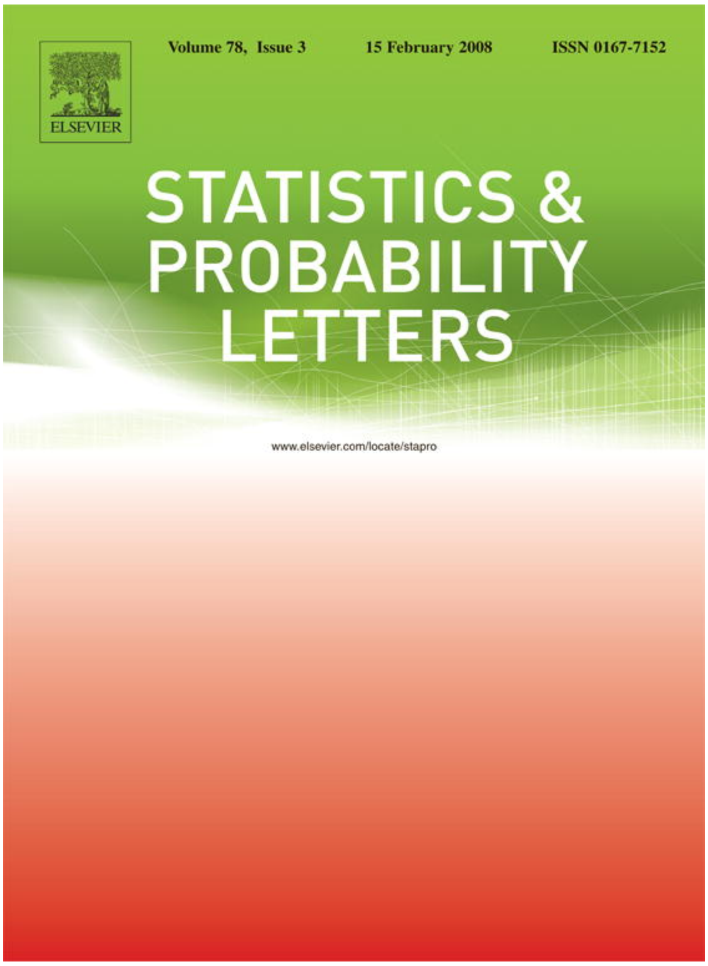
Volume 78, Issue 3
15 February 2008
ISSN 0167-7152
ELSEVIER
STATISTICS & PROBABILITY LETTERS
www.elsevier.com/locate/stapro

Available online at www.sciencedirect.com
ScienceDirect
Statistics & Probability Letters 78 (2008) 286–295
STATISTICS & PROBABILITY LETTERS
www.elsevier.com/locate/stapro

On estimation of the shape parameter of the gamma distribution

A. Zaigraev*, A. Podraza-Karakulska

Faculty of Mathematics and Computer Science, N. Copernicus University, Chopin str. 12/18, 87-100 Toruń, Poland

Received 8 December 2006; received in revised form 17 May 2007; accepted 24 July 2007
Available online 31 July 2007

Abstract

The problem of estimation of an unknown shape parameter under the sample drawn from the gamma distribution, where the scale parameter is also unknown, is considered. A new estimator, called the maximum likelihood scale invariant estimator, is proposed. It is established that both the bias and the variance of this estimator are less than that of the usual maximum likelihood estimator. A property of the psi function is also obtained.

MSC: primary 62F10; secondary 33B15

Keywords: Gamma distribution; Maximum likelihood scale invariant estimator; Maximum likelihood estimator; Psi function

1. Introduction

Let a sample $x = (x_1, x_2, \ldots, x_n)$ be drawn from the gamma distribution $\Gamma(\alpha, \sigma)$ with an unknown shape parameter $\alpha > 0$ and an unknown scale parameter $\sigma > 0$, whose density function has the form

$$p(u; \alpha, \sigma) = \frac{u^{\alpha-1} e^{-u/\sigma}}{\sigma^\alpha \Gamma(\alpha)}, \quad u > 0.$$

Consider the problem of estimation of α. One of the most popular estimators is the well-known maximum likelihood estimator (ML-estimator) (e.g. Barndorff-Nielsen, 1978, Sections 9.3, 9.4; Bowman and Shenton, 1988; Crain, 1976; Dang and Weerakkody, 2000). Let

$$\mathbf{p}(x; \alpha, \sigma) = \sigma^{-n\alpha} (\Gamma(\alpha))^{-n} \left(\prod_{j=1}^{n} x_j \right)^{\alpha-1} \exp\left(-\frac{1}{\sigma} \sum_{k=1}^{n} x_k \right)$$

*Corresponding author.
E-mail address: alzaig@mat.uni.torun.pl (A. Zaigraev).

