e-mail: **imozas@elx.uned.es**
http://telefonica.net/web/imm/

EXÁMENES RESUELTOS DE LOS CURSOS 98-99 Y 99-00

JUNIO. 1999

PRIMERA PARTE: CUESTIONES TEÓRICO-CONCEPTUALES

1. La estimación de un parámetro a partir de una muestra se puede comparar al tiro al blanco con fusil. En este paralelismo:

El centro de la diana representa el verdadero valor del parámetro

Cada disparo representa una estimación (muestra) concreta

El fusil es el estimador (es decir, la fórmula de estimación)

En el planteamiento de este símil ¿Cuando diremos que el fusil será eficiente?

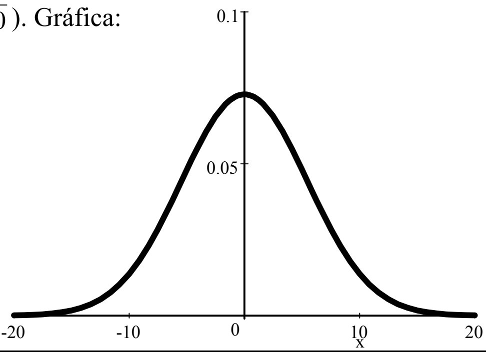
> Debe ser en primer lugar insesgado, es decir, el valor esperado debe ser el centro de la diana (es decir, cuando se apunte al centro se espera que dé en el centro) y además la varianza del estimador debe ser la mínima, es decir, el fusil no tiene que tener ninguna desviación.

2. ¿Guarda alguna relación el concepto de Estimador Eficiente y la Cota de Cramer-Rao?

> La cota de Cramer-Rao proporciona un límite inferior para la varianza del estimador. Si un estimador es eficiente, su varianza es la mínima y coincide con la cota de Cramer-Rao

3. Si tenemos tres distribuciones normales: $X_1 \to N(5, 2)$; $X_2 \to N(-3, 1)$; $X_3 \to N(2, 5)$; ¿Cómo se distribuye la variable aleatoria $U = X_1 + X_2 - X_3$? Representar gráficamente la función de densidad de la variable U.

> U se distribuye $N(0, \sqrt{30})$. Gráfica:



4. ¿Está de acuerdo en que la característica esencial de un Contraste no Paramétrico está en no requerir el conocimiento de la distribución de la población de partida? Señale los Contrastes No Paramétricos más significativos y que objetivos persiguen

> Efectivamente, en los contrastes no paramétricos no se requiere conocer la distribución de la población (y por tanto los estadísticos que se utilizan deben tener distribución independiente de la distribución de la población).
>
> Los contrastes no paramétricos más significativos son: **1)** de aleatoriedad, sobre la aleatoriedad de las muestras; **2)** de localización, sobre medidas de posición (cuantiles); **3)** de comparación de poblaciones, sobre las distribuciones poblacionales.

SEGUNDA PARTE: PROBLEMAS

1. Suponiendo que la cotización de cierre diaria de las acciones del Banco Santander Central Hispano tiene una distribución uniforme entre los 20 y 21 Euros, se pide:

a) ¿Cuál es la probabilidad de que un día la cotización de cierre supere los 20,90 Euros?

b) ¿Cuál es el porcentaje de días que presentaron una cotización de cierre entre 20,40 y 20,60 Euros?

e-mail: **imozas@elx.uned.es**
http://telefonica.net/web/imm/

c) Si sumáramos dos distribuciones uniformes U (20;21) y U (22;24) que fuesen independientes, demuestre si la distribución resultante presenta o no carácter de distribución uniforme

d) Entre los días en los que la cotización de cierre ha sido superior a los 20,50 Euros ¿Cual es el porcentaje de los mismos en que la cotización ha oscilado entre 20,80 y 20,90 Euros?

a) P[X > 20,9] = 0,1; **b)** P[20,4 < X < 20,6] = 0,2 → el 20%; **c)** si X e Y son uniformes (20; 21) y (22; 24) respectivamente, la función generatriz de X+Y sería

$$g_{X+Y}(t) = \frac{e^{21t} - e^{20t}}{t} \cdot \frac{e^{24t} - e^{22t}}{2t},$$

que no corresponde a la función generatriz de una variable aleatoria uniforme.; **d)** P[20,8 < X < 20,9/ X> 20,5] = $\frac{P[20,8 < X < 20,9]}{P[X > 20,5]} = \frac{0,1}{0,5} = 0,2$ → el 20%;

