Tema 4.- Esperanza condicionada

Asignatura: PROBABILIDAD

Grado en Ingeniería Informática y Matemáticas

(3er Curso - 1er semestre)

©Prof. Dr. José Luis Romero Béjar

Departamento de Estadística e Investigación Operativa
Facultad de Ciencias (Despacho 10)

Periodo de docencia: 12/09/2022 a 20/12/2022

Outline

Outline

Definición de esperanza condicionada

Sean X e Y variables aleatorias sobre el mismo espacio de probabilidad tales que existen $E[X]$ y $E[Y]$. Se consideran las distribuciones condicionadas de Y a X y de X a Y (**ambas distribuciones poseen momento no centrado de orden uno finito** porque existen para las marginales).

Se definen la **esperanza condicionada** de X a Y, denotada por $E[X/Y]$, y la **esperanza condicionada** de Y a X, denotada por $E[Y/X]$, **como las variables aleatorias** que toman el valor, en el caso **discreto** y **continuo**:

- **Caso discreto**

$$E[Y/X=x] = \sum_{y \in Supp(p_{Y/X=x})} y p_{Y/X=x}(y), \quad \forall x \in E_X,$$

cuando la v.a. X toma el valor x.

- **Caso continuo**

$$E[Y/X=x] = \int_{Supp(f_{Y/X=x})} y f_{Y/X=x}(y) dy, \quad \forall x \in E_X,$$

donde $Supp(f)$ denota el soporte de la función f, cuando la v.a. X toma el valor x.

Esperanza condicionada de una función

Del mismo modo, se define la **esperanza condicionada de una función** $g_1 : E_X \longrightarrow \mathbb{R}$, medible, de la v.a. X, condicionada a un valor concreto y de Y, denotada por $E[g_1(X)/Y = y]$, así como, para una función medible $g_2 : E_Y \longrightarrow \mathbb{R}$, la **esperanza condicionada** de $g_2(Y)$ a un valor concreto x de X, denotada por $E[g_2(Y)/X = x]$.

Más concretamente, por ejemplo, para la esperanza condicionada de $g_2(Y)$, condicionada a un valor concreto x de X, se tiene:

- **Caso discreto**

$$E[g_2(Y)/X = x] = \sum_{y \in Supp(p_{Y/X=x})} g_2(y) p_{Y/X=x}(y), \quad \forall x \in E_X.$$

- **Caso continuo**

$$E[g_2(Y)/X = x] = \int_{Supp(f_{Y/X=x})} g_2(y) f_{Y/X=x}(y) dy, \quad \forall x \in E_X.$$

donde $Supp(f)$ denota el soporte de la función f.

Propiedades de la esperanza condicionada

Se notará, por simplicidad, en lo que sigue, por $E[X/Y]$ y $E[Y/X]$ a las esperanzas condicionadas de X a Y, y de Y a X, respectivamente.

(i). Si $X \geq 0$ c.s., entonces si existe $E[X/Y]$, se tiene que $E[X/Y] \geq 0$, y $E[X/Y] = 0$, si y sólo si $P(X = 0) = 1$.

(ii). Si existe la $E[X/Y]$, entonces $|E[X/Y]| \leq E[|X|/Y]$.

(iii). **Linealidad**: $\forall i = 1, \ldots, n$, $\exists E[X_i/Y]$, entonces $\exists E[\sum_{i=1}^{n} a_i X_i + b/Y] = \sum_{i=1}^{n} a_i E[X_i/Y] + b$, $\forall (a_1, \ldots, a_n) \in \mathbb{R}^n$, $b \in \mathbb{R}$.

(iv). **Conservación del orden** Si $\exists E[X_i/Y]$, $i = 1, 2$, y $X_1 \leq X_2$, c.s., entonces $E[X_1/Y] \leq E[X_2/Y]$.

(v). Si existe $E[g(X)/Y]$, entonces $E[E[g(X)/Y]] = E[g(X)]$, para cualquier función medible g, y para cualesquiera variables aleatorias, X e Y, sobre el mismo espacio probabilístico base.

Demostración

i) Si $X > 0 \Rightarrow$ Si $\exists E[X/Y] > 0$ c.s. y $E[X/Y] = 0$ sii X es degenerada en 0

Caso discreto:

$E[X/Y=y] = \sum_{x \in E_x} x \cdot P_{X/Y}(x) > 0$ c.s. pq $X > 0$ c.s.

