de disco duro y 32 KB de memoria RAM, se encargaba de convertir los datos recibidos directamente por el radio-telescopio a una serie de caracteres alfanuméricos. El software, diseñado por Bob Dixon y Jerry Ehman, era bastante sofisticado, ya que hacía continuos chequeos del funcionamiento del equipo y era capaz de ejecutar varios algoritmos de búsqueda simultáneamente, incluidos unos algoritmos de búsqueda capaces de aislar señales pulsantes o continuas. Además, sirvió para solucionar la falta de espacio en los registros de impresora y el ahorro de tinta ya que se estaban rastreando cincuenta canales en la frecuencia del hidrógeno neutro (1420 MHz). Cada fila representaba los resultados de los datos recogidos durante aproximadamente doce segundos de búsqueda. Eran necesarios diez segundos para obtener las intensidades

de todos los canales, y aproximadamente dos segundos para que la computadora procesara los datos recibidos. Las columnas representaban las intensidades para los cincuenta canales en rastreo, de 10kHz de ancho de banda cada uno, con el canal n. 1 situado en el extremo izquierdo y el canal n. 50 situado en el extremo derecho.

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

- 1 En un primer momento se dividía en seis porciones la señal recibida en cada canal, de las cuales se separaban $1/6$ del valor actual y $5/6$ del valor anterior y eran separadas para eliminar el ruido de base.
- 2 En el siguiente paso el resto era dividido por la desviación estándar computada sobre 60 períodos (porciones de señal), $1/60$ del valor actual más $59/60$ del valor anterior.
- 3 El número calculado en el primer paso era dividido por el número calculado en el segundo. Esta operación daba la relación de ruido de la señal.
- 4 Después se tomaba la parte entera de esta relación de ruido de la señal.
- 5 Por último, el número entero era impreso con las siguientes modificaciones: Si el valor era un 0, se representaba mediante un espacio en blanco; los valores entre el 1 y el 9 eran impresos tal cual; los enteros del 10 al 35 eran representados con las letras mayúsculas que van de la A a la Z respectivamente. Si alguna señal tenía una intensidad de 36 o superior, el programa simplemente empezaba de nuevo desde 0. Así, por ejemplo, el valor 39 sería convertido a 4 (39-35).

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

- 6: los valores entre 6,0 y 6,999...
- E: los valores entre 14,0 y 14,999...
- Q: los valores entre 26,0 y 26,999...
- U: los valores entre 30,0 y 30,999...
- J: los valores entre 19,0 y 19,999...
- 5: los valores entre 5,0 y 5,999...

El intervalo más intenso recibido (la "U" significa que la señal era treinta veces más intensa que el ruido de fondo. Mucho de este ruido de fondo llega al receptor sin que se vea alterado, pero algunos ruidos pueden provenir de los árboles, de la hierba u otros objetos circundantes, y algo proviene del remanente del Big Bang, explosión que se estima habría ocurrido hace 13.700 millones de años.

1420,4056 MHz - Hidrógeno neutro

¿Por qué en esta frecuencia? Pues porque es la del elemento más abundante en el Universo. Hay millones de frecuencias posibles en todo el espectro radio-eléctrico, pero se piensa que cualquier civilización inteligente lo suficientemente avanzada como para estudiar el universo, debería conocer la radio-astronomía y por tanto hacer investigaciones radioastronómicas. Si esto es así deberían conocer la frecuencia natural de emisión del hidrógeno neutro, que al ser el elemento más abundante del universo proporciona un canal óptimo para la emisión y recepción de señales.

Ya tenemos el canal, pero ¿en qué tipo de onda podemos esperar recibir un posible mensaje? Hay varios tipos de ondas que se diferencian por sus características a la hora de imprimirlas un mensaje, como son la modulación de frecuencia (FM), la modulación de amplitud (AM), modulación de fase, modulación digital, banda lateral única, etc... Pero de entre todas ellas destaca una en particular por su capacidad de concentrar gran cantidad de energía en el menor ancho de banda. Esta es la conocida como onda continua o CW (Continuous Wave) que por ser de una frecuencia fija y estable es la onda óptima para salvar las grandes distancias interestelares a la vez que es capaz de ser escuchada a niveles muy bajos de señal (el código Morse se emite en CW).

Ancho de banda y frecuencia

El ancho de banda de la señal es menor a 10 kHz. Para la frecuencia se han dado dos valores diferentes: 1420.356 MHz y 1420.456 MHz; en cualquier caso, esas frecuencias están próximas a la frecuencia de transición hiperfina del hidrógeno. Esa frecuencia forma parte del espectro de radio donde está prohibida la emisión por tratados internacionales.

Se sugieren dos posibles coordenadas ecuatoriales:

- A.R. = $19^h 22^m 22^s \pm 5^s$
- A.R. = $19^h 25^m 12^s \pm 5^s$

Ambas coordenadas tienen declinación = $-27^\circ 03' 20''$.

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.

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 señal WOW!, 72 segundos, como su forma, corresponderían a un origen extraterrestre. Hay tres hipótesis que intentan explicar el fenómeno:

- 1 La señal proviene de una emisión de radio de un satélite artificial que atravesó esa órbita en ese instante.
- 2 La emisión de radio fue producto de un acontecimiento astronómico de enorme potencia.
- 3 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 señal, 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.

11.3 Reseña del artículo "Is Economics the Physical Science?"

En las últimas décadas físicos han realizado investigaciones en economía, la gente involucrada en este tipo de temas lo llama econofísica. Recientemente también se han escrito libros acerca de este tema, se realizan tesis de doctorado sobre economía, y gran parte de los profesores físicos en Europa se especializan en econofísica. A pesar de que los economistas son mejor entrenados que los físicos en análisis estadísticos, los físicos han llegado a este campo con distintos objetivos y filosofía, dado que los físicos manejan con las leyes universales y los economistas analizan la diferencia en el almacenamiento de distintos lugares, etc. Los físicos han llegado a este campo con ojos frescos y distintas hipótesis [7].

La mayoría del trabajo de los físicos en economía concierne en leyes de potencia, la primera ley de potencia de distribución en cualquier campo fue observada en economía, pero los físicos han metido más conceptos útiles, como la varianza y la desviación estándar además de métodos analíticos.

Los físicos han descubierto que la volatilidad es solo uno de los procesos severos de gran-memoria en mercado. Uno de los más sorprendentes de estos concierne en fluctuaciones en fuente y demanda.

El principio de la eficiencia del mercado toma distintas formas. Un mercado es informalmente eficiente si toda la información disponible refleja los precios; este es arbitrariamente eficiente si es imposible para los inversores hacer exceso de beneficios. Una de las consecuencias de informalmente eficiente es que los precios no deberían de ser predecibles. En realidad el término de precio inpredecible no es malo, aún las mejores estrategias de comercio explotan pero sólo muy debidamente los niveles de predictibilidad.

Un gran problema de trabajo empírico utiliza métodos y analogías desde la física. Por ejemplo, teoría de matrices aleatorias, desarrollada en física nuclear, y correlaciones ultramétricas han provisto utilidad para entender la correlación entre los movimientos de los precios de distintas empresas.

El estudio de la economía tiene un lazo con el estudio de las personas, por lo que se han implementado métodos de inteligencia artificial, simulación en computadoras sobre el comportamiento humano todo esto para una búsqueda racional. Los físicos se han unido con los economistas para buscar nuevas teorías que no sean completamente una selección racional.

El enfoque de cero inteligencia se remonta a la obra de Herbert Simon, premio Nobel de Economía y uno de los pioneros en la inteligencia artificial. El principal modelo campeón en años recientes ha sido de los físicos, quienes han utilizado analogías a mecanismos estáticos a desarrollar nuevos modelos de mercado.

Como otra alternativa a la selección racional, econofísicos han desarrollado modelos altamente simplificados de estrategias de interacción para capturar la esencia del comportamiento colectivo en un mercado financiero.

Finanzas no es la única área de economía en la cual los físicos están activos. Esperamos que en los próximos años la física y algunos departamentos de economía diseñen un curso básico de la enseñanza de los elementos esenciales de la física y la economía. Los físicos seguirán contribuyendo a la economía en una variedad de temas que van desde la macroeconomía de la microestructura del mercado, y sus contribuciones tendrán incremento en las implicaciones para la formulación de políticas económicas. Un área de oportunidad, en la que la aplicabilidad de la física podría no ser evidente a priori, se refiere a la construcción de índices económicos, tales como el Índice de Precios al Consumidor o el Promedio Industrial Dow Jones. Esos índices proporcionan sólo resúmenes escalares de crudo de fenómenos muy complejos, sino que también juegan un papel importante en la toma de decisión en economía. Por ejemplo, el pago de pensiones y salarios se basan en el CPI. Índices económicos son actualmente construidos utilizando esencialmente métodos ad hoc. Creemos que la exactitud de tales índices podría mejorarse mediante la reflexión cuidadosa en términos de análisis dimensional, combinado con mejores datos de análisis que se correlaciona precios y otros factores de apropiarse de las cualidades relacionadas con los índices. La construcción de índices económicos precisos en última instancia, se relaciona con la comprensión de por qué la economía exhibe tantos comportamientos libres de escala, tales como la distribución de la riqueza o el tamaño de las empresas.

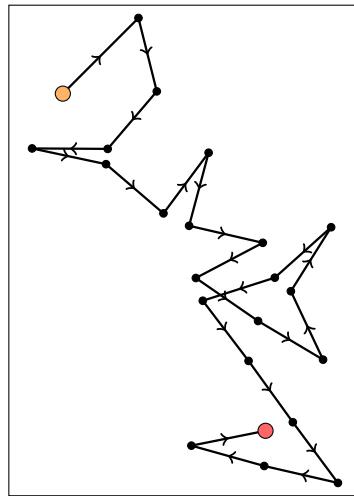
Para arrojar luz sobre esta cuestión, los econofísicos necesitan entender mejor las dimensiones naturales de la vida económica, y el análisis dimensional sistemático será una herramienta muy útil. Métodos dimensionales y de escala eran una piedra angular en la elucidación de fenómenos complejos como la turbulencia en los fluidos, y todos los componentes que hacen de flujo de fluido correlaciones de largo plazo complejo, no lineal, y el caos, es probable que sean factores importantes en la economía.

Capítulo 12

12.1 Caminata aleatoria

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.

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 Márkov. Varias propiedades de los paseos aleatorios incluyen distribuciones dispersas, tiempos de primer cruce y rutas de encuentro.



Caminatas aleatorias en grafos

En el estudio de la teoría general de las caminatas aleatorias, aparece con bastante frecuencia que el espacio donde se requiere realizar la caminata, puede ser modelado como cierto grafo. La situación usual es como sigue: Dado un grafo G y comenzando en uno de sus vértices, seleccionamos de alguna manera al azar uno de sus vecinos y nos movemos a éste; entonces nosotros seleccionamos un vecino de éste último vértice y nos movemos de nuevo, etc. La sucesión aleatoria de vértices obtenidos de esta forma es una caminata aleatoria sobre el grafo G . La teoría relacionada con

caminatas aleatorias se desarrolla en el marco general de la teoría de los procesos estocásticos, más exactamente con la relacionada con las cadenas de Márkov, y no solo eso; no hay mucha diferencia entre la teoría de las caminatas aleatorias en grafos y la teoría de las cadenas de Márkov finitas ya que cada cadena de Márkov de estas, puede ser vista como una caminata aleatoria sobre cierto grafo dirigido. De manera similar, las cadenas de Márkov reversibles pueden ser vistas como caminatas aleatorias en grafos no dirigidos, y las cadenas de Márkov simétricas, como caminatas aleatorias en grafos regulares.

Caminatas aleatorias en grafos surgen en muchos modelos en matemáticas y en física. De hecho, ésta es una de esas nociones que empiezan a aparecer en todas partes una vez se empieza a buscarlas. Por ejemplo, considere la disposición de una baraja de cartas, construya un grafo cuyos vértices sean todas las permutaciones de las cartas de la baraja de tal manera que dos permutaciones son adyacentes si y solo si una se obtiene a partir de la otra cambiando la posición de dos de las cartas. Barajar el mazo de cartas, corresponde a una caminata aleatoria en este grafo.

Recientemente caminatas aleatorias en grafos más generales, aunque finitos, han recibido mayor atención, y los aspectos estudiados son más cuantitativos: cuánto se debe caminar hasta llegar a la posición inicial, hasta llegar a un vértice dado o hasta pasar por todos los vértices del grafo. Las caminatas aleatorias también están relacionadas con otras ramas de la teoría de grafos; propiedades básicas de las caminatas aleatorias son determinadas por el espectro del grafo y también por la resistencia eléctrica de la red asociada de manera natural con éste; es por esto que gran parte de la terminología correspondiente a tales caminatas se da en términos de la teoría de redes eléctricas lo cual resulta ser bastante útil dado que es posible extraer resultados de tal teoría a la de caminatas aleatorias en grafos y viceversa. Todas esas conexiones son fructíferas y dan tanto herramientas para el estudio como oportunidades para encontrar nuevas aplicaciones.

12.2 Simulación caminata aleatoria

12.2.1 Código en ROOT

```
//Autor: Itzael Jimenez Aranda
//Instituto: CIC-IPN
//Materia: Probabilidad , procesos aleatorios e inferencia
void random_walk(){
    TRandom3 *r = new TRandom3(0);
    TCanvas *c1 = new TCanvas("h_caminata_azar_gauss_1"
                            , "Lanzamientos"
                            , 20
                            , 10
                            , 700
                            , 700);
    c1->SetGrid();

    TCanvas *c2 = new TCanvas("h_rw_pasos_derecha_gauss_1"
                            , "Distribucion"
                            , 800
                            , 10
                            , 700
                            , 700);
    c2->SetGrid();

    TH1F* hrw= new TH1F("Datos"
                        , "Posicion_final"
                        , 100
                        , -20
                        , 20);
    hrw->SetLineColor(209);
    hrw->SetLineWidth(4);

    TH1F* hrwr= new TH1F("Datos"
                        , "Pasos_a_la_derecha"
                        , 100
                        , 0
                        , 20);
    hrwr->SetLineColor(209);
    hrwr->SetLineWidth(4);

    int paso = 0;
    int N_pasos = 0;
    int N_pasos_d = 0;
    int conteo = 0;
    int repeticiones = 100000;
```

```

for(int i = 0; i < repeticiones; i++){
    while(N_pasos < 20){
        paso = (r->Integer(10))+1;
        if(paso < 6 ){
            conteo--;
            N_pasos++;
        }
        else if(paso > 5){
            conteo++;
            N_pasos++;
            N_pasos_d++;
        }
    }
    hrw->Fill(conteo);
    hrwr->Fill(N_pasos_d );
    conteo = 0;
    paso = 0;
    N_pasos = 0;
    N_pasos_d = 0;
}
}

TLatex *x1label = new TLatex();
x1label->SetTextFont(43); x1label->SetTextSize(23);

TLatex *y1label = new TLatex();
y1label->SetTextFont(43); y1label->SetTextSize(23);

c1->cd();
hrw->Scale(1/hrw->Integral());
hrw->Draw();
hrw->Fit("gaus",0);
x1label ->DrawLatex(22, 0, "m");
y1label ->DrawLatex(-21, 0.19200, "#frac{n}{N}");
double mean, dev;
mean = hrw->GetMean();
dev = hrw->GetStdDev();
cout << mean << endl;
cout << dev << endl;
c2->cd();
hrwr->Scale(1/hrwr->Integral());
hrwr->Draw();
hrwr->Fit("gaus",0);
x1label ->DrawLatex(21, 0, "m");
y1label ->DrawLatex(-0.5, 0.19200, "#frac{n_{1}}{N}");
}
}

```

En los siguientes histogramas se muestra la distribución de la posición final cuando la probabilidad de dar un paso a la derecha es la misma que dar un paso a la izquierda y la distribución de cuantos pasos en total se dieron a la derecha.

