Universidad Autónoma de Aguascalientes

Departamento de Estadística

Academia de Métodos Estadísticos Avanzados A2

Formulario para los cursos:
Estadística I
Probabilidad y estadística
Estadística descriptiva y probabilidad

Author:
Dr. José Antonio Guerrero Díaz de León

2018

1 Estadística descriptiva

Dada una muestra $\{x_1, x_2, \ldots, x_n\}$, se definen

Frecuencias

absolutas	n_i
relativas	$f_i = \dfrac{n_i}{n}$
porcentajes	$p_i = f_i \times 100$
absolutas acumuladas	$N_i = \sum_{k \le i} n_k$
relativas acumuladas	$F_i = \dfrac{N_i}{n}$

Media $\bar{x} = \dfrac{\sum_{i=1}^{n} x_i}{n} = \dfrac{\sum_{k=1}^{r} x_k n_k}{n}$

Mediana $\tilde{x} = \begin{cases} x_{\left(\frac{n+1}{2}\right)} & n \text{ impar} \\ \dfrac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}}{2} & n \text{ par} \end{cases}$

Varianza

$$S_n^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n} = \frac{\sum_{k=1}^{r}(x_k - \bar{x})^2 n_k}{n}$$

$$S_{n-1}^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{\sum_{k=1}^{r}(x_k - \bar{x})^2 n_k}{n-1}$$

Rango $R = x_{(n)} - x_{(1)}$

Cuantiles $C_p,\ 0 \le p \le 1$

$C_p = (1 - \triangle r)x_{(\boxed{r})} + \triangle r\, x_{(\boxed{r}+1)}$

donde $r = p(n-1) + 1$,

$\triangle r$=parte decimal, $\boxed{r}$=parte entera

- Cuartiles $Q_k = C_{\frac{k}{4}}$, $k = 1, 2, 3$
- Deciles $D_k = C_{\frac{k}{10}}$, $k = 1, 2, \ldots, 9$
- Percentiles $P_k = C_{\frac{k}{100}}$, $k = 1, 2, \ldots, 99$

Rango intercuartilico $RIQ = Q_3 - Q_1$

Coeficiente de simetría $\alpha_3 = \dfrac{M_3}{S_n^3}$

con $M_3 = \dfrac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3$

Coeficiente de curtosis $\alpha_4 = \dfrac{M_4}{S_n^4}$

con $M_4 = \dfrac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4$

Diagrama de caja y brazos

$CI_i =, Q_1 - 1.5 \times RIQ$

$CI_s = Q_3 + 1.5 \times RIQ$

2 Teoría de eventos

Conmutatividad	**Distributividad**
$A \cup B = B \cup A$	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
$A \cap B = B \cap A$	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
Leyes de Morgan	**Asociatividad**
$(A \cup B)^c = A^c \cap B^c$	$(A \cup B) \cup C = A \cup (B \cup C)$
$(A \cap B)^c = A^c \cup B^c$	$(A \cap B) \cap C = A \cap (B \cap C)$
Complementario	
$A \cap A^c = \emptyset$	$\emptyset^c = \Omega$
$A \cup A^c = \Omega$	$(A^c)^c = A$
$\Omega^c = \emptyset$	$A - B = A \cap B^c$

3 Probabilidad

Axiomas de Kolmogorov

- $P(A) \ge 0$
- $P(\Omega) = 1$
- Si A y B son excluyentes

  $P(A \cup B) = P(A) + P(B)$

Probabilidad condicional

$P(A|B) = \dfrac{P(A \cap B)}{P(B)}$, si $P(B) > 0$

Regla del producto

$P(A \cap B) = P(A|B)P(B)$

Leyes de suma

- $P(A^c) = 1 - P(A)$
- $P(A - B) = P(A) - P(A \cap B)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Probabilidad total

- $P(B) = P(B|A)P(A) + P(B|A^c)P(A^c)$
- $P(B) = \sum_{i=1}^{k} P(B|A_i)P(A_i)$ donde $\{A_1, A_2, \ldots, A_k\}$ forman una partición de Ω

