Práctica 5 - Bidimensionales discretas

domingo, 12 de mayo de 2024 21:11

Experimentos en que cada resultado tiene 2 o más variables.

Si ambas variables son discretas, (x,y) es discreta. Lo mismo si son continuas

Cumple que:

$P(x_i, y_j) > 0$

$\sum p(x_i, y_j) = 1$

$(x_i, y_j) \in R_{xy}$

Función de probabilidad puntual conjunta (fdp conjunta)

Y/X	0	1	2	3	4	5	$F_Y(y)$
0	0	0.01	0.03	0.05	0.07	0.09	0.25
1	0.01	0.02	0.04	0.05	0.06	0.08	0.26
2	0.01	0.03	0.05	0.05	0.05	0.06	0.25
3	0.01	0.02	0.04	0.06	0.06	0.05	0.24
$F_X(x)$	0.03	0.08	0.16	0.21	0.24	0.28	1

Funciones de distribución marginales $\longrightarrow$ Se consiguen sumando columnas para la de X y filas para la de y

- $F_x(x) = P(X=x) = \sum_{y \in R_y} f(x,y)$
- $F_y(y) = P(Y=y) = \sum_{x \in R_x} f(x,y)$

Funciones de probabilidad condicionales

Ejemplo: Supongamos que deseamos conocer la probabilidad de que la linea I produzca tres articulos sabiendo que la linea II ha fabricado dos. (con lo de arriba)

$$P(x=3|Y=2) = \frac{p(x=3, y=2)}{p(y=2)} = \frac{p(3,2)}{q(2)} = \frac{0.05}{0.25}$$

División entre la función de probabilidad conjunta y la marginal del "dado"

- $E(x) = \sum_{x \in R_X} x . f_x(x)$ $\longrightarrow$ Sumar el valor de x por la marginal puntual
- $E(y) = \sum_{y \in R_Y} y . f_y(y)$ $\longrightarrow$ Idem pero con y xd
- $E(x^2) = \sum_{x \in R_X} x^2 f_x(x)$
- $V(x) = E(x^2) . E(x)^2$
- $Dt(x) = \sqrt{v(x)}$
- $E(xy) = \sum_{x \in R_x} \sum_{y \in R_y} xyf(x,y)$ $\longrightarrow$ Esperanza de una bidimensional discreta $\longrightarrow$ Valor de x. valor de y. valor de la celda+
- $Cov(x,y) = E(x.y) - E(x).E(y)$ $\longrightarrow$ Puede dar +,-,0 $\longrightarrow$ $\sigma_{x,y}$

Propiedades de operaciones con esperanza y varianza

- $E(x+y) = E(x) + E(y)$
- $E(x.y) = E(x).E(y)$ $\longrightarrow$ Solo si son INDEPENDIENTES
- $V(x+y) = V(x) + V(y) + \underbrace{2.cov(xy)}$

  Si son independientes no es necesario

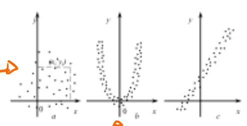
COMBINACIÓN LINEAL DE V.A NORMALES INDEPENDIENETS

- $E(aX+bY+c) = a.E(x) + b.E(y) + C$
- $V(aX+bY+c) = a^2.v(x) + b^2.v(y) + 2.a.b.cov(x,y)$ $\longrightarrow$ Varianza NO es lineal!!!!!

Coeficiente de correlación

$$P_{xy} = \frac{cov(x,y)}{\sigma_x . \sigma_y}$$

Se usa para ver si hay algún grado de relación lineal entre x e y



No hay   relación parábola   relación lineal

Como saber si dos variables son independientes

El valor de cada celda tiene que ser igual al producto de las marginales

$f(x,y)=f_x(x).f_y(y)$

Suma de variables aleatorias independientes (según distribución)

$X+Y \sim P(\lambda_1+\lambda_2)$
$X+Y \sim B(n_1+n_2, p)$ → Suma de tamaño muestral
$X+Y \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$ → Suma de esperanzas, suma de varianzas
$X-Y \sim N(\mu_1-\mu_2$, suma de las varianzas siempreeeeee $)$

Promedio de variables aleatorias normales independientes

$$\bar{X}=\frac{\sum_{i=1}^{n} X_i}{n}$$ → Cuando tengo n variables aleatorias donde $X_i \sim N(\mu,\sigma^2)$ para todo i

La media es μ y la varianza $\frac{\sigma^2}{n}$

COMBINACIÓN LINEAL DE V.A NORMALES INDEPENDIENETS

$$\bar{X} \sim N\left(\mu,\frac{\sigma^2}{n}\right) \qquad Z=\frac{\bar{X}-\mu}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0,1)$$

SIEMPRE EXACTO

PROMEDIO MUESTRAL, MEDIA MUESTRAL → $\bar{X}$

Suma de variables aleatorias normales independientes

$\sum_{i=1}^{n} X_i$ → Cuando tengo n variables aleatorias donde $X_i \sim N(\mu, \sigma^2)$ para todo i

La media es ? y la varianza

En este puedo definir una nueva variable y decir la dist. De esa

$$Z=\frac{\sum_{i=1}^{n} X_i-\mu n}{\sqrt{n\sigma^2}} \sim N(0,1)$$

$n\sim(nu,o^2)$

Teorema central del limite (TCL)

V.A INDEPENDIENTES

- Si n>=30, puedo tratarla como una normal aunque no lo sea (cuando tenemos variables aleatorias independientes con = distribución)

Aproximación de la distribución binomial a la normal

- Si np >= 10 y n(1-p) >= 10

SIEMPRE aproximado
SOLO se pone cuando saltas de p() a fi.

$$Z_n=\frac{X-n\mu}{\sqrt{n\sigma^2}}=\frac{X-n.p}{\sqrt{n.p(1-p)}} \approx N(0,1) \quad \text{si } n \text{ es lo suficientemente grande}$$

Tengo que corregir por continuidad
- Si $P(X=k)$ → $p(k-\frac{1}{2} \le X \le k+\frac{1}{2})$
- Si $P(X<=k)$ → $p(X \le k+\frac{1}{2})$
- Si $P(X>=k)$ → $p(X \ge k-\frac{1}{2})$

Ejemplos de uso

$$P(X\le 8) \approx P(X\le 8.5)=P\left(\frac{X-np}{\sqrt{np(1-p)}} \le \frac{8.5-10}{\sqrt{25\times 0.4\times 0.6}}\right) \approx \Phi(-0.61)=0.2709$$

corrección por continuidad

$$P(X=8) \approx P(7.5\le X\le 8.5)=P\left(\frac{7.5-10}{\sqrt{6}} \le \frac{X-10}{\sqrt{6}} \le \frac{8.5-10}{\sqrt{6}}\right)=P\left(-1.02 \le \frac{X-10}{\sqrt{6}} \le -0.61\right)$$

Aproximación de Poisson a distribución normal

Si λ>=30 → $\frac{X-\lambda}{\sqrt{\lambda}} \approx N(0,1)$

Y también se corrige por continuidad

Que pasa si $n<30$ y no sé distribución? NO SE PUEDE RESOLVER

Siempre que no sepa distribución/dice aproximada, tcl