12.2.2 Histograma de posición final con paso a la derecha de 0.5 de probabilidad

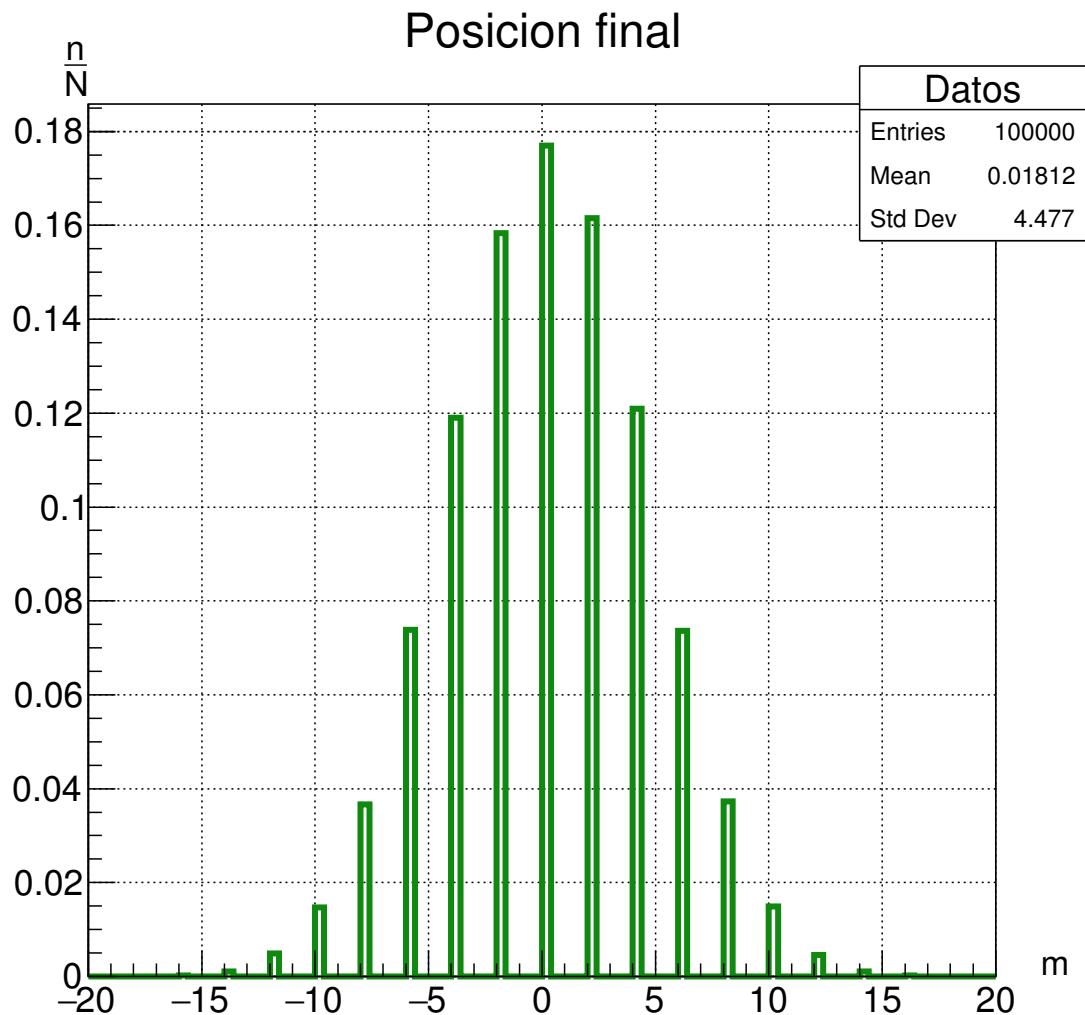


Figura 62: Posición final con paso a la derecha de 0.5 de probabilidad

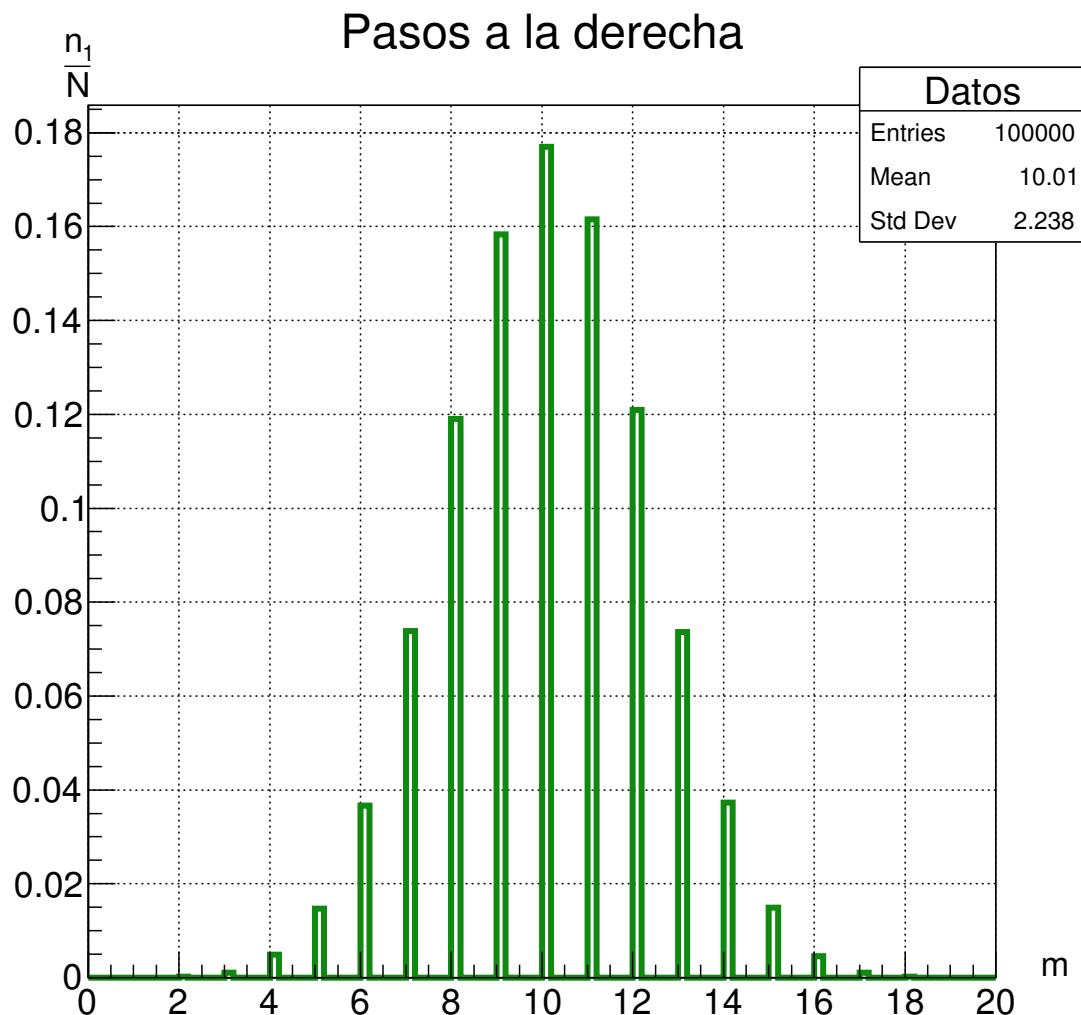
12.2.3 Histograma de los pasos totales a la derecha con paso a la derecha de 0.5 de probabilidad

Figura 63: Pasos totales a la derecha con paso a la derecha de 0.5 de probabilidad

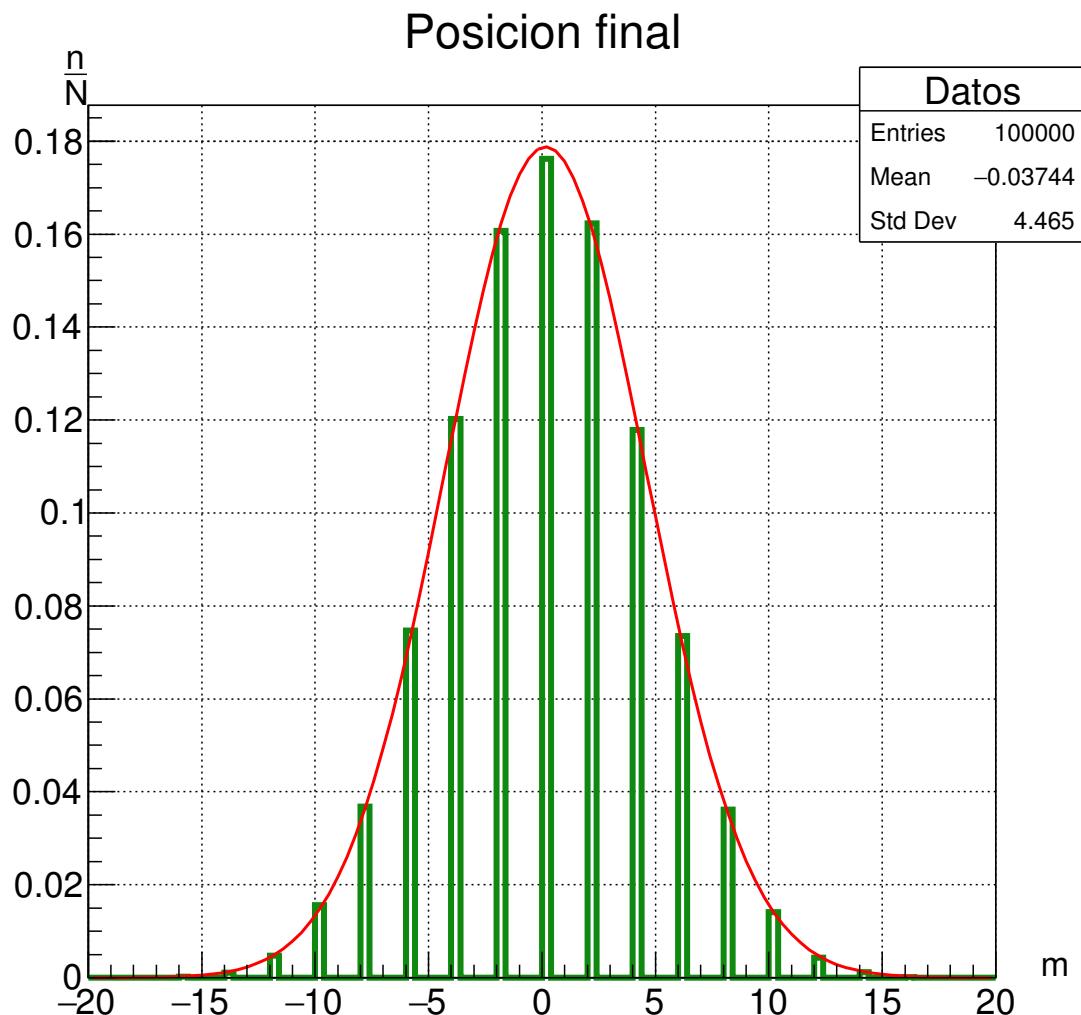
12.2.4 Ajuste de una Gaussiana de la posición final con paso a la derecha de 0.5 de probabilidad

Figura 64: Ajuste de una Gaussian de la posición final con paso a la derecha de 0.5 de probabilidad

12.2.5 Ajuste de una Gaussiana de los pasos totales a la derecha con paso a la derecha de 0.5 de probabilidad

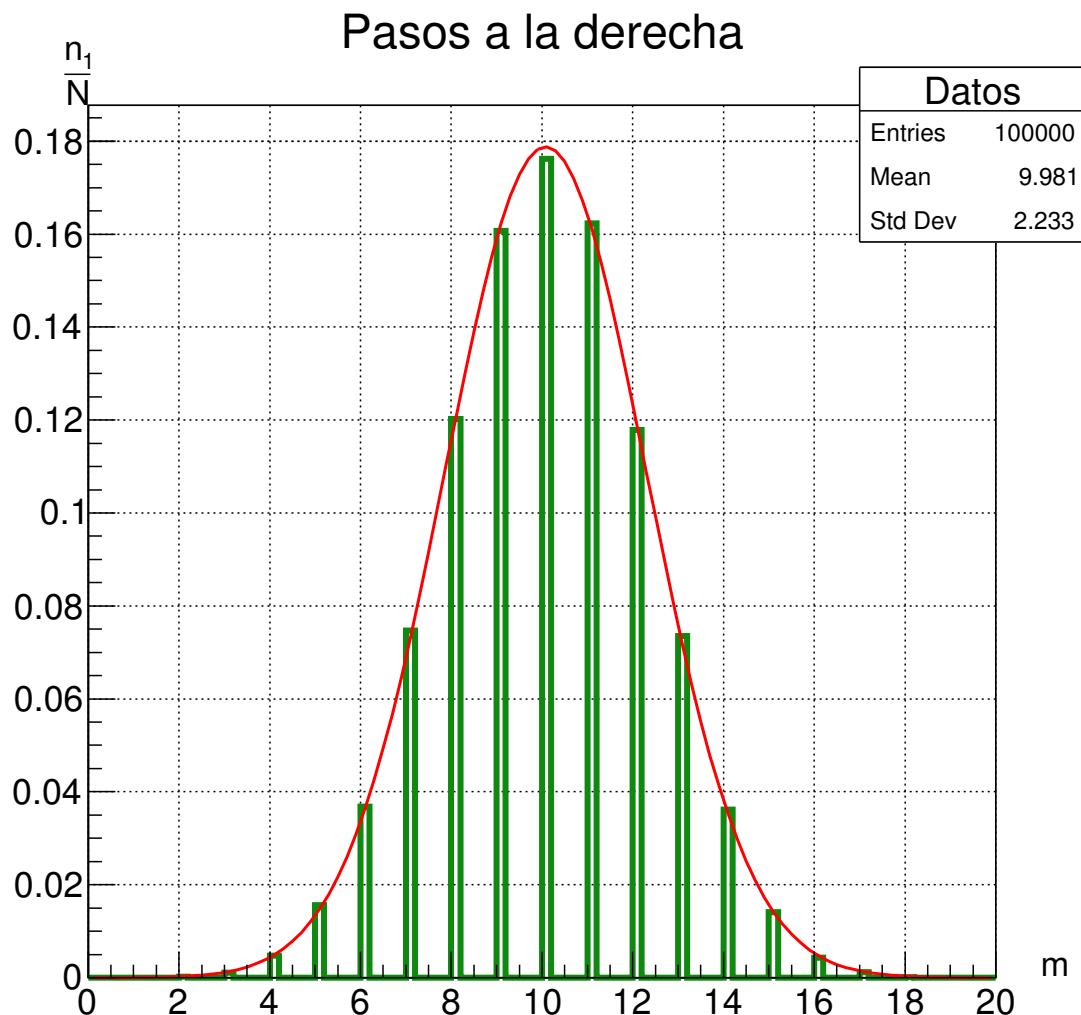


Figura 65: Ajuste de una Gaussiana de los pasos totales a la derecha con paso a la derecha de 0.5 de probabilidad

En los siguientes histogramas se muestra la distribución de la posición final cuando la probabilidad de dar un paso a la derecha es de 0.6 y la distribución de cuantos pasos en total se dieron a la derecha.

