**Grupo 4.**

Integrantes:

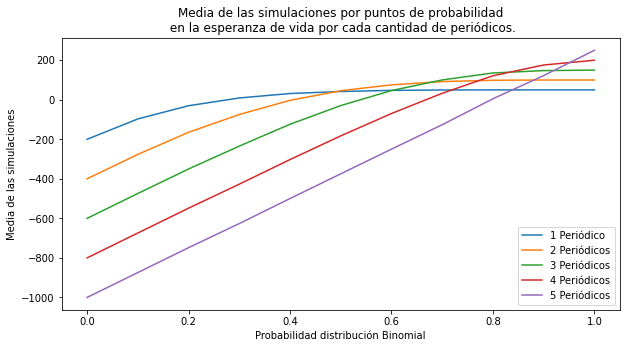
* Luisa María De La Hortúa Pulido 201712785
* David Santiago Moreno Medina 201728270

1. **Modelo del vendedor de periódicos.**
2. Modifique el modelo para que la demanda sea una variable aleatoria binomial entre 0 y 5. Modifique el parámetro p (probabilidad de éxito) y describa su impacto.

Al tener una variable aleatoria binomial con parámetros 0 y 5, se sabe que se trabajarán con periódicos que oscilarán entre 0 y 5 periódicos. El parámetro más importante para generar variaciones, en este caso es p (probabilidad de éxito) o en el contexto es la probabilidad de que venda el periódico. Al momento de generar variaciones se obtiene la siguiente tabla descriptiva de las medias por cada cantidad de periódicos comprados:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **p** | **media x=1** | **media x=2** | **media x=3** | **media x=4** | **media x=5** |
| 0.0 | -200 | -400 | -600 | -800 | -1000 |
| 0.1 | -97.025 | -276.525 | -474.250 | -674.175 | -874.175 |
| 0.2 | -30.175 | -164.650 | -350.225 | -548.550 | -748.325 |
| 0.3 | 9.150 | -74.700 | -234.475 | -427.050 | -626.300 |
| 0.4 | 31.575 | -2.925 | -123.900 | -302.575 | -499.725 |
| 0.5 | 41.900 | 45.775 | -29.550 | -182.725 | -374.750 |
| 0.6 | 47.450 | 75.475 | 46.575 | -69.675 | -249.800 |
| 0.7 | 49.450 | 92.050 | 100.350 | 32.625 | -126.375 |
| 0.8 | 49.950 | 98.625 | 135.525 | 121.475 | 4.950 |
| 0.9 | 50.000 | 99.950 | 147.600 | 175.850 | 122.275 |
| 1.0 | 50.000 | 100.000 | 150.000 | 200.000 | 250.000 |

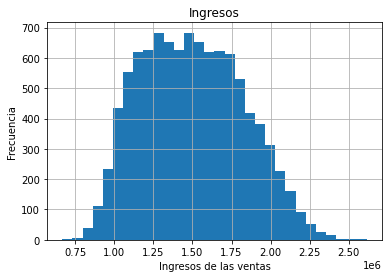
Se puede evidenciar como solo desde la probabilidad de un 30% se evidencia un escenario con alguna ganancia para el vendedor, lo cual es un comportamiento esperado, ya que va a vender mayor porcentaje de los periódicos. Si se llegará a tener una probabilidad de éxito del 70%, se le aconsejaría comprar 3 periódicos para vender y poder generar una ganancia promedio de $100.35. Adicionalmente, se puede ver la siguiente gráfica,



Donde muestra como es el comportamiento de las medias de las simulaciones para cada una de las probabilidades evaluadas y se evidencia que cuando se tiene una probabilidad del 100%, lo mejor es comprar los 5 periódicos porque si o si los 5 periódicos se van a vender. En el resto de los casos, no sería recomendable la compra de 5 periódicos y cuando la probabilidad es inferior toma los valores de un 0% a un 20% lo mejor es no comprar periódicos, ya que no sería rentable en ninguno de los casos analizados.

