# EXAMEN AGOSTO 2003
# PROBABILIDAD Y ESTADÍSTICA

## DATOS DEL ESTUDIANTE

| Apellidos y Nombres | Cédula |
|---|---|
| | |

| |
|---|
| Nombre del docente de teórico: |
| Nombre del docente de práctico: |

- Si concurre a más de un teórico o un práctico, incluya los nombres de todos sus docentes.
- Si no recuerda el nombre del docente ponga el horario.
- Si no concurre a clase, responda "NO", pero no deje su casillero en blanco.

*ATENCION: No se corregirá ningún parcial en el que al menos uno de los casilleros de información anteriores esté en blanco;* ***llénelos a todos***

---

## Talón para el estudiante - Probabilidad y Estadística - Ex. Agosto 2003

- La duración total de la prueba es de **4 horas**.
- Los resultados serán publicados en la web del curso (http://imerl.fing.edu.uy/pye) y en las carteleras del IMERL el

  **Jueves 21 de Agosto a las 20.00 horas.**

- La muestra de la prueba será el

  **Viernes 22 de Agosto a las 18.00 horas.**

  en salón a anunciar.

**BUENA SUERTE!!**

***Este talón no se entrega, contiene información para el estudiante.***

## Ejercicio 1 (Total: 50 puntos)

1. Considere tres cajas: la Caja 1 contiene 2 bolas rojas y 2 azules, la Caja 2 contiene 3 bolas rojas y 1 azul y la Caja 3 está inicialmente vacía. Se extraen al azar *de manera independiente* una bola de la Caja 1 y una bola de la Caja 2 y se las coloca en la Caja 3.

   (a) **(8 puntos)** Calcule la probabilidad de que en la Caja 3 queden dos bolas rojas.

   (b) **(8 puntos)** Calcule la probabilidad de que en la Caja 3 queden una bola roja y otra azul.

   (c) **(12 puntos)** Si se extrae al azar una bola de la caja 3, ¿cuál es la probabilidad de que ésta sea roja?

   (d) **(8 puntos)** Dado que al extraer una bola al azar de la caja 3, ésta resultó ser roja, ¿cuál es la probabilidad de que la restante bola en la caja también sea roja?

2. Se necesita decidir entre dos hipótesis relativas a la Caja 3:

$$\begin{cases} H_0: & \text{las dos bolas son rojas} \\ H_1: & \text{una de las bolas es roja y la otra azul.} \end{cases}$$

   El criterio para decidir entre una u otra hipótesis es el siguiente: se hacen dos extracciones *con reposición* de la Caja 3 y se decide $H_0$ en caso de que en las dos extracciones se obtenga una bola roja; se decide $H_1$ en caso contrario.

   (a) **(7 puntos)** ¿Cuál es la probabilidad de decidir $H_0$ siendo $H_1$ verdadera?

   (b) **(7 puntos)** ¿Cuál es la probabilidad de decidir $H_1$ siendo $H_0$ verdadera?

## Ejercicio 2 (Total: 50 puntos)

1. **(14 puntos)** A los efectos de calibrar una máquina en una fábrica se tomó la siguiente muestra que corresponde al error en la medida de los *diámetros* de ciertos tornillos producidos por la máquina. Se puede suponer que la muestra es *iid* y con distribución normal.

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $X$ | -1.19 | -2.2 | 0.99 | -0.52 | 0.33 | 0.23 | 0.02 | -1 | -0.95 | -0.37 |

   Esta muestra cumple que $\overline{X}_n = -0.466$ y $s_n = 0.9151$. ¿Es razonable suponer que la media es $\mu = 0$ y la varianza es $\sigma^2 = 1$? Justifique.

2. Para los mismos tornillos se midió también el error cometido en el *largo* de los mismos, obteniéndose la siguiente muestra:

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $Y$ | 1.2 | 1.41 | -1.49 | 0.32 | 0.36 | -1.51 | -1.11 | 1.62 | -0.2 | 0.58 |

   Esta muestra cumple que $\overline{Y}_n = 0.118$ y $s_n = 1.167$.

   (a) **(8 puntos)** Analizar si la muestra $Y$ es *iid*.

   (b) **(8 puntos)** ¿$X$ e $Y$ son independientes entre sí?

   (c) **(14 puntos)** Se supone que $Y$ tiene distribución normal al igual que $X$. ¿Es esto razonable? Justifique.

   (d) **(6 puntos)** Asumiendo el resultado de la parte anterior, estime el error medio mediante un intervalo de confianza al nivel 10%.

