



Prueba Diagnóstica Maestría en Estadística Aplicada

Repositorio con el proyecto y los cálculos realizados: https://github.com/Ludwing-MJ/Prueba_Diagnostica
Tema 1 (40 puntos)

La siguiente distribución de frecuencias representa el tiempo (en segundos) que los cajeros de un almacén necesitaron para servir a una muestra de clientes.

Tiempo (en segundos)	fi
20	29
30	39
40	49
50	59
60	69
70	79
80	89
90	99
100	109
110	119
120	129

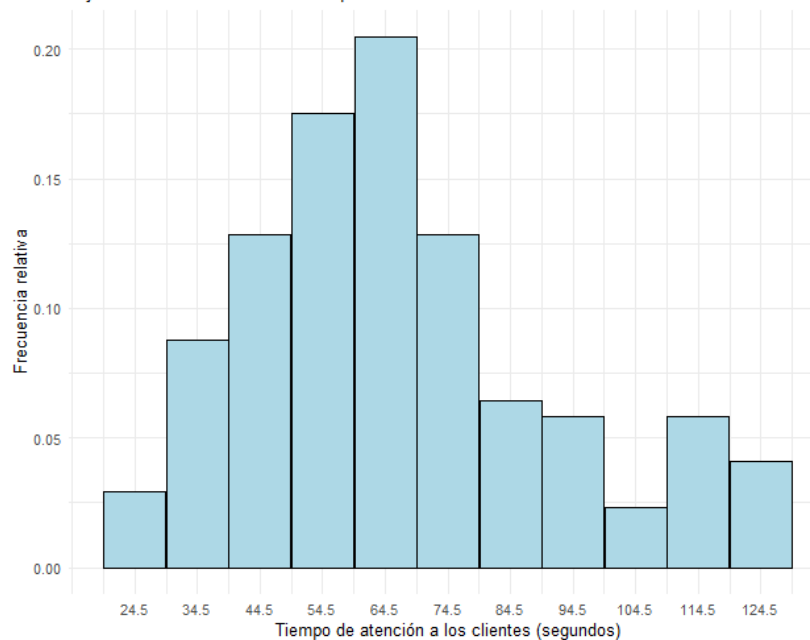
- a. ¿Cuál es el intervalo de clase?

R// Intervalo de clase: 9 segundos

- b. Elabore un histograma.

Histograma del Tiempo de Atención al cliente

Distribución de frecuencias del tiempo (en segundos) que los cajeros de un almacén necesitaron para servir a una muestra de clientes





ESCUELA DE ESTUDIOS DE
POSTGRADO
FACULTAD DE INGENIERÍA

c. ¿Cuál es el tiempo promedio que los cajeros necesitan?

R// En promedio los cajeros necesitan 67.01462 segundos para servir una muestra.

d. ¿Cuál es el coeficiente de variación relativa de Pearson?

R// El coeficiente variación relativa de Pearson es de: 36.9824 %

e. ¿Cuál es el valor del coeficiente de asimetría $S_{k3} = \frac{(Q3-Me)-(Me-Q1)}{(Q3-Q1)}$

R// El valor del coeficiente de asimetría corresponde a: 0.0693871.

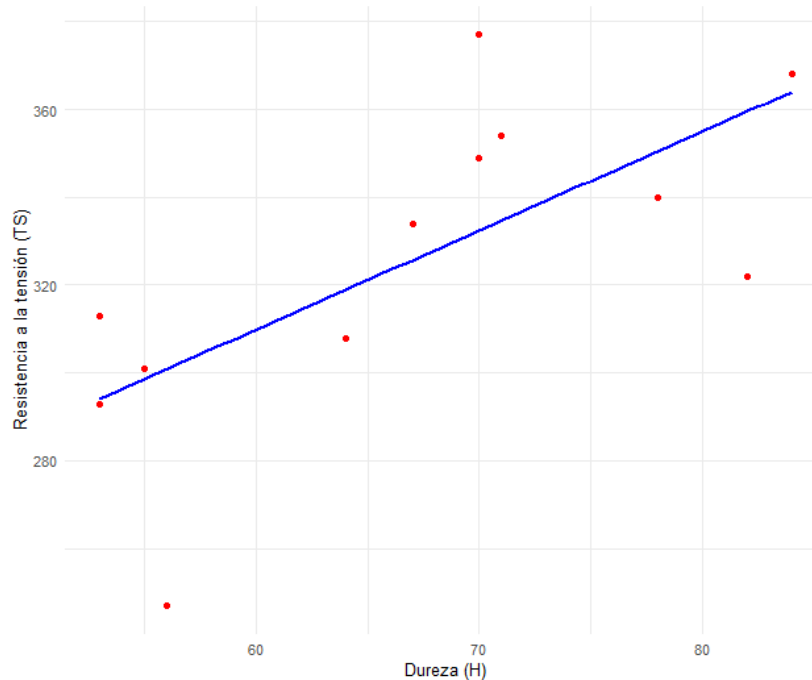
Tema 2 (30 puntos)

