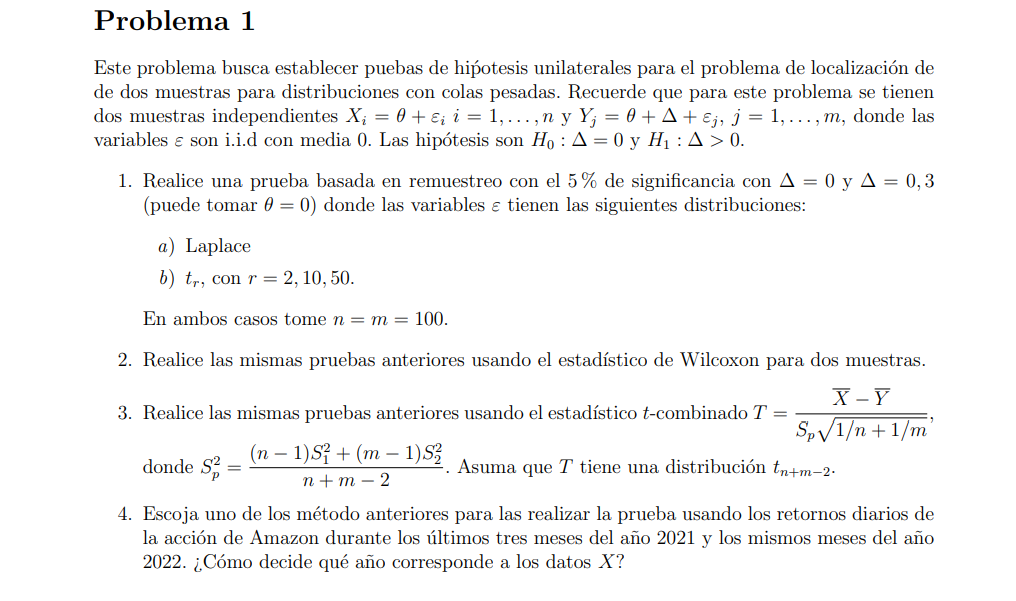
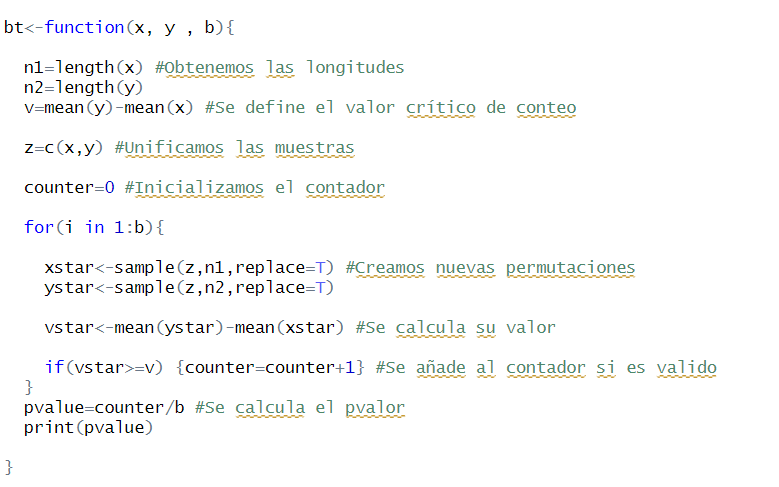
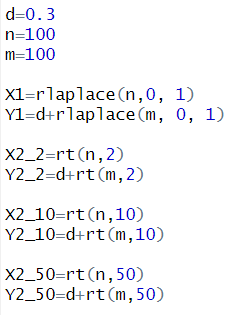
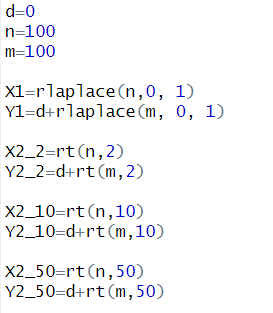
## Codigo: <https://github.com/Gabriel-Sanchez-106/Proyecto_Estadistica.git>

## Problema 1:

Enunciado:



1. El código de Bootstrap empleado fue el siguiente: 

Los resultados obtenidos a partir de los vectores de datos: 

