RELACIÓN 4 DE EJERCICIOS

3.- a) Sea una población que se distribuye normalmente con media μ y varianza 1. Se intenta contrastar H_0: $\mu=0$ contra H_1: $\mu=1$. A un tamaño dado del 5% determine el test más potente. Aplíquelo suponiendo que se dispone de una m.a.s. de tamaño 5, con los siguientes valores: 3, 1, -2, -5, 4. Calcule la probabilidad de cometer error del tipo II.
b) Si la población es normal y σ es desconocida, ¿se podría aplicar el mismo resultado que en el apartado anterior para contrastar H_0: $\mu=\mu_0$ contra H_1: $\mu=\mu_1$?.¿Conoce alguna otra forma de obtener un contraste para esta situación?
c) Demuestre cuál es el test uniformemente más potente para contrastar H_0: $\mu=10$ frente a H_1: $\mu=8$ en una $N(\mu,3)$ si n=20 y $\alpha=0.05$. Obtenga la potencia de dicho contraste.
d) Sea una población normal, una muestra suya compuesta de 10 individuos y el contraste habitual para H_0: $\mu=1$ frente H_1: $\mu\neq 1$ de tamaño 10%. ¿Qué tipo de error puede ocurrir, según cuál sea el resultado del test, si μ toma cada uno de los tres valores siguientes: 0, 1 y 2?. Halle en cada caso una expresión de la probabilidad de que ocurra.

a) Según el teorema de Neyman-Pearson el test u.m.p tiene como región crítica

$C=\{(x_1,x_2,\dots x_n)\in\mathbb{R}^n\ /\ \Lambda_C(x_1,x_2,\dots x_n)\leq K\}$ y verifica

$\underset{\mu=0}{P}[(x_1,x_2,\dots x_n)\in C]=\alpha$

La función de densidad de la distribución $N(\mu,1)$ es:

$$f(x)=\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{1}{2}(x-\mu)^2}$$

. Si $x_1,x_2,\dots x_n$ es una m.a.s.

la función de verosimilitud es:

$$L_\mu(x_1,x_2,\dots x_n)=\prod_{i=1}^{n}f_\mu(x_i)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{1}{2}(x_i-\mu)^2}$$

$$L_\mu(x_1,x_2,\dots,x_n)=\left(\frac{1}{\sqrt{2\pi}}\right)\cdot e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

Como:

$$\Lambda_C(x_1,x_2,\dots x_n)=\frac{L_{\mu=0}(x_1,x_2,\dots x_n)}{L_{\mu=1}(x_1,x_2,\dots x_n)}=\frac{\left(\frac{1}{\sqrt{2\pi}}\right)^n\cdot e^{-\frac{1}{2}\sum_{i=1}^{n}x_i}}{\left(\frac{1}{\sqrt{2\pi}}\right)^n\cdot e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-1)^2}}$$

Por tanto:

$$\Lambda_C(x_1,x_2,\dots x_n)=e^{\frac{1}{2}\sum_{i=1}^{n}(x_i-1)^2-x_i^2}=e^{\frac{n}{2}-\sum_{i=1}^{n}x_i}$$

finalmente podemos poner que:

$$\Lambda_C(x_1, x_2, \dots x_n) = e^{\frac{n}{2} - n\bar{x}}$$

La región crítica es

$$C = \left\{ (x_1, x_2, \dots, x_n) \in \mathbb{R}^n \,/\, e^{\frac{n}{2} - n\bar{x}} \leq k \right\}$$

Ahora tenemos que calcular k.

$P_{\mu=0}\left[e^{\frac{n}{2} - n\bar{x}} \leq k \right] = 0'05$, siendo $\bar{x} \rightsquigarrow N(0, \frac{1}{\sqrt{n}})$

ya que si $X \rightsquigarrow N(\mu, \sigma)$ por el lema de Fisher

$\bar{x} \rightsquigarrow N\left(0, \frac{1}{\sqrt{n}}\right) \Rightarrow \sqrt{n}\,\bar{x} \rightsquigarrow N(0,1)$

$$e^{\frac{n}{2} - n\bar{x}} \leq k \Leftrightarrow \frac{n}{2} - n\bar{x} \leq \ln k \Leftrightarrow \frac{n}{2} - \ln k \leq n\bar{x}$$

$$\frac{\frac{n}{2} - \ln k}{n} \leq \bar{x} \Leftrightarrow \frac{\sqrt{n}}{n}\left(\frac{n}{2} - \ln k\right) \leq \sqrt{n}\,\bar{x}$$

$$\frac{\frac{n}{2} - \ln k}{\sqrt{n}} < \sqrt{n}\,\bar{x}$$

Luego $P_{\mu=0}\left[e^{\frac{n}{2} - n\bar{x}} \leq k \right] = 0'05 \Rightarrow P\left[\sqrt{n}\,\bar{x} > \frac{\frac{n}{2} - \ln k}{\sqrt{n}} \right] =$

