Tema 4. Parte I. Esperanza Condicionada: Regresión y Correlación

Doble Grado en Ingeniería Informática y Matemáticas

Úrsula Torres Parejo

Departamento de Estadística e I.O. Universidad de Granada

Contenido.

- Esperanza condicionada de una variable aleatoria
- Esperanza condicionada de una función medible
- Propiedades y teoremas de la esperanza condicionada
- Momentos condicionados bidimensionales
- Teorema de descomposición de la varianza

Esperanza condicionada.

Sean X e Y variables aleatorias definidas sobre un mismo espacio de probabilidad, tales que $\exists\ E[X]$. Se define la esperanza de X dada Y (o esperanza de X condicionada a Y) y se denota por $E[X|Y]$ como la variable aleatoria que toma el valor $E[X|Y = y_0]$ cuando $Y = y_0$ y se calcula como sigue:

- **Caso discreto**
  $E[X|Y = y_0] = \sum_x x \cdot P[X = x|Y = y_0]$ con $P[Y = y_0] > 0$
- **Caso continuo**
  $E[X|Y = y_0] = \int_{-\infty}^{\infty} x \cdot f_{X|Y=y_0}(x)dx$ con $f_Y(y) > 0$

Notas:

- $E[X|Y]$ es una función de Y
- $E[X|Y = y_0]$ no es más que la media de X considerando como distribución la condicionada de X dado $Y = y_0$
- De forma similar se puede definir la esperanza de Y dada X, supuesta la existencia de E[Y].

Esperanza condicionada de una función medible.

Sean X e Y variables aleatorias definidas sobre un mismo espacio de probabilidad y sea $g : (\mathbb{R}, \mathcal{B}) \to (\mathbb{R}, \mathcal{B})$ una función medible tal que $\exists\ E[g(X)]$. Se define la esperanza condicionada de g(X) dado Y y se denota por $E[g(X)|Y]$ como la variable aleatoria que toma el valor $E[g(X)|Y = y_0]$ cuando $Y = y_0$ y se calcula como sigue:

- **Caso discreto**
  $E[g(X)|Y = y_0] = \sum_x g(x) \cdot P[X = x|Y = y_0]; P[Y = y_0] > 0$
- **Caso continuo**
  $E[g(X)|Y = y_0] = \int_{-\infty}^{\infty} g(x) \cdot f_{X|Y=y_0}(x)dx \quad \text{con } f_Y(y) > 0$

Nota: De forma similar se puede definir $E[g(Y)|X]$

Propiedades de la esperanza condicionada.

1. $E[c/Y] = c$
2. **Linealidad:** Sean X e Y variables aleatorias definidas en el mismo espacio de probabilidad tales que $\exists\ E[X]$ y $a, b \in \mathbb{R} \Rightarrow \exists\ E[(aX + b)|Y] = aE[X|Y] + b$
3. Sean $X_1, ..., X_n$ variables aleatorias tales que $\exists\ E[X_i], i = 1, 2, ..., n \Rightarrow \forall a_1, ..., a_n \in \mathbb{R}$ $\exists\ E[(a_1X_1 + ... + a_nX_n)|Y] = a_1E[X_1|Y] + ... + a_nE[X_n|Y]$
4. Si $X \geq 0$ y $\exists\ E[X] \Rightarrow E[X|Y] \geq 0$ y $E[X|Y] = 0 \Leftrightarrow P[X = 0] = 1$
5. **Conservación del orden**: Si X_1 y X_2 son variables aleatorias tales que $\exists\ E[X_1], E[X_2]$ y $X_1 \leq X_2 \Rightarrow E[X_1|Y] \leq E[X_2|Y]$

Teoremas.

Teorema 1
Si X e Y son variables aleatorias independientes y $\exists\ E[g(X)]$ siendo g una función medible $\Rightarrow$

$$E[g(X)|Y] = E[g(X)]$$

En particular, $E[X|Y] = E[X]$

Teorema 2
Si X e Y son variables aleatorias definidas en el mismo espacio de probabilidad tales que $\exists\ E[g(X)]$ siendo g una función medible $\Rightarrow$

$$\exists\ E[E[g(X)|Y]] = E[g(X)]$$

En particular, $E[E[X|Y]] = E[X]$

Momentos condicionados bidimensionales.

Sean X e Y variables aleatorias definidas sobre el mismo espacio de probabilidad:

- Si $\exists\ E[X^n],\ n \in \mathbb{N} \Rightarrow$ a $E[X^n|Y]$ se le llama momento condicionado no centrado de orden n de X dada Y.
- Si $\exists\ E[X^n],\ n \in \mathbb{N} \Rightarrow$ a $E[(X - E[X|Y])^n|Y]$ se le llama momento condicionado centrado de orden n de X dada Y.

Ejemplo de cálculo:

$E[X^n|Y = y_0] = \sum_x X^n \cdot P[X = x|Y = y_0]$

$E[X^n|Y = y_0] = \int_{-\infty}^{\infty} X^n \cdot f_{X|Y=y_0}(x)dx$

Casos particulares:

- Momento condicionado no centrado de orden uno: Esperanza condicionada
- Momento condicionado centrado de orden dos: Varianza condicionada:
  $Var[X|Y] = E[(X - E[X|Y])^2|Y] = E[X^2|Y] - E[X|Y]^2$

Teorema de descomposición de la varianza.

Si $\exists\ E[X^2] \Rightarrow \exists\ Var[E[X|Y]]$ y $E[Var[X|Y]]$ y además:

$$Var[X] = Var[E[X|Y]] + E[Var[X|Y]]$$

Corolario

En las condiciones del teorema:

- $Var[X] \geq Var[E[X|Y]]$
- $Var[X] \geq E[Var[X|Y]]$