Comparing some iterative methods of parameter estimation for censored gamma data

A dissertation submitted to
The University of Waikato
in partial fufilment of the requirements
of the degree of

Master of Science

by

Benjamin Milo Bolstad

The University of Waikato

February 1998

Abstract

Maximum likelihood estimation is a great deal more complicated then method of moment estimation for the gamma distribution. The EM algorithm can be used to find maximum likelihood estimates in situations where there is censoring taking place. We will attempt to construct some EM like methods for use with the method of moments in situations where there is censored data. Then we shall compare these moment estimates, by way of simulation, with the EM estimates.

Acknowledgements

I would like to thank Dr Murray Jorgenson, my supervisor, for providing me with the topic and proof reading some of my drafts.

Thank you to all the staff members of the Department of Statistics for providing me a relatively enjoyable education over the last two years.

Thank you to my family for putting up with me and supporting me through my studies.

Thank you to the graduate students (even the mathematicians) for keeping their interruptions and computer questions to a minimum.

Contents

Chapter 1

Introduction

The purpose of this work is to explore several different methods of estimating the parameters of gamma distribution data in censored data situations. As we will see the moment estimators for the gamma distribution are a great deal simpler in form than the maximum likelihood estimators. This suggests that we could find a method for censored data that used the method of moments which would be simpler than the EM algorithm. Because the EM algorithm is related to MLE we cannot use it with the method of moments. We will attempt to develop some iterative methods in the spirit of the EM algorithm for the method of moment estimators.

In the second chapter we will introduce maximum likelihood, the moments method, EM algorithm and define what we mean by censoring.

In the third chapter we will look at the the maximum likelihood and method of moments estimation for the parameters of the gamma distribution when we have observed a random sample upon which no censoring has taken place.

The fourth chapter contains three proposed methods of estimating the parameters of the gamma distribution when censoring has taken place. One method is the EM algorithm, the other two are different approaches to the method of moments.

The fifth chapter contains our description and analysis of the results of some simulations. We compare the three types of estimators we proposed in the fourth chapter in terms of efficiency and iterations to convergence.

The sixth chapter contains some concluding remarks on the results of our simulation trials and on which of the three methods seems to be superior.

We also have an appendix where we present some of the numerical results from our simulations.

Chapter 2

Preliminaries

2.1 Maximum Likelihood Estimation

Maximum Likelihood Estimation is based upon maximising what is known as the likelihood function. The joint density function of a set of random variables $X_1, \ldots, X_n$ evaluated at $x_1, \ldots, x_n$, which we may call $f\left(x_1, \ldots, x_n; \theta\right)$, is referred to as a likelihood function $L\left(\theta\right)$.

The value $\hat{\theta}$ at which $L\left(\theta\right)$ is maximised is called the maximum likelihood estimate (MLE) of θ. That is

$$f(x_1, \ldots, x_n; \hat{\theta}) = \max_{\theta \epsilon \Omega} f\left(x_1, \ldots, x_n; \theta\right) \tag{2.1}$$

where Ω is the parameter space.

If $L\left(\theta\right)$ is differentiable then the MLE will be the solution of the equation

$$\frac{d}{d\theta} L\left(\theta\right) = 0 \tag{2.2}$$

Note also that any value which maximises the likelihood function should also maximise the log-likelihood $l\left(\theta\right) = \ln L\left(\theta\right)$ and since this is usually computationally more convenient we usually solve

$$\frac{d}{d\theta} l(\theta) = 0 \tag{2.3}$$

which is known as the score equation.

2.2 Method of Moments Estimation

Method of Moments Estimation provides us a method which is easy to apply and widely applicable. Our Method of Moment Estimators (MMEs) are the solutions of a series of equations where we equate the theoretical moments with the corresponding sample moments.

We define our theoretical moments (around the origin) such that the jth moment is given by

$$\mu_j(\theta_1, \ldots, \theta_k) = \mathrm{E}\left[X^j\right] \tag{2.4}$$

Our sample moments are such that the jth sample moment is given as

$$\mathrm{M}_j = \frac{\sum_{i=1}^{n} X_i^j}{n} \tag{2.5}$$

based upon a random sample $X_1, \ldots, X_n$ from a distribution $f(x; \theta_1, \ldots, \theta_n)$

If we have k parameters say $(\theta_1, \ldots, \theta_k)$ then we require k equations to solve. Thus our estimates for the parameters $\hat{\theta}_1, \ldots, \hat{\theta}_k$ are the solutions of the equations

$$M_j = \mu_j(\theta_1, \ldots, \theta_k) \qquad j = 1, \ldots, k \tag{2.6}$$

Note that there is no requirement that we use the first k moments. We could choose any k distinct theoretical moments and equate them with the corresponding sample moment to provide us with the required number of equations.

2.3 EM algorithm and Maximum Likelihood

The Expectation-Maximisation (EM) algorithm is an iterative procedure for computing Maximum Likelihood Estimators in situations where the observed data can be considered to be incomplete in some manner. If we had the complete data then computation of the MLEs would be straight forward.

Rather than having the full data $\mathbf{x}$ we have only have our observed data $\mathbf{y}$. The EM algorithm provides us a method to maximise incomplete data likelihood $L(\theta; \mathbf{y})$.

The EM algorithm consists of two steps. The Expectation step (E-Step) where we calculate the expectation of the complete data log-likelihood ($\log L_c(\theta; \mathbf{x})$) conditional on the observed data $\mathbf{y}$ and current estimates of the parameters $\theta^{(k)}$. And the Maximisation step (M-step) where we maximise the expectation of the complete data likelihood obtained in the E-step to find a new parameter estimate $\theta^{(k+1)}$.

More formally we define the two steps of the EM algorithm as

E-Step Calculate $Q\left(\theta; \theta^{(k)}\right)$, where

$$Q\left(\theta; \theta^{(k)}\right) = E_{\theta^{(k)}}\left[\log L_c(\theta; \mathbf{x}) | \mathbf{y}\right] \tag{2.7}$$

M-Step Choose $\theta^{(k+1)}$ to be any value of θ in the allowable parameter space that maximises $Q\left(\theta; \theta^{(k)}\right)$ ie

$$Q\left(\theta^{(k+1)}; \theta^{(k)}\right) \geq Q\left(\theta; \theta^{(k)}\right) \tag{2.8}$$

for all θ in the parameter space.

The E and M steps are repeated alternately until some convergence criterion is met. For example until

$$L\left(\theta^{(k+1)}\right) - L\left(\theta^{(k)}\right) \tag{2.9}$$

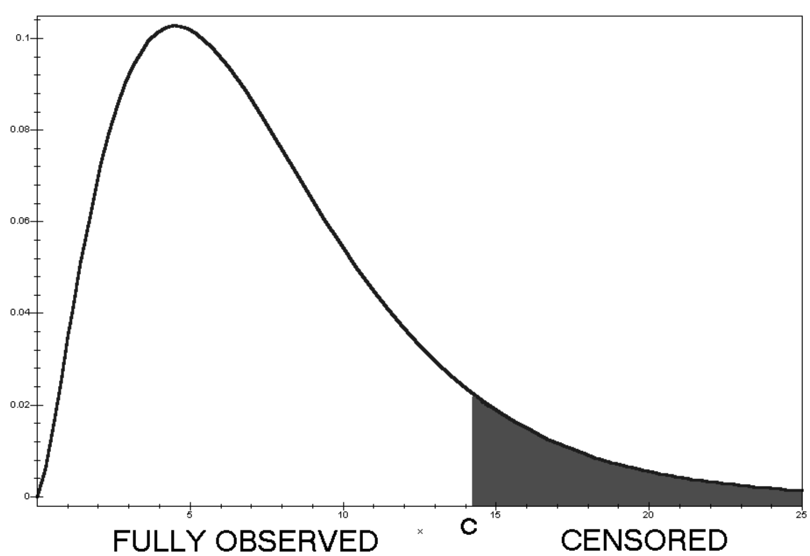
is less than some specified amount. Dempster et al (1977) show that the incomplete data likelihood is not decreased after an EM iteration and so the above condition ensures convergence of a sequence of likelihood values bounded above.

2.4 Censoring

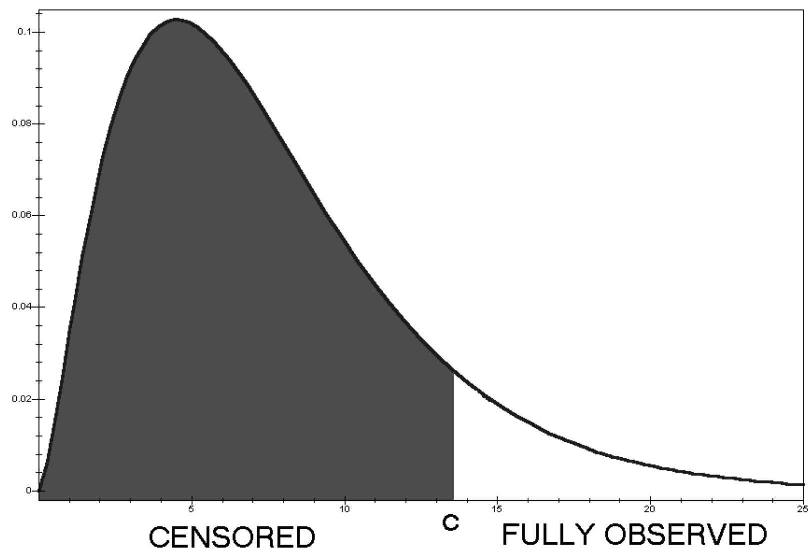
Censoring can occur to data for a number of reasons. For example right censored data could arise in a survival time study when we still have test subjects alive at the end of the study period or left censored data in a chemical analysis when our method has some sort of detection limit.

We shall consider the situation where we have a random sample of size n from some distribution $f(x;\theta)$. Of this we observe m values completely, that is we observe the true value, and $n-m$ values as the value of some censoring point c. In the case of right censoring all the observed data will be less than this value c and in the case of left censoring all the observed data will be greater than c.

The following diagram shows right censoring. Values falling in the unshaded area (to the left of c) are fully observed. Values falling in the shaded area (to the right of c) are censored, we only observe the value c for these observations.



Similarly this diagram illustrates left censored data were we fully observe values in the unshaded area (to the right of c) and observe the value c for values falling in the shaded area.