Regla de Bayes

- $P(A|B) = \dfrac{P(B|A)P(A)}{P(B)}$
- $P(A|B) = \dfrac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}$
- $P(A_i|B) = \dfrac{P(B|A_i)P(A_i)}{\sum_{j=1}^{k} P(B|A_j)P(A_j)}$

Independencia

A y B son independientes sí y sólo si

- $P(A|B) = P(A)$
- $P(B|A) = P(B)$
- $P(A \cap B) = P(A)P(B)$

4 Distribuciones discretas

En las propiedades, $X \sim f = f_X$ y $Y \sim f_y$ y $c \in \mathbb{R}$

Función de distribución

$f(x) \geq 0$

$\sum_x f(x) = 1$

Si $X \sim f$, $P(X = x) = f(x)$

Función de distribución acumulada

$$F(x) = P(X \leq x) = \sum_{y \leq x} f(y)$$

- $0 \leq F(x) \leq 1$
- Si $x \leq y$ entonces $F(x) \leq F(y)$

Valor esperado

$$E[X] = \sum_x x f(x)$$

- $E[c] = c$
- $E[cX] = cE[X]$
- $E[X + c] = E[X] + c$
- $E[X + Y] = E[X] + E[Y]$

$$E[g(X)] = \sum_x g(x) f(x)$$

Varianza

$$\begin{aligned} V[X] &= E[(X - E[X])^2] \\ &= \sum_x (x - E[X])^2 f(x) \end{aligned}$$

- $V[c] = 0$
- $V[cX] = c^2 V[X]$
- $V[X + c] = V[X]$
- Si X y Y son independientes, $V[X+Y] = V[X]+V[Y]$

$$SD[X] = \sqrt{V[X]}$$

Momentos

$$m_r = E[X^r] = \sum_x x^r f(x)$$

- $m_1 = E[X]$
- $m_2 = V[X] + E[X]^2$
- $V[X] = m_2 - m_1^2$

Función generadora de momentos

$$M_X(t) = E[e^{tX}] = \sum_x e^{tx} f(x)$$

- $m_k = \left. \dfrac{\mathrm{d}^k M(t)}{\mathrm{d}t^k} \right|_{t=0}$
- Si $Y = cX$, $M_Y(t) = M_X(ct)$
- Si $Y = X + c$, $M_Y(t) = e^{ct} M_X(t)$
- Si $Y = \sum_{i=1}^{n} X_i$,
  $M_Y(t) = M_{X_1}(t) \cdots M_{X_n}(t)$
- Si X y Y son dos variables tales que $M_X(t) = M_Y(t)$, entonces $f_X(x) = f_Y(x)$

5 Distribuciones continuas

En las propiedades, $X \sim f = f_X$, $Y \sim f_y$ y $c \in \mathbb{R}$

Función de densidad

$f(x) \geq 0$
$\int f(x)\mathrm{d}x = 1$
$P(a < X < b) = \int_a^b f(x)\mathrm{d}x$

Función de distribución acumulada

$F(x) = P(X \leq x) = \int_{-\infty}^x f(y)\mathrm{d}y$

- $0 \leq F(x) \leq 1$
- Si $x \leq y$ entonces $F(x) \leq F(y)$
- $F'(x) = f(x)$

Valor esperado

$E[X] = \int x f(x)\mathrm{d}x$

- $E[c] = c$
- $E[cX] = cE[X]$
- $E[X + c] = E[X] + c$
- $E[X + Y] = E[X] + E[Y]$

$E[g(X)] = \int g(x) f(x)\mathrm{d}x$

Varianza

$$\begin{aligned} V[X] &= E[(X - E[X])^2] \\ &= \int (x - E[X])^2 f(x)\mathrm{d}x \end{aligned}$$

- $V[c] = 0$
- $V[cX] = c^2 V[X]$
- $V[X + c] = V[X]$
- Si X y Y son independientes, $V[X+Y] = V[X]+V[Y]$

