

TRABAJO PRÁCTICO INTEGRADOR

1. Conceptos teóricos:

Población: conjunto de individuos, personas o cosas, sobre las cuales se desea información o de las cuales se estudia alguna característica. Son las unidades estadísticas poseedoras de las características bajo estudio.

Muestra: la muestra es una porción representativa de la población sobre la cual se trabaja para obtener valores estadísticos que se infieren a la población. En este caso, la muestra consiste en 500 usuarios que utilizaron el programa en el mes de enero y febrero de 2021

Unidad Estadística: está formada por cada uno de los elementos individuales que en forma conjunta constituyen la población en estudio.

2. Tablas de frecuencia

a) Saldo (5 intervalos)

| Variable | Clase | LI | LS | MC | FA | FR | FAA | FRA |
|----------|-------|------------|------------|----------|-----|--------|-----|--------|
| Saldo | 1 | [-75,0000 | 58,0000) | -8,5000 | 3 | 0,0060 | 3 | 0,0060 |
| Saldo | 2 | [58,0000 | 191,0000) | 124,5000 | 63 | 0,1260 | 66 | 0,1320 |
| Saldo | 3 | [191,0000 | 324,0000) | 257,5000 | 238 | 0,4760 | 304 | 0,6080 |
| Saldo | 4 | [324,0000 | 457,0000) | 390,5000 | 166 | 0,3320 | 470 | 0,9400 |
| Saldo | 5 | [457,0000 | 590,0000] | 523,5000 | 30 | 0,0600 | 500 | 1,0000 |

b) Notificaciones activas

| Variable | Clase | Categorías | FA | FR | FAA | FRA |
|----------|-------|------------|-----|--------|-----|--------|
| Noti | 1 | 0 | 384 | 0,7680 | 384 | 0,7680 |
| Noti | 2 | 1 | 116 | 0,2320 | 500 | 1,0000 |

| Frecuencias absolutas | | |
|-----------------------|-------|------------|
| Noti | Total | Porcentaje |
| 0 | 384 | 76,8000 |
| 1 | 116 | 23,2000 |
| Total | 500 | 100,0000 |

3. Medidas

a) Saldo

| Variable | n | Media | D.E. | Var(n-1) | CV | Mín | Máx | Mediana | Asimetría | Kurtosis |
|----------|-----|----------|---------|-----------|---------|----------|----------|----------|-----------|----------|
| Saldo | 500 | 299,3943 | 96,7825 | 9366,8552 | 32,3261 | -60,6200 | 564,6300 | 299,3950 | -0,0578 | 0,0491 |

n: representa la muestra

Media: Representa el promedio que tiene el saldo para una muestra de 500

Mediana: Expresa que el 50% de la muestra de la población tiene 299,40 o menos y el otro 50% tiene 299,40 o más

D.E: Expresa que los valores pueden variar entre +- 96,78 respecto de la media

Var(n-1): La varianza representa que tan dispersos se van a encontrar los valores respecto de la media. A mayor varianza, los valores se encuentran más dispersos y a menor varianza los valores se encuentran menos dispersos respecto de la media

CV :La desviación estándar representa el 32.33% de la media. A mayor valor, mayor heterogeneidad y a menor valor mayor homogeneidad

Rango: es la diferencia entre el valor máximo y el min, en este caso es de: 625,25

Asimetría: Como la asimetría tiene un valor de -0,06, esto indica que es una asimetría por izquierda, es decir, que la mayoría de los datos están en la parte izquierda respecto de la media (sesgo izquierda) y que: media < mediana < moda

Kurtosis: Como el valor de kurtosis es 0,05, indica que presenta un elevado grado de concentración alrededor de los valores centrales de la variable y la curva toma una forma leptocúrtica.

b) Notificaciones activas

| Variable | n | Media | D.E. | Var (n-1) | CV | Mín | Máx | Mediana | Asimetría | Kurtosis |
|----------|-----|--------|--------|-----------|----------|--------|--------|---------|-----------|----------|
| Noti | 500 | 0,2320 | 0,4225 | 0,1785 | 182,1257 | 0,0000 | 1,0000 | 0,0000 | 1,2736 | -0,3876 |

n: representa la muestra de las notificaciones, tomando el valor de 500

media: representa el valor promedio que tiene las notificaciones con una muestra de 500

mediana: expresa que el 50% de la muestra de la población tiene 0 o menos y el otro 50% tiene 0 o más.

