

Técnicas Cuantitativas II

INTRODUCCIÓN

Algunos modelos continuos de variable aleatoria
Introducción a la inferencia Estadística

Índice

Algunos modelos
continuos de variable
aleatoria

Concepto de muestra y
estadístico

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Apéndice

Algunos modelos continuos de variable aleatoria

Concepto de muestra y estadístico

Algunos ejemplos de estadísticos

Valor esperado y varianza de la media muestral

Valor esperado de la varianza y cuasivarianza muestral

Apéndice

Algunos modelos continuos de variable aleatoria: uniforme, exponencial, gamma, beta, normal y asociadas a la normal: χ^2 , t-Student, F-Snedecor

Definición 1 Dada X una variable aleatoria continua que toma valores en el intervalo real (a, b) , se dice que se distribuye de forma uniforme, lo que se notará por $X \sim U(a, b)$, si y sólo si, su función de densidad responde a la expresión

$$f(x) = \begin{cases} \frac{1}{b-a} & , \text{ si } x \in (a, b) \\ 0 & , \text{ en otro caso} \end{cases}.$$

Observa que cualquier punto del intervalo son igualmente probables!
Su función de distribución será

$$F(x) = \begin{cases} 0 & , \text{ si } x < a \\ \frac{x-a}{b-a} & , \text{ si } a \leq x < b \\ 1 & , \text{ si } x \geq b \end{cases}.$$

Además, las principales características de esta distribución son:

- Esperanza matemática: $E[X] = \frac{a+b}{2}$.
- Varianza: $Var(X) = \frac{(b-a)^2}{12}$.

Un ejercicio interesante sería calcular cada una de ellas.

Para $a = 0$ y $b = 1$, $U(0, 1)$ es particularmente sencilla y se conoce como *distribución uniforme standard*.

Definición 2 Dada X una variable aleatoria continua que toma valores en $\mathbb{R}^+$, se dice que X se distribuye según una distribución exponencial de parámetro $\theta > 0$, lo que se notará por $X \sim Exp(\theta)$, si y sólo si su función de densidad responde a la siguiente expresión

$$f(x) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}} & , \text{ si } x > 0 \\ 0 & , \text{ en otro caso } \end{cases} .$$

Su función de distribución será:

$$F(x) = \begin{cases} 0 & , \text{ si } x < 0 \\ 1 - e^{-\frac{x}{\theta}} & , \text{ si } x \geq 0 \end{cases} .$$

Además, las principales características de la distribución exponencial son:

- Esperanza matemática: $E[X] = \theta$.
- Varianza: $Var(X) = \theta^2$.

Adviértase que para obtener los momentos de primer y segundo orden es necesario integrar por partes.

Índice

Algunos modelos
continuos de variable
aleatoria

Uniforme

Exponencial

Gamma

Beta

Normal

Chi-Cuadrado

t-Student

F-Snedecor

Concepto de muestra y
estadístico

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Apéndice

TC II

La distribución exponencial se puede expresar también (usando el cambio $\theta = \frac{1}{\lambda}$) como sigue:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & , \text{ si } x > 0 \\ 0 & , \text{ en otro caso} \end{cases} .$$

En tal caso, su función de distribución viene dada por

$$F(x) = \begin{cases} 0 & , \text{ si } x < 0 \\ 1 - e^{-\lambda x} & , \text{ si } x \geq 0 \end{cases} .$$

Además, $E[X] = \frac{1}{\lambda}$ y $Var(X) = \frac{1}{\lambda^2}$.

Definición 3 Dado $\alpha > 0$, se define la función Gamma de α como

$$\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} e^{-x} dx.$$

Además se verifican las siguientes propiedades:

- $\int_0^{+\infty} x^{\alpha-1} e^{-\frac{x}{\theta}} dx = \Gamma(\alpha) \cdot \theta^\alpha, \quad \alpha, \theta > 0.$
- $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1), \quad \alpha > 0.$

Como consecuencia de la segunda propiedad: $\Gamma(n) = (n - 1)!, \quad n \in \mathbb{N}.$

Definición 4 Dada X una variable aleatoria continua que toma valores en $\mathbb{R}^+$, se dice que se distribuye según una gamma de parámetros $\alpha, \theta > 0$, si y sólo si, su función de densidad responde a la siguiente expresión

$$f(x) = \begin{cases} \frac{1}{\Gamma(\alpha)\theta^\alpha} x^{\alpha-1} e^{-\frac{x}{\theta}} & , \text{ si } x > 0 \\ 0 & , \text{ en otro caso} \end{cases}.$$

Se nota $X \sim G(\alpha, \theta).$

Adviértase que para $\alpha = 1$, la distribución $G(\alpha, \theta)$ pasa a ser una $Exp(\theta).$

Índice

Algunos modelos
continuos de variable
aleatoria

Uniforme

Exponencial

Gamma

Beta

Normal

Chi-Cuadrado

t-Student

F-Snedecor

Concepto de muestra y
estadísticoAlgunos ejemplos de
estadísticosValor esperado y
varianza de la media
muestralValor esperado de la
varianza y cuasivarianza
muestral

Apéndice

El momento de primer orden de la distribución gamma es $E[X] = \alpha \cdot \theta$.
En efecto,

$$\begin{aligned}
 E[X] &= \int_{-\infty}^{+\infty} x f(x) dx \\
 &= \int_0^{+\infty} x \cdot \frac{1}{\Gamma(\alpha)\theta^\alpha} \cdot x^{\alpha-1} e^{-\frac{x}{\theta}} dx \\
 &= \frac{1}{\Gamma(\alpha)\theta^\alpha} \int_0^{+\infty} x^\alpha e^{-\frac{x}{\theta}} dx \\
 &= \frac{1}{\Gamma(\alpha)\theta^\alpha} \cdot \Gamma(\alpha + 1) \cdot \theta^{\alpha+1} \\
 &= \frac{\theta^{\alpha+1}}{\theta^\alpha} \cdot \frac{\alpha\Gamma(\alpha)}{\Gamma(\alpha)} \\
 &= \alpha \cdot \theta,
 \end{aligned}$$

donde se han usado las propiedades relacionadas con la función gamma.

Índice

Algunos modelos
continuos de variable
aleatoria

Uniforme

Exponencial

Gamma

Beta

Normal

Chi-Cuadrado

t-Student

F-Snedecor

Concepto de muestra y
estadístico

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Apéndice

El momento de segundo orden de la distribución gamma es $Var(X) = \alpha\theta^2$.
En efecto,

$$\begin{aligned}
 E[X^2] &= \int_{-\infty}^{+\infty} x^2 f(x) dx \\
 &= \int_0^{+\infty} x^2 \cdot \frac{1}{\Gamma(\alpha)\theta^\alpha} \cdot x^{\alpha-1} e^{-\frac{x}{\theta}} dx \\
 &= \frac{1}{\Gamma(\alpha)\theta^\alpha} \int_0^{+\infty} x^{\alpha+1} e^{-\frac{x}{\theta}} dx \\
 &= \frac{1}{\Gamma(\alpha)\theta^\alpha} \cdot \Gamma(\alpha+2) \cdot \theta^{\alpha+2} \\
 &= \frac{\theta^{\alpha+2}}{\theta^\alpha} \cdot \frac{\alpha(\alpha+1)\Gamma(\alpha)}{\Gamma(\alpha)} \\
 &= \alpha \cdot (\alpha+1) \cdot \theta^2,
 \end{aligned}$$

donde se han usado las mismas propiedades que antes.

Por tanto, como $Var(X) = E[X^2] - (E[X])^2$, entonces

$$Var(X) = \alpha \cdot (\alpha+1) \cdot \theta^2 - \alpha^2 \theta^2 = \alpha^2 \theta^2 + \alpha \theta^2 - \alpha^2 \theta^2 = \alpha \theta^2.$$

Definición 5 Dados $a, b > 0$, se define la función beta como

$$\beta(a, b) = \int_0^1 x^{a-1} \cdot (1-x)^{b-1} dx.$$

Dicha función verifica que $\beta(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$.

Definición 6 Dada X variable aleatoria continua que toma valores en el intervalo $(0, 1)$, se dice que se distribuye según una beta de parámetros a y b , si y sólo si, su función de densidad responde a la siguiente expresión

$$\begin{aligned} f(x) &= \begin{cases} \frac{1}{\beta(a, b)} \cdot x^{a-1} \cdot (1-x)^{b-1} & , \text{ si } 0 < x < 1 \\ 0 & , \text{ en otro caso} \end{cases} \\ &= \begin{cases} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot x^{a-1} \cdot (1-x)^{b-1} & , \text{ si } 0 < x < 1 \\ 0 & , \text{ en otro caso} \end{cases}. \end{aligned}$$

Además, se nota $X \sim \beta(a, b)$. Observa que $\beta(1, 1) = U(0, 1)$.

