Año 2011

TECNÓLOGO INFORMÁTICO

PROBABILIDAD Y ESTADÍSTICA

CURSO TEÓRICO

Parte 1) Técnicas de Conteo

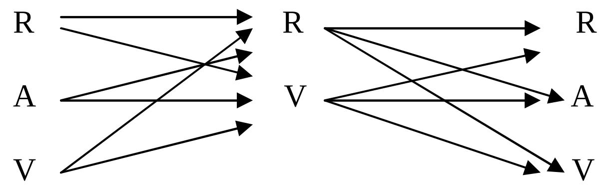
Regla de la Suma:"Si 2 acontecimientos no se verifican simultáneamente, (en este caso se dice que son excluyentes), uno de ellos se cumple en m casos y el otro en n ocasiones, entonces el acontecimiento definido en que uno de ellos es verdadero se cumple en (m+n) veces."

Ejemplo: Se arroja un dado al aire, ¿cuántos casos posibles hay para que el resultado sea un 5 o un par? Son acontecimientos excluyentes, si sale un 5, el resultado no es par y si sale par no es posible que el resultado haya sido un 5. Las chances para que el resultado sea un 5 son 1/6, mientras que para que sea par es ½, por tanto para que sea par o haya salido un 5 tenemos 1/6 + ½ = 2/3 de casos favorables.

Regla del Producto: "Si un acontecimiento se verifica en m ocasiones y por cada una de estas, otro acontecimiento se cumple en n casos, entonces el acontecimiento dado porque ambos sean verdaderos simultáneamente, será cierto en m.n veces"

Ejemplo: Se tienen 3 semáforos, 2 en esquinas contiguas y otro en medio de ellos. Este último tiene rota la luz amarilla ¿Cuántas señales de luces diferentes pueden realizarse, entre los 3 semáforos? Con el primer semáforo 3 señales, pero por cada una de estas, en el segundo semáforo, dos señales, finalmente, por dada una de estas 2, con el 3º, podremos realizar 3 señales adicionales, por tanto, en total tendremos 3*2*3 = 18 señales.

Esquema: Semáforo 1 Semáforo 2 Semáforo 3



En otras palabras: (RRR), (RRA), (RRV), (RVR), (RVA), (RVV)
(ARR), (ARA), (ARV), (AVR), (AVA), (AVV)
(VRR), (VRA), (VRV), (VVR), (VVA), (VVV)

Arreglos o Variaciones de m en n sin repetición:"Se llaman Arreglos de m elementos tomados de n, sin repetición, al nº de grupos que se puedan formar a partir de los m elementos dados agrupándolos de a n, de modo tal que cada grupo difiera de otro en: a) el orden de los elementos; b) en un elemento por lo menos."

Ejemplo: Dados los elementos {a, b, c, d}, formar los arreglos de 4 en 1, en 2, en 3 y en 4.

1) A_1^4: {a}; {b}; {c}; {d}, o sea A_1^4 = 4, de la misma forma entonces, $A_1^m = m$.-

2) A_2^4 : {a, b}; {a, c}; {a, d}; es decir, cada A_1^4 determina 3 A_2^4, por tanto A_2^4 = 4*3 = 12

   De la misma forma $A_2^m = m*(m-1)$.-

3) A_3^4: {a, b, c}; {a, b, d}; es decir, cada A_2^4 determina 2 A_3^4, por tanto A_3^4 = 4*3*2 = 24

   De esa forma, $A_3^m = m*(m-1)*(m-2)$.-

4) A_4^4: {a, b, c, d}; o sea que cada A_3^4 determina 1 A_4^4, por lo tanto A_4^4 = 4*3*2*1 = 24

   Análogamente, $A_4^m = m*(m-1)*(m-2)*(m-3)$

De lo anterior deducimos que:

$$\boxed{A_n^m = m*(m-1)*(m-2)*(m-3)...........*[m-(n-1)]}$$

Lo cual indica que para calcular los A_n^m, se realiza un producto de naturales sucesivos, decrecientes, a partir del índice m, existiendo tantos factores como indica el sub índice.

Ejemplo: $A_4^5 = 5*4*3*2 = 120$

Factorial de un natural: "Dado un natural n, se llama factorial de n y se denota n! al producto de los n primeros y sucesivos naturales es decir n*(n-1)*(n-2)…..*2*1 ⇔ n ≥ 1 y por otra parte, en caso que n = 0, se indica que 0! = 1"

$$n! = \begin{cases} 1 \Leftrightarrow n = 0 \\ n*(n-1)*(n-2).....*2*1 \Leftrightarrow n \geq 1 \end{cases}$$

Tomando en cuenta que el nº de elementos disponibles m, debe ser ≥ que el nº de elementos con que se forman los grupos, o sea n, supongamos para fijar ideas que m > n. Desde este punto de vista, si formamos el m!, a medida que los factores de la forma (m – 1), (m – 2), se van generando, observamos que los sustraendos van creciendo de a una unidad a partir de 1, por tanto antes de llegar a (m – 1), con lo cual en el producto final habremos llegado al factor 1, estos sustraendos deben necesariamente pasar por n, así que podríamos escribir el siguiente desarrollo:

$$m! \;=\; \underbrace{m*(m-1)*(m-2)*(m-3)*(m-4)\ldots*[m-(n-1)]}_{\mathrm{I}}*\underbrace{(m-n)*[m-(n+1)]\ldots*4*3*2*1}_{\mathrm{II}}$$

El producto señalado como I, según lo ya expuesto, representa los A_n^m, mientras que el producto señalado como II, de acuerdo a la definición de factorial de un natural, representa el (m –n)!, por tanto podemos expresar:

m! = (m –n)! * A_n^m, de donde surge que: $\boxed{A_n^m = \frac{m!}{(m-n)!}}$ que representa la misma fórmula ya hallada, pero escrita utilizando factoriales.

Permutaciones de m elementos: “Se llaman Permutaciones de m elementos, al nº de grupos que se pueden formar con los m elementos agrupándolos de a m, de este modo, como todos los grupos estarán compuestos por los mismos elementos, el único criterio que permite distinguir un grupo de otro es el orden, en que estos se hallan dispuestos en la em-upla”

De la definición anterior se desprende que las P(m), (permutaciones de m elementos), coinciden con los A_m^m, por tanto será:

$$\boxed{P(m) = A_m^m = \frac{m!}{(m-m)!} = \frac{m!}{0!} = \frac{m!}{1} = m!\text{, por lo tanto las } P(m) = m!}$$

Ejemplo: Con las letras de la palabra FORMULA, ¿cuántas palabras de 4 letras, con o sin sentido se pueden formar?

Es claro que se trata de Arreglos, pues al cambiar el orden de las letras se obtiene otra palabra e incluso al cambiar una sola letra, por tanto la solución serán los $A_4^7 = 7*6*5*4 = 840$.

Ejemplo: Se tienen 3 libros de Matemática, 2 de Física y 1 de Química, en un estante. ¿De cuántas maneras diferentes se pueden ubicar los libros en el estante? Es claro que se ubican en diferente formas cambiando el orden entre 6 libros que se disponen, por tanto la solución son las P(6) = 6! = 720.

Combinaciones de m en n, sin repetición: "Se llaman Combinaciones de m elementos tomados de a n, al nº de grupos que se pueden formar a partir de los m elementos dados, agrupándolos de a n, de modo que, cada grupo difiera de otro, en por lo menos 1 elemento"

Queda entonces absolutamente claro, que la única diferencia entre Arreglos y Combinaciones es que en estas últimas, el orden en que se ubican los elementos en una en-upla, no representa un criterio que permita diferenciar un grupo de otro.

Ejemplo: Dados los elementos {a, b, c, d}, formar las Combinaciones de 4 en 2.