D.E: representa que los valores pueden variar entre +/- 0,42 respecto de la media

Var(n-1): representa que tan dispersos se van a encontrar los valores respecto de la media, a mayor varianza, los valores se encuentran más dispersos, y a menor varianza los valores se encontrarán menos dispersos respecto de la media

CV: representa 182,13% de la media. a mayor valor mayor heterogeneidad y a menor valor, mayor homogeneidad

rango: es la diferencia entre el valor máximo y el mínimo, en este caso es de: 1

Asimetría: como la asimetría tiene un valor de 1,27 esto indica que es asimetría por derecha, es decir, que la mayoría de los datos están en la parte derecha respecto a la media (sesgo derecha) y que media > mediana > moda.

kurtosis: como el valor de kurtosis es -0,39 indica que representa un reducido grado de concentración alrededor de los valores centrales de la variable y la curva toma una forma platicúrtica

4. a) El parámetro de interés es la media poblacional y la variable de estudio es el Saldo, el cual representa el monto de dinero que posee el usuario

b) Varianza poblacional desconocida, muestra mayor a 30, distribución que tiende a la normal

$$K(\hat{\theta}; \theta) = \frac{\bar{x} - \mu}{\frac{\hat{\sigma}}{\sqrt{n}}} \sim N(0,1)$$

$$e = z_{1-\frac{\alpha}{2}} \frac{\hat{\sigma}}{\sqrt{n}} \quad L_i = \bar{x} - e \quad L_s = \bar{x} + e$$

$$1 - \alpha = \Pr \left\{ \bar{x} - z_{1-\frac{\alpha}{2}} \frac{\hat{\sigma}}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{1-\frac{\alpha}{2}} \frac{\hat{\sigma}}{\sqrt{n}} \right\}$$

c) Intervalo de confianza del 95%

| Variable | Parámetro | Estimación | E.E. | n | LI(95%) | LS(95%) |
|----------|-----------|------------|--------|-----|----------|----------|
| Saldo | Media | 299,3943 | 4,3282 | 500 | 290,8905 | 307,8981 |

La media poblacional estimada con respecto al saldo es de 299,39.

Con una confianza del 95% se puede concluir que la media poblacional respecto del saldo se encuentra entre los valores 290,89 y 307,90.

Además se deduce que cada 100 intervalos que se construyen tomando todas las muestras posibles de la población, 95 de cada 100 de ellos contendrán al parámetro bajo estudio siendo este el saldo que cada usuario posee.

d) Intervalo de confianza del 99%

| Variable | Parámetro | Estimación | E.E. | n | LI(99%) | LS(99%) |
|----------|-----------|------------|--------|-----|----------|----------|
| Saldo | Media | 299,3943 | 4,3282 | 500 | 288,2027 | 310,5859 |

La media poblacional estimada con respecto al saldo es de 299,39.

Con una confianza del 99% se puede concluir que la media poblacional respecto del saldo se encuentra entre los valores 288,20 y 310,59.

Además se deduce que cada 100 intervalos que se construyen tomando todas las muestras posibles de la población, 99 de cada 100 de ellos contendrán al parámetro bajo estudio siendo este el saldo que cada usuario posee.

Conclusión final: al aumentar el nivel de confianza desde 95% a 99% los límites del intervalo se amplían por lo tanto el error aceptable en relación al ejercicio anterior aumentó y por consecuencia el riesgo de caer fuera de ese intervalo disminuyó.

