

Matemáticas aplicadas a las Ciencias Sociales II

Variables Aleatorias. Distribuciones Binomial y Normal

Índice

1. Variables aleatorias	2
2. Distribución de probabilidad para variables aleatorias discretas	3
3. Parámetros de una variable aleatoria discreta	6
4. El experimento binomial	7
5. La distribución binomial	8
6. Distribución de probabilidad para variables aleatorias continuas	12
7. Parámetros de una variable aleatoria continua	13
8. La distribución normal	14
9. Distribución normal estándar. Tipificación de la variable	15
10. Aplicaciones de la distribución normal	18
10.1. Teorema central del límite	19
10.2. Aproximación de la binomial a la normal	19
10.2.1. Corrección por continuidad	20
11. Ejercicios	21
12. Tabla de la distribución binomial	26
13. Tabla de la distribución normal estándar	27

1. Variables aleatorias

Sabemos que existen numerosos experimentos cuyo resultado no puede predecirse antes de ser realizados, que se deben al azar. A estos experimentos los denominamos experimentos aleatorios. A cada prueba realizada de un experimento aleatorio podemos asociarle un valor numérico de la siguiente forma:

- Si el resultado del experimento es cuantitativo, asociaremos a cada realización del experimento el valor obtenido.
- Si el resultado es cualitativo, haremos corresponder a cada realización del experimento un número elegido según un criterio fijado de antemano.

Ejemplos:

- ✓ En el experimento “*número de llamadas que llegan a una centralita en diez minutos*”, el resultado es cuantitativo; luego el número de llamadas será el valor que asociamos a cada realización del experimento.
- ✓ En el experimento “*observar si el valor obtenido al lanzar un dado es mayor que cuatro*”, el resultado es cualitativo: ser mayor que cuatro o no ser mayor que cuatro. En este caso podemos adoptar como criterio para asociar un valor numérico a cada realización del experimento: el valor cero si el número obtenido es menor o igual que cuatro, y el valor uno, en caso contrario.

Sea E el espacio muestral asociado a un experimento aleatorio. Se llama **variable aleatoria** a toda aplicación que asocia a cada elemento del espacio muestral un número real.

Normalmente, las variables aleatorias suelen designarse con las letras mayúsculas X , Y , Z , $\dots$, y al conjunto de valores de $\mathbb{R}$ asignados a los elementos de E se le llama *recorrido* de la variable aleatoria.

Ejemplos:

- ✓ En el caso de la variable aleatoria “*número de llamadas recibidas en una centralita en diez minutos*”, el recorrido es $0, 1, 2, 3, \dots$.
- ✓ La variable aleatoria que al lanzar un dado le asocia el número cero si el número que aparece es menor o igual que cuatro, y el número uno en caso contrario, tiene como recorrido los valores 0 y 1 .

En estos dos ejemplos, las variables aleatorias consideradas tienen recorridos finitos; sin embargo, también es posible encontrar variables aleatorias con recorridos infinitos.

Ejemplos:

- ✓ Consideremos el experimento “*Realización consecutiva de dos sorteos de la lotería primitiva hasta obtener los números 19 y 37*”. La variable aleatoria que a cada realización del experimento le asocia el número veces que se ha repetido el mismo, podrá tomar como valor cualquier número entero mayor o igual a uno; así, el recorrido de esta variable es infinito.
- ✓ En el experimento “*medir el tiempo empleado por una atleta en correr los cien metros lisos*”, la variable aleatoria que asocia a cada realización del experimento esta medida, puede tomar infinitos valores.

Una variable aleatoria es **discreta** si toma un número finito de valores o un número infinito numerable.

Una variable aleatoria es **continua** si toma valores en un intervalo de la recta real; es decir, si su recorrido es infinito no numerable.

Esta clasificación es necesaria, ya que si tratamos de asociar a cada valor de la variable aleatoria la probabilidad de que se presente dicho valor, tenemos que:

- Si la variable aleatoria es discreta, siempre es posible asociar a cada uno de sus valores la probabilidad de que se produzca este valor, de forma que esta probabilidad verifique lo siguiente:
 - No siempre se anula.
 - La suma de las probabilidades asociadas a cada uno de los valores de las variables es uno.
- Si la variable aleatoria es continua, no podemos asociar a cada valor que pueda tomar esta variable la probabilidad de que éste ocurra, de forma que esta probabilidad cumpla las dos condiciones anteriores. Esto es debido a que si en un intervalo de la recta real tratamos de hallar la probabilidad de un número de este intervalo con un valor concreto a , observaremos que, según la regla de Laplace, esta probabilidad es el número de casos favorables, que es uno (el punto a), dividido entre el número de casos posibles (infinito) y, por tanto, tendremos que concluir que esta probabilidad es cero, de lo que se deduciría que éste es un suceso imposible; sin embargo, que la variable aleatoria tome el valor a no es un suceso imposible.

2. Distribución de probabilidad para variables aleatorias discretas

En el apartado anterior hemos indicado que si $\{x_1, x_2, \dots\}$ son los valores que puede tomar una variable aleatoria discreta X , siempre es posible hacer corresponder a cada uno de estos valores la probabilidad de que al realizar el experimento obtengamos este valor.

Se llama **función de probabilidad** de una variable aleatoria discreta X a la aplicación que a cada valor x_i de la variable le hace corresponder la probabilidad de que la variable tome dicho valor.

$$f(x_i) = P(X = x_i)$$

Ejemplo:

- ✓ En el experimento que consiste en tirar tres veces a una diana, la variable aleatoria que asocia a cada realización de este experimento el número de aciertos en el blanco puede tomar los valores 0, 1, 2 ó 3.

Si sabemos que la persona que lanza la flecha tiene una probabilidad 0,4 de acertar, podremos calcular la función de probabilidad de esta variable utilizando la teoría de probabilidades: $P(X = 0) = 0,6^3 = 0,216$, $P(X = 1) = 3 \cdot 0,6^2 \cdot 0,4 = 0,432$, $P(X = 2) = 3 \cdot 0,6 \cdot 0,4^2 = 0,288$, $P(X = 3) = 0,4^3 = 0,064$

Obsérvese que, como ya se había comentado anteriormente, la suma de las probabilidades anteriores es igual a 1:

$$P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) = 0,216 + 0,432 + 0,288 + 0,064 = 1$$

Como consecuencia de las propiedades de la probabilidad de un suceso, es fácil deducir que la función de probabilidad verifica:

- Como $0 \leq P(X = x_i) \leq 1$, se tiene que $0 \leq f(x_i) \leq 1$.
- Si $\{x_1, x_2, \dots\}$ son los valores toma la variable aleatoria X , se cumple que $\sum P(X = x_i) = 1$, ya que esta suma es, en realidad, la probabilidad del suceso seguro. Por tanto también se verifica que $\sum f(x_i) = 1$.

La función de probabilidad de una variable aleatoria discreta se describe, normalmente, mediante una tabla de valores. Su representación gráfica puede realizarse sobre unos ejes de coordenadas, indicando en el eje OX los distintos valores de la variable aleatoria $\{x_1, x_2, \dots\}$ y levantando sobre cada valor x_i barras de altura proporcional a $f(x_i)$.

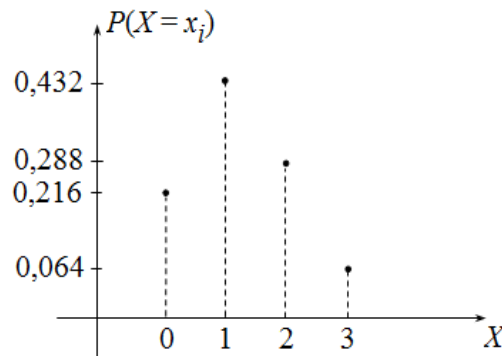


Figura 1: Función de probabilidad

Ejemplo:

- ✓ La representación gráfica de la variable aleatoria del experimento considerado en el ejemplo anterior es la de la Figura 1 y su descripción en una tabla de valores es la que tienes a continuación:

x_i	0	1	2	3
$P(X = x_i)$	0,216	0,432	0,288	0,064

En ocasiones nos puede interesar conocer no sólo la probabilidad de que la variable X tome el valor concreto x_i , sino la probabilidad de que la variable tome valores menores o iguales que x_i . Para ello utilizaremos la función de distribución, que se define de la siguiente forma.