Es obvio que $E[X/Y=y] = 0$ sii $x = 0$ c.s. $\Leftrightarrow P[X=0] = 1$

ii) Si $\exists E[X/Y] \Rightarrow |E[X/Y]| \leq E[|X|/Y]$

Caso continuo:

$$|E[X/Y=y]| = \left| \int_{x \in Supp(f_{X/Y=y})} x \cdot f_{X/Y=y}(x)\, dx \right| \leq \int_{x \in Supp(f_{X/Y=y})} |x \cdot f_{X/Y=y}(x)| dx = \int_{x \in Supp(f_{X/Y=y})} |x| \cdot f_{X/Y}(x)\, dx = E[|X|/Y]$$

iii) Si $\forall i = 1, \dots, n$ $E[X_i/Y] \Rightarrow \exists E[\underbrace{\Sigma a_i X_i + b}_{f(x)}/Y] = \Sigma a_i E[X_i/Y] + b$ Denoto $X = (X_1, \dots, X_n)$

Caso continuo:

$$E[(\sum_{i=1}^{n} a_i X_i + b)/Y=y] = \int_{\mathbb{R}^n} (\sum_{i=1}^{n} a_i X_i + b) \cdot f_{X/Y=y}(x)\, dx =$$

$$= \sum_{i=1}^{n} a_i \int_{\mathbb{R}^n} X_i \cdot f_{X/Y=y}(x)\, dx + b \int_{\mathbb{R}^n} f_{X/Y=y}(x)\, dx = \sum_{i=1}^{n} a_i \cdot E[X_i/Y=y] + b$$

v) $E[\underbrace{E[g(X)/Y]}_{f(y)}] = E[g(X)]$

$$E[E[g(X)/Y]] \overset{\text{Caso continuo}}{=} \int_{\mathbb{R}} E[g(X)/Y=y] \cdot f_Y(y)\, dy = \int_{\mathbb{R}} \int_{\mathbb{R}} g(x) \cdot f_{X/Y=y}(x)\, dx \cdot f_Y(y)\, dy = \int g(x) \cdot f(x,y)\, dx dy = E(g(X))$$

Propiedades de la esperanza condicionada

(vi). Si X e Y son independientes, para cualquier función medible g, tal que existe $E[g(X)]$, se tiene $E[g(X)/Y] = E[g(X)]$, puesto que las distribuciones condicionadas coinciden con las marginales. → Es una línea

(vii). $E[Xg(Y)/Y] = g(Y)E[X/Y]$, para cualquier función medible g.

Las demás las hacemos nosotros

(viii). Para cualquier función h medible $E[h(X)/X] = h(X)$.

(ix). $E[X/Y] = E[c/Y] = c$, para cualquier variable degenerada X sobre el mismo espacio de probabilidad de Y, con $P(X = c) = 1$.

Todas las propiedades enunciadas anteriormente **se obtienen de forma directa, a partir de las definiciones** proporcionadas sobre esperanza condicionada.

vi) $E[g(X)/Y] = \int\limits_{x \in Supp(f_{X/Y})} g(x) \cdot f_{X/Y}(x)\, dx \overset{\text{Indep.}}{=} \int g(x) \cdot f_X(x)\, dx = E[g(X)]$

vii) $E[Xg(Y)/Y] = \int\limits_{x \in Supp(f_{X/Y=y})} x \cdot g(y) \cdot f_{X/Y}(x)\, dx = g(y) \cdot \int\limits_{x \in Supp(f_{X/Y=y})} x \cdot f_{X/Y}(x)\, dx = g(y) \cdot E[X/Y]$

viii) $E[h(X)/X] = \int\limits_{y \in Supp(f_{Y/X=x})} h(x) \cdot f_{Y/X}(y)\, dy = h(x) \cdot \underbrace{\int f_{Y/X}(y)\, dy}_{1}$

ix) $E[X/Y] = \sum\limits_{x \in Supp(P_{X/Y})} x \cdot P_{X/Y}(x) \overset{\text{Por ser degenerada, es indep.}}{=} \sum\limits_{\substack{x \in E_X \\ y \in E_Y}} x \cdot P_X(x) \cdot P_Y(y) = \sum\limits_{y \in E_Y} c \cdot P_Y(y) = c \cdot \underbrace{\sum\limits_{y \in E_Y} P_Y(y)}_{1}$

