Tema 11: Contrastes no paramétricos

AUTOMÓVILES **Una empresa fabricante de automóviles está llevando a cabo un estudio de mercado para analizar la posible relación entre el *tipo de vehículo adquirido* (Deportivo, Familiar y Utilitario) y el *sexo del comprador*. A partir de una muestra aleatoria simple de 100 clientes que han adquirido alguno de sus automóviles en el último mes se ha obtenido la siguiente información:**

Sexo \ Tipo de vehículo	Deportivo	Familiar	Utilitario
Hombre	8	24	30
Mujer	14	12	12

a) Desde el departamento de marketing se afirma que el tipo de vehículo adquirido es independiente del sexo del comprador. ¿Avala la información muestral esta afirmación?

b) La empresa también está realizando un estudio sobre la calidad de los vehículos que fabrica. Dentro del mismo se ha seleccionado una muestra aleatoria de 40 propietarios para los cuales se ha observado la duración de sus coches (en miles de Km.), obteniendo g_1=0.1637 y g_2=-0.8027.
b1) Se desea contrastar la normalidad de la variable a partir de la muestra. ¿Cuál sería la conclusión?
b2) Desde la empresa se afirma que la duración de los vehículos presenta una varianza no mayor de 500. ¿Qué se podría decir sobre esta afirmación, sabiendo que s^2=142?

SOLUCIÓN

a) Desde el departamento de marketing se afirma que el tipo de vehículo adquirido es independiente del sexo del comprador. ¿Avala la información muestral esta afirmación?

Para contrastar la independencia entre el tipo de vehículo comprado y el sexo del cliente aplicaremos la siguiente discrepancia:

$$d_{IND} = \sum_{i=1}^{r}\sum_{j=1}^{s}\frac{\left(n_{ij}-\frac{n_{i.}n_{.j}}{n}\right)^2}{\frac{n_{i.}n_{.j}}{n}} \approx \chi^2_{(r-1)(s-1)}$$

Las frecuencias teóricas son las siguientes:

Sexo \ Tipo de vehículo	Deportivo	Familiar	Utilitario
Hombre	**13,64**	**22,32**	**26,04**
Mujer	**8,36**	**13,68**	**15,96**

obteniéndose para la muestra observada la discrepancia siguiente:

$$d^*_{IND} = \frac{(8-13{,}64)^2}{13{,}64} + \frac{(24-22{,}32)^2}{22{,}32} + \ldots + \frac{(12-15{,}96)^2}{15{,}96} = 8{,}055$$

El nivel crítico asociado es p= 0,0178, por lo que existe evidencia para rechazar la hipótesis nula y se podría afirmar que el modelo de vehículo adquirido no es independiente del sexo del comprador.

b) La empresa también está realizando un estudio sobre la calidad de los vehículos que fabrica. Dentro del mismo se ha seleccionado una muestra aleatoria de 40 propietarios para los cuales se ha observado la duración de sus coches (en miles de Km.), obteniendo g_1=0.1637 y g_2=-0.8027.
b1) Se desea contrastar la normalidad de la variable a partir de la muestra. ¿Cuál sería la conclusión?

El contraste de normalidad de Jarque-Bera se basa en la expresión de la discrepancia definida por:

$$d_{JB} = \frac{n}{6}\left(g_1^2 + \frac{1}{4}g_2^2\right)$$

expresión que bajo la hipótesis nula de normalidad sigue una distribución chi-cuadrado con 2 grados de libertad. Para la muestra observada la discrepancia es d* = 1,2527, siendo el nivel crítico asociado:

$$p = P(d_{JB} > d^*_{JB}/H_0) = P(\chi^2_2 > 1{,}2527) = 0{,}5345$$

por lo que no rechazaremos la hipótesis de normalidad para la duración de los vehículos.

b2) Desde la empresa se afirma que la duración de los vehículos presenta una varianza no mayor de 500. ¿Qué se podría decir sobre esta afirmación, sabiendo que s^2=142?
Se trata de resolver el siguiente contraste:

$$H_0 : \sigma^2 \leq 500$$
$$H_1 : \sigma^2 > 500$$

siendo la discrepancia:

$$d_{S^2} = \frac{(n-1)S^2}{\sigma^2} \approx \chi^2_{n-1}$$

La discrepancia observada en la muestra, donde n = 40 y s^2=141,9 es:

$$d^*_{S2} / H_0 = \frac{39(141{,}9)}{500} = 11{,}068$$

siendo el nivel crítico asociado:

$$p = P(d_{S2} > d^*_{S2} / H_0) = P\left(\chi^2_{39} > 11{,}068\right) = 0{,}999$$

por lo que no existe evidencia muestral para rechazar la hipótesis nula.