$SD[X] = \sqrt{V[X]}$

Covarianza

$Cov[X, Y] = E[(X - E[X])(Y - E[Y])]$

- $Cov[X, Y] = E[XY] - E[X]E[Y]$
- $V[X + Y] = V[X] + V[Y] + 2Cov[X, Y]$

Función gamma

- $\Gamma(x) = (x - 1)\Gamma(x - 1)$
- $\Gamma(n) = (n - 1)!$ si $n \in \mathbb{N}$
- $\Gamma(\frac{1}{2}) = \sqrt{\pi}, \quad \Gamma(\frac{1}{3}) = 2.6789, \quad \Gamma(\frac{1}{4}) = 3.6256$

Estandarización

- $Z = \dfrac{X - \mu}{\sigma} \sim N(0, 1),$

donde $\mu = E[X]$ y $\sigma = \sqrt{V[X]}$

Momentos

$m_k = E[X^k] = \int x^k f(x)\mathrm{d}x$

- $m_1 = E[X]$
- $m_2 = V[X] + E[X]^2$
- $V[X] = m_2 - m_1^2$

Función generadora de momentos

$M_X(t) = E[e^{tX}] = \int e^{tx} f(x)\mathrm{d}x$

- $m_k = \left.\dfrac{\mathrm{d}^k M(t)}{\mathrm{d}t^k}\right|_{t=0}$
- Si $Y = cX$, $M_Y(t) = M_X(ct)$
- Si $Y = X + c$, $M_Y(t) = e^{ct} M_X(t)$
- Si $Y = \sum_{i=1}^n X_i$,
  $M_Y(t) = M_{X_1}(t) \cdots M_{X_n}(t)$
- Si $M_X(t) = M_Y(t)$, entonces $f_X(x) = f_Y(x)$

6 Distribucions conjuntas

Propiedades

- $f_{X,Y}(x, y) \geq 0$
- $\int\int f_{X,Y}(x, y)\mathrm{d}y\mathrm{d}x = 1$

Valor esperado

- $E[g(X, Y)] = \int\int g(X, Y) f_{X,Y}(x, y)\mathrm{d}y\mathrm{d}x$

Acumulada

$F(a, b) = \int_{-\infty}^a \int_{-\infty}^b f(x, y)\mathrm{d}y\mathrm{d}x$

Marginales

- $f_X(x) = \int f_{X,Y}(x, y)\mathrm{d}y$
- $f_Y(y) = \int f_{X,Y}(x, y)\mathrm{d}x$

Condicionales

- $f_{X|Y=y}(x) = \dfrac{f_{X,Y}(x, y)}{f_Y(y)}$
- $f_{Y|X=x}(y) = \dfrac{f_{X,Y}(x, y)}{f_X(x)}$

Nombre de la distribución	Función de distribución de probabilidad	Dominio	Parámetros	Media $E[X]$	Varianza $V[X]$	Generadora de momentos $M_X(t)$
Uniforme	$P(X=x)=\frac{1}{N}$	$x \in \{x_1, x_2, \ldots, x_N\}$		$\bar{x}$	S_N^2	$\frac{1}{N}\sum_{i=1}^{n} e^{tx_i}$
	$P(X=x)=\frac{1}{N}$ $N=b-a+1$	$x=a,\ldots,b$	a : Límite inferior b : Límite superior	$\frac{a+b}{2}$	$\frac{N^2-1}{12}$	$\frac{e^{at}-e^{(b+1)t}}{N(1-e^t)}$
Bernoulli	$P(X=x)=p^x(1-p)^{1-x}$	$x=0,1$	p : Probabilidad de éxito	p	$p(1-p)$	$pe^t+(1-p)$
Binomial	$P(X=x)=\binom{N}{x}p^x(1-p)^{N-x}$	$x=0,\ldots,N$	p : Probabilidad de éxito N : Número de ensayos	Np	$Np(1-p)$	$(pe^t+(1-p))^N$
Geométrica	$P(X=x)=p(1-p)^{x-1}$	$x=1,2,\ldots$	p : Probabilidad de éxito	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{pe^t}{1-(1-p)e^t}$
	$P(Y=y)=p(1-p)^y$ $Y=X-1$	$y=0,1,\ldots$	p : Probabilidad de éxito	$\frac{1-p}{p}$	$\frac{1-p}{p^2}$	$\frac{p}{1-(1-p)e^t}$
Binomial negativa	$P(X=x)=\binom{x-1}{r-1}p^r(1-p)^{x-r}$	$x=r,r+1,\ldots$	p : Probabilidad de éxito r : Número de éxitos deseados	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$	$\left(\frac{pe^t}{1-(1-p)e^t}\right)^r$
Poisson	$P(X=x)=\frac{e^{-\lambda}\lambda^x}{x!}$	$x=0,1,\ldots$	λ : tasa	λ	λ	$e^{\lambda(e^t-1)}$
Hipergeométrica	$P(X=x)=\frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$	$x=A,\ldots,B$ $A=\max(0,n+K-N)$ $B=\min(n,K)$	N : Tamaño de la población n : Tamaño de la muestra K : Tamaño de subconjunto de interés	$n\frac{K}{N}$	$\frac{nK(N-K)(N-n)}{N^2(N-1)}$	No es práctica

