DATOS NO AGRUPADOS

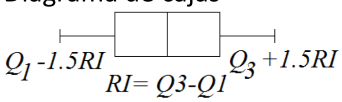
Media $\bar{X} = \sum_{i=1}^{n} \frac{x_i}{n}$

Varianza $S^2 = \sum_{i=1}^{n} \frac{(x_i-\bar{x})^2}{n-1}$

Cuantiles i.a=(n+1)(%C)

$x_{i.a} = x_i + 0.a(x_{i+1} - x_i)$

Diagrama de cajas



DATOS AGRUPADOS

Media $\bar{X} = \sum_{i=1}^{k} \frac{f_i m_i}{n}$; k:#clases fi= frecuencia

Varianza $S^2 = \sum_{i=1}^{k} \frac{f_i(x_i-\bar{x})^2}{n-1}$

Cuantiles = $L_i + \frac{j\left(\frac{n}{part}\right)-F_{i-1}}{f_i}(A)$

Se la analiza donde se encuentre la mayor frecuencia f.

Li=Limite inferior del intervalo A=Amplitud

Moda = $L_i + \frac{\Delta a}{\Delta S+\Delta a}$(A) ; ΔS= $f_i - f_{i+1}$; Δa= $f_i - f_{i-1}$

PROBABILIDADES

P(E)=$\frac{\#casos\ favorables}{\#\ total\ de\ casos}$;

$0 \leq P(E) \leq 1$

P(Ω)=1 $\sum_{x \epsilon s} f(x)=1$

Probabilidad condicional

P(A/B)=$\frac{P(A\cap B)}{P(B)}=\frac{N(A\cap B)}{N(B)}$

Probabilidad total

P(A)=$\sum_{i=1}^{n} P(A_i)P(B/A_i)$

Teorema de Bayes

P(E_r|A)=$\frac{P(A|E_r)P(E_r)}{P(A)}$

COVARIANZA

$S_{xy}=\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$

Matriz.Varian.Cova.

$[S_{xy}]=\begin{bmatrix} S_x^2 & S_{xy} \\ S_{yx} & S_y^2 \end{bmatrix}$;

Sxy=Syx

Coeficiente de correlación lineal

$\sigma_{xy}=\frac{S_{xy}}{S_x S_y}$; $-1\leq \sigma_{xy} \leq 1$

Matriz.Coef.Corre.Lin.

$[R_{xy}]=\begin{bmatrix} 1 & \sigma_{xy} \\ \sigma_{yx} & 1 \end{bmatrix}$

Variables Aleatorias Discretas

VALOR ESPERADO

E(g(x))=$\sum_{x\epsilon s} g(x)f(x)$

E(a)=a E(ag(x))=aE(g(x))

Siendo a una constante

Función generadora de momentos (depende solo de t)

M_x(t)=E(e^{tx})= $\sum_{x\epsilon s} e^{tx}f(x)$

Media $\boldsymbol{\mu}$=E(x)=$\sum_{x\epsilon z} xf(x)$

M'$_x$(t=0)=E(x)

Varianza σ²=V(x)=E[(x-$\boldsymbol{\mu}$)²]

σ²=E(x²)-[E(x)]² , V(a)=0

V(ax)=a²V(x)

M''$_x$(t=0)=E(x²)

M_x^n(t=0)=E(x^n)

$\sum_{x\epsilon s} f(x)$=1

VARIABLES ALEATORIAS DISCRETAS

Uniforme x~U(1,N)

f(x)= $\frac{1}{N}$;

Sx={1,2,…N}

$\boldsymbol{\mu}$=E(x)=$\frac{N+1}{2}$

σ²=$\frac{N^2-1}{12}$

Misma probabilidad para cada elemento de Sx

Binomial x~b(x;n,p)

f(x)=$\binom{n}{x} p^x(1-p)^{n-x}$

Sx={0,1,2,3….n}

$\boldsymbol{\mu}$=np σ²=np(1-p)

