PROBABILIDAD Y ESTADÍSTICA

EXAMEN, 7 de febrero de 2011

DATOS DEL ESTUDIANTE

No. de Examen	Nombre y Apellido	Cédula

- **La duración del examen es de 4 horas.**
- **Publicación de resultados: Lunes 21 de febrero, 20:00 hs.**
- **Muestra de exámenes: Miércoles 23 de febrero, 14:00 hs.**

Problema 1

Dos jugadores A y B juegan el siguiente juego.

- Al principio cada uno apuesta un peso y luego cada uno tira una moneda equilibrada.
- Si el jugador A saca cruz y el jugador B cara, termina el juego, gana A y se lleva el dinero.
- Si el jugador B saca cruz y el jugador A cara, termina el juego, gana B y se lleva el dinero.
- En cualquier otro caso hay empate. Si empatan en la primera jugada, en la segunda cada uno apuesta dos pesos y así sucesivamente cada vez que hay empate A y B duplican lo apostado anteriormente antes de la siguiente jugada (de modo que el pozo se duplica cada vez).

El juego sigue hasta que haya un ganador o hasta que en la vigésima jugada haya empate, en cuyo caso los jugadores se dividen el pozo, recuperando cada uno lo apostado. Para todo entero entre 0 y n se consideran los eventos:

- A_n : "el juego se termina en la n-ésima jugada y gana A";
- B_n : "el juego se termina en la n-ésima jugada y gana B";
- N_n : "hay empate en la n-ésima jugada".

Sean $x_n = P(A_n)$, $y_n = P(B_n)$ y $z_n = P(N_n)$.

1. Calcule x_1, y_1 y z_1.
2. Pruebe que $x_n = \left(\frac{1}{2}\right)^{n+1}$, $y_n = \left(\frac{1}{2}\right)^{n+1}$ y $z_n = \left(\frac{1}{2}\right)^{n}$.
3. Se considera la variable aleatoria $X =$"cantidad de pesos en el pozo de la última jugada".
   (a) ¿Cuál es el recorrido de X? Justifique.
   (b) Calcule $P(X = 2^k)$ para $k = 1, \ldots, 20$.
   (c) Calcule $E(X)$ usando la definición de valor esperado.

Problema 2

Se considera F_1 la función de distribución de una variable aleatoria $N(1,1)$ (media 1 y varianza 1) y F_2 la función de distribución de una variable aleatoria $N(2,4)$ (media 2 y varianza 4). Se define la función F como:

$$F(t) = 0,2F_1(t) + 0,8F_2(t).$$

1. Pruebe que F es una función de distribución.
2. Escriba la función F en función de la función Φ, siendo Φ la función de distribución de una variable aleatoria $N(0,1)$.
3. Además se considera el siguiente experimento. En una planilla A se tiene una lista de datos iid que se sabe siguen una distribución $N(1,1)$. En otra planilla B se tiene otra lista de datos iid que se sabe siguen una distribución $N(2,4)$. Se sabe además que los datos de la planilla A y de la planilla B son independientes entre sí. Se sortea al azar, de manera independiente, n veces, un número de dos posibles (1 y 2), con probabilidad $0,2$ de que salga el 1 y $0,8$ de que salga el 2. Si sale el número 1, se toma un dato de la planilla A y si sale el 2, se toma de la planilla B. Justifique que la muestra de n datos así obtenida sigue la distribución F.
4. Se considera la siguiente muestra:

4,87	7,82	1,86	1,82	4,44	3,03	4,72	2,06	0,4	0,29	5,93	-0,69

   (a) Estudie si es iid (realice dos pruebas de hipótesis).
   (b) Indique si la muestra se ajusta a la distribución F definida antes (realice una prueba de hipótesis).

Problema 3

Se sabe que la variable aleatoria X que modela el tiempo de vida de determinados ventiladores tiene la siguiente densidad.

$$f(x) = \begin{cases} cxe^{-x^2} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$

Se dice que un ventilador es de excelente calidad si su tiempo de vida supera las dos unidades de tiempo.

