EYP 1027 Modelos Probabilísticos Clase 13-1

Profesor: Reinaldo B. Arellano-Valle

Departamento de Estadística Pontificia Universidad Católica de Chile

Segundo Semestre 2020



Contenido I

- 1 Esperanza de Funciones de un Vector Aleatorio
 - Aplicaciones

Sea $g(x_1,\ldots,x_n)$ una función real valorada definida sobre el espacio muestral o recorrido $\mathcal X$ de un vector aleatorio (X_1,\ldots,X_n) . Entonces, $g(X_1,\ldots,X_n)$ es un variable aleatoria cuya esperanza (provisto que exista) puede calcularse a partir de la distribución conjunta de X_1,\ldots,X_n como sigue:

Teorema 1.1

$$\mathbb{E}\{g(X_1,\ldots,X_n)\} = \begin{cases}
\sum \cdots \sum_{(x_1,\ldots,x_n)\in\mathbb{R}^n} g(x_1,\ldots,x_n) f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) & \text{c.d.} \\
\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1,\ldots,x_n) f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) dx_1 \cdots dx_n & \text{c.c.}
\end{cases}$$

(provisto que las sumatorias y las integrales convergan)

Aplicaciones

I) Esperanza de funciones lineales $Si\ X_1,\ldots,X_n$ son va's con esperanza finita, entonces

$$\mathsf{E}\left(\sum_{i=1}^n a_i X_i + b\right) = \sum_{i=1}^n a_i \mathsf{E}(X_i) + b,$$

donde a_1, \ldots, a_n, b son constantes reales.

Por ejemplo, para la media muestral $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$, se tiene que,

$$\mathsf{E}(\bar{X}) = \mathsf{E}\left(\frac{1}{n}\sum_{i=1}^{n}X_{i}\right) = \frac{1}{n}\sum_{i=1}^{n}\mathsf{E}(X_{i}).$$

Ejemplo 1.1

1) Sean X_1 y X_2 va's independientes, donde $X_1 \sim U(0,1)$, con fdp $f_{X_1}(x) = I_{(0,1)}(x)$, y $X_2 \sim N(0,1)$, con fdp $f_{X_2}(x) = \phi(x)$ donde $\phi(x) = (2\pi)^{-1/2} \exp(-x^2/2)$ para $x \in \mathbb{R}$. Entonces,

$$E(X_1) = \int_0^1 x dx = \frac{1}{2}$$
 y $E(X_2) = \int_{-\infty}^\infty x \phi(x) dx = 0$.

Así, tenemos, por ejemplo, que $\mathrm{E}(2X_1+X_2)=2\mathrm{E}(X_1)+\mathrm{E}(X_2)=1$ y $\mathrm{E}(2X_1-X_2-1)=2\mathrm{E}(X_1)-\mathrm{E}(X_2)-1=0.$

Ejemplo 1.2

2) Sean X_1, X_2, X_3 va's con fdp conjunta dada por

$$f_{X_1, X_2, X_3}(x_1, x_2, x_3) = \begin{cases} \frac{3}{4}(x_1^2 + x_2^2 + x_3^2) & \text{si } 0 < x_1, x_2, x_3 < 1, \\ 0 & \text{eoc.} \end{cases}$$

Calcule la esperanza de $\bar{X} = (X_1 + X_2 + X_3)/3$.

Alternativa 1. Calcule primero la medias marginales,

$$E(X_1) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 f_{X_1, X_2, X_3}(x_1, x_2, x_3) dx_3 dx_2 dx_1$$

$$= \int_0^1 \int_0^1 \int_0^1 \frac{3}{4} x_1 (x_1^2 + x_2^2 + x_3^2) dx_3 dx_2 dx_1$$

$$= \frac{3}{4} \int_0^1 x_1 \int_0^1 \int_0^1 (x_1^2 + x_2^2 + x_3^2) dx_3 dx_2 dx_1$$
5

$$= \frac{3}{4} \int_0^1 x_1 \int_0^1 \left(x_1^2 + x_2^2 + \frac{1}{3} \right) dx_2 dx_1$$

$$= \frac{3}{4} \int_0^1 x_1 \left(x_1^2 + \frac{1}{3} + \frac{1}{3} \right) dx_1$$

$$= \frac{3}{4} \int_0^1 \left(x_1^3 + \frac{2}{3} x_1 \right) dx_1$$

$$= \frac{3}{4} \left(\frac{1}{4} + \frac{1}{3} \right) = \frac{7}{16}.$$

Por analogogía, se tiene que $\mathsf{E}(X_3)=\mathsf{E}(X_2)=\mathsf{E}(X_1)=7/16\Longrightarrow \mathsf{E}(\bar{X})=7/16.$

Tarea: i) Calcule $\mathsf{E}(X_1X_2X_3)$; ii) Calcule las fdp's marginales y estudie si X_1, X_2, X_3 son va's iid.