Índice

Algunos modelos
continuos de variable
aleatoria

Uniforme

Exponencial

Gamma

Beta

Normal

Chi-Cuadrado

t-Student

F-Snedecor

Concepto de muestra y
estadísticoAlgunos ejemplos de
estadísticosValor esperado y
varianza de la media
muestralValor esperado de la
varianza y cuasivarianza
muestral

Apéndice

El momento de primer orden de la distribución beta es $E[X] = \frac{a}{a+b}$.
En efecto,

$$\begin{aligned}
 E[X] &= \int_{-\infty}^{+\infty} x f(x) dx \\
 &= \int_0^1 x \cdot \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot x^{a-1} \cdot (1-x)^{b-1} dx \\
 &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 x^a \cdot (1-x)^{b-1} dx \\
 &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \frac{\Gamma(a+1)\Gamma(b)}{\Gamma(a+1+b)} \\
 &= \frac{\Gamma(a+b)a\Gamma(a)}{\Gamma(a)(a+b)\Gamma(a+b)} \\
 &= \frac{a}{a+b},
 \end{aligned}$$

donde se han usado las propiedades relacionadas con las funciones beta y gamma.

El momento de segundo orden de la distribución beta es

$$Var(X) = \frac{ab}{(a+b)^2(a+b+1)}.$$

En efecto,

$$\begin{aligned} E[X^2] &= \int_{-\infty}^{+\infty} x^2 f(x) dx = \int_0^1 x^2 \cdot \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot x^{a-1} \cdot (1-x)^{b-1} dx \\ &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 x^{a+1} \cdot (1-x)^{b-1} dx = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \frac{\Gamma(a+2)\Gamma(b)}{\Gamma(a+2+b)} \\ &= \frac{\Gamma(a+b)(a+1)a\Gamma(a)}{\Gamma(a)(a+b+1)(a+b)\Gamma(a+b)} = \frac{a(a+1)}{(a+b)(a+b+1)}, \end{aligned}$$

donde también se han usado las propiedades relacionadas con las funciones beta y gamma.

Con ello, puesto que $Var(X) = E[X^2] - (E[X])^2$ se tiene que

$$Var(X) = \frac{a(a+1)}{(a+b)(a+b+1)} - \frac{a^2}{(a+b)^2} = \frac{ab}{(a+b)^2(a+b+1)}.$$

Índice

Algunos modelos
continuos de variable
aleatoria

Uniforme

Exponencial

Gamma

Beta

Normal

Chi-Cuadrado

t-Student

F-Snedecor

Concepto de muestra y
estadístico

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Apéndice

La distribución normal o de Gauss ajusta gran cantidad de variables no solo económicas o sociológicas, sino también biológicas, agrícolas, etc. Esta gran versatilidad de la distribución normal tiene un gran apoyo en el Teorema Central del Limite, de manera que siempre que un determinado efecto sea producido por un gran número de causas independientes, la distribución de éste será aproximadamente normal. Esta pontecialidad hace que sea la distribución más importante y más utilizada.

Definición 7 Se dice que una variable aleatoria continua, X , se distribuye según una Normal de parámetros μ y σ^2 , lo cual denotaremos por $X \sim N(\mu, \sigma^2)$, si, y sólo si, su función de densidad es

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}, \sigma > 0, \mu \in \mathbb{R}.$$

Los momentos de primer y segundo orden de la distribución Normal son $E[X] = \mu$ y $Var(X) = \sigma^2$, es decir, el primer parámetro desconocido, μ , coincide con la media poblacional, mientras que el segundo, σ^2 , coincide con la varianza poblacional. Evidentemente, la raíz cuadrada positiva de la varianza, σ , será la desviación típica.

Para calcular la media se parte de la definición correspondiente:

$$\begin{aligned}
 E[X] &= \int_{-\infty}^{+\infty} x f(x) dx = \int_{-\infty}^{+\infty} x \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
 &= \left[\begin{array}{l} \frac{x-\mu}{\sigma} = y \iff x = \mu + \sigma y \\ dx = \sigma dy \end{array} \right] \\
 &= \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{+\infty} (\mu + \sigma y) e^{-\frac{y^2}{2}} \sigma dy \\
 &= \frac{1}{\sqrt{2\pi}} \left[\mu \int_{-\infty}^{+\infty} e^{-\frac{y^2}{2}} dy + \sigma \int_{-\infty}^{+\infty} y e^{-\frac{y^2}{2}} dy \right] \\
 &= \frac{1}{\sqrt{2\pi}} \left[\mu \sqrt{2\pi} - \sigma \cdot 0 \right] = \mu,
 \end{aligned}$$

donde se ha tenido en cuenta

$$\int_{-\infty}^{+\infty} e^{-\frac{y^2}{2}} dy = \sqrt{2\pi},$$

y que $\int_a^b e^{g(x)} g'(x) dx = \left[e^{g(x)} \right]_a^b$.

De modo análogo actuamos para calcular la varianza:

$$\begin{aligned}
 Var(X) &= E[(X - E[X])^2] = E[(X - \mu)^2] = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x) dx \\
 &= \int_{-\infty}^{+\infty} (x - \mu)^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
 &= \left[\begin{array}{l} \frac{x-\mu}{\sigma} = y \iff x = \mu + \sigma y \\ dx = \sigma dy \end{array} \right] \\
 &= \int_{-\infty}^{+\infty} \frac{\sigma^2 y^2}{\sigma\sqrt{2\pi}} e^{-\frac{y^2}{2}} \sigma dy = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} y^2 e^{-\frac{y^2}{2}} dy \\
 &= \left[\begin{array}{l} u = y \iff du = dy \\ dv = ye^{-\frac{y^2}{2}} dy \iff v = -e^{-\frac{y^2}{2}} \end{array} \right] \\
 &= \frac{\sigma^2}{\sqrt{2\pi}} \left[\left[-ye^{-\frac{y^2}{2}} \right]_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} -e^{-\frac{y^2}{2}} dy \right] \\
 &= \frac{\sigma^2}{\sqrt{2\pi}} (0 + \sqrt{2\pi}) = \sigma^2.
 \end{aligned}$$

Índice

Algunos modelos
continuos de variable
aleatoria

Uniforme

Exponencial

Gamma

Beta

Normal

Chi-Cuadrado

t-Student

F-Snedecor

Concepto de muestra y
estadístico

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Apéndice

También es importante destacar la propiedad de reproductividad de la distribución Normal: La combinación lineal de distribuciones normales sigue siendo una distribución normal, cuya media es la combinación lineal de las medias de cada distribución, y cuya varianza es la combinación lineal con los coeficientes al cuadrado de las varianzas de cada distribución.

Esto es, si $Y = \sum_{i=1}^n a_i \cdot X_i$ donde $X_i \sim N(\mu_i, \sigma_i^2)$ y $a_i \in \mathbb{R}$, entonces

$$Y \sim N \left(\sum_{i=1}^n a_i \cdot \mu_i, \sum_{i=1}^n a_i^2 \cdot \sigma_i^2 \right).$$

Un caso particular, que será usado en los siguientes capítulos, es aquel en el que se trabaja con sólo dos variables, esto es:

$$\left. \begin{array}{l} X \sim N(\mu_1, \sigma_1^2) \\ Y \sim N(\mu_2, \sigma_2^2) \end{array} \right\} \Rightarrow Z = aX + bY \sim N(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2).$$

De igual forma, una propiedad interesante (usada, por ejemplo para calcular probabilidades) es que la distribución normal es simétrica con respecto a su media.

