DISTRIBUCIONES TEÓRICAS DE PROBABILIDAD

Variable aleatoria. Funciones de probabilidad y de densidad de probabilidad. Función de distribución. Esperanza y varianza. Modelos para variables discretas: Binomial y Poisson. Modelo para variables continuas: Normal. Convergencias.

INTRODUCCIÓN

“all the world believes it (Gaussian distribution) firmly, because the mathematicians imagine that it is a fact of observation, and the observers that it is a theorem of mathematics” - Lippmann a Poincaré, 1912[1]

A. El concepto de DISTRIBUCIÓN se refiere al reparto de los individuos de una población según una característica. Supongamos que la característica es el sexo (mujer vs hombre). Si en la población hay tantas personas que son mujeres como hombres se dice que la distribución de sexo en la población es uniforme, y por tanto la probabilidad de que un individuo seleccionado al azar sea mujer vale 0,5:

Uniforme: P (M) = 0,5 ; P (H) = 0,5

Si el reparto no es uniforme, entonces la probabilidad de que el individuo sea mujer valdrá diferente de 0,5, por ejemplo:

No-uniforme: P (M) = 0,7 ; P (H) = 0,3

Sea como sea el reparto, al final la suma de todas las probabilidades ha de ser siempre igual a uno:

Suma Total = P(M) + P(H) = 1

B. El reparto de individuos según una característica depende de la población que se está estudiando. Por ejemplo, la distribución por sexos (hombres y mujeres) es

[1] En: Encyclopedia of Statistical Sciences. Vol 4: Laws of error. Ver bibliografía.

diferente en distintos colectivos profesionales. Lo mismo ocurre con características numéricas como la altura, cuya distribución cambia entre población joven y adulta.

C. Hay que distinguir ente distribución de frecuencias y distribución de probabilidad. La primera es un reparto empírico observado en una colección de datos. El histograma es una representación gráfica de la distribución de frecuencias observadas. Una distribución de probabilidad es un reparto teórico de la población y por tanto es una función matemática. La representación gráfica será una curva como por ejemplo la campana de Gauss.

D. En teoría de probabilidad se usa el concepto de variable aleatoria como una manera de conseguir que cualquier característica estudiada sean numérica.

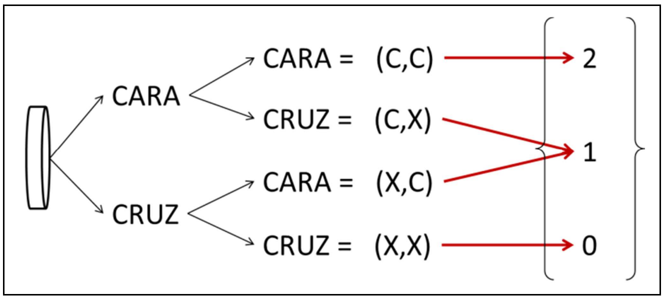
E = “seleccionar al azar una persona”
S = {Hombre , Mujer}
VA: Hombre = 0 ; Mujer = 1

E. VARIABLE ALEATORIA, VA, es una función matemática que asigna diferentes números reales a cada resultado de una experiencia aleatoria.

VA: espacio muestral (S) $\rightarrow$ conj. Reales (R)

F. Ejemplo:

E = “lanzar al aire una moneda 2 veces”
S = {(c,c); (c,x); (x,c); (x,x)}
VA = “contar el número de caras”



G. RECORRIDO de una VA: Conjunto R de todos los valores posibles de la variable. En el ejemplo anterior R = {0, 1, 2}.

H. La intención al definir una V.A. es poder convertir cualquier espacio muestral en un conjunto de números. Esto se puede hacer de distintas maneras. Ejemplos:

- Asignación arbitraria
  - E = “lanzar una moneda”
  - S = {cara, cruz}
  - VA=”valor de la cara obtenida: c=0; x=1"
  - R = {0,1}
- Asignación identidad
  - E = “lanzar un dado de 6 caras”
  - S = {1, 2, 3, 4, 5, 6}
  - VA= “puntos de la cara”
  - R = {1, 2, 3, 4, 5, 6}

- Asignación por medida

E = "extraer al azar una persona de una población"

S = {Pepe, Luís, Marta, Carlos,......, Ana}

VA="altura de una persona"

R = {148,152.5, 167.3,...,190.4}

I. Las variables aleatorias se identifican con letras mayúsculas- X, Y, Z,…- y los valores particulares con letras minúsculas- x, y, z, … Ejemplo:

E = "extraer al azar una persona de una población"
Resultado posible = Miguel
Variable aleatoria = Altura = X
Valor de X en Miguel = Altura de Miguel = X(Miguel)= 173,6 cm = x

J. Sobre una experiencia aleatoria con un espacio muestral dado se puede definir más de una variable aleatoria, cada una con un recorrido distinto. Ejemplo:

E = "lanzar dos monedas"
S = {(c,c); (c,x); (x,c); (x,x)}; c=0, x=1
X = "contar número de caras"; R = {0, 1, 2}
Y = "sumar los dos resultados"; R = {0, 1, 2}
Z = "multiplicar los dos resultados": R = {0,1}
W = "sumar los dos resultados, la 2ªtirada vale doble"; R = {0, 1, 2, 3}

K. Tipos de variables aleatorias

- DISCRETA ("discontinua"). Variable con un conjunto de resultados limitados (recorrido finito o infinito numerable).
  X = número de piezas dentales con caries
  Y = número de urgencias atendidas en Barcelona en un mes
- CONTINUA. Variable con un conjunto ilimitado de resultado, es decir, el recorrido es infinito. Ejemplos:
  X = la altura de un individuo
  Y = el salario de un trabajador

FUNCIONES DE PROBABILIDAD Y DE DENSIDAD DE PROBABILIDAD

A. La teoría de la probabilidad tuvo un origen histórico anterior al de estadística. Tradicionalmente se afirma que empezó en Francia en 1654, a raíz de la correspondencia epistolar entre dos matemáticos (Pascal y Fermat) que debatían sobre la división de apuestas en los juegos de azar[2].

B. En los juegos de azar de los casinos las probabilidades bajas se recompensan con premios más altos. En el juego "lanzar dos dados de seis caras y sumar los puntos de ambos" la probabilidad de obtener un siete es mayor (6/36) que la de obtener un dos

[2] En: Encyclopedia of Statistical Sciences. Vol 7: Probability, History of. Ver bibliografía

(1/36) ya que hay más combinaciones para sumar siete. Así apostar por dos (“snake eyes”) es más arriesgado, pero puede ser más lucrativo.

C. Una distribución teórica de probabilidad propone un reparto de los valores de una variable aleatoria en una población. A los valores más abundantes se les asigna mayor probabilidad porque tienen mayor expectativa de aparecer al realizar la experiencia aleatoria que los valores más escasos.

