Facultad de Ingeniería
IMERL
PROBABILIDAD Y ESTADÍSTICA
Curso 2015
Segundo Parcial, 2 de julio de 2015
Resumen de la solución

Parte de Múltiple Opción (Total: 20 puntos)

Respuesta correcta: 5 puntos	Respuesta incorrecta: -1 punto	Respuesta en blanco: 0 punto

- Se realiza una encuesta a 100 personas preguntándoles cuántas personas viven en su casa.
  El resultado fue el siguiente:

Resultado	1	2	3	4	5
Frecuencia	11	23	34	19	13

  Sea X la variable aleatoria que cuenta cuántas personas viven en una casa elegida al azar en la ciudad en la que se hizo el estudio. Se realiza la prueba de hipótesis: $\begin{cases} H_0 : E(X) = 3 \\ H_1 : E(X) > 3 \end{cases}$. Si la región crítica es $R = \{\overline{X}_n > 3.25\}$.
  El valor aproximado de la probabilidad de error de tipo I es (dar el valor más cercano a α):

  RESPUESTA CORRECTA $\alpha \cong 0.017$

- Se dispone de una muestra i.i.d. $X_1, X_2, \cdots, X_n$ de datos normales con media μ y desvío σ (conocido) y se considera la prueba de hipótesis $\begin{cases} H_0 : \mu = 0 \\ H_1 : \mu > 0 \end{cases}$, con región crítica $R = \{\overline{X}_n > 1,645\,\sigma/\sqrt{n}\}$. Sabiendo que se cumple $\mu = 1$, la probabilidad de error tipo II (β) de la prueba es:

  RESPUESTA CORRECTA $\beta = \Phi\,(1,645\, - \sqrt{n}/\sigma)$.

- Sean $X_1, \cdots, X_n, \cdots$ variables aleatorias *i.i.d.* exponenciales de parámetro λ.
  Para cada $i = 1, 2, \cdots$ se definen: $Y_i = \overline{X}_i = \frac{X_1+\cdots+X_i}{i}$, $Z_i = min\{X_1, \cdots, X_i\}$.
  Considerar las siguientes afirmaciones:

  i. Y_i y Z_i son independientes para todo i.
  ii. Y_i converge casi seguro a una constante cuando i tiende a infinito.
  iii. $E(\frac{Z_1+\cdots+Z_n}{n}) = \frac{1}{n\lambda}$.

  RESPUESTA CORRECTA Sólo la afirmación *ii* es correcta.

- Sean $X_1, \ldots, X_n$ variables aleatorias i.i.d. $\mathcal{U}(\theta - 1, \theta + 1)$, con $\theta \in R$, $X_1^* = min\{X_1, \ldots, X_n\}$, $X_n^* = max\{X_1, \ldots, X_n\}$ y $\hat{\theta}_{MM}$, $\hat{\theta}_{MV}$ los estimadores de θ por el Método de los Momentos y Máxima Verosimilitud respectivamente. Entonces se cumple que:

  RESPUESTA CORRECTA $\hat{\theta}_{MM} = \bar{X}_n$ y $\hat{\theta}_{MV}$ es cualquier valor en el intervalo $[X_n^* - 1, X_1^* + 1]$.

De modo que para el **Cod. 273** las respuestas correctas de la parte múltiple opción son:

PREGUNTA	**1**	**2**	**3**	**4**
RESPUESTA	B	D	E	E

De modo que para el **Cod. 354** las respuestas correctas de la parte múltiple opción son:

PREGUNTA	**1**	**2**	**3**	**4**
RESPUESTA	C	F	A	F

Parte de desarrollo (Total: 40 puntos)

(En esta parte no se penalizan los errores con puntajes negativos)

Ejercicio 1. (28 puntos) Sea X una variable aleatoria absolutamente continua con densidad dada por

$$f(x,\theta) = \begin{cases} \frac{2x}{\theta} e^{-x^2/\theta} & \text{si } x \geq 0; \\ 0 & \text{si } x < 0, \end{cases}$$

en donde $\theta > 0$ es un parámetro desconocido.

(a) Mostrar que $E(X^2) = \theta$.

$$E(X^2) = \int_0^{+\infty} x^2 f(x,\theta)\, dx = \int_0^{+\infty} x^2 \frac{2x}{\theta} e^{-x^2/\theta}\, dx = (partes) =$$

$$x^2 e^{-x^2/\theta} \mid_0^{+\infty} + \int_0^{+\infty} 2x e^{-x^2/\theta}\, dx = 0 + \theta \int_0^{+\infty} \frac{2x}{\theta} e^{-x^2/\theta}\, dx = \theta.$$

(b) Sea $X_1, \ldots, X_n$ una muestra aleatoria simple de X. Hallar $\hat{\theta} = \text{EMV}(\theta)$ el estimador de máxima verosimilitud de θ.

$$\log\left(\prod_{i=1}^n f(X_i,\theta)\right) = \sum_{i=1}^n \log(X_i) - n\log(\theta) - \sum_{i=1}^n X_i^2/\theta.$$

que alcanza su máximo en

$$\hat{\theta} = \sum_{i=1}^n X_i^2/n.$$

(c) ¿Es $\hat{\theta}$ consistente? Justificar.
Sí. Por la ley fuerte de los grandes números $\hat{\theta}$ converge casi seguramente a $E(X^2) = \theta$.

(d) Mostrar que $\hat{\theta}$ es insesgado.

$$E(\hat{\theta}) = E\left(\sum_{i=1}^n X_i^2/n\right) = \sum_{i=1}^n E(X_i^2)/n = n\theta/n = \theta.$$

(e) Supongamos que se dispone de una muestra $X_1, \cdots, X_n$ con n suficientemente grande. Dar un intervalo asintótico de confianza $1-\alpha$ para θ para $\alpha \in (0,1)$.

Se considera la muestra $Y_1, \cdots, Y_n$ tal que $Y_i = X_i^2$. El intervalo de confianza en cuestión es:

$$[\overline{Y} - z_{1-\alpha/2}.s_Y/\sqrt{n}\,,\, \overline{Y} + z_{1-\alpha/2}.s_Y/\sqrt{n}]$$

donde

$$\overline{Y} = \sum_{i=1}^{n} Y_i/n,$$

$$s_Y^2 = \sum_{i=1}^{n} (Y_i - \overline{Y})^2/n$$

y $z_{1-\alpha/2}$ es el percentil $1-\alpha/2$ de la distribución normal.

Ejercicio 2. (12 puntos)
La siguiente tabla contiene una muestra de tamaño $n = 10$, de una variable X que tiene distribución exponencial de parámetro $\lambda > 0$ desconocido.

1.97	0.95	2.97	0.23	1.33	1.29	2.70	0.49	0.24	0.11

(a) Realizar **dos** pruebas de hipótesis para determinar si la muestra es i.i.d..

Prueba de Rachas up-down: R=7, p-valor=0,4524.
Prueba de Spearman: s=-0,5273, p-valor=0,062.
Se acepta la hipótesis de aleatoriedad.

(b) Realizar **una** prueba de hipótesis para determinar si la muestra se ajusta a una distribución exponencial.

Prueba de exponencialidad de Lilliefors : $KSL = 0,1502 < 0,341$.
Se acepta la hipótesis de exponencialidad.

(c) Sabiendo que para una muestra de tamaño 100 de dicha distribución, exactamente 40 de los valores son mayores a 1, estimar el parámetro λ.

$$e^{-\hat{\lambda}} = 40/100$$

por lo tanto

$$\hat{\lambda} = -\log(0,4) = 0.9163.$$