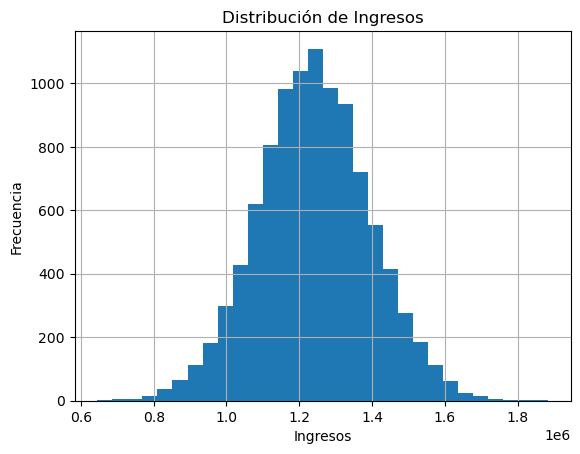
1. **Modelo de venta de artículos al por mayor.**
2. Modifique el modelo para que el precio se distribuya normal con media 80 y desviación estándar 10 y cada venta consista de entre 15000 y 16000 unidades.
3. Genere resultados para el nuevo escenario y compare con el anterior a través de una diferencia de medias (intervalo de confianza y/o prueba de hipótesis).

Para poder comparar los dos escenarios, se va a examinar el primer escenario de manera individual. Donde el primer escenario corresponde a un modelo donde se utiliza un modelo para que el precio se distribuya normal con media 100 y desviación estándar 10 y cada venta consista de entre 10000 y 20000 unidades. Generando la siguiente distribución de los ingresos.

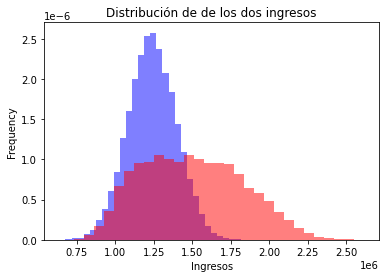


Donde el primer escenario tiene la distribución como se muestra en la gráfica anterior, con una media cercana a 1499588 y una desviación estándar de 325823.4. El intervalo de confianza para la media (1493200.8960353208, 1505974.4830555713).

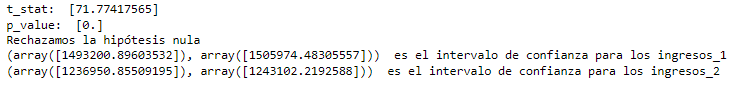
Al generar las modificaciones solicitadas por el enunciado, donde el modelo para que el precio se distribuya normal con media 80 y desviación estándar 10 y cada venta consista de entre 15000 y 16000 unidades. Se puede evidenciar que la distribución que siguen los ingresos en este caso sí cambia la distribución y es comporta como la gráfica que se presenta a continuación.



Se evidencia, que es un comportamiento que presenta una curtosis positiva o leptocúrtica, indicando que los valores están más centrados en un solo valor y no como la distribución anterior. Esta distribución tiene una media cercana a 1240027 y una desviación estándar de 156906.4. Los intervalos de confianza para la media de esta distribución están entre (1236950.8550919467, 1243102.2192588015). Para saber si nuestras medias son parecidas, se va a mostrar una gráfica donde vemos que las dos variables son bastante diferentes, los datos agrupados en la primera variable están mucho más centrados a la media mientras que en la segunda variable hay mayor dispersión.



Para saber estadísticamente si la diferencia es significativa, se realiza una prueba de medias utilizando el t-test

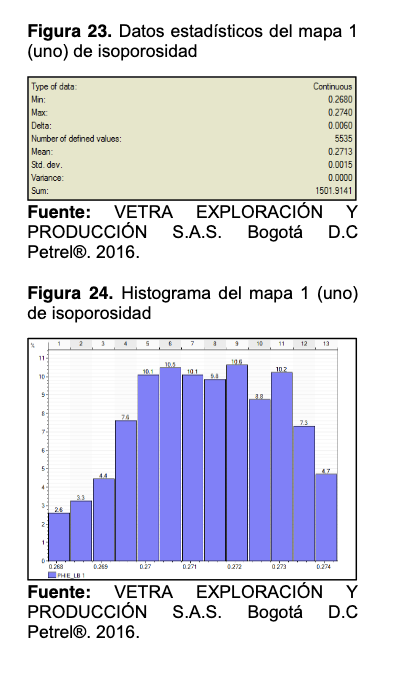


Observamos que el valor de p es significativamente pequeño, lo que nos lleva a concluir que las dos medias no son iguales o similares. Al examinar los intervalos de confianza al 95%, notamos que no se superponen en ningún punto.

