Tarea 3

MA3403 - Probabilidades y Estadística
Profesor: Fernando Lema
Auxiliares: Antonio Lizama & José Cereceda
12 de noviembre de 2013

P1. Sea (X, Y) un vector aleatorio. Se define el coeficiente de correlación ρ_{XY} entre x e Y como:

$$\rho_{XY} = \frac{\mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y)))}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}.$$

a) Muestre que si X e Y son independientes, entonces $\rho_{XY} = 0$.

b) Muestre, por medio de un ejemplo, que si $\rho_{XY} = 0$ no necesariamente X e Y son independientes.

c) Muestre que $-1 \leq \rho_{XY} \leq 1$.

d) Muestre que $Y = a + bX$ si y solo si $\rho^2_{XY} = 1$. **Indicación:** Ver Meyer.

P2. Sean $X_1 \sim N(\mu, \sigma^2),\ X_2 \sim N(\mu, \sigma^2)$, independientes. Considere

$$Z(t) = X_1 cos(wt) + X_2 sen(wt),\ V(t) = \frac{\partial Z(t)}{\partial t}$$

a) Determine la distribución de $Z(t)$ y $V(t)$ (t fijo).

b) Muestre que $\rho_{ZV} = 0$.

P3. La duración (en horas) de dos aparatos eléctricos son una v.a. (d_1, d_2) con distribución $N(43, 36)$ y $N(45, 9)$ respectivamente.

a) Si usted debe elegir uno de dichos aparatos, ¿cuál escogería?, ¿por qué?

b) Si se instalan dos equipos (uno de cada tipo) de tal forma que uno de ellos funciona el otro falla. Calule la probabilidad que el sistema total dure más de 80 hrs.

c) Si se elige sólo equipos de tipo d_1 y se instalan de tal forma que el i-ésimo comienza a funcionar cuando el $(i-1)$-ésimo falla, determine cuántos equipos se deben instalar para que el sistema funcione más de 750 hrs. con probabilidad 0.99. Suponga independencia en la falla de los equipos.

P4. Se considera que el tiempo de reacción frente a un estímulo luminoso es una v.a. T distribuida normalmente con media 0.65 segundos. De acuerdo con los resultados de una investigación se estima que bajo el efecto de cierta dosis de alcohol, el tiempo de reacción frente al mismo estímulo puede expresarse como $T* = 1{,}4T - 0{,}02$. Además, se indica que la probabilidad que un individuo que ha ingerido alcohol reaccione antes de 1 segundo es de 0.9.

a) ¿Cuál es la probabilidad de que un sujeto que no ha ingerido alcohol reaccione antes de 0.7 segundos?

b) Si se escogen 10 sujetos de forma independiente que no han ingerido alcohol. ¿cuál es la probabilidad que a lo más 2 reaccionen después de 0.7 segundos?

P5. *a*) Sean $X_i \sim U(0, 1),\ i = 1, ..., n$, independientes. Determine usando la función generadora de momentos la distribución de la v.a.:

$$Y = -2ln\left(\prod_{i=1}^{n} X_i\right)$$

¿Cuál es la distribución para n grande? Calcule $\mathbb{P}(Y > 60)$, si $n = 50$.

b) Sea X v.a. tq $\mathbb{E}(X^k) = (k+1)2^k,\ k = 1, 2, ...$ Determine, usando la *f.g.m.* la distribución de X.

P6. *a*) Una partícula parte del origen de la recta real y a cada seguna da saltos de tamaño 2 con probabilidad 1/6, de tamaño 1 con probabilidad 1/6, de tamaño -1 con probabilidad 1/3, y se queda en el lugar con probabilidad 1/3.

1) Calcule la probabilidad de que a los 36 segundos la partícula se encuentre a más de 10 unidades del origen.
2) Si X_n es la posición de la partícula a los n segundos, ¿qué distribución tiene X_n para n grande? ¿Hacia donde converge la partícula?
3) Determine cuántos segundos deben pasar para que con probabilidad de 0.999 la partícula se encuentre en el lado positivo de la recta.

b) Considere 2 trenes (A y B) deben transportar N personas diarias (entre los 2). Si cada persona al azar e independientemente elige un tren para viajar, determinaremos la cantidad de asientos, m, que debe llevar cada tren para asegurar con probabilidad 0.99 que todas las personas irán sentadas. Para esto:

1) Si X_A : número de personas que suben al tren A. ¿Qué distribución tiene X_A? ¿y para N grande?
2) Plantee condiciones sobre X_A para asegurar que todos los pasajeros viajen sentados (tanto en el tren A, como en el B).
3) Imponga la condición de probabilidad y determine m considerando N grande.

