MA-34A Probab. y Procesos Estocásticos **2 de Julio, 1996**

Pauta Examen

Prof. Cátedras: F. Farias, M. Kiwi, y A. Maass
Prof. Auxiliares: D. Hojman, J.P. Risco, y P.P. Romagnoli

TIEMPO: 3.5 HRS.

PROBLEMA 1: Sean U, V variables aleatorias independientes $Normal(0,1)$. Se define el vector aleatorio

$$\vec{X} = \begin{pmatrix} X \\ Y \end{pmatrix} = A \cdot \begin{pmatrix} U \\ V \end{pmatrix} + \vec{\mu} \,,$$

donde $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ es una matriz invertible real y $\vec{\mu} = \begin{pmatrix} \mu \\ \nu \end{pmatrix} \in I\!R^2$.

1. (2.5 pts.) Probar que si $M = AA^T$, entonces la densidad conjunta del vector $\vec{X}$ es

$$f_{X,Y}(x,y) = \frac{1}{2\,\pi} \cdot \frac{1}{\sqrt{det M}} \cdot e^{-\frac{1}{2}(\vec{x}-\vec{\mu})^T M^{-1}(\vec{x}-\vec{\mu})} \,, \qquad \text{donde } \vec{x} = \begin{pmatrix} x \\ y \end{pmatrix} .$$

<u>Nota</u>: La densidad de una variable aleatoria $Normal(0,1)$ es $\frac{1}{\sqrt{2\,\pi}} \cdot e^{-\frac{1}{2}t^2}$.

2. (1.5 pts.) Identifique o calcule $\mathbf{E}(X)$, $\mathbf{E}(Y)$, y $\begin{pmatrix} \mathbf{V}(X) & \mathbf{Cov}(X,Y) \\ \mathbf{Cov}(X,Y) & \mathbf{V}(Y) \end{pmatrix}$.

3. (2.0 pts.) Suponga que el vector aleatorio $\begin{pmatrix} S \\ T \end{pmatrix}$ tiene densidad conjunta

$$f_{S,T}(s,t) = \frac{1}{2\,\pi} \cdot e^{-\frac{1}{2}(2s^2+t^2+2st)} \,.$$

Determinar las marginales $f_S(s)$ y $f_T(t)$. ¿Son las variables S y T independientes?

SOLUCIÓN:

1. Si U, V son variables aleatorias independientes $Normal(0,1)$, se tiene que su densidad conjunta queda dada por:

$$f_{U,V}(u,v) = \frac{1}{2\,\pi} \cdot e^{-\frac{1}{2}(u^2+v^2)} = \frac{1}{2\,\pi} \cdot e^{-\frac{1}{2}\vec{u}^T \cdot \vec{u}} \,, \qquad \text{donde } \vec{u} = \begin{pmatrix} u \\ v \end{pmatrix} .$$

Por otro lado, como $\vec{X} = \vec{g}(U) = A \cdot \vec{U} + \vec{\mu}$ es una transformación invertible de $\vec{U}$ se tiene, usando el método del Jacobiano, que si $\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ entonces la densidad conjunta de $\vec{X}$ es:

$$\begin{aligned} f_{X,Y}(x,y) &= f_{U,V}(g^{-1}(\vec{x})) \cdot \frac{1}{|det(J_g(g^{-1}(\vec{x})))|} \\ &= \frac{1}{2\,\pi\,|det(A)|} \cdot e^{-\frac{1}{2}(A^{-1}(\vec{x}-\vec{\mu}))^T(A^{-1}(\vec{x}-\vec{\mu}))} \\ &= \frac{1}{2\,\pi\,\sqrt{det(M)}} \cdot e^{-\frac{1}{2}(\vec{x}-\vec{\mu})^T M^{-1}(\vec{x}-\vec{\mu})}\,. \end{aligned}$$

La segunda igualdad es consecuencia de que $g^{-1}(\vec{x}) = A^{-1}(\vec{x} - \vec{\mu})$ y puesto que $J_g(g^{-1}(\vec{x})) = A$. La última igualdad es consecuencia del hecho que $(A^{-1})^T \cdot A = (A \cdot A^T)^{-1} = M^{-1}$ y $det(M) = det(A \cdot A^T) = det(A) \cdot det(A^T) = (det(A))^2$.

2. Por linealidad del operador esperanza y dado que $\mathbf{E}(U) = \mathbf{E}(V) = 0$, sique que $\mathbf{E}(X) = \mathbf{E}(A_{1,1}U + A_{1,2}V + \mu) = A_{1,1}\,\mathbf{E}(U) + A_{1,2}\,\mathbf{E}(V) + \mu = \mu$. Análogamente se deduce que $\mathbf{E}(Y) = \nu$.