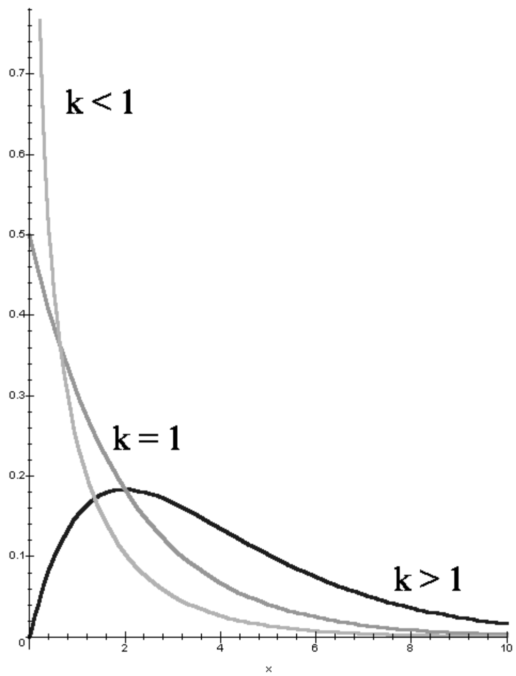


2.5 Two parameter Gamma distribution

A random variable X follows a two parameter gamma distribution if its probability density function is given by

$$f(x;\kappa,\theta)=\frac{1}{\theta^{\kappa}\Gamma(\kappa)}x^{\kappa-1}\exp\left(-\frac{x}{\theta}\right) \tag{2.10}$$

for $0<x<\infty$ where $\kappa>0$ and $\theta>0$. The parameter κ is a shape parameter and the parameter θ is a scale parameter. There are specifically three basic shapes of the distribution depending on whether $\kappa>1$, $\kappa=1$ and $\kappa<1$ as we will now illustrate.



If $\kappa = 1$ we have the special case of the exponential distribution. If $\kappa < 1$ the distribution has an asymptote at 0. The Gamma distribution is a Pearson type III distribution. Also note that setting $\theta = 2$ and $\kappa = \nu/2$ gives the chi-square distribution with ν degrees of freedom.

The mean of the distribution is given by

$$\mathrm{E}[X] = \kappa\theta \tag{2.11}$$

and the variance is given by

$$\mathrm{Var}[X] = \kappa\theta^2 \tag{2.12}$$

Chapter 3

Estimation for uncensored gamma distribution

3.1 Maximum Likelihood Estimators for the two parameter Gamma distribution

To find the maximum likelihood estimates for the gamma distribution let us restate the gamma distribution density function as

$$f(x;\theta,\kappa) = \frac{1}{\theta^{\kappa}\Gamma(\kappa)} x^{\kappa-1} \exp\left(-\frac{x}{\theta}\right) \tag{3.1}$$

where $\theta > 0$ and $\kappa > 0$ and defined for $0 < x < \infty$.

Also lets consider that we have a random sample $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ from a gamma distribution, then our likelihood function for this data given by the product of the density functions which is

$$L(\theta,\kappa;\mathbf{x}) = \prod_{i=1}^{n} \frac{1}{\theta^{\kappa}\Gamma(\kappa)} x_i^{\kappa-1} \exp\left(-\frac{x_i}{\theta}\right)$$

Then taking logs, so as to get the log-likelihood, gives us

$$\begin{aligned} l(\theta, \kappa; \mathbf{x}) &= \ln\left(L(\theta, \kappa; \mathbf{x})\right) \\ &= \ln\left(\prod_{i=1}^{n} \frac{1}{\theta^{\kappa}\Gamma(\kappa)} x_i^{\kappa-1} \exp\left(-\frac{x_i}{\theta}\right)\right) \\ &= -n\kappa\ln(\theta) - n\ln(\Gamma(\kappa)) + (\kappa-1)\sum_{i=1}^{n}\ln(x_i) - \sum_{i=1}^{n}\left(\frac{x_i}{\theta}\right) \end{aligned} \tag{3.2}$$

Then differentiating with respect to each of the parameters κ and θ we get

$$\frac{\partial l}{\partial \kappa} = -n\ln(\theta) - n\Psi(\kappa) + \sum_{i=1}^{n}\ln(x_i) \tag{3.3}$$

(where $\Psi(\kappa) = \frac{d}{d\kappa}\log\Gamma(\kappa)$) and

$$\frac{\partial l}{\partial \theta} = -\frac{n\kappa}{\theta} + \sum_{i=1}^{n}\frac{x_i}{\theta^2} \tag{3.4}$$

now setting (3.3) and (3.4) to zero and solving for θ and κ will give us the maximum likelihood estimators. Thus the maximum likelihood estimates are given by the solutions of

$$\kappa\theta = \bar{x} \tag{3.5}$$

$$n\ln(\theta) + n\Psi(\kappa) = \sum_{i=1}^{n}\ln(x_i) \tag{3.6}$$

Harter et al (1965) have looked at maximum likelihood estimation for the gamma distribution.

3.1.1 Numerical methods for solving the score equations

Newton-Raphson

Solving equation (3.5) for θ results in

$$\theta = \frac{\bar{x}}{\kappa} \tag{3.7}$$

where $\bar{x}$ is the arithmetic mean. Substituting this into the second equation to get an expression in terms of κ only we get

$$-n \ln\left(\frac{\bar{x}}{\kappa}\right) - n\Psi\left(\kappa\right) + \sum_{i=1}^{n} \ln(x_i) = 0$$

which we can easily rearrange to give

$$\ln(\kappa) - \Psi(\kappa) - \ln\left(\frac{\bar{x}}{\tilde{x}}\right) = 0 \tag{3.8}$$

where $\tilde{x} = \left(\prod_{i=1}^{n} x_i\right)^{1/n}$ is the geometric mean. Now (3.8) can not be solved in a closed form but we could solve it numerically by using Newton Raphson iterations. ie

$$\kappa_{n+1} = \kappa_n - \frac{\ln(\kappa_n) - \Psi(\kappa_n) - \ln(\bar{x}/\tilde{x})}{1/\kappa_n - \Psi'(\kappa_n)}$$

performed until $|\kappa_{n+1} - \kappa_n| < \epsilon$ where ϵ is some defined error tolerance and $\epsilon > 0$.

So we see that the maximum likelihood estimates $\hat{\theta}$ and $\hat{\kappa}$ are given by equation (3.7) and the solution of equation (3.8).

The Newton-Raphson method has been explored in Choi and Wette (1969).

A fixed point method

Again we take $\theta = \frac{\bar{x}}{\kappa}$ but to solve (3.8) we use a fixed point method rather than Newton-Raphson iterations. We will define a fixed point method such that if we wish to solve the equation

$$\mathrm{F}\left(\kappa\right)=0$$

we shall use an iteration of the form

$$\kappa_{n+1}=\mathrm{f}\left(\kappa_{n}\right)$$

It can be shown (using the mean value theorem) that for the above iteration to converge we require that $|\mathrm{f}'\left(k\right)|< 1$.

From the previous section we have that

$$\mathrm{F}\left(\kappa\right)=\ln(\kappa)-\Psi(\kappa)-\ln\left(\frac{\bar{x}}{\tilde{x}}\right)=0$$

which we rearrange to give us

$$\kappa=\exp\left(\Psi\left(\kappa\right)\right)\frac{\bar{x}}{\tilde{x}}$$

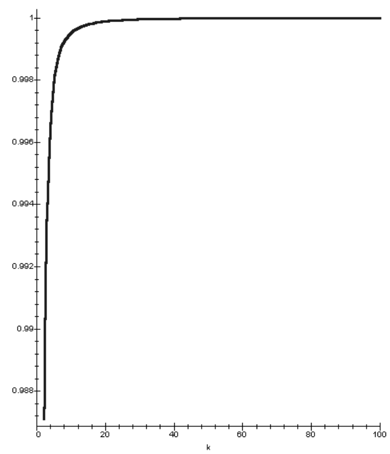
that is we have

$$\mathrm{f}\left(\kappa\right)=\exp\left(\Psi\left(\kappa\right)\right)\frac{\bar{x}}{\tilde{x}}$$

Differentiating we get

$$\mathrm{f}'\left(\kappa\right)=\Psi'\left(\kappa\right)\exp\left(\Psi\left(\kappa\right)\right)\frac{\bar{x}}{\tilde{x}}$$

The diagram that follows illustrates this derivative function



Now it can be shown that for $\kappa_1 < \kappa_2$

$$\mathrm{f}'(\kappa_1) < \mathrm{f}'(\kappa_2)$$

ie that the derivative function is a strictly increasing function.

It can also be shown that

$$\mathrm{f}'(\kappa) > 0$$

for $\forall \kappa > 0$. That is that the derivative function is always positive.

Finally it can be shown that

$$\lim_{\kappa \to \infty} \mathrm{f}'(\kappa) = \frac{\bar{x}}{\tilde{x}}$$

Thus to ensure convergence $\forall \kappa \in \Re$ we require that

$$\mathrm{f}'(\kappa) < 1$$

which means essentially that we require

$$\bar{x} < \tilde{x}$$

for convergence from any starting value.

If on the other hand we have $\bar{x} \geq \tilde{x}$ our method will only converge if we choose a starting κ_1 such that

$$\exp\left(\Psi\left(\kappa_1\right)\right)\Psi'\left(\kappa_1\right) < \frac{\tilde{x}}{\bar{x}}$$

and that value $\hat{\kappa}$ also statisfies the above inequality.

Thus our fixed point method would be

$$\kappa_{n+1} = \frac{\bar{x}}{\tilde{x}}\exp\left(\Psi\left(\kappa_n\right)\right) \tag{3.9}$$

Other Methods

Choi and Wette (1969) consider a maximum likelihood scoring method and note that although this proposed technique involves more computation than the Newton Raphson method it is appealing because it provides a statistical criterion for stopping the iterations.

3.1.2 Approximate maximum likelihood estimates

Several authors have considered approximate solutions to the maximum likelihood estimation equations.

