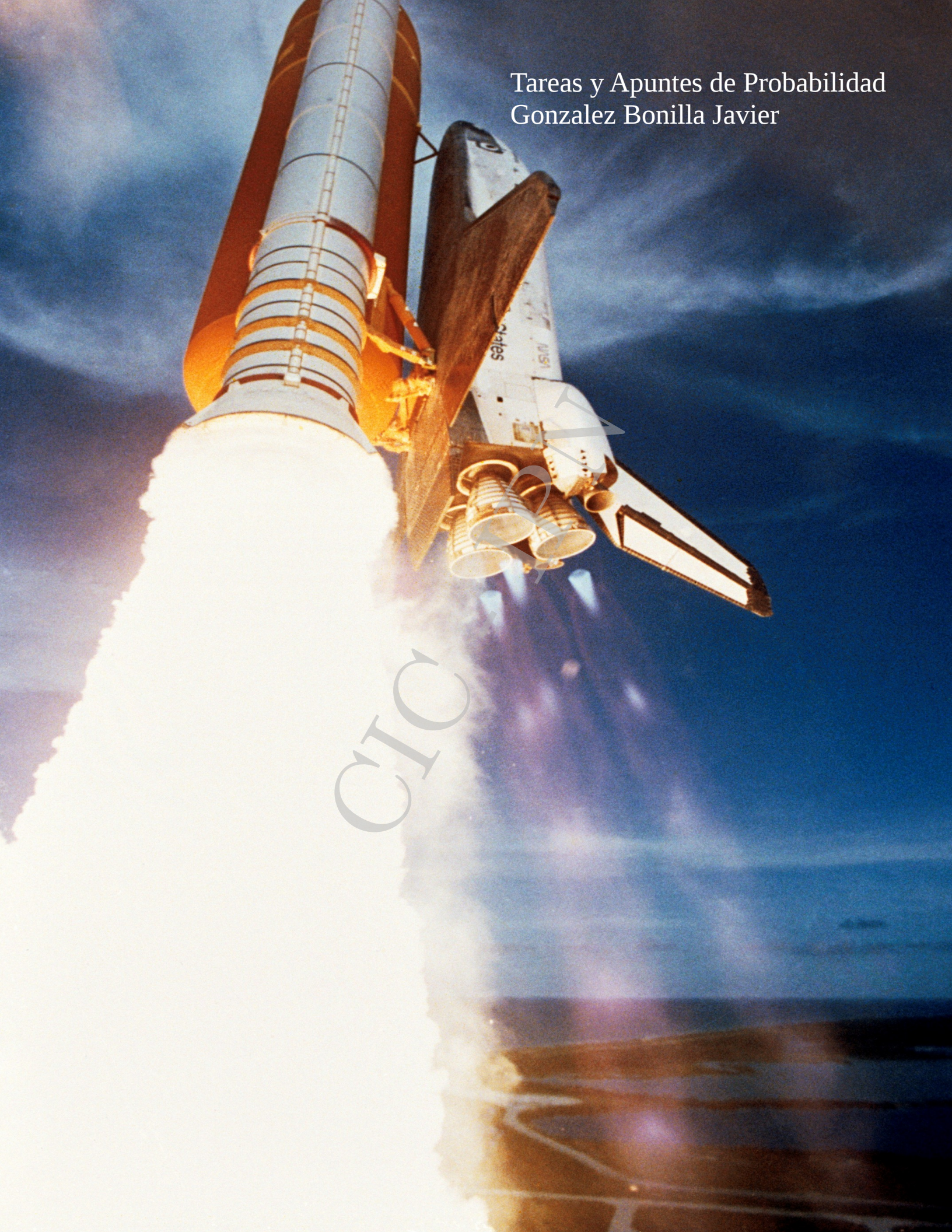


Tareas y Apuntes de Probabilidad
Gonzalez Bonilla Javier



Índice

| | |
|--|-----------|
| 1. Tarea 1 | 7 |
| 1.1. Experimentos | 7 |
| 1.1.1. Aleatorio | 7 |
| 1.1.2. Determinista | 7 |
| 1.2. Tipos de Probabilidad | 8 |
| 1.2.1. Probabilidad Clásica | 8 |
| 1.2.2. Probabilidad Geométrica | 8 |
| 1.2.3. Probabilidad Frecuentista | 8 |
| 1.2.4. Probabilidad Subjetiva | 9 |
| 1.2.5. Probabilidad Axiomática | 9 |
| 2. Tarea 2 | 10 |
| 2.1. Ejemplos de Permutaciones. | 10 |
| 2.2. Ejemplos de combinaciones. | 12 |
| 2.3. Generalización del problema de caminos posibles para el caso de un rectángulo | 14 |
| 2.4. Generalización del problema de caminos posibles para el caso de un cubo | 16 |
| 3. Tarea 3 | 18 |
| 3.1. Ejemplos de Permutaciones Circulares | 18 |
| 3.2. Problemas 1.6, 1.7 y 1.8 de [1] | 21 |
| 3.3. Números de Catalán | 22 |
| 3.4. Tensor de Levy-Civita | 23 |
| 4. Tarea 4 | 24 |
| 4.1. Graficaciones en RooT | 24 |
| 4.1.1. Graficar en ROOT $\ln(x!)$ | 24 |
| 4.1.2. Graficar en ROOT $x\ln(x) - x$ | 26 |
| 4.1.3. Graficar en ROOT la diferencia de las 2 funciones anteriores. | 27 |
| 4.2. ¿Que es un histograma? | 28 |
| 4.3. Fórmula de Stirling. | 30 |
| 4.4. Principio Multiplicativo y Aditivo | 31 |
| 4.5. Numeros grandes | 33 |
| 4.6. Aportaciones de Srinivasa Aiyangar Ramanujan | 33 |
| 4.6.1. Número de Ramanujan | 33 |
| 5. Tarea 5 | 34 |
| 5.1. Simulación de un dado de 3 caras | 34 |
| 5.2. Causas de la explosión del Challenger | 35 |
| 5.3. Tipos de infinito | 35 |
| 5.3.1. \aleph_0 | 35 |
| 5.3.2. \aleph_1 | 35 |
| 5.3.3. Mas allá de \aleph_1 | 36 |
| 5.4. ¿Como se compacta el mp3? | 36 |
| 5.5. Examen 4 | 37 |
| 5.6. Máquina enigma | 40 |

| | |
|--|-----------|
| 6. Tarea 6 | 43 |
| 6.1. Realizar un programa en gcc controlado por make que resuelva una ecuación cuadrática. | 43 |
| 6.2. Realizar un histograma en ROOT que se llene con la distribución de Gauss. | 44 |
| 6.3. Usar root para obtener el histograma 1.2.3 de [2]. | 46 |
| 6.4. Mínimos cuadrados. | 48 |
| 6.5. Densidad de probabilidad | 52 |
| 7. Tarea 7 | 54 |
| 7.1. Eigen valores y eigen vectores. | 54 |
| 7.2. Análisis de componentes principales. | 56 |
| 7.3. Demostraciones del capítulo 1 de [2] | 60 |
| 7.4. Mínimos cuadrados ejemplo | 62 |
| 8. Tarea 8 | 71 |
| 8.1. Ejemplos de cálculo de eigenvalores y eigenvectores (con multiplicidad) | 71 |
| 8.2. Con los datos del ejercicio anterior ilustrar la propiedad 8 | 74 |
| 8.3. Multiplicidad en eigenvectores y eigenvalores. | 74 |
| 8.3.1. Multiplicidad algebraica | 74 |
| 8.4. Matriz singular. | 74 |
| 8.5. Traza de una matriz. | 75 |
| 8.6. Traspuesta de una matriz. | 75 |
| 8.6.1. Propiedades | 75 |
| 8.7. Teorema de Cayley-Hamilton. | 76 |
| 8.7.1. Ejemplo | 77 |
| 9. Tarea 9 | 78 |
| 9.1. Teorema de Cayley-Hamilton | 78 |
| 10. Tarea 10 | 79 |
| 10.1. Graficación en R | 79 |
| 10.1.1. Ejemplo 1 | 79 |
| 10.1.2. Ejemplo 2 | 79 |
| 10.1.3. Ejemplo 3 | 80 |
| 10.1.4. Ejemplo 4 | 80 |
| 10.1.5. Ejemplo 5 | 81 |
| 10.2. Matrices en R | 83 |
| 10.2.1. Funciones sobre matrices | 84 |
| 10.2.2. Operaciones con matrices | 86 |
| 10.3. Ecuación cuadrática root (clases) | 86 |
| 11. Tarea 11 | 87 |
| 11.1. Correlación de Pearson | 87 |
| 12. Tarea 12 | 91 |
| 12.1. Aproximación a $\ln(x)$ por series de Taylor | 91 |
| 13. Tarea 13 | 94 |
| 13.1. Ejemplos de mínimos cuadrados en Mathematica | 94 |
| 13.2. Cadenas de Markov | 97 |
| 13.2.1. Matriz de transición: | 97 |
| 13.2.2. Ejemplos de Cadenas de Markov en Mathematica | 98 |
| 13.2.3. Otros Ejemplos de Matrices de transición y sus respectivos grafos | 99 |

| | |
|---|------------|
| 13.3. Histograma | 100 |
| 13.3.1. Ejemplo | 101 |
| 13.4. Box and whisker plots | 101 |
| 13.4.1. Ejemplo | 102 |
| 14.Tarea 14 | 104 |
| 14.1. Demostraciones de la clase de Caminata al azar de [3] | 104 |
| 14.1.1. Ecuación 46 de [3] | 104 |
| 14.1.2. Ecuación 54 de [3] | 104 |
| 14.1.3. ecuación 55 de [3] | 105 |
| 14.1.4. Ecuación 56 de [3] | 105 |
| 14.1.5. Ecuación 57 de [3] | 105 |
| 14.1.6. Ecuación 58 de [3] | 106 |
| 14.1.7. Ecuación 59 de [3] | 106 |
| 14.1.8. Ecuación 60 de [3] | 107 |
| 14.1.9. Ecuación 61 de [3] | 107 |
| 14.1.10.ecuación 62 de [3] | 107 |
| 15.Tarea 15 | 109 |
| 15.1. Congruencia Zeller | 109 |
| 15.1.1. Algoritmo | 109 |
| 15.1.2. Codigo en matemática | 110 |
| 16.Tarea 16 | 111 |
| 16.1. Función de Distribución de Energía | 111 |
| 16.1.1. Distribución Maxwell-Boltzmann | 111 |
| 16.1.2. Distribución Fermi-Dirac | 112 |
| 16.1.3. Distribución Bose Einstein | 113 |
| 17.Tarea 17 | 114 |
| 17.1. La última oveja de pie | 114 |
| 17.1.1. Solución 1 | 114 |
| 17.1.2. Solución 2 | 115 |
| 17.1.3. Solución 3 | 116 |
| 18.Tarea 18 | 118 |
| 18.1. Series Geométricas | 118 |
| 18.1.1. Suma | 118 |
| 18.1.2. Convergencia | 118 |
| 18.2. Ejemplos | 118 |

Índice de cuadros

| | |
|--|----|
| 1. Tabla de Frecuencia. | 28 |
| 2. Tabla de Frecuencia 2. | 30 |
| 3. Datos. | 50 |
| 4. Datos para la función de densidad de probabilidad | 52 |

Índice de figuras

| | | |
|-----|--|-----|
| 1. | Probabilidad geométrica en una línea. | 8 |
| 2. | Probabilidad de que P se encuentre en S . | 8 |
| 3. | Caminos posibles para un cuadrado. | 14 |
| 4. | camino posibles para un rectángulo compuesto por 2 cuadrados. | 14 |
| 5. | Generalización del número de caminos posibles a cualquier vértice. | 15 |
| 6. | Caminos posibles en un cubo. | 16 |
| 7. | Caminos posibles en un paralelepípedo. | 17 |
| 8. | Triangulación de polígonos. | 22 |
| 9. | Gráfica de $\ln(x)$. | 24 |
| 10. | Gráfica de $x\ln(x) - x$. | 26 |
| 11. | Gráfica de $\ln(x) - (x\ln(x) - x)$. | 27 |
| 12. | Histograma tomado de [4]. | 28 |
| 13. | Histograma obtenido a través de la tabla de frecuencias 1. | 29 |
| 14. | Histograma obtenido a través de la tabla de frecuencias 1. | 30 |
| 15. | Histogramas del lanzamiento de un dado de 3 caras. | 34 |
| 16. | Máquina enigma | 40 |
| 17. | Rodillos de la Máquina enigma | 41 |
| 18. | Conexión interna de la Máquina enigma | 42 |
| 19. | Histograma que se ajusta a la distribución de Gauss | 44 |
| 20. | Histograma 1.2.3 de [2] | 46 |
| 21. | Graficación del error de los puntos en el ajuste | 48 |
| 22. | Graficación de los datos de la tabla 3 | 50 |
| 23. | Recta calculada con mínimos cuadrados. | 51 |
| 24. | Histograma y la función de densidad de probabilidad. | 52 |
| 25. | Cálculo de la probabilidad con histograma y con la función. | 53 |
| 26. | Gráfica del ejemplo 1 de mínimos cuadrados | 63 |
| 27. | Gráfica del ejemplo 2 de mínimos cuadrados | 64 |
| 28. | Gráfica del ejemplo 3 de mínimos cuadrados | 66 |
| 29. | Gráfica del ejemplo 4 de mínimos cuadrados | 68 |
| 30. | Gráfica del ejemplo 5 de mínimos cuadrados | 70 |
| 31. | Ejemplo 1 de mínimos cuadrados con R | 79 |
| 32. | Ejemplo 2 de mínimos cuadrados en R | 80 |
| 33. | Ejemplo 3 de mínimos cuadrados en R | 81 |
| 34. | Ejemplo 4 de mínimos cuadrados en R | 82 |
| 35. | Ejemplo 5 de mínimos cuadrados en R | 83 |
| 36. | Gráfica generada con números aleatorios | 88 |
| 37. | Gráfica generada a partir de una ecuación cuadrática | 89 |
| 38. | Gráfica de una recta | 90 |
| 39. | Aproximación de la función $\ln(x)$ por serie de Taylor. | 93 |
| 40. | Puntos aleatorios alrededor de la función x . | 94 |
| 41. | Ajuste con una función $f(x) = x$. | 95 |
| 42. | Ajuste con una función $f(x) = x^2$. | 95 |
| 43. | Ajuste con una función $f(x) = x^3$. | 96 |
| 44. | Ajuste con una función $f(x) = x^4$. | 96 |
| 45. | Gráfica de puntos y sus aproximaciones. | 97 |
| 46. | Grafo que representa la matriz de transiciones 57 | 98 |
| 47. | Grafo que representa a la matriz de transiciones 58 | 99 |
| 48. | Grafo que representa la matriz de transiciones 59 | 100 |
| 49. | Histogramas generados por MATHEMATICA. | 101 |

| | |
|--|-----|
| 50. Diagrama Box and Whiskers. | 102 |
| 51. Diagrama Box and Whiskers. | 103 |

CIC-IPN

1. Tarea 1

1.1. Experimentos

1.1.1. Aleatorio

Un experimento aleatorio, es un experimento en el cual, los resultados no son esencialmente los mismos, a pesar de que las condiciones sean aproximadamente idénticas. A continuación se describen 10 ejemplos de ellos.

1. Lanzar una moneda.
2. Lanzar una moneda.
3. El número de Tornillos defectuosos que produce una máquina.
4. Medir el tiempo de vida de una lámpara.
5. Girar un ruleta.
6. Sacar una bola de una urna que contiene bolas numeradas.
7. Sacar una carta de una baraja.
8. Contar el número de estrellas fugases que pueden observarse en una noche.
9. Contar el número de accidentes automovilísticos en el D.F.
10. Contar el número de personas que mueren en una semana en México.

1.1.2. Determinista

Un experimento determinista es aquel que dadas ciertas condiciones, los resultados son siempre los mismos. A continuación se describen 10 ejemplos de ellos.

1. Dejar caer un objeto al suelo.
2. Sacar una bola roja de una urna que tiene 5 bolas rojas.
3. El sol sale por el este.
4. La tolerancia en el maquinado de un eje es de $\pm 1mm$ y la máquina que lo elabora tiene una precisión de $0,01mm$.
5. Si un número natural es par, el siguiente número será impar.
6. Si se calienta acero inoxidable por arriba de a $1430^{\circ}C$, este se funde.
7. Al a calentar agua a $100^{\circ}C$ a una presión de 1 atm, se evapora.
8. Calcular el volumen de un cubo de 3 cm por lado.
9. La resistencia de un trozo metal es menor antes de ser forjado.
10. Medir la longitud de una regla de 30cm.

1.2. Tipos de Probabilidad

1.2.1. Probabilidad Clásica

- Si un suceso puede ocurrir en h maneras diferentes, de un número total de n maneras posibles, todos igualmente factibles, entonces la probabilidad del suceso es $\frac{h}{n}$.
- Sea A un evento para un experimento con un espacio muestral S , la probabilidad sencilla de A es:

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

1.2.2. Probabilidad Geométrica

La probabilidad de un evento representa la posibilidad de que ese evento ocurra. La probabilidad geométrica describe la posibilidad de que un punto esté en una parte de un segmento de línea o en una parte de una región. Si se considera el segmento de línea de la figura 1 \overline{AC} , y suponga un punto x escogido al azar, que está sobre la línea. Entonces, la probabilidad de x esté en $\overline{AB} = \frac{AB}{AC}$.



Figura 1: Probabilidad geométrica en una línea.

Considere la región R y S de la figura 2 y considere que un punto P es escogido al azar.

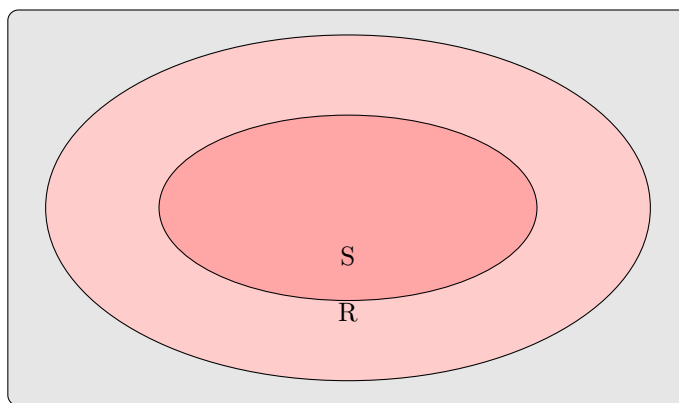


Figura 2: Probabilidad de que P se encuentre en S .

Entonces la probabilidad de que P esté en S es igual a $P(S) = \frac{\text{Área de } S}{\text{Área de } R}$.

1.2.3. Probabilidad Frecuentista

Si después de n repeticiones de un experimento, donde n es muy grande, un suceso ocurre h veces, entonces la probabilidad del suceso es:

$$P(h) = \frac{h}{n}.$$

Esto también se llama "probabilidad empírica del suceso".

1.2.4. Probabilidad Subjetiva

El lenguaje de probabilidad a menudo se utiliza en conexión con situaciones que son inherentemente irrepetibles. Algunos ejemplos incluyen: “Las probabilidades de un tratado de paz son buenas”, o “Es probable que el contrato le sea otorgado a nuestra compañía”, o “Como su mejor mariscal de campo está lesionado, espero que no anoten mas de 10 puntos contra nosotros”. En tales situaciones se desearía, asignar probabilidades numéricas a varios resultados y eventos. Por lo tanto se debe de adoptar una interpretación alternativa de estas probabilidades. Como diferentes observadores pueden tener información y opiniones previas con respecto a tales situaciones experimentales, las asignaciones de probabilidad ahora pueden definir de un individuo a otro. Las interpretaciones en tales situaciones se conocen por lo tanto como subjetivas

1.2.5. Probabilidad Axiomática

El objetivo de la probabilidad es asignar a cada evento de A un número $P(A)$ llamado probabilidad del evento A , el cuál da una precisa medida de la oportunidad de que el evento A ocurra. Los axiomas de probabilidad son los siguientes:

- Axioma 1: Para cada evento A , $P(A) \geq 0$.
- Axioma 2: $P(S) = 1$.
- Axioma 3: Si A_1, A_2, A_3, \dots es una colección infinita de eventos disjuntos, entonces $P(A_1 \cup A_2 \cup A_3 \cup \dots) = \sum_{i=1}^{\infty} P(A_i)$.

2. Tarea 2

2.1. Ejemplos de Permutaciones.

- 1.- ¿De cuántas maneras diferentes pueden sentarse 10 personas en una banca en la que solo caben 4?

Respuesta:

En el primer puesto se puede sentar cualquiera de las 10 personas, después, solo hay 9 formas para ocupar el segundo puesto, 8 para formas para el el tercero y finalmente 7 para el cuarto. Entónces, el número de formas distintas en las que 10 personas pueden sentarse en 4 puestos es:

$${}_{10}P_4 = 10 * 9 * 8 * 7 = 5040.$$

- 2.- Se quieren sentar 5 hombres y 4 mujeres en una fila, de modo que las mujeres ocupen los sitios pares, ¿De Cuántas formas pueden sentarse ?

Respuesta:

Los hombres pueden sentarse de ${}_5P_5$ formas, y las mujeres de ${}_4P_4$ formas. Cada ordenación de los hombres puede asociarse con cada ordenación de las mujeres.

Ordenaciones posibles = ${}_5P_5 * {}_4P_4 = 5 * 4 = 2880$.

- 3.- Cuatro libros distintos de matemáticas, seis diferentes de física y dos diferentes de química se colocan en un estante. ¿De cuántas formas distintas es posible ordenarlos si (a) los libros de cada asignatura deben de estar todos juntos, (b) solamente los libros de matemáticas deben de estar juntos?

Respuestas:

- a) Los libros de Matemáticas pueden ordenarse de ${}_4P_4 = 4!$ formas diferentes, los libros de física pueden ordenarse de ${}_6P_6 = 6!$ formas diferentes, los libros de química pueden ordenarse de ${}_2P_2 = 2$ formas diferentes, y finalmente cada paquete de libros de cada materia puede ordenarse de ${}_3P_3 = 3!$ formas diferentes, entonces, el número de formas de ordenar los libros está determinada por:

$${}_6P_6 \cdot {}_4P_4 \cdot {}_2P_2 \cdot {}_3P_3 = 6!4!2!3! = 207360.$$

- b) Si se considera al paquete de libros de matemáticas como un solo elemento, entonces se podría asumir que el número de libros conformado por libros de física y de química es de 8, mas un elemento más de los libros de matemáticas, entonces son 9 elementos; el número de formas en las que pueden ordenarse es ${}_9P_9$. No hay que olvidar que los libros de matemáticas pueden ordenarse entre ellos mismos de ${}_4P_4$ formas diferentes. Entonces el número de formas de ordenar los libros esta determinado por:
- $${}_9P_9 \cdot {}_4P_4 = 9!4! = 8709120.$$

- 4.- Se ordenan en una fila 5 bolas rojas, 2 bolas blancas y 3 bolas azules. Si las bolas de igual color no se distinguen entre sí, ¿De cuántas formas posibles pueden ordenarse?

Respuesta:

Supóngase que hay N diferentes ordenaciones. Multiplicando N por el número de ordenaciones de las 5 bolas rojas entre sí, de las 2 bolas blancas entre sí y de las 3 bolas azules entre sí, se obtiene el número de ordenaciones de 10 bolas si todas ellas fuesen distintas, es decir $10!$. Entonces $5!2!3!N = 10! \rightarrow N = \frac{10!}{5!3!2!} = 2520$.

- 5.- ¿De cuántas formas pueden sentarse 7 personas alrededor de una mesa, si (a) pueden sentarse de cualquier forma, (b) si dos personas determinadas no deben estar una al lado de la otra?

Respuestas:

- a) Considérese una de ellas sentada en cualquier parte. Entonces las 6 restantes pueden sentarse de ${}_6P_6 = 6! = 720$ formas, que es el total de casos que se dan en la ordenación de 7 personas en un círculo.
- b) Considérense las dos personas que no deben juntas como una sola. Entonces hay 6 personas para sentarse en círculo, que lo pueden hacer de ${}_5P_5 = 5!$ formas. Pero las dos personas consideradas como una sola pueden ordenarse entre sí de ${}_2P_2 = 2!$ formas. Así pues, el número de ordenaciones de 6 personas sentadas alrededor de una mesa con 2 determinadas de ellas sentadas juntas es de ${}_5P_5 \cdot {}_2P_2 = 5! \cdot 2! = 240$, Entonces, mediante (a), se tiene el número total de formas en que 6 personas pueden sentarse alrededor de una mesa, entonces si se quiere eliminar los casos en los que dos de las personas no están sentadas juntas se le debe sustraer al total de formas de ordenar a las 6 personas, el número de casos en los que esa dos personas estan juntas.
 $720 - 240 = 480$ formas.

- 6.- Una placa de circuito impreso tiene ocho lugares diferentes en las que un componente se puede colocar. Si cuatro diferentes componentes se van a colocar en la tarjeta, ¿Cuántos diseños diferentes son posibles?

Respuesta:

Cada diseño consiste en seleccionar una ubicación de los ocho lugares para el primer componente, siete para el segundo componente, seis para el tercer componente, y una ubicación en los cinco restantes para el cuarto componente. Por lo tanto:

$${}_8P_4 = \frac{8!}{4!} = 8 \cdot 7 \cdot 6 \cdot 5 = 1680 \text{ diferentes diseños.}$$

- 7.- Considere una operación de mecanizado en la que una pieza de chapa metálica necesita dos agujeros de diámetro perforados idénticos y dos muescas cortadas tamaño idéntico. Denotamos una operación de perforación como d y una operación de entallar como n. En la determinación programa para un taller mecánico, podríamos estar interesados en el número de diferentes secuencias posibles de las cuatro operaciones.

Respuesta:

El número de posibles secuencias para dos operaciones de perforación y dos operaciones de entallar es:

$$\frac{4!}{2!2!} = 6.$$

Existen 6 secuencias fácilmente resumidas a $S = \{ddnn, dndn, dnnd, nddn, ndnd, nndd\}$

- 8.- Una pieza está marcada con cuatro líneas gruesas, tres líneas medias y dos líneas delgadas. Si cada orden de las nueve líneas representa una diferente etiqueta, ¿Cuántas etiquetas diferentes pueden ser generadas usando este esquema

Respuesta:

El número de posibles etiquetas es: $\frac{9!}{4!3!2!} = 2520$.

- 9.- En el palo de señales de un barco se pueden izar tres banderas rojas, dos azules y cuatro verdes. ¿Cuántas señales distintas pueden indicarse con la colocación de las nueve banderas?

Respuesta:

$${}_9P_{2,3,4} = \frac{9!}{2!3!4!} = 1260.$$

- 10.- ¿De cuántas formas diferentes pueden sentarse 6 alumnos en un salón de clases con 25 pupitres?

Respuesta:

El primer alumno puede elegir sentarse en los primeros 25 lugares, el segundo en 24, el tercero en 23, el cuarto en 22, el quinto en 21 y finalmente el sexto en 20, entonces el numero de formas en las que se pueden sentar es de ${}_{25}P_6 = 25 \cdot 24 \cdot 23 \cdot 22 \cdot 21 \cdot 20 = 127512000$.

2.2. Ejemplos de combinaciones.

- 1.- ¿De cuántas maneras pueden 10 objetos dividirse en dos grupos de 4 y 6 respectivamente ?

Respuesta:

Esto es lo mismo que el numero de ordenaciones o permutaciones de 10 objetos, de los cuales 4 de ellos son iguales y los otros 6 de ellos también son iguales entre sí. Entonces se puede usar la expresión ${}_nP_{n_1, n_2} = {}_{10}P_{4,6} = \frac{10!}{6!4!} = 210$

El problema es equivalente a encontrar el número de grupos de 4 objetos que se pueden formar con 10 objetos dados, (o de 6 objetos con 10 objetos dados), no teniendo en cuenta el orden de los objetos dentro del grupo. En general el número de grupos distintos de r objetos que se pueden formar con n objetos dados, se le llama número de combinaciones de los n objetos tomados de r en r y se denota por ${}_nC_r$ ó $\binom{n}{r}$, y esto viene dado por

$${}_nC_r = \binom{n}{r} = \frac{n!}{r!(n-r)!} = \frac{n(n-1)\dots(n-r+1)}{r!} = \frac{{}_nP_r}{r!},$$

Entonces si $n = 10$ y $r = 6$ ó $r = 4$, entonces:

$$\binom{n}{r} = \binom{10}{6} = \binom{10}{4} = 210.$$

- 2.- De un total de 5 matemáticos y 7 físicos, se forma un comité de 2 matemáticos y 3 físicos. ¿De cuántas formas puede formarse, si (a) puede pertenecer a el cualquier matemático y físico, (b) un físico determinado debe pertenecer al comité, (c) dos matemáticos determinados no pueden estar en el comité?

Respuestas:

- 2 matemáticos de un total de 5 pueden elegirse de ${}_5C_2$ formas y 3 físicos pueden escogerse de un total de 7 de ${}_7C_3$. Entonces, el número de selecciones posibles es ${}_5C_2 * {}_7C_3 = 10 * 35 = 350$.
- 2 matemáticos de un total de 5 pueden elegirse de ${}_5C_2$ formas y como un físico determinado está en el grupo, esto nos deja que 2 físicos de un nuevo total de 6 se pueden escoger de ${}_6C_2$, entonces, el número de selecciones posibles es de ${}_5C_2 * {}_6C_2 = 10 * 15 = 150$.
- Ya que se eliminan a dos matemáticos del comité, entonces, 2 matemáticos se pueden escoger de un nuevo total de 3 de ${}_3C_2$ y 3 físicos se pueden escoger de un total de 7 de ${}_7C_3$ formas, entonces el número de selecciones posibles es de ${}_3C_2 * {}_7C_3 = 3 * 35 = 105$.

- 3.- ¿Cuántas ensaladas pueden prepararse con lechuga, escarola, endibia, berro y achicoria?

Respuesta:

Primero puede seleccionarse 1 de las 5 verduras (1 ingrediente), luego 2 de las 5 verduras (2 ingredientes), ..., 5 de las 5 verduras (5 ingredientes). Entonces el número requerido de ensaladas es:

$${}_5C_1 + {}_5C_2 + {}_5C_3 + {}_5C_4 + {}_5C_5 = 5 + 10 + 10 + 5 + 1 = 31 = 2^5 - 1.$$

- 4.- Con 7 consonantes y 5 vocales diferentes, ¿cuántas palabras pueden formarse, que consten 4 consonantes y 3 vocales? No es necesario que las palabras tengan significado.

Respuesta:

Las 4 consonantes pueden ser elegidas de 7C_4 y las 3 vocales pueden ser elegidas de 5C_3 . Entonces ya teniendo las 7 letras escogidas, estas se pueden ordenarse entre sí de ${}^7P_7 = 7!$.

Entonces el número de palabras es de: ${}^7C_4 * {}^5C_3 * {}^7P_7 = 35 * 10 * 5040 = 1764000$.

- 5.- Un circuito impreso tiene 8 diferentes lugares en los que un componente puede ser colocado. Si 5 componentes idénticos serán colocados en la tarjeta, ¿Cuántos diseños diferentes son posibles.

Respuesta:

Cada diseño es un subconjunto de los 8 lugares que contendrán a los componentes, entonces el número de posibles combinaciones es $\binom{8}{5} = \frac{8!}{5!(8-5)!} = 56$.

- 6.- Un lote de 50 piezas manufacturadas contiene 3 piezas defectuosas y 47 no defectuosas. Una muestra de 6 piezas son seleccionadas de las 50. Las partes seleccionadas no son remplazadas. Esto significa que cada pieza puede ser seleccionada únicamente una vez, y la muestra es un subconjunto de 50 piezas. ¿Cuántas diferentes muestras existen de tamaño seis que contengan exactamente 2 piezas defectuosas?

Respuesta:

Un subconjunto que contenga exactamente 2 piezas defectuosas puede ser formado por escoger primero dos piezas defectuosas de las 3 defectuosas que existen $\binom{3}{2} = \frac{3!}{2!1!} = 3$ formas diferentes

El segundo paso es seleccionar las 4 partes no defectuosas de las 47 piezas no defectuosas que existen, y esta definida por $\binom{47}{4} = \frac{47!}{4!(47-4)!} = 178365$ formas diferentes

Entonces por la regla de la multiplicación, el número de subconjuntos de seis que contengan exactamente 2 piezas defectuosas es de $3 * 178365 = 535095$.

Si se requiere conocer la probabilidad de que suceda el evento descrito arriba, es necesario dividir el número de subconjuntos de 6 piezas que contengan 2 elementos defectuosos entre el número de subconjuntos de 6 piezas que se pueden formar con 50 piezas

$$\binom{50}{6} = \frac{50!}{6!(50-6)!} = 15890700$$

Entonces la probabilidad de que suceda el evento descrito anteriormente es de $\frac{535095}{15890700} = ,034$.

- 7.- ¿De cuantas formas posibles se pueden combinar letras A,B,C,D para obtener pares diferentes?

Respuesta:

$$\binom{4}{2} = \frac{4!}{2!2!} = 6.$$

- 8.- Obtenga el número de manos posibles de 3 cartas de una baraja de 52 cartas.

