



Grado de Medicina – Universidad de Barcelona
Bioestadística básica, Epidemiología y Introducción a la Investigación (2016/17)
Begoña Campos – Departamento de Fundamentos Clínicos

DISTRIBUCIONES TEÓRICAS DE PROBABILIDAD

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

INTRODUCCIÓN

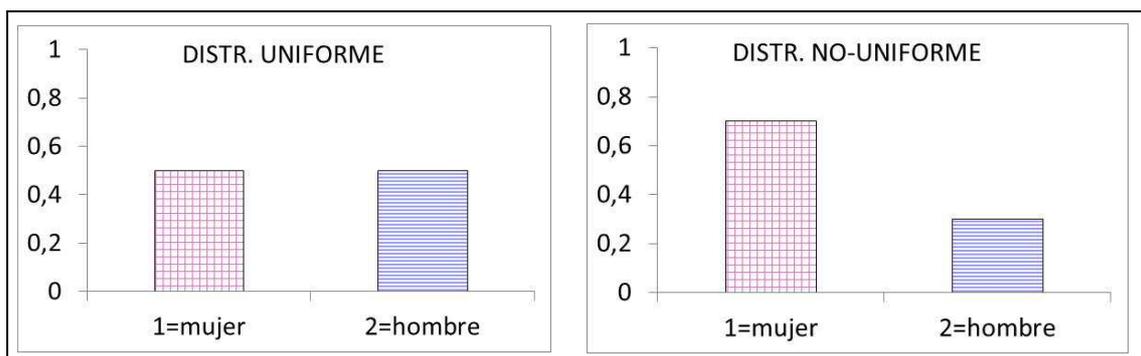
“all the world believes it (Gaussian distribution) firmly, because the mathematicians imagine that it is a fact of observation, and the observers that it is a theorem of mathematics”
(Lippmann a Poincaré, 1912)¹

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. Si en la población hay tantas 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: $P(M) = 0,5$; $P(H) = 0,5$ Suma = 1

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

No-uniforme: $P(M) = 0,7$; $P(H) = 0,3$ Suma = 1



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

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 técnica gráfica para representar la distribución de frecuencias de una variable

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

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

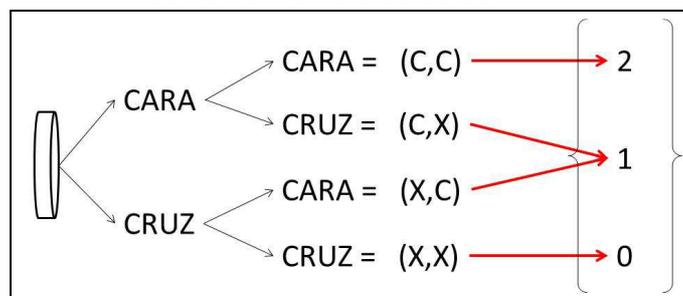
VA: espacio muestral (S) → conj. Reales (R)

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. 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 = $X(\text{Miguel}) = 173,6\text{cm} = x$

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"

$S = \{(c,c); (c,x); (x,c); (x,x)\}; \quad c=0, \quad x=1$

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 2ª tirada 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 hijos de una mujer

Y = número de partos en taxis de Barcelona en un año

- CONTINUA. Variable con un conjunto de resultados ilimitado, 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 separado de la estadística. Tradicionalmente se afirma que empezó en Francia en 1654 con la correspondencia por carta entre dos matemáticos (Pascal y Fermat) que debatían sobre la división de apuestas en los juegos de azar².

B. En el juego de lanzar dos dados de seis caras y sumar los puntos de ambos, los resultados no son equiprobables. Obtener un siete es más probable ($6/36$) que obtener un dos ($1/36$) porque hay más maneras de conseguir ese resultado. Para compensar, el premio asociado al dos ("snake eyes") es mayor.

C. Una distribución teórica de probabilidad describe el reparto de los valores de una variable aleatoria en una población. Los valores que son más abundantes tendrán mayor probabilidad de aparecer al realizar la experiencia aleatoria que los valores más escasos.

