TEMA 2

CARACTERÍSTICAS DE VARIABLES ALEATORIAS MULTIDIMENSIONALES

Esperanza matemática de vectores aleatorios

Definición:
Sea $X = (X_1, \ldots, X_n) : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}^n, \mathcal{B}^n, P_X)$ un vector aleatorio. Diremos que existe la esperanza de X si existe la esperanza de cada componente X_i y, en tal caso, se define

$$\mathrm{E}X = (\mathrm{E}X_1, \ldots, \mathrm{E}X_n)$$

también llamado **vector de medias** del vector aleatorio X.
La definición no supone ningún concepto adicional ya que la verificación de existencia y el cálculo de esperanzas de vectores se reduce al de variables unidimensionales. También, como en el caso unidimensional, puede calcularse la esperanza de cualquier función del vector a partir de su distribución.

Esperanza de una función (unidimensional) de un vector aleatorio

Sea $X = (X_1, \ldots, X_n) : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}^n, \mathcal{B}^n, P_X)$ y $g : (\mathbb{R}^n, \mathcal{B}^n) \longrightarrow (\mathbb{R}, \mathcal{B})$ medible, de forma que $g(X_1, \ldots, X_n)$ es una variable aleatoria unidimensional.

- Si X es de tipo discreto, existe la esperanza de $g(X) = g(X_1, \ldots, X_n)$ si y sólo si

  $$\sum_{x_1,\ldots,x_n} |g(x_1, \ldots, x_n)| P\{X_1 = x_1, \ldots, X_n = x_n\} < \infty$$

  y en caso de existir

  $$\mathrm{E}[g(X_1, \ldots, X_n)] = \sum_{x_1,\ldots,x_n} g(x_1, \ldots, x_n) P\{X_1 = x_1, \ldots, X_n = x_n\}$$

- Si X es de tipo continuo, existe la esperanza de $g(X) = g(X_1, \ldots, X_n)$ si y sólo si

  $$\int_{\mathbb{R}^n} |g(x_1, \ldots, x_n)| f_X(x_1, \ldots, x_n) dx_1 \cdots dx_n < \infty$$

  y en caso de existir

  $$\mathrm{E}[g(X_1, \ldots, X_n)] = \int_{\mathbb{R}^n} g(x_1, \ldots, x_n) f_X(x_1, \ldots, x_n) dx_1 \cdots dx_n$$

Nota: Según el resultado anterior, la esperanza de cada X_i puede calcularse bien a partir de la conjunta, o bien a partir de la marginal.

$$\mathrm{E}X_i = \begin{cases} \sum_{x_i} x_i P\{X_i = x_i\} = \sum_{x_1,\dots,x_n} x_i P\{X_1 = x_1, \dots, X_n = x_n\} & \text{discreto} \\ \int_{\mathbb{R}} x_i f_{X_i}(x_i) dx_i = \int_{\mathbb{R}^n} x_i f_X(x_1, \dots, x_n) dx_1 \cdots dx_n & \text{continuo} \end{cases}$$

Esta propiedad permita generalizar algunas propiedades de la esperanza que se estudiaron en el caso unidimensional, como la linealidad y la conservación del orden.

Propiedades

Teorema 1
Si $X_1, \dots, X_n$ son variables aleatorias unidimensionales sobre $(\Omega, \mathcal{A}, P)$ tales que $\exists EX_i, \ \forall i = 1, 2, \dots, n$, entonces

$$\exists E[a_1 X_1 + \cdots + a_n X_n] = \sum_{i=1}^n a_i \, EX_i, \quad \forall a_1, \dots, a_n \in \mathbb{R}.$$

Dem.- Lo probamos en el caso continuo, y en el discreto es análogo.

- $\exists E\left[\sum_{i=1}^n a_i X_i\right] \iff \int_{\mathbb{R}^n} \left|\sum_{i=1}^n a_i x_i\right| f_X(x_1, \dots, x_n) dx_1 \cdots dx_n < \infty$

$$\int_{\mathbb{R}^n} \left|\sum_{i=1}^n a_i x_i\right| f_X(x_1, \dots, x_n) dx_1 \cdots dx_n \le \int_{\mathbb{R}^n} \sum_{i=1}^n |a_i| \, |x_i| \, f_X(x_1, \dots, x_n) dx_1 \cdots dx_n$$
$$= \sum_{i=1}^n |a_i| \int_{\mathbb{R}^n} |x_i| \, f_X(x_1, \dots, x_n) dx_1 \cdots dx_n = \sum_{i=1}^n a_i E|X_i| < \infty$$

- $E\left[\sum_{i=1}^n a_i X_i\right] = \int_{\mathbb{R}^n} \left(\sum_{i=1}^n a_i x_i\right) f_X(x_1, \dots, x_n) dx_1 \cdots dx_n$

$$= \sum_{i=1}^n a_i \int_{\mathbb{R}^n} x_i f_X(x_1, \dots, x_n) dx_1 \cdots dx_n = \sum_{i=1}^n a_i \, EX_i$$

Teorema 2: Conservación del orden
Si X_1 y X_2 son variables aleatorias unidimensionales sobre $(\Omega, \mathcal{A}, P)$ tales que $X_1 \le X_2$, entonces $EX_1 \le EX_2$ (siempre que $\exists EX_1, EX_2$).

Dem.- Igual que antes, suponemos que (X_1, X_2) es continuo y $f_X(x_1, x_2)$ su función de densidad. Notando que, puesto que $X_1 \le X_2$, $f_X(x_1, x_2) = 0$ si $x_1 > x_2$, se tiene

$$EX_1 = \int_{\{(x_1,x_2)/x_1 \le x_2\}} x_1 f_X(x_1, x_2) dx_1 dx_2 \le \int_{\{(x_1,x_2)/x_1 \le x_2\}} x_2 f_X(x_1, x_2) dx_1 dx_2 = EX_2$$

Teorema 3: Teorema de multiplicación

Si $X_1, \dots, X_n$ son independientes y $\exists EX_i, \ \forall i = 1, \dots, n$, entonces

$$\exists E[X_1 \cdots X_n] = E[X_1] \cdots E[X_n]$$

Dem

Lo probamos sólo en el caso continuo, el discreto es análogo.