P7. La nota de los alumnos del curso MA3403 puede ser considerada como una va normal de media μ y varianza σ^2. Si se sabe que se aprueba con nota igual o superior a 4, se reprueba con nota menor a $3,7$ y queda pendiente con nota mayor o igual a $3,7$ y menor a 4.

a) Suponga que $\mu = 4{,}2$, $\sigma = 0{,}8$ y el total de alumnos es 110. Si se elige un grupo de 10 alumnos, sin reposición, calcule la probabilidad de que "a lo más dos de ellos estén reprobados y al menos 9 estén aprobados".

b) Si del curso se eliminan los reprobados, ¿cuál es la densidad de nota de los restantes? Calcule su esperanza.

c) Suponga que μ es desconocido y se escoge una muestra de alumnos de tamaño n (de manera independiente). Determine n de tal forma que la media muestral $\overline{X}$ difiera de la media poblacional μ en menos de 0,5 con probabilidad 0,95.

P8. *a*) Se dice $X_1, X_2, \ldots, X_n$ converge a C en media cuadrática si

$$\lim_{n\to\infty} \mathbb{E}((X_n - C)^2) = 0$$

Sean $X_1, \ldots, X_n$ tales que

$$\begin{aligned} \mathbb{P}\left(X_n = \frac{1}{n}\right) &= 1 - \frac{1}{n^2} \\ \mathbb{P}(X_n = n) &= \frac{1}{n^2} \end{aligned}$$

¿Existe C tal que X_n converge en probabilidad a C? ¿Existe C tal que X_n converge en media cuadrática a C?

b) Sean X, Y va's tales que $\mathbb{E}(X) = \mathbb{E}(Y) = \theta$, $\mathrm{Var}(X) = \sigma^2$ y $\mathrm{Var}(Y) = k\sigma^2$. Sean $X_1, \ldots, X_n$ e $Y_1, \ldots, Y_m$ m.a.s. de X e Y respectivamente y $\hat{\theta} = \alpha\overline{X} + \beta\overline{Y}$. Determine α y β (constantes ambos) para que $\hat{\theta}$ sea estimador insesgado y de varianza mínima.

c) Se desea estudiar la variable X: número de vehículos por hora que cruzan con luz roja un determinado semáforo. La variable X se puede modelar por una distribución de Poisson(λ). Para estimar λ, un estudiante decide tomar una muestra de tamaño 2 y usar $\hat{\lambda} = \frac{1}{2}(X_1 + X_2)$. Otro estudiante (un poco más flojo) recibe la siguiente información: Parece que podría ser 1, por lo que decide usar $\hat{\lambda} = 1$. En términos del error cuadrático medio, determine cuando un estimador es preferible a otro.

P9. Se sospecha que cierto río está altamente contaminado con bacterias, superando el nivel permitido (que es de 200 por Unidad de Volumen (UV)). Se toma una muestra de tamaño 10 (ie, 10 UV). Contando la cantidad de bacterias los resultados fueron los siguientes: 175, 190, 215, 198, 184, 207, 210, 193, 196, 180. Suponga que la cantidad de bacterias por UV tiene distribución $\mathcal{N}(\mu, \sigma^2)$.

a) Suponga que $\sigma^2 = 169$. Determine el intervalo de confianza para μ con conianza 0,99.

b) Un estudiante dice que no hay problemas de contaminación ya que había calculado un intervalo de confianza simétrico para μ obteniendo $(190,6; 199)$. ¿Qué confianza tiene este intervalo? ¿Qué puede concluir de *a*) y de lo dicho por el estudiante?

c) Determine un intervalo de confianza del 95 % para σ^2.

P10. *a*) Sea $X \sim exp(\alpha)$, $\theta = 1/\alpha$ parámetro que se desea estudiar. Sea $X_1, ..., X_n$ m.a.

1) Se propone el estimador de θ

$$\hat{\theta} = k\bar{X}$$

Determine k para minimizar el error cuadrático medio.

2) Determine el *E.M.V.* de θ y para n grande indique su distribución de probabilidades.

b) Se dice que X tiene distribución de Maxwell *ssi*

$$f_X(x) = \frac{\sqrt{2}x^2 e^{-x^2/2\sigma^2}}{\sqrt{\pi}\sigma^3} \quad x > 0$$

Encuentre el *EMV* de σ basado en una m.a.s. $X_1, ..., X_n$, con distribución de Maxwell.