0167-7152/$ - see front matter
doi:10.1016/j.spl.2007.07.003

be the corresponding likelihood function. The ML-estimators of α and σ are determined by the equations:

$$\begin{cases} \ln\sigma + \Psi(\alpha) = \sum_{j=1}^{n} \ln x_j/n, \\ \alpha - \sum_{k=1}^{n} x_k/(n\sigma) = 0, \end{cases}$$

where $\Psi(\alpha):=(\mathrm{d}/\mathrm{d}\alpha)\ln\Gamma(\alpha):=(\ln\Gamma(\alpha))'$ is the so-called Euler psi (digamma) function.

From those equations one can obtain the ML-estimators α^* and σ^* of α and σ, respectively. Namely, α^* is the root of the equation

$$g(\alpha):=\ln\alpha - \Psi(\alpha) = \ln\bar{x} - \frac{1}{n}\sum_{j=1}^{n}\ln x_j$$

while

$$\sigma^* = \frac{\bar{x}}{\alpha^*}.$$

Here, $\bar{x}$ is the sample mean, i.e.

$$\bar{x} = \frac{1}{n}\sum_{k=1}^{n} x_k.$$

Observe that the function g is strictly decreasing and takes values in $(0,\infty)$ (e.g. Alzer, 1997, Theorem 1). Therefore, the estimator α^* is well defined and unique.

Observe also that the estimator α^* is scale invariant. Furthermore, one can easily see that

$$\mathrm{E}_{(\alpha,\sigma)}\left(\ln\bar{x} - \frac{1}{n}\sum_{j=1}^{n}\ln x_j\right) = \Psi(n\alpha) - \Psi(\alpha) - \ln n = g(\alpha) - g(n\alpha) := g_n(\alpha).$$

The question arises: *why would not one take the root of the equation*:

$$g_n(\alpha) = \ln\bar{x} - \frac{1}{n}\sum_{j=1}^{n}\ln x_j \tag{1}$$

as an estimator of an unknown shape parameter α? Such an estimator would coincide with that based on the method of moments.

It turns out that such a choice has quite a deep reasoning. Since in our scheme σ becomes a nuisance parameter, it is natural to apply the maximum likelihood principle to the measure defined on the σ-algebra of the scale invariant sets generated by the underlying gamma distribution. As it is known (e.g. Hájek et al., 1999, Subsection 3.2.2; Nagaev, 1996, Section 8.3), the density corresponding to this measure, with respect to that generated by the standard normal distribution, is given as follows:

$$\mathbf{q}(x;\alpha) = \frac{\int_0^\infty t^{n-1}\mathbf{p}(tx;\alpha,\sigma)\,\mathrm{d}t}{\int_0^\infty t^{n-1}\mathbf{s}(tx)\,\mathrm{d}t} = \frac{2\pi^{n/2}\Gamma(n\alpha)(\sum_{i=1}^{n}x_i^2)^{n/2}(\prod_{i=1}^{n}x_i)^{\alpha-1}}{\Gamma(n/2)(\Gamma(\alpha))^n(\sum_{i=1}^{n}x_i)^{n\alpha}},$$

where

$$\mathbf{s}(x) = (2\pi)^{-n/2}\exp\left(-\frac{1}{2}\sum_{k=1}^{n}x_k^2\right).$$

Then by direct calculations one can obtain that the maximum likelihood scale invariant estimator (IML-estimator) $\alpha^{**} \in \arg\max_{\alpha>0}\mathbf{q}(x;\alpha)$ is the root of the Eq. (1).

The estimator α^{**} is also scale invariant, well defined and unique since the function g_n is strictly decreasing and takes values in $(0,\infty)$ (see Lemma 1 in Section 3) (see also Fig. 1).

