Introducción

El presente manual está pensado como libro de texto de la asignatura Métodos Estadísticos y Numéricos, optativa del proyecto curricular de 2º de Bachillerato LOGSE. Está en su totalidad rodado y trabajado durante un curso académico con alumnos de dicho curso en el Colegio Marista Cristo Rey de La Coruña, con excelente aprovechamiento.

Además de ello y en la esperanza de que pueda ser de utilidad como manual de laboratorio o libro auxiliar en los cursos de Estadística de los primeros cursos de carreras científicas y técnicas, incluye capítulos que exceden el diseño curricular base de 2º de Bachillerato, adentrándose por un extremo en técnicas simples de interpolación y probabilidad y por otro en problemas de programación lineal de mayor nivel o utilización de herramientas informáticas habituales en la resolución de cálculos estadísticos.

En todos los temas se ha procurado reducir al mínimo indispensable el contenido teórico, del mismo modo que se ha pretendido enfatizar especialmente la importancia de comprender las herramientas que en el texto se compendian y el auténtico conocimiento y aprehensión de procedimientos y protocolos de trabajo en técnicas estadísticas. Para ello, el autor ha centrado su trabajo de forma especial en la creación y/o clarificación de tales protocolos, intentando que el lector, ante un problema de análisis estadístico concreto, pueda observar la situación, sus datos y con mínimas dosis de intuición, adentrarse en la solución del problema o en la realización del estudio necesario. O bien ante una incógnita o una hipótesis por verificar, sea mínimamente autosuficiente a la hora de plantear su estudio.

Es por ello que cada capítulo se inicia con una apartado titulado "¿Para qué?", donde escuetamente se expone un caso particular de aplicación de las herramientas del capítulo o una pequeña y simple glosa sobre la utilidad y uso de la misma.

De la misma forma, cada capítulo finaliza con la biografía de un personaje relevante dentro del mundo de la Estadística, lo que pretende mostrar dos cosas fundamentales: lo variopinto de sus personalidades y sus ambiciones personales. Lo cual dará idea de lo profundamente multidisciplinar que es la materia del texto, no tanto en su tratamiento (puramente matemático) como en sus fines y el amplio campo de investigación que esta materia presenta, pues se presta a todo tipo de especulación científica.

El pequeño diccionario de estadística que se incluye al final del libro no es más que un intento de soslayar esos pequeños "vuelos de ángel" que todos (el ladrón a todos los ve de su condición) tenemos en alguna ocasión y pretende, como es lógico, ser de auxilio en cualquier instante si el lector tropieza con algún término del texto cuyo significado propio no alcance a entender o recordar.

El autor, modestamente, ha pretendido crear una obra que si bien carece de originalidad en cuanto al contenido, ya que obviamente no ha realizado investigación sino bibliográfica en la materia y de buceo en sus notas de profesor, sí es considerablemente novedosa en cuanto al planteamiento general (supone el autor que igual de modestamente que admite su falta de descubrimientos en la materia, es lícito que se precie del mucho trabajo realizado), pues no pretende tanto mostrar o enseñar la Estadística como materia "reglada", como enseñar a "hacer estadística" con auténtica conexión con el mundo real, el método científico y el planteamiento de problemas.

Para este autor, como profesor, y esta es ya una reflexión que excede los límites normales de una introducción, carece de sentido que los alumnos adquieran formulismos de rigor matemático importante, (que por otra parte y personalmente encuentra apasionantes) y prefiere que ante un problema científico, sepan reaccionar de la forma adecuada, identificándolo, planteándolo, diseccionándolo y llegando a su solución aunque para ello deban consultar la fórmula concreta o detalles del protocolo en las fuentes a su alcance.

En este sentido, de apego al mundo e investigación reales, se programan en el libro algo más de una docena de "trabajos de campo" que en el año de rodaje del texto que se comentaba al principio de esta introducción fueron realizados en grupos, a razón de uno por cuatrimestre, incluyendo como tarea de terminación la defensa de los mismos ante los compañeros de curso.

Algunos de dichos trabajos de campo son en realidad el resultado de sugerencias realizadas por los propios alumnos, para evitar, todo sea dicho, elegir de entre la docena de ellos que este profesor les propuso; no hace falta que se avise al lector de cuáles son aquellos (los propuestos por el alumnado); observarán que la toma de datos y la elaboración de los mismos, es en todos los casos y por decirlo suavemente, "algo más liviana" que en los doce originales.
El autor tiene en preparación una obra complementaria de la presente, con la solución detallada del centenar largo de problemas propuestos y otro centenar de problemas nuevos de nivel de complejidad superior.

Teodoro Rodríguez

Tabla de contenidos

OPERATIVIDAD BÁSICA

¿Para qué?

Para recordar determinados procedimientos matemáticos sin cuyo conocimiento no se podrá comprender el contenido de la materia y refrescar técnicas de planteamiento y resolución de problemas elementales, que facilitarán el acceso a los propios de los temas que trataremos.

Operatividad

Se propone este capítulo inicial como un breve repaso de contenidos matemáticos de tipo práctico, fundamentales para el estudio del texto, con los que el lector podrá identificar sus "carencias matemáticas" y preparar las correspondientes sesiones de reciclaje o repaso.

Ecuaciones

1. $x^2+2x-3=0$	2. $10^{x^2-11x+30}=2^2\cdot 5^2$
3. $9x^2-24x+16=0$	4. $8^{x^2+3x+2}=1$
5. $3x^2+8x+5=0$	6. $3^{x-1}+3^x+3^{x+1}=117$
7. $x^4-12x^2+27=0$	8. $3^{2(x+1)}-28\cdot 3x+3=0$
9. $x^4+100=29x^2$	10. $3^x+\frac{1}{3^{x-1}}=4$
11. $x^2+\frac{324}{x^2}=45$	12. $2^{x+2}=0{,}5^{2x-1}$
13. $4=5x^2-x^4$	14. $\left(\frac{2}{7}\right)^5=3{,}5^{x+1}$
15. $x-\sqrt{2x+3}=6$	16. $\frac{\log(16-x^2)}{\log(3x-4)}=2$
17. $3-x=2\sqrt{x}$	18. $\log 8+(x^2-5x+7)\log 3=\log 24$
19. $3\sqrt{3x-1}=2\sqrt{3(2x-1)}$	20. $2\log x=\log\frac{x}{2}-\frac{7}{4}$
21. $\sqrt{2x+3}-\sqrt{x-2}=2$	22. $\log x^3=\log 6+2\log x$
23. $2+\sqrt{2x+2}=x-1$	24. $(x^2-x-3)\log 4=3\log\frac{1}{4}$
25. $2^x\cdot 3^x=12\cdot 18$	26. $\log x+\log 50=\log 1000$
27. $4^x\cdot 5^{x-1}=1600$	28. $\log(25-x^3)-3\log(4-x)=0$
29. $9^{2x+3}=3^{2x+5}$	30. $\log x=1+\log(22-x)$

Sistemas

1. $\begin{cases} y+z=3+x \\ x+z=y+z \\ x+y=z+1 \end{cases}$	2. $\begin{cases} 2^x \cdot 2^{3y+1} = 4 \\ 4x-3y=19 \end{cases}$
3. $\begin{cases} 5x+3=4y \\ -10x+8y=-2 \end{cases}$	4. $\begin{cases} 3x-y+2z=4 \\ 4x-2y-z=1 \\ -x+3y-3z=-1 \end{cases}$

Problemas

1. Resuelve gráficamente la siguiente ecuación: $2x^2+2x+4=6$

2. Uno de los lados de un rectángulo es triple que el otro. Si su superficie es de 75 Uds², ¿cuáles son sus dimensiones?. Calcula la longitud de su diagonal.

3. La suma de las tres cifras de un número es 6; si se intercambian la 1ª y la 2ª, el número aumenta en 90 Uds., pero si se intercambian la 2ª y la 3ª, el número aumenta en 9 Uds. Calcula dicho número.

4. Determina dos números consecutivos, tales que la suma de sus inversos, sea 11/30.

5. La suma de las dos cifras de un número es 9. Si a dicho número le restamos 45, el número que resulta es igual al que se obtiene cambiando de orden las cifras del número original. Calcula dicho número.

6. La edad de Julia es doble que la de Javier. Hace 7 años, la suma de sus edades era igual a la edad actual de Julia. ¿Qué edad tienen ahora Julia y Javier?.

7. Se desea mezclar harina de 55 ptas./k con otra de 40 ptas./k, de modo que la mezcla resulte a 45 ptas./k. ¿Cuántos kilos de cada clase de harina hay que mezclar para obtener 300 kilos de mezcla al precio mencionado?.

Intervalos y entornos

Representa gráfica y algebraicamente los siguientes intervalos y entornos:

a) $E^*(-3,4)$
b) $E^+(2,4)$
c) $(0,+\infty)$

Inecuaciones

1. $x-2>\frac{x}{3}$	2. $\frac{2x+6}{x-5}\geq 0$
3. $3x-\frac{x}{6}+2>3+\frac{5x}{2}$	4. $\frac{x-8}{3x-6}<0$
5. $x^2-2x+1<0$ *Analítica...y...gráficamente*	6. $x\leq 4-2y$
7. $\begin{cases}3x+\frac{1}{5}<4x+\frac{3}{2}\\ 25x+\frac{5x}{3}>102x-\frac{4}{7}\end{cases}$	8. $\begin{cases}x-y>6\\ x>2y+3\end{cases}$
9. $\frac{3x^2+4}{(x-1)(x-2)}\geq 0$	10. $\begin{cases}2x-y\geq 2\\ 3x+y>2\end{cases}$

Cálculo de áreas

1. Determina el área encerrada entre la función $y=x^2-5x+6$ y el eje OX
2. Determina el área encerrada entre la función $y=x^2-5x+6$ y la recta $y=2$
3. Determina el área encerrada entre las curvas $y=x^2+5$ e $y=-x^2+3$
4. Determina el área encerrada entre las curvas $y=x^2+5$ e $y=-x^2+9$

INTERPOLACIÓN

¿Para qué?

Para calcular un dato desconocido entre otros que conocemos, para hacer previsiones elementales de futuro basadas en datos del pasado y presente

Interpolación

Definición

Interpolar es encontrar un valor desconocido, entre varios conocidos, de una serie de datos.
Matemáticamente, dados los puntos (x_0,y_0) (x_1,y_1).......(x_n,y_n) conocidos, se llama función de interpolación a la función y=f(x) más sencilla que para $x=x_0$, $x=x_1$, $x=x_2$.....$x=x_n$, toma los valores y_1, y_2,.....y_n respectivamente.
Una vez obtenida la función de interpolación, y=f(x), denominamos interpolar a obtener el valor de f(x) en el punto x=a, comprendido entre los valores menor y mayor de la serie conocida.

Interpolación lineal

Se trata de realizar la interpolación entre dos puntos, tomando como referencia una función lineal (una recta) que pasa por ellos.

-Dados los puntos (x_0,y_0) y (x_1,y_1), con abscisas distintas, existe una única función polinómica de primer grado que pasa por ellos y cuya expresión algebraica es y=mx+b.

-Para determinar dicha función de interpolación, resolvemos el sistema

$$\begin{cases} y_0 = mx_0 + b \\ y_1 = mx_1 + b \end{cases}$$

Interpolación cuadrática

Se trata de realizar la interpolación entre dos puntos, tomando como referencia una función polinómica de 2º grado, cuya gráfica será por lo tanto una parábola.

