Tarea 4 de Modelos Probabilistas Aplicados

Distribución de Poisson

5271

29 de septiembre de 2020

1. Introducción

En este trabajo se presenta un acercamiento a la simulación de variables aleatorias de Poisson a partir de variables aleatorias de distribución Uniforme, Exponencial y Normal. Además se presenta la aproximación de la distribución de Poisson a la distribución Binomial que seguían los largos de palabras de libro "The Adventures of Sherlock Holmes"[2]. Este libro se encuentra disponible en la biblioteca virtual gratuita Project Gutenberg, con el siguiente enlace: https://www.gutenberg.org.

2. Proceso de Poisson

La distribución de Poisson es una distribución de probabilidad discreta que expresa, a partir de una frecuencia de ocurrencia media, la probabilidad de que ocurra un determinado número de eventos durante cierto período de tiempo o espacio. Se centra en la probabilidad de ocurrencia de eventos con probabilidades muy pequeñas. Se especifica por un parámetro lambda (λ). Este parámetro es igual a la media y la varianza de la distribución.

3. Aproximación mediante el uso de variables aleatorias Exponenciales

Los momentos en que se incrementa el proceso de Poisson se denominan tiempos de llegada o tiempos de ocurrencia , ya que en los modelos estocásticos clásicos representan las llegadas o ocurrencias de algo, como llegadas de clientes a la fila de un banco. Las diferencias entre tiempos consecutivos se denominan tiempos entre llegadas. Los tiempos entre llegadas de un proceso de Poisson homogéneo lo forman variables aleatorias exponenciales independientes , un resultado conocido como el Teorema del intervalo . A partir de esta relación se pueden generar variables aleatorias exponenciales $E_1, E_2, ...N_n$ y N es el número entero más pequeño tal que:

