

Examen 3

- 1) Se quieren probar dos tipos de alimentos para los 75 pingüinos de un zoológico cuyo peso se distribuye normalmente. Se separan en dos grupos, uno formado por 40 pingüinos y otro por 35. Al cabo de un mes son pesados y se obtiene para primer grupo un peso medio de 13 kg y desviación típica de 0.7, y para el segundo grupo un peso medio de 11 kg y desviación típica de 0.3. ¿Se puede afirmar que están mejor alimentados los del primer grupo que los del segundo? Considera un nivel de confianza del 99%.

Soluciones:

Contraste de hipótesis para comparar 2 medias (independientes), unilateral a la izquierda. Se conoce la distribución: Normal. Se desconocen las varianzas poblacionales pero tenemos muestras grandes.

$$H_0: \mu_1 \leq \mu_2$$

$$H_1: \mu_1 > \mu_2$$

Datos muestrales:

Grupo 1

$$n_1 = 40$$

$$\bar{x}_1 = 13$$

$$s_1 = 0.7$$

Grupo 2

$$n_2 = 35$$

$$\bar{x}_2 = 11$$

$$s_2 = 0.3$$

Estadístico de contraste:

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{13 - 11}{\sqrt{\frac{0.7^2}{40} + \frac{0.3^2}{35}}} = \frac{2}{\sqrt{0.01225 + 0.00257143}} = \frac{2}{0.0097} = 206$$

Valor crítico:

$$Z_{\alpha=0.01} = 2.33$$

Entonces, como es un contraste unilateral a la derecha y $\hat{Z} = 206 > Z_{\alpha=0.01} = 2.33$, el valor del estadístico cae dentro de la Región de Rechazo, por lo cual rechazamos H_0 , es decir, que sí se puede afirmar que están mejor alimentados los del primer grupo que los del segundo.

- 2) Se consideran los siguientes modelos de regresión lineal. Los números entre paréntesis debajo de la pendiente del modelo indican el p-valor del contraste de hipótesis de significación de ese coeficiente:

Modelo 1 $Y=2+3X$ (p-valor=0.14) $R^2 = 90.5$	Modelo 2 $Y=0.5-21X$ (p-valor=0.023) $R^2 = 87.9$	Modelo 3 $Y=0.5-9X$ (p-valor=0.01) $R^2 = 93$
--	--	--

- ¿A qué es igual la pendiente en el modelo 1?
- ¿Cómo interpretarías el coeficiente de la pendiente en el modelo 1?
- Analiza la significancia del coeficiente de la pendiente en los tres modelos, fijándote en el p-valor de cada uno.
- Según la significancia obtenida en el inciso anterior, ¿qué modelo(s) elegirías como posible(s) candidato(s) y cuál o cuáles desecharías?
- Según lo obtenido hasta el momento, ¿con qué modelo te quedarías mirando el valor del coeficiente de determinación? Explica tu respuesta.

Soluciones:

- En el modelo 1, la pendiente es $\beta_1 = 3$.
- Cuando la X aumenta en 1 unidad, la Y aumenta en 3 unidades.
- El contraste de significación de la pendiente es el siguiente:

$$H_0: \beta_1 = 0$$

$$H_1: \beta_1 \neq 0$$

Para el modelo 1, el p-valor es mayor que el nivel de significación $\alpha = 0.05 = 5\%$, por lo cual no se rechaza la hipótesis nula de que la pendiente es cero, con lo cual no sería significativo y no podemos usar ese modelo. Para los modelos 2 y 3, el $p - valor < \alpha$, con lo cual se rechaza la hipótesis nula y se acepta la alternativa de que el β_1 es significativamente distinto de cero. Esos modelos sí los podríamos usar.

- Desecharíamos el modelo 1, y nos quedaríamos con el 2 y con el 3.
- Entre los dos modelos que no hemos desechado, es decir, los modelos 2 y 3, mirando el coeficiente de determinación R^2 , el modelo que tiene mayor valor es el modelo 3 con $R^2 = 93\%$, que significa que ese modelo explica un 93% de la variabilidad, y es mejor que el otro, que explica un menor porcentaje.