Respuesta:

Para obtener el número de manos posibles de 3 cartas, observamos que en este caso el orden de las cartas no importa, entonces, se procede a utilizar una combinación para realizar el cálculo. ${}_{52}C_3 = \binom{52}{3} = \frac{52*51*50}{3*2*1} = 22100$ manos diferentes.

9.- ¿De cuantas formas distintas se pueden seleccionar 6 preguntas de un total de 15?

Respuesta:

Debido a que en este problema el orden no interesa, entonces el número de formas distintas es $\binom{15}{6} = {}_{15}C_6 = 5005$ formas distintas.

10.- ¿Cuántos comités diferentes de 3 hombres y 4 mujeres pueden formarse con 8 hombres y 6 mujeres?

Respuesta:

$$\binom{8}{3} * \binom{6}{4} = {}_8C_3 * {}_6C_4 = 840 \text{ comités diferentes}$$

2.3. Generalización del problema de caminos posibles para el caso de un rectángulo

Los caminos posibles para un rectángulo que está segmentado internamente en cuadrados del mismo tamaño se pueden obtener mediante el conteo de cada uno de los caminos, o implementando un caso general para su solución. Primero, se obtiene el conteo de caminos posibles para el caso de un cuadrado simple y posteriormente de un rectángulo pequeño compuesto por 2 cuadrados, y se define que la trayectoria del camino solo puede ir de arriba hacia abajo y de izquierda a derecha.

Para el caso del cuadrado, obtenemos un número total de caminos igual a: 2 para llegar de A a C



Figura 3: Caminos posibles para un cuadrado.

Se procede a contar el número de caminos posibles para un rectángulo compuesto por 2 cuadrados: Se comprueba

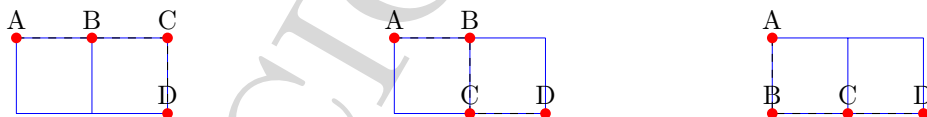


Figura 4: caminos posibles para un rectángulo compuesto por 2 cuadrados.

gráficamente que para un rectángulo compuesto por dos cuadrados hay 3 caminos diferentes para llegar de A a D. Con las 2 comprobaciones gráficas que se realizaron, se puede hacer una relación con el triángulo de Pascal de la siguiente forma:

Para el cuadrado, se pasan obligatoriamente por 2 vértices y cuenta solamente con 1 lado, entonces podemos hacer la relación de ${}_2C_1 = 2$ y si se observa bien en el elemento 2,2 del triángulo de Pascal, tenemos el mismo resultado. Para el rectángulo ocurre algo semejante, se pasan obligatoriamente por 3 vértices y el rectángulo mide 2 cuadrados de ancho por uno de largo, si hacemos la combinación ${}_3C_2 = {}_3C_1 = 3$, y si se observa bien, los elementos 3, 1 y 3, 1 tienen el mismo resultado, entonces, a partir de esto, se puede crear una regla general para cualquier paralelogramo rectangular (compuesto obviamente de paralelogramos rectangulares más pequeños) mediante la siguiente ecuación:

$$\text{Caminos} = \binom{n}{r}. \quad (1)$$

Donde:

n es el número de aristas que se deben cruzar para llegar al punto deseado

r es el número de aristas de los cuadrados que componen cualquier lado del paralelogramo. Esto se cumple por:

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

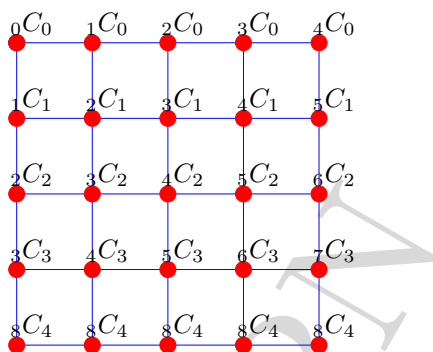


Figura 5: Generalización del numero de caminos posibles a cualquier vértice.

2.4. Generalización del problema de caminos posibles para el caso de un cubo

El caso del cubo es semejante al caso del cuadrado y del rectángulo. Generalizando un espacio de tres dimensiones, si se trata de un arreglo cúbico de $k \times k \times k$ dimensiones, tendremos que el total de caminos posibles (de menor longitud) para ir de un extremo A al extremo B más alejado es:

$$N = \frac{(3kl!)}{(k!)^3}. \quad (2)$$

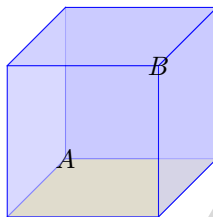


Figura 6: Caminos posibles en un cubo.

Si se trata de un paralelepípedo rectángulo de menor extensión es de:

$$N = \frac{(k + l + m)!}{k!l!m!}. \quad (3)$$

Ejemplo:

Si $k = 1, l = 1$ y $m = 2$, entonces:

$$N = \frac{(1 + 1 + 2)!}{1!1!2!} = 12.$$

Por lo tanto, existen 12 posibles caminos entre A y B

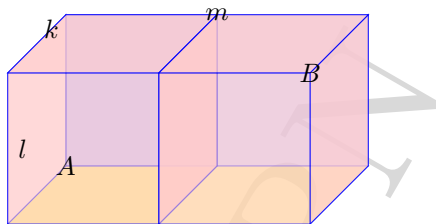


Figura 7: Caminos posibles en un paralelepípedo.

3. Tarea 3

3.1. Ejemplos de Permutaciones Circulares

- 1 En una mesa circular estan sentadas 10 personas, ¿De cuántas formas distintas se pueden sentar en la mesa?

Respuesta:

Dado que estan en una mesa circular y se desea saber de cuantas maneras distintas se pueden ordenar, se utiliza una permutacion circular para resolver el problema.

$$PC_n = (n - 1)! = (10 - 1)! = 9! = 362880.$$

- 2 En una mesa circular, estan sentadas 10 personas, ¿Cuántas formas diferentes hay de sentar a las personas en la mesa si 3 de esas personas se sientan juntas?

Respuesta:

Si consideramos a las 3 personas como un solo elemento en vez de que fueran 3, se puede aplicar el mismo procedimiento que para el problema 1.

$$PC_n = (n - 1)! = (7 - 1)! = 6! = 720.$$

Pero, esas 3 personas pueden acomodarse de diferente forma en esos 3 lugares, el número de formas en las que se pueden ordenar es $3! * 720 = 4320$.

- 3 Del problema anterior, ¿Cual es la probabilidad de que 3 personas especificas se sienten juntas en una mesa de 10 lugares?

Para calcular la probabilidad de que 3 personas se sienten en una mesa circular, se debe relacionar el número de diferentes formas en las que 3 personas se pueden sentar juntas y el número total de formas en las que las 10 personas se sientan en la mesa, y esto seria igual a:

$$P(3personas) = \frac{PC_7 * 3!}{PC_{10}} = \frac{6! * 3!}{9!} = \frac{6! * 3!}{9 * 8 * 7 * 6!} = \frac{3!}{9 * 8 * 7} = \frac{1}{84} = 01,1\%.$$

- 4 Ana y Alejandro invitan a cenar a 3 amigos y 3 amigas ¿Cuántas formas tienen de colocarse en una mesa redonda?

Respuesta:

Son 4 mujeres y 4 hombres, en total son 8 personas, y para calcular de cuantas maneras diferentes pueden sentarse, se utiliza una permutacion circular PC_n

$$PC_8 = (8 - 1)! = 7! = 5040.$$

- 5 Del problema anterior, ¿En cuántas formas están juntos Ana y Alejandro

Respuesta:

$$2 * PC_7 = (n - 1)! * 2! = 6! = 720 * 2 = 1440.$$

- 6 Del problema anterior, ¿De cuántas maneras se pueden sentar las 8 personas sin que Alejandro y Ana se sienten juntas.

Respuesta:

Sabemos que el número de formas en las que Ana y Alejandro se sientan juntos es de 1440, entonces, se deben de remover el número de formas en la que se sientan juntos al número total de formas en las que se sientan las 8 personas.

Respuesta:

$$PC_8 - 2 * PC_7 = 5040 - 1440 = 3600.$$

- 7 Del problema 4, en cuántas ormas pueden sentarse los 4 hombres y las 4 mujeres sin que 2 hombres se sienten juntos.

Respuesta:

Ya que no se quiere que 2 hombres se sienten juntos, la única manera de hacerlo es hacer que los hombres y las mujeres se sienten de forma alternada y para esto las mujeres se sentaran de $PC_3 = 2$ formas diferentes, ahora, los hombres pueden se sentados en los restantes 3 asientos en $3! = 6$ formas diferentes. Notese que la fórmula de la permutacion circular no se ha usado debido a que despues de que las mujeres han sido sentadas, ellas se quedan fijas, y podemos acomodar a los hombres en los lugares correspondientes. Entonces la respuesta es:

$$PC_3 * 3 P_3 = 2! * 3! = 12.$$

- 8 In how many ways can five people A, B, C, D, and E be seated around a circular table if:

a) A and B must sit next to each other?

There are 4 groups ? AB, C, D, E.

$$(4 - 1)! = 3!,$$

then we must take into account that A and B can switch places:

$$3!2! = 6(2) = 12.$$

b) A and B must not sit next to each other?

We can consider the total number of ways five people can sit at a circular table and then subtract the number of ways two people can sit together as found in the first part of this question:

$$(5 - 1)! - 12 = 4! - 12$$

$$= 24 - 12$$

$$= 12.$$

c) A and B must be together and C and D must be together?

There are 3 groups: AB, CD and E. Take into account that AB and CD can switch places:

$$2!2!2! = 8.$$

9 ¿En cuantas formas pueden 4 parejas sentarse en una mesa circular?

Respuesta:

$$2! * PC_4 = 2! * (4 - 1)! = 2! * 3! = 2 * 6 = 12.$$

10 Si 6 personas van a sentarse juntas, pero no deseamos que dos de ellas se sienten juntas, ¿de cuántas formas diferentes se pueden sentar?

Respuesta

La forma en la que dos personas se sienten juntas es:

$$2! * PC_5 = 2!4! = 48,$$

y el total de formas en las que se pueden sentar las 6 personas es

$$PC_6 = 5! = 120,$$

entonces el número de formas posibles de que esas dos personas no se sienten juntas es

$$120 - 48 = 72.$$

3.2. Problemas 1.6, 1.7 y 1.8 de [1]

- 1.6 A palindrome is a finite sequence of characters that reads the same forwards and backwards [GNU DUNG] Find the numbers of 7-digit and 8-digit palindromes, under the restriction that no digit may appear more than twice. By the mirror-symmetry of a palindrome (of length n), only the first $\lfloor (n+1)/2 \rfloor$ positions need be considered. In our case this number is 4 for both lengths. Since the first digit may not be 0, there are 9 ways to fill the first position. There are then $10 - 1 = 9$ ways to fill the second position; $10 - 2 = 8$ ways for the third; $10 - 3 = 7$ ways for the fourth. Thus there are $(9)(9)(8)(7) = 4536$ palindromic numbers of either length.
- 1.7 Prove that a palindromic (decimal) number of even length is divisible by 11. The inductive proof exploits the fact that when the first and last characters are stripped from a palindrome, a palindrome remains. Thus, let N be a palindromic number of length $2k$. If $k = 1$, the theorem obviously holds. If ≥ 2 we have:

$$\begin{aligned} N &= a_{2k-1}10^{2k-1} + a_{2k-2}10^{2k-2} + \dots + a_k10^k + a_k10^{k-1} + \dots + a_{2k-2}10^1 + a_{2k-1}10^0 \\ &= a_{2k-1}(10^{k-1} + 10^0) + (a_{2k-2}10^{2k-2} + \dots + a_{2k-2}10^1) \\ &= a_{2k-1}P + Q. \end{aligned}$$

Here,

$$P = \underbrace{100\dots,001}_{\text{length } 2k} = 11 \times \underbrace{9090\dots,9091}_{\text{length } 2k-2}.$$

And either $Q = 0$ (divisible by 11) or, for some $1 \leq r \leq k-1$

$$Q = 10^r \{ \text{palindrome of length } (1k - r) \} = 10^r \{ 11R \}.$$

where the last step follows from the induction hypothesis. Therefore, N is divisible by 11, and the proof is complete.

- 1.8 In a binary palindrome the first digit is 1 and each succeeding digit may be 0 or 1. Count the binary palindromes of length n . See Problem 1.6. Here we have $\lfloor (n+1)/2 \rfloor - 1 = \lfloor (n-1)/2 \rfloor$ free positions, so the desired number is

$$2^{\lfloor \frac{n-1}{2} \rfloor}.$$

3.3. Números de Catalán

Los números de Catalán forman una secuencia de números naturales que aparece en varios problemas de conteo que habitualmente son recursivos. Obtienen su nombre del matemático belga Eugène Charles Catalan (1814-1894). El n -ésimo número de Catalan se obtiene, aplicando coeficientes binomiales, a partir de la siguiente fórmula:

$$C_n = \frac{(2n)!}{(n+1)!n!} = \frac{1}{n+1} \binom{2n}{n} \quad n \geq 0. \quad (4)$$

Encuentre el número de maneras en que el interior de un polígono convexo de lados, puede ser dividido en triángulos, dibujando diagonales que no se crucen, para $n \geq 3$

Solución:

Hay exactamente una manera de triangular un triángulo

$$T_3 = C_1 = 1.$$

Hay exactamente dos maneras de triangular un cuadrado

$$T_4 = C_2 = 2.$$

Hay cinco diferentes maneras de triangular un pentágono

$$T_5 = C_3 = 5.$$

Hay 14 diferentes maneras de triangular un hexágono

$$T_6 = C_4 = 14.$$

Hay 42 maneras de triangular un polígono de 7 lados

$$T_7 = C_5 = 42.$$

Hay 132 maneras de triangular un polígono de 8 lados

$$T_8 = C_6 = 132.$$

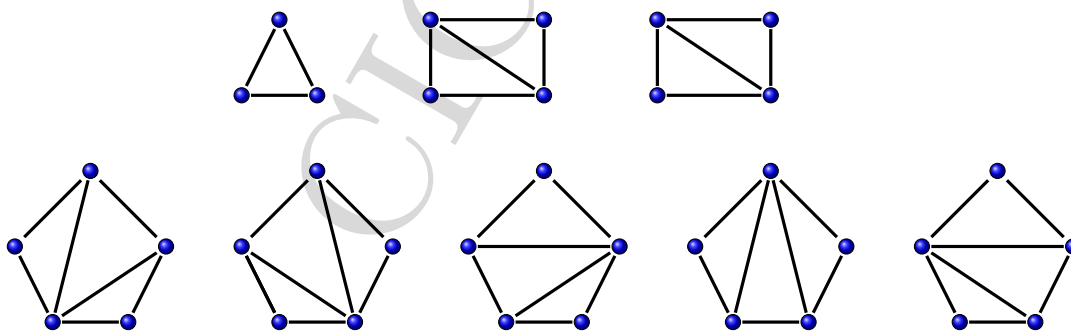


Figura 8: Triangulación de polígonos.

Supongamos que tenemos un conjunto de $n+1$ números para multiplicarlos juntos, lo que significa que hay n multiplicaciones para realizar. Sin cambiar el orden de los números, se pueden multiplicar de las siguientes maneras o en los ordenes especificados en la figura.

| | | |
|---------|---|------------|
| $n = 0$ | (a) | 1 manera. |
| $n = 1$ | $(a * b)$ | 1 manera. |
| $n = 2$ | $((a * b) * c), (a * (b * c))$ | 2 maneras. |
| $n = 3$ | $((((a * b) * c) * d), ((a * b) * (c * d)), ((a * (b * c)) * d), (a * ((b * c) * d)), (a * (b * (c * d))))$ | 5 maneras. |

3.4. Tensor de Levy-Civita

En matemáticas ,y en particular en calculo tensorial, se define el símbolo de Levi-Civita, tambien llamadosímbolo de permutación o tensor de Levi, como sigue:

$$|\overline{A} \times \overline{B}| = \sum_{i,j,k=1}^3 \epsilon_{i,j,k} A_j B_k. \quad (5)$$

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

$$\begin{aligned} (A \times B)_1 &= \epsilon_{1,j,k} A_j B_k = \epsilon_{1,1,k} A_1 B_k + \epsilon_{1,2,k} A_2 B_k + \epsilon_{1,3,k} A_3 B_k \\ &= \epsilon_{1,1,1} A_1 B_1 + \epsilon_{1,1,2} A_1 B_2 + \epsilon_{1,1,3} A_1 B_3 + \epsilon_{1,2,1} A_2 B_1 + \epsilon_{1,2,2} A_2 B_2 + \dots \\ &\quad \dots + \epsilon_{1,2,3} A_2 B_3 + \epsilon_{1,3,1} A_3 B_1 + \epsilon_{1,3,2} A_3 B_2 + \epsilon_{1,3,3} A_3 B_3 \\ (A \times B)_1 &= \epsilon_{2,1,3} A_1 B_3 + \epsilon_{2,3,1} A_3 B_1 = A_2 B_3 - A_3 B_2 \end{aligned}$$

$$\begin{aligned} (A \times B)_2 &= \epsilon_{2,j,k} A_j B_k = \epsilon_{2,1,k} A_1 B_k + \epsilon_{2,2,k} A_2 B_k + \epsilon_{2,3,k} A_3 B_k \\ &= \epsilon_{2,1,1} A_1 B_1 + \epsilon_{2,1,2} A_1 B_2 + \epsilon_{2,1,3} A_1 B_3 + \epsilon_{2,2,1} A_2 B_1 + \epsilon_{2,2,2} A_2 B_2 + \dots \\ &\quad \dots + \epsilon_{2,2,3} A_2 B_3 + \epsilon_{2,3,1} A_3 B_1 + \epsilon_{2,3,2} A_3 B_2 + \epsilon_{2,3,3} A_3 B_3 \\ (A \times B)_2 &= \epsilon_{2,3,1} A_3 B_1 + \epsilon_{2,1,3} A_1 B_1 = A_3 B_1 - A_1 B_3 \end{aligned}$$

$$\begin{aligned} (A \times B)_3 &= \epsilon_{3,j,k} A_j B_k = \epsilon_{3,1,k} A_1 B_k + \epsilon_{3,2,k} A_2 B_k + \epsilon_{3,3,k} A_3 B_k \\ &= \epsilon_{3,1,1} A_1 B_1 + \epsilon_{3,1,2} A_1 B_2 + \epsilon_{3,1,3} A_1 B_3 + \epsilon_{3,2,1} A_2 B_1 + \epsilon_{3,2,2} A_2 B_2 + \dots \\ &\quad \dots + \epsilon_{3,2,3} A_2 B_3 + \epsilon_{3,3,1} A_3 B_1 + \epsilon_{3,3,2} A_3 B_2 + \epsilon_{3,3,3} A_3 B_3 \\ (A \times B)_3 &= \epsilon_{3,1,2} A_1 B_2 + \epsilon_{3,2,1} A_2 B_1 = A_1 B_2 - A_2 B_1 \end{aligned}$$

$$|\overline{A} \times \overline{B}| = (A_2 B_3 - A_3 B_2) \hat{i} + (A_3 B_1 - A_1 B_3) \hat{j} + (A_1 B_2 - A_2 B_1) \hat{k}$$

4. Tarea 4

4.1. Graficaciones en ROOT

4.1.1. Graficar en ROOT $\ln(x!)$



Figura 9: Gráfica de $\ln(x)$.

```
void Grafical() {
    //Author: Javier
    TCanvas *c1 = new TCanvas("c1", "Grafica_de_la_funcion_ln(x!)", 200, 10, 700, 500);

    c1->SetFillColor(42);
    c1->SetGrid();

    const Int_t n = 170;
    Double_t x[n], y[n];
    Double_t fact;
    fact=1;
    for (Int_t i=1; i<n; i++) {
        x[i] = i;
        fact=fact*i;
        y[i] =log(fact);
        printf("_i_ %d _f_ %f \n", i, x[i], y[i]);
    }
    TGraph *gr = new TGraph(n,x,y);
    gr->SetLineColor(2);
    gr->SetLineWidth(4);
    gr->SetMarkerColor(4);
    gr->SetMarkerStyle(21);
    gr->SetTitle("Grafica_1");
    gr->GetXaxis()->SetTitle("X");
    gr->GetYaxis()->SetTitle("ln(x!)");
    gr->Draw("ACP");
    // TCanvas::Update() draws the frame, after which one can change it
    c1->Update();
    c1->GetFrame()->SetFillColor(21);
    c1->GetFrame()->SetBorderSize(12);
    c1->Modified();
}
```


}

CIC-IPN

4.1.2. Graficar en ROOT $x \ln(x) - x$.



Figura 10: Gráfica de $x \ln(x) - x$.

```
void Grafica2() {
    //Author: Javier
    TCanvas *c1 = new TCanvas("c1","Grafica_de_la_funcion_xln(x)-x",200,10,700,500);
    c1->SetFillColor(42);
    c1->SetGrid();
    const Int_t n = 170;
    Double_t x[n], y[n];
    Double_t fact;
    fact=1;
    for (Int_t i=1;i<n;i++) {
        x[i] = i;
        y[i] = (x[i]*log(x[i]))-x[i];
        printf("_i_ %d_ %f_ %d_\n",i,x[i],y[i]);
    }
    TGraph *gr = new TGraph(n,x,y);
    gr->SetLineColor(2);
    gr->SetLineWidth(4);
    gr->SetMarkerColor(4);
    gr->SetMarkerStyle(21);
    gr->SetTitle("Grafica_2");
    gr->GetXaxis()->SetTitle("x");
    gr->GetYaxis()->SetTitle("xln(x)-x");
    gr->Draw("ACP");
    // TCanvas::Update() draws the frame, after which one can change it
    c1->Update();
    c1->GetFrame()->SetFillColor(21);
    c1->GetFrame()->SetBorderSize(12);
    c1->Modified();
}
```

4.1.3. Graficar en ROOT la diferencia de las 2 funciones anteriores.

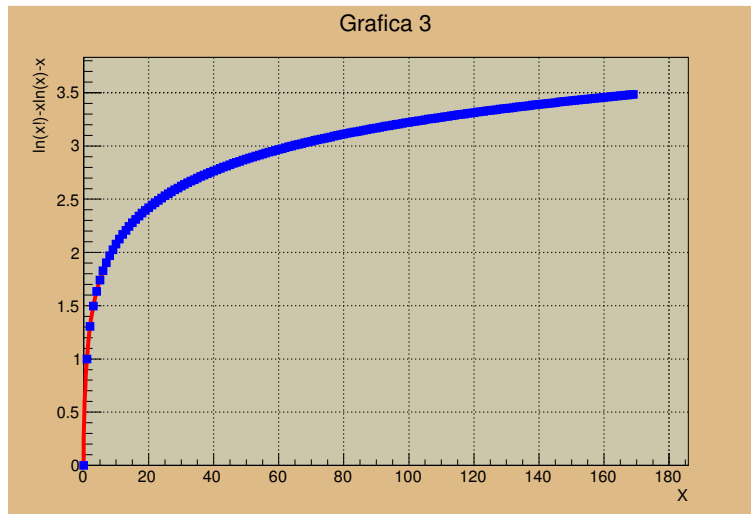


Figura 11: Gráfica de $\ln(x!) - (x\ln(x) - x)$.

```
void Grafica3() {
    //Author: Javier
    TCanvas *c1 = new TCanvas("c1", "Grafica_de_la_funcion_ln(x!)-(xln(x)-x)", 200, 10, 700, 500);
    c1->SetFillColor(42);
    c1->SetGrid();
    const Int_t n = 170;
    Double_t x[n], y[n];
    Double_t fact;
    fact=1;
    for (Int_t i=1; i<n; i++) {
        x[i] = i;
        fact=fact*i;
        y[i] =log(fact)-((x[i]*log(x[i]))-x[i]));
        printf("i_%d_%f_%f\n", i, x[i], y[i]);
    }
    TGraph *gr = new TGraph(n, x, y);
    gr->SetLineColor(2);
    gr->SetLineWidth(4);
    gr->SetMarkerColor(4);
    gr->SetMarkerStyle(21);
    gr->SetTitle("Grafica_3");
    gr->GetXaxis()->SetTitle("X");
    gr->GetYaxis()->SetTitle("ln(x!)-xln(x)-x");
    gr->Draw("ACP");
    // TCanvas::Update() draws the frame, after which one can chan$
    c1->Update();
    c1->GetFrame()->SetFillColor(21);
    c1->GetFrame()->SetBorderSize(12);
    c1->Modified();
}
```

4.2. ¿Que es un histograma?

- a) De [4], un histograma es una aproximación a una función de densidad de probabilidad, vease fig12. Para cada intervalo del histograma, el area de las barras es igual a la frecuencia relativa (proporción o porcentaje) de las mediciones en el intervalo. La frecuencia relativa es una estimación de la probabilidad de que una medición caiga en el intervalo. Del mismo modo, el area bajo $f(x)$ en cualquier intervalo es igual a la verdadera probabilidad de que una medición caiga en el intervalo.

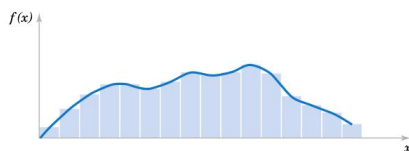


Figura 12: Histograma tomado de [4].

¿Que es una función de densidad de probabilidad?

Para una variable continua aleatoria X , una función de densidad de probabilidad es una función tal que

(1)

$$f(x) \geq 0$$

(2)

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

(3)

$$P(a \leq X \leq b) = \int_a^b f(x) dx = \text{area under } f(x)$$

- b) De [5] Si una muestra o población es grande, es difícil observar diferentes características o computar datos estadísticos, tales como la media, desviación típica, etc. Por esta razón es útil organizar o agrupar los datos. Supongase una muestra que consiste de las estaturas de 100 estudiantes de la universidad XYZ. Ordenamos los datos en clases o categorías, y determinamos el número de individuos que pertenecen a cada clase, denominada la frecuencia de la clase, la ordenación resultante (Cuadro 1) se le conoce como distribución de frecuencia o tabla de frecuencia.

| Estatura | Numero Al. |
|----------|------------|
| 60-62 | 5 |
| 63-65 | 48 |
| 66-68 | 12 |
| 69-71 | 27 |
| 72-74 | 8 |
| Total | 100 |

Cuadro 1: Tabla de Frecuencia.

La primera clase o categoría consiste de estaturas de 60 hasta 62 pulgadas, lo que se llama un intervalo de clase. Puesto que 5 estudiantes tienen estaturas correspondientes a esta clase, la correspondiente frecuencia de clase es 5. Ya que una estatura registrada como 60 pulgadas realmente está entre 59.5 y 60.5 pulgadas, mientras que una registrada como 62 pulgadas realmente está entre 61.5 y 62.5 pulgadas, podríamos haber

registrado el intervalo de la clase como 59.5-62.5. El siguiente intervalo de clase sería entonces 62.5-65.5, etc. En el intervalo de clase 59.5-62.5, los números 59.5 y 62.5 se conocen como límites reales de clase. El ancho del intervalo j -ésimo, denotado por c_j , que comúnmente es el mismo para todas las clases (en cuyo caso se denota por c), es la diferencia entre el límite real superior e inferior. En este caso $c = 62.5 - 59.5 = 3$.

El punto medio del intervalo de clase, que puede tomarse como representativo de la clase, se llama marca de clase. en la tabla anterior, la marca de clase correspondiente a intervalo 60-62 es 61.

Una representación gráfica para la distribución de frecuencia de la tabla 1, puede sumistrarse por un histograma, como se muestra en la figura , o por un polígono de frecuencias uniendo los puntos medios de los techos del histograma.

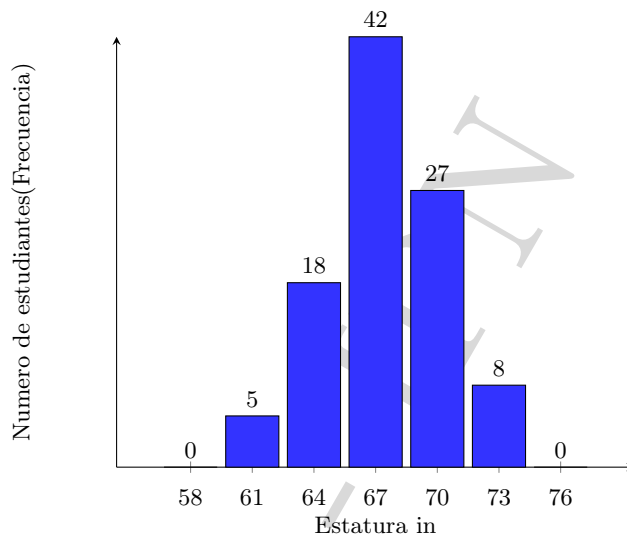


Figura 13: Histograma obtenido a través de la tabla de frecuencias 1.

c) Otro ejemplo de Histogramas

Prepare la tabla de frecuencia compuesto de cinco intervalos para el conjunto de los siguientes 20 datos:

5, 7, 8, 3, 7, 7, 1, 9, 6, 8, 5, 6, 7, 8, 7, 9, 6, 8, 6, 6

Dato mayor = 9, Dato menor = 1. Por tanto, el ancho es: $\frac{9-1}{5} 1,6 \cong 2$

Como el dato menor es 1, seleccionaremos a 2 como la límite superior del primer intervalo. Entonces, el límite superior del segundo intervalo será $2 + 2 = 4$; el del segundo será $4 + 2 = 6$ y así sucesivamente, obtenemos que las límites superiores csaspara los restantes intervalos son: 8, 10, 12.

Por tanto, la tabla de frecuencia es:

Y entonces, el histograma queda de la siguiente manera:

Cuadro 2: Tabla de Frecuencia 2.

| Intervalos | Frecuencia |
|------------|------------|
| 0-2 | 1 |
| 2-4 | 1 |
| 4-6 | 7 |
| 6-8 | 9 |
| 8-10 | 2 |

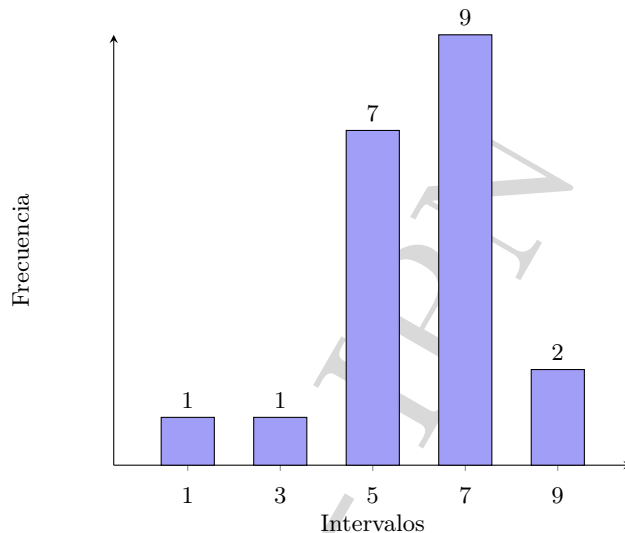


Figura 14: Histograma obtenido a través de la tabla de frecuencias 1.

4.3. Fórmula de Stirling.

La fórmula de Stirling es la siguiente:

$$\lim_{n \rightarrow \infty} \frac{n!}{n^n e^{-n} \sqrt{2\pi n}} = 1. \quad (6)$$

Despejando a $n!$ de la ecuación 6 nos queda:

$$n! \approx n^n e^{-n} \sqrt{2\pi n}. \quad (7)$$

Eso significa que para calcular un límite en el que $n!$ es un factor del numerador o del denominador de la sucesión podemos sustituirlo por ella. Esta sustitución suele ser muy útil en los casos en los que la presencia de $n!$ como factor nos dificulta operar dentro de la sucesión. Vamos a ver una demostración de esta equivalencia para la que usaremos la función gamma de Euler:

$$\Gamma(n+1) = n! = \int_0^\infty x^n e^{-x} dx. \quad (8)$$

Realizando el cambio de variable $x = nt$ sobre la ecuación anterior

$$\Gamma(n+1) = n^{n+1} \int_0^\infty t^n e^{-nt} dt$$

Ahora usando el cambio de variable siguiente:

$$t = 1 + \frac{s}{\sqrt{n}},$$

sobre la ecuación anterior nos queda:

$$\begin{aligned}\Gamma(n+1) &= n! \sqrt{n} = \int_{-\sqrt{n}}^{\infty} \left(1 + \frac{s}{\sqrt{n}}\right)^n e^{-s\sqrt{n}} ds \\ &= n^n e^{-n} \sqrt{n} \int_{-\sqrt{n}}^{\infty} \left(1 + \frac{s}{\sqrt{n}}\right)^n e^{-s\sqrt{n}} ds \\ &= n^n e^{-n} \sqrt{n} \int_{-\sqrt{n}}^{\infty} e^{n \ln \left(1 + \frac{s}{\sqrt{n}}\right) - s\sqrt{n}} ds\end{aligned}\quad (9)$$

Teniendo en cuenta ahora el desarrollo en serie de Taylor de la función $\ln 1 + x$ es:

$$\ln 1 + x = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots$$

El exponente de la exponencia que aparece en la integral queda de la forma:

$$-\frac{s^2}{2} + \frac{s^3}{3\sqrt{n}} - \frac{s^4}{4\sqrt{n}} + \dots$$

Haciendo el límite cuando $n \rightarrow \infty$ se obtiene:

$$\lim_{n \rightarrow \infty} -\frac{s^2}{2} + \frac{s^3}{3\sqrt{n}} - \frac{s^4}{4\sqrt{n}} + \dots = -\frac{s^2}{2}\quad (10)$$

Volviendo a la ecuación 9 y sustituyendo la ecuación 10 y haciendo $n \rightarrow \infty$ en la integral, se obtiene:

$$\begin{aligned}\lim_{n \rightarrow \infty} \Gamma(n+1) &= n^n e^{-n} \sqrt{n} \lim_{n \rightarrow \infty} \int_{-\sqrt{n}}^{\infty} e^{n \ln \left(1 + \frac{s}{\sqrt{n}}\right) - s\sqrt{n}} ds \\ &= n^n e^{-n} \sqrt{n} \int_{-\infty}^{\infty} e^{-\frac{s^2}{2}} ds \\ &= n^n e^{-n} \sqrt{n} \sqrt{2\pi} \\ &= n^n e^{-n} \sqrt{2\pi n}\end{aligned}\quad (11)$$

La aplicación de esta fórmula se usa en la forma logarítmica, como se muestra a continuación.

$$\ln n! = \ln \left(n^n e^{-n} \sqrt{2\pi n} \right)$$

Desarrollando y despreciando el último término la aproximación queda:

$$\ln n! = n \ln n - n\quad (12)$$

4.4. Principio Multiplicativo y Aditivo

Principio Multiplicativo.

