Probabilidad y Estadística
Examen, 16 de julio de 2011

Datos del estudiante

Nro. de parcial	Nombre y apellido	Cédula

- **La duración del examen es de 4 horas.**
- **Publicación de resultados: viernes 29 de julio, 20:00 hs.**
- **Muestra de exámenes: martes 2 de agosto, 14:00 hs.**

Ejercicio 1
En un programa de televisión hay un concurso que consiste en sortear una llave de un total de M (que se desconoce) y probar si abre la puerta de un auto. Los integrantes del público van pasando sucesivamente hasta que alguno extrae la llave ganadora. Las llaves utilizadas no se reponen.

1. ¿Es cierto que el primer jugador tiene mayor probabilidad de ganar que el segundo? ¿Cuáles son dichas probabilidades?
2. Sea X el número de intentos hasta obtener la llave ganadora. Halle $E(X)$ y $Var(X)$.
3. Halle un estimador consistente e insesgado de M (pruebe que cumple ambas condiciones). Calcule su error cuadrático medio.
4. Se emiten 400 programas. En promedio, el número de intentos hasta obtener la llave ganadora fue 9. Estime M sabiendo que no varía de programa en programa. Estime $Var(X)$ y halle un intervalo de confianza 90 % aproximado para M.
5. Construya el estimador de máxima verosimilitud de M.

Puede ser útil recordar que $\sum_{i=1}^{M} i = M(M+1)/2$ *y* $\sum_{i=1}^{M} i^2 = M(M+1)(2M+1)/6$.

Ejercicio 2
Los siguientes datos corresponden errores de medición en un termómetro (en milésimas de grado).

0,178	0,799	-0,027	0,355	-0,677	0,286	0,721	0,553

1. Estudiar si puede asumirse que los datos son iid. Realizar dos tests de hipótesis al 10 % e indicar el p-valor.
2. Estudiar si puede asumirse que los datos tienen distribución uniforme en el intervalo $[-1, 1]$. Realizar un test de hipótesis al 10 % e indicar el p-valor.

3. Estudiar si puede asumirse que los datos tienen la siguiente densidad:

$$f(t) = \begin{cases} t+1 & \text{si } t \in [-1,0] \\ -t+1 & \text{si } t \in [0,1] \end{cases}.$$

   a) Calcular y graficar la función de distribución.

   b) Realizar un test de hipótesis al 10 % e indicar el p-valor.

Ejercicio 3

Se tiene el siguiente esquema de una red. Cada nodo está conectado con los que comparte una

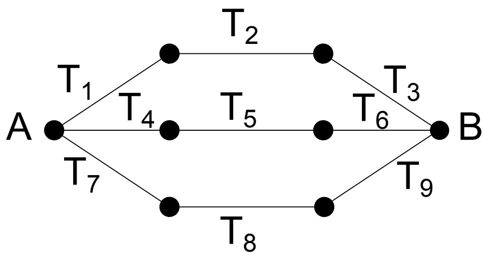

Figura 1: Red.

arista durante un tiempo aleatorio. El sistema comienza con todos los nodos conectados y se modela el tiempo que dura la conexión (en horas) mediante variables aleatorias exponenciales i.i.d., es decir que $T_i \sim \exp(\lambda)$ para todo $i = 1, \ldots, 9$. Para que A y B permanezcan conectados deben estar funcionando los enlaces 1, 2 y 3, o los enlaces 4, 5 y 6 o los enlaces 7, 8 y 9, y la red de la figura 1 es equivalente a la red de la figura 2.

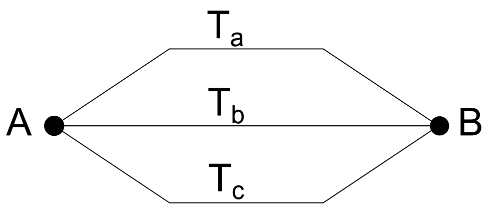

Figura 2: Red simplificada.

1. Calcular las distribuciones de T_a, T_b y T_c (los tiempos de conexión entre A y B en cada camino).
2. Sea T el tiempo de conexión entre A y B. Hallar $P(T \leq x)$ para $x \geq 0$ en función de λ.
3. Hallar λ para que $P(T > 1) \geq 0,999$.
4. Sabiendo que los enlaces 4 y 7 dejaron de funcionar en el transcurso de la primera hora calcular la probabilidad de que el tiempo de conexión T sea mayor que 2 horas.
5. Calcular $E(T)$ (el tiempo medio de conexión entre A y B) en función de λ.

Examen, 16 de julio de 2011 - Soluciones

Ejercicio 1

1. Consideramos los sucesos: A: el primer jugador es el ganador, B : el segundo jugador es el ganador.

   La probabilidad de que el primer jugador acierte la llave es $P(A) = \dfrac{1}{M}$, y la probabilidad de que el segundo jugador acierte la llave es $P(B) = P(B|A^c)P(A^c) = \dfrac{1}{M-1}\dfrac{M-1}{M} = \dfrac{1}{M}$.