Esperanza condicionada de vectores aleatorios

Sean $\mathrm{X} = (X_1, \ldots, X_n)$ y $\mathrm{Y} = (Y_1, \ldots, Y_m)$ dos vectores aleatorios sobre el mismo espacio de probabilidad. Se define, suponiendo que **las sumas e integrales, correspondientes al caso discreto y continuo, respectivamente, son absolutamente convergentes**,

$$E[\mathrm{X}/\mathrm{Y}] = (E[X_1/\mathrm{Y}], \ldots, E[X_n/\mathrm{Y}])$$

- Caso **discreto**, es decir, para (X_i, Y) un vector aleatorio discreto:

$$E[X_i/\mathrm{Y} = \mathrm{y}] = \sum_{x_i \in Supp(p_{X_i/\mathbf{Y}=\mathbf{y}})} x_i p_{X_i/\mathrm{Y}=\mathrm{y}}(x_i), \quad \forall \mathrm{y} \in E_{\mathrm{Y}}, \quad i = 1, \ldots n.$$

- Caso **continuo**, es decir, si (X_i, Y) es un vector aleatorio continuo,

$$E[X_i/\mathrm{Y} = \mathrm{y}] = \int_{Supp(f_{X_i/\mathbf{Y}=\mathbf{y}})} x_i f_{X_i/\mathrm{Y}=\mathrm{y}}(x_i)\, dx_i, \quad \forall \mathrm{y} \in E_{\mathrm{Y}} \quad i = 1, \ldots n.$$

Esperanza condicionada de una función unidimensional medible de un vector aleatorio

- Caso **discreto**: para E_X tal que $P(X \in E_X) = 1$, y E_Y tal que $P(Y \in E_Y) = 1$, y $g : E_X \longrightarrow \mathbb{R}$, una función medible,

$$E[g(X)/Y = y] = \sum_{x \in Supp(p_{\mathbf{X}/\mathbf{Y}=\mathbf{y}})} g(x)P(X = x/Y = y), \quad \forall y \in E_Y,$$

  donde (X, Y) es un vector aleatorio **discreto**.

- Caso **continuo**: para (X, Y) un vector aleatorio **continuo**, y $g : E_X \longrightarrow \mathbb{R}$, medible,

$$E[g(X)/Y = y] = \int_{Supp(f_{\mathbf{X}/\mathbf{Y}=\mathbf{y}})} g(x)f_{X/Y=y}(x)dx, \quad \forall y \in E_Y.$$

Se supone, como en el apartado anterior, que **las sumatorias e integrales anteriores son absolutamente convergentes**.

Outline

Momentos condicionados no centrados

Nótese que las expresiones que a continuación se formulan, se obtienen de forma directa del apartado anterior considerando $n = m = 1$, $g(X) = X^k$, y $k \in \mathbb{N}$.

Momentos condicionados no centrados de orden $k \in \mathbb{N}$ **:**

$$\textbf{Caso discreto} \quad E[X^k / Y = y] = \sum_{x \in Supp(p_{X/Y=y})} x^k p_{X/Y=y}(x) \quad \forall y \in E_Y$$

$$\textbf{Caso continuo} \quad E[X^k / Y = y] = \int_{Supp(f_{X/Y=y})} x^k f_{X/Y=y}(x) dx \quad \forall y \in E_Y.$$

Momentos condicionados centrados

Nótese que las expresiones que a continuación se formulan, se obtienen de forma directa del apartado anterior considerando $n = m = 1$, y $g(X) = [X - E(X/Y)]^k$, $k \in \mathbb{N}$.

Momentos condicionados centrados, respecto a la media, de orden $k \in \mathbb{N}$ **:**

Caso discreto

$$E\left[(X - E[X/Y])^k / Y = y\right] = \sum_{x \in Supp(p_{X/Y=y})} (x - E[X/Y])^k p_{X/Y=y}(x)$$

$\forall y \in E_Y$

Caso continuo

$$E\left[(X - E[X/Y])^k / Y = y\right] = \int_{Supp(f_{X/Y=y})} (x - E[X/Y])^k f_{X/Y=y}(x)dx$$

$\forall y \in E_Y$.

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

Caso especial: varianza condicionada

Supongamos que X, tal que $\exists E[X^2]$, e Y son dos variables aleatorias sobre el mismo espacio de probabilidad, entonces:

(i). $\exists \mathrm{Var}(X/Y)$, y $\mathrm{Var}(X/Y) \geq 0$.