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

$$N_1 \times N_2 \times \dots \times N_r \text{ maneras o formas}$$

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

Principio Aditivo.

Si se desea llevar a efecto una actividad, la cuál tiene formas alternativas para ser realizada, donde la primera de esas alternativas puede ser realizada de M maneras o formas, la segunda alternativa puede realizarse de N maneras o formas y la última de las alternativas puede ser realizada de W maneras o formas, entonces esa actividad puede ser llevada a cabo de:

$$M + N + \dots + W \text{ maneras o formas}$$

CIC-IPN

4.5. Numeros grandes

Un gúgol es aproximadamente igual al factorial de 70, nota 1 y sus únicos factores primos son 2 y 5 (cien veces cada uno). Escrito en el sistema binario ocupa 333 bits.

$$1 \text{ gugol} = 10^{100}$$

Un gúgolplex (googolplex en inglés) es un uno seguido de un gúgol de ceros, esto es, 10 elevado a la gugol-ésima potencia

$$1 \text{ gugolplex} = 10^{10^{100}}$$

De la misma forma que el gúgolplex es un uno seguido de gúgol ceros, el gúgolduplex (googolduplex en inglés) es un uno seguido de gúgolplex ceros. Es decir:

$$1 \text{ gugolduplex} = 10^{10^{10^{100}}}$$

4.6. Aportaciones de Srinivasa Aiyangar Ramanujan

Ramanujan ha aportado la siguiente fórmula:

$$1 + \frac{1}{1 \cdot 3} + \frac{1}{1 \cdot 3 \cdot 5} + \frac{1}{1 \cdot 3 \cdot 5 \cdot 7} + \frac{1}{1 \cdot 3 \cdot 5 \cdot 7 \cdot 9} + \cdots + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{\ddots}}}}}}} = \sqrt{\frac{e \cdot \pi}{2}} \quad (13)$$

Se trata de una especie de obra de arte matemática donde se conecta una serie matemática infinita y una fracción continua para aportar así una relación entre dos célebres Constantes matemáticas.

4.6.1. Número de Ramanujan

Se denomina número de Hardy-Ramanujan a todo entero natural que se puede expresar como la suma de dos cubos de dos maneras diferentes.

Hardy comenta la historia detras de estos numeros.

Recuerdo que fui a verle una vez, cuando él ya estaba muy enfermo, en Putney. Había tomado yo un taxi que llevaba el número 1729 y señalé que tal número me parecía poco interesante, y yo esperaba que él no hiciera sino un signo desdeñoso. - "No me respondió- este es un número muy interesante; es el número más pequeño que podemos descomponer de dos maneras diferentes como suma de dos cubos.

$$9^3 + 10^3 = 1^3 + 12^3 = 1729$$

$$2^3 + 16^3 = 9^3 + 15^3 = 4104$$

$$10^3 + 27^3 = 19^3 + 24^3 = 20683$$

5. Tarea 5

5.1. Simulación de un dado de 3 caras

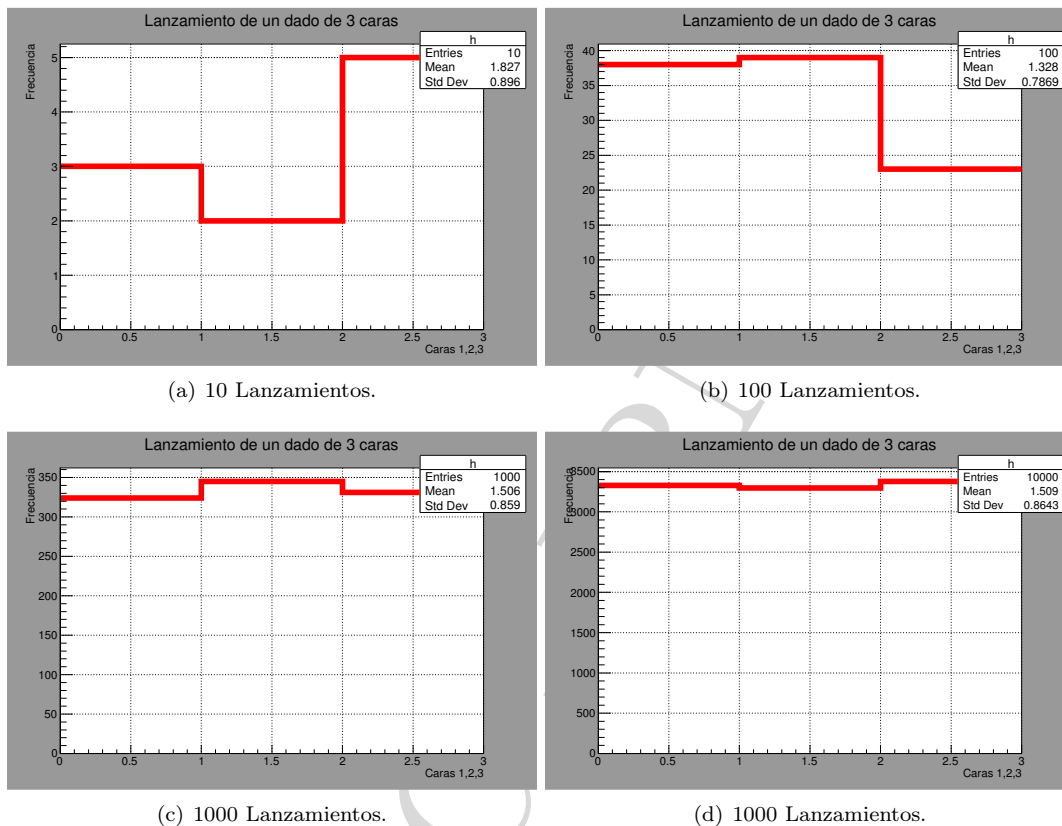


Figura 15: Histogramas del lanzamiento de un dado de 3 caras.

```
void hist_gen()
{
    TCanvas *c1 = new TCanvas("c1", "Simulacion_del_lanzamiento_de_un_dado_de_3_caras"
        ,200,100,700,500);
    TH1F* h1=new TH1F("h", "histograma", 3,0,3);
    h1->SetLineColor(2);
    h1->SetLineWidth(8);
    h1->SetMarkerColor(50);
    h1->SetMarkerStyle(50);
    h1->SetTitle("Lanzamiento_de_un_dado_de_3_caras");
    h1->GetXaxis()->SetTitle("Caras_1,2,3");
    h1->GetYaxis()->SetTitle("Frecuencia");
    TF1 fun1 ("fun1", "2", 0, 3);
    c1->SetFillColor(15);
    c1->SetGrid();
    h1->SetMinimum(0);
    for (int i=0; i<10000; i++){
        h1->Fill(fun1.GetRandom());
    }
}
```

```

h1->Draw();
c1->Update();
c1->GetFrame()->SetFillColor(10);
c1->GetFrame()->SetBorderSize(20);
c1->Modified();
}

```

5.2. Causas de la explosión del Challenger

- Condiciones climáticas no contempladas (Corrientes de chorro)
- Condiciones climáticas ignoradas (Bajas Temperaturas)
- Fallas en las juntas de los tanques de combustible sólido (las juntas no se dilataron lo suficientemente rápido debido a las bajas temperaturas del ambiente)
- Los Factores mencionados fueron los causantes de la fuga de hidrógeno líquido, que causó finalmente la explosión del transbordador.

5.3. Tipos de infinito

En teoría de conjuntos, álef \aleph , primera letra del alfabeto hebreo) es un signo empleado para referirse a ciertos números transfinitos que de hecho resultan ser números ordinales iniciales y por tanto números cardinales.

En análisis matemático aparecen frecuentemente álef 0 y álef 1, aunque pueden definirse números transfinitos arbitrariamente grandes más allá de estos dos. El cardinal álef 0 representa la cantidad de elementos de un conjunto infinito del como el de los números naturales, de hecho este cardinal es el número transfinito más pequeño. Georg Cantor, que inauguró la teoría de conjuntos, demostró que existían diferentes tipos de infinitos incommensurables entre sí, y por tanto no todos los conjuntos infinitos eran equipotentes. De hecho, Cantor demostró que el conjunto de los números reales tenía "más elementos" que los números enteros (si bien ninguno de los dos conjuntos es finito, ambos diferían en su grado de "infinitud"). El número de elementos de la recta real se representó como álef 1.

Puede probarse rigurosamente que dada la clase formada por todos los números ordinales, existe un único isomorfismo (de orden) entre esta clase y la clase de los cardinales transfinitos. Este isomorfismo, denotado como \aleph , se emplea en teoría de conjuntos para construir cardinales transfinitos arbitrariamente grandes. De hecho, dicho isomorfismo es un epimorfismo (isomorfismo suprayectivo) y, por tanto, matemáticamente todos los cardinales transfinitos resultan ser un cardinal de tipo álef.

5.3.1. \aleph_0

El más pequeño de todos los números transfinitos (cardinales), y el más simple de entender conceptualmente es \aleph_0 (se lee como álef sub cero o álef cero), este cardinal es el número de elementos del conjunto de los números naturales. En análisis matemático puede definirse de manera sencilla e intuitiva la clase de conjuntos numerables (conjuntos cuyo cardinal es \aleph_0). Cualquier conjunto que pueda ponerse en correspondencia biunívoca con los números naturales es un conjunto numerable. En términos prácticos, esto significa que los elementos de un conjunto numerable pueden etiquetarse como 1, 2, 3 ... de tal manera que a cada elemento de dicho conjunto le corresponda un número natural (y nada más que un número natural).

Más formalmente, dentro de la teoría axiomática de conjuntos de Zermelo-Fraenkel, el axioma del infinito postula la existencia de un conjunto infinito que puede equipararse fácilmente con los números naturales cuyo cardinal resulta ser \aleph_0 .

5.3.2. \aleph_1

En matemáticas, se define \aleph_1 como el menor cardinal mayor que \aleph_0 , es decir, el menor cardinal mayor que el cardinal del conjunto de los números naturales. Es decir, \aleph_1 es el sucesor de \aleph_0 , lo cual se escribe $\aleph_1 = \aleph_0^+$.

En análisis matemático, se interpreta usualmente al cardinal \aleph_1 como la cantidad de números reales, asumiendo como cierta la hipótesis del continuo. Para justificar esto se parte del teorema de Cantor. Este teorema afirma que el cardinal de $P(N)$ es mayor que \aleph_0 , donde $|P(N)| = |R|$ es el cardinal del conjunto potencia de los números naturales, que es exactamente el mismo que el cardinal de los números reales.

5.3.3. Mas allá de \aleph_1

El teorema de Cantor sobre el conjunto potencia afirma que para cualquier conjunto A se cumple que:

$$|A| < |P(A)|$$

Esto abre la posibilidad de que existan cardinales transfinitos mayores que \aleph_1 . La hipótesis del continuo generalizada de hecho permite ordenar los cardinales transfinitos de manera sencilla, ya que en esencia afirma que:

$$\forall n \geq 0 : |A| = \aleph_n \rightarrow |P(A)| = \aleph_{n+1}$$

5.4. ¿Como se compacta el mp3?

MPEG-1 Audio Layer III o MPEG-2 Audio Layer III, más comúnmente conocido como MP3 es un formato de compresión de audio digital patentado que usa un algoritmo con pérdida para conseguir un menor tamaño de archivo. Es un formato de audio común usado para música tanto en ordenadores como en reproductores de audio portátil.

Los archivos MPEG-1 corresponden a las velocidades de muestreo de 32, 44.1 y 48 kHz.

Los archivos MPEG-2 corresponden a las velocidades de muestreo de 16, 22.05 y 24 kHz.

MP3 fue desarrollado por el Moving Picture Experts Group (MPEG) para formar parte del estándar MPEG-1 y del posterior y más extendido MPEG-2. Un MP3 creado usando una compresión de 128kbit/s tendrá un tamaño de aproximadamente unas 11 veces menor que su homónimo en CD. Un MP3 también puede comprimirse usando una mayor o menor tasa de bits por segundo, resultando directamente en su mayor o menor calidad de audio final, así como en el tamaño del archivo resultante.

En esta capa existen varias diferencias respecto a los estándares MPEG-1 y MPEG-2, entre las que se encuentra el llamado banco de filtros para que el diseño tenga mayor complejidad. Esta mejora de la resolución frecuencial empeora la resolución temporal introduciendo problemas de pre-eco que son predichos y corregidos. Además, permite calidad de audio en tasas tan bajas como 64 kbps.

Banco de Filtros.

El banco de filtros utilizado en esta capa es el llamado banco de filtros híbrido polifase/MDCT. Se encarga de realizar el mapeado del dominio del tiempo al de la frecuencia tanto para el codificador como para los filtros de reconstrucción del decodificador. Las muestras de salida del banco están cuantificadas y proporcionan una resolución en frecuencia variable, 6x32 o 18x32 subbandas, ajustándose mucho mejor a las bandas críticas de las diferentes frecuencias. Usando 18 puntos, el número máximo de componentes frecuenciales es: $32 \times 18 = 576$. Dando lugar a una resolución frecuencial de: $24000/576 = 41,67$ Hz (si $f_s = 48$ kHz.). Si se usan 6 líneas de frecuencia la resolución frecuencial es menor, pero la temporal es mayor, y se aplica en aquellas zonas en las que se espera efectos de pre-eco (transiciones bruscas de silencio a altos niveles energéticos).

La Capa III tiene tres modos de bloque de funcionamiento: dos modos donde las 32 salidas del banco de filtros pueden pasar a través de las ventanas y las transformadas MDCT y un modo de bloque mixto donde las dos bandas de frecuencia más baja usan bloques largos y las 30 bandas superiores usan bloques cortos. Para el caso concreto del MPEG-1 Audio Layer 3 (que concretamente significa la tercera capa de audio para el estándar MPEG-1) especifica cuatro tipos de ventanas: (a) NORMAL, (b) transición de ventana larga a corta (START), (c) 3 ventanas cortas (SHORT)

Modelo Psicoacústico

La compresión se basa en la reducción del margen dinámico irrelevante, es decir, en la incapacidad del sistema auditivo para detectar los errores de cuantificación en condiciones de enmascaramiento. Este estándar divide la señal en bandas de frecuencia que se aproximan a las bandas críticas, y luego cuantifica cada subbanda en función del umbral de detección del ruido dentro de esa banda. El modelo psicoacústico es una modificación del empleado en el esquema II, y utiliza un método denominado predicción polinómica. Analiza la señal de audio y calcula la cantidad de ruido que se puede introducir en función de la frecuencia, es decir, calcula la cantidad de enmascaramiento o umbral de enmascaramiento en función de la frecuencia.

El codificador usa esta información para decidir la mejor manera de gastar los bits disponibles. Este estándar provee dos modelos psicoacústicos de diferente complejidad: el modelo I es menos complejo que el modelo psicoacústico II y simplifica mucho los cálculos. Estudios demuestran que la distorsión generada es imperceptible para el oído experimentado en un ambiente óptimo desde los 192 kbps y en condiciones normales.[cita requerida] Para el oído no experimentado, o común, con 128 kbps o hasta 96 kbps basta para que se oiga "bien" (a menos que se posea un equipo de audio de alta calidad donde se nota excesivamente la falta de graves y se destaca el sonido de "fritura" en los agudos). Las personas que tienen experiencia en la parte auditiva de archivos digitales de audio, especialmente música, desde 192 hasta 256 kbps basta para oír bien, pero la compresión en 320 kbps es la óptima para cualquier escucha. [cita requerida]. La música que circula por Internet, en su mayoría, está codificada entre 128 y 192 kbps, aunque hoy debido al aumento de ancho de banda es cada vez más frecuente compartir archivos en calidad máxima de compresión.

5.5. Examen 4

- 1 ¿De cuántas maneras diferentes pueden sentarse 10 personas en una banca en la que solo caben 4?
- 2 Mencione un ejemplo de un experimento Aleatorio.
- 3 Se quieren sentar 5 hombres y 4 mujeres en una fila, de modo que las mujeres ocupen los sitios pares, ¿De Cuántas formas pueden sentarse ?
- 4 ¿De cuántas formas pueden sentarse 7 personas alrededor de una mesa, si dos personas determinadas no deben estar una al lado de la otra?
- 5 ¿De cuántas formas diferentes pueden sentarse 6 alumnos en un salón de clases con 25 pupitres?
Respuesta
- 6 Mencione un ejemplo de un experimento determinístico.
- 7 ¿De cuantas maneras pueden 10 objetos dividirse en dos grupos de 4 y 6 respectivamente ?
- 8 De un total de 5 matemáticos y 7 físicos, se forma un comité de 2 matemáticos y 3 físicos. ¿De cuántas formas puede formarse, si (a) puede pertenecer a él cualquier matemático y físico, (b) un físico determinado debe pertenecer al comité, (c) dos matemáticos determinados no pueden estar en el comité?
- 9 ¿Cuántos comités diferentes de 3 hombres y 4 mujeres pueden formarse con 8 hombres y 6 mujeres?
- 10 En una mesa circular están sentadas 10 personas, ¿De cuántas formas distintas se pueden sentar en la mesa?

Respuestas:

- 1 En el primer puesto se puede sentar cualquiera de las 10 personas, después, solo hay 9 formas para ocupar el segundo puesto, 8 para formas para el el tercero y finalmente 7 para el cuarto. Entonces, el número de formas distintas en las que 10 personas pueden sentarse en 4 puestos es:

$${}_{10}P_4 = 10 * 9 * 8 * 7 = 5040.$$

- 2 Experimento aleatorio.

- 3 Respuesta:

Los hombres pueden sentarse de ${}_5P_5$ formas, y las mujeres de ${}_4P_4$ formas. Cada ordenación de los hombres puede asociarse con cada ordenación de las mujeres. Así pues

$$\text{Ordenaciones posibles} = {}_5P_5 * {}_4P_4 = 5 * 4 = 2880.$$

- 4 Considérense las dos personas que no deben juntas como una sola. Entonces hay 6 personas para sentarse en círculo, que lo pueden hacer de ${}_5P_5 = 5!$ formas. Pero las dos personas consideradas como una sola pueden ordenarse entre sí de ${}_2P_2 = 2!$ formas. Así pues, el número de ordenaciones de 6 personas sentadas alrededor de una mesa con 2 determinadas de ellas sentadas juntas es de ${}_5P_5 * {}_2P_2 = 5! * 2! = 240$.

Entonces, mediante (a), se tiene el número total de formas en que 6 personas pueden sentarse alrededor de una mesa, entonces si se quiere eliminar los casos en los que dos de las personas no estén sentadas juntas se le debe sustraer al total de formas de ordenar a las 6 personas, el número de casos en los que esa dos personas están juntas.

- 5 El primer alumno puede elegir sentarse en los primeros 25 lugares, el segundo en 24, el tercero en 23, el cuarto en 22, el quinto en 21 y finalmente el sexto en 20, entonces el número de formas en las que se pueden sentar es de ${}_{25}P_6 = 25 * 24 * 23 * 22 * 21 * 20 = 127512000$.

- 6 Experimento determinístico

- 7 Respuesta:

Esto es lo mismo que el número de ordenaciones o permutaciones de 10 objetos, de los cuales 4 de ellos son iguales y los otros 6 de ellos también son iguales entre sí. Entonces se puede usar la expresión ${}_nP_{n_1, n_2} = {}_{10}P_{4,6} = \frac{10!}{6!4!} = 210$.

El problema es equivalente a encontrar el número de grupos de 4 objetos que se pueden formar con 10 objetos dados, (o de 6 objetos con 10 objetos dados), no teniendo en cuenta el orden de los objetos dentro del grupo. En general el número de grupos distintos de r objetos que se pueden formar con n objetos dados, se le llama número de combinaciones de los n objetos tomados de r en r y se denota por ${}_nC_r$ ó $\binom{n}{r}$, y esto viene dado por:

$${}_nC_r = \binom{n}{r} = \frac{n!}{r!(n-r)!} = \frac{n(n-1)\dots(n-r+1)}{r!} = \frac{{}_nP_r}{r!}.$$

Entonces si $n = 10$ y $r = 6$ ó $r = 4$, entonces:

$$\binom{n}{r} = \binom{10}{6} = \binom{10}{4} = 210.$$

- 8 a) 2 matemáticos de un total de 5 pueden elegirse de ${}_5C_2$ formas y 3 físicos pueden escogerse de un total de 7 de ${}_7C_3$. Entonces, el número de selecciones posibles es ${}_5C_2 * {}_7C_3 = 10 * 35 = 350$
- b) 2 matemáticos de un total de 5 pueden elegirse de ${}_5C_2$ formas y como un físico determinado está en el grupo, esto nos deja que 2 físicos de un nuevo total de 6 se pueden escoger de ${}_6C_2$, entonces, el número de selecciones posibles es de ${}_5C_2 * {}_6C_2 = 10 * 15 = 150$

- c) Ya que se eliminan a dos matemáticos del comité, entonces, 2 matemáticos se pueden escoger de un nuevo total de 3 de ${}_3C_2$ y 3 físicos se pueden escoger de un total de 7 de ${}_7C_3$ formas, entonces el número de selecciones posibles es de ${}_3C_2 * {}_7C_3 = 3 * 35 = 105$

9 Respuesta

$$\binom{8}{3} * \binom{6}{4} = {}_8C_3 * {}_6C_4 = 840 \text{ comités diferentes}$$

10 Respuesta:

Dado que están en una mesa circular y se desea saber de cuantas maneras distintas se pueden ordenar, se utiliza una permutación circular para resolver el problema.

$$PC_n = (n - 1)! = (10 - 1)! = 9! = 362880$$

CIC-IPN

5.6. Máquina enigma

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

La máquina Enigma fue un dispositivo electromecánico, lo que significa que usaba una combinación de partes mecánicas y eléctricas. El mecanismo estaba constituido fundamentalmente por un teclado similar al de las máquinas de escribir cuyas teclas eran interruptores eléctricos, un engranaje mecánico y un panel de luces con las letras del alfabeto. (Vease imagen 5.6)

La parte eléctrica consistía en una batería que encendía una lámpara de una serie de ellas, que representan cada



Figura 16: Máquina enigma

una de las diferentes letras del alfabeto. Se puede observar en la parte inferior de la imagen adjunta el teclado, y las lámparas son los círculos que aparecen encima de éste.

El corazón de la máquina Enigma era mecánico y constaba de varios rotores conectados entre sí. Cada rotor es un disco circular plano con 26 contactos eléctricos en cada cara, uno por cada letra del alfabeto vease imagen (5.6). Cada contacto de una cara está conectado o cableado a un contacto diferente de la cara contraria. Por ejemplo, en

un rotor en particular, el contacto número 1 de una cara puede estar conectado con el contacto número 14 en la otra cara y el contacto número 5 de una cara con el número 22 de la otra. Cada uno de los cinco rotores proporcionados con la máquina Enigma estaba cableado de una forma diferente y los rotores utilizados por el ejército alemán poseían un cableado distinto al de los modelos comerciales.

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



Figura 17: Rodillos de la Máquina enigma

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

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

avance de cada uno de los rotores, era un parámetro configurable por el operario.

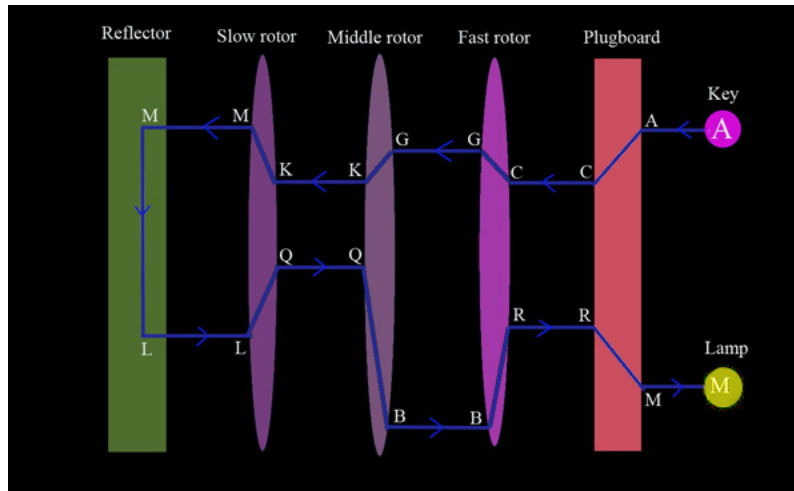


Figura 18: Conexión interna de la Máquina enigma

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

6. Tarea 6

6.1. Realizar un programa en gcc controlado por make que resuelva una ecuación cuadrática.

Programa en C

[illegible]

Makefile

```
ecuacion: qad.c
          gcc -o ecuacion qad.c -lm
clean:
          rm -v *.o ecuacion
```

6.2. Realizar un histograma en ROOT que se llene con la distribución de Gauss.

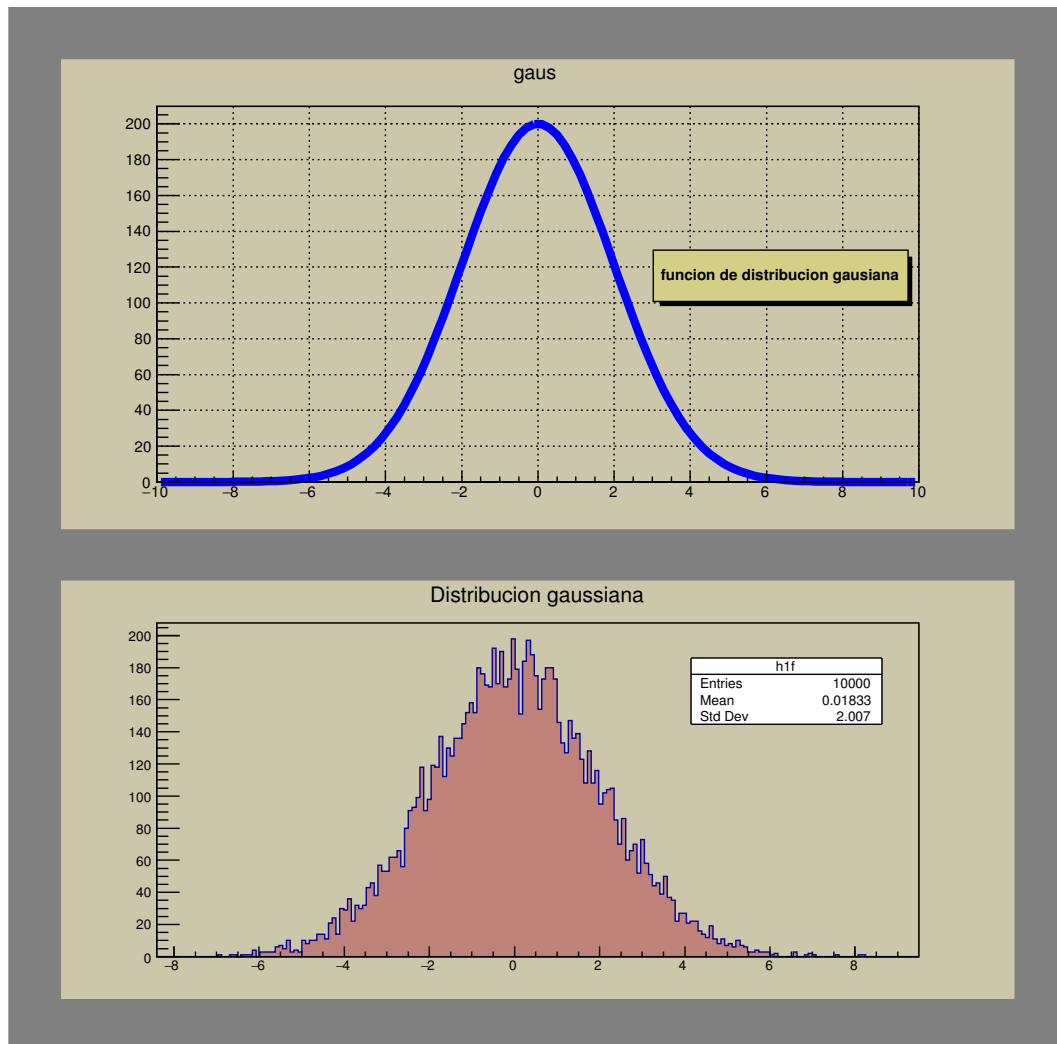


Figura 19: Histograma que se ajusta a la distribución de Gauss

```
void gauss() {
    //Fill a 1-D histogram from a parametric function
    // To see the output of this macro, click begin\_html <a href="gif/fillrandom.gif">here</a>.
    end_html
    //Author: Rene Brun

    TCanvas *c1 = new TCanvas("c1", "The_FillRandom_example", 200, 10, 700, 900);
    c1->SetFillColor(14);

    auto pad1 = new TPad("pad1", "The_pad_with_the_function", 0.05, 0.50, 0.95, 0.95, 21);
    auto pad2 = new TPad("pad2", "The_pad_with_the_histogram", 0.05, 0.05, 0.95, 0.45, 21);
    pad1->Draw();
    pad2->Draw();
    pad1->cd();

    gBenchmark->Start("fillrandom");
}
```

```

//
// A function (any dimension) or a formula may reference
// an already defined formula
//
auto form1 = new TFormula("form1","abs(sin(x)/x)");
auto sqroot = new TF1("sqroot","gaus_",-10,10);//+ [3]*form1
sqroot->SetParameters(200,0,2,20);
pad1->SetGridx();
pad1->SetGridy();
pad1->GetFrame()->SetFillColor(42);
pad1->GetFrame()->SetBorderMode(-1);
pad1->GetFrame()->SetBorderSize(5);
sqroot->SetLineColor(4);
sqroot->SetLineWidth(6);
sqroot->Draw();
auto lfunction = new TPaveLabel(5,39,9.8,46,"funcion_de_distribucion_gausiana");
lfunction->SetFillColor(41);
lfunction->Draw();
c1->Update();

//
// Create a one dimensional histogram (one float per bin)
// and fill it following the distribution in function sqroot.
//
pad2->cd();
pad2->GetFrame()->SetFillColor(42);
pad2->GetFrame()->SetBorderMode(-1);
pad2->GetFrame()->SetBorderSize(5);
auto h1f = new TH1F("h1f","Distribucion_gaussiana",200,10,10);
h1f->SetFillColor(45);
h1f->FillRandom("sqroot",10000);
h1f->Draw();
c1->Update();
//
// Open a ROOT file and save the formula, function and histogram
//
TFile myfile("fillrandom.root","RECREATE");
form1->Write();
sqroot->Write();
h1f->Write();
gBenchmark->Show("fillrandom");
}

```

6.3. Usar root para obtener el histograma 1.2.3 de [2].

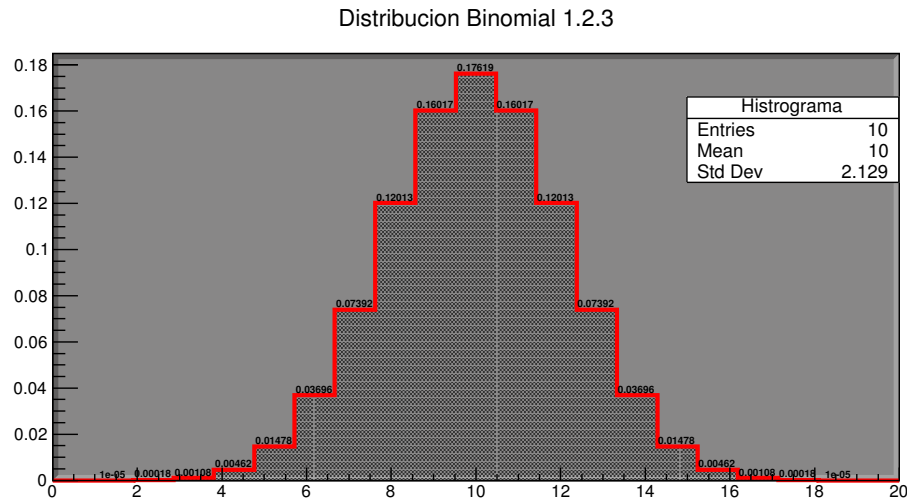


Figura 20: Histograma 1.2.3 de [2]

```
void funcion_binomial(){
    TCanvas *c1 = new TCanvas("c1", "Distribucion_binomial_plot_1.2.3", 200, 52, 700, 500);
    c1->Range(-0.375, -0.525, 3.375, 4.725);
    c1->SetFillColor(24);
    c1->SetBorderMode(0);
    c1->SetBorderSize(2);
    c1->SetFrameBorderMode(0);
    c1->SetFrameFillColor(35);
    c1->SetFrameBorderMode(0);
    c1->SetFrameBorderSize(10);

    TH1F *h__1 = new TH1F("Histograma", "Distribucion_Binomial_1.2.3", 21, 0, 20);
    h__1->SetBinContent(1, 0);
    h__1->SetBinContent(2, .00001);
    h__1->SetBinContent(3, .00018);
    h__1->SetBinContent(4, .00108);
    h__1->SetBinContent(5, .00462);
    h__1->SetBinContent(6, .01478);
    h__1->SetBinContent(7, .03696);
    h__1->SetBinContent(8, .07392);
    h__1->SetBinContent(9, .12013);
    h__1->SetBinContent(10, .16017);
    h__1->SetBinContent(11, .17619);
    h__1->SetBinContent(12, .16017);
    h__1->SetBinContent(13, .12013);
    h__1->SetBinContent(14, .07392);
    h__1->SetBinContent(15, .03696);
    h__1->SetBinContent(16, .01478);
    h__1->SetBinContent(17, .00462);
    h__1->SetBinContent(18, .00108);
    h__1->SetBinContent(19, .00018);
    h__1->SetBinContent(20, .00001);
    h__1->SetBinContent(21, 0);
    h__1->SetMinimum(0);
    h__1->SetEntries(10);
}
```

```
TPaveStats *ptstats = new TPaveStats(0.78,0.775,0.98,0.935,"brNDC");
ptstats->SetName(" stats");
ptstats->SetBorderSize(1);
ptstats->SetFillColor(0);
ptstats->SetTextAlign(12);
ptstats->SetTextFont(42);

TText *AText = ptstats->AddText("h");
AText->SetTextSize(0.0368);
AText = ptstats->AddText(" Entries_=_10_");
AText = ptstats->AddText(" Mean_=_1.359");
AText = ptstats->AddText(" Std_Dev_=_0.6569");
ptstats->SetOptStat(1111);
ptstats->SetOptFit(0);
ptstats->Draw();
c1->Modified();
c1->cd();
c1->SetSelected(c1);
gStyle->SetPaintTextFormat("g");
h__1->SetMarkerSize(1);
c1->cd(1);
h__1->Draw("HIST_TEXT0");
}
```

6.4. Mínimos cuadrados.

Técnica de análisis numérico enmarcada dentro de la optimización matemática.