Mx(t)=[e^tp+(1-p)]^n

Se fija n

X:#de sucesos ocurrido

Geométrica x~g(x;1,p)

f(x)=p$(1-p)^{x-1}$,

Sx={1,2,3,…n}

$\boldsymbol{\mu}$=$\frac{1}{p}$ σ²= $\frac{1-p}{p^2}$

Mx(t)=e^tp$\left[\frac{1}{1-(1-p)e^t}\right]$

X:#de repeticiones hasta que ocurra el 1er éxito

Binomial Negativa x~bn(x;r,p)

f(x)=$\binom{x-1}{r-1} p^r(1-p)^{x-r}$, Sx={r, r+1, r+2,…}

$\boldsymbol{\mu}$=$\frac{r}{p}$ σ²= $\frac{r(1-p)}{p^2}$ Mx(t)=$(e^t p)^r\left[\frac{1}{1-(1-p)e^t}\right]^r$

Hasta que ocurra el r-ésimo éxito.

X:#de repeticiones , para que el r-ésimo suceso ocurra

Hipergeométrica x~h(x;a,N,n)

f(x)=$\frac{\binom{N-a}{n-x}\binom{a}{x}}{\binom{N}{n}}$, Sx={0,1,2…k} , k=min{n;a}

$\boldsymbol{\mu}$=$\frac{an}{N}$ σ²= $\frac{an(N-a)(N-n)}{N^2(N-1)}$ N: # eleme. Pob. Obje

n:# eleme. Muestra a:# elem. Carac. Interés

x: # elem.q cumplen con las características

Poisson x~p(x;λ)

f(x)=$\frac{e^{-\lambda}\lambda^x}{x!}$,

Sx={0,1,2…}

$\boldsymbol{\mu}$=σ²=λ

Mx(t)=$e^{\lambda(e^t-1)}$

Conteo en un tiempo

Bernoulli x~ber(p)

Éxito= p

Fracaso= 1-p

$\boldsymbol{\mu}$=p

σ²=p(1-p)

X: 1 , 0

Multinomial

Sx={1,2…}

f($x_1, x_2 .. x_k$)=$\frac{n!}{x_1!..x_k!}p_1^{x_1}..p_k^{x_k}$

x_i =0,1,2…n

$\sum_{i=1}^{k} x_i$=n

VARIABLES ALEATORIAS CONTINUAS

Uniforme x~U(α,β)

f(x)=$\frac{1}{\beta-\alpha}$, Sx={xϵR/α<x<β} $\boldsymbol{\mu}$=$\frac{\alpha+\beta}{2}$

σ²=$\frac{(\beta-\alpha)^2}{12}$ Mx(t)=$\frac{(e^{t\beta}-e^{t\alpha})}{t(\beta-\alpha)}$

Normal x~N(μ,σ²)

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

Mx(t)=$e^{\mu t+\frac{t^2\sigma^2}{2}}$

Normal Estándar x~N(0,1)

f(x)=$\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$, Sx={xϵR}

Mx(t)=$e^{\frac{t^2}{2}}$

Estandarización

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

P_3→P(x≤P_3)=0.03

P(x<C)=%C

Gamma x~G(α,β)

f(x)=$\frac{x^{\alpha-1}e^{-x/\beta}}{\Gamma(\alpha)\beta^\alpha}$, Sx={xεR/x>0}

Γ(α)=(α-1)! μ=αβ σ²=αβ²

Mx(t)=$\frac{1}{(1-\beta t)^2}$, t<$\frac{1}{\beta}$

$$\begin{cases} \alpha = 1, \beta = \beta \ \rightarrow Exponencial \\ \alpha = \frac{n}{2}, \beta = 2 \rightarrow Ji-cuadrado \\ \alpha = n, \beta = \beta \rightarrow Erlang \end{cases}$$

Beta x~B(α,β)

f(x)=$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$

Sx={xϵR/x>0} , $\boldsymbol{\mu}$=$\frac{\alpha}{\alpha+\beta}$

σ²=$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

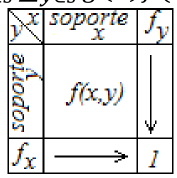
Weiball x~W(α,β)

f(x)=$\frac{\alpha x^{\alpha-1}e^{-\left(\frac{x}{\beta}\right)^\alpha}}{\beta^\alpha}$,x>0