(ii). $\mathrm{Var}(X/Y) = E[X^2/Y] - (E[X/Y])^2$.

(iii). **Descomposición de la varianza**: si $\exists \mathrm{Var}(E[X/Y])$, y $\exists E[\mathrm{Var}(X/Y)]$, entonces:

$$\mathrm{Var}(X) = \mathrm{Var}(E[X/Y]) + E[\mathrm{Var}(X/Y)].$$

$\mathrm{Var}(E[X/Y]) = E[(E[X/Y] - E[E[X/Y]])^2]$

Demostración:

$g(W) = (W - E(W))^2$

Propiedad V

$$\begin{aligned}
\mathrm{Var}(X) &= E\left[(X-E[X])^2\right] = E\left[E\left[(X-E[X])^2/Y\right]\right] \text{ (Prop.(v))} \\
&= E\left[E\left[(X^2+(E[X])^2-2XE[X])/Y\right]\right] \\
&= E\left[E[X^2/Y]+(E[X])^2-2E[X]E[X/Y]\right] \text{ (Linealidad)} \\
&= E\left[\underbrace{E[X^2/Y]-(E[X/Y])^2}+\underbrace{(E[X/Y])^2+(E[X])^2-2E[X]E[X/Y]}\right] \\
&= E[\mathrm{Var}(X/Y)] + E\left[(E[X/Y]-E[X])^2\right] \\
&= E[\mathrm{Var}(X/Y)] + \mathrm{Var}(E[X/Y]). \text{ (Prop. (v) y def. momento centrado condicionado de orden dos.)}
\end{aligned}$$

$E[E[X^2/Y]] = E[X^2]$

2 y $E[X]$ no dependen de x, luego se pueden sacar

$E[X] = E[E[X/Y]]$ (Considerando g la identidad)

$$\mathrm{Var}(X/Y) = E[(X-E[X/Y])^2/Y] = \int_{x\in \mathrm{Supp}(f_{X/Y})} (x - E[X/Y])^2 \cdot f_{X/Y}(x)\,dx = \int_{x\in \mathrm{Supp}(f_{X/Y})} x^2 \cdot f_{X/Y}(x) + E[X/Y]^2 \underbrace{\int f_{X/Y}(x)\,dx}_{1} - 2\cdot E[X/Y]\cdot \underbrace{\int_{x\in \mathrm{Supp}(f_{X/Y})} x \cdot f_{X/Y}(x)\,dx}_{E[X/Y]} =$$

$$= E[X^2/Y] + E[X/Y]^2 - 2\cdot E[X/Y]^2 = E[X^2/Y] - E[X/Y]^2$$

Outline

Regresión mínimo cuadrática bidimensional

El **problema de regresión** consiste en **aproximar una v.a.** Y **mediante una función de otra variable** X, es decir,

$$Y \simeq \varphi(X),$$

donde Y es la variable dependiente, explicada o endógena; X es la variable independiente, explicativa o exógena y φ es la función de regresión.

Consideramos la aproximación óptima de Y mediante la v.a. $\varphi(X)$, que proporciona la solución al problema de optimización, definido a partir de la **función de pérdida** $E[(Y - \varphi(X))^2]$, conocida como el **error cuadrático medio** (E.C.M) asociado a la estimación $\widehat{Y} = \varphi(X)$ de la v.a. Y.

Es decir, se considera φ_{opt}, donde

$$\varphi_{\text{opt}}(X) = \min_{\varphi} E[(Y - \varphi(X))^2].$$

Buscar en los apuntes de EDIP por qué se alcanza ahí el mínimo.

La **función** φ_{opt}, **que minimiza el E.C.M.**, en la aproximación de Y mediante una función de X **viene dada** por:

$$\varphi_{\text{opt}}(X) = E[Y/X].$$

Regresión mínimo cuadrática bidimensional

Por tanto,

$$
\begin{aligned}
\text{E.C.M.}(\varphi(X)) &= E\left[(Y - E[Y/X])^2\right] \\
(v) &= E\left[E\left[(Y - E[Y/X])^2/X\right]\right] = \underbrace{E\left[\mathrm{Var}(Y/X)\right]}_{\mathrm{Var}\,Y - \mathrm{Var}(E[Y/X])}.
\end{aligned}
$$