-Dados 3 puntos (x_0,y_0) , (x_1,y_1), (x_2,y_2) con abscisas distintas y no alineados, existe una única función polinómica de 2º grado que pasa por ellos y cuya expresión algebraica es $y=ax^2+bx+c$

-Para determinar la función de interpolación, resolvemos el sistema

$$\begin{cases} y_0 = ax_0^2 + bx_0 + c \\ y_1 = ax_1^2 + bx_1 + c \\ y_2 = ax_2^2 + bx_2 + c \end{cases}$$

Interpolación polinómica general

Generalizando los hechos anteriores, podemos decir que que si conocemos n+1 puntos de abscisas diferentes, se puede determinar una función de interpolación general de grado n, de la forma $y = a_n x^n + a_{n-1}x^{n-1} + ... + a_2 x^2 + a_1 x + a_0$

-Para determinar tal función, utilizamos el hecho de que los n+1 puntos pertenecen a dicha función y planteamos el siguiente sistema de ecuaciones

$$\begin{cases} y_0 = a_n x_0^n + a_{n-1}x_0^{n-1} + ... + a_2 x_0^2 + a_1 x_0 + a_0 \\ y_1 = a_n x_1^n + a_{n-1}x_1^{n-1} + ... + a_2 x_1^2 + a_1 x_1 + a_0 \\ \text{..} \\ y_n = a_n x_n^n + a_{n-1}x_n^{n-1} + ... + a_2 x_n^2 + a_1 x_n + a_0 \end{cases}$$

Resuelto el cual podemos, con los coeficientes resultantes de la solución, a_0, a_1,.......a_n , construir la función de interpolación.

Método de Lagrange

El método consiste en encontrar una función polinómica de grado n, $y = P_o(x) + P_1(x) + ... + P_n(x)$, conocidos n+1 puntos (x_0,y_0).......(x_n,y_n).

-Dicha función de interpolación es, en realidad, la suma de n+1 polinomios de grado n, cada uno de los cuales debe tomar el valor y_i para $x=x_i$ y anularse en los demás valores conocidos de x.

-Estos n+1 polinomios son:

$$P_0(x) = \frac{(x-x_1)(x-x_2)(x-x_3)..........(x-x_n)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)..........(x_0-x_n)}.y_0$$

que efectivamente toma el valor y_0 al sustituir x por x_0 y se anula al sustituir x por cualquiera de los otros valores conocidos, al haber en cada caso, una de las diferencias del numerador que se hace cero.

Análogamente, tendremos:

$$P_0(x) = \frac{(x-x_0)(x-x_1)(x-x_2)..........(x-x_{n-1})}{(x_n-x_0)(x_n-x_1)(x_0-x_2)..........(x_0-x_{n-1})}.y_n$$

-La función $y = P_o(x) + P_1(x) + ... + P_n(x)$ se denomina función de interpolación de Lagrange para n+1 puntos conocidos.

Extrapolación

Es el proceso que consiste en determinar, mediante la misma función de interpolación, el valor de la función para valores de x que se encuentran fuera del intervalo conocido.
Matemáticamente, si (x_0,y_0) (x_1,y_1).......(x_n,y_n) son n+1 puntos, tales que verifican $x_0<x_1$.............$<x_n$ y su función de interpolación es y=f(x), se denomina extrapolar a obtener, para un valor x=a tal que $a<x_0$ o $a>x_n$, el correspondiente valor de y.

Estimación

1. Definición Estimar un número es sustituir dicho número por otro próximo o aproximado al mismo.
2. Definición Se dice que "a" es una estimación de un número real x, con un error menor que r, si el número x pertenece al intervalo (a-r,a+r).
3. Definición El número r, se denomina Cota de Error de la estimación.

Interpolación. Problemas

1. La madre de un bebé realiza las siguientes anotaciones en su tabla de ganancia de peso:

Nacimiento	mes 1	mes2	mes3	mes4
3,5Kg.	4Kg		5,1Kg	5,4Kg

   Calcula su peso a los dos meses, cuando por un catarro no pudo llevar a cabo la anotación.

2. Sabiendo que $\sqrt{4} = 2 \; y \; \sqrt{9} = 3$, calcula la raíz cuadrada de 6 mediante interpolación lineal.

3. Obtén la función de interpolación cuadrática que pasa por los puntos siguientes:
   (0,-1), (1,-3), (3,-13)

4. Obtén mediante interpolación lineal y dados los puntos (-2,3), (3,13), la ordenada para el punto x=1.

5. Obtén mediante la aplicación del método de interpolación del polinomio de Lagrange, la función de interpolación que pasa por los puntos (-1,0), (1,4), (2,5), (3,7). Determina el valor de dicha función para x=0.

6. El número de chocolatinas de la marca Acme, vendidas por la máquina expendedora de un centro escolar, ha variado según su precio de la forma que se muestra en la siguiente tabla:

Precio	155	160	165	170
Unidades	200	230	215	185

   Halla la función de interpolación que ajuste los datos y calcula las ventas si como se pretende, el precio sube a 200 ptas.

7. Dada la siguiente sucesión, obtén la función de interpolación que representa el término general S_n, sabiendo que se trata de una función lineal:
   3, 5, 7, 11. El primer término corresponde a la abscisa x=0.

8. De una tabla de logaritmos se han tomado los siguientes datos:
   $\log 312=2{,}494$
   $\log 314=2{,}496$
   Calcula por interpolación lineal $\log 313$.

9. Sabiendo que sen 47º=0,7313 y sen 49º=0,7547, calcula por interpolación lineal sen 48º.

10. Calcula el valor esperado para las abscisas $x=1$ y $x=2$ del polinomio interpolador que tiene por ecuación $y=x^2+x+1$.

JOSEPH LOUIS LAGRANGE

Procedía de una ilustre familia parisiense que tenía profundo arraigo en Cerdeña, y algún rastro de noble linaje italiano. Pasó sus primeros años en Turín, su activa madurez en Berlín, y sus últimos años en París, donde logró su mayor fama. Una especulación insensata, llevada a cabo por su padre, abandonó a Lagrange a sus propios recursos a una edad temprana, pero este cambio de fortuna no resultó ser una gran calamidad, "pues de otro modo -dijo él- tal vez nunca hubiera descubierto mi vocación".

En la escuela, sus intereses infantiles eran Homero y Virgilio y cuando una memoria de Halley le cayó en las manos, se alumbró la chispa matemática. Como Newton, pero a una edad aún más temprana, llegó al corazón de la materia en un espacio de tiempo increíblemente corto. A los dieciséis años de edad fue nombrado profesor de matemáticas en la Escuela Real de Artillería de Turín, donde el tímido muchacho, que no poseía recursos de oratoria y era de muy pocas palabras, mantenía la atención de hombres bastante mayores que él. Su encantadora personalidad atraía su amistad y entusiasmo. Pronto condujo un joven grupo de científicos que fueron los primeros miembros de la Academia de Turín. Lagrange se transfiguraba cuando tenía una pluma en sus manos y desde un principio, sus escritos fueron la elegancia misma.

A los diecinueve años de edad, obtuvo fama resolviendo el llamado problema isoperimétrico, que había desconcertado al mundo matemático durante medio siglo. Comunicó su demostración en una carta a Euler, el cual se interesó enormemente por la solución, de modo especial en cuanto concordaba con un resultado que él mismo había hallado. Euler, con admirable tacto y amabilidad, respondió a Lagrange ocultando deliberadamente su propia obra, de manera que todo el honor recayera sobre su joven amigo. En realidad Lagrange no sólo había resuelto un problema, también había inventado un nuevo método, un nuevo cálculo de variaciones, que sería el tema central de la obra de su vida. Este cálculo pertenece a la historia del mínimo esfuerzo, que comenzó en los espejos reflectores de Herón y continuó cuando Descartes reflexionó sobre la curiosa forma de sus lentes ovales. Lagrange podía demostrar que los postulados newtonianos de materia y movimiento, un tanto modificados, se adaptaban al amplio principio de economía de la naturaleza. El principio ha conducido a los resultados aún más fructíferos de Hamilton y Maxwell, así como a resultados de interés en Teoría de la Relatividad y Mecánica Cuántica.

Lagrange estaba dispuesto a apreciar el trabajo sutil de los demás, pero estaba igualmente capacitado para descubrir un error. En una temprana memoria sobre las matemáticas del sonido, señaló defectos incluso en la obra de Newton. Otros matemáticos le reconocían, sin envidia, primero como su compañero y más tarde, como el mayor matemático viviente.

Le gustaba la música, decía que le aislaba y le ayudaba a pensar, que interrumpía la conversación general: “La escucho durante los tres primeros compases, luego no distingo nada, pero me entrego a mis pensamientos. De esta manera he resuelto muchos problemas difíciles“.

DISTRIBUCIONES UNIDIMENSIONALES

Para qué?

Para aprender a organizar los datos en tablas, facilitar la elaboración y manipulación de los mismos, para interpretar los datos en función de sus valores medios y sus variaciones respecto a dichos valores. Para situarnos en el camino de la toma de decisiones sobre una colección de datos.

Distribuciones Unidimensionales

Tratamiento de la información. Conceptos fundamentales

1. *Población:* Conjunto sobre el que realizamos el estudio
2. *Muestra:* Subconjunto de la población, seleccionado por características o criterios que lo hacen representativo de la misma.
3. *Individuo:* Elemento de la población
4. *Carácter:* Variable que estudiamos, puede ser cualitativo o cuantitativo.
5. *Modalidad:* Diferentes valores que puede tomar la variable
6. *Variable estadística:* Aplicación inyectiva de las modalidades de un carácter en $\mathbb{R}$; puede ser:
   - Discreta: Toma un número finito de valores
   - Continua: Toma un número infinito de valores

Estudio estadístico. Pasos

a) Selección de la población y el carácter a estudiar
b) Selección de la muestra, si la población lo permite
c) Recogida de datos
d) Elaboración de tablas
e) Realización de gráficos
f) Cálculo de parámetros

Tablas

Son una forma de organización de la información recogida. Si consideramos que la variable estudiada tiene k modalidades posibles, asociamos a cada una de ellas la frecuencia absoluta o f_i, que representa el número de individuos que presenta dicha modalidad y la frecuencia relativa $(f_r)_i = \frac{f_i}{N}$, donde N es el número total de individuos en la población.
La frecuencia acumulada, que se hace corresponder a un determinado valor de la variable es simplemente la suma de las frecuencias absolutas de dicho valor y todos los precedentes, estando la distribución ordenada de menor a mayor

- Si la población es contínua o discreta con gran número de modalidades, se agrupan éstas en las denominadas clases o intervalos de clase

- En este caso, se toma como valor representativo del intervalo, la denominada "marca de clase" que no es otra cosa que el valor central del intervalo.
- A dicho valor, se asocian tanto las frecuencias absolutas como las densidades de frecuencia, siendo éstas, las frecuencias absolutas divididas por la amplitud del intervalo.

Recapitulando, tendremos:

Intervalos	$[a_i, b_i]$
Marcas de clase	$m_i = \frac{a_i + b_i}{2}$
Frecuencias absolutas	f_i
Frecuencias relativas	$\frac{f_i}{N}$
Densidades de Frecuencia	$d_i = \frac{f_i}{b_i - a}$

Parámetros estadísticos de centralización

Se trata de parámetros que calculamos a la búsqueda de valores medios, centrales o representativos de la población. Son, a saber:

1. *Moda:* Es valor o modalidad de la variable estadística que se presenta con mayor frecuencia absoluta. Se representa por M_o.
2. *Media:* Con sus diferentes modalidades, que veremos más adelante, es el valor promedio de la población, se designa de forma genérica como $\bar{x}$ seguido de un subíndice que indica la variedad.
3. *Mediana:* Es el valor de la variable que divide a la población en dos subpoblaciones iguales, conteniendo cada una de ellas el 50% de los datos, se representa como M_e.
4. *Cuantiles:* Llamamos cuantil de orden i de una distribución al valor x_i, tal que mayora al i% de la población. Los de orden unidad se denominan percentiles, los de orden 10, Deciles, los de orden 25 se denominan Cuartiles.