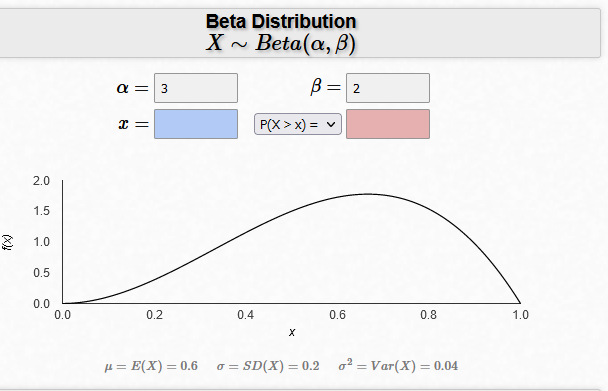
Para saber cuál de las dos distribuciones generaría que el 90% de los ingresos esté sobre el valor más alto, se tiene que verificar el percentil 10 en cada una de las distribuciones y comparar ambos valores. Logrando ver que con la primera distribución en el percentil 10 se tiene un valor de 1078045.34 y en la segunda distribución en el percentil 10 se tiene un valor de 1039613.19. Mostrando que con la primera distribución el 90% de los ingresos tienen un valor por encima de 1078045.34, lo que muestra una distribución un poco menos riesgosa en términos de ingresos para el vendedor, pero con una variabilidad mayor.

1. **Caso** **Macuay**

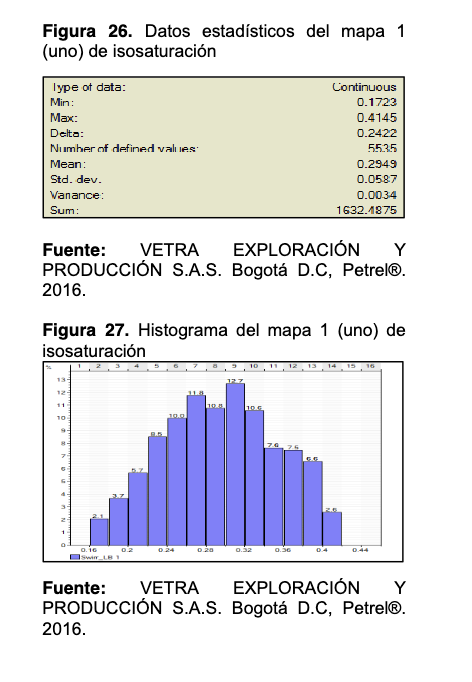
El análisis de entrada se basó en el proyecto de grado citado(<https://repository.uamerica.edu.co/bitstream/20.500.11839/122/1/5111626-2016-2-IP.pdf>), que incluye varias variables de interés relacionadas con el problema. Estas variables están asociadas con sus respectivas distribuciones de probabilidad. El estudio realizado en el proyecto proporciona una base sólida para comprender las características de las variables en cuestión. Para La primera variable que tenemos es la porosidad.



Es una elección adecuada utilizar una distribución beta con parámetros alfa=3 y beta=2 para modelar la variable de porosidad en el rango de 0.26 a 0.27. La distribución beta es una elección común cuando se desea representar datos con asimetría hacia la derecha, ya que, en este caso, los parámetros alfa y beta permiten controlar la forma de la distribución.

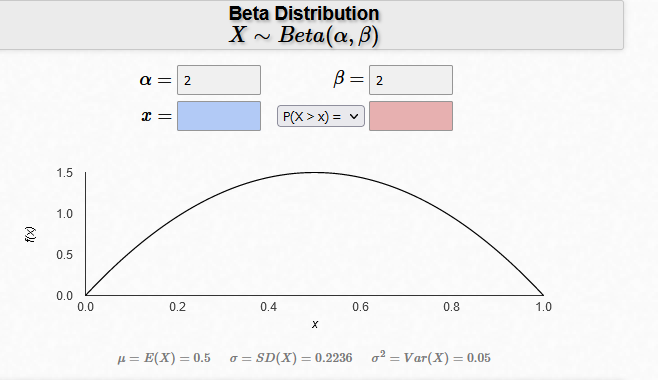


Ahora se tiene la variable de saturación

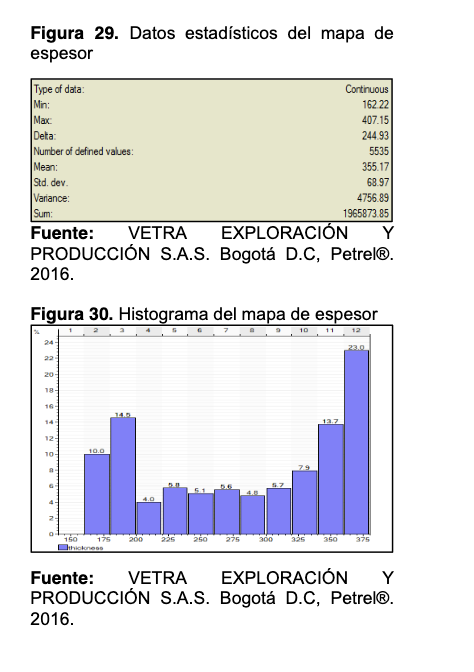


Se observa que esta variable no presenta colas largas ni características pronunciadas en sus extremos. Se considera que sería una propuesta interesante ajustar una distribución beta con parámetros alfa=2 y beta=2 para modelarla.

