EYP 1025-1027 Modelos Probabilísticos

Clase 13

Profesor: Reinaldo B. Arellano-Valle

Departamento de Estadística
Pontificia Universidad Católica de Chile

Contenido I

Esperanza de Funciones de un Vector Aleatorio

Dado un vector aleatorio $(X_1, \ldots, X_n)$ en $(\Omega,\ \mathcal{A},\ P)$ y una función real valorada $g : \mathbb{R}^n \longrightarrow \mathbb{R}$, entonces la transformación $Y = g(X_1, \ldots, X_n)$ define una variable aleatoria en $(\Omega,\ \mathcal{A},\ P)$, ya que para cada $\omega \in \Omega$, se tiene que $Y(\omega) = g(X_1(\omega), \ldots, X_n(\omega)) \in \mathbb{R}$.

Aunque, en general, es importante conocer en la distribución de la variable aleatoria $g(X_1, \ldots, X_n)$, podemos prenscindir de ella para el calculo de su esperanza (cuando existe).

Teorema 1.1

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

(provisto que las sumatorias y las integrales convergan)

I) **Esperanza de funciones lineales**

Si $g(x_1, \ldots, x_n) = \sum_{i=1}^{n} a_i x_i + b$, donde $a_1, \ldots, a_n, b$ son constantes reales, entonces

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

provisto que $\mathsf{E}(|X_i|) < \infty$ para todo $i = 1, \ldots, n$.

Por ejemplo, si $\bar{X} = \dfrac{1}{n} \sum_{i=1}^{n} X_i$ es la media muestral de $X_1, \ldots, X_n$, entonces

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

Esperanza de Funciones de Vectores Aleatorios

En particular, si todas las X_i's tienen la misma media, digamos $\mathsf{E}(X_i) = \mu$ para $i = 1, \ldots, n$, entonces

$$\mathsf{E}(\bar{X}) = \frac{1}{n}\sum_{i=1}^{n} \mu = \mu \quad \forall\, n = 1, 2, \ldots. \quad (*)$$

Esto ocurre, por ejemplo, cuando $X_1, \ldots, X_n$ corresponde a una muestra aleatoria (m.a.) de una distribución F (población) con media μ, lo cual equivale a decir que $X_1, \ldots, X_n \overset{\text{iid}}{\sim} F$, donde F una distribucón con media μ.

El resultado $(*)$ indica que la media muestral $\bar{X}$ es un predictor insegado de la media poblacional μ.

Ejemplo 1.1

Sean $X_1, \ldots, X_n \overset{iid}{\sim} \exp(\lambda)$. Sabemos que a media de la distribución exponencial es

$$\mu = \int_0^\infty x \underbrace{\lambda e^{-\lambda x}}_{f_X(x)} dx = \frac{1}{\lambda} \implies \mathrm{E}(\bar{X}) = \frac{1}{\lambda}.$$

Esperanza de Funciones de Vectores Aleatorios

II) Función generadora de momentos (fgm) multivariada

Si $g(x_1, \ldots, x_n) = \prod_{i=1}^{n} e^{t_i x_i} = e^{\sum_{i=1}^{n} t_i x_i}$, entonces $\mathrm{E}\{g(X_1, \ldots, X_n)\}$ define la fgm conjunta de $X_1, \ldots, X_n$, provisto, obviamente, que ella exista; es decir:

Definición 1.1

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

$$M_{X_1,\ldots,X_n}(t_1, \ldots, t_n) = \mathrm{E}\left(e^{\sum_{i=1}^{n} t_i X_i}\right)$$
$$= \begin{cases} \sum \cdots \sum_{(x_1,\ldots,x_n) \in \mathbb{R}^n} e^{\sum_{i=1}^{n} t_i x_i} f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) & \text{c.d.} \\ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{\sum_{i=1}^{n} t_i x_i} f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) dx_1 \cdots dx_n & \text{c.c.} \end{cases}$$

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 de la fgm multivariada:

(i) $M_{X_1,\ldots,X_k}(t_1,\ldots,t_k) = M_{X_1,\ldots,X_k,X_{k+1},\ldots,X_n}(t_1,\ldots,t_k,0,\ldots,0)$ para todo $k = 1,\ldots,n-1$.