D. El reparto de probabilidad de una variable discreta se define mediante la FUNCIÓN DE PROBABILIDAD que asigna a cada valor posible de la variable aleatoria un número p en el intervalo [0,1]:

FP: $x \rightarrow P(x) = p$ si x está en el recorrido

$x \rightarrow P(x) = 0$ si el valor x no está recorrido

Dos propiedades importantes son:

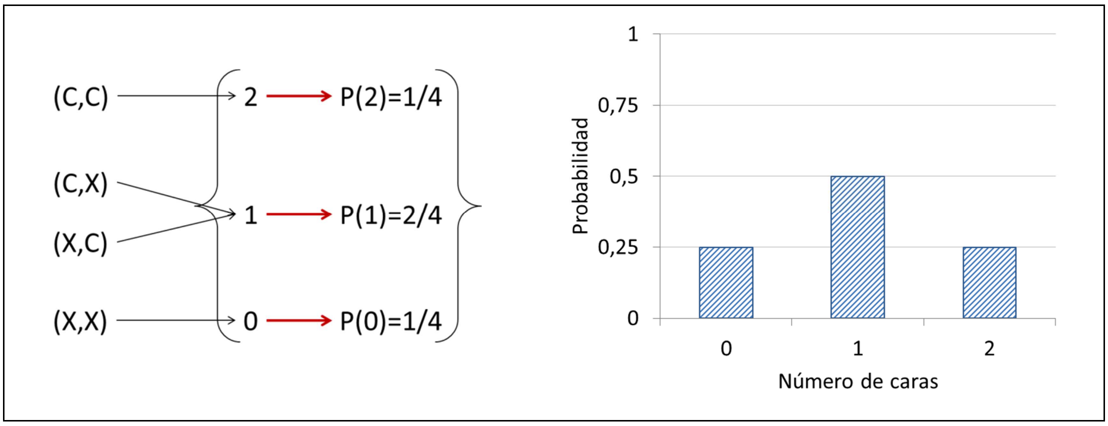
- Los valores p son siempre positivos o nulos: $p \geq 0$
- La suma de todas las p’s ha de ser igual a uno: $\Sigma p = 1$

E. Ejemplo:

E = “lanzar al aire una moneda 2 veces”

S = {(c,c); (c,x); (x,c); (x,x)}

VA = X= “contar el número de caras”



F. Problema. ¿Cuál es la función de probabilidad del juego “Chuck-a-luck”? [Un jugador escoge un número entre 1 y 6. Se lanzan 3 dados. Si el número sale en los tres dados el jugador ganará 3$, si aparece sólo en dos ganará 2$ y si aparece en sólo uno ganará 1$, pero si no aparece en ninguno dado tendrá que pagar 1$]. ¿Es justo?

G. En las variables aleatorias continuas no se puede usar la función de probabilidad porque la probabilidad en un punto x, sea cual sea, es siempre cero[3]:

$$P(X=x) = 0$$

En consecuencia en variables continuas sólo tiene sentido hablar de probabilidad referida a intervalos, que pueden ser tan pequeños como se quiera:

$P(X \in [a, b]) \neq 0$; siendo a y b números reales cualesquiera distintos

[3] Una forma intuitiva de justificarlo es partir de la idea de que el recorrido es infinito, por tanto repartir 1 entre todos los valores del recorrido queda así: $\lim(R \rightarrow +\infty) \left[{}^{1}/_{R}\right] = 0$

H. Para explicar el reparto de probabilidad de una variable continua se utiliza la FUNCIÓN DE DENSIDAD DE PROBABILIDAD (f) que es una curva con las siguientes propiedades:

- En cualquier punto de la recta real la función es positiva o cero, pero sin restricción de ser menor o igual que uno:

$$f(x) \geq 0\,; \quad \forall x\, real$$

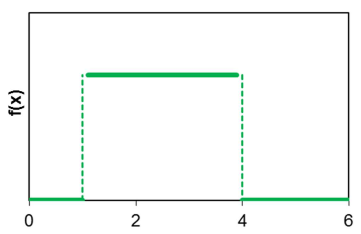
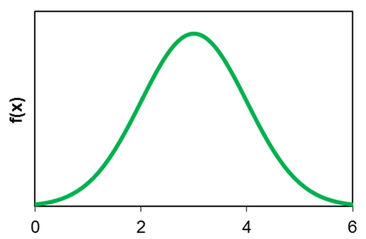
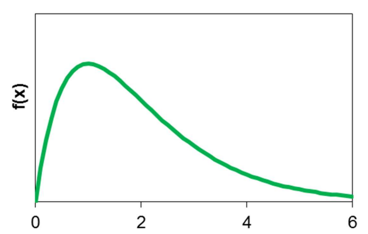
- El área total bajo la curva de f vale uno:

$$\int_{-\infty}^{+\infty} f v\, dv = 1$$

La función f no es una probabilidad. Hay que integrarla para obtener la probabilidad.

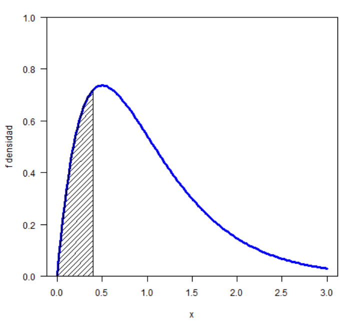
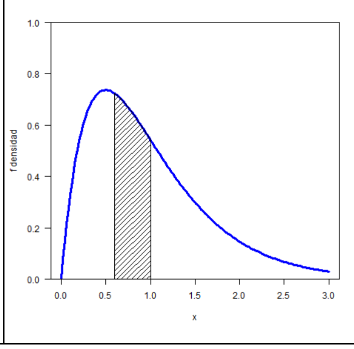
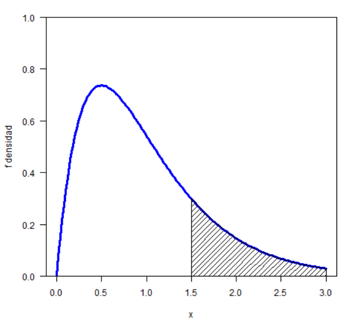
I. Existen muchas curvas que cumplen los criterios para ser función de densidad, por ejemplo la que corresponde a la distribución rectangular (continua uniforme):

$$f(x) = \begin{cases} (b-a)^{-1}, & a < x < b \\ 0\ , & en\ otro\ caso \end{cases}$$

uniforme	no-uniforne simétrica	no-uniforme asimétrica
		

J. La función de densidad de probabilidad servirá para calcular la probabilidad en un intervalo resolviendo la integral en ese intervalo.

$$P(a < X \leq b) = \int_a^b f(x)dx \quad \leq 1$$

(a,b) = (-∞, 0.4)	(a,b) = (0.6, 1.0)	(a,b) = (1.0, +∞)
		

A. La FUNCIÓN de DISTRIBUCIÓN (F) de una variable aleatoria nos informa de la probabilidad acumulada por debajo de un valor real t. La definición es la misma tanto para variables discretas como continuas:

$$F:\ t \quad \rightarrow \quad F(t) = P(X \leq t)$$

Las propiedades son:

- Está acotada en Y: el mínimo es cero y el máximo uno:

  $0 \leq F(t) \leq 1$ para todo t real
- Es no-decreciente: al desplazarnos por el eje X hacia la derecha la función crece o se mantiene constante:

  $F(a) \leq F(b)$ si $a < b$, para todo a y b real
- Es siempre continua por la derecha (acercarse a un punto desde la derecha). Por la izquierda puede ser continua o discontinua.
- La probabilidad en un intervalo cualquiera (a, b) se obtiene por diferencia de funciones de distribución:

  $P(a < X \leq b) = F(b) - F(a)$ si $a < b$, para todo a y b real

B. Formalmente las variables aleatorias se clasifican en discretas o continuas por su función de distribución. Las primeras presentan puntos de discontinuidad y las segundas son continuas.

