Análisis de datos
Respuestas a los ejercicios del segundo control

a.

Se utiliza un modelo de regresión lineal simple. Variable predictora (X), «Longitud de la víbora». Variable explicada (Y), «Número de ondulaciones».

$$Y = \beta_0 + \beta_1 X + U; \qquad U \sim \mathrm{N}(0, \sigma)$$

Requisitos del modelo: linealidad, normalidad, homocedasticidad, independencia.
En el diagrama de dispersión se observa que la relación si parece aproximadamente lineal.
En el histograma es difícil decidir la normalidad debido al pequeño número de datos.

b.

Intervalo de confianza del 95 % para β_1.
En la tabla de «Coeficientes» se observa que $\hat{\beta}_1 = 0'933$ y que el error típico en esta estimación es $0'122$. En las tablas obtenemos $t_{10;0'025} = 2'228$ por tanto:

$$\mathrm{IC}_{95\,\%}(\beta_1) = (0'933 \pm 2'228 \cdot 0'122) = (0'661, 1'205)$$

c.

Hipótesis del contraste: nula: $\mathrm{H}_0 \equiv \beta_1 = 0\,;$ alternativa: $\mathrm{H}_1 \equiv \beta_1 \neq 0$
Decisión: El p-valor del contraste de la regresión (tabla ANOVA, se ve $'000$) es menor que $0'0005$, por tanto al nivel de significación $\alpha = 0'05$ se rechaza la hipótesis nula H_0.
Conclusión: la longitud de la víbora tiene una influencia significativa sobre el número de ondulaciones.

d.

Estimación puntual $\overline{Y}_0$ del número medio de oscilaciones de las víboras cuya longitud es de 58 cm:

$$\overline{Y}_0 = -15'801 + 0'933 \cdot 58 = 38'313$$

En las tablas proporcionadas se lee:

$$S_R^2 = 2'889\,; \qquad nv_x = (n-1)s_x^2 = 11 \cdot (4'1845)^2 = 192'61\,; \qquad \overline{X} = 55'858$$

Por tanto:

$$\mathrm{IC}_{95\,\%}(\overline{Y}\,|_{X=58}) = \left(38'313 \pm 2'228 \cdot \sqrt{2'889\left(\frac{1}{12} + \frac{(58-55'858)^2}{192'61}\right)}\right) = (38'313 \pm 2'228 \cdot 0'5564)$$

$$\mathrm{IC}_{95\,\%}(\overline{Y}\,|_{X=58}) = (37'073, 39'553)$$