a) $\exists E[X_1 \cdots X_n] \;\Leftrightarrow\; \int_{\mathbb{R}^n} |x_1 \cdots x_n| f_X(x_1, \dots, x_n)\, dx_1 \cdots dx_n < \infty$

$$\int_{\mathbb{R}^n} |x_1 \cdots x_n| f_X(x_1, \dots, x_n)\, dx_1 \cdots dx_n = \int_{\mathbb{R}^n} |x_1 \cdots x_n| f_{X_1}(x_1) \cdots f_{X_n}(x_n)\, dx_1 \cdots dx_n$$

$$\int_{\mathbb{R}} |x_1| f_{X_1}(x_1)\, dx_1 \cdots \int_{\mathbb{R}} |x_n| f_{X_n}(x_n)\, dx_n < \infty \; (\exists EX_i,\ \forall i)$$

b) $E[X_1 \cdots X_n] = \int_{\mathbb{R}^n} x_1 \cdots x_n f_X(x_1, \dots, x_n)\, dx_1 \cdots dx_n$

$$= \int_{\mathbb{R}^n} x_1 \cdots x_n f_{X_1}(x_1) \cdots f_{X_n}(x_n)\, dx_1 \cdots dx_n$$

$$\int_{\mathbb{R}} x_1 f_{X_1}(x_1)\, dx_1 \cdots \int_{\mathbb{R}} x_n f_{X_n}(x_n)\, dx_n = E[X_1] \cdots E[X_n]$$

Momentos de variables aleatorias n-dimensionales

Momentos no centrados (centrados respecto al origen):

Sea $X = (X_1, \dots, X_n)$ un vector aleatorio y $h_1, \dots, h_n$ enteros no negativos. Se define el momento no centrado de orden $(h_1, \dots, h_n)$ como

$$E[X_1^{h_1} \cdots X_n^{h_n}]$$

siempre que exista.

$$m_{h_1,\dots,h_n} = E[X_1^{h_1} \cdots X_n^{h_n}] = \begin{cases} \sum_{x_1,\dots,x_n} x_1^{h_1} \cdots x_n^{h_n} P[X_1 = x_1, \dots, X_n = x_n] \\ \\ \int_{\mathbb{R}^n} x_1^{h_1} \cdots x_n^{h_n} f_X(x_1, \dots, x_n) dx_1 \cdots dx_n \end{cases}$$

Notemos que si se hace $h_1 = \cdots = h_{i-1} = h_{i+1} = \cdots = h_n = 0$, se obtienen los momentos no centrados de la variable X_i ($\forall i = 1, \dots, n$) (suelen denominarse momentos marginales).

Momentos centrados (respecto a las medias):

Sea $X = (X_1, \dots, X_n)$ un vector aleatorio y $h_1, \dots, h_n$ enteros no negativos. Si $\exists EX_i,\ \forall i = 1, \dots, n$, se define el momento centrado de X de orden $(h_1, \dots, h_n)$ como

$$E[(X_1 - EX_1)^{h_1} \cdots (X_n - EX_n)^{h_n}]$$

siempre que exista.

$$\mu_{h_1,\dots,h_n} = E[(X_1 - EX_1)^{h_1} \cdots (X_n - EX_n)^{h_n}]$$

Casos particulares: Vectores bidimensionales:

$$\begin{array}{lll} m_{10} = E[X_1] & m_{01} = E[X_2] & \\ m_{20} = E[X_1^2] & m_{02} = E[X_2^2] & m_{11} = E[X_1X_2] \\ \mu_{10} = 0 & \mu_{01} = 0 & \\ \mu_{20} = Var[X_1] & \mu_{02} = Var[X_2] & \mu_{11} = E[(X_1 - EX_1)(X_2 - EX_2)] \end{array}$$

A $\mu_{11} = E[(X_1 - EX_1)(X_2 - EX_2)]$ se le denomina covarianza de X_1 y X_2

$$Cov(X_1, X_2) = \mu_{11} = E[(X_1 - EX_1)(X_2 - EX_2)]$$

Para un vector aleatorio de dimensión arbitraria se define su **matriz de covarianzas** como

$$\Sigma_X = ((cov(X_i, X_j)))_{i,j=1,\dots,n}$$

siempre que existan todas las covarianzas. Esta matriz es simétrica y, en el caso bidimensional,

$$\Sigma_X = \begin{pmatrix} Var(X_1) & cov(X_1, X_2) \\ cov(X_1, X_2) & Var(X_2) \end{pmatrix}$$

Propiedades de los momentos de segundo orden

1) Si $\exists EX^2$ y EY^2, entonces $\exists E[X \cdot Y]$.

2) Si $\exists E[X \cdot Y]$ entonces $\exists cov(X, Y) = E[XY] - EXEY$.

Dem

$$cov(X, Y) = E[(X - EX)(Y - EY)] = E[XY - XEY - YEX + (EX)(EY)] = E[XY] - (EX)(EY)$$

3) Sean $X_1, X_2, \dots, X_n$ variables aleatorias definidas sobre el mismo espacio de probabilidad tales que $\exists E[X_i^2]$, $\forall i = 1, 2, \dots, n$. Entonces, $\forall a_1, \dots, a_n \in \mathbb{R}$

$$\exists Var\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 Var X_i + \sum_{\substack{i,j=1 \\ i \neq j}}^n a_i a_j cov(X_i, X_j)$$

Dem

$$Var\left(\sum_{i=1}^n a_i X_i\right) = E\left[\left(\sum_{i=1}^n a_i(X_i - EX_i)\right)^2\right] =$$

$$E\left[\sum_{i=1}^n a_i^2 (X_i - EX_i)^2 + \sum_{\substack{i,j=1 \\ i \neq j}}^n a_i a_j (X_i - EX_i)(X_j - EX_j)\right]$$

La esperanza anterior existe por existir la esperanza de cada sumando por 1) y 2) y es igual a la expresión dada.

4) $\exists cov(X, Y) \;\Rightarrow\; \exists cov(aX + b, cY + d) = ac\, cov(X, Y), \quad \forall a, b, c, d \in \mathbb{R}$

Dem

cov(aX+b,cY+d)=E[(aX+b-aEX-b)(cY+d-cEY-d)]=acE[(X-EX)(Y-EY)]

Como consecuencia inmediata del Teorema de Multiplicación de esperanzas se tiene

5) X e Y independientes y $\exists cov(X,Y)$, entonces $cov(X,Y)=0$.

Dem

$E[XY]=(EX)(EY) \;\Rightarrow\; cov(X,Y)=E[XY]-(EX)(EY)=0$

El recíproco de este resultado no es cierto, existen variables no independientes con covarianza nula.

