# ESTAD _ T8. 1. ESPERANZA CONDICIONADA.
2. PROPIEDADES.
3. LINEA GENERAL de REGRESIÓN.
4. REGRESIÓN MÍNIMO CUADRÁTICA.
(cubrir)5. PROPIEDADES.

---

## 1 _ ESPERANZA CONDICIONADA.

En ocasiones, cuando un experimento aleatorio da lugar a una v.a. bidimensional, es interesante estudiar la distrib. de una de ellas condicionada a un valor concreto de la otra. Al igual que en el caso bidimensional, nos podemos plantear el cálculo de las características de la distrib. condicionada.

Definición: Sea $(\xi, \eta)$ una v.a. bidimensional y sea $y$ un valor posible de $\eta$ ( $P(\eta = y) \neq 0$ v.a. discreta ó $f_\eta(y) \neq 0$ v.a. continua). Llamamos esperanza de $\xi$ condicionada por $\eta = y$ al valor, si existe

$$E[\xi/\eta = y] = \sum_i x_i P(\xi = x_i/\eta = y) \quad \text{para v.a. discretas}$$

ó bien

$$E[\xi/\eta = y) = \int_{x=-\infty}^{x=+\infty} x \cdot f_{\xi/\eta}(x)\, dx \quad \text{para v.a. continuas}$$

Análogamente, se puede definir $E[\eta / \xi = x]$

Recordemos que $P(\xi = x_i/\eta = y) = \dfrac{P(\xi = x_i, \eta = y)}{P(\eta = y)}$

$$f_{\xi/\eta}(x) = \frac{f_{\xi,\eta}(x,y)}{f_\eta(y)}$$

Observaciones

01 → $E[\xi/\eta = y]$ existe si $E[|\xi|/_{\eta=y}]$ es finita.

Al igual que en el caso unidim., la esperanza existe si la suma infinita ó la integral es absolutamente convergente.

02 → La interpretación de $E[\xi/\eta=y]$ es obvia: es el valor esperado de $\xi$ para un valor concreto de $\eta$.

03 → $E[\xi/\eta=y] = g(y)$, $\forall y$.

Para cada valor posible de $\eta$, existe una distrib. condic. de $\xi$, y por tanto un valor concreto de la esperanza condicionada, por lo que la esperanza condicionada es una función de $y$.

04 → se puede calcular la esperanza de una función de $\xi$ de manera natural.

$$E[h(\xi)/_{\eta=y}] = \sum_i h(x_i) \cdot P(\xi=x_i/_{\eta=y})$$ , caso discreto

ó bien

$$E[h(\xi)/_{\eta=y}] = \int_{x=-\infty}^{+\infty} h(x) \cdot f_{\xi/\eta}(x)\, dx$$ , caso continuo

05 → Generalización a una condición compuesta

$E[\xi/_{\eta>y}]$ ó $E[\xi/_{\eta\leq y}]$ ó $E[\xi/_{y_1\leq\eta\leq y_2}]$ . . .

## 2 – PROPIEDADES de la esperanza condicionada

La esperanza condicionada tiene, con la salvedad de la condición, las mismas propiedades de la esperanza matemática de una v.a. unidimensional + otras específicas. Las vemos:

### Generales

P1 – La esperanza condicionada de una cte es ella misma.

$$E[\xi/\eta = y] = K \quad \text{si} \quad P(\xi = K) = 1$$

P2 – La esperanza condicionada verifica la condición de linealidad, siempre que existan las esperanzas.

$$E\left[(a \cdot g(\xi_1) + b \cdot h(\xi_2))/_{\eta = y}\right] = a\, E[g(\xi_1)/_{\eta = y}] + b\, E[h(\xi_2)/_{\eta = y}]$$

P3 – $g(\xi) \leq h(\xi) \Rightarrow E[g(\xi)/_{\eta = y}] \leq E[h(\xi)/_{\eta = y}]$

siempre que existan dichas esperanzas.