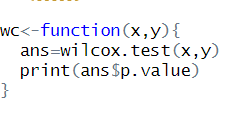
**Resultados:**

Arrojaban p-valores esperados en el caso (d=0) con valores entre 0.2 y 0.9 aceptando siempre la hipótesis nula. En el caso (d=0.3) es interesante que a mayor fuese el grado de libertad en el caso de la distribución t, mejor sería el p-valor para rechazar la hipótesis nula, dando de forma frecuente para el caso de 50 constantemente valores cercanos a 0.01. Aún así, hubo excepciones donde esta prueba fallo al rechazar la hipotesis nula en algunos casos, arrojando p-valores de hasta 0.15. En este punto, cabe mencionar que el número de iteraciones empleadas para las permutaciones fueron (b=15000), ya que entre menos iteraciones mayor era la presencia de p-valores erroneos.

**Discusión y Conclusiones:**

En este caso podemos notar que la prueba es altamente efectiva, sin embargo, está restringida especialmente por el número de permutaciones que se realicen. Además, cabe destacar que es posible inferir una relación inversa entre los grados de libertad de la distribucion t y la probabilidad de un error en el p-valor obtenido. En el caso de la distribución laplace, si bien no se remarca nada especialmente, cuando aumenta el número de muestras tambien aumenta la efectividad de estos p-valores.

Podemos concluir entonces que la efectividad de la prueba depende meramente en el número de permutaciones que se permitan.

1. En este caso, como no se especificó ningún requerimiento de entrada de la prueba se empleó la librería de “stat” con a función “wilcoxon.test(..)” de la siguiente forma: 

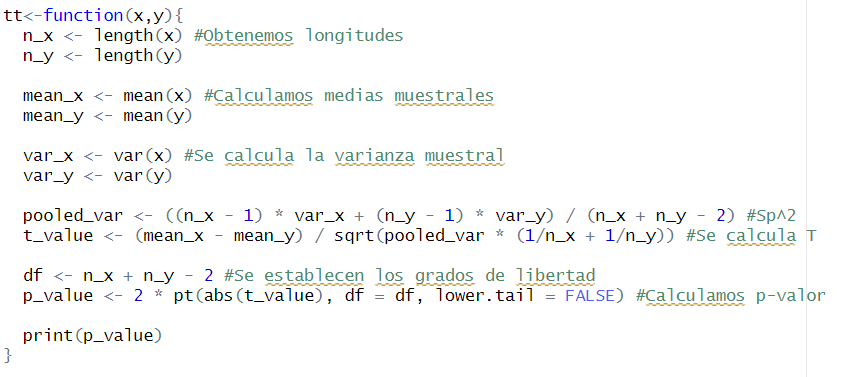
**Resultados:**

Antes que nada, cabe mencionar que en esta prueba se emplearon los mismos vectores de datos que en el inciso anterior. Para el caso (d=0) en la situación de los vectores laplace se obtuvieron valores que permitian aceptar la hipotesis nula con p-valores desde 0.2, increiblemente tambien se obtuvo el p-valor 1. Respecto a los casos de la distribución t, tambien se obtuvieron los p-valores correctos, con valores desde 0.2 a 0.9, cabe mencionar que a mayor fuese el grado de libertad mayores p-valores se obtuvieron consistentemente . Por otra parte, para (d=0.3), en el caso de laplace, se obtenían p-valores que permitian aceptar la hipótesis nula en 1 de cada 5 casos. Respecto a los casos de la distribucion t, para r=2 y 10 tambien se mantuvo esta inconsistencia casi en la misma proporcion. Finalmente, para r=50, tambien hubo casos donde se presentó esta incosistencia, aún así fueron mucho menos frecuentes que en los anteriores, en general se obtuvieron p-valores excelentes con un máximo de 0.02.

**Discusión y Conclusiones:**

En este caso, es evidente que la eficiencia de la prueba no depende de los vectores de muestras y en general parece ser guiada por la cantidad de datos, por lo tanto la única conlusión en este caso es que este test no es viable con el tamaños de muestras pequeños empleados.