$$\sum_{k=1}^{n} E_k > 1 \tag{1}$$

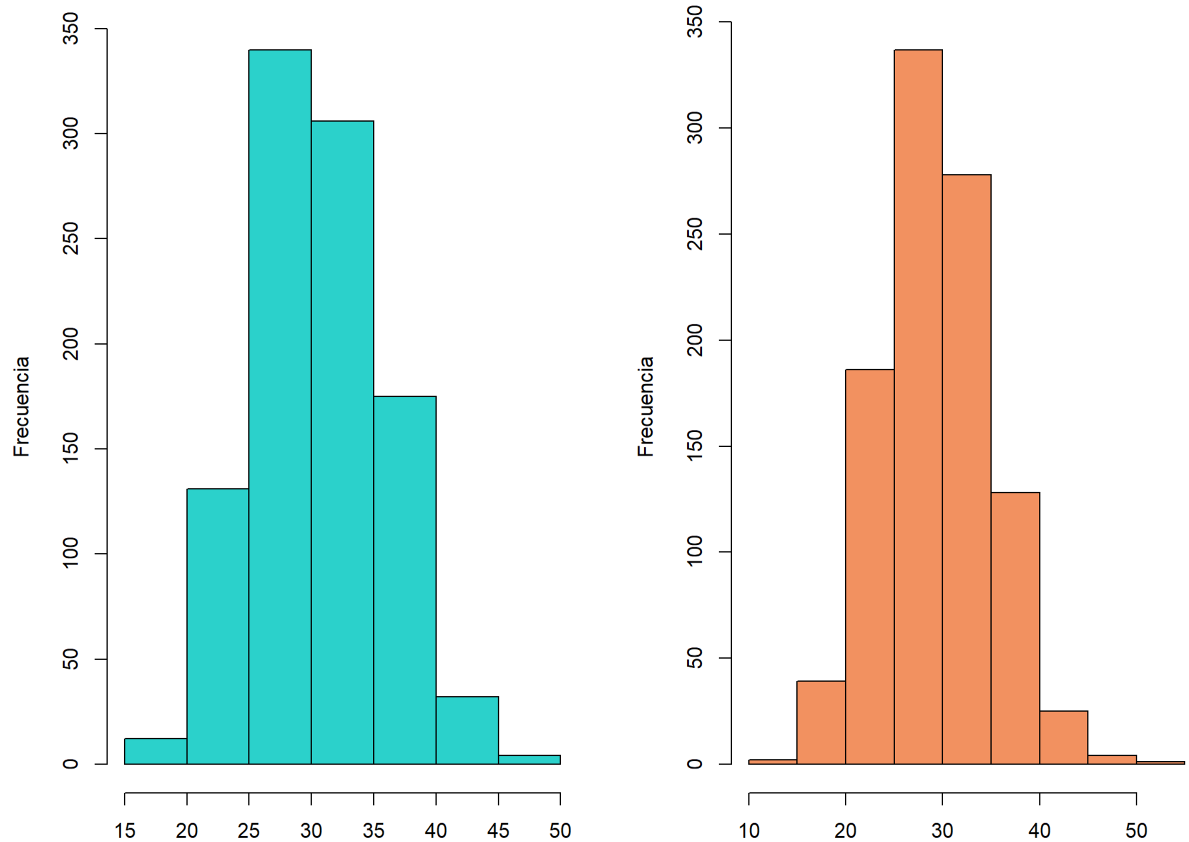

Figura 1: Histogramas de las poblaciones simuladas, a la izquierda la aproximación a partir de la exponencial y a la derecha la creada a partir de la biblioteca *rpois*

Entonces N es Poisson (λ), Lema 3.2 Capítulo 10 [1]

A continuación realizaremos una comparación entre las variables aleatorias de Poisson creadas a partir de la suma de las variables aleatorias exponenciales y las creadas a partir de la biblioteca *rpois*.

En la figura 1 de la página 2 se muestra los histogramas de las dos poblaciones simuladas. En la figura 2 de la página 3, se muestran superpuestos los diagramas de densidad de ambas poblaciones

Para calcular la diferencia entre las distribuciones se realiza el cálculo de la distancia Kolmogorov—-Smirnov, que se define como la distancia vertical máxima entre las funciones de distribución acumulada

