

# LABS03:ANÁLISIS DE REGRESIÓN. ECUACIÓN DE REGRESIÓN

Aranda Huerta, Milene || Escriba Flores, Daniel

2025-04-02

## Contents

<b>A. Preparacion de Datos - Diagrama de dispersion</b>	<b>1</b>
<b>B.Regresion Lineal</b>	<b>3</b>
<b>C. Pronostico</b>	<b>4</b>
Interpretacion . . . . .	5

Contexto:

Cargo, empresa que realiza entrega de pedidos a todo el Perú, desea llevar a cabo un estudio para ver la relación entre la distancia de recorrido de un embarque y el tiempo que demora en llegar.

Embarque	Distancia (millas)	Tiempo de envío (días)	Embarque	Distancia (millas)	Tiempo de envío (días)
1	656	5	11	862	7
2	853	14	12	679	5
3	646	6	13	835	13
4	783	11	14	607	3
5	610	8	15	665	8
6	841	10	16	647	7
7	785	9	17	685	10
8	639	9	18	720	8
9	762	10	19	652	6
10	762	9	20	828	10

Considerando a la distancia de envío como la variable independiente y al tiempo de envío como la variable dependiente, haga lo siguiente:

## A. Preparacion de Datos - Diagrama de dispersion

Traslade la información de la tabla mostrada a un documento de texto, asígnele el nombre ‘entregas.txt’ y elabore un diagrama de dispersión para las variables numéricas.

## Lectura de datos

```
#Leemos el archivo redactado

data = read.table("entrega.txt" , sep=" ", header=T)
data = data[-1]
data

##      distancia.millas T.envio.dias
## 1           656           5
## 2           853          14
## 3           646           6
## 4           783          11
## 5           610           8
## 6           841          10
## 7           785           9
## 8           639           9
## 9           762          10
## 10          762           9
## 11          862           7
## 12          679           5
## 13          835          13
## 14          607           3
## 15          665           8
## 16          647           7
## 17          685          10
## 18          720           8
## 19          652           6
## 20          828          10

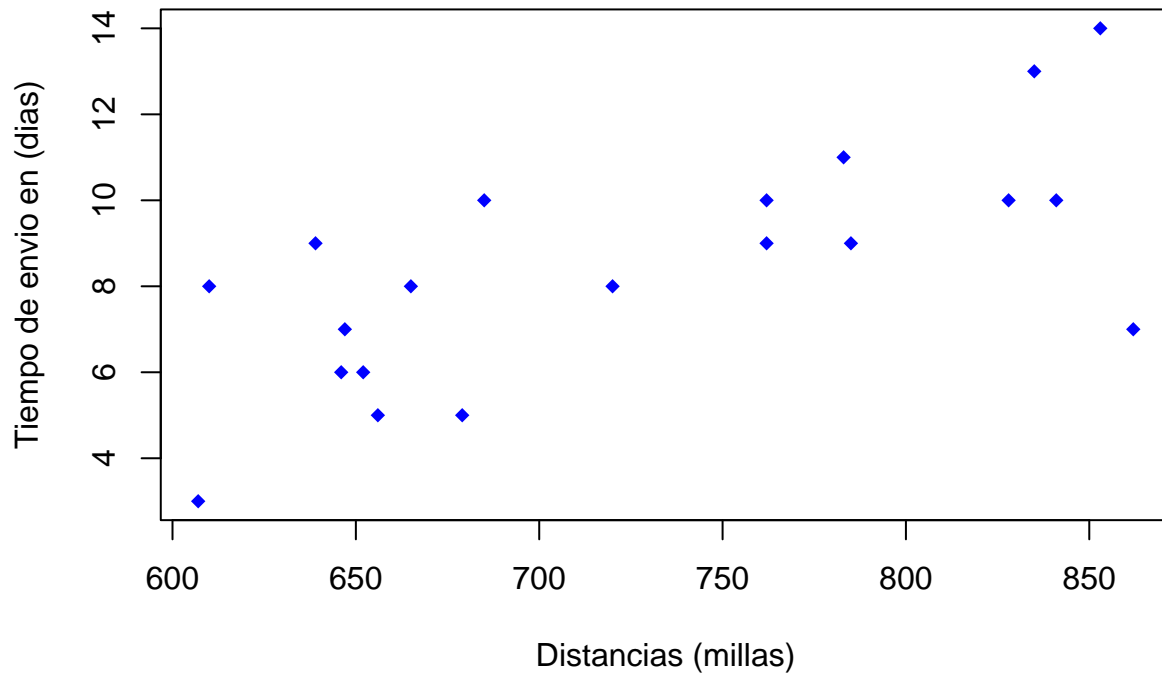
##      distancia.millas T.envio.dias
## Min.      :607.0      Min.      : 3.00
## 1st Qu.:650.8      1st Qu.: 6.75
## Median :702.5      Median : 8.50
## Mean   :725.9      Mean   : 8.40
## 3rd Qu.:795.8      3rd Qu.:10.00
## Max.   :862.0      Max.   :14.00
```

## Diagrama de dispersion

```
x = data$distancia.millas
y = data$T.envio.dias

plot(x=x, y=y,
     pch=18, col="blue",
     xlab="Distancias (millas)",
     ylab="Tiempo de envio en (dias)",
     main="Diagrama de dispersion")
```

## Diagrama de dispersion



## B.Regresion Lineal

Realice un análisis de regresión, escriba la ecuación y grafique la recta de regresión.