Dada una variable aleatoria discreta X , su **función de distribución** es la aplicación que a cada valor x de la variable le asigna la probabilidad de que ésta tome valores menores o iguales que x :

$$F(x) = P(X \leq x)$$

De la definición es fácil deducir que si la función de probabilidad de una variable aleatoria discreta X es

X	x_1	x_2	$\dots$	x_n
$f(x_i) = P(X = x_i)$	p_1	p_2	$\dots$	p_n

entonces, la función de distribución de dicha variable será:

$$F(x) = \begin{cases} 0 & \text{si } x < x_1 \\ p_1 & \text{si } x_1 \leq x < x_2 \\ p_1 + p_2 & \text{si } x_2 \leq x < x_3 \\ \dots & \dots \\ p_1 + p_2 + \dots + p_n & \text{si } x_{n-1} \leq x < x_n \\ 1 & \text{si } x_n \leq x \end{cases}$$

Ejemplo:

- ✓ La función de distribución de la variable aleatoria que hemos considerado en los dos ejemplos anteriores es:

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ 0,216 & \text{si } 0 \leq x < 1 \\ 0,648 & \text{si } 1 \leq x < 2 \\ 0,936 & \text{si } 2 \leq x < 3 \\ 1 & \text{si } 3 \leq x \end{cases}$$

Y su representación gráfica es la que se muestra en la Figura 2:

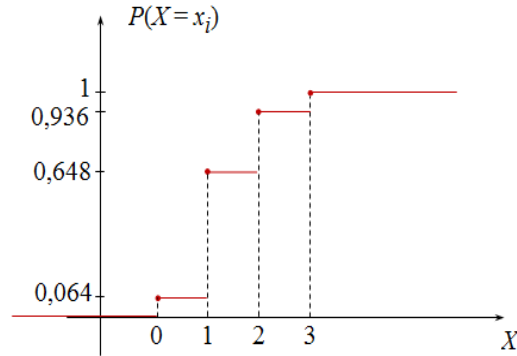


Figura 2: Representación gráfica de la función de distribución

La función de distribución $F(x)$ de cualquier variable aleatoria discreta X verifica las siguientes propiedades:

- Es una función escalonada, ya que es constante en cada intervalo.
- Es una función creciente, porque es una suma acumulativa de probabilidades que son siempre positivas.
- En cada intervalo $[x_i, x_{i+1})$, $F(x)$ es continua a la derecha de x_i y discontinua en x_{i+1} , $\forall i \in 0, 1, 2, \dots, n-1$.
- Como $F(x)$ es una probabilidad, se cumple que $0 \leq F(x) \leq 1$.
- Si $x_i < x_j$, entonces $F(x_j) - F(x_i) = P(x_i < x \leq x_j)$.

3. Parámetros de una variable aleatoria discreta

Sea una variable aleatoria discreta X cuya función de probabilidad es:

X	x_1	x_2	$\dots$	x_n
$f(x_i) = P(X = x_i)$	p_1	p_2	$\dots$	p_n

A partir de esta distribución podemos definir una serie de medidas características de la variable aleatoria denominadas parámetros. La primera de estas medidas es la media o valor esperado de la variable X ; además definiremos la varianza y la desviación típica, que nos servirán para medir el grado de dispersión de dicha variable respecto de la media. Dada una variable aleatoria discreta X :

- La **media** o **esperanza matemática** de la variable X , μ , viene dada por la expresión:

$$\mu = x_1p_1 + x_2p_2 + \dots + x_np_n = \sum_{i=1}^n x_ip_i$$

- La **varianza** de la variable X , σ^2 , viene dada por la expresión:

$$\sigma^2 = (x_1 - \mu)^2 \cdot p_1 + (x_2 - \mu)^2 \cdot p_2 + \dots + (x_n - \mu)^2 \cdot p_n = \sum_{i=1}^n (x_i - \mu)^2 p_i = \sum_{i=1}^n x_i^2 p_i - \mu^2$$

- La **desviación típica** de la variable X , σ , es la raíz cuadrada de la varianza:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\sum_{i=1}^n (x_i - \mu)^2 p_i}$$

Ejemplo:

- ✓ Volvamos a considerar la variable aleatoria que asocia a cada experimento consistente en la realización de tres tiros a una diana, el número de aciertos. Si queremos hallar el número de aciertos esperado y el grado de dispersión respecto a la media, calcularemos la media y la varianza o la desviación típica. Para simplificar tanto el cálculo de la media como el de la varianza utilizaremos la siguiente tabla:

x_i	p_i	$x_i \cdot p_i$	$x_i^2 \cdot p_i$
0	0,216	0	0
1	0,432	0,432	0,432
2	0,288	0,576	1,152
3	0,064	0,192	0,576
Suma	1	1,2	2,16

Así $\mu = \sum_{i=1}^4 x_i p_i = 1,2$ (valor esperado de aciertos en tres lanzamientos) y, por otro lado,

$$\sigma^2 = \sum_{i=1}^4 x_i^2 p_i - \mu^2 = 2,16 - 1,2^2 = 0,72 \Rightarrow \sigma = \sqrt{0,72} = 0,85.$$

4. El experimento binomial

Son numerosos los estudios políticos, sociológicos, industriales, educativos... en los que se trata de precisar la proporción de población que posee un determinado atributo. Así, es frecuente que un partido político quiera saber el número de personas que son favorables a su candidato, que un sociólogo desee estudiar la proporción de viviendas que poseen calefacción, que un canal de televisión intente conocer el número de telespectadores que lo prefieren o que el Ministerio de Educación esté interesado en saber qué proporción de profesores está a favor de la introducción de las nuevas tecnologías en el aula. Todos estos experimentos son experimentos binomiales.

Llamamos **experimento binomial** al que cumple las siguientes propiedades:

- El experimento consiste en n ensayos idénticos.
- Cada ensayo tiene dos resultados posibles: *éxito* y *fracaso*.
- La probabilidad de éxito en un ensayo permanece constante a lo largo de todo el experimento y la denotamos por p , de forma que la probabilidad de fracaso en cada ensayo es $q = 1 - p$.
- Los ensayos son independientes, es decir, el resultado obtenido en uno de ellos no influye para nada en el resultado de cualquier otro.

Por tanto, para determinar si un experimento es o no binomial deberemos asegurarnos de que verifica estas cuatro propiedades.

Ejemplo

- ✓ Supongamos que en un país existe una proporción p de individuos a favor del acceso libre de inmigrantes y una proporción $q = 1 - p$ en contra. Seleccionamos una muestra de mil personas y les preguntamos si están a favor o en contra de dicha medida. ¿Será éste un experimento binomial?
 - El experimento consiste en mil ensayos iguales.
 - Cada ensayo sólo tiene dos resultados posibles: a favor o en contra
 - La probabilidad p de estar a favor de esta medida no varía de un ensayo a otro. Aunque para que la probabilidad realmente no varíe de un individuo a otro, la selección debe realizarse reemplazando los individuos de la muestra, esto es, teniendo en cuenta la posibilidad de volverlos a seleccionar. Pero si el tamaño de la población es infinito o muy grande en comparación con el tamaño de la muestra, los resultados obtenidos con una selección sin reemplazamiento son similares a los obtenidos en una selección con reemplazamiento.
 - El resultado de cada ensayo o encuesta es independiente de los resultados obtenidos en otros ensayos pues, si la encuesta se realiza correctamente, la respuesta de una persona no influye sobre la respuesta de cualquier otra.

Por tanto podemos concluir que sí que se trata de un experimento binomial, pues cumple las cuatro propiedades asociadas al mismo.

5. La distribución binomial

Si consideramos la variable aleatoria X que asocia a cada realización de un experimento binomial el número de éxitos obtenidos, estaremos ante una variable aleatoria binomial.

La variable aleatoria X sigue una **distribución binomial o de Bernoulli** con parámetros n y p , $B(n, p)$, si cumple las siguientes condiciones:

- Se han realizado n ensayos de un experimento binomial.
- En cada ensayo, la probabilidad de éxito es p .
- La variable aleatoria X asocia a cada realización del experimento binomial el número de ensayos con éxito.

Ejemplo

- ✓ Supongamos que lanzamos seis veces una moneda y que consideramos éxito el obtener cara. La variable aleatoria que asocia a cada seis lanzamientos de la moneda el número de caras obtenido es una variable aleatoria binomial en la que $n = 6$, pues lanzamos la moneda seis veces (seis ensayos), y $p = \frac{1}{2} = 0,5$ porque la probabilidad de obtener cara en un lanzamiento es 0,5.

Es evidente que una variable aleatoria binomial con parámetros n y p , $B(n, p)$ es una variable aleatoria discreta, pues sólo puede tomar como valores $0, 1, 2, 3, \dots, n$ y, por tanto, tendrá asociada una función de probabilidad y una función de distribución.