Thom (1958) give the following approximate maximum likelihood solutions

$$\hat{\kappa} = \frac{1 + \sqrt{1 + 4M/3}}{4M}$$

$$\hat{\theta} = 3\bar{x}\left(\sqrt{1 + 4M/3} - 1\right)$$

where

$$M = \ln\left(\frac{\bar{x}}{\tilde{x}}\right)$$

Greenwood and Durand (1960) provide better rational approximations for $\hat{\kappa}$. These approximations are

$$\hat{\kappa} = \begin{cases} \dfrac{0.5000876 + 0.1648852M - 0.0544274M^2}{M} & 0 < M \leq 0.5772 \\ \dfrac{8.898919 + 9.059950M + 0.9775373M^2}{M\left(17.79728 + 11.968477M + M^2\right)} & 0.5772 < M \leq 17 \\ \dfrac{1}{M} & M > 17 \end{cases} \tag{3.10}$$

where again

$$M = \ln\left(\frac{\bar{x}}{\tilde{x}}\right)$$

Shenton and Bowman (1977) note that for the two lower ranges the errors (in comparison to the m.l. estimators) are 0.008 percent and 0.0054 percent respectively. The solution for θ is still defined as

$$\hat{\theta} = \frac{\bar{x}}{\hat{\kappa}}$$

3.1.3 Other Methods

Gilchrist(1981) uses the density of the statistic $(x_2/x_1, x_3/x_1, \ldots, x_n/x_1)$ where $x_1, x_2, \ldots, x_n$ is the sample to find a form of maximum likelihood estimate for κ.

3.2 Method of Moment Estimators for the two parameter Gamma distribution

To find the method of moment estimators we must equate the moments for the population distribution with the sample moments.

Firstly lets consider the moment generating function for the gamma distribution.

$$\begin{aligned}
\mathrm{M}_X(t) &= \mathrm{E}\left[e^{tx}\right] \\
&= \int_0^\infty e^{tx}\frac{1}{\theta^\kappa\Gamma(\kappa)}x^{\kappa-1}e^{\left(-\frac{x}{\theta}\right)}dx \\
&= \frac{1}{\theta^\kappa\Gamma(\kappa)}\int_0^\infty x^{\kappa-1}e^{\left(tx-\frac{x}{\theta}\right)}dx
\end{aligned}$$

let $u = -\left(t - 1/\theta\right)x$ then $du = -\left(t-1/\theta\right)dx$ and so

$$\begin{aligned}
\mathrm{M}_X(t) &= \frac{1}{\theta^\kappa\Gamma(\kappa)}\int_0^\infty e^{-u}\frac{u^{\kappa-1}}{\left(1/\theta - t\right)^\kappa}du \\
&= \frac{\left(1/\theta - t\right)^{-\kappa}}{\theta^\kappa\Gamma(\kappa)}\int_0^\infty e^{-u}u^{\kappa-1}du \\
&= \frac{\left(1/\theta - t\right)^{-\kappa}}{\theta^\kappa\Gamma(\kappa)}\Gamma(\kappa) \\
&= \theta^{-\kappa}\{\left(1/\theta - t\right)^{-\kappa} \\
&= \left(1-t\theta\right)^{-\kappa}
\end{aligned}$$

Now we can easily compute the population moments for the gamma distribution. The first population moment is

$$\begin{aligned}
\mu_1'(\theta,\kappa) &= \mathrm{E}\left[X\right] \\
&= \mathrm{M}_X'(t)|_{t=0} \\
&= \kappa\theta\left(1-t\theta\right)^{-\kappa-1}|_{t=0} \\
&= \kappa\theta \qquad (3.11)
\end{aligned}$$

and the second population moment is

$$\begin{aligned}
\mu_2'(\theta,\kappa) &= \mathrm{E}\left[X^2\right] \\
&= \mathrm{M}_X''(t)|_{t=0} \\
&= \frac{d}{dt}\mathrm{M}_X'(t)|_{t=0} \\
&= \kappa\theta\left(-\kappa-1\right)\left(1-t\theta\right)^{-\kappa-2}\left(-\theta\right)|_{t=0} \\
&= \kappa\theta^2\left(\kappa+1\right) \qquad (3.12)
\end{aligned}$$

The kth sample moment is given by

$$\mathrm{M}_k' = \frac{\sum_{i=1}^n x_i^k}{n}$$

so it is easy to see that the first and second sample moments are

$$\mathrm{M}_1' = \frac{\sum_{i=1}^n x_i}{n} \qquad (3.13)$$

and

$$\mathrm{M}_2' = \frac{\sum_{i=1}^n x_i^2}{n} \qquad (3.14)$$

Now to find the MMEs we just equate (3.11) with (3.13) and (3.12) with (3.14) and then solve the resulting system of equations for our two parameters θ and κ.

Taking the first set of equations (those for the first sample and population moments) we get

$$\frac{\sum_{i=1}^n x_i}{n} = \theta\kappa \qquad (3.15)$$

which we may rearrange to form the estimator for κ

$$\hat{\kappa} = \frac{\bar{x}}{\hat{\theta}} \tag{3.16}$$

which we can substitute into the second equation, which is

$$\frac{\sum_{i=1}^{n} x_i^2}{n} = \kappa\theta^2 \left(\kappa + 1\right) \tag{3.17}$$

to get

$$\frac{\sum_{i=1}^{n} x_i^2}{n} = \frac{\bar{x}}{\theta}\theta^2 \left(\frac{\bar{x}}{\theta} + 1\right)$$

and rearranging this to make theta the subject we get

$$\theta = \frac{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2}{n\bar{x}} \tag{3.18}$$

So equations (3.16) and (3.18) give us our method of moments estimators $\hat{\kappa}$ and $\hat{\theta}$. Note that we can also rearrange (3.18) to give

$$\hat{\theta} = \frac{\left[\left(n-1\right)/n\right] S^2}{\bar{x}} \tag{3.19}$$

where S is the sample standard deviation.

There is no requirement that we use the first two moment equations. More generally we might wish to consider the two moment equations

$$\theta^a \frac{\Gamma\left(\kappa + a\right)}{\Gamma\left(\kappa\right)} = \frac{\sum x_i^a}{n} \tag{3.20}$$

and

$$\theta^b \frac{\Gamma\left(\kappa + b\right)}{\Gamma\left(\kappa\right)} = \frac{\sum x_i^b}{n} \tag{3.21}$$

where a and b are some chosen constants $a \neq b$.

We can see the relative simplicity of the moment estimators compared to the maximum likelihood estimators. This suggests that perhaps it would be more convenient to use method of moments estimation equations in a censored data situation. We will explore this in the next chapter.

3.3 Comparisons between moment and maximum likelihood estimates

Fisher (1921) showed that the method of moment may be inefficient for estimating the parameters of Pearson type III distributions (of which the gamma distribution is one) and suggested the use of maximum likelihood.

Kendall and Stuart (1977) pp70 -72 show that the method of moments when applied to the gamma distribution can have efficiency as low as 22 percent when compared to the maximum likelihood estimates.

Chapter 4

Estimation for censored gamma distribution

4.1 Maximum Likelihood for censored Gamma data

In the previous chapter we explored the maximum likelihood situation for uncensored gamma distribution data. In this section we explore maximum likelihood estimation for censored gamma distribution data.

Consider that we observe $\mathbf{y} = (y_1, \ldots, y_m, y_{m+1}, \ldots, y_n)$ where y_i if fully observed for $i = 1, \ldots, m$ and observed as $y_i = c$ for $i = m+1, \ldots, n$. That is the last $n-m$ observations are censored at some point c. We may then write the density function as

$$g\left(y; \kappa, \theta\right) = \begin{cases} \frac{1}{\theta^{\kappa}\Gamma(\kappa)} x^{\kappa-1} \exp\left(-\frac{x}{\theta}\right) & 0 \leq y < c \\ \delta\left(y-4\right)\left(1 - F\left(c; \kappa, \theta\right)\right) & y \geq c \end{cases} \tag{4.1}$$

where $F(c; \kappa, \theta)$ is the gamma distribution CDF given as

$$\begin{aligned} F\left(c; \kappa, \theta\right) &= \int_0^c \frac{1}{\theta^{\kappa}\Gamma\left(\kappa\right)} x^{\kappa-1} \exp\left(-\frac{x}{\theta}\right) dx \\ &= \frac{1}{\theta^{\kappa}\Gamma\left(\kappa\right)} \int_0^c x^{\kappa-1} \exp\left(-\frac{x}{\theta}\right) dx \end{aligned}$$

then making the substitution $u = \frac{x}{\theta}$ we get

$$F(c;\kappa,\theta) = \frac{\gamma(\kappa, c/\theta)}{\Gamma(\kappa)} \tag{4.2}$$

It is easy to see then that the likelihood in terms of the observed data $\mathbf{y}$ is then

$$\begin{aligned} L(\theta,\kappa) &= \prod_{i=1}^{n} g(y_i;\kappa,\theta) \\ &= \prod_{i=1}^{m} \frac{1}{\theta^{\kappa}\Gamma(\kappa)} x^{\kappa-1} \exp\left(-\frac{y_i}{\theta}\right) \prod_{i=m+1}^{n} \frac{\Gamma(\kappa, c/\theta)}{\Gamma(\kappa)} \end{aligned}$$

which implies that the log-likelihood is then

$$l(\kappa,\theta) = -mk\ln(\theta) - n\ln\Gamma(\kappa) + (\kappa-1)\sum_{i=1}^{m}\ln(y_i) - \frac{1}{\theta}\sum_{i=1}^{m} y_i + (n-m)\ln\Gamma(\kappa, c/\theta) \tag{4.3}$$

to get the maximum likelihood estimates we the differentiate with respect to κ and θ to get the score equations

$$\frac{\partial l}{\partial \kappa} = -m\ln\theta - n\Psi(\kappa) + \sum_{i=1}^{m}\ln y_i + \frac{n-m}{\Gamma(\kappa, c/\theta)}\frac{\partial}{\partial\kappa}\Gamma(\kappa, c/\theta) = 0 \tag{4.4}$$

and

$$\frac{\partial l}{\partial \theta} = -\frac{m\kappa}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{m} y_i + \frac{n-m}{\Gamma(\kappa, c/\theta)}\frac{\partial}{\partial\theta}\Gamma(\kappa, c/\theta) = 0 \tag{4.5}$$

the maximum likelihood estimates for θ and κ are the solutions of these equations. Rather than solve these equations directly we will instead consider an iterative solution to these equations using the EM method in the next section.

Maximum Likelihood for censored gamma distributions has been considered by Harter and Moore (1965) who looked, as we have done, at a single point of censoring. Cohen and Norgaard (1977) have also considered maximum likelihood estimation for censored gamma distributions, but they consider progressive censoring. That is they have considered the situation where there is more than one point at which censoring is taking place. For example in a lifetime study when some specimens are removed for some reason earlier than others.

4.2 EM algorithm for censored data from a Gamma distribution

As can be seen from (3.2) the sufficient statistics for the gamma distribution are given by $\sum_{i=1}^{n} x_i$ and $\sum_{i=1}^{n} \ln(x_i)$. We use these to calculate our maximum likelihood estimates $\hat{\kappa}$ and $\hat{\theta}$.