$= 0'05 \Rightarrow P\left[n\bar{x} \leq \frac{\frac{n}{2} - \ln k}{\sqrt{n}} \right] = 1 - 0'05 = 0'95 \Rightarrow$

$$\Rightarrow \frac{\frac{n}{2} - \ln k}{\sqrt{n}} = z_{0'05} = 1'65 \Rightarrow \frac{n}{2} - 1'65\sqrt{n} = \ln k$$

$$k = e^{\frac{n}{2} - 1'65\sqrt{n}}$$

La región crítica es

$$C = \left\{ (x_1, x_2, \dots x_n) \in \mathbb{R}^n \,/\, e^{\frac{n}{2} - n\bar{x}} \leq e^{\frac{n}{2} - 1'65\sqrt{n}} \right\}$$

Si simplificamos esta expresión la región crítica es:

$C = \left\{ (x_1, x_2, \dots x_n) \in \mathbb{R}^n \,/\, \bar{x} \geq \frac{1'65}{\sqrt{n}} \right\}$

Si tomamos la muestra $(3, 1, -2, -5, 4)$ de tamaño $n=5$

$\Rightarrow \bar{x} = \frac{1}{5} = 0'2$

como $\frac{1'65}{\sqrt{5}} = 0'7379$ $\Big\} \Rightarrow \bar{x} \not\geq 0'7379$ no

se puede rechazar la hipótesis nula. Aceptamos que $\mu = 0$. con un nivel de significación del 5%

Para calcular la probabilidad de cometer error de tipo II

$P[\text{error de tipo II}] = P_{\mu=1}[(x_1, x_2, \dots x_n) \notin C] =$

$= P_{\mu=1}\left[\bar{x} < \frac{1'65}{\sqrt{n}}\right] = P_{\mu=1}\left[\bar{x} < \frac{1'65}{\sqrt{5}}\right] =$

$= P_{\mu=1}[\bar{x} < 0'7379]$. En este caso $\bar{x} \rightsquigarrow N\left(1, \frac{1}{\sqrt{5}}\right)$

$\bar{x} \rightsquigarrow N(1, 0'448)$

$P[\bar{x} < 0'7379] = P\left[\frac{\bar{x}-1}{0'448} \leq \frac{0'7379-1}{0'448}\right] = P[Z \leq -0'546]$

$= P[Z \geq 0'546] = 1 - P[Z \leq 0'546] = 1 - 0'7054 = 0'2946$

Luego $P[\text{error tipo II}] = 0'2946$

b) Si no se conoce σ, como según el Lema de Fisher $\bar{x} \rightsquigarrow N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$ y $\Lambda_c(x_1, x_2, \dots x_n) = e^{\frac{n}{2} - n\bar{x}}$,

la región crítica es $C = \{(x_1, x_2, \dots x_n) \in \mathbb{R}^n / e^{\frac{n}{2} - n\bar{x}} \leq k\}$

El nivel de significación del test es $\alpha = 0'05$ por lo que

$P_{\mu=0}\left[e^{\frac{n}{2} - n\bar{x}} \leq k\right] = 0'05$, para hallar k necesitamos conocer la distribución de $\bar{x}$ y esto es imposible sin el valor σ. No se puede resolver el ejercicio de la misma forma que anteriormente.

Se podría pensar en obtener el test u.m.p. por medio del cociente de verosimilitudes generalizado y aplicando después el teorema de Wilks pero esto requiere utilizar muestras grandes, nuestra muestra es de tamaño $n=5 << 30$, no nos sirve el método.

No tenemos más remedio que acudir a intervalos de confianza. Como se trata de una normal con σ desconocida el intervalo de confianza de nivel α es:

$$IC_{1-\alpha}(\mu) = \left(\bar{x} - t_{n-1,\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n-1}}, \; \bar{x} + t_{n-1,\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n-1}}\right)$$

es decir que

$$IC_{95\%}(\mu) = \left(0'2 - t_{4,0'025} \cdot \frac{3'31}{2}, \; 0'2 + t_{4,0'025} \cdot \frac{3'31}{2}\right)$$

$$S^2 = \frac{\sum x_i^2}{n} - \bar{x}^2 = 11 - 0'2^2 = 10'96 \Rightarrow S = 3'31 \; ; \; t_{4,0'025} = 2'776$$

$$IC_{95\%}(\mu) = (-6'99, \; 4'7678)$$

Al ser $\mu = 0 \in IC_{95\%}$ no podemos rechazar la hipótesis nula, aceptamos que $\mu = 0$.

