Journal of the Korean Data & Information Science Society
2010, **21**(2), 345–354

한국데이터정보과학회지

Likelihood ratio in estimating gamma distribution parameters

Mezbahur Rahman[1] · S. M. Muraduzzaman[2]

[1]Minnesota State University, Mankato, USA and BRAC University
[2]Bangladesh Institute of Health Science

Received 2 February 2010, revised 17 March 2010, accepted 22 March 2010

Abstract

The Gamma Distribution is widely used in Engineering and Industrial applications. Estimation of parameters is revisited in the two-parameter Gamma distribution. The parameters are estimated by minimizing the likelihood ratios. A comparative study between the method of moments, the maximum likelihood method, the method of product spacings, and minimization of three different likelihood ratios is performed using simulation. For the scale parameter, the maximum likelihood estimate performs better and for the shape parameter, the product spacings estimate performs better. Among the three likelihood ratio statistics considered, the Anderson-Darling statistic has inferior performance compared to the Cramer-von-Misses statistic and the Kolmogorov-Smirnov statistic.

Keywords: Di-gamma function, grid search method.

1. Introduction

The random variable X has a Gamma distribution with two parameters β and α if it has a probability density function of the form:

$$f(x;\beta,\alpha)=\frac{x^{\alpha-1}e^{-x/\beta}}{\Gamma(\alpha)\beta^{\alpha}};\quad \beta>0,\quad \alpha>0, \tag{1.1}$$

where α is known as the shape parameter and β as the scale parameter. The distribution function of the Gamma distribution (1.1) can be written as

$$F(x;\beta,\alpha)=\int_0^x \frac{t^{\alpha-1}e^{-t/\beta}}{\Gamma(\alpha)\beta^{\alpha}}dt;\quad \beta>0,\quad \alpha>0. \tag{1.2}$$

The random variables $X_{1:n}, X_{2:n}, \cdots, X_{n:n}$ are defined as an ordered random sample from the Gamma distribution (1.1).

[1] Corresponding author: Minnesota State University, Mankato, USA and BRAC University, Bangladesh (visiting on sabbatical leave). E-mail: mezbahur.rahman@mnsu.edu
[2] Bangladesh Institute of Health Science, Dhaka. E-mail: murad@hcdp-bd.org

In the literature, estimation of parameters in the two parameter Gamma distribution is discussed extensively. Readers are referred to the following references: Harter and Moore (1965), Choi and Wette (1969), Wilks (1990), Lee (1992), Dang and Weerakkody (2000), and Evans *et al.*, (2000). Rahman *et al.* (2007) implemented the method of product spacings in estimating parameters in a two parameter Gamma distribution. Recently, Zhang and Wu (2005) developed three different versions of likelihood ratio tests and implemented for testing normality. These tests are more powerful than the traditional counterparts, the Kolmogorov-Smirnov test, the Cramer von-Mises test, and the Anderson-Darling test.

In this paper, it is shown that the likelihood ratio test statistics can also be successfully used in estimating papameters. As the likelihood ratio tests use the distribution functions instead of the density function, might have advantages in certain situations.

1.1. Likelihood-ratio tests

Let X be a continuous random variable with distribution function $F(x)$, and $X_1, X_2, \dots, X_n$ be a random sample from $F(x)$ with order statistics $X_{(1)}$, $X_{(2)}$, ..., $X_{(n))}$. One may wish to test the null hypothesis

$$H_0 : F(x) = F_0(x) \text{ for all } x \in (-\infty, \infty)$$

$$H_1 : F(x) \neq F_0(x) \text{ for some } x \in (-\infty, \infty),$$

where $F_0(x)$ is a hypothetical distribution function which is completely specified.

Zhang and Wu (2005) developed three different versions of likelihood-ratio tests and implemented for the tests for normality along with some power comparisons. The tests are as follows:

Likelihood-ratio Kolmogorov-Smirnov statistic $LK =$

$$\max_{i \in \{1,2,\cdots,n\}} \left\{(i-0.5)\log\frac{i-0.5}{nF_0(X_{(i)})} + (n-i+0.5)\log\frac{n-i+0.5}{n[1-F_0(X_{(i)})]}\right\}, \tag{1.3}$$

Likelihood-ratio Cramer-von-Mises statistic

$$LC = \sum_{i=1}^{n} \left\{\log\frac{F_0(X_{(i)})^{-1}-1}{(n-0.5)/(i-0.75)-1}\right\}^2 \tag{1.4}$$

and Likelihood-ratio Anderson-Darling statistic

$$LA = -\sum_{i=1}^{n} \left\{\frac{\log F_0(X_{(i)})}{n-i+0.5} + \frac{\log[1-F_0(X_{(i)})]}{i-0.5}\right\}. \tag{1.5}$$

Zhang and Wu (2005) showed that these tests are more powerful than the traditional Kolmogorov-Smirnov test, Cramer-von-Mises test, and Anderson-Darling test, respectively.