When the data is censored the EM algorithm provides a method by which we can get our maximum likelihood estimates. The EM algorithm provides us a method by which we can maximise the incomplete data likelihood (or equivalently solve the incomplete data score equations) iteratively.

In this case we may estimate our sufficient statistics based on those values that were observed (the uncensored data) and the expected value of the sufficient statistic based on out current estimates of the parameters and the point at which the data has been censored. This is the E-step of the EM algorithm.

Using the estimates of the sufficient statistics calculated in the E-step we may then calculate new values of the parameters, as given by (3.7) and (3.8). This is the M-step of the EM algorithm. Repeating the E and M-steps of the algorithm will lead to progressively better estimates for $\hat{\kappa}$ and $\hat{\theta}$.

For the gamma distribution we must estimate the sufficient statistics based on the observed data and the expected value of each of the sufficient statistics assuming values for θ and κ.

Firstly we will consider a gamma distribution right censored at c. In this case the form of the data is given by $(y_1, y_2, \ldots, y_m, y_{m+1}, \ldots, y_n)$ where $(y_1, y_2, \ldots, y_m)$ are the first m data points, the uncensored data (have value less than c), and $(y_{m+1}, \ldots, y_n)$ are $n-m$ data values for which the data has been censored and we observe $y_i = c$.

In this case to estimate $\sum_{i=1}^{n} x_i$ take

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{m} y_i + (n-m)\,\mathrm{E}\left[x|\theta, \kappa, x > c\right] \tag{4.6}$$

and to estimate $\sum_{i=1}^{n} \ln(x_i)$ take

$$\sum_{i=1}^{n} \ln(x_i) = \sum_{i=1}^{m} \ln(y_i) + (n-m)\,\mathrm{E}\left[\ln(x)\,|\theta, \kappa, x > c\right] \tag{4.7}$$

Calculating the expected value required for (4.6) we find

$$\mathrm{E}\left[x|\theta, \kappa, x > c\right] = \int_0^c x f\,(x|\,\kappa, \theta, x > c)dx$$

where

$$f\,(x|\kappa, \theta, x > c) = \frac{\frac{1}{\theta^\kappa \Gamma(\kappa)} x^{\kappa-1} \exp\left(-\frac{x}{\theta}\right)}{1 - F(c;\kappa,\theta)}$$

where $F(c;\kappa,\theta)$ is the gamma CDF at c, the censoring point. Thus

$$\begin{aligned}
\mathrm{E}\left[x|x > c, \kappa, \theta\right] &= \frac{\displaystyle\int_c^\infty x \frac{1}{\theta^\kappa \Gamma(\kappa)} x^{\kappa-1} \exp\left(-\frac{x}{\theta}\right) dx}{\displaystyle\int_c^\infty \frac{1}{\theta^\kappa \Gamma(\kappa)} x^{\kappa-1} \exp\left(-\frac{x}{\theta}\right) dx} \\
&= \frac{\displaystyle\int_c^\infty x^{(1+k)-1} \exp\left(-\frac{x}{\theta}\right) dx}{\displaystyle\int_c^\infty x^{\kappa-1} \exp\left(-\frac{x}{\theta}\right) dx}
\end{aligned}$$

now put $u = \frac{x}{\theta} \Rightarrow du = \frac{dx}{\theta}$ so

$$\mathrm{E}\left[x|x > c, \theta, \kappa\right] = \theta \frac{\displaystyle\int_{c/\theta}^{\infty} u^{(1+\kappa)-1} \exp(-u) du}{\displaystyle\int_{c/\theta}^{\infty} u^{\kappa-1} \exp(-u) du} \tag{4.8}$$

$$= \theta \frac{\Gamma(\kappa+1, c/\theta)}{\Gamma(\kappa, c/\theta)} \tag{4.9}$$

where $\Gamma(\kappa + 1, c/\theta)$ is the the incomplete gamma function (we use the notation of Abramowitz and Stegun (1964)).

Similarly for the expected value required for (4.7) we get

$$\mathrm{E}\left[\ln(x)|\theta, \kappa, x > c\right] = \int_c^{\infty} \ln(x)\, f(x|\, \kappa, \theta, x > c) dx$$

where $f(y|\kappa, \theta, y > c)$ is as was given before. So

$$\mathrm{E}\left[\ln(x)|\theta, \kappa, x > c\right] = \frac{\displaystyle\int_c^{\infty} \ln(x) \frac{1}{\theta^{\kappa}\Gamma(\kappa)} x^{\kappa-1} \exp\left(-\frac{x}{\theta}\right) dx}{\displaystyle\int_c^{\infty} \frac{1}{\theta^{\kappa}\Gamma(\kappa)} x^{\kappa-1} \exp\left(-\frac{x}{\theta}\right) dx}$$

$$= \frac{\displaystyle\int_c^{\infty} \ln(x)\, x^{\kappa-1} \exp\left(-\frac{x}{\theta}\right) dx}{\displaystyle\int_c^{\infty} x^{\kappa-1} \exp\left(-\frac{x}{\theta}\right) dx}$$

now put $u = \frac{y}{\theta} \Rightarrow du = \frac{dx}{\theta}$ so

$$\begin{aligned}
\mathrm{E}\left[\ln(x)\,|\theta,\kappa,x>c\right] &= \frac{\displaystyle\int_{c/\theta}^{\infty} \ln(u\theta)\, u^{\kappa-1}\exp(-u)\,du}{\displaystyle\int_{c/\theta}^{\infty} u^{\kappa-1}\exp(-u)\,du} \\
&= \frac{1}{\Gamma(\kappa, c/\theta)} \int_{c/\theta}^{\infty} \ln(u\theta)\, u^{\kappa-1}\exp(-u)\,du \\
&= \frac{1}{\Gamma(\kappa, c/\theta)} \left(\ln(\theta)\int_{c/\theta}^{\infty} u^{\kappa-1}\exp(-u)\,du \right.\\
&\left. + \int_{c/\theta}^{\infty} \ln(u)\, u^{\kappa-1}\exp(-u)\,du\right) \\
&= \ln(\theta) + \frac{1}{\Gamma(\kappa, c/\theta)} \int_{c/\theta}^{\infty} \frac{d}{d\kappa}\left[u^{\kappa-1}\exp(-u)\right] du \\
&= \ln(\theta) + \frac{1}{\Gamma(\kappa, c/\theta)} \frac{d}{d\kappa}\Gamma(\kappa, c/\theta) \qquad (4.10)
\end{aligned}$$

Using these values for our sufficient statistics and equations (3.7) and (3.8) we can then update our estimates of the parameters. Then we may repeat the process with the updated parameter estimates to get even better parameter estimates. The process is repeated until convergence. ie we solve

$$\begin{aligned}
& -n\ln\left(\theta_{(t+1)}\right) - n\Psi\left(\kappa_{(t+1)}\right) + \sum_{i=1}^{m} \ln(y_i) + \\
& \qquad (n-m)\left(\ln\theta_{(t)} + \frac{1}{\Gamma\left(\kappa_{(t)}, c/\theta_{(t)}\right)} \frac{\partial}{\partial\kappa}\Gamma\left(\kappa_{(t)}, c/\theta_{(t)}\right)\right) = 0 \qquad (4.11)
\end{aligned}$$

and

$$-\frac{n\kappa_{(t+1)}}{\theta_{(t+1)}} + \frac{1}{\theta_{(t+1)}^2}\left(\sum_{i=1}^{m} y_i + (n-m)\,\theta_{(t)}\frac{\Gamma(\kappa_{(t)}+1, c/\theta_{(t)})}{\Gamma(\kappa_{(t)}, c/\theta_{(t)})}\right) = 0$$

for $\kappa_{(t+1)}$ and $\theta_{(t+1)}$ given current estimates $\kappa_{(t)}$ and $\theta_{(t)}$. Note that if we take $\theta_{(t)} = \theta_{(t+1)} = \theta$ and $\kappa_{(t)} = \kappa_{(t+1)} = \kappa$ it is not to hard to show that these equations are the incomplete data score equations we derived in the previous section. Blight (1970) has considered censoring for exponential families. This paper uses an argument based around reorganising the incomplete data score into the form of the complete data score and then solving iteratively. This method turns out to be the EM method.

Following similar working to above we can show that for data left censored, at c, we use

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{m} y_i + (n-m)\,\mathrm{E}\left[x|\theta,\kappa,x<c\right]$$

and

$$\sum_{i=1}^{n} \ln\left(x_i\right) = \sum_{i=1}^{m} \ln\left(y_i\right) + (n-m)\,\mathrm{E}\left[\ln(x)|\theta,\kappa,x<c\right]$$

as our estimates of the sufficient statistics where

$$\mathrm{E}\left[x|\theta,\kappa,x>c\right] = \theta\frac{\gamma\left(\kappa+1,c/\theta\right)}{\gamma\left(\kappa,c/\theta\right)}$$

and

$$\mathrm{E}\left[\ln\left(x\right)|\theta,\kappa,x>c\right] = \ln\left(\theta\right) + \frac{1}{\gamma\left(\kappa,c/\theta\right)}\frac{d}{dk}\gamma\left(\kappa,c/\theta\right)$$

where $\gamma\left(k,c/\theta\right)$ is the complementary incomplete gamma fucntion (again using the notation of Abramowitz and Stegun (1964)).