1. El codigo empleado para este inciso fue:

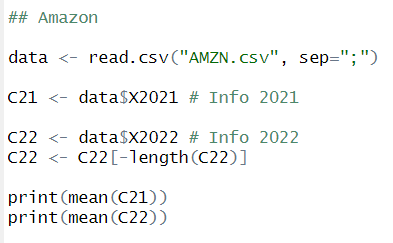


**Resultados:**

En el caso (d=0) se obtuvieron para todos los vectores de datos, los valores esperados, sin embargo, para el caso (d=0.3) en general todos los vectores de datos arrojaban en frecuencias similares p-valores que permitian tanto aceptar como rechazar la hipótesis nula, este margen de error disminuyó a mayores tamaños muestrales.

**Discusion y conclusiones:**

Es posible razonar que este test requiere especialmente más muestras que los demás, ya que al depender del estadístico de varianza, se hace habitual que los 100 datos obtenidos no den una aproximación de la verdadera varianza. Por lo tanto la unica conclusión posible es que el test requiera más muestras para funcionar óptimamente.

1. Los datos de las acciones en los intervalos requeridos vienen en el codigo completo en un archvio llamado “AMZN.csv” para seleccionar cual de los 2 intervalos corresponde a X, como el valor esperado de los epsilon es 0, por hipótesis, solo requeririamos calcular la media muestral de los datos, de la siguiente forma: 

**Resultados:**

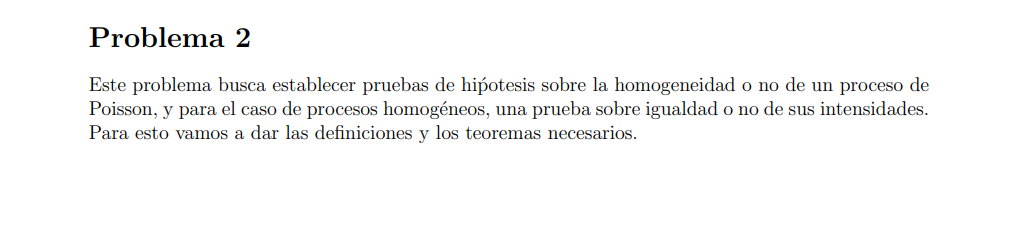
Donde obtenemos que los datos correpondientes a 2022 serían equivalentes a X en el enunciado. Ahora apliquemos la prueba del inciso 2) a los datos, y comprobemos que el hipotetico delta es mayor a 0, ya que se obtiene el p-valor: 2.509298e-22.

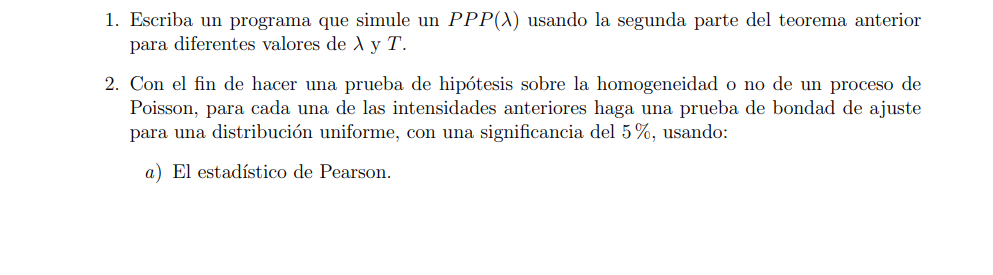
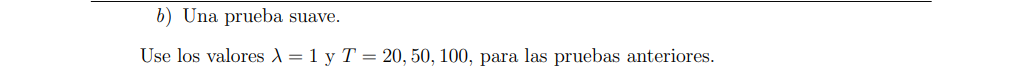
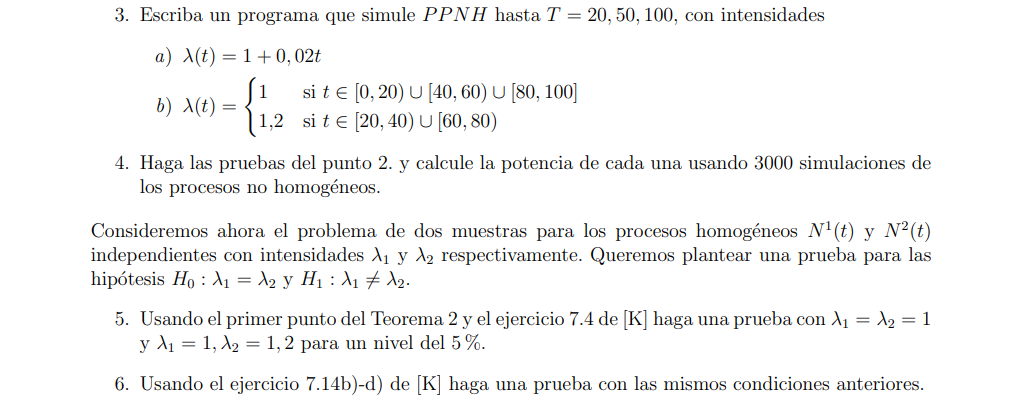
**Discusion y conclusiones:**