Índice

Algunos modelos continuos de variable aleatoria

Uniforme

Exponencial

Gamma

Beta

Normal

Chi-Cuadrado

t-Student

F-Snedecor

Concepto de muestra y estadístico

Algunos ejemplos de estadísticos

Valor esperado y varianza de la media muestral

Valor esperado de la varianza y cuasivarianza muestral

Apéndice

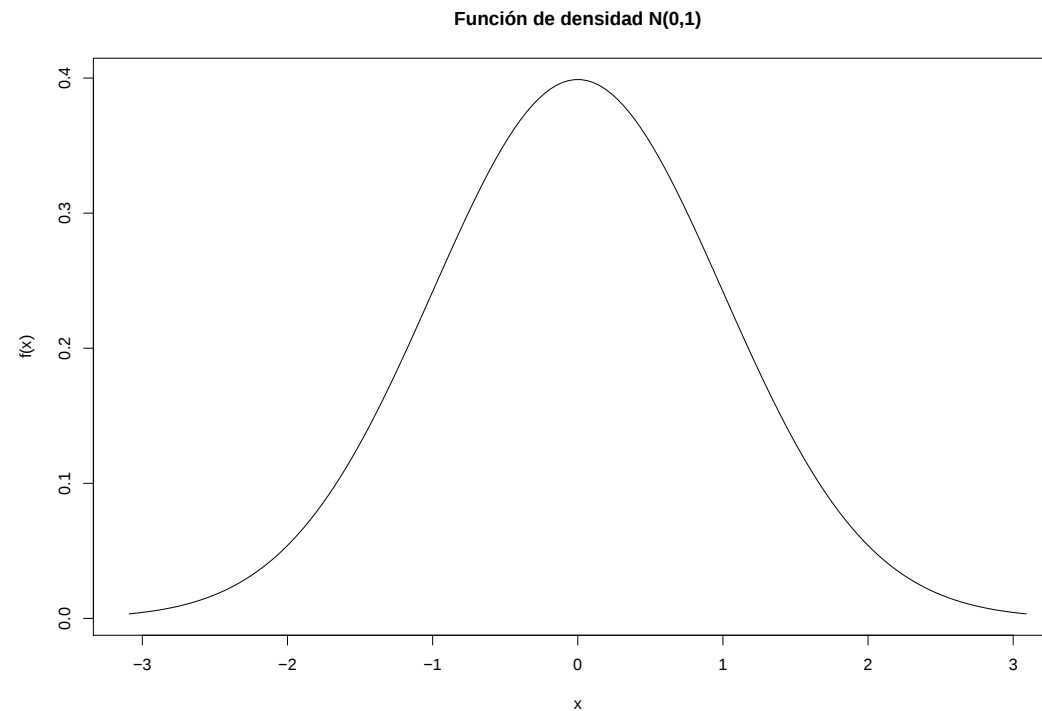


Figure 1: Representación de la función de densidad de una $N(0,1)$

Para la distribución Normal existen tablas con los valores de las distintas probabilidades, si bien, en este caso dichas tablas corresponden a un caso particular de la distribución Normal en el que $\mu = 0$ y $\sigma^2 = 1$.

Definición 8 Diremos que una variable aleatoria continua, Z , se distribuye como una normal $(0, 1)$, que denotaremos como $Z \sim N(0, 1)$, si se distribuye como una distribución normal de media 0 y varianza 1.

Las principales características de la distribución normal tipificada son:

- Función de densidad: $f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \quad \forall z \in \mathbb{R}.$
- Momentos de primer y segundo orden: $E[Z] = 0, Var(Z) = 1.$

Veamos a continuación un resultado que relaciona cualquier variable que se distribuya según una Normal con una Normal tipificada mediante un proceso conocido como tipificación:

Sea X una variable distribuida según una normal de parámetros μ y σ^2 . Entonces, la variable aleatoria tipificada $Z = \frac{X - \mu}{\sigma}$, se distribuye según una normal $(0, 1)$.

En efecto, dada una variable aleatoria continua, X , a partir de la que se obtiene otra, Y , mediante la siguiente relación lineal $Y = aX + b$, con $a > 0$. Sabemos entonces que se verifica:

$$f_Y(y) = \frac{1}{a} \cdot f_X\left(\frac{y-b}{a}\right),$$

conocida como transformación lineal de una variable aleatoria.

Puesto que en nuestro caso

$$Z = \frac{X - \mu}{\sigma} = \frac{1}{\sigma} \cdot X - \frac{\mu}{\sigma},$$

identificando $a = \frac{1}{\sigma} > 0$ y $b = -\frac{\mu}{\sigma}$, se verifica que

$$\begin{aligned} f_Z(z) &= \sigma \cdot f_X\left(\frac{z - (-\mu/\sigma)}{1/\sigma}\right) = \sigma \cdot f_X\left(\sigma\left(z + \frac{\mu}{\sigma}\right)\right) \\ &= \sigma \cdot f_X(\sigma z + \mu) = \sigma \cdot \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(\sigma z + \mu - \mu)^2}{2\sigma^2}} \\ &= \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{\sigma^2 z^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{z^2}{2}}, \end{aligned}$$

Índice

Algunos modelos
continuos de variable
aleatoria

Uniforme

Exponencial

Gamma

Beta

Normal

Chi-Cuadrado

t-Student

F-Snedecor

Concepto de muestra y
estadístico

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Apéndice

donde se ha usado la expresión de la densidad de una distribución Normal evaluada en el punto $\sigma z + \mu$.

Luego, como la función de densidad de la variable aleatoria Z corresponde a la de una normal tipificada, y ésta caracteriza a la variable aleatoria, entonces se verifica que $Z \sim N(0, 1)$.

A partir de este resultado, cualquier variable normal puede ser tipificada y tratada como una $N(0, 1)$, por lo que conocida la función de densidad de una distribución normal se puede calcular todo tipo de probabilidades asociadas a dicha variable, usando las tablas de probabilidades asociadas a una $N(0, 1)$.

Dichas tablas responden a la siguiente estructura:

- En la primera columna tenemos la parte entera y el primer decimal del punto x para el que se quiere calcular la probabilidad.
- En la primera fila tenemos el segundo decimal de dicho punto.
- Finalmente, en el interior de la tabla tenemos las probabilidades que determinan los puntos anteriores.

Ejemplo 1 Las calificaciones de los alumnos de estadística de primero de GADE se distribuyen según una variable aleatoria, X , que puede suponerse que se ajusta a una distribución aproximadamente normal, con una media de 6 puntos y una desviación típica de 3 puntos. Se pide hallar la probabilidad de aprobar.

En primer lugar, para poder usar las tablas de las probabilidades tenemos que tener una normal tipificada. Es decir, hay que transformar la $N(6, 9)$ que modeliza la variable considerada en una $N(0, 1)$. Con tal objetivo usaremos el teorema anterior:

$$X \sim N(6, 9) \implies Z = \frac{X - 6}{3} \sim N(0, 1).$$

En tal caso la probabilidad de aprobar viene dada por:

$$P[X > 5] = P\left[\frac{X - 6}{3} > \frac{5 - 6}{3}\right] = P\left[Z > -\frac{1}{3}\right] = P\left[Z < \frac{1}{3}\right] = 0'6293,$$

donde se ha usado que la distribución Normal es simétrica con respecto a su media, que en este caso es el origen. #

Índice

Algunos modelos continuos de variable aleatoria

Uniforme

Exponencial

Gamma

Beta

Normal

Chi-Cuadrado

t-Student

F-Snedecor

Concepto de muestra y estadístico

Algunos ejemplos de estadísticos

Valor esperado y varianza de la media muestral

Valor esperado de la varianza y cuasivarianza muestral

Apéndice

Adviértase que en el ejemplo anterior realmente se ha obtenido una aproximación, ya que la $P\left[Z < \frac{1}{3}\right]$ se encuentra entre $P[Z < 0'33]$ y $P[Z < 0'34]$, habiéndose elegido como probabilidad válida la más cercana. Si bien en estos casos habría que interpolar, una forma de actuar que se repetirá en el curso será la considerada en este caso, es decir, considerar como válida la probabilidad del punto más cercano.

A continuación estudiaremos tres distribuciones básicas, junto a la normal, en el desarrollo de la inferencia estadística y que surgen a partir de la misma: la χ^2 de Pearson, la t de Student y la F de Snedecor.

La distribución chi-cuadrado de Pearson con n grados de libertad, que denotaremos por χ_n^2 , se construye como la suma de n variables aleatorias independientes e idénticamente distribuidas según una normal, de media cero y varianza uno, al cuadrado.

Por tanto, la distribución χ^2 de Pearson responde a la estructura

$$\chi_n^2 = \sum_{i=1}^n X_i^2, \quad X_i \sim N(0, 1), \quad \forall i.$$

Nótese que el parámetro n se denominado grados de libertad (g.l.) y coincide con el número de variables aleatorias independientes usadas. Además, destacaremos que esta distribución no es simétrica y es reproductiva en sus grados de libertad. Esto es, si $X_i \sim \chi_{n_i}^2$, para $i = 1, \dots, k$, entonces

$$\sum_{i=1}^k X_i \sim \chi_{n_1 + \dots + n_k}^2.$$

Por otro lado, existen tablas tabuladas de las probabilidades de dicha distribución. En el interior de estas tablas aparecen para diferentes valores de los grados de libertad, n (primera columna), y diferentes valores de la probabilidad, p (primera fila), los valores que dejan a su izquierda una probabilidad igual a p .