Parámetros estadísticos de dispersión

Informan acerca de la distribución de los datos en torno a los parámetros de centralización, los más importantes son:

1. *Recorrido poblacional:* (de una distribución), es la diferencia entre el mayor y el menor valor de la variable.

2. *Recorrido intercuartílico:* Es la diferencia entre el cuantil del 25%, denominado cuartil 1 o simplemente C1 y el cuantil del 75%, denominado cuartil 3 o simplemente C3.
3. *Desviación media respecto a la media:* Se denota como $D_{\overline{X}}$ y es la media de las desviaciones respecto a la media. La calculamos mediante la expresión:

$$D_{\overline{x}} = \frac{1}{N}\sum_{i=1}^{k} |x_i - \overline{x}|.f_i = \sum_{i=1}^{k} |x_i - \overline{x}|.f_{r_i}$$

4. *Desviación media respecto a la mediana:* Se denota como D_{Me} ; es la media de las desviaciones respecto a la mediana. La calculamos mediante la expresión:

$$D_{Me} = \frac{1}{N}\sum_{i=1}^{k} |x_i - Me|.f_i = \sum_{i=1}^{k} |x_i - Me|.f_{r_i}$$

De forma análoga a los dos casos anteriores, puede calcularse la desviación con respecto a la moda.

5. *Varianza:* Dada una variable x con valores $x_1, x_2, \ldots\ldots\ldots x_n$ y frecuencias absolutas f_i, llamamos varianza de la variable x y la representamos por V(x) o $\sigma^2(x)$ a la media aritmética de los cuadrados de las desviaciones respecto a la media.

$$V(x) = \sigma^2(x) = \frac{|x_1 - \overline{x}|^2 f_i + \ldots\ldots + |x_k - \overline{x}|^2 f_k}{N} = \frac{1}{N}\sum_{i=1}^{k} |x_i - \overline{x}|^2 f_i$$

6. *Desviación Típica:* La denotamos por s o σ y es simplemente la raíz cuadrada de la varianza.

<u>Cálculo de los parámetros estadísticos de centralización</u>

<u>Cálculo de la media</u>

Distinguimos 3 tipos:

- Media aritmética: $\overline{x} = \frac{1}{N}\sum_{i=1}^{k} x_i f_i$
- Media ponderada: Dada la colección de pesos o ponderaciones $\{p_1, \ldots\ldots p_n\}$, la calculamos como:

$$\overline{x}_p = \frac{\sum_i x_i \, p_i}{\sum_i p_i}$$

- Media geométrica: la calculamos como $\overline{x}_g = \sqrt[N]{x^{f_1} + x_2^{f_2} + \ldots\ldots + x_n^{f_n}}$

Cálculo de la moda
Si los datos son simples, la moda será simplemente el dato que aparezca con mayor frecuencia absoluta.
Si por el contrario, los datos aparecen agrupados en clases o intervalos de igual amplitud:

a) Identificamos el intervalo que presenta mayor frecuencia y lo denominamos CLASE MODAL
b) Para calcular el dato concreto de mayor frecuencia, aplicamos la siguiente expresión:

$$M_o = L_i + c\frac{D_{i-1}}{D_{i-1} + D_{i+1}}$$

donde:

L_i = Límite inferior de la clase modal
c = Amplitud del intervalo de clase
D_{i-1} = Diferencia entre las frecuencias absolutas de la clase modal y la clase anterior.
D_{i+1} = Diferencia entre las frecuencias absolutas de la clase modal y la clase posterior.

Cálculo de la mediana
Distinguimos varios casos:

a) La variable es discreta, con datos simples en número impar.
   - Ordenamos los datos de menor a mayor
   - Tomamos el valor central como mediana
b) La variable es discreta, con datos simples en número par.
   - Ordenamos los datos de menor a mayor
   - Tomamos la semisuma de los datos centrales como mediana
c) La variable es discreta pero con datos numerosos.
   - Construimos la tabla de frecuencias acumuladas f_a
   - Tomamos como mediana el primer valor cuya frecuencia acumulada sea superior al 50% de los datos.
   - Si la mitad de los datos coincide exactamente con la frecuencia acumulada de algún valor, la mediana será la semisuma de dicho valor y el siguiente de la tabla.
d) La variable es continua o con datos agrupados en intervalos de igual amplitud.
   - Identificamos la clase mediana siguiendo el procedimiento detallado en c).
   - Determinamos el valor de la mediana mediante la siguiente expresión

$$M_e = L_i + c\frac{\left(N/2\right) - (f_a)_{i-1}}{f_i}$$

donde:

$(f_a)_{i-1}$ = frecuencia absoluta acumulada de la clase anterior a la mediana.
f_i = frecuencia absoluta de la clase mediana
c = amplitud de la clase mediana
L_i = Límite inferior de la clase mediana

Cálculo de los cuantiles

Para calcular los cuantiles, procedemos de forma idéntica al cálculo de la mediana, colocando el porcentaje correspondiente de datos a mayorar en el paréntesis donde en aquella colocábamos N/2, tomando las frecuencias en el intervalo en que se mayora el porcentaje correspondiente de los datos.

Teorema:
Si la distribución es simétrica o ligeramente asimétrica, se verifica que:
- En el intervalo $(\overline{x}-s, \overline{x}+s)$ se encuentra el 68,3% de los datos
- En el intervalo $(\overline{x}-2s, \overline{x}+2s)$ se encuentra el 95,4% de los datos
- En el intervalo $(\overline{x}-3s, \overline{x}+3s)$ se encuentra el 99,7% de los datos

Propiedades especiales de la Media o Esperanza Matemática

a) Si c es una constante $\overline{x(c)} = E(c) = c$
b) La esperanza matemática de la suma de variables aleatorias es la suma de esperanzas matemáticas.
c) La esperanza matemática del producto de una constante por una variable aleatoria, es el producto de la constante por la esperanza matemática de la variable
d) La esperanza matemática del producto de variables aleatorias independientes, es el producto de las esperanzas matemáticas.

Propiedades especiales de la Varianza

a) La varianza de una suma de variables aleatorias independientes es la suma de varianzas
b) La varianza de una constante por una variable aleatoria, es igual al cuadrado de la constante por la varianza de la variable.
c) La varianza de una constante es cero.

Distribuciones Unidimensionales. Problemas.

1. Dada la siguiente distribución:

x_i	2	3	8	12	17
n_i	2	2	3	3	1

   Calcular:
   a) Media aritmética
   b) Media Geométrica
   Comprobar la relación existente entre ellas

2. Calcular la media aritmética de los siguientes valores, su moda, mediana, desviación media respecto a la mediana, desviación típica y varianza, agrupándolos primero en intervalos de longitud 5 y posteriormente en intervalos de longitud 10.

49	48	43	42	49	41	42	43	44	44	51	43
53	54	51	59	58	57	56	54	51	54	53	64
62	64	63	62	61	62	68	58	67	66	69	

3. En una campaña de inspección medioambiental se ha estudiado el número de deficiencias de una serie de industrias, obteniéndose la siguiente distribución:

Defs.	1	2	3	4	5	6	7	8	9	10	11	12
Frec.	20	20	22	15	14	10	12	6	5	2	3	2

   a) Determinar todas las medidas de centralización excepto cuantiles.
   b) Determinar todas las medidas de dispersión
   c) Representar gráficamente Número deficiencias/ Frecuencias
   d) Extráiganse las conclusiones más relevantes.

4. Las calificaciones de la asignatura Inglés, de un curso pasado, se distribuyen de acuerdo con la siguiente tabla, para alumnos que se presentaron a examen en Junio.

Calificación	Valor	Alumnos (nº)
Insuficiente	0	110
Suficiente	1	90
Bien	2	23
Notable	3	12
Sobresaliente	4	2

   Calcúlense media y mediana, varianza, moda, y cuartíl 3

5. En la tabla siguiente, se dan las medidas originales de Cavendish de 1798 sobre la densidad de la Tierra (relativa al agua), utilizando una balanza de torsión.

5.50	5.61	5.88	5.07	5.26	5.55	5.36	5.29	5.58	5.65
5.57	5.53	5.62	5.29	5.44	5.34	5.79	5.10	5.27	5.39
5.42	5.47	5.63	5.34	5.46	5.30	5.75	5.68	5.85	

a) Calcular media, mediana, moda y desviación típica.
b) Cavendish se equivocó al escribir el tercer valor de la tabla, tomando 5.88 en lugar del valor medido 4.88 ¿Cómo cambian la media y la mediana cuando se corrige dicho valor?
c) Después de la sexta medida, Cavendish cambió el cable de la balanza por uno más rígido y consideró que ello afectaba las medidas; calcular por separado media, mediana y desviación típica para las seis primeras medidas (después de cambiar 5.88 por 4.88) y para las 23 restantes. Comentar el resultado.

NOTA: El verdadero valor de la densidad de la Tierra es 5.51

6. Se ha realizado un estudio de salinidad de ciertas disoluciones, midiéndose en ellas la concentración de ClNa en mg/l y obteniéndose los siguientes resultados:

71	73	162	140	125	96	82	200	23	104
72	72	168	21	62	78	101	108	139	120
142	83	18	169	42	92	119	80	149	150
44	167	95	108	178	83	47	71	81	118

Agrupar en intervalos y calcular media, mediana, moda; representar gráficamente la distribución con un gráfico a elegir.

7. En 1789 Michelson obtuvo las medidas de la velocidad de la luz que se dan en la tabla. Los valores tabulados +299.000 dan la velocidad de la luz en km/s.

850	740	900	1070	930	850	950	980	980	880
1000	980	930	650	760	810	1000	1000	960	960
960	940	960	940	880	800	850	880	900	840
830	790	810	880	880	830	800	790	760	800
880	880	880	860	720	720	620	860	970	950
880	910	850	870	840	840	850	840	840	840
890	810	810	820	800	770	760	740	750	760
910	920	890	860	880	720	840	850	850	780
890	840	780	810	760	810	790	810	820	850
870	870	810	740	810	940	950	800	810	870

Agrupar en intervalos de 50 km/s y determinar media y desviación típica

8. En la Asignatura de Métodos Estadísticos, las calificaciones de 30 alumnos de una clase de 2º de Bachillerato, fueron las siguientes:
   6; 2,7; 8,4; 5; 5; 10; 3,7; 3,6; 2; 9; 6,7; 8; 3,4; 7,7; 1,6; 10; 6; 4,6; 0,8; 1,4; 2,5; 6; 6,2; 3,7; 7,4; 2,7; 1,2; 6,3; 10; 5,2.
   a) Determinar media, mediana, moda, varianza y desviación típica.
   b) Realizar agrupación en intervalos de longitud 2 y calcular de nuevo todas las medidas anteriores.

9. Dada la siguiente distribución de frecuencias, calcúlese la media aritmética, la media geométrica, la mediana, la moda y determínese empíricamente una relación de orden entre ellas.

x_i	1	3	6	7	8	9
f_i	2	4	3	2	3	1

10. En una encuesta entre 37 estudiantes, acerca del número de horas mensuales de estudio, se han obtenido los siguientes datos:

Horas	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40
f_i	1	1	2	7	11	9	3	2

   Determínense C_1, C_2, C_3

11. En idénticas condiciones a las del problema anterior, determínense D_3 y D_7.

12. Usando los mismos datos del problema 3, determínense P_{65} y P_{40}.

13. Usando los datos de la encuesta de estudiantes, determínense el recorrido y el recorrido intercuartílico de la población.

14. Calcúlense varianza y desviación típica de la distribución de las horas de estudio en la población de estudiantes citada en el problema 3.

15. Calcúlense moda y mediana de la distribución del problema 3 y compárense, comentando el resultado y su relación con la media.

