



<b>CURSO 2005/2006.</b>	<b>JUNIO. 1ª semana</b>
<b>Código de la Carrera 42</b>	<b>Código de la Asignatura 209.</b>

### PRIMERA PARTE: CUESTIONES TEÓRICO-CONCEPTUALES

1. Explicar brevemente como afecta a la varianza de una distribución un cambio de escala.

**Respuesta.-**

La varianza queda multiplicada por el cuadrado del factor de escala. En efecto, si  $Y = kX \rightarrow \text{Var}(Y) = E(Y - E(Y))^2 = E(kX - kE(X))^2 = k^2 E(X - E(X))^2 = k^2 \text{Var}(X)$

2. Comente brevemente que utilidad tiene la desigualdad de Chebychev.

**Respuesta.-**

$$P[|X - E(X)| \geq k] \leq \frac{\sigma^2}{k^2} \leftrightarrow P[|X - E(X)| < k] > 1 - \frac{\sigma^2}{k^2}$$

Proporciona una cota superior para la probabilidad de que la variable se encuentre a más de una distancia  $k$  dada, de la media poblacional. Equivalentemente, proporciona una cota inferior para la probabilidad de que la distancia de la variable a su media sea menor que cierto número dado  $k$ . En este caso se usa para hallar intervalos de confianza para la media poblacional, cuando se conoce la desviación típica y no se conoce la distribución de la población.

3. Indicar qué significado tiene el término grados de libertad.

Una distribución  $\chi^2$  con  $n$  grados de libertad es la suma de los cuadrados de  $n$  variables aleatorias normales  $N(0, 1)$  independientes. El término grados de libertad significa pues el número de variables independientes de las que se compone  $\chi^2$ .

En relación con este concepto puede comentarse que cuando mediante muestreo se desea estimar la varianza poblacional en una población normal con media desconocida se usa

el estadístico  $\frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2}$  que se distribuye según una  $\chi^2$  (Teorema de Fisher),

pero, a pesar de que las  $X_i$  sean independientes, sin embargo  $\sum_{i=1}^n (X_i - \bar{X}) = 0$ , es decir, existe una relación de dependencia entre las variables  $(X_i - \bar{X})$  y es por eso que los grados de libertad son  $n-1$ .

4. Indicar cual es la diferencia entre hipótesis simples y compuestas a la hora de formular un contraste de hipótesis.

Una hipótesis paramétrica es simple cuando se refiere a un valor puntual del parámetro. Una hipótesis es compuesta cuando el contraste se refiere a una región del espacio paramétrico

Por ejemplo, en un contraste como el siguiente:

$$H_0: \mu = \mu_0$$

$$H_1: \mu \neq \mu_0$$

la hipótesis nula es simple mientras que la alternativa es compuesta

### PROBLEMAS

1.- En base a la información obtenida sobre el comportamiento de los consumidores de una tienda se sabe que las ventas de esa tienda tienen una distribución uniforme de entre 20 y 30 mil euros diarios. Determinar la posibilidad que tiene esa tienda de facturar más de 4.800 mil euros en 200 días.

### Solución.-

Determinaremos la probabilidad que tiene la tienda de facturar más de 4.800 miles de euros en 200 días.

Sea  $X_i$  la variable aleatoria “miles de euros vendidos el día  $i$ ”. Sabemos que  $E(X_i) = \frac{20+30}{2} = 25$  y  $\text{Var}(X_i) = \sigma_i^2 = \frac{(30-20)^2}{12} = \frac{25}{3} \rightarrow \sigma_i = \frac{5}{\sqrt{3}}$ .

Consideremos la variable  $Y = \sum_{i=1}^{200} X_i$ , para la que se cumplirá que  $E(Y) = 200 \cdot 25 = 5000$  y  $\text{Var}(Y) = \sigma^2 = 200 \cdot \frac{25}{3} = \frac{5000}{3} \rightarrow \sigma = \sqrt{\frac{5000}{3}} = \frac{100}{\sqrt{6}}$

en virtud del Teorema central del límite, la variable  $Z = \frac{Y-5000}{\frac{100}{\sqrt{6}}}$  se distribuye

aproximadamente normal  $N(0, 1)$ . Se tendrá:

$$P(Y > 4800) = P\left(Z > \frac{4800-5000}{\frac{100}{\sqrt{6}}}\right) = P(Z > -4,9) \cong 1.$$

**2.-** El gasto medio por hogar, en España, en Artículos de vestir y calzado para el año 2000 se distribuye según una Normal con desviación típica de 908,51€. De una muestra aleatoria de 100 familias se obtuvo un gasto medio de 1.484,87€.

**a)** Determinar los intervalos de confianza del 90% y del 95% para el gasto medio por hogar en Artículos de vestir y calzado.

**b)** Calcular el tamaño de la muestra necesario para obtener un intervalo de confianza del 90%, para el gasto medio por hogar en Artículos de vestir y calzado, con una amplitud de 500€.

### Solución.-

**a)** El intervalo de confianza para la media de una población normal  $N(\mu, \sigma)$ , conocida  $\sigma$ , es:

- para el 90% de confianza:  $\left[\bar{X} - \frac{1,64\sigma}{\sqrt{n}}, \bar{X} + \frac{1,64\sigma}{\sqrt{n}}\right]$ , que en nuestro caso es  $\left[1484,87 - \frac{1,64 \cdot 908,51}{\sqrt{100}}, 1484,87 + \frac{1,64 \cdot 908,51}{\sqrt{100}}\right] = [1335,87 ; 1633,87]$

- para el 95% de confianza:  $\left[\bar{X} - \frac{1,96\sigma}{\sqrt{n}}, \bar{X} + \frac{1,96\sigma}{\sqrt{n}}\right]$ , que en nuestro caso es  $\left[1484,87 - \frac{1,96 \cdot 908,51}{\sqrt{100}}, 1484,87 + \frac{1,96 \cdot 908,51}{\sqrt{100}}\right] = [1306,80 ; 1662,94]$ .

**b)** La amplitud de un intervalo de confianza del 90% es  $\frac{2 \cdot 1,64\sigma}{\sqrt{n}}$ . Para los datos del problema:  $\frac{2 \cdot 1,64 \cdot 908,51}{\sqrt{n}} = 500$ , y resolviendo la ecuación se obtiene  $n = 36$ .