C. En las variables discretas la función de distribución para t se obtiene por sumas de probabilidades asociadas a valores de X inferiores a t. La representación gráfica tiene forma de escalera.

$$F(t) = \Sigma\, p(x) \quad ; \quad x \leq t$$

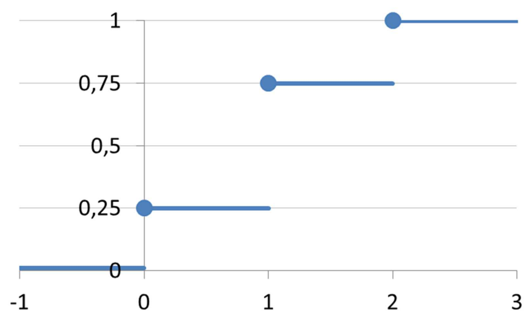
D. Ejemplo

E = “lanzar al aire una moneda 2 veces”

S = {(c,c); (c,x); (x,c); (x,x)}

VA = X = “contar el número de caras”

t	x	P(x)	F(t)
< 0	-	-	0
0 a 1	0	0,25	0,25
1 a 2	1	0,50	0,75
>=2	2	1,0	1,0



E. En las variables continuas la función de distribución para t se obtiene integrando la función de densidad de probabilidad desde menos infinito (-∞) hasta el valor t. La representación gráfica tiene forma en “s” (escalones mínimos).

$$F(t) = \int_{-\infty}^{t} f(x)\, dx$$

[nota: ergo, la función de densidad es la derivada de la función de distribución]

F. La probabilidad acumulada F corresponde al área de la cola inferior de la función de densidad.

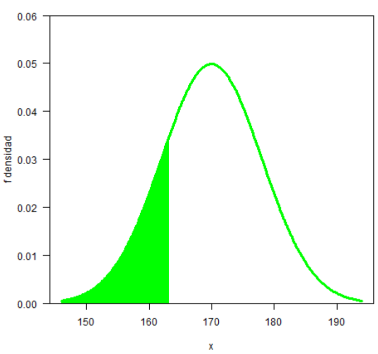
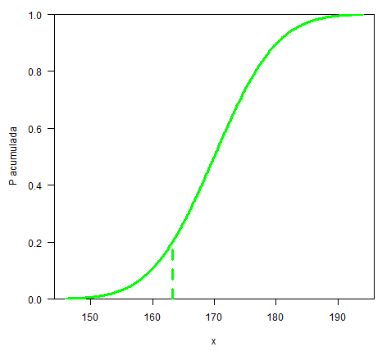
G. Ejemplo

E = "extraer al azar una persona de una población"

S = {Pepe, Luís, Marta, Carlos,......, Ana}

VA="altura de una persona"

P(Altura $\leq$163,27) = F(163,27) = 0,2

Función de densidad (f)	Función de Distribución (F)
	

H. Se llama percentil de una variable aleatoria X al valor x que acumula una determinada probabilidad (%). En el ejemplo anterior x=163,27 es el percentil 20%. El percentil 50% acumula la mitad de la distribución y corresponde siempre a la mediana.

I. Resumen

V.A. Discreta		V.A. Continua
⇩		⇩
F. de PROBABILIDAD P(X=x)		F. de DENSIDAD f(x)
⤷	Función de DISTRIBUCIÓN F(t) = P(X<=t)	⤶

ESPERANZA Y VARIANZA

A. Por "valores esperados" se entiende tanto la media como la variancia de una variable aleatoria referida a una población de estudio.

B. Ambos conceptos son los mismos para variables discretas que para continuas, pero cambia la forma de cálculo. Los valores esperados de las discretas se basan en la función de probabilidad y se resuelven con sumatorios. En las variables continuas se basan en las funciones de densidad de probabilidad y hay que integrar.

C. ESPERANZA de una variable aleatoria, media teórica, es un valor alrededor del cual se distribuye el conjunto de valores ponderando por la probabilidad y se denota:

$$E(X) = \mu_X$$

D. Ejemplo. Sea X= "contar el número de caras después de lanzar al aire una moneda 2 veces"

$$Esperanza(X) = \sum_R [x * p(x)] = 0 * \frac{1}{4} + 1 * \frac{1}{2} + 2 * \frac{1}{4} = 1$$

E. La esperanza de una variable se localiza entre el mínimo y el máximo del recorrido, pero no tiene que coincidir necesariamente con un valor de la variable.

F. La esperanza se interpreta como centro de gravedad actuando las probabilidades como masas, es decir, es el punto de equilibrio. En los juegos de azar, una esperanza cero indica que las ganancias y las pérdidas están compensadas, por tanto que el juego es justo. Sin embargo, una esperanza negativa indica que está sesgado a favor de la banca, como es el caso del juego "Chuck-a-luck".

G. VARIANCIA de una variable aleatoria, variancia teórica, es una medida de la dispersión de todos los valores alrededor del valor central y se denota:

$$V(X) = E([X - E(X)]^2) = \sigma_X{}^2$$

Se la considera un valor esperado porque es la esperanza de la variable $[X\text{-}E(X)]^2$ y por ello se puede interpretar como promedio ponderado de distancias cuadráticas.

H. Ejemplo. Sea X= "contar el número de caras después de lanzar al aire una moneda 2 veces"

$$Variancia(X) = \sum_R [(x - \mu)^2 * p(x)] = (0 - 1)^2 * \frac{1}{4} + (1 - 1)^2 * \frac{1}{2} + (2 - 1)^2 * \frac{1}{4} = \frac{1}{2}$$

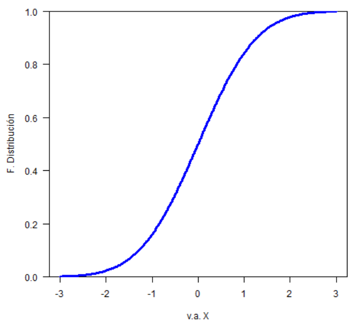
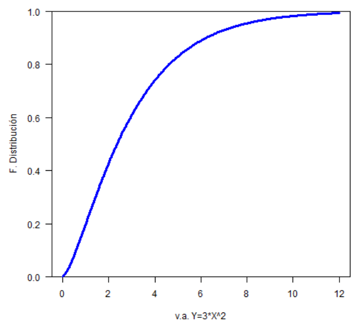
I. La variancia de una variable toma siempre un valor positivo o nulo.

J. La esperanza tiene la misma unidad de medida que la variable X, mientras que la variancia tiene la misma unidad elevada al cuadrado.

K. Una variable aleatoria X puede transformarse aplicando una función g y dando como resultado una nueva variable aleatoria Y. Un caso sencillo es la multiplicación de

X por una constante, como ocurre cuando se aplica un factor de conversión de unidades para pasar, por ejemplo, de centímetros a pulgadas. La nueva variable Y tiene su propia función de distribución.

Por ejemplo: $g(X) = 3 * X^2 = Y$

F. distribución de X	F. distribución de Y
	

L. También se pueden aplicar operaciones aritméticas, como la suma, a dos variables aleatorias. Sean X y Z dos variables aleatorias cualesquiera, con igual o diferente función de distribución, entonces g(X,Z) da lugar a una nueva variable aleatoria Y. Un ejemplo sencillo sería definir X="lanzar un dado azul de 6 caras" y Z="lanzar un dado verde de 6 caras", entonces Y=X+Z="suma de los puntos del dado azul con los del dado verde". Mientras que X y Z tiene distribución de probabilidad uniforme, la nueva variable Y tiene distribución no uniforme.