Alternativa 2. Aplicando el Teorema 1.1 directamente,

$$\begin{split} \mathsf{E}(\bar{X}) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{x_1 + x_2 + x_3}{3} f_{X_1, X_2, X_3}(x_1, x_2, x_3) dx_3 dx_2 dx_1 \\ &= \frac{1}{3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 f_{X_1, X_2, X_3}(x_1, x_2, x_3) dx_3 dx_2 dx_1 \longrightarrow \frac{1}{3} \mathsf{E}(X_1) \\ &+ \frac{1}{3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_2 f_{X_1, X_2, X_3}(x_1, x_2, x_3) dx_3 dx_2 dx_1 \longrightarrow \frac{1}{3} \mathsf{E}(X_2) \\ &+ \frac{1}{3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_3 f_{X_1, X_2, X_3}(x_1, x_2, x_3) dx_3 dx_2 dx_1 \longrightarrow \frac{1}{3} \mathsf{E}(X_3) \end{split}$$

$$\begin{split} &=\frac{1}{3}\int_0^1\int_0^1\int_0^1x_1\frac{3}{4}x_1(x_1^2+x_2^2+x_3^2)dx_3dx_2dx_1 \longrightarrow \frac{1}{3}\mathsf{E}(X_1) \\ &+\frac{1}{3}\int_0^1\int_0^1\int_0^1x_2\frac{3}{4}x_1(x_1^2+x_2^2+x_3^2)dx_3dx_2dx_1 \longrightarrow \frac{1}{3}\mathsf{E}(X_2) \\ &+\frac{1}{3}\int_0^1\int_0^1\int_0^1x_3\frac{3}{4}x_1(x_1^2+x_2^2+x_3^2)dx_3dx_2dx_1 \longrightarrow \frac{1}{3}\mathsf{E}(X_3) \\ &=\frac{7}{16}. \end{split}$$

II) Fgm conjunta

Definición 1.1

La fgm de conjunta de X_1, \ldots, X_n se difine como,

$$M_{X_1,\dots,X_n}(t_1,\dots,t_n) = \mathbf{E}\left(e^{\sum_{i=1}^n t_i X_i}\right)$$

$$= \begin{cases} \sum \dots \sum_{(x_1,\dots,x_n) \in \mathbb{R}^n} e^{\sum_{i=1}^n t_i x_i} f_{X_1,\dots,X_n}(x_1,\dots,x_n) & \text{c.d.} \end{cases}$$

$$= \begin{cases} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{\sum_{i=1}^n t_i x_i} f_{X_1,\dots,X_n}(x_1,\dots,x_n) dx_1 \dots dx_n & \text{c.c.} \end{cases}$$

provisto que la esperanza exista para todo $(t_1, \ldots, t_n) \in \mathbb{R}^n$ tal que $|t_k| < h_k$, algún $(h_1, \ldots, h_k) \in \mathbb{R}^n$, con $h_k > 0$ para todo $k = 1, \ldots, n$.

Propiedades

- (i) $M_{X_1,\dots,X_k}(t_1,\dots,t_k)=M_{X_1,\dots,X_k,X_{k+1},\dots,X_n}(t_1,\dots,t_k,0,\dots,0)$ para todo $k=1,\dots,n-1$.
- (ii) $\frac{\partial^{k_1+\dots+k_n} M_{X_1,\dots,X_n}(t_1,\dots,t_n)}{\partial t_1^{k_1}\dots \partial t_n^{k_n}} |_{t_1=\dots=t_n=0} = \mathsf{E}(X_1^{k_1}\times\dots\times X_n^{k_n}).$
- (iii) X_1,\ldots,X_n son va's independientes si, y sólo si,

$$M_{X_1,...,X_n}(t_1,...,t_n) = \prod_{i=1}^n M_{X_i}(t_i)$$

para todo (t_1, \ldots, t_n) donde las fgm's existen.

(iv)
$$M_{\sum_{i=1}^{n} a_i X_i + b}(t) = e^{bt} M_{X_1, \dots, X_n}(a_1 t, \dots, a_n t).$$

Así, si X_1, \ldots, X_n son va's independientes, entonces,

$$M_{\sum_{i=1}^{n} a_i X_i + b}(t) = e^{bt} \prod_{i=1}^{n} M_{X_i}(a_i t).$$

En particular, si $X_1, \ldots, X_n \stackrel{iid}{\sim} M(t)$, entonces, para la fgm de $\sum_{i=1}^n X_i$, se tiene que,

$$M_{\sum_{i=1}^{n} X_i}(t) = \{M(t)\}^n$$
.