C_2^4: {a, b}; {a, c}; {a, d}; {b, c}; {b, d}; {c, d}; lo que determina, entre otras cosas que $C_2^4 = 6$, es importante notar que si a cada una de las C_2^4, les aplicamos un cambio de orden se obtienen otras C_2^4, es decir: {b, a}; {c, a}; {d, a}; {c, b}; {d, b}; {d, c}; que naturalmente, también son 6. Vamos ahora a agrupar estas 12 parejas en forma rectangular, considerando que en su conjunto representan los A_2^4.

$$\left\{ \begin{matrix} \{a,b\} & \{a,c\} & \{a,d\} & \{b,c\} & \{b,d\} & \{c,d\} \\ \{b,a\} & \{c,a\} & \{d,a\} & \{c,b\} & \{d,b\} & \{d,c\} \end{matrix} \right\}$$

Si observamos detenidamente, veremos que las columnas del cuadro, representan las C_2^4, mientras que las filas son las P(2). Por lo tanto como en su conjunto determinan los A_2^4, podemos escribir:

A_2^4 = P(2) * C_2^4 de donde surge que: $C_2^4 = \dfrac{A_2^4}{P(2)}$ y generalizando resulta que: $C_n^m = \dfrac{A_n^m}{P(n)}$, de donde:

$$\boxed{C_n^m = \frac{m!}{n!*(m-n)!}}$$

Que es la fórmula a partir de la cual calculamos combinaciones de m en n sin repetición.

Ejemplo: Se extraen 4 cartas con reposición de un mazo de 40 cartas, ¿cuántas cuaternas diferentes se pueden extraer? En un juego de cartas el orden NO interviene, por tanto son $C_4^{40} = 91340$

Arreglos de m elementos con Repetición de orden m y agrupados de a n

Supongamos ahora que queremos agrupar los dígitos 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, en grupos de 3, pudiendo reiterarse, es decir que un número podría ser 111, o 995, etc. Parece claro que se trata de arreglos, ya que tanto el orden como la diferencia en por lo menos 1 dígito, implica grupos distintos, pero como un mismo dígito puede reiterarse, incluso hasta 3 veces, entonces NO es posible utilizar la expresión que hemos deducido, sobre la hipótesis de que no haya repetición, ya que por ejemplo los 2 ejemplos indicados NO serían contados, obviamente entre muchos otros.

¿Cómo encarar esta contabilización? Bueno, una forma simple de hacerlo es fijar la atención en un primer dígito, cualquiera sea este, por ejemplo un 2, está claro que como 2º dígito para acompañar a nuestro 2, tenemos 10 posibilidades, pero análogamente, para cada una de estas 10, tendremos 10 posibilidades para el 3º dígito del nº a formar. La regla del producto nos dice que con el 2 como primer dígito, podríamos formar 100 números de 3 cifras, pero como para el primer dígito también tenemos 10 posibilidades, entonces en definitiva podremos registrar 1000 números de 3 cifras, con repetición. Observemos que $1000 = 10^3$, lo cual no parece ser una casualidad.

De disponer de m elementos y con ellos querer agruparlos de a n, con m repeticiones entonces los arreglos de m en n con índice de repetición m vienen dados por:

$$\boxed{AR_n^m = m^n}$$

Ejemplo: Una empresa de ómnibus emite boletos numerados de 5 dígitos, con los números 1, 2, 3, 4, 5, 6. ¿Cuántos boletos diferentes puede numerar? Bueno son arreglos con repetición de orden 6, agrupados de a 5, entonces son $6^5 = 7776$ boletos.

Permutaciones de m elementos con Repetición de r de ellos

Supongamos que tenemos 4 elementos a, b, c y d. Pero de estos elementos el b coincide con el c, es decir tenemos 2 elementos que se repiten, o sea a, b, b, d. Formemos primero las permutaciones de los 4 elementos a, b, c, d:

abcd	*abdc*	**acbd**	*acdb*	*adbc*	*adcb*
bacd	*badc*	*bcad*	*bcda*	*bdac*	*bdca*
cabd	*cadb*	*cbad*	*cbda*	*cdab*	*cdba*
dabc	*dacb*	*dbac*	*dbca*	*dcab*	*dcba*

Es fácil advertir que las permutaciones de 4 que hemos formado con la condición de que b = c, y que la única diferencia que tengan es la permutación en el orden de estos 2 elementos serán por tanto idénticas. Vamos a marcar con negritas 2 de estas permutaciones.

Con esto vemos que por cada una de las 24 permutaciones, hay una idéntica a ella, lo que no permite distinguirlas como diferentes. Esto lleva a que si b = c, entonces tendremos sólo $\frac{24}{2!} = \frac{24}{2} = 12$ permutaciones diferentes. Observemos también que, si además de ser b = c, fuera b = c = d, entonces, la permutación indicada en negrita **abcd**, sería igual además de a la acbd, a todas las que se obtienen permutando entre sí las 3 letras b, c y d, manteniendo la a en la misma posición. En este caso serían 3! por lo ya visto para permutaciones sin repetición. Por lo tanto, en vez de 24 permutaciones diferentes tendríamos $\frac{24}{3!} = \frac{24}{6} = 4$ permutaciones de 4 elementos con 1 elemento cuyo orden de repetición es 3. Más general entonces, el nº de Permutaciones diferentes con m elementos en donde sólo uno de ellos se repite r veces, viene dado por: $P(m)_r^1 = \frac{m!}{r!}$ y más general aún, si se tienen h elementos, con índices de repetición r_1, r_2, r_3,….., r_h, en los m disponibles, el nº de permutaciones diferentes que se pueden formar vienen dado por:

$$\boxed{P(m)_{r1,r2,\ldots,rh}^{h} = \frac{m!}{r1!*r2!*r3!*\ldots*rh!}}$$

Ejemplo: Una persona intenta recordar una clave personal que ha olvidado, la misma es de 6 letras y se forma con las letras a, b y c, en las que cada una de ellas se repite 2 veces. ¿Puedes decirle cuántas alternativas tiene? S claro que disponemos de 6 lugares o elementos, de los cuales uno se repite 2 veces, otro también y finalmente el tercero igualmente, entonces m = 6; h = 3 y r_1=r_2=r_3=2, por lo tanto resulta:

$$P(6)_{2,2,2}^{3} = \frac{6!}{2!*2!*2!} = \frac{720}{8} = 90$$

Combinaciones de m en n, con repetición

"Se llaman combinaciones con repetición de m elementos tomados de a n, al nº de grupos que se puedan formar con los m elementos dados, tomándolos de a n, permitiendo que estos se hallen reiterados"

La expresión que más comúnmente se utiliza para realizar los cálculos respectivos es:

$$\boxed{CR_n^m = C_n^{m+n-1} = \frac{(m+n-1)!}{n!*(m-1)!}}$$

Justificación:

Podemos suponer que los m elementos dados son los números naturales 1, 2, 3,….., m, de no ser así es siempre posible una biyección entre los m elementos dados y el conjunto anterior. Como en las combinaciones no importa el orden, podemos suponer que cada en-upla que se genere se encuentra ordenada, según el orden natural. Por tanto, cada una de esas en-uplas será un sub conjunto del considerado inicialmente, en donde: $a_1 \le a_2 \le a_3 \le \ldots\ldots \le a_n$ en donde pueden haber elementos repetidos. Pero para cada una de esas en-uplas vale la siguiente desigualdad estricta:

$a_1 < a_2 + 1 < a_3 + 2 < a_4 + 3 < \ldots\ldots\ldots < a_n + n - 1$, cada una de esta secuencia numérica está comprendida entre los números naturales 1 y m + n -1 y de acuerdo a lo expuesto, existe la misma cantidad de en-uplas de estos últimos que de los primeros, lo que explica la igualdad del último recuadro.

Parte 2) Teoría de Probabilidad

2.1. Fenómenos Determinísticos y Aleatorios

Según la naturaleza de los fenómenos podemos distinguir entre:

Fenómenos determinísticos: "Son aquellos que en igualdad de condiciones, producen el mismo resultado. También se les conoce como Experimentos Científicos"

Fenómenos o Experimentos Aleatorios: "Son los que, en igualdad de condiciones, nada garantiza, a priori, obtener el mismo resultado."

La estadística y más concretamente la probabilidad se encarga de este segundo tipo de fenómenos, intentado dar una medida de la incertidumbre respecto a la ocurrencia o no, de estos.

2.2. Álgebra de Sucesos

Dado un experimento aleatorio, (por ejemplo lanzar un dado) cualquiera, se define:

Espacio Muestral: "Es el conjunto de todos los posibles resultados del experimento. Se nota como Ω."

