

Informe de resultados

TP Final - PEIA

Integrantes:

Sebastian Carreras

Lucas Fajardo

Marcelo Mas Valdecantos

1. *Análisis de ventas de cada mes en "Santa Ana". Para significancias del 95% y 99%.*

Fecha	IC_95_lower	IC_95_upper	IC_99_lower	IC_99_upper
2023-01-31	16092.602493	18067.463549	15750.421171	18409.644871
2023-02-28	17854.315835	20468.461039	17396.387014	20926.389861
2023-03-31	21020.975441	22951.704652	20686.440779	23286.239314
2023-04-30	19109.945756	21009.700987	18779.659415	21339.987328
2023-05-31	20345.360835	22095.514783	20042.114188	22398.761429
2023-06-30	21329.911547	23260.547603	20994.256349	23596.202801
2023-07-31	19895.991677	21821.431982	19562.373416	22155.050242
2023-08-31	20921.390168	22982.521985	20564.260835	23339.651318
2023-09-30	21889.839921	23927.102824	21535.646867	24281.295879
2023-10-31	21211.753456	23136.682380	20878.223802	23470.212034
2023-11-30	21311.828744	23147.737761	20992.642534	23466.923970
2023-12-31	18984.463062	20812.301732	18667.756095	21129.008698

Se realizaron intervalos de confianza estadísticos utilizando el valor crítico de la distribución t de Student, ya que se asume que la varianza poblacional es desconocida.

Estos resultados nos permiten tener una aproximación sobre un parámetro, que en este caso es el rango en el que se encuentran la media de ventas de cada mes

Si se repitiera este análisis muchas veces con diferentes muestras de datos, el 95% de las veces la media verdadera estaría dentro de este intervalo.

Sin embargo, esto no significa que las ventas diarias estén entre estos valores, sino que la media de las ventas diarias del mes se espera que esté en ese rango.

Los intervalos de confianza calculados para el supermercado 'Santa Ana' reflejan que las ventas son relativamente estables en la mayoría de los meses, con pocas fluctuaciones que pudieran señalar eventos atípicos.

2. *Prueba ANOVA para determinar si las ventas esperadas de todas las tiendas son iguales.*

Antes de comenzar con el análisis, se verifica que las varianzas sean homogéneas; esto se hace con un contraste de homogeneidad de varianzas de G de Cochran ya que nuestros datos son balanceados

$$\left. \begin{array}{l} H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 = \sigma_5^2 \\ H_1: \text{Al menos una igualdad no es cierta} \end{array} \right\} \begin{array}{l} \rightarrow \beta \\ \rightarrow \alpha = 0.05 \end{array}$$

Para este contraste el estadístico y el punto crítico quedan definidos como:

$$G_{\text{expt}} = \frac{\max(S_i^2)}{\sum_{i=1}^k S_i^2}$$

$$P_c = G_{\text{teo}} = G_{n,k,\alpha}^1$$

Con la siguiente condición de rechazo:

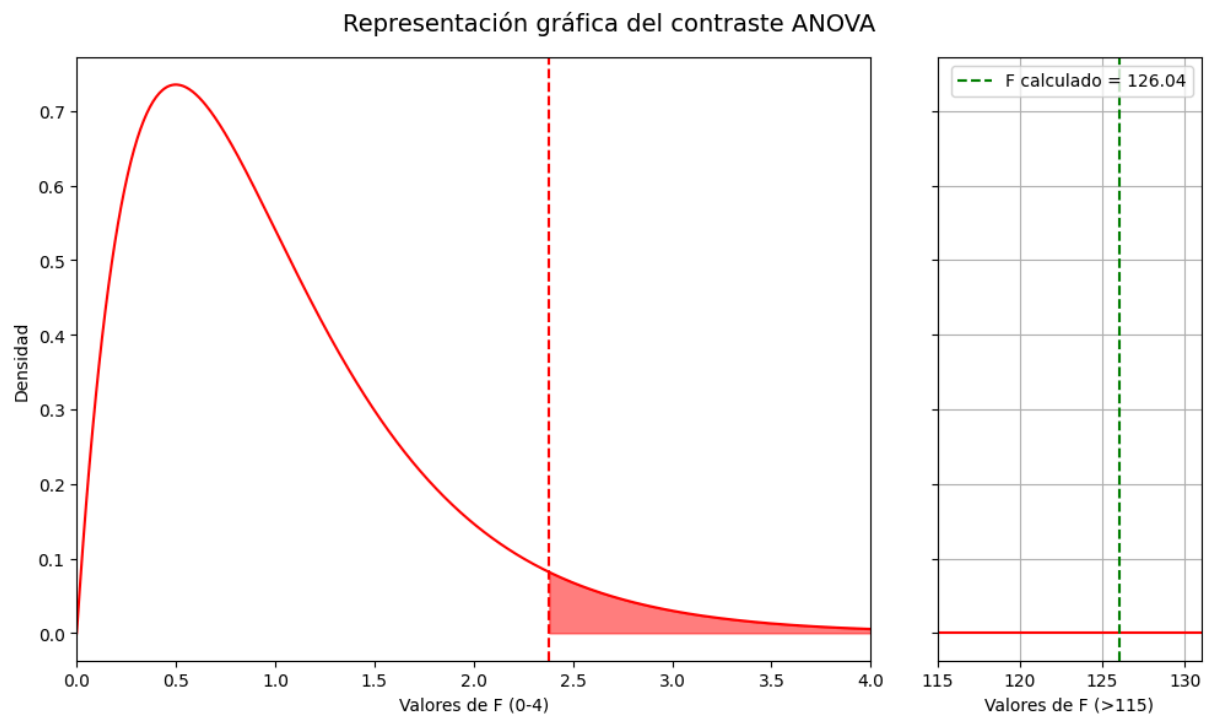
$$G_{\text{expt}} = \frac{\max(S_i^2)}{\sum_{i=1}^k S_i^2} > G_{n,k,\alpha} = G_{\text{teo}} \Rightarrow \text{SE RECHAZA } H_0$$

Nuestro estadístico G: 0.2077 que se compara con un G de tabla para para k=5 y n=145 para una significancia del 5%; Se elige este valor porque es el más próximo de las tablas encontradas, al igual que ejercicio visto en clase. El G de tabla tiene un valor de 0.2513. Concluimos en que no se rechaza la hipótesis nula, se cumple el requisito de homogeneidad en las varianzas.

Una vez se comprueba la homogeneidad, se procede a hacer el test ANOVA:

$$\left. \begin{array}{l} H_0: \mu_0 = \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 \\ H_1: \text{Al menos una igualdad no es cierta} \end{array} \right\} \begin{array}{l} \rightarrow \beta \\ \rightarrow \alpha = 0.05 \end{array}$$

Un estadístico F con un valor de: 126.0367 y un p valor asociado asociado muy pequeño, calculado con los grados de libertad entre grupos (4) y los grados de libertad dentro de los grupos (1820), nos dan a entender que se rechaza la hipótesis nula y se concluye que las ventas esperadas no son iguales en todas las tiendas



3. *Prueba de hipótesis para determinar si la tienda con mayor ventas es diferente estadísticamente significativa a tienda con menor ventas*

Se sabe que:

Promedio de ventas por tienda:

Santa Ana	20990.659056
La Floresta	21713.716026
Los Cedros	18068.391082
Palermo	20787.278462
Córdoba	17900.271075

La tienda con mayor promedio de ventas es: La Floresta

La tienda con menor promedio de ventas es: Córdoba

El test se construye mediante el siguiente estadístico:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S_{pool}^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

$$S_{pool}^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$$

Es una prueba t clásica con varianza agrupada a dos colas. Se utiliza esta prueba debido a que se asume que las varianzas son iguales por el test realizado en el punto anterior. Además es útil para trabajar con muestras con varianza poblacional desconocida, como se asume en este caso

Se plantea la siguiente hipótesis nula: No existe una diferencia significativa en las ventas promedio entre las dos tiendas.