Ejemplo: Sea X una variable aleatoria tal que $EX=EX^3=0$ (por ejemplo una $\mathcal{N}(0,1)$) y sea $Y=X^2$. Es evidente que X e Y no son independientes y, sin embargo $E[XY]=E[X^3]=0$ y $(EX)(EY)=0$. Por tanto, $cov(X,Y)=0$.

6) Si $X_1,\ldots,X_n$ son independientes y $\exists EX_i^2,\ \forall i=1,\ldots,n$, entonces

$$\forall a_1,\ldots,a_n\in\mathbb{R},\ \exists Var(a_1X_1+\ldots a_nX_n)=a_1^2VarX_1+\ldots+a_n^2VarX_n$$

Por tanto, la varianza de sumas de variables aleatorias independientes (o de diferencias) es la suma de las varianzas.

Desigualdad de Cauchy-Schwarz

Teorema

Sean X,Y variables aleatorias tales que $\exists EX^2,\ EY^2$, entonces

- $(E[XY])^2\le E[X^2]E[Y^2]$
- Si X o Y es degenerada en cero o las dos son degeneradas, se da la igualdad.
- Si X e Y son no degeneradas, se da la igualdad si y sólo si

$$\exists a,b\in\mathbb{R}, \text{ambos no nulos, tal que } P[aX+bY=0]=1$$

La desigualdad de Cauchy-Schwarz da lugar a la siguiente expresión para los momentos centrados:

Corolario

Sean X,Y variables aleatorias tales que $\exists EX^2,\ EY^2$, entonces

- $(cov(X,Y))^2\le(VarX)(VarY)$

- Si una de las variables es degenerada, se da la igualdad.
- Si las variables son no degeneradas, se da la igualdad si y sólo si

$$\exists a \neq 0, b \neq 0, c \in \mathbb{R} \text{ tal que } P[aX + bY = c] = 1$$

Función generatriz de momentos

Definición

Sea $X = (X_1, \ldots, X_n)$ un vector aleatorio tal que $\exists E\left[\mathrm{e}^{t_1 X_1 + \cdots t_n X_n}\right], \forall t_1, \ldots, t_n \,/\, t_i \in (-t_0^i, t_1^i)$ con $t_0^i, t_1^i \in \mathbb{R}^+$. Se define la función generatriz de momentos de X como

$$M_X : (-t_0^1, t_1^1) \times \cdots \times (-t_0^n, t_1^n) \longrightarrow \mathbb{R}^+$$

dada por

$$M_X(t_1, \ldots, t_n) = E\left[\mathrm{e}^{t_1 X_1 + \cdots t_n X_n}\right]$$

La función generatriz de momentos que, como en el caso unidimensional, puede no existir, tiene propiedades análogas a las de tal caso. Destacamos, como allí, el Teorema de Unicidad y la Relación con los momentos.

Teorema de unicidad

La función generatriz de momentos de un vector aleatorio, si existe, es única y determina unívocamente su distribución (y, por tanto las marginales).

Relación entre f.g.m. y momentos

Si $\exists M_{(X_1,\ldots,X_n)}$, existen todos los momentos y se pueden calcular como

$$E\left[X_1^{h_1} \cdot \cdots \cdot X_n^{h_n}\right] = \left.\frac{\partial^{h_1 + \cdots h_n} M_{(X_1,\ldots,X_n)}(t_1, \ldots, t_n)}{\partial t_1^{h_1} \cdots \partial t_n^{h_n}}\right|_{t_1=0,\ldots,t_n=0}$$

A continuación probamos que si existe la fgm de un vector existe la de cualquier subvector.

Teorema

Si $X = (X_1, \ldots, X_n)$ tiene función generatriz de momentos $M_X(t_1, \ldots, t_n)$, existe la fgm de cada subvector (y, por tanto, de cada componente) y puede calcularse a partir de la conjunta por

$$M_{(X_{i_1},\ldots,X_{i_k})}(t_{i_1}, \ldots, t_{i_k}) = M_{(X_1,\ldots,X_n)}(0, \ldots, 0, t_{i_1}, 0, \ldots, 0, t_{i_k}, 0, \ldots, 0)$$

para $t_{i_1}, \ldots, t_{i_k}$ en los intervalos correspondientes.

Dem

$$M_{(X_{i_1},\ldots,X_{i_k})}(t_{i_1},\ldots,t_{i_k}) = E\left[\mathrm{e}^{t_{i_1}X_{i_1}+\cdots+t_{i_k}X_{i_k}}\right] =$$

$$E\left[\mathrm{e}^{0\cdot X_1+\cdots+0\cdot X_{i_1-1}+t_{i_1}X_{i_1}+0\cdot X_{i_1+1}+\cdots+0\cdot X_{i_k-1}+t_{i_k}X_{i_k}+0\cdot X_{i_k+1}+\cdots+0\cdot X_n}\right]$$

Ejemplo

Sea (X,Y) un vector aleatorio con función de densidad

$$f(x,y) = \mathrm{e}^{-x-y}, \quad x>0,\ y>0$$

Calcular la función generatriz de momentos del vector, la de cada componente, las medias, varianzas y covarianza.

$$\begin{aligned} M_{(X,Y)}(t_1,t_2) &= E[\mathrm{e}^{t_1X+t_2Y}] = \int_0^\infty\int_0^\infty \mathrm{e}^{t_1x+t_2y}\mathrm{e}^{-x-y}dx\,dy \\ &= \left(\int_0^\infty \mathrm{e}^{t_1x}\mathrm{e}^{-x}dx\right)\left(\int_0^\infty \mathrm{e}^{t_2y}\mathrm{e}^{-y}dy\right) = \left(\int_0^\infty \mathrm{e}^{x(t_1-1)}dx\right)\left(\int_0^\infty \mathrm{e}^{y(t_2-1)}dy\right) \\ &= \frac{1}{1-t_1}\frac{1}{1-t_2}, \quad t_1<1,\ t_2<1 \end{aligned}$$

$$M_X(t_1) = M_{(X,Y)}(t_1,0) = \frac{1}{1-t_1}, \quad t_1<1$$

$$M_Y(t_2) = M_{(X,Y)}(0,t_2) = \frac{1}{1-t_2}, \quad t_2<1$$

$$EX = \left.\frac{\partial M(t_1,t_2)}{\partial t_1}\right|_{t_1=t_2=0} = \left.\frac{1}{1-t_2}\ \frac{1}{(1-t_1)^2}\right|_{t_1=t_2=0} = 1$$