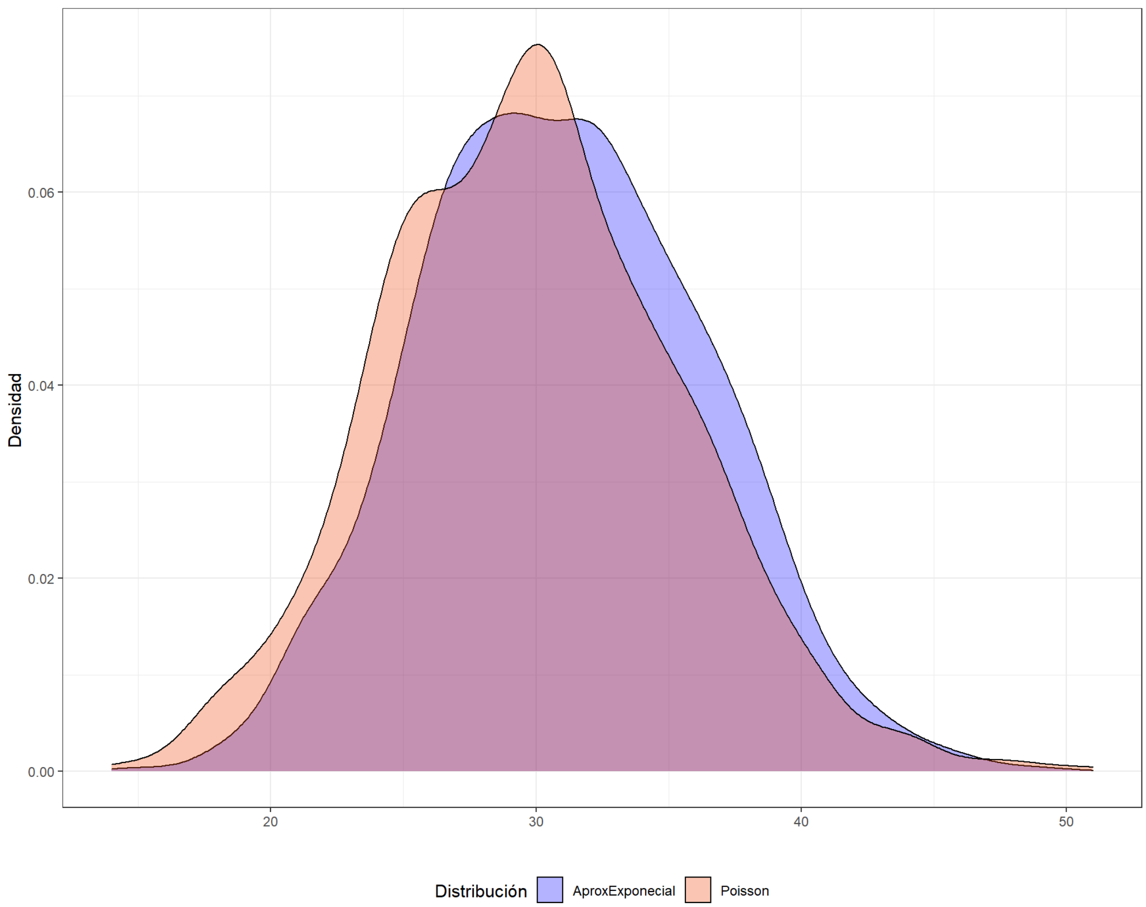

Figura 2: Diagramas de densidad de ambas poblaciones

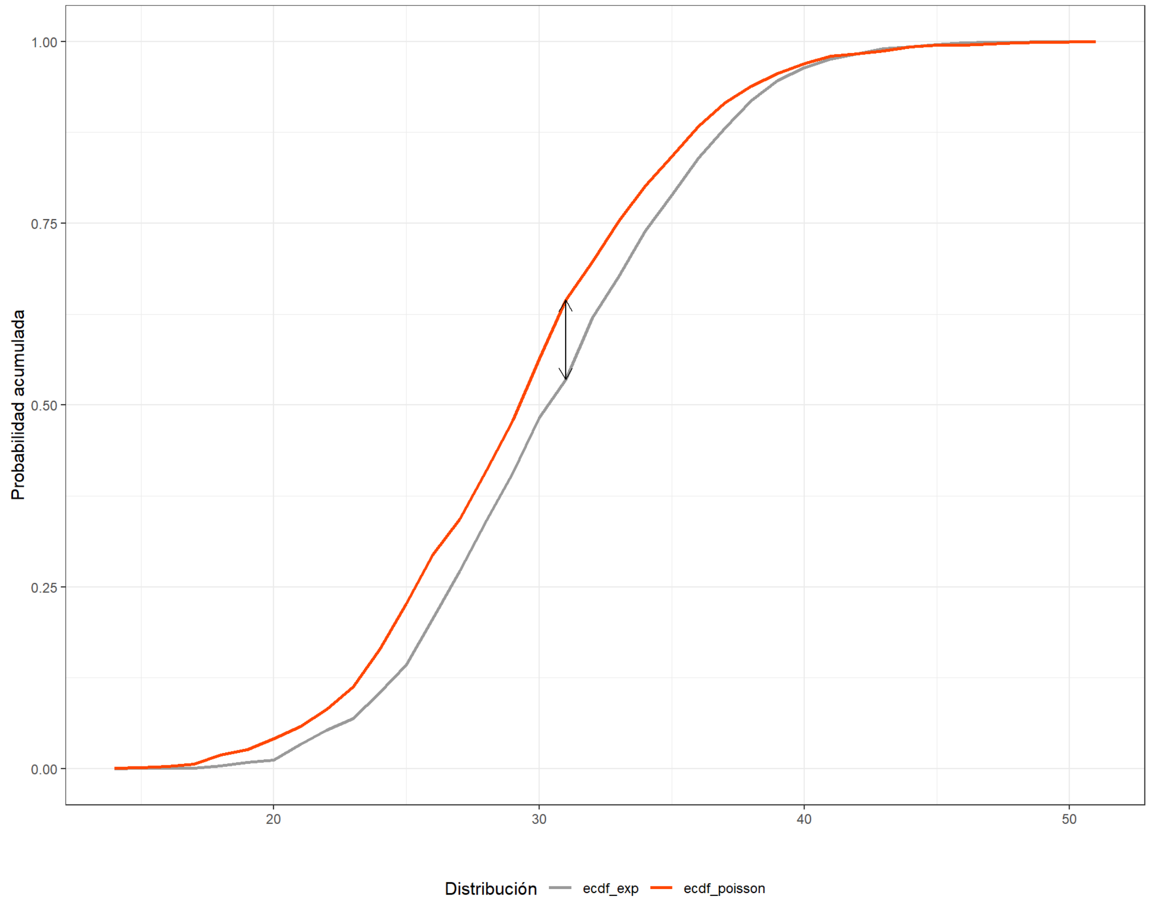

Figura 3: Distancia de Kolmogorov—-Smirnov entre ambas poblaciones

empíricas de dos muestras, donde el valor de la distancia es 0,11. Esto se puede observar en la figura 3 de la página 4.