The organization of the paper is as follows: Different estimation procedures are presented in Section 2. In Section 3, a comparison study is conducted using simulation. An application is presented in Section 4. Finally, a concluding summary is presented in Section 5.

2. Estimation procedures

2.1. Method of moment estimates (MME)

The method of moment estimates for β and α are respectively,

$$\hat{\beta}_M = \frac{S^2}{\bar{X}}$$

and

$$\hat{\alpha}_M = \left(\frac{\bar{X}}{S}\right)^2$$

where $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$.

2.2. Maximum likelihood estimates (MLE)

The maximum likelihood estimates for β and α are respectively,

$$\hat{\beta}_L = \frac{\bar{X}}{\hat{\alpha}_L}$$

with $\hat{\alpha}_L$ found as the solution of the following non-linear equation

$$\log\hat{\alpha}_L - \Psi(\hat{\alpha}_L) = \log\left[\bar{X}/\left(\prod_{i=1}^{n} X_i\right)^{\frac{1}{n}}\right] \tag{2.1}$$

where $\Psi(\alpha) = \frac{\Gamma'(\alpha)}{\Gamma(\alpha)}$ and $\Gamma'(\alpha)$ is the derivative of $\Gamma(\alpha)$ with respect to α. $\Psi(\alpha)$ is also known as the Di-gamma function.

The solution of (2.1) can easily be obtained using the Newton-Raphson method with $\hat{\alpha}_M$ as the starting value for $\hat{\alpha}_L$.

Alternately, the log-likelihood function

$$\ln L(x;\beta,\alpha) = (\alpha - 1)\sum_{i=1}^{n}\ln x_i - \frac{1}{\beta}\sum_{i=1}^{n} x_i - n\ln\Gamma(\alpha) - n\alpha\ln\beta, \tag{2.2}$$

where 'ln' stands for the natural logarithm is maximized using a grid search approach in the range of $\left(\max\left\{0, \hat{\beta}_M - 2\hat{\sigma}_{\hat{\beta}_M}\right\}, \hat{\beta}_M + 2\hat{\sigma}_{\hat{\beta}_M}\right)$ and $(\max\{0, \hat{\alpha}_M - 2\hat{\sigma}_{\hat{\alpha}_M}\}, \hat{\alpha}_M + 2\hat{\sigma}_{\hat{\alpha}_M})$ can be used. Here, usual maximization involving the Di-gamma function and it's derivative are avoided to keep consistency with the methods discussed below. In simulation, the estimates of the standard errors $\hat{\sigma}_{\hat{\beta}_M}$ and $\hat{\sigma}_{\hat{\alpha}_M}$ are obtained by implementing the method of moments estimates in separate prior to implementing the other estimates. In application to a data, a suitable range around the moment estimates can be used instead.

2.3. Method of product spacings (MPS)

The method of product spacings (MPS) was concurrently introduced by Cheng and Amin (1983) and Ranneby (1984). Let

$$D_i = \int_{x_{i-1:n}}^{x_{i:n}} f(x;\theta)dx, \quad i = 1, 2, \cdots, n+1,$$

where $x_{0:n}$ is the lower limit and $x_{n+1:n}$ is the upper limit of the domain of the density function $f(x;\theta)$, and θ can be vector-valued. Also, $X_{1:n}, X_{2:n}, \cdots, X_{n:n}$ are defined as an ordered random sample from $f(x;\theta)$. Clearly, the spacings sum to unity, that is $\sum D_i = 1$. The MPS method is, quite simply, to choose θ to maximize the geometric mean of the spacings,

$$G = \left(\prod_{i=1}^{n+1} D_i\right)^{\frac{1}{n+1}}$$

or, equivalently, its logarithm

$$H = \ln G.$$

MPS estimation gives consistent estimators under much more general conditions than MLEs. MPS estimators are asymptotically normal and are asymptotically as efficient as MLEs when these exist. For detailed goodness properties of MPS estimators, readers are referred to Cheng and Amin (1983), Ranneby (1984), Cheng and Iles (1987), Shah and Gokhale (1993), Rahman and Pearson (2002) and the references therein. Using the density function (1) and the cdf (2), H can be written as follows:

$$\begin{aligned} H &= \tfrac{1}{n+1}\left[\ln F(X_{1:n};\beta,\alpha) + \ln\left\{1 - F(X_{n:n};\beta,\alpha)\right\}\right] \\ &\quad + \tfrac{1}{n+1}\left[\sum_{i=1}^{n-1} \ln\left\{F(X_{i+1:n};\beta,\alpha) - F(X_{i:n};\beta,\alpha)\right\}\right] \end{aligned} \tag{2.3}$$