Ejemplo 1.3

1) Si
$$X_1, \ldots, X_n \stackrel{iid}{\sim} Ber(p) \Longrightarrow M(t) = 1 - p + pe^t$$
. Luego,

$$M_{\sum_{i=1}^{n} X_i}(t) = (1 - p + pe^t)^n \iff Y = \sum_{i=1}^{n} X_i \sim Bin(n, p).$$

2) Si
$$X_1, \ldots, X_n \stackrel{iid}{\sim} \exp(\lambda) \Longrightarrow M(t) = (1 - t/\lambda)^{-1}, \ t < \lambda$$
. Luego,

$$M_{\sum_{i=1}^{n} X_i}(t) = (1 - t/\lambda)^{-n} \iff Y = \sum_{i=1}^{n} X_i \sim Gama(n, \lambda).$$

III) Covarianza y Correlación Sean X e Y va's definidas en el mismo espacio de probabilidad, y sean $\mu_X = \mathsf{E}(X)$, $\mu_Y = \mathsf{E}(Y)$, $\sigma_X^2 = \mathsf{Var}(X)$ y $\sigma_Y^2 = \mathsf{Var}(Y)$. Suponga que $0 < \sigma_Y^2 < \infty$ y $0 < \sigma_Y^2 < \infty$.

Definición 1.2

La convarianza entre X e Y se define como,

$$\operatorname{Cov}(X,Y) = \operatorname{E}\{(X - \mu_X)(Y - \mu_Y)\}\$$

$$= \begin{cases} \sum_{x \in \mathbb{R}} \sum_{y \in \mathbb{R}} (x - \mu_X)(y - \mu_Y) f_{X,Y}(x,y) & \text{c.d.} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) f_{X,Y}(x,y) dy dx & \text{c.c.} \end{cases}$$

La correlación entre X e Y se define como,

$$\rho_{XY} := \operatorname{Correlación}(X,Y) = \frac{\operatorname{Cov}(X,Y)}{\sqrt{\operatorname{Var}(X)\operatorname{Var}(Y)}}.$$

Propiedades

- (i) Cov(X,Y) = E(XY) E(X)E(Y) (formula alternativa)
- (ii) $Cov(X, X) = Var(X) \ge 0$ (operador positivo definido)
- (iii) Cov(X, Y) = Cov(Y, X) (simetría)
- (iv) $\operatorname{Cov}(X,c) = \operatorname{Cov}(c,X) = 0$ para cualquier constante $c \in \mathbb{R}$
- (v) Si X e Y son va's independientes, entonces $\operatorname{Cov}(X,Y)=0$
- (vi) $\operatorname{Cov}(aX+b,Y)=a\operatorname{Cov}(X,Y)$ y $\operatorname{Cov}(aX+b,X)=a\operatorname{Var}(X)$ donde a y b son constante.
- $(\mathsf{vii}) \quad |\mathsf{Cov}(X,Y)| \leq \sqrt{\mathsf{Var}(X)\mathsf{Var}(Y)}.$
- Las propiedades (vi)+(ii) $\Longrightarrow |\rho_{XY}| \le 1$ con igualdad ssi Y = aX + b algún $a \ne 0$ y b.

Ejemplo 1.4

1) Sean X=3Z+1 e $Y=-\frac{1}{3}Z-1$, donde $Z\sim N(0,1)$. Ya que, $\mathrm{E}(Z)=0$ y $\mathrm{Var}(Z)=1$, entonces:

$$E(X) = 3E(Z) + 1 = 1, \quad E(Y) = -3E(Z) - 1 = -1,$$

$$Var(X) = 3^{2}Var(Z) = 9, \quad Var(Y) = (-\frac{1}{3})^{2}Var(Z) = \frac{1}{9},$$
$$Cov(X, Y) = Cov(3Z + 1, Y)$$

$$= 3\operatorname{Cov}(Z, Y) \quad (\operatorname{por} (\operatorname{vi}))$$
$$= 3\operatorname{Cov}(Z, -\frac{1}{3}Z - 1)$$

$$= 3(-\frac{1}{3})\operatorname{Var}(Z) \quad \text{(por (vi))}$$

$$= -\operatorname{Var}(Z) = -1 \implies \rho_{XY} = -1/\sqrt{9\frac{1}{9}} = -1.$$