(ii) $\frac{\partial^{k_1+\cdots+k_n} M_{X_1,\ldots,X_n}(t_1,\ldots,t_n)}{\partial t_1^{k_1}\cdots\partial t_n^{k_n}} \big|_{t_1=\cdots=t_n=0} = \mathsf{E}(X_1^{k_1} \times \cdots \times X_n^{k_n})$.

(iii) $X_1,\ldots,X_n$ son va's independientes si, y sólo si,

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

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

(iv) Si $Y = \sum_{i=1}^{n} a_i X_i + b$, entonces,

$$M_Y(t) = e^{bt} M_{X_1,\ldots,X_n}(a_1 t,\ldots,a_n t).$$

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

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

En particular, si $X_1, \ldots, X_n \overset{iid}{\sim} M(t)$ e $Y = \sum_{i=1}^{n} X_i$, entonces

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

Ejemplo 1.2

Sean $X_1, \ldots, X_n \stackrel{iid}{\sim} Ber(p)$. Pruebe que $Y = \sum_{i=1}^{n} X_i \sim Bin(n, p)$.
Demostración Ya que $X_1, \ldots, X_n \stackrel{iid}{\sim} Ber(p)$, entonces

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

Luego, la fgm de $Y = \sum_{i=1}^{n} X_i$ es,

$$\begin{aligned} M_Y(t) &= M_{\sum_{i=1}^{n} X_i}(t) \\ &= \{M(t)\}^n \\ &= (1 - p + pe^t)^n \longrightarrow \text{ fgm de la distribución } Bin(n, p). \end{aligned}$$

Por la unicidad de la fgm se concluye que $Y = \sum_{i=1}^{n} X_i \sim Bin(n, p)$.

Ejemplo 1.3

Sean que X_1 e X_2 va's independientes, con $X_i \sim N(\mu_i, \sigma^2)$, $i = 1, 2$. Encuentre la distribución de $Y = X_1 - X_2$.

Solución Ya que $X_i \sim N(\mu_i, \sigma^2)$, entonces $M_{X_i}(t) = e^{\mu_i t + \frac{1}{2}\sigma^2 t^2}$ para $i = 1, 2$. Además, como X_1 y X_2 son va's independientes, entonces la fgm de $Y = X_1 - X_2$ es,

$$\begin{aligned} M_{X_1 - X_2}(t) &= M_{X_1}(t) M_{X_2}(-t) \\ &= e^{\mu_1 t + \frac{1}{2}\sigma^2 t^2} e^{-\mu_2 t + \frac{1}{2}\sigma^2 t^2} \\ &= e^{(\mu_1 - \mu_2)t + \frac{1}{2}(2\sigma^2)t^2}, \forall\ t \in \mathbb{R}. \end{aligned}$$

Como $M_{X_1 - X_2}(t) = e^{(\mu_1 - \mu_2)t + \frac{1}{2}(2\sigma^2)t^2}$ es la fgm de una distribución normal con media $\mu_1 - \mu_2$ y varianza $2\sigma^2$, se concluye que $X_1 - X_2 \sim N(\mu_1 - \mu_2, 2\sigma^2)$.

III) **Covarianza y Correlación**

Si $n = 2$, con $X_1 = X$ y $X_2 = Y$, sean $\mu_X = \mathsf{E}(X)$, $\mu_Y = \mathsf{E}(Y)$, $\sigma_X^2 = \mathsf{Var}(X)$ y $\sigma_Y^2 = \mathsf{Var}(Y)$, y suponga que $0 < \sigma_X^2 < \infty$ y $0 < \sigma_Y^2 < \infty$. Entonces, la función $g(x, y) = (x - \mu_X)(y - \mu_Y)$ permite definir la covarianza entre las va's X e Y; es decir:

Definición 1.2

La convarianza entre dos variables aleatorias X e Y se define como,

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

Notación alternativa e interpretación:

$$\sigma_{XY} := \mathsf{Cov}(X, Y)$$

Interpretación:

$$\begin{aligned}
\sigma_{XY} > 0 &\implies \text{asociación } +, \\
\sigma_{XY} < 0 &\implies \text{asociación } -, \\
\sigma_{XY} = 0 &\implies \text{no hay asociación lineal.}
\end{aligned}$$