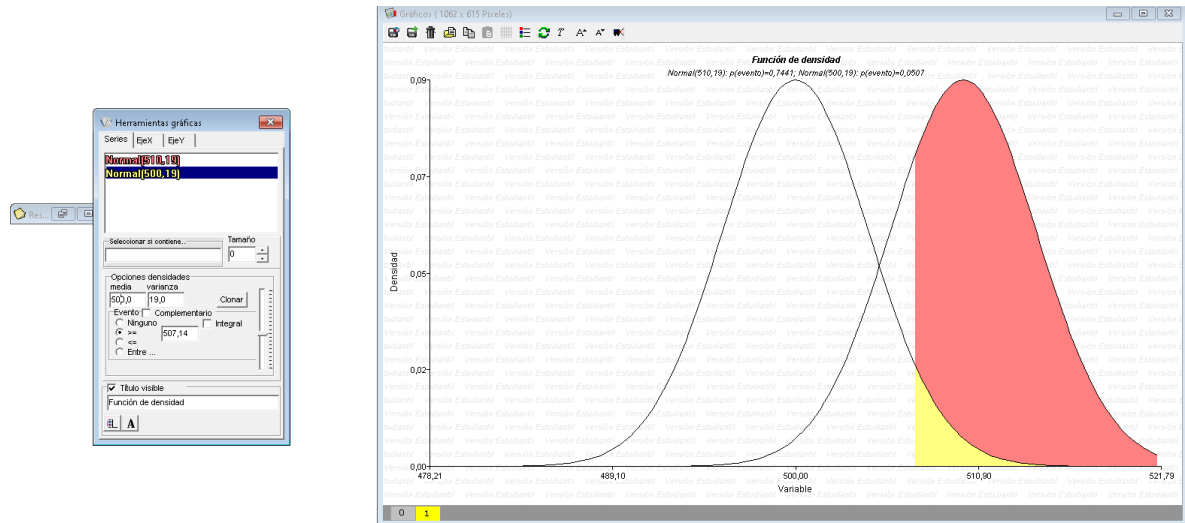
5.

| Variable | n | Media | DE | LI(95) | T | p(Unilateral D) |
|----------|-----|----------|---------|----------|----------|-----------------|
| Saldo | 500 | 299,3943 | 96,7825 | 292,2617 | -46,3480 | >0,9999 |

Conclusión: como el valor p es mayor a α , ($0,05 < 0,9999$), la hipótesis nula es aceptada. Esto indica que la hipótesis alternativa planteada para el investigador de que el saldo

promedio es mayor a \$500 es rechazada ya que como se puede ver, el valor promedio de los saldos es de 299,39.

6. Contraste al punto 5



Normal(510,19): p(evento)=0,7441; Normal(500,19): p(evento)=0,0507

Al realizar el punto 5 se tiene como conclusión de que la hipótesis alternativa será rechazada ya que supera por demás al promedio que se obtiene de los saldos, pero si se quieren analizar dos hipótesis alternativa sin tener en cuenta la hipótesis nula se obtiene el gráfico adjunto previamente.

En ese gráfico se pueden obtener α , $1-\alpha$, β y $1-\beta$ con los datos que se tenían siendo las medias 500 y 510, para una dódima lateral derecha con un punto crítico de 507,14 y una desviación estándar muestral de 18,73372 y 19 respectivamente para las medias mencionadas anteriormente.

Se eligió la dódima lateral derecha ya que en el enunciado se menciona que el promedio será mayor a un número dado.

α = Probabilidad de rechazar la hipótesis nula dado que esta es cierta = 0,0495

$1-\alpha$ = Probabilidad de aceptar la hipótesis nula dado que esta es cierta = 0,2559

β = Probabilidad de aceptar la hipótesis nula dado que es falsa = 0,2559

$1-\beta$ = Probabilidad de rechazar la hipótesis nula dado que es falsa = 0,7441

7. a) Tamaño de muestra para el punto 2

| Noti | Total | Porcentaje |
|-------|-------|------------|
| 0 | 384 | 76,80 |
| 1 | 116 | 23,20 |
| Total | 500 | 100,00 |