Calculando

Se desarrolla el cuadrado
+ Linealidad esperanza

$$
\begin{aligned}
E\left[(Y - E[Y/X])^2\right] &= E[Y^2] + E[(E[Y/X])^2] - 2E[YE[Y/X]] \\
&= E[Y^2] + E[(E[Y/X])^2] - 2E[E[YE[Y/X]/X]] \\
&= E[Y^2] + E[(E[Y/X])^2] - 2E[(E[Y/X])^2] \\
&= E[Y^2] - E\left[(E(Y/X))^2\right]
\end{aligned}
$$

Regresión mínimo cuadrática bidimensional

Alternativamente,

$$\begin{aligned}\text{E.C.M.}(\varphi(X)) &= E\left[\text{Var}(Y/X)\right] \\ &= E[E[Y^2/X] - (E(Y/X))^2]] = E[Y^2] - E\left[(E(Y/X))^2\right]\end{aligned}$$

Según lo que se vio, en el apartado anterior, considerando la fórmula de descomposición de la varianza, se tiene

$$\text{Var}(Y) = E\left[\text{Var}(Y/X)\right] + \text{Var}(E[Y/X]) = \text{E.C.M.}(\varphi(X)) + \text{Var}(\varphi_{\text{opt}}(X)).$$

Outline

Curvas de regresión mínimo cuadrática de Y sobre X

Se consideran, para diferentes valores x de la variable X, los correspondientes valores estimados de la v.a. Y, mediante la función φ_{opt}, es decir,

$$y = \varphi_{\text{opt}}(x) = E[Y/X = x], \quad \forall x \in E_X; \quad P(X \in E_X) = 1,$$

que recibe el nombre de **curva de regresión de Y sobre X**.

- El **E.C.M asociado a la curva de regresión** de Y sobre X es $E[\text{Var}(Y/X)]$.
- Para un **valor concreto**, **el E.C.M. asociado** a $E[Y/X = x]$ es $\text{Var}(Y/X = x)$.

Curvas de regresión mínimo cuadrática de X sobre Y

De forma análoga, para diferentes valores y de la variable Y, los correspondientes valores estimados de la v.a. X, mediante la función φ_{opt}, es decir,

$$x = \varphi_{\text{opt}}(y) = E[X/Y = y], \quad \forall y \in E_Y; \quad P(Y \in E_Y) = 1,$$

que recibe el nombre de **curva de regresión de** X **sobre** Y.

- El **E.C.M asociado a la curva de regresión** de X sobre Y es $E[\text{Var}(X/Y)]$.
- Para un **valor concreto**, **el E.C.M. asociado** a $E[X/Y = y]$ es $\text{Var}(X/Y = y)$.

Curvas de regresión mínimo cuadrática: algunas propiedades

(i). Si Y **depende funcionalmente de** X, es decir, $Y = f(X)$, la **curva de regresión** de Y sobre X, $E[Y/X = x]$, $x \in E_X$, **coincide con la curva de dependencia** $y = f(x)$, $x \in E_X$.

(ii). Si hay **dependencia funcional recíproca** entre X e Y, es decir, $Y = f(X)$ y $X = f^{-1}(Y)$, ambas **curvas de regresión coinciden con las curvas de dependencia**: $y = f(x)$, $x \in E_X$, y $x = f^{-1}(y)$, $y \in E_Y$.

(iii). Si X e Y son **independientes**, las curvas de regresión **son rectas paralelas** a los ejes: $y = E[Y]$ y $x = E[X]$.
En este caso **estimamos una variable, sin observar la otra, mediante su esperanza y el E.C.M. es su varianza.**

Cuadro resumen para predecir Y (equivalentemente para predecir X)

	Sin observar X	**Observando** X	**Para valor concreto** $X = x$
Predicción para Y	$E[Y]$	$E[Y/X]$	$E[Y/X = x]$
E.C.M.	$Var[Y]$	$E[Var[Y/X]]$	$Var[Y/X = x]$

Outline

Razones de correlación

La **correlación** estudia la bondad del ajuste de la función de regresión encontrada mediante el método de mínimos cuadrados, esto es, **en qué medida la función de regresión explica a una variable a partir de la otra**.