$\boldsymbol{\mu}$=βΓ(1+1/α)

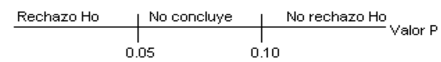
σ²=β²{Γ(1+2/α)-[Γ(1+1/α)]²}

VARIABLES ALEATORIAS CONJUNTAS

2 V.A.Discretas	Marginales	Valores Esperados	Cov(x,y)=E(xy)-E(x)E(y)
$f(x,y)=P(X=x,Y=y)=f_{xy}$ $0\leq f(x,y)\leq 1$ $\sum_{x\in s}\sum_{y\in s} f(x,y)=1$ $E(g(x))=\sum_{x\in s} g(x) f_x$ $f(x/y)=\frac{f(x,y)}{f_y}$,condicio.	$f_x=\sum_{y\in s} f(x,y)$ $f_y=\sum_{x\in s} f(x,y)$ 3VariablesDiscretas $f_z=\sum_{x\in s}\sum_{y\in s} f(x,y,z)$	E(g(x))= $\sum_{x\in s}\sum_{y\in s} g(x) f(x,y)$  	Cov(x,a)=0 Cov(x,x)=Var(x) Cov(x,y)=Cov(y,x) Cov(ax,by)=abCov(x,y) Cov(x+a,y+b)= cov(x,y) Cov(ax,y)=Cov(x,ay)=aCov(x,y) Cov(ax+by)=a^2var(x)+b^2var(y)-2abcov(x,y) Cov(ax+by,cz)=acCov(x,z)+bcCov(y,z)

Error Tipo I (**α**) : P(Rechazar Ho ; Ho es verdadera)
Error Tipo II (**β**) : P(No rechazar Ho ; H1 es verdadero)
Potencia de la prueba (1-β)= 1 – P(β)= P(Rechazar Ho ; H1 es verdadero)
Nivel de confianza: 1-α= P(No rechazar Ho ; H1 es verdadero)

VALOR P= P(Dist.Muestral <,> Estadist.de prueba)

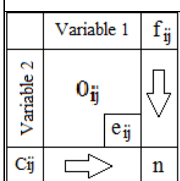


PRUEBAS PAREADAS

$D_i=x_1-x_2$ | $\bar{D}=\sum_{i=1}^{n}\frac{D_i}{n}$ | $S_D^2=\sum_{i=1}^{n}\frac{(D_i-\bar{D})^2}{n-1}$ |E.P: ZóT=$\frac{\bar{D}-D}{S_D/\sqrt{n}} \sim t_{(\alpha/2,n-1)}$

Estimador insesgado: $E(\hat{\Theta})=\Theta$
Estimador eficiente: $Var(\Theta_1)<Var(\Theta_2)$
Estimador consistente: $\lim_{n\to\infty} Var(\hat{\theta})=0$
Sesgo de un estimador: $B=E(\hat{\Theta})-\Theta$

TABLAS DE CONTINGENCIA

Ho: variable 1 es independiente de la variable 2 Vs H1: ¬Ho (negación de Ho)	Estadístico de Prueba	Región de Rechazo
	$\chi^2=\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\left[\frac{(0_{ij}-e_{ij})^2}{e_{ij}}\right]$ v=(f-1)(c-1) grados de libertad	$\chi^2>\chi^2_{(\alpha,v)}$, v=(f-1)(c-1) g.l Si no hay α, utilizamos el valor p.
 0_{ij}:Observaciones dadas $e_{ij}=(P_{ij}n)=\frac{f_i c_j}{n}$, $e_{ij}\geq 5$ f_i=número de filas c_j=número de columnas	Si $e_{ij}<5$ puedo unir a mi criterio las filas o columnas que desee. e_{ij}: Valor que se espera en la celda	Determinar si los datos de nuestra muestra son independientes.

