UNIVERSIDAD CATOLICA BOLIVIANA “SAN PABLO”

MAESTRIA EN CIENCIA DE DATOS, TERCERA VERSION



Materia: ANALISIS ESTADÍSTICO I

**Examen**

**Maestrante: Ramón Wilder Serdán Cárdenas**

Noviembre 2021

La Paz – Bolivia

**PREGUNTAS**

1. Con la base de datos: datos\_ p\_18.xlsx, se pide:

Realizar todos los gráficos estadísticos, e interprete. (se evalúa interpretación)

Supuesto: La descripción de los campos es la siguiente:

* **Precio** es el precio de autos en pesos.
* **Ingreso** es el ingreso expresado en pesos.
* **Autos** es el número de autos en una escala (miles).
* **Ventas** número de autos vendidos en una escala (miles).

**Histogramas de las variables cuantitativas**

Los precios de los autos se encuentran en 2 rangos, el primero entre 110 pesos y 150 pesos y el segundo entre 180 pesos y 200 pesos. El ingreso tiene relación con los precios puesto que se encuentran también en dos rangos diferenciados. Las variables “Autos” y “Ventas” muestran una distribución aparentemente normal.

|  |  |
| --- | --- |
| **Precio**  C:\Users\wilde\AppData\Local\Microsoft\Windows\INetCache\Content.Word\Rplot.png | **Ingreso**  C:\Users\wilde\AppData\Local\Microsoft\Windows\INetCache\Content.Word\Rplot01.png |
| **Autos**  **C:\Users\wilde\AppData\Local\Microsoft\Windows\INetCache\Content.Word\Rplot02.png** | **Ventas**  **C:\Users\wilde\AppData\Local\Microsoft\Windows\INetCache\Content.Word\Rplot03.png** |

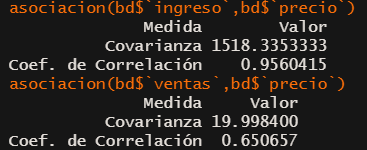
**Boxplot**

Ninguna de las variables presentan valores atípicos salvo las Ventas en miles de pesos que presenta un valor inferior que se podría tomar como atípico.

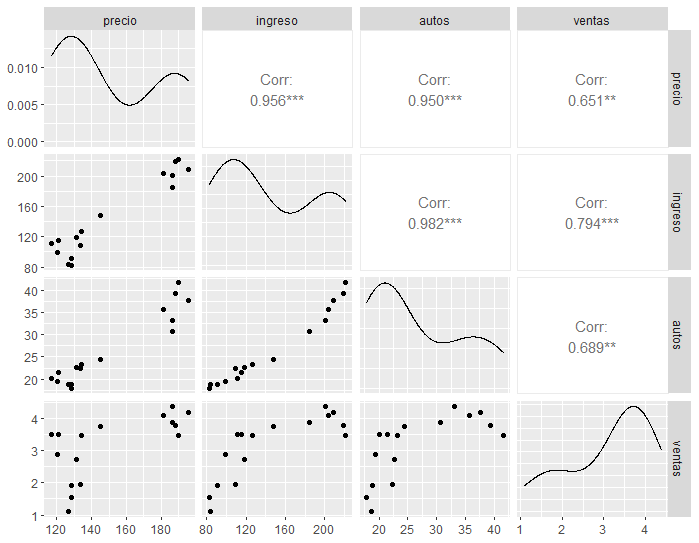
|  |  |
| --- | --- |
| **Precio**  **C:\Users\wilde\AppData\Local\Microsoft\Windows\INetCache\Content.Word\Rplot04.png** | **Ingreso**  **C:\Users\wilde\AppData\Local\Microsoft\Windows\INetCache\Content.Word\Rplot05.png** |
| **Autos**  **C:\Users\wilde\AppData\Local\Microsoft\Windows\INetCache\Content.Word\Rplot06.png** | **Ventas**  **C:\Users\wilde\AppData\Local\Microsoft\Windows\INetCache\Content.Word\Rplot07.png** |

**Covarianza y el coeficiente de correlación**

Como se aprecia en el siguiente cuadro, las variables “Ingreso” y “Precio” muestran un grado de asociación alto.



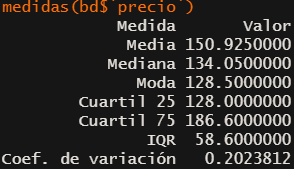
Asimismo, se corroboraron los resultados con el siguiente gráfico utilizando la librería GGally en el que también se muestran los gráficos de dispersión y correlaciones entre todas las variables de la base de datos.



2.- Determinar los estadísticos de las 4 variables e interpretar

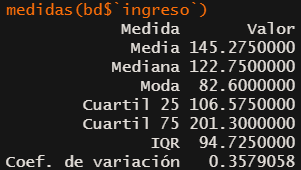
Variable Precio

El precio expresado en pesos no se encuentra muy disperso en virtud al coeficiente de variación.

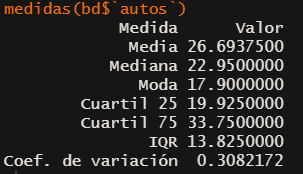
****

Ingreso

El ingreso, que se encuentra en función del precio, tampoco muestra mucha dispersión.

