



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

Tarea-Examen 1

Integrantes:

Yonathan Berith Jaramillo Ramírez. 419004640

Dalia Viridiana Martínez Retiz. 309213293

Alejandro Sánchez Campos. 310131786

Profesor: Ruth Selene Fuentes García

Ayudante: Carlos Alberto Arriaga Solórzano



13 Septiembre, 2023

De la 1 a la 5 las respuestas fueron resueltas en el repo:
<https://github.com/Yony6041/statistics>

1. Resume la informacion de los siguientes datos como te parezca mas conveniente.
2. Las calificaciones de un grupo de 18 estudiantes en una escala de 100 puntos: 31, 35, 37, 40, 40, 51, 54, 55, 57, 58, 60, 60, 62, 62 65, 67, 75, 89.
3. Los dias laborables que han perdido 20 trabajadores: 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 2, 3, 3, 4, 4, 5, 5, 8, 45
4. La estatura de 20 mujeres en metros que se estudia con relacion a una condicion medica: 1.52, 1.60, 1.57, 1.52, 1.60, 1.75, 1.73, 1.63, 1.55, 1.63, 1.65, 1.55, 1.65, 1.60, 1.68 , 2.50, 1.52, 1.65, 1.60, 1.65
5. Los datos contenidos en la base mtcars en R

Base_descriptivo.csv

y para la variables post carga 2 hrs y colesterol total, ¿considera usted que su comportamiento es similar en toda la muestra?

	Post carga 2 hrs	Colesterol total
count	428.000000	428.000000
mean	123.640187	199.859813
std	61.346304	40.963444
min	46.000000	106.000000
25%	89.000000	171.000000
50%	109.000000	196.000000
75%	135.000000	224.250000
max	479.000000	420.000000

Basado en el análisis descriptivo, se puede observar que las variables "Post carga 2 hrs" y "Colesterol total" muestran características distintas:

- **Escala y Rango:** Las dos variables operan en diferentes escalas y rangos. Mientras que "Post carga 2 hrs" tiene un rango de 46 a 479, "Colesterol total" varía de 106 a 420.

- **Media y Desviación Estándar:** La media y la desviación estándar también son diferentes para ambas variables. La media es aproximadamente 123.64 para "Post carga 2 hrs" y aproximadamente 199.86 para "Colesterol total". La desviación estándar es aproximadamente 61.35 y 40.96 respectivamente.

Estas diferencias en escala, rango, y estadísticas descriptivas sugieren que el comportamiento de estas dos variables no es similar en toda la muestra.

6. Suponga que X_1, \dots, X_n son variables aleatorias iid distribuidas Binomial(3,p). Obtenga la distribución para $Y = \sum_{i=1}^n X_i$. Considere $Z = Y/(3n)$, ¿Cuales son los posibles valores que toma Z , ¿ Como se distribuye Z ? Note: que Z es una variable aleatoria discreta.

Suponga que X_1, \dots, X_n son variables aleatorias iid distribuidas Binomial(3,p). La variable aleatoria $Y = \sum_{i=1}^n X_i$ también se distribuirá de manera binomial pero con parámetros modificados, es decir $Y \sim \text{Binomial}(3n, p)$.

Ahora, consideremos $Z = \frac{Y}{3n}$. Los posibles valores que Z puede tomar son $\{0, \frac{1}{3n}, \frac{2}{3n}, \dots, 1\}$.

Z es una variable aleatoria discreta que puede interpretarse como una media muestral de n observaciones de una Binomial(3, p), y por lo tanto su distribución se acercará a una distribución normal para n grande, según el teorema del límite central. Sin embargo, para n pequeño, Z no tiene una forma de distribución simple.

7. Considere dos variables aleatorias independientes $u \sim X_p^2$ y $v \sim X_q^2$, muestre que $f = \frac{u/p}{v/q}$ se distribuye $F_{p,q}$

Considere dos variables aleatorias independientes $u \sim \chi_p^2$ y $v \sim \chi_q^2$. Queremos mostrar que $f = \frac{u/p}{v/q}$ se distribuye $F_{p,q}$.

Podemos reescribir f como:

$$f = \frac{u/p}{v/q} = \frac{u \cdot q}{v \cdot p}$$

Ahora, $\frac{u}{p}$ y $\frac{v}{q}$ son variables χ^2 divididas por sus grados de libertad, y por lo tanto se distribuyen como una distribución gamma.

Usando la propiedad de que la división de dos variables gamma independientes sigue una distribución F , podemos concluir que f se distribuye como $F_{p,q}$.