c) Test u.m.p

$$\begin{aligned} H_0 &: \mu = 10 \\ H_1 &: \mu = 8 \end{aligned} \quad \text{en una } N(\mu, 3) \quad n = 20 \quad \alpha = 5\%$$

Sea $x_1, x_2, \ldots x_n$ una m.a.s. Tomamos el cociente de verosimilitudes:

$$\Lambda_c(x_1, x_2, \ldots, x_n) = \frac{L_{\mu=10}(x_1, x_2, \ldots x_n)}{L_{\mu=8}(x_1, x_2, \ldots x_n)}$$

como la función de densidad de X es:

$$f(x) = \frac{1}{3\sqrt{2\pi}} \cdot e^{-\frac{1}{18}(x-\mu)^2}$$, se tiene que:

$$L_{\mu=10}(x_1,x_2,\ldots x_n)=\prod_{i=1}^{n}\frac{1}{3\sqrt{2\pi}}\cdot e^{-\frac{1}{18}(x_i-10)^2}=\left(\frac{1}{3\sqrt{2\pi}}\right)^n\cdot e^{-\frac{1}{18}\sum_{i=1}^{n}(x_i-10)^2}$$

De la misma manera:

$$L_{\mu=8}(x_1,x_2,\ldots,x_n)=\left(\frac{1}{3\sqrt{2\pi}}\right)^n\cdot e^{-\frac{1}{18}(x_i-8)^2}$$. Entonces:

$$\Lambda_C(x_1,x_2,\ldots x_n)=e^{-\frac{1}{18}\sum_{i=1}^{n}[(x_i-8)^2-(x_i-10)^2]}$$

$$\Lambda_C(x_1,x_2,\ldots x_n)=e^{2n-\frac{2}{9}n\bar{x}}=e^{2n(1-\bar{x})}$$

La región crítica es:

$$C=\{(x_1,x_2,\ldots x_n)\in\mathbb{R}^n \,/\, e^{2n(1-\bar{x})}\leq K\}$$

Según el lema de Fisher si $X \rightsquigarrow N(\mu,3)$ y $n=20$

$$\Rightarrow \bar{X}\rightsquigarrow N\left(\mu,\frac{3}{\sqrt{20}}\right)\Rightarrow \frac{\bar{X}-\mu}{\frac{3}{\sqrt{20}}}\rightsquigarrow N(0,1)$$

$$e^{2n(1-\bar{x})}\leq K \iff 2n(1-\bar{x})\leq \ln K \Rightarrow 1-\bar{x}\leq\frac{\ln K}{2n}$$

$$\Rightarrow \bar{x}\geq 1-\frac{\ln K}{2n}$$. Luego:

$$\underset{\mu=10}{P}[(x_1,x_2,\ldots x_n)\in C]=1-\alpha \iff \underset{\mu=10}{P}\left[\bar{X}\geq 1-\frac{\ln K}{2n}\right]=1-\alpha$$

$$\iff \underset{\mu=10}{P}\left[\frac{\bar{X}-\mu}{\frac{3}{\sqrt{20}}}\geq\frac{1-\frac{\ln K}{2n}-\mu}{\frac{3}{\sqrt{20}}}\right]\iff \underset{\mu=10}{P}\left[Z\geq-\frac{9+\frac{\ln K}{2n}}{\frac{3}{\sqrt{20}}}\right]=$$

$$=1-\alpha \iff P\left[Z\leq\frac{9+\ln K/2n}{3/\sqrt{20}}\right]=1-\alpha \Rightarrow$$

$$\Rightarrow \frac{9+\frac{\ln K}{2n}}{\frac{3}{\sqrt{20}}}=z_\alpha.$$ Como $\alpha=0'05$

$$\frac{9+\frac{\ln K}{2n}}{\frac{3}{\sqrt{20}}}=z_{0'05}=1'65\Rightarrow 9+\frac{\ln K}{40}=\frac{1'65\cdot 3}{\sqrt{20}}$$

$$\frac{\ln k}{40} = \frac{4'95}{\sqrt{20}} - 9 \Rightarrow \ln k = 40\left[\frac{4'95}{\sqrt{20}} - 9\right] \Rightarrow$$

$$\Rightarrow k = e^{40\left[\frac{4'95}{\sqrt{20}} - 9\right]} = -1'85 \cdot 10^{18}$$

La región crítica es:

$$C = \{(x_1, x_2, \dots x_n) \in \mathbb{R}^n \,/\, \Lambda_c(x_1, x_2, \dots x_n) \leq -1'85 \cdot 10^{18}\}$$ y

esto es equivalente a poner:

$$C = \{(x_1, x_2, \dots x_n) \in \mathbb{R}^n \,/\, 40(1-\bar{x}) \leq 40\left[\frac{4'95}{\sqrt{20}} - 9\right]\} =$$

$$C = \{(x_1, x_2, \dots x_n) \in \mathbb{R}^n \,/\, \bar{x} - 1 \geq 9 - \frac{4'95}{\sqrt{20}}\} =$$

$$C = \{(x_1, x_2, \dots x_n) \in \mathbb{R}^n \,/\, \bar{x} \geq 8'893\}$$

d)