La elección de esta distribución se justifica por el hecho de que la variable no exhibe colas largas y su comportamiento se asemeja a una parábola, con un aumento y disminución gradual en sus valores.

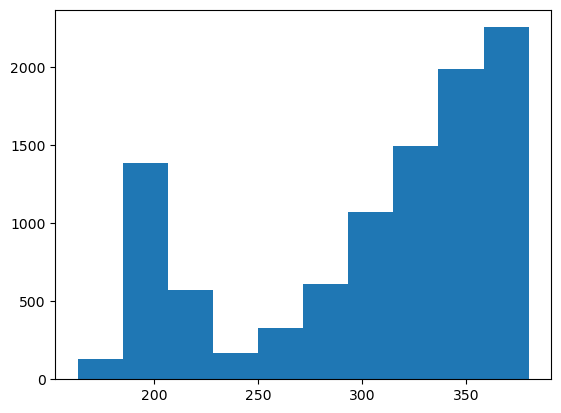


Para la última variable llamada espesor se tiene lo siguiente:



Para esta variable, se observa que presenta dos comportamientos en los datos. Para modelarla, ajustaremos una distribución multinomial que capture ambos comportamientos identificados en los datos. En primer lugar, notamos un comportamiento que se asemeja a una distribución normal con una varianza muy pequeña. Luego, se identifica un segundo comportamiento que se asemeja a una distribución normal, pero con una tendencia a agrupar los datos hacia la derecha, indicando un alto sesgo a la izquierda.

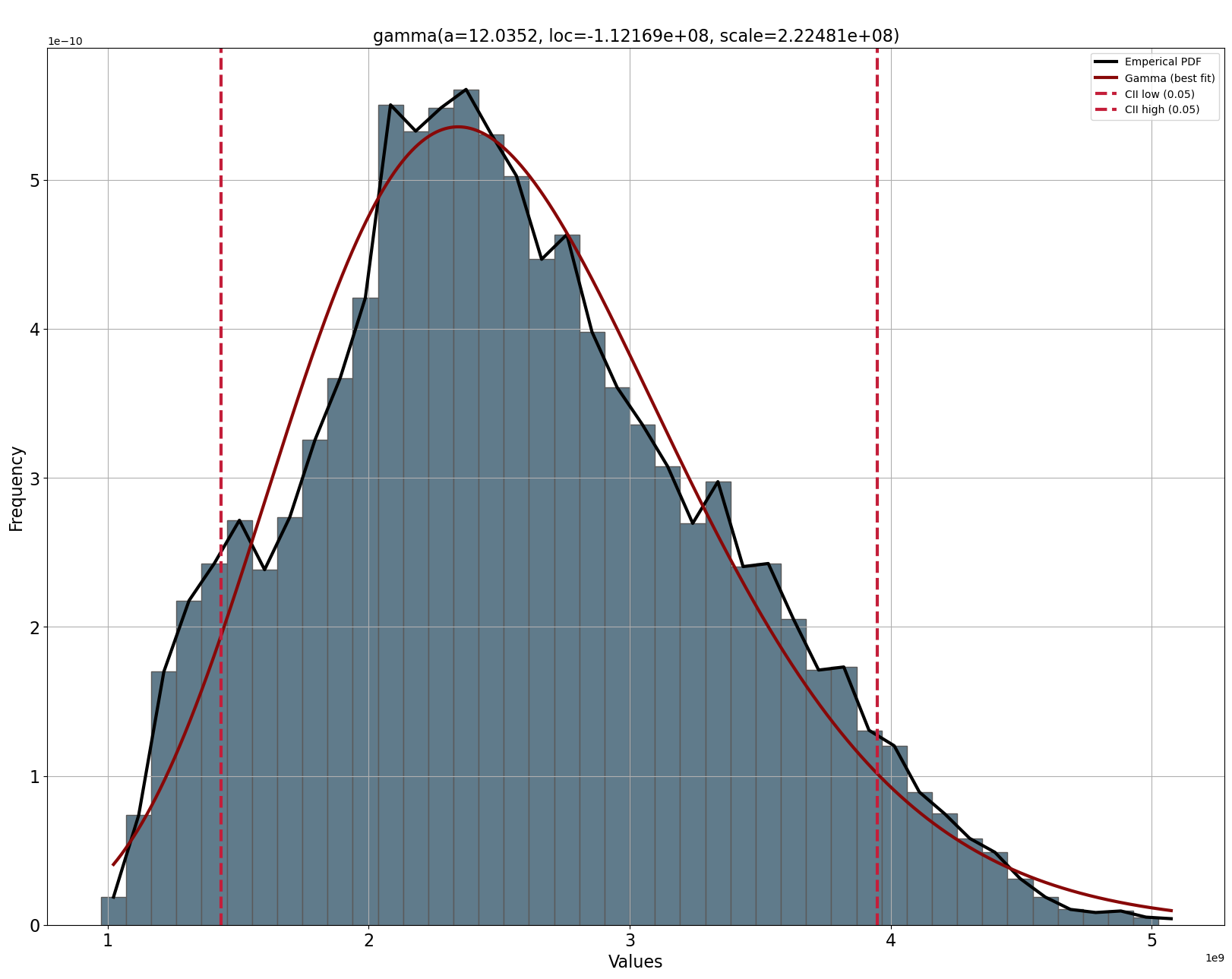
Para explorar y comparar esta hipótesis, se generan dos variables aleatorias en Python y se representan gráficamente. Se observa que estas variables aleatorias se asemejan mucho más a la variable original en los datos, lo que respalda la idea de que una distribución multinomial con dos comportamientos puede ser un modelo adecuado para esta variable.