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

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 $[0,1]$:

$$\begin{aligned} \text{FP: } x &\rightarrow P(x) = p && \text{si } x \text{ está en el recorrido} \\ x &\rightarrow P(x) = 0 && \text{si el valor } x \text{ no está recorrido} \end{aligned}$$

Dos propiedades importantes:

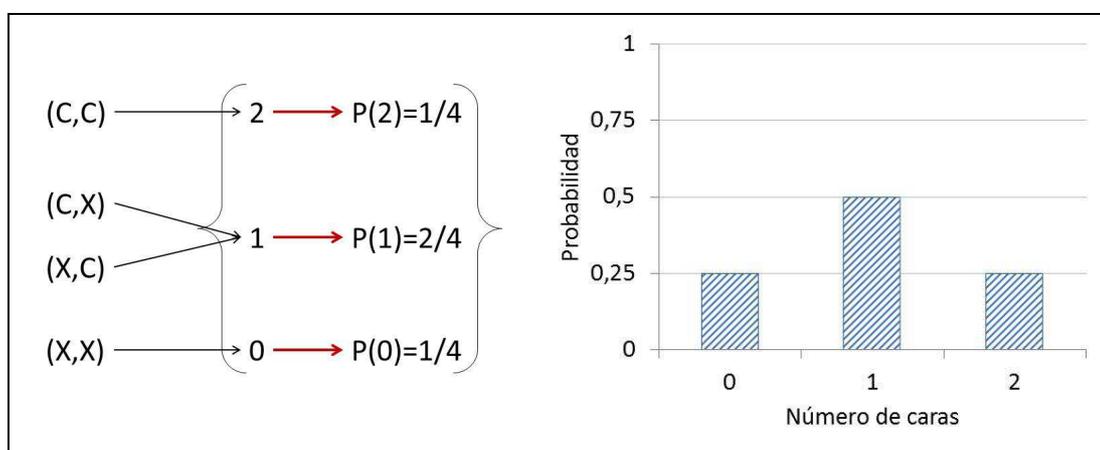
- Los valores p son siempre positivos o nulos: $p \geq 0$
- La suma de todas las p 's ha de ser igual a uno: $\sum p = 1$

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"



D. 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?

E. 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 cero³:

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

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

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

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

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: $\lim_{R \rightarrow +\infty} [1/R] = 0$

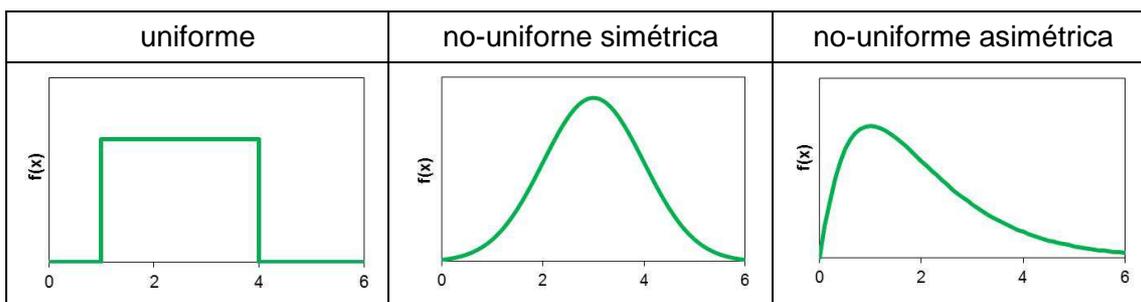
- En cualquier punto de la recta real la función es positiva o cero (pero no necesariamente menor que 1):

$$f(x) \geq 0; \quad \forall x \text{ real}$$

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

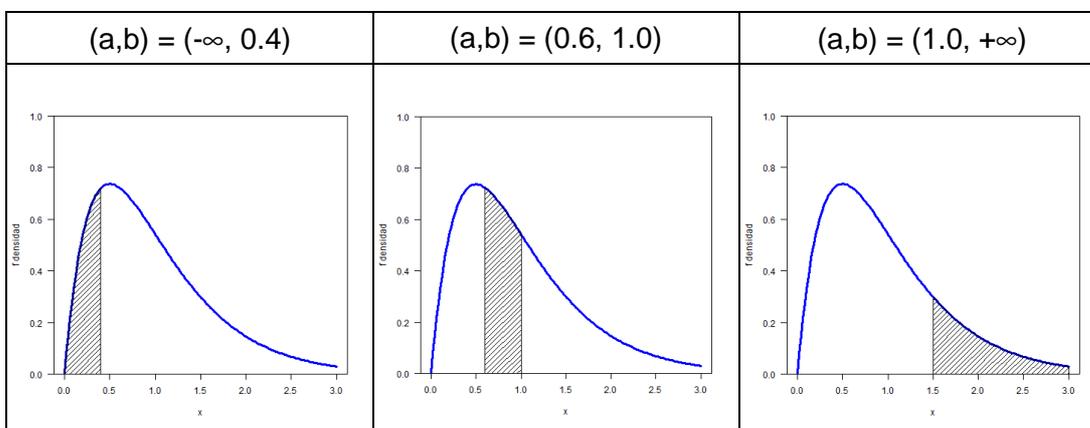
$$\int_{-\infty}^{+\infty} f(x) dx = 1$$

G. Existen muchas curvas que cumplen los criterios para ser función de densidad, por ejemplo las siguientes:



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

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



FUNCIÓN DE DISTRIBUCIÓN

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:

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

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:

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

B. 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.

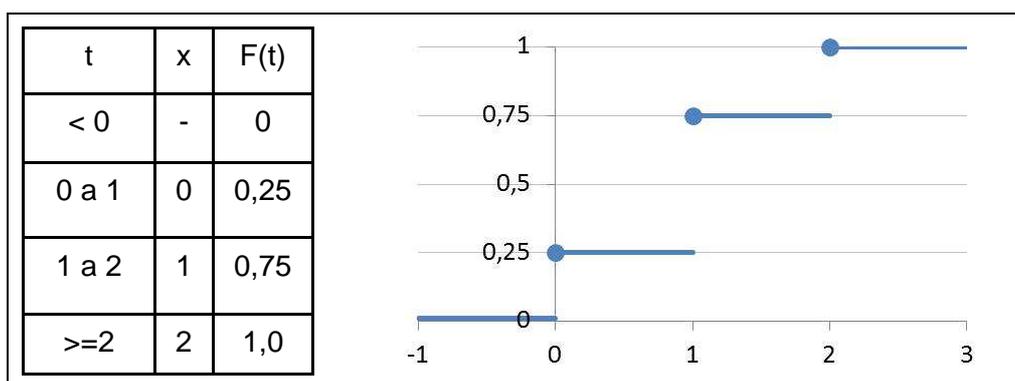
$$F(t) = \sum p(x) \quad ; \quad x \leq t$$

C. 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"



D. En las variables continuas la función de distribución para t se obtiene integrando la función de densidad de probabilidad desde $-\infty$ 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]

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

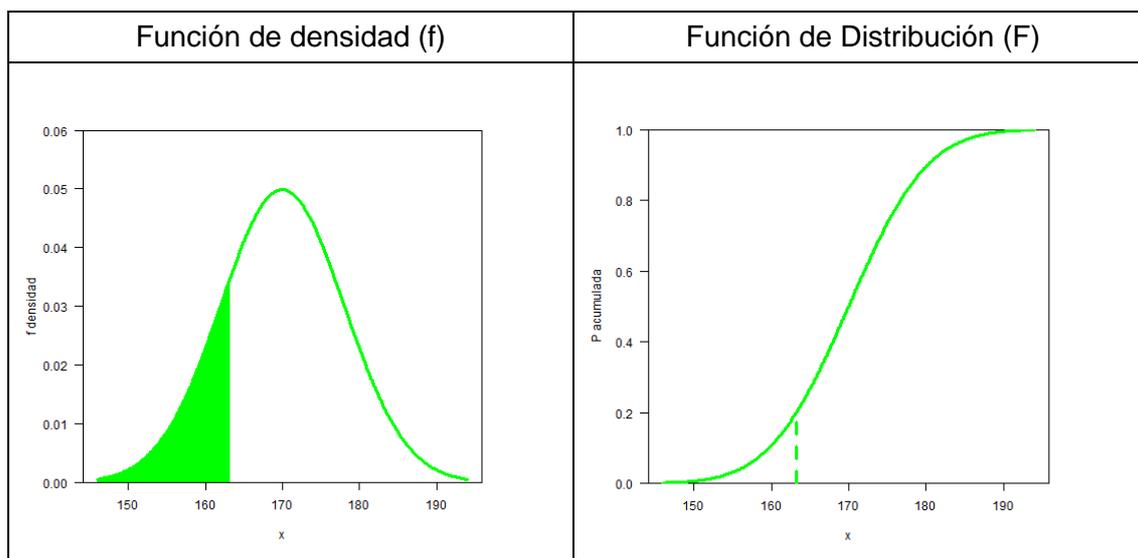
E. Ejemplo

E = “extraer al azar una persona de una población”

$S = \{\text{Pepe, Luís, Marta, Carlos,....., Ana}\}$

$VA = \text{“altura de una persona”}$

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



F. 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.

ESPERANZA Y VARIANZA

A. Se llaman “valores esperados” de una variable aleatoria a su media y variancia en la población.

B. El concepto es el mismo para variables discretas y continuas, pero cambia la forma de cálculo. Los valores esperados de las discretas derivan de la función de probabilidad y se resuelven con sumatorios. En las continuas se aplica la función de densidad y hay que integrar.

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

$$E(X) = \mu$$

D. El valor de la esperanza se localiza entre el mínimo y el máximo de X y no tiene que coincidir necesariamente con un valor del recorrido.