2. Se desea investigar la demanda de un cierto producto. Para ello se consulta a 30 personas, preguntándoles el número de veces que efectúan la adquisición de dicho producto por semana. Las respuestas obtenidas fueron:

Ninguna vez .... 9 de los consultados
Una vez .... 10 de los consultados
Dos veces .... 7 de los consultados
Tres veces .... 3 de los consultados
Cuatro veces .... 1 de los consultados

Se establece la hipótesis de que el modelo de distribución de probabilidad que corresponde a la demanda de la que procede la muestra sigue una distribución de tipo Poisson. Se pregunta si se acepta, o se rechaza, la hipótesis establecida, contrastándola al nivel de significación del 5 %.

Suponiendo la hipótesis, estimemos λ mediante la media muestral, para la que se obtiene un valor λ = 1,23. Construimos la siguiente tabla:

Número de adquisiciones	Frecuencias observadas	Probabilidades teóricas	Frecuencias esperadas	Desviaciones	
x_i	**n_i**	**p_i**	**$n \cdot p_i$**	**$(n_i-np_i)^2$**	**$(n_i-np_i)^2/np_i$**
0	9	0,2913	8,74	0,068	0,0078
1	10	0,3593	10,78	0,607	0,0563
2	7	0,2216	6,65	0,125	0,0188
3	3	0,0911	2,73	0,071	0,0262
4	1	0,0281	0,84	0,025	0,0294
	30		**29,74**		**0,1384**

La variable aleatoria $\chi^2_{exp} = \sum_{i=1}^{5} \frac{(n_i - np_i)^2}{np_i}$ se distribuye como una χ^2_3, luego la región crítica al nivel de significación 5% viene dada por la condición: χ^2 >7,815. Puesto que χ^2_{exp}=0,1384, se acepta la hipótesis.

JUNIO 1999 RESERVA

SEGUNDA PARTE: PROBLEMAS

1. Los gastos de transporte que realiza una oficina oscilan uniformemente entre 100.000 y 140.000 pesetas al mes. Se pregunta:

a) ¿Cuál es la probabilidad de que en un mes determinado el gasto en transporte sea exactamente 120.000 pesetas?

b) Calcule la desviación típica del gasto mensual.

c) Estimándose que el gasto en transporte es excesivo, se pretende llevar a cabo un control para comprobar la necesidad de dicho gasto. Para ello se observa aleatoriamente el gasto mensual durante 3 años ¿Cual es la probabilidad de que el gasto mensual medio, durante esos 3 años se superior a 130.000 pesetas?

> **a)** cero; **b)** $\sigma = \frac{2\sqrt{3}}{3}\cdot 10^4$; **c)** Siendo X_i el gasto mensual observado, consideramos la variable $\overline{X} = \frac{1}{36}\sum_{i=1}^{36} X_i$ para la que se tiene E($\overline{X}$) = μ = 120000 y DT($\overline{X}$) = $\frac{\sigma}{\sqrt{36}} = \frac{\sqrt{3}}{9}\cdot 10^4$. Del teorema central del límite se deduce que $\overline{X}$ se distribuye aproximadamente N(12000, $\frac{\sqrt{3}}{9}\cdot 10^4$) luego P[$\overline{X}$>130000] = P[Z>$3\sqrt{3}$] ≅ 0.