By maximizing (2.3) for different values of β and α, the MPS estimates can be obtained as $\hat{\beta}_P$ and $\hat{\alpha}_P$. The Newton-Raphson method can be used in solving when the two first derivatives are equal to zero. The MME's are used as the starting values. The first derivatives of H with respect to β and α are respectively,

$$\begin{aligned} H'_\beta &= \tfrac{1}{n+1}\left[\tfrac{F'_\beta(X_{1:n};\beta,\alpha)}{F(X_{1:n};\beta,\alpha)} + \sum_{i=1}^{n-1} \tfrac{F'_\beta(X_{i+1:n};\beta,\alpha) - F'_b(X_{i:n};\beta,\alpha)}{F(X_{i+1:n};\beta,\alpha) - F(X_{i:n};\beta,\alpha)} \right. \\ &\quad \left. - \tfrac{F'_\beta(X_{n:n};\beta,\alpha)}{1 - F(X_{n:n};\beta,\alpha)}\right] \end{aligned} \tag{2.4}$$

and

$$\begin{aligned} H'_\alpha &= \tfrac{1}{n+1}\left[\tfrac{F'_\alpha(X_{1:n};\beta,\alpha)}{F(X_{1:n};\beta,\alpha)} + \sum_{i=1}^{n-1} \tfrac{F'_\alpha(X_{i+1:n};\beta,\alpha) - F'_\alpha(X_{i:n};\beta,\alpha)}{F(X_{i+1:n};\beta,\alpha) - F(X_{i:n};\beta,\alpha)} \right. \\ &\quad \left. - \tfrac{F'_\alpha(X_{n:n};\beta,\alpha)}{1 - F(X_{n:n};\beta,\alpha)}\right] \end{aligned} \tag{2.5}$$

where

$$F'_\beta(x;\beta,\alpha) = \frac{\alpha}{\beta}\left[F(x;\beta,\alpha+1) - F(x;\beta,\alpha)\right],$$

$$F'_\alpha(x;\beta,\alpha) = E_x(\ln x;\beta,\alpha) - F(x;\beta,\alpha)(\ln\beta + \Psi(\alpha)),$$

and

$$E_x(\ln x;\beta,\alpha)=\int_0^x \ln t\frac{t^{\alpha-1}e^{-t/\beta}}{\Gamma(\alpha)\beta^\alpha}dt.$$

The second derivatives of H with respect to β and α are respectively,

$$\begin{aligned}H''_{\beta\beta} \quad = \quad & \tfrac{1}{n+1}\Big[\tfrac{F(X_{1:n};\beta,\alpha)F''_{\beta\beta}(X_{1:n};\beta,\alpha)-\left\{F'_\beta(X_{1:n};\beta,\alpha)\right\}^2}{\{F(X_{1:n};\beta,\alpha)\}^2} \\ & +\sum_{i=1}^{n-1}\tfrac{\{F(X_{i+1:n};\beta,\alpha)-F(X_{i:n};\beta,\alpha)\}\left\{F''_{\beta\beta}(X_{i+1:n};\beta,\alpha)-F''_{\beta\beta}(X_{i:n};\beta,\alpha)\right\}}{\{F(X_{i+1:n};\beta,\alpha)-F(X_{i:n};\beta,\alpha)\}^2} \\ & -\tfrac{\left\{F'_\beta(X_{i+1:n};\beta,\alpha)-F'_\beta(X_{i:n};\beta,\alpha)\right\}^2}{\{F(X_{i+1:n};\beta,\alpha)-F(X_{i:n};\beta,\alpha)\}^2} \\ & -\tfrac{\{1-F(X_{n:n};\beta,\alpha)\}F''_{\beta\beta}(X_{n:n};\beta,\alpha)+\left\{F'_\beta(X_{n:n};\beta,\alpha)\right\}^2}{\{1-F(X_{n:n};\beta,\alpha)\}^2}\Big],\end{aligned} \tag{2.6}$$

$$\begin{aligned}H''_{\beta\alpha} \quad = \quad & \tfrac{1}{n+1}\Big[\tfrac{F(X_{1:n};\beta,\alpha)F''_{\beta\alpha}(X_{1:n};\beta,\alpha)-F'_\beta(X_{1:n};\beta,\alpha)F'_\alpha(X_{1:n};\beta,\alpha)}{\{F(X_{1:n};\beta,\alpha)\}^2} \\ & +\sum_{i=1}^{n-1}\tfrac{\{F(X_{i+1:n};\beta,\alpha)-F(X_{i:n};\beta,\alpha)\}\left\{F''_{\beta\alpha}(X_{i+1:n};\beta,\alpha)-F''_{\beta\alpha}(X_{i:n};\beta,\alpha)\right\}}{\{F(X_{i+1:n};\beta,\alpha)-F(X_{i:n};\beta,\alpha)\}^2} \\ & -\tfrac{\left\{F'_\beta(X_{i+1:n};\beta,\alpha)-F'_\beta(X_{i:n};\beta,\alpha)\right\}\left\{F'_\alpha(X_{i+1:n};\beta,\alpha)-F'_\alpha(X_{i:n};\beta,\alpha)\right\}}{\{F(X_{i+1:n};\beta,\alpha)-F(X_{i:n};\beta,\alpha)\}^2} \\ & -\tfrac{\{1-F(X_{n:n};\beta,\alpha)\}F''_{\beta\alpha}(X_{n:n};\beta,\alpha)+F'_\beta(X_{n:n};\beta,\alpha)F'_\alpha(X_{n:n};\beta,\alpha)}{\{1-F(X_{n:n};\beta,\alpha)\}^2}\Big]\end{aligned} \tag{2.7}$$