4.- Una población se distribuye según una distribución de Bernouilli de parámetro p. Se sabe que p=0.5 ó 0.75. Determine el test más potente que le ayude a elegir entre una u otra hipótesis tomando como hipótesis nula la asociada al menor valor de p. Aplíquelo al caso de un experimento en el que p es la probabilidad de tener éxito y al ser repetido 90 veces de forma independiente, produce 65 éxitos. ¿Qué hipótesis rechazaría en esa situación a un tamaño del 5%?

La función de probabilidad de $X \rightarrow b(p)$ es:

$$P[X=x] = p^x \cdot q^{1-x} = \left(\frac{p}{q}\right)^x \cdot q$$

El contraste de hipótesis que hemos de realizar es

$H_0: \ p = 0'5$
$H_1: \ p = 0'75$

El cociente de verosimilitudes es:

$$\Lambda_c = \frac{L_{p=0'5}(x_1, x_2, \dots x_n)}{L_{p=0'75}(x_1, x_2, \dots x_n)} = \frac{\prod_{i=1}^{n} \left(\frac{0'5}{0'5}\right)^{x_i} \cdot 0'5}{\prod_{i=1}^{n} \left(\frac{0'75}{0'25}\right)^{x_i} \cdot 0'25}$$

$$\Lambda_c(x_1, x_2, \dots x_n) = \frac{0'5^n}{3^{\sum x_i} \cdot 0'25^n} = \frac{\frac{1}{2^n}}{3^{\sum x_i} \cdot \frac{1}{(2^2)^n}} = \frac{(2^n)^2}{2^n \cdot 3^{\sum x_i}}$$

$$\Lambda_c(x_1, x_2, \dots x_n) = \frac{2^n}{3^{\sum_1^n x_i}}$$

La región crítica es:

$$C = \left\{ (x_1, x_2, \dots x_n) \in \mathbb{R}^n / \ \frac{2^n}{3^{\sum_1^n x_i}} \leq k \right\}$$ verificando

$P_{p=0'5}\left[(x_1, x_2, \dots x_n) \in C\right] = \kappa$. Como la distribución de Bernouilli no es continua no podemos asegurar que esta región corresponda al test u.m.p ya que este puede que no exista.

$$\frac{2^n}{3^{\sum x_i}} \leq K \Rightarrow n \ln 2 - \sum_{i=1}^{n} x_i \ln 3 \leq \ln K$$

$$\sum_{i=1}^{n} x_i \ln 3 \geq n \ln 2 - \ln K \Rightarrow \sum x_i \geq \frac{\ln \frac{2^n}{K}}{\ln 3}$$

Así la región crítica es:

$$C = \left\{ (x_1, x_2, \dots x_n) \in \mathbb{R}^n \,/\, \sum_{i=1}^{n} x_i \geq \frac{\ln \frac{2^n}{K}}{\ln 3} \right\}$$

$$\underset{p=0'5}{P}\left[\sum_{i=1}^{n} x_i \geq \frac{\ln \frac{2^n}{K}}{\ln 3} \right] = \alpha \text{ (nivel de significación)}$$

Como $\sum_{i=1}^{n} x_i \rightsquigarrow B[n, p]$, según el teorema de Moivre se puede aproximar $\sum_{i=1}^{n} x_i \rightsquigarrow N(n \cdot p, \sqrt{n \cdot p \cdot q})$ es decir que

$\sum_{i=1}^{n} x_i \rightsquigarrow N(0'5\, n, 0'5 \sqrt{n})$. En nuestro caso si $n = 90$:

$\sum_{i=1}^{n} x_i \rightsquigarrow N(45, 4'74)$. Sea $K' = \frac{\ln \frac{2^n}{K}}{\ln 3}$ y $\alpha = 0'05$. Entonces ha de ser:

$$P\left[\sum_{i=1}^{n} x_i \leq K' \right] = \alpha \Rightarrow P\left[\frac{\sum x_i - 45}{4'74} \leq \frac{K' - 45}{4'74} \right] = 0'05$$

$\Rightarrow P\left[Z \leq \frac{K'-45}{4'74} \right] = 0'05$. Esto nos indica que $\frac{K'-45}{4'74}$ es negativo, por tanto:

$$P\left[Z \geq \frac{45 - K'}{4'74} \right] = 0'05 \Rightarrow P\left[Z \leq \frac{45-K'}{4'74} \right] = 0'95$$

con lo que $\frac{45-k'}{4'74} = 1'65$ (se busca en las tablas) $\Rightarrow$

$\Rightarrow k' = 37'163$. Luego la región crítica es:

$C = \{(x_1, x_2, \ldots x_n) \in \mathbb{R}^n \,/\, \sum_{i=1}^{n} x_i \leq 37'173\}$ y esta región crítica pertenece al u.m.p. si consideramos que $\sum_{i=1}^{n} x_i \rightsquigarrow N(0'5n, 0'5\sqrt{n})$.