2. Para discutir la conveniencia de aumentar sus instalaciones, una empresa desea estimar la demanda que espera recibir. Para ello selecciona a 10 de sus clientes habituales, al azar, observando que el número de unidades demandadas en último año por estos, se distribuye en la forma siguiente:

Nº de Unidades	Nº de clientes
1.000	1
1.002	2
1.004	1
1.006	2
1.008	1
1.010	2
1.012	1

Suponiendo que la demanda siga comportándose de manera similar en el siguiente período, se pide:

a) Las estimaciones de la demanda media y de la desviación típica.

b) Si se toma como desviación típica de la población la obtenida en el apartado anterior, determinar un intervalo de confianza para la demanda media del 95 % en los siguientes casos:

- Sin efectuar hipótesis sobre la distribución de la demanda
- Suponiendo que la demanda se comporte con arreglo a una distribución de tipo normal,

> **a)** Estimaremos por el método de los momentos, obteniéndose: $\hat{\mu}$=1006 y $\hat{\sigma} = \sqrt{14{,}4}$; **b)** aplicaremos la desigualdad de Chebychev a la variable $\overline{X}$, para la cual se tiene E($\overline{X}$) = = μ y DT($\overline{X}$) = $\frac{\sigma}{\sqrt{n}} = \sqrt{1{,}44} = 1{,}2$; así pues, de la desigualdad de Chebychev, P[$\overline{X} - k\frac{\sigma}{\sqrt{n}} \le \mu \le \overline{X} + k\frac{\sigma}{\sqrt{n}}$] $\ge 1-\frac{1}{k^2}$ y sustituyendo $1-\frac{1}{k^2}$= 0,95, obtenemos el intervalo [1000,63, 1011,37]; para la N(0,1) , P[–1,96 < Z < 1,96] = 0,95 luego si X es N(μ, σ) $\rightarrow \overline{X}$ es N($\mu, \frac{\sigma}{\sqrt{n}}$) y se obtiene el intervalo [1003,65 , 1008, 35].

e-mail: **imozas@elx.uned.es**
http://telefonica.net/web/imm/

SEPTIEMBRE 1999

PRIMERA PARTE: CUESTIONES TEÓRICO-CONCEPTUALES

1. Partiendo de la Función Generatriz de Momentos de una distribución de Poisson P(λ), demostrar cuánto vale la media de esta distribución.

$g(t) = e^{\lambda(e^t-1)} \rightarrow g'(t) = e^{\lambda(e^t-1)} \cdot \lambda \cdot e^t \rightarrow \mu = g'(0) = \lambda.$

2. Demostrar que si dos variables aleatorias $X_1 \rightarrow N(9, 4)$ y $X_2 \rightarrow N(8, 3)$ son normales e independientes, la variable aleatoria suma $U= X_1 + X_2$ es también una distribución normal de características: $U \rightarrow N(17, 5)$

La función generatriz de una v. a. N(μ, σ) es $g(t) = e^{t\mu+\frac{1}{2}t^2\sigma^2}$. La función generatriz de U será: $E(e^{tU}) = E\left(e^{t(X_1+X_2)}\right) = E\left(e^{tX_1}\right). E\left(e^{tX_2}\right) = e^{9t+\frac{1}{2}t^2 16} \cdot e^{8t+\frac{1}{2}t^2 9} = e^{17t+\frac{1}{2}t^2 25}$ que corresponde a una normal N(17, 5).

3. ¿Para qué sirve la Cota de Cramer-Rao en la teoría de la estimación?.

La cota de Cramer-Rao proporciona un límite inferior para la varianza del estimador $\hat{\theta}$, es decir:

$$\mathrm{Var}(\hat{\theta}) \geq \frac{1}{nE\left[\left(\frac{\partial \ln f(x,\theta)}{\partial\theta}\right)^2\right]}$$

Si $\hat{\theta}$ es insesgado y su varianza coincide con la cota de Cramer-Rao, diremos que es eficiente.

4. Partiendo de muestras obtenidas por muestreo aleatorio simple de tamaño n, y con un nivel de confianza prefijado, obtenga de forma razonada el intervalo de confianza para la media de una población normal, con varianza poblacional conocida.

