

Inferencia Estadística

FACULTAD
DE
CIENCIAS
UNIVERSIDAD DE GRANADA



Los Del DGIIM, losdeldgiim.github.io

Doble Grado en Ingeniería Informática y Matemáticas
Universidad de Granada

se crean derivados de estos datos originales y no para fines comerciales.

Inferencia Estadística

Los Del DGIIM, losdeldgiim.github.io

Granada, 2025

Índice general

1. Ejercicios de clase

Esta sección tiene el propósito de recoger todos los ejercicios propuestos en clase por parte de la profesora y que fueron resueltos por los alumnos en pizarra.

1.1. Estadísticos muestrales

Ejercicio 1.1.1. Obtener la función masa de probabilidad conjunta de una m.a.s. de $X \rightsquigarrow B(k_0, p)$ y la función de densidad de una m.a.s. de $X \rightsquigarrow U(a, b)$.

Recordamos que si $X \rightsquigarrow B(k_0, p)$, entonces:

$$P[X = x] = \binom{k_0}{x} p^x (1-p)^{n-x} \quad \forall x \in \{0, \dots, k_0\}$$

Por lo que si tenemos una m.a.s. de n variables independientes e idénticamente distribuidas a X , (X_1, \dots, X_n) , su función de densidad vendrá dada por:

$$\begin{aligned} P[X_1 = x_1, \dots, X_n = x_n] &\stackrel{\text{indep.}}{=} \prod_{i=1}^n P[X_i = x_i] \stackrel{\text{id. d.}}{=} \prod_{i=1}^n P[X = x_i] \\ &= \prod_{i=1}^n \binom{k_0}{x_i} p^{x_i} (1-p)^{k_0-x_i} = p^{\sum_{i=1}^n x_i} (1-p)^{nk_0 - \sum_{i=1}^n x_i} \prod_{i=1}^n \binom{k_0}{x_i} \\ &\quad \forall x_i \in \{0, \dots, k_0\} \end{aligned}$$

Si ahora $X \rightsquigarrow U(a, b)$ para ciertos $a, b \in \mathbb{R}$ con $a < b$, entonces:

$$f_X(x) = \frac{1}{b-a} \quad \forall x \in [a, b]$$

de donde:

$$f_{(X_1, \dots, X_n)}(x_1, \dots, x_n) \stackrel{\text{indep.}}{=} \prod_{i=1}^n f_{X_i}(x_i) \stackrel{\text{id. d.}}{=} \prod_{i=1}^n f_X(x_i) = \prod_{i=1}^n \frac{1}{b-a} = \frac{1}{(b-a)^n} \quad \forall x \in [a, b]$$

Ejercicio 1.1.2. Para cada realización muestral, $(x_1, \dots, x_n) \in \mathcal{X}^n$, F_{x_1, \dots, x_n}^* es una función de distribución en \mathbb{R} . En particular es una función a saltos, con saltos de amplitud $1/n$ en los sucesivos valores muestrales ordenados de menor a mayor, supuestos que sean distintos, y de saltos múltiples en el caso de que varios valores muestrales coincidieran.

En las condiciones del enunciado, es decir, suponiendo que x_1, \dots, x_n están ordenados de menor a mayor y son distintos, entonces es fácil ver que:

$$F_{x_1, \dots, x_n}^*(x) = \begin{cases} 0 & \text{si } x < x_1 \\ 1/n & \text{si } x_1 \leq x < x_2 \\ \vdots & \\ 1 & \text{si } x > x_n \end{cases} \quad \forall x \in \mathbb{R}$$

Por lo que es claro que F_{x_1, \dots, x_n}^* es no decreciente, continua por la derecha, con límite 0 en $-\infty$ y con límite 1 en $+\infty$.

Ejercicio 1.1.3. $\forall x \in \mathbb{R}$, $F_{X_1, \dots, X_n}^*(x)$ es una variable aleatoria tal que $nF_{X_1, \dots, X_n}^*(x) \rightsquigarrow B(n, F(x))$ y:

$$E[F_{X_1, \dots, X_n}^*(x)] = F(x), \quad \text{Var}[F_{X_1, \dots, X_n}^*(x)] = \frac{F(x)(1 - F(x))}{n}$$

donde $F(x)$ es la función de distribución de X .

Recordamos que:

$$F_{X_1, \dots, X_n}^*(x) = \frac{1}{n} \sum_{i=1}^n I_{]-\infty, x]}(X_i) \quad \forall x \in \mathbb{R}$$

Fijado $x \in \mathbb{R}$, tenemos que $I_{]-\infty, x]}(X) \rightsquigarrow B(1, P[X \leq x]) \equiv B(1, F(x))$, por lo que por la propiedad reproductiva de la binomial tenemos que:

$$nF_{X_1, \dots, X_n}^*(x) \rightsquigarrow B(n, F(x))$$

Por lo que:

$$nE[F_{X_1, \dots, X_n}^*(x)] = E[nF_{X_1, \dots, X_n}^*(x)] = nF(x)$$

de donde:

$$E[F_{X_1, \dots, X_n}^*(x)] = F(x)$$

Para la varianza:

$$n^2 \text{Var}[F_{X_1, \dots, X_n}^*(x)] = \text{Var}[nF_{X_1, \dots, X_n}^*(x)] = nF(x)(1 - F(x))$$

de donde:

$$\text{Var}[F_{X_1, \dots, X_n}^*(x)] = \frac{F(x)(1 - F(x))}{n}$$

Ejercicio 1.1.4. Para valores grandes de n , en virtud del Teorema Central del Límite:

$$F_{X_1, \dots, X_n}^*(x) \rightsquigarrow \mathcal{N}\left(F(x), \frac{F(x)(1 - F(x))}{n}\right)$$

