# SOLUCIÓN PROYECTO DE AULA INGENIERIA DE SISTEMAS

**CAROLINA TOBÓN PINO**

**JUAN PABLO MURILLO**

**SEBASTIAN MUÑOZ ZAPATA**

**SANTIAGO YEPES ZULETA**

**ESTADISTICA PARA INGENIERIA**

**UNIVERSIDAD DE MEDELLIN**

**MEDELLIN, ANTIOQUIA**

**2022**

**INTRODUCCIÓN**

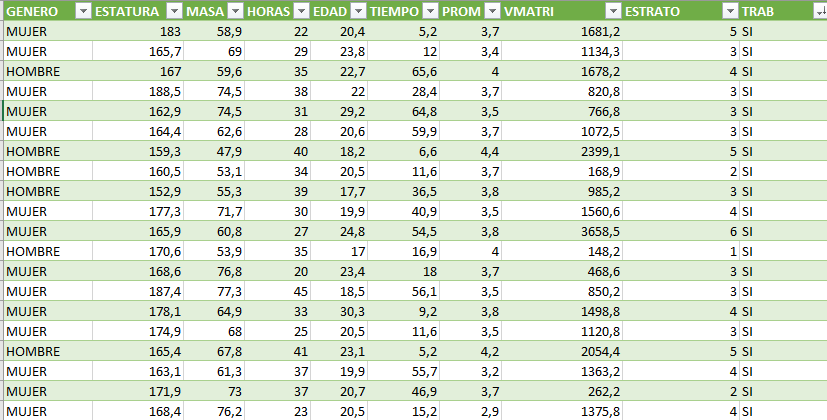
Este trabajo contiene el desarrollo de 4 puntos propuestos en la materia de estadística para ingeniería en el proyecto de aula, en el cual, 3 puntos son teóricos y el otros es un desarrollo de una aplicación de calculadora de las distribuciones (Normal, t, chi cuadrado, binomial, normal estándar), este desarrollo nos permite encontrar tanto el valor de x dada una probabilidad y los parámetros de la distribución y también dado una x encontrar la probabilidad.

En el desarrollo de los puntos teóricos se manejaron conceptos que se vieron desde el inicio de semestre como estadística descriptiva, probabilidades hasta temas de final de semestre como de pruebas de hipótesis, para hacer todas las soluciones se adjuntaron las fórmulas utilizadas para las mismas y en el punto 1 se adjuntará un documento en Excel que permitirá visualizar como se sacaron las tablas.

Finalmente, el desarrollo de este proyecto nos permitió profundizar en varios conceptos vistos anteriormente en la materia y poder aplicarlos en problemas de la vida cotidiana.

## PUNTO 1:

1. La base de datos corresponde a una muestra aleatoria de tamaño 200, que contiene las siguientes variables: GENERO (HOMBRE o MUJER), ESTATURA (en cm. del estudiante), MASA (en kg del estudiante), HORAS (número de horas dedicadas semanalmente a estudiar, fuera del horario de clases), EDAD (en años), TIEMPO (en min. Requerido para ir de su casa a la universidad), PROM (promedio obtenido en el semestre anterior), VMATRI (en miles de pesos, valor de matrícula pagada en el semestre actual), ESTRATO y TRAB (trabaja SI o NO).

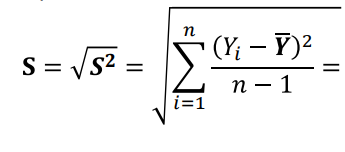


1. Del archivo Datos proyecto de aula, seleccione dos variables continuas calcule: Media, Mediana y desviación estándar e interprete sus resultados.

**Fórmulas Utilizadas:**

|  |  |
| --- | --- |
| **Media:** | **Mediana:** |
|  |  |

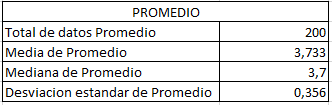
**Desviación estándar:**

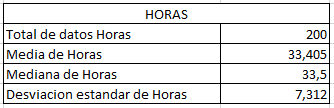
****

****

****

****





Se puede interpretar con la variable horas, que el promedio de horas que los estudiantes dedican a estudiar fuera de clase son 33 sin embargo la desviación estándar tiene un valor alto por lo que se puede concluir que estos valores son muy dispersos ya que hay estudiantes que estudian muy pocas horas y otros que estudian muchas más.