4.3 Moment Estimators for censored Gamma data

Suppose that our Gamma distribution data has been censored. Let $\mathbf{y}$ be our observed data and $\mathbf{x}$ be the hypothetical complete data. We will firstly consider the case of right censored data so that we only observe the correct values (ie $y_i = x_i$) when $x_i < c$ for some c which is our censoring point and we observe $y_i = c$ when $x_i >= c$. We may write the density function for y_i as

$$\mathrm{g}\left(y;\kappa,\theta\right) = \begin{cases} \frac{1}{\theta^{\kappa}\Gamma(\kappa)} y^{\kappa-1} \exp\left(-\frac{y}{\theta}\right) & y < c \\ \delta\left(y-c\right) \frac{\Gamma(\kappa,c/\theta)}{\Gamma(\kappa)} & y \geq c \end{cases} \tag{4.12}$$

We will suppose that y_i for $i = 1, \ldots, m$ is the fully observed (uncensored) data and $i = m+1, \ldots, n$ is the censored data. Given the above probability density (4.12) we could calculate a set of incomplete data method of moment estimating equations. That is we set the theoretical moment with the observed data moment. ie we take the equations

$$\mathrm{E}\left[y\right] = \frac{\sum_{i=1}^{m} y_i + (n-m)c}{n} \tag{4.13}$$

and

$$\mathrm{E}\left[y^2\right] = \frac{\sum_{i=1}^{m} y_i^2 + (n-m)c^2}{n} \tag{4.14}$$

as our moment equations and solve for kappa and theta. Where the left hand sides are given by

$$\int_0^{\infty} y\mathrm{g}\left(y;\kappa,\theta\right) dy = \theta\frac{\gamma\left(\kappa+1,c/\theta\right)}{\Gamma\left(\kappa\right)} + c\frac{\Gamma\left(\kappa,c/\theta\right)}{\Gamma\left(\kappa\right)} \tag{4.15}$$

and

$$\int_0^\infty y^2 \mathrm{g}\left(y; \kappa, \theta\right) dy = \theta^2 \frac{\gamma\left(\kappa + 2, c/\theta\right)}{\Gamma\left(\kappa\right)} + c^2 \frac{\Gamma\left(\kappa, c/\theta\right)}{\Gamma\left(\kappa\right)} \tag{4.16}$$

respectively.

Thus to get estimates for theta and kappa we are required to solve

$$\theta \frac{\gamma\left(\kappa + 1, c/\theta\right)}{\Gamma\left(\kappa\right)} + c \frac{\Gamma\left(\kappa, c/\theta\right)}{\Gamma\left(\kappa\right)} = \frac{\sum_{i=1}^m y_i + (n-m)c}{n} \tag{4.17}$$

and

$$\theta^2 \frac{\gamma\left(\kappa + 2, c/\theta\right)}{\Gamma\left(\kappa\right)} + c^2 \frac{\Gamma\left(\kappa, c/\theta\right)}{\Gamma\left(\kappa\right)} = \frac{\sum_{i=1}^m y_i^2 + (n-m)c^2}{n} \tag{4.18}$$

By similar argument to the above we could show that for left censored data the incomplete moment equations are

$$\theta \frac{\Gamma\left(\kappa + 1, c/\theta\right)}{\Gamma\left(\kappa\right)} + c \frac{\gamma\left(\kappa, c/\theta\right)}{\Gamma\left(\kappa\right)} = \frac{\sum_{i=1}^m y_i + \left(n-m\right) c}{n}$$

and

$$\theta^2 \frac{\Gamma\left(\kappa + 2, c/\theta\right)}{\Gamma\left(\kappa\right)} + c^2 \frac{\gamma\left(\kappa, c/\theta\right)}{\Gamma\left(\kappa\right)} = \frac{\sum_{i=1}^m y_i + \left(n-m\right) c^2}{n}$$

Now if we wished to solve (4.17) and (4.18) for θ and κ using Newton-Raphson iterations we see that this would require us to calculate the derivatives of the incomplete gamma functions. This seems overly complicated compared to the simple solutions that the moment equations had in the complete data case. This is why we will now consider a simple iterative method.

4.4 Moment Estimators for censored Gamma data using an iterative method

If we compare the incomplete data moment equations ((4.17) and (4.18)) which we derived in the previous section with the complete data moment equations (3.15) and (3.17) we see that the complete data equations are a great deal simpler. In the spirit of Heyde and Morton (1996), we attempt to construct an iterative method for solving the incomplete data moment equations. Using the proposed method we essentially equate a complete data score (Q) with the corresponding incomplete data score (Q^*). This allows us to work out the manipulation needed to give us the incomplete data moment equation in the form of the complete data moment equation. We then propose an iterative method based upon this rearrangement. This is perhaps analogous to considering the EM method as merely the rearrangement of the incomplete data score into the complete data score equations.

Let us define our complete data moment estimating equations by that

$$\mathrm{Q}_1(\theta,\kappa) = \theta\frac{\Gamma(\kappa+1)}{\Gamma(\kappa)} - \frac{\mathrm{fx}_1(\kappa,\theta,\mathbf{y})}{n} = 0 \tag{4.19}$$

and

$$\mathrm{Q}_2(\theta,\kappa) = \theta^2\frac{\Gamma(\kappa+2)}{\Gamma(\kappa)} - \frac{\mathrm{fx}_2(\kappa,\theta,\mathbf{y})}{n} = 0 \tag{4.20}$$

where $\mathrm{fx}_1(\kappa,\theta,y)$ and $\mathrm{fx}_2(\kappa,\theta,y)$ are functions of the observed data and the parameters.

Let us also define our incomplete data moment estimating estimating equations so that

$$\mathrm{Q}_1^*(\kappa,\theta;\mathbf{y}) = \theta\frac{\gamma(\kappa+1,c/\theta)}{\Gamma(\kappa)} + c\frac{\Gamma(\kappa,c/\theta)}{\Gamma(\kappa)} - \frac{\sum_{i=1}^m y_i + (n-m)c}{n} = 0 \tag{4.21}$$

and

$$\mathrm{Q}_2^*\left(\kappa,\theta;\mathbf{y}\right)=\theta^2\frac{\gamma\left(\kappa+2,c/\theta\right)}{\Gamma\left(\kappa\right)}+c^2\frac{\Gamma\left(\kappa,c/\theta\right)}{\Gamma\left(\kappa\right)}-\frac{\sum_{i=1}^m y_i^2+\left(n-m\right)c^2}{n}=0 \tag{4.22}$$

We need the solutions of the complete equations to be the solutions of the incomplete data equations. This suggest we take

$$\mathrm{Q}_1\left(\kappa,\theta\right)=\mathrm{Q}_1^*\left(\kappa,\theta;\mathbf{y}\right) \tag{4.23}$$

and

$$\mathrm{Q}_2\left(\kappa,\theta\right)=\mathrm{Q}_2^*\left(\kappa,\theta;\mathbf{y}\right) \tag{4.24}$$

and solve for the values of $\mathrm{fx}_1\left(\kappa,\theta,\mathbf{y}\right)$ and $\mathrm{fx}_2\left(\kappa,\theta,\mathbf{y}\right)$.

Taking the first equation we get

$$\begin{aligned}
\theta\frac{\Gamma\left(\kappa+1\right)}{\Gamma\left(\kappa\right)}-\frac{\mathrm{fx}_1\left(\kappa,\theta,\mathbf{y}\right)}{n}&=\theta\frac{\gamma\left(\kappa+1,c/\theta\right)}{\Gamma\left(\kappa\right)}+c\frac{\Gamma\left(\kappa,c/\theta\right)}{\Gamma\left(\kappa\right)}-\frac{\sum_{i=1}^m y_i+(n-m)c}{n}\\
\mathrm{fx}_1\left(\kappa,\theta,\mathbf{y}\right)&=n\theta\left(\frac{\Gamma\left(\kappa+1\right)-\gamma\left(\kappa+1,c/\theta\right)}{\Gamma\left(\kappa\right)}\right)-nc\frac{\Gamma\left(\kappa,c/\theta\right)}{\Gamma\left(\kappa\right)}\\
&\quad+\sum_{i=1}^m y_i+(n-m)c\\
\mathrm{fx}_1\left(\kappa,\theta,\mathbf{y}\right)&=n\theta\frac{\Gamma\left(\kappa+1,c/\theta\right)}{\Gamma\left(\kappa\right)}-nc\frac{\Gamma\left(\kappa,c/\theta\right)}{\Gamma\left(\kappa\right)}+\sum_{i=1}^m y_i+(n-m)c
\end{aligned} \tag{4.25}$$

and similarly taking the second equation we get

$$\begin{aligned}\theta^2\frac{\Gamma(\kappa+2)}{\Gamma(\kappa)} - \frac{\mathrm{fx}_2(\kappa,\theta,\mathbf{y})}{n} &= \theta^2\frac{\gamma(\kappa+2,c/\theta)}{\Gamma(\kappa)} + c^2\frac{\Gamma(\kappa,c/\theta)}{\Gamma(\kappa)} - \frac{\sum_{i=1}^m y_i^2 + (n-m)c^2}{n} \\ \mathrm{fx}_2(\kappa,\theta,\mathbf{y}) &= n\theta^2\frac{\Gamma(\kappa+2,c/\theta)}{\Gamma(\kappa)} - nc^2\frac{\Gamma(\kappa,c/\theta)}{\Gamma(\kappa)} \\ &\quad + \sum_{i=1}^m y_i^2 + (n-m)c^2 \end{aligned} \tag{4.26}$$

thus at each step we evaluate $\mathrm{fx}_1\left(\kappa_{(t)},\theta_{(t)},\mathbf{y}\right)$ and $\mathrm{fx}_2\left(\kappa_{(t)},\theta_{(t)},\mathbf{y}\right)$ given our current estimates for $\theta_{(t)}$ and $\kappa_{(t)}$. Then we solve

$$\kappa_{(t+1)}\theta_{(t+1)} = \frac{\mathrm{fx}_1\left(\kappa_{(t)},\theta_{(t)},\mathbf{y}\right)}{n} \tag{4.27}$$

and

$$\theta_{(t+1)}^2\kappa_{(t+1)}\left(\kappa_{(t+1)}+1\right) = \frac{\mathrm{fx}_2\left(\kappa_{(t)},\theta_{(t)},\mathbf{y}\right)}{n} \tag{4.28}$$

for $\kappa_{(t+1)}$ and $\theta_{(t+1)}$.

Note that if we were considering left censored data that the appropriate functions would be

$$\mathrm{fx}_1(\kappa,\theta,\mathbf{y}) = n\theta\frac{\gamma(\kappa+1,c/\theta)}{\Gamma(\kappa)} - nc\frac{\gamma(\kappa,c/\theta)}{\Gamma(\kappa)} + \sum_{i=1}^m y_i + (n-m)c \tag{4.29}$$

and

$$\mathrm{fx}_2(\kappa,\theta,\mathbf{y}) = n\theta^2\frac{\gamma(\kappa+2,c/\theta)}{\Gamma(\kappa)} - nc^2\frac{\gamma(\kappa,c/\theta)}{\Gamma(\kappa)} + \sum_{i=1}^m y_i^2 + (n-m)c^2 \tag{4.30}$$

4.5 A "pseudo" EM approach to moment estimation

In the previous section we considered an iterative method for solving the incomplete data moment equations and drew the analogy with the EM algorithm in the sense that we took the incomplete data form and manipulated it into the easier to solve complete data form. In this section we shall again try to draw a similarity with the EM algorithm. But in this case we shall consider the situation where we substitute in any missing values with expected values based upon current parameter estimates.