Se obtiene el siguiente estadístico y p-valor:

Estadístico t: 17.021

P-valor: 0.00 (valor pequeño)

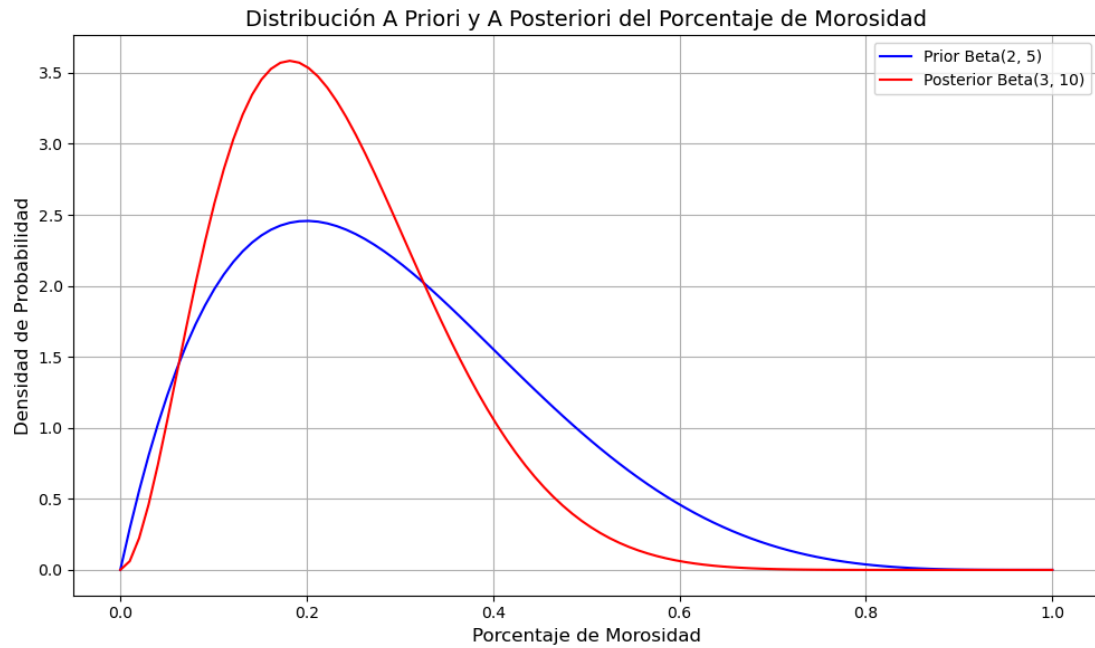
Por ello rechazamos la hipótesis nula y que la diferencia entre las tiendas es significativa con una significancia del 5%.

Como conclusión podemos decir que la prueba de hipótesis realizada entre las tiendas con mayor y menor promedio de ventas confirma que existe una diferencia significativa en las ventas promedio, lo que podría indicar factores diferenciadores en el rendimiento de cada tienda.

4. Determinación de la distribución a posteriori del parámetro p de porcentaje de morosidad

Se asumió una distribución Beta como a priori, y se ajustaron los parámetros basados en los datos observados de morosidad. Posteriormente, se calculó la distribución a posteriori del parámetro p, junto con su media y varianza.

La distribución a posteriori del porcentaje de morosidad fue ajustada utilizando una distribución Beta. La media de la distribución posterior fue de ~0.231, y la varianza fue de ~0.0127.



Media de la distribución posterior: 0.23076923076923078

Varianza de la distribución posterior: 0.01267962806424345

La distribución a posteriori del porcentaje de morosidad proporciona una estimación clara del comportamiento de la morosidad en las tiendas, con una media del 23.1%. La varianza relativamente baja indica que la estimación es precisa y que el porcentaje de morosidad no presenta grandes fluctuaciones.