BONDAD DE AJUSTE

X:variable aleatoria poblacional f(x):Densidad o Distribución especificada o supuesta para x Ho: f(x)=fo(x) vs H1: ¬Ho (negación de Ho)	**Por Ji-Cuadrado (Discretas)** $\chi^2=\sum_{i=1}^{n}\frac{(0_i-e_i)^2}{e_i}$, v=(f-1) g.lib. $e_i=P_i*n$	*Cuando hay datos agrupados Determina si los datos de la muestra dada se ajustan a la distribución especificada o propuesta
●XϵSx, misma probabilidad (uniforme discreta) ●Conteo en un tiempo o espacio (Poisson)	●Datos pertenecientes al tiempo de espera (Exponencial)	●Datos pert. Peso, edad (normal) ●Tiempo de vida útil (weiball)
Por Kolmogorov – Smirnov (Continuas) *Datos no agrup.		
Sn(x): Distribución empírica (1/n para cada elemento)-Utiliza la muestra **Fo(x)=**P(x≤xo) Acumulada, utiliza la población	Estadístico de Prueba D=max|Sn(x)-Fo(x)|	Región de rechazo $D>D_n$, n:tamaño de la muestra. Si no hay α utilizamos el valor p.

REGRESIÓN LINEAL SIMPLE

Modelo de R.L.S Estimado

$\hat{y}_i=\hat{\beta}_o+\hat{\beta}_1 x_i$

Modelo matricial de β estimados

$$\begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix}\begin{bmatrix} \hat{\beta}_o \\ \hat{\beta}_1 \end{bmatrix}=\begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}$$

Tabla ANOVA

Fuentes de variación	Grados de libertad	Suma de cuadrados	Cuadrados medios	Estadístico de prueba
REGRESIÓN	1	SCR	SCR/1	F*=$\frac{(SCR/1)}{(SCE/n-2)}$
ERROR	n-2	SCE	S^2=SCE/(n-2)	
TOTAL	n-1	SCT		

●SCE=$\sum_{i=1}^{n}(y_i-\hat{y}_i)^2$
●SCR=$\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2$
●SCT=SCR+SCE
●Error estimado→ $e_i=y_i-\hat{y}_i$
●Potencia de la prueba r^2
$r^2=\frac{SCR}{SCT}$, $0\leq r^2\leq 1$

	Estadístico de prueba	Región de Rechazo	Intervalo de Confianza
H_0: β_o=0 Vs 1) H_1: $\beta_o \neq b_0$ 2) H_1: $\beta_o < b_0$ 3) H_1: $\beta_o > b_0$	σ^2 desconocida→estimador S^2 $T=\frac{\widehat{\beta}_o-\beta_o}{\sqrt{S^2_{\widehat{\beta}_o}}}$, v=(n-2)g.li. , $S^2_{\widehat{\beta}_o}=S^2\left[\frac{\sum_{i=1}^{n} x_i^2}{nS_{xx}}\right]$	1) $t < -t_{\alpha/2}$ V $t > t_\alpha$ 2) $t < -t_\alpha$ 3) $t > t_\alpha$	$\hat{\beta}_o - t_{\alpha/2}\sqrt{S^2_{\widehat{\beta}_o}} < \beta_o < \hat{\beta}_o + t_{\alpha/2}\sqrt{S^2_{\widehat{\beta}_o}}$
H_0: β_1=0 Vs 1) H_1: $\beta_1 \neq 0$ 2) H_1: $\beta_1 < 0$ 3) H_1: $\beta_1 > 0$	σ^2 desconocida→estimador S^2 $T=\frac{\widehat{\beta}_1-\beta_1}{\sqrt{S^2_{\widehat{\beta}_1}}}$, v=(n-2)g.li. , $S^2_{\widehat{\beta}_1}=\frac{S^2}{S_{xx}}$	1) $t < -t_{\alpha/2}$ V $t > t_\alpha$ 2) $t < -t_\alpha$ 3) $t > t_\alpha$	$\hat{\beta}_1 - t_{\alpha/2}\sqrt{S^2_{\widehat{\beta}_1}} < \beta_1 < \hat{\beta}_1 + t_{\alpha/2}\sqrt{S^2_{\widehat{\beta}_1}}$
Ho: $\varepsilon \sim N(0,\sigma^2)$ Vs H1: ¬Ho	D_n= max|Sn(x_i)-Fo(x_i)| Kolmogorov-Smirnov		