and

$$\begin{aligned}H''_{\alpha\alpha} \quad = \quad & \tfrac{1}{n+1}\Big[\tfrac{F(X_{1:n};\beta,\alpha)F''_{\alpha\alpha}(X_{1:n};\beta,\alpha)-\left\{F'_\alpha(X_{1:n};\beta,\alpha)\right\}^2}{\{F(X_{1:n};\beta,\alpha)\}^2} \\ & +\sum_{i=1}^{n-1}\tfrac{\{F(X_{i+1:n};\beta,\alpha)-F(X_{i:n};\beta,\alpha)\}\left\{F''_{\alpha\alpha}(X_{i+1:n};\beta,\alpha)-F''_{\alpha\alpha}(X_{i:n};\beta,\alpha)\right\}}{\{F(X_{i+1:n};\beta,\alpha)-F(X_{i:n};\beta,\alpha)\}^2} \\ & -\tfrac{\left\{F'_\alpha(X_{i+1:n};\beta,\alpha)-F'_\alpha(X_{i:n};\beta,\alpha)\right\}^2}{\{F(X_{i+1:n};\beta,\alpha)-F(X_{i:n};\beta,\alpha)\}^2} \\ & -\tfrac{\{1-F(X_{n:n};\beta,\alpha)\}F''_{\alpha\alpha}(X_{n:n};\beta,\alpha)+\left\{F'_\alpha(X_{n:n};\beta,\alpha)\right\}^2}{\{1-F(X_{n:n};\beta,\alpha)\}^2}\Big],\end{aligned} \tag{2.8}$$

where

$$F''_{\beta\beta}(x;\beta,\alpha)=\frac{\alpha(\alpha+1)}{\beta^2}\left[F(x;\beta,\alpha+2)-2F(x;\beta,\alpha+1)+F(x;\beta,\alpha)\right],$$

$$\begin{aligned}F''_{\beta\alpha}(x;\beta,\alpha) \quad = \quad & \tfrac{\alpha}{\beta}\left[E_x(\ln x;\beta,\alpha+1)-E_x(\ln x;\beta,\alpha)\right] \\ & -\tfrac{\alpha}{\beta}F(x;\beta,\alpha+1)(\ln\beta+\Psi(\alpha)-\tfrac{1}{\beta}F(x;\beta,\alpha)(1-\alpha\ln\beta-\alpha\Psi(\alpha)),\end{aligned}$$

$$\begin{aligned}F''_{\alpha\alpha}(x;\beta,\alpha) \quad = \quad & E_x((\ln x)^2;\beta,\alpha)-2E_x(\ln x;\beta,\alpha)(\ln\beta+\Psi(\alpha)) \\ & +F(x;\beta,\alpha)\left[(\ln\beta)^2+2\ln\beta\Psi(\alpha)-\Psi'(\alpha)+\Psi(\alpha)\right],\end{aligned}$$

with $\Psi'(\alpha)$ being the derivative of $\Psi(\alpha)$, and

$$E_x((\ln x)^2;\beta,\alpha)=\int_0^x (\ln t)^2\frac{t^{\alpha-1}e^{-t/\beta}}{\Gamma(\alpha)\beta^\alpha}dt.$$

Then, the multivariate Newton-Raphson iteration is performed as

$$\begin{bmatrix} \hat{\beta}_P^{(l+1)} \\ \hat{\alpha}_P^{(l+1)} \end{bmatrix} = \begin{bmatrix} \hat{\beta}_P^{(l)} \\ \hat{\alpha}_P^{(l)} \end{bmatrix} - \begin{bmatrix} H_{\beta\beta}^{\prime\prime(l)} & H_{\beta\alpha}^{\prime\prime(l)} \\ H_{\beta\alpha}^{\prime\prime(l)} & H_{\alpha\alpha}^{\prime\prime(l)} \end{bmatrix}^{-1} \begin{bmatrix} H_{\beta}^{\prime(l)} \\ H_{\alpha}^{\prime(l)} \end{bmatrix}, \tag{2.9}$$

where l is the index for the iterations.

Alternately, the expression H in (2.3) is maximized using a grid search approach in the range of $\left(\max\left\{0, \hat{\beta}_M - 2\hat{\sigma}_{\hat{\beta}_M}\right\}, \hat{\beta}_M + 2\hat{\sigma}_{\hat{\beta}_M}\right)$ and $(\max\{0, \hat{\alpha}_M - 2\hat{\sigma}_{\hat{\alpha}_M}\}, \hat{\alpha}_M + 2\hat{\sigma}_{\hat{\alpha}_M})$ can be used. Here, usual maximization involving the Di-gamma function and it's derivative are avoided to keep consistency with the methods discussed below. In simulation, the estimates of the standard errors $\hat{\sigma}_{\hat{\beta}_M}$ and $\hat{\sigma}_{\hat{\alpha}_M}$ are obtained by implementing the method of moments estimates in separate prior to implementing the other estimates. In application to a data, a suitable range around the moment estimates can be used instead.