En el ejemplo del lanzamiento de un dado sería: $\Omega = \{1, 2, 3, 4, 5, 6\}$. Si el experimento aleatorio fuera arrojar una moneda al aire, entonces $\Omega = \{C, N\}$. Si se tratara de hacer girar un disco, en torno a una flecha fija que señala un punto del perímetro de la circunferencia que limita al disco, entonces $\Omega = \{P / P \in \mathscr{C}\}$, siendo $\mathscr{C}$, dicha circunferencia. Salvo el último caso, se trata de espacios muestrales finitos y por tanto numerables, o sea discretos. El último ejemplo de Ω, es un conjunto infinito no numerable y por tanto continuo. Estas definiciones son importantes, pues determinan el tipo de variable aleatoria cuantitativa se deberá utilizar en cada caso, así como conceptualmente, que diferencias existen entre una variable discreta y una continua. Más adelante definiremos con más precisión este concepto.

Suceso: "Es un subconjunto de elementos del espacio muestral. Se define suceso seguro como el que ocurre siempre y suceso imposible al que no ocurre nunca. El suceso seguro coincide con el espacio muestral y el imposible con el vacío"

Al ser los sucesos subconjuntos, sobre ellos están definidas las siguientes operaciones:

Unión: "Dados dos sucesos A y B, se define el suceso unión A U B como el suceso que contiene todos los elementos de A y B." Es fácil advertir las siguientes propiedades:

1. Asociativa: (A U B) U C = A U (B U C).
2. Conmutativa: A U B = B U A.
3. Idempotencia: A U A = A.

Intersección: "Dados dos sucesos A y B, el suceso intersección ($A \cap B$) es el que contiene a la vez todos los elementos que están en A y en B." Al igual que en la Unión, se verifican las propiedades siguientes:

1. Asociativa: $(A \cap B) \cap C = A \cap (B \cap C)$.
2. Conmutativa: $(A \cap B) = B \cap A$.
3. Idempotente: $A \cap A = A$.

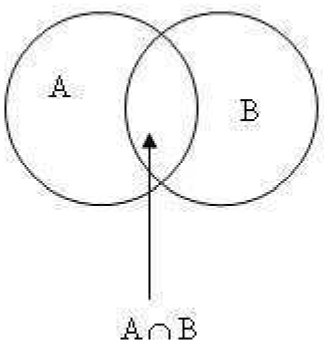
Otras propiedades mixtas que también se cumplen por parte de los Sucesos son:

1. Distributiva: $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.
2. Simplificativa: $A \cup (A \cap B) = A$ y $A \cap (B \cup A) = A$.

Complemento de un Suceso dado:"Se denota el suceso complemento de A como el suceso A^c que es aquél que contiene todos los elementos de Ω, que no están en A."

Sucesos iguales:" Se dice que dos sucesos son iguales si $A \subseteq B$ y $B \subseteq A$"

Una manera de observar los sucesos teóricos es mediante los diagramas de Venn, que son representaciones del tipo:



Leyes de De Morgan: 1) $(A \cup B)^c = (A^c \cap B^c)$ y 2) $(A \cap B)^c = (A^c \cup B^c)$

Sucesos Excluyentes:"Dos sucesos son excluyentes o incompatibles $\Leftrightarrow$ su intersección es el conjunto vacío: $A \cap B = \emptyset$". Esto quiere decir que 2 sucesos excluyentes no pueden ser ciertos simultáneamente, aunque si pueden ser ambos falsos. Por otra parte, 2 sucesos complementarios son excluyentes, por tanto nunca pueden ser ambos ciertos, pues son excluyentes, pero además y a diferencia de aquellos, tampoco pueden ser falsos simultáneamente.

2.2.1. σ Álgebra de sucesos

Dado un espacio muestral Ω, se define una **σ** algebra de sucesos $\mathscr{A}$, a una clase de subconjuntos de Ω, tal que:

1. Si A ∈ $\mathscr{A}$ ⇒ $A^c \in \mathscr{A}$.

2. Si A1, . . . ,An ∈ $\mathscr{A}$ $\Rightarrow (\overset{n}{\underset{1}{U}} A_i) \in \mathscr{A}$

3. El Ø ∈ $\mathscr{A}$.

2.3. Definición Axiomática de Probabilidad – Axiomas de Kolmogorov

Sea un espacio muestral Ω y $\mathscr{A}$ una σ álgebra de sucesos definida sobre él, definiremos probabilidad como una función P definida en $\mathscr{A}$, sobre los Reales, es decir P($\mathscr{A}$) -> R, de modo tal que $\forall A \in \mathscr{A}$ -> P(A) ∈ R.

Axiomas de Probabilidad:

Ax. 1) $\forall A \in \mathscr{A}$ -> $P(A) \geq 0$.

Ax. 2) P (Ω) = 1.

Ax. 3) Sean $A_1, A_2, \ldots, A_n$ sucesos disjuntos 2 a 2 es decir (Ai ∩ Aj) = Ø, $\forall_i \neq_j$, (también decimos que son una en-upla de sucesos excluyentes o incompatibles), entonces vale que:

$$P(\overset{n}{\underset{1}{U}} A_i) = \sum_{1}^{n} (Ai)$$

2.3.1. Propiedades

1. Si A ∈ $\mathscr{A}$, entonces P(A) = 1 − P (A^c).

Demo:

Por definición es (A U A^c) = Ω y P (Ω) = 1 por el Axioma 2, entonces P(A U A^c) = P (Ω) = P(A) + P (A^c) = 1 ⇒ P(A) = 1 − P (A^c)

2. **La probabilidad del conjunto vacío es 0, P (Ø) = 0**.

Demo:

Como $Ø^c$ = Ω entonces, por la propiedad anterior es P(Ω) = 1 − P(Ø) y como P(Ω) = 1, entonces surge que P (Ø) = 0

3. Sean A y B ∈ $\mathscr{A}$, entonces P(A U B) = P(A) + P(B) − P(A ∩ B).

Demo:

Podemos descomponer A y B de la forma: A = (A ∩ B) U (B^c ∩ A) y B = (A ∩ B) U (A^c ∩ B) por lo que (A U B) = (A ∩ B) U (B^c ∩ A) U (A^c ∩ B), pasando a probabilidades se tiene que:

P(A U B) = P(A ∩ B) + P(B^c ∩ A) + P(A^c ∩ B) esto por el Axioma 3
P(A) = P(A ∩ B) + P(B^c ∩ A) y P(B) = P(A ∩ B) + P(A^c ∩ B) ahora sumando miembro a miembro:

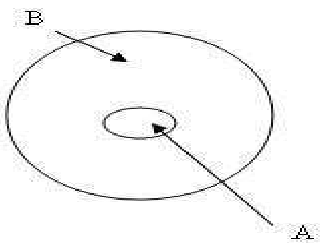
P(A) + P (B) = P (A^c ∩ B) + P(A^c ∩ B) + 2 P(A ∩ B) y ahora, restando en cada miembro P(A ∩ B)

P (A) + P (B) − P(A ∩ B) = P(A^c ∩ B) + P(A^c ∩ B) + P(A ∩ B) = P(A U B)

4. Sean A, B ∈ $\mathscr{A}$, tal que A ⊆ B, entonces, P(A) ≤ P (B).

Demo:

B = A U (B ∩ A^c) ⇒ P(B) = P(A) + P(B ∩ A^c) y como por definición P(B ∩ A^c) ≥ 0, resulta que P(B) ≥ P(A).



5. Sea A ∈ $\mathscr{A}$, entonces 0 ≤ P(A) ≤1.

Demo:

Que la P (A) ≥0, no es preciso justificarlo pues basta echar mano del ax. 1, y por otra parte, usando la propiedad 4) como es A ⊆ Ω ⇒debe ser P (A) ≤1, lo que prueba la prueba la propiedad 5.

2.3.2. Probabilidad Clásica o de Laplace

Se denomina Probabilidad Clásica o de Laplace a la definición de probabilidad dada por el cociente entre casos favorables entre posibles, es decir, dado un experimento la probabilidad del suceso A es el cociente entre el número de veces que se presente ese suceso entre el número de resultados posibles:

$$\boxed{P(A) = \frac{card(A)}{card(\Omega)}}$$

Esta definición de probabilidad es muy sencilla e inmediata en muchos casos, utilizándose en algunos casos técnicas de conteo
.