M. La esperanza y variancia de una variable Y obtenida por transformación de otras se puede obtener de las originales aplicando las siguientes propiedades:

- Sea X una variable aleatoria constante, X=k, entonces
  E (X) =k ; V(X) = 0
- Sea Y = k∗X, donde k es una constante, entonces
  E (Y) =k*∗E(X) ; V(Y) = k^2 ∗ V(X)
- Sea Y = X+Z, suma de dos variables aleatorias
  E (Y) = E(X) + E(Z)
  V (Y) = V(X) + V(Z) si X y Z son independientes
- Sea Y = X-Z, resta de dos variables aleatorias
  E (Y) = E(X) - E(Z)
  V (Y) = V(X) + V(Z) [¡!] si X y Z son independientes
- Sea Y = X∗Z, multiplicación de dos variables aleatorias
  E (Y) = E(X) ∗ E(Z) si X y Z son independientes

MODELO DISCRETO: BINOMIAL

A. Un modelo de distribución de probabilidad es un reparto teórico que se expresa mediante una fórmula matemática.

B. Hay modelos de probabilidad para variables discretas y continuas. Entre los más conocidos están[4]:

V. Discretas	V. Continuas
Bernoulli	Normal
Binomial	Ji-cuadrado
Poisson	t-Student
	F de Fisher

C. Cada modelo se distingue por la función de distribución (F), cuya forma gráfica dependerá de la ecuación general y de los valores particulares que tengan sus coeficientes, llamados parámetros. Un ejemplo análogo es decir que una curva parabólica es la representación de una ecuación de segundo grado y que la posición del vértice depende de los coeficientes.

D. El modelo BERNOULLI es el más simple para variables discretas y se aplica cuando la experiencia aleatoria es del tipo:

E = "Realizar un ensayo con resultado dicotómico: éxito o fracaso"

y la variable aleatoria se define de forma arbitraria con indicadores

X = "resultado del ensayo: 0=fracaso; 1=éxito" ; R = {0, 1}

E. Ejemplo

E = "escoger al azar una persona de un grupo" → éxito = "ser zurdo"

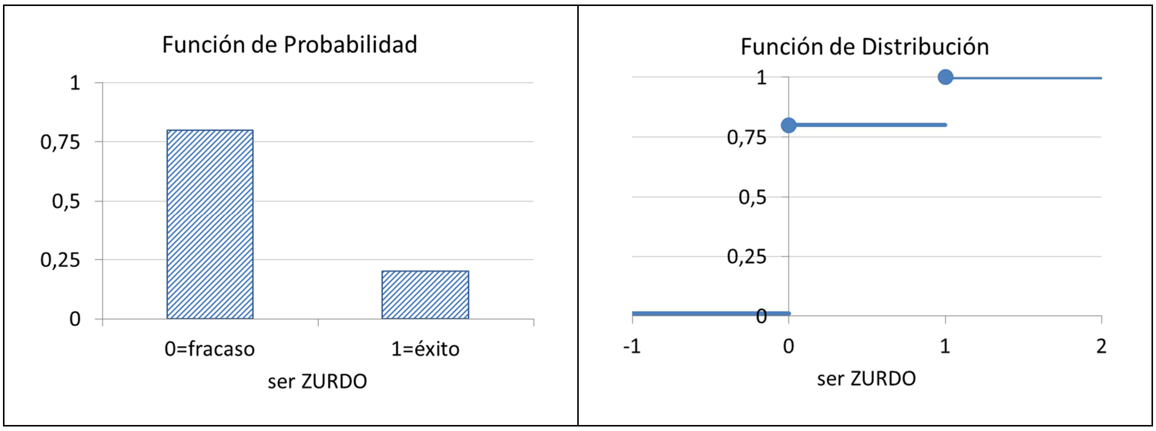
X = "mano de escritura: 0=derecha; 1=izquierda"; R = {0, 1}

F. Función de probabilidad y de distribución de Bernoulli

Probabilidad de éxito = p

F. Probabilidad: P(x=0) = 1-p; P(x=1) = p

F. Distribución: F(-1)= 0; F(0) = 1-p; F(1)= 1



[4] NIST/SEMATECH e-Handbook of Statistical Methods < http://www.itl.nist.gov/div898/handbook/eda/section3/eda366.htm >

G. Lo valores esperados del modelo Bernoulli son:

Esperanza (X) = Σ [x∗p(x)] = [0*(1-p)]+[1*(p)] = p

Varianza (X) = Σ [(x-E(X))^2 ∗ p(x)] = [(0-p)^2*(1-p)]+[(1-p)^2∗(p)] = p*(1-p)

H. El modelo BINOMIAL[5] explica el reparto de probabilidad de los resultados que se pueden observar al realizar una experiencia aleatoria del tipo:

E = “repetir n veces un ensayo con resultado dicotómico”

y la variable aleatoria es el recuento de los éxitos

X = ”número de éxitos en n repeticiones del ensayo”

R = {0,1…, n}

I. Ejemplo

E = “escoger al azar 3 persona de un grupo” [6] → éxito = “ser zurdo”

S = {(d, d, d); (d, d, z); (d, z, d); (z, d, d); (d, z, z); (z, d, z); (z, z, d); (z, z, z)}

X = ”número de zurdos en 3 personas escogidas al azar”

R = {0, 1, 2, 3}

J. Para indicar que una variable aleatoria X sigue una distribución de probabilidad Binomial se utiliza la expresión:

X ~ B (n,p)

siendo n y p los dos parámetros del modelo:

n = número de repeticiones

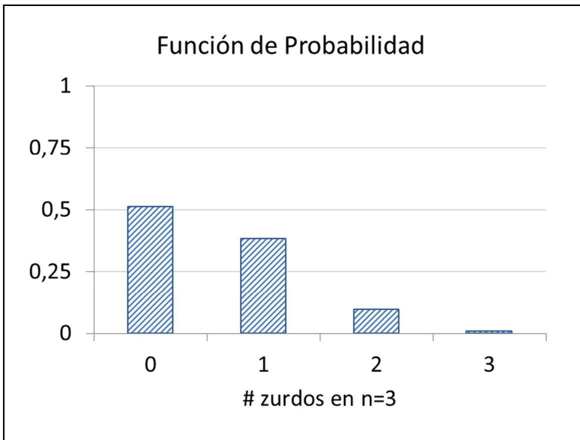
p = probabilidad de éxito (1-p= q =probabilidad de fracaso)

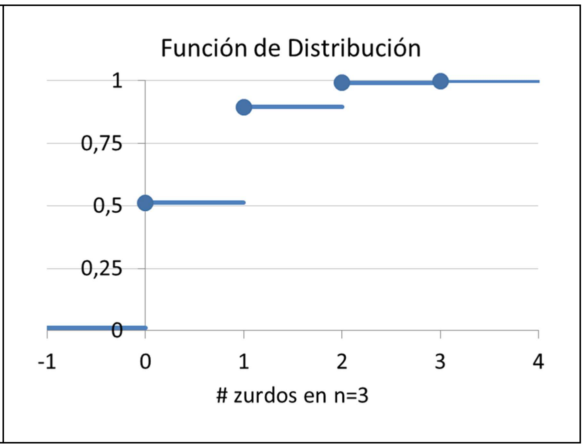
K. La función de probabilidad de una Binomial se genera con la fórmula:

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad ; \quad x = 0,1,\ldots,n$$

siendo el paréntesis vertical el número de combinaciones posibles[7]

$$\binom{n}{x} = C(n,x) = \frac{n!}{x!\,(n-x)!}$$





[5] El nombre se debe al Teorema del Binomio: ΣP(x) = 1 = (p+(1-p))^n

[6] Después de cada extracción tiene que haber reemplazo, pero el efecto es desdeñable si el grupo es muy grande.