Para hallar la *función de probabilidad* tendremos que calcular la probabilidad de que en n ensayos haya i éxitos y $n - i$ fracasos, con i variando desde 0 hasta n . Esta probabilidad se calcula fácilmente utilizando la teoría de combinatoria y conocimientos de probabilidad adquiridos en un tema anterior. Así obtenemos que la función de probabilidad de una distribución binomial $B(n, p)$ es:

i	$P(X = i)$
0	$\binom{n}{0} \cdot (1 - p)^n$
1	$\binom{n}{1} \cdot p \cdot (1 - p)^{n-1}$
2	$\binom{n}{2} \cdot p^2 \cdot (1 - p)^{n-2}$
...	...
k	$\binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k}$
...	...
$n - 1$	$\binom{n}{n-1} \cdot p^{n-1} \cdot (1 - p)$
n	$\binom{n}{n} \cdot p^n$

A partir de aquí es fácil deducir que la *función de distribución* es:

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ P(X \leq x) = \binom{n}{0} \cdot (1-p)^n + \dots + \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} & \text{si } x_1 \leq x < x_2 \\ 1 & \text{si } n \leq x \end{cases}$$

siendo k el mayor entero no superior a x .

En el caso de una distribución binomial es sencillo determinar la media, la varianza y la desviación típica conocidos los valores de n y p . No se realizará la demostración aquí, pero estos parámetros vienen dados por las siguientes expresiones:

- **Media:** $\mu = np$.
- **Varianza:** $\sigma^2 = npq = np(1-p)$.
- **Desviación típica:** $\sigma = \sqrt{npq} = \sqrt{np(1-p)}$.

Ejemplo

✓ La función de probabilidad de la variable aleatoria del ejemplo anterior es:

i	$P(X = i)$
0	$\binom{6}{0} \cdot \left(\frac{1}{2}\right)^6 = \left(\frac{1}{2}\right)^6 = \frac{1}{64}$
1	$\binom{6}{1} \cdot \left(\frac{1}{2}\right)^1 \cdot \left(\frac{1}{2}\right)^5 = 6 \cdot \left(\frac{1}{2}\right)^6 \frac{6}{64} = \frac{3}{32}$
2	$\binom{6}{2} \cdot \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^4 = 15 \cdot \left(\frac{1}{2}\right)^6 \frac{15}{64}$
3	$\binom{6}{3} \cdot \left(\frac{1}{2}\right)^3 \cdot \left(\frac{1}{2}\right)^3 = 20 \cdot \left(\frac{1}{2}\right)^6 \frac{20}{64} = \frac{5}{16}$
4	$\binom{6}{4} \cdot \left(\frac{1}{2}\right)^4 \cdot \left(\frac{1}{2}\right)^2 = 15 \cdot \left(\frac{1}{2}\right)^6 \frac{15}{64}$
5	$\binom{6}{5} \cdot \left(\frac{1}{2}\right)^5 \cdot \left(\frac{1}{2}\right)^1 = 6 \cdot \left(\frac{1}{2}\right)^6 \frac{6}{64} = \frac{3}{32}$
6	$\binom{6}{6} \cdot \left(\frac{1}{2}\right)^6 = \left(\frac{1}{2}\right)^6 = \frac{1}{64}$

A partir de la función de probabilidad obtenemos que la función de distribución es:

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{1}{64} & \text{si } 0 \leq x < 1 \\ \frac{1}{64} + \frac{6}{64} = \frac{7}{64} & \text{si } 1 \leq x < 2 \\ \frac{1}{64} + \frac{6}{64} + \frac{15}{64} = \frac{22}{64} & \text{si } 2 \leq x < 3 \\ \frac{1}{64} + \frac{6}{64} + \frac{15}{64} + \frac{20}{64} = \frac{42}{64} & \text{si } 3 \leq x < 4 \\ \frac{1}{64} + \frac{6}{64} + \frac{15}{64} + \frac{20}{64} + \frac{15}{64} = \frac{57}{64} & \text{si } 4 \leq x < 5 \\ \frac{1}{64} + \frac{6}{64} + \frac{15}{64} + \frac{20}{64} + \frac{15}{64} + \frac{63}{64} = \frac{63}{64} & \text{si } 5 \leq x < 6 \\ \frac{1}{64} + \frac{6}{64} + \frac{15}{64} + \frac{20}{64} + \frac{15}{64} + \frac{63}{64} + \frac{1}{64} = \frac{64}{64} = 1 & \text{si } 6 \leq x \end{cases}$$

Los parámetros de esta variable aleatoria son:

- **Media:** $\mu = np = 6 \cdot \frac{1}{2} = 3$
- **Varianza:** $\sigma^2 = npq = 6 \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{3}{2} = 1,5$
- **Desviación típica:** $\sigma = \sqrt{npq} = \sqrt{\frac{3}{2}} = \frac{\sqrt{6}}{2} \simeq 1,225$

La media se puede interpretar como el número esperado de caras al lanzar seis veces una moneda (a la media también se le llama **esperanza matemática**). Por eso se obtiene que la media es 3 (sólo hay que pensar que si este experimento se repite “muchas veces”, cualquiera concluiría que el promedio de caras es 3).

La desviación típica mide la media de las desviaciones a la media. Es como decir que los resultados del experimento se encuentran, mayormente, a razón de $\pm\sigma$ alrededor de la media. Dicho de otro modo, “la mayor parte de los resultados obtenidos” se encuentran en el intervalo $(\mu - \sigma, \mu + \sigma)$. En este ejemplo es bastante probable que, la “mayoría de las veces” que se repite el experimento “lanzar seis veces una moneda”, el número de caras que se obtienen esté en el intervalo $(3 - 1,225, 3 + 1,225) = (1,775, 4,225)$, es decir, que la “mayoría de las veces” que se repite el experimento, saldrán 2, 3 ó 4 caras.

Al estudiar los intervalos de confianza se verá exactamente cuál es la probabilidad de que los resultados del experimento se encuentren en tal intervalo.

Por cierto, al final del tema hay una tabla de la distribución binomial para hallar probabilidades sin necesidad de hacer cálculos con números combinatorios.

6. Distribución de probabilidad para variables aleatorias continuas

Como se mencionó al principio de este tema, existen experimentos en los que la variable aleatoria puede tomar cualquier valor de un intervalo de la recta real, por ejemplo, el peso de un niño al nacer, la altura de los jugadores de baloncesto, el tiempo de duración de un anuncio de televisión... En estos casos decimos que la variable aleatoria asociada a este experimento es una variable aleatoria continua.

En el caso de variables aleatorias continuas no es posible conocer el valor exacto que toman en un punto; así, si nos dicen que el peso de un niño al nacer es de tres kilos y cien gramos, todo lo que podremos afirmar es que el peso real de este niño está en torno a los tres kilos y cien gramos.

Para realizar un estudio estadístico sobre una variable aleatoria continua dividiremos el recorrido de la variable en intervalos y representaremos las medidas obtenidas mediante un histograma.

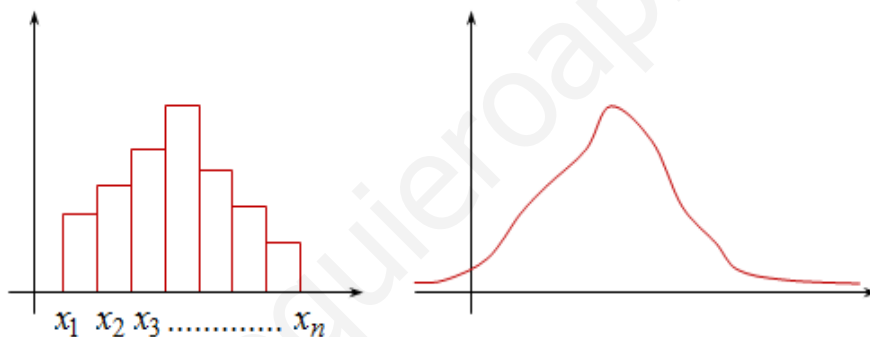


Figura 3: Histograma y curva aproximativa

Si consideramos intervalos más pequeños, obtendremos un nuevo histograma y, si repetimos este proceso sucesivamente, el histograma tenderá a aproximarse a una curva que describe el comportamiento de la variable aleatoria. Esta curva corresponde a la gráfica de una función $f(x)$ denominada función de densidad y representa la distribución de probabilidad de las variables aleatorias continuas.