Dados un conjunto de pares ordenados: variable independiente, variable dependiente, se intenta encontrar la función continua, que mejor se aproxime a los datos de acuerdo con el criterio de mínimo error cuadrático.

Se asume que se tiene una cantidad n de puntos (x_i, y_i) y se requiere aproximar la tendencia de estos mediante una recta de la forma

$$f(x) = y = a + bx.$$

Se pretende hacer que todos los puntos pasen por la recta, de esta forma se obtienen n ecuaciones de recta de la forma:

$$\hat{y}_1 = a + bx_1$$

$$\hat{y}_2 = a + bx_2$$

$$\vdots$$

$$\hat{y}_n = a + bx_n$$

En forma matricial:

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{bmatrix}. \quad (14)$$

Pero para todo x_i habrá un \hat{y}_i diferente de y_i por que la recta no pasa realmente por todos los puntos. Entonces se tiene un error para cada punto. El error se mide en forma de distancias entre y_i y \hat{y}_i . El error para cada x_i se define:

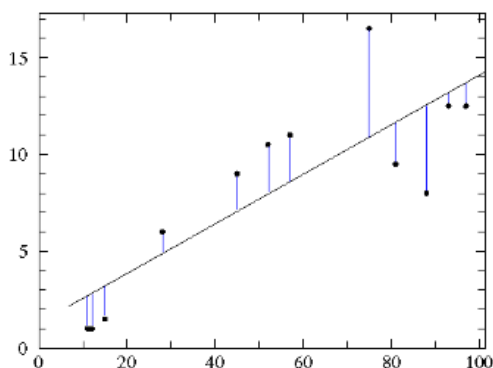


Figura 21: Graficación del error de los puntos en el ajuste

$$e_i^2 = (y_i - \hat{y}_i)^2 = (y_i - (a + bx_i))^2.$$

De forma matricial se obtiene:

$$\begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} = \begin{bmatrix} y_1 - (a + bx_1) \\ y_2 - (a + bx_2) \\ \vdots \\ y_n - (a + bx_n) \end{bmatrix}. \quad (15)$$

Se busca los valores de a y b de forma que se obtenga el error mínimo

$$\left\| \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} \right\| = \sqrt{e_1^2 + e_2^2 + \dots + e_n^2} \rightarrow \varepsilon^2 = \sum_{i=1}^n [e_i^2]. \quad (16)$$

Se intenta minimizar el error entre la recta y todos los puntos. Se busca este mínimo por medio del cálculo diferencial. Para ello se optimiza la función de error ε^2 , lo cual implica derivarla con respecto a cada una de las variables a y b . Derivando la ecuación (3)

$$\begin{aligned} \frac{\partial \varepsilon^2}{\partial a} &= \frac{\partial}{\partial a} \left(\sum_{i=1}^n [y_i - (a + bx_i)]^2 \right) = \sum_{i=1}^n \left[\frac{\partial}{\partial a} [y_i - (a + bx_i)]^2 \right], \\ \frac{\partial \varepsilon^2}{\partial b} &= \frac{\partial}{\partial b} \left(\sum_{i=1}^n [y_i - (a + bx_i)]^2 \right) = \sum_{i=1}^n \left[\frac{\partial}{\partial b} [y_i - (a + bx_i)]^2 \right]. \end{aligned}$$

Desarrollando e igualando a cero para encontrar el mínimo de la función:

$$\begin{aligned} \frac{\partial \varepsilon^2}{\partial a} &= -2 \sum_{i=1}^n [y_i - (a + bx_i)] = 0, \\ \frac{\partial \varepsilon^2}{\partial b} &= -2 \sum_{i=1}^n [y_i - (a + bx_i)] x_i = 0. \end{aligned}$$

Distribuyendo la sumatoria a cada termino y sabiendo q a y b son constantes:

$$\sum_{i=1}^n y_i = an + b \sum_{i=1}^n x_i, \quad (17)$$

$$\sum_{i=1}^n x_i y_i = a \sum_{i=1}^n x_i + b \sum_{i=1}^n x_i^2. \quad (18)$$

En forma matricial:

$$\begin{bmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \end{bmatrix} \Rightarrow Bx = z. \quad (19)$$

Se busca despejar los coeficientes a y b . Se procede a calcular la matriz inversa por determinantes para despejar dichas variables.

$$\det |B| = \det \begin{bmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{bmatrix} = n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2. \quad (20)$$

La matriz inversa queda de la forma:

$$B^{-1} = \begin{bmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{bmatrix}^{-1} = \frac{1}{\det |B|} \begin{bmatrix} \sum_{i=1}^n x_i^2 & -\sum_{i=1}^n x_i \\ -\sum_{i=1}^n x_i & n \end{bmatrix}, \quad (21)$$

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \frac{\sum_{i=1}^n (x_i^2)}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} & \frac{-\sum_{i=1}^n x_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \\ \frac{-\sum_{i=1}^n x_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} & \frac{n}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \end{bmatrix} \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \end{bmatrix}, \quad (22)$$

$$a = \frac{\sum_{i=1}^n (x_i^2) \sum_{i=1}^n y_i - \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}, \quad (23)$$

$$b = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}. \quad (24)$$

Reescribiendo la ecuación (23) y (24).

$$a = \frac{\bar{y} \sum_{i=1}^n (x_i^2) - \bar{x} \sum_{i=1}^n x_i y_i}{n \sum_{i=1}^n x_i^2 - n \bar{x}^2}, \quad (25)$$

$$b = \frac{\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y}}{n \sum_{i=1}^n x_i^2 - n \bar{x}^2} \quad (26)$$

Realice el ajuste por mínimos cuadrados del siguiente conjunto de datos.

| x | y | xy | x ² |
|----|----|-----|----------------|
| 8 | 3 | 24 | 64 |
| 2 | 10 | 20 | 4 |
| 11 | 3 | 33 | 121 |
| 6 | 6 | 36 | 36 |
| 5 | 8 | 40 | 25 |
| 4 | 12 | 48 | 16 |
| 12 | 1 | 12 | 144 |
| 9 | 4 | 36 | 81 |
| 6 | 9 | 54 | 36 |
| 1 | 14 | 14 | 1 |
| 64 | 70 | 317 | 528 |

Cuadro 3: Datos.

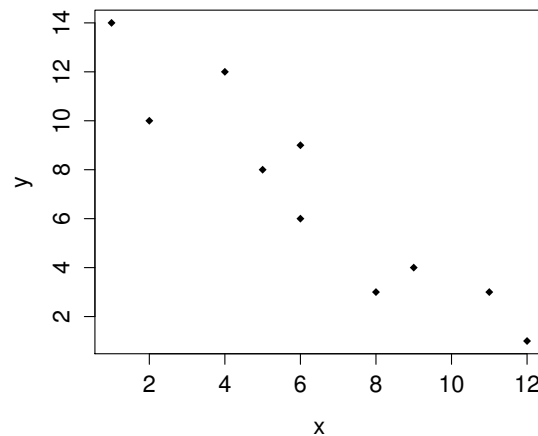


Figura 22: Graficación de los datos de la tabla 3

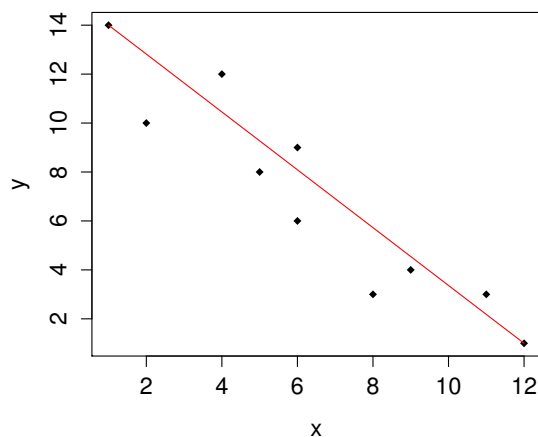


Figura 23: Recta calculada con mínimos cuadrados.

Calculamos \bar{x} y \bar{y} .

$$\bar{x} = \frac{64}{10} = 6,4$$

$$\bar{y} = \frac{70}{10} = 7$$

Ahora calculamos a y b .

$$a = \frac{\bar{y} \sum_{i=1}^n (x_i^2) - \bar{x} \sum_{i=1}^n x_i y_i}{n \sum_{i=1}^n x_i^2 - n \bar{x}^2} = \frac{(7 * 528) - (6,4 * 317)}{528 - (10 * (6,4^2))} = 14,0810$$

$$b = \frac{\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y}}{n \sum_{i=1}^n x_i^2 - n \bar{x}^2} = b = \frac{317 - (10 * 6,4 * 7)}{528 - (10 * (6,4^2))} = -1,1064$$

6.5. Densidad de probabilidad

En la teoría de la probabilidad, la función de densidad de probabilidad, función de densidad, o 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 (FDPs) es no-negativa a lo largo de todo su dominio y su integral sobre todo el espacio es de valor unitario.

$$f(x) \geq 0; \forall x.$$

$$\int_{-\infty}^{\infty} f(x)dx = 1.$$

Si se desea calcular la probabilidad de que la variable aleatoria x se encuentre entre los valores a y b o sea $P(a \leq x \leq b)$.

Se tienen los siguientes datos sobre la probabilidad de que elegir un trabajador que tenga de 0 a 7 años laborando para una empresa:

| Años labor | Probabilidad |
|------------|--------------|
| 0-1 | .1 |
| 1-2 | .26 |
| 2-3 | .28 |
| 3-4 | .2 |
| 4-5 | .11 |
| 5-6 | .04 |
| 6-7 | .01 |

Cuadro 4: Datos para la función de densidad de probabilidad

A partir de la tabla (4) se construye el histograma y la función de distribución y obtenemos que una de a a b probabilidad es el área bajo la curva entre dichos puntos.

La curva a la derecha es la gráfica de cualquier función f , que se llama una función de densidad de probabilidad.

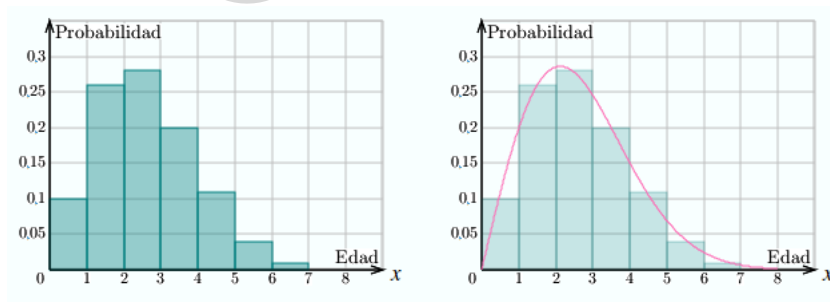


Figura 24: Histograma y la función de densidad de probabilidad.

Tomamos para el dominio de f el intervalo $[0+)$, pues este intervalo es el rango de los valores posibles que pueden

tomar X . Además, usamos x para referir a valores particulares de X , así que no es coincidencia que aquellos valores son mostrados en el eje- x . Por lo general, una función de densidad de probabilidad tendrá cualquier (posiblemente no acotado) intervalo como su dominio. Si se desea calcular la probabilidad de que un trabajador tenga entre 2 a 3.5 años entonces se tiene que:

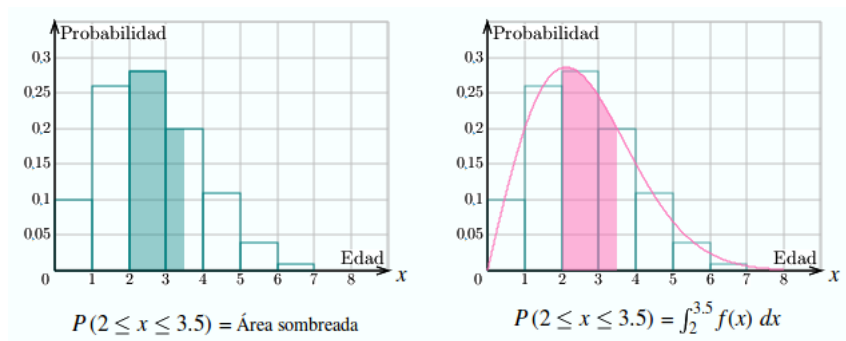


Figura 25: Cálculo de la probabilidad con histograma y con la función.

7. Tarea 7

7.1. Eigen valores y eigen vectores.

En álgebra lineal, los vectores propios, autovectores o eigenvectores de un operador lineal son los vectores no nulos que, cuando son transformados por el operador, dan lugar a un múltiplo escalar de sí mismos, con lo que no cambian su dirección. Este escalar λ recibe el nombre valor propio, autovalor, valor característico o eigenvalor. A menudo, una transformación queda completamente determinada por sus vectores propios y valores propios. Un espacio propio, autoespacio, eigenspacio o subespacio fundamental asociado al valor propio λ es el conjunto de vectores propios con un valor propio común.

La palabra alemana eigen, que se traduce en español como propio, se usó por primera vez en este contexto por David Hilbert en 1904 (aunque Helmholtz la usó previamente con un significado parecido). Eigen se ha traducido también como inherente, característico o el prefijo auto-, donde se aprecia el énfasis en la importancia de los valores propios para definir la naturaleza única de una determinada transformación lineal. Las denominaciones vector y valor característicos también se utilizan habitualmente.

Si se quiere calcular los valores propios de una matriz dada y ésta es pequeña, se puede calcular simbólicamente usando el polinomio característico. Sin embargo, a menudo resulta imposible para matrices extensas, caso en el que se debe usar un método numérico.

Cálculo de los valores propios.

Una herramienta importante para encontrar valores propios de matrices cuadradas es el polinomio característico: decir que λ es un valor propio de A es equivalente a decir que el sistema de ecuaciones lineales $Av = \lambda v \rightarrow Av - \lambda v = 0$ (factorizando por v queda) $(A - \lambda I)v = 0$ (donde I es la matriz identidad) tiene una solución no nula v (un vector propio), y de esta forma es equivalente al determinante:

$$\det(A - \lambda I) = 0. \quad (27)$$

La función $p(\lambda) = \det(A - \lambda I)$ es un polinomio de λ pues los determinantes se definen como sumas de productos. éste es el polinomio característico de A : los valores propios de una matriz son los ceros de su polinomio característico.

Todos los valores propios de una matriz A pueden calcularse resolviendo la ecuación $p_A(\lambda) = 0$.

Si A es una matriz $n \times n$, entonces p_A tiene grado n y A tiene como máximo n valores propios.

El teorema fundamental del álgebra dice que esta ecuación tiene exactamente n raíces (ceros), teniendo en cuenta su multiplicidad. Todos los polinomios reales de grado impar tienen un número real como raíz, así que para n impar toda matriz real tiene al menos valor propio real. En el caso de las matrices reales, para n par e impar, los valores propios no reales son pares conjugados.

Cálculo de los vectores propios.

Una vez que se conocen los valores propios λ , los vectores propios se pueden hallar resolviendo el sistema de ecuaciones homogéneo:

$$(A - \lambda I)v = 0 \quad (28)$$

Una forma más sencilla de obtener vectores propios sin resolver un sistema de ecuaciones lineales se basa en el teorema de Cayley-Hamilton que establece que cada matriz cuadrada satisface su propio polinomio característico. Así, si $\lambda_1, \lambda_2, \dots, \lambda_n$ son los valores propios de A se cumple que:

$$(A - \lambda_1 I)(A - \lambda_2 I) \dots (A - \lambda_n I) = 0 \quad (29)$$

por lo que los vectores columna de $(A - \lambda_2 I) \dots (A - \lambda_n I)$ son vectores propios de λ_1 .

Ejemplo:

Considérese la matriz:

$$\begin{bmatrix} 0 & 1 & -1 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix},$$

que representa un operador lineal $R^3 \rightarrow R^3$. Si se desea computar todos los valores propios de A , se podría empezar determinando el polinomio característico:

$$\begin{aligned} p(x) = \det(A - xI) &= \det \begin{bmatrix} -x & 1 & -1 \\ 1 & 1-x & 0 \\ -1 & 0 & 1-x \end{bmatrix}, \\ &= -x^3 + 2x^2 + x - 2, \end{aligned}$$

y debido a que $p(x) = -(x-2)(x-1)(x+1)$ se ve que los valores propios de A son 2, 1 y -1. El teorema de Cayley-Hamilton establece que cada matriz cuadrada satisface su propio polinomio característico. Es decir: $(A - 2I)(A - I)(A + I) = 0$

Efectivamente, para el caso del valor propio 2, se puede comprobar que:

$$(A - I)(A + I) = \begin{bmatrix} -1 & 1 & -1 \\ 1 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & -1 \\ 1 & 2 & 0 \\ -1 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix},$$

de donde $(1, 1, -1)$ es un vector propio asociado a 2.

$$A \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ -2 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}.$$

7.2. Análisis de componentes principales.

Cuando se recoge la información de una muestra de datos, lo más frecuente es tomar el mayor número posible de variables. Sin embargo, si tomamos demasiadas variables sobre un conjunto de objetos, por ejemplo 20 variables, tendremos que considerar $\binom{20}{2} = 180$ posibles coeficientes de correlación; si son 40 variables dicho número aumenta hasta 780. Evidentemente, en este caso es difícil visualizar relaciones entre las variables. Otro problema que se presenta es la fuerte correlación que muchas veces se presenta entre las variables: si tomamos demasiadas variables (cosa que en general sucede cuando no se sabe demasiado sobre los datos o sólo se tiene ánimo exploratorio), lo normal es que estén relacionadas o que midan lo mismo bajo distintos puntos de vista. Por ejemplo, en estudios médicos, la presión sanguínea a la salida del corazón y a la salida de los pulmones están fuertemente relacionadas. Se hace necesario, pues, reducir el número de variables. Es importante resaltar el hecho de que el concepto de mayor información se relaciona con el de mayor variabilidad o varianza. Cuanto mayor sea la variabilidad de los datos (varianza) se considera que existe mayor información.

Estas técnicas fueron inicialmente desarrolladas por Pearson a finales del siglo XIX y posteriormente fueron estudiadas por Hotelling en los años 30 del siglo XX. Sin embargo, hasta la aparición de los ordenadores no se empezaron a popularizar. Para estudiar las relaciones que se presentan entre p variables correlacionadas (que miden información común) se puede transformar el conjunto original de variables en otro conjunto de nuevas variables incorreladas entre sí (que no tenga repetición o redundancia en la información) llamado conjunto de componentes principales. Las nuevas variables son combinaciones lineales de las anteriores y se van construyendo según el orden de importancia en cuanto a la variabilidad total que recogen de la muestra. De modo ideal, se buscan $m < p$ variables que sean combinaciones lineales de las p originales y que estén incorreladas, recogiendo la mayor parte de la información o variabilidad de los datos. Si las variables originales están incorreladas de partida, entonces no tiene sentido realizar un análisis de componentes principales. El análisis de componentes principales es una técnica matemática que no requiere la suposición de normalidad multivariante de los datos, aunque si esto último se cumple se puede dar una interpretación más profunda de dichos componentes.

Cálculo de las componentes principales.

Se considera una serie de variables (x_1, x_2, \dots, x_p) sobre un grupo de objetos o individuos y se trata de calcular, a partir de ellas, un nuevo conjunto de variables y_1, y_2, \dots, y_p , incorreladas entre sí, cuyas varianzas vayan decreciendo progresivamente. Cada y_j (donde $j = 1, \dots, p$) es una combinación lineal de las x_1, x_2, \dots, x_p originales, es decir:

$$\begin{aligned} y_j &= a_{j1}x_1 + a_{j2}x_2 + \dots + a_{jp}x_p \\ &= a'_{jx}, \end{aligned} \quad (30)$$

siendo $a'_{jx} = (a_{1j}, a_{2j}, \dots, a_{pj})$ un vector de constantes, y

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}. \quad (31)$$

Obviamente, si lo que queremos es maximizar la varianza, como veremos luego, una forma simple podría ser aumentar los coeficientes a_{ij} . Por ello, para mantener la ortogonalidad de la transformación se impone que el módulo del vector $a'_{jx} = (a_{1j}, a_{2j}, \dots, a_{pj})$ sea 1.

Es decir:

$$a_{pj}a_j = \sum_{k=1}^p a_{kj}^2 = 1.$$

El primer componente se calcula eligiendo a_1 de modo que y_1 tenga la mayor varianza posible, sujeta a la restricción de que $a_1^T a_1 = 1$. El segundo componente principal se calcula obteniendo a_2 de modo que la variable obtenida, y_2 esté incorrelada con y_1 . Del mismo modo se eligen y_1, y_2, \dots, y_p , incorrelados entre sí, de manera que las variables aleatorias obtenidas vayan teniendo cada vez menor varianza.

Proceso de extracción de factores:

Queremos elegir a_1 de modo que se maximice la varianza de y_1 sujeta a la restricción de que $a_1^T a_1 = 1$

$$\text{Var}(y_1) = \text{Var}(a_1^T x) = a_1^T \Sigma a_1.$$

El método habitual para maximizar una función de varias variables sujeta a restricciones el método de los multiplicadores de Lagrange. El problema consiste en maximizar la función $a_1^T \Sigma a_1$ sujeta a la restricción $a_1^T a_1 = 1$. Se puede observar que la incógnita es precisamente a_1 (el vector desconocido que nos da la combinación lineal óptima). Así, construyo la función L :

$$L(a_1) = a_1^T \Sigma a_1 - \lambda(a_1^T a_1 - 1).$$

y busco el máximo, derivando e igualando a 0:

$$\frac{\partial L}{\partial a_1} = 2\Sigma a_1 - 2\lambda I a_1 = 0 \rightarrow (\Sigma - \lambda I)a_1 = 0$$

Esto es, en realidad, un sistema lineal de ecuaciones. Por el teorema de Roché-Frobenius, para que el sistema tenga una solución distinta de 0 la matriz $(\Sigma - \lambda I)$ tiene que ser singular. Esto implica que el determinante debe ser igual a cero:

$$|\Sigma - \lambda I| = 0,$$

y de este modo, λ es un autovalor de Σ . La matriz de covarianzas σ es de orden p y si además es definida positiva, tendrá p autovalores distintos, $\lambda_1, \lambda_2, \dots, \lambda_p$ tales que, por ejemplo, $\lambda_1 > \lambda_2 > \dots > \lambda_p$. Se tiene que, desarrollando la expresión anterior,

$$\begin{aligned} (\Sigma - \lambda I)a_1 &= 0 \\ \Sigma a_1 - \lambda I a_1 &= 0 \\ \Sigma a_1 &= \lambda I a_1 \end{aligned}$$

Entonces:

$$\begin{aligned} \text{Var}(y_1) &= \text{Var}(a_1^T x) = a_1^T \Sigma a_1 = a_1^T \lambda I a_1 = \\ &= \lambda a_1^T a_1 = \lambda * 1 = \lambda. \end{aligned} \tag{32}$$

Luego, para maximizar la varianza de y_1 se tiene que tomar el mayor autovalor, digamos λ_1 , y el correspondiente autovector a_1 . En realidad, a_1 es un vector que nos da la combinación de las variables originales que tiene mayor varianza, esto es, si $a_1 = (a_{11}, a_{12}, \dots, a_{1p})$, entonces:

$$y_1 = a_1^T x = a_{11}x_1 + a_{12}x_2 + \dots + a_{1p}x_p.$$

El segundo componente principal, digamos $y_2 = a_2^T x$, se obtiene mediante un argumento parecido. Además, se quiere que y_2 esté incorrelado con el anterior componente y_1 , es decir, $\text{Cov}(y_2, y_1) = 0$. Por lo tanto:

$$\begin{aligned}
Cov(y_2, y_1) &= Cov(a'_2 x, a'_1 x) = \\
&= a'_2 * E[(x - \mu)(x - \mu)'] * a_1 = \\
&= a'_2 \Sigma a_1.
\end{aligned}$$

es decir, se requiere que $a'_2 \Sigma a_1 = 0$.

Como se tenía que $\Sigma a_1 = \lambda a_1$, lo anterior es equivalente a:

$$a'_2 \Sigma a_1 = a'_2 \lambda a_1 = \lambda a'_2 a_1 = 0, \quad (33)$$

Esto equivale a que $a'_2 a_1 = 0$, es decir, que los vectores sean ortogonales. De este modo, tendremos que maximizar la varianza de y_2 , es decir, $a'_2 \Sigma a_2$, sujeta a las siguientes restricciones:

$$\begin{aligned}
a'_2 a_2 &= 1, \\
a'_2 a_1 &= 0,
\end{aligned} \quad (34)$$

Se toma la función:

$$L(a_2) = a'_2 \Sigma a_2 - \lambda(a'_2 a_2 - 1) - \delta a'_2 a_1,$$

y se deriva:

$$\frac{\partial L(a_2)}{\partial a_2} = 2\Sigma a_2 - 2\lambda a_2 - \delta a_1 = 0,$$

Si se multiplica por a'_1 , entonces:

$$2a'_1 \Sigma a_2 - \delta = 0,$$

porque:

$$\begin{aligned}
a'_1 a_2 &= a'_2 a_1 = 0 \\
a'_1 a_1 &= 1
\end{aligned}$$

Luego:

$$\delta = 2a'_1 \Sigma a_2 = 2a'_2 \Sigma a_1 = 0, \quad (35)$$

ya que $Cov(y_2, y_1) = 0$.

De este modo, $\frac{\partial L(a_2)}{\partial a_2}$ queda finalmente como:

$$\begin{aligned}
\frac{\partial L(a_2)}{\partial a_2} &= 2\Sigma a_2 - 2\lambda a_2 - \delta a_1 = 2\Sigma a_2 - 2\lambda a_2 \\
&= (\Sigma - \lambda I)a_2 = 0,
\end{aligned}$$

Usando los mismos razonamientos que antes, elegimos λ como el segundo mayor autovalor de la matriz Σ con su autovector asociado a_2 .

Los razonamientos anteriores se pueden extender, de modo que al j -ésimo componente le correspondería el j -ésimo autovalor. Entonces todos los componentes y (en total p) se pueden expresar como el producto de una matriz formada por los autovectores, multiplicada por el vector x que contiene las variables originales x_1, \dots, x_p

$$y = Ax$$

donde

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix}, A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \dots & a_{pp} \end{pmatrix}, X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix},$$

Como:

$$\text{Var}(y_1) = \lambda_1$$

$$\text{Var}(y_2) = \lambda_2$$

...

$$\text{Var}(y_p) = \lambda_p,$$

La matriz de covarianzas de y será:

$$\begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \lambda_p \end{pmatrix},$$

porque y_1, \dots, y_p se han construido como variables incorreladas.

Se tiene que:

$$\Lambda = \text{Var}(Y) = A' \text{Var}(X) A = A' \Sigma A,$$

o bien

$$\Sigma = A \Lambda A$$

ya que A es una matriz ortogonal (porque $a_i' a_i = 1$ para todas sus columnas) por lo que $AA' = I$

7.3. Demostraciones del capítulo 1 de [2]

- Demostrar $\overline{\Delta u} = \overline{u - \bar{u}} = \bar{u} - \bar{u} = 0$.

Partiendo de la ecuación

$$\Delta u = u - \bar{u}, \quad (36)$$

El promedio de (36) se denota como

$$\overline{\Delta u} = \overline{u - \bar{u}}, \quad (37)$$

Sabiendo que el promedio es:

$$\bar{u} = \sum_{i=1}^M P(u_i)(u_i), \quad (38)$$

Y aplicando el concepto de la ecuación (38) tenemos

$$\overline{\Delta u} = \sum_{i=0}^M P(u_i)(u_i - \bar{u}), \quad (39)$$

$$\begin{aligned} \overline{\Delta u} &= \sum_{i=0}^M P(u_i)u_i - \sum_{i=0}^M P(u_i)\bar{u} \\ &= \sum_{i=0}^M P(u_i)u_i - \bar{u} \sum_{i=0}^M P(u_i). \end{aligned} \quad (40)$$

Sabemos que $\sum_{i=1}^M P(u_i) = 1$, por lo que sustituyendo en (40)

$$\begin{aligned} \overline{\Delta u} &= \bar{u} - (1)\bar{u} \\ \overline{\Delta u} &= \bar{u} - \bar{u} = 0 \quad \square. \end{aligned} \quad (41)$$

- Demostrar $\overline{(\Delta u)^2} = \sum_{i=0}^M P(u_i)(u_i - \bar{u})^2 \geq 0$

Partiendo de la ecuación (36), y elevando ambos lados de la ecuación al cuadrado

$$(\Delta u)^2 = (u - \bar{u})^2, \quad (42)$$

$(\Delta u)^2$ es siempre positivo o mayor que 0, entonces

$$(\Delta u)^2 = (u - \bar{u})^2 \geq 0, \quad (43)$$

Así que, obteniendo el promedio de $(\Delta u)^2$

$$\begin{aligned} \overline{(\Delta u)^2} &= \overline{(u - \bar{u})^2} \geq 0 \\ \overline{(\Delta u)^2} &= \sum_{i=1}^M P(u_i)(u_i - \bar{u})^2 \geq 0 \quad \square. \end{aligned} \quad (44)$$

- Demostrar $\overline{(u - \bar{u})^2} = \overline{u^2} - \bar{u}^2$.

Partiendo de la ecuación (42), se puede desarrollar el binomio del lado derecho de la ecuación

$$(\Delta u)^2 = (u - \bar{u})^2 = u^2 - 2u\bar{u} + (\bar{u})^2, \quad (45)$$

Obteniendo el promedio de la función se tiene

$$\begin{aligned} \overline{(u - \bar{u})^2} &= \overline{u^2 - 2u\bar{u} + (\bar{u})^2} \\ &= \overline{u^2} - 2\bar{u}^2 + \bar{u}^2 \\ \overline{(u - \bar{u})^2} &= \overline{u^2} - \bar{u}^2 \quad \square. \end{aligned} \quad (46)$$

- Demostrar $\overline{u^2} \geq \bar{u}^2$.

