

# Análisis de los Datos de la Federación Nacional de Cafeteros: Exploración con Intervalos de Confianza y Pruebas de Hipótesis

Santiago Montejo, Angel Granados, Wilson Jerez

Mayo 2024

## Introducción

En éste ensayo, nuestro objetivo es analizar los datos proporcionados por la Federación Nacional de Cafeteros para determinar el intervalo de confianza adecuado para el valor promedio de las facturas de las empresas extranjeras que realizan compras. Utilizaremos este intervalo de confianza como base para establecer una prueba de hipótesis que nos permita analizar el parámetro de interés. La variable en cuestión, “Valor\_factura\_USD”, representa el monto total de las transacciones en dólares estadounidenses y es crucial para evaluar el rendimiento comercial de la federación en el mercado internacional del café. Exploraremos cómo la Prueba de Hipótesis puede ayudarnos a comprender mejor si hay diferencias significativas en el valor promedio de estas transacciones, lo que podría influir en las estrategias de negociación y comercio de la federación.

## Selección de la Variable

La variable seleccionada para este análisis es “Valor\_factura\_USD”, que representa el monto total de las facturas en dólares estadounidenses de las empresas extranjeras que realizan compras a la Federación Nacional de Cafeteros. Esta variable es fundamental porque proporciona una medida directa del rendimiento comercial de la federación en el mercado internacional del café. Su análisis permite evaluar la salud financiera de la federación, identificar tendencias en las transacciones comerciales y tomar decisiones estratégicas para optimizar las operaciones comerciales.

```
DestinoDeExportación <- read.csv('DestinoDeExportación.csv')
Valor_factura_USD <- DestinoDeExportación$Valor_factura_USD
```

## Selección del Intervalo de Confianza

Para determinar el intervalo de confianza más adecuado de la variable seleccionada, utilizamos la siguiente función:

```
calcular_intervalo_confianza <- function(nivel_confianza, alpha, coeficiente_confianza, estimador, error) {
  # Calcula el valor crítico según el nivel de confianza y alpha
  valor_critico <- qnorm(1 - alpha / 2)

  # Calcula el error estándar
  margen_error <- valor_critico * error_estandar

  # Calcula el intervalo de confianza
  limite_inferior <- estimador - margen_error
  limite_superior <- estimador + margen_error
}
```

```

    return(list(limite_inferior = limite_inferior, limite_superior = limite_superior))
}

# Calcular el estimador (media) y el error estándar
media <- mean(Valor_factura_USD, na.rm = TRUE)
print(media)

## [1] 555284.7

error_estandar <- sd(Valor_factura_USD, na.rm = TRUE) / sqrt(sum(!is.na(Valor_factura_USD)))
print(error_estandar)

## [1] 7682.327

# Parámetros para el cálculo del intervalo de confianza
nivel_confianza <- 0.95
alpha <- 0.05
coeficiente_confianza <- 1.96 # Para un nivel de confianza del 95%

# Llamada a la función para calcular el intervalo de confianza
resultado_intervalo_confianza <- calcular_intervalo_confianza(nivel_confianza, alpha, coeficiente_confianza)

# Mostrar resultado
print(resultado_intervalo_confianza)

## $limite_inferior
## [1] 540227.6
##
## $limite_superior
## [1] 570341.7

```

## Formulación de la Hipótesis

En este análisis, formulamos las siguientes hipótesis:

- Hipótesis Nula (H0): El valor promedio de las facturas de las empresas extranjeras que compran café a la Federación Nacional de Cafeteros es igual a \$555,284.7 USD.

```
H0 <- media
```

- Hipótesis Alternativa (H1): El valor promedio de las facturas de las empresas extranjeras que compran café a la Federación Nacional de Cafeteros es diferente de \$555,284.7 USD.

La hipótesis nula establece una referencia a partir de la cuál se puede evaluar si hay suficiente evidencia para rechazarla en favor de la hipótesis alternativa. En el contexto del análisis de datos, estas hipótesis nos permiten investigar si el valor promedio de las facturas es significativamente diferente del valor de referencia establecido, lo que puede tener implicaciones importantes para la gestión financiera y comercial de la federación.