La media muestral $\overline{X}$ se distribuye $N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$, luego $Z = \frac{\overline{X}-\mu}{\sigma}\sqrt{n}$ se distribuye N(0, 1). Dado un nivel de confianza 1–α, sea $z_{\frac{\alpha}{2}}$ tal que $P(Z > z_{\frac{\alpha}{2}}) = \frac{\alpha}{2}$, luego :

$P\left(-z_{\frac{\alpha}{2}} < \frac{\overline{X}-\mu}{\sigma}\sqrt{n} < z_{\frac{\alpha}{2}}\right)$ = 1– α, de donde se deduce que $\mu \in \left[\overline{X} - \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}, \overline{X} + \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}\right]$ con probabilidad 1–α

SEGUNDA PARTE: PROBLEMAS

1.- El volumen de agua consumido por una ciudad se puede representar por una variable aleatoria X y oscila entre 5.000 y 6.000 unidades al mes. La función de densidad de probabilidad del citado consumo toma siempre un valor constante, de valor igual a c en el intervalo (5.000, 6.000) y valores nulos en el resto. Calcular razonadamente.

a) ¿Cuál es el volumen esperado de agua consumida durante un mes?.

b) Significado y valor del coeficiente de variación del agua mensual consumida.

c) ¿Cuál seria el volumen esperado de consumo de agua durante un año?.

d) ¿Qué probabilidad existe de que el consumo de agua en un mes sea superior a 5.100 unidades e inferior a 5900?

e-mail: **imozas@elx.uned.es**
http://telefonica.net/web/imm/

e) ¿Tiene sentido aplicar la Desigualdad de Chebychev para tratar de resolver el punto anterior?.

Se trata de una variable aleatoria uniforme con función de densidad:

$$f(x) = \begin{cases} 0{,}001, & 5000 \le x \le 6000 \\ 0, & \text{resto} \end{cases}$$

Se tiene: **a)** E(X) = $\frac{a+b}{2}$ =5500; **b)** CV = $\frac{\sqrt{Var(X)}}{E(X)} = \frac{288{,}67}{5500}$ = 0,052; **c)** Si U = $\sum_{i=1}^{12} X_i$ es el consumo durante un año → E(U) = 12·5500 = 66000; **d)** P[5100 < X < 5900] = 0,8; **e)** No porque conocemos la función de densidad.

2.- Supongamos que el Banco de España decide efectuar una investigación sobre los rendimientos obtenidos por la Banca española con un determinado producto financiero.

Para ello selecciona una muestra aleatoria simple de 9 bancos, y además dispone de la información de que los rendimientos del producto en cuestión, en todo el conjunto bancario, se distribuyen según una distribución normal de media 6% y desviación típica del 3%. En base a ello se pide:

1º) ¿Cual es la probabilidad de que el rendimiento medio muestral se mantenga entre el 5% y el 7% ?.

2º) ¿Cuál es la probabilidad de que la varianza muestral sea superior a 9 ? .

3º) El valor de K tal que $P(S^2 > K) = 0{,}98$.

4º) Suponiendo, ahora, que la desviación típica para todo el conjunto bancario fuera desconocida, y conociésemos que la desviación típica de la muestra de 9 bancos es del 2%, se pide obtener la probabilidad de que la media muestral sea superior al 8%.

1º) $\overline{X}$ es N(6, 1) y se obtiene que P[5 < $\overline{X}$ < 7] = 0,6827; **2º)** $\frac{(n-1)S^2}{\sigma^2} = \frac{8S^2}{9}$ es χ^2_8, luego $P(S^2 > 9) = P\left(\frac{8S^2}{9} > \frac{8}{9}9\right) = P(\chi^2_8 > 8) = 0{,}4335$; **3º)** $0{,}98 = P(S^2 > k) = P\left(\frac{8S^2}{9} > \frac{8k}{9}\right)$ y de las tablas se obtiene que $\frac{8k}{9}$ =2,0325 → k = 2,2865; **4º)** $\frac{\overline{X}-\mu}{S}\sqrt{n} = \frac{\overline{X}-6}{2}.3$ es t_8 →

$P[\overline{X} > 8] = P\left[\frac{\overline{X}-6}{2}\cdot 3 > 3\right] = 0{,}008$

SEPTIEMBRE 1999 (RESERVA)

PRIMERA PARTE: CUESTIONES TEÓRICO-CONCEPTUALES

1.- ¿Puede ser negativa la potencia de un contraste?. Si β es la probabilidad de cometer un error de tipo II, ¿cómo será la potencia del contraste?