**Solución Ejercicio 1**

1. (a) Sean los sucesos $A =$"La bola elegida de la caja 1 es roja" y $B =$"La bola elegida de la caja 2 es roja". Es claro que $\mathbf{P}(A) = \frac{1}{2}$ y $\mathbf{P}(B) = \frac{3}{4}$. Si $C =$"En la caja 3 quedan 2 bolas rojas" se tendrá que $C = A \cap B$. Como las extracciones son independientes, se tiene que:

$$\mathbf{P}(C) = \mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B) = \frac{1}{2}\frac{3}{4} = \frac{3}{8}$$

(b) Con la misma idea que antes, el suceso $D =$"En la caja 3 queda una bola roja y una azul" se escribe como $D = (A \cap B^c) \cup (A^c \cap B)$ y la unión es disjunta. Como $A$ y $B$ independientes, entonces $A$ con $B^c$ y $A^c$ con $B$ también lo son. Por lo tanto:

$$\mathbf{P}(D) = \mathbf{P}(A \cap B^c) + \mathbf{P}(A^c \cap B) = \mathbf{P}(A)\mathbf{P}(B^c) + \mathbf{P}(A^c)\mathbf{P}(B) = \frac{1}{2}\left(1 - \frac{3}{4}\right) + \left(1 - \frac{1}{2}\right)\frac{3}{4} = \frac{1}{2}$$

(c) Sea $E =$"En la caja 3 hay dos bolas azules". Entonces $C \cup D \cup E = \Omega$ y la unión es disjunta, por lo que el suceso $F =$"Se extrae una bola roja de la caja 3" se descompone como:

$$\begin{aligned} \mathbf{P}(F) &= \mathbf{P}(F \cap C) + \mathbf{P}(F \cap D) + \mathbf{P}(F \cap E) \\ &= \mathbf{P}(F \mid C)\mathbf{P}(C) + \mathbf{P}(F \mid D)\mathbf{P}(D) + \mathbf{P}(F \mid E)\mathbf{P}(E) \end{aligned}$$

Ahora, $\mathbf{P}(F \mid C) = 1$ pues dado que hay dos bolas rojas, es seguro extraer una roja al elegir al azar. Del mismo modo $\mathbf{P}(F \mid D) = \frac{1}{2}$ pues hay una roja y una azul, y $\mathbf{P}(F \mid E) = 0$. Por lo tanto, utilizando las partes anteriores:

$$\mathbf{P}(F) = 1\,\frac{3}{8} + \frac{1}{2}\frac{1}{2} + 0 = \frac{5}{8}$$

(d) $G =$"La bola no extraída es roja". Se pide:

$$\mathbf{P}(G \mid F) = \frac{\mathbf{P}(G \cap F)}{\mathbf{P}(F)}$$

Observando que $G \cap F = A \cap B$ (la única forma de que tanto la extraída como la no extraída sean rojas es que ambas sean rojas), y utilizando las partes (a) y (c) se tiene que:

$$\mathbf{P}(G \mid F) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(F)} = \frac{3/8}{5/8} = \frac{3}{5}$$

2. (a) Sea $A_i$="La bola de la extracción $i$ es roja". Entonces se decide $H_0$ si y solo si ocurre $A_1 \cap A_2$. Por lo tanto:

$$\beta = \mathbf{P}_{H_1}(A_1 \cap A_2) = \mathbf{P}_{H_1}(A_1)\mathbf{P}_{H_1}(A_2) = \frac{1}{2}\frac{1}{2} = \frac{1}{4}$$

donde hemos usado la independencia de las extracciones y el hecho de que si hay una bola de cada color ($H_1$), la probabilidad de extraer la roja es $\frac{1}{2}$.

(b) En este caso:

$$\alpha = \mathbf{P}_{H_0}((A_1 \cap A_2)^c) = 0$$

ya que en la hipótesis de que las dos bolas de la caja son rojas ($H_0$) no puede ocurrir otra cosa que extraer dos bolas rojas en el muestreo.

**Solución Ejercicio 2**

1. Se aplican los test paramétricos para muestras gaussianas.

*Test para* $\sigma^2$: Como $s_n^2 = 0.8374 < 1$ se plantea:

$$\begin{cases} H_0: & \sigma^2 = 1 \\ H_1: & \sigma^2 < 1 \end{cases}$$

El estadístico es $E = (n-1)\frac{s_n^2}{\sigma_0} = 9\frac{0.8374}{1} = 7.5366$ y la región crítica $\mathcal{R} = \left\{E \leq \chi^2_{1-\alpha}(n-1)\right\}$. De la tabla $\chi^2$ con 9 grados de libertad se tiene que $0.1 < \alpha^* < 0.5$ por lo que se acepta $H_0$.