La matriz pedida corresponde a la matriz de covarianza del vector aleatorio normal multivariado $\vec{X}$, esto es $M = A \cdot A^T$.

3. Se tiene que

$$2\,s^2 + t^2 + 2\,s\,t = \begin{pmatrix} s & t \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} s \\ t \end{pmatrix},$$

i.e., si $M^{-1} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ entonces

$$f_{S,T}(s,t) = \frac{1}{2\,\pi}\, e^{-\frac{1}{2}\begin{pmatrix} s & t \end{pmatrix} M^{-1} \begin{pmatrix} s \\ t \end{pmatrix}}\,.$$

Entonces, $f_{S,T}(s,t)$ es una normal multivariada. Luego, S y T son normales. La esperanza de S es 0 y la de T es también 0, y sus varianzas son las que aparecen en la diagonal de $M = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$, i.e., $\mathbf{V}(S) = 1$ y $\mathbf{V}(T) = 2$. Luego, las marginales son

$$f_S(s) = \frac{1}{\sqrt{2\,\pi}}\, e^{-\frac{1}{2}s^2}\,, \qquad f_T(t) = \frac{1}{\sqrt{2\,\pi}}\,\frac{1}{\sqrt{2}}\, e^{-\frac{1}{4}t^2}\,.$$

Como $f_S(s) \cdot f_T(t) \neq f_{S,T}(s,t)$, entonces S, T son dependientes.

PROBLEMA 2: Sean $X_1, \ldots, X_n$ variables aleatorias independientes tales que $X_i \sim$ *Binomial*(n, p_i), i.e. $I\!P(X_i = k) = \binom{n}{k} p_i^k (1-p_i)^{n-k}$, para $k = 0, \ldots, n$. En este problema probaremos que si $\delta > 0$, $X = \sum_{i=1}^{n} X_i$, y $\mu = \mathbf{E}(X)$ entonces

$$I\!P(X \geq (1+\delta)\mu) \;\; \leq \;\; \frac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu}}\,, \qquad \forall\, t \geq 0\,. \qquad (*)$$

1. (2.0 pts.) Pruebe que la función generadora de momentos de X_i es

$$\phi_{X_i}(t) \;=\; \mathbf{E}\left(e^{tX_i}\right) \;=\; \left(1 + p_i\left(e^t - 1\right)\right)^n .$$

Luego, como $1 + x \leq e^x$ sigue que $\phi_{X_i}(t) \leq e^{np_i(e^t-1)}$.

2. (2.0 pts.) Pruebe que $\forall\, t \geq 0$ se tiene,

$$I\!P(X \geq (1+\delta)\mu) \;\; \leq \;\; \frac{\prod_{i=1}^{n} \mathbf{E}\left(e^{tX_i}\right)}{e^{t(1+\delta)\mu}} .$$

<u>Indicación</u>: Observe que si $t \geq 0$, entonces $a \geq b$ ssi $e^{ta} \geq e^{tb}$.

3. (2.0 pts.) Concluya la desigualdad (*).

SOLUCIÓN:

1. Observar que

$$\mathbf{E}(e^{tX_i}) \;=\; \sum_{j=0}^{n} \binom{n}{j} e^{tj} p_i^j (1-p_i)^{n-j} \;=\; \sum_{j=0}^{n} \binom{n}{j} (e^t p_i)^j (1-p_i)^{n-j} .$$

Pero por el Teorema del Binomio esta última expresión es igual a $(p_i e^t + 1 - p_i)^n$.

2. Aplicando la indicación se tiene que $X \geq (1+\delta)\mu$ ssi $e^{tX} \geq e^{t(1+\delta)\mu}$. Luego,

$$I\!P(X \geq (1+\delta)\mu) = I\!P(e^{tX} \geq e^{t(1+\delta)\mu}) \leq \frac{\mathbf{E}(e^{tX})}{e^{t(1+\delta)\mu}} = \frac{\mathbf{E}(e^{t\sum_{i=1}^{n} X_i})}{e^{t(1+\delta)\mu}} ,$$

donde la desigualdad se obtiene obsevando que e^{tX} es una variable aleatoria no-negativa y aplicando la desigualdad de Markov.

3. Como $X_i \sim Binomial(n, p_i)$ se tiene que su esperanza es np_i. Luego, por la linealidad del operador esperanza,

$$\mu \;=\; \mathbf{E}(X) \;=\; \mathbf{E}(\sum_{i=1}^{n} X_i) \;=\; \sum_{i=1}^{n} \mathbf{E}(X_i) \;=\; \sum_{i=1}^{n} np_i \,.$$

Por otro lado $\mathbf{E}(e^{tX_i}) \leq e^{np_i(e^t-1)}$, luego

$$\prod_{i=1}^{n} \mathbf{E}(e^{X_i}) \;=\; e^{(e^t-1)\sum_{i=1}^{n} np_i} \;=\; e^{(e^t-1)\mu} \,.$$

La última igualdad y el punto anterior implican la desigualdad pedida.