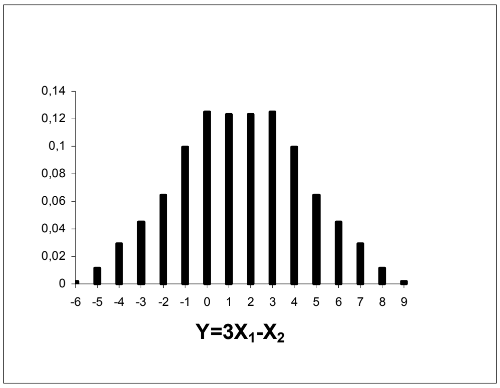
No puede ser negativa por que es la probabilidad de rechazar H_0. Si β = probabilidad de cometer error de tipo II → potencia = 1 – β

2.- Si se tienen dos variables aleatorias independientes X_1, y X_2 distribuidas según una B(3;0,5) y B(6;0,5) respectivamente, calcular la distribución que seguirá Y para $Y=3X_1 - X_2$

La función generatriz $g_Y(t) = E[e^{Yt}] = E\left[e^{3X_1t}\right]\cdot E\left[e^{-X_2t}\right] = g_{X_1}(3t)\cdot g_{X_2}(-t) =$
$= (0{,}5 + 0{,}5\cdot e^{3t})^3\cdot(0{,}5 + 0{,}5\cdot e^{-t})^6$ que no corresponde a ninguno de los tipos notables de distribuciones. La función de probabilidad sería: $P(Y = n) = \left(\sum_{3i-j=n} \binom{3}{i}\binom{6}{j}\right)(0{,}5)^9$, n = –6, –5, –4, ..., 9. Por ejemplo $P(Y=3) = P(X_1=1;X_2=0)+P(X_1=2;X_2=3)+ P(X_1=3;X_2=6)=$

e-mail: imozas@elx.uned.es
http://telefonica.net/web/imm/

$= \left(\binom{3}{1}\binom{6}{0} + \binom{3}{2}\binom{6}{3} + \binom{3}{3}\binom{6}{6}\right)0,5^9 = 64\cdot 0,5^9$. Etc.. Se obtiene la siguiente representación gráfica:



3.- Razone que distribución utilizaría para crear un intervalo de confianza para μ, en una población normal si se conoce σ^2.

La distribución normal, pues la media muestral $\overline{X}$ se distribuye $N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$ lo cual nos permite construir un intervalo de confianza para μ.

4.- Indique, razonándolo, cuál es la relación existente entre las distribuciones Binomial, Poisson y Normal.

Se demuestra que $\lim_{n\to\infty}\binom{n}{x}p^x(1-p)^{n-x} = \frac{\lambda^x}{x!}e^{-\lambda}$, donde λ = np. Se considera que la aproximación es adecuada si n ≥30 y p ≤ 0,1. Si X es B(n, p) entonces (T. de Moivre) $\frac{X-np}{\sqrt{npq}}$ tiende a ser N(0, 1) cuando n →∞ (se considera adecuada la aproximación si $p \le \frac{1}{2}$ y np > 5 ó si $p > \frac{1}{2}$ y nq > 5. Finalmente, si X es Poisson de parámetro $\lambda \ge 10$, $\frac{X-\lambda}{\sqrt{\lambda}}$ se aproxima a una N(0, 1)

SEGUNDA PARTE: PROBLEMAS

1. El consumo en bienes de capital de las empresas del sector Industrial, se distribuye normalmente con una desviación típica de 13,3 millones de ptas. La creencia general al respecto es que el consumo en bienes de capital no es inferior a 30 millones de ptas. Para contrastar este supuesto, se realiza un estudio; en el cual, tras la obtención de una muestra aleatoria de 120 empresas del sector, se observa un consumo medio en bienes de capital de 45 millones de ptas. ¿Sería admisible mantener la creencia general como cierta con un 2% de significación?

Consideremos las hipótesis: $H_0 : \mu \ge 30$ y H_1: $\mu < 30$. $Z = \frac{\overline{X}-\mu}{\sigma}\sqrt{n}$ es N(0, 1) y se tiene que 0,02 = P[Z < –2,05] de donde se obtiene que la región crítica viene dada por la condición: $\overline{X} < 27,51$, luego debemos admitir H_0.