Índice

Algunos modelos continuos de variable aleatoria

Uniforme

Exponencial

Gamma

Beta

Normal

Chi-Cuadrado

t-Student

F-Snedecor

Concepto de muestra y estadístico

Algunos ejemplos de estadísticos

Valor esperado y varianza de la media muestral

Valor esperado de la varianza y cuasivarianza muestral

Apéndice

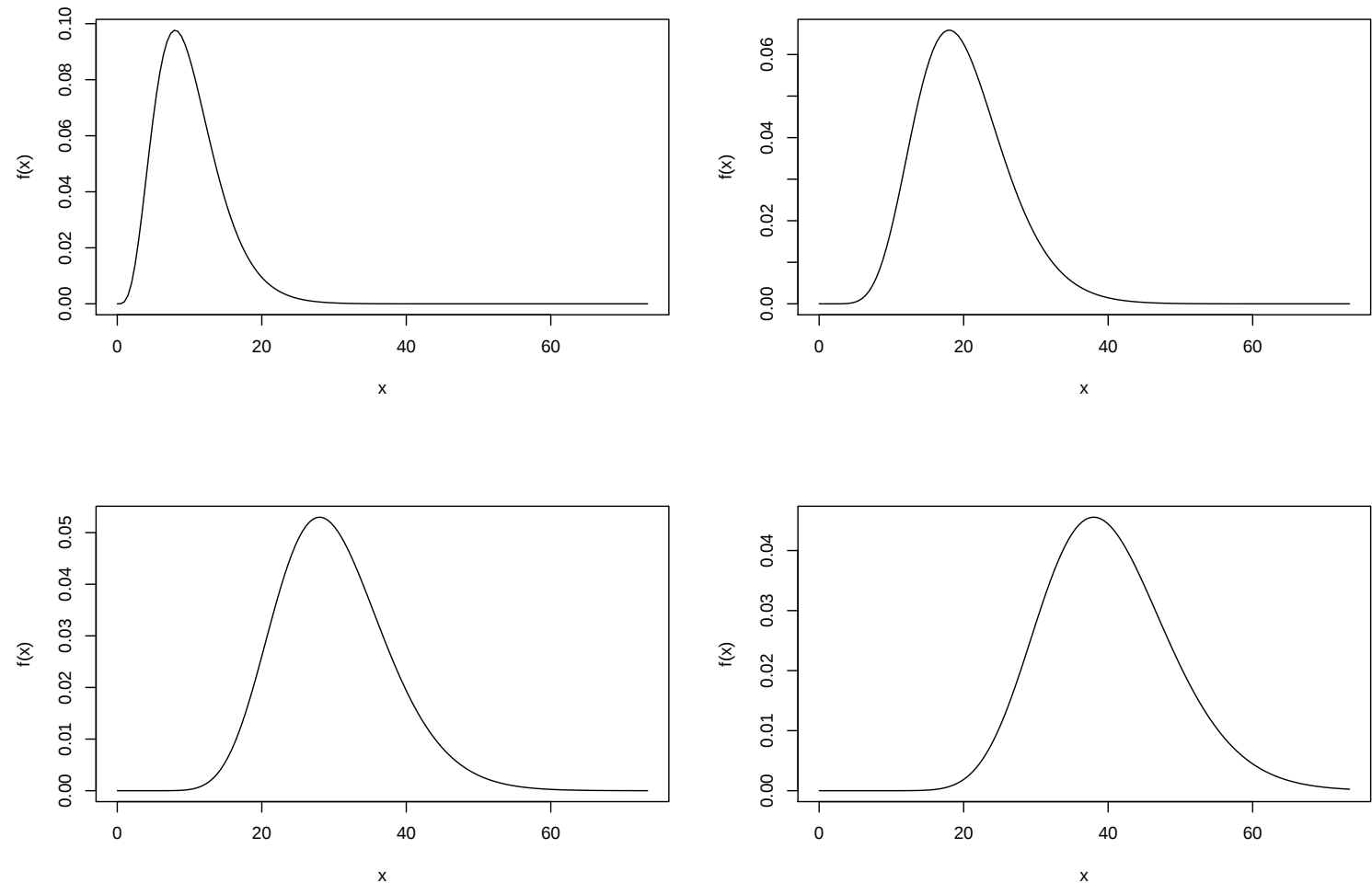


Figure 2: Representación de la función de densidad de una chi-cuadrado con 10, 20, 30 y 40 grados de libertad

Ejemplo 2 Dada $X \sim \chi_{10}^2$, determine el valor de x que haga que $P[X \leq x] = 0'005$ y $P[X \geq x] = 0'005$.

Buscando en la primera columna los grados de libertad (10) y en la primera columna el orden del cuantil (0'005 y 0'995, respectivamente) se tiene que

$$P[X \leq x] = 0'005 \implies x = 2'16,$$

$$P[X \geq x] = 0'005 \implies P[X \leq x] = 0'995 \implies x = 25'19.$$

Además, si $X \sim \chi_{20}^2$ e $Y \sim \chi_{14}^2$, calcule $P[X \leq 12'44]$, $P[Y \geq 21'06]$ y $P[Z \leq 24]$ siendo $Z = X + Y$.

Buscando en la primera columna los grados de libertad (20, 14 y 34, respectivamente) y en el interior de la tabla los distintos valores, se tiene que:

$$P[Y \geq 21'06] = 1 - P[Y \leq 21'06] = 1 - 0'9 = 0'1,$$

$$P[X \leq 12'44] = 0'1, \quad P[Z \leq 24] = 0'1.$$

Donde para calcular la última probabilidad se debería de haber interpolado. Sin embargo, como hicimos anteriormente, hemos considerado como válida la probabilidad del punto más cercano (en este caso 23'95). #

Índice

Algunos modelos
continuos de variable
aleatoria

Uniforme

Exponencial

Gamma

Beta

Normal

Chi-Cuadrado

t-Student

F-Snedecor

Concepto de muestra y
estadístico

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Apéndice

La distribución t -Student con n grados de libertad, que denotaremos por t_n , se construye como el cociente entre una variable aleatoria distribuida según una normal $(0, 1)$ y la raíz cuadrada de una chi-cuadrado de n grados de libertad dividida entre sus grados de libertad, siendo ambas distribuciones independientes.

Por tanto, la distribución t de Student responde a la siguiente estructura:

$$t_n = \frac{X}{\sqrt{\frac{Y}{n}}}, \quad X \sim N(0, 1), \quad Y \sim \chi_n^2.$$

Además, es importante destacar que:

- para tamaños muestrales altos se aproxima la distribución t de Student por la normal tipificada,
- que esta distribución es simétrica y
- que las probabilidades de esta distribución también se encuentran tabuladas de manera totalmente análoga a la de la chi-cuadrado.

Índice

Algunos modelos continuos de variable aleatoria

Uniforme

Exponencial

Gamma

Beta

Normal

Chi-Cuadrado

t-Student

F-Snedecor

Concepto de muestra y estadístico

Algunos ejemplos de estadísticos

Valor esperado y varianza de la media muestral

Valor esperado de la varianza y cuasivarianza muestral

Apéndice

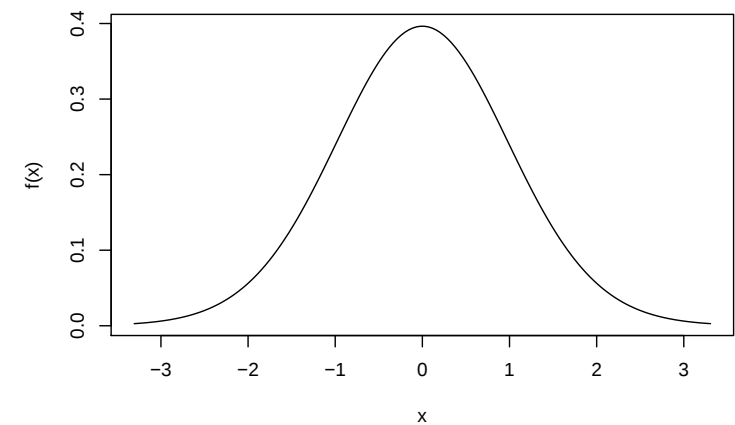
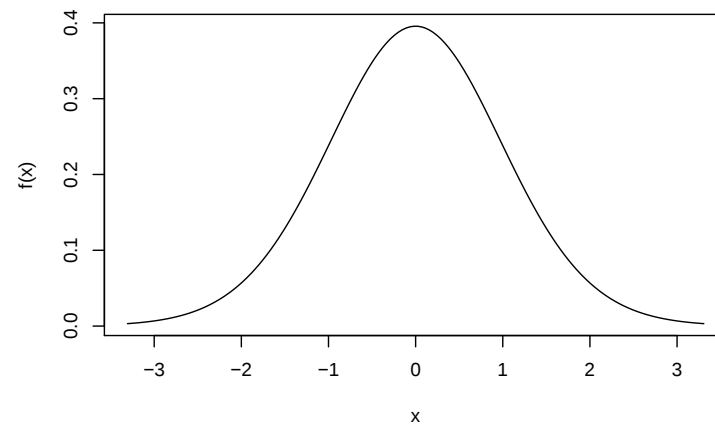
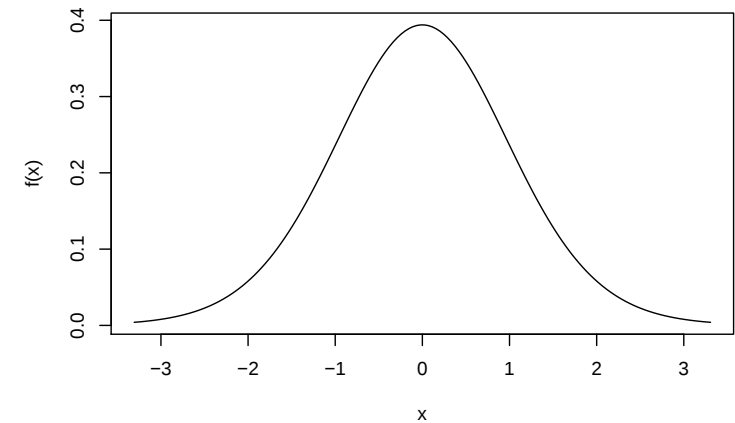
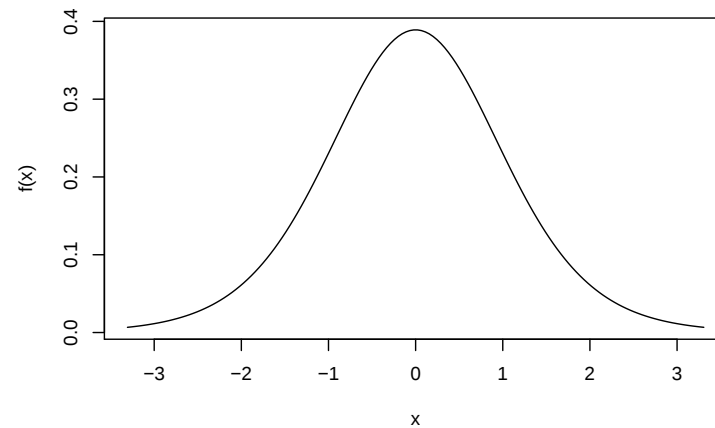