### Específicas

P4 – Si las v.a. son independientes, la esperanza condicionada coincide con la esperanza marginal.

$$\xi \text{ indep } \eta \Rightarrow E[\xi/_{\eta = y}] = E[\xi] \quad , \text{siempre que existan}$$

Dem: Utilizando la propiedad $f(x/y) = f(x)$ para continuas sale directamente. Se haría de manera análoga para discretas.

P5 – $E_\eta\, E_\xi(\xi/\eta) = E_\xi(\xi)$

donde $E_\xi \equiv$ esperanza respecto a la v.a. $\xi$

$E_\eta \equiv$ esperanza respecto a la v.a. $\eta$

Dem: $$E_\eta E_\xi(\xi/\eta) = \int_\eta \left( \int_\xi x\, f(x/y)\, dx \right) f_2(y)\, dy =$$

$$= \int_\eta \left( \int_\xi x \frac{f(x,y)}{\cancel{f_2(y)}} \cdot \cancel{f_2(y)}\, dy dx \right) =$$

$$= \int_\xi x \underbrace{\left( \int_\eta f(x,y)\, dy \right)}\, dx = \int_\xi x \cdot f_1(x)\, dx = E_\xi[\xi]$$

## 3_ LÍNEA GENERAL de REGRESIÓN

Con frecuencia, estamos interesados en el estudio de la relación existente entre 2 (ó más) variables aleatorias de cara a explicar el comportamiento de una de ellas a partir de la otra (u otras) y, en última instancia, predecir su valor.

En el campo económico, los análisis relacionados con la regresión provocaron, década de los 30, el nacimiento de una rama científica de la Estadística, la Econometría, herramienta básica en la investigación económica.

Dada una v.a. bidimensional $(\xi, \eta)$, el objetivo de la regresión general es buscar una función de una de las var. que explique lo mejor posible la otra.

Objetivo: encontrar $h$ tq $\boxed{\eta = h(\xi)}$

Para cada valor $x$ de $\xi$, el comportamiento de $\eta$ viene explicitado por su f. distrib. condicionada:

$$F(y/_{\xi = x}) = F(y/x)$$

El criterio de regresión general consiste en asignar a cada valor $x$ de $\xi$, el valor medio o esperanza de $\eta$ de cada distrib. condicionada, con lo cual la curva de regresión I será el lugar geométrico de las medias condicionadas para cada $\xi = x$

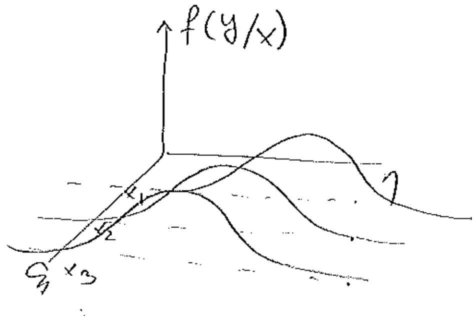
---

El término "regresión" se debe a Galton en sus estudios de Biometría (las alturas de los hijos regresan al valor medio) y a partir de allí se utilizó impropiamente para designar las técnicas de análisis de las relaciones entre var. estadísticas.

Para cada ~~x~~ $\xi = x$,

$$\bar{y} = E[\eta / \xi = x] = \int_{-\infty}^{+\infty} y \cdot f(y/x)\, dy = \frac{1}{f_1(x)} \int_{-\infty}^{+\infty} y \cdot f(x,y)\, dy$$

$$= m_1(x)$$




Análogamente, para cada $\eta = y$,

$$\bar{x} = E[\xi / \eta = y] = \ldots = m_2(y)$$

- Las curvas de regresión I proporcionan las funciones explicativas $\left\{ \begin{matrix} m_1(x) \\ m_2(y) \end{matrix} \right\}$ que ofrecen la "mejor representación o estimación posible" de la variable explicada.

- Minimiza el ECM para todos y cada uno de los valores x de $\xi$

$$\min ECM[\eta] = \min E[\eta - h_1(x) / \xi = x]^2 = \ldots = E[\eta - \underset{\text{media}}{m_1(x)}]^2$$

en virtud de la P2 de la varianza.

→ Por tanto, $\eta = m_1(x)$ es la función que mejor se ajusta a la masa de probabilidad conjunta y ofrece una capacidad explicativa con menor error.

