CURSO 2002/2003.	**JUNIO. Principal.**
Código de la Carrera 42	**Código de la Asignatura 209.**

PRIMERA PARTE: CUESTIONES TEÓRICO-CONCEPTUALES

1. Explique para que sirven los momentos de una variable aleatoria.

Permiten cuantificar medidas de dispersión y de forma para comparar distribuciones

2.¿Qué relación existe entre la distribución Binomial y la de Bernoulli?

La variable aleatoria X sigue una distribución de Bernoulli B(1, p) si su función de probabilidad es: $P(X = x) = \begin{cases} \mathbf{p}, \text{ si } x = 1 \\ \mathbf{q = 1 - p}, \text{ si } x = 0 \\ \mathbf{0}, \text{ para el resto de valores de } x \end{cases}$, siendo $0 < p < 1$.

Una variable X sigue una distribución binomial B(n, p) si es suma de n variables de Bernoulli B(1, p) independientes. En ese caso. la función de probabilidad es P(X = x) =

$$= \begin{cases} \binom{\mathbf{n}}{\mathbf{x}} \mathbf{p}^{\mathbf{x}} \mathbf{q}^{\mathbf{n-x}}, \text{ si } x = 0, 1, 2, ..., n \\ \mathbf{0}, \text{ para el resto de valores de } x \end{cases}$$

3.¿Porqué se considera más aconsejable la utilización de los estimadores máximo verosímiles frente a los estimadores obtenidos por el método de los momentos?

Por que, en general, los estimadores obtenidos por el método de la máxima verosimilitud, son más eficientes que los obtenidos por el método de los momentos

4.¿Por qué se utilizan los contrastes de significación sabiendo que normalmente no son contrastes uniformemente más potentes?

Por que los contrastes de significación tienen muy buenas propiedades cuando el tamaño de la muestra es grande.

SEGUNDA PARTE:PROBLEMAS

1. Se sabe que el 10% de los hogares que reciben un catálogo de venta a distancia realizan un pedido. En el próximo mes se van a enviar catálogos a 1000 hogares. Calcular la probabilidad de que en el próximo mes:

a) realicen un pedido al menos 400 hogares.

b) que realicen un pedido más de 50 hogares y menos de 200.

c) el número esperado de hogares que realicen un pedido.

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

La variable X = "nº de hogares que realizan un pedido" es $B(1000; 0{,}1) \cong N(100, \sqrt{90})$. Puesto que los valores significativos de X son enteros, haremos la correspondiente corrección por continuidad. Así pues:

a) $P[X \geq 400] = P[X \geq 399{,}5] = P\left[Z \geq \dfrac{299{,}5}{\sqrt{90}}\right] = P[Z \geq 31{,}57] \cong 0.$

b) $P[50 < X < 200]=P[50{,}5 \leq X \leq 199{,}5]=P\left[\dfrac{-49{,}5}{\sqrt{90}} \leq Z \leq \dfrac{99{,}5}{\sqrt{90}}\right]=P[-5{,}22 \leq Z \leq 10{,}49] \cong 1.$

c) $E[X] = np = 100$

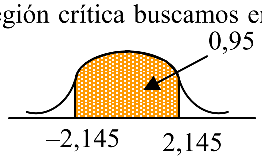
d) En virtud del teorema de Moivre, si X es B(n, p), entonces, cuando $n \to \infty$, la variable $\dfrac{X - np}{\sqrt{npq}}$ tiende hacia una distribución N(0, 1).

Se considera buena aproximación cuando $np > 5$ y $p \leq 1/2$ o bien cuando $nq > 5$ y $p > 1/2$.

2. Le han pedido que realice un estudio para evaluar si las ventas medias de su compañía en los establecimientos minoristas son mayores o son menores que las ventas medias del sector.

Se sabe que las ventas medias del sector son de 60101,2 € por establecimiento y compañía. Para llevar a cabo el estudio Ud. toma una muestra aleatoria de 15 establecimientos, obteniendo que las ventas medias de su compañía por establecimiento son de 48081,0 € con una desviación típica de 30050,6 €. Con esta información y suponiendo normalidad en la distribución y para un nivel de significación del 5 %, evalúe estadísticamente la situación de su empresa.

Estableceremos las hipótesis: H_0: μ =60101,2 € ; H_1: $\mu \neq 60101{,}2$. El estadístico $\frac{\bar{X}-\mu}{S}\sqrt{n}$ es t_{n-1}, y en nuestro caso $\begin{cases} \bar{X} = 48081 \\ \mu = 60101{,}2 \\ S = 30050{,}6 \\ n = 15 \end{cases}$. Para determinar la región crítica buscamos en las tablas x tal que $P[\,|t_{14}|>x] = 0{,}05$, obteniéndose x = 2,145.



Puesto que $\frac{48081-60101{,}2}{30050{,}6}\sqrt{15} = -1{,}5492$ se encuentra en la región de aceptación, no podemos descartar la hipótesis nula, es decir, con los datos que poseemos, no podemos afirmar que la venta media de la compañía en cuestión sea mayor o menor que 60101,2 €.

CURSO 2002/2003.	**JUNIO. Reserva.**
Código de la Carrera 42	**Código de la Asignatura 209.**

PRIMERA PARTE: CUESTIONES TEÓRICO CONCEPTUALES

1. Explique el concepto de variable aleatoria.

Sea (E, $\mathcal{A}$, P) un espacio probabilístico, donde E es el espacio muestral, P una función de probabilidad y $\mathcal{A}$ la σ-álgebra de sucesos sobre los que se ha definido la probabilidad. Sea X: E $\to \mathbb{R}$ una aplicación que a cada elemento de E (a cada suceso elemental) le hace corresponder un número real. Ahora, podemos escribir los sucesos haciendo uso de la aplicación X.