Esperanza de Funciones de Vectores Aleatorios

Propiedades de la covarianza: Para dos va's X e Y definidas sobre el mismo espacio de probabilidades, se tiene que:

(i) $\text{Cov}(X, Y) = \text{E}(XY) - \text{E}(X)\text{E}(Y)$ (formula alternativa)

(ii) $\text{Cov}(X, X) = \text{Var}(X) \geq 0$ (operador positivo definido)

(iii) $\text{Cov}(X, Y) = \text{Cov}(Y, X)$ (simetría)

(iv) $\text{Cov}(X, c) = \text{Cov}(c, X) = 0$ para cualquier constante $c \in \mathbb{R}$

(v) Si X e Y son va's independientes, entonces $\text{Cov}(X, Y) = 0$

Nota: Si X e Y son independientes, entonces $g(X)$ y $h(Y)$ también son va's idependientes, para cualquier funciones g y h

(vi) $\text{Cov}(aX + b, Y) = a\text{Cov}(X, Y)$ y $\text{Cov}(aX + b, X) = a\text{Var}(X)$, donde a y b son constantes

(vii) $|\text{Cov}(X, Y)| \leq \sqrt{\text{Var}(X)\text{Var}(Y)}$ (desigualdad de Couchy-Schwarz)

Definición 1.3

La correlación (coeficiente de correlación) entre dos variables aleatorias X e Y, se define como,

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}.$$

La Propiedad (vii) $\Longrightarrow |\rho_{XY}| \leq 1$, con igualdad si, y sólo si $Y = aX + b$ para $a \neq 0$, de acuerdo con y la Propiedad (vi).

Más formalmente:

Teorema 1.2

Suponga que X es una variable aleatoria tal que $0 < \sigma_X^2 < \infty$. Si existen constantes $a \neq 0$ y b tal que $Y = aX + b$, entonces, $\rho_{XY} = 1$ si $a > 0$ (asociación lineal $+$) y $\rho_{XY} = -1$ si $a < 0$ (asociación lineal $-$).

Demostración 1.1

Como $\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X)$ y $\mathrm{Cov}(X, aX + b) = a\mathrm{Var}(X)$, la definición de correlación implica que,

$$\rho(X, aX + b) = \frac{\mathrm{Cov}(X, aX + b)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(aX + b)}} = \frac{a}{\sqrt{a^2}} = \begin{cases} 1 & \text{si } a > 0, \\ -1 & \text{si } a < 0. \end{cases}$$

IV) **Varianza de funciones lineales**

Considerando nuevamente la funcion lineal $Y = \sum_{i=1}^{n} a_i x_i + b$, ahora podemos calcular su varianza:

Sean $X_1, \dots, X_n$ va's con varianzas finitas y $a_1, \dots, a_n, b$ constantes en $\mathbb{R}$. Vimos que

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

Ahora, podemos también afirmar que,

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

Esperanza de Funciones de Vectores Aleatorios

En particular, si las va's $X_1, \ldots, X_n$ son independientes, entonces $\mathsf{Cov}(X_i, X_j) = 0\ \forall\ i \neq j$, y de $(**)$ se tiene que,

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

Por ejemplo, colocando para $n = 2$, $X_1 = X$ y $X_2 = Y$, entonces de $(**)$ se tiene que,

$$\mathsf{Var}(X + Y) = \mathsf{Var}(X) + \mathsf{Var}(Y) + 2\mathsf{Cov}(X, Y),$$
$$\mathsf{Var}(X - Y) = \mathsf{Var}(X) + \mathsf{Var}(Y) - 2\mathsf{Cov}(X, Y).$$

En particular,

$$\mathsf{Cov}(X, Y) = 0 \Longrightarrow \mathsf{Var}(X \pm Y) = \mathsf{Var}(X) + \mathsf{Var}(Y).$$