Sea (X_1, \dots, X_n) una m.a.s. de n muestras, sea:

$$S_n = \sum_{i=1}^n I_{]-\infty, x]}(X_i) \quad \forall n \in \mathbb{N}$$

Por el Teorema Central del Límite tenemos que:

$$\frac{S_n - E[S_n]}{\sqrt{Var[S_n]}} \xrightarrow{n \rightarrow \infty} \mathcal{N}(0, 1) \implies S_n \xrightarrow{n \rightarrow \infty} \mathcal{N}\left(F(x), \frac{F(x)(1 - F(x))}{n}\right)$$

Como $S_n \rightsquigarrow B(n, F(x))$, entonces tenemos que:

$$\begin{aligned} E[S_n] &= nF(x) \\ Var[S_n] &= nF(x)(1 - F(x)) \end{aligned}$$

Por lo que:

$$F_{X_1, \dots, X_n}^*(x) = \frac{1}{n} S_n \xrightarrow{n \rightarrow \infty} \mathcal{N}\left(F(x), \frac{F(x)(1 - F(x))}{n}\right)$$

Ejercicio 1.1.5. Dada una muestra aleatoria simple formada por las observaciones (3, 8, 5, 4, 5), obtener su función de distribución muestral y realizar la representación gráfica.

Aplicando la definición de la función de distribución muestral obtenemos que:

$$F_{(3,8,5,4,5)}^*(x) = \begin{cases} 0 & \text{si } x < 3 \\ 1 & \text{si } 3 \leq x < 4 \\ 2 & \text{si } 4 \leq x < 5 \\ 4 & \text{si } 5 \leq x < 8 \\ 5 & \text{si } x \geq 8 \end{cases}$$



Figura 1.1: Representación gráfica de $F_{(3,8,5,4,5)}^*(x)$.

Ejercicio 1.1.6. Sea X una variable aleatoria con distribución $B(1, p)$ con $p \in (0, 1)$. Se toma una muestra de tamaño 5, $(X_1, X_2, X_3, X_4, X_5)$, y se obtiene la siguiente observación (0, 1, 1, 0, 0). Determinar el valor de los estadísticos estudiados en la observación.

Aplicando las fórmulas vistas en clase obtenemos:

- Media: 0,4.
- Varianza: 0,24.

- Cuasivarianza: 0,3.
- $x_{(1)} = 0, x_{(2)} = 0, x_{(3)} = 0, x_{(4)} = 1, x_{(5)} = 1.$

Ejercicio 1.1.7. Sea (X_1, \dots, X_n) una m.a.s. y $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$, entonces:

$$M_{\bar{X}}(t) = (M_X(t/n))^n$$

$$M_{\bar{X}}(t) = E[e^{t\bar{X}}] = E\left[e^{\frac{t}{n} \sum_{i=1}^n X_i}\right] = M_{\sum_{i=1}^n X_i}\left(\frac{t}{n}\right) \stackrel{\text{indep.}}{=} \prod_{i=1}^n M_{X_i}\left(\frac{t}{n}\right) \stackrel{\text{id. d.}}{=} \left(M_X\left(\frac{t}{n}\right)\right)^n$$

Ejercicio 1.1.8. Obtener la distribución muestral de \bar{X} para (X_1, \dots, X_n) una m.a.s. de $X \rightsquigarrow \mathcal{N}(\mu, \sigma^2)$.

$$M_{\bar{X}}(t) = \left(M_X\left(\frac{t}{n}\right)\right)^n = \left(e^{\mu t + \frac{\sigma^2 t^2}{2n}}\right)^n = e^{\mu t + \frac{\sigma^2 t^2}{2}}$$

Luego $\bar{X} \rightsquigarrow \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$, ya que la función generatriz de momentos caracteriza la distribución.

Proposición 1.1. Si tenemos una m.a.s. (X_1, \dots, X_n) , entonces:

$$\begin{aligned} F_{X_{(n)}}(x) &= (F_X(x))^n \quad \forall x \in \mathbb{R} \\ F_{X_{(1)}}(x) &= 1 - (1 - F_X(x))^n \end{aligned}$$

Demostración. Para la distribución del máximo:

$$\begin{aligned} F_{X_{(n)}}(x) &= P[X_{(n)} \leq x] = P[X_1 \leq x, \dots, X_n \leq x] \stackrel{\text{indep.}}{=} \prod_{i=1}^n P[X_i \leq x] \\ &\stackrel{\text{id. d.}}{=} \prod_{i=1}^n P[X \leq x] = (F_X(x))^n \end{aligned}$$

Para la del mínimo:

$$\begin{aligned} F_{X_{(1)}}(x) &= P[X_{(1)} \leq x] = 1 - P[X_{(1)} > x] = 1 - P[X_1 > x, \dots, X_n > x] \\ &\stackrel{\text{indep.}}{=} 1 - \prod_{i=1}^n P[X_i > x] \stackrel{\text{id. d.}}{=} 1 - (P[X > x])^n = 1 - (1 - F_X(x))^n \end{aligned}$$

□

Ejercicio 1.1.9. Obtener las distribuciones muestrales de $X_{(1)}$ y $X_{(n)}$ para $X \rightsquigarrow U(a, b)$.

Si $X \rightsquigarrow U(a, b)$, entonces:

$$F_X(x) = \frac{x - a}{b - a} \quad \forall x \in [a, b]$$

Por lo que aplicando la Proposición superior:

$$F_{X_{(n)}}(x) = (F_X(x))^n = \left(\frac{x-a}{b-a}\right)^n \quad \forall x \in [a, b]$$

$$F_{X_{(1)}}(x) = 1 - (1 - F_X(x))^n = 1 - (1 - F_X(x))^n = 1 - \left(1 - \frac{x-a}{b-a}\right)^n$$

$$= 1 - \left(\frac{b-x}{b-a}\right)^n \quad \forall x \in [a, b]$$

1.2. Distribuciones en el muestreo de poblaciones normales

Proposición 1.2. Sea $X \rightsquigarrow \mathcal{N}(0, 1)$, entonces $X^2 \rightsquigarrow \chi^2(1)$.