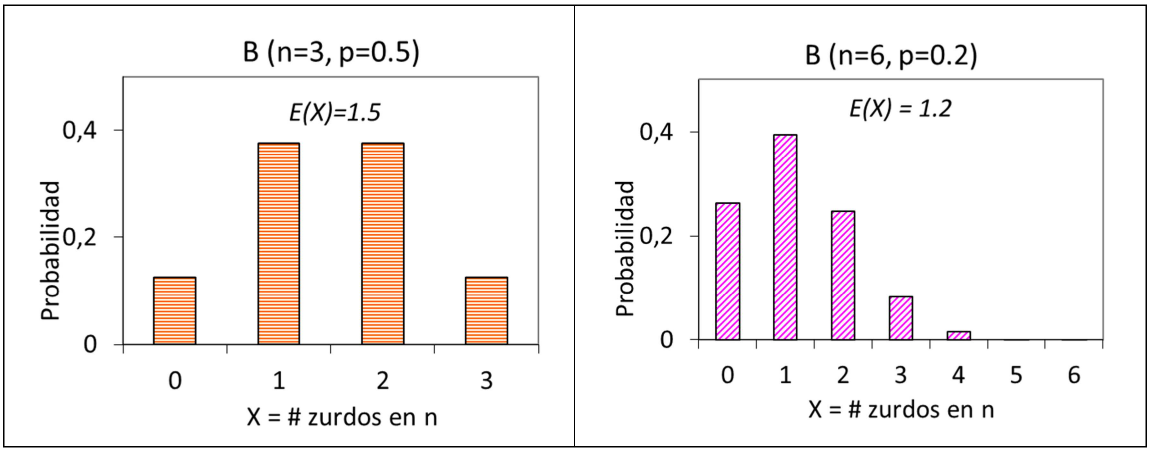
[7] En las combinaciones se tienen en cuenta todos los resultados que dan lugar al mismo recuento, sin distinguir el orden: 1 zurdo es equivalente a (d,d,z) = (d,z,d) = (d,z,d)

L. Para que el uso de un modelo sea válido como calculadora de probabilidades se han de cumplir las siguientes condiciones:

- el mismo ensayo se repite n veces, siempre de la misma manera
- en cada ensayo sólo hay dos resultados posibles: se observa la característica de interés (éxito) o no (fracaso).
- los ensayos son independientes, es decir, el resultado de un ensaya no condiciona el del siguiente.
- la probabilidad de éxito, p, se mantiene constante en todos los ensayos.

M. La forma de la función de probabilidad de un modelo Binomial depende de los valores de los parámetros:

- el número de barras es igual a (n+1), en consecuencia al aumentar n la altura de las barras disminuirá para un mismo valor de p.
- la simetría de la función sólo ocurre para p=0,5, independientemente del valor de n, mientras que para cualquier otro valor de p el reparto de probabilidad es asimétrico. En el ejemplo X="número de zurdos en 3 personas escogidas al azar" la probabilidad de X=0 (ningún zurdo) es mayor que la de X=3 (todos zurdos) si sólo una quinta parte de la población es zurda, es decir, p=0,2.



N. La función de distribución Binomial se construye por suma de probabilidades:

$$P(X \leq x) = \sum_{i=0}^{x} P(X = x)$$

Ñ. Los valores esperados del modelo Binomial son:

Esperanza (X) = Σ [x∗p(x)] = n∗p

Variancia (X) = Σ [(x-E(X))^2 ∗ p(x)] = n∗p∗(1-p)

si X ~ B (n=3, p=0,5), entonces E(X)=3*0,5=1,5 y V(X)=3*0,5(1-0,5)=0,75

O. Cualquier Binomial se puede generar por sumas de n variables Bernoulli con la misma probabilidad de éxito

si X ~ Bernoulli (p=0,2), entonces Y= [X + X + X] ~ B (n=3,p=,2)

P. A partir de lo anterior se afirma que la Binomial es aditiva, es decir, que la suma de variables Binomiales con distintos números de repeticiones, pero p común, es también Binomial:

Sean $X_1 \sim B(n_1, p)$ y $X_2 \sim B(n_2, p)$; $Y = [X_1 + X_2] \sim B(n_1 + n_2, p)$

MODELO DISCRETO: POISSON

A. El modelo de POISSON[8] explica el reparto de probabilidad de una variable aleatoria que se ajuste a una expresión del tipo:

X = "número de acontecimientos raros en un intervalo de tiempo o espacio"

R = {0,1, 2, 3, …}

En esencia, se trata de un recuento de resultados sin límite superior, y con probabilidad de ocurrir muy baja.

B. Ejemplos

- número de soldado prusianos muertos por coz de caballo en un año[9]

- número mujeres que dan a luz durante un vuelo de avión en un año

- número de nuevos casos de cáncer de mama en varones en una región

C. Para indicar que una variable aleatoria X sigue una distribución de probabilidad Poisson se utiliza la expresión:

$X \sim \mathscr{P}(\text{lambda})$

siendo lambda (λ) el único parámetro de la distribución

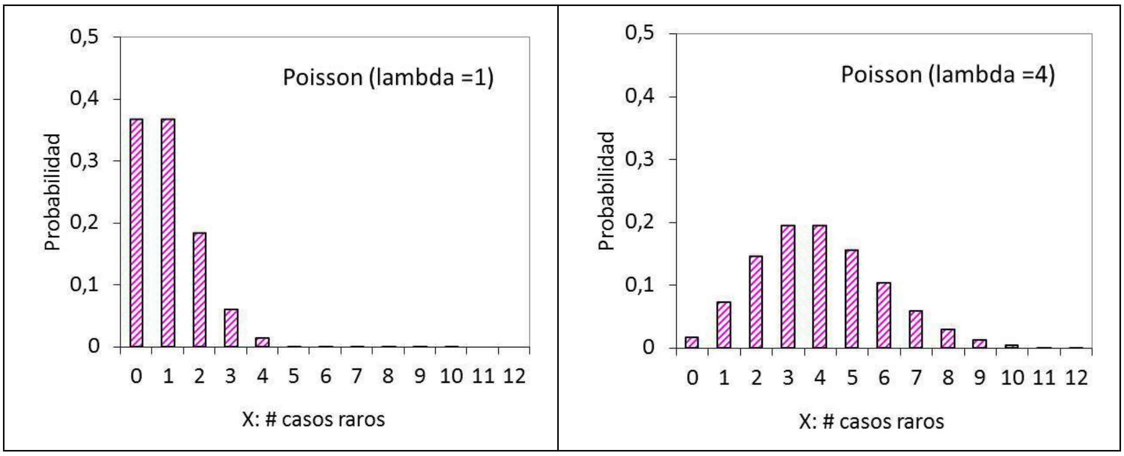
D. La función de probabilidad de una Poisson se genera con la fórmula:

$$P(X=x) = \frac{\lambda^x}{x!} exp(-\lambda); \qquad x = 0, 1, 2, 3, \cdots$$

E. La forma de la función de probabilidad dependerá sólo del valor de lambda. Teóricamente el recorrido de la variable es una secuencia infinita de números enteros, por eso el número de barras no es un número concreto. Ahora bien, la altura de las barras, o sea la probabilidad, va disminuyendo a medida que los valores de X incrementan hacia más infinito.