Ejemplo 21. En una urna hay 6 bolas de color verde, 3 bolas de color rojo y dos bolas de color amarillo. Si se extrae una bola al azar, calcular:

1. La probabilidad de ser amarilla. 2) Ídem. que no sea Roja. 3) Ídem. sea verde o roja

1) P(amarillo) = $\frac{2}{11}$; 2) P(no ser roja) = P(ser amarilla o verde) = $\frac{8}{11}$; P(verde o roja) = $\frac{9}{11}$

2.3.3. Definición Frecuentista

Consiste en realizar un experimento un número elevado de veces y observar las frecuencias relativas con que aparecen los sucesos, asignando como probabilidad de un suceso esta frecuencia relativa:

$$\boxed{P(A) = \lim_{n\to\infty} \frac{casosfavorables}{casostotales}}$$

2.4. Probabilidad Condicional

Si arrojamos un dado al aire y nos preguntan ¿cuál es la probabilidad de que salga un 5?, la respuesta es 1/6. Pues tenemos un solo caso favorable y 6 casos posibles, pero si nos preguntaran sobre la base de saber quel resultado dado es par ¿cuál es la probabilidad de que salga un 5? Entonces dicha probabilidad será 0, a diferencia de 1/6, como había sucedido inicialmente. El ejemplo nos ilustra sobre cómo, el hecho de saber que tal suceso ha ocurrido, afecta la probabilidad de otro. El estudio de la probabilidad de un suceso, en el conocimiento de que otro ha sido dado, conduce al cálculo de la Probabilidad Condicional.

Dado un suceso B con probabilidad no nula, se llama probabilidad de A condicionado a B a la probabilidad de que ocurra A sabiendo que ha ocurrido B, y se calcula como:

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

Si se realizara un experimento aleatorio que da como posibles resultados los sucesos A y B, entonces, si se sabe que B ha ocurrido, el número de casos favorables para A, se reduce al card(A∩B) y por otro lado, el número de casos posibles se reduce al card(B), ya que se sabe que este ha ocurrido. Por tanto dividiendo en numerador y denominador por el card (Ω), surge la fórmula recuadrada más arriba.

Propiedades:

1) P(A/B) ≥ 0 obvio por axioma 1.

2. P (Ω/B) = 1 obvio por axioma 2.

Una familia A_1, A_2,……., A_n de sucesos de **Ω**, es una partición de este, ⇔ verifican simultáneamente:

$$1)\ \bigcup_{1}^{n} A_i = \Omega \ \text{ y } \ 2)\ A_i \cap A_j = \emptyset \ \ \forall i \neq j$$

3) $P(\bigcup_{1}^{n} A_i) = \sum_{1}^{n} P(A_i / B)$, para toda familia $A_i \in \mathscr{A}$, que sea una partición de dicha σ Álgebra.

4. $P(A^c/B) = 1 - P(A/B)$.

Ejemplo 2.2.

En una ciudad el 40% de las personas tienen el pelo rubio, el 25% tiene los ojos azules y el 15% pelo rubio y ojos azules. Seleccionada una persona al azar, calcular la probabilidad de:

1. Que tenga el pelo rubio, si tiene los ojos azules.

Sean los sucesos: A=”pelo rubio”; B = “ojos azules”, por lo tanto se trata de calcular la probabilidad de que ocurra A, dado que ocurrió B, por tanto:

P(A/B) = 0,15/0,25 = 0,6

2. Que tenga los ojos azules, si tiene el pelo rubio. Acá es la Probabilidad inversa es decir P(B/A) = 0,15/0,4 = 0,375

3. Que tenga una de estas características, ojos azules o pelo rubio. En este caso se trata de calcular la P(A U B) = P(A) + P(B) – P(A ∩ B) = 0,4 + 0,25 − 0,15 = 0,5

2.5. Independencia de Sucesos

Dos sucesos A y B son independientes ⇔ la ocurrencia o no, de uno de ellos, no influye en el otro, por lo que P(A/B) = P(A) =) P(A ∩ B) / P(B) = P(A) =) P(A ∩ B) = P(A)P(B). Esta última, es decir que la probabilidad de la intersección sea el producto de las probabilidades es la condición natural para determinar si 2 sucesos son independientes.

Ejemplo 2.3.

Se arroja un dado al aire. Sea A : “sale Par” y B : “sale ≤4”, ¿son A y B sucesos independientes?

P(A) = ½ y P(B) = 2/3, por lo tanto P(A)*P(B) = 1/3; A ∩ B : “2 y 4” entonces P(A ∩ B) = 1/3 y los sucesos son independientes. Sea ahora C: “sale < 4”, ¿son A y C sucesos independientes?, como la P(C) = ½ entonces P(A)*P(C) = ¼, mientras que A ∩ C: “2” con P(A ∩ C) = 1/6 de donde los sucesos A y C NO son independientes.

Este resultado se puede extender a un mayor número de sucesos, por ejemplo si tenemos A1, . . . ,An, si estos son independientes =) P(A1 ∩ A2∩ . . .∩ An) = P(A1)P(A2) . . . P(An).

2.6. Teorema de la probabilidad total y Teorema de Bayes

1) Teorema de la probabilidad total: "Dados n sucesos A1, . . . ,An disjuntos, tales que formen una partición Ω, se cumple que:

$$\boxed{P(B) = \sum_{1}^{n} P(B / A_i) * P(A_i)}$$

Demo:

P(B) = P(B ∩ Ω) = P (B ∩ (U A_i) = P (U(B ∩ A_i)) = $\sum_{1}^{n} P($B ∩ A_i) = $\sum_{1}^{n} P(B/A_i) * P(A_i)$, que es lo que queríamos demostrar.

2) Teorema de Bayes

Se considera una partición A_i de Ω, sea B un suceso arbitrario, en ese caso P(A_i/B) = $\frac{P(A_i \cap B)}{P(B)}$ por definición, pero sustituyendo la P(B) por la expresión vista en el Teorema de la Probabilidad Total y teniendo en cuenta que por la conmutatividad de las intersecciones es P(A_i ∩ B) = P(B ∩ A_i) = P(A_i)*P(B/A_i), resulta la fórmula de Bayes:

$$\boxed{P(A_i / B) = \frac{P(A_i) * P(B / A_i)}{\sum_{1}^{n} P(B / A_i) * P(A_i)}}$$

Ejemplo 4.4

Consideramos el experimento consistente en elegir una moneda entre dos y lanzarla. La probabilidad de obtener cara con la primera moneda es 0,4 y con la segunda 0,7. Se pide:

1. La probabilidad de que el resultado del lanzamiento sea cruz. Supondremos que la probabilidad de escoger entre una moneda u otra es la misma 0,5, por lo que, aplicando el teorema de la probabilidad total, tendremos que: P(moneda A) = 0,5 P(cara/moneda A) = 0,4 P(cruz/moneda A) = 0,6. P(moneda B) = 0,5 P(cara/moneda B) = 0,7 P(cruz/moneda B) = 0,3. P(cruz) = P(moneda A)P(cruz/moneda A) + P(moneda B)P(cruz/moneda B) = 0,5 * 0,4 + 0,5 * 0,6 = 0,45.

2. Probabilidad que fuese lanzada la segunda moneda si se conoce que el resultado del lanzamiento ha sido cruz. Si usamos el teorema de Bayes, resulta:

P(moneda B/cruz) = P(moneda B)P(cruz/moneda B)/(P(moneda A)P(cruz/moneda A) + P(moneda B)P(cruz/moneda B) = 0,15/0,45 = 0,333

Parte 3) Variables Aleatorias

Cualquier experimento (por ejemplo lanzar una moneda) tiene asociado un conjunto de posibles resultados, llamado espacio muestral, no siempre de tipo cuantitativo (cara o cruz), de tal manera que nos sería útil cuantificar los resultados, para poder trabajar con ellos. Para esto surge el concepto de variable aleatoria.

Espacio probabilístico:"Se define un espacio probabilístico como el conjunto (Ω, $\mathscr{A}$, P)."

