ESTIMACIÓ DELS PARÀMETRES DE MODELS DE CUES

- **TEST DE χ^2 PER ALS PROCESSOS D'ARRIBADA/SERVEI.**
- **INTERVALS DE CONFIANÇA PER A λ, μ, ρ.**
- **PRÀCTICA 2.**

Pràctica 2

Objectiu: Es disposa de una mostra dels temps entre arribades a un S.E. i dels temps de servei del servidor d'aquest S.E. En ambdós casos la grandària de la mostra es de 500 observacions. Se sap que corresponen a distribucions exponencials de temps. Es pretén:

a) Verificar mitjançant el test de χ^2 que efectivament corresponen a una distribució exponencial.

b) Obtenir els intervals de confiança per a la taxa d'arribades per unitat de temps (paràmetre λ de la distribució exponencial del procés de arribades) i per al factor de càrrega ρ del S.E.

Intervals de confiança per a les taxes λ y μ i per al factor de càrrega $\rho = \lambda/\mu$.

Es disposa de dues mostres $t_1, t_2, \ldots, t_n$ y $s_1, s_2, \ldots, s_m$ per als processos de arribada i de servei (temps distribuïts exponencialment).

Se vol determinar un interval de confiança de probabilitat 1-α per a les taxes d'arribada λ , de servei μ i per al factor de càrrega ρ d'un S.E. M/M/1 partint de les dues mostres.

El estimador màxim versemblant per a λ i μ és:

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

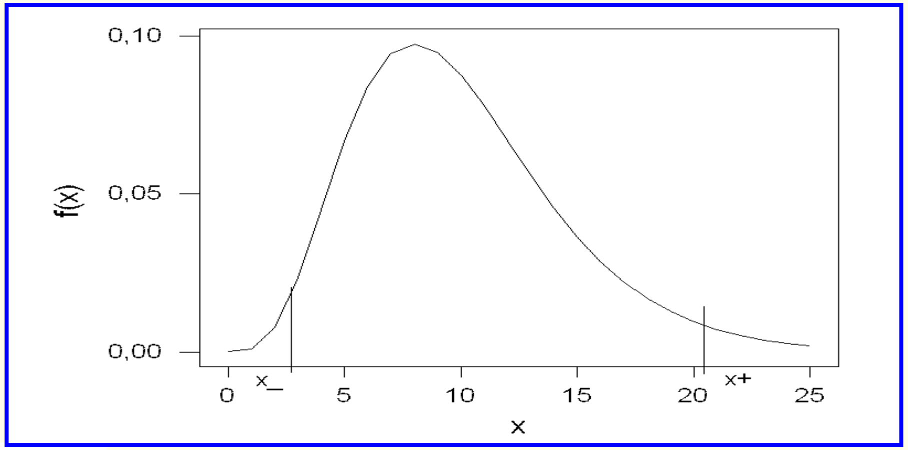
Universitat de Barcelona

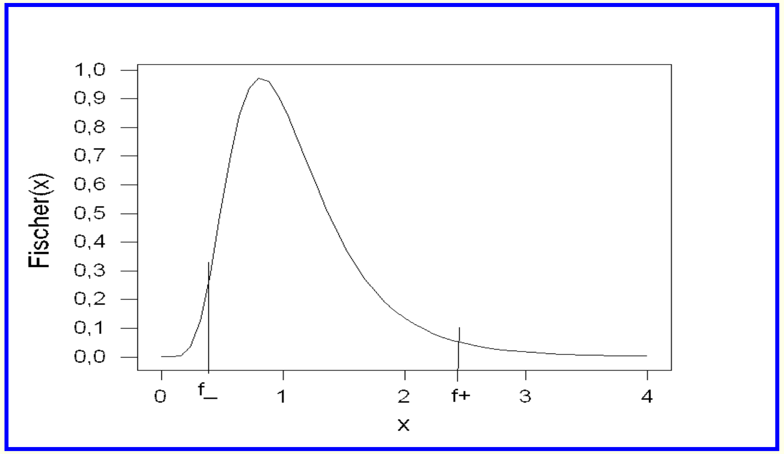
UPC

Ja que T_n es distribueix segons una llei n-Erlang de paràmetre $\theta = \lambda/k$ (o també una Gamma(λ, n)),