Una función $y = f(x)$ es la **función de densidad** de una variable aleatoria continua X si cumple que:

- $f(x)$ es mayor o igual que 0 para todo x del intervalo en el que está definida.
- El área bajo la curva es igual a 1.
- La probabilidad de que la variable tome valores en el intervalo $[x_1, x_2]$ es el área bajo la curva correspondiente a dicho intervalo.

Como consecuencia de la tercera propiedad de la definición de función de densidad y del cálculo integral (el área bajo una curva se puede calcular, en ciertas condiciones, mediante una integral definida), podemos obtener la probabilidad de que la variable aleatoria X tome valores entre x_1 y x_2 :

$$P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} f(x)dx$$

Al igual que en el caso discreto, a partir de la definición de probabilidad podemos definir la función de distribución de una variable aleatoria continua.

Dada una variable aleatoria continua X , su función de distribución es la aplicación que a cada valor x de la variable le asigna la probabilidad de que ésta tome valores menores o iguales que x :

$$F(x) = P(X \leq x)$$

Si la función de densidad toma valores distintos de cero en el intervalo $[a, b]$, entonces la función de distribución asociada será:

$$F(x) = \begin{cases} 0 & \text{si } x < a \\ \int_a^b f(x)dx & \text{si } a \leq x < b \\ 1 & \text{si } b \leq x \end{cases}$$

7. Parámetros de una variable aleatoria continua

Al igual que en el caso de variables aleatorias discretas, para las variables aleatorias continuas podemos hallar una serie de parámetros que nos servirán para caracterizarlas. Estos parámetros son la media o el valor esperado de la variable aleatoria X , la varianza y la desviación típica.

Si X es una variable aleatoria continua que toma valores en el intervalo $[a, b]$ y $f(x)$ es su función de densidad:

- La **media** o **esperanza matemática**, μ , de la variable X es:

$$\mu = \int_a^b xf(x)dx$$

- La **varianza** σ^2 de X viene dada por:

$$\sigma^2 = \int_a^b (x - \mu)^2 f(x)dx$$

- La **desviación típica** σ de X es la raíz cuadrada de la varianza:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\int_a^b (x - \mu)^2 f(x)dx}$$

La media representa el promedio del conjunto de todas las posibles observaciones; por tanto, es una *medida de centralización* que trata de describir con un único número la localización general de un conjunto de valores. La varianza y la desviación típica son, sin embargo, *medidas de dispersión*, ya que miden el grado de dispersión de los datos respecto de la media (véase la observación al final del último ejemplo de la sección dedicada a la distribución binomial).

8. La distribución normal

La distribución más importante dentro de las distribuciones continuas es la distribución normal. Ésta es una distribución teórica, ya que su función de densidad viene dada por una fórmula matemática y nunca coincidirá exactamente con los datos observados empíricamente; sin embargo, debe su importancia a que:

- Numerosas variables aleatorias siguen este modelo. Por ejemplo, el peso de los niños al nacer, la altura de los jugadores de baloncesto, la duración de las pilas de una determinada marca, etc.
- Algunas distribuciones pueden aproximarse, como veremos al estudiar el teorema central del límite, a distribuciones normales.
- Las distribuciones muestrales de varios estadísticos, que se estudian en el tema de inferencia estadística, y que son de gran importancia para la obtención de información de una población a partir de una muestra, siguen distribuciones normales si los tamaños de las muestras son grandes.

Una variable aleatoria continua X sigue una distribución normal de media μ y desviación típica σ , y la denotaremos $N(\mu, \sigma)$, si se cumple que:

- a) El recorrido de la variable X es toda la recta real.
- b) Su función de densidad es:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Una distribución normal queda totalmente descrita conociendo su media μ y su desviación típica σ . La gráfica de este tipo de distribuciones tiene forma de una campana que se conoce con el nombre de *campana de Gauss*. A partir de esta gráfica podemos deducir que la distribución normal verifica que:

- Su campo de existencia es toda la recta real.

- Es simétrica respecto de la recta $x = \mu$ (recta vertical que pasa por $x = \mu$).
- La función posee un máximo en el punto de abscisa $x = \mu$ y dos puntos de inflexión en los puntos de abscisa $x_1 = \mu - \sigma$ y $x_2 = \mu + \sigma$.
- El eje de abscisas es una asíntota horizontal de la curva.
- En toda distribución normal, el área encerrada bajo la curva es uno y se distribuye en intervalos de la siguiente forma:
 - En el intervalo $(\mu - \sigma, \mu + \sigma)$, el área encerrada es 0,6826, es decir el 68,26 % del total.
 - En el intervalo $(\mu - 2\sigma, \mu + 2\sigma)$, el área encerrada es 0,9544, esto es, el 95,44 % del total.
 - En el intervalo $(\mu - 3\sigma, \mu + 3\sigma)$, el área encerrada es 0,9973, que equivale al 99,73 % del total.

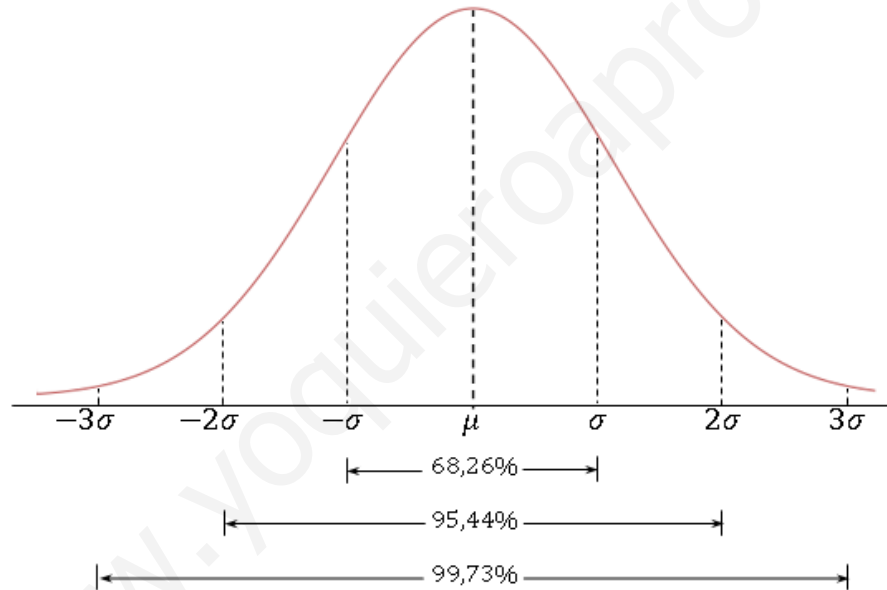


Figura 4: Campana de Gauss

9. Distribución normal estándar. Tipificación de la variable

Entre todas las distribuciones normales hay una de interés especial, aquella cuya media es cero y cuya desviación típica es uno.

Se conoce como **distribución normal estándar** a la distribución normal $N(\mu, \sigma)$ en la que $\mu = 0$ y $\sigma = 1$; es decir, a la distribución $N(0, 1)$.

La función de densidad de esta distribución es:

$$f(z) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{z^2}{2}}$$

La gran ventaja de la distribución normal estándar, $N(0, 1)$, y que desde ahora notaremos con la letra Z , es que se encuentra tabulada, y ello nos permite calcular fácilmente las probabilidades asociadas a los distintos valores de la variable.

Existen diversos modelos de tablas para la distribución normal estándar; al final del tema encontrarás una de fácil manejo. En ella se exponen las áreas comprendidas desde $-\infty$ hasta el punto z ; siendo $z > 0$.

El valor del área encerrada por la curva desde $-\infty$ hasta z , o sea, la probabilidad de que la variable Z tome un valor menor o igual que z ($P(Z \leq z)$), es el valor que aparece en la intersección de la fila que contiene las unidades y las décimas de z , y la columna que contiene sus centésimas.

Ejemplo

✓ Para calcular $P(Z \leq 0,67)$ tendremos que buscar en la tabla la fila correspondiente a 0,6 (unidades y décimas) y la columna correspondiente a 0,07 (centésimas). La intersección de la fila y la columna mencionadas es 0,7486 y, por tanto: $P(Z \leq 0,67) = 0,7486$.

Es importante observar que $P(Z \leq z) = P(Z < z)$ pues al tratarse de una variable aleatoria continua la probabilidad de tomar exactamente un valor es igual a cero: $P(Z = z) = 0$ (en la gráfica se correspondería exactamente al área de un segmento, que es igual a 0).