[8] En honor al matemático francés Siméon Denis Poisson, por ser quien la derivó en 1837.

[9] Ejemplo real e histórico que fue la primera aplicación de este modelo



F. Para que la aplicación del modelo sea válida hay que partir de la idea de que el intervalo de tiempo t, o espacio, se puede dividir en n subintervalos de tamaño t/n=dt que no se solapan y asumir que

- la probabilidad de observar un acontecimiento en un subintervalo es proporcional a su tamaño.
- el número de acontecimientos por unidad de tiempo es constante a lo largo de todo el intervalo (estacionalidad)
- la ocurrencia de una acontecimiento en un subintervalo no influye en lo que sucederá en el siguiente (independencia).

G. Los valores esperados del modelo Poisson son:

Esperanza (X) = λ

Variancia (X) = λ

De aquí se concluye que λ es el valor promedio de observaciones en el periodo de tiempo, o espacio, estudiado.

H. La igualdad entre esperanza y variancia se aplica para detectar epidemias. Se sospecha de un efecto de contagio si al estudiar los casos nuevos de una enfermedad la igualdad no se cumple.

I. La función de distribución de Poisson se construye por suma de probabilidades:

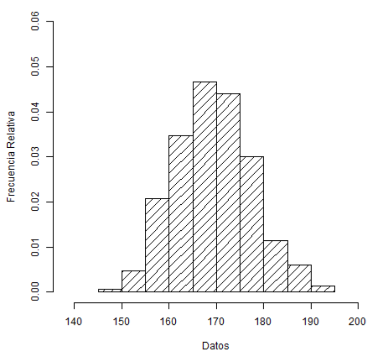
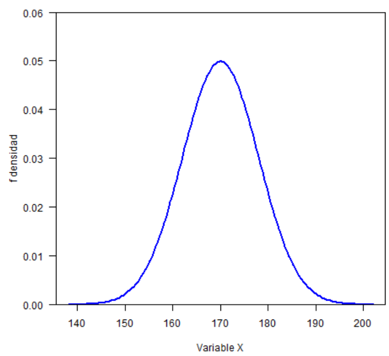
$$P(X \leq x) = \sum_{i=0}^{x} P(X = x)$$

J. Se dice que la Poisson es aditiva porque al sumar dos variables con distribuciones de Poisson de lambda cualquiera la nueva variable también sigue una Poisson:

Sean $X_1 \sim \mathcal{P}(\lambda_1)$ y $X_2 \sim \mathcal{P}(\lambda_2)$; $[X_1 + X_2] \sim \mathcal{P}(\lambda_1 + \lambda_2)$

MODELO CONTINUO: NORMAL

A. La distribución NORMAL es un objeto matemático que ha sido estudiado durante casi 300 años recibiendo varios nombres. La idea fue concebida por Abraham de Moivre al buscar una manera precisa de calcular las probabilidades de distribuciones Binomiales con gran número de repeticiones y que aparece publicada en su “Doctrine of Chance” en 1733[10]. Anteriormente astrónomos como Galileo Galilei ya habían razonado que los errores de observación se distribuían simétricamente, pero no sería hasta más tarde que se usaran modelos de distribución. En 1809 Carl F. Gauss publicó un estudio de datos astronómicos que incluía una ley de los errores con expresión exponencial cuadrática. Es por ello que la función de densidad de la normal también se conoce como campana de Gauss[11]. Un año más tarde Pierre-S. Laplace publicó una demostración que generalizaba los resultados de de Moivre y su segunda ley de los errores, que era equivalente a la de Gauss.

Histograma Distribución de frecuencias	Campana de Gauss Distribución de masa de probabilidad
	

B. En ciencias de la vida el modelo normal toma protagonismo hacia finales del siglo XIX con los estudios de Sir Francis Galton sobre la variación de las medidas cuantitativas de características biológicas. Este científico polifacético fue el precursor de la escuela biométrica y uno de los iniciadores del uso del término “curva normal”. Más tarde su discípulo Karl Pearson, uno de los pioneros de la bioestadística, contribuyó a su expansión al mencionarlo en sus artículos. Ejemplos de variables aleatorias continuas que se modelan con la Normal son:

- *altura de la población de hombres*
- *concentración de colesterol en una población de personas sanas*

C. La función de DENSIDAD, f, del modelo de distribución Normal se expresa mediante la fórmula:

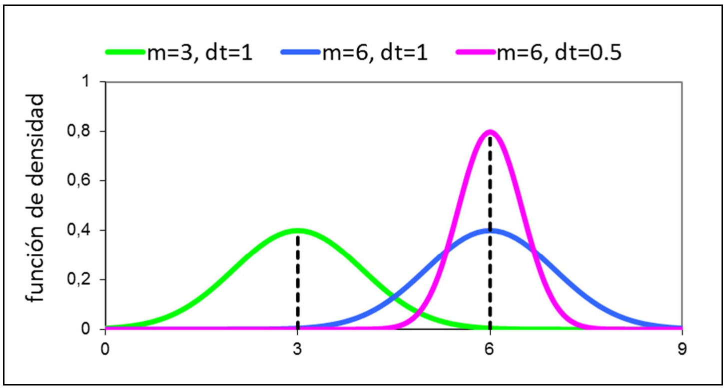
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right); \quad x \in (-\infty, +\infty)$$

[10] En: Stigler, S M. “The seven pillars of statistical wisdom”. Harvard University Press. 2016
[11] En: “Encyclopedia of Statistical Science”, vol 4: Laws of errors. Ver bibliografía.

D. La CAMPANA de Gauss es la representación gráfica de la función de densidad.

E. Los parámetros del modelo son μ y σ. El primero es la esperanza de X, $\mu=E(X)$, y decide la posición de la campana a los largo del eje horizontal. El segundo es la raíz cuadrada de la variancia X, $\sigma=\sqrt{V(X)}$, o sea la desviación típica, y determina la anchura de la campana. Para decir que una variable aleatoria X sigue una distribución Normal se utiliza la expresión:

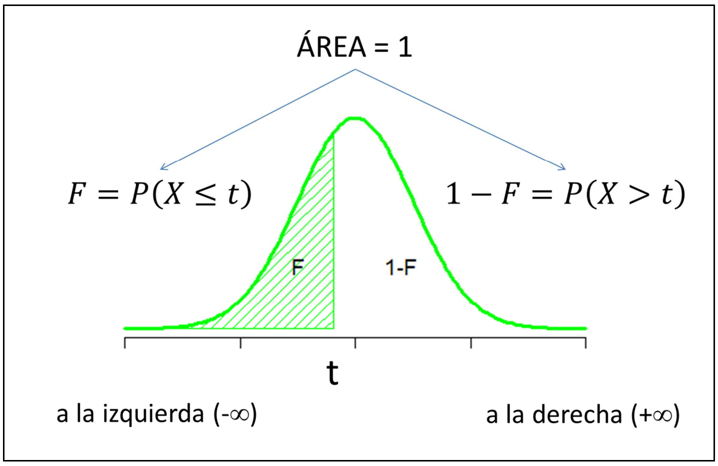
$$X \sim n\ (\mu, \sigma)$$



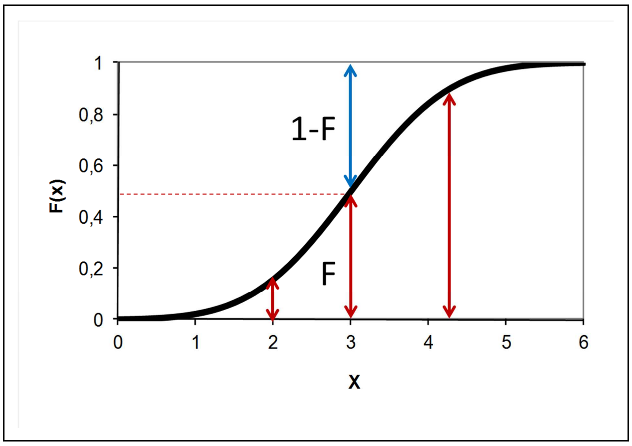
F. Características de la función de densidad normal:

- el área total bajo la curva de la función vale 1, por definición de densidad.
- la curva presenta un pico (máximo) en el centro, cuya altura puede ser mayor que 1 si la desviación típica es muy pequeña.
- la curva es simétrica respecto al eje vertical que pasa por el centro.
- la media, mediana y moda coinciden en el mismo valor.
- la función tiene dos puntos de inflexión en las abscisas $(\mu\text{-}1\sigma)$ y $(\mu+1\sigma)$ respectivamente, con lo cual la desviación típica coincide con la distancia horizontal desde el eje central a los puntos de inflexión.
- las colas son infinitas, es decir, la curva toca el eje X en $-\infty$ y $+\infty$.

G. La función de DISTRIBUCIÓN, F, de la Normal se obtiene integrando la función de densidad. El valor de F para cualquier valor real t, F(t), es el área bajo la campana desde $-\infty$ hasta t y se ha de interpretar como la probabilidad acumulada de que la variable aleatoria tome un valor por debajo de t.



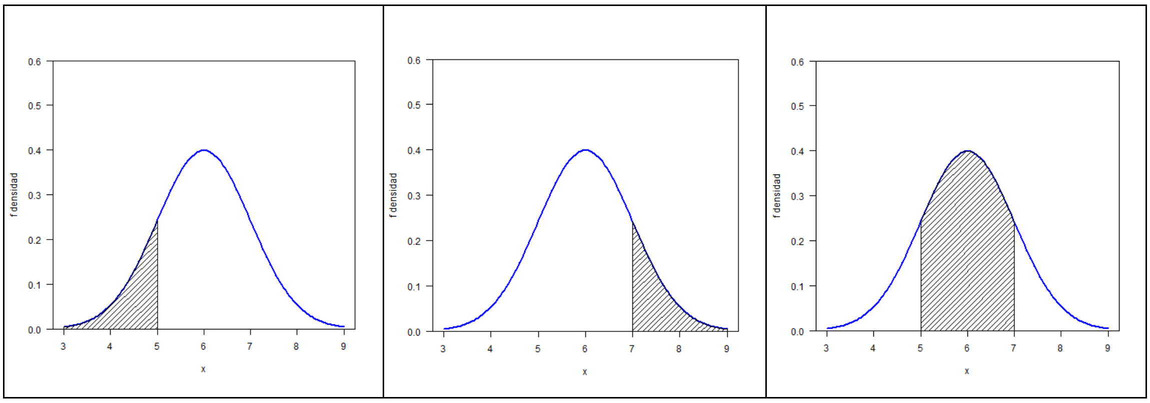
H. La integral de f no tiene solución analítica y se resuelve por métodos numéricos. La representación gráfica de F es una línea en forma de “ese” acotada por un mínimo en cero y un máximo en uno.



I. Los PERCENTILES de una variable con distribución Normal quedan determinados por la función F. Por ejemplo el percentil 90% es el valor de x cuya F vale 0.9. La media de una normal, debido a su simetría, acumula el 50% de la probabilidad y es equivalente al percentil 50%.

J. En el modelo Normal el cálculo de la probabilidad acumulada aprovecha la simetría de la campana para obtener:

- cola inferior: $P\ (X \leq \mu\text{-}c) = F(\mu\text{-}c)$
- cola superior: $P\ (X \geq \mu+c) = 1\text{-}\ F(\mu+c) = F(\mu\text{-}c)$
- intervalo centrado: $P(\mu\text{-}c < X \leq \mu+c) = 1\ \text{-}2*F(\mu\text{-}c)$



K. Cualquier campana de Gauss, es decir, sean cuales sean los valores de media y desviación típica, siempre cumple las siguientes proporciones:

área entre $(\mu\text{-}1\sigma)$ y $(\mu+1\sigma)$ $\approx$ 68%

área entre $(\mu\text{-}2\sigma)$ y $(\mu+2\sigma)$ $\approx$ 95%

área entre $(\mu\text{-}3\sigma)$ y $(\mu+3\sigma)$ $\approx$ 99%

Esta idea es la que se aplica en la definición de “intervalos de referencia” para interpretar resultados de bioquímica clínica o del coeficiente de inteligencia.

L. Cualquier variable aleatoria puede ser transformada mediante operaciones matemáticas en una nueva variable con propiedades iguales o no a la original. La tipificación es un proceso para transformar cualquier variable X en una variable reducida, es decir, que la media sea cero y la variancia uno. Consiste en restar a X su media y dividirla por su desviación típica.

$$Y = \frac{X - E(X)}{DT(X)} = \frac{1}{\sigma} * (X - \mu) \quad \Rightarrow \quad E(Y) = 0;\ DT(Y) = 1;\ V(Y) = 1$$

La tipificación de un variable X que no es normal, no la convierte en normal. Sin embargo, si la variable X original sigue este modelo, entonces la variable nueva sí heredará la distribución normal, pero con los parámetros modificados. Una variable con distribución normal de parámetros cero y uno se la conoce como "normal tipificada" o sencillamente "zeta".

$$X \sim N(\mu, \sigma) \quad \rightarrow tipificación \quad \rightarrow Z \sim N(0,1)$$

M. El proceso de tipificación en la normal es importante para el cálculo de probabilidades, porque la probabilidad acumulada asociada a un valor x de una normal cualquiera se conserva al transformarlo a zeta:

$$F(x) = P(X \leq x) = P\left(Z \leq \frac{x - \mu}{\sigma}\right) = F(z)$$

La función de distribución N(0,1) está tabulada convenientemente en lo que se conoce como tabas de la Z. Percentiles importantes de la misma son:

F(z=1,285) = 0,90 F(z=1,645) = 0,95 F(z=1,96) = 0,975

N. Teorema de la Adición. La suma de dos variables aleatorias Normales independientes entre sí da lugar a una nueva variable aleatoria con distribución también Normal de parámetros

$$\mu_Y = E(X_1 + X_2) \qquad \sigma_y = \sqrt{V(X_1) + V(X_2)}$$

CONVERGENCIAS

A. Convergencia es un concepto del análisis matemático que tiene que ver con lo que ocurre con una sucesión que se lleva al límite. Si la sucesión está compuesta por variables aleatorias y el límite es otra variable aleatoria con una distribución aproximadamente equivalente, entonces se habla de convergencia de distribuciones. En el fondo esto significa que la forma de una distribución evoluciona al cambiar el valor de los parámetros de tal manera que acaba pareciéndose a otras distintas.