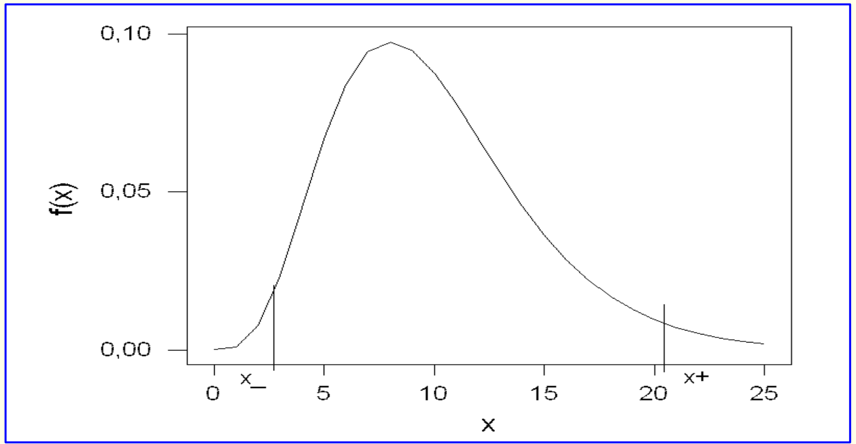
$$E\left[2\lambda \sum_{i=1}^{n} t_i\right] = E\left[2\lambda T_n\right] = 2n \Rightarrow 2\lambda T_n \sim Gamma\left(\frac{1}{2}, n\right) = \chi^2_{2n}$$

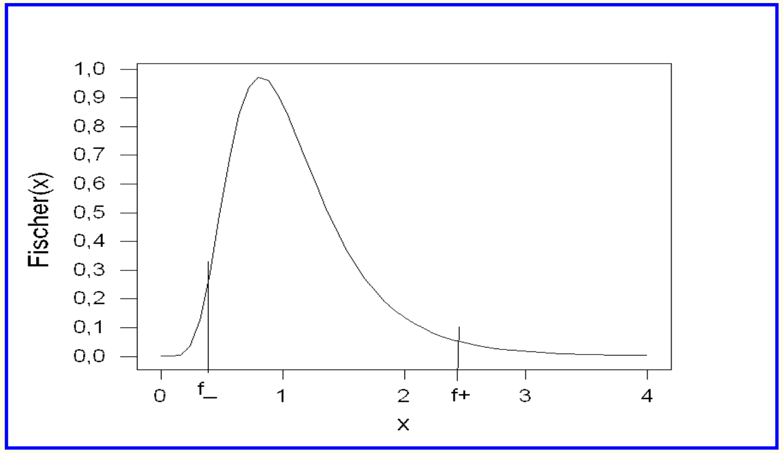
$$\boxed{2n\frac{\lambda}{\hat{\lambda}} \sim \chi^2_{2n}, \quad 2m\frac{\mu}{\hat{\mu}} \sim \chi^2_{2m} \quad \Rightarrow \quad \frac{\lambda/\hat{\lambda}}{\mu/\hat{\mu}} = \frac{\rho}{\hat{\rho}} \sim F_{2n,2m}}$$





Universitat de Barcelona





I.C. al 1-α per a λ: $\left[\frac{\hat{\lambda}}{2n}x_- \ , \ \frac{\hat{\lambda}}{2n}x_+\right]$

I.C. al 1-α per a ρ $\left[\hat{\rho}\, f_- \ , \ \hat{\rho}\, f_+\right]$

Taula amb les grandàries de mostra necessàries per a obtenir un I.C. del 95% i sa amplitud.

