Computational Statistics
Project 2

Tsadimas Anargyros
f3612318

For the Maximum Likelihood, from the pdf i will construct the Likelihood function. Then the log Likelihood and differenciate with respect to θ.
Set the derivative to zero to find θ.
Then i will use Newton - Raphson method to calculate $\hat{\theta}$ numerically.

pdf $f(x;\theta) = \frac{\theta^3}{\theta+1} x(1+\frac{x}{2}) \exp(-\theta x)$

Likelihood $L(\theta) = L(\theta|x) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n}\left(\frac{\theta^3}{\theta+1} x(1+\frac{x}{2}) \exp(-\theta x)\right) =$

$$= \left(\frac{\theta^3}{\theta+1}\right)^n \prod_{i=1}^{n} x_i \cdot \prod_{i=1}^{n} (1+\frac{x}{2}) \exp(-\theta \sum_{i=1}^{n} x_i)$$

log Likelihood $l(\theta) = \log(L(\theta)) = \log\left[\left(\frac{\theta^3}{\theta+1}\right)^n \cdot \prod_{i=1}^{n} x_i \cdot \prod_{i=1}^{n}(1+\frac{x}{2}) \exp(-\theta \sum_{i=1}^{n} x_i)\right]$

$$= \log\left(\frac{\theta^3}{1+\theta}\right)^n + \log \prod_{i=1}^{n} x_i + \log \prod_{i=1}^{n} (1+\frac{x}{2}) + \log\left(\exp(-\theta \sum_{i=1}^{n} x_i)\right) =$$

$$= n(\log\theta^3 - \log(1+\theta)) + \sum_{i=1}^{n} \log x_i + \sum_{i=1}^{n} \log(1+\frac{x}{2}) - \theta \sum_{i=1}^{n} x_i =$$

$$= n(3\log\theta - \log(1+\theta)) + \sum(\log x_i + \log(1+\frac{x}{2})) - \theta \sum_{i=1}^{n} x_i$$

take derivatives

$$\frac{\partial \ell(\theta)}{\partial \theta} = 0 \iff n\left(\frac{3}{\theta} - \frac{1}{1+\theta}(1+\theta)'\right) - \sum_{i=1}^{n} x_i = 0$$

$$\frac{3}{\theta} - \frac{1}{1+\theta} = \frac{\sum x_i}{n} = \bar{x} \iff \bar{x} = \frac{3(\theta+1) - \theta}{\theta(1+\theta)}$$

$$\iff \bar{x} = \frac{2\theta+3}{\theta^2+\theta} \iff \bar{x}(\theta^2+\theta) = 2\theta+3$$

$$\iff \bar{x}\theta^2 + \bar{x}\theta - 2\theta - 3 = 0 \iff \bar{x}\theta^2 + (\bar{x}-2)\theta - 3 = 0$$

$$\frac{\partial^2 \ell(\theta)}{\partial \theta^2} = 0 \iff 2\cdot\bar{x}\theta + (\bar{x}-2) = 0$$

Because this is challenging to solve analytically i will use N-R method as i mentioned before: I set $F(\theta) = \bar{x}\theta^2 + (\bar{x}-2)\theta - 3$ and set it to zero, to solve it numerically.

1) I initallize $\theta^{(0)}$, i set it to be 1

2) $\theta^{(r+1)} = \theta^{(r)} - \frac{F(\theta^{(r)})}{F'(\theta^{(r)})}$, 3) Repeat while $|\theta^{(r+1)} - \theta^{(r)}| < \varepsilon$

After running the algorithm for various initial values i found the $\theta = 0.797472$ the only acceptable value for MLE.

Then for the standard error estimation using parametric bootstrap, i will use rejection method to produce a sample from the original pdf. my envelope function will be Gamma(2, 0.6) and $M = 1.1$ both of which I found graphically in R trying out different values.

1) I take random sample from g ~ Gamma(2, 0.6)
2) I take $y \sim unif(0, M * g(x))$
3) if $y < f(x, \hat{\theta})$ keep x.
else, go back to step 1.