Demostración. Sea $Y = X^2 = h(X)$, entonces $X = \pm\sqrt{Y} = h^{-1}(y)$, por lo que:

$$f_Y(y) = f_X(h_1^{-1}(y)) \left| \frac{dh_1^{-1}(y)}{dy} \right| + f_X(h_2^{-1}(y)) \left| \frac{dh_2^{-1}(y)}{dy} \right|$$

Como $X \rightsquigarrow \mathcal{N}(0, 1)$, entonces:

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad \forall x \in \mathbb{R}$$

De donde:

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(\sqrt{y})^2}{2}} \left| \frac{1}{2\sqrt{y}} \right| + \frac{1}{\sqrt{2\pi}} e^{-\frac{(-\sqrt{y})^2}{2}} \left| \frac{-1}{2\sqrt{y}} \right| = \frac{1}{\sqrt{2\pi y}} e^{-\frac{y}{2}} \quad \forall y > 0$$

Por lo que $Y \rightsquigarrow \chi^2(1)$. □

Ejercicio 1.2.1. Calcula el valor de k o la probabilidad inducida:

a) $P[\chi^2(10) \geq k] = 0,005.$

$k = 25,1881.$

b) $P[\chi^2(45) \leq k] = 0,005.$

$$P[\chi^2(45) \geq k] = 0,995 \implies k = 24,3110$$

c) $P[\chi^2(14) \geq 21,06]$

$0,1$

d) $P[\chi^2(20) \leq 12,44]$

$$P[\chi^2(20) \leq 12,44] = 1 - P[\chi^2(20) \geq 12,44] = 1 - 0,9 = 0,1$$

Ejercicio 1.2.2. Calcula el valor de k o la probabilidad inducida:

a) $P[t(26) \geq k] = 0,05$

$$k = 1,7056$$

b) $P[t(20) \leq k] = 0,25$

$$k = -0,6870$$

c) $P[t(26) \geq k] = 0,9$

$$k = -1,3150$$

d) $P[t(21) \geq 1,721]$

$$0,05$$

e) $P[t(11) \leq 0,697]$

$$0,75$$

f) $P[t(8) \leq -2,306]$

$$0,025$$

Ejercicio 1.2.3. Calcula el valor de k o la probabilidad inducida:

a) $P[F(7, 3) \leq k] = 0,95$

$$k = 8,89$$

b) $P[F(8, 4) \geq k] = 0,01$

$$0,01 = 1 - P[F(8, 4) \leq k] \implies P[F(8, 4) \leq k] = 0,99 \implies k = 14,8$$

c) $P[F(2, 2) \leq 19]$

$$0,95$$

d) $P[F(3, 5) \geq 12,1]$

$$P[F(3, 5) \geq 12,1] = 1 - P[F(3, 5) \leq 12,1] = 1 - 0,99 = 0,01$$

e) $P[F(60, 40) \leq k] = 0,05$

$$k = 0,627$$

1.2.1. Varias demostraciones

Tenemos una (X_1, \dots, X_n) m.a.s. con $X \rightsquigarrow \mathcal{N}(\mu, \sigma^2)$, si tomamos:

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

Demostraciones importantes que pueden caer.

Proposición 1.3. En dichas condiciones, veamos que:

$\bar{X}, (X_1 - \bar{X}, \dots, X_n - \bar{X})$ son independientes

Demostración. Para ellos, usaremos la caracterización por la función generatriz de momentos conjunta:

$$M_{\bar{X}, X_1 - \bar{X}, \dots, X_n - \bar{X}}(t, t_1, \dots, t_n) \stackrel{?}{=} M_{\bar{X}}(t) M_{X_1 - \bar{X}, \dots, X_n - \bar{X}}(t_1, \dots, t_n)$$

$$\begin{aligned} M_{\bar{X}, X_1 - \bar{X}, \dots, X_n - \bar{X}}(t, t_1, \dots, t_n) &= E[e^{(t, t_1, \dots, t_n) \cdot (\bar{X}, X_1 - \bar{X}, \dots, X_n - \bar{X})}] \\ &= E \left[e^{t\bar{X} + \sum_{i=1}^n (X_i - \bar{X})t_i} \right] \\ &= E \left[e^{\frac{t}{n} \sum_{i=1}^n X_i + \sum_{i=1}^n X_i t_i - \sum_{i=1}^n \bar{X} t_i} \right] \\ &= E \left[e^{\frac{t}{n} \sum_{i=1}^n X_i + \sum_{i=1}^n X_i t_i - \bar{X} \sum_{i=1}^n t_i} \right] \\ &= E \left[e^{\frac{t}{n} \sum_{i=1}^n X_i + \sum_{i=1}^n X_i t_i - \frac{1}{n} \sum_{i=1}^n X_i \sum_{i=1}^n t_i} \right] \\ &= E \left[e^{\frac{t}{n} \sum_{i=1}^n X_i + \sum_{i=1}^n X_i t_i - \sum_{i=1}^n X_i \frac{1}{n} \sum_{i=1}^n t_i} \right] \\ &= E \left[e^{\frac{t}{n} \sum_{i=1}^n X_i + \sum_{i=1}^n X_i t_i - \sum_{i=1}^n X_i \bar{t}} \right] \\ &= E \left[e^{\sum_{i=1}^n X_i \left(\frac{t}{n} + t_i - \bar{t} \right)} \right] \\ &= E \left[\prod_{i=1}^n e^{X_i \left(\frac{t}{n} + t_i - \bar{t} \right)} \right] \\ &\stackrel{\text{indep.}}{=} \prod_{i=1}^n E \left[e^{X_i \left(\frac{t}{n} + t_i - \bar{t} \right)} \right] = \prod_{i=1}^n M_{X_i} \left(\frac{t}{n} + t_i - \bar{t} \right) \\ &\stackrel{(*)}{=} \prod_{i=1}^n e^{\left(\frac{t}{n} + t_i - \bar{t} \right) \mu + \left(\frac{t}{n} + t_i - \bar{t} \right)^2 \frac{\sigma^2}{2}} \\ &= e^{\sum_{i=1}^n \left[\left(\frac{t}{n} + t_i - \bar{t} \right) \mu + \frac{\sigma^2}{2} \left(\frac{t^2}{n^2} + (t_i - \bar{t})^2 + 2 \frac{t}{n} (t_i - \bar{t}) \right) \right]} \\ &= e^{\sum_{i=1}^n \frac{\mu t}{n} + \sum_{i=1}^n \mu (t_i - \bar{t}) + \frac{\sigma^2}{2} \left(\sum_{i=1}^n \frac{t^2}{n^2} + \sum_{i=1}^n (t_i - \bar{t})^2 + 2 \sum_{i=1}^n \frac{t}{n} (t_i - \bar{t}) \right)} \\ &= e^{\cancel{\mu t} + \cancel{\mu \sum_{i=1}^n t_i} + \cancel{\mu \sum_{i=1}^n \bar{t}} + \frac{\sigma^2}{2} \left(\frac{nt^2}{n^2} + \sum_{i=1}^n (t_i - \bar{t})^2 + 2 \frac{t}{n} \sum_{i=1}^n (t_i - \bar{t}) \right)} \\ &= e^{\mu t + \frac{\sigma^2 t^2}{2n} + \frac{\sigma^2}{2} \sum_{i=1}^n (t_i - \bar{t})^2} \end{aligned}$$