- La curva de regresión $m_1(x)$ tendrá la forma funcional que corresponda a los puntos que son esperanzas condicionadas, que dependerá de cómo esté distribuida la masa de probab. conjunta, $F(x,y)$.

## Regresión lineal general

En el caso concreto de que a las funciones $\left\{ \begin{matrix} m_1(x) \\ m_2(y) \end{matrix} \right\}$ se les imponga una restricción funcional, que pertenezcan a la familia ~~recta~~ lineal, hay que encontrar los coef. de las rectas de regresión.

$$\bar{y} = E[\eta/\xi = x] = m_1(x) = b_0 + b_1 x$$
$$\bar{x} = E[\xi/\eta = y] = m_2(y) = b_0' + b_1' y$$

En el caso de la regresión de $\eta/\xi$ (análogo para $\xi/\eta$):

$$\bar{y} = E[\eta/\xi = x] = \frac{1}{f_1(x)} \int_{-\infty}^{+\infty} y f(x,y)\, dy = b_0 + b_1 x$$

$$\int_{-\infty}^{+\infty} y f(x,y)\, dy = b_0 f_1(x) + b_1 x f_1(x)$$

Integrando ambos miembros respecto a $x$:

$$(*) \quad \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y f(x,y)\, dy\, dx = b_0 \underbrace{\int_{-\infty}^{+\infty} f_1(x) dx}_{=1} + b_1 \int_{-\infty}^{+\infty} x f_1(x) dx$$

Multiplicando ambos miembros por $x$ antes de integrar:

$$(**) \quad \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x y f(x,y)\, dy\, dx = b_0 \int_{-\infty}^{+\infty} x f_1(x) dx + b_1 \int_{-\infty}^{+\infty} x^2 f_1(x) dx$$

$(*)$ y $(*)$ dan lugar a un sistema de ecuaciones que, en términos de momentos respecto al origen, quedo:

$$\left. \begin{aligned} \alpha_{01} &= b_0 + b_1 \alpha_{10} \\ \alpha_{11} &= b_0 \alpha_{10} + b_1 \alpha_{20} \end{aligned} \right\} \text{sist.}$$

---

Para resolverlo por reducción, multiplicamos la primera ecuación por $(-\alpha_{10})$ y sumamos:

$$-\alpha_{01}\cdot\alpha_{10} = -\cancel{b_0\alpha_{10}} - b_1\,\alpha_{10}^2$$

$$\alpha_{11} = \cancel{b_0\alpha_{10}} + b_1\,\alpha_{20}$$

---

$$\alpha_{11} - \alpha_{01}\cdot\alpha_{10} = (\alpha_{20} - \alpha_{10}^2)\cdot b_1$$

Despejando $b_1$:

$$b_1 = \frac{\alpha_{11} - \alpha_{01}\cdot\alpha_{01}}{\alpha_{20} - \alpha_{10}^2} = \frac{\mu_{11}}{\mu_{20}} = \frac{\mathrm{Cov}(\xi,\eta)}{\mathrm{Var}(\xi)}$$

Sustituyendo $b_1$ en la 1ª ecuación original:

$$b_0 = \alpha_{01} - b_1\alpha_{10}$$

Resumiendo, para $\eta/\xi$:

$$\left.\begin{array}{l}\alpha_{01} = b_0 + b_1\alpha_{10}\\ \alpha_{11} = b_0\alpha_{10} + b_1\alpha_{20}\end{array}\right\} \xrightarrow{\text{Soluc.}} \begin{array}{l} b_1 = \dfrac{\mathrm{Cov}(\xi,\eta)}{\mathrm{Var}(\xi)} \\ b_0 = \alpha_{01} - b_1\alpha_{10}\end{array}$$

Análogamente, para $\xi/\eta$:

$$\xrightarrow{\text{Soluc.}} \begin{array}{l} b'_1 = \dfrac{\mathrm{Cov}(\xi,\eta)}{\mathrm{Var}(\eta)} \\ b'_0 = \alpha_{10} - b'_1\alpha_{01}\end{array}$$

## 4. REGRESIÓN MÍNIMO CUADRÁTICA.

La regresión mínimo-cuadrática se diferencia de la regresión general en dos aspectos.