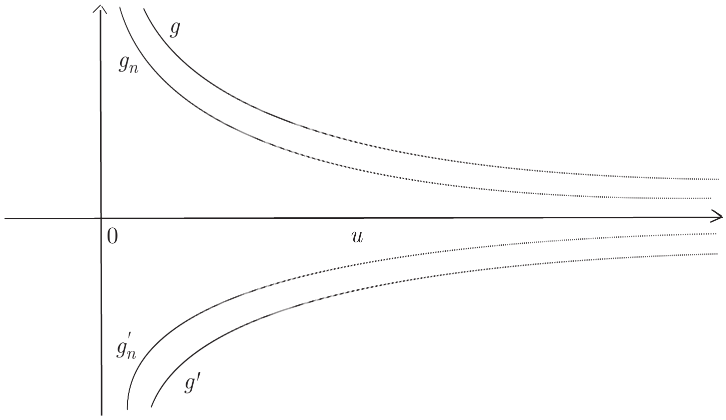

Fig. 1. Graphs of the functions g, g_n, g', g_n' for $n = 2$.

It is worth noting that the scale invariance of the maximum likelihood estimator of a shape parameter is a quite common property in the case when the scale is also unknown. Indeed, consider the likelihood function

$$\mathbf{p}(x;\alpha,\sigma) = \sigma^{-n}\prod_{j=1}^{n} p(\sigma^{-1}x_j;\alpha,1),$$

where α is a shape parameter taking values in (α_-,α_+). Assume that

$$\max_{\alpha\in(\alpha_-,\alpha_+),\ \sigma>0}\mathbf{p}(x;\alpha,\sigma) = \max_{\alpha\in(\alpha_-,\alpha_+)}\max_{\sigma>0}\mathbf{p}(x;\alpha,\sigma).$$

Let

$$\hat{\sigma}(x;\alpha)\in\arg\max_{\sigma>0}\mathbf{p}(x;\alpha,\sigma).$$

Observe that for any $\lambda>0$

$$\mathbf{p}(\lambda x;\alpha,\sigma) = \lambda^{-n}\mathbf{p}(x;\alpha,\sigma/\lambda),$$

whence

$$\hat{\sigma}(\lambda x;\alpha) = \lambda\hat{\sigma}(x;\alpha).$$

Thus,

$$\begin{aligned}(\alpha^*,\sigma^*) \in \arg\max_{\alpha\in(\alpha_-,\alpha_+),\ \sigma>0}\mathbf{p}(x;\alpha,\sigma) &= \arg\max_{\alpha\in(\alpha_-,\alpha_+)}\mathbf{p}(x;\alpha,\hat{\sigma}(x;\alpha))\\ &= \arg\max_{\alpha\in(\alpha_-,\alpha_+)}\left((\hat{\sigma}(x;\alpha))^{-n}\prod_{j=1}^{n}p((\hat{\sigma}(x;\alpha))^{-1}x_j;\alpha,1)\right).\end{aligned}$$

It is evident that $\alpha^*(\lambda x) = \alpha^*(x)$, i.e. the estimator α^* is scale invariant. Therefore, it is reasonable to apply the method presented here also for other distributions.

The goal of this paper is to compare two estimators of α: the ML-estimator and the IML-estimator. We show that both the bias and the variance of the IML-estimator α^{**} are less than that of the ML-estimator α^*.

The paper is organized as follows. In Section 2 we establish the main result while all the auxiliary results are proved in Section 3.

2. Main result

Theorem. *If a sample $x = (x_1, x_2, \ldots, x_n)$ is drawn from $\Gamma(\alpha,\sigma)$ distribution, then*

$$\mathrm{E}\alpha^* > \mathrm{E}\alpha^{**} > \alpha, \quad \mathrm{Var}\,\alpha^* > \mathrm{Var}\,\alpha^{**}. \tag{2}$$

Proof. Since the functions g and g_n are strictly decreasing, the inverse functions g^{-1} and g_n^{-1} are well defined. According to the definitions,

$$\alpha^* = g^{-1}(T(x)), \quad \alpha^{**} = g_n^{-1}(T(x)),$$

where

$$T(x) := \ln \bar{x} - \frac{1}{n}\sum_{j=1}^{n} \ln x_j.$$

Denote

$$f(u) := g^{-1}(u) - g_n^{-1}(u), \quad u > 0.$$

The function f is positive. Indeed, since

$$g_n(g^{-1}(u)) = g(g^{-1}(u)) - g(ng^{-1}(u)) = u - g(ng^{-1}(u)) < u$$

and the function g_n is strictly decreasing, we obtain

$$g^{-1}(u) > g_n^{-1}(u), \quad u > 0.$$

Therefore, $\mathrm{E}\alpha^* > \mathrm{E}\alpha^{**}$.