12.2.6 Histograma de posición final con paso a la derecha de 0.6 de probabilidad

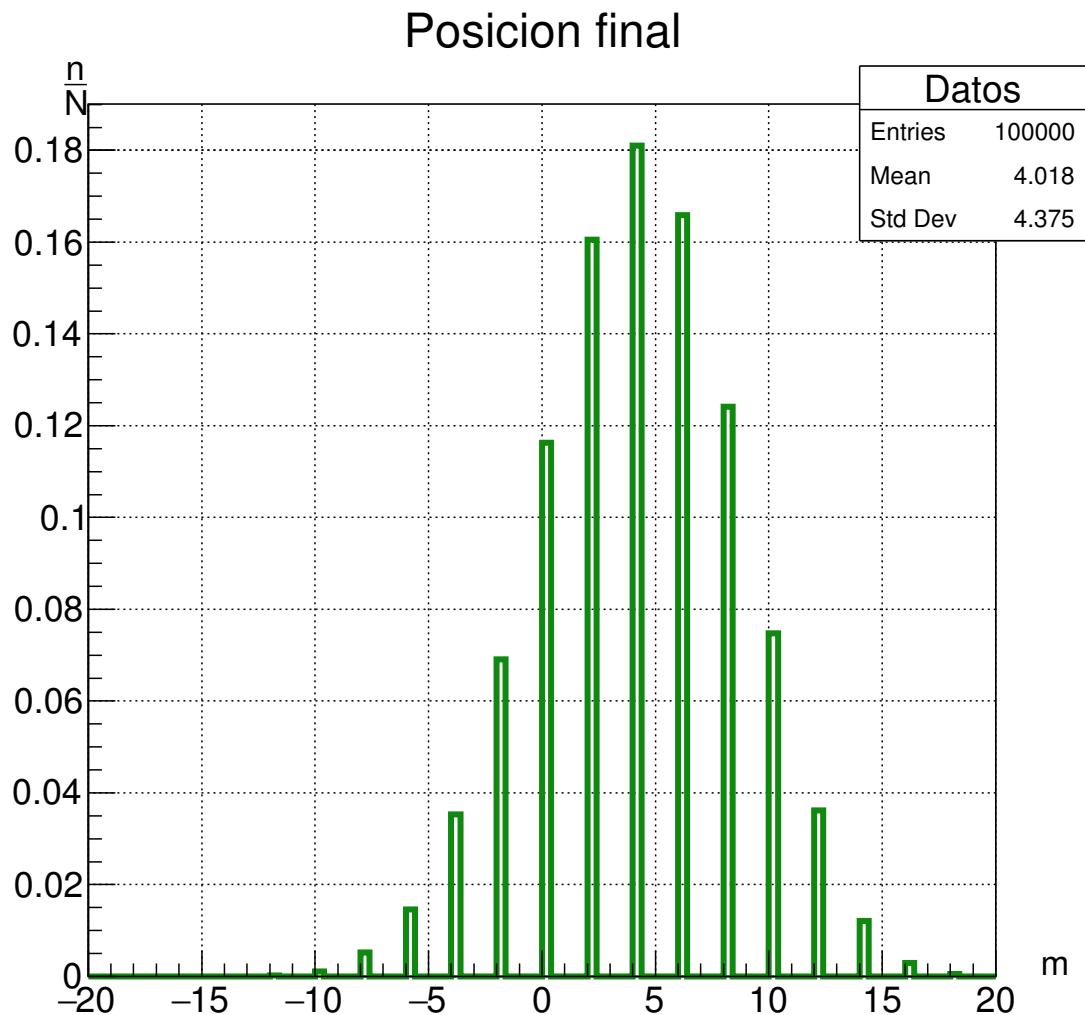


Figura 66: Posición final con paso a la derecha de 0.6 de probabilidad

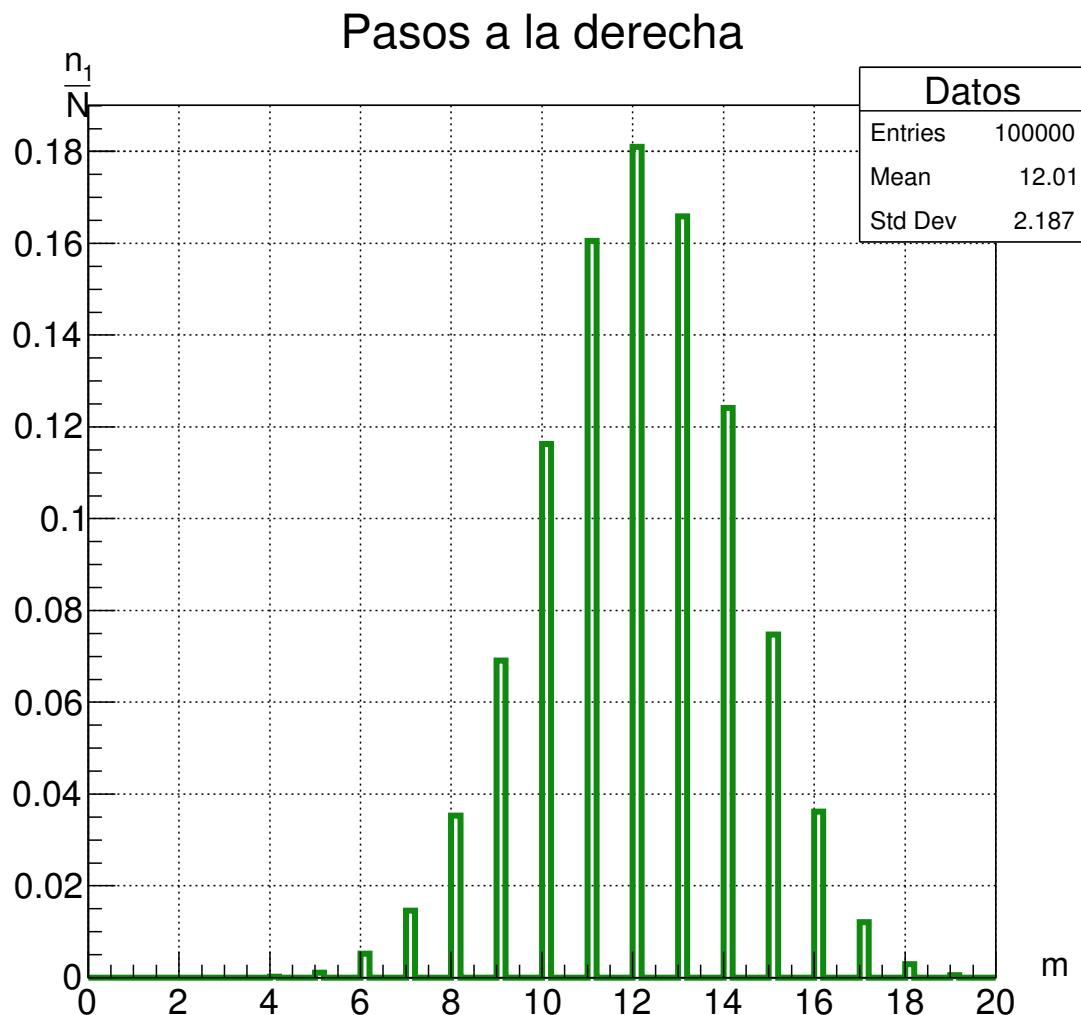
12.2.7 Histograma de los pasos totales a la derecha con paso a la derecha de 0.6 de probabilidad

Figura 67: Pasos totales a la derecha con paso a la derecha de 0.6 de probabilidad

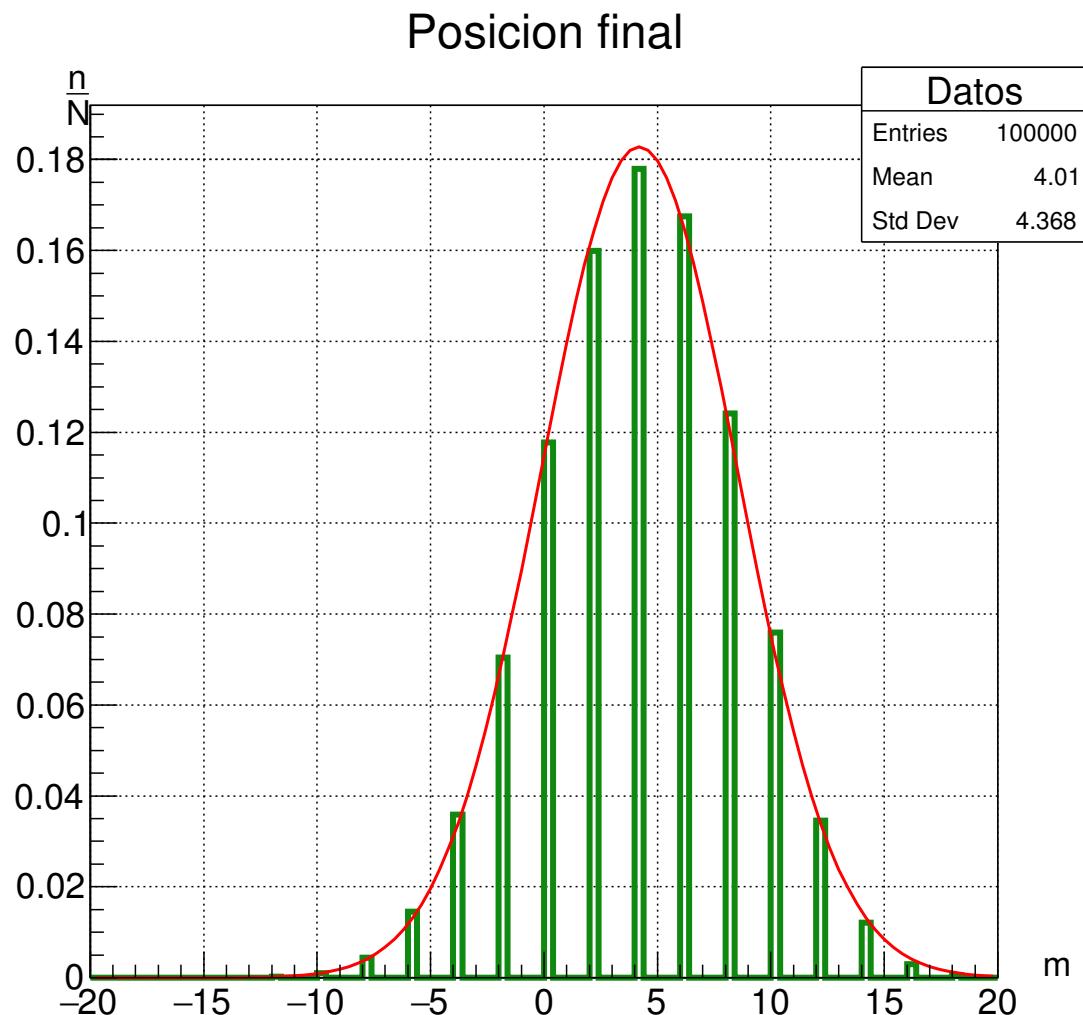
12.2.8 Ajuste de una Gaussiana de la posición final con paso a la derecha de 0.6 de probabilidad

Figura 68: Ajuste de una Gaussiana de la posición final con paso a la derecha de 0.6 de probabilidad

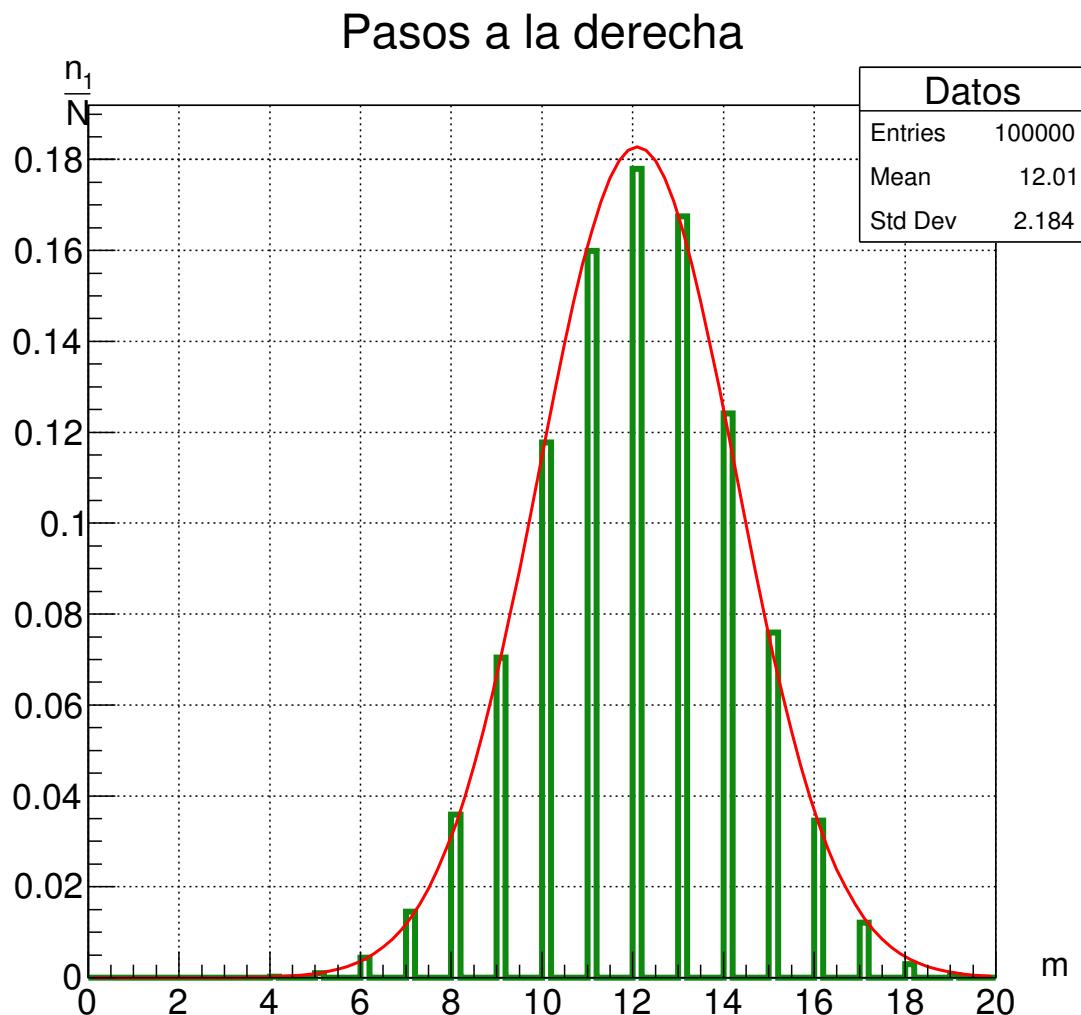
12.2.9 Ajuste de una Gaussiana de los pasos totales a la derecha con paso a la derecha de 0.6 de probabilidad

Figura 69: Ajuste de una Gaussiana de los pasos totales a la derecha con paso a la derecha de 0.6 de probabilidad

En los siguientes histogramas se muestra la distribución de la posición final cuando la probabilidad de dar un paso a la derecha es de 0.4 y la distribución de cuantos pasos en total se dieron a la derecha.