16. Un alumno tiene cuatro calificaciones en una asignatura, que son respectivamente: 0, 10, 2, 8; otro alumno en la misma asignatura ha obtenido: 5, 5, 4,5, 5,5. Obviamente, uno es muy regular y otro muy irregular. Justifíquese dicha apreciación mediante los parámetros estadísticos correspondientes.

BLAISE PASCAL (1623-1662)

Filósofo y científico francés. A los 11 años participaba en reuniones científicas de Mersenne, origen de la futura academia de las ciencias, descubrió y demostró por su propia cuenta la proposición del libro I de los elementos de Euclídes.

Sus trabajos relacionados con problemas muy diversos de las matemáticas y de la física, incidieron decisivamente en el posterior desarrollo de la ciencia de su tiempo. En 1640 publicó "Essai pour les coniques", donde además de recoger diversos teoremas de Descartes y los griegos, demuestra el llamado teorema del hexágono de Pascal. De este teorema se derivaron cerca de 400 proposiciones acerca de las cónicas. En 1649 verificó las conclusiones de Torricelli sobre la presión atmosférica con la celebre experiencia del Puy de Dome. Su más completa contribución a la hidrostática esta contenida en el Traité de l'equilibre des liqueurs (1653) Estudio que comprende un calculo de la presión atmosférica y diversas observaciones relativas a la estática de los fluidos, entre las cuales el famoso Principio de Pascal sobre la transmisión de las presiones en los líquidos.

Se puede considerar a Pascal como fundador del calculo de probabilidades, formulado en 1654 como "geometría del azar " en su correspondencia con Fermat.

El Traité du triangle arithmétique, que prolonga estas ideas, muestra la importancia del triángulo en cuestiones tales como el cálculo combinatorio y el cálculo de la potencia de un binomio.

A partir de 1654 interrumpió casi por completo su labor científica aunque aportó en 1658 otro trabajo fundamental: "Les lettres de A. Dettonville" conteniendo algunos de sus descubrimientos en geometría, donde resuelve varios problemas relacionados con la cicloide que él mismo había propuesto a los más importantes científicos de su tiempo y que ninguno de ellos había logrado desentrañar.

Su pensamiento se encuentra determinado por su condición de científico que desconfía de la razón para abarcar los problemas últimos de la vida y por su profunda religiosidad en la que encuentra la salvación para no caer en la filosofía de lo absurdo. Todo ello le conduce a admitir 2 principios de conocimiento: L'esprite géometrique (razón), orientada a las razones científicas y L'esprite de finnesse (Corazón), en el que se dan en forma de intuiciones los principios básicos para la comprensión de la vida e incluso aquellos principios fundamentales de que arranca toda ciencia.

Según Pascal, tanto la razón como el corazón son dos formas igualmente válidas de conocer, y tal vez el segundo es superior a la abstracción racional, como lo expuso al decir: "Conocemos la verdad no sólo con la razón, sino también con el corazón " y " el corazón tiene sus razones que la razón no conoce ".

VARIABLE ESTADÍSTICA BIDIMENSIONAL

¿Para qué?

Para trabajar dos variables simultáneamente o dos caracteres de una misma variable, establecer valores medios y variaciones respecto a dichos valores medios. Para establecer relaciones entre dichas variables o dichos caracteres y adentrarnos en el interesante camino de la búsqueda de la causalidad, la influencia de unos factores en otros y seguir avanzando por el camino de la toma de decisiones acerca de una colección de datos.

Variable estadística bidimensional

Si estudiamos dos caracteres de una misma variable, estaremos estudiando una variable estadística bidimensional.

- Si A y B son los dos caracteres en estudio, a partir de A queda definida una variable unidimensional x, de valores x_1,.......x_n y otra y, a partir de B, de valores y_1,......,y_m.
- Partiendo de tal situación, se define la variable bidimensional (x,y) como aquella que toma los valores (x_i,y_i) con $1 \leq i \leq n$ y $1 \leq j \leq m$.
- De esta forma, cada individuo de la población, está definido por un par de valores de la variable (x_i,y_i), lo que permite la elaboración de una tabla de frecuencias de doble entrada:

X\Y	y_1	y_2	...	...	...	y_m	
x_1	f_{11}	f_{12}				f_{1m}	A_1
x_2	f_{21}	f_{22}				f_{2m}	A_2
...	...	...					...
...	...	...					...
...	...	...					...
x_n	f_{n1}	f_{2n}				f_{nm}	A_n
	B_1	B_2	...	...	...	B_m	N

- En dicha tabla, definimos:

  f_{ij} = Frecuencia absoluta conjunta

  $A_i = \sum_j f_{ij}$ Frecuencia absoluta marginal de la variable X

  $B_j = \sum_i f_{ij}$ Frecuencia absoluta marginal de la variable Y

  $N = \sum_i^n \sum_j^m f_{ij} = \sum_{i=1}^n A_i = \sum_{j=1}^m B_j = N$ Número de individuos de la población

  $f_{r_{ij}}$ Frecuencia relativa del par (x_i,y_j), verificando $\sum_{i=1}^n \sum_{j=1}^m (f_r)_{ij} = 1$

En la tabla, si consideramos 1ª y última columnas, obtenemos la distribución unidimensional de la variable x. Así mismo, si consideramos 1ª y última filas, obtenemos la distribución unidimensional de la variable y.
Estas distribuciones son denominadas Distribuciones marginales.
Si consideramos la primera y una columna intermedia, obtenemos la distribución de la variable x, condicionada al valor o modalidad y_i de la variable y.

<u>Parámetros estadísticos</u>

Podemos trabajar las distribuciones bidimensionales bien en base a las distribuciones marginales, bien en base a las frecuencias conjuntas. En el primer caso, usamos las expresiones ya conocidas, ligeramente modificadas para adaptarlas a la nomenclatura de la tabla bidimensional:

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{n} x_i A_i ;; \bar{y} = \frac{1}{N}\sum_{j=1}^{m} y_j B_j ;; N = \sum_{i=1}^{n} A_i = \sum_{j=1}^{m} B_j$$

Si lo que usamos son las frecuencias relativas conjuntas, debemos usar las expresiones

$$\bar{x} = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i (f_{r_{ij}});; \bar{y} = \sum_{i=1}^{n}\sum_{j=1}^{m} y_j (f_{r_{ij}});. \quad \textit{Ya que} \quad A_i = \sum_{j=1}^{m} f_{ij} ;; B_j = \sum_{i=1}^{n} f_{ij} .$$

El punto $\left(\bar{x}, \bar{y}\right)$ recibe el nombre de <u>centro de gravedad de la distribución</u>.

Conocidas las medias, pueden calcularse las varianzas de las variables x e y, que serán, respectivamente:

$$V(x) = \sigma^2(x) = \frac{1}{N}\sum_{i=1}^{n}\sum_{j=1}^{m} (x_i - \bar{x})^2 \cdot A_i \qquad ; V(y) = \sigma^2(y) = \frac{1}{N}\sum_{i=1}^{n}\sum_{j=1}^{m} (y_j - \bar{y})^2 \cdot B_j$$

Si se utilizan frecuencias relativas conjuntas, las fórmulas a utilizar serán:

$$\sigma^2(x) = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i^2 (f_{r_{ij}}) - \bar{x}^2 \qquad \sigma^2(y) = \sum_{i=1}^{n}\sum_{j=1}^{m} y_j^2 (f_{r_{ij}}) - \bar{y}^2$$

En ambos casos, las desviaciones típicas serán:

$$\sigma(x) = \sqrt{\sigma^2(x)} \qquad \sigma(y) = \sqrt{\sigma^2(y)}$$

La COVARIANZA es un parámetro específico de variables bidimensionales, también denominado varianza conjunta. Es la media aritmética de los productos de las desviaciones típicas de cada variable respecto a su media.

$$S_{xy} = \frac{1}{N}\sum_{i=1}^{n}\sum_{j=1}^{m} (x_i - \bar{x})(y_i - \bar{y}) f_{ij}$$

$$S_{xy} = \frac{1}{N}\sum_{i=1}^{n}\sum_{j=1}^{m} x_i y_j f_{ij} - \overline{x}\,\overline{y}$$

Correlación

Es la técnica que intenta buscar una dependencia o relación entre las dos variables de una distribución bidimensional.

- Pueden darse los siguientes casos:
  - Correlación de tipo funcional
  - Correlación de tipo lineal o curvilínea
  - Correlación positiva cuando $Sig(\Delta x)=Sig(\Delta y)$
  - Correlación negativa cuando $Sig(\Delta x)\neq Sig(\Delta y)$
  - Correlación NULA cuando no existe relación entre las variables.

Correlación lineal

Una vez decidido que lo que se desea buscar entre las dos variables es una relación lineal, es preciso determinar formal y funcionalmente dicha relación. Para ello se utiliza el coeficiente de correlación lineal, denotado como “r”.

$$r = \frac{S_{xy}}{S_x S_y}$$

Descriptivamente, se analiza como sigue:

- $-1<r<1$. si $r>1$ o $r<-1$ el cálculo es incorrecto
- Si $r=1$ o $r=-1$ existe una dependencia funcional lineal entre x e y
- Si $r<0$ existe correlación negativa entre x e y
- Si $r>0$ existe correlación positiva entre x e y
  - En estos dos últimos casos, la correlación será tanto más fuerte cuanto más próximo a ± 1 sea el valor de r
- Si $r=0$, no existe correlación entre x e y

Regresión. Regresión lineal

Se trata de un método de ajuste matemático que permite encontrar la forma funcional que relaciona a las dos variables de una distribución bidimensional; de forma que, conocido el valor de una de ellas, pueda determinarse el valor de la otra.
Aunque dicha forma funcional puede tomar prácticamente cualquier forma, nosotros estudiamos únicamente la forma lineal y la determinamos mediante el denominado “Método de mínimos cuadrados”.
Analíticamente, se aplica como sigue:

1. Se determina el coeficiente de correlación lineal y se adopta la siguiente regla de decisión:
   a. $r\approx 0$ No tiene sentido intentar la correlación
   b. $r\approx \pm 1$ Nuestras estimaciones serán muy aproximadas a los datos reales
   c. $r=\pm 1$ Los valores estimados serán exactamente iguales a los reales

2. Determinamos el coeficiente de regresión de “y sobre x” o “x sobre y” según corresponda.

$\frac{S_{xy}}{S_x^2} = r_x$ Coeficiente de regresión de y sobre x, que usaremos para calcular una recta de regresión en la que x sea variable independiente e y variable dependiente

$\frac{S_{xy}}{S_x^2} = r_x$ Coeficiente de regresión de x sobre y, que usaremos para calcular una recta de regresión en la que y sea variable independiente y x variable dependiente.

3. Calculamos la recta de regresión

$$y - \overline{y} = \frac{S_{xy}}{S_x^2}(x - \overline{x}) \quad \text{Re}\,cta \quad y \; sobre \; x$$

$$x - \overline{x} = \frac{S_{xy}}{S_y^2}(y - \overline{y}) \quad \text{Re}\,cta \quad x \; sobre \; y$$

<u>Regla mnemotécnica</u>: Ante la duda en la elección de la recta de regresión, aplicar: “El que va “sobre” es dependiente”.

