

**(1994-
2024)**

30 años de la
Consagración Constitucional
de la Autonomía y Autarquía
Universitaria en Argentina.



UNIVERSIDAD NACIONAL DEL LITORAL

FACULTAD DE INGENIERÍA Y CIENCIAS HÍDRICAS

ESTADÍSTICA (II)

PROBABILIDAD Y ESTADÍSTICA (IIA)

RESPUESTAS DE GUÍA DE PRÁCTICA

UNIDAD 7 - REGRESIÓN Y CORRELACIÓN

Responsable de cátedra: Prof. Juan Pablo Taulamet

Equipo de cátedra: *Auxiliares:* Lic. María José Llop (JTP) - Ing. Franco Nardi (Ay. 1°)

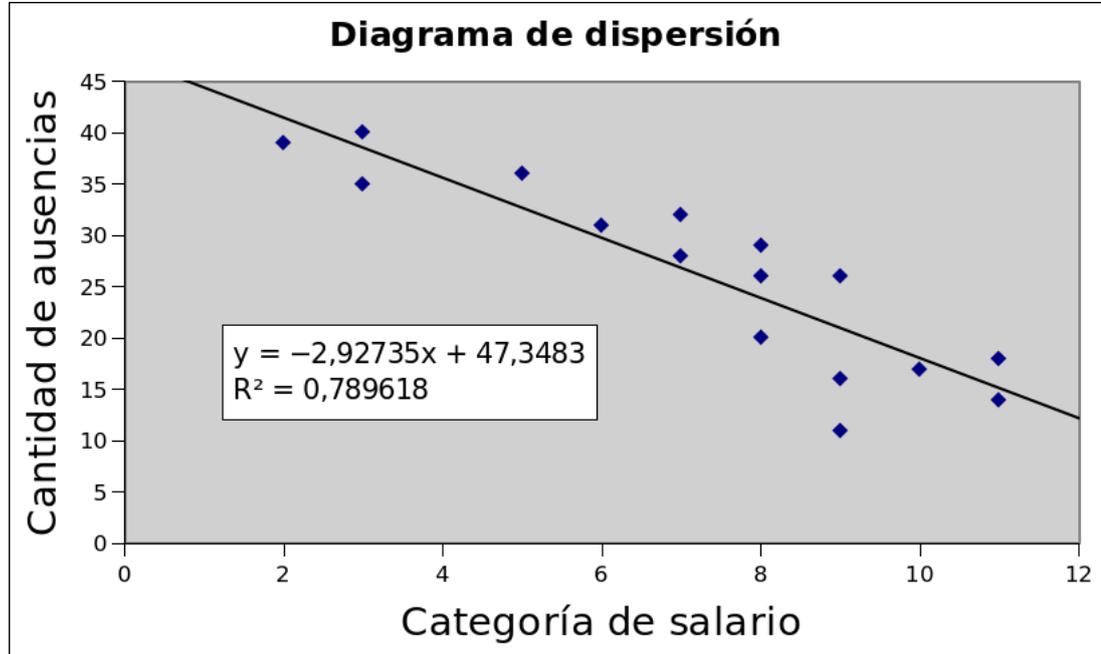
Ayudantes: AIA Cristian Bottazzi - Téc. Eliana García

Carreras: Ingeniería en Informática - Ingeniería en Inteligencia Artificial

AÑO ACADÉMICO 2024

Ejercicio 1

a)



Existe una relación lineal con tendencia decreciente.

b) Llamamos $X =$ categoría, $Y =$ ausencias.

Modelo planteado para las observaciones: $Y_i = \alpha + \beta X_i + \varepsilon_i$ con ε_i independiente de X_i , $E(\varepsilon_i) = 0$ y $V(\varepsilon_i) = \sigma^2$, $i = 1, \dots, n$. Este modelo implica $E(Y_i | X_i = x_i) = \alpha + \beta x_i$, $i = 1, \dots, n$.

Modelo estimado: $\hat{Y} = 47.348 - 2.9274x$

c) Coeficiente de correlación: $r \approx -0.89$ lo que implica una fuerte relación lineal inversamente proporcional. Coeficiente de determinación: $r^2 = 0.789$ indica que casi el 80% de la variabilidad de Y es explicada por el modelo propuesto.