Se definen **las razones de correlación** de X sobre Y, $\eta^2_{X/Y}$ y de Y sobre X, $\eta^2_{Y/X}$, como sigue:

$\mathrm{Var}(X) = E(\mathrm{Var}(X/Y)) + \mathrm{Var}(E[X/Y])$

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

Observaciones:

- La razón de correlación es una **medida adimensional** que representa la **proporción de varianza de la variable dependiente que queda explicada** por la función de regresión.
- En este sentido se puede interpretar como una **medida de la bondad del ajuste** de la distribución a la curva de regresión correspondiente.

$$\frac{Var(E[aX/Y])}{Var(aX)} = \frac{a^2 Var(E[X/Y])}{a^2 Var(X)} = \frac{Var(E[X/Y])}{Var(X)}$$

Propiedades

(i). Para cualquier $a > 0$, se tiene $\eta^2_{aX/Y} \overset{\shortparallel}{=} \eta^2_{X/Y}$

(ii). Para cualquier $a > 0$, se tiene $\eta^2_{aY/X} = \eta^2_{Y/X}$

(iii). $0 \leq \eta^2_{X/Y} \leq 1$ y $0 \leq \eta^2_{Y/X} \leq 1$ (Ejercicio voluntario)

(iv). Si $\eta^2_{X/Y} = 0$ y $\eta^2_{Y/X} = 0$, entonces las curvas de regresión de X/Y e Y/X coinciden respectivamente con las rectas $x = E[X]$ e $y = E[Y]$.

(v). Recíprocamente, si las curvas de regresión de X/Y e Y/X coinciden con las rectas $x = E[X]$ e $y = E[Y]$, respectivamente, entonces $\eta^2_{X/Y} = 0$ y $\eta^2_{Y/X} = 0$.

(vi). $\eta^2_{X/Y} = 1 \Leftrightarrow X$ dependen funcionalmente de Y.

(vii). $\eta^2_{Y/X} = 1 \Leftrightarrow Y$ dependen funcionalmente de X.

(viii). $\eta^2_{Y/X} = \eta^2_{X/Y} = 1 \Leftrightarrow$ hay dependencia funcional recíproca entre X e Y.

(iv). Si $\eta^2_{X/Y} = 0$ y $\eta^2_{Y/X} = 0$, entonces las curvas de regresión de X/Y e Y/X coinciden respectivamente con las rectas $x = E[X]$ e $y = E[Y]$.

(v). Recíprocamente, si las curvas de regresión de X/Y e Y/X coinciden con las rectas $x = E[X]$ e $y = E[Y]$, respectivamente, entonces $\eta^2_{X/Y} = 0$ y $\eta^2_{Y/X} = 0$.

(vi). $\eta^2_{X/Y} = 1 \iff X$ dependen funcionalmente de Y.

(vii). $\eta^2_{Y/X} = 1 \iff Y$ dependen funcionalmente de X.

(viii). $\eta^2_{Y/X} = \eta^2_{X/Y} = 1 \iff$ hay dependencia funcional recíproca entre X e Y.

(iv) $\eta^2_{X/Y} = 0 \Rightarrow Var(E[X/Y]) = 0 \Rightarrow$ Todos los valores de $E[X/Y]$ son iguales $\Rightarrow$ La curva de regresión es $E[X/Y] = c$: Es una recta paralela al eje.

Tenemos que probar que $E[X/Y] = E[X]$. Para ello aplicamos la propiedad (v) de la esperanza condicionada :

Como $\exists\, E[X/Y]$, se tiene que $E[E[X/Y]] = E[X]$

$$\Rightarrow E[E[X/Y]] = E[c] = c = E[X/Y] = E[X]$$

(v) Supongamos que $E[X/Y] = E[X] = c \Leftrightarrow Var(E[X/Y]) = 0$
$\Rightarrow \eta^2_{X/Y} = 0$

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

Por la descomposición de la varianza :

$$Var(E[X/Y]) = Var(X) - E[Var(X/Y)] = Var(X)$$

$$\Rightarrow E[Var(X/Y)] = 0 \overset{0 \leq Var \leq 1}{\Rightarrow} Var(X/Y) = 0 \Rightarrow$$

$\Rightarrow$ Todos los valores de X para un $Y = y$ son iguales

Para cada $y \in E_Y$, se define $f(y) = x$ (valor de X para ese y), con lo que $X = f(Y)$, es decir, X depende funcionalmente de Y.