E. 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”.

F. Ejemplo.

E = “lanzar al aire una moneda 2 veces”

X = “contar el número de caras”

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

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

$$V(X) = E([X - E(X)]^2) = \sigma^2$$

H. El valor de la variancia es siempre positivo o nulo.

I. La variancia se puede interpretar como promedio ponderado de distancias cuadráticas.

J. Ejemplo.

E = “lanzar al aire una moneda 2 veces”

X = “contar el número de caras”

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

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

L. Una variable aleatoria X puede modificarse aplicando transformaciones del tipo multiplicar por una constante. Esto sucede cuando se aplican factores de conversión de unidades para pasar, por ejemplo, de centímetros a pulgadas. La esperanza y variancia de una variable transformada Y se pueden derivar de las originales con 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
 $E(Y) = E(X) + E(Z)$
 $V(Y) = E(X) + E(Z)$ si X y Z son independientes
- Sea $Y = X - Z$, resta de dos variables aleatorias
 $E(Y) = E(X) - E(Z)$
 $V(Y) = E(X) + E(Z)$ [i!] 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 ecuación.

B. Hay modelos de probabilidad para variables discretas y continuas. Entre los más conocidos están⁴:

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

C. Cada modelo tiene una función de distribución (F) distinta, cuya forma dependerá de la ecuación general y de los valores particulares que tengan sus coeficientes, llamados parámetros.

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: fracaso=0; éxito=1"; $R = \{0, 1\}$

E. Ejemplo

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

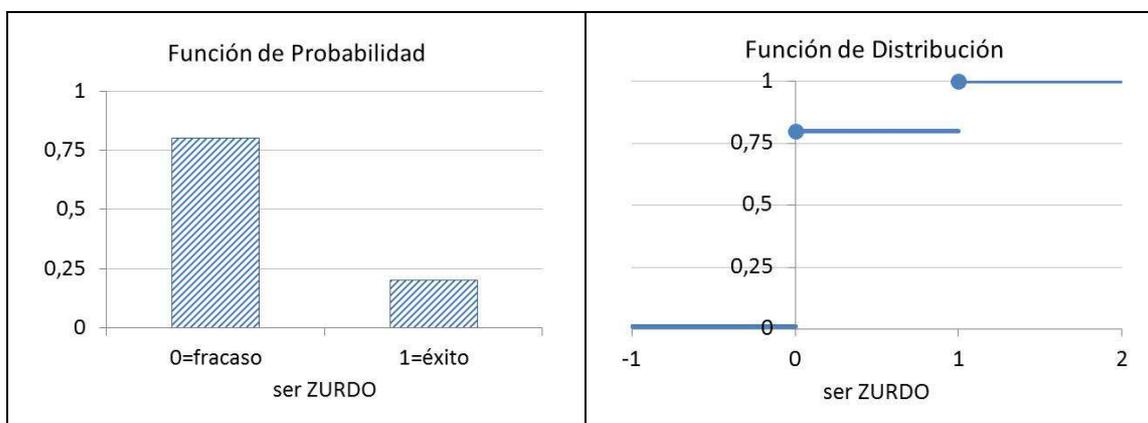
X = "mano de escritura: derecha=0; izquierda=1"; $R = \{0, 1\}$

F. Función de probabilidad y de distribución de Bernoulli

Probabilidad de éxito = p

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

F.D. $F(-1) = 0$; $F(0) = 1-p$; $F(1) = 1$



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:

$$\text{Esperanza (X)} = p$$

$$\text{Variancia (X)} = p \cdot (1-p)$$

H. El modelo BINOMIAL⁵ 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"

$$R = \{0, 1, \dots, n\}$$

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 decir que una variable aleatoria X sigue una distribución de probabilidad Binomial se utiliza la expresión:

$$X \sim B(n, p)$$

siendo n y p los dos parámetros del modelo:

n = número de repeticiones

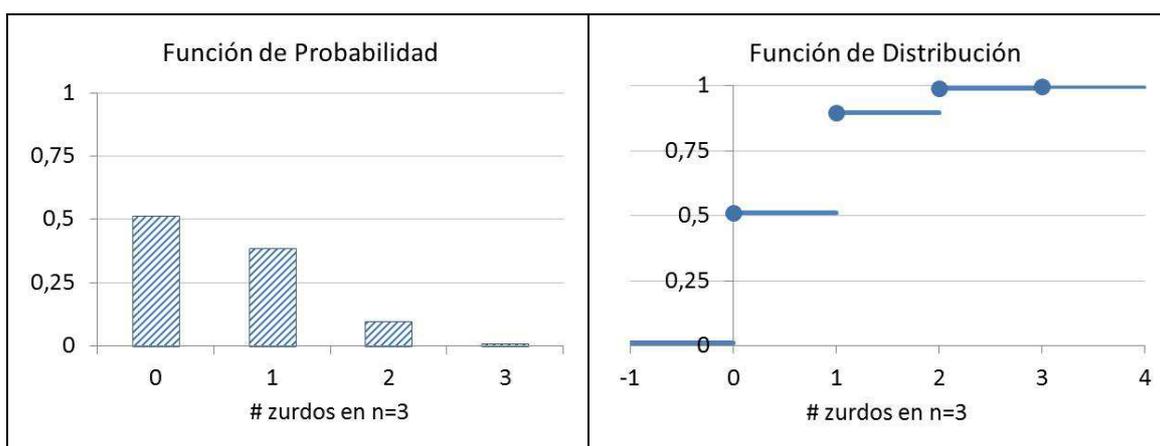
p = probabilidad de éxito (1-p= q =probabilidad de fracaso)