Figure 3: Representación de la función de densidad de una t-Student con 10, 20, 30 y 40 grados de libertad

Ejemplo 3 Sea X una variable aleatoria que se distribuye como una t de Student con 8 grados de libertad. Determine el valor de x tal que $P[X \leq x] = 0'95$ y $P[-x \leq X \leq x] = 0'99$.

- Buscando directamente en las tablas: $P[X \leq x] = 0'95 \implies x = 1'86$.
- Puesto que

$$\begin{aligned}
 P[-x \leq X \leq x] &= P[X \leq x] - P[X \leq -x] \\
 &= P[X \leq x] - P[X > x] \\
 &= P[X \leq x] - (1 - P[X \leq x]) \\
 &= 2P[X \leq x] - 1,
 \end{aligned}$$

entonces $2P[X \leq x] - 1 = 0'99$, con lo que $P[X \leq x] = 0'995$, y entonces buscando en las tablas $x = 3'355$.

Adviértase que en el cálculo de la segunda probabilidad hemos usado que la distribución t -Student es simétrica. ‡

Índice

Algunos modelos
continuos de variable
aleatoria

Uniforme

Exponencial

Gamma

Beta

Normal

Chi-Cuadrado

t-Student

F-Snedecor

Concepto de muestra y
estadístico

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Apéndice

Ejemplo 4 Sea X una variable aleatoria que se distribuye según una t de Student de 8 grados de libertad. Calcule las probabilidades siguientes: $P[X \leq 1'86]$, $P[X > 4'5]$, $P[X \leq -1'86]$ y $P[-1'86 \leq X \leq 1'86]$.

Buscando directamente en las tablas:

- $P[X \leq 1'86] = 0'95$.
- $P[X > 4'5] = 1 - P[X \leq 4'5] = 1 - 0'999 = 0'001$.
- Teniendo en cuenta la simetría de la distribución t de Student: $P[X \leq -1'86] = P[X > 1'86] = 1 - P[X \leq 1'86] = 1 - 0'95 = 0'05$.
- Igualmente, teniendo en cuenta la simetría y usando los apartados anteriores: $P[-1'86 \leq X \leq 1'86] = P[X \leq 1'86] - P[X \leq -1'86] = 0'9$.

#

La distribución F de Snedecor con n y m grados de libertad, que denotaremos por $F_{n,m}$, se construye a partir del cociente de dos variables aleatorias independientes y distribuidas según chi-cuadrados con n y m grados de libertad, respectivamente, divididas entre sus correspondientes grados de libertad.

Por tanto, la distribución F de Snedecor responde a la siguiente estructura:

$$F_{n,m} = \frac{\frac{X}{n}}{\frac{Y}{m}}, \quad X \sim \chi_n^2, Y \sim \chi_m^2.$$

Destacaremos de esta distribución que no es simétrica y que puesto que sólo disponemos de tablas para los valores de $p = 0'9, 0'95, 0'975, 0'99$ y $0'999$, la siguiente propiedad será muy útil en el caso en el que queramos trabajar con probabilidades complementarias a las anteriores:

$$F_{m,n}(1 - \alpha) = \frac{1}{F_{n,m}(\alpha)},$$

donde $F_{n,m}(\alpha)$ denota al punto de una F de Snedecor con n y m grados de libertad que deja a su izquierda una probabilidad α .

Índice

Algunos modelos continuos de variable aleatoria

Uniforme

Exponencial

Gamma

Beta

Normal

Chi-Cuadrado

t-Student

F-Snedecor

Concepto de muestra y estadístico

Algunos ejemplos de estadísticos

Valor esperado y varianza de la media muestral

Valor esperado de la varianza y cuasivarianza muestral

Apéndice

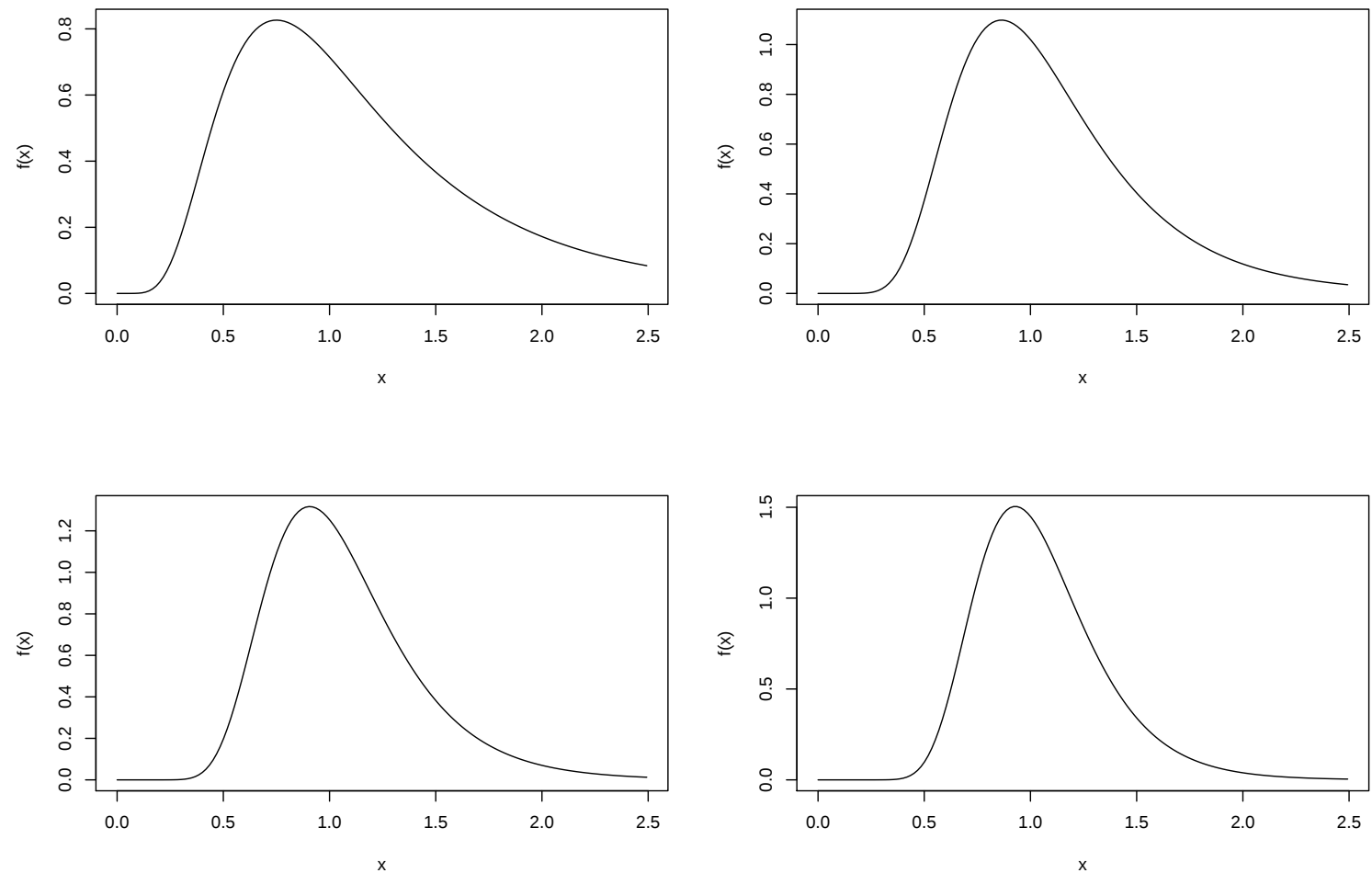


Figure 4: Representación de la función de densidad de una F-Snedecor con (20, 10), (40, 20), (60, 30) y (80, 40) grados de libertad

En las tablas de las probabilidades de la distribución F de Snedecor aparecen en la primera fila y en la primera columna los grados de libertad n y m , respectivamente. En el interior de cada una de las tablas se tienen, para los diferentes valores de n y m , los valores tales que a su izquierda dejan una probabilidad igual a p .