$$EY = \left.\frac{\partial M(t_1,t_2)}{\partial t_2}\right|_{t_1=t_2=0} = \left.\frac{1}{1-t_1}\ \frac{1}{(1-t_2)^2}\right|_{t_1=t_2=0} = 1$$

$$EX^2 = \left.\frac{\partial^2 M(t_1,t_2)}{\partial t_1^2}\right|_{t_1=t_2=0} = \left.\frac{1}{1-t_2}\ \frac{2}{(1-t_1)^3}\right|_{t_1=t_2=0} = 2 = EY^2$$

$$VarX = VarY = 1$$

$$E[XY] = \left.\frac{\partial^2 M(t_1,t_2)}{\partial t_1\,\partial t_2}\right|_{t_1=t_2=0} = \frac{\partial}{\partial t_2}\left[\frac{1}{1-t_2}\frac{1}{(1-t_1)^2}\right]_{t_1=t_2=0} = \left.\frac{1}{(1-t_1)^2}\frac{1}{(1-t_2)^2}\right|_{t_1=t_2=0} = 1$$

$$cov(X,Y) = E[XY]-(EX)(EY) = 0$$

A continuación establecemos una nueva caracterización de independencia en términos de funciones generatrices de momentos y que se obtiene como consecuencia del Teorema de Multiplicación.

Caracterización de independencia por fgm

Sea $X = (X_1, \dots, X_n)$ un vector aleatorio tal que $\exists M_{X_i}(t_i),\ t_i \in I_i,\ i = 1, \dots, n$, siendo I_i un intervalo abierto que contiene al 0, $\forall i = 1, \dots, n$.

$$X_1, \dots, X_n \text{ son independientes } \Leftrightarrow \exists M_X(t_1, \dots, t_n) = M_{X_1}(t_1) \cdots M_{X_n}(t_n)\ \forall (t_1, \dots, t_n) \in I_1 \times \cdots \times I_n.$$

Dem

$\Rightarrow]$ $X_1, \dots, X_n$ independientes $\Rightarrow$ $\mathrm{e}^{t_1 X_1}, \dots, \mathrm{e}^{t_n X_n}$ independientes y además $\exists E\left[\mathrm{e}^{t_i X_i}\right],\ i = 1, \dots, n$. Aplicando el Teorema de Multiplicación

$$\exists E\left[\prod_{i=1}^{n} \mathrm{e}^{t_i X_i}\right] = \prod_{i=1}^{n} E\left[\mathrm{e}^{t_i X_i}\right]$$

$\Leftarrow]$ Lo probamos en el caso continuo y el discreto se haría de forma análoga.

$$\int_{\mathbb{R}^n} \mathrm{e}^{t_1 x_1 + \cdots t_n x_n} f_X(x_1, \dots, x_n)\, dx_1 \cdots dx_n = \left(\int_{\mathbb{R}} \mathrm{e}^{t_1 x_1} f_{X_1}(x_1)\, dx_1\right) \cdots \left(\int_{\mathbb{R}} \mathrm{e}^{t_n x_n} f_{X_n}(x_n)\, dx_n\right)$$

$$\int_{\mathbb{R}^n} \mathrm{e}^{t_1 x_1 + \cdots t_n x_n} f_{X_1}(x_1) \cdots f_{X_n}(x_n)\, dx_1 \cdots dx_n, \quad \forall t_1, \dots, t_n$$

Entonces $f_X(x_1, \dots, x_n)$ y $f_{X_1}(x_1) \cdots f_{X_n}(x_n)$ son dos funciones de densidad sobre $\mathbb{R}^n$ a las que corresponde la misma fgm. Por el teorema de unicidad de la fgm dichas densidades coinciden y, por tanto, $X_1, \dots, X_n$ son independientes.

Reproductividad de distribuciones

Ver presentación multimedia.

Esperanza condicionada

Variables aleatorias unidimensionales.

Definición.- Sean X e Y variables aleatorias unidimensionales definidas en el mismo espacio de probabilidad $(\Omega, \mathcal{A}, P)$ y supongamos que $\exists \mathrm{E}X$. Se define la **esperanza condicionada de** X **dada** Y (o esperanza de X condicionada a Y), y se nota por $\mathrm{E}[X/Y]$ como la variable aleatoria que toma el valor $\mathrm{E}[X/Y = y]$ cuando la variable aleatoria Y toma el valor y, donde $\mathrm{E}[X/Y = y]$ se define a su vez como

$$
\mathrm{E}[X/Y=y]=\begin{cases}\displaystyle\sum_x x\mathrm{P}[X=x/Y=y] & \text{si } (X,Y) \text{ es de tipo discreto} \\ & \text{y } \mathrm{P}[Y=y]>0 \\ \displaystyle\int_{-\infty}^{+\infty} x f_{X/Y=y}(x)dx & \text{si } (X,Y) \text{ es de tipo continuo} \\ & \text{y } f_2(y)>0\end{cases}
$$

Notas.-

1) $\mathrm{E}[X/Y]$ es una función de Y.

2) $\mathrm{E}[X/Y=y]$ no es más que la media de la variable aleatoria X considerando como distribución la condicionada de X dado $Y=y$.

3) De forma similar se puede definir la esperanza de Y dada X, supuesta la existencia de la $\mathrm{E}Y$.

Esperanza condicionada de una función medible

Sean X e Y variables aleatorias unidimensionales definidas en el mismo espacio de probabilidad $(\Omega,\mathcal{A},P)$ y sea $h:(\mathbb{R},\mathcal{B})\longrightarrow(\mathbb{R},\mathcal{B})$ una función medible tal que $\exists \mathrm{E}h(X)$. Se define la **esperanza condicionada de** $h(X)$ **dado** Y, y se nota $\mathrm{E}[h(X)/Y]$, como la variable aleatoria que toma el valor $\mathrm{E}[h(X)/Y=y]$ cuando la variable aleatoria Y toma el valor y, donde $\mathrm{E}[h(X)/Y=y]$ se define a su vez como

$$
\mathrm{E}[h(X)/Y=y]=\begin{cases}\displaystyle\sum_x h(x)\mathrm{P}[X=x/Y=y] & \text{si } (X,Y) \text{ es de tipo discreto} \\ & \text{y } \mathrm{P}[Y=y]>0 \\ \displaystyle\int_{-\infty}^{+\infty} h(x) f_{X/Y=y}(x)dx & \text{si } (X,Y) \text{ es de tipo continuo} \\ & \text{y } f_2(y)>0\end{cases}
$$

Nota.-
De forma similar se definiría la $\mathrm{E}[h(Y)/X]$.