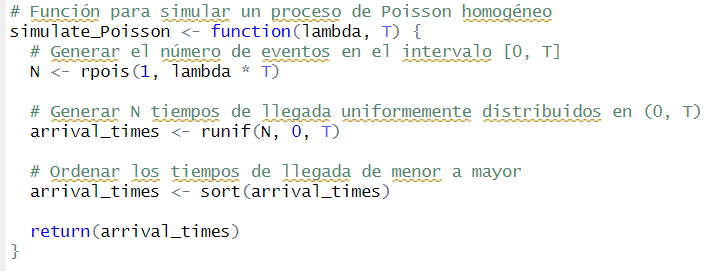
Podemos concluir entonces que los datos no comparten localización.

## Problema 2:

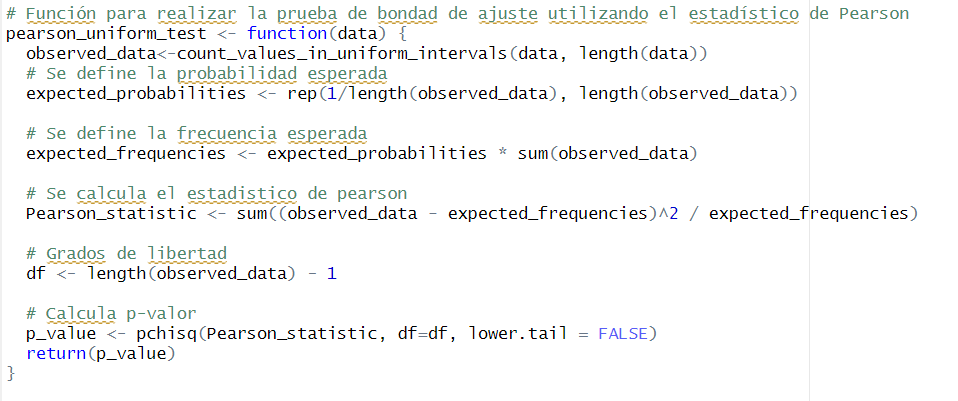
Enunciado:

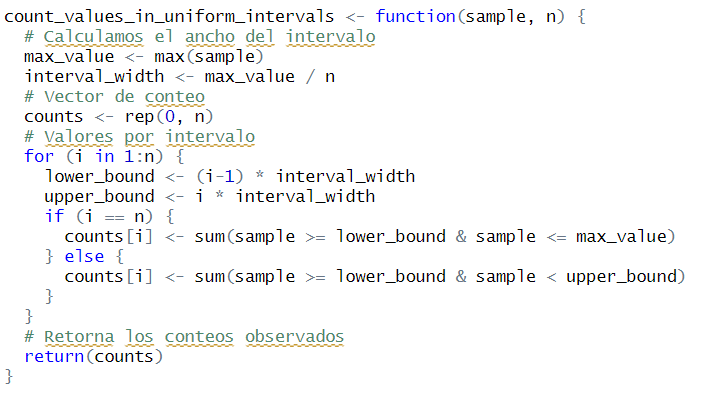


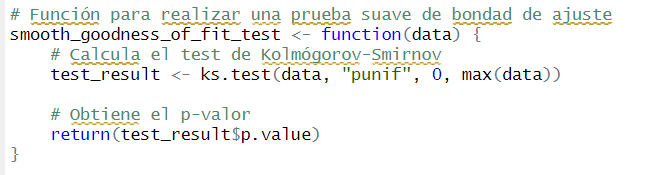
1. Para simular una muestra, se empleará la función: 

(Acá no hay resultados ni conclusiones desde mi perspectiva)

1. A) Para hacer la prueba del estadístico Pearson se empleará el siguiente código: 



B) Para hacer la prueba de bondad suave se usará el siguiente código:

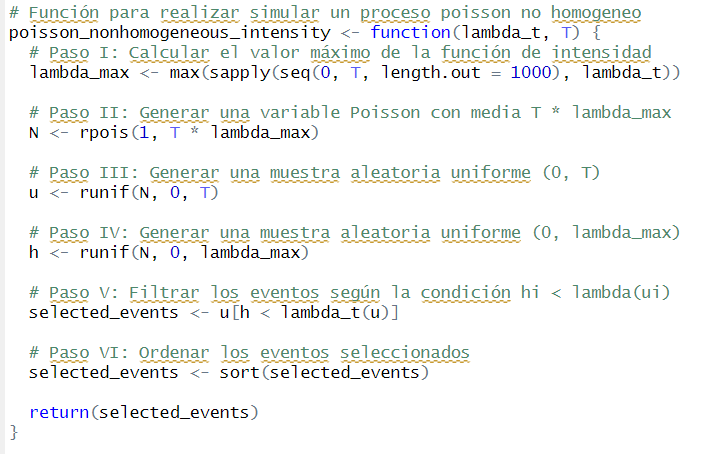


**Resultados:**

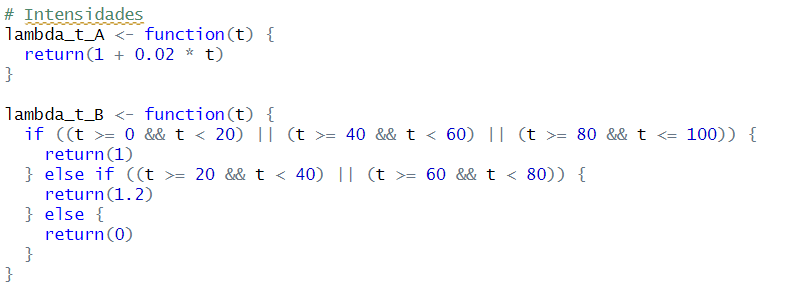
Durante la experimentación con estas pruebas en muestras de procesos Poisson homogéneos con los valores estipulados (20,50,100) se obtuvieron p-valores esperados, sin embargo, hubo cierto margen de error para T=20 y T=50 con la prueba suave, donde se obtuvieron pvalores por debajo del nivel (0.05) en 1 de cada 10 muestras aproximadamente, sin embargo, esta anomalía no ocurrió para T=100

**Discusión y Conclusiones:**

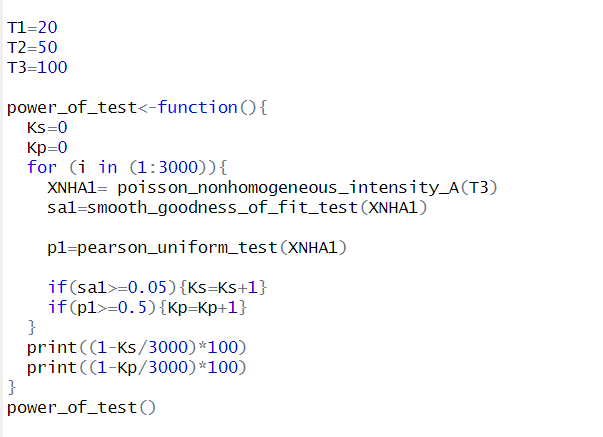
Según lo observado, podemos concluir que el estadístico Pearson es óptimo para tamaños de muestra relativamente pequeños, ya que el caso Kolmogórov-Smirnov solo parece fiable para valores de T grandes.

1. El programa correspondiente siguiendo las instrucciones dada, será: 

Con las funciones de intensidad:



1. **Resultados:**

Se empleó el siguiente código, modificando para examinar el caso de las muestras 3.a o 3.b con distintos tiempos (se creo una funcion auxiliar para cada distribucion no homogenea, es la misma función vista anteriormente, solo que especificando cada caso por un error de lógica presente en R, más información en el codigo): 

Para las muestras 3.a:

Para T=20 y T= 50 se obtuvieron valores bastante bajos de potencia, de entre 50%-60% para el estadístico de Pearson y de menos del 37% para el caso de Kolmogorov-Smirov. Para T=100 sin embargo, la situación fue distinta, ya que la potencia de la prueba suave aumentó a valores entre 96% y 99%, mientras que el caso Pearson quedó situado en 78%