2.4. Likelihood-ratio Kolmogorov-Smirnov statistic (KSE)

The likelihood-ratio Kolmogorov-Smirnov statistic estimate for the parameters β and α, $\hat{\beta}_K$ and $\hat{\alpha}_K$ are found by minimizing the Likelihood-ratio Kolmogorov-Smirnov statistic $LK =$

$$\max_{i \in \{1,2,\cdots,n\}} \left\{(i-0.5)\log\frac{i-0.5}{nF(X_{(i)};\theta)} + (n-i+0.5)\log\frac{n-i+0.5}{n[1-F(X_{(i)};\theta)]}\right\}, \tag{2.10}$$

where $F(x;\beta,\alpha)$ is defined in (1.2). To minimize (2.10), a grid search in the range of $\left(\max\left\{0, \hat{\beta}_M - 2\hat{\sigma}_{\hat{\beta}_M}\right\}, \hat{\beta}_M + 2\hat{\sigma}_{\hat{\beta}_M}\right)$ and $(\max\{0, \hat{\alpha}_M - 2\hat{\sigma}_{\hat{\alpha}_M}\}, \hat{\alpha}_M + 2\hat{\sigma}_{\hat{\alpha}_M})$ can be used. Here, usual maximization involving the Di-gamma function and it's derivative are avoided to keep consistency with the methods discussed below. In simulation, the estimates of the standard errors $\hat{\sigma}_{\hat{\beta}_M}$ and $\hat{\sigma}_{\hat{\alpha}_M}$ are obtained by implementing the method of moments estimates in separate prior to implementing the other estimates. In application to a data, a suitable range around the moment estimates can be used instead.

2.5. Likelihood-ratio Cramer-von-Mises statistic (CVE)

The likelihood-ratio Cramer-von-Mises statistic estimate for the parameters β and α, $\hat{\beta}_C$ and $\hat{\alpha}_C$ are found by minimizing the Likelihood-ratio Cramer-von-Mises statistic

$$LC = \sum_{i=1}^{n}\left\{\log\frac{F(X_{(i)})^{-1}-1}{(n-0.5)/(i-0.75)-1}\right\}^2 \tag{2.11}$$

where $F(x;\beta,\alpha)$ is defined in (1.2). To minimize (2.11), a grid search in the range of $\left(\max\left\{0, \hat{\beta}_M - 2\hat{\sigma}_{\hat{\beta}_M}\right\}, \hat{\beta}_M + 2\hat{\sigma}_{\hat{\beta}_M}\right)$ and $(\max\{0, \hat{\alpha}_M - 2\hat{\sigma}_{\hat{\alpha}_M}\}, \hat{\alpha}_M + 2\hat{\sigma}_{\hat{\alpha}_M})$ can be used. Here, usual maximization involving the Di-gamma function and it's derivative are avoided to keep consistency with the methods discussed below. In simulation, the estimates of the standard errors $\hat{\sigma}_{\hat{\beta}_M}$ and $\hat{\sigma}_{\hat{\alpha}_M}$ are obtained by implementing the method of moments estimates in separate prior to implementing the other estimates. In application to a data, a suitable range around the moment estimates can be used instead.

2.6. Likelihood-ratio Anderson-Darling statistic (ADE)

The likelihood-ratio Anderson-Darling statistic estimate for the parameters β and α, $\hat{\beta}_A$ and $\hat{\alpha}_A$ are found by minimizing the Likelihood-ratio Anderson-Darling statistic

$$LA = -\sum_{i=1}^{n} \left\{ \frac{\log F(X_{(i)})}{n-i+0.5} + \frac{\log[1-F(X_{(i)})]}{i-0.5} \right\} \tag{2.12}$$

where $F(x;\beta,\alpha)$ is defined in (1.2). To minimize (2.12), a grid search in the range of $\left(\max\left\{0,\hat{\beta}_M - 2\hat{\sigma}_{\hat{\beta}_M}\right\}, \hat{\beta}_M + 2\hat{\sigma}_{\hat{\beta}_M}\right)$ and $(\max\{0,\hat{\alpha}_M - 2\hat{\sigma}_{\hat{\alpha}_M}\}, \hat{\alpha}_M + 2\hat{\sigma}_{\hat{\alpha}_M})$ can be used. Here, usual maximization involving the Di-gamma function and it's derivative are avoided to keep consistency with the methods discussed below. In simulation, the estimates of the standard errors $\hat{\sigma}_{\hat{\beta}_M}$ and $\hat{\sigma}_{\hat{\alpha}_M}$ are obtained by implementing the method of moments estimates in separate prior to implementing the other estimates. In application to a data, a suitable range around the moment estimates can be used instead.