1. Halle el valor de la constante c y el tiempo medio de vida de los ventiladores.
2. Calcule la probabilidad de que un ventilador sea de excelente calidad.
3. Si se sabe que el tiempo de vida de un ventilador superó dos unidades de tiempo (es decir que el ventilador es de excelente calidad), calcule la probabilidad de que supere cuatro unidades de tiempo.
4. En un supermercado se dispone de 5 ventiladores. Sea Y la variable aleatoria que cuenta la cantidad de ventiladores de excelente calidad en dicho supermercado. Halle y grafique la función de probabilidad de Y
5. En un lote de mil ventiladores 12 resultaron de excelente calidad. Indique si es razonable suponer en este caso que la probabilidad de que un ventilador sea de excelente calidad es la hallada en la pregunta 2. Justifique mediante un intervalo de confianza o una prueba de hipótesis.

EXAMEN, 7 de febrero de 2011 - SOLUCIONES

Problema 1

1. $x_1 = 1/4$, $y_1 = 1/4$, $z_1 = 1/2$.

2. En cada jugada la probabilidad de empate es 1/2 (ambos sacan cara o ambos sacan cruz), y la probabilidades de que ganen A y B son 1/4 cada una. Para que ocurra A_n debe haber $n-1$ empates (con probabilidad $(1/2)^{n-1}$) y en la última jugada debe salir una cruz y una cara (con probabilidad 1/4). Idem para B_n, pero en la última debe salir una cara y una cruz. De donde:
$$x_n = y_n = \left(\frac{1}{2}\right)^{n-1}\left(\frac{1}{4}\right) = \left(\frac{1}{2}\right)^{n+1}.$$
z_n es la probabilidad de n empates, es decir:
$$z_n = \left(\frac{1}{2}\right)^{n}.$$

3. Como en cada jugada se duplica la apuesta en caso de empate el recorrido es $\{2, 4, 8, \ldots, 2^{20}\}$.

4. Para $k = 1, \ldots, 19$
$$P(X = 2^k) = x_k + y_k = 2\left(\frac{1}{2}\right)^{k+1} = \left(\frac{1}{2}\right)^{k}.$$
$$P(X = 20) = z_19 = \left(\frac{1}{2}\right)^{19}.$$
Observar que $\sum_{k=1}^{19}\left(\frac{1}{2}\right)^k + \left(\frac{1}{2}\right)^{19} = \frac{1/2-(1/2)^{20}}{1-1/2} + \left(\frac{1}{2}\right)^{19} = 1.$

5. $$E(X) = \sum_{k=1}^{20} 2^k P(X = 2^k) = \sum_{k=1}^{19} 2^k \left(\frac{1}{2}\right)^k + 2^{20}\left(\frac{1}{2}\right)^{19} = 21.$$

Problema 2

1. F es un función de distribución.
   - $0 \le F_1(t) \le 1$, $0 \le F_2(t) \le 1$, de donde $0 \le F(t) \le 1$.
   - F es continua por ser combinación lineal de funciones continuas (y por lo tanto continua por la derecha).
   - F es creciente pues $F\prime_1(t) > 0$, $F\prime_2(t) > 0$, de donde $F\prime(t) = 0,2F_1\prime(t) + 0,8F_2\prime(t) > 0$.
   - Por ser F_1 y F_2 funciones de distribución
$$\lim_{t\to-\infty} F(t) = 0,2\lim_{t\to-\infty} F_1(t) + 0,8\lim_{t\to-\infty} F_2(t) = 0,$$
$$\lim_{t\to+\infty} F(t) = 0,2\lim_{t\to+\infty} F_1(t) + 0,8\lim_{t\to+\infty} F_2(t) = 0,2 + 0,8 = 1.$$