Ejemplos

1.- *Una urna contiene tres bolas rojas y dos blancas. Se extrae aleatoriamente una muestra de dos bolas.*
a) Con reemplazamiento.
b) Sin reemplazamiento.
Sea $X=0$ si la primera bola es blanca y $X=1$ si es roja. Sea $Y=0$ si la segunda bola es blanca e $Y=1$ si es roja.
Calcular $E[X/Y]$ y $E[Y/X]$.

a) Extracción con reemplazamiento. La distribución conjunta de X e Y es

Y \ X	0	1	
0	4/25	6/25	10/25
1	6/25	9/25	15/25
	10/25	15/25	

$X/Y = 0$

$$\mathrm{P}[X = 0/Y = 0] = \frac{\mathrm{P}[X = 0, Y = 0]}{\mathrm{P}[Y = 0]} = \frac{4/25}{10/25} = \frac{2}{5}$$

$$\mathrm{P}[X = 1/Y = 0] = \frac{\mathrm{P}[X = 1, Y = 0]}{\mathrm{P}[Y = 0]} = \frac{6/25}{10/25} = \frac{3}{5}$$

$$\mathrm{E}[X/Y = 0] = 0\frac{2}{5} + 1\frac{3}{5} = \frac{3}{5}$$

$X/Y = 1$

$$\mathrm{P}[X = 0/Y = 1] = \frac{\mathrm{P}[X = 0, Y = 1]}{\mathrm{P}[Y = 1]} = \frac{6/25}{15/25} = \frac{2}{5}$$

$$\mathrm{P}[X = 1/Y = 1] = \frac{\mathrm{P}[X = 1, Y = 1]}{\mathrm{P}[Y = 1]} = \frac{9/25}{15/25} = \frac{3}{5}$$

$$\mathrm{E}[X/Y = 1] = 0\frac{2}{5} + 1\frac{3}{5} = \frac{3}{5}$$

Como $\mathrm{E}[X/Y]$ es una variable aleatoria que toma el valor $\mathrm{E}[X/Y = 0]$ si $Y = 0$ y el valor $\mathrm{E}[X/Y = 1]$ si $Y = 1$, en este caso $\mathrm{E}[X/Y]$ es la variable aleatoria degenerada en el punto $3/5$.

$Y/X = 0$

$$\mathrm{P}[Y = 0/X = 0] = \frac{\mathrm{P}[X = 0, Y = 0]}{\mathrm{P}[X = 0]} = \frac{4/25}{10/25} = \frac{2}{5}$$

$$\mathrm{P}[Y = 1/X = 0] = \frac{\mathrm{P}[X = 0, Y = 1]}{\mathrm{P}[X = 0]} = \frac{6/25}{10/25} = \frac{3}{5}$$

$$\mathrm{E}[Y/X = 0] = 0\frac{2}{5} + 1\frac{3}{5} = \frac{3}{5}$$

$Y/X = 1$

$$\mathrm{P}[Y = 0/X = 1] = \frac{\mathrm{P}[X = 1, Y = 0]}{\mathrm{P}[X = 1]} = \frac{6/25}{15/25} = \frac{2}{5}$$

$$\mathrm{P}[Y = 1/X = 1] = \frac{\mathrm{P}[X = 1, Y = 1]}{\mathrm{P}[X = 1]} = \frac{9/25}{15/25} = \frac{3}{5}$$

$$\mathrm{E}[Y/X=1] = 0\frac{2}{5} + 1\frac{3}{5} = \frac{3}{5}$$

En este caso $\mathrm{E}[Y/X]$ es también la variable aleatoria degenerada en el punto $3/5$.

b) Extracción sin reemplazamiento. La distribución conjunta de X e Y es

$Y \backslash X$	0	1	
0	2/20	6/20	8/20
1	6/20	6/20	12/20
	8/20	12/20	

$X/Y=0$

$$\mathrm{P}[X=0/Y=0] = \frac{\mathrm{P}[X=0,Y=0]}{\mathrm{P}[Y=0]} = \frac{2/20}{8/20} = \frac{1}{4}$$

$$\mathrm{P}[X=1/Y=0] = \frac{\mathrm{P}[X=1,Y=0]}{\mathrm{P}[Y=0]} = \frac{6/20}{8/20} = \frac{3}{4}$$

$$\mathrm{E}[X/Y=0] = 0\frac{1}{4} + 1\frac{3}{4} = \frac{3}{4}$$

$X/Y=1$

$$\mathrm{P}[X=0/Y=1] = \frac{\mathrm{P}[X=0,Y=1]}{\mathrm{P}[Y=1]} = \frac{6/20}{12/20} = \frac{1}{2}$$

$$\mathrm{P}[X=1/Y=1] = \frac{\mathrm{P}[X=1,Y=1]}{\mathrm{P}[Y=1]} = \frac{6/20}{12/20} = \frac{1}{2}$$

$$\mathrm{E}[X/Y=1] = 0\frac{1}{2} + 1\frac{1}{2} = \frac{1}{2}$$

$\mathrm{E}[X/Y]$ es una variable aleatoria que toma el valor $\mathrm{E}[X/Y=0]=3/4$ si $Y=0$ (con probabilidad $8/20$) y el valor $\mathrm{E}[X/Y=1]$ si $Y=1$ (con probabilidad $12/20$).

$Y/X=0$

$$\mathrm{P}[Y=0/X=0] = \frac{\mathrm{P}[X=0,Y=0]}{\mathrm{P}[X=0]} = \frac{2/20}{8/20} = \frac{1}{4}$$

$$\mathrm{P}[Y=1/X=0] = \frac{\mathrm{P}[X=0,Y=1]}{\mathrm{P}[X=0]} = \frac{6/20}{8/20} = \frac{3}{4}$$

$$\mathrm{E}[Y/X=0] = 0\frac{1}{4} + 1\frac{3}{4} = \frac{3}{4}$$

$Y/X=1$

$$P[Y=0/X=1]=\frac{P[X=1,Y=0]}{P[X=1]}=\frac{6/20}{12/20}=\frac{1}{2}$$

$$P[Y=1/X=1]=\frac{P[X=1,Y=1]}{P[X=1]}=\frac{6/20}{12/20}=\frac{1}{2}$$

$$E[Y/X=1]=0\frac{1}{2}+1\frac{1}{2}=\frac{1}{2}$$

En este caso E$[Y/X]$ coincide con E[X/Y].

