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

EXÁMENES RESUELTOS SEPTIEMBRE 2002

PRINCIPAL

PRIMERA PARTE: CUESTIONES TEÓRICO-CONCEPTUALES

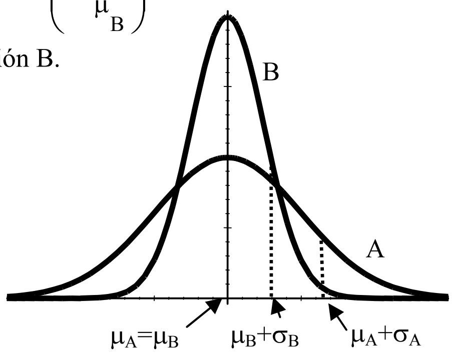
1. Comente brevemente qué utilidad tiene la desigualdad de Chebychev.

La desigualdad de Chebychev: $P[\,|X-\mu| \geq k] \leq \dfrac{\sigma^2}{k^2}$ o equivalentemente $P[\,|X-\mu| \leq k] \geq 1 - \dfrac{\sigma^2}{k^2}$ permite acotar probabilidades cuando se dispone de una variable aleatoria con media y varianza conocidas pero de distribución desconocida. Si no se conoce la media μ, entonces podemos obtener un intervalo de confianza con un nivel de confianza $\geq 100(1-\alpha)\%$ pues, siendo $\overline{X}$ la media muestral (tamaño de la muestra n) $\Rightarrow P[\,|\overline{X} - E(\overline{X})| \leq k] \geq 1 - \dfrac{Var(\overline{X})}{k^2} = 1 - \dfrac{\sigma^2}{nk^2} = 1-\alpha$, de donde se obtiene el intervalo $\left[\overline{X} - \dfrac{\sigma}{\sqrt{n\alpha}}, \overline{X} + \dfrac{\sigma}{\sqrt{n\alpha}}\right]$

2. Qué interpretación tiene el hecho de que en dos distribuciones normales Ay B con medias iguales $\mu_A = \mu_B$, la varianza de A sea superior a la de B. Representar gráficamente las distribuciones de densidad de ambas poblaciones.

Al ser iguales las medias, entonces el coeficiente de variación de A $\left(=\dfrac{\sigma_A}{\mu_A}\right)$, es superior al coeficiente de variación de B $\left(=\dfrac{\sigma_B}{\mu_B}\right)$, es decir, la dispersión de la población A es superior a la dispersión de la población B.

Gráficas:



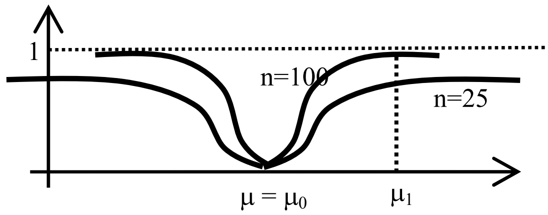
3. ¿Cuándo podemos decir que un estimador es consistente?

Sea $X = X_i$, $i = 1, 2, 3, ..., n$ una variable aleatoria. Sea θ un parámetro y $\hat{\theta}_i = g(X_1, ..., X_i)$, $i = 1, 2, 3, ..., n$, una sucesión de estimadores. Diremos que la sucesión es consistente si $\lim\limits_{n\to\infty} P\left[\,|\hat{\theta}_n - \theta| < \varepsilon\right] = 1$ (es decir, converge en probabilidad hacia θ). Cada elemento de la sucesión es un estimador consistente.

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

4. ¿Qué efecto tiene un incremento del tamaño muestral sobre la potencia de un contraste?

Para un nivel de significación α fijo, al aumentar el tamaño muestral disminuye el error de tipo II, β = P[Aceptar H_0/H_0 falsa], por tanto, **aumenta** la potencia del contraste $1 - \beta =$ = P[Rechazar H_0/H_0 falsa]. Si representamos la curva de potencia para dos tamaños muestrales, p. ej. n = 25 y n = 100, se obtienen curvas del tipo:



SEGUNDA PARTE: PROBLEMAS

1.- Una compañía de seguros tiene en la actualidad 3.000 pólizas contratadas, que cubren al asegurado de un accidente doméstico. La incidencia de este tipo de accidentes es del 3 por mil. En caso de accidente la compañía indemniza al asegurado con 1.000 €.

a) Probabilidad de que ocurran menos de 3 siniestros.

b) Probabilidad de que ocurran al menos 5 siniestros.

c) Calcular la esperanza matemática de la indemnización.

d) Justificar la distribución de probabilidad utilizada para la resolución del problema.

La variable X = nº de siniestros es binomial B(3000; 0,003) que se distribuye aproximadamente igual que una Poisson P(9). Así pues (de las tablas):
a) $P[X < 3] = 0{,}0062$; **b)** $P[X \geq 5] = 1 - P[X \leq 4] = 0{,}945$; **c)** $E(X) = 9 \Rightarrow$ la esperanza de la indemnización será de 9000 €; **d)** puesto que $n = 3000 > 30$ y $p = 0{,}003 < 0{,}1$, la distribución binomial se puede aproximar por una distribución de Poisson de parámetro $\lambda =$ $= np = 9$, ya que $\lambda < 10$.

2. - Las ventas medias semanales de la revista X según un estudio publicado por una auditoria de medios es de 20.000 ejemplares. La editorial que vende la revista X discrepa sobre ese dato y por ello encarga un estudio de mercado para contrastar el mismo, del cual y con una muestra tomada durante 10 semanas se obtiene que su venta media semanal es de 22.000 ejemplares y que su varianza es de 9.000.000. Con esta información y suponiendo un comportamiento normal de la variable estudiada, la empresa quiere contrastar a un nivel de significación del 0,05 si su venta media semanal realmente es mayor que la publicada.