It seems natural to try to estimate these parameters just as we did in the maximum likelihood case. That is we will substitute the expected value in at an "E-step" and then solve for new parameter estimates at a "M-step".

But noting that the two maximum likelihood sufficient statistics are $\sum x$ and $\sum \ln(x)$ it seems reasonable that we should take equation (3.15) (which we have previously used) and another more general equation for the a'th moment, as the two moment equations which we will solve.

That is we will take

$$\mathrm{E}\left[x|\kappa,\theta\right] = \frac{\sum_{i=1}^{n} x_i}{n}$$

and

$$\mathrm{E}\left[x^a|\kappa,\theta\right] = \frac{\sum_{i=1}^{n} x_i^a}{n}$$

as the two equations which we solve. Where the first equation is given by

$$\kappa\theta = \frac{\sum_{i=1}^{n} x_i}{n}$$

as before. We can derive the left hand side of the second equation

$$\begin{aligned}
\mathrm{E}\left[x^a|\kappa,\theta\right] &= \int_0^\infty x^a \frac{1}{\theta^\kappa \Gamma\left(\kappa\right)} x^{\kappa-1} \exp\left(-\frac{x}{\theta}dx\right) \\
&= \frac{1}{\theta^\kappa \Gamma\left(\kappa\right)} \int_0^\infty x^{\kappa+a-1} \exp\left(-\frac{x}{\theta}\right) dx
\end{aligned}$$

again we substitute in $u = \frac{x}{\theta}$ to get that

$$\begin{aligned}
\mathrm{E}\left[x^a|\kappa,\theta\right] &= \frac{1}{\theta^\kappa \Gamma\left(\kappa\right)} \theta^{a+\kappa} \int_0^\infty u^{a+\kappa-1} \exp\left(-u\right) du \\
&= \frac{\theta^a}{\Gamma\left(\kappa\right)} \Gamma\left(\kappa + a\right)
\end{aligned}$$

The “E-step” of this method will then consist of us evaluating

$$\begin{aligned}
\mathrm{E}\left[x^a|\theta_{(t)},\kappa_{(t)},x>c\right] &= \frac{\displaystyle\int_c^\infty x^a \frac{1}{\theta_{(t)}^{\kappa_{(t)}} \Gamma\left(\kappa_{(t)}\right)} x^{\kappa_{(t)}-1} \exp\left(-\frac{x}{\theta_{(t)}}\right) dx}{\displaystyle\int_c^\infty \frac{1}{\theta_{(t)}^{\kappa_{(t)}} \Gamma\left(\kappa_{(t)}\right)} x^{\kappa_{(t)}-1} \exp\left(-\frac{x}{\theta_{(t)}}\right) dx} \\
&= \frac{\displaystyle\int_c^\infty x^{a+\kappa_{(t)}-1} \exp\left(-\frac{x}{\theta_{(t)}}\right) dx}{\displaystyle\int_c^\infty x^{\kappa_{(t)}-1} \exp\left(-\frac{x}{\theta_{(t)}}\right) dx}
\end{aligned}$$

using the usual substitution $u = \frac{x}{\theta}$ we then get

$$\begin{aligned}
\mathrm{E}\left[x^a|\theta_{(t)},\kappa_{(t)},x>c\right] &= \theta_{(t)}^a \frac{\displaystyle\int_{c/\theta_{(t)}}^\infty u^{a+\kappa_{(t)}-1} \exp\left(-u\right) du}{\displaystyle\int_{c/\theta_{(t)}}^\infty u^{\kappa_{(t)}-1} \exp\left(-u\right) du} \\
&= \theta_{(t)}^a \frac{\Gamma\left(\kappa_{(t)} + a, c/\theta_{(t)}\right)}{\Gamma\left(\kappa_{(t)}, c/\theta_{(t)}\right)}
\end{aligned}$$

At the t+1'th "M-step" we will then solve the equations

$$\kappa_{(t+1)}\theta_{(t+1)} = \frac{1}{n}\left(\sum_{1}^{m} x_i + (n-m)\,\theta_{(t)}\frac{\Gamma\left(\kappa_{(t)}+1, c/\theta_{(t)}\right)}{\Gamma\left(\kappa_{(t)}, c/\theta_{(t)}\right)}\right) \tag{4.31}$$

$$\frac{\theta_{(t+1)}{}^{a}\Gamma\left(\kappa_{(t+1)}+a\right)}{\Gamma\left(\kappa_{(t+1)}\right)} = \frac{1}{n}\left(\sum_{1}^{m} x_i^a + (n-m)\,\theta_{(t)}^a\frac{\Gamma\left(\kappa_{(t)}+a, c/\theta_{(t)}\right)}{\Gamma\left(\kappa_{(t)}, c/\theta_{(t)}\right)}\right) \tag{4.32}$$

for $\kappa_{(t+1)}$ and $\theta_{(t+1)}$.

Similarly for left censored data we would solve the equations

$$\kappa_{(t+1)}\theta_{(t+1)} = \frac{1}{n}\left(\sum_{1}^{m} x_i + (n-m)\,\theta_{(t)}\frac{\gamma\left(\kappa_{(t)}+1, c/\theta_{(t)}\right)}{\gamma\left(\kappa_{(t)}, c/\theta_{(t)}\right)}\right)$$

$$\frac{\theta_{(t+1)}{}^{a}\Gamma\left(\kappa_{(t+1)}+a\right)}{\Gamma\left(\kappa_{(t+1)}\right)} = \frac{1}{n}\left(\sum_{1}^{m} x_i^a + (n-m)\,\theta_{(t)}^a\frac{\gamma\left(\kappa_{(t)}+a, c/\theta_{(t)}\right)}{\gamma\left(\kappa_{(t)}, c/\theta_{(t)}\right)}\right)$$

Note that although we have considered a general a'th moment. Choosing $a = 2$ provides us with a set of equations which can be easily solved.

Chapter 5

Simulations

5.1 About the simulations

We have written a program which takes random samples from a specified gamma distribution. We perform censoring on the sample and then apply each of the three methods of estimation to the data (ie we apply EM, incomplete moments and 'pseudo'-EM moments approaches). We have used samples of size 100, 1000 and 10000. We have looked at censoring 10%, 50% and 80% of the data. We have looked at both left and right tail censoring. In each case we have taken 10000 random samples. Note that we have used the first and second moment equations as these provide algebraically easy equations to solve.

5.2 Results

Tables giving fuller detail of the results can be found in appendix A.

In all cases the EM algorithm gave the the most efficient estimates (it always had the smallest Mean Square Error (MSE)). We found that the efficiences of the two moments methods relative to the EM estimates varied from around 90% when we had only 10% data censored down to around 50% when estimating the parameters on an

80% left censored distribution.

For right censored data we found that at the lower levels of censoring that the incomplete moment method generally gave more efficient estimates than the pseudo-EM moment method but at the higher levels of censoring the pseudo-EM moment method was superior. For left censored data we found that the pseudo-EM moment method gave more efficient estimates than the incomplete moment method in all cases.

In all cases we found that the MSE approached the variance of the estimate as the sample size n increased. We also found that the variances decreased as the sample size increased.

Looking in terms of iterations to convergence we generally found that for right censored data that the EM method provided the method with the least number of iterations and the incomplete moments method gave the estimates that took the longest to converge. We also found that both moments method had a lot more variability in the number of iterations to convergence than the EM method had. For left censored data we found that in some cases the pseudo-EM moments method was quicker than the EM algorithm but again it was more variable. The incomplete moments method was usually slower and always more variable.

From our trials we found that the region of convergence for both the moments methods seemed to be fairly small and that if we started our sequences from values too different from the theoretical parameter values our sequences of parameter estimates appeared to quickly diverge. This raises great concerns about the wide applicability of these methods when we don't have a reasonably good idea of what the parameter values should be.

Chapter 6

Conclusions

As we saw in chapter three the method of moments estimators had a much simpler form for the full data then the maximum likelihood estimators. Because of this we attempted to create EM like methods for method of moment estimators in censored data situations.

Unfortunately, as we found in chapter four, our attempts to construct a iterative methods in the spirit of the EM algorithm for the method of moment estimators did not yield any appreciably simpler formulae. Our simulations in chapter five showed us that the two moment methods that we proposed did yield more efficient estimators and the methods tended to converged slower than the EM algorithm. We also found through our trials that convergence for the two proposed methods was sensitive to the choice of starting values. Because of this the moment methods are not recommended.

Of the two moments methods the pseudo-EM method generally gave the more efficent estimates and it was generally faster than the incomplete moments method.

We see that the EM algorithm is much not more complicated than the two proposed moments methods, it is more well behaved in convergence terms, and gave more efficent estimates. Thus this method would be the recommended method to use when we are trying to estimate the parameters of a gamma distribution when censoring has taken place.

Bibliography

[1] Abramowitz, M and Stegun, I. A. (1964) *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. U.S. Govt. Print. Off.

[2] Atkinson, K.E. (1978) *An Introduction to Numerical Analysis*. John Wiley and Sons Inc.

[3] Bain, L. J. *Gamma Distribution*. Encyclopedia of Statistical Science. V3. pp292-298. John Wiley and Sons Inc.

[4] Bain, L.J and Englehardt, M (1991) *Introduction to probability and mathematical statistics*. PWS-KENT.

[5] Blight, B.J.N. (1970) *Estimation from a censored sample for the exponential family*. Biometrika. **57**, 2 ,pp389-395.

[6] Bowman, K. O. and Shenton, L.R. (1982) *Properties of Estimators for the Gamma Distribution*. Communications in Statistics - Simulation and Computation. **11**, 4, pp377-519.

[7] Choi, S.C. and Wette, R (1969) *Maximum Likelihood Estimation of the Parameters of the Gamma Distribution and Their Bias*. Technometrics. **11**, 4, pp683-690.

[8] Cohen, A.C. and Norgaard, N.J. (1977) *Progressively censored sampling in the three parameter gamma dsitribution*. Technometrics, **19**,3,pp333-340.

[9] Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977) *Maximum Likelihood from Incomplete Data via the EM Algorithm*. Journal of the Royal Statistical Society B. **39**, pp1-23.

[10] Fisher, R.A. (1921) *On the mathematical foundations of theoretical statistics*. Philosophical Transactions of the Royal Society London A,**222**, pp309-368.

[11] Gilchrist, R. (1981) *Estimation of the parameters of the gamma distribution by means of a maximal invariant*. Communications in Statistics, A, 10, pp1095-1110.

