

A lo largo de esta lista de tareas sugerimos explorar los comandos `ecdf()`, `plot(ecdf())`, `curve()` y la familia de comandos `pgamma()`, teniendo especial cuidado en este último en la forma de especificar los argumentos:

`pgamma(p, shape=alfa, rate=lambda)` o `pgamma(p, shape=alfa, scale=1/lambda)`.

**Análisis de datos: una breve explicación de los datos.**

(fuente: [http://astrostatistics.psu.edu/datasets/GRB\\_afterglow.html](http://astrostatistics.psu.edu/datasets/GRB_afterglow.html))

Los Gamma-ray bursts (GRBs) son uno de los fenómenos más exóticos de los que estudia la astronomía moderna. Fueron descubiertos por casualidad durante los 60's por los satélites americanos y rusos y aparecieron como explosiones con emisiones de rayos gamma con una duración de 0.1-100 segundos en lugares aleatorios del cielo. A partir de mediados de los 90's se empezó a ver a los GRBs como un efecto secundario del nacimiento de un agujero negro. Los GRBs de larga duración se asocian la emisión de un chorro de material energético a velocidades cercanas a la de la luz (relativistas) de una estrella supergigante antes de su colapso final y explosión de supernova. Los GRBs de corta duración se asocian a dos estrellas neutrónicas que se espiralan y emergen como un agujero negro, otra vez eyectando un chorro de material energético a velocidades relativistas.

Como el chorro se expande y se enfría, según cálculos astrofísicos, un resplandor de radiación en longitudes de onda larga, tales como rayos-X, bandas visibles y de radio, deberían ser detectados y deberían decaer en tiempos de escala horas-días-semanas.

**Los datos:** las observaciones corresponden al decaimiento de los rayos-X de GRB 050525a obtenidos por el Telescopio de rayos-X (XRT) a bordo del satélite Swift (A. J. Blustin y 64 coautores, *Astrophys. J.* 637, 901-913 2006. Disponible en

<http://arxiv.org/abs/astro-ph/0507515>

El conjunto de datos corresponden a 63 mediciones del brillo en la escala espectral 0.4-4.5 keV a tiempos que van entre 2 minutos y 5 días después de la explosión. Las columnas registran las siguientes variables: 1) el tiempo de observación en segundos, 2) X-ray flux (en unidades de 10-11 ergcm<sup>2</sup>s, 2-10 keV) y 3) mediciones del error del flux basadas en la relación señal-ruido. En el análisis que sigue ignoraremos la variable **tiempo**.

Para acceder a los datos puede realizar el comando:

```
read.table("http://astrostatistics.psu.edu/datasets/GRB_afterglow.dat",
header=T, skip=1)
```

. y la variable de nuestro interés, **flux**, se halla en la segunda columna.

**Estimación bajo modelo Gamma  $\Gamma(\alpha, \lambda)$ :**

Distribución Gamma de parámetros  $\alpha > 0, \lambda > 0, X_i \sim \Gamma(\alpha, \lambda)$ .

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}.$$

Los estimadores de los parámetros  $\alpha$  y  $\lambda$  que son fáciles de computar son los obtenidos por el Método de Momentos y corresponden a:

$$\begin{aligned}\tilde{\alpha} &= \frac{n\bar{X}_n^2}{\sum_{i=1}^n (X_i - \bar{X}_n)^2} \\ \tilde{\lambda} &= \frac{n\bar{X}}{\sum_{i=1}^n (X_i - \bar{X}_n)^2}\end{aligned}$$

Por otro lado, la verosimilitud, cuando  $x_i \geq 0 \forall i$ , resulta

$$\begin{aligned}L(\alpha, \lambda; \mathbf{x}) &= \prod_{i=1}^n f(x_i, \alpha, \lambda) = \prod_{i=1}^n \frac{\lambda^\alpha}{\Gamma(\alpha)} x_i^{\alpha-1} e^{-\lambda x_i} \\ &= \frac{\lambda^{n\alpha}}{\Gamma(\alpha)^n} e^{-\lambda \sum_{i=1}^n x_i} \left\{ \prod_{i=1}^n x_i \right\}^{\alpha-1}.\end{aligned}$$

Por lo tanto, se obtiene que la log-verosimilitud es:

$$\ell(\alpha, \lambda; \mathbf{x}) = n\alpha \log(\lambda) - n \log(\Gamma(\alpha)) + (\alpha - 1) \sum_{i=1}^n \log(x_i) - \lambda \sum_{i=1}^n x_i. \quad (1)$$

En este caso, no hay fórmula cerrada para obtener los valores que maximizan esta expresión y se utilizan métodos numéricos para la aproximación de los mismos.

1. Implementar una función **MME.gamma** que tenga por argumento un conjunto de datos  $\mathbf{x} = (x_1, \dots, x_n)$  y devuelva los estimadores de momentos de  $\alpha$  y  $\lambda$ .
2. Suponiendo que los datos de **flux** provienen de una población con distribución Gamma, estimar  $\alpha$  y  $\lambda$  usando la función **MME.gamma**.
3. Estimar  $P(X \leq 40)$  usando las estimaciones obtenidas y el método plug-in, es decir reemplazando cada parámetro por su valor estimado. Comparar con las estimaciones obtenidas en la Guía de clase.
4. Graficar la empírica asociada a los datos **flux** y superponer la función de distribución acumulada Gamma con los parámetros que considere pertinentes. Comparar con los gráficos obtenidos en la Guía de clase.
5. Realizar un histograma para los datos de **flux** y superponer la función de densidad gamma con los parámetros que considere pertinentes.

6. Bajo el mismo supuesto, calcular los estimadores de máxima verosimilitud de  $\alpha$  y  $\lambda$ . Explorar el comando `fitdistr` de la librería MASS. Comparar con los valores obtenidos en ítems anteriores.

Repetir los tres ítems anteriores y comparar los resultados obtenidos.

7. Estimar por el método de Máxima Verosimilitud el flux medio a partir de los datos. ¿Con qué valor podría comparar esta estimación?
8. Estimar por el método plug-in la mediana de flux. Explorar el comando `qgamma()`