Problema 3: Sean $(X_i)_{i=1}^{\infty}$ variables aleatorias independientes e identicamente distribuidas con esperanza μ desconocida y varianza $\sigma^2 = 4$. Se define $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. En este problema queremos diseñar una estrategia para confirmar o rechazar la hipótesis de que la esperanza de estas variables es μ_0 (un valor real conocido).

1. (1.5 pts.) Calcular $\mathbf{E}(\bar{X}_n)$ en función de μ, y $\mathbf{V}(\bar{X}_n)$.

2. (1.5 pts.) Si $\mu = \mu_0$ y $c > 0$, justifique que $\lim_{n\to\infty} \mathbb{P}(|\bar{X}_n - \mu_0| \geq c) = 0$.

3. (2.0 pts.) Suponga en este punto que las variables X_i son normales, que $n = 100$, y que $\mu = \mu_0$. Encontrar el menor valor $c > 0$ tal que

$$\mathbb{P}(|\bar{X}_n - \mu_0| \geq c) \;\leq\; 0.1$$

<u>Nota</u>: $Z \sim Normal(0,1)$ entonces $\mathbb{P}(Z \leq 1.64) = 0.95$

Uno rechaza la hipótesis que $\mu = \mu_0$ con probabilidad 0.1 cuando $|\bar{X}_n - \mu_0| \geq c$.

4. (1.0 pts.) Proponga un método para confirmar o rechazar la hipótesis $\mu = \mu_0$ cuando las variables aleatorias X_i no son necesariamente normales como en el punto anterior. Justifique.

Solución:

1. Por lineilidad de la esperanza, $\mathbf{E}(\bar{X}_n) = \frac{1}{n}\mathbf{E}(\sum_{i=1}^{n} X_i) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{E}(X_i) = \frac{1}{n}n\mu = \mu$.

Por otro lado, $\mathbf{V}(\bar{X}_n) = \frac{1}{n^2}\mathbf{V}(\sum_{i=1}^{n} X_i) = \frac{1}{n^2}\sum_{i=1}^{n}\mathbf{V}(X_i) = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n} = \frac{4}{n^2}$, donde la segunda igualdad es consecuencia del hecho que las variables aleatorias X_i son independientes.

2. Como las variables aleatorias X_i son i.i.d. y con esperanza y varianza finitas, se cumplen las hipótesis de la ley de los grandes números en su versión débil, luego $\bar{X}_n$ converge en $I\!P$-probabilidad a μ, i.e., como $\mu = \mu_0$

$$\lim_{n\to\infty} I\!P(|\bar{X}_n - \mu_0| \geq c) \;=\; 0\,, \qquad \forall\, c > 0\,.$$

3. Como los X_i son variables aleatorias independientes de esperanza μ y varianza σ^2 entonces $\bar{X}_n$ es una variable aleatoria normalmente distribuida de esperanza $\mathbf{E}(\bar{X}_n) = \mu$ y varianza $\mathbf{V}(\bar{X}_n) = \frac{\sigma^2}{n}$. Luego, por propiedades de la distribución normal,

$$Z \;=\; \frac{\bar{X}_n - \mu}{\sqrt{\frac{\sigma^2}{n}}} \;\sim\; Normal(0,1)\,.$$

Luego,

$$I\!P(|\bar{X}_n - \mu_0| \geq c) \;=\; I\!P\left(|Z| \geq \frac{c\sqrt{n}}{\sigma}\right) \;=\; 2\left(1 - I\!P\left(Z \leq \frac{c\sqrt{n}}{\sigma}\right)\right)\,.$$

Para determinar el c pedido, resolvemos la siguiente ecuación con incognita c:

$$2\left(1 - I\!P\left(Z \leq \frac{c\sqrt{n}}{\sigma^2}\right)\right) \;=\; 0.1\,.$$

i.e., $I\!P(Z \leq \frac{c\sqrt{n}}{\sigma}) = 0.95$, luego de la indicación sigue que, $c\sqrt{n} = 1.64\,\sigma$. Concluimos que el valor pedido es $c = 1.64\sqrt{\frac{4}{100}} = 0.328$.

4. Cuando n es suficientemente grande, podemos usar el mismo método que en el punto anterior puesto que sabemos, por el Teorema Central del Limite, que $\bar{X}_n$ sigue una distribución aproximadamente normal. La normalidad de la distribución de $\bar{X}_n$ era la componente clave del análisis del punto anterior.