Si en esta repetición de 90 veces el experimento de Bernoulli alcanzó 65 éxitos $\Rightarrow \sum_{i=1}^{90} x_i = 65 \Rightarrow (x_1, x_2, \ldots x_n) \notin C$ $\Rightarrow$ No se puede rechazar la hipótesis nula y por tanto admitimos que $p = 0'5$.

5.– a) Considere la familia de distribuciones B(n,p) con n conocida. ¿Es una familiaexponencial? ¿Tiene un cociente de verosimilitudes monótono para el parámetro p en algún estadístico? Razone las respuestas.

b) Los artículos que salen de la línea de producción de una empresa A se van empaquetando en cajas de 100 unidades que son enviadas a una segunda empresa B, que utiliza dichos artículos en la fabricación de nuevos productos más elaborados. La empresa A asegura que el número medio de artículos defectuosos por caja es de 3, pero en la empresa B piensan que puede ser superior. Para contrastar estas aseveraciones, los responsables de la empresa B escogen 7 cajas al azar y examinan todos los artículos, obteniendo la siguiente cantidad de defectuosos por caja: 2,5,4,1,3,2,5. En estas condiciones, **b1)** ¿se aceptará la afirmación del proveedor para un tamaño de test del 5%? y **b2)** ¿se mantendría la decisión para un tamaño de test inferior al anterior?

a) $B(n,p)$ pertenece a la familia exponencial ya que su función de probabilidad es

$$P[X=x] = \binom{n}{x} p^x \cdot (1-p)^{n-x} = \binom{n}{x} \cdot \left(\frac{p}{1-p}\right)^x \cdot (1-p)^n$$

Podemos poner entonces:

$P[X=x] = \binom{n}{x} \cdot e^{x \ln \frac{p}{1-p}} \cdot (1-p)^n$. Esta función es de la forma $P[X=x] = a(p) \cdot b(x) \cdot e^{c(p) \cdot d(x)}$ ya que:

$a(p) = (1-p)^n \quad b(x) = \binom{n}{x} \quad c(p) = \ln \frac{p}{1-p} \quad d(x) = x.$

Por una propiedad vista en teoría sabemos que si $c(p) = \ln \frac{p}{1-p}$ es monótona entonces la función cociente de verosimilitudes es de la forma

$$\Lambda_c(x_1, x_2, \dots x_n) = h[T(x_1, x_2, \dots x_n), p_0, p_1)$$ es

monótona respecto a T, además la monotonía de Λ_c es contraria a la de $c(p)$

$$\frac{dc(p)}{dp} = \frac{p-1}{p} \cdot \frac{1}{(p-1)^2} = \frac{1}{p(p-1)} > 0 \Rightarrow c(p) \text{ es cre-}$$

ciente respecto a $p \Rightarrow \Lambda_c(x_1, x_2, \dots x_n)$ es decreciente respecto a:

$$T(x_1, x_2, \dots x_n) = \sum_{i=1}^{n} d(x_i) = \sum_{i=1}^{n} x_i$$

luego la distribución $B(n, p)$ tiene la propiedad c.v.m respecto al estadístico $T(x_1, x_2, \dots x_n) = \sum_{i=1}^{n} x_i$

b) Si $X =$ nº defectuosos por caja en la línea de producción:

$$X \rightsquigarrow B(100, p)$$

El contraste que queremos realizar es:

$H_0: \ p = 0'03$

$H_1: \ p \geq 0'03$

Como $\Lambda_c(x_1, x_2, \dots x_n) = h[T(x_1, x_2, \dots x_n), p_0, p_1]$ es no creciente respecto a $T(x_1, x_2, \dots x_n) = \sum_{i=1}^{n} x_i$, la región crítica es:

$$C = \left\{ (x_1, x_2, \dots x_n) \in \mathbb{R}^n / \sum x_i \geq K \right\}$$

Como $B(100, p)$ es reproductiva, para muestras de tamaño 7:

$$\sum_{i=1}^{7} x_i \rightsquigarrow B(700, p)$$

Por el teorema de Moivre:

$B(700, p) \xrightarrow{L} N(700p, \sqrt{700p(1-p)})$

Queremos hallar k a partir de la expresión

$P_{p=0'03}[\sum x_i \geq k] = 0'05$

En este caso $\sum x_i \rightsquigarrow B(700, 0'03) \xrightarrow{L} N(21, 4'51)$

$P\left[\frac{\sum x_i - 21}{4'5} \geq \frac{k-21}{4'5}\right] = 0'05$ lo cual nos dice

que $P\left[\frac{\sum x_i - 21}{4'5} \leq \frac{k-21}{4'5}\right] = 0'95 \Rightarrow \frac{k-21}{4'5} = 1'65 \Rightarrow$

$\Rightarrow K = 21 + 4'5 \cdot 1'65 \Rightarrow k = 28'425$

La región crítica es:

$C = \{(x_1, x_2, \dots, x_n) \in \mathbb{R}^n \ / \ \sum_{i=1}^{n} x_i \geq 28'425\}$

En la muestra $\sum_{i=1}^{7} x_i = 22 \notin C \Rightarrow$ No se rechaza la hipotesis nula, se acepta que $p = 0'03$.

Si α disminuye $\Rightarrow P\left[\frac{\sum x_i - 21}{4'5} \leq \frac{k-21}{4'5}\right]$ aumenta ya que este valor es $1-\alpha$. Por ejemplo si $\alpha = 0'025$

$P\left[\frac{\sum x_i - 21}{4'5} \leq \frac{k-21}{4'5}\right] = 0'975 \Rightarrow \frac{k-21}{4'5} = 1'96 \Rightarrow$

$\Rightarrow k = 29'82$

Si α disminuye k aumenta, se sigue manteniendo la decisión.

7.– En una centralita de teléfonos, el número de llamadas recibidas por minuto sigue una distribución de Poisson de parámetro λ. En una hora concreta, se han recibido 357. El número de llamadas en minutos diferentes son independientes entre sí.

a) Se quiere contrastar $\lambda=5$ frente a $\lambda=10$. Encuentre el test más potente de tamaño $\alpha=0.05$ y aplíquelo.

b) Haga lo mismo si se quiere contrastar $\lambda=5$ contra $\lambda=\lambda_0$ donde λ_0 es un valor fijo mayor que 5.

c) ¿Existe un test uniformemente más potente para contrastar H_0: $\lambda=5$ contra H_1: $\lambda>5$?

(Indicación: ¿Se puede utilizar el mismo test que en el caso de H_1 simple?).

d) ¿Podría encontrar, por el mismo método, un test uniformemente más potente para la hipótesis H_0: $\lambda\leq 5$ contra la alternativa H_1: $\lambda>5$? ¿Y por otro método?

a) La función de probabilidad de $X \rightarrow P(\lambda)$ es

$P[X=x] = \dfrac{\lambda^x}{x!} \cdot e^{-\lambda}$. Si tomamos una m.a.s.

$x_1, x_2, \ldots x_n$ la función de densidad es

$$L_\lambda(x_1, x_2, \ldots x_n) = \prod_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} = \frac{1}{\prod_{i=1}^{n} x_i!} \cdot e^{-n\lambda} \cdot \lambda^{\sum x_i}$$

La Hipotesis que queremos contrastar es:

$H_0: \lambda = 5$

$H_1: \lambda = 10$

Se trata de hipotesis simples, aplicamos el teorema de Neymann - Pearson:

$$\Lambda_c(x_1, x_2 \ldots x_n) = \frac{L_{\lambda=5}(x_1, x_2 \ldots x_n)}{L_{\lambda=10}(x_1, x_2 \ldots x_n)} = \frac{\cancel{\frac{1}{\prod x_i!}} \cdot e^{-5n} \cdot 5^{\sum x_i}}{\cancel{\frac{1}{\prod x_i!}} e^{-10n} \cdot 10^{\sum x_i}}$$

$$\Lambda_c(x_1, x_2, \ldots x_n) = e^{5n} \cdot \frac{\cancel{5^{\sum x_i}}}{2^{\sum x_i} \cdot \cancel{5^{\sum x_i}}} = e^{5n} \cdot 2^{-\sum x_i}$$

El test u.m.p, si existe, tiene por región crítica

$$C = \{(x_1, x_2, \ldots x_n) / e^{5n} \cdot 2^{-\sum x_i} \leq K\}$$

$e^{5n} \cdot 2^{-\sum x_i} \leq k \iff 5n - \sum x_i \ln 2 \leq \ln k$

$\sum x_i \geq \dfrac{5n - \ln k}{\ln 2} = k'$

Podemos poner que la región crítica es:

$C = \left\{ (x_1, x_2, \cdots x_n) \in \mathbb{R}^n / \sum x_i \geq k' \right\}$

Además se ha de verificar que:

$P_{\lambda=5}\left[\sum x_i \geq k'\right] = 0'05$

Si $x_i \rightsquigarrow P(\lambda) \Rightarrow \sum_{i=1}^{n} x_i \rightsquigarrow P(n\lambda)$ ya que la distribución de Poisson es reproductiva. El tamaño de la muestra es 60, al ser n grande podemos aplicar el T.C.L. y obtenemos

$P(300) \rightsquigarrow N(300, \sqrt{300})$, es decir:

$\sum x_i \rightsquigarrow N(300, \sqrt{300})$. Como queremos calcular

$P\left[\sum x_i \geq k'\right] = 0'05 \Rightarrow P\left[\sum x_i \leq k'\right] = 0'95 \Rightarrow$

$P\left[\dfrac{\sum x_i - 300}{\sqrt{300}} \leq \dfrac{k' - 300}{\sqrt{300}}\right] = 0'95 \Rightarrow \dfrac{k'-300}{\sqrt{300}} = 1'65$

$k' = 300 + 1'65\sqrt{300} = 328'58$

Entonces la región crítica es:

$C = \left\{ (x_1, x_2, \cdots x_n) \in \mathbb{R}^n / \sum_{i=1}^{n} x_i \geq 328'58 \right\}$. En la muestra $\sum_{i=1}^{n} x_i = 357 \geq 328'58 \Rightarrow$ se rechaza la hipótesis nula y se acepta la hipótesis alternativa. Aceptamos que $\lambda = 10$.

b) El contraste que queremos hacer es:

$H_0: \lambda = 5$

$H_1: \lambda = \lambda_0 \quad (\lambda_0 > 5)$

Como se trata de hipótesis simples, aplicamos de nuevo el teorema de Neyman-Pearson:

$$\Lambda_c(x_1, x_2, \dots x_n) = \frac{\cancel{\frac{1}{\prod_{i=1}^{n} x_i!}} \cdot e^{-5n} \cdot 5^{\sum x_i}}{\cancel{\frac{1}{\prod_{i=1}^{n} x_i!}} \cdot e^{-\lambda_0 n} \cdot \lambda_0^{\sum x_i}}$$

$$\Lambda_c(x_1, x_2, \dots x_n) = e^{(\lambda_0 - 5)n} \cdot \left(\frac{5}{\lambda_0}\right)^{\sum x_i}$$

La región crítica es:

$$C = \left\{ (x_1, x_2, \dots x_n) \in \mathbb{R}^n \,/\, e^{(\lambda_0 - 5) \cdot n} \cdot \left(\frac{5}{\lambda_0}\right)^{\sum x_i} \leq k \right\}$$

$$e^{(\lambda_0 - 5)n} \cdot \left(\frac{5}{\lambda_0}\right)^{\sum x_i} \leq k \iff (\lambda_0 - 5)n + \sum x_i \cdot \ln\left(\frac{5}{\lambda_0}\right) \leq \ln k$$

$$\sum_{i=1}^{n} x_i \geq \frac{\ln k + (5 - \lambda_0)n}{\ln\left(\frac{5}{\lambda_0}\right)} = k' \text{ ya que } \ln\left(\frac{5}{\lambda_0}\right) < 0 \text{ al ser } \lambda_0 > 5$$

La región crítica es por tanto

$C = \left\{ (x_1, x_2, \dots x_n) \in \mathbb{R}^n \,/\, \sum x_i \geq k' \right\}$ es la misma que anteriormente $k' = 328'58$

$$C = \left\{ (x_1, x_2, \dots x_n \in \mathbb{R}^n \,/\, \sum_{i=1}^{n} x_i \geq 328'58 \right\}$$

c) Se puede emplear el mismo test que en el apartado b) puesto que queremos contrastar:

$H_0: \lambda = 5$

$H_1: \lambda > 5$. Este contraste es equivalente a:

$H_0: \lambda = 5$
$H_1: \lambda = \lambda_0$ donde $\lambda_0 > 5$

El test u.m.p. es el que tiene región crítica

$C = \{(x_1, x_2, \dots x_n) / \sum_{i=1}^{n} x_i \le K\}$

Como $x_i \rightsquigarrow P(\lambda) \Rightarrow \sum_{i=1}^{n} x_i \rightsquigarrow P(n\lambda)$. Si el nivel de significación es α debe verificarse

$P_{\lambda=5}\left[\sum_{i=1}^{n} x_i \le K\right] = \alpha$ y en este caso $\sum_{i=1}^{n} x_i \rightsquigarrow P(5n)$

$K = P_{5n,\alpha}$ (K es el valor de la variable X para el que la probabilidad en la distribución de Poisson de parámetro $5n$ es α)
La región crítica es