2.- *Sea (X,Y) un vector aleatorio con función de densidad*

$$f(x,y)=2 \qquad 0<x<y<1$$

Calcular E$[X/Y]$ y E$[Y/X]$.

- E$[X/Y]$ es la variable aleatoria que toma los valores E$[X/Y=y_0]$ si $Y=y_0$

$$E[X/Y=y_0]=\int x f_{X/Y=y_0}(x)dx$$

Calculemos previamente la distribución condicionada

$$f_2(y)=\int_0^y 2dx=2y \qquad 0<y<1$$

Luego, si $0<y_0<1$

$$fX/Y=y_0(x)=\frac{2}{2y_0}=\frac{1}{y_0} \qquad 0<x<y_0$$

Por tanto, si $0<y_0<1$

$$E[X/Y=y_0]=\int_0^{y_0} x\frac{1}{y_0}dx=\frac{1}{y_0}\frac{x^2}{2}\Big|_0^{y_0}=\frac{y_0^2}{2y_0}=\frac{y_0}{2}$$

Y, en consecuencia

$$E[X/Y]=\frac{Y}{2}$$

- E$[Y/X]$ es la variable aleatoria que toma los valores E$[Y/X=x_0]$ si $X=x_0$

$$E[Y/X=x_0]=\int y f_{Y/X=x_0}(y)dy$$

Calculemos previamente la distribución condicionada

$$f_1(x) = \int_x^1 2dy = 2(1-x) \qquad 0 < x < 1$$

Luego, si $0 < x_0 < 1$

$$f_{Y/X=x_0}(y) = \frac{2}{2(1-x_0)} = \frac{1}{1-x_0} \qquad x_0 < y < 1$$

Por tanto, si $0 < x_0 < 1$

$$\mathrm{E}[Y/X = x_0] = \int_{x_0}^1 y \frac{1}{1-x_0} dy = \frac{1}{1-x_0} \frac{y^2}{2}\bigg|_{x_0}^1 = \frac{1-x_0^2}{2(1-x_0)} = \frac{1+x_0}{2}$$

Y, en consecuencia

$$\mathrm{E}[Y/X] = \frac{1+X}{2}$$

3.- *Sea (X,Y) un vector aleatorio con función de densidad*

$$f(x,y) = 1 \qquad (x,y) \in T(0,0),(2,0),(1,1)$$

Calcular $E[Y/X]$

Marginal de X

$$f_1(x) = \begin{cases} \displaystyle\int_0^x dy = x & 0 < x < 1 \\ \displaystyle\int_0^{2-x} dy = 2-x & 1 < x < 2 \end{cases}$$

Condicionada $Y/X = x_0$

Si $0 < x_0 < 1$, $\quad f_{Y/X=x_0}(y) = \dfrac{1}{x_0} \qquad 0 < y < x_0$

Si $1 < x_0 < 2$, $\quad f_{Y/X=x_0}(y) = \dfrac{1}{2-x_0} \qquad 0 < y < 2-x_0$

$\mathrm{E}[Y/X = x_0]$

$$\text{Si } 0 < x_0 < 1, \quad \mathrm{E}[Y/X = x_0] = \int_0^{x_0} y \frac{1}{x_0} = \frac{1}{x_0} \frac{y^2}{2} \bigg|_0^{x_0} = \frac{x_0}{2}$$

$$\text{Si } 1 < x_0 < 2, \quad \mathrm{E}[Y/X = x_0] = \int_0^{2-x_0} y \frac{1}{2-x_0} = \frac{1}{2-x_0} \frac{y^2}{2} \bigg|_0^{2-x_0} = \frac{2-x_0}{2}$$

Luego

$$\mathrm{E}[Y/X] = \begin{cases} X/2 & 0 < X < 1 \\ \\ (2-X)/2 & 1 < X < 2 \end{cases}$$

Propiedades de la esperanza condicionada

Dado que los valores de la esperanza condicionada son esperanzas, las propiedades son análogas.

1) $\mathrm{E}[c/Y] = c, \quad c \in \mathbb{R}$.

2) Sean X e Y variables aleatorias definidas en el mismo espacio de probabilidad tal que $\exists \mathrm{E}[X]$ y $a, b \in \mathbb{R}$, entonces $\exists \mathrm{E}[aX + b/Y] = a\mathrm{E}[X/Y] + b$.

3) Sean $X_1, X_2, \ldots, X_n$ variables aleatorias tales que $\exists \mathrm{E}X_i,\ i = 1, 2, \ldots, n$, entonces $\forall a_1, a_2, \ldots, a_n \in \mathbb{R}$, $\exists \mathrm{E}[(a_1X_1 + a_2X_2 + \ldots + a_nX_n/Y] = a_1\mathrm{E}[X_1/Y] + a_2\mathrm{E}[X_2/Y] + \ldots + a_n\mathrm{E}[X_n/Y]$.

4) Si $X \geq 0$ y $\exists \mathrm{E}X$, entonces $\mathrm{E}[X/Y] \geq 0$.

5) Si X_1 y X_2 son variables aleatorias tales que $\exists \mathrm{E}X_1, \mathrm{E}X_2$ y $X_1 \leq X_2$, entonces $\mathrm{E}[X_1/Y] \leq \mathrm{E}[X_2/Y]$

Teorema.

Si X e Y son independientes y $\exists \mathrm{E}[h(X)]$ siendo h una transformación medible, entonces

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

En particular, $\mathrm{E}[X/Y] = \mathrm{E}[X]$

Demostración.-

Es debida a que en el caso de independencia, la distribución de X condicionada a Y coincide con la marginal de X. (Comentar ejemplo 1)

Dado que E[X/Y] es una variable aleatoria (función de Y), podemos considerar su esperanza y se obtiene el siguiente resultado:

Teorema.