2. $F_1(t) = \Phi(t-1)$ y $F_2(t) = \Phi\left(\frac{t-2}{2}\right)$, luego
$$F(t) = 0,2\Phi(t-1) - 0.8\Phi\left(\frac{t-2}{2}\right).$$

3. Sea Y el número sorteado (Si $Y = 1$ se elige de la lista A y si $Y = 2$ se elige de B). Entonces $P(Y = 1) = 0,2$ y $P(Y = 2) = 0,8$). Sea X la variable del experimento en 3.

   $$P(X \leq t) = P(X \leq 3|Y = 1)P(Y = 1) + P(X \leq 3|Y = 2)P(Y = 2) = 0,2F_1(t) + 0,8F_2(t).$$

4. (a) Rachas: $R = 8$, p-valor $\alpha^* = 0.5547$, por lo tanto se acepta que los datos sean iid.
   Correlación de rangos de Spearman: $r_S = -0,4825$, p-valor $\alpha^* < 0,1$ por lo tanto se rechaza que los datos sean iid. Observar que si una de las pruebas acepat y otra rechaza no podemos asumir que los datos sean iid.

   (b) Si bien no podemos asumir que los datos sean iid, se realiza la prueba de Kolmogorov Smirnov. Kolmogorov Smirnov: $D = 0,3276$, p-valor $0,1 < \alpha^* < 0,2$, por lo tanto se acepta que los datos tengan distribución F.

Problema 3

1. $$1 = \int_0^{\infty} cxe^{-x^2} = -\frac{c}{2}e^{-x^2}\Big|_0^{+\infty} = \frac{c}{2},$$

   de donde $c = 2$. Además si X es el tiempo de vida de los ventiladores, X tiene que

   $$E(X) = \int_0^{+\infty} 2x^2 e^{-x^2} dx,$$

   haciendo partes con x y $2xe^{-x^2}$

   $$E(X) = -xe^{-x^2}|_0^{+\infty} = \int_0^{+\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{2}.$$

   (Para calcular $\int_0^{+\infty} e^{-x^2} dx$ se puede hacer el cambio de variable $u = x/\sqrt{2}$ y usar los valores de la integral de la densidad de una variable normal estándar.)

2. $$P(X > 2) = \int_2^{+\infty} 2xe^{-x^2} dx = e^{-x^2}\Big|_0^{+\infty} = e^{-4}.$$

3. $$P(X > 4|X > 2) = \frac{P(X > 4, X > 2)}{P(X > 2)} = \frac{P(X > 4)}{P(X > 2)} = \frac{e^{-16}}{e^{-4}} = e^{-12}.$$

4. $Y \sim \mathrm{Bin}(5, \mathrm{e}^{-4})$, de donde $P(Y = k) = C_k^5 e^{-4k}(1 - e^{-4})^{5-k}$, $k = 0, \ldots, 5$.

5. Sea p la probabilidad de tener un ventilador de excelente calidad. Un estimador de p es $\widehat{p} = 0,012$. Se considera la siguiente prueba de hipótesis de proporciones:

   $$\begin{cases} p = e^{-4} \approx 0,0183 \\ p < e^{-4} \end{cases}$$

   La región crítica al nivel α es $\left\{ \frac{\sqrt{n}(\widehat{p} - e^{-4})}{\sqrt{\widehat{p}(1-\widehat{p})}} \leq -z_\alpha \right\}$ y el p-valor es $\alpha^* = 0,034$, de donde se puede asumir que $p < e^{-4}$. También se puede realizar el intervalo de confianza al nivel α, $\left[\widehat{p} - z_{\alpha/2}\frac{\sqrt{\widehat{p}(1-\widehat{p})}}{\sqrt{n}}, \widehat{p} + z_{\alpha/2}\frac{\sqrt{\widehat{p}(1-\widehat{p})}}{\sqrt{n}}\right]$. Para el nivel 0,05 el intervalo es $[0,0052; 0,0187]$, que contiene al valor e^{-4} por lo que según este criterio se podría asumir que la probabilidad es e^{-4}.