Sabemos que:

$$\begin{aligned} M_{\bar{X}}(t) &= M_{(\bar{X}, X_1 - \bar{X}, \dots, X_n - \bar{X})}(t, 0, \dots, 0) = e^{\mu t + \frac{\sigma^2 t^2}{2n}} \\ M_{(X_1 - \bar{X}, \dots, X_n - \bar{X})}(t_1, \dots, t_n) &= M_{(\bar{X}, X_1 - \bar{X}, \dots, X_n - \bar{X})}(0, t_1, \dots, t_n) \\ &= e^{\frac{\sigma^2}{n} \sum_{i=1}^n (t_i - \bar{t})^2} \end{aligned}$$

Por lo que es cierto que el producto de las funciones generatrices de momentos es la generatriz de mmoentos conjunta, luego las variables son independientes. \square

Corolario 1.3.1. Como corolario de la Proposición anterior, tenemos que:

- Se vio ya, y se saca de la demostración de arriba.
- Lema de Fisher: \bar{X} y S^2 son independientes.

Como S^2 es función del vector de la Proposición anterior, tenemos que es independiente con \bar{X} , ya que las funciones de variables independientes son independientes.

- $\frac{(n-1)S^2}{\sigma^2} \rightsquigarrow \chi^2(n-1)$

Para demostrarlo:

$$\sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma} \right)^2 \rightsquigarrow \chi^2(n)$$

Ahora, queremos ver que:

$$\frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2} \rightsquigarrow \chi^2(n-1)$$

Para ello:

$$\begin{aligned} \frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2} &= \frac{\sum_{i=1}^n (X_i - \bar{X} + \bar{X} - \mu)^2}{\sigma^2} \\ &= \frac{\sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{i=1}^n (\bar{X} - \mu)^2 + 2 \sum_{i=1}^n (X_i - \bar{X})(\bar{X} - \mu)}{\sigma^2} \\ &= \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2} + n \frac{(\bar{X} - \mu)^2}{\sigma^2} \end{aligned}$$

Y como:

$$n \frac{(\bar{X} - \mu)^2}{\sigma^2} \rightsquigarrow \chi^2(1)$$

Buscamos ver lo que sigue lo de la derecha ($A = B + C$). Para ello, usaremos la función generatriz de momentos. tenemos que $B = f(S^2)$ y $C = f(\bar{X})$, luego B y C son independientes, por lo que:

$$M_{A=B+C}(t) \stackrel{indep.}{=} M_B(t) M_C(t) = M_B(t) \frac{1}{(1-2t)^{\frac{1}{2}}} \quad t < \frac{1}{2}$$

Y sabemos que:

$$M_A(t) = \frac{1}{(1-2t)^{\frac{n}{2}}}$$

De donde:

$$M_B(t) = \frac{M_A(t)}{M_C(t)} = \frac{\frac{1}{(1-2t)^{\frac{n}{2}}}}{\frac{1}{(1-2t)^{\frac{1}{2}}}} = \frac{1}{(1-2t)^{\frac{n-1}{2}}} \quad t < \frac{1}{2}$$

Por lo que $B \rightsquigarrow \chi^2(n-1)$

■ $\frac{\bar{X} - \mu}{S/\sqrt{n}} \rightsquigarrow t(n-1)$

Para ello, al igual que la χ^2 , lo más sencillo es ir a la construcción de t :

$$\left. \begin{array}{l} X \rightsquigarrow \mathcal{N}(0, 1) \\ Y \rightsquigarrow \chi^2(n) \\ \text{indep} \end{array} \right\} \frac{X}{\sqrt{Y/n}} \rightsquigarrow t(n)$$

Como:

$$\begin{aligned} \bar{X} \rightsquigarrow \mathcal{N}(\mu, \sigma^2) &\implies \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \rightsquigarrow \mathcal{N}(0, 1) \\ \frac{(n-1)S^2}{\sigma^2} &\rightsquigarrow \chi^2(n-1) \end{aligned}$$

que son independientes por el Lema de Fisher. Si aplicamos la construcción:

$$\frac{\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{(n-1)S^2}{\sigma^2(n-1)}}} = \frac{\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}}{\frac{S}{\sigma}} = \frac{\bar{X} - \mu}{S/\sqrt{n}} \rightsquigarrow t(n-1)$$

2. Relaciones de Ejercicios

2.1. Estadísticos muestrales

Ejercicio 2.1.1. Sea (X_1, \dots, X_n) una muestra aleatoria simple de una variable aleatoria X . Dar el espacio muestral y calcular la función masa de probabilidad de (X_1, \dots, X_n) en cada uno de los siguientes casos:

- a) $X \rightsquigarrow \{B(k_0, p) : p \in (0, 1)\}$ Binomial.