Si X e Y son variables aleatorias definidas en el mismo espacio de probabilidad tal que $\exists \mathrm{E}[h(X)]$ siendo h una transformación medible, se tiene

$$\exists \mathrm{E}\left[\mathrm{E}[h(X)/Y]\right] = \mathrm{E}h(X)$$

En particular, $\mathrm{E}\left[\mathrm{E}[X/Y]\right] = \mathrm{E}X$

EJEMPLOS

1.- *Para el ejemplo 3 anterior*

$$\mathrm{E}Y = \int_0^1 \int_y^{2-y} y dx dy = \int_0^1 y(2-2y)dy = 2\int_0^1 (y-y^2)dy =$$

$$= 2\left(\frac{y^2}{2} - \frac{y^3}{3}\right)\Big|_0^1 = 2\left(\frac{1}{2} - \frac{1}{3}\right) = 2\frac{1}{6} = \frac{1}{3}$$

Por otro lado

$$\mathrm{E}[\mathrm{E}[Y/X]] = \int_0^1 \frac{x}{2} x dx + \int_1^2 \frac{2-x}{2}(2-x)dx = \frac{x^3}{6}\Big|_0^1 - \frac{(2-x)^3}{6}\Big|_1^2 = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

Momentos condicionados

Sean X e Y variables aleatorias definidas sobre el mismo espacio de probabilidad y vamos a considerar la $\mathrm{E}[h(X)/Y]$ para funciones particulares h.

- Si $\exists \mathrm{E}X^n$ a $\mathrm{E}[X^n/Y]$ se le llama **momento condicionado no centrado de orden n de X dada Y**.
- Si $\exists \mathrm{E}X^n$ a $\mathrm{E}[(X - \mathrm{E}[X/Y])^n/Y]$ se le llama **momento condicionado centrado de orden n de X dada Y**. Observemos que $(X - \mathrm{E}[X/Y])^n$ es una función $h(X,Y)$ y de aquí que

$$\mathrm{E}[(X - \mathrm{E}[X/Y])^n/Y = y] = \begin{cases} \sum_x (x - \mathrm{E}[X/Y=y])^n \mathrm{P}[X = x/Y = y] \\ \int (x - \mathrm{E}[X/Y=y])^n f_{X/Y=y}(x)dx \end{cases}$$

CASOS PARTICULARES

- Momento condicionado no centrado de orden uno = **Esperanza condicionada**.
- Momento condicionado centrado de orden dos = **Varianza condicionada**

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

que se puede probar que verifica

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

En efecto,

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

$$\sum_x x^2 \mathrm{P}[X=x/Y=y] + (\mathrm{E}[X/Y=y])^2 \sum_x \mathrm{P}[X=x/Y=y] -$$

$$-2(\mathrm{E}[X/Y=y]) \sum_x x\mathrm{P}[X=x/Y=y] = \mathrm{E}[X^2/Y=y] - (\mathrm{E}[X/Y=y])^2$$

A continuación establecemos una importante propiedad de la varianza condicionada que se usará en el problema de regresión.
Teorema: Descomposición de la varianza.

Si $\exists \mathrm{E}X^2$ entonces $\exists \mathrm{Var}[\mathrm{E}[X/Y]]$ y $\exists \mathrm{E}[\mathrm{Var}[X/Y]]$ y además

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

Corolario.

En las condiciones del Teorema

$$\mathrm{Var}X \geq \mathrm{Var}[\mathrm{E}[X/Y]]$$

$$\mathrm{Var\ X} \geq \mathrm{E}[\mathrm{Var}[X/Y]]$$

dado que los dos términos del lado derecho de la expresión del Teorema son no negativos.

EJEMPLOS

1.- URNA. Consideramos sólo el caso de la extracción sin reemplazamiento ya que si las extracciones se hacen con reemplazamiento, las variables X e Y son independientes y $\mathrm{Var}(X/Y) = \mathrm{Var}X$ y $\mathrm{Var}(Y/X) = \mathrm{Var}Y$.

Extracción sin reemplazamiento. La distribución conjunta de X e Y es

Y \ X	0	1	
0	2/20	6/20	8/20
1	6/20	6/20	12/20
	8/20	12/20	

$\mathrm{Var}[X/Y]$ es una variable aleatoria que toma los valores $\mathrm{Var}[X/Y=0]$ si $Y=0$ y $\mathrm{Var}[X/Y=1]$ si $Y=1$

$X/Y = 0$

$$\mathrm{P}[X = 0/Y = 0] = \frac{1}{4}$$

$$\mathrm{P}[X = 1/Y = 0] = \frac{3}{4}$$

$$\mathrm{E}[X/Y = 0] = \frac{3}{4}$$

$$\mathrm{E}[X^2/Y = 0] = \frac{3}{4}$$

con lo que

$$\mathrm{Var}[X/Y = 0] = \frac{3}{4} - \left(\frac{3}{4}\right)^2 = \frac{3}{4} - \frac{9}{16} = \frac{3}{16}$$

$X/Y = 1$

$$\mathrm{P}[X = 0/Y = 1] = \frac{1}{2}$$

$$\mathrm{P}[X = 1/Y = 1] = \frac{1}{2}$$

$$\mathrm{E}[X/Y = 1] = \frac{1}{2}$$

$$\mathrm{E}[X^2/Y = 1] = \frac{1}{2}$$

con lo que

$$\mathrm{Var}[X/Y = 1] = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

Como $\mathrm{Var}[X/Y]$ es una variable aleatoria que toma el valor $\mathrm{Var}[X/Y = 0]$ si $Y = 0$ y el valor $\mathrm{Var}[X/Y = 1]$ si $Y = 1$, en este caso $\mathrm{Var}[X/Y]$ toma el valor $3/16$ si $Y = 0$ (con probabilidad $8/20$) y el valor $1/4$ si $Y = 1$ (con probabilidad $12/20$). En este caso

$$\mathrm{E}[\mathrm{Var}[X/Y]] = \frac{3}{16}\frac{8}{20} + \frac{1}{4}\frac{12}{20} = \frac{72}{320} = \frac{9}{40}$$