2) Si $Z \sim U(-1,1)$, entonces,

$$\mathsf{E}(Z) = \frac{1 + (-1)}{2} = 0$$
 y $\mathsf{Var}(Z) = \frac{(1 - (-1))^2}{12} = \frac{1}{3}$

Luego, de acuerdo con el ejemplo anterior,

$$\begin{split} \mathsf{E}(X) &= 1, \quad \mathsf{E}(Y) = -1, \\ \mathsf{Var}(X) &= 9\mathsf{Var}(Z) = 3, \quad \mathsf{Var}(Y) = \frac{1}{9}\mathsf{Var}(Z) = \frac{1}{27}, \\ \mathsf{Cov}(X,Y) &= -\mathsf{Var}(Z) = -\frac{1}{3} \implies \rho_{XY} = -1. \end{split}$$

IV) Varianza de funciones lineales Sean X_1, \ldots, X_n va's y a_1, \ldots, a_n, b constantes reales. Si todas las X_i 's tienen esperanza finita, vimos que,

$$\mathsf{E}\left(\sum_{i=1}^{n} a_{i} X_{i} + b\right) = \sum_{i=1}^{n} a_{i} \mathsf{E}(X_{i}) + b. \quad (*)$$

Además, si todas las X_i tienen varianza finita, entonces,

$$\operatorname{Var}\left(\sum_{i=1}^n a_i X_i + b\right) = \sum_{i=1}^n a_i^2 \operatorname{Var}(X_i) + 2 \sum_{\substack{1 \leq i,j \leq n \\ i < j}} a_i a_j \operatorname{Cov}(X_i, X_j). \quad (**)$$

En particular, si las va's X_1, \ldots, X_n son independientes, entonces (**) se reduce para,

$$\operatorname{Var}\left(\sum_{i=1}^n a_i X_i + b\right) = \sum_{i=1}^n a_i^2 \operatorname{Var}(X_i).$$

Por ejemplo, si X_1,\ldots,X_n son va's iid con media μ y varianza σ^2 , entonces la media y la varianza de la media muestral $\bar{X}=\sum_{i=1}^n X_i/n$, estan dadas por,

$$\mathsf{E}(\bar{X}) = \mu \quad \mathsf{y} \quad \mathsf{Var}(\bar{X}) = \frac{\sigma^2}{n}.$$

Además, la desigualdad de Chebyshev, implica que, para todo $\varepsilon>0$,

$$0 \leq P(|\bar{X} - \mu| > \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2} \longrightarrow 0, \quad \text{cuando } n \longrightarrow \infty,$$

es decir, la probabilidad de que \bar{X} difiera de μ en una cantidad positiva (ε) arbitrariamente pequeña es despreciable para n lo suficientemente grande.

Ejemplo 1.5

a) Sean $X, Y \stackrel{\text{iid}}{\sim} N(0, 1) \Longrightarrow E(X) = E(Y) = 0$, Var(X) = Var(Y) = 1 y Cov(X, Y) = 0. Luego,

$$Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y) = 1 + 1 + 2 \times 0 = 2.$$

Análogamente,

$$Var(X - Y) = Var(X) + Var(Y) - 2Cov(X, Y) = 1 + 1 - 2 \times 0 = 2.$$

Además, como E(X+Y)=E(X-Y)=0, usando la propiedad (i) de la covarianza, se tiene que,

Cov
$$(X + Y, X - Y) = E\{(X + Y)(X - Y)\} - 0 \times 0$$

$$= E(X^2 - Y^2) = E(X^2) - E(Y^2)$$

$$= 1 - 1 = 0 \implies \rho_{X+Y,X-Y} = 0.$$

b) Si $X,Y \stackrel{\text{iid}}{\sim} \exp(1) \Longrightarrow X+Y$ y X-Y también tienen correlación nula. En efecto, en este caso

$$\mathsf{E}(X) = \mathsf{E}(Y) = \mathsf{Var}(X) = \mathsf{Var}(Y) = 1 \quad \mathsf{y} \quad \mathsf{Cov}(X,Y) = 0.$$

Luego,

$$\mathsf{E}(X+Y) = 2, \quad \mathsf{E}(X-Y) = 0, \quad \mathsf{Var}(X+Y) = \mathsf{Var}(X-Y) = 2$$

у

$$\begin{split} \mathsf{Cov}(X+Y,X-Y) &= \mathsf{E}\{(X+Y)(X-Y)\} - 2 \times 0 = \mathsf{E}(X^2) - \mathsf{E}(Y^2) \\ &= \mathsf{Var}(X) + \mathsf{E}(X)^2 - \{\mathsf{Var}(Y) + \mathsf{E}(Y)^2\} \\ &= 2 - 2 = 0 \quad \Longrightarrow \ \rho_{X+Y,X-Y} = 0. \end{split}$$