Now let us show that for any $\alpha > 0$

$$\mathrm{E}\alpha^{**} = \mathrm{E}g_n^{-1}(T(x)) > \alpha.$$

Observe that the function g_n^{-1} is strictly convex since from Lemma 1 in Section 3 it follows that

$$\frac{\partial^2 g_n^{-1}(u)}{\partial u^2} = -\frac{g_n''(g_n^{-1}(u))}{(g_n'(g_n^{-1}(u)))^3} > 0, \quad u > 0.$$

Since $\mathrm{E}T(x) = g_n(\alpha)$, by the Jensen inequality we get

$$\mathrm{E}g_n^{-1}(T(x)) > g_n^{-1}(\mathrm{E}T(x)) = g_n^{-1}(g_n(\alpha)) = \alpha.$$

It remains to prove the inequality for variances. From the evident equalities

$$\alpha^* - \mathrm{E}\alpha^{**} = \alpha^* - \alpha^{**} + \alpha^{**} - \mathrm{E}\alpha^{**}, \quad \alpha^* - \mathrm{E}\alpha^{**} = \alpha^* - \mathrm{E}\alpha^* + \mathrm{E}\alpha^* - \mathrm{E}\alpha^{**}$$

one can obtain

$$\mathrm{Var}\,\alpha^* + (\mathrm{E}\alpha^* - \mathrm{E}\alpha^{**})^2 = \mathrm{Var}\,\alpha^{**} + \mathrm{E}(\alpha^* - \alpha^{**})^2 + 2\mathrm{E}(\alpha^* - \alpha^{**})(\alpha^{**} - \mathrm{E}\alpha^{**}).$$

Therefore,

$$\mathrm{Var}\,\alpha^* - \mathrm{Var}\,\alpha^{**} = \mathrm{Var}\,(\alpha^* - \alpha^{**}) + 2\mathrm{E}(\alpha^* - \alpha^{**})(\alpha^{**} - \mathrm{E}\alpha^{**}).$$

In order to show the second part of (2), it is enough to establish that

$$\mathrm{E}(\alpha^* - \alpha^{**})(\alpha^{**} - \mathrm{E}\alpha^{**}) > 0.$$

Denote $c := \mathrm{E}\alpha^{**}$ and observe that

$$\begin{aligned}\mathrm{E}(\alpha^* - \alpha^{**})(\alpha^{**} - \mathrm{E}\alpha^{**}) &= \mathrm{E}(g^{-1}(T(x)) - g_n^{-1}(T(x)))(g_n^{-1}(T(x)) - c) \\ &= \int_0^\infty (g^{-1}(u) - g_n^{-1}(u))(g_n^{-1}(u) - c)p_{T(x)}(u)\,\mathrm{d}u,\end{aligned}$$

where $p_{T(x)}$ is the density of $T(x)$.

Let

$$\Delta(u) := g_n^{-1}(u) - c, \quad u > 0.$$

Observe that

$$\Delta(u) > 0, \quad 0 < u < g_n(c),$$

$$\Delta(u) = 0, \quad u = g_n(c),$$

$$\Delta(u) < 0, \quad u > g_n(c).$$

Since, as it is proved in Lemma 3 of Section 3, the function f is strictly decreasing, we obtain

$$\int_0^{g_n(c)} f(u)\Delta(u)p_{T(x)}(u)\,\mathrm{d}u > f(g_n(c))\int_0^{g_n(c)} \Delta(u)p_{T(x)}(u)\,\mathrm{d}u$$

and

$$\int_{g_n(c)}^{\infty} f(u)\Delta(u)p_{T(x)}(u)\,\mathrm{d}u > f(g_n(c))\int_{g_n(c)}^{\infty} \Delta(u)p_{T(x)}(u)\,\mathrm{d}u.$$

Therefore,

$$\begin{aligned}\mathrm{E}(\alpha^* - \alpha^{**})(\alpha^{**} - \mathrm{E}\alpha^{**}) &= \int_0^{\infty} f(u)\Delta(u)p_{T(x)}(u)\,\mathrm{d}u > f(g_n(c))\int_0^{\infty} \Delta(u)p_{T(x)}(u)\,\mathrm{d}u \\ &= f(g_n(c))\mathrm{E}(\alpha^{**} - \mathrm{E}\alpha^{**}) = 0. \qquad \square\end{aligned}$$

