

PROBABILITAT, PROCESSOS ESTOCÀSTICS I ESTADÍSTICA

Examen Parcial

21 de maig de 2015

1. Una variable bidimensional contínua (X, Y) té densitat conjunta:

$$f(x, y) = 8xy, \quad 0 < y < x < 1.$$

- (a) Calcula les densitats marginals de X i Y . Són X i Y independents?
 (b) Calcula la densitat de Y condicionada a X i l'esperança condicionada $E[Y|X=x]$.
 (c) Quina és la millor estimació possible del valor de Y si sabem que $X = \frac{1}{2}$? Compara aquest valor amb el que s'obtidria estimant Y per una constant. Sense fer més càlculs, què podem dir de la estimació lineal de Y donada X ?

Solució:

(a)

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy = \int_0^x 8xy dy = 8x \left[\frac{y^2}{2} \right]_0^x = 4x^3, \quad 0 < x < 1.$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx = \int_y^1 8xy dx = 8y \left[\frac{x^2}{2} \right]_y^1 = 4y(1 - y^2), \quad 0 < y < 1.$$

$f_X(x)f_Y(y) = 16x^3y(1 - y^2)$ és diferent de $f(x, y)$, així que X i Y no són independents.

(b)

$$f(y|x) = \frac{f(x, y)}{f_X(x)} = \frac{8xy}{4x^3} = \frac{2y}{x^2}, \quad 0 < y < x.$$

$$E[Y|X=x] = \int_{-\infty}^{\infty} y f(y|x) dy = \int_0^x y \frac{2y}{x^2} dy = \frac{2}{x^2} \left[\frac{y^3}{3} \right]_0^x = \frac{2}{3}x$$

- (c) La millor estimació possible és la no lineal $\hat{Y} = E[Y|X] = \frac{2}{3}X$. Si $X = \frac{1}{2}$ la millor estimació és $\frac{2}{3} \cdot \frac{1}{2} = \frac{1}{3}$.

Estimant $\hat{Y} = C$ constant, és $C = E[Y] = \int_0^1 y \cdot 4y(1 - y^2) dy = 4 \left[\frac{y^3}{3} - \frac{y^5}{5} \right]_0^1 = \frac{8}{15}$. Aquesta estimació no té gaire sentit quan sabem que $X = \frac{1}{2}$ ja que $\frac{8}{15} = 0,53$ mentre que en aquest cas Y varia entre 0 i $\frac{1}{2}$.

L'estimació lineal $\hat{Y} = aX + b$ és amb $a = \frac{2}{3}$ i $b = 0$ ja que la no lineal ja té aquesta forma.

2. El senyal que s'envia a un canal de comunicació, X , i el soroll afegit per aquest canal, Y , són variables aleatòries conjuntament gaussianes, amb $m_X = 10$, $m_Y = 2$, $\sigma_X = 2$, $\sigma_Y = 1$ i $\rho = \frac{1}{4}$.

- (a) Escriu la densitat de la variable que dona la sortida del canal: $Z = X + Y$.
 (b) Calcula la millor estimació lineal homogènia de X donada Z , usant el principi d'ortogonalitat.
 (c) Troba el valor de la constant γ que fa que les variables Z i $T = X + \gamma Y$ siguin incorrelades. Podem dir llavors alguna cosa sobre la independència de Z i T ?

Solució:

Calculem primer $E[X] = m_X = 10$, $E[Y] = m_Y = 2$, $E[X^2] = \sigma_X^2 + m_X^2 = 104$, $E[Y^2] = \sigma_Y^2 + m_Y^2 = 5$, $E[XY] = C[X, Y] + m_X m_Y = \rho \sigma_X \sigma_Y + m_X m_Y = 20,5$.