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

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

siendo el paréntesis vertical el número de combinaciones posibles⁷

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



⁵ El nombre se debe al Teorema del Binomio: $\sum P(x) = 1 = (p+(1-p))^n$

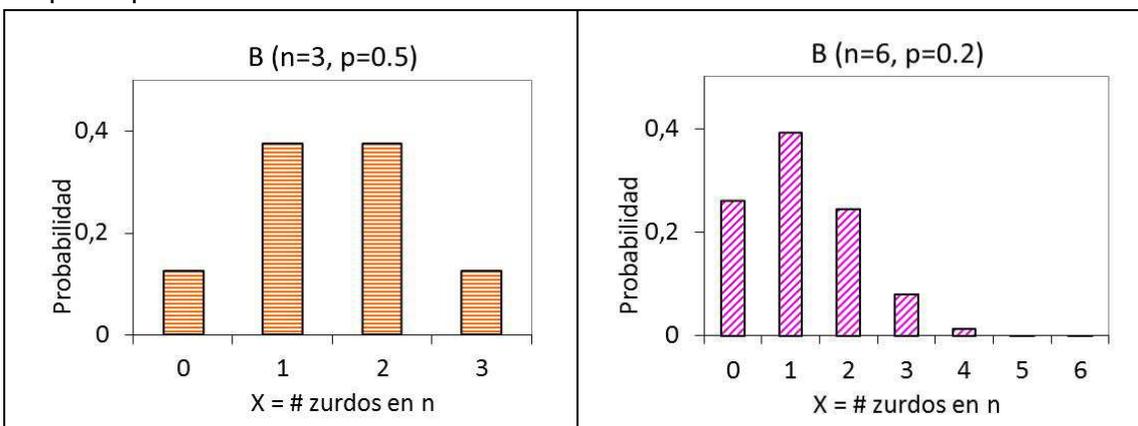
⁶ Después de cada extracción tiene que haber reemplazo, pero es desdeñable si el grupo es muy grande.

⁷ En las combinaciones se tienen en cuenta todos los resultados que dan el mismo recuento, sin importar el orden: (d,d,z) = (d,z,d) = (d,z,d).

L. Para que las probabilidades calculadas con el modelo sean válidas, se han de cumplir las siguientes condiciones:

- el mismo ensayo se repite n veces 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 uno no condiciona el siguiente.
- la probabilidad de éxito, p, se mantiene constante en todos los ensayos.

M. La forma de la función de probabilidad dependerá de los valores de los parámetros. El número de barras es igual a (n+1) y por tanto a mayor n más bajas son las barras. La simetría del conjunto sólo ocurre para p=0.5, mientras que para cualquier otro valor de p el reparto es asimétrico.



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) = n*p
- Variancia (X) = n*p*(1-p)

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

si $X \sim \text{Bernoulli}(p=0,2)$, entonces $[X + X + X] \sim B(n=3,p=0,2)$

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

MODELO DISCRETO: POISSON

A. El modelo de POISSON⁸ explica el reparto de probabilidad de una variable aleatoria que se ajuste a una expresión del tipo:

$X = \text{"número de acontecimientos raros en un intervalo de tiempo o espacio"}$

$R = \{0, 1, \dots, \infty\}$

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

B. Ejemplos

- número de soldado prusianos muertos por coz de caballo en un año⁹
- 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 decir que una variable aleatoria X sigue una distribución de probabilidad Poisson se utiliza la expresión:

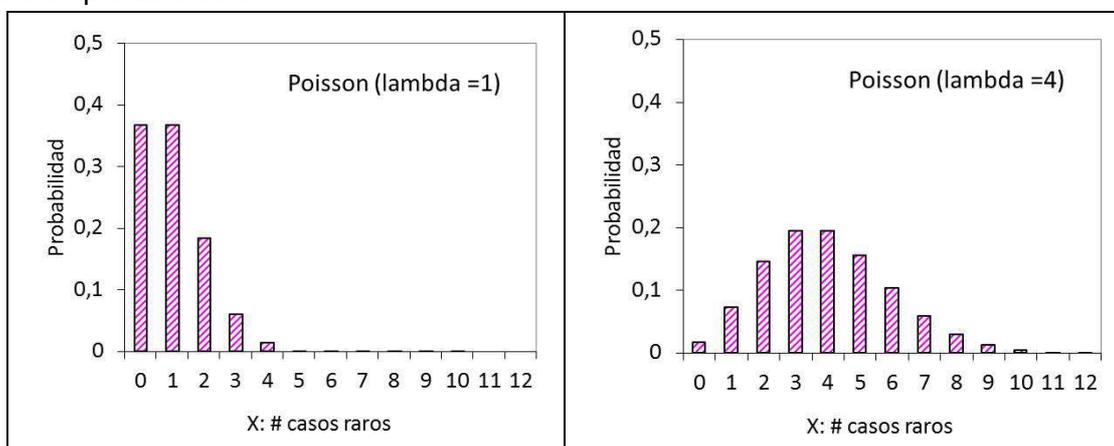
$X \sim \mathcal{P}(\lambda)$

siendo λ (λ) el único parámetro de la distribución

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

$$P(X = x) = \frac{\lambda^x}{x!} \exp(-\lambda); \quad x = 0, 1, \dots, \infty$$

E. La forma de la función de probabilidad dependerá sólo del valor λ . Teóricamente el recorrido de la variable acaba en infinito por eso el número de barras es siempre ilimitado.



F. Para que la aplicación del modelo sea válida se ha 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

- 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)

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



- la ocurrencia de un acontecimiento en un subintervalo no influye en lo que sucederá en el siguiente (independencia).

G. Los valores esperados del modelo Poisson son:

$$\text{Esperanza } (X) = \lambda$$

$$\text{Variancia } (X) = \lambda$$

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

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.

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

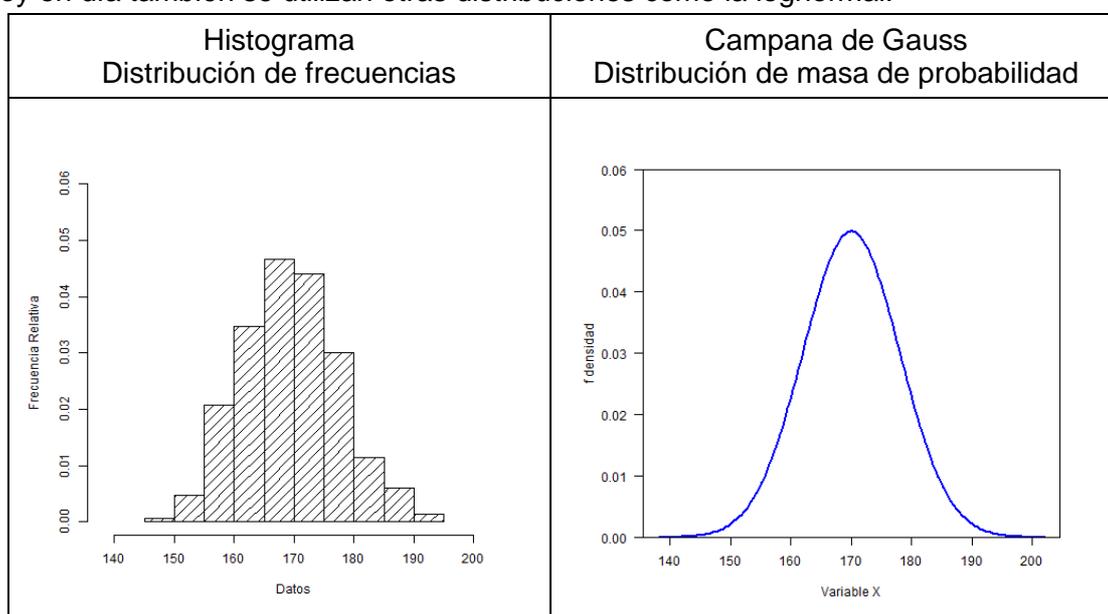
$$P(X \leq x) = \sum_{i=0}^x P(X = x)$$

I. 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:

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

MODELO CONTINUO: NORMAL

A. La distribución NORMAL es un objeto matemático. Fue la solución encontrada por A. de Moivre para aproximar cálculos de probabilidad en juegos de azar a principios del siglo XVIII. Posteriormente fue empleada para explicar los errores de medida en física y astronomía. Galileo ya había razonado que los errores de observación se distribuían simétricamente y tendían a agregarse en torno a su valor real. K.F. Gauss la utilizó en un trabajo que publicó en 1809 y su nombre acabó identificando la distribución porque la usó con a menudo cuando analizaba datos astronómicos. El protagonismo de esta distribución en las ciencias de la vida se debe al desarrollo de la Biometría. Hacia finales del siglo XIX se generalizó el uso de esta distribución de probabilidad para justificar la distribución de frecuencias de características biológicas. Hoy en día también se utilizan otras distribuciones como la lognormal.