$C = \{(x_1, x_2, \dots x_n) / \sum_{i=1}^{n} x_i \le P_{5n,\alpha}\}$. Al ser la variable discreta no está garantizado que exista el test u.m.p.

d) Se trata de contrastar

$H_0: \lambda \le 5$
$H_1: \lambda > 5$

Las hipótesis no son simples, son compuestas, por tanto el test de los apartados anteriores no sirven.
Veamos si la distribución de Poisson cumple la propiedad c.v.m. Si $X \rightsquigarrow P(\lambda)$, pertenece a la familia exponencial.
La función de probabilidad de la distribución de Poisson es:

$P[X = x] = \frac{\lambda^x}{x!} \cdot e^{-\lambda}$. Podemos poner entonces:

$P[X = x] = \frac{1}{x!} \cdot e^{-\lambda} \cdot e^{x \ln \lambda} = a(\lambda) \cdot b(x) \cdot e^{c(\lambda) \cdot d(x)}$

donde $a(\lambda) = e^{-\lambda}$ $b(x) = \frac{1}{x!}$ $c(\lambda) = \ln \lambda$ $d(x) = x$

teniendo en cuenta que $\lambda > 0$.

Como $c(\lambda) = \ln \lambda$ es creciente ya que $\frac{d\, c(\lambda)}{d\lambda} = \frac{1}{\lambda} > 0$

$\Rightarrow \Lambda_C(x_1, x_2, \dots x_n) = h[T(x_1, x_2 \dots x_n], 5, \lambda_1]$ es decreciente respecto a $T = \sum_{i=1}^{n} x_i$. Por el teorema del c.v.m:

Si h es no creciente y el contraste es $\begin{cases} H_0: & \lambda \leq 5 \\ H_1 & \lambda > 5 \end{cases}$

La región crítica del u.m.p es:

$$C = \{(x_1, x_2, \dots x_n) \in \mathbb{R}^n \,/\, \sum_{i=1}^{n} x_i \geq k\}$$

Teniendo en cuenta que $\sum_{i=1}^{n} x_i \rightarrow P[n\lambda] = P[5n]$ si $\lambda = 5$

se debe verificar

$$P[\sum x_i \geq k] = \alpha \Rightarrow P[\sum_{i=1}^{n} x_i \leq k] = 1-\alpha \Rightarrow$$

$$\Rightarrow k = P_{5n, 1-\alpha}$$

8.– Una universidad que ha de comprar una partida de lámparas para retroproyector, acude a una empresa que asegura para ellas una duración media –o esperanza de vida– mínima de 600 horas con una desviación típica de 36.5. La duración de las lámparas sigue una distribución normal. Antes de realizar el pedido pone a prueba una muestra de 36 de ellas, y decide que lo efectuará si comprueba que lo afirmado por la empresa es aceptable. En todo caso, el dato acerca de la desviación típica sí que es considerado válido.

a) Fijado un riesgo de α=5%, ¿cuál es la duración media que debe arrojar la muestra para hacer el pedido?

b) ¿Cuál es la probabilidad de que el pedido sea hecho, si la esperanza de vida de las lámparas es 585? ¿Qué tipo de error mide esa probabilidad? ¿Se contradice su valor con el 5% de riesgo que se fijó para calcular el test?

a) $X \rightsquigarrow N(\mu, 36'5)$ en la muestra $\bar{x} = 600$

Queremos contrastar:

$H_0: \mu \geq 600$

$H_1: \mu < 600$

La distribución $N(\mu, \sigma)$ con σ conocida pertenece a la familia exponencial ya que:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2\sigma^2}(x-\mu)^2} = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2\sigma^2}(x^2 - 2\mu x + \mu^2)}$$

La función de densidad es:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2\sigma^2}\mu^2} \cdot e^{-\frac{1}{2\sigma^2}x^2} \cdot e^{\frac{1}{\sigma^2}\mu x} = a(\mu)\cdot b(x)\cdot e^{c(\mu)\cdot d(x)}$$

donde: $a(\mu) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}\mu^2}$ $\quad b(x) = e^{-\frac{1}{2\sigma^2}x^2}$

$c(\mu) = \frac{1}{\sigma^2}\mu \quad d(x) = x$

Además $c(\mu)$ es creciente ya que $\frac{d\,c(\mu)}{d\mu} = \frac{1}{\sigma^2} > 0$

con lo que:

$\Lambda_c(x_1, x_2, \dots x_n) = h[T(x_1, x_2, \dots x_n), \mu_0, \mu_1]$ es decreciente respecto a $T(x_1, x_2, \dots, x_n] = \sum_{i=1}^{n} x_i$

según el teorema para c.v.m. la región crítica es:

$$C = \{(x_1, x_2, \dots x_n) \in \mathbb{R}^n \,/\, \sum_{i=1}^{n} x_i \leq K\}$$

según el