PROBABILIDAD Y ESTADÍSTICA
EXAMEN DICIEMBRE 2005

DATOS DEL ESTUDIANTE

No. de Examen	Apellidos y Nombres	Cédula

- **La duración del examen es de 3 horas y media.**
- **El puntaje mínimo para aprobar es 50 puntos.**
- **Publicación de resultados: 29 de diciembre - 18:00 hs**
- **Muestra de exámenes: 30 de diciembre - 14:00 hs**

Problema 1 (30 puntos)

Considere una caja que inicialmente contiene 10 bolillas: 6 azules y 4 rojas. Considere además el procedimiento que se describe a continuación. Se elige una bolilla al azar de la caja: si la bolilla es azul se la pinta de rojo y si la bolilla es roja se la pinta de azul. Luego se devuelve la bolilla a la caja.

Se denota con N_1 el número de bolillas rojas al final del procedimiento.

a) **(6 puntos)** Calcule los posibles valores de la variable aleatoria N_1 y sus correspondientes probabilidades.

Suponga ahora que se repite nuevamente el procedimiento anterior. Se denota con N_2 el número de bolillas rojas que quedan al final.

b) **(10 puntos)** Calcule los posibles valores de la variable aleatoria N_2 y sus correspondientes probabilidades.

c) **(6 puntos)** Calcule $\mathbf{E}(N_2)$ y $\mathbf{Var}(N_2)$.

d) **(8 puntos)** Calcule la probabilidad de que $N_1 = 5$ dado que $N_2 = 4$.

Problema 2 (34 puntos)

Considere un círculo de radio X, donde X es una variable aleatoria exponencial de parámetro λ. Se denota con A el área del círculo.

a) **(9 puntos)** Calcule la función de distribución de la variable aleatoria A.

b) **(9 puntos)** Calcule $\mathbf{E}(A)$.

Considere ahora un nuevo círculo de radio Y, donde Y es una variable aleatoria *independiente* de X, con distribución exponencial de parámetro λ (el mismo parámetro que la variable X). Suponga que los centros de los círculos están separados por una distancia D.

c) **(16 puntos)** Pruebe que la probabilidad de que los dos círculos se superpongan es

$$\mathbf{P}(\text{Superposición}) = (1 + \lambda D)e^{-\lambda D}.$$

Problema 3 (36 puntos)

El tiempo de funcionamiento de una componente electrónica se modela con una variable aleatoria continua T, cuya densidad de probabilidad es:

$$f_\theta(t) = \begin{cases} \frac{t}{\theta^2}\, e^{\frac{-t}{\theta}} & \text{si } t \geq 0 \\ 0 \quad \text{si } t < 0 \end{cases} \tag{1}$$

donde $\theta > 0$.

Considere $T_1, T_2, \ldots, T_n$ una muestra i.i.d. correspondiente a los tiempos de funcionamiento de las componentes electrónicas en cuestión.

a) **(10 puntos)** Calcule el estimador de máxima verosimilitud para θ.

b) **(6 puntos)** Estime θ por el método de los momentos.

Se dispone ahora de una muestra de tiempos de vida de dichas componentes

18	16	43	30	17	73	59	63	22	39

Nota: En las pruebas de hipótesis utilice el siguiente criterio de decisión: se acepta la hipótesis nula si el p-valor es superior a 0.1

c) **(6 puntos)** Estudie la aleatoriedad de la muestra (realice una sola prueba).

d) **(14 puntos)** Implemente el test de Kolmogorov-Smirnov para verificar si los datos ajustan a la distribución correspondiente a la densidad definida en (1) con $\theta = 19$.

SOLUCIÓN

Problema 1

a) Los posibles valores son $N_1 = 3$ y $N_1 = 5$. Además:

$$\mathbf{P}(N_1 = 3) = \frac{4}{10} \quad \text{y} \quad \mathbf{P}(N_1 = 5) = \frac{6}{10}.$$

b) Los posibles valores son $N_2 = 2$, $N_2 = 4$ y $N_2 = 6$. Por otra parte:

$$\mathbf{P}(N_2 = 2) = \frac{4}{10} \times \frac{3}{10} = \frac{12}{100}$$

$$\mathbf{P}(N_2 = 4) = \frac{6}{10} \times \frac{5}{10} + \frac{4}{10} \times \frac{7}{10} = \frac{58}{100}$$

$$\mathbf{P}(N_2 = 6) = \frac{6}{10} \times \frac{5}{10} = \frac{30}{100}$$

c) De los resultados anteriores resulta:

$$\mathbf{E}(N_2) = 2\frac{12}{100} + 4\frac{58}{100} + 6\frac{30}{100} = \frac{436}{100} = 4.36$$

