A lo largo de esta lista de tareas sugerimos explorar los comandos `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`

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 po 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 ergcm2s, 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.

Parte 1

1. Estimar $P(X \leq 40)$. Informar el standard error (que también hemos llamado incertidumbre) asociado a esta estimación.
2. Estimar a partir de los datos el flux medio. Informar el standard error (que también hemos llamado incertidumbre) asociado a esta estimación.
3. Estimar a partir de los datos la mediana de flux. Informar el standard error (que también hemos llamado incertidumbre) asociado a esta estimación.
4. Estimar a partir de los datos la varianza de flux. Informar el standard error (que también hemos llamado incertidumbre) asociado a esta estimación.
5. ¿ Se asume algún modelo para obtener las estimaciones previas?
6. Realizar un histograma para los datos **flux**. ¿Los datos parecen tener alguna distribución conocida? Explorar el comando `density`.

Parte 2: Estimación bajo modelo exponencial: $\mathcal{E}(\lambda)$

Distribución exponencial de parámetro $\lambda > 0$, $X_i \sim \mathcal{E}(\lambda)$.

$$f(x) = \lambda e^{-\lambda x} \mathcal{I}_{\{x \geq 0\}}, \quad F(t) = 1 - \exp\{-\lambda t\}\ , \quad \text{para } t \geq 0.$$

En tal caso, $\mathbb{E}(X) = 1/\lambda$ y $\mathbb{V}(X) = 1/\lambda^2$. Recordemos que, de la Ley de los Grandes Números, podemos deducir que $\widehat{\lambda}_n = 1/\bar{X}_n$ es un estimador consistente de λ. De hecho, se puede demostrar que este es también el estimador de máxima verosimilitud.

7. Estimar $P(X \leq 40)$ asumiendo que los datos provienen de una distribución exponencial. Comparar con la estimación obtenida en la Sección 1. Informar el standard error (que también hemos llamado incertidumbre) asociado a esta estimación.
8. Realizar un histograma para los datos de **flux** y superponer la función de densidad exponencial con el parámetro que considere pertinente.
9. Estimar el flux medio asumiendo que los datos provienen de una distribución exponencial. Informar el standard error (que también hemos llamado incertidumbre) asociado a esta estimación.
10. Si $X \sim \mathcal{E}(\lambda)$, entonces su mediana resuelve la ecuación $1 - e^{-\lambda t} = 0.5$, y por consiguiente vale $-\log(0.5)/\lambda = \log(2)/\lambda$. Estimar la mediana de flux asumiendo que los datos provienen de una distribución exponencial. Informar el standard error (que también hemos llamado incertidumbre) asociado a esta estimación.
11. Estimar la varianza de flux, asumiendo que los datos provienen de una distribución exponencial. Informar el standard error (que también hemos llamado incertidumbre) asociado a esta estimación.

Error Cuadrático Medio

Generaremos datos provenientes de una familia exponencial: $X \sim \mathcal{E}(\lambda)$. Procuraremos estimar

$$\theta = P(X \leq 40).$$

12. Calcular $\mathbb{E}(X)$ y $\mathbb{V}(X)$.

13. Obtenga una fórmula para el parámetro de interés en función de λ.

14. De aquí en más consideraremos $\lambda = 0{,}03$. Utilizaremos $\widehat{\theta}_n$ para denotar el estimador utilizado en la Parte 1, mientras que con $\widetilde{\theta}_n$ denotamos el estimador obtenido en la Parte 2, asumiendo que los datos provienen de una distribución exponencial.

    Recordemos que si θ es un parámetro de interés, el error cuadrático medio de un estimador $\widehat{\theta}_n = \widehat{\theta}_n(X_1, \ldots, X_n)$ está dado por

    $$\text{ECM} = \mathbb{E}\{(\widehat{\theta}_n - \theta)^2\}. \tag{1}$$

    Para obtener el ECM necesitamos conocer la distribución de $\widehat{\theta}_n$. Sin embargo, cuando simulamos y generamos datos, podemos estimar el ECM con su versión empírica,(ECME) haciendo

    $$\text{ECME} = \frac{1}{Nrep} \sum_{k=1}^{Nrep} (\widehat{\theta}_{n,k} - \theta)^2, \tag{2}$$

    siendo que $\widehat{\theta}_{n,k}$ la estimación obtenida en la k-ésima replicación.

15. Representar en una tabla el error cuadrático medio empírico, ECME, de los estimadores $\widehat{\theta}_n$ y $\widetilde{\theta}_n$ para muestras de tamaño n=50, n=150, n=200 y n=500, utilizando Nrep=1000 replicaciones en cada caso.

    ¿Qué estimador prefiere bajo este escenario?