Variable aleatoria: "Dado un espacio probabilístico (Ω, $\mathscr{A}$, P) a la función X: Ω ⟶ R, de modo que si A ∈ Ω $\Rightarrow P(A) \in R$ y tal que $X^{-1}(-\infty, x) = \{A \in \Omega / X(A) \leq x\}$ se le denomina variable aleatoria. Tal vez no sea muy feliz la denominación utilizada porque en realidad, una variable aleatoria es una función, como se indica de Ω ⟶ R.

Una variable aleatoria sería la función definida sobre Ω que transforma los posibles resultados del experimento aleatorio, en valores reales. Por ejemplo, sea el experimento de lanzar un dado, una variable aleatoria sería aquella función: X(cara) = 0 y X(cruz) = 1, calculando las probabilidades como P(X = 0) y P(X = 1).

Se dice que una variable aleatoria es:

Discreta:"Cuando los valores que puede tomar son contables (finito o infinito numerable)."

Continua: "Cuando los valores son intervalos definidos sobre un conjunto infinito no numerable"

3.2. Función de probabilidad. Función de distribución de Probabilidad

Variable Aleatoria Discreta

Si la variable es de tipo discreta, esta se presentará con n valores x1, . . . , xn y P[X = xi] = pi con pi ≥ 0 y Σ (pi) = 1 ∀ i = 1, . . . n.
De tal forma que a {(xi, pi)} ∀ i = 1, . . . , n se le llama función de probabilidad o función cuantía de probabilidad.

Ejemplo 3.1. Se lanza una moneda al aire 3 veces. Sea X=número de caras en tras lanzamientos, obtener su f. c. p:
En este caso Ω = {CCC, CCN, C N C, NCC, NN C, NCN, C NN, NNN}, por lo que la f. c. p sería:

$$f(x) = \begin{cases} P(X=0) = 1/8 \\ P(X=1) = 3/8 \\ P(X=2) = 3/8 \\ P(X=3) = 1/8 \end{cases}$$

Igualmente, se define la función de distribución de la v.a. X como la probabilidad que la variable X sea menor o igual que x:

$F(X) : R \longrightarrow [0, 1]\ X \longrightarrow F(x) = P[X \leq x]$ cumpliendo que:

1. $F(-\infty) = 0$ y $F(\infty) = 1$.

2. Es una función no decreciente $xi \leq xj \Rightarrow F(xi) \leq F(xj)$.

3. Es continua a la derecha.

4. $P[a < x \leq b] = F(b) - F(a)$.

Ejemplo 3.2. La función de distribución del ejemplo 3.1 sería:

$$F(X) = \begin{cases} 0 \Leftrightarrow x < 0 \\ 1/8 \Leftrightarrow 0 \leq x < 1 \\ ½ \Leftrightarrow 1 \leq x < 2 \\ 7/8 \Leftrightarrow 2 \leq x < 3 \\ 1 \Leftrightarrow x \geq 3 \end{cases}$$

3.2.2. Variable Aleatoria Continua

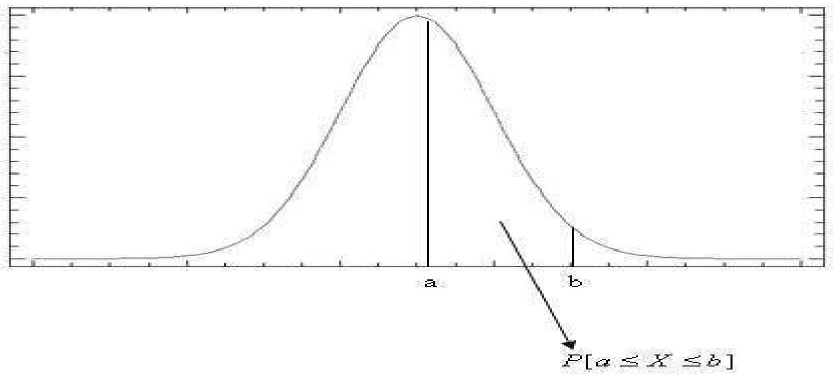
Si X es una Variable Aleatoria Continua sus probabilidades vienen definidas por la función de densidad, la cual se define como: $f : R \longrightarrow R+,\ x \longrightarrow f(x)$ cumpliendo:

1. $f(x) \geq 0 \quad \forall\, x / -\infty < x < \infty$.

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

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

Es decir las probabilidades no son más que el área de una función entre los valores a y b:



P[X = a] = $\int_a^a f(x)dx$ = 0. Se define entonces la función de Distribución F(x) = P(X$\leq$x), y como esta última probabilidad es $\int_{-\infty}^{x} f(t)dt$, entonces será F'(x) = f(x), eventualmente a diferencia de una constante.

Ejemplo 3.4. Sea X una v.a. continua con función de densidad dada por:

$$f(x) = \begin{cases} 0 \Leftrightarrow x < 0 \\ x \Leftrightarrow 0 \leq x < 1 \\ 2 - x \Leftrightarrow 1 \leq x < 2 \end{cases}$$

La función de distribución vendrá definida por trozos, de tal manera que:

Para x < 0 $\Rightarrow$F(x) = P(X<0) = $\int_{-\infty}^{0} 0dt = 0$;

Para 0 $\leq x < 1 \Rightarrow$F(x) = P(0<X<1) = 0 + $\int_0^x xdx = \frac{x^2}{2} = \frac{x^2}{2}$

Para 1 $\leq$x < 2 $\Rightarrow$ F(x) = P(1<X<2) = 0 + $\int_0^1 xdx + \int_1^x (2-x)dx = \left.\frac{x^2}{2}\right]_0^1 + \left[2x - \frac{x^2}{2}\right]_1^x$ = 2x - $\frac{x^2}{2}$ - 1

Para x $\geq$ 2 F(x) = P(X<2) = $\int_0^1 xdx + \int_1^2 (2-x)dx = \left.\frac{x^2}{2}\right]_0^1 + \left[2x - \frac{x^2}{2}\right]_1^2 = 1$

Por tanto resulta finalmente:

$$F(x) = \begin{cases} 0 \Leftrightarrow x < 0 \\ \frac{x^2}{2} \Leftrightarrow 0 \leq x < 1 \\ 2x - \frac{x^2}{2} - 1 \Rightarrow 1 \leq x < 2 \\ 1 \Rightarrow x \geq 2 \end{cases}$$

3.3 Características de una Variable Aleatoria

3.3.1. Esperanza Matemática

Se define la Esperanza Matemática como el valor promedio o esperado de una variable. Dependiendo de como sea la variable se calcula:

Variable Aleatoria Discreta:

$$\boxed{E[X] = \sum_{1}^{n} (x_i p_i)}$$

Variable Aleatoria Continua:

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

En ambos casos se verifica que:

1. E [aX + b] = aE [X] + b

2. Sean X e Y v.a., entonces E [X + Y] = E[X] + E[Y].

3. Sean X e Y v.a. independientes, E [XY] = E[X] * E[Y].

3.3.2. Varianza

La varianza es el momento centrado de segundo orden:

$$Var(X) = E[(X - E[X])^2] = E[X^2] - [E(X)]^2$$

y verifica que:

1. $Var[AX + B] = A^2 * Var[X]$.

2. $Var[X] \geq 0$.

Ejemplo 3.5. Calcular la media y la varianza del ejemplo 3.1.

En primer lugar la media sería:

$$E[X] = \sum_{0}^{3} x_i\, p_i = 0*1/8 + 1*3/8 + 2*3/8 + 3*1/8 = 12/8 = 1,5$$

Para calcular la varianza, hemos de obtener primero el momento de orden dos,

$E[X^2] = \sum_{0}^{3} (x_i^2\, p_i) = 0*1/8 + 1*3/8 + 4*3/8 + 9*1/8 = 24/8 = 3$; como E(X) = 1,5 $\Rightarrow$ $[E(X)]^2 = 2,25$

$Var(X) = E[X^2] - E[X]^2 = 3 - 2,25 = 0,75$

Parte 4) Funciones de Probabilidad

4.1 Probabilidad de Bernoulli

Un experimento aleatorio que arroje sólo 2 tipos de resultados, que podríamos denominar éxito y fracaso, con valores 1 y 0, tiene una variable aleatoria discreta de valores 0 y 1. Supongamos que la P(X=1) = p y P(X=0) = q, entonces será p = 1 – q y se dice que dicha función de probabilidad es de Bernoulli, de parámetro p, es decir B(p).