E(X + Y) = 1, E(X - Y) = -1,

c) Si $X \sim N(0,1)$ e $Y \sim N(1,1)$ son va's independientes, entonces,

$$\label{eq:Var} \begin{aligned} \operatorname{Var}(X+Y) &= \operatorname{Var}(X-Y) = 2, \\ \operatorname{Cov}(X+Y,X-Y) &= \operatorname{E}\{(X+Y)(X-Y)\} - \operatorname{E}(X+Y)\operatorname{E}(X-Y) \end{aligned}$$

 $= 1 - (1+1) - 1 \times (-1) = 0 \implies \rho_{X+YX-Y} = 0.$

Ejemplo 1.6

- a) Sean $X,Y \stackrel{\text{iid}}{\sim} N(0,1)$. Pruebe que X+Y y X-Y son va's iid N(0,2).
- b) Sean $X, Y \stackrel{\text{iid}}{\sim} \exp(1)$. Las va's X + Y y X Y son independientes? Indique las distribuciones marginales de X + Y y X Y.
- c) Sean $X \sim N(0,1)$ e $Y \sim N(1,1)$ va's independientes. Obtenga las fgm's conjunta y marginales de X + Y y X Y. (Tarea!)

Solución problema a)

Sabemos que $M_X(t)=M_Y(t)=e^{t^2/2}$ para todo $t\in\mathbb{R}.$ Entonces, por definición la fgm conjunta de X+Y y X-Y esta dada por,

$$\begin{split} M_{X+Y,X-Y}(t,s) &= \mathsf{E} \left(e^{t(X+Y)+s(X-Y)} \right) \\ &= \mathsf{E} \left(e^{(t+s)X+(t-s)Y} \right) \\ &= \mathsf{E} \left(e^{(t+s)X} e^{(t-s)Y} \right) \\ &= \mathsf{E} \left(e^{(t+s)X} \right) \mathsf{E} \left(e^{(t-s)Y} \right) \quad (X \text{ e } Y \text{ son independientes}) \\ &= M_X(t+s) M_Y(t-s) \\ &= e^{(t+s)^2/2} e^{(t-s)^2/2} \\ &= e^{(2t^2+2s^2)/2} \quad \forall \, (s,t) \in \mathbb{R}^2. \end{split}$$

Además, las fgm's marginales son,

$$M_{X+Y}(t) = M_{X+Y,X-Y}(t,0) = e^{(2t^2)/2}, \quad t \in \mathbb{R},$$

У

$$M_{X-Y}(s) = M_{X+Y,X-Y}(0,s) = e^{(2s^2)/2}, \quad s \in \mathbb{R}.$$

Es claro que,

$$M_{X+Y}(t) = M_{X-Y}(t) = e^{(2t^2)/2}$$

 $\implies X + Y$ y X - Y tienen la misma distribución N(0,2).

También es claro que

$$M_{X+Y,X-Y}(t,s) = M_{X+Y}(t)M_{X-Y}(s) \quad \forall (s,t) \in \mathbb{R}^2$$

 $\Longrightarrow X+Y$ y X-Y son va's independientes, y como ambas tienen distribución N(0,2), entonces las va's X+Y y X-Y son iid N(0,2).

Solución problema b)

En este caso, $M_X(t) = M_Y(t) = (1-t)^{-1}$ la cual esta definida para t < 1. Procediendo como en el caso a), la fgm conjunta de X + Y y X - Y es,

$$M_{X+Y,X-Y}(t,s) = M_X(t+s)M_Y(t-s)$$

$$= \{1 - (t+s)\}^{-1}\{1 - (t-s)\}^{-1}, \quad t+s < 1, \ t-s < 1$$

$$= \{(1-t)^2 - s^2\}^{-1}, \quad t < 1, \ -1 + t < s < 1 + t.$$

Las fmg's marginales son:

$$M_{X+Y}(t)=(1-t)^{-2}$$
 para $t<1\Longrightarrow X+Y\sim Gama(2,1)$, y $M_{X-Y}(s)=(1-s^2)^{-1}$ para $|s|<1\Longrightarrow$ tiene distribución doble exponencial, con fdp $f_{X-Y}(z)=\frac{1}{2}\,e^{-|z|}$ para $z\in\mathbb{R}.$

Luego, $M_{X+Y,X-Y}(t,s) \neq M_{X+Y}(t)M_{X-Y}(s) \Longrightarrow X+Y$ y X-Y no son va's independientes.