2. En una empresa de construcción el número mensual medio de días de baja de los trabajadores debidas a accidentes laborales por obra, se distribuye según una Poisson con varianza 15,6. Calcular la probabilidad de que en cuatro obras independientes, se acumulen más de 68 días, en total, de baja en el periodo de un mes.

$Y = \sum_{i=1}^{4} X_i$ es de Poisson de parámetro 4λ = 62,4 $\Rightarrow \dfrac{Y - 62,4}{\sqrt{62,4}}$ es N(0,1) y se obtiene P[Y ≥ 68,5] = **0,2200** (se ha efectuado la corrección por continuidad)

JUNIO 2000

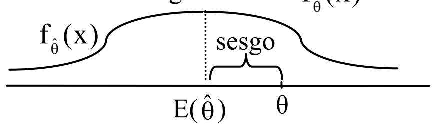
PRIMERA PARTE: CUESTIONES TEÓRICO-CONCEPTUALES

1.- Indique el significado de la desigualdad de Chebychev.

$P\left[\left|X - E(X)\right| \geq k\right] \leq \dfrac{\sigma^2}{k^2}$; proporciona una cota superior para la probabilidad de que la distancia de un valor de la variable a su esperanza supere a cierto número k. Una consecuencia es : $P\left[\left|X - E(X)\right| \geq k\sigma\right] \leq \dfrac{1}{k^2}$; por ejemplo, $P\left[\left|X - E(X)\right| \geq 2\sigma\right] \leq \dfrac{1}{4}$, $P\left[\left|X - E(X)\right| \geq 3\sigma\right] \leq \dfrac{1}{9}$, etc.

2.- Explique el concepto de sesgo en una estimación, de forma gráfica.

Se denomina sesgo de un estimador $\hat{\theta}$ a la diferencia $E(\hat{\theta}) - \theta$. Si esta diferencia es cero, diremos que el estimador es insesgado. Sea $f_{\hat{\theta}}(x)$ la función de densidad de $\hat{\theta}$. Gráficamente:



3.- ¿Cuándo se puede considerar a un contraste de hipótesis como el de máxima potencia?.

Cuando el tamaño β del error de tipo II sea cero

4.- Distinga entre Parámetro, Estadístico y Estimador.

Dada una población con una variable aleatoria X y una función de probabilidad o de densidad, se llama **parámetro** a cualquier <u>constante</u> que caracterice dicha función (por ejemplo, E(X), Var(X), etc.).

Elegida una muestra $X_1, X_2, ..., X_n$, se llama **estadístico**, a cualquier <u>variable aleatoria</u> que sea función de la muestra: $g(X_1, X_2, ..., X_n)$.

Si θ es un parámetro desconocido, un **estimador** será un estadístico que elegiremos para estimar el parámetro (por ejemplo si θ = E(X) desconocido, podemos elegir como estimador $\overline{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$).

SEGUNDA PARTE: PROBLEMAS

1.- Los resultados obtenidos en una investigación de mercado han determinado que el 25% de los consumidores son clientes habituales de los productos de la empresa R. Si se eligiese al azar a 10 consumidores, calcular: **a)** La probabilidad de que entre los 10 consumidores se encuentren:

- un máximo de 3 clientes.
- Entre 4 y 7 clientes.

b) El número esperado de clientes; **c)** La varianza de la distribución

a) La variable aleatoria X = "nº de clientes" es B(10; 0,25); puesto que n·p = 2,5 < 5 no procedería aproximar por una normal. Así pues:

P[X≤ 3] = $\sum_{i=0}^{3}\binom{10}{i} 0,25^i \cdot 0,75^{10-i}$ = 0,7759; P[4 ≤ X≤ 7] = P[X≤7] – P[X≤3] = 0,9996 – 0,7759 = = 0,2237; **b)** E[X] = n·p = 2,5 ; **c)** Var[X] = npq = 1,875

e-mail: **imozas@elx.uned.es**
http://telefonica.net/web/imm/

2.- Una aseguradora realiza un análisis de los costes de tramitación por accidentes. De este análisis se obtiene que los costes son 6.000 ptas de media. Este coste se consideró poco competitivo, para lo cual se rehicieron los procesos de tramitación. Una vez instaurados los nuevos procesos y con el objeto de evaluar su eficacia, se eligió una muestra aleatoria de 26 expedientes, y se les realizó un nuevo análisis de los costes. De este análisis se obtuvieron los siguientes resultados: la media muestral de los costes fue de 5.600 ptas y la desviación típica de 1.000 ptas. Para un nivel de significación del 0,01, ¿se puede aceptar que el nuevo proceso ha sido eficiente, o no (la reducción de costes se debe a variaciones muestrales no significativas)?.