## Contexto de Aplicación

La Prueba de Hipótesis es una herramienta esencial para evaluar afirmaciones sobre el valor promedio de las facturas de las empresas extranjeras que compran café a la Federación Nacional de Cafeteros. Este análisis ayuda a la federación a entender si hay diferencias significativas en el monto promedio de estas transacciones, lo que puede influir en sus estrategias comerciales y financieras.

Por ejemplo, la Prueba de Hipótesis se aplica para:

- Determinar el impacto de las promociones y descuentos en el valor promedio de las facturas.
- Evaluar si las variaciones estacionales afectan el monto promedio de las transacciones.
- Comparar el rendimiento comercial con competidores en el mercado internacional del café.

En síntesis, la Prueba de Hipótesis proporciona una metodología sólida para entender cómo los cambios en el entorno comercial pueden influir en el valor promedio de las facturas, permitiendo a la Federación Nacional de Cafeteros tomar decisiones informadas y estratégicas.

## Procedimiento para Enfrentar las Hipótesis

La aplicación de la Prueba de Hipótesis en el contexto del análisis de las ventas hacia el exterior en el Departamento de Analítica de la Federación Nacional de Cafeteros implica los siguientes pasos detallados:

1. **Establezca las hipótesis nula y alternativa.**
2. **Elija un estadístico de prueba adecuado.**
3. **Calcule el valor P con base en los valores del estadístico de prueba.**
4. **Saque conclusiones con base en el valor P y los conocimientos del sistema científico.**

Este procedimiento proporciona un marco estructurado para evaluar afirmaciones sobre el valor promedio de las facturas de las empresas extranjeras que compran café a la Federación Nacional de Cafeteros, lo que permite tomar decisiones informadas y estratégicas en el ámbito comercial y financiero.

## Uso de la Evidencia Muestral

En el análisis de las ventas hacia el exterior en la Federación Nacional de Cafeteros, la evidencia muestral juega un papel clave:

### 1. Recolección de Datos:

- Se recopila una muestra aleatoria e imparcial de 47 empresas, representando el 10% de las ventas estimadas en 2024.

```
DestinoDeEsportación24 <- subset(DestinoDeEsportación, Año == 2024)
tamano_muestra <- round(0.1 * nrow(DestinoDeEsportación24))
set.seed(42)
DesExRandom24 <- DestinoDeEsportación24[sample(nrow(DestinoDeEsportación24),
size = tamano_muestra, replace = FALSE), ]
```

### 2. Análisis de la Muestra:

- Se analizan estadísticas descriptivas como el promedio y la desviación estándar del valor de las facturas.

```
promedioExrandom24 <- mean(DesExRandom24$Valor_factura_USD)
DesviaciónExRandom24 <- sd(DesExRandom24$Valor_factura_USD)
```

### 3. Prueba de Hipótesis:

- Se utiliza la muestra para realizar la Prueba de Hipótesis sobre el valor promedio de las facturas.

### 4. Interpretación de los Resultados:

- Los resultados de la prueba guían la toma de decisiones, determinando si el valor promedio difiere significativamente de una referencia establecida.

### 5. Consideraciones sobre la Recolección y Análisis de Datos:

- La recolección de datos debe ser imparcial y representativa, mientras que el análisis debe seguir procedimientos rigurosos para garantizar resultados confiables.

Este proceso garantiza una evaluación fundamentada del rendimiento comercial, permitiendo decisiones estratégicas basadas en datos sólidos.

## Posibilidades de Errores

En el contexto de la Prueba de Hipótesis, existen dos tipos de errores que pueden ocurrir:

### 1. Error Tipo I:

- Ocurre cuando se rechaza incorrectamente la hipótesis nula, aunque sea verdadera.
- En el análisis de las ventas hacia el exterior de la Federación Nacional de Cafeteros, un error tipo I implicaría concluir que el valor promedio de las facturas es significativamente diferente de \$555,284.7 USD. cuando en realidad no lo es.

### 2. Error Tipo II:

- Ocurre cuando se acepta incorrectamente la hipótesis nula, aunque sea falsa.
- En el mismo análisis, un error tipo II podría implicar no detectar diferencias significativas en el valor promedio de las facturas cuando realmente existen.