In the bootstrap i produce a new sample every time and calculate and keep the theta in a vector.
finally i calculate the standard error from the standard deviation of the bootstrap thetas vector.

$s.e. = 0.063$

To examine if the data come from the given distribution we need a test statistic for Goodness of fit, for each simulated sample.
H_0: the data come from the specific distribution.
H_1: not H_0.

i compare the theoreticall and empeirical Cdf.

The estmated p-value is $0.998 > 0.05$
so we do not reject the null hypothesis.
The data could plausibly have come from the specified distribution.

to find the CDF:

$$f(x,\theta) = \frac{\theta^3}{\theta+1} x \left(1+\frac{x}{2}\right) e^{-\theta x}$$

$$F_X(x,\theta) = \int_0^x f(t,\theta)\,dt = \int_0^x \frac{\theta^3}{\theta+1} t e^{-\theta t} dt + \int_0^x \frac{\theta^3}{\theta+1}\frac{t^2}{2} e^{-\theta t} dt =$$

$$= \frac{\theta^3}{\theta+1}\left(\int_0^x t e^{-\theta t} dt + \int_0^x \frac{t^2}{2} e^{-\theta t} dt\right) = \circledast$$

$$= \frac{\theta^3}{\theta+1}(I_1 + I_2)$$

$$I_1 = \int_0^x t e^{-\theta t} = \frac{x \cdot e^{-\theta x}}{-\theta} - \int_0^x \frac{e^{-\theta t}}{-\theta} dt = -\frac{x}{\theta} \cdot e^{-\theta x} + \frac{1}{\theta}\int_0^x e^{-\theta t} dt =$$

$$= -\frac{x e^{-\theta x}}{\theta} - \frac{1}{\theta^2} \cdot e^{-\theta t}\Big|_0^x = \frac{-x \cdot e^{-\theta x}}{\theta} - \frac{1}{\theta^2} \cdot e^{-\theta x} + \frac{1}{\theta^2}$$

$$I_2 = \int_0^x \frac{t^2}{2} e^{-\theta t} dt = \frac{1}{2}\int_0^x t^2 \cdot e^{-\theta t} dt = ①$$

$$= \frac{1}{2}\left(x^2 \cdot \frac{e^{-\theta x}}{-\theta} - \int_0^x \frac{e^{-\theta t}}{-\theta} \cdot 2t\,dt\right) = \frac{1}{2}\left(-\frac{x^2}{\theta} e^{-\theta x} + \frac{2}{\theta}\int_0^x e^{-\theta t} dt\right) =$$

$$= \frac{1}{2}\left[-\frac{x^2}{\theta} e^{-\theta x} + \frac{2}{\theta}\left(\frac{-x e^{-\theta x}}{\theta} - \frac{1}{\theta^2} e^{-\theta x} + \frac{1}{\theta^2}\right)\right] =$$

$$= \frac{1}{2}\left(-\frac{x^2}{\theta}e^{-\theta x} - \frac{2x}{\theta^2}\cdot e^{-\theta x} - \frac{2}{\theta^3}\cdot e^{-\theta x} + \frac{2}{\theta^3}\right) =$$

$$= -\frac{x^2}{2\theta}\cdot e^{-\theta x} - \frac{x}{\theta}e^{-\theta x} - \frac{1}{\theta^3}e^{-\theta x} + \frac{1}{\theta^3} \qquad 2$$

So 1, 2 $\circledast$ $= \frac{\theta^3}{\theta+1}\left(-\frac{x}{\theta}\cdot e^{-\theta x} - \frac{1}{\theta^2}\cdot e^{-\theta x} + \frac{1}{\theta^2} - \frac{x^2}{2\theta}\cdot e^{-\theta x} - \frac{x}{\theta}\cdot e^{-\theta x} - \frac{1}{\theta^3}\cdot e^{-\theta x} + \frac{1}{\theta^3}\right)$

So this is our CDF.