Distribuciones continuas

Departamento de Estadística, UAA

Nombre	**Función de densidad**	**Parámetros**	**Media** $E[X]$	**Varianza** $V[X]$	**Función eneradora de momentos** $M_X(t)$
Uniforme	$f_X(x) = \frac{1}{b-a}$ $a \leq x \leq b$	a: localización (límite inferior) b: localización (límite superior) $\infty < a, b < \infty$ $a < b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{bt}-e^{at}}{(b-a)t}$
Exponencial	$f_X(x) = \lambda e^{-\lambda x}$ $x > 0$	λ: tasa $\lambda > 0$ $\lambda = 1/\theta$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda - t}$ $t < \lambda$
	$f_X(x) = \frac{1}{\theta}e^{-\frac{x}{\theta}}$ $x > 0$	θ: escala $\theta > 0$	θ	θ^2	$\frac{1}{1-t\theta}$ $t < \frac{1}{\theta}$
Normal	$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ $-\infty < x < \infty$	μ: localización (central) $-\infty < \mu < \infty$ σ: escala$\sigma > 0$	μ	σ^2	$e^{\mu t + \frac{t^2\sigma^2}{2}}$
Gamma	$f_X(x) = \frac{\lambda^\beta}{\Gamma(\beta)}x^{\beta-1}e^{-\lambda x}$ $x > 0$	λ: tasa β: forma $\beta, \lambda > 0$	$\frac{\beta}{\lambda}$	$\frac{\beta}{\lambda^2}$	$\left(\frac{\lambda}{\lambda - t}\right)^\beta$ $t < \lambda$
	$f_X(x) = \frac{1}{\theta^\beta\Gamma(\beta)}x^{\beta-1}e^{-\frac{x}{\theta}}$ $x > 0$	θ: escala β: forma $\beta, \theta > 0$	$\beta\theta$	$\beta\theta^2$	$\left(\frac{1}{1-t\theta}\right)^\beta$ $t < \frac{1}{\theta}$
Weibull	$f_X(x) = \beta\alpha^\beta x^{\beta-1}e^{-(\alpha x)^\beta}$ $x > 0$	α: tasa β: forma $\alpha, \beta > 0$ $\alpha = \frac{1}{\theta}$	$\frac{1}{\alpha}\Gamma\left(1+\frac{1}{\beta}\right)$	$\frac{1}{\alpha^2}\left[\Gamma\left(1+\frac{2}{\beta}\right) - \Gamma\left(1+\frac{1}{\beta}\right)^2\right]$	$\sum_{k=0}^{\infty}\frac{t^k}{k!\alpha^k}\Gamma\left(1+\frac{1}{\beta}\right)$ $\beta \geq 1$
	$f_X(x) = \frac{\beta}{\theta}\left(\frac{x}{\theta}\right)^{\beta-1}e^{-\left(\frac{x}{\theta}\right)^\beta}$ $x > 0$	θ; escala β: forma $\theta, \beta > 0$	$\theta\Gamma\left(1+\frac{1}{\beta}\right)$	$\theta^2\left[\Gamma\left(1+\frac{2}{\beta}\right) - \Gamma\left(1+\frac{1}{\beta}\right)^2\right]$	$\sum_{k=0}^{\infty}\frac{t^k\theta^k}{k!}\Gamma\left(1+\frac{1}{\beta}\right)$ $\beta \geq 1$