Problema 1 (3 puntos)

Sea $\mathbf{X} = (X_1, \dots, X_n)$ una muestra aleatoria i.i.d. proveniente de la distribución Beta de parámetros $(\theta, 1)$, con función de densidad de probabilidad dada por $f(x|\theta) = \theta x^{\theta-1}$, $0 < x < 1$, donde $\theta > 0$.

- (a) Muestre que el test más potente de tamaño α para contrastar $H_0 : \theta = \theta_0$ versus $H_1 : \theta = \theta_1$, con $\theta_1 > \theta_0$, rechaza H_0 para valores grandes del estadístico $T(\mathbf{X}) = \sum_{i=1}^n \log X_i$, i.e., tiene una región crítica de la forma $Re^* = \{\mathbf{x} \in [0, 1]^n : T(\mathbf{x}) \geq k_\alpha\}$.
- (b) Demuestre que $-2\theta \log X_i \sim \chi_2^2$. Deduzca la distribución de $-2\theta \sum_{i=1}^n \log X_i$.
- (c) ¿Cuál es su decisión en relación a H_0 ? considerando $\theta_0 = 2$, $\theta_1 = 2,8$, $\alpha = 0,05$, $n = 10$ y $\sum \log x_i = -5,846$ y un error de tipo I de 5%. ¿Cuál es el error de tipo II? ¿El p-valor del test es mayor o menor que 5%?

Problema 2 (3 puntos)

Sean $\mathbf{X} = (X_1, \dots, X_n)$, $\mathbf{Y} = (Y_1, \dots, Y_m)$ muestras aleatorias mutuamente independientes provenientes de las distribuciones $N(\mu_1, \sigma^2)$ y $N(\mu_2, \sigma^2)$, respectivamente, donde σ^2 es desconocida.

- (a) Proponga un test para contrastar $H_0 : \mu_1 \geq \mu_2$ versus $H_1 : \mu_1 < \mu_2$.
- (b) ¿Cuál es su decisión con un error de tipo I de 5%? considerando los datos $n = 50$, $m = 60$, $\bar{x}=3.2$, $\bar{y}=3.7$, $\frac{1}{n} \sum (x_i - \bar{x})^2=0.62$ y $\frac{1}{m} \sum (y_i - \bar{y})^2=0.79$.
- (c) ¿Como cambia la región crítica cuando n y m decrecen?

Problema 3 (6 puntos)

Un laboratorio requiere determina la concentración de ADN de una bacteria en un extracto dado. Considerando que se cometen errores de medición, se replica 20 veces la medición. En la tabla adjunta se presenta los datos obtenidos de manera aleatoria e independiente. Se sabe que las mediciones se ajustan a una distribución $\mathcal{N}(\mu, \sigma^2)$ donde μ y σ^2 son desconocidos y μ es el valor real de la concentración de ADN.

Replica	Concentración	Replica	Concentración	Replica	Concentración	Replica	Concentración
1	0.503	2	0.249	3	1.047	4	0.670
5	0.882	6	0.858	7	0.805	8	0.549
9	0.563	10	0.706	11	0.503	12	0.536
13	1.168	14	0.813	15	0.496	16	0.790
17	0.518	18	0.463	19	0.735	20	1.038

La media y la desviación estándar son respectivamente: $\bar{x} = 0,695$ y $s = 0,232$.

- (a) Encuentre el intervalo más pequeño tal que el valor real μ de la concentración pertenezca a dicho intervalo con una probabilidad de 95%. ¿Por qué $\bar{x}$ pertenece a este intervalo?
- (b) Si queremos un intervalo de largo igual al obtenido en (a) con un nivel de confianza de 95% ¿cual debe ser el tamaño de la muestra? Si queremos un intervalo de largo igual al obtenido en (a) con una muestra de tamaño igual a 20 ¿cual debe ser el nivel de confianza?

- (c) Para el laboratorio es más arriesgado sobreestimar el valor real μ de la concentración. Encuentre un intervalo $[a,b]$ de modo que la concentración de ADN pertenezca a dicho intervalo con un nivel de confianza de 95 % y tal que $IP(\mu < a) = \frac{1}{4}IP(\mu > b)$.
- (d) Para demostrar al laboratorio la baja calidad de las mediciones, se solicita a usted que encuentre un intervalo al cual pertenezca σ^2 , el valor real de la varianza de la medición, con un nivel de confianza de 90 %.
- (e) Sea una variable aleatoria con distribución de Poisson de parámetro $\lambda > 0$ desconocido. Se toma una muestra aleatoria simple $\mathbf{X} = (X_1, \dots, X_n)$ con $n > 0$ y se obtiene $x_i = 0 \quad \forall i$. Dado $\lambda = \lambda_o > 0$, calcule $IP(x_1 = 0, \dots, x_n = 0 | \lambda = \lambda_o)$. Pruebe que si λ_o es tal que $e^{-n\lambda_o} \geq \alpha_o$ entonces $\lambda_o \leq -\frac{1}{n} \ln(\alpha_o)$. Concluya entonces que $\lambda_o \in [0, -\frac{1}{n} \ln(0,05)]$ si $IP(x_1 = 0, \dots, x_n = 0 | \lambda = \lambda_o) \geq 0,05$. Comente los alcances de este resultado.