12.2.10 Histograma de posición final con paso a la derecha de 0.4 de probabilidad

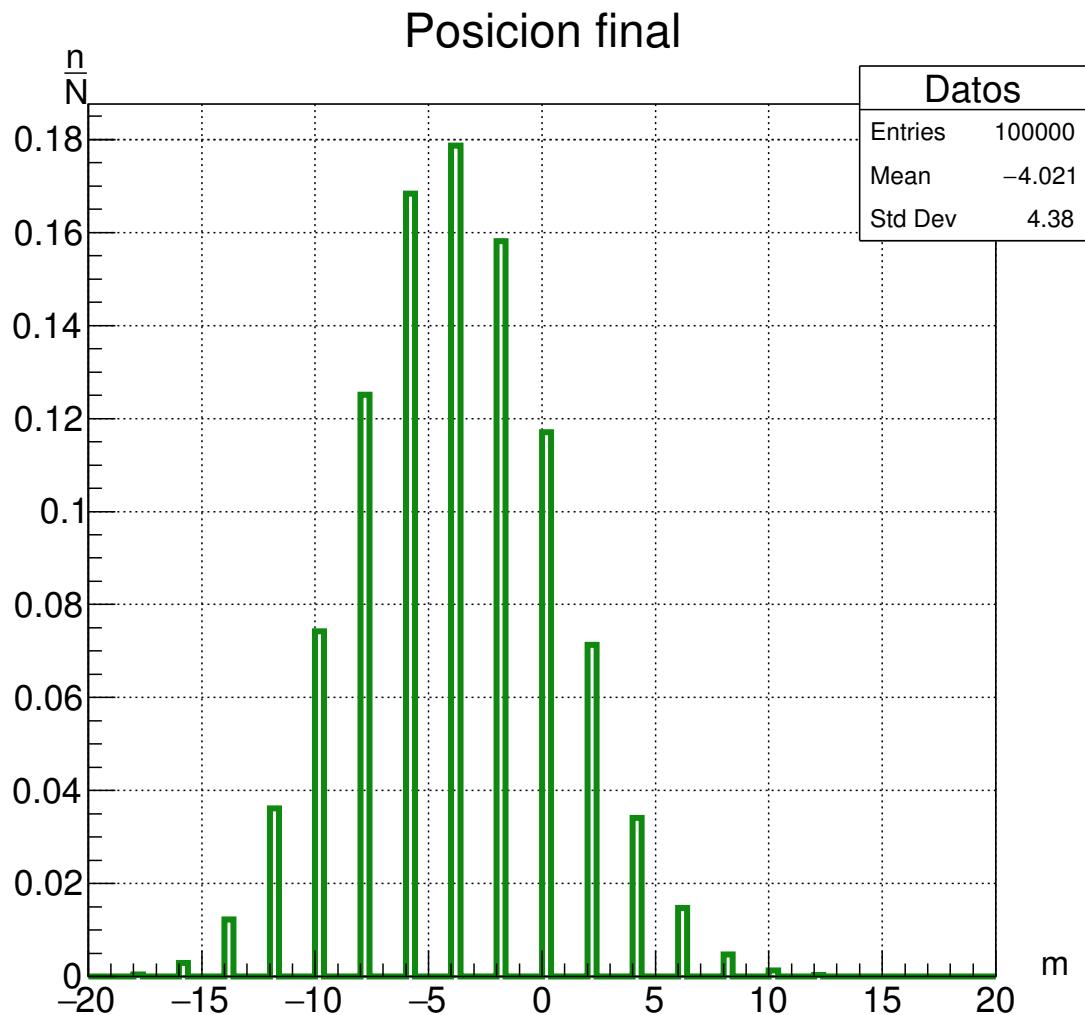


Figura 70: Posición final con paso a la derecha de 0.4 de probabilidad

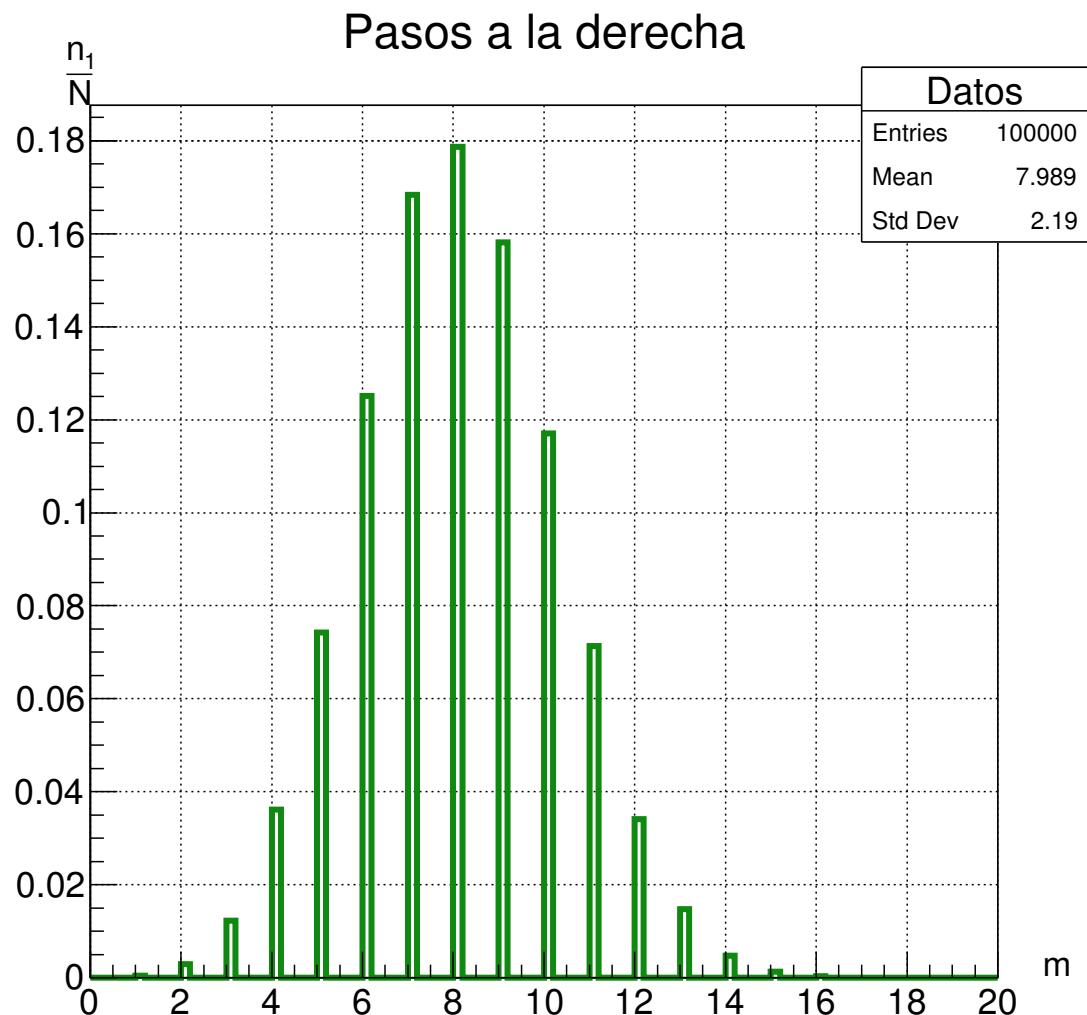
12.2.11 Histograma de los pasos totales a la derecha con paso a la derecha de 0.4 de probabilidad

Figura 71: Pasos totales a la derecha con paso a la derecha de 0.4 de probabilidad

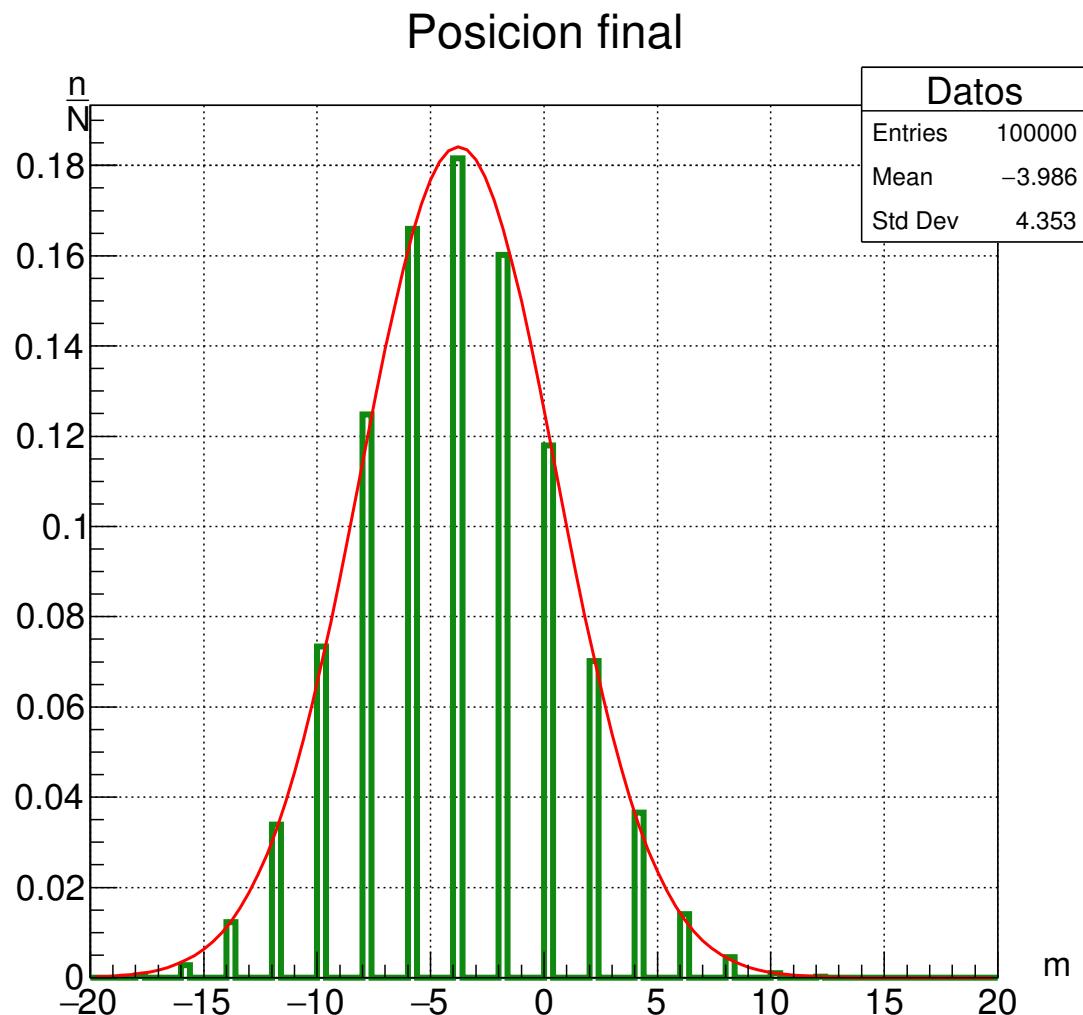
12.2.12 Ajuste de una Gaussiana de la posición final con paso a la derecha de 0.4 de probabilidad

Figura 72: Ajuste de una Gaussiana de la posición final con paso a la derecha de 0.4 de probabilidad

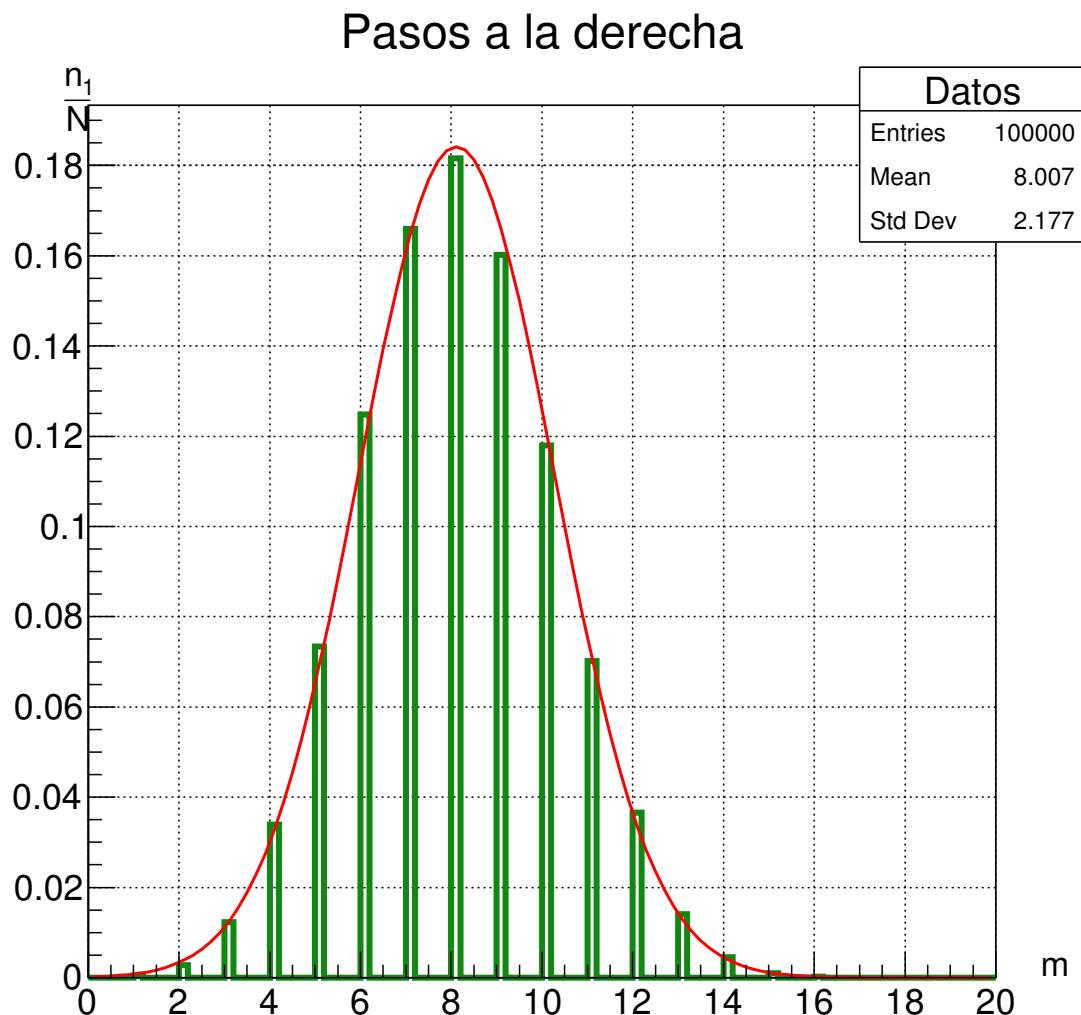
12.2.13 Ajuste de una Gaussiana de los pasos totales a la derecha con paso a la derecha de 0.4 de probabilidad

Figura 73: Ajuste de una Gaussiana de los pasos totales a la derecha con paso a la derecha de 0.4 de probabilidad

12.3 Demostraciones caminata aleatoria

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

$$\begin{aligned}
 \overline{(u - \bar{u})^2} &= \overline{u^2} - \bar{u}^2 \\
 &= \sum_{i=1}^M P(u_i)(u_i - \bar{u})^2 \\
 &= \sum_{i=1}^M P(u_i)(u_i^2 - 2u_i\bar{u} + \bar{u}^2) \\
 &= \sum_{i=1}^M P(u_i) \cdot u_i^2 - 2\bar{u} \sum_{i=1}^M P(u_i) \cdot u_i + \bar{u}^2 \\
 &= \sum_{i=1}^M P(u_i) \cdot u_i^2 - 2\bar{u}\bar{u} + \bar{u}^2 \\
 &= \overline{u^2} - \bar{u}^2
 \end{aligned}$$

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

$$\begin{aligned}
 \overline{n_1} &= Np. \\
 \sum_{i=1}^N n_1 p
 \end{aligned}$$

Ecuación 33

$$\begin{aligned}
 \overline{n_2} &= Nq. \\
 \sum_{i=1}^N n_2 q
 \end{aligned}$$

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

$$\begin{aligned}
 n_1^2 p^{n_1} &= n_1 p \frac{\partial}{\partial p} (p^{n_1}) = (p \frac{\partial}{\partial p})^2 (p^{n_1}) \\
 &= p^2 \frac{\partial^2}{\partial p^2} (p^{n_1}) \\
 &= p^2 \frac{\partial}{\partial p} (n_1 p^{n_1 - 1}) \\
 &= n_1 \cdot p^2 \frac{\partial}{\partial p} (p^{n_1 - 1}) \\
 &= n_1 \cdot p^2 ((n_1 - 1)p^{n_1 - 2}) \\
 &= n_1(n_1 - 1)(p^{n_1 - 2 + 2}) \\
 &= n_1(n_1 - 1)(p^{n_1})
 \end{aligned}$$

12.4 Simulación del movimiento Browniano

El movimiento browniano es el movimiento aleatorio que se observa en algunas partículas microscópicas que se hallan en un medio fluido (por ejemplo, polen en una gota de agua). Recibe su nombre en honor al escocés Robert Brown, biólogo y botánico que descubrió este fenómeno en 1827 y observó que pequeñas partículas de polen se desplazaban en movimientos aleatorios sin razón aparente. En 1785, el mismo fenómeno había sido descrito por Jan Ingenhousz sobre partículas de carbón en alcohol.