Efectuamos las hipótesis H_0: $\mu \geq 6000$ y H_1: $\mu < 6000$; la región crítica viene definida por la condición : $t_{25} < -2{,}485$; puesto que $\frac{5600-6000}{1000}\sqrt{26} = -2{,}0396$, luego no podemos rechazar H_0 y por tanto, la reducción de costes se debe a variaciones muestrales no significativas.

JUNIO 2000 (RESERVA)

PRIMERA PARTE: CUESTIONES TEÓRICO-CONCEPTUALES

1 . ¿Qué se entiende por Potencia del Contraste.

Cuando se efectúa el contraste de una hipótesis nula H_0 con una alternativa H_1, se llama error de tipo II a la probabilidad de aceptar H_0 siendo falsa, probabilidad que representaremos por β. Se llama potencia del contraste a $1 - \beta$ = probabilidad de rechazar H_0, siendo falsa.

2. ¿Qué debe cumplir un estimador insesgado para que se le pueda catalogar como óptimo?

Que su varianza sea la mínima, es decir, que alcance la cota de Cramer-Rao:

$$\mathrm{Var}(\theta) = \frac{1}{E\left[\left(\frac{\partial \ln dF_n}{\partial\theta}\right)^2\right]}$$

3. Si la variable aleatoria U está distribuida uniformemente en $-4 \leq u \leq 4$, determinar $P(|U - 2| < 2)$.

Será $f(u) = \begin{cases} \frac{1}{8}, & -4 \leq u \leq 4 \\ 0, & \text{resto} \end{cases} \rightarrow P[|U-2| < 2] = 4 \cdot \frac{1}{8} = 0{,}5.$

4. Explique qué condiciones debe cumplir una función f(x) de una variable aleatoria continua X para que se la pueda considerar como función de densidad.

1º) $f(x) \geq 0,\ \forall x \in \mathbb{R}$; 2º) $\int_{-\infty}^{+\infty} f(x)dx = 1$

SEGUNDA PARTE: PROBLEMAS

1.- La media de coches vendidos al día de una determinada marca es de 5. Suponiendo que las ventas siguen una distribución de Poisson, calcular la probabilidad de que: **a)** en un día se vendan más de 6 coches; **b)** en una semana (6 días hábiles) se vendan 20 coches.

e-mail: **imozas@elx.uned.es**
http://telefonica.net/web/imm/

$P[X = x] = \frac{\lambda^x}{x!} e^{-\lambda}$ y $E[X] = \lambda = 5$. **a)** $P[X > 6] = 1 - \sum_{i=0}^{6} P[X = i] = 0{,}238$.

b) la variable $Y = \sum_{i=1}^{6} X_i$ es de Poisson con $E[Y] = 6\lambda = 30$, luego $P[Y = y] =$

$= \frac{30^y}{y!} e^{-30} \Rightarrow P[Y = 20] = \frac{30^{20}}{20!} e^{-30} = 0{,}01341$

2.- Se supone, que al incrementar el precio de un producto por encima de 5 € (debido las elasticidades demanda precio respecto a los bienes sustitutivos) se disminuirá su consumo. Para comprobarlo se realizó un experimento aleatorio en cinco hipermercados y resultó que el precio de reacción (al que empezaron a disminuir las compras) fue de 4,8 € con una desviación típica de 0,3 €. ¿son los resultados estadísticamente significativos al nivel 0,05?

Suponemos que el precio de reacción se distribuye $N(\mu, \sigma)$. Pondremos H_0: $\mu \geq 5$ y H_1: $\mu < 5$. La región crítica a un nivel de significación 0,05, viene dada por la condición: $t_4 < -2{,}132$; como $\frac{4{,}8 - 5}{0{,}3}\sqrt{5} = -1{,}4907$, se acepta H_0, es decir, la discrepancia no es estadísticamente significativa.