[12] Greenwood, J.A. and Durand, D (1960) *Aids for fitting the Gamma distribution by Maximum Likelihood*. Technometrics. **2**, pp55.

[13] Harter, H.C. and Moore, A.H. (1965) *Maximum likelihood estimation of the parameters of the gamma and weibel distributions from complete and censored samples*. Technometrics, **7**, pp639-643

[14] Heyde, C.C. and Morton, R (1996) *Quasi-likelihood and Generalizing the EM Algorithm*. Journal of the Royal Statistical Society B, **58**, 2, pp317-327.

[15] Kendall, M and Stuart, A (1977) *The Advanced Theory of Statistics Vol 2* Charles Griffin & Company Limited.

[16] Laird, N (1985) *Missing Information Principle*. Encyclopedia of Statistical Science. V5. pp548-552. John Wiley and Sons Inc.

[17] Little, R.J.A. and Rubin, D.B. (1987) *Statistical analysis with missing data*. John Wiley and Sons Inc.

[18] McLachlan, G.J. and Krishnan,T (1996) *The EM Algorithm and Extensions*. John Wiley and Sons Inc.

[19] Moore, R.J. (1982) *Algorithm AS187 Derivatives of the Incomplete Gamma Integral*. Applied Statistics. **31**, 3, pp330-333.

[20] NAG (1993) *The NAG Fortran Library Manual, Mark 16*. The Numerical Algorithms Group Limited.

[21] Shenton, L.R. and Bowman, K.O. (1977) *Maximum likelihood estimation in small samples*. Charles Griffen & Company ltd.

[22] Thom, H.C.S. (1958) *A note on the gamma distribution*. Monthly Weather Review. **86**, 4, pp117-122

Appendix A

Results

This appendix contains the results of the simulations. What follows is a series of data tables containing the results of each of the simulations. We have recorded the mean, variance and MSE of each estimator based upon our sample of 10000 estimates. We have looked at both right and left tail censored. We have looked at 10%, 50% and 80% levels of censoring to represent low, medium and high amounts of censoring. We have also looked at samples of size 100, 1000 and 10000 to represent small, medium and large samples. To ensure convergence we have started our iterations from the true parameter values.

In the tables we use RIGHT or LEFT to indicate which tail was censored. We have used the following set of abbreviations for the methods.

EM EM algorithm

IM Incomplete Moments Method

PS 'Pseudo'- EM Moments Method

A.1 Gamma $\kappa = 2.5$,$\theta = 2.0$

A.1.1 n=100

Estimates for κ					
Tail	% censored	Method	Mean	Var	MSE
RIGHT	10	EM	2.5597	0.1339	0.1374
		IM	2.5279	0.1556	0.1564
		PS	2.5562	0.1777	0.1808
	50	EM	2.6206	0.2619	0.2764
		IM	2.5843	0.2903	0.2974
		PS	2.6197	0.3104	0.3247
	80	EM	2.8196	0.8295	0.9315
		IM	2.7650	0.9540	1.0241
		PS	2.8092	0.9001	0.9956
LEFT	10	EM	2.5587	0.1438	0.1473
		IM	2.6051	0.2217	0.2327
		PS	2.6054	0.1979	0.2090
	50	EM	2.6429	0.3826	0.4030
		IM	2.7507	0.6568	0.7196
		PS	2.6934	0.4461	0.4834
	80	EM	3.0545	2.3225	2.6297
		IM	3.4052	4.0303	4.8493
		PS	3.1376	2.4850	2.8912

Estimates for θ					
Tail	% censored	Method	Mean	Var	MSE
RIGHT	10	EM	1.9986	0.1011	0.1011
		IM	2.0379	0.1249	0.1264
		PS	2.0161	0.1344	0.1346
	50	EM	1.9978	0.2486	0.2486
		IM	2.0610	0.3224	0.3261
		PS	2.0218	0.3125	0.3129
	80	EM	2.0243	0.9516	0.9521
		IM	2.1906	1.4239	1.4601
		PS	2.0905	1.2365	1.2445
LEFT	10	EM	1.9944	0.0974	0.0974
		IM	1.9795	0.1432	0.1436
		PS	1.9739	0.1295	0.1302
	50	EM	1.9845	0.1843	0.1845
		IM	1.9658	0.2938	0.2950
		PS	1.9635	0.2127	0.2140
	80	EM	1.9464	0.4660	0.4688
		IM	1.9061	0.7429	0.7516
		PS	1.9155	0.4899	0.4970

Number of iterations							
Tail	% censored	Method	Mean	Median	Stdev	Min	Max
RIGHT	10	EM	8.9985	9	0.0383	8	9
		IM	28.8670	29	5.3254	10	52
		PS	11.9927	12	1.1513	4	15
	50	EM	33.9574	34	0.7782	31	37
		IM	352.3146	336	101.6509	3	839
		PS	60.9131	61	8.7850	11	109
	80	EM	144.7734	144	7.9897	107	185
		IM	3572.7121	3295	1814.4008	3	10001
		PS	358.6349	339	92.2012	4	1079
LEFT	10	EM	9.0757	9	0.2646	9	10
		IM	8.1623	8	1.0609	3	13
		PS	5.8940	6	0.5021	3	7
	50	EM	36.0349	36	0.9026	32	40
		IM	32.8403	32	6.0109	2	60
		PS	14.6055	14	2.0073	4	22
	80	EM	148.9143	149	4.9100	125	167
		IM	173.9111	163	57.6051	5	458
		PS	48.6768	44	12.6076	12	106

A.1.2 n=1000

Estimates for κ					
Tail	% censored	Method	Mean	Var	MSE
RIGHT	10	EM	2.5586	0.1355	0.1389
		IM	2.5306	0.1580	0.1589
		PS	2.5595	0.1815	0.1850
	50	EM	2.6133	0.2596	0.2724
		IM	2.5781	0.2891	0.2952
		PS	2.6171	0.3082	0.3219
	80	EM	2.7990	0.8480	0.9373
		IM	2.7380	0.9837	1.0403
		PS	2.7931	0.9098	0.9956
LEFT	10	EM	2.5606	0.1447	0.1483
		IM	2.6066	0.2216	0.2329
		PS	2.6068	0.1981	0.2094
	50	EM	2.6471	0.3970	0.4186
		IM	2.7501	0.6636	0.7261
		PS	2.6956	0.4562	0.4944
	80	EM	3.0385	2.2167	2.5065
		IM	3.3835	3.7754	4.5555
		PS	3.1220	2.3570	2.7436

Estimates for θ					
Tail	% censored	Method	Mean	Var	MSE
RIGHT	10	EM	2.0001	0.1039	0.1039
		IM	2.0367	0.1297	0.1310
		PS	2.0153	0.1417	0.1419
	50	EM	2.0038	0.2465	0.2465
		IM	2.0683	0.3244	0.3290
		PS	2.0235	0.3120	0.3125
	80	EM	2.0500	0.9759	0.9783
		IM	2.2322	1.5027	1.5564
		PS	2.1089	1.2561	1.2679
LEFT	10	EM	1.9937	0.0994	0.0995
		IM	1.9788	0.1441	0.1445
		PS	1.9733	0.1307	0.1314
	50	EM	1.9845	0.1892	0.1894
		IM	1.9661	0.2912	0.2923
		PS	1.9632	0.2130	0.2143
	80	EM	1.9539	0.4731	0.4752
		IM	1.9062	0.7170	0.7257
		PS	1.9196	0.4870	0.4935

Number of iterations							
Tail	% censored	Method	Mean	Median	Stdev	Min	Max
RIGHT	10	EM	9.0000	9	0.0000	9	9
		IM	21.5246	22	2.5290	3	27
		PS	10.3115	11	0.9276	3	12
	50	EM	37.4760	37	0.9535	36	40
		IM	199.3644	204	34.3990	3	301
		PS	44.0936	45	5.4441	5	56
	80	EM	153.1502	153	1.9373	148	158
		IM	1614.8037	1685	459.8888	5	2987
		PS	220.3977	227	33.9559	4	307
LEFT	10	EM	11.0000	11	0.0000	11	11
		IM	9.1838	9	0.8400	4	11
		PS	6.2205	6	0.5441	4	7
	50	EM	43.2367	43	1.1247	41	46
		IM	39.9266	41	4.8442	2	54
		PS	16.3411	17	1.6522	5	21
	80	EM	169.6356	170	1.6965	164	174
		IM	207.4745	210	34.3153	2	315
		PS	54.1815	54	7.2457	4	75

A.1.3 n=10000

Estimates for κ					
Tail	% censored	Method	Mean	Var	MSE
RIGHT	10	EM	2.5010	1.266E-03	1.267E-03
		IM	2.5009	1.496E-03	1.497E-03
		PS	2.5013	1.735E-03	1.737E-03
	50	EM	2.5009	2.179E-03	2.180E-03
		IM	2.5007	2.444E-03	2.444E-03
		PS	2.5011	2.672E-03	2.673E-03
	80	EM	2.5023	5.100E-03	5.105E-03
		IM	2.5013	5.330E-03	5.331E-03
		PS	2.5028	5.906E-03	5.913E-03
LEFT	10	EM	2.5010	1.326E-03	1.327E-03
		IM	2.5013	2.091E-03	2.093E-03
		PS	2.5014	1.862E-03	1.864E-03
	50	EM	2.5021	3.163E-03	3.167E-03
		IM	2.5025	5.586E-03	5.592E-03
		PS	2.5023	3.723E-03	3.728E-03
	80	EM	2.5042	1.131E-02	1.133E-02
		IM	2.5072	2.114E-02	2.119E-02
		PS	2.5049	1.237E-02	1.239E-02

Estimates for θ					
Tail	% censored	Method	Mean	Var	MSE
RIGHT	10	EM	1.9994	1.055E-03	1.055E-03
		IM	1.9997	1.221E-03	1.221E-03
		PS	1.9993	1.379E-03	1.380E-03
	50	EM	2.0001	2.434E-03	2.434E-03
		IM	2.0006	2.815E-03	2.816E-03
		PS	2.0002	2.918E-03	2.917E-03
	80	EM	2.0009	8.590E-03	8.590E-03
		IM	2.0027	8.754E-03	8.760E-03
		PS	2.0009	9.874E-03	9.874E-03
LEFT	10	EM	1.9994	9.952E-04	9.955E-04
		IM	1.9993	1.450E-03	1.450E-03
		PS	1.9993	1.331E-03	1.332E-03
	50	EM	1.9991	1.877E-03	1.878E-03
		IM	1.9994	2.954E-03	2.954E-03
		PS	1.9991	2.177E-03	2.177E-03
	80	EM	1.9994	4.827E-03	4.827E-03
		IM	1.9995	8.065E-03	8.065E-03
		PS	1.9992	5.247E-03	5.247E-03