Para concluir si ambas poblaciones viene de una misma distribución se procede a aplicar la prueba de Cucconi, que es una prueba no paramétrica para comparar conjuntamente la tendencia central y la variabilidad (detectando cambios de ubicación y escala) en dos muestras. Lo s resultados son para el estadístico $C = 12{,}287$ y $p - valor = 0$, se rechaza la hipótesis nula que ambas poblaciones viene de una misma distribución.

3.1. Aproximación mediante el uso de variables aleatorias Exponenciales

Para reducir los cálculos, se reformula el método que utiliza variables aleatorias exponenciales, por el productos de variables aleatorias uniformes, debido a identidades logarítmicas. Se usan variables aleatorias uniformes estándar $U_1, U_2, ...U_n$ y N es el número entero más pequeño tal que:

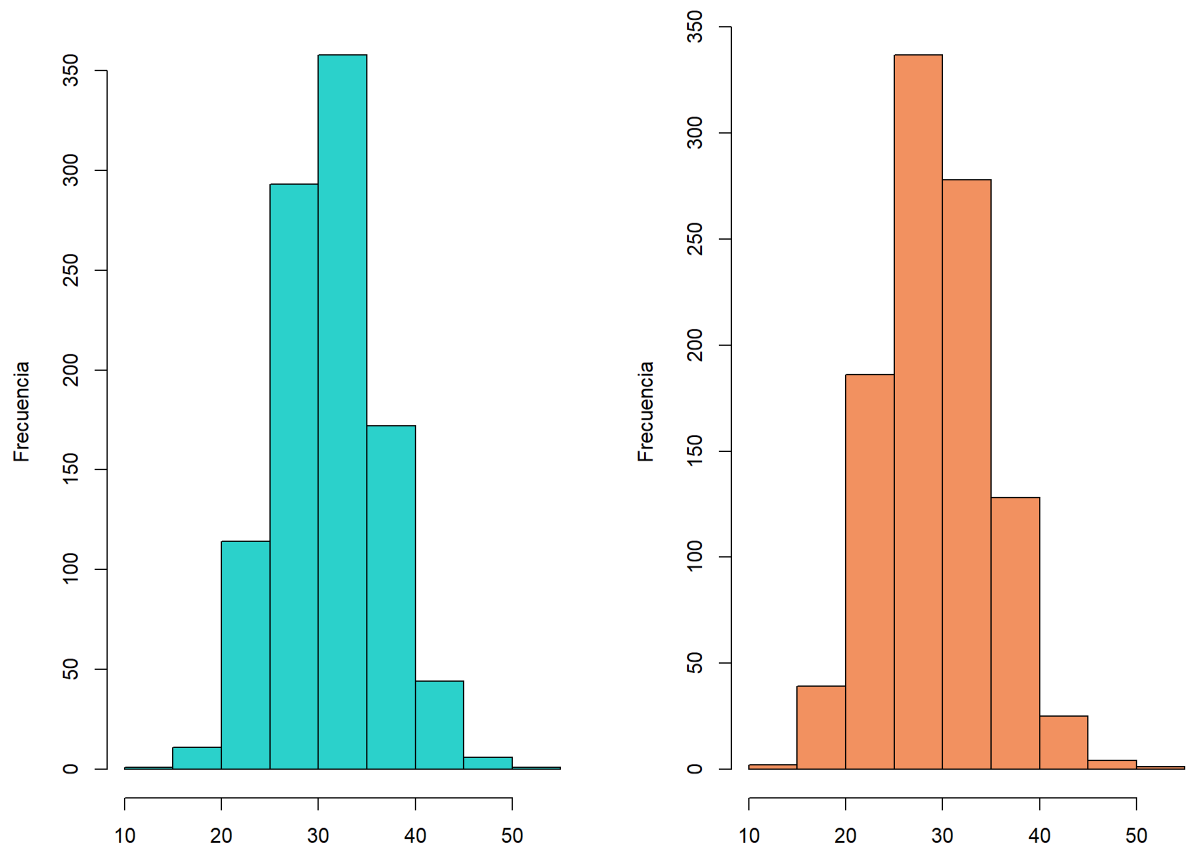

Figura 4: Histogramas de las poblaciones simuladas, a la izquierda la aproximación a partir de la Uniforme y a la derecha la creada a partir de la biblioteca *rpois*

$$\prod_{k=1}^{n} U_k > e^{-\lambda} \tag{2}$$

Entonces N es Poisson (λ), Lema 3.3 Capítulo 10 [1]

A continuación realizaremos una comparación entre las variables aleatorias de Poisson creadas a partir de la suma del producto de variables aleatorias uniformes y las creadas a partir de la biblioteca *rpois*.