SEPTIEMBRE 2000

PRIMERA PARTE: CUESTIONES TEÓRICO-CONCEPTUALES

1- Indicar las diferencias entre una función de cuantía y una función de densidad.

Si X es una variable aleatoria discreta, entonces $P[X = x]$ es la función de cuantía. Si X es una variable continua, y $F(x) = P[X \leq x]$ es su función de distribución, entonces $f(x) = \frac{dF(x)}{dx}$ es la función de densidad; es decir, $P[X \leq x] = \int_{-\infty}^{x} f(t)dt$

2.- Indicar la relación existente entre una normal N(0; 1) y una normal N(5;2).

Si X es N(5, 2) $\Rightarrow Z = \frac{X-5}{2}$ es N(0, 1) $\leftrightarrow X = 2Z + 5$

3.- Explicar el porqué de la utilización del estadístico $\frac{(n-1)S^2}{\sigma}$, para el cálculo de un intervalo de confianza para la varianza poblacional σ^2, si la media poblacional μ es desconocida.

El estadístico $\frac{(n-1)S^2}{\sigma}$ es χ^2_{n-1} y, conocido n, podemos encontrar un intervalo (las dos colas del mismo tamaño) con un nivel de confianza $1 - \alpha$, de forma que $P\left[a < \frac{(n-1)S^2}{\sigma^2} < b\right] = 1 - \alpha$ y de aquí obtener el intervalo para σ^2 .

4.- Indicar la región de rechazo, de aceptación y el valor crítico de una hipótesis H_0 cualesquiera, con un nivel de significación del 0,05 para una prueba de dos colas en una normal N(0, 1).

De las tablas $P[-1{,}96 < Z < 1{,}96] = 0{,}95$. Luego:
Región de rechazo: $]-\infty, -1{,}96] \cup [1{,}96, +\infty[$
Región de aceptación: $[-1{,}96, 1{,}96]$
Valores críticos: $-1{,}96$ y $1{,}96$

SEGUNDA PARTE: PROBLEMAS

1.- Un hipermercado acepta llevar a cabo una política de fidelización al cliente basada en admitir todas las devoluciones de los clientes. La media de clientes que devuelven artículos es de 11,5 y tiene una desviación típica de 3.2 clientes al día. ¿Cual es el porcentaje de días en que hay : **a)** menos de 8 clientes que devuelven artículos?; **b)** entre 12 y 14 clientes?; **c)** ningún cliente devuelve artículos?

> Supondremos que X = "nº de clientes que devuelven artículos un día" se distribuye N(11,5 ; 3,2), luego, efectuando la correspondiente corrección por continuidad: **a)** $P[X \le 7,5] = P[Z \le -1,25] = 0,1056 \to 10,56\%$; **b)** $P[11,5 < X < 14,5] = 0,3257 \to 32\%$; **c)** $P[-0,5 < X < 0,5] = 0,0002 \cong 0\%$

2.- Los ingresos anuales medios de los economistas con un año de experiencia, en las empresas del sector servicios son de 30.000 € con una desviación típica de 3.000 €. Una empresa del sector desea saber si en su organización esta tipología de trabajador cobra más o menos que los 30.000 € de media del sector.

Se utiliza como hipótesis alternativa que la media no es 30.000 € y se debe realizar con un nivel de significación de 0,10. Se seleccionó una muestra de 120 economistas con un año de experiencia. Los ingresos medios resultaron ser de 30.500 € ¿se debería rechazar la hipótesis nula?

> Supondremos que los ingresos de los economistas se distribuyen de forma normal. Consideraremos H_0: $\mu = 30000$; H_1: $\mu \neq 30000$. La variable $Z = \frac{\bar{X} - \mu}{\sigma}\sqrt{n}$ es N(0,1) y $P[|Z| > z_0] = 0,1 \to z_0 = 1,64$. Como $\frac{30500 - 30000}{3000}\sqrt{120} = 1,8257 > 1,64$, debe rechazarse H_0.