Variable estadística bidimensional. Problemas

1. Tomamos un grupo de observación de 10 muestras de agua en las que medimos la concentración en nitratos en mg/l y temperatura en ºC. Tenemos los siguientes datos:

mg/l	70	65	80	60	75	85	70	65	80	85
ºC	36.5	36.5	37.0	36.0	37.0	37.5	37.0	36.0	37.5	37.0

   Hallar las ecuaciones lineales que representan este fenómeno y realizar su representación gráfica. ¿Qué temperatura tendría una muestra con una concentración en nitratos de 72 mg/l

2. En un estudio dietético se pretende estudiar si existe o no relación entre la cantidad total, inyectada durante un mes, de determinado preparado farmacéutico y el aumento de peso provocado en esa persona. El estudio está realizado sobre personas de edad y altura similares. En una muestra de 10 se obtienen los siguientes resultados:

Δp	1	2	2	3	2	4	3	5	6	2
cc prod.	15	20	25	25	19	30	30	35	40	18

   A partir de estos datos, ¿puede inferirse una relación entre ambas variables?. Caso de respuesta afirmativa, ¿puede expresarse esa relación de forma matemática?

3. Se ha ajustado una recta de regresión de la variable y sobre x, resultando:
$$y+5=0,6x$$
   Teniendo en cuenta que para un grupo de 20 observaciones, las variables x e y presentan respectivamente unas medias aritméticas de 50 y 25, y varianzas 0,64 y 0,36, calcúlese el coeficiente de correlación.

4. En 1929, Hubble presentó las primeras medidas sobre distancias y velocidades radiales de las galaxias, demostrando que había una correlación positiva entre las magnitudes y que por consiguiente, el Universo se encontraba en expansión. La siguiente lista presenta parte de los datos originales de Hubble:

d(Mpc)	0.5	0.5	0.8	0.9	0.9	1.0	1.1	1.1	1.4	1.7	2.0	2.0
v(km/s)	290	270	300	650	150	920	450	500	500	960	800	1090

a) Calcular media y varianza de las variables distancia y velocidad. Calcular asimismo la covarianza.
b) Usando los resultados del apartado anterior, calcular y representar la recta de regresión de v sobre d. El coeficiente de correlación obtenido se denomina constante de Hubble.

5. Una clínica de adelgazamiento quiere conocer la eficacia de uno de sus tratamientos. La siguiente tabla da datos de 16 personas sometidas al mismo, siendo x el peso inicial en Kg e y el peso perdido en un año en la misma unidad:

x	225	235	173	223	200	199	129
y	15	44	31	39	6	16	21
x	140	156	146	195	155	185	150
y	5	12	-3	19	10	24	-3

a) ¿Podemos afirmar que el tratamiento en un año ha sido efectivo?
b) Represéntese el diagrama de dispersión de las dos variables
c) Calcular y representar la recta de regresión
d) ¿Existe relación entre el peso inicial y la pérdida de peso al final del año?

6. La siguiente tabla da la temperatura media y la precipitación en una ciudad durante los meses de Julio de los años 1975-1984. Hallar el coeficiente de correlación entre ambas variables y , si mereciese la pena, calcúlese y represéntese la recta de regresión.

Año	T(ºC)	P(l)
1975	25.6	158.2
1976	22.1	92.5
1977	24.2	86.9
1978	22.6	72.1
1979	24.1	46.5
1980	23.1	71.6
1981	23.9	102.6
1982	24.1	65.0
1983	23.2	30.0
1984	21.3	106.4

7. Se quiere determinar, si existe, la relación entre la demanda química de oxígeno de cierta agua y la tasa de procesos diarreicos de la población que la consume. Represéntese la nube de puntos, calcúlese la covarianza, el coeficiente de

correlación y determínese si existe dependencia estadística entre ambas variables, calculando la recta de regresión en dicho caso.

DQO	Casos/10^5 hab
253	1095
303	3161
309	1030
348	1100
243	904
257	1154
286	1313
262	2826
337	1825
231	2378
275	1102

THOMAS BAYES

Se sabe que Thomas Bayes nació en Londres, Inglaterra, en 1702, pero no se ha encontrado registro de la fecha exacta de su nacimiento. Su padre fue uno de los primeros seis ministros presbiterianos que fueron ordenados en Inglaterra. La educación de Thomas fue privada, un hecho que se antoja necesario para el hijo de un ministro presbiteriano de aquellos tiempos. Parece ser que de Moivre fue su maestro particular, pues se sabe que por ese entonces ejercía como profesor en Londres.

Bayes fue ordenado ministro presbiteriano y asistió a su padre en Holborn. Al final de la década iniciada en 1720 fue nombrado pastor en Turnbridge Wells (Kent, Inglaterra).

Aunque trató de retirarse de su puesto eclesiástico en 1749, permaneció en él hasta 1752; una vez retirado siguió viviendo en Turnbridge Wells hasta el día de su muerte, el 17 de abril de 1761. Sus restos descansan en el cementerio londinense de Bunhill Fields.

Teólogo, matemático y miembro de la Royal Society desde 1742, Bayes fue el primero en utilizar la probabilidad inductivamente y establecer una base matemática para la inferencia probabilística (la manera de calcular, a partir de la frecuencia con la que un acontecimiento ocurrió, la probabilidad de que ocurrirá en el futuro).

Los únicos trabajos que se sabe que Thomas Bayes publicó en vida son: Divine Providence and Government Is the Happiness of His Creatures (1731) y An Introduction to the Doctrine of Fluxions, and a Defence of The Analyst (1736), que fueron blanco de críticas por parte del obispo Berkeley, quien sustentaba sus ideas en los fundamentos lógicos del cálculo de Newton.

En 1763 se publicó póstumamente Essay Towards Solving a Problem in the Doctrine of Chances, donde el reverendo Bayes abordó el problema de las causas a través de los efectos observados, y donde se enuncia el teorema que lleva su nombre.

Los métodos estadísticos que suponen el parámetro de prueba como una variable aleatoria postulando además una función de densidad para dicho parámetro se conocen como métodos de Bayes.

Las técnicas de Bayes permiten abordar en forma diferente el área de "toma de decisiones", formulándola en términos de pérdidas o ganancias económicas y no en términos de la probabilidad de tomar la decisión correcta.

COMBINATORIA

¿Para qué?

Para saber cuántas combinaciones de Lotería Primitiva o Quiniela tenemos que hacer para garantizarnos un premio, para saber de cuantas formas diferentes puede darse un fenómeno o cuantas posibilidades de elección diferentes tenemos, cuando nos enfrentamos a un problema con múltiples situaciones. Para abrir la puerta, de forma confiada y segura, al campo de la Probabilidad

Combinatoria

Variaciones Ordinarias

Definición: Se denominan variaciones ordinarias de m elementos, tomados de n en n, con $n \leq m$, a los diferentes grupos formados por n elementos, de forma que:

- Los n elementos son distintos (No hay repeticiones)
- Dos grupos son distintos si se diferencian en algún elemento (al menos 1) o en el orden en que están colocados.

Se designan como $V_{m,n}$ y se calculan como:

$$V_{m,n}=m\cdot(m-1)\cdot(m-2)\cdot\ldots\ldots\cdot(m-n+1)$$

Variaciones con repetición

Definición: Se denominan variaciones con repetición de m elementos tomados de n en n, con $n \leq m$, a los diferentes grupos formados por n elementos, de forma que :

- Los elementos que forman cada grupo pueden estar repetidos
- Dos grupos son distintos si se diferencian en algún elemento o en el orden en que están colocados.

Se designan como $VR_{m,n}$ y se calculan como:

$$VR_{m,n}=m^n$$

Permutaciones Ordinarias

Definición: Se denominan permutaciones de m elementos a las agrupaciones de dichos elementos de forma que :

- En cada agrupación aparecen los m elementos sin repetirse ninguno
- Dos grupos son diferentes si el orden de colocación de alguno de los elementos es diferente.

Se designan por P_m y se calculan como

$$P_m=m!=m\cdot(m-1)\cdot(m-2)\cdot\ldots\ldots 2\cdot 1$$

Ello viene derivado del hecho de que las permutaciones ordinarias son en realidad variaciones ordinarias de m elementos, tomados de m en m.

Existe además, el caso de las permutaciones circulares, en el que la ordinalidad, (primero...último) carece de sentido, teniéndose en consideración, tan sólo el orden de los elementos. El ejemplo tipo es el grupo de personas sentadas en torno a una mesa.
Se denotan como PC_m y su número es igual al de las permutaciones de (m-1) elementos.
Así pues:

$PC_m = P_{m-1} = (m-1)!$

Permutaciones con repetición

Definición: Se denominan permutaciones con repetición de m elementos, donde el primero se repite a_1 veces, el segundo a_2 veces y el último a_m veces, a los distintos grupos que pueden formarse con los m elementos, de forma que :

- Cada grupo se compone de a_1 veces el primer elemento, a_2 veces el segundo elemento .... y a_m veces el último.
- Dos grupos se diferencian en el orden de colocación de alguno de sus elementos
- Se denotan como $P_n^{a_1, a_2 \ldots a_m}$ y se calculan como

$$P_n^{a_1, a_2 \ldots a_m} = \frac{n!}{P_{a_1} \cdot P_{a_2} \cdot \ldots\ldots P_{a_m}} = \frac{n!}{a_1! \cdot a_2! \cdot \ldots\ldots a_m!}$$

Combinaciones

Definición: Se denominan combinaciones de m elementos tomados de n en n, con n≤m, a todas las agrupaciones posibles que pueden formarse con los m elementos, de forma que :

- Cada agrupación está formada por n elementos distintos entre si, es decir, no hay repeticiones.
- Dos agrupaciones son distintas, si se diferencian al menos en un elemento, sin tener en cuenta el orden de los mismos

Se denotan como $C_{m,n}$ y se calculan como:

$$C_{m,n} = \frac{V_{m,n}}{P_n} = \frac{m(m-1)(m-2)\ldots\ldots(m-n+1)}{n!} = \frac{m!}{n!(m-n)!}$$

Resolución de problemas combinatorios

Ver esquema al final del capítulo.

Números combinatorios

Se denomina número combinatorio de índice m y orden n, al número de combinaciones de m elementos, tomados de n en n, tales que m≥n≥0.

Se denota como $\binom{m}{n}$, se lee "m sobre n" y se calcula como:

$$\binom{m}{n} = \frac{m!}{n!(m-n)!}$$

Los números combinatorios, en lo referente a su cálculo, cumplen determinadas propiedades de interés, de las que señalaremos las siguientes:

1. $\binom{m}{n-1} + \binom{m}{n} = \binom{m+1}{n}$
2. $\binom{m}{n} = \binom{m}{m-n}$
3. $\binom{m}{0} + \binom{m}{1} + \binom{m}{2} + ... + \binom{m}{n} = 2^m$

<u>Triángulo de Pascal</u>

Es una herramienta de cálculo que sirve para obtener, de forma rápida y simple, el valor de los números combinatorios elementales.