Los siguientes datos son para dureza (H) y resistencia a la tensión (HS) del aluminio vaciado en troqueles.

TS (y)	293	349	368	301	340	308	354	313	322	334	377	247
H (x)	53	70	84	55	78	64	71	53	82	67	70	56

a) Elabore un diagrama de dispersión.

Diagrama de Dispersión: Dureza (H) vs Resistencia a la tensión (TS)



b) Encontrar la ecuación de la línea regresión para estimar la resistencia a la tensión, partiendo de la dureza.

R// Ecuación:

TS = Resistencia a la tensión

H = Dureza

$$TS = 174.6900 + 2.25 * H$$



> summary(regresion_ej2)

Call:

```
lm(formula = TS_y ~ H_x, data = ej2)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-53.897	-10.591	3.178	17.129	44.551

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	174.6900	51.1773	3.413	0.00662 **
H_x	2.2537	0.7555	2.983	0.01373 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 27.6 on 10 degrees of freedom

Multiple R-squared: 0.4709, Adjusted R-squared: 0.418

F-statistic: 8.899 on 1 and 10 DF, p-value: 0.01373

c) Calcular el coeficiente de correlación.

R// El coeficiente de correlación lineal simple para la resistencia a la tensión del aluminio partiendo de la dureza es de: 0.6862096.

```
> cor.test(ej2$H_x, ej2$TS_y)
```

Pearson's product-moment correlation

data: ej2\$H_x and ej2\$TS_y

t = 2.9832, df = 10, p-value = 0.01373

alternative hypothesis: true correlation is not equal to 0

95 percent confidence interval:

0.1852706 0.9040723

sample estimates:

cor

0.6862096

d) Estime el valor de la resistencia a la tensión cuando la dureza es 60.

R// Cuando la dureza es 60 la resistencia a la tensión sería de: 309.69

Nota final del ejercicio: El ajuste del modelo de regresión lineal simple es bajo al tener un coeficiente de correlación de 0.69 y un coeficiente de determinación ajustado de 0.418, considerando estos valores para mejorar el ajuste del modelo se puede analizar la transformación de variables para obtener un modelo linealizado o optar por un modelo polinómico, al no haber instrucción/ planteamiento relacionado al ajuste buscado en el modelo se asumió un modelo de regresión lineal simple sin usar transformaciones.



Tema 3 (20 puntos)

La producción por hora de los trabajadores en una fábrica se considera distribuida normalmente con media de 240 unidades y desviación típica de 20 unidades. Considérese que en esta fábrica trabajan en la producción 10 mil trabajadores.

- a. ¿cuántos trabajadores tienen una producción de más de 250 unidades por hora?

R// 3,085 trabajadores tienen una producción de más de 250 unidades por hora

```
> # TEMA 3: Distribución normal
> ej3_media <- 240
> ej3_sd <- 20
> ej3_n <- 10000
>
> # a. ¿cuántos trabajadores tienen una producción de más de 250 unidades por hora?
> p_ej3_a <- pnorm(250, ej3_media, ej3_sd, F) ; p_ej3_a
[1] 0.3085375
>
> t_250 <- p_ej3_a*ej3_n ; t_250
[1] 3085.375
```

- b. Si cualquier trabajador que produzca menos de 200 unidades por hora debe recibir entrenamiento posterior, ¿cuántos recibirán entrenamiento?

R// 228 trabajadores tienen que recibir entrenamiento posterior ya que producen menos de 200 unidades por hora.

```
> # b. Si cualquier trabajador que produzca menos de 200 unidades por hora debe
> # recibir entrenamiento posterior, ¿cuántos recibirán entrenamiento?
> p_ej3_b <- pnorm(200, ej3_media, ej3_sd, T) ; p_ej3_b
[1] 0.02275013
> t_200 <- p_ej3_b * ej3_n ; t_200
[1] 227.5013
~
```