Por ejemplo, si x es un número real, escribiremos:

$[X = x] = \{e\in E \;/\; X(e) = x\}$, es decir, el suceso formado por aquellos sucesos elementales, cuya imagen es x;

$[X \le x] = \{e\in E \;/\; X(e) \le x\}$, es decir, el suceso formado por aquellos sucesos elementales, cuya imagen es $\le$ x;

$[a \le X \le b] = \{e\in E \;/\; a \le X(e) \le b\}$, es decir, el suceso formado por aquellos sucesos elementales, cuya imagen es $\ge$ a y $\le$ b; etc.

Diremos que la aplicación X es una **variable aleatoria** si $[X \le x] \in \mathcal{A}$, es decir si existe la probabilidad de los sucesos de la forma $[X \le x]$, para cualquier $x\in \mathbb{R}$.

2. Explique brevemente cual es la utilidad del teorema de Moivre.

El teorema de Moivre establece que, si X es binomial B(n, p), entonces, cuando $n \to \infty$, la variable $\frac{X-np}{\sqrt{npq}}$ tiende hacia una distribución N(0, 1).

Es útil pues, para calcular probabilidades en el caso de una variable binomial B(n, p) cuando n sea "grande". Se considera buena aproximación cuando np > 5 y p≤ 1/2 o bien cuando nq > 5 y p > 1/2.

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

Debe tenerse en cuenta que los valores significativos (con probabilidad no nula) de una variable binomial, son **enteros**, (se trata de una variable discreta) y por tanto, si aproximamos con una variable normal, que es continua, debemos hacer una corrección por continuidad, esto es, un valor entero i debe sustituirse por el intervalo $\left[i - \frac{1}{2}, i + \frac{1}{2} \right]$

3. Razone de forma intuitiva cuando un estimador es suficiente.

Un estimador es suficiente para estimar un parámetro θ cuando utiliza toda la información contenida en la muestra aleatoria, respecto a θ y ningún otro estadístico contiene más información sobre θ. Por ejemplo, la media muestral sería un estimador suficiente de la media poblacional puesto que en su cálculo intervienen todos los valores de la muestra. Si embargo la mediana no sería un estimador suficiente de la media poblacional.

4. ¿Qué se entiende por nivel de significación de un contraste?

Supongamos que se desea efectuar un contraste de un parámetro θ que toma valores en un espacio paramétrico Ω. Sea $\Omega_0 \subset \Omega$ y consideremos las hipótesis:

$$H_0: \theta \in \Omega_0$$
$$H_1: \theta \in \Omega_1 = \Omega - \Omega_0$$

Se denomina error de tipo I a la probabilidad P(rechazar H_0/H_0 es cierta).

Se denomina **nivel de significación** α o también **tamaño de la región crítica**, al máximo valor del error de tipo I cuando $\theta \in \Omega_0$:

$$\alpha = \max_{\theta \in \Omega_0} \text{ P(rechazar } H_0/H_0 \text{ es cierta).}$$

PROBLEMAS

1. En un estudio reciente se observó que el 10% de los conductores da positivo en los controles de alcoholemia efectuados durante los fines de semana. Si se selecciona una muestra aleatoria de 10 conductores, indicar:

a) ¿Cual es la probabilidad de que exactamente cinco conductores den positivo?

b) ¿Cual es la probabilidad de que den positivo menos de tres conductores?

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

a) $P[X=5] = \binom{10}{5} 0,1^5 0,9^5 \cong 0,0015$

b) $P[X < 3] = \binom{10}{0} 0,1^0 0,9^{10} + \binom{10}{1} 0,1^1 0,9^9 + \binom{10}{2} 0,1^2 0,9^8 \cong 0,93$

c) Se ha usado la fórmula de probabilidad binomial por que la variable X = "nº de conductores que dan positivo" es B(10; 0,1).

2. Una empresa de telefonía, que pretende implantarse en Argentina y con objeto de estudiar la viabilidad del proyecto, quiere obtener un intervalo de confianza del 90% para el tiempo medio de duración de las llamadas de teléfonos móviles a fijos en dicho país. De una muestra sobre 20 llamadas se obtiene que el tiempo medio de duración es de 10 minutos con una desviación típica de 3 minutos. Determinar el tamaño de la muestra necesario para obtener un intervalo de ±l minuto sabiendo que la distribución poblacional es normal.

Siendo X = "tiempo de duración de una llamada" se tiene que $\frac{\bar{X} - \mu}{S}\sqrt{n}$ es t_{n-1} y $\left[\bar{X} - t_{\alpha/2} \frac{S}{\sqrt{n}}, \bar{X} + t_{\alpha/2} \frac{S}{\sqrt{n}} \right]$ es el intervalo de confianza de nivel 100(1–α)%. La amplitud del intervalo es $L = 2t_{\alpha/2} \frac{S}{\sqrt{n}} \leftrightarrow n = 4 \frac{S^2 t^2_{\alpha/2}}{L^2}$, estimándose $t_{\alpha/2}$ y S de la muestra dada

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

(muestra piloto). En nuestro caso se tiene: $P[t_{19} > t_{0,05}] = 0,05 \rightarrow t_{0,05} = 1,729$, $S = 3$ y $L = 2$ de donde:

$$n = 4\frac{9 \cdot 1,729^2}{2^2} = 26,904969 \cong 27$$