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389

Figura 5: Búsqueda en la tabla de la distribución normal estándar

Con la tabla podremos calcular cualquier probabilidad si tenemos en cuenta que el área total encerrada por la curva es la unidad y que ésta es simétrica respecto del eje de ordenadas (eje OY). Veamos los distintos casos que se nos pueden presentar, considerando que $z > 0$:

- $P(Z \leq z)$ es el valor obtenido directamente al leer la tabla, tal y como se ha visto en el ejemplo anterior.
- $P(Z \geq z) = 1 - P(Z \leq z)$.

- c) $P(Z \leq -z) = P(Z \geq z) = 1 - P(Z \leq z)$.
- d) $P(Z \geq -z) = 1 - P(Z \leq -z) = 1 - P(Z \geq z) = 1 - (1 - P(Z \leq z)) = P(Z \leq z)$.
- e) $P(z_1 \leq Z \leq z_2) = P(Z \leq z_2) - P(Z \leq z_1)$.

Obsérvese que en todos los casos se puede reducir el cálculo de la probabilidad al cálculo de una probabilidad del tipo $P(Z \leq z)$, que son las que podemos buscar fácilmente en la tabla de la distribución normal estándar. Además no es conveniente memorizar estos casos, sino que es mucho más recomendable tener en mente la gráfica de la función de probabilidad de la distribución normal estándar y sus propiedades, para deducir cada uno de ellos.

Por último se insiste de nuevo en que si, en cualquiera de estos casos, en lugar de las desigualdades dadas: $\leq$ o $\geq$ aparecen las desigualdades en sentido estricto: $<$ o $>$, el cálculo se efectuaría de la misma forma pues, como ya se ha comentado anteriormente, en las distribuciones continuas se tiene que $P(X = x) = 0$.

Ejemplo

✓ Si Z sigue la distribución normal estándar $N(0, 1)$, entonces:

- $P(Z \geq 0,67) = 1 - P(Z \leq 0,67) = 1 - 0,7486 = 0,2514$.
- $P(Z \leq -0,67) = 1 - P(Z \leq 0,67) = 1 - 0,7486 = 0,2514$.
- $P(Z \geq -0,67) = P(Z \leq 0,67) = 0,7486$.
- $P(-0,67 \leq Z \leq 0,67) = P(Z \leq 0,67) - P(Z \leq -0,67) = 0,7486 - 0,2514 = 0,4972$.

Análogamente:

- $P(Z > 0,67) = 1 - P(Z < 0,67) = 1 - 0,7486 = 0,2514$.
- $P(Z < -0,67) = 1 - P(Z < 0,67) = 1 - 0,7486 = 0,2514$.
- $P(Z > -0,67) = P(Z < 0,67) = 0,7486$.
- $P(-0,67 < Z < 0,67) = P(Z < 0,67) - P(Z < -0,67) = 0,7486 - 0,2514 = 0,4972$.

Para calcular probabilidades en una distribución normal cualquiera, $X \rightsquigarrow N(\mu, \sigma)$, vamos a poder utilizar los cálculos que acabamos de hacer para la distribución normal estándar $Z \rightsquigarrow N(0, 1)$, que, como ya se ha comentado, es la única distribución normal tabulada.

Para ello tendremos que transformar, mediante un proceso conocido como *tipificación de la variable*, la variable X de distribución $N(\mu, \sigma)$ en la variable Z de distribución $N(0, 1)$.

La tipificación de X consiste en realizar el siguiente cambio de variable:

$$Z = \frac{X - \mu}{\sigma}$$

Si en el cambio de variable anterior despejamos X tenemos:

$$X = \sigma Z + \mu$$

De esta forma se verifica que

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

Obsérvese que, en el último paso, no cambia el sentido de la desigualdad ya que $\sigma > 0$.

Ejemplo

- ✓ Si una variable X sigue una distribución normal $N(10, 2)$, para calcular, por ejemplo, la probabilidad $P(X \leq 11)$, procedemos a tipificar la variable X , con lo que $Z = \frac{X - 10}{2}$. De esta forma deducimos que:

$$P(X \leq 11) = P\left(Z \leq \frac{11 - 10}{2}\right) = P(Z \leq 0,5) = 0,6915$$

10. Aplicaciones de la distribución normal

Al comenzar el estudio de esta distribución, decíamos que son numerosos los fenómenos que en su representación se aproximaban a la gráfica de una curva normal y, por tanto, podíamos estudiarlos a partir del conocimiento de esta distribución.

Ejemplo

- ✓ Una de las primeras aplicaciones de la curva normal fue realizada por el matemático y astrónomo F.W. Bessel en 1818. Bessel comprobó que los errores de 300 medidas astronómicas se ajustaban a una campana de Gauss.

Si suponemos que esta campana es la de una normal de media cero y desviación típica ocho grados, podremos calcular, por ejemplo, qué tanto por ciento de los errores cometidos son menores de dos grados.

Para ello basta con calcular $P(-2 \leq X \leq 2)$ donde X sigue una distribución normal $N(0, 8)$. Tipificando la variable X , tenemos que $Z = \frac{X - 0}{8} = \frac{X}{8}$, donde ya sabemos Z sigue una distribución normal estándar $N(0, 1)$. Por tanto:

$$\begin{aligned} P(-2 \leq X \leq 2) &= P\left(-\frac{2}{8} \leq Z \leq \frac{2}{8}\right) = P(-0,25 \leq Z \leq 0,25) = \\ &= P(Z \leq 0,25) - P(Z \leq -0,25) = 2 \cdot P(Z \leq 0,25) - 1 = 0,1974 \end{aligned}$$

Es decir, el 19,74 % de las observaciones tendrá un error inferior a 2 grados.

Otra de las aplicaciones de la distribución normal es la que nos permite aproximar una serie de distribuciones a distribuciones normales. Esta aproximación se deduce del teorema central del límite.

10.1. Teorema central del límite

El **teorema central del límite** dice que si $X_1, X_2, \dots, X_n$ son n variables aleatorias independientes con medias $\mu_1, \mu_2, \dots, \mu_n$ y varianzas $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$, entonces la variable $Y = X_1 + X_2 + \dots + X_n$, cuando n es “grande”, se aproxima a una distribución normal con media $\mu_1 + \mu_2 + \dots + \mu_n$ y desviación típica $\sqrt{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2}$.

$$Y = X_1 + X_2 + \dots + X_n \rightsquigarrow N\left(\mu_1 + \mu_2 + \dots + \mu_n, \sqrt{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2}\right)$$

10.2. Aproximación de la binomial a la normal

Utilizando el teorema central de límite anterior, podremos considerar una variable X que siga una distribución binomial $B(n, p)$, como una suma de n variables o ensayos independientes X_i con media p y varianza $p(1 - p)$, pues suponemos que X_i toma el valor uno si el ensayo es un éxito, y toma el valor cero si es un fracaso, es decir:

x_i	p_i	$x_i \cdot p_i$	$x_i^2 \cdot p_i$
1	p	p	p
0	$1 - p$	0	0
	1	p	p

Los parámetros (media y desviación típica) de esta variable son:

$$\mu = \sum_{i=1}^2 x_i p_i = 1 \cdot p + 0 \cdot (1 - p) = p$$

$$\sigma^2 = \sum_{i=1}^2 x_i^2 p_i - \mu^2 = (1^2 \cdot p + 0^2 \cdot (1 - p)) - p^2 = p - p^2 = p(1 - p)$$

Por tanto, como consecuencia del teorema central del límite, y si n es grande, podemos aproximar una variable aleatoria X que siga una distribución binomial $B(n, p)$ por una distribución normal cuya media y desviación típica son:

$$\mu = \mu_1 + \mu_2 + \dots + \mu_n = p + p + \dots \text{ (n veces) } \dots + p = np$$

$$\sigma = \sqrt{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2} = \sqrt{p(1 - p) + p(1 - p) + \dots \text{ (n veces) } \dots + p(1 - p)} = \sqrt{np(1 - p)}$$

Resumiendo:

$$X \rightsquigarrow B(n, p) \text{ y } n \text{ “grande”} \Rightarrow X \rightsquigarrow N(np, \sqrt{np(1 - p)})$$

Abraham de Moivre demostró que esta aproximación es buena si n es mayor que 30 y p no está próximo a cero ni a uno. En general se considera que la aproximación es buena si:

$$np \geq 5 \text{ y } n(1-p) \geq 5$$

Ejemplo

- ✓ Según un estudio realizado en la Comunidad de Murcia, el 75 % del voluntariado de la región son mujeres. Con motivo de preparar la fiesta del Día Internacional del Voluntario se seleccionan 1000 al azar de entre los 11000 voluntarios existentes en la actualidad. ¿Qué probabilidad hay de que al menos 260 sean hombres?