3. Simulation results

One thousand samples are generated for two different parameter settings $\{(\beta = 2.0, \alpha = 4.0), (\beta = 4.0, \alpha = 2.0)$ and $(\beta = 0.5, \alpha = 0.5)\}$ and for two different sample sizes ($n = 20$ and $n = 50$). Means (MN), standard deviations (SD), mean of the absolute biases (AB), biases (BS), and mean squared errors (MS) are computed and displayed in Table 3.1-3.2.

MATLAB software is used in all computations and the codes are readily available from the author.
From Table 3.1, in estimating β, usually, the standard deviation is the lowest for $\hat{\beta}_C$, the mean absolute bias is the lowest for $\hat{\beta}_L$, the biases are usually the lowest for $\hat{\beta}_L$ with some exceptions, and the mean squared errors are lower for $\hat{\beta}_L$ except the case $\alpha = 0.5$ and $\beta = 0.5$ when $\hat{\beta}_C$ has the lowest mean squared errors.
In Table 3.2, in estimating α, $\hat{\alpha}_A$ has the largest standard deviation, bias, and mean squared error. $\hat{\alpha}_P$ has the lowest standard deviation, absolute bias, and mean squared error. And $\hat{\alpha}_P$ also has lower biases except in larger samples when $\hat{\alpha}_K$ has lower biases.

4. Application

The following data in Table 3.3 represents failure times of machine parts from manufacturer A and are taken from $http://v8doc.sas.com/sashtml/stat/chap29/sect44.htm$:
For this data, $\hat{\beta}_M = 483.22$, $\hat{\beta}_L = 522.42$, $\hat{\beta}_P = 522.42$, $\hat{\beta}_K = 520.82$, $\hat{\beta}_C = 512.02$, $\hat{\beta}_A = 522.42$, $\hat{\alpha}_M = 0.9700$, $\hat{\alpha}_L = 0.9208$, $\hat{\alpha}_P = 0.9208$, $\hat{\alpha}_K = 0.8911$, $\hat{\alpha}_C = 0.8911$, and $\hat{\alpha}_A = 0.8911$.

5. Summary and cocluding remarks

When both the parameters are of interest, the maximum likelihood estimate is suggested. When only the shape parameter is of concern, the maximization of the product spacings

Table 3.1 Simulation results for estimating β

	$\hat{\beta}_M$	$\hat{\beta}_L$	$\hat{\beta}_P$	$\hat{\beta}_K$	$\hat{\beta}_C$	$\hat{\beta}_A$
			$\beta = 2.0$ $\alpha = 4.0$ $n = 20$			
MN	1.8833	1.8945	2.4055	2.0185	1.8011	1.9655
SD	0.7079 (6)	0.5853 (2)	0.6333 (5)	0.6212 (4)	0.5789 (1)	0.6089 (3)
AB	0.5639 (5)	0.4935 (1)	0.6209 (6)	0.5102 (3)	0.5133 (4)	0.5045 (2)
BS	-0.1167 (4)	-0.1055 (3)	0.4055 (6)	0.0185 (1)	-0.1989 (5)	-0.0345 (2)
MS	0.5148 (5)	0.3538 (1)	0.5655 (6)	0.3862 (4)	0.3746 (3)	0.3719 (2)
			$\beta = 2.0$ $\alpha = 4.0$ $n = 50$			
MN	1.9521	1.9603	2.2287	2.0115	1.8892	1.9529
SD	0.4400 (6)	0.3796 (1)	0.4036 (4)	0.4043 (5)	0.3816 (2)	0.3920 (3)
AB	0.3503 (5)	0.3070 (1)	0.3812 (6)	0.3307 (4)	0.3235 (3)	0.3196 (2)
BS	-0.0479 (4)	-0.0397 (2)	0.2287 (6)	0.0115 (1)	-0.1108 (5)	-0.0471 (3)
MS	0.1959 (5)	0.1457 (1)	0.2152 (6)	0.1636 (4)	0.1579 (3)	0.1559 (2)
			$\beta = 4.0$ $\alpha = 2.0$ $n = 20$			
MN	3.7173	3.7705	4.7802	4.0013	3.5770	3.9024
SD	1.5864 (6)	1.2392 (2)	1.3200 (5)	1.3038 (4)	1.2229 (1)	1.2740 (3)
AB	1.2696 (5)	1.0567 (1)	1.2773(6)	1.0811 (3)	1.0958 (4)	1.0627 (2)
BS	-0.2827 (4)	-0.2295 (3)	0.7802 (6)	0.0013 (1)	-0.4230 (5)	-0.0976 (2)
MS	2.5965 (6)	1.5883 (1)	2.3510 (5)	1.7000 (4)	1.6743 (3)	1.6327 (2)
			$\beta = 4.0$ $\alpha = 2.0$ $n = 50$			
MN	3.9393	3.9615	4.5279	4.0739	3.8134	3.9422
SD	1.0561 (6)	0.8210 (2)	0.8798 (5)	0.8689 (4)	0.8190 (1)	0.8305 (3)
AB	0.8230 (5)	0.6581 (1)	0.8316(6)	0.7013 (4)	0.6819 (3)	0.6686 (2)
BS	-0.0607 (3)	-0.0385 (1)	0.5279 (6)	0.0739 (4)	-0.1866 (5)	-0.0578 (2)
MS	1.1191 (6)	0.6756 (1)	1.0528 (5)	0.7604 (4)	0.7055 (3)	0.6930 (2)
			$\beta = 0.5$ $\alpha = 0.5$ $n = 20$			
MN	0.4368	0.4846	0.6451	0.5261	0.4494	0.5080
SD	0.2722 (6)	0.1913 (2)	0.2181 (5)	0.2090 (4)	0.1835 (1)	0.2004 (3)
AB	0.2024 (5)	0.1597 (1)	0.2133(6)	0.1723 (4)	0.1610 (2)	0.1646 (3)
BS	-0.0632 (5)	-0.0154 (2)	0.1451 (6)	0.0261 (3)	-0.0506 (4)	0.0080 (1)
MS	0.0781 (6)	0.0368 (2)	0.0686 (5)	0.0444 (4)	0.0362 (1)	0.0402 (3)
			$\beta = 0.5$ $\alpha = 0.5$ $n = 50$			
MN	0.4739	0.4959	0.5806	0.5126	0.4712	0.4951
SD	0.1783 (6)	0.1302 (2)	0.1413 (4)	0.1415 (5)	0.1269 (1)	0.1317 (3)
AB	0.1402 (6)	0.1053 (1)	0.1312(5)	0.1145 (4)	0.1068 (3)	0.1066 (2)
BS	-0.0261 (4)	-0.0041 (1)	0.0806 (6)	0.0126 (3)	-0.0288 (5)	-0.0049 (2)
MS	0.0325 (6)	0.0170 (2)	0.0265 (5)	0.0202 (4)	0.0169 (1)	0.0174 (3)