El espacio muestral en este caso es \mathcal{X}^n , donde:

$$\mathcal{X} = \{0, 1, \dots, k_0\}$$

Recordamos que si $X \rightsquigarrow B(k_0, p)$, entonces:

$$P[X = x] = \binom{k_0}{x} p^x (1-p)^{k_0-x} \quad \forall x \in \mathcal{X}$$

Por tanto, para nuestra m.a.s. tendremos la función masa de probabilidad:

$$\begin{aligned} P[X_1 = x_1, \dots, X_n = x_n] &\stackrel{\text{indep.}}{=} \prod_{i=1}^n P[X_i = x_i] \stackrel{\text{id. d.}}{=} \prod_{i=1}^n P[X = x_i] \\ &= \prod_{i=1}^n \binom{k_0}{x_i} p^{x_i} (1-p)^{k_0-x_i} = p^{\sum_{i=1}^n x_i} (1-p)^{nk_0 - \sum_{i=1}^n x_i} \prod_{i=1}^n \binom{k_0}{x_i} \\ &\quad \forall (x_1, \dots, x_n) \in \mathcal{X}^n \end{aligned}$$

- b) $X \rightsquigarrow \{\mathcal{P}(\lambda) : \lambda \in \mathbb{R}^+\}$ Poisson.

El espacio muestral de X es:

$$\mathcal{X} = \mathbb{N} \cup \{0\}$$

Recordamos que si $X \rightsquigarrow \mathcal{P}(\lambda)$, entonces:

$$P[X = x] = e^{-\lambda} \frac{\lambda^x}{x!} \quad \forall x \in \mathcal{X}$$

Por tanto:

$$\begin{aligned} P[X_1 = x_1, \dots, X_n = x_n] &\stackrel{\text{indep.}}{=} \prod_{i=1}^n P[X_i = x_i] \stackrel{\text{id. d.}}{=} \prod_{i=1}^n P[X = x_i] \\ &= \prod_{i=1}^n e^{-\lambda} \frac{\lambda^{x_i}}{x_i!} = e^{-n\lambda} \prod_{i=1}^n \frac{\lambda^{x_i}}{x_i!} = e^{-n\lambda} \cdot \frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} \quad \forall (x_1, \dots, x_n) \in \mathcal{X}^n \end{aligned}$$

c) $X \rightsquigarrow \{BN(k_0, p) : p \in (0, 1)\}$ Binomial Negativa.

El espacio muestral de X es:

$$\mathcal{X} = \mathbb{N} \cup \{0\}$$

Recordamos que si $X \rightsquigarrow BN(k_0, p)$, entonces:

$$P[X = x] = \binom{x + k_0 - 1}{x} (1 - p)^x p^{k_0} \quad \forall x \in \mathcal{X}$$

Por tanto:

$$\begin{aligned} P[X_1 = x_1, \dots, X_n = x_n] &\stackrel{\text{indep.}}{=} \prod_{i=1}^n P[X_i = x_i] \stackrel{\text{id. d.}}{=} \prod_{i=1}^n P[X = x_i] \\ &= \prod_{i=1}^n \binom{x_i + k_0 - 1}{x_i} (1 - p)^{x_i} p^{k_0} = p^{nk_0} (1 - p)^{\sum_{i=1}^n x_i} \prod_{i=1}^n \binom{x_i + k_0 - 1}{x_i} \\ &\quad \forall (x_1, \dots, x_n) \in \mathcal{X}^n \end{aligned}$$

d) $X \rightsquigarrow \{G(p) : p \in (0, 1)\}$ Geométrica.

El espacio muestral de X es:

$$\mathcal{X} = \mathbb{N} \cup \{0\}$$

Recordamos que $G(p) \equiv BN(1, p)$, por lo que si sustituimos en la fórmula obtenida en la Binomial Negativa $k_0 = 1$:

$$P[X_1 = x_1, \dots, X_n = x_n] = p^n (1 - p)^{\sum_{i=1}^n x_i} \quad \forall (x_1, \dots, x_n) \in \mathcal{X}^n$$

e) $X \rightsquigarrow \{P_N : N \in \mathbb{N}\}$, $P_N(X = x) = \frac{1}{N}$, $x = 1, \dots, N$.

El espacio muestral ya nos lo dan: $\mathcal{X} = \{1, \dots, N\}$. Calculemos la masa de probabilidad:

$$\begin{aligned} P[X_1 = x_1, \dots, X_n = x_n] &\stackrel{\text{indep.}}{=} \prod_{i=1}^n P[X_i = x_i] \stackrel{\text{id. d.}}{=} \prod_{i=1}^n P[X = x_i] \\ &= \prod_{i=1}^n \frac{1}{N} = \left(\frac{1}{N}\right)^n \quad \forall (x_1, \dots, x_n) \in \mathcal{X}^n \end{aligned}$$

Ejercicio 2.1.2. Sea (X_1, \dots, X_n) una muestra aleatoria simple de una variable aleatoria X . Dar el espacio muestral y calcular la función de densidad de (X_1, \dots, X_n) en cada uno de los siguientes casos:

a) $X \rightsquigarrow \{U(a, b) : a, b \in \mathbb{R}, a < b\}$ Uniforme.