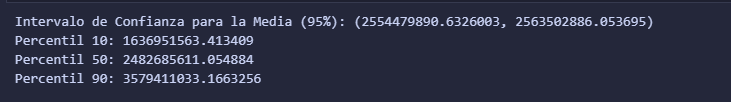
2)

* La variable "a" será una constante.
* La variable "H" se generará como una combinación de dos distribuciones normales. Una de estas distribuciones tendrá una alta concentración alrededor de la media y una desviación estándar baja, mientras que la otra mostrará un sesgo hacia la izquierda, como se mencionó anteriormente.
* La variable "porosidad" se modelará utilizando una distribución beta con parámetros 3 y 2, y los valores estarán en el rango de 0.26 a 0.28.
* La variable "saturación" se generará utilizando una distribución beta con parámetros 2 y 2.
* Las demás variables se tomarán como distribuciones uniformes.

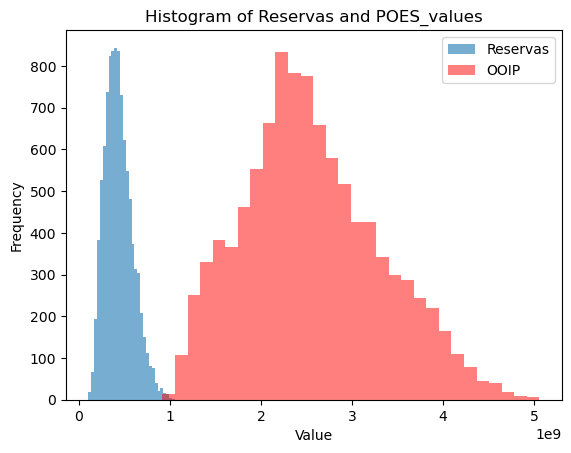
1. Para calcular la distribución de probabilidad, se utilizó el paquete distfit y se probaron distintas distribuciones que se ajusten a los datos.

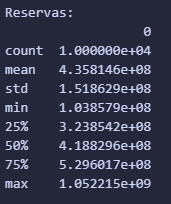


En la gráfica se observa que hay una distribución que se acerca muy bien al comportamiento de los datos, la cual es una variable que distribuye gamma. Para obtener el intervalo de confianza, al 95 porciento se toma la variable gamma y se obtiene lo siguiente:



4)





En resumen, basándose en estos estadísticos descriptivos (De reservas), se puede concluir que las reservas varían considerablemente en términos de magnitud y dispersión. La media y la mediana están en el rango de 400 a 500 millones de unidades de medida, lo que sugiere que estos valores son representativos de las reservas. Sin embargo, la variabilidad es notable, ya que las reservas varían desde aproximadamente 103.9 millones hasta 1,052.2 millones de unidades de medida. La desviación estándar de 151.9 millones indica esta variabilidad en los datos.