En cuanto a construcción, sus propiedades fundamentales son:

- Todas las filas son simétricas
- Todas las filas empiezan y acaban por 1
- Salvo los extremos, todos los números se obtienen sumando los dos que tiene encima.
- La fila n-sima está formada por n+1 números.

$$\begin{array}{ccccccccc} & & & & 1 & & & & \\ & & & 1 & & 1 & & & \\ & & 1 & & 2 & & 1 & & \\ & 1 & & 3 & & 3 & & 1 & \\ 1 & & 4 & & 6 & & 4 & & 1 \end{array}$$

Su correspondencia con los números combinatorios, es la siguiente:

$$\binom{0}{0}$$

$$\binom{1}{0} \quad \binom{1}{1}$$

$$\binom{2}{0} \quad \binom{2}{1} \quad \binom{2}{2}$$

$$\binom{3}{0} \quad \binom{3}{1} \quad \binom{3}{2} \quad \binom{3}{3}$$

$$\binom{4}{0} \quad \binom{4}{1} \quad \binom{4}{2} \quad \binom{4}{3} \quad \binom{4}{4}$$

Binomio de Newton

Es una forma sistemática de desarrollar el cuadrado de un binomio, teniendo en cuenta que en el desarrollo matemático de los mismos, aparecen una serie de regularidades, tales como:

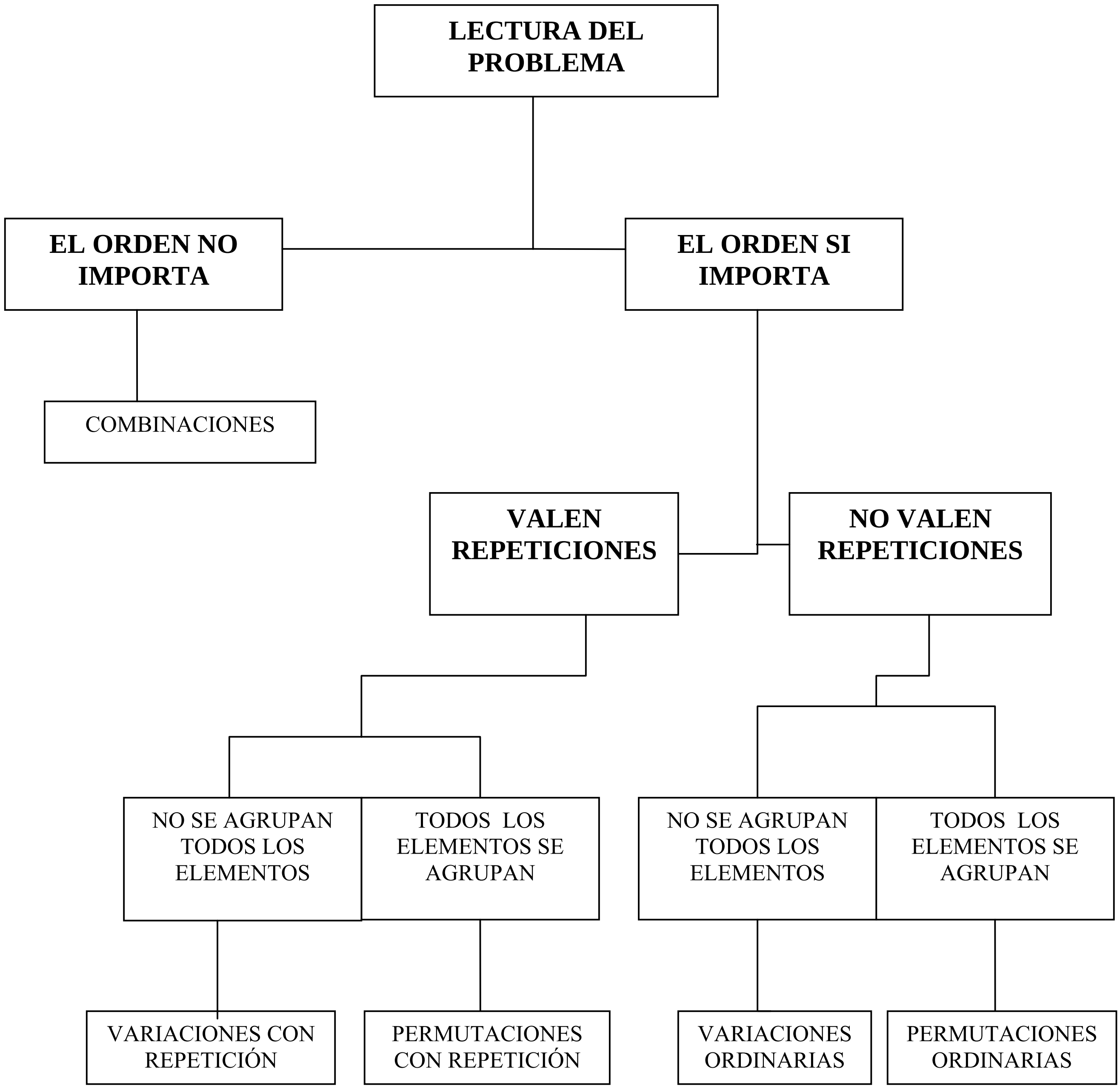
- En el desarrollo de $(a+b)^n$, aparecen n+1 términos
- Los coeficientes del desarrollo de $(a+b)^n$ coinciden con los números de la fila n-sima de del Triángulo de Pascal
- Los coeficientes de a, van disminuyendo de n a 0 de uno en uno.
- Los coeficientes de b van aumentando de 0 a n de uno en uno.
- En cada término del desarrollo, la suma de exponentes de a y b es igual a n.

Generalizando:

La expresión $(a+b)^n$ se conoce como Binomio de Newton y su desarrollo es el siguiente:

$$(a+b)^2 = \binom{n}{0} a^n b^0 + \binom{n}{1} a^{n-1} b^1 + \dots\dots\dots + \binom{n}{n-1} a^1 b^{n-1} + \binom{n}{n} a^0 b^n$$

Resolución de problemas combinatorios.



Combinatoria. Problemas

1. Teniendo en cuenta que hay veinte equipos jugando en primera división de la liga profesional de fútbol, ¿de cuántas formas diferentes se pueden organizar los tres primeros puestos al final de la misma?

2. Si una clase tiene 27 alumnos ¿de cuántas formas diferentes pueden organizarse los cargos representativos de la misma, delegado y subdelegado?

3. Resuelve $V_{x,2}=12$

4. ¿Cuántos números diferentes de cinco cifras pueden formarse con los números 1,2,3,4 ¿Cuántos de ellos serán mayores de 12.000 y menores de 13.000?

5. ¿Cuántas columnas debemos realizar para asegurarnos un pleno al 15 en una quiniela?

6. ¿De cuántas formas diferentes pueden ordenarse 10 cuentas en un collar abierto, extendido sobre una mesa? . Si colocamos el collar en torno al cuello de la chica y suponemos que el broche es tal que permite el paso de las cuentas de un lado al otro del mismo, ¿son las mismas formas diferentes, o son menos?

7. ¿Cuántos números de 7 cifras existen, que estén formados por dos doses, tres treses y dos cincos?

8. ¿Cuántas palabras diferentes, con sentido o sin él, pueden formarse con las letras de la palabra ALTRUISMO?

9. Una fábrica de caramelos, con cinco líneas de fabricación de caramelos de un solo sabor, quiere ampliar su producción a caramelos de dos y tres sabores. ¿cuántos caramelos diferentes podrá fabricar?

10. ¿Cuánto suman todos los elementos de la fila 10 del Triángulo de Pascal?

11. Usa la fórmula del binomio de Newton para desarrollar los siguientes:
    a) $(6a+4b)^7$
    b) $(3a+b/5)^4$
    c) $(2ab-b^2)^3$

12. ¿De cuántas formas distintas pueden sentarse 10 personas en la fila de un cine?

13. Sigue el esquema de resolución de problemas combinatorios y, sin resolverlos, identifica el tipo de los siguientes :
    a) Sentar a cuatro personas en torno a una mesa
    b) Número de combinaciones posibles de lotería primitiva
    c) Número de columnas para asegurar pleno de quiniela
    d) Formas de entregar 4 caramelos a seis niños.

14. ¿Cuántas quinielas distintas con 7 unos, 5 equis y tres doses pueden rellenarse?

15. Considera un señalero de la armada, simplifiquemos sus movimientos posibles a un plano y dentro de ese plano, a tres posiciones para cada una de sus manos. ¿cuántas señales diferentes puede hacer?

ISAAC NEWTON

Nació el día de Navidad del antiguo calendario en 1642 (correspondiente al 4 de Enero de 1643 del nuevo calendario), año en que moría Galileo, en el pueblecito de Woolsthorpe, unos 13 Km. al sur de Grantham, en Lincolnshire. Fue un niño prematuro y su padre murió antes de su nacimiento, a los treinta y siete años. Isaac fue educado por su abuela, preocupada por la delicada salud de su nieto. Su madre, mujer ahorrativa y diligente, se casó de nuevo cuando su hijo no tenía más que tres años. Newton frecuentó la escuela del lugar y, siendo muy niño, manifestó un comportamiento completamente normal, con un interés marcado por los juguetes mecánicos.

El reverendo William Ayscough, tío de Newton y diplomado por el Trinity College de Cambridge, convenció a su madre de que lo enviara a Cambridge en lugar de dejarlo en la granja familiar para ayudarla. En junio de 1661, a los dieciocho años, era pues alumno del Trinity College, y nada en sus estudios anteriores permitía entrever o incluso esperar la deslumbrante carrera científica del fundador de la mecánica y la óptica. Por otra parte, el Trinity College tenía fama de ser una institución sumamente recomendable para aquellos que se destinaban a las órdenes. Afortunadamente, esta institución le brindó hospitalidad, libertad y una atmósfera amistosa que le permitieron tomar contacto verdadero con el campo de la ciencia.

Al comienzo de su estancia en Cambridge, se interesó en primer lugar por la química, y este interés, según se dice, se manifestó a lo largo de toda su vida. Durante su primer año de estudios, y probablemente por primera vez, leyó una obra de matemáticas sobre la geometría de Euclídes, lo que despertó en él el deseo de leer otras obras. Parece también que su primer tutor fue Benjamin Pulleyn, posteriormente profesor de griego en la Universidad. En 1663, Newton leyó la Clavis mathematicae de Oughtred, la Geometría a Renato Descartes de Van Schooten, la Óptica de Kepler, la Opera mathematica de Vieta, editadas por Van Schooten y, en 1644, la Aritmética de Wallis que le serviría como introducción a sus investigaciones sobre las series infinitas, el teorema del binomio, ciertas cuadraturas. También a partir de 1663 Newton conoció a Barrow, quien le dio clase como primer profesor lucasiano de matemáticas. En la misma época, Newton entró en contacto con los trabajos de Galileo, Fermat, Huygens y otros, a partir probablemente de la edición de 1659 de la Geometría de Descartes por Van Schooten.

Desde finales de 1664, Newton parece dispuesto a contribuir personalmente al desarrollo de las matemáticas. Aborda entonces el teorema del binomio, a partir de los trabajos de Wallis, y el cálculo de fluxiones. Después, al acabar sus estudios de bachiller, debe volver a la granja familiar a causa de una epidemia de peste bubónica. Retirado con su familia durante los años 1665-1666, conoce un período muy intenso de descubrimientos: descubre la ley del inverso del cuadrado, de la gravitación, desarrolla su cálculo de fluxiones, generaliza el teorema del binomio y pone de manifiesto la naturaleza física de los colores. Sin embargo, Newton guarda silencio sobre sus descubrimientos y reanuda sus estudios en Cambridge en 1667.

De 1667 a 1669, emprende activamente investigaciones sobre óptica y es elegido fellow del Trinity College. En 1669, Barrow renuncia a su cátedra lucasiana de matemáticas y Newton le sucede y ocupa este puesto hasta 1696. El mismo año envía a Collins, por medio de Barrow, su Analysis per aequationes numero terminorum infinitos. Para Newton, este manuscrito representa la introducción a un potente método general, que desarrollará más tarde: su cálculo diferencial e integral.

En 1672 publicó una obra sobre la luz con una exposición de su filosofía de las ciencias, libro que fue severamente criticado por la mayor parte de sus contemporáneos, entre ellos Robert Hooke (1638-1703) y Huygens, quienes sostenían ideas diferentes sobre la naturaleza de la luz. Como Newton no quería publicar sus descubrimientos, no le faltaba más que eso para reafirmarle en sus convicciones, y mantuvo su palabra hasta 1687, año de la publicación de sus Principia, salvo quizá otra obra sobre la luz que apareció en 1675.

Desde 1673 hasta 1683, Newton enseñó álgebra y teoría de ecuaciones, pero parece que asistían pocos estudiantes a sus cursos. Mientras tanto, Barrow y el astrónomo Edmond Halley (1656-1742) reconocían sus méritos y le estimulaban en sus trabajos. Hacia 1679, verificó su ley de la gravitación universal y estableció la compatibilidad entre su ley y las tres de Kepler sobre los movimientos planetarios.