Dado que $\overline{(\Delta u)^2} \geq 0$ y $\overline{(\Delta u)^2} = \overline{u^2} - \bar{u}^2$, como se mostró anteriormente. Entonces:

$$\overline{u^2} - \bar{u}^2 \geq 0, \quad (47)$$

Resolviendo la desigualdad se demuestra que:

$$\overline{u^2} \geq \bar{u}^2 \quad \square. \quad (48)$$

7.4. Mínimos cuadrados ejemplo

- 1) Use el método de mínimos cuadrados para determinar la ecuación de la recta que mejor se ajusta para los datos.

| | | | | | | | | | | |
|---|---|----|----|---|---|----|----|---|---|----|
| x | 8 | 2 | 11 | 6 | 5 | 4 | 12 | 9 | 6 | 1 |
| y | 3 | 10 | 3 | 6 | 8 | 12 | 1 | 4 | 9 | 14 |

Se calculan las medias de los valores de x y los valores de y , la suma de los cuadrados de los valores de x , y la suma de cada valor de x multiplicado por su valor correspondiente y .

| x | y | xy | x^2 |
|---------------|---------------|-----------------|------------------|
| 8 | 3 | 24 | 64 |
| 2 | 10 | 20 | 4 |
| 11 | 3 | 33 | 121 |
| 6 | 6 | 36 | 36 |
| 5 | 8 | 40 | 25 |
| 4 | 12 | 48 | 16 |
| 12 | 1 | 12 | 144 |
| 9 | 4 | 36 | 81 |
| 6 | 9 | 54 | 36 |
| 1 | 14 | 14 | 1 |
| $\sum x = 64$ | $\sum y = 70$ | $\sum xy = 317$ | $\sum x^2 = 528$ |

Se calcula la pendiente.

$$\begin{aligned}
 m &= \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}} \\
 &= \frac{317 - \frac{(64)(70)}{10}}{528 - \frac{(64)^2}{10}} \\
 &\approx -1,1
 \end{aligned}$$

Para calcular la intersección de y primero se calcula la la media de x y y :

$$\begin{aligned}
 \bar{x} &= \frac{\sum x}{n} = \frac{64}{10} = 6,4, \\
 \bar{y} &= \frac{\sum y}{n} = \frac{70}{10} = 7,0.
 \end{aligned}$$

Finalmente se calcula la intersección en y para formar la ecuación de la recta que más se ajusta a los datos.

$$\begin{aligned}
 b &= \bar{y} - m\bar{x} \\
 &= 7,0 - ((-1,1)(6,4)) \\
 &= 7,0 + 7,04 \approx 14,0
 \end{aligned}$$

Por lo tanto, la ecuación es $y = -1,1 + 14$

Ejemplo 1 de minimos cuadrados

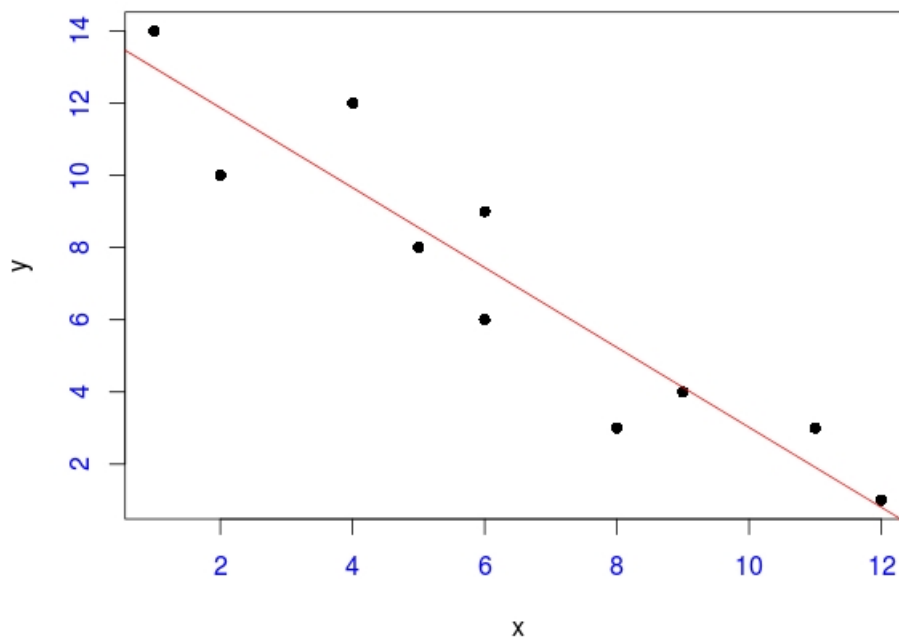


Figura 26: Gráfica del ejemplo 1 de mínimos cuadrados

- 2) De acuerdo a la información mostrada a continuación, determina ¿cuáles serán los costos en una jornada de trabajo de 40 horas?

| Mes | Costo(Y) | Horas(X) |
|---------|----------|----------|
| Enero | 400 | 10 |
| Febrero | 500 | 12.5 |
| Marzo | 500 | 17.5 |
| Abril | 600 | 20 |
| Mayo | 1500 | 50 |
| Junio | 900 | 30 |
| Total | 4400 | 140 |

Solución

| Mes | Costo(Y) | Horas(X) | (X)(Y) | X^2 |
|---------|----------|----------|--------|-------|
| Enero | 400 | 10 | 4000 | 100 |
| Febrero | 500 | 12.5 | 6250 | 156 |
| Marzo | 500 | 17.5 | 8750 | 306 |
| Abril | 600 | 20 | 12000 | 400 |
| Mayo | 1500 | 50 | 75000 | 2500 |
| Junio | 900 | 30 | 27000 | 900 |
| Total | 4400 | 140 | 133000 | 4363 |

$$b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$$

$$a = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2}$$

$$b = \frac{6(133000) - (140)(4400)}{6(4363) - (140)^2} = 27,67$$

$$a = \frac{(4400)(4363) - (140)(133000)}{6(4363) - (140)^2} = 87,75$$

$$Y = a + bx = 87,75 + 27,67x$$



Ejemplo 2 de mínimos cuadrados

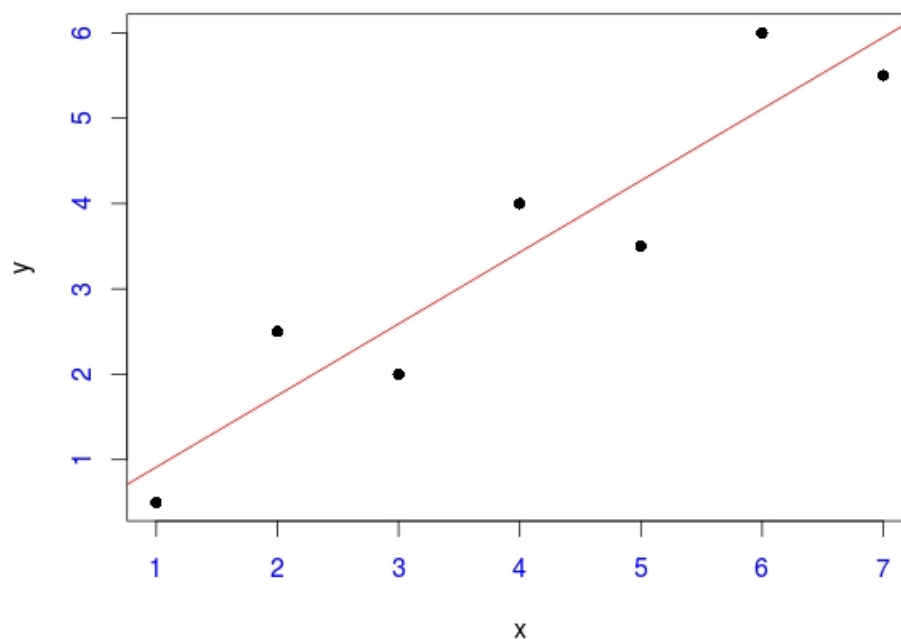


Figura 27: Gráfica del ejemplo 2 de mínimos cuadrados

3) Ajústese una línea recta a los valores x y y de las primeras dos columnas de la siguiente tabla:

| X | Y | (X)(Y) | X^2 |
|-------------------------|-----|--------|-------|
| 1 | 0.5 | .5 | 1 |
| 2 | 2.5 | 5 | 4 |
| 3 | 2 | 6 | 9 |
| 4 | 4 | 16 | 16 |
| 5 | 3.5 | 17.5 | 25 |
| 6 | 6.0 | 36 | 36 |
| 7 | 5.5 | 38.5 | 49 |
| $\Sigma \rightarrow 28$ | 24 | 119.5 | 140 |

$$b = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2},$$

$$a = \frac{(\Sigma y)(\Sigma x^2) - (\Sigma x)(\Sigma xy)}{n(\Sigma x^2) - (\Sigma x)^2},$$

$$b = \frac{7(119,5) - (28)(38,5)}{7(140) - (28)^2} = 0,8392$$

$$a = \frac{(24)(140) - (24)(119,5)}{7(140) - (28)^2} = 0,071$$

$$Y = a + bx = 87,75 + 27,67x$$

Ejemplo 3 de mínimos cuadrados

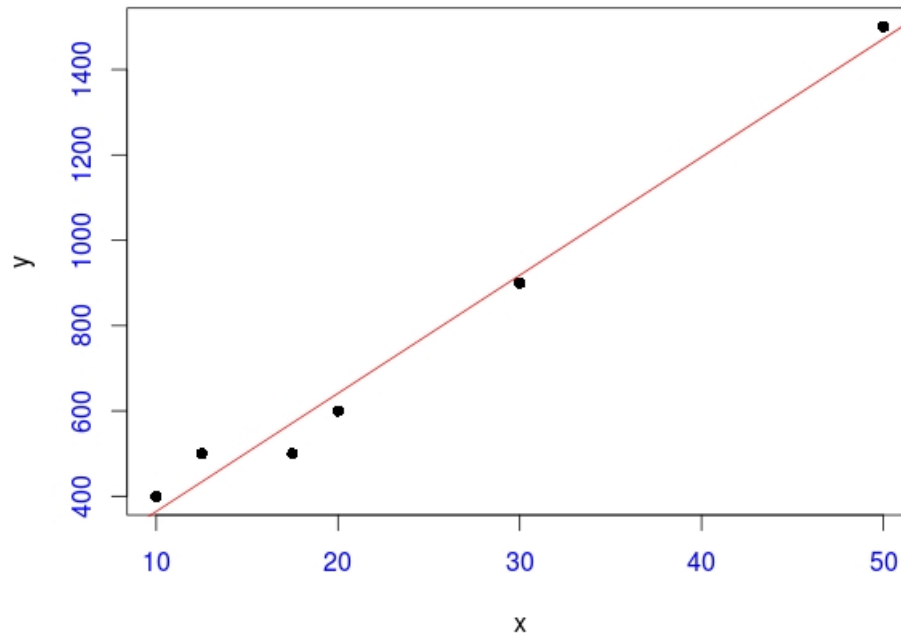


Figura 28: Gráfica del ejemplo 3 de mínimos cuadrados

4)

| | | | | | | |
|---|----|----|-----|-----|-----|-----|
| x | 1 | 2 | 3 | 4 | 5 | 6 |
| y | 10 | 40 | 120 | 300 | 800 | 500 |

Solución

Estimando la ecuación de regresión exponencial:

$$\begin{aligned}
 Y &= Ae^{Bx}, \\
 \ln(Y) &= \ln(Ae^{Bx}), \\
 \ln(Y) &= Y', \\
 \ln(A) &= A', \\
 Y' &= A' + Bx.
 \end{aligned}$$

$$\begin{aligned}
 B &= \frac{n \sum xY' - (\sum x)(\sum Y')}{n \sum x^2 - (\sum x)^2}, \\
 A' &= \bar{Y'} - B\bar{x}.
 \end{aligned}$$

| | x | Y | $Y' = \ln(Y)$ | xY' | x^2 |
|----------|----|-----|---------------|---------|-------|
| | 1 | 10 | 2.303 | 2.303 | 1 |
| | 2 | 40 | 3.689 | 7.378 | 4 |
| | 3 | 120 | 4.787 | 14.362 | 9 |
| | 4 | 300 | 5.704 | 22.815 | 16 |
| | 5 | 800 | 6.685 | 33.423 | 25 |
| | 6 | 500 | 6.215 | 37.288 | 36 |
| Σ | 21 | | 29.382 | 117.569 | 91 |

$$B = \frac{6 \cdot 117,569 - 21 \cdot 29,382}{6 \cdot 91 - 21^2}$$

$$B = 0,84$$

$$A' = \frac{29,382}{6} - 0,84 \cdot \frac{21}{6}$$

$$A' = 1,95 = \ln(A) \rightarrow A = e^{1,95}$$

$$A = 7,03$$

$$Y = 7,03 \cdot e^{0,84x}.$$

Ejemplo 4 de minimos cuadrados

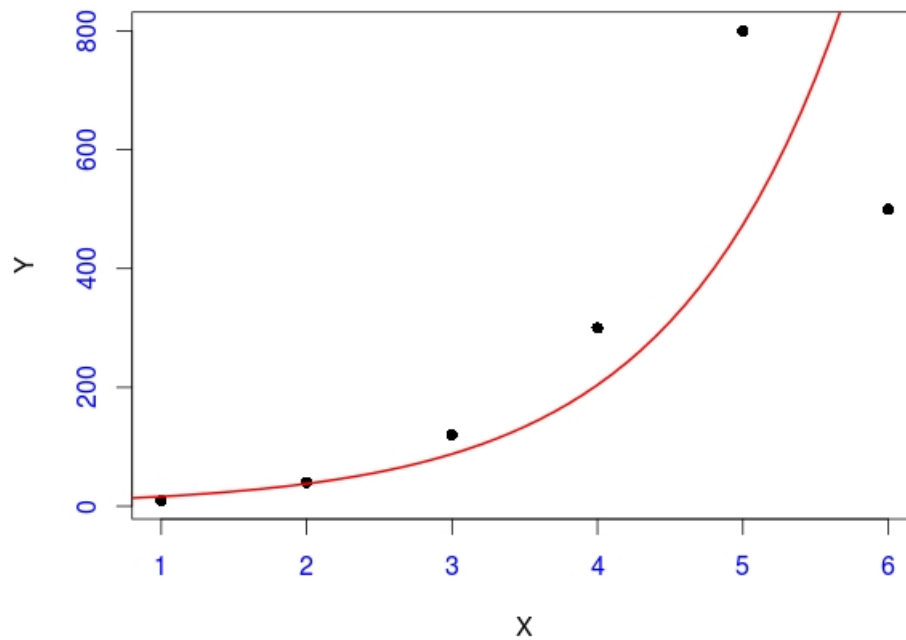


Figura 29: Gráfica del ejemplo 4 de mínimos cuadrados

5) Ejemplo 5.

| | | | | | | |
|---|----|----|-----|-----|-----|-----|
| x | 1 | 2 | 3 | 4 | 5 | 6 |
| y | 10 | 40 | 120 | 300 | 800 | 500 |

Solución

Estimando la ecuación de regresión exponencial:

$$\begin{aligned}
 Y &= Ae^{Bx}, \\
 \ln(Y) &= \ln(Ae^{Bx}), \\
 \ln(Y) &= Y', \\
 \ln(A) &= A', \\
 Y' &= A' + Bx.
 \end{aligned}$$

$$B = \frac{n \sum xY' - (\sum x)(\sum Y')}{n \sum x^2 - (\sum x)^2},$$

$$A' = \bar{Y}' - B\bar{x}.$$

| X | Y | Y' | XY' | X ² |
|----|-------|------|-------|----------------|
| 0 | 126.6 | 4.84 | 0 | 0 |
| 1 | 101.8 | 4.62 | 4.62 | 1 |
| 2 | 71.6 | 4.27 | 8.54 | 4 |
| 4 | 101.6 | 4.62 | 18.48 | 16 |
| 6 | 68.1 | 4.22 | 25.32 | 36 |
| 8 | 62.9 | 4.14 | 33.13 | 64 |
| 9 | 45.5 | 3.81 | 34.35 | 81 |
| 10 | 41.9 | 3.73 | 37.35 | 100 |
| 11 | 46.3 | 3.83 | 42.18 | 121 |
| 12 | 34.1 | 3.52 | 42.35 | 144 |
| 13 | 38.2 | 3.64 | 47.35 | 169 |
| 14 | 41.7 | 3.73 | 52.22 | 196 |
| 15 | 24.7 | 3.20 | 48.1 | 225 |
| 16 | 41.5 | 3.72 | 59.61 | 256 |
| 18 | 36.6 | 3.60 | 64.8 | 324 |
| 19 | 19.6 | 2.97 | 56.53 | 361 |
| 20 | 22.8 | 3.12 | 62.53 | 400 |
| 21 | 29.6 | 3.38 | 71.14 | 441 |
| 22 | 23.5 | 3.15 | 69.45 | 484 |
| 24 | 15.3 | 2.72 | 65.46 | 576 |
| 25 | 13.4 | 2.59 | 64.88 | 625 |
| 26 | 26.8 | 3.98 | 85.49 | 676 |
| 27 | 9.8 | 2.28 | 61.62 | 729 |
| 28 | 18.8 | 2.93 | 82.14 | 784 |
| 29 | 25.9 | 3.25 | 94.37 | 841 |
| 30 | 19.3 | 2.96 | 88.8 | 900 |

$$B = \frac{6 \cdot 117,569 - 21 \cdot 29,382}{6 \cdot 91 - 21^2},$$

$$B = 0,84,$$

$$A' = \frac{29,382}{6} - 0,84 \cdot \frac{21}{6},$$

$$A' = 1,95 = \ln(A) \rightarrow A = e^{1,95},$$

$$A = 7,03.$$

$$Y = 7,03 \cdot e^{0,84x}$$

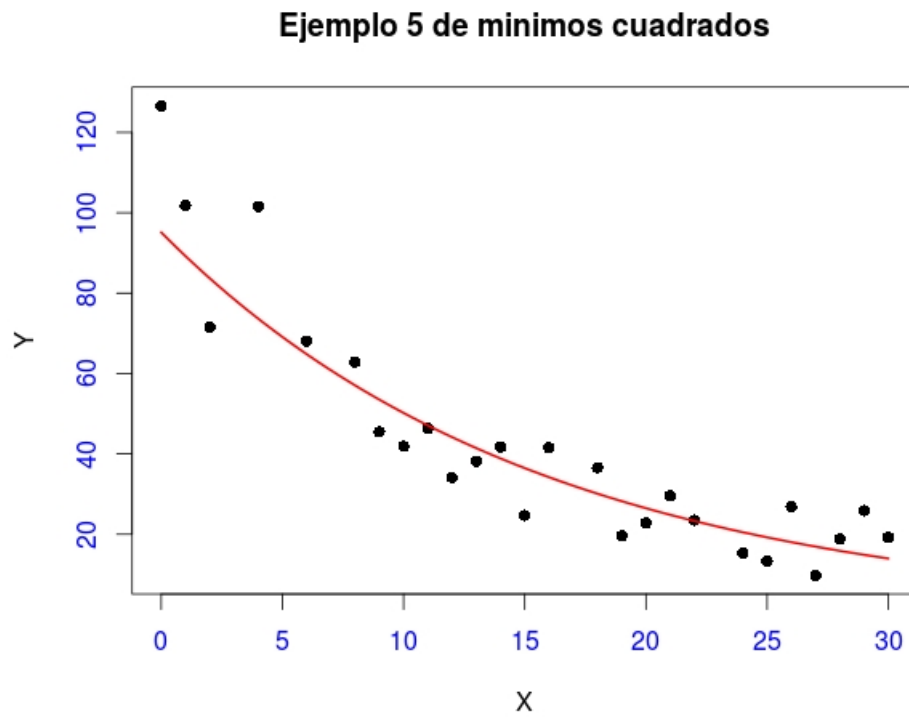


Figura 30: Gráfica del ejemplo 5 de mínimos cuadrados

8. Tarea 8

8.1. Ejemplos de cálculo de eigenvalores y eigenvectores (con multiplicidad)

- (1) Sea $A = \begin{bmatrix} 5 & -3 \\ 3 & -1 \end{bmatrix}$ Calcular sus Eigen valores.

$$\begin{aligned} \det(A - \lambda I) &= \det \left(\begin{bmatrix} 5 & -3 \\ 3 & -1 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right) = \det \left(\begin{bmatrix} 5-\lambda & -3 \\ 3 & -1-\lambda \end{bmatrix} \right) = \\ &= (5-\lambda)(-1-\lambda) + 9 = \lambda^2 - 4\lambda - 5 + 9 = \lambda^2 - 4\lambda + 4 \\ &\lambda_1 = 2, \lambda_2 = 2, \end{aligned}$$

Sustituyendo λ_1 o λ_2 en la ecuacion siguiente:

$$\begin{aligned} (\lambda I - A) \vec{v} &= 0 \\ \begin{bmatrix} 5-2 & -3 \\ 3 & -1-2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \end{aligned}$$

Entonces:

$$\begin{aligned} \begin{bmatrix} 3 & -3 \\ 3 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \\ \vec{v}_1 = \vec{v}_2 &= \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \end{aligned}$$

- (2) Sea $A = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}$ Calcular sus Eigen valores.

$$\begin{aligned} \det(A - \lambda I) &= \det \left(\begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right) = \det \left(\begin{bmatrix} 1-\lambda & 2 \\ 3 & -4-\lambda \end{bmatrix} \right) = \\ &= (1-\lambda)(-4-\lambda) - 6 = \lambda^2 - 4\lambda - 5 + 9 = \lambda^2 + 3\lambda - 10, \\ &\lambda_1 = -5, \lambda_2 = 2. \end{aligned}$$

Sustituyendo λ_1 en la ecuacion siguiente:

$$\begin{aligned} (\lambda I - A) \vec{v} &= 0, \\ \begin{bmatrix} 1+5 & 2 \\ 3 & -4+5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \end{aligned}$$

Entonces:

$$\begin{aligned} \begin{bmatrix} 6 & 2 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \\ \vec{v}_1 &= \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}, \end{aligned}$$

Sustituyendo λ_2 en la ecuacion siguiente:

$$\begin{aligned} (\lambda I - A) \vec{v} &= 0, \\ \begin{bmatrix} 1-2 & 2 \\ 3 & -4-2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \end{aligned}$$

Entonces:

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

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

(3) Sea $A = \begin{bmatrix} 1 & -4 \\ 4 & -7 \end{bmatrix}$ Calcular sus Eigen valores.

$$\begin{aligned} \det(A - \lambda I) &= \det \left(\begin{bmatrix} 1 & -4 \\ 4 & -7 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right) = \det \left(\begin{bmatrix} 1-\lambda & -4 \\ 4 & -7-\lambda \end{bmatrix} \right) = \\ &= (1-\lambda)(-7-\lambda) - 16 = (\lambda+3)^2, \\ \lambda_1 &= -3, \lambda_2 = -3, \end{aligned}$$

Sustituyendo λ_1 o λ_2 en la ecuacion siguiente:

$$\begin{aligned} (\lambda I - A) \vec{v} &= 0, \\ \begin{bmatrix} 1+3 & -4 \\ 4 & -7+3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \end{aligned}$$

Entonces:

$$\begin{aligned} \begin{bmatrix} 4 & -4 \\ 4 & -4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \\ \vec{v}_2 = \vec{v}_1 &= \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \end{aligned}$$

(4)

$$A = \begin{bmatrix} 2 & -2 & 3 \\ 10 & -4 & 5 \\ 5 & -4 & 6 \end{bmatrix}$$

, Entonces,

$$\det(\lambda I - A) = \det \left(\begin{bmatrix} \lambda-2 & 2 & -3 \\ -10 & \lambda+4 & -5 \\ -5 & 4 & \lambda-6 \end{bmatrix} \right).$$

El determinante de la matriz esta dado por el polinomio característico:

$$\det(\lambda I - A) = \lambda^3 - 4\lambda^2 + 5\lambda - 2 = 0,$$

Cuyas raíces son:

$$\lambda_1 = 1; \lambda_2 = 1; \lambda_3 = 2,$$

si sustituimos λ_1 y λ_3 en $(\lambda I - A)$, entonces los vectores propios son:

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 5 \\ 3 \end{bmatrix},$$

$$\vec{v}_2 = \begin{bmatrix} 4 \\ 15 \\ 10 \end{bmatrix}.$$

(5)

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

Entonces,

$$\det(\lambda I - A) = \det \begin{pmatrix} \lambda - 4 & 0 & 0 & 0 \\ 0 & \lambda - 6 & -1 & 0 \\ 0 & 0 & \lambda - 6 & -1 \\ 0 & 0 & 0 & \lambda - 6 \end{pmatrix}.$$

EL determinante de la matriz esta dado por el polinomio característico:

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

Cuyas raíces son:

$$\lambda_1 = 6; \lambda_2 = 6; \lambda_3 = 6; \lambda_4 = 4,$$

Si sustituimos λ_1 , λ_2 , λ_3 y λ_4 en $(\lambda I - A)$, entonces los vectores propios son:

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

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

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

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

8.2. Con los datos del ejercicio anterior ilustrar la propiedad 8

(1)

$$\det \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix} = (1 \times -4) - (2 \times 3) = -10 = \lambda_1 \times \lambda_2 = 2 \times -5.$$

(2)

$$\det \begin{bmatrix} 5 & -3 \\ 3 & -1 \end{bmatrix} = (5 \times -1) - (-3 \times 3) = 4 = \lambda_1 \times \lambda_2 = 2 \times 2.$$

(3)

$$\det \begin{bmatrix} 1 & -4 \\ 4 & -7 \end{bmatrix} = (-7 \times 1) - (-4 \times 4) = 9 = \lambda_1 \times \lambda_2 = -3 \times -3.$$

(4)

$$\det \begin{bmatrix} 2 & -2 & 3 \\ 10 & -4 & 5 \\ 5 & -4 & 6 \end{bmatrix} = 2 = \lambda_1 \times \lambda_2 \times \lambda_3 = 1 \times 1 \times 2 = 2.$$

(5)

$$\det \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 6 & 1 & 0 \\ 0 & 0 & 6 & 1 \\ 0 & 0 & 0 & 6 \end{bmatrix} = 864 = \lambda_1 \times \lambda_2 \times \lambda_3 \times \lambda_4 = 6 \times 6 \times 6 \times 4 = 36 \times 24 = 864.$$

8.3. Multiplicidad en eigenvectores y eigenvalores.

8.3.1. Multiplicidad algebraica

Sean V un espacio vectorial de dimension finita. La multiplicidad algebraica de λ se define como la multiplicidad de la raíz λ del polinomio característico C_T , esto es como:

$$\max\{k \in \{1, \dots, n\} : (x - \lambda)^k \text{ divide a } C_T\}$$

8.4. Matriz singular.

Matriz cuadrada cuyo determinante es igual a cero. Una matriz singular no tiene matriz inversa. Se dice que una matriz A es invertible o no singular si existe una matriz B de tal manera que

$$AB = BA = I$$

donde I es la matriz identidad. De tal forma que B es única. Y esto sucede si $AB_1 = B_1A = I$ y $AB_2 = B_2A = I$, entonces;

$$B_1 = B_1I = B_1(AB_2) = (B_1A)B_2 = IB_2 = B_2$$

Etonces se llama a B la inversa de A y se denota por A^{-1}

8.5. Traza de una matriz.

La traza de una matriz cuadrada A de $n \times n$ está definida como la suma de los elementos de la diagonal principal de A . Es decir:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}$$

$$\text{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn},$$

donde a_{ij} representa el elemento que está en la fila i -ésima y en la columna j -ésima de A .

Debido al especial comportamiento de la traza de una matriz al cambiar de base puede definirse unívocamente la traza de una aplicación lineal, independientemente de cual sea la base elegida. Si un espacio vectorial de dimensión finita está dotado de un producto escalar, y se tiene una base ortonormal entonces la traza de un endomorfismo de dicho espacio viene dada por:

$$\text{tr } f := \sum_k \langle f(e_k), e_k \rangle,$$

Puede comprobarse que si A_f es la matriz de dicha aplicación respecto a dicha base la cantidad anterior es igual a la traza de la matriz A . Y de hecho si B_f es la matriz de la misma aplicación respecto a cualquier otra base ortonormal se tiene:

$$\text{tr } f = \text{tr } A_f = \text{tr } B_f.$$

8.6. Traspuesta de una matriz.

Sea A una matriz con m filas y n columnas. La **matriz traspuesta**, denotada con A^t . Está dada por:

$$(A^t)_{ij} = A_{ji}, \quad 1 \leq i \leq n, \quad 1 \leq j \leq m,$$

En donde el elemento a_{ji} de la matriz original A se convertirá en el elemento a_{ij} de la matriz traspuesta A^t .

Ejemplos:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^t = \begin{bmatrix} a & c \\ b & d \end{bmatrix},$$

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}^t = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}.$$

8.6.1. Propiedades

Para toda matriz A

$$(A^t)^t = A.$$

Sean A y B matrices con elementos pertenecen a un anillo \mathcal{A} y sea $c \in \mathcal{A}$

$$(A + B)^t = A^t + B^t$$

$$(cA)^t = cA^t.$$

Si el producto de las matrices A y B está definido,

$$(AB)^t = B^t A^t.$$

Si A es una matriz cuadrada cuyas entradas son números reales, entonces:

$$A^t A.$$

es semidefinida positiva.

Una matriz cuadrada A es simétrica si coincide con su transpuesta:

$$A^t = A.$$

Una matriz cuadrada A es antisimétrica si su transpuesta coincide con su inverso aditivo.

$$A^t = -A.$$

Si los elementos de la matriz A son números complejos y su transpuesta coincide con su conjugada, se dice que la matriz es hermítica.

$$A^t = \bar{A}, \quad A = (\bar{A})^t = A^\dagger.$$

y antihermítica si

$$A^t = -\bar{A}.$$

8.7. Teorema de Cayley-Hamilton.

El teorema de Cayley-Hamilton asegura que todo endomorfismo de un espacio vectorial de dimensión finita sobre un cuerpo cualquiera anula su propio polinomio característico.

En términos matriciales, eso significa que :

si A es una matriz cuadrada de orden n y si

$$p(X) = \det(XI_n - A) = X^n + p_{n-1}X^{n-1} + \dots + p_1X + p_0,$$

es su polinomio característico (polinomio de indeterminada X), entonces al sustituir formalmente X por la matriz A en el polinomio, el resultado es la matriz nula:

$$p(A) = A^n + p_{n-1}A^{n-1} + \dots + p_1A + p_0I_n = 0_n.$$

El teorema de Cayley-Hamilton se aplica también a matrices cuadradas de coeficientes en un anillo conmutativo cualquiera.

Un corolario importante del teorema de Cayley-Hamilton afirma que el polinomio mínimo de una matriz dada es un divisor de su polinomio característico, y no solo eso, el polinomio mínimo tiene los mismos factores irreducibles que el polinomio característico.

Consideremos por ejemplo la matriz

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix},$$

El polinomio característico se escribe

$$p(x) = \det \begin{pmatrix} x-1 & -2 \\ -3 & x-4 \end{pmatrix} = (x-1)(x-4) - (-2)(-3) = x^2 - 5x - 2$$

El teorema de Cayley-Hamilton afirma que

$$A^2 - 5A - 2I_2 = 0,$$

y esta relación puede verificarse inmediatamente en ese caso. Además el teorema de Cayley-Hamilton permite calcular las potencias de una matriz de modo más sencillo que por un cálculo directo. Tomemos la relación anterior

$$A^2 - 5A - 2I_2 = 0,$$

$$A^2 = 5A + 2I_2,$$

Así, por ejemplo, para calcular A^4 , podemos escribir

$$A^3 = (5A + 2I_2)A = 5A^2 + 2A = 5(5A + 2I_2) + 2A = 27A + 10I_2,$$

y llegamos a

$$A^4 = A^3A = (27A + 10I_2)A = 27A^2 + 10A = 27(5A + 2I_2) + 10A,$$

$$A^4 = 145A + 54I_2,$$

Podemos utilizar también la relación polinomial inicial $A^2 - 5A - 2I_2 = 0$ para probar la inversibilidad de A y calcular su inverso. En efecto, basta con factorizar una potencia de A donde sea posible y $A(A - 5I) = 2I_2$, lo que demuestra que A admite como inverso

$$A^{-1} = \frac{1}{2}(A - 5I).$$

8.7.1. Ejemplo

$$\begin{pmatrix} 3 & 5 \\ -2 & -4 \end{pmatrix}.$$

La ecuación característica de A está determinada por:

$$\det(A - \lambda I) = \det\left(\begin{pmatrix} 3 & 5 \\ -2 & -4 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) = \det\left(\begin{pmatrix} 3-\lambda & 5 \\ -2 & -4-\lambda \end{pmatrix}\right), \quad (49)$$

de esta manera;

$$\det(A - \lambda I) = (3 - \lambda)(-4 - \lambda) - 5(-2) = \lambda^2 + \lambda - 2, \quad (50)$$

Substituyendo λ por A se obtiene:

$$A^2 + A - 2I = \begin{pmatrix} -1 & -5 \\ 2 & 6 \end{pmatrix} + \begin{pmatrix} 3 & 5 \\ -2 & -4 \end{pmatrix} - 2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \quad (51)$$

9. Tarea 9

9.1. Teorema de Cayley-Hamilton

El teorema de Cayley-Hamilton asegura que todo endomorfismo de un espacio vectorial de dimensión finita sobre un cuerpo cualquiera anula su propio polinomio característico.