B. De Binomial a Poisson. Se puede demostrar matemáticamente que la distribución Poisson (lambda) aparece como un paso al límite de la distribución Binomial(n,p):

$$\lim_{\substack{n \to \infty \\ p \to 0}} B(n,p)(x) = \frac{\lambda^x}{x!} exp(-\lambda) = \wp(\lambda)$$

Esto quiere decir que, cuando n es grande y p suficientemente pequeña, las probabilidades de una variable Binomial (n,p) se pueden obtener, de forma aproximada usando la fórmula de una Poisson (lambda=np). La regla práctica es: n>30 y p<0,1. Así, una Binomial (1000, 0.002) puede ser aproximada por un Poisson (2).

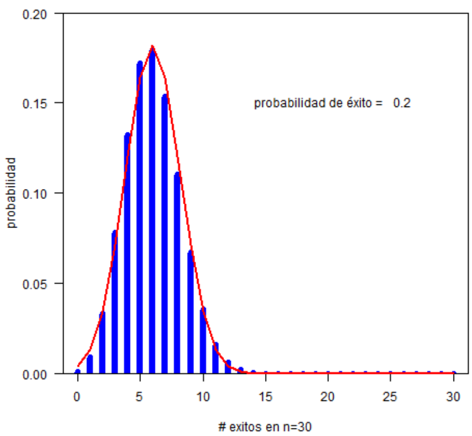
	P(X=0)	P(X=1)	P(X=2)	P(X=3)
B (1000,0.002)	0,13506	0,27067	0,27094	0,18063
Poisson (2)	0,13532	0,27067	0,27067	0,18045

C. De Binomial a Normal. El teorema de DeMovire-Laplace afirma que si X es una variable Binomial construida por suma de n variables Bernoulli, todas con la misma p, entonces la distribución de X converge a una distribución normal en el límite cuando n tiende a infinito.

$$\lim_{n\to\infty} B(n,p)(x) = \int_a^b \frac{1}{\sqrt{2\pi npq}} \exp\left(\frac{-(x-np)^2}{2npq}\right) dx$$

Los parámetros de la Normal han de coincidir con la esperanza y la variancia de la Binomial:

$$\mu = E(Bin) = n * p; \quad \sigma = \sqrt{V(Bin)} = \sqrt{n * p * (1-p)}$$



La aproximación se aplica para valores de n grandes y de p no muy extremos. En la práctica esto se traduce por:

n grande	n > 30
p no pequeño	min (np, nq) >5

Dado que el modelo Normal es contínuo y el Binomial discreto, la aproximación puede mejorarse aplicando un factor de corrección de continuidad (cc) que consiste en sumar 0.5 a los valores de la variable.

	P (X≤20)	P (16 <X≤20)	P (18<X≤22)
B (100, 0.2)	0,559462	0,367124	0,37685
Normal (20,4)	0,50	0,341345	0,382928

D. Teorema del Límite Central (CLT). Este teorema es una generalización del teorema que permite usar la Normal para aproximar una Binomial y es la razón de que el modelo Normal sea tan importante en estadística. Existen varias versiones según las condiciones que se impongan a las variables aleatorias que forman la sucesión. La más completa es la de Lindeberg-Feller que fue resuelta ya entrado el siglo XX, pero la más conocida y utilizada es el caso especial de Lindeberg-Levy:

“Sea $\{X_n , n\geq 1\}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas (no necesariamente normal), con media común y variancia común y finita. La variable suma de la sucesión, S_n, tipificada converge en el límite a una variable con probabilidad acumulada aproximadamente normal con media cero y variancia uno”

$$\lim_{n\to\infty} P\left(a < \frac{S_n - n\mu}{\sigma\sqrt{n}} \leq b\right) \cong F_z(b) - F_z(a)$$

E. Ejemplo. Supongamos 3 dados de seis caras. Se define la variable aleatoria:

X_i = “puntos obtenidos al lanza el dado i una vez”

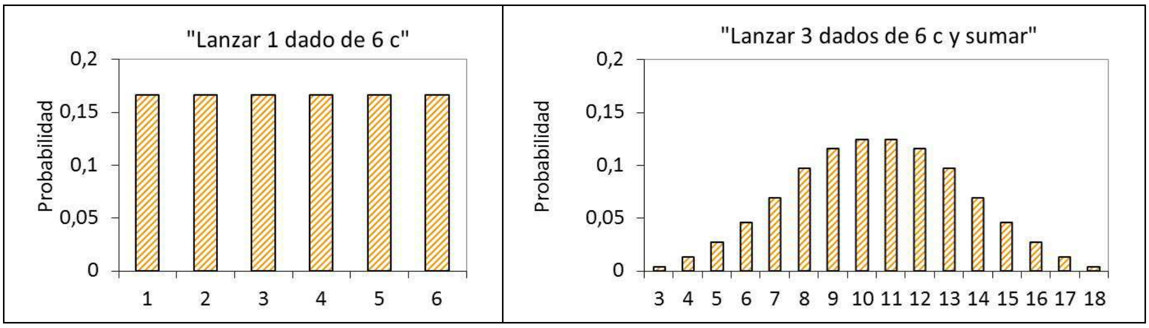
que tiene distribución uniforme, P(X=x)= 1/6, y valores esperados:

$E(X_i) = \mu = 3{,}5 \quad ; \quad V(X_i) = \sigma^2 = 2{,}92$

Se define la suma de la sucesión de $\{X_i \text{ , } n=3\}$

S_3 = “suma de los puntos observados en los tres dados”

La distribución de S_3 no es uniforme, sino que recuerda a una Normal. El teorema dice que si el número de dados fuera infinito, o si usáramos el mismo dado infinitas veces, la distribución de la distribución de puntos sería Normal.



BIBLIOGRAFÍA RECOMENDADA

- Kotz S, Johnson NL. Encyclopedia of Statistical Sciences. Wiley-Interscience 1986.
- Johnson R A, Bhattacharyya GK. Statistics: principles and methods. Hoboken, N.J: Wiley; cop. 2010, 6th ed., International student ed.
- Larson H J. Introduction to probability theory and statistical inference. New York [etc.] : Wiley, cop. 1982, 3rd ed.
- Martínez, MA et al. Bioestadística amigable. Elsevier, 3a ed, 2014.
- Rosner B. Fundamentals of biostatistics. Pacific Grove, Calif. : Brooks/Cole, Cengage Learning, 2011. 7th ed., International ed

GLOSARIO

Aproximación Normal a la Binomial
Asimetría y curtosis de una fdp
Corrección de continuidad
Curva de densidad de probabilidad
Curva de Gauss
Desviación típica (poblacional)
Ensayo de Bernoulli
Esperanza
(valor esperado, media poblacional)
Éxito vs fracaso
Función de densidad de probabilidad
Función de probabilidad acumulada
Función de probabilidad de una v.a. discreta
Independencia de ensayos
Modelo de distribución Binomial
Modelo de distribución de una variable continua
Modelo de distribución Normal
Modelo de distribución Poisson
Modelo de probabilidad
Modelo normal tipificado
Muestreo con/sin reemplazo
Parámetro
Percentiles
Probabilidad en el intervalo (a,b)
Transformación de v.a.
Variable aleatoria (v.a.)
V.a. contínua
V.a. dicotómica
V.a. discreta
Variable tipificada
Variancia / varianza (poblacional)