Ejemplo 5 Sea X una variable aleatoria que sigue una F de Snedecor de 24 y 40 grados de libertad. Determine el valor de x para que $P[X \leq x] = 0'05$.

Como la $F_{24,40}(0'05)$ no viene en las tablas, se utiliza la siguiente propiedad:

$$F_{n_1, n_2}(1 - \alpha) = \frac{1}{F_{n_2, n_1}(\alpha)},$$

que, en este caso, pasa a ser

$$F_{24,40}(0'05) = \frac{1}{F_{40,25}(0'95)} = \frac{1}{1'89} = 0'5291,$$

donde $F_{40,25}(0'95)$ se obtiene buscando en la tabla correspondiente a la probabilidad 0'95 los grados de libertad 40 (primera fila) y 24 (primera columna). #

Índice

Algunos modelos
continuos de variable
aleatoria

Uniforme

Exponencial

Gamma

Beta

Normal

Chi-Cuadrado

t-Student

F-Snedecor

Concepto de muestra y
estadístico

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Apéndice

Ejemplo 6 Sea X una variable aleatoria que se distribuye según una F de Snedecor de 6 y 8 grados de libertad. Calcúlese el valor de x para que $P[X \leq x] = 0'95$ y $P[X \leq x] = 0'99$.

Buscando directamente en las tablas correspondientes a las probabilidades 0'95 y 0'99, respectivamente, los grados de libertad 6 (primera fila) y 8 (primera columna) se obtiene que:

$$P[X \leq x] = 0'95 \implies x = 3'58, \quad P[X \leq x] = 0'99 \implies x = 6'37.$$

#

Índice

Algunos modelos
continuos de variable
aleatoria

Concepto de muestra y
estadístico

Muestra

Estadístico

Estadístico muestral y
parámetro poblacional

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Apéndice

Concepto de muestra y estadístico

Índice

Algunos modelos
continuos de variable
aleatoria

Concepto de muestra y
estadístico

Muestra

Estadístico

Estadístico muestral y
parámetro poblacional

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Apéndice

Un problema fundamental en la estadística consiste en estudiar alguna característica desconocida de una distribución de probabilidad de una variable aleatoria. Aunque dicha cuestión será estudiada en los próximos temas, se presentan a continuación conceptos necesarios para realizar dicho estudio.

Muestra: Subconjunto de elementos de la población. Es fundamental elegir adecuadamente los elementos de la muestra y el tamaño. Normalmente trabajaremos con lo que se conoce como muestreo aleatorio simple, que consiste en que todos los elementos de la población tienen la misma probabilidad de estar en la muestra y son independientes entre ellos.

Parámetro poblacional: Características estocásticas de la población. Los más usados, y por tanto conocidos, son la esperanza matemática o media y la varianza¹, es decir, $E[X]$ y $Var(X)$. El conocimiento del parámetro poblacional permite describir parcial o totalmente la distribución de probabilidad de la variable que estamos investigando.

¹Por ejemplo, en la distribución Binomial son $n \cdot p$ y $n \cdot p \cdot (1 - p)$, en la Poisson es λ y en la Normal son μ y σ^2 .

Índice

Algunos modelos
continuos de variable
aleatoria

Concepto de muestra y
estadístico

Muestra

Estadístico

Estadístico muestral y
parámetro poblacional

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Apéndice

Un **estadístico muestral** es cualquier función real de las variables aleatorias que integran la muestra, la cual no contiene ningún valor o parámetro desconocido.

En general, el estadístico muestral se representa como función de las observaciones muestrales y, por tanto, para cada muestra concreta el estadístico tomará un valor diferente. Es de la forma $T(X_1, \dots, X_n)$ donde $X_1, \dots, X_n$ son las v.a. de las muestras. A medida que vamos tomando muestras distintas se obtienen distintos valores para el estadístico, resultando que efectivamente el estadístico es también una variable aleatoria y, por consiguiente, tendrá su propia distribución a la que llamaremos distribución muestral del estadístico y que se estudiará en el capítulo 3.

Notaremos con letras mayúsculas, $X_1, \dots, X_n$, las v.a. de la muestra y con letras minúsculas, $x_1, \dots, x_n$, a los valores concretos que pudiera tomar cada una de ellas.

Ejemplo 7 Supongamos que se quiere conocer el salario medio de la población femenina granadina. Consultando los datos del Instituto Nacional de Estadística (www.ine.es) se sabe que el 1 de enero de 2009 dicha población asciende a 455.703 individuos. Dada la dificultad para recabar información sobre el salario de cada uno de los elementos, nos planteamos obtener una muestra, por ejemplo, de 50.000 elementos. De dicha muestra se obtendrá un salario medio, pero es evidente que para cada realización de la muestra se obtendrá un salario medio distinto, lo que evidencia el carácter aleatorio del estadístico muestral.

#

Estadístico muestral y parámetro poblacional

Índice

Algunos modelos
continuos de variable
aleatoria

Concepto de muestra y
estadístico

Muestra

Estadístico

Estadístico muestral y
parámetro poblacional

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Apéndice

Resumiendo, se entiende por parámetro poblacional a las características desconocidas de las distribuciones de probabilidad, mientras que el estadístico muestral es un resumen de la información existente en la muestra.

En el siguiente cuadro tenemos algunos parámetros poblacionales junto a sus correspondientes estadísticos muestrales.

Parámetros Poblacionales	Estadísticos Muestrales
Media: $E[X]$	Media muestral: $\bar{X}$
Varianza: $Var(X)$	Varianza muestral: S_n^2
	Cuasivarianza muestral: S_{n-1}^2
Proporción poblacional: p	Proporción muestral: $\hat{p} = \sum_{i=1}^n \frac{A_i}{n}$
	donde $A_i = \begin{cases} 1 & , \text{ ocurre } i \\ 0 & , \text{ en otro caso} \end{cases}$

Índice

Algunos modelos
continuos de variable
aleatoria

Concepto de muestra y
estadístico

Algunos ejemplos de
estadísticos

Media muestral
Varianza y
cuasivarianza muestral

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Apéndice

Algunos ejemplos de estadísticos: media, varianza y cuasivarianza muestral

Dada una muestra aleatoria simple $X_1, X_2, \dots, X_n$ de tamaño n de una variable aleatoria X , se define la **media muestral** como

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = \frac{X_1 + X_2 + \dots + X_n}{n}.$$

Ejemplo 8 Sea X la variable aleatoria definida como tiempo (en horas) transcurrido en realizar el examen final de cierta asignatura. Durante la última convocatoria se anotó el tiempo que tardó cada alumno en realizar dicho examen, obteniéndose la siguiente muestra: 1, 1'7, 1'8, 1'5, 1'85, 1'3, 2, 1'9, 1'6, 1'75, 1'8, 1'9.

En este caso, la media muestral será:

$$\bar{x} = \frac{1 + 1'7 + 1'8 + 1'5 + 1'85 + \dots + 1'75 + 1'8 + 1'9}{12} = \frac{20'1}{12} = 1'675.$$

Es decir, la media de la muestra obtenida en realizar el examen es de 1 hora y 40'5 minutos.

#

Varianza y cuasivarianza muestral

Índice

Algunos modelos
continuos de variable
aleatoria

Concepto de muestra y
estadístico

Algunos ejemplos de
estadísticos

Media muestral

Varianza y
cuasivarianza muestral

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Apéndice

Dada una muestra aleatoria simple $X_1, X_2, \dots, X_n$ de tamaño n de una variable aleatoria X , se define la **varianza muestral** como

$$S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2,$$

mientras que la **cuasivarianza muestral** será

$$S_{n-1}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2.$$

Es evidente que la varianza y cuasivarianza muestral se relacionan mediante la siguiente identidad:

$$n \cdot S_n^2 = (n-1) \cdot S_{n-1}^2.$$

Luego conocida la varianza muestral y el tamaño de la muestra se puede obtener la cuasivarianza muestral y viceversa.

Adviértase que los programas informáticos especializados en estadística calculan siempre la cuasivarianza muestral, ya que como veremos en futuros temas, tiene mejores propiedades.