$\Leftarrow$ Supongamos que X depende funcionalmente de Y: $X = f(Y)$. Entonces, la curva de regresión de X/Y coincide con la curva de dependencia:

$$E[X/Y] = f(Y)$$

Por tanto, $Var(E[X/Y]) = Var(f(Y)) = Var(X) \Rightarrow$
$\Rightarrow \eta^2_{X/Y} = 1$

(viii) Trivial si se prueban las dos anteriores

Outline

Planteamiento de la regresión lineal de Y sobre X (equivalentemente para X sobre Y)

Intentamos resolver el problema de predecir, por mínimos cuadrados, los valores de una variable Y a partir de una función lineal de otra X. En este sentido se busca una función $\varphi^L_{\text{opt}}(X)$ que minimice el E.C.M.

$$\varphi^L_{\text{opt}}(X) = \min_{a,b} E[(Y - aX - b)^2].$$

$(Y-(aX+b))^2$

$Y^2+a^2X^2+b^2+2abX-2aXY-2bY$

El problema se reduce a obtener a y b que minimicen $L = E[(Y - aX - b)^2]$.

$$L = E[(Y - aX - b)^2] = E[Y^2] + a^2E[X^2] + b^2 + 2abE[X] - 2aE[XY] - 2bE[Y]$$

Derivando respecto a y b e igualando a cero, se obtiene el siguiente sistema de ecuaciones denominado sistema de **Ecuaciones Normales**:

$$\begin{aligned} aE[X^2] + bE[X] &= E[XY] \\ b + aE[X] &= E[Y] \end{aligned}$$

Cayó en el examen del año pasado

cuya solución es $a = \frac{Cov(X,Y)}{Var(X)}$ y $b = E[Y] - aE[X]$.

Regresión mínimo-cuadrática, cuando φ^L es una recta ($R_{Y/X}$)

A partir de las soluciones de las ecuaciones normales anteriores se tiene lo siguiente. **En la regresión lineal mínimo cuadrática de Y sobre X,**

$$Y \approx \varphi^L_{\text{opt}}(X) = \min_{a,b} E[(Y - aX - b)^2] = E[Y] + \frac{\text{Cov}(X,Y)}{\text{Var}(X)}(X - E[X]).$$

Esta expresión recibe el nombre de **Recta de Regresión de Y sobre X.**

$$y - E[Y] = \overbrace{\frac{\text{Cov}(X,Y)}{\text{Var}(X)}}(x - E[X]).$$

Coeficiente de regresión. El signo dependerá de Cov(X,Y). Si es negativo, la relación es inversa

El **E.C.M asociado** viene dado por: (Ejercicio voluntario)

$$\begin{aligned} E.C.M.(\varphi^L_{\text{opt}}(X)) &= \text{Var}(Y) - \frac{[\text{Cov}(X,Y)]^2}{\text{Var}(X)} \\ \text{Por tanto, } \text{Var}(Y) &= \text{Var}(\varphi^L_{\text{opt}}(X)) + E.C.M.(\varphi^L_{\text{opt}}(X)). \end{aligned}$$

Regresión mínimo-cuadrática, cuando φ^L es una recta ($R_{X/Y}$)

En la regresión lineal mínimo cuadrática de X sobre Y, de forma similar, se calcula $\varphi^L_{\text{opt}}(Y)$,

$$X \approx \varphi^L_{\text{opt}}(Y) = \min_{a,b} E[(X - aY - b)^2] = E[X] + \frac{\text{Cov}(X,Y)}{\text{Var}(Y)}(Y - E[Y]).$$

Esta expresión recibe el nombre de **Recta de Regresión de X sobre Y**.

$$x - E[X] = \frac{\text{Cov}(X,Y)}{\text{Var}(Y)}(y - E[Y]).$$

El **E.C.M asociado** viene dado por:

$$\begin{aligned} E.C.M.(\varphi^L_{\text{opt}}(Y)) &= \text{Var}(X) - \frac{[\text{Cov}(X,Y)]^2}{\text{Var}(Y)} \\ \text{Por tanto, } \text{Var}(X) &= \text{Var}(\varphi^L_{\text{opt}}(Y)) + E.C.M.(\varphi^L_{\text{opt}}(Y)). \end{aligned}$$

Coeficientes de regresión

Suponiendo la existencia de $E[X^2]$ y $E[Y^2]$, se definen los **coeficientes de regresión** de Y/X, $\gamma_{Y/X}$, y de X/Y, $\gamma_{X/Y}$, como sigue:

$$\gamma_{Y/X} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)}$$

$$\gamma_{X/Y} = \frac{\text{Cov}(X, Y)}{\text{Var}(Y)}$$

Algunas propiedades

(i). Si $\text{Cov}(X, Y) = 0$, entonces las rectas de regresión de Y/X y de X/Y vienen respectivamente dadas por $y = E[Y]$ y $x = E[X]$ (son rectas paralelas a los ejes).
Sustituyendo en la fórmula

(ii). Si la curva de regresión es una recta, coincide con la correspondiente recta de regresión.

