



CURSO 2004/2005.	JUNIO. 2ª semana
Código de la Carrera 42	Código de la Asignatura 209.

PRIMERA PARTE: CUESTIONES TEÓRICO-CONCEPTUALES

1. ¿Qué relación existe entre la distribución t de Student y la Normal?

Respuesta.-

Si X es una variable aleatoria cuya distribución es t-Student con n grados de libertad, existen $n+1$ variables normales $N(0,1)$, $\{Y, X_1, X_2, \dots, X_n\}$, tales que $X = \frac{Y}{\sqrt{\frac{X_1^2 + X_2^2 + \dots + X_n^2}{n}}}$

2. ¿Qué relación existe entre la función de distribución empírica de una muestra y la función de distribución de la variable aleatoria de donde se ha extraído dicha muestra?

Respuesta.-

La función de distribución empírica tiene las mismas propiedades que la función de distribución de la variable aleatoria y se cumple el teorema de Glivenko-Cantelli (también llamado teorema fundamental de la estadística): dado $\varepsilon > 0$, se verifica que:

$$\lim_{n \rightarrow \infty} P \left[\sup_{-\infty < x < +\infty} |F_n(x) - F(x)| \geq \varepsilon \right] = 0$$

es decir, la función de distribución empírica $F_n(x)$ converge en probabilidad a la función de distribución $F(x)$.

3. ¿Qué es y para qué se usa el error cuadrático medio de un estimador?

Respuesta.-

Si θ es un parámetro y $\hat{\theta}$ el estimador, se define el error cuadrático medio de $\hat{\theta}$:

$$ECM(\hat{\theta}) = E[(\hat{\theta} - \theta)^2]$$

Desarrollando la expresión anterior, se obtiene que $ECM(\hat{\theta}) = Var(\hat{\theta}) + (Sesgo(\hat{\theta}))^2$. El objetivo deseable de un estimador es que el ECM y, por lo tanto, los dos sumandos en que se descompone, sean lo mínimos posible, aunque no siempre existirá tal estimador.

4. ¿Qué son los contrastes de significación?

Respuesta.-

Llamamos contrastes de significación aquellos en que la información que tenemos del parámetro θ únicamente nos permite formular la hipótesis $H_0: \theta = \theta_0$, careciendo de información para plantear otra hipótesis alternativa, por lo que esta será $H_1: \theta \neq \theta_0$.

PROBLEMAS

1.- Según los datos del INE para el 2º trimestre del año 2003, el gasto real medio por persona en artículos de vestir y calzado, se distribuye normalmente, y es de 121 euros. Tomamos una muestra de 15 personas para estimar la varianza muestral y obtenemos que esta es de 100 euros². ¿Cuál es la probabilidad de que la media muestral este comprendida entre 117 y 125 euros?

Solución.-

La variable $\frac{\bar{X} - \mu}{S} \sqrt{n}$ se distribuye como una t-Student con $n-1$ grados de libertad. En

$$\text{nuestro caso: } P(117 < \bar{X} < 125) = P\left(\frac{117-121}{10}\sqrt{15} < \frac{\bar{X}-121}{10}\sqrt{15} < \frac{125-121}{10}\sqrt{15}\right) = \\ = P(-1,5492 < t_{14} < 1,5492).$$

En las tablas de la t-Student, para 14 grados de libertad, encontramos que $P(t < 1,523) = 0,925$ y que $P(t < 1,761) = 0,950$. Efectuando una interpolación lineal se obtiene que:

$$P(t < 1,5492) = 0,925 + \frac{0,950 - 0,925}{1,761 - 1,523}(1,5492 - 1,523) \cong 0,92775$$

$$\text{y de aquí: } P(-1,5492 < t_{14} < 1,5492) = 0,92775 - (1 - 0,92775) = 0,8555$$

2.- Con los datos del ejercicio anterior y sabiendo que, en la muestra obtenida en el mismo, la media de euros en artículos de vestir y calzado que hemos obtenido es de 127 euros, se quiere contrastar si la media es realmente la que dice el INE, para un nivel de significación de 0,01.

Solución.-

Los datos para el contraste son: $H_0 = 121$; $H_1 \neq 121$. La región crítica (de dos colas), para un nivel de significación 0,01, que proporciona la tabla de la t-Student con 14 grados de libertad es $|t_{14}| > 2,997$. Como $\frac{\bar{X} - \mu}{S} \sqrt{n} = \frac{127 - 121}{10} \sqrt{15} \cong 2,3238$, no podemos rechazar la media propuesta por el INE.