En la figura 4 de la página 5 se muestra los histogramas de las dos poblaciones simuladas. En la figura 5 de la página 6, se muestran superpuestos los diagramas de densidad de ambas poblaciones

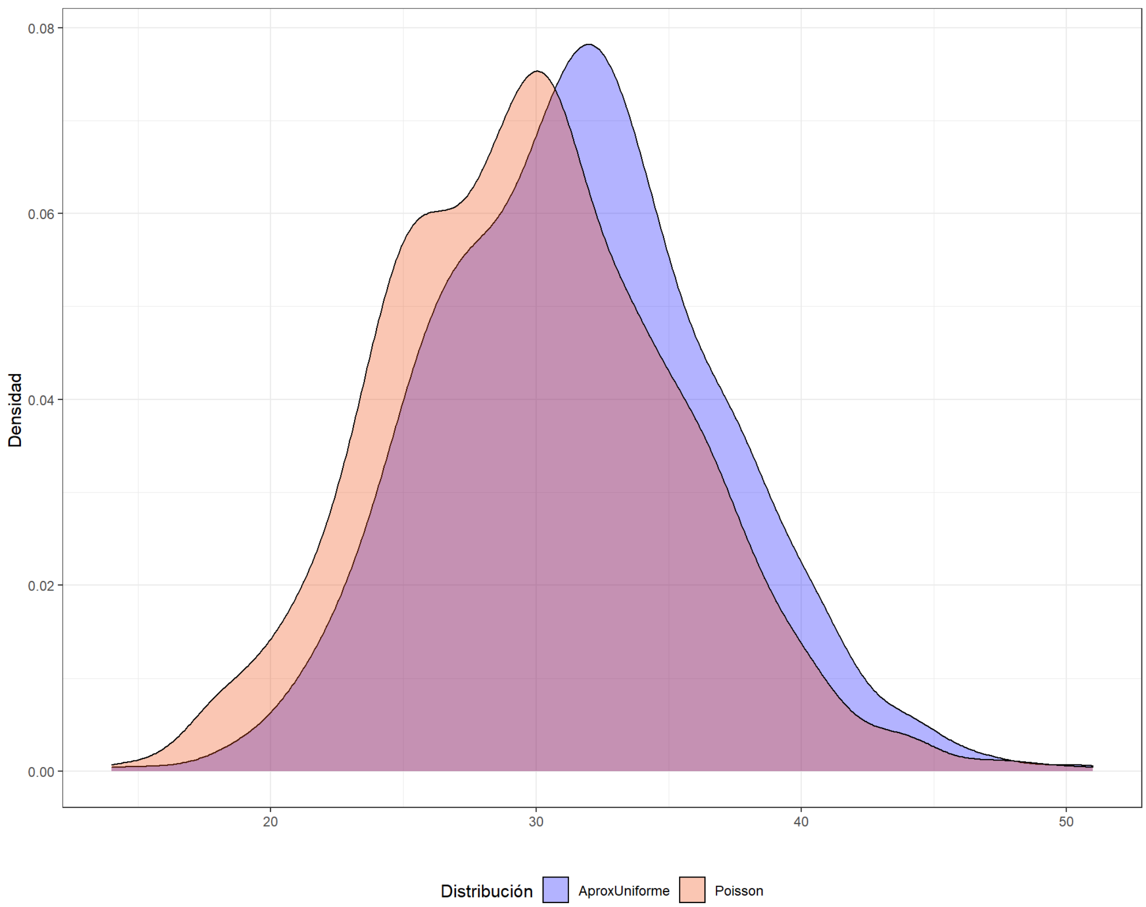

Figura 5: Diagramas de densidad de ambas poblaciones

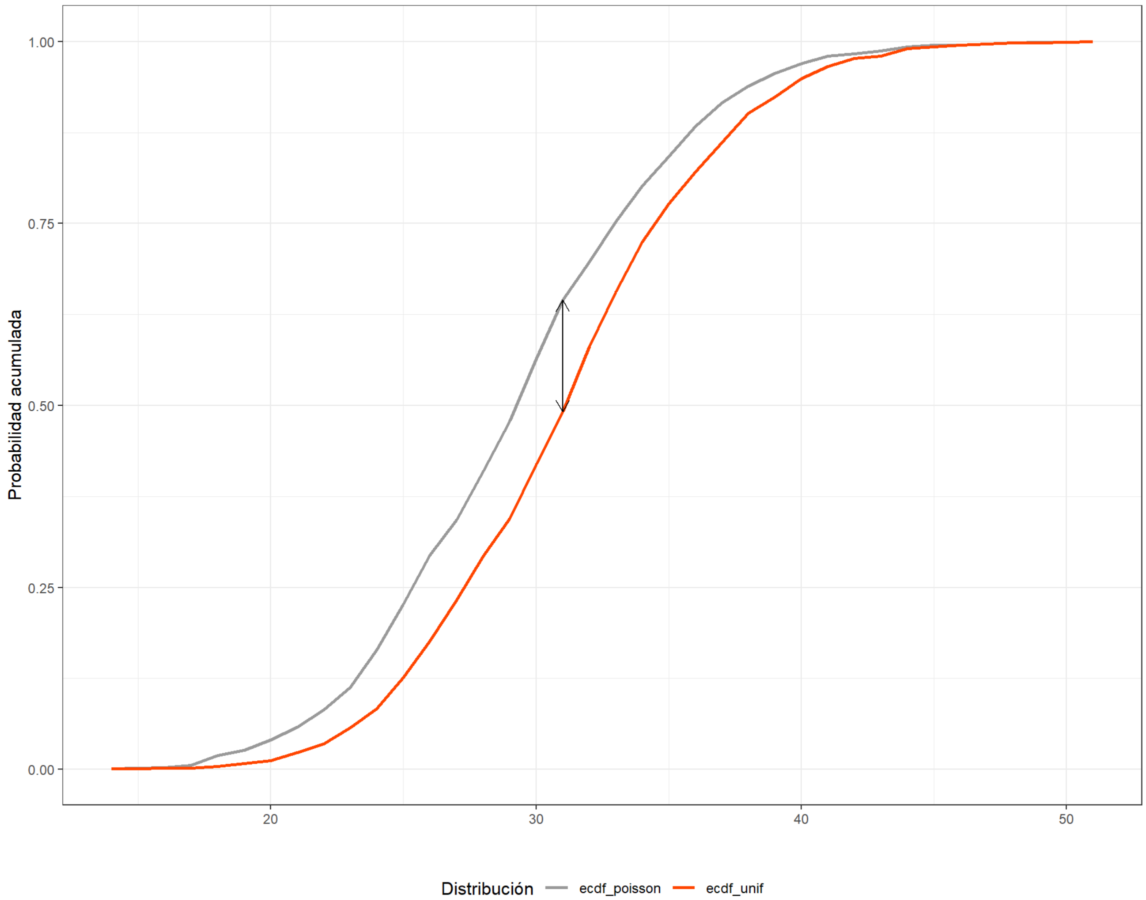

Figura 6: Distancia de Kolmogorov—-Smirnov entre ambas poblaciones

Para calcular la diferencia entre las distribuciones se realiza el cálculo de la distancia Kolmogorov—-Smirnov, que se define como la distancia vertical máxima entre las funciones de distribución acumulada empíricas de dos muestras, donde el valor de la distancia es 0,154. Esto se puede observar en la figura 6 de la página 7.

Para concluir si ambas poblaciones viene de una misma distribución se procede a aplicar la prueba de Cucconi, que es una prueba no paramétrica para comparar conjuntamente la tendencia central y la variabilidad (detectando cambios de ubicación y escala) en dos muestras. Lo s resultados son para el estadístico $C = 27{,}276$ y $p - valor = 0$, se rechaza la hipótesis nula que ambas poblaciones viene de una misma distribución.

El código general se encuentra disponible en el repositorio https://github.com/Albertomnoa/Tareas

Referencias

[1] Luc Devroye. *Non-Uniform Random Variate Generation.* Springer-Verlag, New York, 1986.

[2] Arthur Conan Doyle. *The Adventures of Sherlock Holmes.* George Newnes, United Kingdom, 1892.