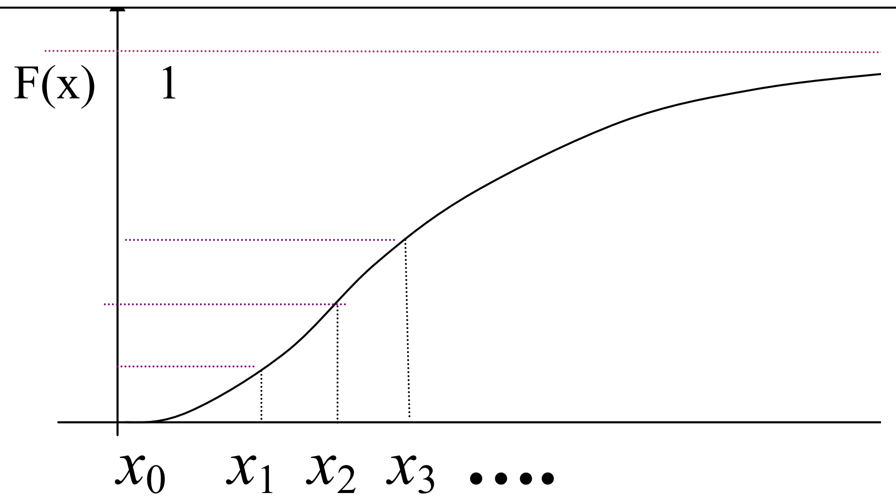
$n=m$	$f-$	$f+$	$e_f(\%)$	$x-/2n$	$x_+/2n$	$e_x(\%)$
10	0,405	2,461	143	0,479	1,708	112
100	0,757	1,321	54	0,813	1,205	38
1000	0,916	1,091	17	0,939	1,062	12
10000	0,972	1,028	5	0,980	1,019	3

TEST DE BONANÇA D'AJUSTAMENT DE χ^2

Es vol comprovar la bondat d'ajustament de una mostra $t_1, t_2, \ldots, t_n$ a una distribució (p.ex: k-Erlang de paràmetres: k etapes i $1/\lambda$=temps mig per etapa=E[t]/k (o Gamma(λ,k)).

a) Es fixen un conjunt de subintervals $[x_i, x_{i+1}]$ en número total N que recobreixen tot l'interval de valors possibles per a la v.a. t de forma que $P(x_i \leq t \leq x_{i+1}) = 1/N$. El número esperat d'elements de la mostra que haurien d'estar compresos en cada subinterval hauria ser constant:

$$n_e = n / N.$$

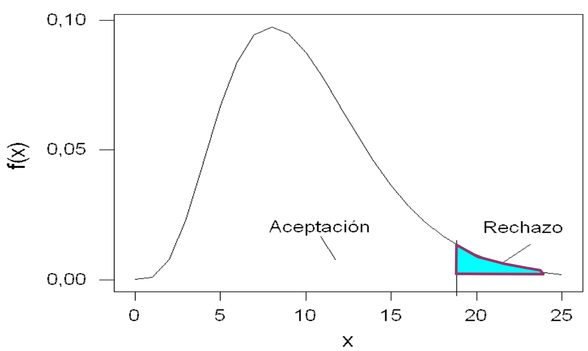


c) Es compte el número d'elements n_i de la mostra que cauen en cada subinterval i.

Es calcula una mesura global de la discrepància entre n_i y n_e :

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

X^2 es distribueix aproximadament segons una llei χ^2_{N-m-1}, sent m el número de paràmetres de la distribució que hagin estat estimats utilitzant la mateixa mostra.