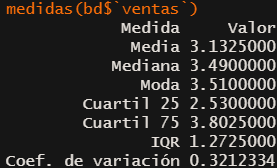
****

Autos

****

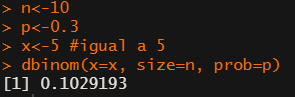
Ventas

La empresa presenta un promedio de ventas de 3,13 miles de pesos con datos ligeramente dispersos.

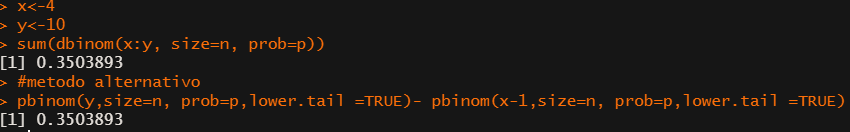
****

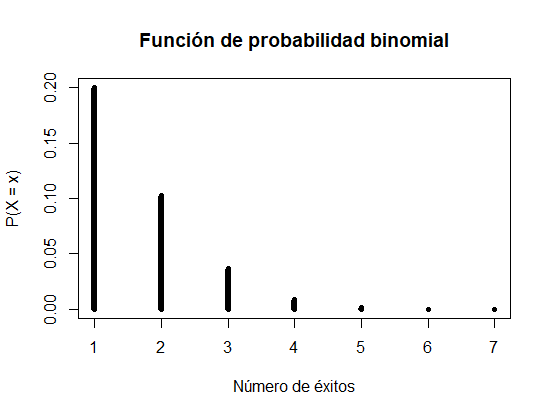
3. Calcular la probabilidad:

1. Cuando X~binomial(10;0.3)
2. P(X=5)

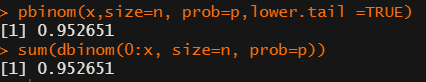


1. P(4<=X<=10)





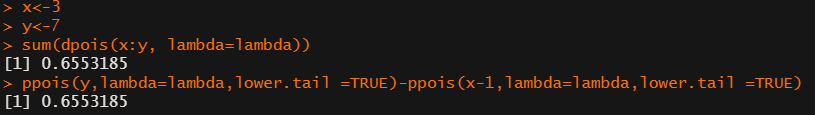
1. P(X<=5)



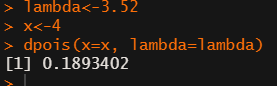
1. Cuando X~Poisson(lambda=3.52)
2. P(X=3)



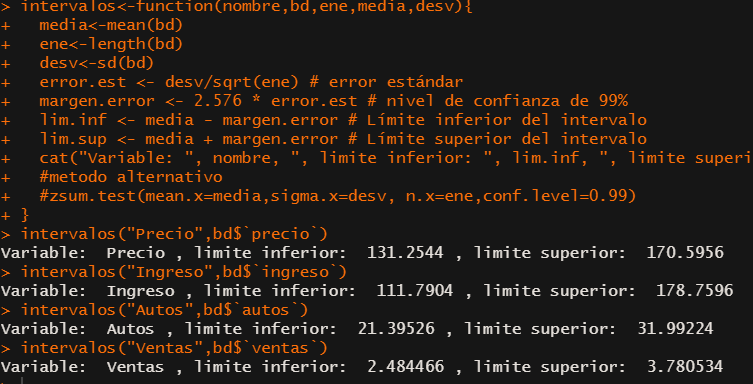
1. P(3<=X<=7)



1. P(X=4)



4. Utilizando los datos de la pregunta 1, se pide determinar los intervalos de confianza para todas las variables con 99% de confianza(interpretar)



Las medias poblacionales en todos los casos se encuentran dentro del intervalo de confianza, lo que permite deducir que la media poblacional se encuentra dentro de los intervalos descritos en el cuadro anterior con un nivel de confianza de 99%.

Asimismo, calculados los p-value de todas las variables nos permiten aceptar la Hipótesis nula de que las medias poblacionales sean iguales a las medias muestrales.

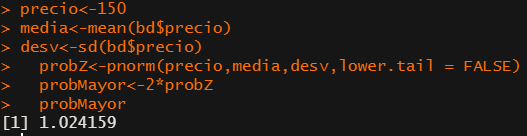
5. Utilizando la base de datos datos\_p\_18.xlsx, probar:

1. H0: mu\_precio =150, con significancia=0.06

Se plantean la hipótesis:

Ho: mu\_precio =150

H1: mu\_precio <>150

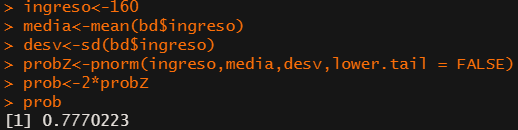


Dado el nivel de significancia, el p-value alto permite **no rechazar** la Hipotesis nula de que la media poblacional sea 150 pesos.

1. H0: mu\_ingreso =160, con significancia=0.03

Ho: mu\_precio =160

H1: mu\_precio <>160



Ya que el nivel de significancia es 0,03, el p-value superior a dicho valor permite que no rechacemos la Hipotesis nula de que la media poblacional del ingreso sea igual a 160 pesos.

6. Aplicar el método bisección o Newton Raphson para la función:

X^2-Cos(x)-1=0

Considerar X\* en [1,2], X0=1.5