B. 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*

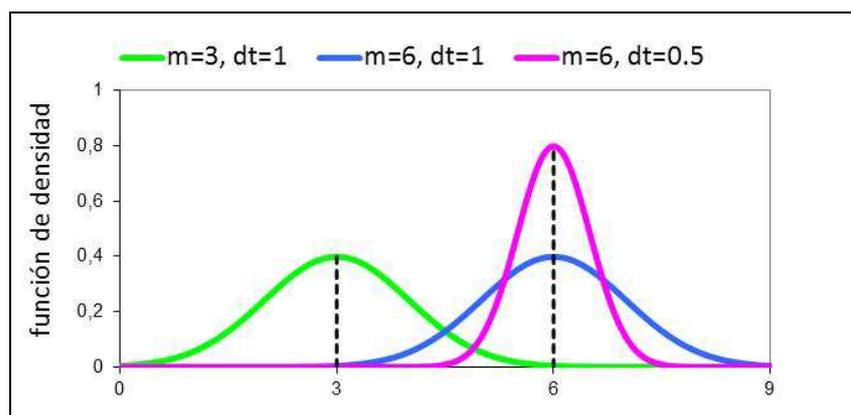
C. La campana de Gauss es la representación gráfica de la función de densidad, f , del modelo Normal cuya fórmula es:

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

D. Los parámetros del modelo son μ y σ . El primero, μ , es la esperanza de X y decide la posición de la campana a los largo del eje horizontal. El segundo, σ , es la raíz cuadrada de la variancia X y determina la anchura de la campana. Para decir que una variable aleatoria X sigue una distribución Normal se utiliza la expresión:

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

$$\text{siendo } \mu = E(X) \text{ y } \sigma = \sqrt{V(X)} = DT(X)$$

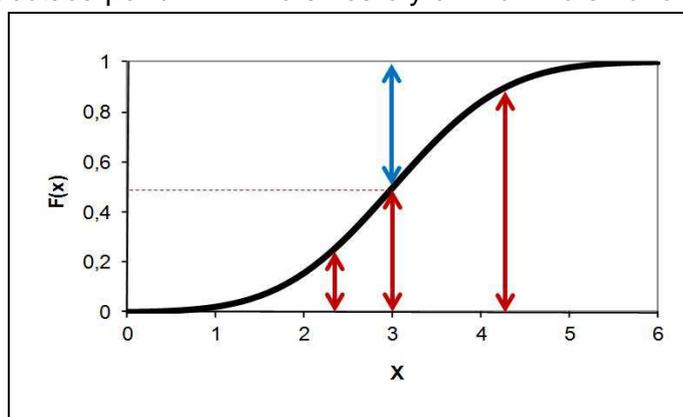


E. Características de la densidad normal

- por definición de densidad, el área total bajo la curva vale 1.
- 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 punto.
- la función tiene dos puntos de inflexión en las abscisas $(\mu-1\sigma)$ y $(\mu+1\sigma)$ respectivamente, con lo cual la desviación típica es la distancia desde el eje central a los puntos de inflexión.
- las colas son infinitas, es decir, la curva toca el eje X en $-\infty$ y $+\infty$.
- cualquier campana de Gauss cumple siempre las siguientes proporciones:
 - área entre $(\mu-1\sigma)$ y $(\mu+1\sigma)$ \approx 68%
 - área entre $(\mu-2\sigma)$ y $(\mu+2\sigma)$ \approx 95%
 - área entre $(\mu-3\sigma)$ y $(\mu+3\sigma)$ \approx 99%

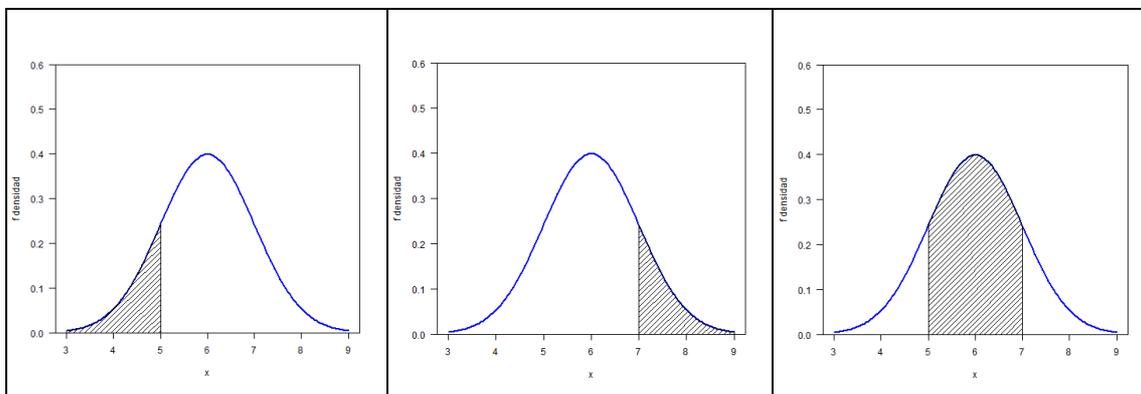
F. Una aplicación directa de este modelo es la definición de “intervalos de referencia” para interpretar resultados de bioquímica clínica o del coeficiente de inteligencia.

