Ejercicio 10

Blanca Cano Camero, Daniel Krell Calvo

2 de Diciembre del 2019

Ejercicio 10

Sean X,Y variables con función de densidad

$f_{(X,Y)}(x,y) = x + y$ en $(x,y) \in [0,1] \times [0,1]$, 0 en el resto del intervalo.

Calcular

(a)Curvas de regresión de Y/X y X/Y, así como los errores cuadráticos medios asociados.

Curvas de regresión de Y/X y X/Y

La curva de regresión Y/X es $Y = E[Y/X]$ Para ello procederé a calcular las respectivas marginales:

$$f_x(x) = \int_0^1 (x+y)dy = [xy + \frac{y^2}{2}]_0^1 = x + \frac{1}{2} \qquad 0 \leq x \leq 1$$
$$f_y(y) = \int_0^1 (y+x)dx = [yx + \frac{x^2}{2}]_0^1 = y + \frac{1}{2} \qquad 0 \leq y \leq 1$$

Veamos ahora las condicionadas:

$$f_{x/y=y_0}(x) = \frac{f_{(x,y)}(x, y_0)}{f_y(y_0)} = \frac{x + y_0}{y_0 + \frac{1}{2}}$$

$$f_{y/x=x_0}(y) = \frac{f_{(x,y)}(x_0, y)}{f_x(x_0)} = \frac{x_0 + y}{x_0 + \frac{1}{2}}$$

A partir de las cuales tenemos que

$$E[Y/X] = \int_0^1 y f_{y/x=x_0}(y)\, dy = \int_0^1 y \frac{x_0 + y}{x_0 + \frac{1}{2}} dy = [\frac{x_0 y^2}{2x_0 + 1} + \frac{y^3}{3(x_0 + \frac{1}{2})}]_{y=0}^{y=1}$$

$$E[Y/X] = \frac{x_0}{2x_0+1} + \frac{1}{3(x_0+\frac{1}{2})} = \frac{3x_0+2}{6x_0+3}$$

De igual manera

$$E[X/Y] = \int_0^1 x f_{x/y=y_0}(x)\, dx = \int_0^1 x \frac{y_0+x}{y_0+\frac{1}{2}} dx = [\frac{y_0 x^2}{2y_0+1} + \frac{x^3}{3(y_0+\frac{1}{2})}]_{x=0}^{x=1}$$

$$E[X/Y] = \frac{y_0}{2y_0+1} + \frac{1}{3(y_0+\frac{1}{2})} = \frac{3y_0+2}{6y_0+3}$$

Cálculo del error cuadrático medio

El **ECM** de la curva de regresión X/Y se calcula como como $E[Var(X/Y)] = E[X^2] - E[(E[X/Y])^2]$

$$E[X^2] = \int_0^1 x^2 f_x(x)\, dx = \int_0^1 x^2(x+\frac{1}{2})\, dx = \frac{5}{12}$$

$$E[(E[X/Y])^2] = \int_0^1 E[X/Y=y]^2 f_y(y)\, dy = \int_0^1 (\frac{3y+2}{6y+3})^2 (y+\frac{1}{2})\, dy =$$

$$= \int_0^1 \frac{(3y+2)^2}{6^2(y+\frac{1}{2})}\, dy = \int_0^1 \frac{9y^2+12y+4}{36y+18}\, dy = \frac{1}{4}\int_0^1 (y+\frac{5}{6}+\frac{1}{36y+18})dy =$$

$$= \frac{1}{4}[\frac{y^2}{2}+\frac{5}{6}y+\frac{\log(36y+18)}{36}]_0^1 = \frac{1}{4}(\frac{1}{2}+\frac{5}{6}+\frac{\log(54)}{36}-\frac{\log(18)}{36}) \simeq 0,34096$$

$$E[(E[X/Y])^2] \simeq 0,3508$$

$$ECM = E[Var(X/Y)] = E[X^2] - E[(E[X/Y])^2] = 0.07571$$

Repetimos el proceso para la curva de regresión Y/X.