Esperanza Matemática y Varianza de una función de Probabilidad B(p)

$$E(X) = \sum_{1}^{2} x_i\, p_i = 1*p + 0*q = p$$

$$Var(X) = \sum_{1}^{2} x_i^2\, p_i - p^2 = p - p^2 = p*(1 - p)$$

4.2 Probabilidad Binomial

Sea un experimento que determina una variable aleatoria B(p). Reiteremos n veces dicho experimento, entonces tendremos n variables aleatorias B(p): $X_1, X_2, X_3, \ldots, X_n$, las cuales provienen de sucesivos experimentos que garantizan la independencia de los sucesos arrojados por cada uno de ellos. Definamos una nueva variable aleatoria que registre el número de éxitos logrados, al cabo de las n repeticiones del B(p). En ese caso supongamos que dicho éxito con probabilidad p, se verificó en r ocasiones, es claro por tanto que en las restantes n – r casos, el resultado dado fue fracaso. Se trata entonces de determinar la probabilidad de la intersección de n resultados independientes, por lo que su probabilidad estará dada, (por la independencia), por el producto de las probabilidades: p*p*p*…*p, en casos y por (1 – p)*(1 – p)*….*(1 – p) en los n-r resultados restantes. De allí se tiene la expresión: $p^r * (1-p)^{n-r}$, ahora bien, esto será así para una determinada forma en que se den estos resultados, pero ¿Cuántas formas se pueden dar para hallar r éxitos en una en-upla? Bueno, por lo que ya vimos, en C_r^n, así que finalmente, la probabilidad de que en n repeticiones de un experimento B(p), se obtengan r éxitos será: $C_r^n\, p^r (1-p)^{n-r}$, o mejor aún:

$$\boxed{P(X=r) = C_r^n\, p^r (1-p)^{n-r}}$$

Que se denomina función de Probabilidad Binomial B(n,p)

Función de Distribución de Probabilidad Binomial

$$F(x) = P(X \le x) = \begin{cases} 0 & \Leftrightarrow x < 0 \\ \sum_{r=0}^{x} (C_r^n p^r (1-p)^{n-r}) & \Leftrightarrow 0 \le x < n \\ 1 & \Leftrightarrow x \ge n \end{cases}$$

Esperanza Matemática y Varianza de una función de Probabilidad Binomial B(n,p)

E(X) = E($X_1, X_2, X_3, ..., X_n$) = por la propiedad 2 ya vista para toda esperanza matemática, a la suma de la esperanza de cada una de las n variables X_i, que como cada una de estas es de B(p), tendrá como valor esperado p, o sea:

E(X) = E($X_1, X_2, X_3, ..., X_n$) = $p + p + p + ... + p$, con n sumandos p, $\Rightarrow$ $\boxed{E(X) = np}$

Análogamente se determina que la varianza es $\boxed{\text{Var(X) = n*p*q}}$

4.3 Probabilidad Hipergeométrica

Se tiene una población de N datos, los cuales pueden agruparse en 2 subconjuntos disjuntos y que su unión coincida con la población total, es decir una familia de 2 subconjuntos de la población que sea, a su vez, una partición del conjunto inicial. Sean N1 y N2 dichos subconjuntos. Si de la población total se extrae sin reposición, una muestra de r elementos, la probabilidad de que h de esos r pertenezcan al subconjunto N1 viene dada por la fórmula de Laplace, es decir el nº de casos favorables sobre el nº de casos posibles. El nº de casos favorables serán las combinaciones de N1 tomadas de h, multiplicadas por las combinaciones de N2 tomadas de r-h y por otra parte, el nº de casos posibles serán las combinaciones de N tomadas de r, por tanto resulta si X mide el nº de elementos de N1, que están presentes en la muestra r, la siguiente función de Probabilidad, denominada Hipergeométrica.

$$\boxed{P(X = r) = \frac{C_h^{N1} * C_{r-h}^{N2}}{C_r^N}}$$

Ejemplo 4.1:

En una urna o recipiente hay un total de *10* objetos, entre los cuales hay una cantidad *3* de objetos que son defectuosos, si se seleccionan de esta urna *4* objetos al azar, y sin reemplazo, ¿cuál es la probabilidad de obtener *2* objetos defectuosos?

Acá N = 10, N1 son los defectuosos y son 3, así que N2 son los NO defectuosos y son 7, la muestra r es de 4 objetos y h = 2, por tanto:

$$P(X=2) = \frac{C_2^3 * C_2^7}{C_4^{10}} = \frac{3*21}{210} = 0{,}3$$

4.4 Probabilidad Geométrica o de Pascal

Sea A un suceso de P(A) = p y sea X la variable aleatoria que expresa el número de fracasos sucesivos que se obtienen, en una reiteración de pruebas de Bernoulli, hasta que ocurra A por primera vez. Decimos que una variable aleatoria X sigue una distribución geométrica de parámetro p, si tiene una función de probabilidad dada por:

$$\boxed{P(X=r)=(1-p)^r * p}$$

Distribución de Probabilidad G(p):

$$F(x) = P(X<r) = \begin{cases} 0 \quad \Leftrightarrow x < 0 \\ \sum_0^r (1-p)^r * p \quad \Leftrightarrow x \geq 0 \end{cases}$$

Esperanza matemática de una X con G(p): $E(X) = \frac{1-p}{p}$

Varianza de una X con G(p): $Var(X) = \frac{1-p}{p^2}$

4.5 Probabilidad de Poisson

Esta distribución se utiliza para modelizar fenómenos del tipo:

1. Número de ocurrencias de un determinado suceso en intervalos fijos de tiempo.

2. Número medio de veces que ocurre un éxito en una unidad de medida de tiempo constante. A la variable X que mide este tipo de ocurrencias, es decir, X = “número de éxitos ocurridos en un determinado periodo de tiempo”, se dice que sigue una distribución de Poisson de parámetro λ, con λ > 0 (λ sería el número medio de ocurrencias), siendo su función de Probabilidad, la siguiente:

$$\boxed{P(X=r) = \frac{e^{-\lambda} * \lambda^r}{r!}}$$

Función de Distribución de Probabilidad:

$$F(X=r) = P(X \leq r) = \sum_{k=1}^{r} \frac{e^{-\lambda} * \lambda^{k}}{k!}$$

Esperanza Matemática E(X) = λ; Varianza Var(X) = λ

4.6 Funciones de Probabilidad de Variable Continua

4.6.1 Función de Probabilidad Normal

Al iniciar el análisis estadístico de una serie de datos, y después de la etapa de detección y corrección de errores, un primer paso consiste en describir la distribución de las variables estudiadas y, en particular, de los datos numéricos. Además de las medidas descriptivas correspondientes, el comportamiento de estas variables puede explorarse gráficamente de un modo muy simple. Consideremos, como ejemplo, los datos de la Figura 1a, que muestra un **histograma** de la tensión arterial sistólica de una serie de pacientes isquémicos ingresados en una unidad de cuidados intensivos. Para construir este tipo de gráfico, se divide el rango de valores de la variable en intervalos de igual longitud, representando sobre cada intervalo un rectángulo con área proporcional al número de datos en ese rango[1]. Uniendo los puntos medios del extremo superior de las barras, se obtiene el llamado **polígono de frecuencias**. Si se observase una gran cantidad de valores de la variable de interés, se podría construir un histograma en el que las bases de los rectángulos fuesen cada vez más pequeñas, de modo que el polígono de frecuencias tendría una apariencia cada vez más suavizada, tal y como se muestra en la Figura 1b. Esta curva suave "asintótica" representa de modo intuitivo la distribución teórica de la característica observada. Es la llamada función de densidad.

Una de las distribuciones teóricas mejor estudiadas en los textos de bioestadística y más utilizada en la práctica es la **Distribución Normal**, también llamada **Distribución Gaussiana**. Su importancia se debe fundamentalmente a la frecuencia con la que distintas variables asociadas a fenómenos naturales y cotidianos siguen, aproximadamente, esta distribución. Caracteres morfológicos (como la talla o el peso), o psicológicos (como el cociente intelectual) son ejemplos de variables de las que frecuentemente se asume que siguen una distribución normal. No obstante, y aunque algunos autores han señalado que el comportamiento de muchos parámetros en el campo de la salud puede ser descrito mediante una distribución normal, puede resultar incluso poco frecuente encontrar variables que se ajusten a este tipo de comportamiento.