Para las muestras 3.b:

En este caso para todos los valores de T se obtuvieron los mismos rangos de potencia, el estadístico Pearson tuvo entre 47% y 51% de potencia, mientras que el caso Kolmogorov-Smirov siempre tuvo una potencia menor al 6.5%

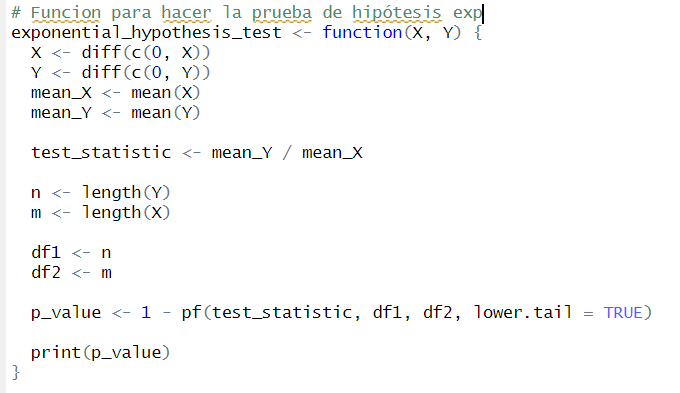
**Discusion y Conclusiones:**

Podemos observar en los porcentajes obtenidos que para las pruebas que en el caso 3.A resulta eficiente emplear el estadístico pearson con muestras relativamente pequeñas, sin embargo, la prueba suave pasa a ser altamente efectiva para muestras más grandes.

Por otra parte para el caso 3.B es evidente que el estadístico Pearson es la mejor elección, sin embargo, este tampoco tiene una potencia alta debido a que la función de intensidad puede parecer quasi-homogenea en algunas situaciones.

Podemos concluir entonces que para el caso 3.A resulta eficiente elegir estadístico de pearson si no hay tantos datos, y la prueba sueva de lo contrario, mientras que en el caso 3.B siempre habría que elegir el estadístico Pearson.

1. Para plantear esta prueba de hipótesis se empleará el siguiente código:

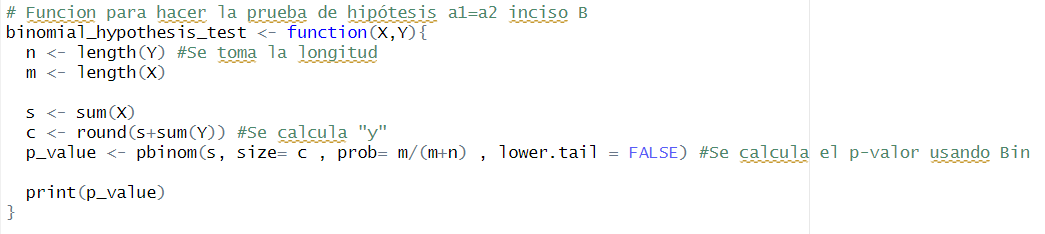


**Resultados:**

Esta prueba funcionó de forma efectiva para aceptar la hipótesis nula, sin embargo, al rechazarla para la segunda situación hubo un margen mínimo.Cabe mencionar que los p-valores tuvieron margenes bastante aleatorios respecto al margen de error es decir si bien eran 1 de cada 5, arrojaban valores cercanos a 0.8, especialmente para T=20

**Discusion y Conclusiones:**

En general, la prueba es util a la hora de probar hipótesis, sin embargo para valores bajos de T la prueba parece tener algún margen de error bastante alto. Se puede concluir que no parece óptimo emplearla, sin embargo si fuese necesario, sería para valores de T grandes.

1. El codigo empleado para esta prueba será: 

**Resultados:**

Esta prueba de hipótesis resultó altamente efectiva para cualquier valor de T, en todas las situaciones aceptó y rechazó la hipótesis nula de forma efectiva. Para T=100 arrojó incluso p= 2.646451e-08 sorprendentemente uno de los valores más bajos que he visto.

**Discusion y Conclusiones:**

Esta prueba de hipótesis resultó altamente efectiva en cualquier caso (e.d cualquier T y Lambda testeado) .Podemos concluir que es extremadamente óptima para realizar pruebas de hipótesis en esta situación.