Esperanza de Funciones de Vectores Aleatorios

Demostración de $(*)$*:* Sea $Y = \sum_{i=1}^{n} a_i X_i + b$. La definición de varianza de una va Y y el resultado en $(*)$ implican que,

$$
\begin{aligned}
\mathsf{Var}(Y) &= \mathsf{E}\left\{(Y - \mathsf{E}(Y))^2\right\} \\
&= \mathsf{E}\left\{\left[\sum_{i=1}^{n} a_i X_i + b - \sum_{i=1}^{n} a_i \mathsf{E}(X)_i - b\right]^2\right\} \\
&= \mathsf{E}\left\{\left[\sum_{i=1}^{n} a_i (X_i - \mathsf{E}(X_i))\right]^2\right\} \\
&= \mathsf{E}\left\{\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j (X_i - \mathsf{E}(X_i))(X_j - \mathsf{E}(X_j))\right\} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \mathsf{E}\left\{(X_i - \mathsf{E}(X_i))(X_j - \mathsf{E}(X_j))\right\}
\end{aligned}
$$

Esperanza de Funciones de Vectores Aleatorios

$$
\begin{aligned}
&= \sum_{i=1}^{n} a_i^2 \mathsf{E}\left\{(X_i - \mathsf{E}(X_i))^2\right\} \\
&+ \sum_{\substack{1 \le i,j \le n \\ i \ne j}} a_i a_j \mathsf{E}\left\{(X_i - \mathsf{E}(X_i))(X_j - \mathsf{E}(X_j))\right\} \\
&= \sum_{i=1}^{n} a_i^2 \mathsf{Var}(X_i) + \sum_{\substack{1 \le i,j \le n \\ i \ne j}} a_i a_j \mathsf{Cov}(X_i, X_j) \\
&= \sum_{i=1}^{n} a_i^2 \mathsf{Var}(X_i) + 2 \sum_{\substack{1 \le i,j \le n \\ i < j}} a_i a_j \mathsf{Cov}(X_i, X_j)
\end{aligned}
$$

Ejemplo 1.4

Sean $X_1, \ldots, X_n$ va's iid, con media μ y varianza σ^2. Entonces, la media y la varianza de la media muestral $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ estan dadas por,

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

Ejemplo 1.5

Suponga que $X_i \overset{\text{ind.}}{\sim} N(\mu_i, \sigma_i^2)$, $i = 1, \dots, n$. Entonces, usando la fgm podemos probar que,

$$Y = \sum_{i=1}^{n} a_i X_i + b \sim N(\underbrace{\sum_{i=1}^{n} a_i \mu_i + b}_{\mathrm{E}(Y)}, \underbrace{\sum_{i=1}^{n} a_i \sigma_i^2}_{\mathrm{Var}(Y)}),$$

de modo que la distribución normal es cerrada bajo transformaciones lineales de va's normales independientes.

Ejemplo 1.6

Supoga que un estudiante debe rendir una prueba verbal y una cuantitativa. Sea X el puntaje en el examen verbal, e Y el puntaje en el examen cuantitativo. Se desea encontrar la correlación entre X e Y. Para ello, se supone que la fdp conjunta de X e Y está dada por,

$$f_{X,Y}(x,y) = \begin{cases} 2xy + 0.5 & \text{para } 0 \leq x \leq 1 \text{ y } 0 \leq y \leq 1 \\ 0 & \text{eoc} \end{cases}$$

Esperanza de Funciones de Vectores Aleatorios

Para calcularemos la covarianza $\mathrm{Cov}(X, Y)$, primero se deben calcular las medias μ_X y μ_Y de X e Y, respectivamente. La simetría en la fdp conjunta, implica que X e Y tienen la misma distribución marginal; por lo tanto, $\mu_X = \mu_Y$, donde,

$$\begin{aligned} \mu_X &= \int_0^1 \int_0^1 \left(2x^2 y + 0.5x\right) dydx \\ &= \int_0^1 \left(x^2 + 0.5x\right) dx = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}, \end{aligned}$$

y $\mu_Y = \frac{7}{12}$. Luego,

$$\mathrm{Cov}(X, Y) = \int_0^1 \int_0^1 \left(x - \frac{7}{12}\right)\left(y - \frac{7}{12}\right)(2xy + 0.5)dydx = \frac{1}{144}.$$

Además, $\sigma_X^2 = \sigma_Y^2 = 11/144$, de modo que $\rho(X, Y) = 1/11$. Aunque, un algo baja, la correlación entre los puntajes es positiva.