En términos matriciales, eso significa que :

si A es una matriz cuadrada de orden n y si

$$p(X) = \det(XI_n - A) = X^n + p_{n-1}X^{n-1} + \dots + p_1X + p_0,$$

es su polinomio característico (polinomio de indeterminada X), entonces al sustituir formalmente X por la matriz A en el polinomio, el resultado es la matriz nula:

$$p(A) = A^n + p_{n-1}A^{n-1} + \dots + p_1A + p_0I_n = 0_n.$$

El teorema de Cayley-Hamilton se aplica también a matrices cuadradas de coeficientes en un anillo conmutativo cualquiera.

Un corolario importante del teorema de Cayley-Hamilton afirma que el polinomio mínimo de una matriz dada es un divisor de su polinomio característico, y no solo eso, el polinomio mínimo tiene los mismos factores irreducibles que el polinomio característico.

Consideremos por ejemplo la matriz

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix},$$

El polinomio característico se escribe

$$p(x) = \det \begin{pmatrix} x-1 & -2 \\ -3 & x-4 \end{pmatrix} = (x-1)(x-4) - (-2)(-3) = x^2 - 5x - 2.$$

El teorema de Cayley-Hamilton afirma que

$$A^2 - 5A - 2I_2 = 0,$$

y esta relación puede verificarse inmediatamente en ese caso. Además el teorema de Cayley-Hamilton permite calcular las potencias de una matriz de modo más sencillo que por un cálculo directo. Tomemos la relación anterior

$$A^2 - 5A - 2I_2 = 0,$$

$$A^2 = 5A + 2I_2,$$

Así, por ejemplo, para calcular A^4 , podemos escribir

$$A^3 = (5A + 2I_2)A = 5A^2 + 2A = 5(5A + 2I_2) + 2A = 27A + 10I_2,$$

y llegamos a

$$A^4 = A^3A = (27A + 10I_2)A = 27A^2 + 10A = 27(5A + 2I_2) + 10A,$$

$$A^4 = 145A + 54I_2.$$

Podemos utilizar también la relación polinomial inicial $A^2 - 5A - 2I_2 = 0$ para probar la inversibilidad de A y calcular su inverso. En efecto, basta con factorizar una potencia de A donde sea posible y $A(A - 5I) = 2I_2$, lo que demuestra que A admite como inverso

$$A^{-1} = \frac{1}{2}(A - 5I),$$

10. Tarea 10

10.1. Graficación en R

10.1.1. Ejemplo 1

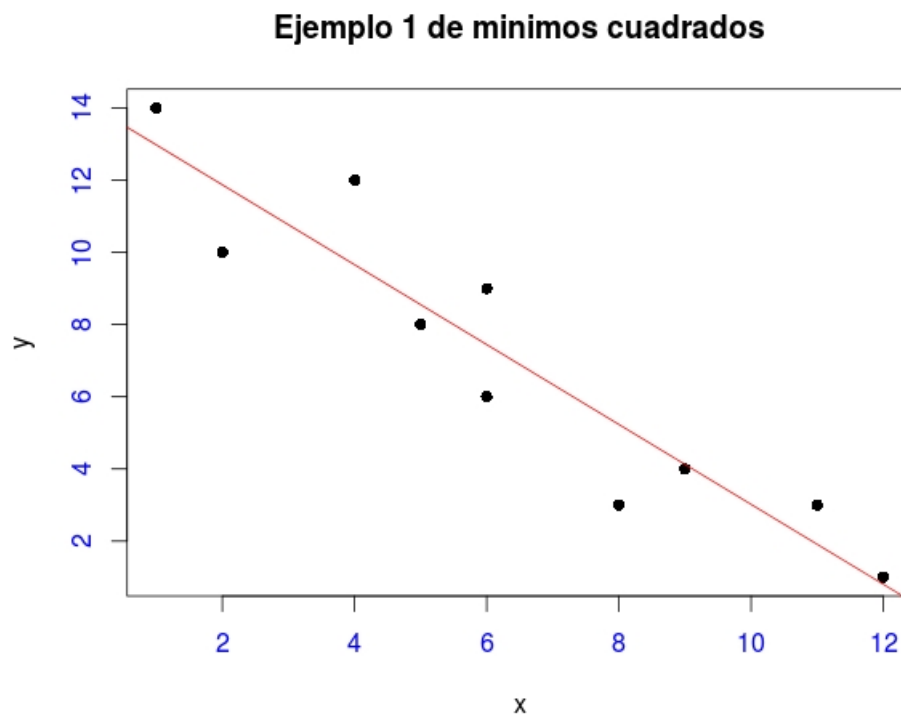


Figura 31: Ejemplo 1 de mínimos cuadrados con R

```
x <- c(8, 2, 11, 6, 5, 4, 12, 9, 6, 1)
y <- c(3, 10, 3, 6, 8, 12, 1, 4, 9, 14)
par(col.lab="black")
par(col.axis="blue")
plot(x,y,pch=16,main="Ejemplo_1_de_minimos_cuadrados")
cor(x,y)
fit <- lm(y ~ x)
lm(formula = y ~ x)
abline(fit,col="red")
summary(fit)
```

10.1.2. Ejemplo 2

```
x <- c(10,12.5,17.5,20,50,30)
y <- c(400,500,500,600,1500,900)
par(col.lab="black")
```

Ejemplo 2 de minimos cuadrados

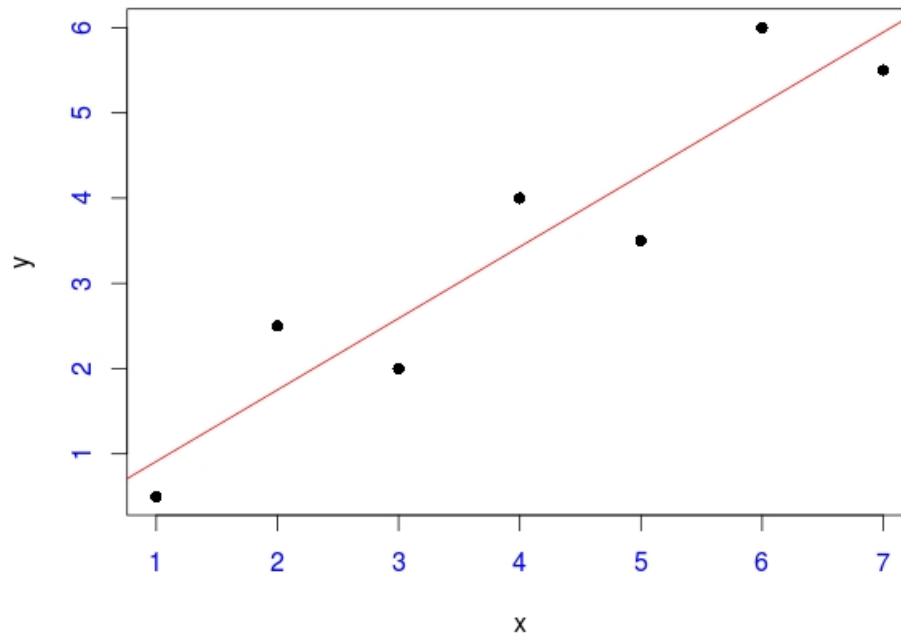


Figura 32: Ejemplo 2 de mínimos cuadrados en R

```
par(col.axis="blue")
par(xlog="TRUE")
plot(x,y,pch=16,main="Ejemplo_2_de_minimos_cuadrados")
cor(x,y)
fit <- lm(y ~ x)
lm(formula = y ~ x)
abline(fit,col="red")
summary(fit)
```

10.1.3. Ejemplo 3

```
x <- c(1, 2,3,4,5,6,7)
y <- c(0.5,2.5,2,4,3.5,6,5.5)
par(col.lab="black")
par(col.axis="blue")

plot(x,y,pch=16,main="Ejemplo_3_de_minimos_cuadrados")
cor(x,y)
fit <- lm(y ~ x)
lm(formula = y ~ x)
abline(fit,col="red")
summary(fit)
```

10.1.4. Ejemplo 4

Ejemplo 3 de mínimos cuadrados

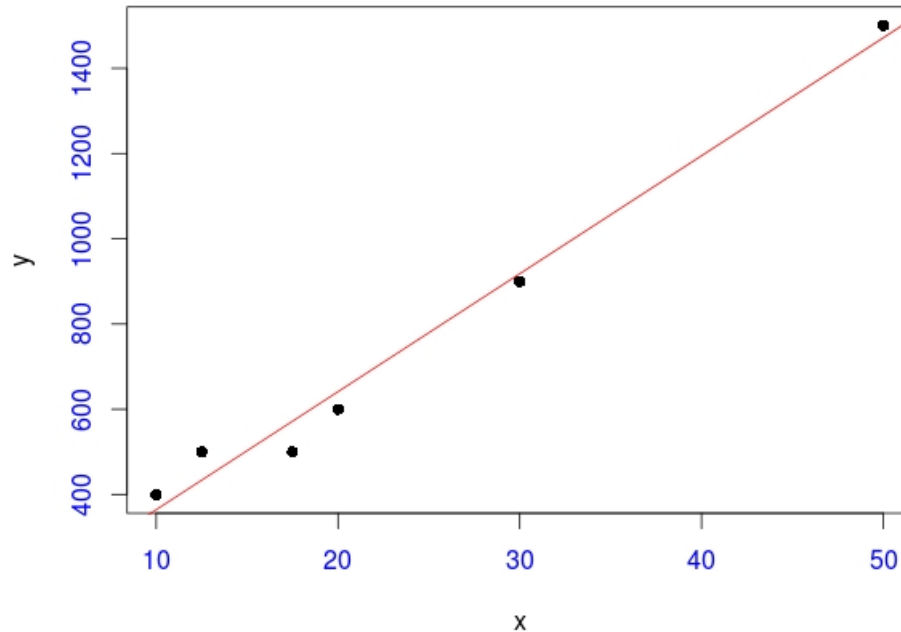


Figura 33: Ejemplo 3 de mínimos cuadrados en R

```
A <- structure(list(X = c(1, 2, 3, 4, 5, 6),
                    Y = c(10, 40, 120, 300, 800, 500)),
               .Names = c("X", "Y"),
               row.names = c(1L, 2L, 3L, 4L, 5L, 6L),
               class = "data.frame")

attach(A)
names(A)
exponential.model <- lm(log(Y) ~ X)
summary(exponential.model)
timevalues <- seq(0, 30, 0.1)
Y.exponential2 <- exp(predict(exponential.model, list(X=timevalues)))
plot(X, Y, pch=16, main = "Ejemplo 4 de mínimos cuadrados")
lines(timevalues, Y.exponential2, lwd=2, col = "red", xlab = "X", ylab = "Y")
```

10.1.5. Ejemplo 5

```
#exponential 2
A <- structure(list(X = c(0, 1, 2, 4, 6, 8, 9, 10, 11, 12, 13,
                        14, 15, 16, 18, 19, 20, 21, 22, 24, 25, 26, 27, 28, 29, 30),
                    Y = c(126.6, 101.8, 71.6, 101.6, 68.1, 62.9, 45.5, 41.9,
                        46.3, 34.1, 38.2, 41.7, 24.7, 41.5, 36.6, 19.6,
                        22.8, 29.6, 23.5, 15.3, 13.4, 26.8, 9.8, 18.8, 25.9, 19.3)),
               .Names = c("X", "Y"), row.names = c(1L, 2L, 3L, 5L, 7L, 9L, 10L, 11L,
                                                    12L, 13L, 14L, 15L, 16L, 17L, 19L,
```

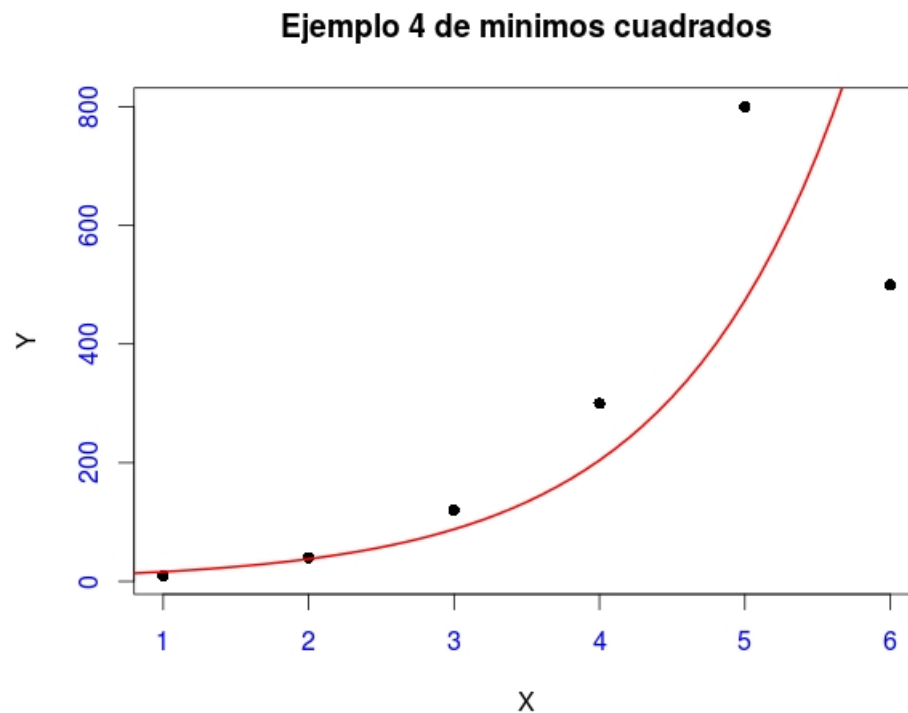


Figura 34: Ejemplo 4 de mínimos cuadrados en R

```

class = "data.frame")
20L, 21L, 22L, 23L, 25L, 26L, 27L,
28L, 29L, 30L,31L),

attach(A)
names(A)
exponential.model <- lm(log(Y)~X)
summary(exponential.model)
timevalues <- seq(0, 30, 0.1)
Y.exponential2 <- exp(predict(exponential.model,list(X=timevalues)))
plot(X, Y,pch=16,main = "Ejemplo_5_de_minimos_cuadrados")
lines(timevalues, Y.exponential2,lwd=2, col = "red", xlab = "X", ylab = "Y")

```

Ejemplo 5 de minimos cuadrados

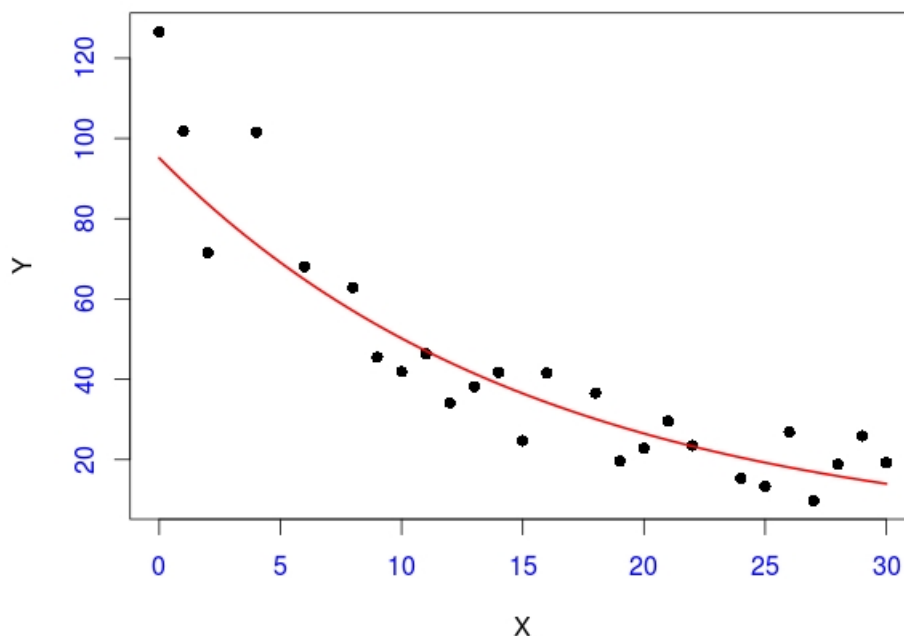


Figura 35: Ejemplo 5 de mínimos cuadrados en R

10.2. Matrices en R

Una matriz en R es un conjunto de objetos indizados por filas y columnas. Un array en R es lo mismo, salvo que puede tener más de dos dimensiones. La forma más simple de crear una matriz en R, es mediante la función `matrix()`, la cual posee cinco argumentos. La sintaxis general de la orden para crear una matriz es la siguiente:

```
matrix(data, nrow, ncol, byrow=F)
```

donde:

| | |
|-------|---|
| data | Datos que forman la matrix |
| nrow | Número de filas que forman la matriz |
| ncol | Número de columnas que forman la matriz |
| byrow | Los datos se forman por fila o columna segun se vayan leyendo |

Ejemplo:

```
> matrix(1:6)
[,1]
[1,] 1
[2,] 2
[3,] 3
[4,] 4
[5,] 5
[6,] 6
```

```
> matrix(1:6,nrow=2)
     [,1] [,2] [,3]
[1,]    1    3    5
[2,]    2    4    6

> matrix(1:6,nrow=2,byrow=T)
     [,1] [,2] [,3]
[1,]    1    2    3
[2,]    4    5    6
```

10.2.1. Funciones sobre matrices

| | |
|-----------|---|
| dim | devuelve las dimensiones de una matriz |
| dimnames | devuelve el nombre de las dimensiones de una matriz |
| colnames | devuelve el nombre de las columnas de una matriz |
| rownames | devuelve el nombre de las filas de una matriz |
| mode | devuelve el tipo de datos de los elementos de una matriz |
| length | devuelve el número total de elementos de una matriz |
| is.matrix | devuelve T si el objeto es una matriz, F si no lo es |
| [,] | accede a elementos dentro de la matriz |
| apply | Aplica una función sobre las filas o columnas de una matriz |
| cbind | A?ade una columna a una matriz dada |
| rbind | A?ade una fila a una matriz dada |

Ejemplos:

```

> x<-matrix(1:6,nrow=3) # Creamos una matriz 3 x 2
> x # Se muestra la matriz x
      [,1] [,2]
[1,]    1    4
[2,]    2    5
[3,]    3    6
> length(x) # N\umero de elementos de x
[1] 6
> mode(x) # Tipo de datos de la matriz x
[1] "numeric"
> dim(x) # Dimensiones de la matriz x
[1] 3 2
> dimnames(x) # Nombre de las dimensiones de la matriz
NULL
> rownames(x) # Nombre de las filas de la matriz
NULL
> colnames(x) # Nombre de las columnas de la matriz
NULL
> is.matrix(x) # El objeto x, ¿es una matriz?
[1] TRUE
> y<-c("blanco","negro") # Creamos un vector de dos palabras
> is.matrix(y) # El objeto y, ¿es una matriz?
[1] FALSE
> x[] # Se muestran todos los elementos de x
      [,1] [,2]
[1,]    1    4
[2,]    2    5
[3,]    3    6
> x[1,2] # Se muestra el elemento 1,2 de x
[1] 4
> x[1,] # Se muestra la primera fila de x
[1] 1 4
> x[,2] # Se muestra la segunda columna de x
[1] 4 5 6

> cbind(x,c(0,0,0)) # Se a?ade una columna de ceros a x
      [,1] [,2] [,3]
[1,]    1    4    0
[2,]    2    5    0
[3,]    3    6    0
> rbind(x,c(0,0)) # Se a?ade una fila de ceros a x
      [,1] [,2]
[1,]    1    4
[2,]    2    5
[3,]    3    6
[4,]    0    0

> datos<-matrix(c(20,65,174,22,70,180,19,68,170),
nrow=3,byrow=T)
> datos
      [,1] [,2] [,3]
[1,]   20   65  174
[2,]   22   70  180
[3,]   19   68  170

> colnames(datos)<-c("edad","peso","altura")
> datos
      edad peso altura
[1,]   20   65   174
[2,]   22   70   180

```

```
[3,] 19 68 170  
> rownames(datos)<-c("paco","pepe","kiko")  
> datos  
      edad peso altura  
paco    20   65   174  
pepe    22   70   180  
kiko    19   68   170
```

10.2.2. Operaciones con matrices

```
A %% B : producto de matrices  
t(A) : transpuesta de la matriz  
A  
solve(A,b) : solución del sistema de ecuaciones Ax=b.  
solve(A) : inversa de la matriz  
A  
svd(A) : descomposición en valores singulares  
qr(A) : descomposición QR  
eigen(A) : valores y vectores propios  
diag(b) : matriz diagonal (  
b es un vector)  
diag(A) : matriz diagonal (  
A es una matriz)  
A %% B == outer(A,B) : producto exterior de dos vectores o  
matrices
```

10.3. Ecuación cuadrática root (clases)

11. Tarea 11

11.1. Correlación de Pearson

En estadística, el coeficiente de correlación de Pearson es una medida de la relación lineal entre dos variables aleatorias cuantitativas. A diferencia de la covarianza, la correlación de Pearson es independiente de la escala de medida de las variables.

De manera menos formal, podemos definir el coeficiente de correlación de Pearson como un índice que puede utilizarse para medir el grado de relación de dos variables siempre y cuando ambas sean cuantitativas.

En el caso de que se esté estudiando dos variables aleatorias x e y sobre una población; el coeficiente de correlación de Pearson se simboliza con la letra $\rho_{x,y}$, siendo la expresión que nos permite calcularlo:

$$\rho_{X,Y} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y}, \quad (52)$$

Donde:

- σ_{XY} es la covarianza de (X, Y)
- σ_X es la desviación típica de la variable X
- σ_Y es la desviación típica de la variable Y

De manera análoga podemos calcular este coeficiente sobre un estadístico muestral, denotado como r_{xy} a:

$$r_{xy} = \frac{\sum x_i y_i - n \bar{x} \bar{y}}{(n-1) s_x s_y} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{(n) \sum y_i^2 - (\sum y_i)^2}}. \quad (53)$$

El valor del índice de correlación varía en el intervalo $[-1, 1]$:

Si $r = 1$, existe una correlación positiva perfecta. El índice indica una dependencia total entre las dos variables denominada relación directa: cuando una de ellas aumenta, la otra también lo hace en proporción constante. Si $0 < r < 1$, existe una correlación positiva. Si $r = 0$, no existe relación lineal. Pero esto no necesariamente implica que las variables son independientes: pueden existir todavía relaciones no lineales entre las dos variables. Si $-1 < r < 0$, existe una correlación negativa. Si $r = -1$, existe una correlación negativa perfecta. El índice indica una dependencia total entre las dos variables llamada relación inversa: cuando una de ellas aumenta, la otra disminuye en proporción constante.

Se realizó un programa en R para ilustrar el comportamiento del coeficiente de Pearson.

```
A=1:20
B<- runif(20, 1,100)
plot(A,B)
r=cor(A,B)
x = seq(1,20)
y = x^2+10*x
plot(x,y)
s=cor(x,y)
x = seq(1,20)
y = x
plot(x,y)
t=cor(x,y)
)
```

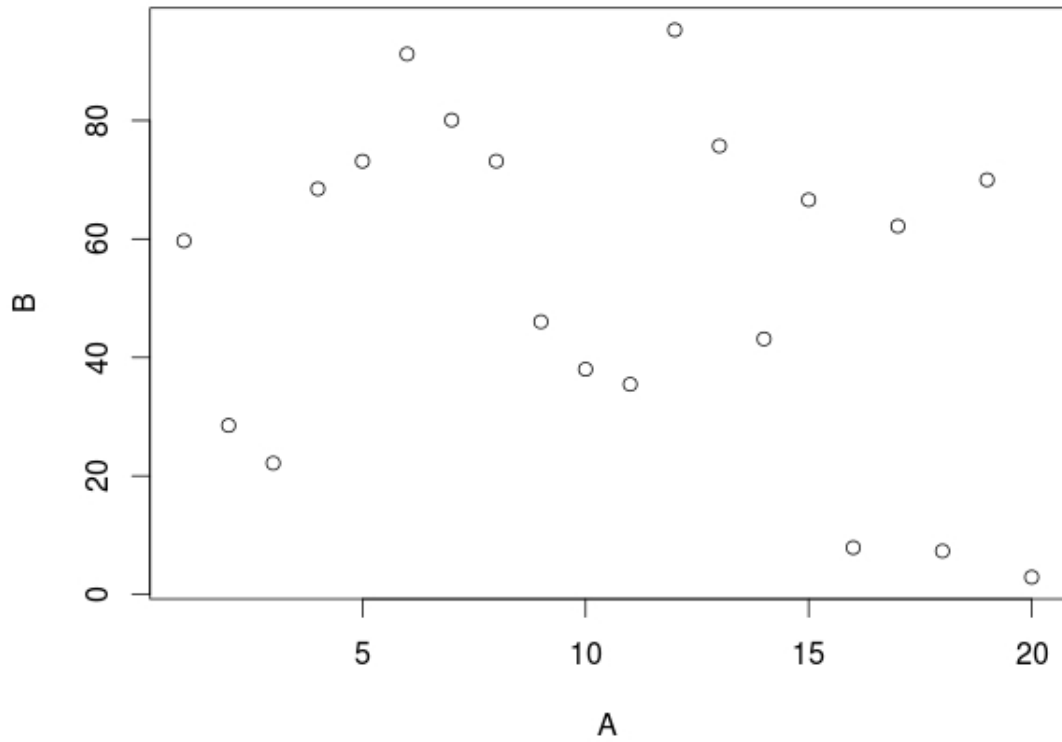


Figura 36: Gráfica generada con números aleatorios

El coeficiente de relación de Pearson que se encontró para la gráfica 36 es:

```
> r  
[1] 0.04246149
```

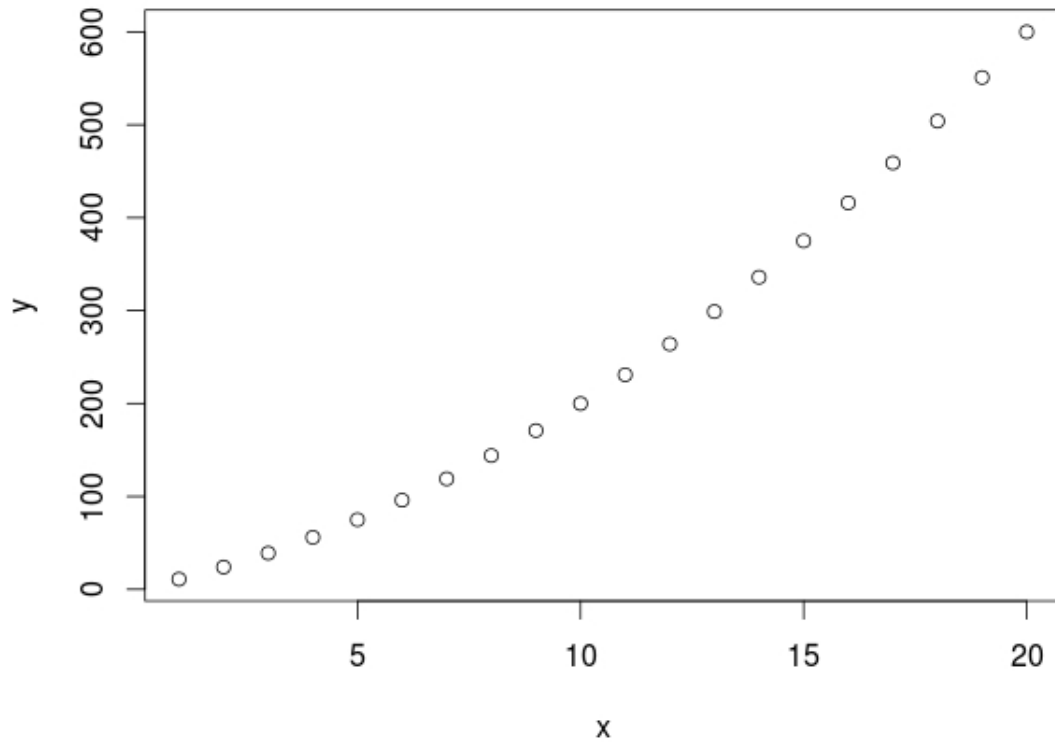



Figura 37: Gráfica generada a partir de una ecuación cuadrática

El coeficiente de relación de Pearson que se encontró para la gráfica 37 es:

```
> s  
[1] 0.986541
```

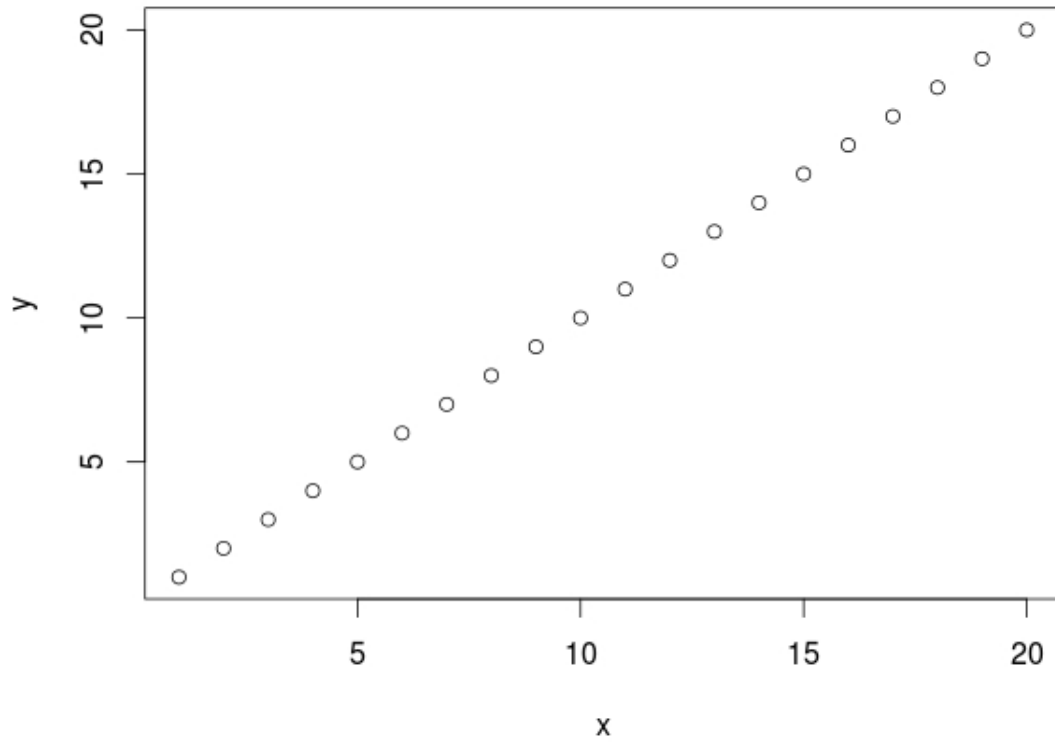


Figura 38: Gráfica de una recta

El coeficiente de relación de Pearson que se encontró para la grafica 38 es:

```
> t
[1] 1
```

12. Tarea 12

12.1. Aproximación a $\ln(x)$ por series de Taylor

La serie de Taylor es una serie funcional y surge de una ecuación en la cual se puede encontrar una solución aproximada a una función. La serie de Taylor proporciona una buena forma de aproximar el valor de una función en un punto en términos del valor de la función y sus derivadas en otro punto.

Por supuesto, para hacer esta aproximación sólo se pueden tomar unas cuantas expresiones de esta serie, por lo que el resto resulta en un error conocido como el término residual, es a criterio del que aplica la serie en número de términos que ha de incluir la aproximación.

Pueden resolver por aproximación funciones trigonométricas, exponenciales, logarítmicas etc...

La serie de Taylor se basa en ir haciendo operaciones según una ecuación general y mientras mas operaciones tenga la serie mas exacto será el resultado que se esta buscando. Dicha ecuación es la siguiente:

$$f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \dots,$$

o expresado de otra forma:

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n, \quad (54)$$

donde:

* $n!$ es el factorial de n

* $f^{(n)}(a)$ denota la n -ésima derivada de f para el valor a de la variable respecto de la cual se deriva.

