Proyecto Final

Fecha de entrega: 30 de mayo de 2023

Debe entregar un informe donde, para cada punto del proyecto, debe incluir:

1. Una introducción del problema.
2. Código y todos los detalles de implementación (valores de tamaño de las muestras, número de iteraciones, valores de parámetros, etc.).
3. Resultados de las simulaciones.
4. Discusión de los resultados y conclusiones.

Problema 1

Este problema busca establecer puebas de hipótesis unilaterales para el problema de localización de de dos muestras para distribuciones con colas pesadas. Recuerde que para este problema se tienen dos muestras independientes $X_i = \theta + \varepsilon_i$ $i = 1, \dots, n$ y $Y_j = \theta + \Delta + \varepsilon_j$, $j = 1, \dots, m$, donde las variables ε son i.i.d con media 0. Las hipótesis son $H_0 : \Delta = 0$ y $H_1 : \Delta > 0$.

1. Realice una prueba basada en remuestreo con el 5 % de significancia con $\Delta = 0$ y $\Delta = 0,3$ (puede tomar $\theta = 0$) donde las variables ε tienen las siguientes distribuciones:

   a) Laplace

   b) t_r, con $r = 2, 10, 50$.

   En ambos casos tome $n = m = 100$.

2. Realice las mismas pruebas anteriores usando el estadístico de Wilcoxon para dos muestras.
3. Realice las mismas pruebas anteriores usando el estadístico t-combinado $T = \dfrac{\overline{X} - \overline{Y}}{S_p\sqrt{1/n + 1/m}}$, donde $S_p^2 = \dfrac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}$. Asuma que T tiene una distribución t_{n+m-2}.
4. Escoja uno de los método anteriores para las realizar la prueba usando los retornos diarios de la acción de Amazon durante los últimos tres meses del año 2021 y los mismos meses del año 2022. ¿Cómo decide qué año corresponde a los datos X?

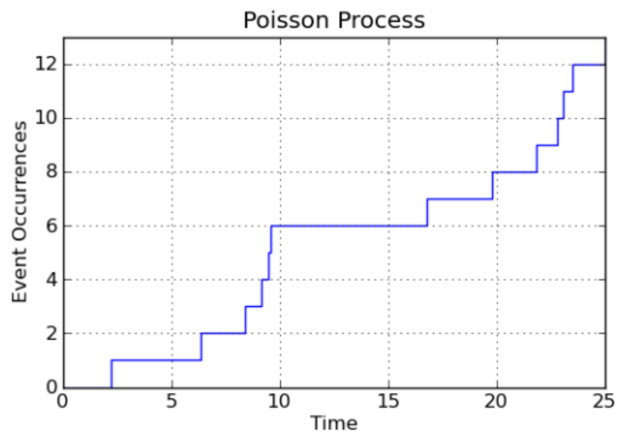
Problema 2

Este problema busca establecer pruebas de hipótesis sobre la homogeneidad o no de un proceso de Poisson, y para el caso de procesos homogéneos, una prueba sobre igualdad o no de sus intensidades. Para esto vamos a dar las definiciones y los teoremas necesarios.

Definición 1 *Se dice que el proceso $\{N(t)\}_{t\geq 0}$ es un proceso de Poisson homogéneo con intensidad $\lambda > 0$, o $PPH(\lambda)$, si:*

- $N(0) = 0$.
- *Los intervalos $N(t) - N(s)$, para $s < t$ son independientes y estacionarios; es decir que si $0 = t_0 < t_1 \ldots < t_n \leq T$, entonces las variables aleatorias $N(t_i) - N(t_{i-1})$ con $i = 1 \ldots n$ son independientes y su distribución depende únicamente del tamaño del intervalo.*
- *El número de eventos en un intervalo de longitud T sigue una distribución de Poisson con media $T\lambda$.*

Un proceso de Poisson se usa por ejemplo para modelar la ocurrencia de eventos como por ejemplo las llamadas a un *call center* o terremotos en alguna región determinada.



Unas propiedades importantes de los $PPH(\lambda)$ están dadas por el siguiente teorema.