El espacio muestral en este caso es \mathcal{X}^n , donde:

$$\mathcal{X} = [a, b]$$

Recordamos que si $X \rightsquigarrow U(a, b)$, entonces:

$$f_X(x) = \frac{1}{b-a} \quad \forall x \in [a, b]$$

Por lo que:

$$\begin{aligned} f_{(X_1, \dots, X_n)}(x_1, \dots, x_n) &\stackrel{\text{indep.}}{=} \prod_{i=1}^n f_{X_i}(x_i) \stackrel{\text{id. d.}}{=} \prod_{i=1}^n f_X(x_i) = \prod_{i=1}^n \frac{1}{b-a} \\ &= \left(\frac{1}{b-a} \right)^n \quad \forall (x_1, \dots, x_n) \in \mathcal{X}^n \end{aligned}$$

b) $X \rightsquigarrow \{\mathcal{N}(\mu, \sigma^2) : \mu \in \mathbb{R}, \sigma^2 \in \mathbb{R}^+\}$ Normal.

El espacio muestral de X es $\mathcal{X} = \mathbb{R}$. Recordamos que si $X \rightsquigarrow \mathcal{N}(\mu, \sigma^2)$, entonces:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \forall x \in \mathbb{R}$$

Por lo que:

$$\begin{aligned} f_{(X_1, \dots, X_n)}(x_1, \dots, x_n) &\stackrel{\text{indep.}}{=} \prod_{i=1}^n f_{X_i}(x_i) \stackrel{\text{id. d.}}{=} \prod_{i=1}^n f_X(x_i) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} \\ &= \left(\frac{1}{\sqrt{2\pi}\sigma} \right)^n \prod_{i=1}^n e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} = \left(\frac{1}{\sqrt{2\pi}\sigma} \right)^n e^{-\sum_{i=1}^n \frac{(x_i-\mu)^2}{2\sigma^2}} \\ &= \left(\frac{1}{\sqrt{2\pi}\sigma} \right)^n e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i-\mu)^2} \quad \forall (x_1, \dots, x_n) \in \mathbb{R}^n \end{aligned}$$

c) $X \rightsquigarrow \{\Gamma(p, a) : p, a \in \mathbb{R}^+\}$ Gamma.

El espacio muestral de X es $\mathcal{X} = \mathbb{R}_0^+$. Recordamos que si $X \rightsquigarrow \Gamma(p, a)$, entonces:

$$f_X(x) = \frac{a^p}{\Gamma(p)} x^{p-1} e^{-ax} \quad \forall x \in \mathbb{R}_0^+$$

Por lo que:

$$\begin{aligned} f_{(X_1, \dots, X_n)}(x_1, \dots, x_n) &\stackrel{\text{indep.}}{=} \prod_{i=1}^n f_{X_i}(x_i) \stackrel{\text{id. d.}}{=} \prod_{i=1}^n f_X(x_i) = \prod_{i=1}^n \frac{a^p}{\Gamma(p)} x_i^{p-1} e^{-ax_i} \\ &= \left(\frac{a^p}{\Gamma(p)} \right)^n \cdot e^{-a \sum_{i=1}^n x_i} \cdot \prod_{i=1}^n x_i^{p-1} \quad \forall (x_1, \dots, x_n) \in \mathcal{X}^n \end{aligned}$$

d) $X \rightsquigarrow \{\beta(p, q) : p, q \in \mathbb{R}^+\}$ Beta.

El espacio muestral de X es $\mathcal{X} = [0, 1]$. Recordamos que si $X \rightsquigarrow \beta(p, q)$, entonces:

$$f_X(x) = \frac{1}{\beta(p, q)} x^{p-1} (1-x)^{q-1} \quad \forall x \in [0, 1]$$

Donde:

$$\beta(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$$

Por tanto:

$$\begin{aligned} f_{(X_1, \dots, X_n)}(x_1, \dots, x_n) &\stackrel{\text{indep.}}{=} \prod_{i=1}^n f_{X_i}(x_i) \stackrel{\text{id. d.}}{=} \prod_{i=1}^n f_X(x_i) = \prod_{i=1}^n \frac{1}{\beta(p, q)} x_i^{p-1} (1-x_i)^{q-1} \\ &= \frac{1}{\beta(p, q)^n} \prod_{i=1}^n x_i^{p-1} (1-x_i)^{q-1} \quad \forall (x_1, \dots, x_n) \in \mathcal{X}^n \end{aligned}$$

$$\text{e) } X \rightsquigarrow \{P_\theta : \theta \in \mathbb{R}^+\}, \quad f_\theta(x) = \frac{1}{2\sqrt{x\theta}}, \quad 0 < x < \theta.$$

Se nos dice que $\mathcal{X} =]0, \theta[$. Calculamos la función de densidad conjunta:

$$\begin{aligned} f_{(X_1, \dots, X_n)}(x_1, \dots, x_n) &\stackrel{\text{indep.}}{=} \prod_{i=1}^n f_{X_i}(x_i) \stackrel{\text{id. d.}}{=} \prod_{i=1}^n f_X(x_i) = \prod_{i=1}^n \frac{1}{2\sqrt{x_i\theta}} \\ &= \frac{1}{(2\sqrt{\theta})^n} \prod_{i=1}^n \frac{1}{\sqrt{x_i}} \quad \forall (x_1, \dots, x_n) \in \mathcal{X}^n \end{aligned}$$

Ejercicio 2.1.3. Se miden los tiempos de sedimentación de una muestra de partículas flotando en un líquido. Los tiempos observados son:

11,5; 1,8; 7,3; 12,1; 1,8; 21,3; 7,3; 15,2; 7,3; 12,1; 15,2;
7,3; 12,1; 1,8; 10,5; 15,2; 21,3; 10,5; 15,2; 11,5

- Construir la función de distribución muestral asociada a a dichas observaciones.