Varianza y cuasivarianza muestral

Índice

Algunos modelos
continuos de variable
aleatoria

Concepto de muestra y
estadístico

Algunos ejemplos de
estadísticos

Media muestral
Varianza y
cuasivarianza muestral

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Apéndice

Ejemplo 9 Para la muestra anterior se obtiene que $S_n^2 = \frac{0'9075}{12} = 0'0756$ y que $S_{n-1}^2 = \frac{0'9075}{11} = 0'0825$ ya que:

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
1	-0'675	0'4556
1'7	0'025	0'0006
1'8	0'125	0'0156
1'5	-0'175	0'0306
1'85	0'175	0'0306
1'3	-0'375	0'1406
2	0'325	0'1056
1'9	0'225	0'0506
1'6	-0'075	0'0056
1'75	0'075	0'0056
1'8	0'125	0'0156
1'9	0'225	0'0506
		0'9075

#

Índice

Algunos modelos
continuos de variable
aleatoria

Concepto de muestra y
estadístico

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
media muestral

Varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Apéndice

Valor esperado y varianza de la media muestral

Valor esperado de la media muestral

Índice

Algunos modelos
continuos de variable
aleatoria

Concepto de muestra y
estadístico

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
media muestral

Varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Apéndice

Dada una variable aleatoria X , con media $E[X]$ y varianza $Var(X)$, y dada una muestra aleatoria simple, $X_1, \dots, X_n$, de tamaño n procedente de dicha población, se puede demostrar que la esperanza de la media muestral es igual a la media de la población, $E[X]$.

Esto es:

$$E[\bar{X}] = E[X].$$

En efecto,

$$\begin{aligned} E[\bar{X}] &= E\left[\frac{1}{n} \sum_{i=1}^n X_i\right] = \frac{1}{n} \sum_{i=1}^n E[X_i] \\ &= \frac{1}{n} \sum_{i=1}^n E[X] = \frac{1}{n} \cdot n \cdot E[X] = E[X], \end{aligned}$$

donde se ha usado que $E[X_i] = E[X]$, para $i = 1, \dots, n$, ya que partimos de una muestra aleatoria simple.

Dada una variable aleatoria X , con media $E[X]$ y varianza $Var(X)$, y dada una muestra aleatoria simple, $X_1, \dots, X_n$, de tamaño n procedente de dicha población, se puede demostrar que la varianza de la media muestral es igual a la varianza poblacional, $Var(X)$, dividida por n . Es decir:

$$Var(\bar{X}) = \frac{Var(X)}{n}.$$

En efecto, partiendo de la definición de varianza y usando la expresión de la media muestral, se tiene que:

$$\begin{aligned} Var(\bar{X}) &= E[(\bar{X} - E[\bar{X}])^2] = E\left[\left(\frac{1}{n} \sum_{i=1}^n X_i - E\left[\frac{1}{n} \sum_{i=1}^n X_i\right]\right)^2\right] \\ &= E\left[\frac{1}{n^2} \left(\sum_{i=1}^n X_i - \sum_{i=1}^n E[X_i]\right)^2\right] \\ &= \frac{1}{n^2} E\left[\left(\sum_{i=1}^n (X_i - E[X_i])\right)^2\right]. \end{aligned}$$

Varianza de la media muestral

Índice

Algunos modelos
continuos de variable
aleatoria

Concepto de muestra y
estadístico

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
media muestral

Varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Apéndice

Por otro lado, el cuadrado de una suma de n elementos responde a la siguiente expresión:

$$\begin{aligned}\left(\sum_{i=1}^n (X_i - E[X_i])\right)^2 &= \left(\sum_{i=1}^n (X_i - E[X_i])\right) \cdot \left(\sum_{i=1}^n (X_i - E[X_i])\right) \\ &= \sum_{i=1}^n (X_i - E[X_i])^2 \\ &\quad + \sum_{i,j=1 \atop i \neq j}^n (X_i - E[X_i])(X_j - E[X_j]),\end{aligned}$$

sin más que desarrollar la sumatoria, multiplicar los dos paréntesis y agrupar de forma conveniente.

En tal caso, teniendo en cuenta que al partir de una muestra independiente se verifica que

$$\text{Cov}[X_i, X_j] = 0, \quad \forall i \neq j, \quad i, j = 1, \dots, n,$$

se tiene que

$$\begin{aligned} \text{Var}(\bar{X}) &= \frac{1}{n^2} E \left[\sum_{i=1}^n (X_i - E[X_i])^2 + \sum_{i,j=1 \atop i \neq j}^n (X_i - E[X_i])(X_j - E[X_j]) \right] \\ &= \frac{1}{n^2} \left(\sum_{i=1}^n E [X_i - E[X_i]]^2 \right. \\ &\quad \left. + \sum_{i,j=1 \atop i \neq j}^n E [(X_i - E[X_i])(X_j - E[X_j])] \right) \\ &= \frac{1}{n^2} \left(\sum_{i=1}^n \text{Var}(X_i) + \sum_{i,j=1 \atop i \neq j}^n \text{Cov}[X_i, X_j] \right) \\ &= \frac{1}{n^2} \cdot n \cdot \text{Var}(X) = \frac{\text{Var}(X)}{n}, \end{aligned}$$

donde se ha usado

$$\text{Var}(X_1) = \cdots = \text{Var}(X_n) = \text{Var}(X),$$

por ser un conjunto de variables idénticamente distribuidas.

Índice

Algunos modelos
continuos de variable
aleatoria

Concepto de muestra y
estadístico

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Valor esperado de la
varianza y
cuasivarianza muestral

Apéndice

Valor esperado de la varianza y cuasivarianza muestral

Valor esperado de la varianza y cuasivarianza

Índice

Algunos modelos
continuos de variable
aleatoria

Concepto de muestra y
estadístico

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Valor esperado de la
varianza y
cuasivarianza muestral

Apéndice

Dada una variable aleatoria X , con valor esperado $E[X]$ y varianza $Var(X)$, y dada una muestra aleatoria simple, $X_1, \dots, X_n$, de tamaño n procedente de dicha población, la esperanza de la cuasivarianza muestral es igual a la varianza de la población, $Var(X)$.

Es decir:

$$E[S_{n-1}^2] = Var(X).$$

Además, también se verifica que

$$E[S_n^2] = \frac{n-1}{n} \cdot Var(X).$$

En efecto, partiendo de la expresión

$$S_{n-1}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2,$$

si se introduce en la sumatoria la media, $E[X]$, agrupamos de forma conveniente y desarrollamos el cuadrado se obtiene

Valor esperado de la varianza y cuasivarianza

Índice

Algunos modelos
continuos de variable
aleatoria

Concepto de muestra y
estadístico

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Valor esperado de la
varianza y
cuasivarianza muestral

Apéndice

$$\begin{aligned}\sum_{i=1}^n (X_i - \bar{X})^2 &= \sum_{i=1}^n (X_i - E[X] + E[X] - \bar{X})^2 \\&= \sum_{i=1}^n [(X_i - E[X]) - (\bar{X} - E[X])]^2 \\&= \sum_{i=1}^n [(X_i - E[X])^2 + (\bar{X} - E[X])^2 - 2(X_i - E[X])(\bar{X} - E[X])] \\&= \sum_{i=1}^n (X_i - E[X])^2 + \sum_{i=1}^n (\bar{X} - E[X])^2 \\&\quad - 2(\bar{X} - E[X]) \sum_{i=1}^n (X_i - E[X]) \\&= \sum_{i=1}^n (X_i - E[X])^2 + n(\bar{X} - E[X])^2 - 2(\bar{X} - E[X])n(\bar{X} - E[X]) \\&= \sum_{i=1}^n (X_i - E[X])^2 - n(\bar{X} - E[X])^2,\end{aligned}$$

donde se ha usado también que $\sum_{i=1}^n X_i = n \cdot \bar{X}$.

Valor esperado de la varianza y cuasivarianza

Índice

Algunos modelos
continuos de variable
aleatoria

Concepto de muestra y
estadístico

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Valor esperado de la
varianza y
cuasivarianza muestral

Apéndice

Luego, tomando esperanzas en la expresión anterior y usando que la esperanza matemática es lineal

$$\begin{aligned} E \left[\sum_{i=1}^n (X_i - \bar{X})^2 \right] &= E \left[\sum_{i=1}^n (X_i - E[X])^2 - n (\bar{X} - E[X])^2 \right] \\ &= \sum_{i=1}^n E [(X_i - E[X])^2] - n E [(\bar{X} - E[X])^2] . \end{aligned}$$

Entonces, teniendo en cuenta la definición de varianza, $Var(X_i) = E [(X_i - E[X_i])^2]$, y que, como hemos demostrado con anterioridad, $E[X] = E[\bar{X}]$, se tiene que

$$\begin{aligned} E \left[\sum_{i=1}^n (X_i - \bar{X})^2 \right] &= \sum_{i=1}^n Var(X_i) - n E [(\bar{X} - E[\bar{X}])^2] \\ &= \sum_{i=1}^n Var(X) - n Var(\bar{X}) = n Var(X) - n \cdot \frac{Var(X)}{n} \\ &= n Var(X) - Var(X) = (n - 1) \cdot Var(X), \end{aligned}$$