1 _ Se selecciona a priori la forma funcional que represente la relación entre $\xi$ y $\eta$.
Se trata de encontrar, entre todas las funciones de la misma familia, la que mejor represente a $\eta$.

2 _ El criterio de optimización es el de los mínimos cuadrados, es decir, la función que buscamos debe verificar que el cuadrado del error cometido sea mínimo.

En el caso lineal:

$\hat{\eta} = \beta_0 + \beta_1 \xi \longleftarrow$ recta (estimación lineal de $\eta$).

$e = \eta - \hat{\eta} = \eta - (\beta_0 + \beta_1 \xi) \longleftarrow$ error cometido

objetivo: ~~tener~~ encontrar $\beta_0$ y $\beta_1$ que minimicen la esperanza de la v.a. error al cuadrado.

$$\min_{\beta_0,\beta_1} E[e^2] = \min E[\eta - (\beta_0 + \beta_1 \xi)]^2 = \min \Psi(\beta_0, \beta_1)$$

$$\text{C.N.} \longrightarrow \left.\begin{array}{l} \dfrac{\partial \Psi(\beta_0,\beta_1)}{\partial \beta_0} = 2E[(\eta - (\beta_0 + \beta_1 \xi))(-1)] = 0 \\ \dfrac{\partial \Psi(\beta_0,\beta_1)}{\partial \beta_1} = 2E[(\eta - (\beta_0 + \beta_1 \xi))(-\xi)] = 0 \end{array}\right\} \xrightarrow{\text{soluc.}}$$

$$\longrightarrow \left\{\begin{array}{l} -E(\eta) + \beta_0 + \beta_1 E(\xi) = 0 \\ -E(\xi\eta) + \beta_0 E(\xi) + \beta_1 E(\xi^2) = 0 \end{array}\right\} \longrightarrow \left\{\begin{array}{l} E[\eta] = \beta_0 + \beta_1 E[\xi] \\ E[\xi\eta] = \beta_0 E[\xi] + \beta_1 E[\xi^2] \end{array}\right.$$

$$\longrightarrow \left\{\begin{array}{l} \alpha_{01} = \beta_0 + \beta_1 \alpha_{10} \\ \alpha_{11} = \beta_0 \alpha_{10} + \beta_1 \alpha_{20} \end{array}\right\} \longrightarrow$$ sistema de ec's normales cuya soluc. es

---

Soluc. de C.N. $\longrightarrow \begin{cases} \beta_1 = \dfrac{\alpha_{11} - \alpha_{10}\alpha_{01}}{\alpha_{20} - \alpha_{10}^2} = \dfrac{\mu_{11}}{\mu_{20}} = \dfrac{\mathrm{Cov}(\xi,\mu)}{\mathrm{Var}(\xi)} \\ \beta_0 = \alpha_{01} - \beta_1\alpha_{10} \end{cases}$

C.S. Para que el punto crítico sea mínimo de la f. objetivo, la matriz hessiana en el punto ha de ser def (+).

$$H\psi(\beta_0,\beta_1) = \begin{vmatrix} \dfrac{\partial^2\psi}{\partial\beta_0^2} & \dfrac{\partial^2\psi}{\partial\beta_0\beta_1} \\ \dfrac{\partial^2\psi}{\partial\beta_0\beta_1} & \dfrac{\partial^2\psi}{\partial^2\beta_1} \end{vmatrix} = \begin{vmatrix} \underbrace{2}_{>0} & 2E(\xi) \\ 2E(\xi) & 2E(\xi^2) \end{vmatrix}_{>0} =$$

$$= 4E(\xi^2) - 4(E(\xi))^2 = 4\cdot \mathrm{Var}(\xi) \geq 0 \quad \text{, siempre}$$