El movimiento estocástico de estas partículas se debe a que su superficie es bombardeada incesantemente por las moléculas (átomos) del fluido sometidas a una agitación térmica.

Este bombardeo a escala atómica no es siempre completamente uniforme y sufre variaciones estadísticas importantes. Así, la presión ejercida sobre los lados puede variar ligeramente con el tiempo, y así se genera el movimiento observado.

Tanto la difusión como la ósmosis se basan en el movimiento browniano.

La descripción matemática del fenómeno fue elaborada por Albert Einstein y constituye el primero de sus artículos del que, en la obra de Einstein, se considera el *Annus Mirabilis* ("año maravilloso", en latín), 1905. La teoría de Einstein demostraba la teoría atómica, todavía en disputa a principios del siglo XX, e iniciaba el campo de la física estadística.

12.4.1 Código en R para simulación en 2D

```
N<-1000;  
xdis = rnorm(N, 0 ,1);  
ydis = rnorm(N, 0 ,1);  
xdis = cumsum(xdis);  
ydis = cumsum(ydis);  
plot(xdis, ydis, type="l", col="red"  
      ,main ="Simulacion_de_un_movimiento_Browniano"  
      , xlab="desplazamiento_en_x"  
      , ylab = "desplazamiento_en_y");
```

12.4.2 Gráfica del movimiento Browniano en dos dimensiones

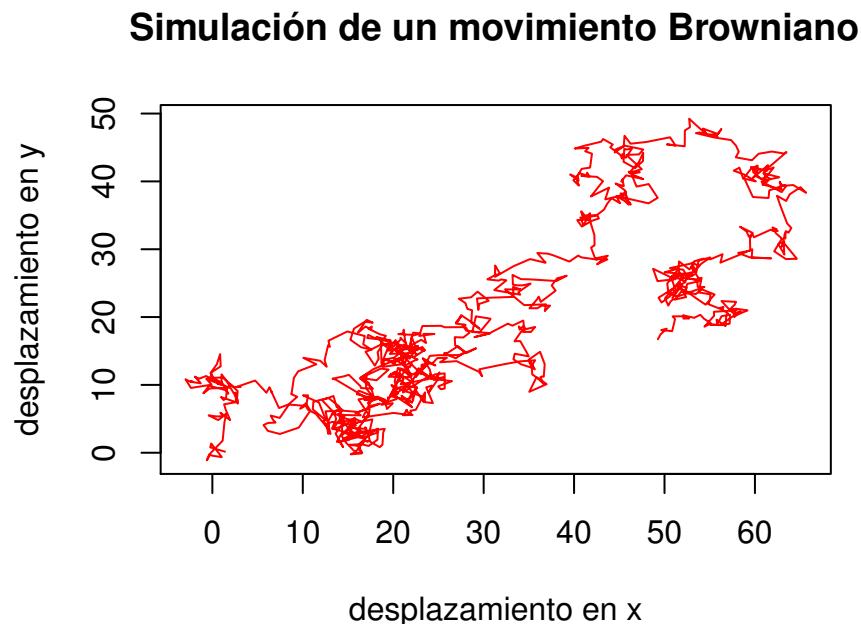


Figura 74: Simulación de un movimiento Browniano en dos dimensiones

12.4.3 Código en R para simulación en 3D

```
library(scatterplot3d)
N<-100000;
xdis = rnorm(N, 0 ,1);
ydis = rnorm(N, 0 ,1);
zdis = rnorm(N, 0 ,1);
xdis = cumsum(xdis);
ydis = cumsum(ydis);
zdis = cumsum(zdis);
scatterplot3d(xdis, ydis, zdis, angle=10
              , color ="red"
              , main ="Simulaci n _de _un _movimiento _Browniano"
              , xlab="desplazamiento _en _x"
              , ylab="desplazamiento _en _y"
              , zlab="desplazamiento _en _z"
              , type="l"
              , grid = FALSE
              , lab.z=3);
```

12.4.4 Gráfica del movimiento Browniano en tres dimensiones

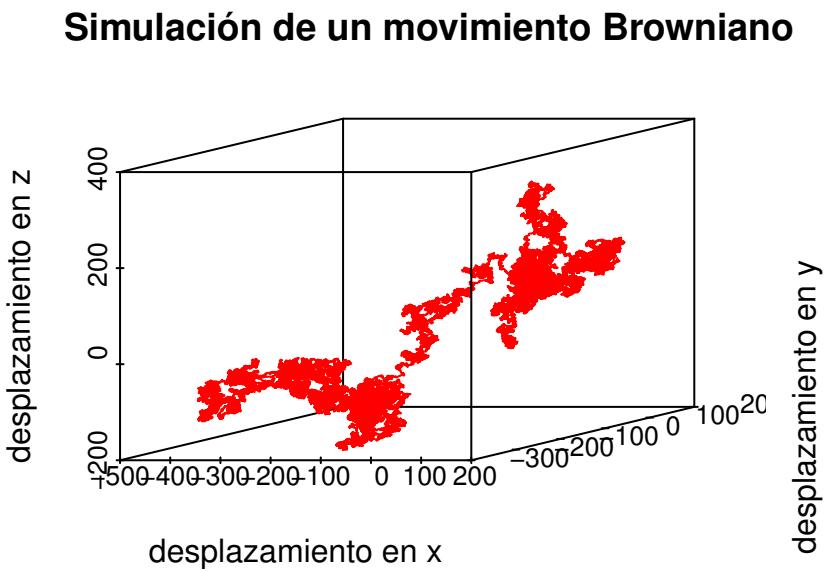


Figura 75: Simulación de un movimiento Browniano en tres dimensiones

Capítulo 13

13.1 Relación de función Gamma en factorial

En matemáticas, la función gamma (denotada como $\Gamma(z)$, donde Γ es la escritura en mayúscula de la letra gamma del alfabeto griego) 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:

$$\Gamma(z) = \int_0^\infty 0t^{z-1}e^{-t}dt \quad (70)$$

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:

$$\Gamma(n) = (n - 1)! \quad (71)$$

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.

13.2 Función Delta de Dirac

En la práctica conviene trabajar con otro tipo de impulso unitario, con una "función" definida con el límite [16],

$$\delta(t - t_0) = \lim_{\epsilon \rightarrow 0} \delta_\epsilon(t - t_0) \quad (72)$$

que se puede caracterizar mediante las dos propiedades siguientes:

$$\delta(t - t_0) = \begin{cases} 0 & t \neq t_0 \\ \infty & t = t_0 \end{cases} \quad (73)$$

$$1 = \int_0^\infty \delta(t - t_0) dt \quad (74)$$

A este impulso unitario se le denomina función **Delta de Dirac**.

Las propiedades (73) y (74) dejan claro que la función delta de Dirac no es una función en el sentido rigurosamente matemático, puesto que la integral (74) (si existe) de una función que es nula en cualquier punto excepto en uno de ellos debe anularse.

Como ya hemos visto antes, podemos desarrollar cualquier vector de estado $|\psi(x)\rangle$ en términos de un conjunto base ortonormal $\{|\phi_n(x)\rangle\}$ mediante,

$$|\psi(x)\rangle = \sum_{n=-\infty}^{\infty} a_n |\phi_n(x)\rangle \quad (75)$$

A los a_n se les denominan **coeficientes del desarrollo**. Aquí se debe cumplir que,

$$\langle \phi_n | \phi_m \rangle = \int_{-\infty}^{\infty} \phi_n^*(x) \phi_m(x) dx = \delta_{nm} \quad (76)$$

que es la condición de ortonormalidad para los elementos del conjunto base.

Es posible calcular los a_n . Al multiplicar ambos miembros (75) por $\langle \phi_m(x)|$ resulta,

$$\langle \phi_m(x) | \psi(x) \rangle = \sum_{n=-\infty}^{\infty} a_n \langle \phi_m(x) | \phi_n(x) \rangle = \sum_{n=-\infty}^{\infty} a_n \delta_{nm} = a_m \quad (77)$$

que al sustituirla en (75) resulta,

$$|\psi(x)\rangle = \sum_{n=-\infty}^{\infty} \langle \phi_n(x') | \psi(x') \rangle |\phi_n(x)\rangle \quad (78)$$

que se denominan **series generalizadas de Fourier**.

Ahora bien, la anterior expresión puede ser escrita como,

$$\begin{aligned} |\psi(x)\rangle &= \sum_{n=-\infty}^{\infty} |\psi(x')\rangle \langle \phi_n(x') | \phi_n(x) \rangle \\ &= \sum_{n=-\infty}^{\infty} |\psi(x')\rangle \int_{-\infty}^{\infty} \phi_n^*(x') \phi_n(x) dx' \\ &= \int_{-\infty}^{\infty} |\psi(x')\rangle \left[\sum_{n=-\infty}^{\infty} \phi_n^*(x') \phi_n(x) \right] dx' \end{aligned}$$

o,

$$|\psi(x)\rangle = \int_{-\infty}^{\infty} |\psi(x')\rangle \delta(x - x') dx' \quad (79)$$

donde,

$$\delta(x - x') = \sum_{n=-\infty}^{\infty} \phi_n^*(x') \phi_n(x) \quad (80)$$

que no es más que la función delta de Dirac unidimensional y su propiedad más importante es la expresada en (79).

Es posible encontrar otras formas de expresar la función delta de Dirac. Considérese el caso de las funciones armónicas,

$$\phi_n(x) = \frac{1}{(2l)^{\frac{1}{2}}} e^{in\pi \frac{x}{l}} \quad (81)$$

Sustituyendo (81) en (80) resulta,

$$\begin{aligned} \delta(x - x') &= \sum_{n=-\infty}^{\infty} \frac{1}{(2l)^{\frac{1}{2}}} e^{-in\pi \frac{x'}{l}} \frac{1}{(2l)^{\frac{1}{2}}} e^{in\pi \frac{x}{l}} \\ &= \frac{1}{2l} \sum_{n=-\infty}^{\infty} e^{in\pi \frac{(x-x')}{l}} \\ &= \frac{1}{2l} \sum_{n=-\infty}^{\infty} \Delta k e^{ik(x-x')} \end{aligned} \quad (82)$$

donde $k = n\pi/l$ y $\frac{1}{2l} = \frac{\Delta k}{2\pi}$ puesto que $\Delta n = 1$. Si $l \rightarrow \infty$ ($\Delta k \rightarrow 0$) en (82) la suma cambia por una integral,

$$\delta(x - x') = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(x-x')} dk \quad (83)$$

o,

$$\delta(x - x') = \frac{1}{2\pi} \lim_{l \rightarrow \infty} \int_{-l}^l e^{ik(x-x')} dk = \lim_{l \rightarrow \infty} \frac{\sin[l(x-x')]}{\pi(x-x')} \quad (84)$$

Las dos ecuaciones anteriores son dos representaciones ampliamente usadas para la función delat de Dirac. En tres dimensiones se escribe,

$$\delta(\vec{r} - \vec{r}') = \delta(x - x')\delta(y - y')\delta(z - z') = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} e^{i\vec{k}(\vec{r}-\vec{r}')} d^3k \quad (85)$$

donde,

$$\delta(\vec{r}) = 0, \quad \vec{r} \neq \vec{0} \quad (86)$$

y,

$$\int_{-\infty}^{\infty} \delta(\vec{r} - \vec{r}') d^3r = 1 \quad (87)$$

La delta de Dirac es par,

$$\delta(\vec{r} - \vec{r}') = \delta(\vec{r}' - \vec{r}) \quad (88)$$

Capítulo 14

14.1 Fork

Una bifurcación (fork en inglés), en el ámbito del desarrollo de software, es la creación de un proyecto en una dirección distinta de la principal u oficial tomando el código fuente del proyecto ya existente. Comúnmente se utiliza el término inglés. Como resultado de la bifurcación se pueden llegar a generar proyectos diferentes que cubren necesidades distintas aunque similares. El término también puede ser usado para representar la ramificación de cualquier trabajo. Un ejemplo de bifurcación es la Enciclopedia Libre escindida de la Wikipedia en español.

Las bifurcaciones de proyectos de software libre surgen de un cisma en los objetivos o un choque de personalidades. En una bifurcación, ambos lados asumen derechos de autor idénticos pero típicamente solo el grupo de mayor tamaño, o el que contiene al arquitecto original, retendrá el nombre original completo y la comunidad de usuarios asociada. Por ello existe una penalización asociada con la bifurcación.

En proyectos de software privativo las bifurcaciones usualmente se llevan a cabo para crear versiones de un mismo programa que funcionen en diversos sistemas operativos, aumentando así la cuota de mercado. Pero suelen retener las mismas características, facilitando así a los usuarios el uso del mismo programa en diferentes sistemas operativos.

14.2 NP (clase de complejidad)

En teoría de la complejidad computacional, NP es el acrónimo en inglés de nondeterministic polynomial time ("tiempo polinomial no determinista"). Es el conjunto de problemas que pueden ser resueltos en tiempo polinómico por una máquina de Turing no determinista.

La clase NP

La importancia de esta clase de problemas de decisión es que contiene muchos problemas de búsqueda y de optimización para los que se desea saber si existe una cierta solución o si existe una mejor solución que las conocidas. En esta clase están el problema del viajante (también llamado "problema del viajante de comercio" o "problema del agente viajero") donde se quiere saber si existe una ruta óptima que pasa por todos los nodos en un cierto grafo y el problema de satisfacibilidad booleana en donde se desea saber si una cierta fórmula de lógica proposicional puede ser cierta para algún conjunto de valores booleanos para las variables.

Dada su importancia, se han hecho muchos esfuerzos para encontrar algoritmos

que decidan algún problema de NP en tiempo polinómico. Sin embargo, pareciera que para algunos problemas de NP (los del conjunto NP-completo) no es posible encontrar un algoritmo mejor que simplemente realizar una búsqueda exhaustiva.

En el artículo de 2002, "PRIMES is in P", Manindra Agrawal con sus estudiantes, encontró un algoritmo que trabaja en tiempo polinómico para el problema de saber si un número es primo. Anteriormente se sabía que ese problema estaba en NP, si bien no en NP-completo, ahora se sabe que también está en P.

El primer problema natural que se demostró que es completo NP fue el problema de satisfacibilidad booleana. Este resultado fue demostrado por Stephen Cook en 1971, y se lo llamó el teorema de Cook. La demostración de Cook de que la satisfacibilidad es un problema NP-completo es muy complicada. Sin embargo, después de que este problema se demostrara que es NP-Completo, es fácil demostrar que muchos otros problemas pertenecen a esta clase. Por lo tanto, una amplia clase de problemas en principio inconexos son reducibles unos a otros, y por lo tanto resultan en "el mismo problema" – un resultado profundo e inesperado.

NP contiene todos los problemas pertenecientes a las clases P y NP-C, y a su vez está contenido en el conjunto de los PSPACE. Aún se desconoce si estas inclusiones son estrictas o no, y si la intersección entre los NP y Co-NP es o no vacía.

En particular, el mayor problema en ciencias de la computación consiste en responder al siguiente problema de decisión: $\text{P} = \text{NP}$?