Se rebutjarà la distribució proposada si

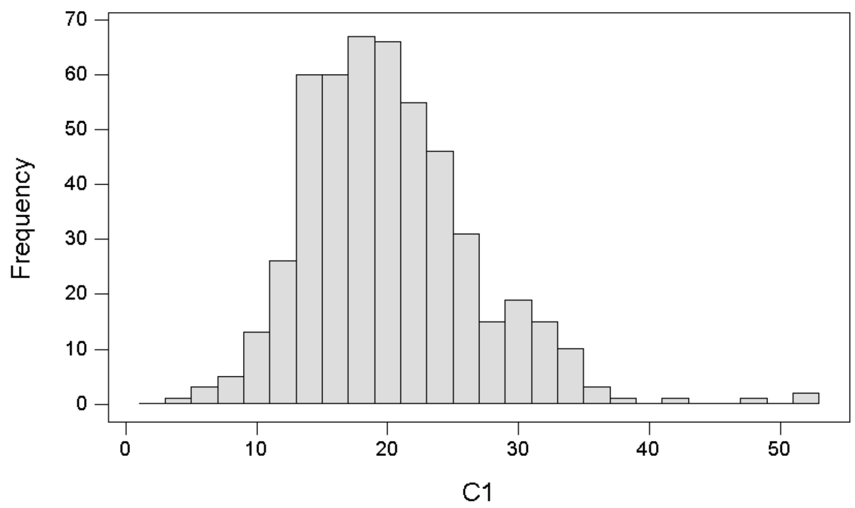
$$P(x \geq X^2) = p\text{-valor} < \alpha$$

Histograma

Suposem, por exemple, que es disposa d'una mostra per a la que les estadístiques bàsiques són:

Variable	N	Mean	Median	TrMean	StDev	SE Mean
sample	500	19,914	19,607	19,702	5,990	0,268

Variable	Minimum	Maximum	Q1	Q3
sample	6,607	41,172	15,660	23,373

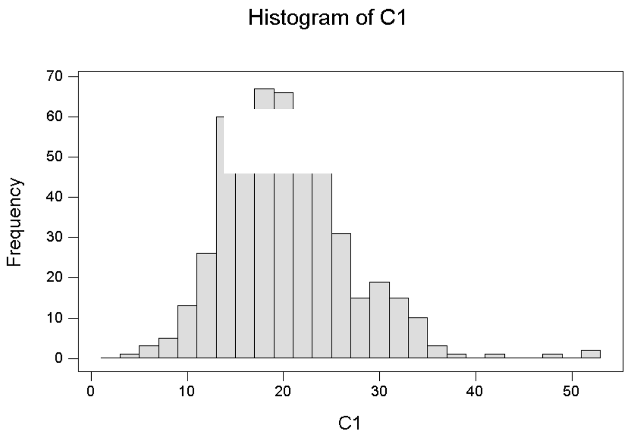


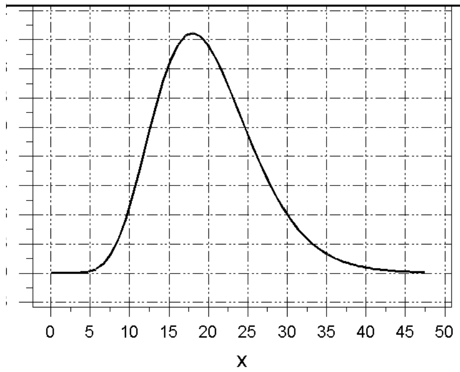
1) Estimació dels paràmetres de la distribució:

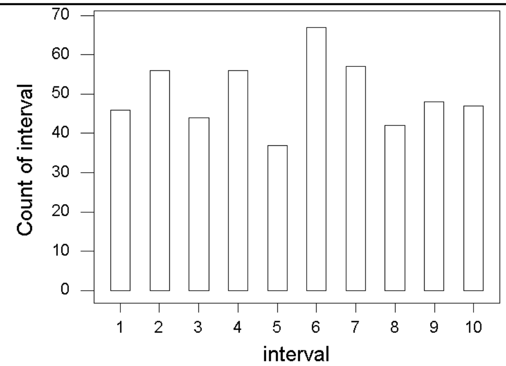
Etapes $k = \left\lceil \left(\frac{\bar{t}}{s_t}\right)^2 \right\rceil = \left\lceil \left(\frac{19,91}{5,99}\right)^2 \right\rceil = \lceil 11,05 \rceil$; s'adopta 11; temps mig per etapa 19,91/11= 1,81036

Proporciona:

- El p-valor i l
- El valor de X^2
- Gràfics:







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)
```