Por tanto, la recta de regresión mínimo cuadrática de $\eta/\xi$ será:

$$\hat{\eta} = \beta_0 + \beta_1\xi \quad \forall\xi. \qquad \begin{cases} \beta_1 = \dfrac{\mathrm{Cov}(\xi,\eta)}{\mathrm{Var}(\xi)} \\ \beta_0 = E(\eta) - \beta_1 E(\xi) \end{cases}$$

que tb. se puede expresar como:

$$\eta - E[\eta] = \frac{\mathrm{Cov}(\xi,\eta)}{\mathrm{Var}(\xi)}(\xi - E[\xi])$$

Análogamente, la recta de regresión mínimo cuadrático de $\xi$ sobre $\eta$ será:

$$\xi - E[\xi] = \frac{\mathrm{Cov}(\xi,\eta)}{\mathrm{Var}(\eta)}(\eta - E[\eta])$$

Observaciones

01 → La intersección de las dos rectas de regresión mínimo-cuadráticas es el centro de gravedad de la distrib. bidim.
$(\alpha_{10}, \alpha_{01}) = (E(\xi), E(\eta))$.

02 → Los coeficientes $\beta_1$ y $\beta_1'$ $\Big|$ $\beta_1 = \dfrac{cov(\xi,\eta)}{Var(\xi)}$, $\beta_1' = \dfrac{cov(\xi,\eta)}{Var(\eta)}$, son las pendientes de las rectas de regresión y reciben el nombre de coeficientes de regresión.
$\Rightarrow$ signo $(\beta_1)$ = signo $(\beta_1')$ = signo $(cov(\xi,\eta))$.

03 → Cuando la regresión general es lineal, los parámetros $b_0$ y $b_1$ coinciden con $\beta_0$ y $\beta_1$.

Tma: La regresión mínimo-cuadrático lineal es la línea de más estricto ajuste a la curva de regresión I, en el sentido que se minimizan las distancias verticales entre cada punto de la recta $\eta^* = \beta_0 + \beta_1\xi$ y la curva de regresión, $m_1(x)$.

Caso particular: Si $m_1(x)$ es un recta, entonces coinciden.

Relación con la correlación:
$Var(\eta) = Var(\eta/\xi) + Var(\text{residual}) \rightarrow 1 = \rho^2 + \dfrac{\sigma_e^2}{\sigma_\eta^2}$

coef. correlac. lineal → $-1 \leq \rho \leq +1$ ⟵ $0 \leq \rho^2 \leq 1$ → coef. determ. general

En el caso lineal, $\rho = \dfrac{cov(\xi,\eta)}{\sigma_\xi \sigma_\eta}$ y $\rho^2 = \beta_1 \cdot \beta_1'$

# 5. REGRESIÓN MÚLTIPLE

Consideremos el vector aleatorio $(\xi_1 \dots \xi_n, \eta)$ donde queremos explicar la variable $\eta$ a partir de la relación existente con las otras variables.

Regresión I múltiple:

$$E[\eta/\xi_1 \dots \xi_n] = \int_{-\infty}^{+\infty} y \cdot f(y/x_1 \dots x_n)\, dy = m_2(x_1 \dots x_n)$$

La hipersuperficie de regresión ofrece la mejor estimación de $\eta$, en el sentido que minimiza el ECM cuando éste se encuentra condicionado a $\xi_1 = x_1 \dots \xi_n = x_n$.

Regresión II múltiple:

Primero hay que elegir la familia de funciones y después obtener los parámetros óptimos que minimicen la esperanza de los errores al cuadrado, $\min E(e)^2$

Si la familia de funciones es lineal, la regresión mínimo-cuadrática está definida por el hiperplano:

$$\eta^* = \beta_0 + \beta_{11}\xi_1 + \dots + \beta_{1n}\xi_n$$

cuya solución es:

$$\hat{\beta}_{1K} = \frac{-|\Sigma_{1K}|}{|\Sigma_{11}|}, \quad K = 1 \dots n$$

$$\beta_0 = E[\eta] - \beta_{11}E[\xi_1] - \dots - \beta_{1n}E[\xi_n]$$

donde $\left.\begin{matrix}|\Sigma_{1K}|\\|\Sigma_{11}|\end{matrix}\right\}$ son los menores complementarios de orden $(1,K)$ y $(1,1)$ respectivamente, de la matriz de varianzas-covarianzas $\Sigma$.