14.3 Reseña capítulo 2 del hombre anumérico

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 ironías 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. Aquí unos ejemplos: Cristóbal Colón descubrió el Nuevo Mundo en 1492 y su compatriota Enrico Fermi descubrió el nuevo mundo del átomo en 1942. La razón entre las alturas de los edificios Sears de Chicago y Woolworth de Nueva York coincide en lo que respecta a las cuatro primeras cifras (1,816 frente a 1.816) con la razón entre las masas del protón y el electrón. Reagan y Gorbachov firmaron el tratado INF el 8 de diciembre de 1987, exactamente siete años después de que John Lennon fuera asesinado [13].

La moraleja vuelve a ser que mientras es probable que ocurra algún hecho improbable, lo es mucho menos que se dé un caso concreto. La conclusión paradójica es que sería muy improbable que los casos improbables no ocurrieran. Si no se concreta con precisión cuál es el acontecimiento a predecir, puede ocurrir un suceso de tipo genérico de muchísimas maneras distintas. Dos extraños, procedentes de puntos op-

uestos 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.

El psicólogo Stanley Milgram 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. Milgram 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 Milgram es más impresionante. Aporta bastante a la explicación de cómo las informaciones confidenciales, los rumores y los chistes corren tan rápidamente entre cierta población.

Otro problema de probabilidad sirve para ilustrar lo corrientes que pueden llegar a ser las coincidencias en otro contexto. El problema se formula a menudo como sigue: si barajamos mil sobres con las direcciones escritas en ellos y mil cartas con las mismas direcciones también, y luego metemos cada carta en un sobre, la probabilidad de que por lo menos una carta vaya en el sobre que le corresponde es también del 63 por ciento. O bien tómense dos mazos de cartas completamente barajadas y puestas boca abajo. Si vamos destapando las cartas de dos en dos, una de cada mazo, ¿cuál es la probabilidad de que el par de cartas coincida por lo menos una vez? El 63 por ciento también.

El ejemplo del cartero que ha de distribuir veintiuna cartas entre veinte buzones nos permitirá ilustrar un principio numérico que a veces sirve para explicar la certeza de un determinado tipo de coincidencias. Como 21 es mayor que 20, puede estar seguro, sin necesidad de mirar previamente las direcciones, que por lo menos uno de los buzones tendrá más de una carta. Este principio de sentido común, que se conoce a veces como principio del casillero o de los cajones de Dirichlet, puede servir a veces para llegar a conclusiones que no son tan obvias. 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 predisposto a creer.

Hay una tendencia general muy fuerte a olvidar los fracasos y concentrarse en los éxitos y los aciertos. Los casinos abonan esta tendencia haciendo que cada vez que alguien gana un cuarto de dólar en una máquina tragaperras, parpadeen las lucecitas y la moneda tintinee en la bandeja de metal. Con tanta lucecita y tant tintineo, no es difícil llegar a creer que todo el mundo está ganando. Las pérdidas y los fracasos son silenciosos. Lo mismo vale para los tan cacareados éxitos financieros frente a los que se arruinan de manera relativamente silenciosa jugando a la bolsa.

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.

Entender bien el significado del valor esperado es útil en el análisis de la mayoría de juegos de casino, así como del no tan conocido juego del chuck-a-luck, que se juega en los carnavales del Medio Oeste e Inglaterra. La explicación del chuck-a-luck que se da para atraer a la gente puede ser muy persuasiva. El que apuesta elige un número de 1 a 6 y el encargado lanza tres dados. Si el número elegido sale en los tres dados, el jugador cobra 3 dólares; si sale en dos de los dados, cobra 2 dólares y si sale en uno de los tres dados, sólo cobra 1 dólar. únicamente en el caso de que el número escogido no salga en ninguno de los dados tendrá que pagar el jugador, y sólo 1 dólar. Con tres dados distintos, el apostador tiene tres posibilidades a su favor; además, a veces gana más de 1 dólar, que es lo máximo que puede perder cada vez. Como diría Joan Rivers: ¿Podemos calcularlo?. (Si no tienes muchas ganas de calcular, sáltate lo que queda hasta el final de la sección).

Está claro que la probabilidad de ganar es independiente del número escogido. Así pues, para concretar, supongamos que el jugador elige siempre el número 4. Como los dados son independientes, la probabilidad de que salga 4 en los tres dados es $1/6 \cdot 1/6 \cdot 1/6 = 1/216$. Por tanto, aproximadamente 1/216 de las veces el jugador ganará 3 dólares. La probabilidad de que salga 4 en dos de los dados es un poco más difícil de calcular, a no ser que se use la distribución binomial de probabilidad de la que hablamos en el Capítulo 1, y que volveré a deducir en el contexto que nos ocupa. Que salga un 4 en dos de los tres dados puede ocurrir de tres maneras distintas y mutuamente excluyentes: X44, 4X4 ó 44X, donde la X significa no 4. La probabilidad del primero es $5/6 \cdot 1/6 \cdot 1/6 = 5/216$. El mismo resultado vale para los otros dos modos restantes. La suma, $15/216$, nos da la probabilidad de que salga 4 en dos de los tres dados, la cual nos da a su vez la probabilidad de que el apostador gane 2 dólares. La probabilidad de sacar un 4 entre los tres dados se calcula de modo análogo, descomponiendo el suceso en los tres modos mutuamente excluyentes en los que este puede ocurrir. La probabilidad de que salga 4XX es $1/6 \cdot 5/6 \cdot 5/6 = 25/216$, y esta es también la probabilidad de que salga X4X ó XX4. Sumándolas nos da $75/216$ como probabilidad de sacar exactamente un 4 entre los tres dados, esto es, la probabilidad de ganar 1 dólar. Para hallar la probabilidad de que al tirar los dados no salga ningún cuatro, buscamos cuánta probabilidad queda. Es decir, restamos $(1/216 + 15/216 + 75/216)$ de 1 (ó 100%), y obtenemos $125/216$. Por tanto, de cada 216 jugadas al chuck-a-luck, el jugador pierde 1 dólar en 125 de ellas. El valor esperado de las ganancias es pues

$$\left(3 \times \frac{1}{216}\right) + \left(2 \times \frac{15}{216}\right) + \left(1 \times \frac{75}{216}\right) + \left(-1 \times \frac{125}{216}\right) = \left(-\frac{17}{216}\right)$$

= -0.08 dolares

con lo que, en promedio, el jugador pierde ocho centavos en cada jugada de ese juego tan prometedor. 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, quizás las probabilidades de éxito sean mayores.

Nuestro modelo supone que nuestra protagonista a la que llamaremos María tiene buenas razones para pensar que se encontrará con N potenciales cónyuges mientras esté en edad núbil. Para algunas mujeres N pueden ser dos, y para otras, doscientos. La pregunta que se plantea María es: ¿Cuándo habría de aceptar al señor X y renunciar a los otros pretendientes que vinieran después, aunque alguno de estos quizás fuera mejor que él? Supondremos que los va conociendo de uno en uno, valora la conveniencia relativa de cada uno de ellos y que, una vez que ha rechazado a uno, lo pierde para siempre. Para concretar más, supongamos que María ha conocido ya a seis hombres y que los ha clasificado así: 3 5 1 6 2 4. Es decir, de los seis hombres, el primero que conoció ocupa el tercer lugar en el orden de preferencia, el segundo en aparecer ocupa el quinto lugar, prefiere el tercero a todos los demás, etc. Si ahora resulta que el séptimo de los hombres que conoce es mejor que todos los demás excepto su favorito, modificará así la clasificación: 4 6 1 7 3 5 2. Después de cada hombre, María reordena la clasificación relativa de sus pretendientes y se

pregunta qué regla habría de seguir para maximizar la probabilidad de escoger al mejor de los N pretendientes que espera tener. En la obtención del mejor sistema se emplea la idea de probabilidad condicional (que presentaremos en el próximo capítulo) y también hay que calcular un poco. El sistema en sí, no obstante, se describe muy fácilmente. Diremos que un pretendiente es un novio si es mejor que todos los candidatos anteriores. María debería rechazar aproximadamente el primer 37% de los candidatos que probablemente vaya a conocer y luego aceptar al primer novio que le salga de entre los pretendientes posteriores (si es que le sale alguno, claro). 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 habría 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 habría 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 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.

La siguiente anomalía deportiva tiene también implicaciones legales. Consideremos dos jugadores de béisbol, Babe Ruth y Lou Gehrig, pongamos por caso. Durante la primera mitad de la temporada, Babe Ruth tiene en el bateo una media de aciertos mayor que Lou Gehrig. Y en la segunda mitad de la temporada vuelve a ocurrir lo mismo. Pero considerando la temporada entera, ocurre que el promedio de aciertos de Lou Gehrig es mejor que el de Babe Ruth. ¿Puede ser cierto? A primera vista parece como si tal situación fuera totalmente imposible, aunque el mero hecho de haber planteado la pregunta pueda de por sí despertar algunas dudas. Lo que podría haber ocurrido es que durante la primera mitad de la temporada Babe Ruth tuviera una media de aciertos de 0,300 y Lou Gehrig de sólo 0,290, pero que Ruth hubiera bateado doscientas veces y Gehrig sólo cien. Mientras que en la segunda mitad de la temporada las medias de aciertos fueran 0,400 para Ruth y sólo 0,390 para Gehrig, pero que Ruth hubiera salido a batear sólo cien veces y Gehrig, doscientas. El resultado global para toda la temporada sería un promedio de aciertos de 0,357 de Gehrig frente a 0,333 de Ruth. La moraleja es que no se pueden sacar promedios de promedios.

Hay un sinfín de otras rarezas semejantes que se derivan de nuestros modos convencionales de medir, expresar y comparar cantidades periódicas, tanto si se trata del cash flow de un gobierno como de las fluctuaciones diarias de la temperatura corporal. Imaginemos a dos jugadores, Pedro y Pablo, que juegan a cara o cruz, tirando una moneda al aire una vez por día. En un momento dado, diremos que Pedro va ganando si hasta aquel momento han salido más caras que cruces, y en caso contrario es Pablo quien va ganando. En cualquier momento, tanto Pedro como Pablo tienen la misma probabilidad de ir ganando, pero sea quien sea el que vaya ganando, este es el que tiene mayor probabilidad de haber estado ganando más rato. Si han tirado la moneda cien veces y acaba ganando Pedro ¡es considerablemente mayor la probabilidad de que este haya estado por delante más del 90 por ciento del tiempo, pongamos, que la de que lo haya estado entre el 45 y el 55 por ciento! Y análogamente, si acaba ganando Pablo, la probabilidad de que este haya estado ganando más del 96 por ciento del tiempo es mucho menor que la de que lo haya estado entre el 48 y el 52 por ciento. Quizás este resultado sea tan contrario a la intuición porque la mayoría de la gente suele pensar como si las desviaciones de la media estuvieran atadas a una banda elástica, de modo que, cuanto mayor fuera la desviación, mayor sería la fuerza recuperadora que tendiese a restaurar la media. La creencia errónea de que el hecho de que hayan salido varias caras seguidas hace más probable que la próxima vez salga cruz se conoce como sofisma del jugador (las mismas ideas valen para la ruleta y los dados).

Capítulo 15

15.1 Matrices semejantes

Dos matrices cuadradas A y B se dice que son semejantes si existe una matriz cuadrada regular P tal que:

$$B = P \cdot A \cdot P^{-1} \quad (89)$$

La semejanza de matrices cuadradas cumple las siguientes propiedades:

1.- La semejanza es una relación de equivalencia.

Comprobamos que cumple las tres propiedades que han de verificar las relaciones de equivalencia.

- Reflexiva. Está claro que toda matriz es semejante a sí misma.
- Simétrica. Si A y B son semejantes existe una matriz cuadrada regular P con:

$$B = PAP^{-1} \Rightarrow B = P^{-1}B(P^{-1})^{-1} \quad (90)$$

Y vemos que B y A son semejantes.

- Transitiva: Si A y B son semejantes por B y C también, entonces existen matrices cuadradas regulares P, P' :

$$\left. \begin{array}{l} B = PAP^{-1} \\ C = P'BP'^{-1} \end{array} \right\} \Rightarrow C = P'PBP^{-1} = P'PB(P'P)^{-1} \quad (91)$$

donde $P'P$ es una matriz cuadrada regular. Por tanto A y C son semejantes.

2. Dos matrices semejantes son equivalentes
3. Dos matrices semejantes tienen la misma dimensión
4. Dos matrices semejantes tienen el mismo determinante

Si $B = PAP^{-1}$

$$|B| = |PAP^{-1}| = |P||A||P|^{-1} = |A| \quad (92)$$

5. Dos matrices semejantes tienen el mismo rango

El recíproco no es cierto. Por ejemplo las siguientes matrices tienen rango 2 pero no son semejantes:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad y \qquad \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$

Ejemplo: Comprobar que las matrices A y B son semejantes considerando la matriz de paso Q .

$$A = \begin{pmatrix} 6 & 1 & -1 \\ 0 & 5 & 1 \\ 2 & -1 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 4 & 0 & 2 \\ 0 & 4 & 2 \\ 0 & 2 & 4 \end{pmatrix} \quad Q = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix}$$

Para poder comprobar que las matrices A y B son semejantes tomando como matriz de paso Q basta verificar que se cumple que $A = Q^{-1}BQ$. Para ello se plantea:

$$\begin{pmatrix} 6 & 1 & -1 \\ 0 & 5 & 1 \\ 2 & -1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} 4 & 0 & 2 \\ 0 & 4 & 2 \\ 0 & 2 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix}$$

$$\begin{pmatrix} 6 & 1 & -1 \\ 0 & 5 & 1 \\ 2 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 3 & 4 \\ -2 & 3 & 2 \\ 2 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix}$$

$$\begin{pmatrix} 6 & 1 & -1 \\ 0 & 5 & 1 \\ 2 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 6 & 1 & -1 \\ 0 & 5 & 1 \\ 2 & -1 & 1 \end{pmatrix}$$

De esta manera, se ha demostrado que las matrices A y B son semejantes, tomando como matriz de paso Q .

Para las matrices A y B del ejemplo anterior, que sabemos que son semejantes con matriz de paso Q , calculemos determinante, rango, traza y polinomios característicos:

$$A = \begin{pmatrix} 6 & 1 & -1 \\ 0 & 5 & 1 \\ 2 & -1 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 4 & 0 & 2 \\ 0 & 4 & 2 \\ 0 & 2 & 4 \end{pmatrix}$$

Calculamos sus determinantes:

$$\det A = \begin{vmatrix} 6 & 1 & -1 \\ 0 & 5 & 1 \\ 2 & -1 & 1 \end{vmatrix} = 6(5 - (-1)) - 1(0 - 2) + (-1)(0 - 10) = 36 + 2 + 10 = 48$$

$$\det B = \begin{vmatrix} 4 & 0 & 2 \\ 0 & 4 & 2 \\ 0 & 2 & 4 \end{vmatrix} = 4(16 - 4) - 0(0 - 0) + 2(0 - 0) = 4 \times 12 = 48$$

Sus determinantes coinciden.

Dado que los determinantes de A y B son distintos a cero, podemos decir que ambas matrices tienen rango 3, por lo tanto, coinciden sus rangos.

Calculamos las trazas:

$$\text{tr}(A) = 6 + 5 + 1 = 12$$

$$\text{tr}(B) = 4 + 4 + 4 = 12$$

Sus trazas coinciden.

Calculamos el polinomio característico:

$$\det(A - \lambda I) = \begin{vmatrix} 6 - \lambda & 1 & -1 \\ 0 & 5 - \lambda & 1 \\ 2 & -1 & 1 - \lambda \end{vmatrix} = \lambda^3 - 12\lambda^2 + 44\lambda - 48$$

$$\det(B - \lambda I) = \begin{vmatrix} 4 - \lambda & 0 & 2 \\ 0 & 4 - \lambda & 2 \\ 0 & 2 & 4 - \lambda \end{vmatrix} = \lambda^3 - 12\lambda^2 + 44\lambda - 48$$

Vemos que ambos polinomios característicos son iguales. Por tanto, ambas matrices tendrán los mismos autovalores y autovectores.