- (a) $m_Z = E[X+Y] = E[X] + E[Y] = 12$. $\sigma_Z^2 = V[X+Y] = V[X] + V[Y] + 2C[X, Y] = \sigma_X^2 + \sigma_Y^2 + 2\rho\sigma_X\sigma_Y = 6$.

$$f_Z(z) = \frac{1}{\sqrt{2\pi}\sigma_Z} e^{-\frac{(z-m_Z)^2}{2\sigma_Z^2}} = \frac{1}{\sqrt{12}\pi} e^{-\frac{(z-12)^2}{12}}.$$

- (b) $\hat{X} = aZ$. Pel principi d'ortogonalitat $E[aZ \cdot Z] = E[XZ]$. Llavors

$$a = \frac{E[XZ]}{E[Z^2]} = \frac{E[X^2] + E[XY]}{E[X^2] + 2E[XY] + E[Y^2]} = \frac{104 + 20,5}{104 + 41 + 5} = 0,83.$$

- (c) Imposem $0 = C[Z, T] = C[X + Y, X + \gamma Y] = V[X] + \gamma V[Y] + (1 + \gamma)C[X, Y] = 4 + \gamma + (1 + \gamma)\frac{1}{2} = \frac{9+3\gamma}{2}$. Llavors ha de ser $\gamma = -3$.

En aquest cas, com les variables són gaussianes, el fet de ser incorrelades implica que Z i T són independents

3. A una botiga, durant cert període de temps, poden entrar 0, 1 o 2 clients amb probabilitats 0,2, 0,5 i 0,3 respectivament. Cada client, independentment dels altres, pot realitzar una compra amb probabilitat 0,4. Considerem les variables $X = \text{“nombre de clients”}$, $Y = \text{“nombre de compres”}$.
- Fes una taula amb la funció de probabilitat conjunta de (X, Y) . Calcula les funcions de probabilitat marginals de X i Y .
 - Calcula l'entropia de X i l'entropia de Y . Quina de les dues variables dóna més informació sobre el resultat de l'experiment? Compara amb l'entropia de la variable conjunta (X, Y) . (Dóna els resultats en bits.)
 - Si codifiquem la variable Y assignant al valor 0 la paraula “1”, al valor 1 la paraula “00” i al valor 2 la paraula “01”, calcula el nombre mitjà de bits emprats en representar el resultat i compara amb l'entropia de Y .

Solució:

- Si $X = 0$ necessàriament $Y = 0$ i $P(X = 0, Y = 0) = 0,2$.
 Si $X = 1$ pot ser $Y = 0$ o $Y = 1$. $P(X = 1, Y = 0) = 0,5 \cdot 0,6 = 0,3$, $P(X = 1, Y = 1) = 0,5 \cdot 0,4 = 0,2$.
 Si $X = 2$ pot ser $Y = 0$, $Y = 1$ o $Y = 2$. $P(X = 2, Y = 0) = 0,3 \cdot 0,6^2 = 0,108$,
 $P(X = 2, Y = 1) = 0,3 \cdot 0,6 \cdot 0,4 \cdot 2 = 0,144$ (el 2 ve de que la compra la pot fer un o altre client), $P(X = 2, Y = 2) = 0,3 \cdot 0,4^2 = 0,048$.

	X = 0	X = 1	X = 2	
Y = 0	0,2	0,3	0,108	0,608
Y = 1	0	0,2	0,144	0,344
Y = 2	0	0	0,048	0,048
	0,2	0,5	0,3	

- A partir de la fórmula general $H = -\sum p_i \ln p_i$ i dividint per $\ln 2$ per a tenir el resultat en bits, trobem:
 $H_X = 1,485$ bits. $H_Y = 1,176$ bits. $H_{X,Y} = 2,409$ bits. X dóna més informació sobre el resultat que Y ja que les probabilitats de X són més uniformes que les de Y . Naturalment, cadascuna dóna menys informació que el parell (X, Y) .
- El nombre mitjà de bits és: $1 \cdot P(Y=0) + 2 \cdot P(Y=1) + 2 \cdot P(Y=2) = 1,392$. És més gran que H_Y ja que H_Y és el nombre òptim de bits per expressar el resultat. El nombre mitjà de bits en un codi qualsevol és major o igual que H_Y .