Si llamamos X a la variable aleatoria que expresa el número de hombres voluntarios, entonces X sigue una distribución $B(1000, 0,25)$. Por tanto, lo que tendremos que hallar es:

$$P(X \geq 260) = P(X = 260) + P(X = 261) + P(X = 262) + \dots + P(X = 1000)$$

Como en este caso se dan las condiciones óptimas para aproximar esta distribución binomial a una normal, es decir, $np = 1000 \cdot 0,25 = 250 > 5$ y $n(1-p) = 1000 \cdot 0,75 = 750 > 5$, podemos simplificar este cálculo aproximando X a una distribución normal $N(250, 13,7)$. De esta forma podremos calcular la probabilidad buscada de forma más sencilla:

$$P(X \geq 260) = P\left(Z \geq \frac{260 - 250}{13,7}\right) = P(Z \geq 0,73) = 1 - P(Z \leq 0,73) = 1 - 0,7673 = 0,2327$$

10.2.1. Corrección por continuidad

La distribución binomial es de variable discreta y, por tanto, tiene sentido calcular probabilidades puntuales. Cuando se aproxima una binomial por una normal, al ser esta última de variable continua no tiene sentido calcular probabilidades puntuales, pues son todas nulas, tal y como ya se ha visto en un apartado anterior.

La aproximación de una variable de distribución binomial X por una normal Y , genera un error que se corrige modificando el intervalo cuya probabilidad se quiere calcular. Así pues:

- $P(X = a) = P(a - 0,5 \leq Y \leq a + 0,5)$.
- $P(X \leq a) = P(Y \leq a + 0,5)$ (para que contenga al punto a).
- $P(X < a) = P(Y \leq a - 0,5)$ (para que no contenga al punto a).
- $P(X > a) = P(Y \leq a + 0,5)$ (para que no contenga al punto a).
- $P(X \geq a) = P(Y \geq a - 0,5)$ (para que contenga al punto a).
- $P(a \leq X \leq b) = P(a - 0,5 \leq Y \leq b + 0,5)$.

Ejemplo

- ✓ Se sabe que el 10 % de los habitantes de una ciudad va regularmente al teatro. Se toma una muestra al azar de 100 habitantes de esta ciudad. ¿Cuál es la probabilidad aproximada de que al menos el 13 % de ellos vayan al teatro?

La distribución $X = \text{"número de personas que van regularmente al teatro"}$ sigue una binomial $B(100, 0,1)$. Como $np = 100 \cdot 0,1 = 10 > 5$ y $nq = 100 \cdot 0,9 = 90 > 5$, se puede aproximar por una normal $Y = N(np, \sqrt{np(1-p)}) = N(10, 3)$. Por tanto:

$$P(X \geq 13) = P(Y \geq 12,5) = P\left(Z \geq \frac{12,5 - 10}{3}\right) = P(Z \leq 0,83) = 0,2033$$

11. Ejercicios

1. Una gran empresa debe reponer las batas de sus 1000 operarios. Se sabe que la talla media es de 170 cm, con una desviación típica de 3 cm. Las batas se confeccionan en tres tallas válidas para estaturas entre 155 y 165 cm, 165 y 175 cm y, finalmente, entre 175 y 185 cm. ¿Cuántas batas de cada talla ha de adquirir?

Solución: Deberá adquirir 48 batas de talla pequeña, 905 de talla media y 48 de talla grande.

2. Las puntuaciones de un examen varían entre cero y diez. Si las notas están normalmente distribuidas con media 6,7 y desviación típica 1,2; encuentra el porcentaje de estudiantes que tuvo una nota comprendida entre 5,5 y 6,5.

Solución: El 27,38 % de estudiantes tuvo una nota comprendida entre 5,5 y 6,5.

3. Se sabe, tras varios sondeos, que en una determinada población únicamente el 15 % es favorable a los tratamientos de psicoterapia. Elegida al azar una muestra de 50 personas, se desea saber:

- a) La probabilidad de que haya, exactamente, una persona favorable a dichos tratamientos.
- b) La probabilidad de que haya más de 5 personas favorables a los tratamientos.

Solución: a) 0,0026; b) 0,8413.

4. Un tirador de dardos acierta ocho de cada diez lanzamientos. Utilizando la aproximación de la binomial a la normal, encuentra la probabilidad de que de 50 lanzamientos acierte 45.

Solución: 0,0287.

5. De una baraja española (40 cartas, 4 palos: 10 de oros, 10 de copas, 10 de espadas y 10 de bastos) se sacan, con devolución, 50. Utilizando la aproximación de la binomial a la normal, encuentra las probabilidades de que:

- a) Más de seis sean de oros.
- b) Ocho sean de oros.

Solución: a) 0,9821; b) 0,0448.

6. Se sabe que la vida media de un electrodoméstico es de 10 años con una desviación típica de 0,7 años. Suponiendo que la vida media sigue una distribución normal, calcula:

- a) La probabilidad de que el electrodoméstico dure más de nueve años.
- b) La probabilidad de que dure entre 9 y 11 años.

Solución: a) 0,9236; b) 0,8472.

7. La probabilidad de que un cierto tipo de piezas de una máquina sea defectuoso es del 6 %. En un almacén se han recibido 2000 piezas. ¿Cuántas habrá defectuosas por término medio? ¿Cuál será la desviación típica?

Solución: El número esperado de piezas defectuosas es de 120. La desviación típica es, aproximadamente, 10,6.

8. En una ferretería venden cajas de clavos; el número de clavos de cada caja sigue una distribución normal $N(200, 10)$.

- a) ¿Qué porcentaje de cajas contiene entre 180 y 220 clavos?
- b) Si se devuelve el importe de las cajas que contienen menos de 180 clavos y compramos dos cajas, ¿cuál es la probabilidad de que tengan que devolvernos el importe de las dos cajas?

Solución: a) 95,44 %; b) 0,00052.

9. Para determinar si aceptamos o no un lote de cien tornillos, seleccionamos diez al azar y observamos si son o no defectuosos, de forma que si en esta selección encontramos más de un tornillo defectuoso, rechazamos el lote. Si la caja contiene el cinco por ciento de tornillos defectuosos:

- a) ¿Qué probabilidad hay de que rechacemos la caja?
- b) ¿Cuál es la probabilidad de que la caja sea aceptada?

Solución: a) 0,0862; b) 0,9138.

10. Se elige al azar una familia de seis hijos y se observa el número de hijos varones. Calcula la probabilidad de que la familia tenga:

- a) Dos hijos varones.

- b) Los mismos hijos que hijas.
- c) Alguna hija.

Solución: a) 0,2344; b) 0,3125; c) 0,9844.

11. En un club universitario deciden admitir al 20 % de aspirantes con las mejores puntuaciones tras realizar una serie de pruebas. Si la nota de estas pruebas sigue una distribución normal $N(100, 15)$, ¿a partir de qué puntuación se admite a los aspirantes?

Solución: a partir de 112,6.

12. Supongamos que la probabilidad de que una persona sufra un resfriado es 0,3. ¿Qué probabilidad hay de que en un grupo de 200 personas haya 50 resfriadas?

Solución: 0,0195.

13. Supongamos que el índice de audiencia de un determinado programa es del 53 %. Si elegimos al azar a 1.000 telespectadores, ¿cuál es la probabilidad de que al menos 550 estén viendo este programa?

Solución: 0,102.

14. Una ONG extranjera está vendiendo lotería para obtener dinero que emplear en sus objetivos. Se venden mil papeletas a cien u.m. (unidades monetarias) cada una y el premio es una cantidad en metálico de 25000 u.m. Un profesor de Matemáticas decide comprar dos papeletas, ¿cuál es la ganancia esperada?

Solución: aproximadamente -150 u.m.

15. La probabilidad de nacimiento de una niña es 0,48. Si sabemos que en una clínica se han producido 9 nacimientos, ¿cuál es la probabilidad de que todos hayan sido del mismo sexo?

Solución: 0,00413.

16. Un examen consta de 10 preguntas de tipo test. Para cada pregunta se ofrecen cuatro posibles respuestas, de las que solamente una es correcta. Si un estudiante responde todas las preguntas al azar:

- a) ¿Cuál es la probabilidad de que apruebe el examen, es decir, de contestar correctamente a cinco o más preguntas?
- b) ¿Cuál es la probabilidad de que falle todas las preguntas?
- c) ¿Cuál es la probabilidad de obtener una calificación de notable, es decir, siete u ocho respuestas correctas?
- c) ¿Cuál será el número medio de respuestas correctas?