Para el cálculo de la serie de Taylor de la función $\ln(x)$ centrada en $a = 1$ se procede por calcular las derivadas y a evaluar las funciones en $a = 1$. A continuación se observan

$$\begin{aligned} f(x) &= \ln(x) \\ \frac{df(x)}{dx} &= \frac{1}{x} = x^{-1} \\ \frac{d^2 f(x)}{dx^2} &= -x^{-2} \\ \frac{d^3 f(x)}{dx^3} &= 2x^{-3} \\ \frac{d^4 f(x)}{dx^4} &= -6x^{-4} \\ \frac{d^5 f(x)}{dx^5} &= 24x^{-5} \\ &\vdots \end{aligned}$$

Si evaluamos cada una de las derivadas en 1 entonces obtenemos:

$$\begin{aligned} f(1) &= \ln(1) = 0 \\ f'(1) &= 1 \\ f''(1) &= -1^{-2} = -1 \\ f'''(1) &= 2(1)^{-3} = 2 \\ f^{(4)}(1) &= -6(1)^{-4} = -6 \\ f^{(5)}(1) &= 24(1)^{-5} = 24 \\ &\vdots \end{aligned}$$

Sustituyendo en la formula general se obtiene:

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \frac{f^{(4)}(a)}{4!}(x-a)^4 + \dots$$

$$f(x) = f(1) + \frac{f'(1)}{1!}(x-1) + \frac{f''(1)}{2!}(x-1)^2 + \frac{f^{(3)}(1)}{3!}(x-1)^3 + \frac{f^{(4)}(1)}{4!}(x-1)^4 + \dots$$

$$f(x) = 0 + \frac{1}{1}(x-1) + \frac{-1}{2}(x-1)^2 + \frac{2}{6}(x-1)^3 + \frac{-6}{24}(x-1)^4 + \frac{24}{120}(x-1)^5 + \dots$$

$$f(x) = (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \frac{(x-1)^4}{4} + \frac{(x-1)^5}{5} - \dots$$

En forma de sumatoria:

$$\sum_{n=1}^{\infty} \frac{(-1)^n (x-1)^n}{n}. \quad (55)$$

Se realizó un programa en R para verificar el resultado de la serie.

```
x<-seq(0,3,.01)
y<-c(log(x))
y1<-c(x-1)
y2<-c((x-1) -((1/2)*(x-1)^2))
y3<-c((x-1) -((1/2)*(x-1)^2) +((1/3)*(x-1)^3) )
y4<-c((x-1) -((1/2)*(x-1)^2) +((1/3)*(x-1)^3) -((1/4)*(x-1)^4))
y5<-c((x-1) -((1/2)*(x-1)^2) +((1/3)*(x-1)^3) -((1/4)*(x-1)^4) +((1/5)*(x-1)^5))
y6<-c((x-1) -((1/2)*(x-1)^2) +((1/3)*(x-1)^3) -((1/4)*(x-1)^4) +((1/5)*(x-1)^5)-((1/6)*(x-1)^6))
plot(x,y,'l')
lines(x,y1,col='cyan')
lines(x,y2,col='cyan')
lines(x,y3,col='tomato')
lines(x,y4,col='pink3')
lines(x,y5,col='blue2')
lines(x,y6,col='brown3')
```

La gráfica obtenida mediante la serie de Tylor es la siguiente:

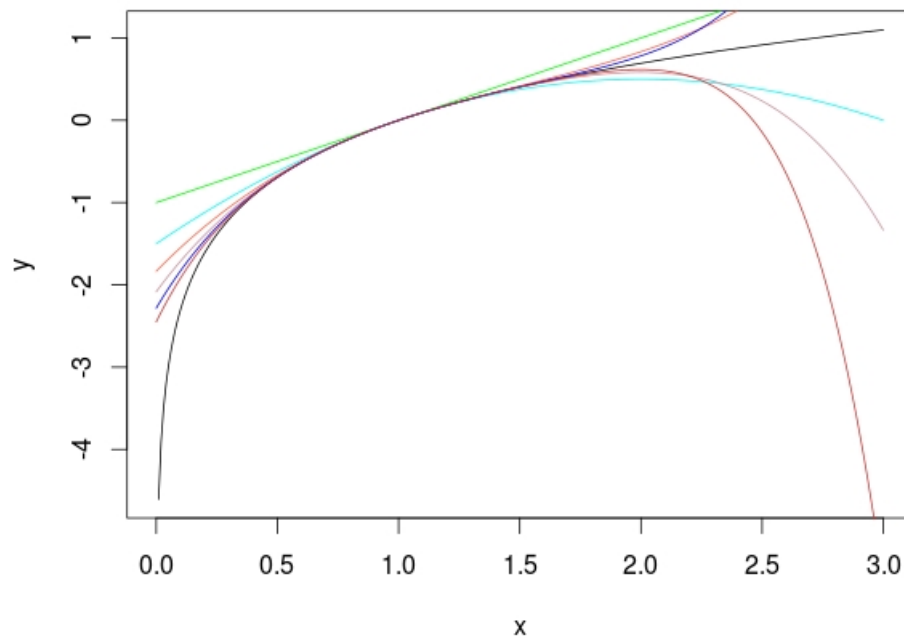


Figura 39: Aproximación de la función $\ln(x)$ por serie de Taylor.

13. Tarea 13

13.1. Ejemplos de mínimos cuadrados en Mathematica

Se genera un conjunto de puntos aleatorios alrededor de la función x^2 , y posteriormente se grafica

```
datos = Table[{x, x^2 + Random[Real, {-2, 2}]}, {x, 1, 5, .1}];
g1 = ListPlot[datos]
```

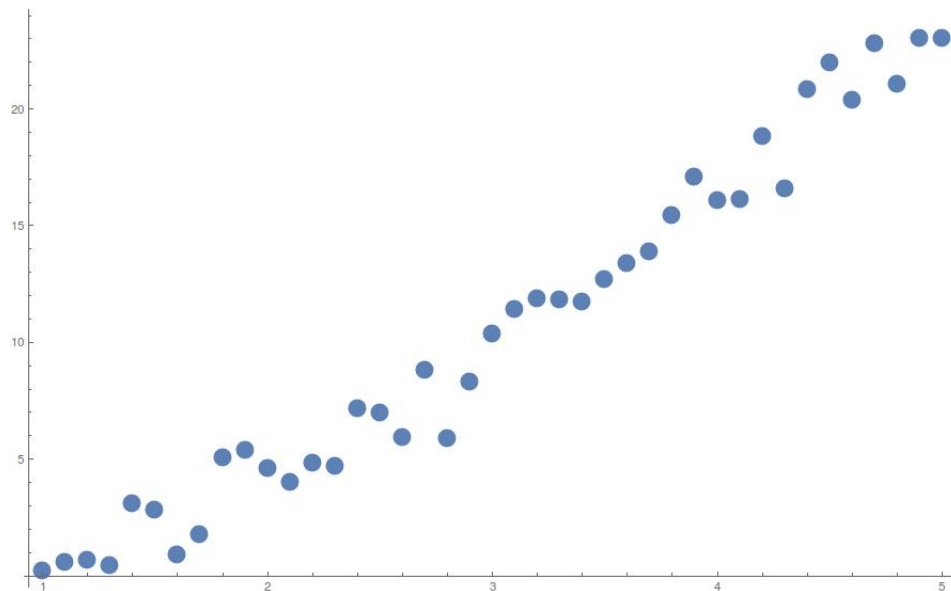


Figura 40: Puntos aleatorios alrededor de la función x .

Posteriormente se obtiene una aproximación mediante mínimos cuadrados y la función Fit de matemática.

El código que se utilizó para hacer las aproximaciones y las gráficas es el siguiente:

```
sol1 = Fit[datos, {1, x}, {x}]
g2 = Plot[sol1, {x, -2, 7}]
sol2 = Fit[datos, {1, x, x^2}, {x}]
g3 = Plot[sol2, {x, -2, 7}]
sol3 = Fit[datos, {1, x, x^2, x^3}, {x}]
g4 = Plot[sol3, {x, -2, 7}]
sol4 = Fit[datos, {1, x, x^2, x^3, x^4}, {x}]
g5 = Plot[sol4, {x, -2, 7}]
Show[g1, g2, g3, g4]
```

A continuación se muestran las gráficas pertenecientes a cada una de las aproximaciones.

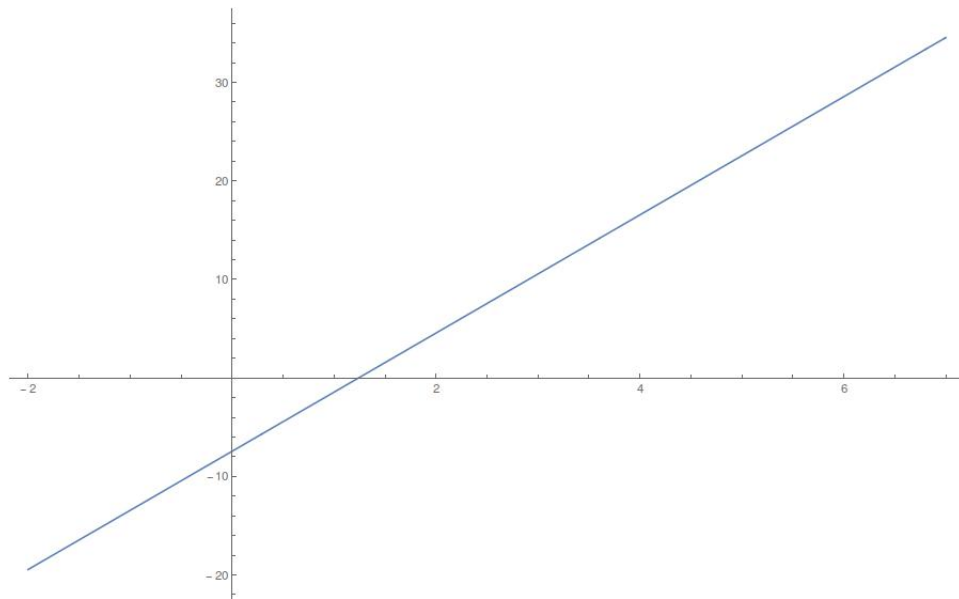


Figura 41: Ajuste con una función $f(x) = x$.

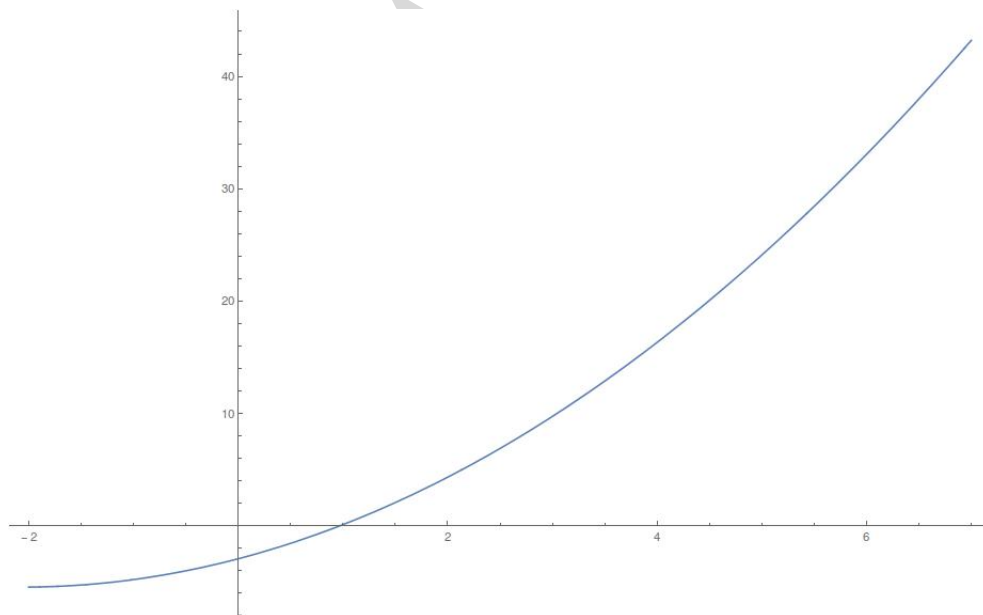


Figura 42: Ajuste con una función $f(x) = x^2$.

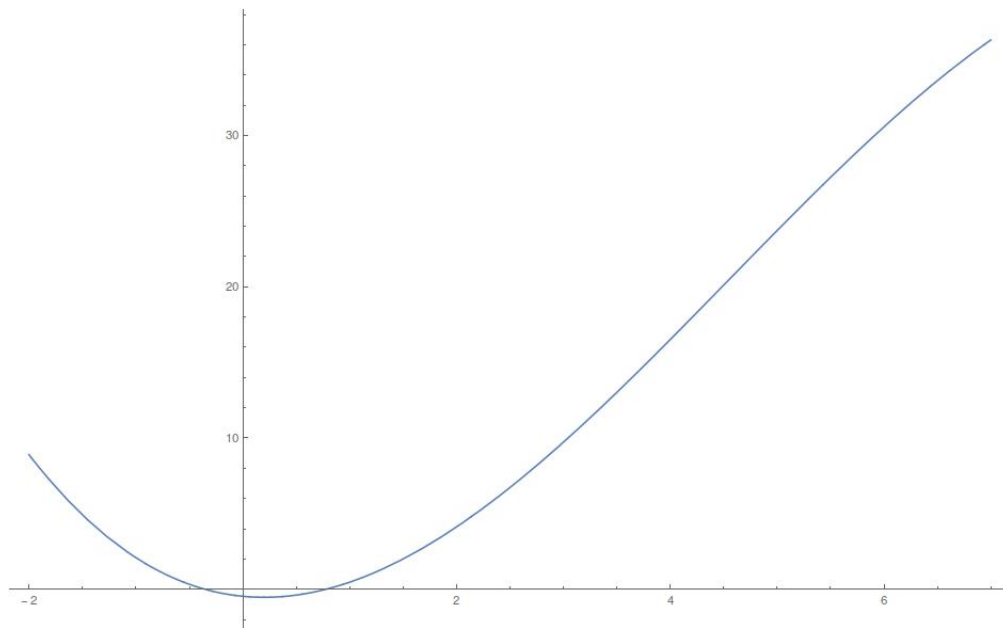


Figura 43: Ajuste con una función $f(x) = x^3$.

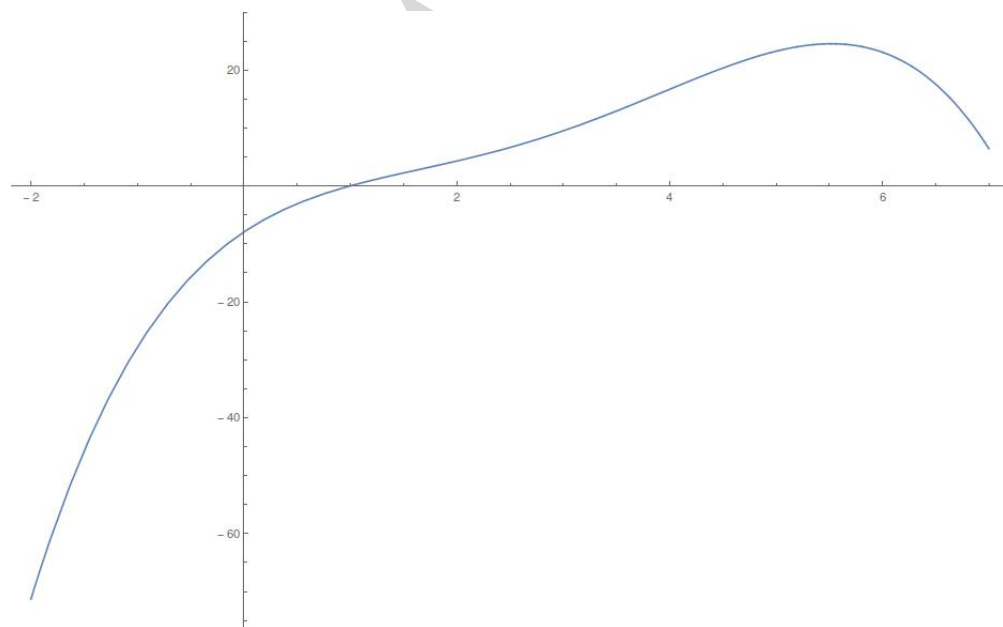


Figura 44: Ajuste con una función $f(x) = x^4$.

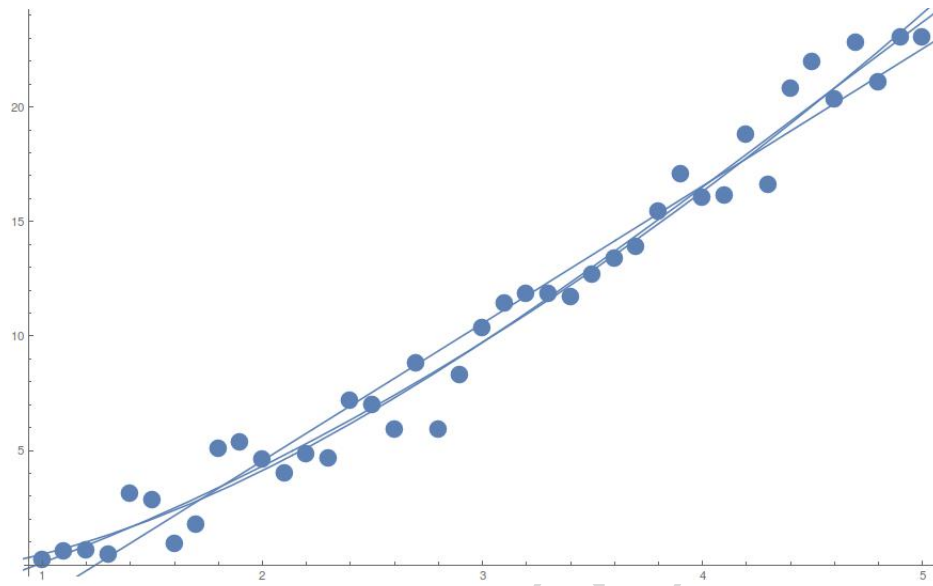


Figura 45: Gráfica de puntos y sus aproximaciones.

13.2. Cadenas de Markov

Una cadena de markov consta de unos estados $E_1 E_2 E_3 E_4 \dots$. En. que inicialmente en un tiempo 0 o paso 0 se le llama estado inicial, además de esto consta de una matriz de transición que significa la posibilidad de que se cambie de estado en un próximo tiempo o paso.

13.2.1. Matriz de transición:

Una matriz de transición para una cadena de Markov de n estado es una matriz de $n \times n$ con todos los registros no negativos y con la propiedad adicional de que la suma de los registros de cada columna (o fila) es 1. Por ejemplo: las siguientes son matrices de transición

$$\begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0,8 & 0,2 & 0. \\ 0,18 & 0,75 & 0,16 \\ 0,02 & 0,5 & 0,84 \end{pmatrix} \begin{pmatrix} 0,7 & 0,1 & 0. & 0,01 \\ 0,3 & 0,6 & 0,45 & 0,05 \\ 0. & 0,15 & 0,15 & 0,55 \\ 0. & 0,15 & 0,4 & 0,39 \end{pmatrix}. \quad (56)$$

13.2.2. Ejemplos de Cadenas de Markov en Mathematica

Supongamos que el clima de una determinada región sólo puede ser soleado(s1) o nublado (s2) y que las condiciones del clima en mañanas sucesivas forman una cadena de Markov con probabilidades de transición estacionarias. La matriz de transición está dada por:

$$P = \begin{pmatrix} 0,7 & 0,3 \\ 0,6 & 0,4 \end{pmatrix}. \quad (57)$$

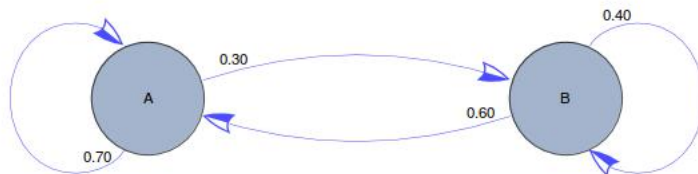


Figura 46: Grafo que representa la matriz de transiciones 57

Si un día concreto está nublado, ¿cuál es la probabilidad de que esté nublado el día siguiente? $P(2, 2) = ,4$
 Suponemos que la probabilidad de que el miércoles haga sol es 0.2 y la probabilidad de que esté nublado es 0.8. Calcular: 1. Probabilidad de que esté nublado el jueves. 2. Probabilidad de que esté nublado el viernes. 3. Probabilidad de que esté nublado el sábado.

- (1) 1.-Miércoles: $v = (0,2, 0,8) \rightarrow w = vP$
 $w = v.P$

$$(0,2 \quad 0,8) * \begin{pmatrix} 0,7 & 0,3 \\ 0,6 & 0,4 \end{pmatrix} = (0,62 \quad 0,38).$$

- (2) $w = vP = (0,62, 0,38) \rightarrow vP^2 = vPP = wP$.

$$w = (0,62, 0,38) * \begin{pmatrix} 0,7 & 0,3 \\ 0,6 & 0,4 \end{pmatrix} * \begin{pmatrix} 0,7 & 0,3 \\ 0,6 & 0,4 \end{pmatrix} = (0,662, 0,338).$$

- (3) $vP^3 = vP^2P$.

$$(0,662, 0,338) * \begin{pmatrix} 0,7 & 0,3 \\ 0,6 & 0,4 \end{pmatrix} * \begin{pmatrix} 0,7 & 0,3 \\ 0,6 & 0,4 \end{pmatrix} * \begin{pmatrix} 0,7 & 0,3 \\ 0,6 & 0,4 \end{pmatrix} = (0,66620, 33338).$$

13.2.3. Otros Ejemplos de Matrices de transición y sus respectivo grafos

(1) Para la matriz de transición siguiente,

$$A = \begin{pmatrix} 0,1 & 0,4 & 0,2 & 0,1 & 0,1 & 0,1 \\ 0,2 & 0,3 & 0,2 & 0,1 & 0,1 & 0,1 \\ 0,1 & 0,2 & 0,3 & 0,1 & 0,1 & 0,1 \\ 0,1 & 0,1 & 0,2 & 0,2 & 0,2 & 0,1 \\ 0,1 & 0,1 & 0,1 & 0,3 & 0,3 & 0,2 \\ 0,1 & 0,1 & 0,1 & 0,4 & 0,4 & 0,2 \end{pmatrix}, \quad (58)$$

el grafo que la representa es:

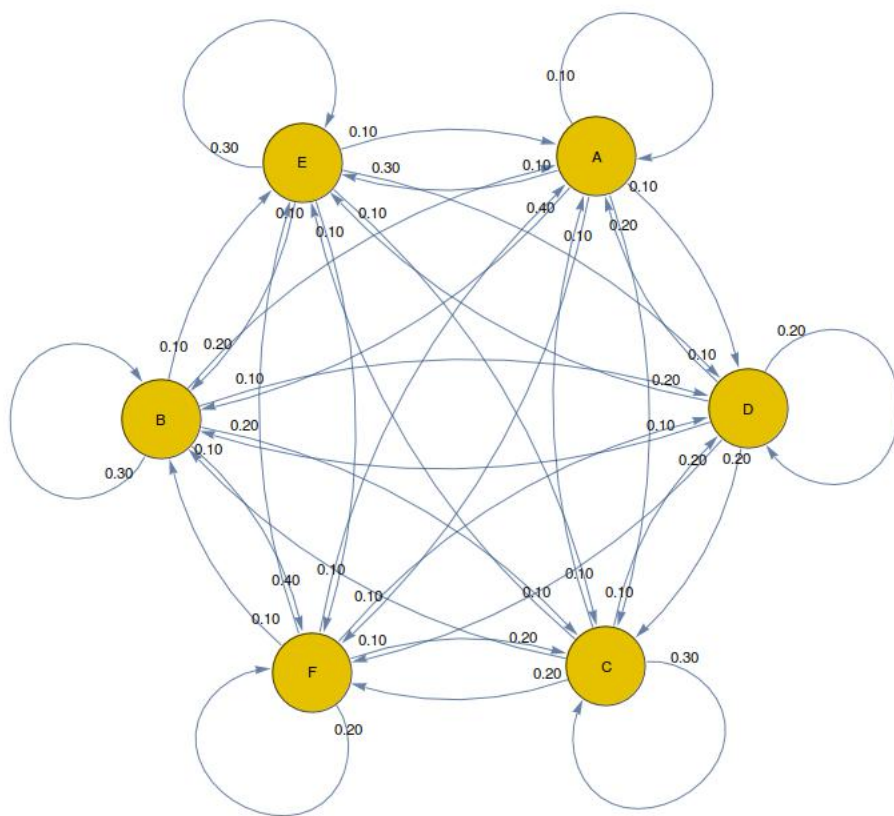


Figura 47: Grafo que representa a la matriz de transiciones 58

(2) Para la matriz de transición siguiente,

$$B = \begin{pmatrix} 0 & 0,25 & 0,5 & 0 & 0,25 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0,5 & 0,5 \\ 0,25 & 0,25 & 0 & 0 & 0 & 0,5 \\ 0,25 & 0 & 0,25 & 0 & 0,5 & 0 \end{pmatrix} \quad (59)$$

el grafo que la representa es:

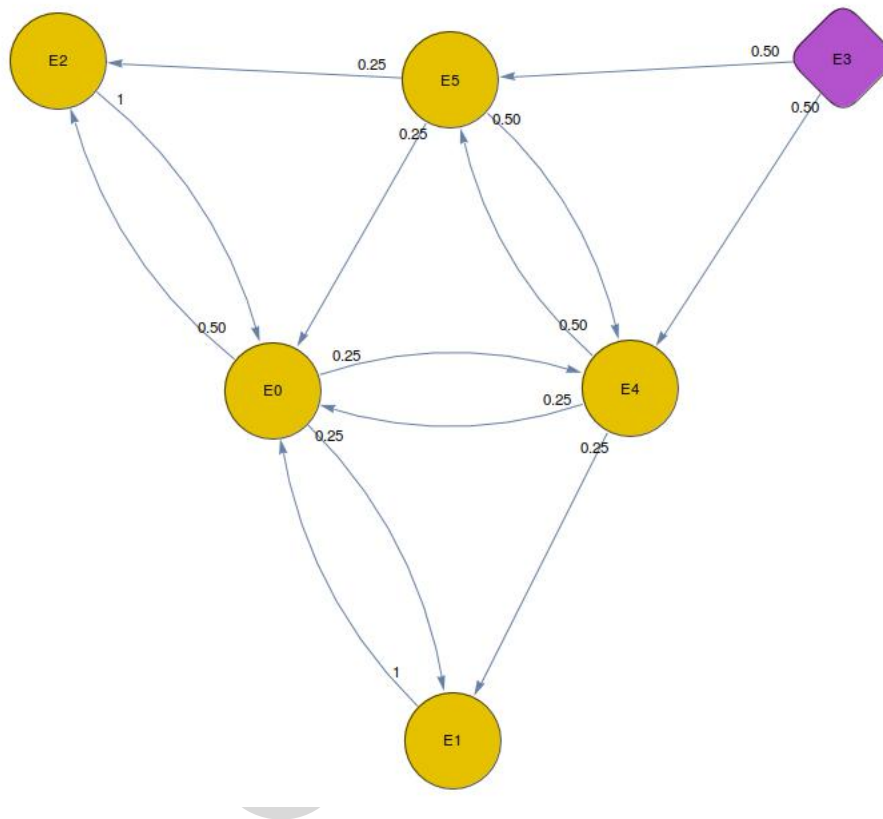


Figura 48: Grafo que representa la matriz de transiciones 59

13.3. Histograma

Los histogramas se usan para resaltar la diferencia entre las clases que se han agrupado los datos. Por tanto, para su construcción se necesita primero agrupar los datos en una tabla la cual se conoce como una tabla de frecuencia. En el caso de datos numéricos continuos los datos se agrupan en intervalos o bins, como se dice en inglés. La frecuencia de un intervalo es el número de datos que se encuentran en él.

Los intervalos deben poseer las siguientes características:

- 1 Todos deben ser del mismo ancho (class width).
- 2 No deben solapar.

3 Todos los datos deben caer en uno de las intervalos.

4 Deben haber un total de entre 5 y 15 intervalos.

Para preparar la tabla de frecuencia para un histograma, es necesario primero establecer el número de intervalos que se desea tener. Luego, se debe determinar el ancho común de los intervalos (interval width). Para esto, se calcula la diferencia del dato mayor y el dato menor, y se divide entre el número de intervalos deseados. Este resultado se redondea al entero mayor más cercano. Por ejemplo, si se desea tener 10 intervalos y encontramos que el dato mayor es 35 y el menor es 12, El ancho común se determina:

$$ancho_{bin} = \frac{35 - 12}{10} = 2,3 \approx 3$$

13.3.1. Ejemplo

En Mathematica se usa el comando "BinCounts" para determinar el ancho de las barras.

```
a = RandomInteger[{1, 6}, 100];
b = Sort[Tally[a]] [[All, 2]] //
  MatrixForm (* muestra la frecuencia de los datos *)
BinCounts[a, {1, 6, 1}] (* para crear el histograma se requiere crear los bins por medio de clases *)
Histogram[a]
a = RandomVariate[NormalDistribution[0, 1], 1000];
BinCounts[a, {-3, 3, .5}]
Histogram[a]
```

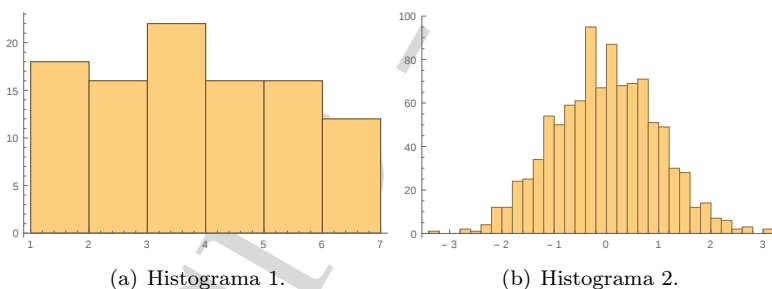


Figura 49: Histogramas generados por MATHEMATICA.

13.4. Box and whisker plots

Los diagramas de Caja-Bigotes (boxplots o box and whiskers) son una presentación visual que describe varias características importantes, al mismo tiempo, tales como la dispersión y simetría.

Para su realización se representan los tres cuartiles y los valores mínimo y máximo de los datos, sobre un rectángulo, alineado horizontal o verticalmente.

Una gráfica de este tipo consiste en una caja rectangular, donde los lados más largos muestran el recorrido intercuartílico. Este rectángulo está dividido por un segmento vertical que indica donde se posiciona la mediana y por lo tanto su relación con los cuartiles primero y tercero (recordemos que el segundo cuartil coincide con la mediana). Esta caja se ubica a escala sobre un segmento que tiene como extremos los valores mínimo y máximo de la variable. Las líneas que sobresalen de la caja se llaman bigotes. Estos bigotes tienen un límite de prolongación, de modo que cualquier dato o caso que no se encuentre dentro de este rango es marcado e identificado individualmente.

13.4.1. Ejemplo

Se tienen los siguientes datos. 20, 23, 24, 24, 24, 25, 29, 31, 31, 33, 34, 36, 36, 37, 39, 39, 40, 40, 41, 45

Para calcular el primer cuartil se toma en cuenta que el 25 % de los valores de la distribución. Para este caso $N = 20$
 $N/4 = 5$

$$Q_1 = (24 + 25)/2 = 24,5$$

El segundo cuartil, es el valor de la variable que ocupa el lugar central en un conjunto de datos ordenados. Como $N/2 = 10$; la mediana es la media aritmética de dicho valor y el siguiente:

$$Q_2 = (33 + 34)/2 = 33,5$$

El Tercer Cuartil, es el valor que sobrepasa al 75 % de los valores de la distribución. En nuestro caso, como $3N/4 = 15$, resulta:

$$Q_3 = (39 + 39)/2 = 39$$

El bigote de la izquierda representa al colectivo de edades (X_{\min} , Q_1)

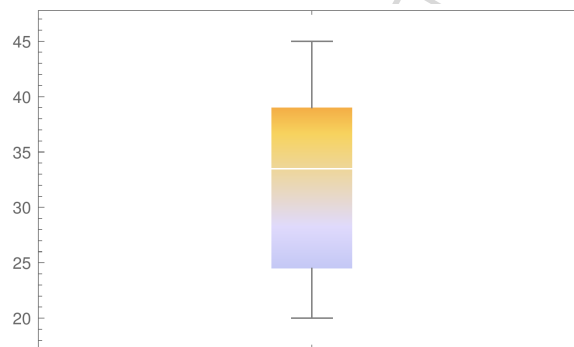


Figura 50: Diagrama Box and Whiskers.

La primera parte de la caja a (Q_1 , Q_2)

La segunda parte de la caja a (Q_2 , Q_3)

El bigote de la derecha esta dado por (Q_3 , X_{\max})

En Mathematica se usa el comando "BoxWhiskerChart" para crear el diagrama. requiere de una lista de datos. Se puede graficar mas de una lista.

```
a = RandomInteger[{1, 6}, 100];  
b = RandomInteger[{1, 10}, 500];  
c = RandomInteger[{1, 10}, 500];  
d = RandomInteger[{1, 10}, 500];  
e = RandomInteger[{1, 10}, 500];  
BoxWhiskerChart[{a, b, c, d, e}]
```

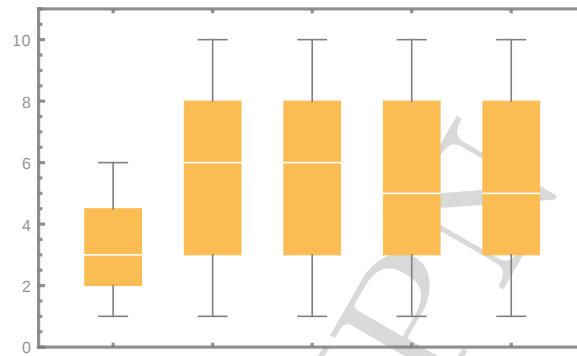


Figura 51: Diagrama Box and Whiskers.