El uso extendido de la distribución normal en las aplicaciones estadísticas puede explicarse, además, por otras razones. Muchos de los procedimientos estadísticos habitualmente utilizados asumen la normalidad de los datos observados. Aunque muchas de estas técnicas no son demasiado sensibles a desviaciones de la normal y, en general, esta hipótesis puede obviarse cuando se dispone de un número suficiente de datos, resulta recomendable contrastar siempre si se puede asumir o no una distribución normal. La simple exploración visual de los datos puede sugerir la forma de su distribución. No obstante, existen otras medidas, gráficos de normalidad y contrastes de hipótesis que pueden ayudarnos a decidir, de un modo más riguroso, si la muestra de la que se dispone procede o no de una distribución normal. Cuando los datos no sean normales, podremos o bien transformarlos o emplear otros métodos estadísticos que no exijan este tipo de restricciones (los llamados métodos no paramétricos).

A continuación se describirá la distribución normal, su ecuación matemática y sus propiedades más relevantes, proporcionando algún ejemplo sobre sus aplicaciones a la inferencia estadística.

Figura 1. Histograma de los valores de tensión arterial sistólica para dos muestras de pacientes isquémicos ingresados en una unidad de cuidados intensivos.

Figura 1a.- Valores de tensión arterial sistólica en una muestra de 1000 pacientes isquémicos ingresados en UCI.

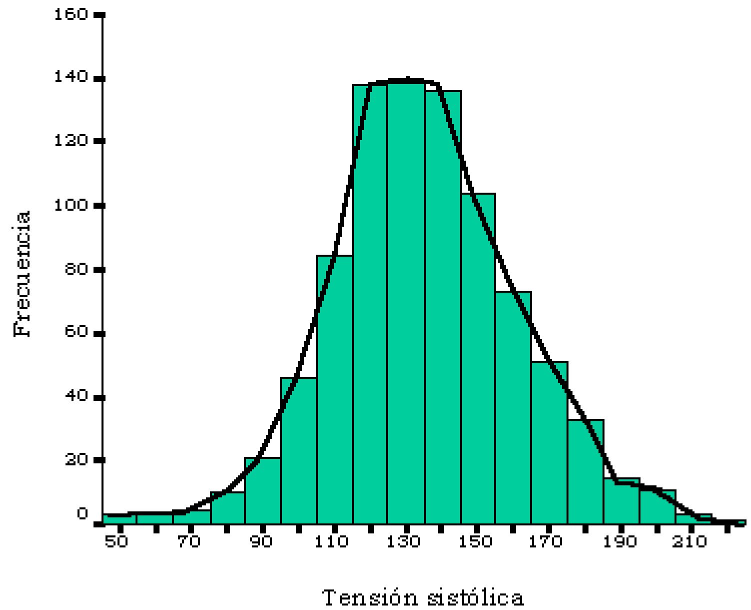
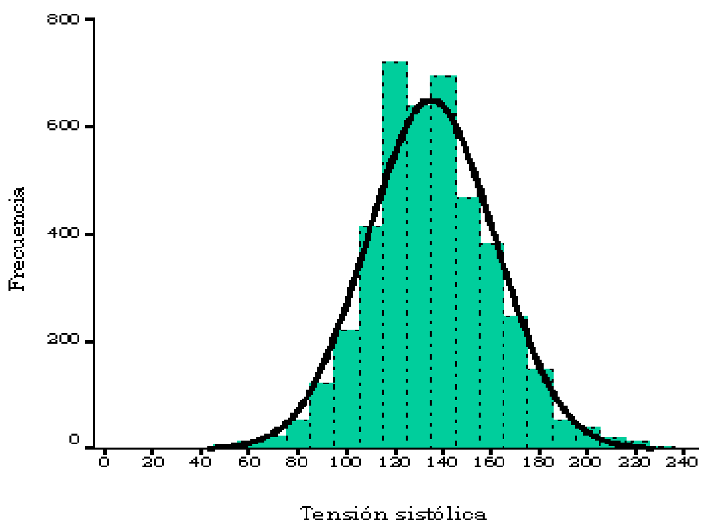


Figura 1b.- Valores de tensión arterial sistólica de una muestra de 5000 pacientes ingresados en UCI.



La Función de Probabilidad Normal

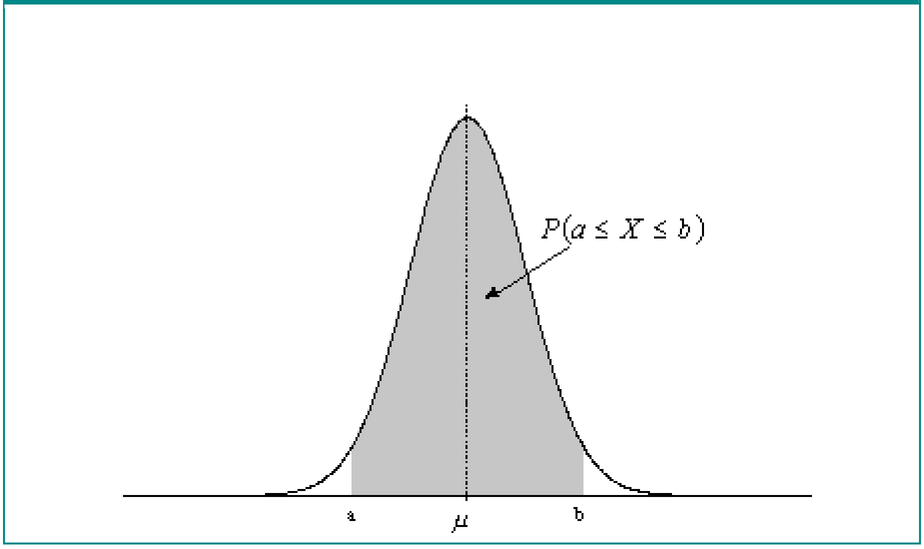
La distribución normal fue reconocida por primera vez por el francés Abraham De Moivre (1667-1754). Posteriormente, Carl Friedrich Gauss (1777-1855) elaboró desarrollos más profundos y formuló la ecuación de la curva; de ahí que también se la conozca, más comúnmente, como la **"campana de Gauss"**. La distribución de una variable normal está completamente determinada por dos parámetros, su media y su desviación estándar, denotadas generalmente por μ y σ. Con esta notación, la densidad de la normal viene dada por la ecuación:

$$\boxed{f(x) = \frac{1}{\sigma\sqrt{2\pi}} * e^{\frac{-1}{2}\left[\frac{(x-\eta)}{\sigma}\right]^2}}$$

que determina la curva en forma de campana que tan bien conocemos, ver figura 2. Así, se dice que una característica X sigue una distribución normal de media μ y varianza σ^2, y se denota como $X \approx N(\mu,\sigma)$, si su función de densidad viene dada por la expresión recuadrada.

Al igual que ocurría con un histograma, en el que el área de cada rectángulo es proporcional al número de datos en el rango de valores correspondiente si, tal y como se muestra en la figura 2, en el eje horizontal se levantan perpendiculares en dos puntos *a* y *b*, el área bajo la curva delimitada por esas líneas indica la probabilidad de que la variable de interés, *X*, tome un valor cualquiera en ese intervalo. Puesto que la curva alcanza su mayor altura en torno a la media, mientras que sus "ramas" se extienden asintóticamente hacia los ejes, cuando una variable siga una distribución normal, será mucho más probable observar un dato cercano al valor medio que uno que se encuentre muy alejado de éste.

Figura 2. Gráfica de una distribución normal y significado del área bajo la curva.



Propiedades de la distribución normal:

La distribución normal posee ciertas propiedades importantes que conviene destacar: Tiene una única moda, que coincide con su media y su mediana. La curva normal es asintótica al eje de abscisas. Por ello, cualquier valor entre $-\infty$ y $+\infty$ es teóricamente posible. El área total bajo la curva es, por tanto, igual a 1.