approach is suggested. In goodness-of-fit applications, the minimization of the Anderson-Darling type statistic to be avoided. This paper gives a simpler grid search approach which avoids using the Di-gamma function and it's derivative in both the maximum likelihood method and the maximization of the product spacings method.

Table 3.2 Simulation results for estimating α

	$\hat{\alpha}_M$	$\hat{\alpha}_L$	$\hat{\alpha}_P$	$\hat{\alpha}_K$	$\hat{\alpha}_C$	$\hat{\alpha}_A$
		$\beta = 2.0$	$\alpha = 4.0$	$n = 20$		
MN	4.7844	4.6058	3.6584	4.3635	4.8383	7.1151
SD	1.6919 (5)	1.3716 (3)	1.0845 (1)	1.3602 (2)	1.4266 (4)	4.2117 (6)
AB	1.3154 (6)	1.1101 (4)	0.9235 (1)	1.0530 (2)	1.2432 (5)	1.0592 (3)
BS	0.7844 (4)	0.6058 (3)	-0.3416 (1)	0.3635 (2)	0.8383 (5)	3.1151 (6)
MS	3.4779 (5)	2.2484 (3)	1.2928 (1)	1.9824 (2)	2.7378 (4)	27.4422 (6)
		$\beta = 2.0$	$\alpha = 4.0$	$n = 50$		
MN	4.2928	4.2261	3.7512	4.1439	4.3829	4.4538
SD	0.9309 (5)	0.7960 (2)	0.7053 (1)	0.8156 (3)	0.8400 (4)	1.2449 (6)
AB	0.7387 (6)	0.6381 (2)	0.6242 (1)	0.6516 (3)	0.7126 (5)	0.6689 (4)
BS	0.2928 (4)	0.2261 (2)	-0.2488 (3)	0.1439 (1)	0.3829 (5)	0.4538 (6)
MS	0.9522 (5)	0.6847 (2)	0.5594 (1)	0.6860 (3)	0.8522 (4)	1.7558 (6)
		$\beta = 4.0$	$\alpha = 2.0$	$n = 20$		
MN	2.4782	2.3216	1.8708	2.2097	2.4416	5.6093
SD	0.9644 (5)	0.7361 (3)	0.5736 (1)	0.7167 (2)	0.7773 (4)	2.5154 (6)
AB	0.7615 (6)	0.5898 (4)	0.4726(1)	0.5529 (2)	0.6648 (5)	0.5647 (3)
BS	0.4782 (5)	0.3216 (3)	-0.1292 (1)	0.2097 (2)	0.4416 (4)	1.6093 (6)
MS	1.1588 (5)	0.6453 (3)	0.3457 (1)	0.5576 (2)	0.7991 (4)	19.3544 (6)
		$\beta = 4.0$	$\alpha = 2.0$	$n = 50$		
MN	2.1663	2.1039	1.8749	2.0588	2.1780	4.2436
SD	0.5177 (5)	0.3948 (2)	0.3486 (1)	0.3993 (3)	0.4140 (4)	1.0683 (6)
AB	0.4150 (6)	0.3134 (2)	0.3058(1)	0.3161 (3)	0.3453 (5)	0.3207 (4)
BS	0.1663 (4)	0.1039 (2)	-0.1251 (3)	0.0588 (1)	-0.1780 (5)	2.2436 (6)
MS	0.2956 (5)	0.1666 (3)	0.1371 (1)	0.1629 (2)	0.2030 (4)	6.1748 (6)
		$\beta = 0.5$	$\alpha = 0.5$	$n = 20$		
MN	0.6902	0.5605	0.4697	0.5375	0.5872	5.0421
SD	0.2780 (5)	0.1683 (3)	0.1324 (1)	0.1613 (2)	0.1804 (4)	2.5984 (6)
AB	0.2485 (6)	0.1275 (4)	0.1070(1)	0.1216 (2)	0.1411 (5)	0.1221 (3)
BS	0.1902 (5)	0.0605 (3)	-0.0303 (1)	0.0375 (2)	0.0872 (4)	4.5421 (6)
MS	0.1135 (5)	0.0320 (3)	0.0185 (1)	0.0274 (2)	0.0401 (4)	27.3822 (6)
		$\beta = 0.5$	$\alpha = 0.5$	$n = 50$		
MN	0.5749	0.5207	0.4748	0.5128	0.5371	3.8665
SD	0.1583 (5)	0.0896 (2)	0.0784 (1)	0.0916 (3)	0.0949 (4)	0.4950 (6)
AB	0.1359 (6)	0.0705 (2)	0.0676 (1)	0.0710 (3)	0.0769 (5)	0.0714 (4)
BS	0.0749 (5)	0.0207 (2)	-0.0252 (3)	0.0128 (1)	0.0371 (4)	3.3665 (6)
MS	0.0307 (5)	0.0085 (2)	0.0068 (1)	0.0086 (3)	0.0104 (4)	11.5784 (6)

