Práctica 2- Estimación de los parámetros de entrada de colas. Test X^2

Objetivo: Se dispone de una muestra de los tiempos entre llegadas a un S.E. y de los tiempos de servicio del servidor de este S.E. En ambos casos el tamaño de la muestra es de 500 observaciones. Se sabe que corresponden a distribuciones exponenciales de tiempo. Se pretende:

a) Verificar mediante el test de χ^2 que efectivamente corresponden a una distribución exponencial.

b) Obtener los intervalos de confianza para la tasa de llegadas por unidad de tiempo (parámetro λ de la distribución exponencial del proceso de llegadas) y para el factor de carga ρ del S.E.

1. Intervalos de confianza para las tasas λ y μ y para el factor de carga $\rho = \lambda/\mu$.

Se dispone de dos muestras $t_1, t_2, \ldots, t_n$ y $s_1, s_2, \ldots, s_m$ para los procesos de llegada y de servicio (tiempos distribuidos exponencialmente). Se quiere encontrar un intervalo de confianza de probabilidad 1-α para las tasas de llegada λ y de servicio μ a partir de las dos muestras.

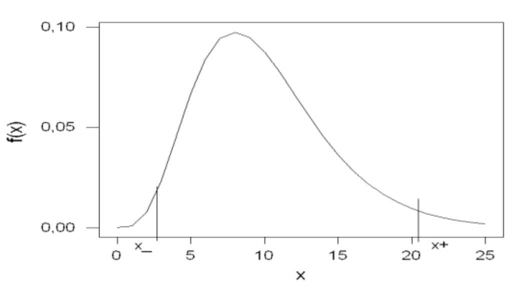
El estimador máximo verosímil para λ y μ es:

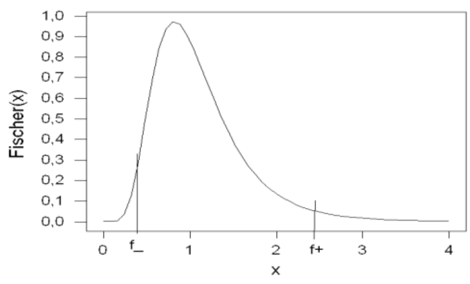
$$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} t_i} = \frac{n}{T_n}, \quad \hat{\mu} = \frac{m}{\sum_{i=1}^{m} s_i} = \frac{m}{S_m}$$

Dado que T_n se distribuye según una ley n-Erlang de parámetro $\theta = \lambda/k$ (o también una Gamma(λ, n)),

$$E\left[2\lambda \sum_{i=1}^{n} t_i\right] = E\left[2\lambda T_n\right] = 2n \Rightarrow 2\lambda T_n \sim Gamma\left(\frac{1}{2}, n\right) = \chi^2_{2n}$$

$$\boxed{2n\frac{\lambda}{\hat{\lambda}} \sim \chi^2_{2n}, \quad 2m\frac{\mu}{\hat{\mu}} \sim \chi^2_{2m} \quad \Rightarrow \quad \frac{\lambda/\hat{\lambda}}{\mu/\hat{\mu}} = \frac{\rho}{\hat{\rho}} \sim F_{2n,2m}}$$





Intervalo de confianza al 1-α para λ : $\left[\frac{\hat{\lambda}}{2n}x_- \;,\; \frac{\hat{\lambda}}{2n}x_+\right]$

Intervalo de confianza al 1-α para ρ : $\left[\hat{\rho}\, f_- \;,\; \hat{\rho}\, f_+\right]$

La siguiente tabla ilustra los tamaños de muestra necesarios para obtener intervalos de confianza del 95% y la amplitud de los mismos.

n=m	***f-***	***f+***	***e_f(%)***	***x-/2n***	***x_+/2n***	***e_x(%) .***
10	0,405	2,461	143	0,479	1,708	112
100	0,757	1,321	54	0,813	1,205	38
1000	0,916	1,091	17	0,939	1,062	12
10000	0,972	1,028	5	0,980	1,019	3

2. Test de Bondad de Ajuste de χ^2

Utilizad la macro "x2.mtb" para efectuar un test de bondad de ajuste de χ^2 a una distribución k-Erlang a partir de una muestra de tiempos para los procesos de llegadas y/o de servicios a un S.E.

Dada una muestra $t_1, t_2, \ldots, t_n$ para la que se quiera comprobar su bondad de ajuste a una distribución k-Erlang de parámetros: k etapas y 1/λ=tiempo medio por etapa=E[t]/k (o también Gamma(λ,k)), el test de χ^2 procede como sigue:

a) Se fijan un conjunto de subintervalos $[x_i, x_{i+1}]$ en número total N que recubren todo el intervalo de valores posibles para la v.a. t de forma que $P(x_i \le t \le x_{i+1}) = 1/N$. De esta forma el número esperado de elementos de la muestra que deberían estar comprendidos en cada subintervalo debería ser constante: $n_e = n/N$. **Pueden ser 10 subintervalos o más.**

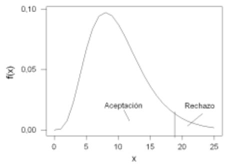
b) Se cuentan el número de elementos n_i de la muestra que caen en cada subintervalo i.

c) Se calcula una medida global de la discrepancia entre n_i y n_e :

$$X^2 = \sum_{i=1}^{N} \frac{(n_i - n_e)^2}{n_e}$$

La variable X^2 se distribuye según una ley χ^2_{N-m-1}, siendo m el número de parámetros de la distribución que hayan sido estimados utilizando la misma muestra.