Es simétrica con respecto a su media μ. Según esto, para este tipo de variables existe una probabilidad de un 50% de observar un dato mayor que la media, y un 50% de observar un dato menor.

La distancia entre la línea trazada en la media y el punto de inflexión de la curva es igual a una desviación típica (σ). Cuanto mayor sea σ, más aplanada será la curva de la densidad.
El área bajo la curva comprendida entre los valores situados aproximadamente a dos desviaciones estándar de la media es igual a 0.95. En concreto, existe un 95% de posibilidades de observar un valor comprendido en el intervalo $(\mu - 1.96\sigma; \mu + 1.96\sigma)$.

La forma de la campana de Gauss depende de los parámetros μ y σ. La media indica la posición de la campana, de modo que para diferentes valores de μ la gráfica es desplazada a lo largo del eje horizontal. Por otra parte, la desviación estándar determina el grado de apuntamiento de la curva. Cuanto mayor sea el valor de σ, más se dispersarán los datos en torno a la media y la curva será más plana. Un valor pequeño de este parámetro indica, por tanto, una gran probabilidad de obtener datos cercanos al valor medio de la distribución.
Como se deduce de este último apartado, no existe una única distribución normal, sino una familia de distribuciones con una forma común, diferenciadas por los valores de su media y su varianza. De entre todas ellas, la más utilizada es la **distribución normal estándar**, que corresponde a una distribución de media 0 y varianza 1. Así, la expresión que define su densidad se puede obtener de la expresión recuadrada líneas arriba, resultando:

$$\boxed{f(z) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-z^2}{2}\right); \quad -\infty < z < \infty}$$

Es importante conocer que, a partir de cualquier variable *X* que siga una distribución $N(\mu, \sigma)$, se puede obtener otra característica *Z* con una distribución normal estándar, sin más que efectuar la transformación:

$$** \quad \boxed{Z = \frac{X - \mu}{\sigma}}$$

Esta propiedad resulta especialmente interesante en la práctica, ya que para una distribución $N(0,1)$ existen tablas publicadas , como la que se explicita más adelante, a partir de las que se puede obtener de modo sencillo la probabilidad de observar un dato menor o igual a un cierto valor z, y que permitirán resolver preguntas de probabilidad acerca del comportamiento de variables de las que se sabe o se asume que siguen una distribución aproximadamente normal.

Ejemplo:

Consideremos, por ejemplo, el siguiente problema: supongamos que se sabe que el peso de los sujetos de una determinada población sigue una distribución aproximadamente normal, con una media de 80 Kg y una desviación estándar de 10 Kg. ¿Podremos saber cuál es la probabilidad de que una persona, elegida al azar, tenga un peso superior a 100 Kg?

Denotando por *X* a la variable que representa el peso de los individuos en esa población, ésta sigue una distribución $N(80,10)$. Si su distribución fuese la de una normal estándar podríamos utilizar la Tabla para calcular la probabilidad que nos interesa. Como éste no es el caso, resultará entonces útil transformar esta característica, según el cambio de variable dado por la ecuación **

$$\boxed{Z = \frac{X - 80}{10}}$$

para poder utilizar dicha tabla. Así, la probabilidad que se desea calcular será:

$$P(X > 100) = P\left(Z > \frac{100 - 80}{10}\right) = P(Z > 2)$$

Como el área total bajo la curva es igual a 1, se puede deducir que:

$$P(Z > 2) = 1 - P(Z \leq 2)$$

Esta última probabilidad puede ser fácilmente obtenida a partir de la Tabla, resultando ser $P(Z \leq 2) = 0.9772$. Por lo tanto, la probabilidad buscada de que una persona elegida aleatoriamente de esa población tenga un peso mayor de 100 Kg , es de 1–0.9772=0.0228, es decir, aproximadamente de un 2.3%.
De modo análogo, podemos obtener la probabilidad de que el peso de un sujeto esté entre 60 y 100 Kg:

$$P(60 \leq X \leq 100) = P\left(\frac{60 - 80}{10} \leq Z \leq \frac{100 - 80}{10}\right) = P(-2 \leq Z \leq 2)$$

De la figura 2, tomando a=-2 y b=2, podemos deducir que:

$$P(-2 \leq Z \leq 2) = P(Z \leq 2) - P(Z \leq -2)$$

Por el ejemplo previo, se sabe que $P(Z \leq 2) = 0.9772$. Para la segunda probabilidad, sin embargo, encontramos el problema de que las tablas estándar no proporcionan el valor de $P(Z \leq z)$ para valores negativos de la variable. Sin embargo, haciendo uso de la simetría de la distribución normal, se tiene que:

$$P(Z \leq -2) = P(Z \geq 2) = 1 - P(Z \leq 2) = 1 - 0.9772 = 0.0228$$

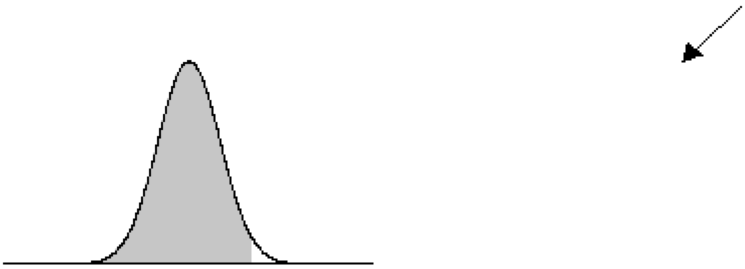
Finalmente, la probabilidad buscada de que una persona elegida al azar tenga un peso entre 60 y 100 Kg., es de 0.9772-0.0228=0.9544, es decir, aproximadamente de un 95%.

No obstante, es fácil observar que este tipo de situaciones no corresponde a lo que habitualmente nos encontramos en la práctica. Generalmente no se dispone de información acerca de la distribución teórica de la población, sino que más bien el problema se plantea a la inversa: a partir de una muestra extraída al azar de la población que se desea estudiar, se realizan una serie de mediciones y se desea extrapolar los resultados obtenidos a la población de origen. En un ejemplo similar al anterior, supongamos que se dispone del peso de n=100 individuos de esa misma población, obteniéndose una media muestral de $\bar{X}=75$ Kg, y una desviación estándar muestral $S=12$ Kg, querríamos extraer alguna conclusión acerca del valor medio real de ese peso en la población original. La solución a este tipo de cuestiones se basa en un resultado elemental de la teoría estadística, el llamado teorema central del límite. Dicho axioma viene a decirnos que las medias de muestras aleatorias de cualquier variable siguen ellas mismas una distribución normal con igual media que la de la población y desviación estándar la de la población dividida por $\sqrt{n}$.

En nuestro caso, podremos entonces considerar la media muestral $\bar{X} \approx N\left(\mu, \sigma/\sqrt{n}\right)$, con lo cual, a partir de la propiedad (iii) se conoce que aproximadamente un 95% de los posibles valores de $\bar{X}$ caerían dentro del intervalo $\left(\mu - \frac{1.96\sigma}{\sqrt{n}}; \mu + \frac{1.96\sigma}{\sqrt{n}}\right)$. Puesto que los valores de μ y σ son desconocidos, podríamos pensar en aproximarlos por sus análogos muestrales, resultando $\left(78 - \frac{1.96 \times 12}{\sqrt{100}}; 78 + \frac{1.96 \times 12}{\sqrt{100}}\right) = (75.6; 80.3)$. Estaremos, por lo tanto, un 95% seguros de que el peso medio real en la población de origen oscila entre 75.6 Kg y 80.3 Kg. Aunque la teoría estadística subyacente es mucho más compleja, en líneas generales éste es el modo de construir un intervalo de confianza para la media de una población.

Tabla 1. Áreas bajo la curva normal estándar. Los valores de la tabla que no se muestran en negrita representan la probabilidad de observar un valor menor o igual a z. La cifra entera y el primer decimal de z se buscan en la primera columna, y el segundo decimal en la cabecera de la tabla.

$$P(Z \le z) = \int_{-\infty}^{z} f(t)dt$$



Segunda cifra decimal del valor de z

z	**0.00**	**.01**	**.02**	**.03**	**.04**	**.05**	**.06**	**.07**	**.08**	**.09**
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830

1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.4878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998