

**BIOESTADÍSTICA**  
**EXAMEN 24 DE FEBRERO DE 2006**

DATOS DEL ESTUDIANTE

Nombre	Cédula

- La duración del examen es 3 horas.
- El puntaje mínimo para aprobar es 50 puntos.

**Problema 1 (30 + 5 puntos)**

Considere el triplete **TTC** que codifica para la *fenilalanina*. Una mutación puntual consiste en el reemplazo de una de las bases del triplete por otra base distinta<sup>1</sup>

- a) **(6 puntos)** ¿Cuántas son las posibles mutaciones que puede sufrir el triplete **TTC**? Haga una lista de todos los tripletes resultantes.

Se sabe que la *leucina* es codificada por los tripletes: **TTA**, **TTG**, **CTT**, **CTC**, **CTA** y **CTG**. Suponga en lo que sigue que todas las mutaciones descritas en la parte a) tienen la misma probabilidad y suponga además que el triplete original **TTC** sufre una mutación puntual.

- b) **(8 puntos)** ¿Cuál es la probabilidad de que el triplete resultante de la mutación codifique para la *leucina*?
- c) **(8 puntos)** ¿Cuál es la probabilidad de que el triplete resultante de la mutación codifique para la *leucina*, dado que la mutación se produjo en la tercera posición?
- d) **(8 puntos)** ¿Cuál es la probabilidad de que la mutación se haya producido en la primera posición, dado que el triplete resultante de la mutación codifica para la *leucina*?
- e) **Parte extra (5 puntos):** Suponga ahora que, luego de la primera mutación, una segunda mutación puntual afecta al triplete. Sabiendo que la *fenilalanina* sólo es codificada por los tripletes **TTC** y **TTT**, ¿cuál es la probabilidad de que el triplete que resulta de la segunda mutación codifique para la *fenilalanina*?

---

<sup>1</sup>Se recuerda que las posibles bases son **A**, **C**, **G** y **T**.

**Problema 2 (35 puntos)**

En años recientes los granjeros suecos fumigaron sus campos sembrados de cereales con un fungicida que contenía metil de mercurio. Para tener una estimación del grado de contaminación inducido a productos comestibles, se realizó un estudio sobre el nivel de mercurio de los huevos producidos en Suecia. Para tal fin se tomó una muestra aleatoria de 1600 huevos y se observó que 940 de ellos estaban contaminados (esto es, tenían un nivel de metil de mercurio superior al máximo tolerado). En lo que sigue, denotamos por  $p$  a la proporción de huevos contaminados.

- a) **(10 puntos)** Construya un intervalo de confianza aproximado para  $p$  al nivel  $\alpha = 0.10$ .
- b) **(25 puntos)** Realice una prueba de hipótesis al nivel  $\alpha = 0.10$  para decidir entre las siguientes hipótesis:

$$\begin{cases} H_0 : p \geq 0.6 \\ H_1 : p < 0.6 \end{cases}$$

**Problema 3 (35 puntos)**

La mayor parte de las especies de coníferas tienen piñas de polen y piñas de semilla. El polen desprendido por la piña macho es transportado por el viento hasta la piña hembra, donde se fertilizan los huevos. Los siguientes datos corresponden a tiempos transcurridos (medidos en meses) entre la polinización y la fertilización:

1.95	0.96	1.56	1.39	1.89	1.75	1.35	0.27	1.81	1.34
------	------	------	------	------	------	------	------	------	------

**Nota:** En las pruebas de hipótesis utilice el siguiente criterio de decisión: se acepta la hipótesis nula si el  $p$ -valor es superior a 0.10

- a) **(10 puntos)** Realice dos pruebas de hipótesis para decidir si es razonable suponer que los datos son independientes e idénticamente distribuidos.
- b) **(25 puntos)** Implemente la prueba de ajuste de Kolmogorov-Smirnov para decidir si es razonable suponer que los datos ajustan a la siguiente función de distribución:

$$F_o(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{1}{4}x^2 & \text{si } 0 \leq x \leq 2 \\ 1 & \text{si } x > 2. \end{cases}$$

## SOLUCIÓN

### Problema 1

- a) Las posibles mutaciones son 9 (3 mutaciones por cada posición). La lista de tripletes resultantes es la siguiente:

<b>ATC</b>	<b>TAC</b>	<b>TTA</b>
<b>CTC</b>	<b>TCC</b>	<b>TTG</b>
<b>GTC</b>	<b>TGC</b>	<b>TTT</b>

En lo que sigue, todas las probabilidades se calculan dividiendo los *casos favorables* por los *casos posibles*:

- b)  $\frac{1}{3}$   
 c)  $\frac{2}{3}$   
 d)  $\frac{1}{3}$   
 e) Los *casos posibles* son 81, ya que a cada triplete de la parte a) corresponden 9 posibles mutaciones. Para calcular los *casos favorables* observemos que hay 9 mutaciones que transforman los tripletes calculados en a) en el triplete original **TTC** (estas son precisamente las mutaciones inversas de la primera mutación); además hay 2 mutaciones posibles en la tercera posición que dan como resultado el triplete **TTT**. De esta manera la probabilidad de que el triplete que resulta de la segunda mutación codifique para la *fenilalanina* es

$$\frac{11}{81} = 0.1358$$

### Problema 2

- a) Un intervalo de confianza aproximado al nivel  $\alpha$  para la proporción  $p$  es

$$I_{\alpha} = \left[ \hat{p} - \frac{1}{\sqrt{n}} z_{\frac{\alpha}{2}} \sqrt{\hat{p}(1 - \hat{p})}, \hat{p} + \frac{1}{\sqrt{n}} z_{\frac{\alpha}{2}} \sqrt{\hat{p}(1 - \hat{p})} \right]$$

En este caso  $n = 1600$ ,  $\hat{p} = \frac{940}{1600} = 0.5875$  y  $z_{0.05} = 1.65$ , con lo que resulta:

$$I = [0.5672, 0.6078].$$

- b) Para implementar el test de proporciones usamos el estadístico

$$T = \text{cantidad de huevos contaminados}$$

Está claro que si  $H_0$  es verdadera con  $p = p_0 = 0.60$  (que es el valor umbral de  $H_0$ ) entonces  $T$  tiene distribución Binomial de parámetros  $n = 1600$  y  $p_0 = 0.60$ ; de manera que el  $p$ -valor es:

$$\alpha^* = \mathbf{P}(T \leq 940) = \mathbf{P}\left(\frac{T - np_0}{\sqrt{np_0(1-p_0)}} \leq \frac{940 - 960}{19.596}\right) \simeq \mathbf{P}(N(0, 1) \leq -1.02) = 0.154$$

Como  $\alpha^* > 0.10$ , aceptamos  $H_0$ .

### Problema 3

a) **Test de Rachas:** El número de rachas es  $R = 7$  y el  $p$ -valor:  $\alpha^* = 0.4524$ .

**Test de correlación de rangos de Spearman:**  $\sum_{i=1}^{10} (R(X_i) - i)^2 = 216$ , de manera que el estadístico es  $r_s = -0.309$  y  $\alpha^* = 0.193$ .

En ambos casos es razonable suponer que los datos son i.i.d.

b) Para implementar el test de ajuste de Kolmogorov-Smirnov construimos la siguiente tabla:

$X_i$	$X_i^*$	$F_o(X_i^*)$	$\frac{i}{n}$	$ F_o(X_{i+1}^*) - \frac{i}{n} $	$ F_o(X_i^*) - \frac{i}{n} $
1.95	0.27	0.0182	0	0.0182	0.0818
0.96	0.96	0.2304	0.1	0.1304	0.0304
1.56	1.34	0.4489	0.2	<b>0.2489</b>	0.1489
1.39	1.35	0.4556	0.3	0.1556	0.0556
1.89	1.39	0.4830	0.4	0.0830	0.017
1.75	1.56	0.6084	0.5	0.1084	0.0084
1.35	1.75	0.7656	0.6	0.1656	0.0656
0.27	1.81	0.8190	0.7	0.119	0.0190
1.81	1.89	0.8930	0.8	0.093	0.007
1.34	1.95	0.9506	0.9	0.0506	0.0494
			1		

De manera que el estadístico es  $D = 0.2489$  y entonces  $\alpha^* > 0.20$ . Por lo tanto aceptamos  $H_0$ .