Solución: a) 0,0781; b) 0,0563; c) 0,003476; d) 2,5.

17. En un instituto de bachillerato de 680 alumnos la nota media de fin de curso es de 5,2 con una desviación típica de 1,6. Si se supone que la distribución de notas es normal:

- a) ¿Cuántos alumnos tienen la nota media superior a cinco?
- b) ¿Cuál es la probabilidad de que un alumno tenga una nota superior a siete?
- c) Se quiere preseleccionar al 2,5% de los alumnos que han conseguido las notas más altas para hacer un equipo y participar en un concurso cultural. ¿Cuál es la nota mínima (con dos cifras decimales) necesaria para ser preseleccionado?

Solución: a) Aproximadamente 375; b) 0,1292; c) 8,34.

18. La probabilidad de que una persona vacunada contra la gripe contraiga esta enfermedad es de 0,1. Si elegimos 100 personas al azar entre las vacunadas, calcula la probabilidad de que estén enfermas de gripe:

- a) Más de la mitad.
- b) Siete o más y menos de diez.
- c) Alguna persona.

Solución: a) 0; b) 0,3413; c) 0,9987.

19. En la última temporada de entrenamientos, los tiempos empleados por un corredor de 100 metros en recorrer dicha distancia se distribuyeron según una normal de media 10,8 segundos y desviación típica 0,3 segundos:

- a) Para participar en un campeonato ha de recorrer una vez los 100 metros en menos de 10,5 segundos. ¿Qué probabilidad tiene de participar?
- b) Para poder estar en una olimpiada tiene que participar en 25 carreras y obtener una marca de menos de 10,4 segundos en dos de ellas como mínimo. ¿Podrá ir a la olimpiada?
- c) Suponiendo que el atleta está en la olimpiada, ¿qué probabilidad tiene de correr la final si para ello tiene que correr tres eliminatorios en un tiempo inferior a 10,3 segundos?

Solución: a) 0,1587; b) Sí; c) 0,1425.

20. La balanza de una frutería comete errores en cada pesada que se distribuyen según una normal de media 0 gramos y desviación típica 5 gramos:

- a) ¿Cuál es la probabilidad de que en una pesada la balanza marque un peso superior en 10 gramos al verdadero?
- b) ¿Qué porcentaje de veces la balanza cometerá un error de más de 8 gramos a favor del cliente?
- c) ¿Cuál es la probabilidad de que el error sea de menos de 8 gramos?

Solución: a) 0,0228; b) 0,0548; c) 0,8904.

21. Se han sorteado 60 entradas para un concierto entre los alumnos de un instituto. Sabiendo que en dicho instituto el 16 % de los alumnos está matriculado en 2º de bachillerato, se pide:

- a) Probabilidad de que les corresponda una entrada a más de 20 alumnos de 2º de bachillerato.
- b) Probabilidad de que el número de alumnos de 2º de bachillerato con entrada esté comprendido entre 5 y 10, ambos inclusive.

Solución: a) 0,0001; b) 0,5031.

22. En una fábrica se sabe que el 2 % de los tornillos fabricados tienen una longitud inadecuada y que el 4 % tienen el diámetro incorrecto. Si ambos defectos son independientes y los tornillos se empaquetan en cajas de 200, utiliza la distribución binomial y la aproximación normal de la binomial, comparando los resultados, para calcular la probabilidad de que en una caja haya entre 1 y 3 (ambos inclusive) tornillos:

- a) Con la longitud inadecuada.
- b) Con el diámetro incorrecto.
- c) Con la longitud y el diámetro defectuosos.

Solución: Con la binomial: a) 0,4139; b) 0,03924; c) 0,01624. Con aproximación de la binomial a la normal: a) 0,2407; b) 0,0292; c) 0,007028.

23. Una variable aleatoria continua X sigue una distribución normal $N(20, 4)$. Halla el valor de t para que $P(X \leq t) = 0,1526$.

Solución: $t = 15,898$.

24. Un panadero fabrica barras de pan cuyos pesos se distribuyen normalmente según una normal de media 250 gramos. El 5 % de las barras pesa menos de 240 gramos. ¿Cuánto vale su desviación típica?

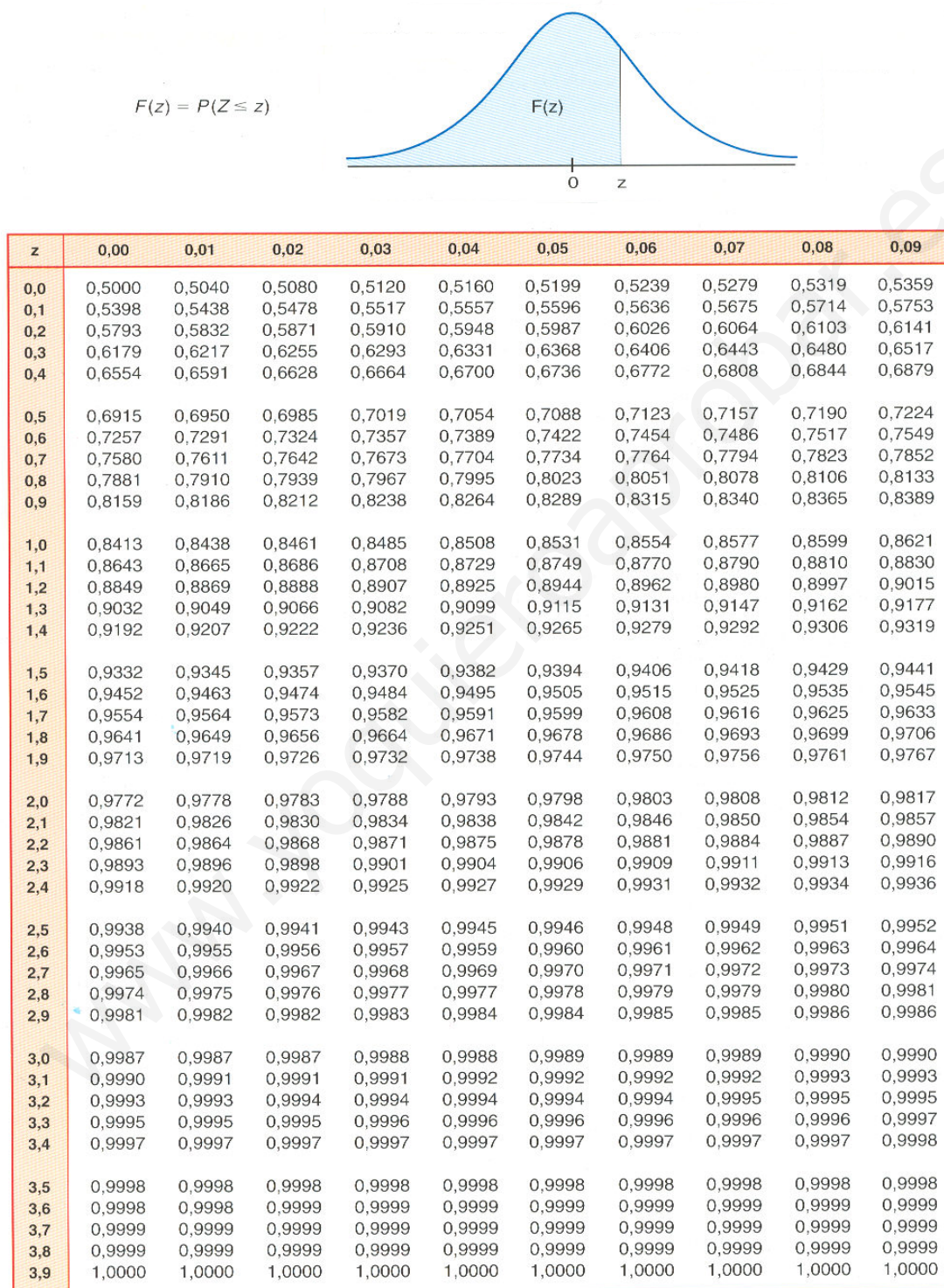
Solución: $\sigma = 6,08$.