Respecto al Promedio se puede ver que la gran mayoría de los estudiantes tienen un buen promedio, e incluso tiene una dispersión de datos muy baja por lo que se puede concluir que la gran mayoría de estudiantes tienen arriba de 3 el promedio, y en general la media es de 3,7.

1. ¿La variable HORAS se comporta igual para los que trabajan y los que no trabajan? Justifique su respuesta.

R// Al realizar el promedio de horas que dedica semanalmente a estudiar fuera del horario de clase una persona que trabaja y otra que no, es que en promedio las personas que no trabajan pueden dedicarle una hora más al estudio con un promedio de 34 horas, mientras que las personas que trabajan tienen un promedio de 33 horas.

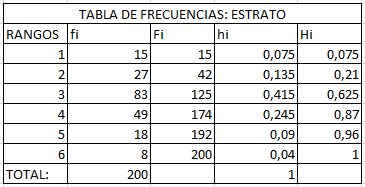
1. Construya una tabla de frecuencias para las variables **ESTRATO** y **TRAB** A partir de la tabla de frecuencias, encuentre la tabla de probabilidad conjunta empírica. A partir de la tabla encuentre las siguientes probabilidades:











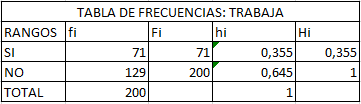
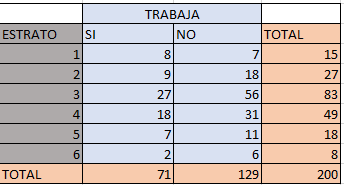


TABLA DE FRECUENCIA DE ESTRATO Y TR



**Fórmulas Utilizadas para sacar la probabilidad:**



1. Encuentre la probabilidad de que seleccionado una persona al azar esta trabaje.

A= Trabaja

P(A)= = 0.35

La probabilidad de que una persona trabaje es del 35%

1. Encuentre la probabilidad de que una persona pertenezca al estrato 3.

A = Estrato

P (A = 3) = = 0.415

La probabilidad de que una persona sea de estrato 3 es de 41%

1. Cuál es la probabilidad del no trabajar y ser del estrato 4.



A= Estrato

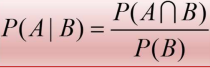
B = No trabaja

P (B y A = 4) = \* = 0.15

P (B y A=4) ==0.15

La probabilidad de que una persona no trabaje y sea de estrato 4 es del 15%

1. Si la persona trabaja, determine la probabilidad de que la persona sea del estrato 2



A= Trabaja

B = Estrato

P (A |B=2) =

La probabilidad de que dado que una persona trabaje sea de estrato 2 es de 0.2325

1. ¿Son la **EDAD** y el **ESTRATO** independientes? Explique.

R// La edad y el estrato son variables independientes ya que sin importar que edad tengas puedes tener cualquier estrato económico, además en el numeral g comprobamos que podíamos hacer el cálculo de la intersección de no trabajar y ser de estrato 4 con la fórmula de la intersección para variables independientes y con la de la intersección de la tabla de frecuencias.

## PUNTO 3:

1. Una máquina que expende bebidas gaseosas está calibrada de modo que descargue un promedio de mililitros por vaso. Si la cantidad de líquido está distribuida normalmente con una desviación estándar igual a ml. Tanto como son desconocidos, pero se cuenta con una muestra de doscientas observaciones para estimarlos. Utilice los estadísticos y como estimadores de y respectivamente, para responder los siguientes literales. Recuerde que
2. ¿Qué porcentaje de vasos contendrá menos de 150 ml?

Teniendo en cuenta, para este y todos los demás numerales, que la distribución es normal, y con un promedio y desviación estándar estimados de:

µ==

σ=S

Debemos hallar la probabilidad de que un vaso contenga menos de 150 ml, por lo tanto, se expresa así:

150) = 0.50869 (Valor encontrado con aplicativo)

= 50.86%

En conclusión, el 50.86% de los 200 vasos, tienen una cantidad menor a 150 ml, y si multiplicamos dicho porcentaje por esa cantidad, nos daría lo siguiente:

200\*50.86%=101.72

Como los vasos son una variable discreta (cantidad entera), tomamos el valor entero del resultado que nos dio, es decir, que de 200 vasos que tenemos, aproximadamente 101 vasos tienen una cantidad menor a 150 ml.