Ambos errores pueden tener consecuencias significativas para la toma de decisiones. Por ejemplo, un error tipo I podría llevar a cambios innecesarios en las estrategias comerciales, mientras que un error tipo II podría resultar en oportunidades perdidas para mejorar el rendimiento comercial. Por lo tanto, es fundamental considerar estos errores y minimizar su probabilidad al diseñar y ejecutar la Prueba de Hipótesis.

## Prueba Seleccionada

- Tipo de Prueba: Dos colas
- Justificación:
  - Se selecciona una prueba de dos colas debido a que la hipótesis alternativa sugiere que el valor promedio de las facturas podría ser diferente del valor de referencia establecido en la hipótesis nula.
  - Esta prueba permite evaluar si hay diferencias significativas tanto en la dirección positiva como en la negativa con respecto al valor de referencia, lo que brinda una comprensión más completa de la variabilidad en el valor promedio de las facturas.
  - Al considerar ambas direcciones, se pueden identificar de manera más efectiva cualquier cambio significativo en el valor promedio de las facturas, lo que ayuda a tomar decisiones más informadas en la gestión financiera y comercial de la federación.

## Uso del p-valor

- **Explicación del p-valor:**
  - Es la probabilidad de obtener un valor del estadístico de prueba igual o más extremo que el valor observado en la muestra, bajo la suposición de que la hipótesis nula es cierta.
  - En este contexto, el p-valor ( $p$ ) indica la probabilidad de observar el valor promedio de las facturas, o uno aún más extremo, si la verdadera media poblacional fuera igual al valor de referencia establecido en la hipótesis nula.
- **Importancia en la Prueba de Hipótesis:**
  - Proporciona una medida de la evidencia en contra de la hipótesis nula, permitiendo tomar decisiones basadas en la evidencia estadística.
  - Un p-valor menor que el nivel de significancia ( $\alpha$ ) predefinido (generalmente 0.05) sugiere rechazar la hipótesis nula en favor de la alternativa.
- **Interpretación del p-valor:**
  - Si es menor que  $\alpha$ , se rechaza la hipótesis nula, indicando que la evidencia muestral es consistente con la hipótesis alternativa.
  - Por ejemplo, un p-valor de 0.02 significa que hay un 2% de probabilidad de observar el valor promedio de las facturas, o uno aún más extremo, bajo la hipótesis nula.

```
facturas <- DesExRandom24$Valor_factura_USD
# Calcular la media muestral
media_muestral <- mean(facturas)
Z <- (media_muestral - H0) / sd(facturas) / sqrt(length(facturas))
# Calcular el p-valor
p_valor <- pt(-abs(Z), df = length(facturas) - 1)

# Comparar el p-valor con el nivel de significación
alpha <- 0.05
if (p_valor < alpha) {
  print("Rechazamos la hipótesis nula. Hay evidencia estadística.")
} else {
  print("No rechazamos la hipótesis nula. No hay evidencia estadística.")
}

## [1] "No rechazamos la hipótesis nula. No hay evidencia estadística."
```

## Conclusión

En este ensayo, se abordó la aplicación de la Prueba de Hipótesis para analizar el valor promedio de las facturas de las empresas extranjeras que compran café a la Federación Nacional de Cafeteros. Se establecieron hipótesis nula y alternativa relacionadas con el valor promedio de las facturas, y se utilizó una muestra aleatoria de 47 empresas para realizar la Prueba de Hipótesis.

**No rechazamos la hipótesis nula. No hay evidencia estadística significativa para concluir que el valor promedio de las facturas es diferente del valor establecido en la hipótesis nula.**

## Referencias

1. Walpole, R. E., Myers, R. H., Myers, S. L., & Ye, K. (2012). Probabilidad y estadística para ingeniería y ciencias (9th ed.). México: Pearson Educación.
2. Wackerly, D. D., Mendenhall, W., & Scheaffer, R. L. (2014). Estadística Matemática con Aplicaciones (7th ed.). México: Cengage Learning Editores.