2. La variable X tiene como recorrido $\{1, 2, \dots, M\}$ y toma cada valor con probabilidad $1/M$ (uniforme discreta).
$$E(X) = \sum_{i=1}^{M} i\frac{1}{M} = \frac{M+1}{2},$$
$$Var(X) = \sum_{i=1}^{M} i^2\frac{1}{M} - \left(\frac{M+1}{2}\right)^2 = \frac{M^2-1}{12}.$$

3. Como $E(X) = \frac{M+1}{2}$ tenemos, por el método de los momentos
$$\overline{X}_n = \frac{\hat{M}+1}{2}, \quad \hat{M} = 2\overline{X}_n - 1.$$
   Es consistente porque por la LFGN $\overline{X}_n \to_n \dfrac{M+1}{2}$ c.s y como $2x-1$ es una función continua $\hat{M} \to_n M$ c.s.

   Además es insesgado: usando linealidad de la esperanza
$$E(\hat{M}) = E(2\overline{X}_n - 1) = 2E(\overline{X}_n) - 1$$
   y también por linealidad
$$E(\overline{X}_n) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{M+1}{2},$$
   por ser iid.

   El error cuadrático medio en este caso (por ser insesgado) es
$$ECM(\hat{M}) = Var(\hat{M}) = 4Var(\overline{X}_n) = \frac{4}{n}\frac{M^2-1}{12} = \frac{M^2-1}{3n},$$
   usando propiedades de la varianza.

4. Para $n = 400$ se tiene que $\overline{X}_n = 9$. Hacemos un intervalo aproximado para $E(X_1) = \frac{M+1}{2}$ de la forma
$$\left[\overline{X}_n - \frac{\hat{\sigma}}{\sqrt{n}}z_{\alpha/2}, \overline{X}_n + \frac{\hat{\sigma}}{\sqrt{n}}z_{\alpha/2}\right]$$
   donde $\hat{\sigma}$ es un estimador de la varianza de X_1. Como $\sigma = \sqrt{\dfrac{M^2-1}{12}}$ tenemos que
$$\hat{\sigma} = \sqrt{\frac{\hat{M}^2-1}{12}} = \sqrt{\frac{\overline{X}_n(\overline{X}_n-1)}{3}}.$$

El intervalo para $\frac{M+1}{2}$ es $[8,58; 9,41]$ de donde el intervalo para M es $[16,17; 17,83]$.

5. Como el recorrido de la variable X es $R_X = \{1, 2, \ldots, M\}$ tenemos que, dada la muestra $X_1, \ldots X_n$ iid la función de verosimilitud es
$$V(M) = \begin{cases} \frac{1}{M^n} & \text{si } X_i \in R_X \text{ para todo } i = 1, \ldots n \\ 0 & \text{si no} \end{cases},$$
de donde el estimador de máxima verosimilitud es $\tilde{M} = \max(X_1, \ldots X_n)$.

Ejercicio 2

1. Para el test de rachas tenemos que $R = 6$ y el p-valor es $\alpha^* = 0,325$. Para el test de correlación de rangos de Spearman tenemos que $R_S = 0,214$ y el p-valor es $\alpha^* = 0,310$. Como el resultado de ambos tests es aceptar que los datos son iid asumimos que son iid.
2. Realizamos el test de Kolmogorov-Smirnov. La función de distribución es
$$F_0(t) = \begin{cases} 0 & t < -1 \\ \frac{1}{2}(t+1) & t \in [-1, 1] \\ 1 & t > 1 \end{cases}.$$
El estadístico es $D = 0,3615$. La región crítica al 10 % está dada por $\{D \geq 0,410\}$ por lo que se acepta H_0. Además el p-valor es $0,1 < \alpha^* < 0,2$.
3. La función de distribución en este caso es
$$G_0(t) = \begin{cases} 0 & t < -1 \\ \frac{1}{2}t^2 + t + \frac{1}{2} & t \in [-1, 0] \\ -\frac{1}{2}t^2 + t + \frac{1}{2} & t \in [0, 1] \\ 1 & t > 1 \end{cases}.$$
El estadístico es $D = 0,412$. La región crítica al 10 % está dada por $\{D \geq 0,410\}$ por lo que se rechaza H_0. Además el p-valor es $0,05 < \alpha^* < 0,1$.

Ejercicio 3

1. $T_a = \min(T_1, T_2, T_3)$ por lo que $T_a \sim \exp(3\lambda)$. Idem T_b y T_c.
2. $T = \max(T_a, T_b, T_c)$ por lo que
$$P(T \leq x) = P(T_a \leq x, T_b \leq x, T_c \leq x) = (1 - e^{-3\lambda x})^3.$$
3. $P(T > 1) = 1 - P(T \leq 1) = 1 - (1 - e^{-3\lambda})^3 \geq 0,999$, de donde $\lambda = 0,035$.
4. $$P(T > 2 | T_b < 1, T_c < 1) = \frac{P(T > 2, T_b < 1, T_c < 1)}{P(T_b < 1, T_c < 1)} = \frac{P(T_a > 2, T_b < 1, T_c < 1)}{P(T_b < 1, T_c < 1)} = P(T_a > 2) = e^{-3\lambda}.$$
5. $$\begin{aligned} E(T) &= \int_0^{+\infty} (1 - P(X \leq x))dx = \int_0^{+\infty} (1 - (1 - e^{-3\lambda x})^3)dx \\ &= \int_0^{+\infty} (3e^{-3\lambda x} - 3e^{-6\lambda x} + e^{-9\lambda x})dx = \frac{11}{18}\frac{1}{\lambda} \end{aligned}$$