1. ¿Cuál es la probabilidad de que un vaso contenga entre 150 y 155 ml?

Teniendo en cuenta que:

b) =P(X<a)-P(X<b-1)

Debemos hallar la probabilidad de que un vaso contenga entre 150 y 155 ml, por lo tanto, X estará entre 150 y 155 ml, se expresa así:

155)=P(X<155)-P(X<149)

=0.9999-0.2026 (Valores encontrados con aplicativo)

=0.7973

= 79.73%

En conclusión, el 79.73% de los vasos se encuentran entre 150 y 155 ml, y si multiplicamos este valor por la cantidad de vasos que tenemos (200), nos da de la siguiente manera:

200\*79.73%=159.46

Como los vasos son una variable discreta (cantidad entera), tomamos el valor entero del resultado que nos dio, es decir, que de 200 vasos que tenemos, aproximadamente 159 vasos tienen entre 150 y 155 ml.

1. Si se usan vasos de 155ml, ¿cuántos de los siguientes 900 vasos se derramarán?

Teniendo en cuenta que la cantidad máxima que puede almacenar el vaso es 155 ml, debemos hallar cuál es la probabilidad de que un vaso se llene más de esa cantidad (155 ml), se expresa así:

P(X>155) =0.00001 (Valor encontrado con aplicativo)

= 0.001%

Después de hallar dicha probabilidad, la multiplicamos por el número de vasos que nos dan (900), y así sabremos cuántos vasos se derramarán:

900\*0.001%=0.009

En conclusión, probablemente ninguno de los siguientes 900 vasos se derramará, ya que teniendo en cuenta que los vasos son una variable discreta (cantidad entera), y el resultado que nos dio es 0.009, dicho resultado se toma como 0.

1. ¿Bajo qué valor estará el 20% de los vasos con menos contenido?

Debemos hallar el valor (X) para el cual están el 20% de los vasos con menos contenido, por lo tanto, la probabilidad que tenemos de encontrar dicho valor es del 20%, y se expresa así:

P (X<?) = 0.20

P(X<148.98) =0.20 (Valor encontrado con aplicativo)

En conclusión, el 20% de los vasos que tienen menos contenido, son los que tienen una cantidad menor a 148.98 ml.

**CONCLUSIONES**

* Para el calculo de probabilidades condicionales hay que tener en cuenta el apropiado uso de la fórmula, porque no es lo mismo calcular que dado que una persona es estrato 5 encontrar la probabilidad de que no trabaje a por el contrario, si nos preguntan la probabilidad de que no trabaje y que sea estrato 5 el calculo sería diferente porque se usa otra fórmula.
* La librería scipy nos ayuda a calcular probabilidades de una manera muy sencilla en Python, lo único que es necesario es conocer algunas propiedades de probabilidades y entender que parámetros recibe cada distribución para poder da una funcionalidad a la librería.

**BIBLIOGRAFIA**

* <https://support.microsoft.com/es-es/office/función-contar-si-e0de10c6-f885-4e71-abb4-1f464816df34>
* <https://uvirtual.udem.edu.co/pluginfile.php/781656/mod_resource/content/2/Pdf%20Probabilidad.pdf>
* <https://es.acervolima.com/pyqt5-qlistwidget-verificar-si-la-clasificacion-esta-habilitada-o-no/>
* <https://scipy.org>
* <https://programmerclick.com/article/68751198231/>

**ANEXOS**

En el documento de la entrega se podrán encontrar todos los anexos correspondientes a los desarrollos hechos en Excel para poder determinar tablas y cálculos.

El lenguaje que se usó fu Python y las librerías trabajadas para el desarrollo de la app fueron:

* Scipy: Para las distribuciones
* Matplotlib: Para graficar distribuciones
* Sys
* Pqt5: Para la interfaz gráfica