12. Tabla de la distribución binomial

n	k	0'01	0'05	0'10	0'15	0'20	0'25	0'30	1/3	0'35	0'40	0'45	0'50
2	0	0'9801	0'9025	0'8100	0'7225	0'6400	0'5625	0'4900	0'4444	0'4225	0'3600	0'3025	0'2500
	1	0'9999	0'9975	0'9900	0'9775	0'9600	0'9375	0'9100	0'8889	0'8775	0'8400	0'7975	0'7500
	2	1'0000	1'0000	1'0000	1'0000	1'0000	1'0000	1'0000	1'0000	1'0000	1'0000	1'0000	1'0000
3	0	0'9703	0'8574	0'7290	0'6141	0'5120	0'4219	0'3430	0'2963	0'2746	0'2160	0'1664	0'1250
	1	0'9997	0'9928	0'9720	0'9392	0'8960	0'8438	0'7840	0'7407	0'7182	0'6480	0'5748	0'5000
	2	1'0000	0'9999	0'9990	0'9966	0'9920	0'9844	0'9730	0'9630	0'9571	0'9360	0'9089	0'8750
	3		1'0000	1'0000	1'0000	1'0000	1'0000	1'0000	1'0000	1'0000	1'0000	1'0000	1'0000
4	0	0'9606	0'8145	0'6561	0'5220	0'4096	0'3164	0'2401	0'1975	0'1785	0'1296	0'0915	0'0625
	1	0'9994	0'9860	0'9477	0'8905	0'8192	0'7383	0'6517	0'5926	0'5630	0'4752	0'3910	0'3125
	2	1'0000	0'9995	0'9963	0'9880	0'9728	0'9492	0'9163	0'8889	0'8735	0'8208	0'7585	0'6875
	3		1'0000	0'9999	0'9995	0'9984	0'9961	0'9919	0'9876	0'9850	0'9744	0'9590	0'9375
	4			1'0000	1'0000	1'0000	1'0000	1'0000	1'0000	1'0000	1'0000	1'0000	1'0000
5	0	0'9510	0'7738	0'5905	0'4437	0'3277	0'2373	0'1681	0'1317	0'1160	0'0778	0'0503	0'0312
	1	0'9990	0'9774	0'9185	0'8352	0'7373	0'6328	0'5282	0'4609	0'4284	0'3370	0'2562	0'1875
	2	1'0000	0'9988	0'9914	0'9734	0'9421	0'8965	0'8369	0'7901	0'7648	0'6826	0'5931	0'5000
	3		1'0000	0'9995	0'9978	0'9933	0'9844	0'9692	0'9547	0'9460	0'9130	0'8688	0'8125
	4			1'0000	0'9999	0'9997	0'9990	0'9976	0'9959	0'9947	0'9898	0'9815	0'9688
	5				1'0000	1'0000	1'0000	1'0000	1'0000	1'0000	1'0000	1'0000	1'0000
6	0	0'9415	0'7351	0'5314	0'3771	0'2621	0'1780	0'1176	0'0878	0'0754	0'0467	0'0277	0'0156
	1	0'9986	0'9672	0'8857	0'7765	0'6553	0'5339	0'4202	0'3512	0'3191	0'2333	0'1636	0'1094
	2	1'0000	0'9978	0'9842	0'9527	0'9011	0'8306	0'7443	0'6804	0'6471	0'5443	0'4415	0'3438
	3		0'9999	0'9987	0'9941	0'9830	0'9624	0'9295	0'8999	0'8826	0'8208	0'7447	0'6562
	4		1'0000	0'9999	0'9996	0'9984	0'9954	0'9891	0'9822	0'9777	0'9590	0'9308	0'8906
	5			1'0000	1'0000	0'9999	0'9998	0'9993	0'9986	0'9982	0'9959	0'9917	0'9844
	6					1'0000	1'0000	1'0000	1'0000	1'0000	1'0000	1'0000	1'0000
7	0	0'9321	0'6983	0'4783	0'3211	0'2097	0'1335	0'0824	0'0585	0'0490	0'0280	0'0152	0'0078
	1	0'9980	0'9556	0'8503	0'7166	0'5767	0'4449	0'3294	0'2634	0'2338	0'1586	0'1024	0'0625
	2	1'0000	0'9962	0'9743	0'9262	0'8520	0'7564	0'6471	0'5706	0'5323	0'4199	0'3164	0'2266
	3		0'9998	0'9973	0'9879	0'9667	0'9294	0'8740	0'8267	0'8011	0'7102	0'6083	0'5000
	4		1'0000	0'9998	0'9988	0'9953	0'9871	0'9712	0'9547	0'9444	0'9037	0'8471	0'7734
	5			1'0000	0'9999	0'9996	0'9987	0'9962	0'9931	0'9910	0'9812	0'9643	0'9375
	6				1'0000	0'9999	0'9998	0'9995	0'9995	0'9994	0'9984	0'9963	0'9922
	7					1'0000	1'0000	1'0000	1'0000	1'0000	1'0000	1'0000	1'0000
8	0	0'9227	0'6634	0'4305	0'2725	0'1678	0'1001	0'0576	0'0390	0'0319	0'0168	0'0084	0'0039
	1	0'9973	0'9428	0'8131	0'6572	0'5033	0'3671	0'2553	0'1951	0'1691	0'1064	0'0632	0'0352
	2	0'9999	0'9942	0'9619	0'8948	0'7969	0'6785	0'5518	0'4682	0'4278	0'3154	0'2201	0'1445
	3	1'0000	0'9996	0'9950	0'9786	0'9437	0'8862	0'8059	0'7413	0'7064	0'5941	0'4770	0'3633
	4		1'0000	0'9996	0'9971	0'9896	0'9727	0'9420	0'9121	0'8939	0'8263	0'7396	0'6367
	5			1'0000	0'9998	0'9988	0'9958	0'9887	0'9803	0'9747	0'9502	0'9115	0'8555
	6				1'0000	0'9999	0'9996	0'9987	0'9974	0'9964	0'9915	0'9819	0'9648
	7					1'0000	1'0000	0'9999	0'9998	0'9998	0'9993	0'9983	0'9961
	8						1'0000	1'0000	1'0000	1'0000	1'0000	1'0000	1'0000
9	0	0'9135	0'6302	0'3874	0'2316	0'1342	0'0751	0'0404	0'0260	0'0207	0'0101	0'0046	0'0020
	1	0'9965	0'9288	0'7748	0'5995	0'4362	0'3003	0'1960	0'1431	0'1211	0'0705	0'0385	0'0195
	2	0'9999	0'9916	0'9470	0'8591	0'7382	0'6007	0'4628	0'3772	0'3373	0'2318	0'1495	0'0898
	3	1'0000	0'9994	0'9917	0'9661	0'9144	0'8343	0'7297	0'6503	0'6089	0'4826	0'3614	0'2539
	4		1'0000	0'9991	0'9944	0'9804	0'9511	0'9012	0'8552	0'8283	0'7334	0'6214	0'5000
	5			0'9999	0'9994	0'9969	0'9900	0'9747	0'9576	0'9464	0'9006	0'8342	0'7461
	6			1'0000	1'0000	0'9997	0'9987	0'9957	0'9917	0'9888	0'9750	0'9502	0'9102
	7					1'0000	0'9999	0'9996	0'9990	0'9986	0'9962	0'9909	0'9805
	8						1'0000	1'0000	0'9999	0'9999	0'9999	0'9992	0'9980
	9							1'0000	1'0000	1'0000	1'0000	1'0000	1'0000
10	0	0'9044	0'5987	0'3487	0'1969	0'1074	0'0563	0'0282	0'0173	0'0135	0'0060	0'0025	0'0010
	1	0'9958	0'9139	0'7361	0'5443	0'3758	0'2440	0'1493	0'1040	0'0860	0'0464	0'0233	0'0107
	2	1'0000	0'9885	0'9298	0'8202	0'6778	0'5256	0'3828	0'2991	0'2616	0'1673	0'0996	0'0547
	3		0'9990	0'9872	0'9500	0'8791	0'7759	0'6496	0'5593	0'5138	0'3823	0'2660	0'1719
	4		0'9999	0'9984	0'9901	0'9672	0'9219	0'8497	0'7869	0'7515	0'6331	0'5044	0'3770
	5		1'0000	0'9999	0'9986	0'9936	0'9803	0'9527	0'9234	0'9051	0'8338	0'7384	0'6230
	6			1'0000	0'9999	0'9991	0'9965	0'9894	0'9803	0'9740	0'9452	0'8980	0'8281
	7				1'0000	0'9999	0'9996	0'9984	0'9966	0'9952	0'9877	0'9726	0'9453
	8					1'0000	1'0000	0'9999	0'9996	0'9995	0'9983	0'9955	0'9893
	9						1'0000	1'0000	1'0000	1'0000	0'9999	0'9997	0'9990
	10										1'0000	1'0000	1'0000

Figura 6: Tabla de la distribución binomial $B(n, p)$

13. Tabla de la distribución normal estándar

Figura 7: Tabla de la distribución normal estándar $N(0, 1)$