Teorema 2 *Sea $\{N(t)\}_{t\geq 0}$ un $PPH(\lambda)$ y sea $T_1, T_2, \ldots$ los tiempos de llegada del proceso. Entonces*

(i) *Las variables S_i, $i = 1 \ldots$, con $S_i = T_i - T_{i-1}$ y $T_0 = 0$, son independientes y siguen una distribución exponencial con media $1/\lambda$.*

(ii) *Condicionado a $N(T) = n$, los tiempos de llegada $T_1, T_2, \ldots, T_n$ tienen la misma distribución que los estadísticos de orden correspondientes a n variables aleatorias uniformemente distribuidas en el intervalo $(0, T)$.*

El teorema anterior permite simular los tiempos de llegada de un $PPH(\lambda)$ de dos formas. Una, a partir de una muestra aleatoria de variables exponenciales con media $1/\lambda$. La otra, fijando un tiempo T, generar una variable Poisson con media $T\lambda$, llamada N, y luego generar N variables uniformes $(0, T)$ y ordenarlas de manera ascendente.

1. Escriba un programa que simule un $PPP(\lambda)$ usando la segunda parte del teorema anterior para diferentes valores de λ y T.
2. Con el fin de hacer una prueba de hipótesis sobre la homogeneidad o no de un proceso de Poisson, para cada una de las intensidades anteriores haga una prueba de bondad de ajuste para una distribución uniforme, con una significancia del 5 %, usando:

   a) El estadístico de Pearson.

b) Una prueba suave.

Use los valores $\lambda = 1$ y $T = 20, 50, 100$, para las pruebas anteriores.

Con el fin de probar la efectividad de las anteriores pruebas, vamos a definir procesos de Poisson no homogéneos.

Definición 3 *Sea $\lambda(t) \geq 0, t \geq 0$ una función y $m(t) = \int_0^t \lambda(s)ds$. Un $PPNH$ con intensidad $\lambda(t)$ es un proceso $\{N(t)\}_{t\geq 0}$ tal que*

- $N(0) = 0$.
- *Los intervalos $N(t) - N(s)$, para $s < t$ son independientes.*
- *La variable aleatoria $N(t) - N(s)$, para $s < t$, tiene una distribución de Poisson con media $m(t) - m(s)$.*

Note que si la función λ es constante tenemos un PPH. Una forma de simular un $PPNH$ con intensidad $\lambda(t)$ es la siguiente:

I. Escoja un valor $T > 0$ y defina como λ^* el valor máximo de la función de intensidad en el intervalo $[0, T]$.

II. Genere una variable Poisson con media $T\lambda^*$, llamada N.

III. Genere una muestra aleatoria $(u_1, \ldots, u_N)$ con distribución uniforme $(0, T)$.

IV. Genere una muestra aleatoria $(h_1, \ldots, h_N)$ con distribución uniforme $(0, \lambda^*)$.

V. Si $h_i < \lambda(u_i)$ seleccione u_i, de lo contrario descártelo.

VI. Ordene los datos seleccionados de manera ascendente.

3. Escriba un programa que simule $PPNH$ hasta $T = 20, 50, 100$, con intensidades

   a) $\lambda(t) = 1 + 0,02t$

   b) $\lambda(t) = \begin{cases} 1 & \text{si } t \in [0, 20) \cup [40, 60) \cup [80, 100] \\ 1{,}2 & \text{si } t \in [20, 40) \cup [60, 80) \end{cases}$

4. Haga las pruebas del punto 2. y calcule la potencia de cada una usando 3000 simulaciones de los procesos no homogéneos.

Consideremos ahora el problema de dos muestras para los procesos homogéneos $N^1(t)$ y $N^2(t)$ independientes con intensidades λ_1 y λ_2 respectivamente. Queremos plantear una prueba para las hipótesis $H_0 : \lambda_1 = \lambda_2$ y $H_1 : \lambda_1 \neq \lambda_2$.

5. Usando el primer punto del Teorema 2 y el ejercicio 7.4 de [K] haga una prueba con $\lambda_1 = \lambda_2 = 1$ y $\lambda_1 = 1, \lambda_2 = 1,2$ para un nivel del 5 %.

6. Usando el ejercicio 7.14b)-d) de [K] haga una prueba con las mismos condiciones anteriores.