Valor esperado de la varianza y cuasivarianza

Índice

Algunos modelos
continuos de variable
aleatoria

Concepto de muestra y
estadístico

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Valor esperado de la
varianza y
cuasivarianza muestral

Apéndice

donde se ha usado además que $Var(X_i) = Var(X)$, ya que se parte de una muestra aleatoria simple de X , y que $Var(\bar{X}) = \frac{Var(X)}{n}$, como se demostró con anterioridad. Finalmente, dividiendo la expresión anterior por $n - 1$, se obtiene que

$$E \left[\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \right] = Var(X),$$

esto es, $E[S_{n-1}^2] = Var(X)$. Mientras que si se divide por n

$$\begin{aligned} E[S_n^2] &= E \left[\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \right] = \frac{1}{n} \cdot E \left[\sum_{i=1}^n (X_i - \bar{X})^2 \right] \\ &= \frac{1}{n} \cdot (n-1) \cdot Var(X) = \frac{n-1}{n} \cdot Var(X). \end{aligned}$$

Índice

Algunos modelos
continuos de variable
aleatoria

Concepto de muestra y
estadístico

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Apéndice

Variables Aleatorias

Distribución binomial

Distribución de Poisson

Aproximaciones de las
distribuciones binomial
y Piosson

Apéndice

Una variable aleatoria, X , es una función $X : \Omega \longrightarrow \mathbb{R}$ tal que su valor está determinado por el azar (sabemos que valores toma pero no se tiene certeza de su ocurrencia).

Podemos clasificar las variables aleatorias en dos familias:

1) Discretas: Cuando X toma un valor finito (1, 2 y 3) o infinito numerable de valores (2, 4, 6, 8, ...). En este caso,

☐ La función de cuantía es: $p_i = P[X = x_i]$ para cada valor x_i que toma X , $i = 1, \dots$

☐ La función de distribución: $P[X \leq x] = \sum_{x_i \leq x} p_i$.

2) Continuas: Cuando X toma un valor infinito no numerable de valores ($\mathbb{R}, \dots$). En este caso,

☐ La función de densidad es: $f(x)$ que verifica que $\int_{-\infty}^{+\infty} f(x)dx = 1$ y que $f(x) \geq 0$. Además $\int_a^b f(x)dx = P[X \in (a, b)]$.

☐ La función de distribución es: $F(x) = \int_{-\infty}^x f(u)du$. Es no decreciente, continua, $F(+\infty) = 1$, $F(-\infty) = 0$ y $F'(x) = f(x)$.

Variables Aleatorias: Esperanza y Varianza

Índice

Algunos modelos
continuos de variable
aleatoria

Concepto de muestra y
estadístico

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Apéndice

Variables Aleatorias

Distribución binomial

Distribución de Poisson

Aproximaciones de las
distribuciones binomial
y Piosson

1) Variables aleatorias discretas:

$$\square E[X] = \sum_i x_i p_i.$$

$$\square Var(X) = E[(X - E[X])^2] = \sum_i (x_i - E[X])^2 p_i.$$

2) Variables aleatorias continuas:

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

$$\square Var(X) = E[(X - E[X])^2] = \int_{-\infty}^{\infty} (x - E[X])^2 f(x) dx.$$

Propiedades:

- $E[X + Y] = E[X] + E[Y],$
- $E[aX + b] = aE[X] + b,$
- $Var(X) = E[X^2] - E[X]^2,$
- $Var(X) \geq 0,$
- $Var(aX + b) = a^2 Var(X).$

A partir de la distribución de Bernouilli, definimos la distribución Binomial como *la suma de éxitos que se obtienen en n repeticiones de un experimento aleatorio tipo Bernoulli*, es decir $X = X_1 + X_2 + \dots + X_n$, donde X_i ($i = 1, 2, \dots, n$) es una variable aleatoria tipo Bernouilli.

Por tanto, la distribución Binomial es la distribución de probabilidad de n variables aleatorias independientes tipo Bernouilli en las que la probabilidad de éxito, p , se mantiene constante en cada ensayo.

Formalmente, la definición de una variable aleatorio binomial es la siguiente:

Definición 9 Una variable aleatoria discreta, X , se dice que sigue una distribución Binomial, lo cual denotaremos como $X \sim B(n, p)$, si, y sólo si, la función de cuantía asociada viene dada por

$$P[X = x] = \binom{n}{x} \cdot p^x \cdot q^{n-x} = \frac{n!}{x! \cdot (n-x)!} \cdot p^x \cdot q^{n-x},$$

donde $x = 0, 1, \dots, n$, $0 \leq p \leq 1$ y $q = 1 - p$.

Además, se verifica que $E[X] = n \cdot p$ y $Var(X) = n \cdot p \cdot q$.

Índice

Algunos modelos
continuos de variable
aleatoria

Concepto de muestra y
estadístico

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Apéndice

Variables Aleatorias

Distribución binomial

Distribución de Poisson

Aproximaciones de las
distribuciones binomial
y Piosson

Una primera interpretación de la distribución de Poisson se obtiene al observar que surge como el límite de la distribución binomial cuando el número de ensayos, n , tiende a infinito ($n \rightarrow \infty$) y la probabilidad de éxito, p , tiende a cero ($p \rightarrow 0$), manteniéndose constante el valor de la media, $n \cdot p$, al que denominamos parámetro de la distribución y lo representamos por λ .

Además de ser útil como aproximación de las probabilidades binomiales, se usa para distribuir variables del tipo número de veces que ha ocurrido un evento en una unidad dada de tiempo o espacio. Por ello a la distribución de Poisson también se le conoce como *número de ocurrencias*, *por término medio*, *de un suceso en un determinado intervalo de tiempo o de área dado*.

Formalmente, la definición es la siguiente:

Definición 10 Una variable aleatoria discreta, X , se dice que sigue una distribución de Poisson, que denotaremos como $X \sim P(\lambda)$, si, y sólo si, la función de cuantía asociada viene dada por

$$P[X = x] = \frac{\lambda^x}{x!} \cdot e^{-\lambda}, \quad x = 0, 1, 2, \dots, \quad \lambda > 0.$$

Además, se verifica que $E[X] = \lambda = Var(X)$.

Aproximaciones de las distribuciones binomial y Piosson

Índice

Algunos modelos
continuos de variable
aleatoria

Concepto de muestra y
estadístico

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Apéndice

Variables Aleatorias

Distribución binomial

Distribución de Poisson

Aproximaciones de las
distribuciones binomial
y Piosson

Distribución original	Condiciones	Aproximación
$B(n, p)$	$n > 20, p < 0'05$	$P(np)$
$B(n, p)$	$n \geq 25, p \geq 0'05, np \geq 5$	$N(np, np(1 - p))$
$P(\lambda)$	$\lambda > 10$	$N(\lambda, \lambda)$

Índice

Algunos modelos
continuos de variable
aleatoria

Concepto de muestra y
estadístico

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Apéndice

Ejercicios Propuestos

Relación de ejercicios propuestos

Índice

Algunos modelos
continuos de variable
aleatoria

Concepto de muestra y
estadístico

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Apéndice

1.1

Se ha comprobado que el tiempo necesario para realizar una transferencia bancaria por internet sigue una distribución normal de media 120 segundos y desviación típica 40 segundos. Se pide:

- (a) ¿Cuál es la probabilidad de que una persona requiera menos de 90 segundos para realizar su transferencia bancaria?
- (b) ¿Cuál es la probabilidad de que una persona requiera entre dos y tres minutos para dicha tarea?
- (c) ¿Cuánto tiempo como mínimo necesitan para realizar la transferencia el 15% de las personas que más tardan?

1.2

Suponga que el peso, en gramos, de los paquetes de folios de una determinada marca sigue una variable aleatoria normal. Durante un periodo largo de tiempo se ha observado que el 5% de los paquetes pesan menos de 315 gramos y que el 15% de los paquetes pesan más de 323. Calcule la media y la varianza del peso de los paquetes de folio de dicha marca. ¿Cuál es la probabilidad de que un paquete de folios pese más de 321 gramos?

Relación de ejercicios propuestos

Índice

Algunos modelos
continuos de variable
aleatoria

Concepto de muestra y
estadístico

Algunos ejemplos de
estadísticos

Valor esperado y
varianza de la media
muestral

Valor esperado de la
varianza y cuasivarianza
muestral

Apéndice

- 1.3** Se sabe que el número de horas semanales que un estudiante navega por internet es una variable aleatoria normal con media 12 y varianza 4. ¿Cuál es la probabilidad de que un estudiante utilice internet más de 14 horas semanales?
- 1.4** Para aprobar unas oposiciones se necesitan obtener 100 puntos, o más, en una prueba. Por experiencias anteriores se sabe que la distribución de los puntos obtenidos por los opositores es una normal de media 110 puntos y desviación típica 15. ¿Qué probabilidad hay de que un opositor apruebe? Si sabemos que hay 1000 opositores y 300 plazas, ¿cuántos puntos se deberán exigir para ajustar el número de plazas al número de opositores aprobados?.