$$E[Y^2] = \int_0^1 y^2 f_y(y)\, dy = \int_0^1 y^2(y+\frac{1}{2})\, dy = [\frac{y^4}{4}+\frac{y^3}{6}]_0^1 = \frac{5}{12}$$

$$E[(E[Y/X])^2] = \int_0^1 E[Y/X=x]^2 f_x(x)\, dx = \int_0^1 (\frac{3x+2}{6x+3})^2 (x+\frac{1}{2})\, dx =$$

$$= \int_0^1 \frac{(3x+2)^2}{6^2(x+\frac{1}{2})}\, dx = \int_0^1 \frac{9x^2+12x+4}{36x+18}\, dx = \frac{1}{4}\int_0^1 (x+\frac{5}{6}+\frac{1}{36x+18})dx =$$

$$= \frac{1}{4}[\frac{x^2}{2} + \frac{5}{6}x + \frac{\log(36x+18)}{36}]_0^1 = \frac{1}{4}(\frac{1}{2} + \frac{5}{6} + \frac{\log(54)}{36} - \frac{\log(18)}{36}) \simeq 0,34096$$

$$E[(E[Y/X])^2] \simeq 0,3508$$

$$ECM = E[Var(Y/X)] = E[Y^2] - E[(E[Y/X])^2] = 0.07571$$

$$ECM = 0.06586$$

(b) Las razones de correlación Y/X y X/Y

Las razones de correlación se definen como

$$\eta^2_{X/Y} = \frac{Var(E[X/Y])}{Var(X)} = 1 - \frac{E[Var(X/Y)]}{Var(X)}$$

Puesto que $E[Var(X/Y)]$ ya la tenemos calculados calcularemos las respectivas varianzas:

$$Var(X) = E[X^2] - E[X]^2$$

$$E[X] = \int_0^1 xf_x(x)dx = \int_0^1 x(x+\frac{1}{2})\,dx = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}$$

Por consiguiente

$$Var(X) = E[X^2] - E[X]^2 = \frac{5}{12} - (\frac{7}{12})^2 = \frac{11}{144} \simeq 0,07639$$

Y finalmente

$$\eta^2_{X/Y} = 1 - \frac{E[Var(X/Y)]}{Var(X)} = 1 - \frac{0.07571}{0.07639} = 0.00890$$

$$\eta^2_{X/Y} = 0.1378800724864219$$

Puesto que los cálculos para $\eta^2_{Y/X}$ son iguales que para $\eta^2_{X/Y}$ podemos afirmar que:

$$\eta^2_{Y/X} = 0.1378800724864219$$

Nota: El ajuste no es muy bueno, ya que es un valore ente cero y uno, que cuanto más próximo esté a uno mejor será.

(c) Rectas de regresión $X/Y, Y/X$

La recta de regresión Y/X se calcula como:

$$Y = aX - b = E[Y] + a(x - E[x]); \text{ donde } a = \frac{Cov(X,Y)}{Var(X)}; \ b = E[Y] - aE[X];$$

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

$$E[XY] = \int_0^1 \int_0^1 xyf_{(X,Y)}(x,y)dydx = \int_0^1 \int_0^1 xy(x+y)dydx = \int_0^1 \frac{x^2}{2} + \frac{x}{3}dx = \frac{1}{3}$$

Por consiguiente

$$cov(X,Y) = \frac{-5}{6}$$

Y la recta de regresión Y/X es: $\varphi(x) = \frac{7}{12} - \frac{120}{11}(x - \frac{7}{12})$

Por la simetría de los cálculos tenemos que la recta de regresión X/Y:

$$\varphi(y) = \frac{7}{12} - \frac{120}{11}(y - \frac{7}{12})$$

(d) Coeficiente de correlación lineal {#d-coeficiente-de-correlación-lineal }

Queremos calcular

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{-120}{11}$$

¡ALGOOO HEMOS CALCULADO MAL! Sabemos que $-1 \leq \rho_{X,Y} \leq 1$ - Si r = 1, existe una correlación positiva perfecta. El índice indica una dependencia total entre las dos variables denominada relación directa. - Si 0 < r < 1, existe una correlación positiva. Si r = 0, no existe relación lineal. Pero esto no necesariamente implica que las variables son independientes: pueden existir todavía relaciones no lineales entre las dos variables. - Si -1 < r < 0, existe una correlación negativa. - Si r = -1, existe una correlación negativa perfecta. El índice indica una dependencia total entre las dos variables llamada relación inversa: cuando una de ellas aumenta, la otra disminuye en proporción constante.