Table 3.3 Failure times

620	470	260	89	388	242	103	100	39	460	284
1285	218	393	106	158	152	477	403	103	69	158
818	947	399	1274	32	12	134	660	548	381	203
871	193	531	317	85	1410	250	41	1101	32	421
32	343	376	1512	1792	47	95	76	515	72	1585
253	6	860	89	1055	537	101	385	176	11	565
164	16	1267	352	160	195	1279	356	751	500	803
560	151	24	689	1119	1733	2194	763	555	14	45
776	1									

References

Cheng, R. C. H. and Amin, N. A. K. (1983). Estimating parameters in continuous univariate distributions with a shifted origin. *Journal of Royal Statistical Society, Series B*, **45**, 394-403.

Cheng, R. C. H. and Iles, T. C. (1987). Corrected maximum likelihood in non-regular problems. *Journal of Royal Statistical Society, Series B*, **49**, 95-101.

Choi, S. C. and Wette, R. (1969). Maximum likelihood estimation of the parameters of the gamma distribution and their bias. *Technometrics*, **11**, 683-690.

Dang, H. and Weerakkody, G. (2000). Bounds for the maximum likelihood estimates in two-parameter gamma distribution. *Journal of Mathematical Analysis and Applications*, **245**, 1-6.

Evans, M., Hastings, N. and Peacock, B. (2000). *Statistical distributions, Third Edition*, John Wiley & Sons, Inc., New York.

Harter, H. L. and Moore, A. H. (1965). Maximum likelihood estimation of the parameters of gamma and Weibull populations from complete and from censored samples. *Technometrics*, **7**, 639-643.

Lee, E. T. (1992). *Statistical methods for survival data analysis, second edition*, John Wiley & Sons, Inc., New York.

Rahman, M. and Pearson, L. M. (2002). Estimation in two-parameter exponential distributions. *Journal of Statistical Computation and Simulation*, **70**, 371-386.

Rahman, M., Pearson, L. M. and Martinovic, U. R. (2007). Method of product spacings in the two-parameter gamma distribution. *Journal of Statistical Research*, **41**, 51-58.

Ranneby, B. (1984). The A spacing method. An estimation method related to the maximum likelihood method. *Scandanavian Journal of Statistics*, **11**, 93-112.

Shah, A. and Gokhale, D. V. (1993). On Maximum product of spacings (MPS) estimation for Burr XII distributions. *Communications in Statistics — Simulation and Computation*, **22**, 615-641.

Wilks, D. S. (1990). Maximum likelihood estimation for the gamma distribution using data containing zeros. *Journal of Climate*, **3**, 1495-1501.

Zhang, J. and Wu, Y. (2005). Likelihood-ratio tests for normality. *Compulational Statistics & Data Analysis*, **49**, 709-721.