*Test para* $\mu$: Como $\overline{X}_n = -0.466 < 0$ se plantea:

$$\begin{cases} H_0: & \mu = 0 \\ H_1: & \mu < 0 \end{cases}$$

El estadístico es $E = \frac{\sqrt{n}(\overline{X}_n - \mu_0)}{s_n} = \frac{\sqrt{10}(-0.466)}{0.9151} = -1.6103$ y la región crítica $\mathcal{R} = \{E \leq -t_\alpha(n-1)\}$. De la tabla $t$ de 9 grados de libertad se tiene que $0.05 < \alpha^* < 0.1$ por lo que se rechaza $H_0$.

Por lo tanto, es razonable suponer que el error en el diámetro tiene media menor que 0 y no se rechaza que su varianza sea 1.

2. (a) Aplicamos los test de aleatoriedad.

*Test de rachas:* El estadístico es $R = 7$ rachas por lo que ante la alternativa de muchas rachas el p-valor (tabla) es $\alpha^* = 0.4524 > 0.1$ por lo que se acepta la hipótesis nula.

*Test de Spearman:*

| $i$ | $R(Y_i)$ |
|---|---|
| 1 | 8 |
| 2 | 9 |
| 3 | 2 |
| 4 | 5 |
| 5 | 6 |
| 6 | 1 |
| 7 | 3 |
| 8 | 10 |
| 9 | 4 |
| 10 | 7 |

Por lo tanto, $\sum_{i=1}^{10}(R(Y_i) - i)^2 = 180$ y $r_s = -0.091$. Ante la alternativa de tendencia decreciente el p-valor sale de la tabla y es $\alpha^* = 0.406$ por lo que se acepta la hipótesis nula.

Con los dos resultados anteriores se decide aceptar la aleatoriedad de la muestra.

(b) Aplicamos el *test de independencia de rangos de Spearman:*

| $i$ | $R(X_i)$ | $R(Y_i)$ |
|---|---|---|
| 1 | 2 | 8 |
| 2 | 1 | 9 |
| 3 | 10 | 2 |
| 4 | 5 | 5 |
| 5 | 9 | 6 |
| 6 | 8 | 1 |
| 7 | 7 | 3 |
| 8 | 3 | 10 |
| 9 | 4 | 4 |
| 10 | 6 | 7 |

Por lo tanto, $\sum_{i=1}^{10}(R(X_i) - R(Y_i))^2 = 288$ y $r_s = -0.745$. Ante la alternativa de dependencia negativa el p-valor sale de la tabla y es $\alpha^* = 0.009$ por lo que se rechaza la hipótesis de independencia de las variables $X$ e $Y$.

Nota: No se puede aplicar test de $\chi^2$ de tablas de contingencia porque solo funciona para $n$ grande y en este caso hay 10 datos.

(c) Aplicamos los test de normalidad.

*Test de D'Agostino:*

El estadístico de D'Agostino es:

$$DA_{10} = \frac{\sum_{i=1}^{10} Y_i^*(i - \frac{10+1}{2})}{\sqrt{10^3(10-1)s_n^2}} = 0.2818$$

Y de la tabla se tiene que $\alpha^* > 0.10$ por lo que se acepta la hipótesis nula.
*Test de Shapiro-Wilk:*
El estadístico de Shapiro-Wilk es:

$$W = \frac{b_n^2}{(n-1)s_n^2}, \quad b_n = \sum_{i=1}^{[n/2]} a_i(X_{n-i+1}^* - X_i^*)$$

Con los $a_i$, $i = 1, \ldots, 5$ de la tabla I de S-W y los datos se obtiene:

$$b_{10} = 3.3423 \quad W = 0.9116$$

y de la tabla II de S-W se tiene que $0.1 < \alpha^* < 0.5$ por lo que se acepta la hipótesis nula.
Con los resultados de estos test, aceptamos que la muestra $Y$ tiene distribución normal.

(d) Con la conclusión de la parte anterior, utilizamos un *I. de C. para muestras gaussianas.* En particular, utilizamos el intervalo para $\mu$ con varianza desconocida dado por:

$$I_\alpha = \left[\overline{Y}_n - \frac{s_n t_{\alpha/2}(n-1)}{\sqrt{n}}, \overline{Y}_n + \frac{s_n t_{\alpha/2}(n-1)}{\sqrt{n}}\right]$$

que con los datos del problema y $\alpha = 0.1$ (por lo que $t_{0.05}(9) = 1.833$) queda:

$$I_{0.1} = [-0.5584, 0.7944]$$