15.2 Problemas de eigen valores y vectores

15.2.1 Ejercicio 1

1) Encontrar los eigen valores y eigen vectores de:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}$$

Solución:

$$\det(A - \lambda I) = \det \left(\begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right) = \det \left(\begin{bmatrix} 1 - \lambda & 2 \\ 3 & -4 - \lambda \end{bmatrix} \right)$$

$$= (1 - \lambda)(-4 - \lambda) - (2 \times 3) = \lambda^2 - 4 - \lambda + 4\lambda - 6 = \lambda^2 + 3\lambda - 10$$

Por lo tanto:

$$\lambda_1 = -5 \quad \lambda_2 = 2$$

Sustituyendo λ_1 en la ecuación siguiente:

$$(\lambda I - A) \vec{v} = 0 \quad (93)$$

$$\begin{bmatrix} 1+5 & 2 \\ 3 & -4+5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Entonces:

$$\begin{bmatrix} 6 & 2 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\vec{v}_1 = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

Ahora sustituyendo λ_2 en la ecuación (93):

$$\begin{bmatrix} 1-2 & 2 \\ 3 & -4-2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Entonces:

$$\begin{bmatrix} -1 & 2 \\ 3 & -6 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\vec{v}_2 = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

15.2.2 Ejercicio 2

2) Encontrar los eigen valores y eigen vectores de:

$$A = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 6 & 1 & 0 \\ 0 & 0 & 6 & 1 \\ 0 & 6 & 0 & 6 \end{bmatrix}$$

Solución:

$$\det(\lambda I - A) = \left(\begin{bmatrix} \lambda - 4 & 0 & 0 & 0 \\ 0 & \lambda - 6 & -1 & 0 \\ 0 & 0 & \lambda - 6 & -1 \\ 0 & 6 & 0 & \lambda - 6 \end{bmatrix} \right)$$

El determinante está dado por el polinomio característico:

$$\det(\lambda I - A) = \lambda^4 - 22\lambda^3 + 180\lambda^2 + 648\lambda - 864 = 0$$

Por lo tanto:

$$\lambda_1 = 6 \quad \lambda_2 = 6 \quad \lambda_3 = 6 \quad \lambda_4 = 4$$

Si sustituimos $\lambda_1, \lambda_2, \lambda_3$ y λ_4 en $(\lambda I - A)$, entonces los vectores propios son:

$$\vec{v}_1 = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \end{bmatrix}$$

$$\vec{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\vec{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\vec{v}_4 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

15.2.3 Ejercicio 3

3) Encontrar los eigen valores y eigen vectores.

$$A = \begin{bmatrix} -1 & 0 & -3 \\ 0 & 2 & 0 \\ -3 & 0 & -1 \end{bmatrix}$$

Solución:

$$(A - \lambda I) = 0$$

$$\begin{aligned}
&= \left(\begin{bmatrix} -1 & 0 & -3 \\ 0 & 2 & 0 \\ -3 & 0 & -1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \\
&= \begin{bmatrix} -1 - \lambda & 0 & -3 \\ 0 & 2 - \lambda & 0 \\ -3 & 0 & -1 - \lambda \end{bmatrix} \quad F_1 = F_1 \cdot \left(\frac{1}{-1 - \lambda} \right) \\
&= \begin{bmatrix} 1 & 0 & \frac{-3}{-1 - \lambda} \\ 0 & 2 - \lambda & 0 \\ -3 & 0 & -1 - \lambda \end{bmatrix} \quad F_3 = F_3 + 3F_1
\end{aligned}$$

Se obtiene la matriz escalonada:

$$\begin{bmatrix} 1 & 0 & \frac{-3}{-1 - \lambda} \\ 0 & 2 - \lambda & 0 \\ 0 & 0 & \frac{-8 - 2\lambda + \lambda^2}{-1 - \lambda} \end{bmatrix} \quad F_3 = F_3 + 3F_1$$

Ahora se encuentra el determinante, multiplicando los elementos de la diagonal.

$$0 = 1 \cdot (2 - \lambda) \cdot \left(\frac{-8 - 2\lambda + \lambda^2}{-1 - \lambda} \right)$$

Factorizando:

$$0 = 1 \cdot (2 - \lambda) \cdot \left(\frac{(\lambda - 2)(\lambda + 4)}{-1 - \lambda} \right)$$

$$\begin{aligned}
(2 - \lambda) &= 0 & (\lambda - 2) &= 0 & (\lambda + 4) &= 0 \\
\lambda &= 2 & \lambda &= 2 & \lambda &= -4
\end{aligned}$$

Por lo tanto los eigen valores son:

$$\boxed{\lambda_1 = 2 \quad \lambda_2 = 4}$$

Ahora calculamos los eigen vectores utilizando la matriz:

$$\begin{bmatrix} -1 - \lambda & 0 & -3 \\ 0 & 2 - \lambda & 0 \\ -3 & 0 & -1 - \lambda \end{bmatrix}$$

Sustituimos los eigen valores obtenidos previamente:

- $\lambda_1 = 2$

$$\begin{bmatrix} -1-2 & 0 & -3 \\ 0 & 2-2 & 0 \\ -3 & 0 & -1-2 \end{bmatrix}$$

$$\begin{bmatrix} -3 & 0 & -3 \\ 0 & 0 & 0 \\ -3 & 0 & -3 \end{bmatrix} \quad F_3 = F_3 - F_1$$

Dado que ya obtuvimos una sola fila con valores podemos obtener los valores de x , y y z . La primer fila corresponde al valor de x , la segunda al de y y la ultima al de z .

$$\begin{bmatrix} -3 & 0 & -3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{aligned} -3x - 3z &= 0 \\ -3x &= 3z \\ -x &= z \end{aligned}$$

Por lo tanto:

$$v_\lambda = \{(x, y, z) | -x = z\}$$

$$v_\lambda = \{(x, y, -x) | x \wedge y \in \mathbb{R}\}$$

Descomponemos el vector anterior:

$$v_\lambda = \{(x, 0, -x) + (0, y, 0) | x \wedge y \in \mathbb{R}\}$$

Sacamos el factor común:

$$v_\lambda = \{x(1, 0, -1) + y(0, 1, 0) | x \wedge y \in \mathbb{R}\}$$

Y así obtenemos los eigen vectores de este eigen valor:

$$v_{p1} = (1, 0, -1)$$

$$v_{p2} = (0, 1, 0)$$

- $\lambda_2 = -4$

$$\begin{pmatrix} -1 - (-4) & 0 & -3 \\ 0 & 2 - (-4) & 0 \\ -3 & 0 & -1 - (-4) \end{pmatrix}$$

$$\begin{pmatrix} 3 & 0 & -3 \\ 0 & 6 & 0 \\ -3 & 0 & 3 \end{pmatrix} \quad F_3 = F_3 - F_1$$

$$\begin{pmatrix} -3 & 0 & -3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\begin{array}{ll} 3x - 3z = 0 & 6y = 0 \\ 3x = 3z & y = 0 \\ x = z & \end{array}$$

Por lo tanto:

$$\begin{aligned} v_\lambda &= \{(x, y, z) | x = z \wedge y = 0\} \\ v_\lambda &= \{(z, 0, z) | z \in \mathbb{R}\} \end{aligned}$$

Sacamos el factor común:

$$v_\lambda = \{z(1, 0, 1) | z \in \mathbb{R}\}$$

Y así obtenemos el eigen vector de este eigen valor:

$$v_{p3} = (1, 0, 1)$$

15.2.4 Ejercicio 4

4) Encontrar los eigen valores y eigen vectores de:

$$A = \begin{bmatrix} 5 & -3 \\ 3 & -1 \end{bmatrix}$$

Solución:

$$\det(A - \lambda I) = \det \left(\begin{bmatrix} 5 & -3 \\ 3 & -1 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right) = \det \left(\begin{bmatrix} 5 - \lambda & -3 \\ 3 & -1 - \lambda \end{bmatrix} \right)$$

$$= (5 - \lambda)(-1 - \lambda) - (-3 \times 3) = \lambda^2 - 4\lambda - 5 + 9 = \lambda^2 - 4\lambda + 4$$

Por lo tanto:

$$\lambda_1 = 2 \quad \lambda_2 = 2$$

Sustituyendo λ_1 o λ_2 en la ecuación (93):

$$\begin{bmatrix} 5 - 2 & -3 \\ 3 & -1 - 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Entonces:

$$\begin{bmatrix} 3 & -3 \\ 3 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\vec{v}_1 = \vec{v}_2 = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

15.2.5 Ejercicio 5

5) Encontrar los eigen valores y eigen vectores de:

$$A = \begin{bmatrix} 2 & -2 & -3 \\ 10 & -4 & 5 \\ 5 & -4 & 6 \end{bmatrix}$$

Solución:

$$\det(\lambda I - A) = \left(\begin{bmatrix} \lambda - 2 & 2 & 3 \\ -10 & \lambda + 4 & -5 \\ -5 & 4 & \lambda - 6 \end{bmatrix} \right)$$

Por lo que el determinante de la matriz está dado por el polinomio característico:

$$\det(\lambda I - A) = \lambda^3 - 4\lambda^2 + 5\lambda - 2 = 0$$

Por lo tanto:

$$\lambda_1 = 1 \quad \lambda_2 = 1 \quad \lambda_3 = 2$$

Si sustituimos λ_1 y λ_3 en $(\lambda I - A)$, entonces los eigen vectores son:

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 5 \\ 3 \end{bmatrix}$$

$$\vec{v}_2 = \begin{bmatrix} 4 \\ 15 \\ 10 \end{bmatrix}$$

15.3 Matriz poco densa

Una matriz poco densa es aquella que posee una proporción alta de ceros entre sus elementos.

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 \\ 0 & 3 & 0 & 2 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 9 & 0 & 0 & 7 & 0 & 0 \end{pmatrix}$$

Las matrices cuadradas poco densas son aquellas que, además de cumplir con la condición de tener una alta proporción de ceros entre sus elementos, tienen igual número de filas que de columnas.

Las matrices cuadradas poco densas tienen ciertas subdivisiones, para iniciar las triangulares:

- Triangular Superior
- Triangular Inferior
- Tridiagonal

Una matriz es Triangular Superior cuando sus elementos distintos de cero se encuentran almacenados sobre la diagonal principal.

$$\begin{pmatrix} 5 & 2 & 4 & 8 & 3 & 2 \\ 0 & 5 & 3 & 7 & 9 & 6 \\ 0 & 0 & 5 & 1 & 6 & 8 \\ 0 & 0 & 0 & 5 & 10 & 7 \\ 0 & 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{pmatrix}$$

Por su parte, una matriz es Triangular Inferior cuando sus elementos distintos de cero se encuentran almacenados debajo de la diagonal principal.

$$\begin{pmatrix} 5 & 0 & 0 & 0 & 0 & 0 \\ 4 & 5 & 0 & 0 & 0 & 0 \\ 1 & 3 & 5 & 0 & 0 & 0 \\ 7 & 8 & 0 & 5 & 0 & 0 \\ 2 & 4 & 5 & 7 & 5 & 0 \\ 9 & 3 & 8 & 7 & 3 & 5 \end{pmatrix}$$

Y finalmente, la matriz Tridiagonal es aquella cuyos elementos distintos de cero se encuentran almacenados en la diagonal principal y en las diagonales contiguas encima y debajo de ésta.

$$\begin{pmatrix} 5 & 2 & 0 & 0 & 0 & 0 \\ 9 & 5 & 3 & 0 & 0 & 0 \\ 0 & 8 & 5 & 1 & 0 & 0 \\ 0 & 0 & 2 & 5 & 7 & 0 \\ 0 & 0 & 0 & 7 & 5 & 6 \\ 0 & 0 & 0 & 0 & 3 & 5 \end{pmatrix}$$

15.4 Problemas de cadena de Markov

15.4.1 Ejercicio 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?

Solución:

Los estados del sistema son S (de estudiar) y T (de no estudiar). La matriz de transición es:

$$P = \begin{matrix} S & T \\ \begin{matrix} S \\ T \end{matrix} & \begin{pmatrix} 0.3 & 0.7 \\ 0.4 & 0.6 \end{pmatrix} \end{matrix}$$

Para averiguar qué sucede a la larga, tenemos que hallar el vector de probabilidad fijo único T de P . $u = (0, 4, 0, 7)$ es un punto fijo de P y así $t = \left(\frac{4}{11}, \frac{7}{11}\right)$ es el vector de probabilidad buscado. Así que a la larga el estudiante estudia $\frac{4}{11}$ de las veces.

15.4.2 Ejercicio 2

Un psicólogo hace los supuestos siguientes que conciernen el comportamiento de ratas sujetas a un régimen especial de alimentación. Para una prueba particular, 80% de las ratas que fueron para la derecha en el experimento previo hicieron lo mismo en esta prueba, y 60% de aquellas que fueron para la izquierda en el experimento previo, fueron para la derecha en esta prueba. Si 50% van a la derecha en la primera prueba, ¿qué se podría predecir para, a) la segunda prueba? b) la tercera prueba? c) la milésima prueba?

Solución:

Los estados del sistema son R (derecha) y L (izquierda). La matriz de transición es:

$$P = \begin{matrix} R & L \\ \begin{pmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{pmatrix} \end{matrix}$$

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 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 26% a 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 . Dado que $u = (0.6, 0.2)$ es un punto fijo de P y así $t = (\frac{3}{4}, \frac{1}{4}) = (0.75, 0.25)$. Por tanto, se espera que en la prueba mil 75% de las ratas irán a la derecha y 25% a la izquierda.