(iii). Si X e Y están linealmente relacionadas, las rectas de regresión coinciden con la recta de dependencia (sólo en este caso, obtenida una recta de regresión, la otra se puede obtener despejando).

(iv). Las rectas de regresión se cortan en el punto $(E[X], E[Y])$.

(v). Los coeficientes de regresión tienen el mismo signo que la covarianza.
Sale de ver la fórmula

(vi). Los coeficientes de regresión cuantifican cuanto se incrementa la variable dependiente por aumentos unitarios en la variable independiente.

Outline

Coeficiente de determinación lineal

Una medida de la asociación lineal de las varibles viene determinada por el **coeficiente de determinación lineal** $\rho^2_{X,Y}$, se define como:

$$\rho^2_{X,Y} = \frac{[\mathrm{Cov}(X,Y)]^2}{\mathrm{Var}(X)\mathrm{Var}(Y)} \left(= \frac{\mathrm{Var}(\varphi^L_{\mathrm{opt}}(X))}{\mathrm{Var}(Y)} = \frac{\mathrm{Var}(\varphi^L_{\mathrm{opt}}(Y))}{\mathrm{Var}(X)} \right).$$

Observación:

Tal y como los razones de correlación, el coeficiente de determinación lineal **es una medida de la bondad del ajuste lineal**, o lo que es lo mismo, mide la proporción de varianza de explicada por el modelo de regresión lineal.

Propiedades del coeficiente de determinación lineal

(i). $\rho^2_{aX+b,cY+d} = \rho^2_{X,Y}$

(ii). $\rho^2_{X,Y} = \gamma_{X/Y}\gamma_{Y/X}$

(iii). $0 \leq \rho^2_{X,Y} \leq 1$

(iv). $\rho^2_{X,Y} = 0 \Leftrightarrow$ Las rectas de regresión de Y/X y de X/Y son respectivamente $y = E[Y]$, y $x = E[X]$.

(v). $\rho^2_{X,Y} = 1 \Leftrightarrow$ Existe dependencia funcional lineal entre las variables X e Y.

(vi). $\rho^2_{X,Y} \leq \eta^2_{Y/X}$ y $\rho^2_{X,Y} \leq \eta^2_{X/Y}$.

Las desigualdades anteriores (prop. (vi).) se convierten en igualdades si y sólo si **las curvas de regresión coinciden con las correspondientes rectas de regresión**.

Coeficiente de correlación lineal

Independientemente de que el coeficiente de determinación lineal ofrece información acerca de la bondad del ajuste del modelo lineal, la **medida de correlación por excelencia** es el **coeficiente de correlación lineal**, ya que también proporciona información, no sólo de la **fuerza de asociación**, sino también de **su sentido**.

Se define el coeficiente de correlación lineal $\rho_{X,Y}$, como:

$$\rho_{X,Y} = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$$

Propiedades del coeficiente de correlación lineal

(i). $-1 \leq \rho_{X,Y} \leq 1$

(ii). $\rho_{aX+b,cY+d} = \pm\rho_{X,Y}$

(iii). $\rho_{X,Y} = 0$ (variables incorreladas linealmente) $\Leftrightarrow$ Las rectas de regresión de Y/X y de X/Y son respectivamente $y = E[Y]$, y $x = E[X]$.

(iv). Si X e Y son independientes, entonces $\rho_{X,Y} = 0$ (el recíproco no es cierto)

(v). $|\rho_{X,Y}| = 1$ $\Leftrightarrow$ Existe dependencia funcional lineal (positiva o negativa, en función del signo) entre X e Y.

(vi). $|\rho_{X,Y}| = 1$ $\Leftrightarrow$ Las dos rectas de regresión coinciden entre sí y con la recta de dependencia, además de con las curvas de regresión.

(vii). Si $\rho_{X,Y} = 1$ la dependencia lineal es exacta y positiva (las dos v.a. crecen o decrecen de forma simultánea).

(viii). Si $\rho_{X,Y} = -1$ la dependencia lineal es exacta y negativa (cuando una v.a crece la otra decrece y viceversa).