G. La función de distribución, F, de la Normal se obtiene integrando la función de densidad, pero no existe solución analítica y se tiene que resolver por métodos numéricos. 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. La representación gráfica es una línea en forma de “ese” acotada por un mínimo en cero y un máximo en uno.



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

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



E. Cualquier variable aleatoria puede ser transformada mediante operaciones matemáticas en una nueva variable con propiedades iguales o no a la original. En particular, el proceso de tipificación que consiste en restar la media y dividir por la desviación típica genera siempre una variable con media cero y desviación típica uno.

$$Y = \frac{X - E(X)}{DT(X)} \Rightarrow E(Y) = 0; DT(Y) = 1$$

Si la variable X original tiene distribución Normal, entonces la variable nueva hereda la distribución. Una variable con distribución Normal de parámetros cero y uno se la conoce como “normal tipificada” o sencillamente “zeta”.

$$X \sim N(\mu, \sigma) \rightarrow \text{tipificación} \rightarrow Z \sim N(0,1)$$

F. Este proceso de tipificación es importante para el cálculo de probabilidades, porque la F de 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:

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

G. 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

$$\mu_Y = E(X_1 + X_2) \quad \sigma_Y = \sqrt{V(X_1) + V(X_2)}$$

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.

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

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

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).

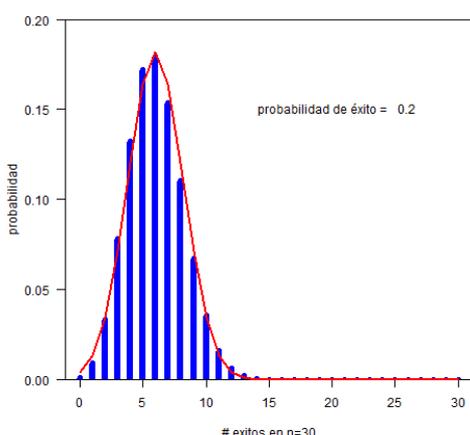
	P(X=0)	P(X=1)	P(X=2)	P(X=3)
B (1000,0.002)	0,13506	0,27067	0,27094	0,18063
Poisson (2)	0,13532	0,27067	0,27067	0,18045

C. De Binomial a Normal. El teorema de DeMoivre-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.

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

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

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



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

$$\begin{array}{ll} n \text{ grande} & n > 30 \\ p \text{ no pequeño} & \min(np, nq) > 5 \end{array}$$

Dado que el modelo Normal es continuo 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	0,37684
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”

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

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

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

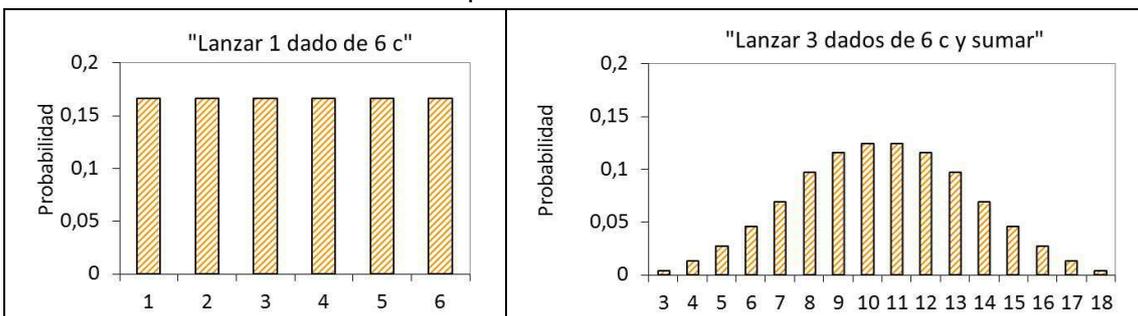
que tiene distribución uniforme, $P(X=x) = 1/6$, y valores esperados:

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

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

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

La distribución de S_n 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. continua
V.a. dicotómica
V.a. discreta
Variable tipificada
Variancia / varianza (poblacional)