15.4.3 Ejercicio 3

Dada la matriz de transición $P = \begin{pmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ con distribución de probabilidad inicial $p^{(0)} = (\frac{1}{3}, \frac{2}{3})$. Definir y hallar:

- a) $p_{21}^{(3)}$
- b) $p^{(3)}$
- c) $p_2^{(3)}$

Solución:

a) $p_{21}^{(3)}$ es la probabilidad de pasar del estado a_2 al estado a_1 en 3 pasos. Esto se puede obtener de la matriz de transición de 3 pasos P^3 ; por tanto, calculamos primero P^3 :

$$P^2 = \begin{pmatrix} 1 & 0 \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix}, \quad P^3 = \begin{pmatrix} 1 & 0 \\ \frac{7}{8} & \frac{1}{8} \end{pmatrix}$$

Entonces $p_{21}^{(3)}$ es el elemento de la segunda fila primera columna de P^3 : $p_{21}^{(3)} = \frac{7}{8}$

b) $p^{(3)}$ es la distribución de probabilidad del sistema después de tres pasos. Puede obtenerse calculando sucesivamente $p^{(1)}$, $p^{(2)}$ y luego $p^{(3)}$:

$$p^{(1)} = p^{(0)}P = \left(\frac{1}{3}, \frac{2}{3}\right) \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \left(\frac{2}{3}, \frac{1}{3}\right)$$

$$p^{(2)} = p^{(1)}P = \left(\frac{2}{3}, \frac{1}{3}\right) \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \left(\frac{5}{6}, \frac{1}{6}\right)$$

$$p^{(3)} = p^{(2)}P = \left(\frac{5}{6}, \frac{1}{6}\right) \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \left(\frac{11}{12}, \frac{11}{12}\right)$$

Sin embargo, como la matriz de transición de 3 pasos P^3 ya se calculó en el inciso a), $p^{(3)}$ también puede obtenerse como sigue:

$$p^{(3)} = p^{(0)}P^3 = \left(\frac{1}{3}, \frac{2}{3}\right) \begin{pmatrix} 1 & 0 \\ \frac{7}{8} & \frac{1}{8} \end{pmatrix} = \left(\frac{11}{12}, \frac{1}{12}\right)$$

c) $p_2^{(3)}$ es la probabilidad de que el proceso esté en el estado a_2 después de 3 pasos; esto es, la segunda componente de la distribución de probabilidad de 3 pasos $p^{(3)}$: $p_2^{(3)} = \frac{1}{12}$

15.4.4 Ejercicio 4

Dada la matriz $P = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 0 \end{pmatrix}$ y la distribución de probabilidad inicial $p^{(0)} = (\frac{2}{3}, 0, \frac{1}{3})$. Hallar:

- a) $p_{32}^{(2)}$ y $p_{13}^{(2)}$
- b) $p^{(4)}$ y $p_3^{(4)}$
- c) El vector al cual se aproxima $p^{(0)}P^n$
- d) La matriz a que se aproxima P^n

Solución:

a) Calculamos primero la matriz de transición de 2 pasos P^2 :

$$P^2 = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{4} & \frac{3}{4} & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$$

Entonces $p_{32}^{(2)} = \frac{1}{2}$ y $p_{13}^{(2)} = 0$, puesto que estos números se refieren a los elementos de P^2 .

b) Para calcular $p^{(4)}$, usamos la matriz de transición de 2 pasos P^2 y la distribución de probabilidad inicial $p^{(0)}$:

$$p^{(2)} = p^{(0)}P^2 = \left(\frac{1}{3}, \frac{2}{3}, 0\right) \quad y \quad p^{(4)} = p^{(2)}P^2 = \left(\frac{1}{4}, \frac{7}{12}, \frac{1}{6}\right)$$

Puesto que $p_3^{(4)}$ es la tercera componente de $p^{(4)}$, $p_3^{(4)} = \frac{1}{6}$

c) $p^{(0)}P^n$ se aproxima al vector de probabilidad fijo único t de P . Para obtener t , hallamos primero un vector fijo $u = (x, y, z)$:

$$(x, y, z) \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 0 \end{pmatrix} = (x, y, z) \quad o \quad \begin{cases} \frac{1}{2}y = x \\ \frac{1}{2}x + \frac{1}{2}y + z = y \\ \frac{1}{2}x = z \end{cases}$$

Hallar una solución no nula del anterior sistema de ecuaciones. Hacemos $x = 1$; luego por la tercera ecuación $x = 2$, y por la primera ecuación $y = 4$. Por tanto, $u = (2, 4, 1)$ es un punto fijo de P y así $t = (\frac{2}{7}, \frac{4}{7}, \frac{1}{7})$. En otras palabras, $p^{(0)}P^n$ se aproxima a $(\frac{2}{7}, \frac{4}{7}, \frac{1}{7})$.

d) P^n se aproxima a la matriz T cuyas filas son cada una el vector de probabilidad fijo de P ; por consiguiente P^n se aproxima:

$$\begin{pmatrix} \frac{2}{7} & \frac{4}{7} & \frac{1}{7} \\ \frac{2}{7} & \frac{4}{7} & \frac{1}{7} \\ \frac{2}{7} & \frac{4}{7} & \frac{1}{7} \end{pmatrix}$$

15.4.5 Ejercicio 5

La región de ventas de un vendedor la componen tres ciudades, A , B y C . Nunca vende en la misma ciudad en días seguidos. Si vende en la ciudad A , entonces el día siguiente vende en la ciudad B . Sin embargo, si vende en una de las dos B o C , entonces al día siguiente está en doble posibilidad tanto para vender en A como en la otra ciudad. A la larga, ¿con qué frecuencia vende en cada ciudad?

Solución:

La matriz de transición del problema es como sigue:

$$P = \begin{matrix} & \begin{matrix} A & B & C \end{matrix} \\ \begin{matrix} A \\ B \\ C \end{matrix} & \begin{pmatrix} 0 & 1 & 0 \\ \frac{2}{3} & 0 & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & 0 \end{pmatrix} \end{matrix}$$

Buscamos el vector de probabilidad fijo único de la matriz P . Hallamos primero un vector fijo $u = (x, y, z)$:

$$(x, y, z) \begin{pmatrix} 0 & 1 & 0 \\ \frac{2}{3} & 0 & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & 0 \end{pmatrix} = (x, y, z) \quad \text{o} \quad \begin{cases} \frac{2}{3}y + \frac{2}{3}z = x \\ x + \frac{1}{3}z = y \\ \frac{1}{3}y = z \end{cases}$$

Hacemos $z = 1$. Luego por la tercera ecuación $y = 3$ y por la primera ecuación $x = \frac{8}{3}$. Así, $u = (\frac{8}{3}, 3, 1)$. También $3u = (8, 9, 3)$ es un vector de probabilidad fijo de P . Multiplicamos $3u$ por $1/(8+9+3) = \frac{1}{20}$ para obtener el vector de probabilidad fijo pedido $t = (\frac{2}{5}, \frac{9}{20}, \frac{3}{20}) = (0.40, 0.45, 0.15)$. Por consiguiente, a la larga vende 40% del tiempo en la ciudad A , 45% del tiempo en B y 15% del tiempo en C .

Capítulo 16

16.1 Clasificaciones de Inteligencia Artificial

16.1.1 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. En 2011, la página Singularity Hub escribió: "A medida que los robots y las inteligencias artificiales estrechas toman un rol en papeles tradicionalmente ocupados por los seres humanos, debemos preguntarnos a nosotros mismos: ¿Es toda esta automatización buena o mala para el mercado de trabajo?"

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". 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 declarado 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..." El Centro Stanford para el Internet y la Sociedad (Stanford Center for Internet and Society) en la siguiente cita, contrasta la IA fuerte con la IA débil en relación con el crecimiento de la IA estrecha que presenta "problemas reales".

La IA débil o "estrecha", en contraste, es una realidad actual. El software controla muchas facetas de la vida cotidiana y, en algunos casos, este control presenta problemas reales. Un ejemplo se dio en mayo de 2010, el "flash crash" que provocó una caída temporal pero enorme en el mercado.

Los siguientes dos extractos de Singularity Hub resumen a la IA débil:

Cuando se llama al banco y se habla con una voz automatizada probablemente estás hablando con una IA... una muy molesta. Nuestro mundo está lleno de estos programas limitados de IA que clasificamos como "débil" o "estrecha" o "aplicada". Estos programas están lejos de los inteligentes, buscadores de

amor, lleno de angustia, las inteligencias artificiales que vemos en la ciencia ficción, pero es temporal. Todas estas inteligencias artificiales estrechas son como los aminoácidos en el lodo de la Tierra.

Poco a poco estamos construyendo una biblioteca de talentos de IA estrechas que cada vez son más impresionantes. El procesamiento y reconocimiento de voz permiten a los equipos convertir sonidos en texto con mayor precisión. Google está usando la IA para subtítular millones de videos en YouTube. Así mismo, la visión del computador está mejorando para que programas como Vitamin d Video pueda reconocer objetos, clasificarlos y entender cómo se mueven. La IA estrecha no sólo está mejorando en el procesamiento del entorno, sino también está entendiendo la diferencia entre lo que un ser humano dice y lo que un ser humano quiere.

- IA débil, un sistema de inteligencia artificial que sólo pretende ser aplicable a un tipo específico de problemas (ajedrez de computadora) y no la intención de mostrar inteligencia humana en general; ver IA fuerte.
- Hipótesis de la IA débil, la posición de la filosofía de la inteligencia artificial es que las máquinas pueden demostrar inteligencia, pero no necesariamente tienen una mente, estados mentales o una conciencia.

16.1.2 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.

16.1.3 Inteligencia Artificial Distribuida

La Inteligencia artificial distribuida es un campo de la IA dedicado al estudio de las técnicas y el conocimiento necesario para la coordinación y distribución del conocimiento y las acciones en un entorno con múltiples agentes.

Podemos distinguir dos áreas principales de desarrollo:

1. Solución cooperativa de problemas distribuidos (SCPD): Estudia como unos conjuntos de módulos (o nodos) cooperan para dividir y compartir el conocimiento de un problema y en el desarrollo de la solución.
2. Sistemas multiagentes (SMA): Estudia la coordinación de la conducta inteligente entre un conjunto de agentes inteligentes autónomos.

La principal diferencia entre ambas áreas estriba en la flexibilidad de la coordinación entre los agentes. En la SCPD, las interacciones y tareas que cada agente realiza, están prefijadas de antemano: hay un plan centralizado de resolución de problemas. Suele haber un miembro que ejerce un control global que centraliza los resultados parciales y datos entre el resto de los componentes del sistema. En contraposición en los SMA, los agentes tienen un grado de autonomía mayor y pueden decidir dinámicamente que interacciones son adecuadas, que tareas deben realizar, quien realiza cada tarea y, además, es posible mantener conocimiento que no es globalmente consistente, incluso los agentes pueden mantener objetivos globales diferentes. Esta definición permite distinguir entre sistemas que se centran en el comportamiento global, con una conducta fija de los agentes (SCPD) y sistemas que se centran en la conducta de los individuos que como resultado, obtenían una conducta del sistema (SMA). Vista como una sociedad, sería escoger entre un estado que planifica y regula las conductas de los individuos (que serán predecibles) o dejar que el sistema se resuelva por la libre iniciativa de los individuos.

Los problemas básicos que estudia la IAD y que son comunes a todos los sistemas son:

1. Como formular, describir, descomponer y asignar problemas y sintetizar los resultados entre un grupo de agentes inteligentes.
2. Como capacitar a los agentes para que se comuniquen e interactúen: qué lenguajes de comunicación o protocolos deben utilizarse, qué y cuándo deben comunicarse, etc.

3. Como asegurar que los agentes actúen coherentemente al tomar decisiones o realizar acciones, como acomodar los efectos globales de las decisiones locales y prevenir interacciones no deseadas.
4. Como capacitar a los agentes para representar y razonar sobre acciones, planes y conocimientos de otros agentes para coordinarse; como razonar sobre el estado de su proceso de coordinación (inicio o terminación).
5. Como reconocer y reconciliar puntos de vista e intenciones conflictivas entre un conjunto de agentes para coordinar sus acciones; como sintetizar los puntos de vista y los resultados.
6. Como utilizar técnicas ingenieriles y desarrollar sistemas con IAD. Como desarrollar plataformas de SMA y metodologías de desarrollo con tenias de IAD.

16.1.4 Ramas de la Inteligencia Artificial

- Sistemas Expertos (Sistemas basados en Conocimiento). Programas computacionales que resuelven problemas que normalmente requieren del conocimiento de un especialista o experto humano. Es un sistema capaz de tomar decisiones inteligentes interpretando grandes cantidades de datos sobre un dominio específico de problemas.
- Aprendizaje y Razonamiento Automático. Máquinas capaces de planificar, tomar decisiones, plantear y evaluar estrategias, aprender a partir de la experiencia, autoreprogramables, etc.
- Robótica. Artefactos autónomos capaces de llevar a cabo diversas tareas mecánicas de manera flexible e inteligente, cumpliendo con un objetivo y ajustándose al entorno cambiante.
- Procesamiento de Lenguaje Natural. Sistemas capaces de reconocer, procesar y emular el lenguaje humano.
- Visión por Computadora (Reconocimiento de patrones). Reconoce y procesa señales, caracteres, patrones, objetos, escenas.

Áreas de Vanguardia: Inteligencia Natural

- Redes Neurales. Crear elementos de procesamiento y organizarlos de acuerdo a un modelo basado en las células del cerebro humano (neuronas). Estos sistemas no se programan, se entrena. Se caracterizan por reconocer objetos partiendo de señales ruidosas.
- Lógica Difusa. Basado en los principios del razonamiento aproximado y el "cálculo con palabras", éstos sistemas logran simplificar y aproximar la descripción del problema de una manera natural, eficiente y robusta. La lógica difusa va más allá de la lógica booleana en cuanto a que acepta valores parciales de verdad, es decir, de 0 a 100%; aceptando con ello expresiones tales

como: "Juan es alto" con un 75% de certeza, o mejor aún, simplificándolo a "Juan no es muy alto"; "El tanque está lleno" con 50% de certeza, o bien, "El tanque está medio lleno o medio vacío".

- Algoritmos genéticos. La ley de la selva de la Naturaleza: "La supervivencia del más apto", ha impulsado la evolución, extinción y supervivencia de los seres vivos. Cada organismo, tiene inscrito y codificado el conocimiento-biológico-genético acumulado durante millones de años. En una computadora, comenzando por una población inicial de organismos-candidatos para la solución de un problema, éstos se recombinan de manera aleatoria (reproducción), luego de seleccionar los mejores, y después de cierto tiempo o número de generaciones se alcanza una solución suficientemente buena para resolver el problema.
- Vehículos Autónomos. Son una amalgama de ramas AI como: robótica, aprendizaje de máquina, sistemas expertos, visión, etc. para lograr que una máquina inteligente pueda maniobrar en un espacio físico singular pero variable, de manera autónoma hasta lograr el objetivo que se le asigne: tomar una muestra de la superficie del planeta Marte, conducir un vehículo hasta cierto destino por una carretera transitada, etc.
- Realidad Virtual. Recrea mundos artificiales en tiempo real que pueden ser captados por diversos canales sensoriales en el espectador, el cuál puede navegar "inmerso" a través de dicho mundo virtual.
- Agentes (Wizards). Son programas "invisibles" tipo espía que analizan las tareas que esté llevando a cabo un usuario, y que dependiendo de las preferencias, costumbres y nivel del usuario, en cuanto se detecte alguna anomalía, el agente "aparece" ante el usuario para ayudarle (dando información), sugiriendo una solución o para ejecutar un conjunto de tareas rutinarias de manera automática.



Figura 76: Ramas de la Inteligencia Artificial

Referencias

- [1] Graphs for divisibility criterion. <http://plover.com/~mjd/misc/math/Divisibility/>, 2016. [Online; accessed 19-June-2016].
- [2] Ana Albarrán, Jorge y García. *Algunas paradojas de la teoría (intuitiva) de conjuntos*. Centro de Investigación en Matemáticas, 2005.
- [3] Joseph K Blitzstein and Jessica Hwang. *Introduction to probability*. CRC Press, 2014.
- [4] 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.
- [5] Anne-C Chardon and JL González. Amenaza, vulnerabilidad, riesgo, desastres, mitigación prevención. primer acercamiento a conceptos, características y metodologías de análisis y evaluación. banco interamericano de desarrollo. universidad nacional de colombia. sede manizales instituto de estudios ambientales. *IDEA*, 2002.
- [6] Morris H Morris H DeGroot et al. *Probability and statistics*. Number 04; QA273, D4 1986. 1986.
- [7] J Doyne Farmer, Eric Smith, and Martin Shubik. Economics: the next physical science? *arXiv preprint physics/0506086*, 2005.
- [8] Francisco José González Gutiérrez. Apuntes de matemática discreta - operaciones con conjuntos, 2004.
- [9] G.Polya. *How to Solve It*. Princeton University Press, 1957.
- [10] Seymour Lipschutz. Probabilidad. serie schaum. *McGraw Hill*. México, 2000.
- [11] Manuel Mateos López. Vínculos matemáticos, los conjuntos, 2001.
- [12] Fernando Suárez Obando. Riesgo y probabilidad: consideraciones conceptuales en la intersección ambiente, genes y cultura. *Revista Salud Bosque*, 2(2):39–54, 2015.
- [13] John Allen Paulos and Josep M Llosa. *El hombre anumérico: el analfabetismo matemático y sus consecuencias*. Tusquets editores, 1990.
- [14] Luis Rincón. Curso elemental de probabilidad y estadística, 2007.
- [15] Luis Rincón. Introducción a los procesos estocásticos. *Departamento de de Matemáticas, Facultad de Ciencias UNAM*, 2008.
- [16] Terenzio Soldoviei. Delta de dirac, 2011.