P11. Sea X una v.a. con densidad $f(x) = \lambda e^{-\lambda(x-c)} \cdot 1_{\{x>c\}}$ y $X_1, \ldots, X_n$ una m.a.s.

a) Encuentre el EMV $\hat{\lambda}$ de λ con c conocido.

b) Encuentre el EMV $\hat{c}$ de c con λ conocido.

c) Muestre que $Y = X - c \sim \exp(\lambda)$ y usando esto determine $\mathbb{E}(\hat{\beta})$, $\text{Var}(\hat{\beta})$ con $\beta = 1/\lambda$.

d) Suponga n grande y $c = 0$. Contruya un intervalo de confianza para β con confianza $1 - \alpha$.

e) Una empresa fabrica equipos cuya duración es una v.a. con densidad exponencial. La empresa afirma que la duración promedio de sus equipos es mayor a 90 hrs para una muestra de tamaño $n = 100$ se obtiene $\overline{X} = 100$. Usando *c*), determine un intervalo de confianza para la duración promedio con confianza 0.95. ¿Qué se puede concluir? ¿Cuál es el máximo nivel de confianza para el cual usted le encuentra razón a lo afirmado por la empresa?

P12. *a*) Sean $X_1, ..., X_n$ observaciones *i.i.d.* del modelo de Rayleigh, cuya función de distribución es $F_X(x|\theta) = 1 - e^{-x^2/(2\theta^2)}$, $x \geq 0$, donde $\theta > 0$ es un parámetro desconocido.

i) Calcule el EMV de θ^2.

ii) Calcule el ECM de $\hat{\theta^2}$.
Indicación: Puede serle útil considerar $Y = X^2$.

b) Si $X \sim U(0, \theta)$, $X_1, ..., X_n$ m.a.s. de X (n grande). Construya un I.C. para θ con confianza $1 - \alpha$ usando el estimador de los momentos de θ.

P13. Un alumno obsesivo llega todas las mañanas y antes de entrar a clases lanza una moneda (la misma) varias veces (siempre la misma cantidad), contando el número de sellos obtenidos. Después de 6 días usted le pide sus registros y él le entrega: 2,0,1,3,2,2.

Estime, mediante el método de los momentos, el número de veces que se lanza la moneda y la probabilidad de sello.

P14. *a*) Sea θ la temperatura de un cuerpo, la cual se desconoce.
Se dispone de dos termómetros tales que el primero entrega una temperatura T_1 con $\mathbb{E}(T_1) = \theta$ y $\text{Var}(T_1) = \sigma_1^2$, y el segundo entrega una tempratura T_2 con $\mathbb{E}(T_2) = \theta$ y $\text{Var}(T_2) = \sigma_2^2$. Sea $T = \alpha T_1 + \beta T_2$. Determine α, β de manera que T sea un estimador insesgado de θ y de varianza mínima. Explique, conceptualmente, que ocurre con T cuando $\sigma_1^2 < \sigma_2^2$ y cuando $\sigma_1^2 \approx \sigma_2^2$.

b) Sea $X \sim$ Geométrica(p), y $X_1, \ldots, X_n$ muestra aleatoria simple. Sea $\lambda = 1/p$. Construya un intervalo de confianza de nivel $(1-\alpha)$ para el parámetro λ, considerano n grande.

c) Sea $X \sim$ Poisson(λ), $X_1, \ldots, X_n$ muestra aleatoria simple. Se propone el estimador de λ

$$\hat{\lambda} = k \sum_{i=1}^{n-1} (X_{i+1} - X_i)^2.$$

Determine k de modo que $\hat{k}$ sea estimador insesgado de λ.

d) La longitud de la ballena azul distribuye $N(\mu, \sigma^2)$, con $\sigma = 7{,}5$. En un estudio estadístico realizado a 25 ejemplares se ha obtenido el intervalo de confianza, no simétrico, $(20{,}6,\ 26{,}2)$, para la longitud media. Calcule el nivel de confianza con el que se ha construido dicho intervalo.

P15. Las estaturas de hombres (H) y mujeres (M) siguen una distribución $N(170, 16)$ y $N(165, 9)$ respectivamente (medidas en cm.).

¿Cuál es el recorrido real de H y M? Si se selecciona al azar y de manera independiente una mujer y un hombre, calcule la probabilidad de que la mujer sea más alta que el hombre.