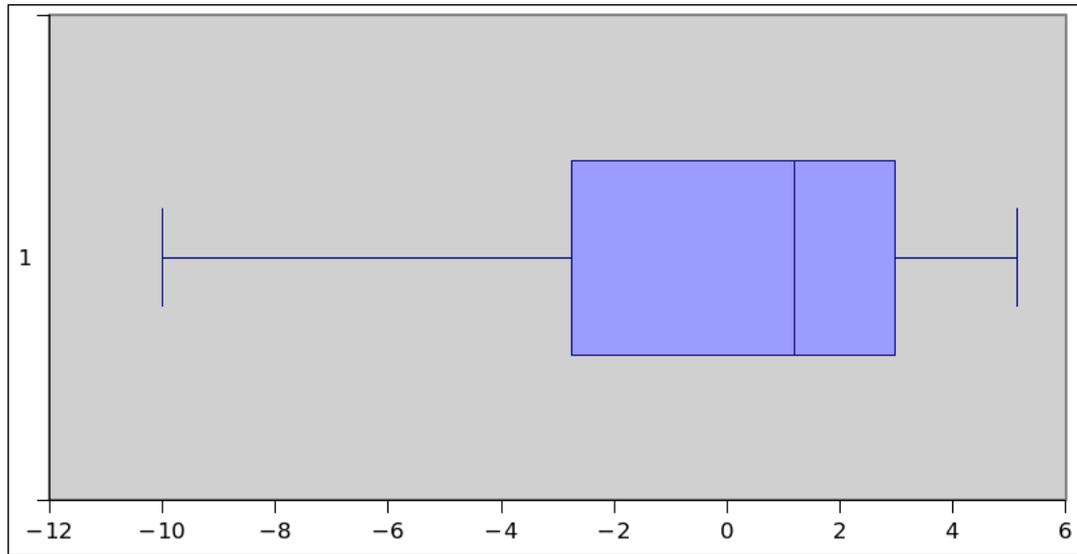
d) $\hat{Y}_{x=10} = 47.348 - 2.927 * 10 = 18.074$. Este valor sirve tanto para estimar $E(Y|x = 10)$ como para pronosticar Y cuando $x = 10$. En este último caso, el error estándar asociado **estimado** es

$$\hat{\sigma} \sqrt{1 + \frac{1}{n} + \frac{(\bar{x} - 10)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$$

donde $\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n - 2}}$. En este caso $\hat{\sigma} = 4.368$. Luego el error estándar estimado asociado al valor pronosticado resulta

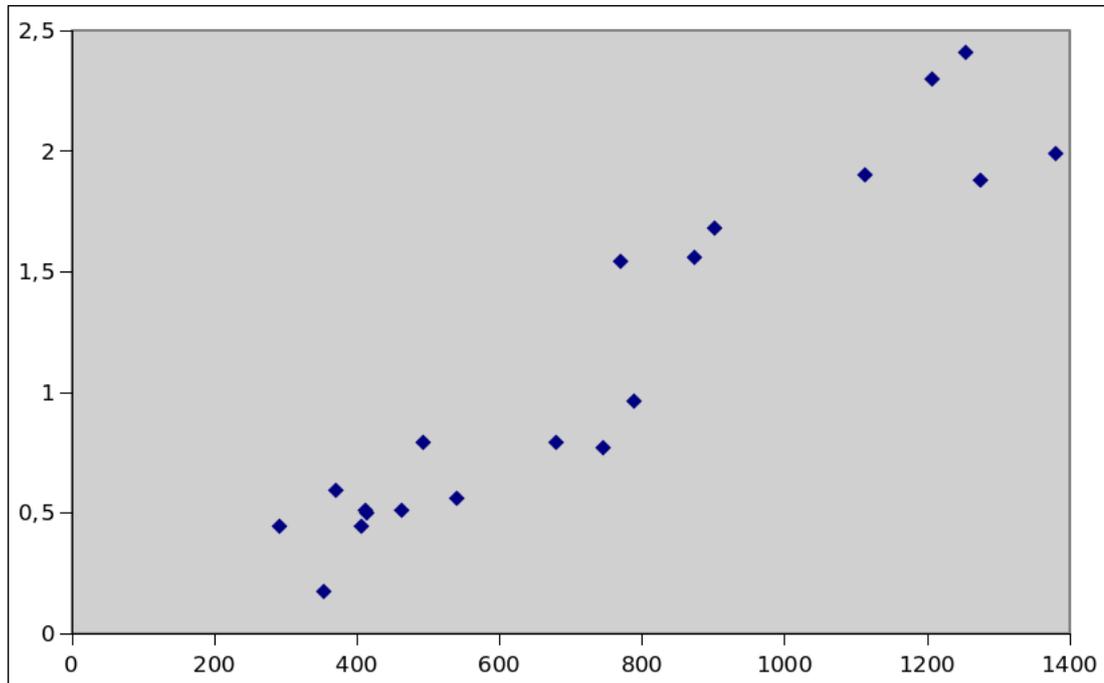
$$4.368 \sqrt{1 + \frac{1}{16} + \frac{(10 - 7.25)^2}{117}} = 4.637.$$

e) Para conversar en clase. Presentamos un gráfico de box plot de los residuos.



Ejercicio 2

a)



b) El modelo estimado resulta $\hat{Y} = -0,2906 + 0,0019 x$. El coeficiente de determinación $r^2 = 0,895$; es decir que aproximadamente el 89,5 % de la variabilidad observada en la demanda en la hora pico (Y) es explicada por la relación propuesta con el uso total de energía mensual (X).

c) La demanda pronosticada para $X = 980$ kWh es de 1,571 kW.

d) La varianza del pronóstico $\sigma^2(Y_i - \hat{Y}_h) = 0,060$.

Ejercicio 3

- a) El modelo estimado resulta aproximado a $\hat{Y} = -0,28x + 12,05$. El coeficiente de determinación $r^2 = 0,903$; es decir que aproximadamente el 90,3 % de la variabilidad observada en la cantidad de clientes (X) es explicada por la relación propuesta con la distancia (Y).
- b) El coeficiente de correlación es -0,95 lo que da cuenta de una fuerte relación lineal inversamente proporcional.
- c) Si la distancia es de 2 km entonces esperamos cerca de 1149 clientes.
- d) Para esperar 500 clientes aproximadamente debería situarse a unos 24,8 km.