Number of iterations							
Tail	% censored	Method	Mean	Median	Stdev	Min	Max
RIGHT	10	EM	10.0000	10	0.0000	10	10
		IM	21.6129	22	3.1264	2	27
		PS	9.8673	10	1.0596	4	12
	50	EM	41.5119	42	0.5498	40	42
		IM	225.6838	238	57.2103	3	337
		PS	45.8689	48	7.3581	3	59
	80	EM	175.1161	175	1.3918	172	180
		IM	1392.2932	1769.5	1048.6980	3	3370
		PS	248.2124	261	59.0964	2	356
LEFT	10	EM	10.9945	11	0.0743	10	11
		IM	6.8625	7	0.6498	2	8
		PS	5.0036	5	0.3556	3	6
	50	EM	43.6506	44	0.7211	42	45
		IM	25.4814	26	3.3371	2	32
		PS	11.7956	12	1.3122	4	14
	80	EM	168.8744	169	1.3138	166	173
		IM	121.1537	125	20.6435	5	165
		PS	35.6248	37	4.9716	2	47

A.2 Gamma $\kappa = 5,\ \theta = 2$

A.2.1 n=100

Estimates for κ					
Tail	% censored	Method	Mean	Var	MSE
RIGHT	10	EM	5.1295	0.6053	0.6220
		IM	5.0752	0.6453	0.6508
		PS	5.1207	0.6901	0.7046
	50	EM	5.2527	1.2385	1.3022
		IM	5.1978	1.3757	1.4145
		PS	5.2406	1.3354	1.3931
	80	EM	5.7124	4.2696	4.7767
		IM	5.6810	5.2753	5.7383
		PS	5.6762	4.3589	4.8158
LEFT	10	EM	5.1260	0.6192	0.6350
		IM	5.1806	0.8332	0.8657
		PS	5.1833	0.7329	0.7664
	50	EM	5.3104	1.6438	1.7401
		IM	5.4545	2.5094	2.7156
		PS	5.3690	1.7748	1.9108
	80	EM	6.0468	8.3423	9.4372
		IM	6.6248	14.2975	16.936
		PS	6.1571	8.7758	10.114

Estimates for θ					
Tail	% censored	Method	Mean	Var	MSE
RIGHT	10	EM	1.9973	0.0998	0.0998
		IM	2.0266	0.1139	0.1146
		PS	2.0082	0.1177	0.1178
	50	EM	1.9999	0.2305	0.2305
		IM	2.0474	0.2908	0.2930
		PS	2.0158	0.2610	0.2612
	80	EM	2.0049	0.7585	0.7585
		IM	2.1232	1.1578	1.1728
		PS	2.0451	0.8797	0.8816
LEFT	10	EM	1.9946	0.0957	0.0957
		IM	1.9873	0.1258	0.1260
		PS	1.9803	0.1109	0.1113
	50	EM	1.9825	0.1903	0.1905
		IM	1.9750	0.2749	0.2755
		PS	1.9684	0.2015	0.2025
	80	EM	1.9512	0.4776	0.4797
		IM	1.9254	0.7285	0.7337
		PS	1.9277	0.4812	0.4864

Number of iterations							
Tail	% censored	Method	Mean	Median	Stdev	Min	Max
RIGHT	10	EM	9.0000	9	0.0000	9	9
		IM	24.4915	24	4.3730	5	42
		PS	11.2202	11	0.9558	4	13
	50	EM	33.6454	34	0.6515	31	36
		IM	237.4755	230	57.3383	5	555
		PS	49.6074	50	5.7598	18	71
	80	EM	138.3584	138	3.6902	124	152
		IM	2105.5266	1910	849.4660	6	8137
		PS	253.3637	249	42.3322	6	462
LEFT	10	EM	10.0000	10	0.0000	10	10
		IM	9.9839	10	1.3949	5	16
		PS	6.7254	7	0.5792	4	8
	50	EM	39.1827	39	1.1036	36	44
		IM	44.8746	45	7.8313	16	89
		PS	18.1064	18	2.2328	9	26
	80	EM	157.3518	157	4.0855	141	174
		IM	246.0227	233	70.7517	8	629
		PS	63.1476	57	14.0392	17	126

A.2.2 n=1000

Estimates for κ					
Tail	% censored	Method	Mean	Var	MSE
RIGHT	10	EM	5.0083	0.0535	0.0535
		IM	5.0024	0.0587	0.0587
		PS	5.0093	0.0637	0.0637
	50	EM	5.0167	0.0992	0.0995
		IM	5.0113	0.1110	0.1111
		PS	5.0162	0.1112	0.1114
	80	EM	5.0551	0.2486	0.2516
		IM	5.0471	0.2901	0.2923
		PS	5.0546	0.2697	0.2726
LEFT	10	EM	5.0087	0.0552	0.0552
		IM	5.0155	0.0768	0.0771
		PS	5.0155	0.0669	0.0671
	50	EM	5.0285	0.1340	0.1348
		IM	5.0448	0.2146	0.2166
		PS	5.0350	0.1467	0.1479
	80	EM	5.0952	0.4574	0.4664
		IM	5.1489	0.8023	0.8243
		PS	5.1060	0.4850	0.4962

Estimates for θ					
Tail	% Censored	Method	Mean	Var	MSE
RIGHT	10	EM	2.0008	0.0097	0.0097
		IM	2.0041	0.0108	0.0108
		PS	2.0013	0.0114	0.0114
	50	EM	2.0023	0.0218	0.0218
		IM	2.0068	0.0255	0.0255
		PS	2.0038	0.0245	0.0245
	80	EM	2.0028	0.0714	0.0715
		IM	2.0146	0.0881	0.0884
		PS	2.0060	0.0783	0.0784
LEFT	10	EM	2.0003	0.0094	0.0094
		IM	1.9992	0.0126	0.0126
		PS	1.9986	0.0113	0.0113
	50	EM	1.9983	0.0189	0.0189
		IM	1.9973	0.0285	0.0285
		PS	1.9967	0.0207	0.0207
	80	EM	1.9937	0.0495	0.0496
		IM	1.9937	0.0822	0.0822
		PS	1.9916	0.0525	0.0526

Number of iterations							
Tail	% censored	Method	Mean	Median	Stdev	Min	Max
RIGHT	10	EM	9.0000	9	0.0000	9	9
		IM	21.5246	22	2.5290	3	27
		PS	10.3115	11	0.9276	3	12
	50	EM	37.4760	37	0.9535	36	40
		IM	199.3644	204	34.3990	3	301
		PS	44.0936	45	5.4441	5	55
	80	EM	153.1502	153	1.9373	148	158
		IM	1614.8037	1685	459.8888	5	2987
		PS	220.3977	227	33.9559	4	307
LEFT	10	EM	11.0000	11	0.0000	11	11
		IM	9.1838	9	0.8400	4	11
		PS	6.2205	6	0.5441	4	7
	50	EM	43.2367	43	1.1247	41	46
		IM	39.9266	41	4.8442	2	54
		PS	16.3411	17	1.6522	5	21
	80	EM	169.6356	170	1.6965	164	174
		IM	207.4745	210	34.3153	2	304
		PS	54.1815	54	7.2457	4	75

A.2.3 n=10000

Estimates for κ					
Tail	% censored	Method	Mean	Var	MSE
RIGHT	10	EM	5.0008	0.0054	0.0054
		IM	5.0000	0.0059	0.0059
		PS	5.0008	0.0064	0.0064
	50	EM	5.0015	0.0099	0.0099
		IM	5.0013	0.0110	0.0110
		PS	5.0010	0.0111	0.0111
	80	EM	5.0068	0.0240	0.0240
		IM	5.0051	0.0276	0.0276
		PS	5.0067	0.0261	0.0261
LEFT	10	EM	5.0010	0.0056	0.0056
		IM	5.0018	0.0076	0.0076
		PS	5.0018	0.0066	0.0066
	50	EM	5.0035	0.0130	0.0130
		IM	5.0055	0.0207	0.0208
		PS	5.0043	0.0142	0.0142
	80	EM	5.0106	0.0443	0.0444
		IM	5.0175	0.0750	0.0753
		PS	5.0121	0.0462	0.0464

Estimates for θ					
Tail	% censored	Method	Mean	Var	MSE
RIGHT	10	EM	2.0003	0.0010	0.0010
		IM	2.0007	0.0011	0.0011
		PS	2.0004	0.0011	0.0011
	50	EM	2.0006	0.0022	0.0022
		IM	2.0008	0.0026	0.0026
		PS	2.0009	0.0025	0.0025
	80	EM	1.9997	0.0071	0.0071
		IM	2.0014	0.0085	0.0085
		PS	2.0001	0.0077	0.0077
LEFT	10	EM	2.0002	0.0010	0.0010
		IM	2.0000	0.0012	0.0012
		PS	1.9999	0.0011	0.0011
	50	EM	1.9998	0.0018	0.0019
		IM	1.9996	0.0028	0.0028
		PS	1.9996	0.0020	0.0020
	80	EM	1.9992	0.0050	0.0050
		IM	1.9986	0.0079	0.0079
		PS	1.9988	0.0053	0.0052

Number of iterations							
Tail	% censored	Method	Mean	Median	Stdev	Min	Max
RIGHT	10	EM	10.0000	10	0.0000	10	10
		IM	19.1484	20	2.3764	4	24
		PS	9.4025	10	0.8900	4	11
	50	EM	41.1500	41	0.5531	40	45
		IM	166.5501	173	32.0708	2	221
		PS	38.8221	40	5.1399	3	48
	80	EM	167.4540	167	2.4600	165	178
		IM	1209.0284	1318	444.0017	4	1947
		PS	187.2667	195	33.9252	3	246
LEFT	10	EM	11.9003	12	0.2997	11	12
		IM	8.3928	9	0.7995	2	10
		PS	5.8139	6	0.4345	3	7
	50	EM	47.4641	47	0.7917	46	51
		IM	35.0332	36	4.7795	3	43
		PS	14.7069	15	1.5822	7	18
	80	EM	182.5744	183	2.6555	176	188
		IM	177.4378	184	29.4090	4	238
		PS	47.5602	49	6.2978	3	61