$Y/X = 0$

$$\mathrm{P}[Y = 0/X = 0] = \frac{1}{4}$$

$$\mathrm{P}[Y = 1/X = 0] = \frac{3}{4}$$

$$E[Y/X = 0] = \frac{3}{4}$$

$$E[Y^2/X = 0] = \frac{3}{4}$$

con lo que

$$\mathrm{Var}[Y/X = 0] = \frac{3}{16}$$

$Y/X = 1$

$$P[Y = 0/X = 1] = \frac{1}{2}$$

$$P[Y = 1/X = 1] = \frac{1}{2}$$

$$E[Y/X = 1] = \frac{1}{2}$$

$$E[Y^2/X = 1] = \frac{1}{2}$$

con lo que

$$\mathrm{Var}[Y/X = 1] = \frac{1}{4}$$

Como $\mathrm{Var}[Y/X]$ es una variable aleatoria que toma el valor $\mathrm{Var}[Y/X = 0]$ si $X = 0$ y el valor $\mathrm{Var}[Y/X = 1]$ si $X = 1$, en este caso $\mathrm{Var}[Y/X]$ toma el valor $3/16$ si $X = 0$ (con probabilidad $8/20$) y el valor $1/4$ si $X = 1$ (con probabilidad $12/20$). En este caso

$$E[\mathrm{Var}[Y/X]] = \frac{3}{16}\frac{8}{20} + \frac{1}{4}\frac{12}{20} = \frac{9}{40}$$

2.- Sea (X, Y) un vector aleatorio con función de densidad

$$f(x, y) = 2 \qquad 0 < x < y < 1$$

Calcular $\mathrm{Var}[X/Y]$, $\mathrm{Var}[Y/X]$, $E[\mathrm{Var}[X/Y]]$ y $E[\mathrm{Var}[Y/X]]$.

- X/Y

  Sabíamos que

$$f_2(y) = 2y \qquad 0 < y < 1$$

  y si $0 < y_0 < 1$

$$f_{X/Y=y_0}(x) = \frac{1}{y_0} \qquad 0 < x < y_0$$

$$\mathrm{E}[X/Y] = \frac{Y}{2}$$

Ahora, si $0 < y_0 < 1$

$$\mathrm{E}[X^2/Y = y_0] = \int_0^{y_0} x^2 \frac{1}{y_0} dx = \frac{1}{y_0} \left. \frac{x^3}{3} \right|_0^{y_0} = \frac{y_0^2}{3}$$

Luego

$$\mathrm{E}[X^2/Y] = \frac{Y^2}{3}$$

$$\mathrm{Var}[X/Y] = \frac{Y^2}{3} - \left(\frac{Y}{2}\right)^2 = \frac{Y^2}{12}$$

$$\mathrm{E}[\mathrm{Var}[X/Y]] = \frac{\mathrm{E}Y^2}{12} = \frac{1}{12}\int_0^1 y^2 2y dy = \frac{1}{6}\int_0^1 y^3 dy = \frac{1}{24} \left. y^3 \right|_0^1 = \frac{1}{24}$$

- Y/X

  Sabíamos que

$$f_1(x) = 2(1-x) \qquad 0 < x < 1$$

  y si $0 < x_0 < 1$

$$f_{Y/X=x_0}(y) = \frac{1}{1-x_0} \qquad x_0 < y < 1$$

$$\mathrm{E}[Y/X] = \frac{1+X}{2}$$

  Ahora, si $0 < x_0 < 1$

$$\mathrm{E}[Y^2/X = x_0] = \int_{x_0}^1 y^2 \frac{1}{1-x_0} dy = \frac{1}{1-x_0} \left. \frac{y^3}{3} \right|_{x_0}^1 = \frac{1-x_0^3}{3(1-x_0)} = \frac{x_0^2 + x_0 + 1}{3}$$

  Luego

$$\mathrm{E}[Y^2/X] = \frac{X^2 + X + 1}{3}$$

$$\mathrm{Var}[Y/X] = \frac{X^2+X+1}{3} - \left(\frac{1+X}{2}\right)^2 = \frac{X^2-2X+1}{12}$$

$$\mathrm{E}[\mathrm{Var}[Y/X]] = \mathrm{E}\left(\frac{X^2-2X+1}{12}\right) = \frac{1}{12}\int_0^1 (x^2-2x+1)2(1-x)dx =$$

$$= \frac{1}{6}\int_0^1 (-x^3+3x^2-3x+1)dx == \frac{1}{6}(-\frac{1}{4}+1-\frac{3}{2}+1) = \frac{1}{4}$$

3.- *Sea (X,Y) un vector aleatorio con función de densidad*

$$f(x,y)=1 \qquad (x,y)\in T\{(0,0),(2,0),(1,1)\}$$

Calcular $Var[Y/X]$ y $E[Var[Y/X]]$.

Sabíamos que la marginal de X era

$$f_1(x) = \begin{cases} \displaystyle\int_0^x dy = x & 0<x<1 \\ \\ \displaystyle\int_0^{2-x} dy = 2-x & 1<x<2 \end{cases}$$

y la condicionada $Y/X=x_0$

Si $0<x_0<1$, $\quad f_{Y/X=x_0}(y) = \dfrac{1}{x_0} \qquad 0<y<x_0$

Si $1<x_0<2$, $\quad f_{Y/X=x_0}(y) = \dfrac{1}{2-x_0} \qquad 0<y<2-x_0$

$$\mathrm{E}[Y/X] = \begin{cases} X/2 & 0<X<1 \\ \\ (2-X)/2 & 1<X<2 \end{cases}$$

Ahora,

si $0<x_0<1$, $\quad \mathrm{E}[Y^2/X=x_0] = \displaystyle\int_0^{x_0} y^2\frac{1}{x_0}dy = \frac{1}{x_0}\frac{y^3}{3}\Big|_0^{x_0} = \frac{x_0^2}{3}$

y si $1<x_0<2$, $\quad \mathrm{E}[Y^2/X=x_0] = \displaystyle\int_0^{2-x_0} y^2\frac{1}{2-x_0} = \frac{1}{2-x_0}\frac{y^3}{3}\Big|_0^{2-x_0} = \frac{(2-x_0)^2}{3}$

$$\mathrm{Var}[Y/X = x_0] = \begin{cases} \frac{x_0^2}{3} - \left(\frac{x_0}{2}\right)^2 = \frac{x_0^2}{12} & 0 < x_0 < 1 \\ \\ \frac{(2-x_0)^2}{3} - \left(\frac{2-x_0}{2}\right)^2 = \frac{(2-x_0)^2}{12} & 1 < x_0 < 2 \end{cases}$$

$$\mathrm{E}[\mathrm{Var}[Y/X]] = \int \mathrm{Var}[Y/X = x] f_1(x) = \int_0^1 \frac{x^2}{12} x dx + \int_1^2 \frac{(2-x)^2}{12}(2-x)dx =$$

$$\left.\frac{x^4}{48}\right|_0^1 - \left.\frac{(2-x)^4}{48}\right|_1^2 = \frac{1}{48} + \frac{1}{48} = \frac{1}{24}$$