Newton descubrió los principios de su cálculo diferencial e integral hacia 1665-1666, y durante el decenio siguiente elaboró al menos tres enfoques diferentes de su nuevo análisis. Desde 1684, su amigo Halley le incita a publicar sus trabajos de mecánica, y finalmente, gracias al sostén moral y económico de este último y de la Royal Society, publica en 1687 sus célebres Philosophiae naturalis principia mathematíca. Los tres libros de esta obra contienen los fundamentos de la física y la astronomía escritos en el lenguaje de la geometría pura. El libro I contiene el método de las "primeras y últimas razones" y, bajo la forma de notas o de escolios, se encuentra como anexo del libro III la teoría de las fluxiones. Aunque esta obra monumental le aportó un gran renombre, resulta un estudio difícil de comprender, y parece que Newton quiso que fuera así con el fin «de evitar ser rebajado por pequeños semisabios en matemáticas». Quiso escapar así a las críticas suscitadas por sus textos sobre la luz.

Después de una larga y atroz enfermedad, Newton murió durante la noche del 20 de marzo de 1727, y fue enterrado en la abadía de Westminster en medio de los grandes hombres de Inglaterra.

"No sé cómo puedo ser visto por el mundo, pero en mi opinión, me he comportado como un niño que juega al borde del mar, y que se divierte buscando de vez en cuando una piedra más pulida y una concha más bonita de lo normal, mientras que el gran océano de la verdad se exponía ante mí completamente desconocido."

PROBABILIDAD

¿Para qué?

Para conocer nuestras posibilidades reales de acertar al azar o ganar un premio en un sorteo. Para determinar las posibilidades de ocurrencia de un fenómeno, en contraposición a otros que puedan presentarse. Para conocer la forma en que la ocurrencia de determinado suceso, afecta a la aparición de otro.

Probabilidad

Experimentos aleatorios. Espacio Muestral. Definiciones

1. *Experimento determinista:* Aquel cuyo resultado está regido por leyes, sean o no de la Naturaleza, resultando por lo tanto predecible.
2. *Experimento aleatorio:* Aquel en el que interviene el azar, resultando en consecuencia impredecible.
3. *Resultado elemental o elemento:* Cada uno de los resultados obtenidos al realizar un experimento.
4. *Espacio muestral:* Es el conjunto de todos los resultados posibles de un experimento aleatorio; se designa por E.
5. *Suceso aleatorio:* También denominado simplemente suceso, es cada uno de los subconjuntos del espacio muestral que pueden formarse con resultados elementales de un experimento aleatorio.
6. *Espacio de sucesos:* Es el conjunto formado por todos los sucesos del experimento. Se denomina S
   Si E tiene k resultados posibles, S tiene 2^k elementos.
   Tipos de sucesos:
   - Suceso elemental: Aquel que contiene un único elemento o resultado elemental.
   - Suceso compuesto: El que está formado por 2 o más resultados elementales.
   - Suceso seguro: El que está formado por todos los resultados posibles del experimento. Coincide por lo tanto con el espacio muestral.
   - Suceso imposible: El que no puede presentarse, se representa por Φ
   - Sucesos iguales: Aquellos que están formados por los mismos sucesos elementales.
   - Suceso incluido: El suceso A se dirá incluido en el suceso B, si todos los resultados elementales de A, están incluidos en el suceso B. Representamos tal situación como $A \subset B$.

Operaciones con sucesos .Álgebra de Boole

1. **Unión de sucesos (OR) (+).** Representamos por $A \cup B$, el suceso caracterizado por la aparición, ocurrencia o cumplimiento del suceso A o el suceso B.

2. **Intersección de sucesos (AND) (·).** Representamos por $A \bigcap B$, el suceso caracterizado por la aparición simultánea de los sucesos A y B.
3. **Diferencia de sucesos (NOT)(-).** Representamos por A-B, el suceso caracterizado por la aparición del suceso A y simultáneamente, la no aparición del suceso B.

Propiedades de la Unión: Dados los sucesos A, B, C $\in$ S, se verifican las siguientes propiedades:

- Asociativa: $(A \bigcup B) \bigcup C = A \bigcup (B \bigcup C)$
- Conmutativa: $A \bigcup B = B \bigcup A$
- Idempotente: $A \bigcup A =$ A
- Complementación: $A \bigcup A^c = E$
- Elemento neutro: $A \bigcup \Phi = A$

Propiedades de la Intersección: Dados los sucesos A, B, C $\in$ S, se verifican las siguientes propiedades:

- Asociativa: $(A \bigcap B) \bigcap C = A \bigcap (B \bigcap C)$
- Conmutativa: $A \bigcap B = B \bigcap A$
- Idempotente: $A \bigcap A = A$
- Complementación: $A \bigcap A^c = \Phi$
- Elemento neutro: $A \bigcap E = A$
- Elemento Absorbente: $A \bigcap \Phi = \Phi$

Propiedades relacionales de Unión e Intersección: Se trata de tres propiedades que relacionan ambas operaciones:

- Simplificativas o de absorción:
  $A \bigcup (B \bigcap A) = A \quad , \quad A \bigcap (B \bigcup A) = A$
- Distributiva de la Unión respecto de Intersección:
  $$A \bigcup (B \bigcap C) = (A \bigcup B) \bigcap (A \bigcup C)$$
- Distributiva de la Intersección respecto de la Unión:
  $$A \bigcap (B \bigcup C) = (A \bigcap B) \bigcup (A \bigcap C)$$

La terna (S,$\bigcup$,$\bigcap$), con las propiedades reseñadas respecto a Unión e Intersección, asociada al espacio muestral E, recibe el nombre de Álgebra de Boole de sucesos aleatorios.

Frecuencia de un suceso

Si A es un suceso cualquiera de un experimento aleatorio y n es el número de pruebas que se realizan, se define la frecuencia absoluta de A y se denota f(A), como el número de veces que el suceso A se ha presentado a lo largo de las n pruebas.

En analogía con la estadística uni y bidimensional, se define la frecuencia relativa y se denota $f_r(A)$, como el número de veces que se ha presentado el suceso A, en relación al número total de pruebas n.

Así pues: $f_r(A) = \frac{f(A)}{n}$

De la propia definición, pueden extraerse varias consecuencias:

$0 \leq f_r(A) \leq 1 \quad ; \quad \sum_i f_r(A_i) = 1$

Definición Axiomática de Probabilidad (Kolmogorov)

Sea K el Álgebra de Boole representada por la terna (S,$\bigcup$,$\bigcap$), llamamos Probabilidad a toda aplicación P definida entre el espacio de los sucesos S y el conjunto de los números reales ℝ, tal que a todo suceso A$\in$S, le asigna un número real P(A) que denominamos probabilidad del suceso A.

$$P : S \to \Re$$
$$A \to P(A)$$

Dicha aplicación, verifica los siguientes axiomas:

A1. La probabilidad del suceso seguro es 1.

P(E)=1

A2. La probabilidad de cualquier suceso es un número no negativo

P(A)$\geq$0

A3. Si los sucesos A y B son incompatibles, la probabilidad del suceso A$\bigcup$B es la suma de probabilidades de los sucesos A y B.

$$A \bigcap B = \Phi \Rightarrow P(A \bigcup B) = P(A) + P(B)$$

Conclusiones directas de la definición axiomática de Probabilidad

1. $A \subset B \Rightarrow P(A) \leq P(B)$
2. $\forall A \in S \Rightarrow 0 \leq P(A) \leq 1$
3. $P(A) + P(A^c) = P(A \bigcup A^c) = P(E) = 1 \Rightarrow P(A^c) = 1 - P(A)$

Teoremas derivados de la definición axiomática de Probabilidad

1. **Th1**.La probabilidad del suceso imposible es 0

Dem.

Φ *y* E *son sucesos incompatibles, verificándose* $\Phi \cup E = E \Rightarrow \Phi$ *y* E *son complementarios*

Por el axioma 3, *tendremos que*:

$P(\Phi \cup E) = P(\Phi) + P(E) = P(E) = 1 \Rightarrow P(\Phi) = 1 - P(E) = 1 - 1 = 0$

2. **Th2.** $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Dem.

Suponemos de partida que $A \cap B \neq \Phi$ *. En consecuencia, podemos expresar A y B como:*

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

Así pues:

$A \cup B = (A \cap \overline{B}) \cup (A \cap B) \cup (\overline{A} \cap B)$ *, con lo que la unión de A y B queda expresada como unión de sucesos incompatibles y estamos en condiciones de aplicar el A3 de Probabilidad. Tendremos pues:*

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

Comparando los tres resultados, tendremos que :

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

Probabilidad condicionada

Supongamos que hemos realizado un determinado experimento aleatorio n veces, de las cuales n_A, ocurrió el suceso A; de ellas, n_{AB} ocurrió el suceso B. Se produce entonces la siguiente relación de frecuencias.

$$f(A) = \frac{n_A}{n}; \quad f\left(B/A\right) = \frac{n_{AB}}{n}; \quad f(A \cap B) = \frac{n_{AB}}{n}$$

dado que obviamente se verifica que :

$$\frac{n_{AB}}{n} = \frac{n_A}{n} \cdot \frac{n_{AB}}{n_A}$$

identificando cada cociente con su frecuencia correspondiente, tendremos que:

$$f(A \cap B) = f(A) \cdot f\left(B/A\right)$$

de donde podremos deducir:

$$f\left(B/A\right) = \frac{f(A \cap B)}{f(A)}$$

Expresión que podemos "traducir" directamente a términos de probabilidad

$$P\left(B/A\right) = \frac{P(A \cap B)}{P(A)}$$

y finalmente, expresar como:

$$P(A \cap B) = P(A) \cdot P\left(B/A\right)$$

Expresión que conocemos con el nombre de **Teorema del Producto**

Sucesos Independientes

En general, es sencillo comprobar que P(B) $\neq$ P(B/A). Diremos que los sucesos A y B son independientes, si se verifica que P(B)=P(B/A).

Th1.Si A es independiente de B, entonces B es independiente de A y además $P(A \cap B) = P(A) \cdot P(B)$

Dem.

B independiente de A $\Rightarrow P(B) = P\left(B/A\right)$. *Usando el Teorema del producto, tendremos que :* $P(A \cap B) = P(A) \cdot P\left(B/A\right) \Rightarrow P(A \cap B) = P(A) \cdot P(B)$

Se verificará además que :

$$P\left(A/B\right) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \cdot P(B)}{P(B)} = P(A)$$

con lo que quedará demostrado que A es independiente de B.

Teorema de Probabilidad Total

Sean $A_1, A_2, \ldots\ldots, A_n$, sucesos que forman una partición de E, es decir:

$$A_1 \cup A_2 \cup \ldots\ldots \cup A_n = E; \quad A_i \cap A_j = \Phi \ \forall i, j$$

Sea B_j cualquier suceso para el que se conozca $P\left(B/A_j\right) \forall j = 1,2,\ldots\ldots n$; supongamos además conocida $P(A_i)$ $\forall$ i=1,.....,n. Con estos datos, podemos calcular P(B).

Dem.