Se rechazará la distribución propuesta si $P(x \geq X^2) = p\text{-}valor < \alpha$



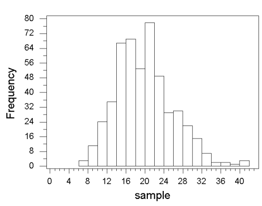
macro chi2test.R

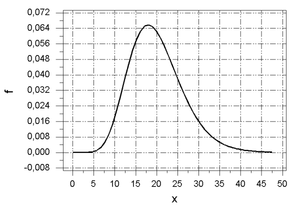
Test X2 per una mostra provinent d'una distr. continua (p.ex. la exponencial)

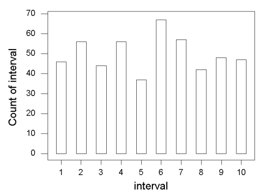
```
sample<-rgamma(240,shape=1,scale=16) # mostra artificialment. 240 elements.
summary(sample) # Es treuen estadístiques bàsiques
mm<-mean(sample) # mitjana
samplerate<-1/mm # prové d'una exponencial convé calcular el seu paràmetre
par(mfrow=c(1,2)) # Es configura com serà el display. En aquest cas serà dues columnes 1 fila
hist(sample,freq=FALSE,breaks=25,col="yellow",main="Equal interval length ... ") # opcio freq=FALSE canvia
l'escala y a densitats i no a freqüencies.
curve(dgamma(x,shape=1,scale=16.398179),col=2,add=T) # dibuixa la distribució contra la que es fa el test
sequence <- seq(0,1,by=0.04) # Crea equència de punts de tall a [0,1] amb distància entre punts de 0.04
# H0: Exponencial lambda = 1/16
perdist <- qexp(sequence,rate=1/16) ; perdist[26]<-80 # Crea la sequència de punts de tall sobre el eix t,
dsample <- cut(sample,breaks=perdist,include.lowest = T)
table(dsample)
summary(dsample)
iobs <- as.vector(table(dsample)) # Obs in the groups defined by nb intervals-percentiles
pexp <- as.vector(rep(1/25,25)) # Expected probability for each groups: sample size/nb intervals
iexp <- 240*pexp # VALOR ESPERAT D'OBSERVACIONS
iobs # VALOR DE Nº D'OBSERVACIONS EN CADA INTERVAL
X2 <- sum(((iobs-iexp)^2)/iexp);X2 # CÀLCUL DE L'ESTADISTIC DE X2
1-pchisq(X2,23) # pvalue = P(H0)=P(Shiq(23)>X2)
chisq.test( iobs,p=pexp ) # EXEMPLE QUE SI FUNCONA
barplot(table(dsample),main="Observed counts",col="green") # ES DIBUIXA UN DIAGRAMA DE BARRES
abline(h=240/25,col="red",lwd=2,lty=2)
```

Proporciona:

- El p-valor en la constante y el resto de constantes bajo las que ha ejecutado.
- El valor de X^2
- Gráficos con el histograma de la muestra, la función de densidad de probabilidad para la distribución y el diagrama de barras para las frecuencias n_i.







Práctica 2. CUESTIONARIO.

Nombre y Apellidos 1: **Curso:**
Nombre y Apellidos 2: **Fecha:**

Fichero para el proceso de llegadas y servicios:

1. Para los ficheros Ann.txt y Smm.txt que se os ha asignado (contienen muestras de tiempos entre llegadas y tiempos de servicio, respectivamente) reportad sus estadísticas básicas. En particular también:

$\boxed{\hat{\lambda} = \qquad}$, $\boxed{\hat{\mu} = \qquad}$, $\boxed{\hat{\rho} = \frac{\hat{\lambda}}{\hat{\mu}} = \qquad}$

2. Calculad el intervalo de confianza del 95% para μ, λ y para ρ, basándose en la tabla del apartado 2 (proceder por interpolación o usad un paquete de software (R, Minitab, …) para los estadísticos f_+, f_-, x_+, x_- para el tamaño muestral con el que trabajéis.

3. Adaptad la macro chi2test.R que se os proporciona para efectuar el test de χ^2 sobre las muestras de los tiempos del proceso de llegadas y servicio. Reportad los gráficos que consideréis adecuados, tales como histogramas de las muestras, diagramas de barras de los intervalos del test χ^2 , etc. Puede aceptarse que la muestra proviene de una distribución exponencial?

$\boxed{X^2 = \qquad}$ $\boxed{p - valor = \qquad}$

$\boxed{X^2 = \qquad}$ $\boxed{p - valor = \qquad}$