14. Tarea 14

14.1. Demostraciones de la clase de Caminata al azar de [3]

14.1.1. Ecuación 46 de [3]

Una Serie de Taylor está definida por la siguiente expresion:

$$f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \dots \quad (60)$$

Si la función que queremos aproximar es $f(n_1) = \ln W(n_1)$, cerca de su máximo

$$\begin{aligned} n_1 &\equiv \tilde{n}_1 + \eta \\ \eta &= n_1 - \tilde{n}_1, \end{aligned}$$

Entonces usando la serie de Taylor de $f(n_1) = \ln W(n_1)$ alrededor del punto \tilde{n}_1 :

$$\ln W(n_1) = \ln W(\tilde{n}_1) + \frac{d \ln W(\tilde{n}_1)}{dn_1} \frac{(n_1 - \tilde{n}_1)}{1!} + \frac{d^2 \ln W(\tilde{n}_1)}{dn_1^2} \frac{(n_1 - \tilde{n}_1)^2}{2!} + \frac{d^3 \ln W(\tilde{n}_1)}{dn_1^3} \frac{(n_1 - \tilde{n}_1)^3}{3!} + \dots,$$

Finalmente, si se sustituye $\eta = n_1 - \tilde{n}_1$ y si se hace $B_k = \frac{d^k \ln W(\tilde{n}_1)}{dn_1^k}$, donde B_k es la k -ésima derivada evaluada en \tilde{n}_1 , donde $k \geq 1$ se obtiene la siguiente expresión:

$$\ln W(n_1) = \ln W(\tilde{n}_1) + B_1\eta + \frac{1}{2}B_2\eta^2 + \frac{1}{6}B_3\eta^3 + \dots \quad (61)$$

14.1.2. Ecuación 54 de [3]

$$\ln W(n_1) = \ln N! - \ln n_1! - \ln(N - n_1)! + n_1 \ln p + (N - n_1) \ln q, \quad (62)$$

Dada la ecuación de la distribución binomial $W(n_1)$:

$$W(n_1) = \frac{N!}{n_1!(N - n_1)!} p^{n_1} q^{N-n_1}$$

Podemos aplicar logaritmo natural a ambos lados de la ecuación, y reducir mediante las propiedades de logaritmos,

$$\ln W(n_1) = \ln \left(\frac{N!}{n_1!(N - n_1)!} p^{n_1} q^{N-n_1} \right), \quad (63)$$

Aplicando la propiedad de la multiplicacion de logaritmos tenemos:

$$\begin{aligned} \ln W(n_1) &= \ln \left(\frac{N!}{n_1!(N - n_1)!} \right) + \ln (p^{n_1} q^{N-n_1}), \\ \ln W(n_1) &= \ln \left(\frac{N!}{n_1!(N - n_1)!} \right) + \ln (p^{n_1}) + \ln (q^{N-n_1}), \end{aligned}$$

Se aplica la propiedad del cociente de logaritmos para el primer sumando de la ecuación anterior, así como la propiedad de los exponentes para los últimos dos términos. Se obtiene entonces:

$$\begin{aligned} \ln W(n_1) &= \ln(N!) - \ln(n_1!(N - n_1)!) + n_1 \ln p + (N - n_1) \ln q \\ \ln W(n_1) &= \ln(N!) - (\ln n_1! + \ln(N - n_1)!) + n_1 \ln p + (N - n_1) \ln q \end{aligned}$$

Finalmente se tiene:

$$\ln W(n_1) = \ln(N!) - \ln n_1! - \ln(N - n_1)! + n_1 \ln p + (N - n_1) \ln q \quad (64)$$

14.1.3. ecuación 55 de [3]

Partiendo de la ecuación 54, se observa que para cuando n es un entero grande de forma que $n \gg 1$, entonces $\ln(n!)$ se puede considerar casi una función continua de n , de forma que se puede aproximar la derivada de la función para la condición anterior con la consideración de que la función solo cambia una pequeña fracción de sí misma en cada paso. Si consideramos como $\ln(n+1!)$ como ese ligero cambio entre $\ln(n!)$ y el siguiente paso, entonces:

$$\frac{d \ln(n!)}{dn} \approx \ln(n+1!) - \ln(n!),$$

Por leyes de logaritmos:

$$\log_b \left(\frac{x}{y} \right) = \log_b(x) - \log_b(y),$$

Entonces:

$$\ln(n+1!) - \ln(n!) = \ln \frac{(n+1)!}{n!},$$

Ya que se hizo la consideración de que n es un entero muy grande de forma que $n \gg 1$ entonces:

$$\ln \frac{(n+1)!}{n!} = \ln(n+1)$$

Por lo que:

$$\frac{d \ln(n!)}{dn} \approx \ln(n+1) \quad (65)$$

14.1.4. Ecuación 56 de [3]

Ya que se considera n como un entero muy grande de forma que $n \gg 1$, entonces, mediante la fórmula de Stirling se tiene:

$$\ln n! \approx n \ln n - n, \quad (66)$$

Derivando ambos lados de la ecuación:

$$\begin{aligned} \frac{d \ln n!}{dn} &\approx \frac{d(n \ln n - n)}{dn}, \\ \frac{d \ln n!}{dn} &\approx \frac{n}{n} + \ln n - 1, \\ \frac{d \ln n!}{dn} &\approx \ln n. \end{aligned} \quad (67)$$

14.1.5. Ecuación 57 de [3]

Dada la ecuación:

$$\ln W(n_1) = \ln(N!) - \ln n_1! - \ln(N - n_1)! + n_1 \ln p + (N - n_1) \ln q, \quad (68)$$

Obtenemos la primera derivada con respecto a n_1 :

$$\begin{aligned} \frac{d \ln W(n_1)}{dn_1} &= \frac{d}{dn_1} (\ln(N!) - \ln n_1! - \ln(N - n_1)! + n_1 \ln p + (N - n_1) \ln q), \\ \frac{d \ln W(n_1)}{dn_1} &= \frac{d \ln(N!)}{dn_1} - \frac{d \ln n_1!}{dn_1} - \frac{d \ln(N - n_1)!}{dn_1} + \frac{dn_1 \ln p}{dn_1} + \frac{d(N - n_1) \ln q}{dn_1} \end{aligned}$$

Cada termino da como resultado lo siguiente:

$$\begin{aligned}\frac{d \ln (N!)}{dn_1} &= 0, \\ \frac{dn_1 \ln p}{dn_1} &= \ln p, \\ \frac{d(N - n_1) \ln q}{dn_1} &= -\ln q,\end{aligned}$$

Recordamos que $\frac{d \ln n!}{dn} \approx \ln n!$, por lo tanto

$$\begin{aligned}\frac{d \ln n_1!}{dn_1} &\approx \ln n_1, \\ \frac{d \ln (N - n_1)!}{dn_1} &\approx \ln (N - n_1),\end{aligned}$$

Finalmente la expresión final es:

$$\frac{d \ln W(n_1)}{dn_1} = -\ln n_1 + \ln (N - n_1) + \ln p - \ln q \quad (69)$$

14.1.6. Ecuación 58 de [3]

Igualando [57] de [3] a 0, se encuentra el valor $n_1 = \tilde{n}_1$, donde W es el máximo, entonces, se obtiene:

$$\begin{aligned}\frac{d \ln W}{dn} &= -\ln(\tilde{n}_1) + \ln(N - \tilde{n}_1) + \ln(p) - \ln(q) = 0, \\ [\ln(N - \tilde{n}_1) - \ln(\tilde{n}_1)] + [\ln(p) - \ln(q)] &= 0,\end{aligned} \quad (70)$$

Por propiedades de los logaritmos:

$$\begin{aligned}\ln\left(\frac{N - \tilde{n}_1}{\tilde{n}_1}\right) + \ln\left(\frac{p}{q}\right) &= 0 \\ \ln\left(\frac{N - \tilde{n}_1}{\tilde{n}_1} \frac{p}{q}\right) &= 0,\end{aligned}$$

14.1.7. Ecuación 59 de [3]

Reorganizando la ecuación 58 de [3]

$$\ln\left(\frac{N - \tilde{n}_1}{\tilde{n}_1} \frac{p}{q}\right) = 0, \quad (71)$$

$$e^{\ln\left(\frac{N - \tilde{n}_1}{\tilde{n}_1} \frac{p}{q}\right)} = e^0,$$

$$\frac{N - \tilde{n}_1}{\tilde{n}_1} \frac{p}{q} = 1,$$

$$(N - \tilde{n}_1)p = \tilde{n}_1 q. \quad (72)$$

14.1.8. Ecuación 60 de [3]

$$\begin{aligned}
(N - \tilde{n}_1)p &= \tilde{n}_1q, \\
Np - \tilde{n}_1p &= \tilde{n}_1q, \\
Np &= \tilde{n}_1q + \tilde{n}_1p, \\
Np &= \tilde{n}_1(q + p), \\
Np &= \tilde{n}_1,
\end{aligned} \tag{73}$$

Recordando que:

$$p + q = 1. \tag{74}$$

14.1.9. Ecuación 61 de [3]

$$\frac{d^2 \ln W(n_1)}{dn_1^2} = -\frac{1}{n_1} - \frac{1}{N - n_1}, \tag{75}$$

Partiendo la primera derivada

$$\begin{aligned}
\frac{d \ln W(n_1)}{dn_1} &= -\ln n_1 + \ln N - n_1 + \ln p - \ln q, \\
\frac{d}{dn_1} \left(\frac{d \ln W(n_1)}{dn_1} \right) &= \frac{d}{dn_1} (-\ln n_1 + \ln N - n_1 + \ln p - \ln q), \\
\frac{d^2 \ln W(n_1)}{dn_1^2} &= -\frac{d(\ln n_1)}{dn_1} + \frac{d \ln(N - n_1)}{dn_1} + \frac{d(\ln p)}{dn_1} - \frac{d(\ln q)}{dn_1},
\end{aligned}$$

Resolviendo las derivadas se tiene que:

$$\begin{aligned}
-\frac{d(\ln n_1)}{dn_1} &= -\frac{1}{n_1}, \\
\frac{d \ln(N - n_1)}{dn_1} &= \frac{d \ln(N - n_1)}{dn_1} \frac{d(N - n_1)}{dn_1} \\
&= \frac{1}{N - n_1} (-1) = -\frac{1}{N - n_1}, \\
\frac{d \ln p}{dn_1} &= 0, \\
\frac{d \ln q}{dn_1} &= 0,
\end{aligned}$$

Sumando los términos obtenidos, comprobamos que:

$$\frac{d^2 \ln W(n_1)}{dn_1^2} = -\frac{1}{n_1} - \frac{1}{N - n_1} \tag{76}$$

14.1.10. ecuación 62 de [3]

Evaluando la ecuación 61 de [3] con lo obtenido en la ecuación 60 de [3] se obtiene:

$$b_2 = -\frac{1}{Np} - \frac{1}{N - Np}, \tag{77}$$

Desarrollando se obtiene:

$$\begin{aligned}b_2 &= -\frac{1}{N} \frac{1}{p} - \frac{1}{N(1-p)}, \\b_2 &= -\frac{1}{N} \frac{1}{p} - \left(\frac{1}{N} \frac{1}{(1-p)} \right), \\b_2 &= -\frac{1}{N} \frac{1}{p} - \left(\frac{1}{N} \frac{1}{q} \right), \\b_2 &= -\frac{1}{N} \left(\frac{1}{p} + \frac{1}{q} \right), \\b_2 &= -\frac{1}{N} \frac{q+p}{pq},\end{aligned}$$

Dado que $p + q = 1$

$$b_2 = -\frac{1}{Npq}. \quad (78)$$

15. Tarea 15

15.1. Congruencia Zeller

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

Para el calendario gregoriano la congruencia de Zeller es

$$h = \left(q + \left\lfloor \frac{(m+1)26}{10} \right\rfloor + K + \left\lfloor \frac{K}{4} \right\rfloor + \left\lfloor \frac{J}{4} \right\rfloor - 2J \right) \text{ mód } 7, \quad (79)$$

donde:

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

q es el día del mes

m es el mes

J es el siglo

K el año de la centuria (año mod 100)

Enero y febrero se cuentan como meses 13 y 14 del año anterior. Obsérvese, que el 2 de enero del 2013, es $m=13$; año=2012 Es oportuno recordar que la función mod es el residuo que queda de la división de dos números. En las implementaciones informáticas en las que el módulo de un número negativo es negativo, la manera más sencilla de obtener un resultado entre 0 y 6 es reemplazar $-2J$ por $+5J$ y $-J$ por $+6J$.

15.1.1. Algoritmo

```
a = (14 - Mes) / 12
y = Anio - a
m = Mes + 12 * a - 2
```

Para el calendario Juliano:

```
d = (5 + dia + y + y/4 + (31*m)/12) mod 7
```

Para el calendario Gregoriano:

```
d = (d\`ia + y + y/4 - y/100 + y/400 + (31*m)/12) mod 7
```

El resultado es un cero (0) para el domingo, 1 para el lunes 6 para el sábado

Ejemplo, ¿En qu\`e d\`ia de la semana cae el 2 de agosto de 1953??

```
' a = (14 - 8) / 12 = 0
' y = 1953 - 0 = 1953
' m = 8 + 12 * 0 - 2 = 6
' d = (2 + 1953 + 1953 / 4 - 1953 / 100 + 1953 / 400 + (31 * 6) / 12) Mod 7
'   = (2 + 1953 + 488 - 19 + 4 + 15) mod 7
'   = 2443 mod 7
'   = 0
' El valor cero(0) corresponde al domingo.
```

15.1.2. Código en matemática

```

InputField[Dynamic[dia], Number]
Dynamic[dia];
InputField[Dynamic[mes], String]
Dynamic[mes];
InputField[Dynamic[anioo], Number]
InputField[Dynamic[d], Number]
Dynamic[anioo];
Dynamic[month];
Dynamic[month =
  Switch[mes, "Enero", 1, "Febrero", 2, "Marzo", 3, "Abril", 4,
    "Mayo", 5, "Junio", 6, "Julio", 7, "Agosto", 8, "Septiembre", 9,
    "Octubre", 10, "Noviembre", 11, "Diciembre", 12]];

Dynamic[a = (14 - month)/12]
Dynamic[y = anioo - a]
Dynamic[m = month + 12*a - 2]
Dynamic[d = Mod[(dia + y + (y/4) - (y/100) + (y/400) + (31*m/12)), 7] + 1]

```

16. Tarea 16

16.1. Función de Distribución de Energía

La función de distribución $f(E)$, es la probabilidad de que una partícula se encuentre en el estado de energía E . La función de distribución es una generalización de las ideas de probabilidad discreta, para el caso donde la energía puede ser tratada como una variable continua. En la naturaleza se encuentran tres funciones de distribución claramente diferentes. El término A en el denominador de cada distribución es un término de normalización, que puede cambiar con la temperatura.

16.1.1. Distribución Maxwell-Boltzmann

La distribución de Maxwell-Boltzmann es la función de distribución clásica, para la distribución de una cantidad de energía entre partículas idénticas pero distinguibles.

$$f(E) = \frac{1}{A e^{\frac{E}{kT}}}, \quad (80)$$

Donde:

$f(E)$: Es la probabilidad de que una partícula tenga energía E

A Es una constante de normalización

k Es la constante de Boltzman

T Es la temperatura de la partícula (temperatura absoluta)

Para una temperatura mayor, es más probable que se pueda encontrar una partícula determinada con la energía E . La probabilidad de ocupar un estado de energía determinado disminuye con la energía. Con el incremento de energía E es menos probable que cualquier partícula dada alcance esa energía, por lo que se encontraran más partículas con energía menores. Se asume que un número ilimitado de partículas puede ocupar cualquier estado de energía.

Además de la presunción de distinguibilidad, la física estadística clásica postula que:

- No hay ninguna restricción sobre el número de partículas que pueden ocupar un estado dado.
- En el equilibrio térmico, la distribución de partículas entre los estados de energía disponibles, se llevará a cabo con la distribución más probable, la cual es consistente con la energía total disponible y el número total de partículas.
- Cada estado específico del sistema tiene la misma probabilidad.

Una de las ideas generales contenidas en estos postulados, es que es poco probable que cualquier partícula consiga una energía muy por encima de la media (es decir, mucho más de lo que le corresponde). Las energías más bajas que la media, se ven favorecidas debido a que hay más maneras de conseguirlas. Si por ejemplo, una partícula tiene una energía de 10 veces el promedio, entonces, reduce el número de posibilidades para la distribución del resto de la energía. Por lo tanto, es poco probable ya que la probabilidad de ocupación de un estado dado, es proporcional al número de formas en que se puede conseguir.

16.1.2. Distribución Fermi-Dirac

La distribución de Fermi-Dirac se aplica a los fermiones, partículas con espín semientero, que obedece el principio de exclusión de Pauli. Cada tipo de función de distribución tiene un término de normalización multiplicando el denominador del exponente, que puede ser dependiente de la temperatura. Para el caso de Fermi-Dirac, ese término se suele escribir:

$$e^{-E_f/kT}, \quad (81)$$

Donde:

E_f es la energía de Fermi

La importancia de la energía de Fermi se ve más claramente estableciendo $T=0$. En el cero absoluto, la probabilidad es $=1$ para energías menores que la energía de Fermi y cero para energías mayores que la energía de Fermi. Se puede imaginar todos los niveles hasta la energía de Fermi llenos, pero ninguna partícula tiene una energía mayor. Esto es totalmente coherente con el principio de exclusión de Pauli, donde cada estado cuántico no puede tener más que una sola partícula.

$$f(E) = \frac{1}{e^{(E-E_F)/kT} + 1}, \quad (82)$$

$f(E)$ Es la probabilidad de que una partícula tenga energía E

Para bajas temperaturas, aquellos estados de energía por debajo de la energía de Fermi E_F tienen una probabilidad de esencialmente 1 y aquellos arriba de la energía de Fermi tienen esencialmente 0

16.1.3. Distribución Bose Einstein

La distribución de Bose-Einstein describe el comportamiento estadístico de las partículas de espín entero (bosones). A bajas temperaturas, los bosones se comportan de manera muy diferente a los fermiones, debido a que un número ilimitado de ellos pueden captar el mismo estado de energía, un fenómeno llamado condensación"

$$f(E) = \frac{1}{Ae^{E/kT} - 1}, \quad (83)$$

Donde:

$f(E)$ Es la probabilidad de que un partícula tenga energía E.

Para fotones $A=1$, la ocupación de estados de energía muy bajos puede incrementar sin límite

CIC-IPN

17. Tarea 17

17.1. La última oveja de pie

There are 100 equally spaced points around a circle. At 99 of the points, there are sheep, and at 1 point, there is a wolf. At each time step, the wolf randomly moves either clockwise or counter-clockwise by 1 point. If there is a sheep at that point, he eats it. The sheep don't move. What is the probability that the sheep who is initially opposite the wolf is the last one remaining?

17.1.1. Solución 1

Imagine a circle with n points equally spaced around the circle. For convenience, we will label the points 0 through $n - 1$ moving around the circle counter-clockwise. At points 1 through $n - 1$, we place sheep, while a wolf is placed at point 0. The wolf randomly chooses to move clockwise or counter-clockwise (by flipping a coin, say) and eats the sheep residing at the adjacent point in that direction. That is, the wolf is equally likely to eat the sheep at points 1 and $n - 1$. The wolf continues in the same manner: at each step, it chooses to move to the adjacent point either clockwise or counter-clockwise, each choice being equally likely. Which sheep is most likely to be eaten last?

At first blush, one would think (or at least I thought) that the sheep opposite the wolf is most likely to be eaten last. However, it turns out that all sheep are equally likely to be the last living sheep. To see why this is the case, consider what must happen in order for sheep i to be eaten last: both of its neighbors must be eaten before i . In particular, the wolf must eat one of its neighbors, then work its way around the circle to its other neighbor before eating i . We can compute the probability that this happens. To simplify matters, we will look at sheep $n-1$ where the wolf is initially at 0 (which is $(n - 1)$'s neighbor on the circle). Then the question becomes, what is the probability that the wolf will eat sheep $n - 2$ before eating $n - 1$? Denote this probability by p . (I suppose that by symmetry we have already shown that all sheep are equally likely to be eaten last because some sheep has to be eaten last, and the only way for a sheep to be eaten last is if both of its neighbors are eaten before it. Nonetheless it is nice to see how to compute p explicitly.)

In order to compute p , it is helpful to rephrase the problem. Consider everything happening on a line instead of a circle. We label points $-1, 0, 1, \dots, n - 2$ on the line. The wolf starts at point 0. The points -1 and $n - 2$ correspond to sheep $n - 1$ and $n - 2$ respectively on the circle. The wolf now moves randomly one step to the right or left and continues eating sheep as before. Now we can reinterpret the event that sheep $n - 1$ is eaten last on the circle as the event that the wolf eats $n - 2$ before -1 on the line (this situation is often referred to as the gambler's ruin). We assume that the wolf stops moving as soon as it gets to either -1 or $n - 2$. Let X_t be the position of the wolf at time t . Then the expected value of X_t is

$$\mathbf{E}[X_t] = 0$$

for all t .

We arrive at this formula inductively. Since the wolf starts at 0, we certainly have $\mathbf{E}[X_0] = 0$. For $t \geq 1$ we consider two cases separately: if $0 \leq X_t \leq n - 3$ then $X_{t+1} = X_t \pm 1$ with both possibilities equally likely. If $X_t = -1$ or $n - 2$, then $X_{t+1} = X_t$. In either event, we have $\mathbf{E}[X_{t+1}] = \mathbf{E}[X_t]$, which closes the induction. Since the wolf will almost surely get to -1 or $n - 2$, we have

$$(-1)(1 - p) + (n - 2)p = \lim_{t \rightarrow \infty} \mathbf{E}[X_t] = 0$$

Solving this equation for p gives $p = 1/(n - 1)$. The same holds for each sheep on the circle. In particular, each sheep is equally likely to be eaten last. Still, one can show that the sheep opposite the wolf (i.e., sheep $n/2$) has the longest expected lifetime, and that sheep's lifetime is $n^2/4$ steps.

17.1.2. Solución 2

Imagina un círculo de n puntos igualmente espaciado en un círculo, etiquetamos los puntos desde 0 hasta $n - 1$, moviéndonos en sentido contrario a las manecillas del reloj.

En los puntos 1 hasta $n - 1$ colocamos una oveja, mientras que un lobo se coloca en el punto 0.

El lobo se mueve aleatoriamente en sentido del reloj o en contra sentido del reloj y se come la oveja que está en el punto adyacente en esa dirección. El lobo continúa de la misma manera: en cada paso, elije moverse al punto adyacente en sentido del reloj o al contrario, cada elección es igualmente probable. ¿Cuál es la oveja que tiene mayor probabilidad de ser comida al último?

En primera instancia, uno pensaría que la oveja al lado opuesto del lobo es la que tiene mayor probabilidad de ser comida al último. Sin embargo, resulta que todas las ovejas tienen la misma probabilidad de ser comidas al último.

Para una cierta oveja i : Ambos vecinos de i deben ser comidos antes que i . En particular, el lobo debe comer sólo uno de los vecinos de i , después el lobo caminará en círculo hasta llegar al segundo vecino de i antes de comerse a i . Podemos calcular la probabilidad de que esto pase. Para simplificar las cosas, veremos a la oveja en $n - 1$ donde el lobo está inicialmente en 0 (el cuál es el vecino de $n - 1$ en el círculo). La pregunta entonces es: "¿Cuál es la probabilidad que el lobo se coma a la oveja $n - 2$ antes de comerse a la oveja $n - 1$?" Llamemos a esta probabilidad P .

Para calcular P , es de ayuda reformular el problema. Consideramos que todo pasa en una línea en lugar de un círculo. Etiquetamos los puntos $-1, 0, 1, \dots, n - 2$ en la línea. El lobo inicia en 0. Los puntos -1 y $n - 2$ corresponden a las ovejas $n - 1$ y $n - 2$ respectivamente. El lobo ahora se mueve aleatoriamente un paso a la derecha o izquierda y continúa comiendo ovejas como antes. Ahora podemos reinterpretar el evento donde la oveja $n - 1$ es comida al último en el círculo como el evento en el que el lobo se come a $n - 2$ antes que -1 en la línea (*Gambler's ruin*). Asumimos que el lobo para de moverse tan pronto como llega a -1 o $n - 2$.

X_t es la posición del lobo en el tiempo t . Entonces el valor esperado de X_t es:

$$E[X_t] = 0 \text{ para todo } t$$

Llegamos a esta fórmula inductivamente. Ya que el lobo inicia en 0, tenemos certeramente $E[X_t] = 0$. Para $t \geq 1$ consideramos dos casos por separado:

- si $0 \leq X_t \leq n - 3$ entonces $X_{t+1} = X_t \pm 1$ con ambas posibilidades igualmente probables.
- si $X_t = -1$ o $n - 2$, entonces $X_{t+1} = X_t$

En cualquier caso, tenemos $E[X_{t+1}] = E[X_t]$, lo que se acerca a la inducción.

Ya que $E[X_t]$ no depende de t , tenemos:

$$\lim_{t \rightarrow \infty} E[X_t] = 0$$

Y ya que el lobo se comerá casi por seguro a -1 o $n - 2$, tenemos:

$$(-1)(1 - P) + (n - 2)P = \lim_{t \rightarrow \infty} E[X_t] = 0$$

Resolviendo esta ecuación para P se tiene $P = 1/(n - 1)$. Lo mismo se aplica para cada oveja en el círculo. En particular, cada oveja tiene la misma probabilidad de ser comida al último.

17.1.3. Solución 3

Probabilidad de avance $(+1) \rightarrow p$

Probabilidad de retroceso $(-1) \rightarrow q = 1 - p$

$P_0 = 0 \rightarrow$ Probabilidad que se coma a n antes que 0

$P_n = 1 \rightarrow$ Probabilidad que se coma a n antes que 0

$$P_i = pP_{i+1} + qP_{i-1} \quad (84)$$

P_i es la probabilidad que el lobo se coma a n antes que 0

Ya que $p + q = 1$, entonces:

$$\begin{aligned} (p + q)P_i &= pP_{i+1} + qP_{i-1} \\ pP_i + qP_i &= pP_{i+1} + qP_{i-1} \\ P_{i+1} - P_i &= \frac{q}{p}(P_i - P_{i-1}) \end{aligned} \quad (85)$$

Para el caso $i = 1$ (Cuando el lobo comienza en 1)

$$P_2 - P_1 = \frac{q}{p}(P_1 - P_0) = \frac{q}{p}P_1$$

Para el caso $i=2$

$$P_3 - P_2 = \frac{q}{p}(P_2 - P_1) = \left(\frac{q}{p}\right)^2 P_1$$

Para el caso general i

$$P_{i+1} - P_i = \left(\frac{q}{p}\right)^i P_1$$

Sumando todas las ecuaciones (desde 1 hasta i)

$$P_{i+1} - P_1 = \sum_{k=1}^i (P_{k+1} - P_k) = \sum_{k=1}^i \left(\frac{q}{p}\right)^k P_1$$

Despejando P_{i+1}

$$P_{i+1} = P_1 + P_1 \sum_{k=1}^i \left(\frac{q}{p}\right)^k = P_1 \sum_{k=0}^i \left(\frac{q}{p}\right)^k \quad (86)$$

$$P_{i+1} = \begin{cases} P_1 \left(\frac{1 - \left(\frac{q}{p}\right)^{i+1}}{1 - \left(\frac{q}{p}\right)} \right) & \text{si } p \neq q \\ P_1(i+1) & \text{si } p = q \end{cases} \quad (87)$$

Suponiendo que $i = n - 1$

$$P_{i+1} = \begin{cases} P_1 \left(\frac{1 - \left(\frac{q}{p}\right)^n}{1 - \left(\frac{q}{p}\right)} \right) & \text{si } p \neq q \\ P_1(n) & \text{si } p = q \end{cases}$$

Y ya que $P_n = 1$

$$\begin{aligned} P_n &= P_i n \\ 1 &= P_1 n \\ P_1 &= \frac{1}{n} \end{aligned} \quad (88)$$

Para un caso general en el que el lobo comienza en la posición i

$$P_{i+1} = \begin{cases} P_1 \left(\frac{1 - \left(\frac{q}{p}\right)^i}{1 - \left(\frac{q}{p}\right)^n} \right) & \text{si } p \neq q \\ \frac{i}{n} & \text{si } p = q \end{cases} \quad (89)$$

CIC-IPN

18. Tarea 18

18.1. Series Geométricas

Una serie geométrica es una serie en la cual la razón entre los términos sucesivos de la serie permanece constante.

Ejemplo:

La serie

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots + \frac{1}{2^n},$$

Es geométrica debido a que cada término sucesivo se obtiene al multiplicar el anterior por $1/2$.

18.1.1. Suma

La suma de una serie geométrica será finita siempre y cuando los términos se aproximen a cero; a medida que se acercan al cero, las cantidades se vuelven insignificamente pequeñas, permitiendo calcular la suma sin importar el hecho que la serie sea infinita. La suma puede ser obtenida utilizando las propiedades autosimilares de la serie.

Para $r \neq 1$, la suma de los primeros n términos de una serie geométrica es:

$$a + ar + ar^2 + ar^3 + \cdots + ar^{n-1} = \sum_{k=0}^{n-1} ar^k = a \frac{1 - r^n}{1 - r}, \quad (90)$$

y su demostración es la siguiente Sea:

$$s = a + ar + ar^2 + ar^3 + \cdots + ar^{n-1}$$

Entonces:

$$rs = ar + ar^2 + ar^3 + ar^4 + \cdots + ar^n,$$

$$s - rs = s(1 - r) = a - ar^n,$$

Finalmente:

$$s = a \frac{1 - r^n}{1 - r}. \quad (91)$$

18.1.2. Convergencia

La serie geométrica real de término inicial $a \in \mathbb{R}$ no nulo y de razón $r \in \mathbb{R}$ es convergente si y solamente si $|r| < 1$.

En tal caso, su suma vale:

$$\sum_{n=0}^{\infty} ar^n = \frac{a}{1 - r} \quad (92)$$

18.2. Ejemplos

1 Determine la formula para la sumatoria de $\sum_{i=0}^n \frac{9^i - 7^i}{11^i}$.

$$\sum_{i=0}^n \frac{9^i - 7^i}{11^i} = \sum_{i=0}^n \left(\frac{9}{11}\right)^i - \sum_{i=0}^n \left(\frac{7}{11}\right)^i,$$

Aplicando la formula

$$\begin{aligned} & \frac{1 - \left(\frac{9}{11}\right)^{n+1}}{1 - \frac{9}{11}} - \frac{1 - \left(\frac{7}{11}\right)^{n+1}}{1 - \frac{7}{11}}, \\ & -\frac{11}{2} \left(\frac{9}{11}\right)^{n+1} + \frac{11}{4} \left(\frac{7}{11}\right)^{n+1}, \end{aligned}$$

2 Determine la formula para la sumatoria de $\sum_{j=1}^n \sum_{i=0}^{\infty} j^{\frac{5}{3}} * \left(1 - \frac{1}{2j^{\frac{1}{3}}}\right)^i$.

$$\sum_{i=0}^{\infty} \left(1 - \frac{1}{2j^{\frac{1}{3}}}\right)^i = \frac{1}{1 - \left(1 - \frac{1}{2j^{\frac{1}{3}}}\right)} = 2j^{\frac{1}{3}},$$

Uniando a a segunda parte

$$\sum_{j=1}^n 2j^{\frac{1}{3}} = \frac{2n(2n+1)(n+1)}{3}.$$

3 La serie geométrica: $\sum_{n=0}^{\infty} \left(\frac{8}{10^2}\right) \left(\frac{1}{10^2}\right)^n$.

$$\sum_{n=0}^{\infty} \left(\frac{8}{10^2}\right) \left(\frac{1}{10^2}\right)^n = \frac{8}{10^2} + \frac{8}{10^4} + \frac{8}{10^6} + \frac{8}{10^8} + \dots,$$

Tiene $r = \frac{1}{10^2}$ y $a = \frac{8}{10^2}$
como $|r| < 1$ la serie converge, y su suma es:

$$\frac{a}{1-r} = \frac{(8/10^2)}{1-(1/10^2)} = \frac{8}{99}.$$

4 La serie geométrica: $\sum_{n=0}^{\infty} \left(\frac{3}{2}\right)^n$.

$$\sum_{n=0}^{\infty} \left(\frac{3}{2}\right)^n = 1 + \frac{3}{2} + \left(\frac{3}{2}\right)^2 + \left(\frac{3}{2}\right)^3 + \dots,$$

Como $|r| \geq 1$ la serie diverge, y no tiene suma.

Referencias

- [1] V. Balakrishnan, *Schaum's Outline of Combinatorics*. McGraw Hill Professional, 1995.
- [2] F. Reif and R. Knacke, *Fundamentals of Statistical and Thermal Physics: Solutions Manual*. McGraw-Hill series in fundamentals of physics, McGraw Hill, 1965.
- [3] M. C. J. Alberto, *Clase 11 de Probabilidad Procesos Aleatorios e inferencia*. CIC-IPN, 2015.
- [4] D. C. Montgomery and G. C. Runger, *Applied statistics and probability for engineers*. John Wiley & Sons, 2010.
- [5] M. R. Spiegel, J. J. Schiller, R. A. Srinivasan, and M. Stephens, *Probabilidad y estadística*. McGraw-Hill, 2010.

CIC-IPN