Consideramos las hipótesis $H_0 : \mu \leq 20.000$ y $H_1: \mu > 20.000$. La variable $\frac{\overline{X} - \mu}{S}\sqrt{n}$ se distribuye $t_{n-1} = t_9$. De las tablas obtenemos la región de rechazo: $\frac{\overline{X} - \mu}{S}\sqrt{n} > 1{,}833$.

Sustituyendo: $\frac{2000}{3000}\sqrt{10} = 2{,}11 > 1{,}833$, luego debe rechazarse H_0, es decir, la venta media semanal es mayor de 20.000 ejemplares.

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

RESERVA

PRIMERA PARTE: CUESTIONES TEÓRICO-CONCEPTUALES

1.- Explique brevemente qué se entiende por función de Distribución.

Sea X una variable aleatoria y x un número real cualquiera. La función $F(x) = P[X \leq x]$ se denomina función de distribución.

2.- Relación entre la distribución Binomial y la distribución de Bernoulli.

Si X_i , i = 1, 2, ..., n son n variables aleatorias de Bernoulli de parámetro p, entonces la variable $X = \sum_{i=1}^{n} X_i$ es una variable aleatoria binomial de parámetros (n, p).
Una variable Bernoulli de parámetro p es una variable binomial de parámetros (1, p)

3.- ¿Que diferencia existe entre los estimadores obtenidos por el método de los momentos y los obtenidos por el método de máxima verosimilitud?

Los primeros son menos eficientes que los segundos, aunque son más fáciles de calcular. Además, el método de los momentos no usa toda la información contenida en la muestra, en particular, no hace uso de la distribución, cosa que sí hace el método de la máxima verosimilitud, por lo que los obtenidos por este segundo método son mejores.

Con más detalle:

	Insesgadez	**Consistencia**	**Normalidad asintótica**	**Eficiencia**	**Eficiencia asintótica**	**Suficiencia**
Momentos	Si, si el parámetro a estimar es un momento poblacional respecto del origen	Sí, con algunas condiciones	Sí	No siempre	No siempre	No siempre
Máxima Verosimilitud	No siempre	Sí	Sí	No siempre	Sí	Si existe un estimador suficiente, el estimador máximo verosímil es función de él

4.- ¿Qué interpretación tiene el que digamos que el intervalo de confianza para el parámetro poblacional μ es de μ ± x (siendo x una cantidad previamente calculada) con un nivel de confianza del 99%?

Si tomamos 100 muestras de tamaño n y construimos para cada una de ellas el correspondiente intervalo, entonces esperamos que el 99 % de ellos contendrá a μ.

SEGUNDA PARTE: PROBLEMAS

1. - Un concesionario de coches, después de realizar un estudio observacional sobre los comportamientos y pautas de los clientes, ha concluido que 6 de cada 100 clientes que entran en el concesionario acaban comprando un coche. Con esta información obtener: **a)** La probabilidad de que se vendan 2 coches si entran 8 clientes. **b)** La probabilidad de que si un

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

día entran 30 clientes se vendan al menos 6 coches. **c)** La probabilidad de que se vendan más de 20 coches si entran 200 clientes. **d)** Justificar la distribución de probabilidad utilizada para la resolución del problema.

> **a)** La variable X = nº de coches que se venden es B(8; 0,06), luego P[X = 2] = $= \binom{8}{2} 0{,}06^2 \cdot 0{,}94^6 = 0{,}0695$; **b)** La variable X es aquí B(30; 0,06) que se distribuye aproximadamente igual que una de Poisson P(1,8), luego P[X ≥ 6] = 1 – P[X ≤ 5] = *(tablas)* = = 1 – 0,9896 = 0,0104; **c)** La variable X es aquí B(200; 0,06) que se distribuye aproximadamente igual que una normal N(12; 3,36); así pues, haciendo la oportuna corrección por continuidad: P[X > 20] = P[X ≥ 20,5] = P[Z ≥ 2,53] = 0,0057; **d)** en el caso b se cumple que n ≥ 30 y p < 0,1 luego la aproximación de la binomial por la de Poisson es adecuada; en el caso c se cumple que np ≥ 5 y p < 0,5, luego la aproximación de la binomial por la normal es adecuada.

2.- Mediante un muestreo aleatorio simple realizado en 15 hoteles de 4 estrellas de la Costa del Sol se ha obtenido que los precios de las habitaciones dobles por noche son:

108 €; 125 €; 112 €; 131 €; 96 €; 126 €; 120 €; 120 €; 78 €; 120 €; 75 €; 141 €; 120 €; 79 €; 120 €

Determinar, (si sabemos de antemano que los precios de las habitaciones siguen una distribución normal), los intervalos de confianza para el precio medio y para la varianza del precio, con un nivel de significación del 0,05.

> La variable $\frac{\overline{X} - \mu}{S}\sqrt{n}$ se distribuye $t_{n-1} = t_{14}$. De las tablas obtenemos que $-2{,}145 \le \frac{\overline{X} - \mu}{S}\sqrt{n} \le 2{,}145$, con probabilidad 0,95. Con los datos del problema se obtiene $\overline{X} = 111{,}4$ y S = 20,297, luego el intervalo es $[100{,}16 \le \mu \le 122{,}64]$.
>
> La variable $\frac{(n-1)S^2}{\sigma^2}$ se distribuye $\chi^2_{n-1} = \chi^2_{14}$. De las tablas obtenemos que $5{,}629 \le \frac{(n-1)S^2}{\sigma^2} \le 16{,}12$, con probabilidad 0,95. Sustituyendo, se obtiene el intervalo $[357{,}79 \le \sigma^2 \le 1024{,}62]$