Si aplicamos la definición de función de distribución muestral obtenemos que esta viene dada por:

$$F_n^*(x) = \begin{cases} 0 & \text{si } x < 1,8 \\ 3/20 & \text{si } 1,8 \leq x < 7,3 \\ 7/20 & \text{si } 7,3 \leq x < 10,5 \\ 9/20 & \text{si } 10,5 \leq x < 11,5 \\ 11/20 & \text{si } 11,5 \leq x < 12,1 \\ 14/20 & \text{si } 12,1 \leq x < 15,2 \\ 18/20 & \text{si } 15,2 \leq x < 21,3 \\ 20/20 & \text{si } x \geq 21,3 \end{cases}$$

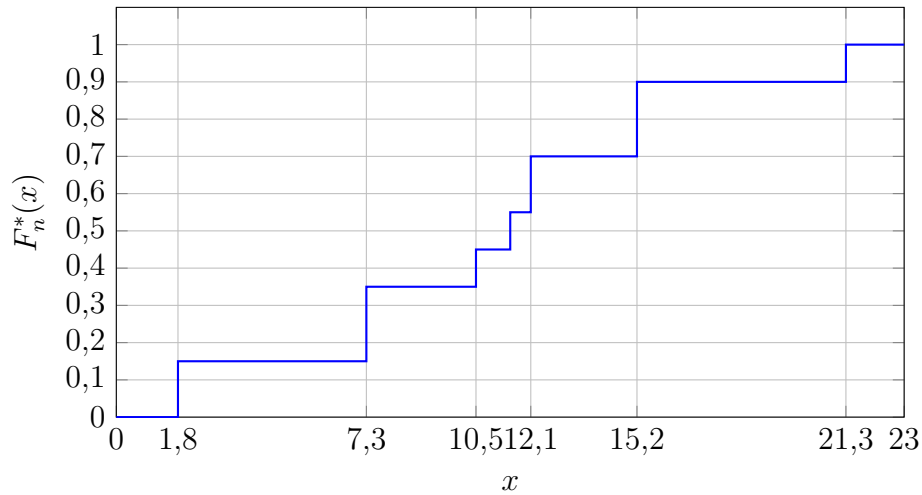


Figura 2.1: Gráfica de la función de distribución muestral.

- Hallar los valores de los tres primeros momentos muestrales respecto al origen y respecto a la media.

Calculamos primero los tres primeros momentos respecto al origen para luego calcular los centrados respecto a la media a partir de ellos:

$$a_1 = \sum_{i=1}^n f_i x_i = 10,915 \quad a_2 = \sum_{i=1}^n f_i x_i^2 = 148,9325$$

$$a_3 = \sum_{i=1}^n f_i x_i^3 = 2280,98365$$

$$b_1 = \sum_{i=1}^n f_i (x_i - \bar{x}) = 0$$

$$\begin{aligned} b_2 &= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^n (x_i^2 - 2x_i\bar{x} + \bar{x}^2) \\ &= \frac{1}{n} \sum_{i=1}^n x_i^2 - \frac{2\bar{x}}{n} \sum_{i=1}^n x_i + \bar{x}^2 = a_2 - 2a_1\bar{x} + a_1^2 = a_2 - a_1^2 \\ &= 148,9325 - 10,915 = 29,795275 \end{aligned}$$

$$\begin{aligned} b_3 &= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3 = \frac{1}{n} \sum_{i=1}^n (x_i^3 - 3x_i^2\bar{x} + 3x_i\bar{x}^2 - \bar{x}^3) \\ &= \frac{1}{n} \sum_{i=1}^n x_i^3 - \frac{3\bar{x}}{n} \sum_{i=1}^n x_i^2 + \frac{3\bar{x}^2}{n} \sum_{i=1}^n x_i - \bar{x}^3 = a_3 - 3a_1a_2 + 3a_1^3 - a_1^3 \\ &= a_3 - 3a_1a_2 + 2a_1^3 = 4,95455925 \end{aligned}$$

- Determinar los valores de los cuartiles muestrales y el percentil 70.

Para ello, primero ordenamos los datos de menor a mayor y los agrupamos en

grupos de $20/4 = 5$ en 5:

1,8; 1,8; 1,8; 7,3; 7,3; 7,3; 7,3; 10,5; 10,5; 11,5; 11,5; 12,1; 12,1;
12,1; 15,2; 15,2; 15,2; 15,2; 21,3; 21,3

Como en los cambios de agrupaciones de números estos se repiten, hemos obtenido el valor de los cuartiles:

$$q_1 = 7,3 \quad q_2 = 11,5 \quad q_3 = 15,2 \quad q_4 = 21,3$$

Para el percentil 70, calculamos:

$$0,7 \cdot 20 = 14$$

Como hemos obtenido un número entero, el percentil 70 será:

$$c_{70} = \frac{X_{(14)} + X_{(15)}}{2} = \frac{12,1 + 15,2}{2} = 13,65$$

En el caso de haber obtenido un número no entero (por ejemplo, 14,2), sería $X_{(15)}$.

Ejercicio 2.1.4. Se dispone de una muestra aleatoria simple de tamaño 40 de una distribución exponencial de media 3, ¿cuál es la probabilidad de que los valores de la función de distribución muestral y la teórica, en $x = 1$, difieran menos de 0,01? Aproximadamente, ¿cuál debe ser el tamaño muestral para que dicha probabilidad sea como mínimo 0,98?

Ejercicio 2.1.5. Se dispone de una muestra aleatoria simple de tamaño 50 de una distribución de Poisson de media 2, ¿cuál es la probabilidad de que los valores de la función de distribución muestral y la teórica, en $x = 2$, difieran menos de 0,02? Aproximadamente, ¿qué tamaño muestral hay que tomar para que dicha probabilidad sea como mínimo 0,99?