Creamos el modelo

```
modelo = lm(y ~ x)
```

```
#summary(modelo)
```

```
modelo$coefficients
```

```
## (Intercept)          x  
## -7.12639554  0.02139064
```

### Escribiendo la ecuación

Con los resultados obtenidos podemos escribir la ecuación obtenida.

La ecuación de regresión lineal será:

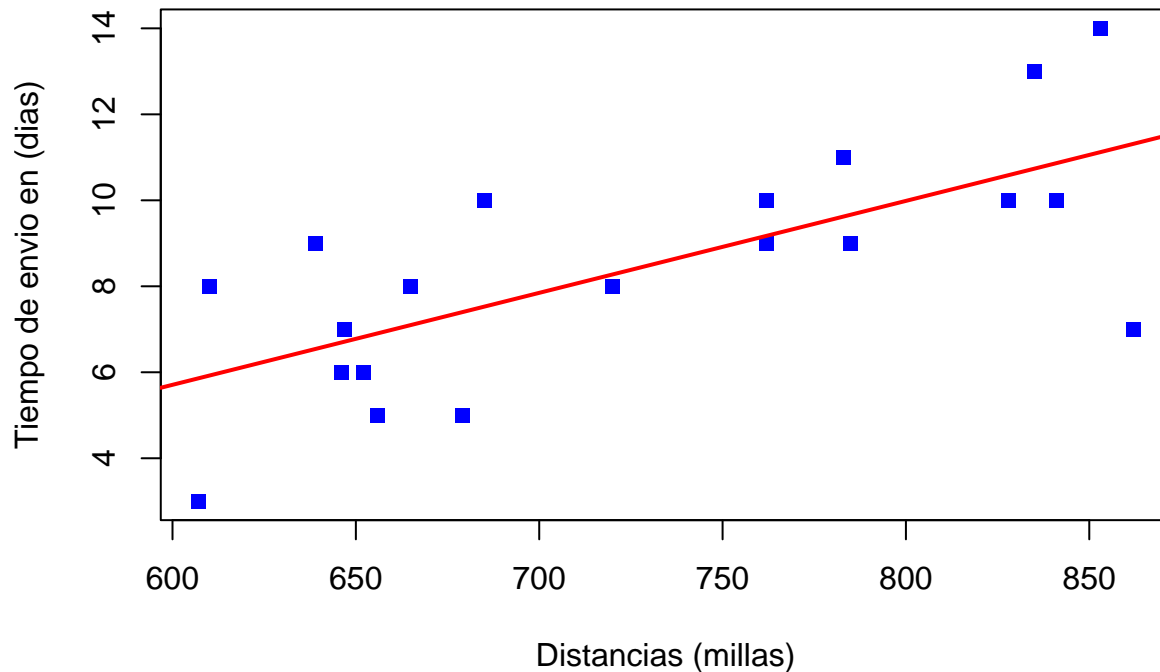
$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

$$y = -7.12639554 + 0.02139064x$$

### Graficamos la recta de regresion

```
plot(y ~ x ,  
     pch=15, col="blue",  
     xlab="Distancias (millas)", ylab="Tiempo de envio en (dias)", main="Diagrama de dispersion")  
  
abline(modelo, col='red', lwd=2)
```

## Diagrama de dispersion



## C. Pronostico

Realice un pronóstico para distancias que van de 900 a 1000 millas en intervalos de 10.

```
distancias_futuras = data.frame(x=seq(900, 1000, 10))

# Hacer las predicciones
dias_est = predict(object = modelo, newdata = distancias_futuras)

# Combinar distancias y predicciones en un solo data frame
resultados = data.frame(Distancia = distancias_futuras$x,
                        Dias_estimados = dias_est)

# Mostrar el data frame

# utilizamos la siguiente libreria, para una mejor visualizacion
library(knitr)
kable(resultados, digits = 5)
```

Distancia	Dias_estimados
900	12.12518
910	12.33909
920	12.55299
930	12.76690
940	12.98081
950	13.19471
960	13.40862

Distancia	Dias_estimados
970	13.62252
980	13.83643
990	14.05034
1000	14.26424

## Interpretacion

- La ecuación que obtuvimos nos indica que por cada milla adicional recorrida, el valor de la variable dependiente aumenta en 0.02139064 unidades. Partiendo de un valor base negativo de -7.12639554, esto permite predecir valores que oscilan entre 12.12 y 14.26 para distancias de 900 a 1000 millas.