Sabemos que :

$$P(B) = P(B \cap E) = P\left[B \cap \left(A_1 \cup A_2 \cup ...... \cup A_n\right)\right] = \sum_i P(B \cap A_i) = \sum_i P(A_i) P\left({B}/{A_i} \right)$$

donde la expresión anterior, se ha obtenido teniendo en cuenta que

$$(B \cap A_j) \cap (B \cap A_i) = \Phi \ \forall i, j \quad \textit{ya que} \quad A_i \cap A_j = \Phi \ \forall i, j$$

El resultado anterior, $P(B) = \sum_i P(A_i) P\left({B}/{A_i} \right)$, se conoce con el nombre de **Teorema de Probabilidad Total.**

Teorema de Bayes

Sabemos, por el Teorema del producto y la propiedad conmutativa de la intersección, que:

$$P(A_i \cap B) = P(A_i) P\left({B}/{A_i} \right) = P(B \cap A_i) = P(B) P\left({A_i}/{B} \right)$$

Lo que podemos expresar simplemente como:

$$P(A_i) P\left({B}/{A_i} \right) = P(B) P\left({A_i}/{B} \right)$$

de donde, despejando

$$P\left({A_i}/{B} \right) = \frac{P(A_i) P\left({B}/{A_i} \right)}{P(B)}$$

Sustituyendo P(B) por su valor según el Teorema de Probabilidad Total, tendremos:

$$P\left({A_i}/{B} \right) = \frac{P(A_i) P\left({B}/{A_i} \right)}{\sum_i P(A_i) P\left({B}/{A_i} \right)}$$

Expresión que se conoce como **Teorema de Bayes** o **Fórmula de Bayes.**

Regla de Laplace

Se trata de una regla de cálculo probabilístico, válida únicamente cuando los sucesos elementales son equiprobables; afirma:

Sea $E = \{A_1, A_2, ...A_n\}$ un espacio muestral formado por n sucesos elementales, de forma que todos ellos tienen la misma probabilidad.

Si el suceso A está formado por m sucesos elementales con $m \leq n$, la probabilidad de A es igual al cociente entre el número de resultados favorables y el número total de resultados posibles.

Probabilidad. Problemas

Probabilidad simple

1. Sacamos al azar dos cartas de una baraja francesa (52 cartas en cuatro palos), calcúlese la probabilidad de:
   a) Sacar dos corazones
   b) Sacar un corazón y una pica
   c) Una de las dos extracciones resulte ser un as.

2. Escogemos al azar 5 bolas de un grupo de 50 entre las que se encuentran 12 bolas negras. Calcúlese la probabilidad de:
   a) Escoger 3 bolas negras
   b) Al menos una de las 5 sea negra
   c) Exactamente 1 sea negra

3. 15 piezas idénticas numeradas del 1 al 15 se extienden alineadas sobre una mesa. Calcúlese la probabilidad de que:
   a)Resulten ordenadas de mayor a menor
   b)Los números 14 y 15 queden situados en las posiciones 14 y 15

4. Un dado está trucado de forma que la probabilidad de obtener 5 es cuádruple de la de obtener 1 y 2 y doble de la de obtener 3, 4 o 6. Calcúlese la probabilidad de obtener 5 al lanzar el dado.

5. Un opositor debe enfrentarse a una prueba en la que será preguntado por un temario de 60 temas, de los que sólo tuvo tiempo de estudiar 20. El examen consta de 4 preguntas elegidas por sorteo. Calcúlese la probabilidad de que conteste correctamente al menos los dos temas que necesita para pasar la calificación de corte.

Probabilidad condicionada

6. Extraemos dos cartas de una baraja española sin efectuar reemplazamiento, lo que quiere decir que la carta extraída en primer lugar no se reincorporará al mazo de cartas antes de extraer la segunda. Calcúlese la probabilidad de que la primera sea bastos y la segunda oros.

7. En una clase el 35 por ciento de los alumnos son varones. El 15 % de los mismos tiene como afición los deportes náuticos. Calcúlese la probabilidad de que al elegir al azar un alumno, este resulte ser varón y con aficiones marineras.

8. Calcúlese la probabilidad de que al elegir consecutivamente y al azar 3 bolas de una bolsa que contiene 25 (de las cuales 10 son blancas y 15 negras), no salga elegida ninguna bola negra. Considérense los casos sin y con reeplazamiento.

9. Elegimos al azar tres pollos de una muestra de la producción de una granja de cría, compuesta por 10 pollos blancos, diez amarillos, diez marrones y diez moteados, calcúlese la probabilidad de:

   a) Extraer 3 blancos
   b) Extraer 3 marrones
   c) No extraer ninguno moteado
   d) Extraer dos moteados y uno amarillo

   Suponemos la extracción consecutiva y sin reemplazamiento.

10. De un lote de 100 artículos, de los cuales 7 son defectuosos, se realizan 5 extracciones consecutivas. Calcúlese la probabilidad de que el primer articulo extraído resulte ser defectuoso. Calcúlese la probabilidad de que el primer artículo defectuoso aparezca en 5º lugar .

Sucesos independientes

11. Calcular la probabilidad de obtener 5 cruces en 5 lanzamientos consecutivos de una moneda.

12. Calcular la probabilidad de obtener 5 cruces en el lanzamiento de 5 monedas.

Teorema de Probabilidad Total

13. Un instituto tiene 860 alumnos, de los cuales, 100 están en 1º de ESO, 120 en 2º de ESO, 120 en 3º de ESO y 120 en 4º de ESO. El resto, se reparte a partes iguales entre 1º y 2º de Bachillerato. Un estudio sobre consumo de bebidas alcohólicas mostró que el porcentaje de alumnos que admitían usarlas era, respectivamente de 1, 3, 6, 15, 25 y 35.
    Calcúlese la probabilidad de que elegido una alumno al azar, éste resulte ser consumidor de alcohol.

Teorema de Bayes

14. En las condiciones iniciales del problema anterior, calcúlese la probabilidad de que elegido un alumno al azar, y resultando ser abstemio, éste resulte ser de 2º de Bachillerato.

15. Un test formula preguntas con 4 posibles respuestas cada una. Una persona tiene una probabilidad 0,3 de conocer la respuesta a la pregunta (ha estudiado, digamos, el 30% de los temas del test) en cuyo caso, responderá la misma acertadamente, y responderá al azar en caso contrario. Si la persona contestó acertadamente, calcúlese la probabilidad de que conociese la respuesta y no fuese una casualidad.

16. Una urna, A, contiene 10 bolas blancas y 5 negras; otra urna, B, contiene 6 bolas blancas y 4 negras. Se elige una urna al azar y después una bola de esa urna. Calcúlese la probabilidad de obtener bola negra. Si realizamos el experimento y la bola resulta ser negra, calcúlese la probabilidad de que proceda de la urna A.

17. El 70% de los clientes de una aseguradora tiene más de 25 años. Un 5% de dichos clientes, tiene un accidente a lo largo del año. En el caso de clientes menores de 25 años, el porcentaje de accidentados anuales sube al 20%. Si elegimos un cliente al azar, calcúlese la probabilidad de que tenga un accidente ese año. Si elegimos un cliente accidentado, calcúlese la probabilidad de que sea menor de 25 años.

18. Lanzamos 2 dados. En el caso de que aparezcan resultados distintos, sumamos las puntuaciones. Calcúlese la probabilidad de que dicha suma sea par.

19. El 40% de la población de una ciudad es menor de 20 años. El 25% son varones y el 15% son varones de menos de 20 años. Escogemos al azar un habitante de la ciudad:
    a) Calcúlese la probabilidad de que sea una mujer mayor de 20 años
    b) Si resulta ser varón, calcúlese la probabilidad de que sea menor de 20 años.
    c) Si es menor de 20 años, calcúlese la probabilidad de que sea varón

20. En una ciudad, el 55% de la población compra el diario A, el 30% compra el diario B y el 20 %, compra ambos diarios. Escogemos una persona al azar, calcúlese:
    a) Probabilidad de que sólo lea el diario A
    b) Probabilidad de que si compra el diario A, compre también el diario B.

21. El total de artículos manufacturados por una fábrica se reparte entre tres máquinas, que denominaremos A,B y C. Dichas máquinas producen, respectivamente, el 20, el 30 y 50% de la producción. Los porcentajes de artículos defectuosos a la salida de máquinas son, respectivamente, el 1; 1,5 y 2,5%. Si elegimos un artículo al azar, calcúlese la probabilidad de que:
    a) Sea defectuoso
    b) Siendo defectuoso, fuese fabricado por la máquina C
    c) No siendo defectuoso, fuese fabricado por la máquina A

22. La probabilidad de que una alarma contra incendios funcione correctamente, es 0,7; si funciona, la probabilidad de que los bomberos lleguen a tiempo, es 0,8. Si por el contrario no funciona, ésta última probabilidad es 0,3.Calcúlese:
    a) Probabilidad de que en caso de incendio, los bomberos lleguen a tiempo.
    b) Probabilidad de que la alarma sonara, si efectivamente los bomberos llegaron a tiempo

23. Un jugador consigue gol en un 75 % de sus lanzamientos de penalty. Si realiza una serie de 3, calcúlese la probabilidad de que al menos 1 sea gol.

24. El diez por ciento de las piezas de recambio de automóviles, que llegan a una compañía distribuidora, tiene algún defecto. En la distribuidora son sometidas a una inspección, que tiene un error del 5%. Si escogemos una pieza al azar, calcúlese la probabilidad de que sea efectivamente defectuosa, si en la inspección fue declarada defectuosa.

25. Una empresa de reparaciones domésticas tiene, en el cuadro de personal, 20 pintores de los que 7 son mujeres, 6 fontaneros y 6 fontaneras y 9 albañiles de los que 3 son mujeres. Si elegimos un empleado al azar, calcúlese la probabilidad de que:
    a) Sea hombre y fontanero
    b) Siendo mujer, sea albañil
    c) Siendo pintor, resulte ser hombre.

PIERRE-SIMON LAPLACE

Matemático y astrónomo francés, nacido en Normandia el 2 de marzo de 1749 en el seno de una humilde familia de campesinos. Sus primeros estudios fueron sufragados por sus convecinos, asistió a la escuela militar de Beaumont y a los 10 años ingresó a la universidad de Caen»; en 1767 con el apoyo de D'Alembert ocupó una plaza de profesor en la Escuela Militar de París. En 1773 fue elegido para la Academia de Ciencias como miembro adjunto y como numerario en 1785. En 1810 sería también elegido miembro de la Academia Francesa.

Sus primeros trabajos científicos fueron de índole exclusivamente matemática. Pero pronto dirigió su atención a la Mecánica celeste que desde la obra de Newton había evolucionado poco y tenia planteados serios problemas. En 1790 publicó la Exposition du système du monde, obra astronómica de divulgación que incluye una teoría cosmogénica según la cual el sol se habría formado por contracciones y enfriamiento de una nebulosa incandescente; los planetas serían fragmentos desprendidos posteriormente por el sol y los satélites procederían de los planetas mediante la misma génesis. Está teoría fue muy difundida y gozó de la aceptación casi general durante mas de un siglo. Pero Laplace estaba, desde 1773 según propia referencia, tratando de resolver el gran problema mecánico del sistema solar y conseguir una relación tan intima entre la teoría y la observación que las ecuaciones empíricas resulten superfluas. Fruto de este esfuerzo fue la Mecánica celeste obra de cinco volúmenes (1799-1825).En ella se enfrenta con el enigma de las anomalías observadas en las órbitas de los planetas, para cuya corrección, según Newton se necesitaba periódicamente la intervención directa del Creador a fin de evitar el caos. Laplace empezó estudiando las irregularidades existentes en los movimientos de Júpiter y Saturno hasta demostrar que eran consecuencia de la atracción mutua y que se compensaba en un ciclo de 92 años; así pues, teniendo tales perturbaciones carácter periódico y no acumulativo como creyera Newton, la estabilidad del conjunto de los dos planetas quedaba indefinidamente asegurada.

En electromagnetismo descubrió las leyes que llevan su nombre.

En 1812 publicó una obra matemática, Teoría analítica de las probabilidades, que daba a conocer las llamadas Transformadas de Laplace, de considerable aplicación en diversas cuestiones matemáticas.A ella siguió en 1814 el Essai philosophique des probabilités.