Sea $n = 50$, tenemos una m.a.s. (X_1, \dots, X_n) idénticamente distribuidas a $X \rightsquigarrow \mathcal{P}(2)$.

$$F_X(2) = \sum_{k=0}^2 e^{-2} \frac{2^k}{k!} = e^{-2} \left(\frac{2^0}{0!} + \frac{2^1}{1!} + \frac{2^2}{2!} \right) = e^{-2}(1 + 2 + 2) = 0,6767$$

$$\begin{aligned} P[|F_{X_1, \dots, X_n}^*(2) - F_X(2)| < 0,02] &= P[-0,02 < F_{X_1, \dots, X_n}^*(2) - 0,6767 < 0,02] \\ &= P[32,835 < 50 \cdot F_{X_1, \dots, X_n}^*(2) < 34,835] \end{aligned}$$

Tomando $Y = 50F_{X_1, \dots, X_n}^*(2) \rightsquigarrow B(50, F_X(2))$, tenemos:

$$P[|F_{X_1, \dots, X_n}^*(2) - F_X(2)| < 0,02] = P[Y = 33] + P[Y = 34] = 0,2397$$

Donde para la última igualdad hemos tenido que consultar la tabla de la Binomial.

Ahora, calculemos el tamaño n para que la probabilidad sea como mínimo 0,99. Sabemos para n grande que:

$$Y = nF_{X_1, \dots, X_n}^*(2) \rightsquigarrow \mathcal{N}(nF_X(2), nF_X(2)(1 - F_X(2)))$$

Por lo que:

$$Z = \frac{n(F_{X_1, \dots, X_n}^*(2) - F_X(2))}{\sqrt{nF_X(2)(1 - F_X(2))}} \rightsquigarrow \mathcal{N}(0, 1)$$

$$\begin{aligned} 0,99 &\leq P[|F_{X_1, \dots, X_n}^* - F_X(2)| < 0,02] = P\left[|Z| < \frac{n \cdot 0,02}{\sqrt{nF_X(2)(1 - F_X(2))}}\right] \\ &= P[|Z| < \sqrt{n} \cdot 0,04276] = 2P[Z < \sqrt{n} \cdot 0,04276] - 1 \\ P[Z < \sqrt{n} \cdot 0,04276] &= 0,995 \implies 2,65 = \sqrt{n} \cdot 0,04276 \implies n = 3639,8 \\ n &\geq 3640 \end{aligned}$$

Ejercicio 2.1.6. Sea $X \rightsquigarrow B(1, p)$ y (X_1, X_2, X_3) una muestra aleatoria simple de X . Calcular la función masa de probabilidad de los estadísticos \bar{X} , S^2 , $\min X_i$ y $\max X_i$.

Como tenemos que:

$$\bar{X} = \frac{1}{3} \sum_{i=1}^3 X_i$$

Si definimos $Y = \sum_{i=1}^3 X_i$, por la propiedad reproductiva de la Binomial, tenemos que $Y \rightsquigarrow B(3, p)$, de donde $\bar{X} \rightsquigarrow B(1, p/3)$.

(X_1, X_2, X_3)	\bar{X}	P	S^2	$X_{(1)}$	$X_{(n)}$
(0,0,0)	0	$(1-p)^3$	0	0	0
(0,0,1)	$1/3$	$(1-p)^2 p$	$1/3$	0	1
(0,1,0)	$1/3$	$(1-p)^2 p$	$1/3$	0	1
(0,1,1)	$1/3$	$(1-p)^2 p$	$1/3$	0	1
(1,0,0)	$1/3$	$(1-p)^2 p$	$1/3$	0	1
(1,0,1)	$2/3$	$(1-p)p^2$	$1/3$	0	1
(1,1,0)	$2/3$	$(1-p)p^2$	$1/3$	0	1
(1,1,1)	1	p^3	0	1	1

Tabla 2.1: Usamos la tabla para cálculos.

Ejercicio 2.1.7. Obtener la función masa de probabilidad o función de densidad de \bar{X} en el muestreo de una variable de Bernoulli, de una Poisson y de una exponencial.

Ejercicio 2.1.8. Calcular las funciones de densidad de los estadísticos $\max X_i$ y $\min X_i$ en el muestreo de una variable X con función de densidad:

$$f_\theta(x) = e^{\theta-x}, \quad x > \theta.$$

Calculamos primero la función de distribución, para calcular con mayor comodidad las funciones de distribución de $X_{(n)}$ y X_1 .

$$F_{\theta}(x) = \int_{\theta}^x f_{\theta}(t) dt = \int_{\theta}^x e^{\theta-t} dt = [-e^{\theta-t}]_0^x = 1 - e^{\theta-x} \quad \forall x \in \mathbb{R}^+$$

Para el máximo:

$$F_{X_{(n)}}(x) = (F_X(x))^n = (1 - e^{\theta-x})^n \implies f_{X_{(n)}} = n(1 - e^{\theta-x})^{n-1} e^{\theta-x} \quad \forall x \in \mathbb{R}^+$$

Para el mínimo:

$$F_{X_{(1)}}(x) = 1 - (1 - F_X(x))^n = 1 - (1 - 1 + e^{\theta-x})^n = 1 - e^{n(\theta-x)} \quad \forall x \in \mathbb{R}^+$$

de donde:

$$f_{X_{(1)}}(x) = n e^{n(\theta-x)-1} \quad \forall x \in \mathbb{R}^+$$

Ejercicio 2.1.9. El número de pacientes que visitan diariamente una determinada consulta médica es una variable aleatoria con varianza de 16 personas. Se supone que el número de visitas de cada día es independiente de cualquier otro. Si se observa el número de visitas diarias durante 64 días, calcular aproximadamente la probabilidad de que la media muestral no difiera en más de una persona del valor medio verdadero de visitas diarias.

Ejercicio 2.1.10. Una máquina de refrescos está arreglada para que la cantidad de bebida que sirve sea una variable aleatoria con media 200 ml. y desviación típica 15 ml. Calcular de forma aproximada la probabilidad de que la cantidad media servida en una muestra aleatoria de tamaño 36 sea al menos 204 ml.