Por otra parte:

$$\mathbf{E}(N_2^2) = 4\frac{12}{100} + 16\frac{58}{100} + 36\frac{30}{100} = \frac{2056}{100} = 20.56$$

de manera que

$$\mathbf{Var}(N_2) = 20.56 - 19.0096 = 1.5504$$

d) De la definición de probabilidad condicional y de los cálculos anteriores, se tiene que:

$$\mathbf{P}(N_1 = 5 | N_2 = 4) = \frac{\mathbf{P}(\{N_1 = 5\} \cap \{N_2 = 4\})}{\mathbf{P}(N_2 = 4)} = \frac{\frac{6}{10} \times \frac{5}{10}}{\frac{58}{100}} = 0.517$$

Problema 2

El área de un círculo de radio X es $A = \pi X^2$.

a) De la definición de función de distribución: $F_A(t) = \mathbf{P}(A \leq t)$. Así que:

- Si $t < 0$ entonces $F_A(t) = 0$.

– Si $t \geq 0$ entonces

$$F_A(t) = \mathbf{P}(A \leq t) = \mathbf{P}(\pi X^2 \leq t) = \mathbf{P}\left(X \leq \sqrt{\frac{t}{\pi}}\right) = 1 - e^{-\lambda\sqrt{\frac{t}{\pi}}},$$

donde en la última igualdad hemos usado la función de distribución de una variable exponencial.

b)

$$\mathbf{E}(A) = \pi\mathbf{E}(X^2) = \pi\frac{2}{\lambda^2},$$

donde hemos usado el hecho de que si X es una variable exponencial de parámetro λ entonces $\mathbf{E}(X^2) = \frac{2}{\lambda^2}$.

c) Es fácil ver que dos círculos de radios X e Y, cuyos centros están separados a una distancia D, se superponen si y sólo si $X + Y \geq D$. De manera que:

$$\mathbf{P}(\text{Superposición}) = \mathbf{P}(X + Y \geq D) = 1 - \mathbf{P}(X + Y \leq D).$$

La última probabilidad involucra una suma de variables exponenciales independientes:

$$\mathbf{P}(X + Y \leq D) = \int_0^D dx \int_0^{D-x} dy\, \lambda^2 e^{-\lambda(x+y)} = 1 - (1 + \lambda D)e^{-\lambda D}.$$

Así que finalmente se obtiene:

$$\mathbf{P}(\text{Superposición}) = (1 + \lambda D)e^{-\lambda D}.$$

Problema 3

a) La función de verosimilitud, para $t_1 > 0, t_2 > 0, \ldots, t_n > 0$, es

$$L(\theta) = \left(\frac{t_1}{\theta^2}e^{\frac{-t_1}{\theta}}\right)\left(\frac{t_2}{\theta^2}e^{\frac{-t_2}{\theta}}\right)\cdots\left(\frac{t_n}{\theta^2}e^{\frac{-t_n}{\theta}}\right) = \frac{t_1 t_2 \ldots t_n}{\theta^{2n}}e^{-\frac{t_1+t_2+\ldots+t_n}{\theta}}$$

Tomando logaritmo

$$\ln(L(\theta)) = \ln(t_1 t_2 \ldots t_n) - 2n\ln(\theta) - \frac{t_1 + t_2 + \ldots + t_n}{\theta}$$

y derivando con respecto a θ, se obtiene

$$\frac{\partial(\ln(L(\theta)))}{\partial\theta} = \frac{-2n}{\theta} + \frac{t_1 + t_2 + \ldots + t_n}{\theta^2} = 0$$

De manera que el estimador de máxima verosimilitud es:

$$\widehat{\theta} = \frac{1}{2}\frac{t_1 + t_2 + \ldots + t_n}{n} = \frac{1}{2}\overline{T}_n.$$

b) Es fácil ver que $\mathbf{E}(T) = 2\theta$, por lo tanto el método de los momentos proporciona el siguiente estimador

$$\mathbf{E}\,(T) = \overline{T}_n \Leftrightarrow \widehat{\theta} = \frac{1}{2}\overline{T}_n,$$

que coincide con el calculado anteriormente.

c) – **Test de Rachas:** $R = 7$ y $p - valor = 0.4524$.
   – **Test de correlación de rangos de Spearman:** $r_s = 0.2727$ y $p - valor > 0.235$.

   Conclusión: No se rechaza la hipótesis de aleatoriedad.

d) La función de distribución correspondiente a la densidad (1) con $\theta = 19$ es

$$F_0(t) = 1 - \left(1 + \frac{t}{\theta}\right) e^{-\frac{t}{\theta}}, \quad \text{si } t \geq 0$$

y $F_0(t) = 0$ si $t < 0$.

   – **Test de Kolmogorov-Smirnov:** $D = 0.2064$ y $p - valor > 0.2$.

   Conclusión: Se acepta la hipótesis de que la muestra ajusta a la densidad (1) con $\theta = 19$.