Remark 1. Since

$$\mathrm{E}(\alpha^* - \alpha)^2 = \mathrm{Var}\,\alpha^* + (\mathrm{E}\alpha^* - \alpha)^2, \quad \mathrm{E}(\alpha^{**} - \alpha)^2 = \mathrm{Var}\,\alpha^{**} + (\mathrm{E}\alpha^{**} - \alpha)^2,$$

from the theorem it follows that

$$\mathrm{E}(\alpha^* - \alpha)^2 > \mathrm{E}(\alpha^{**} - \alpha)^2.$$

Let us apply the Monte-Carlo simulation to confirm the results of the theorem. Given α and σ, we generate 10,000 samples drawn from $\Gamma(\alpha, \sigma)$ distribution for $n = 10, 20, 30, 50, 100$. Next, we solve numerically two

Table 1
Numerical calculations of $v^* = (\mathrm{E}\alpha^* - \alpha)^2$, $v^{**} = (\mathrm{E}\alpha^{**} - \alpha)^2$, $\mathrm{Var}\,\alpha^*$, $\mathrm{Var}\,\alpha^{**}$

n	v^*	v^{**}	$\mathrm{Var}\,\alpha^*$	$\mathrm{Var}\,\alpha^{**}$	v^*	v^{**}	$\mathrm{Var}\,\alpha^*$	$\mathrm{Var}\,\alpha^{**}$
$\alpha = 0.125$					$\alpha = 0.5$			
10	0.00159	0.00126	0.00119	0.00118	0.00410	0.00140	0.02770	0.02760
20	0.00120	0.00097	0.00100	0.00100	0.00258	0.00114	0.02008	0.01941
30	0.00100	0.00084	0.00080	0.00079	0.00120	0.00052	0.01392	0.01347
50	0.00082	0.00072	0.00048	0.00047	0.00056	0.00027	0.00849	0.00821
100	0.00068	0.00064	0.000228	0.000225	0.00011	0.00005	0.00367	0.00361
$\alpha = 1$					$\alpha = 2$			
10	0.0145	0.0042	0.1171	0.1165	0.0536	0.0136	0.4885	0.4876
20	0.0112	0.0048	0.0865	0.0841	0.0463	0.0197	0.3744	0.3705
30	0.0078	0.0038	0.0647	0.0625	0.0320	0.0148	0.2785	0.2706
50	0.0026	0.0012	0.0371	0.0358	0.0128	0.0059	0.1672	0.1615
100	0.0006	0.0003	0.0174	0.0171	0.0029	0.0013	0.0771	0.0756
$\alpha = 4$					$\alpha = 16$			
10	0.216	0.051	2.027	2.026	42.83	12.07	198.44	127.58
20	0.186	0.077	1.541	1.535	3.09	1.21	25.69	25.44
30	0.142	0.062	1.190	1.137	2.04	0.86	19.02	18.62
50	0.062	0.029	0.71	0.69	1.03	0.47	12.50	12.05
100	0.014	0.007	0.335	0.328	0.24	0.11	5.556	5.446
$\alpha = 32$					$\alpha = 128$			
10	174.38	2.43	774.92	560.11	191.55	32.38	2102.56	2099.77
20	11.82	4.38	103.39	102.07	173.42	62.90	1619.45	1615.24
30	8.62	3.59	78.95	76.43	143.47	60.76	1288.96	1247.04
50	3.81	1.69	47.84	46.48	55.77	23.48	781.47	759.32
100	0.89	0.38	21.96	21.53	14.14	5.982	357.58	350.46

equations: $g(\alpha) = T(x)$ and $g_n(\alpha) = T(x)$ with respect to α, and obtain 10,000 values of α^* and α^{**}. Taking their means, we estimate $\mathrm{E}\alpha^*$ and $\mathrm{E}\alpha^{**}$ and then calculate $(\mathrm{E}\alpha^* - \alpha)^2$, $(\mathrm{E}\alpha^{**} - \alpha)^2$, $\mathrm{Var}\,\alpha^*$, $\mathrm{Var}\,\alpha^{**}$. The corresponding programs are written in C++. The results are presented in Table 1.

3. Proofs of auxiliary results

Lemma 1. *For any $n>1$ the function $g_n(u) = \Psi(nu) - \Psi(u) - \ln n$, $u>0$, is strictly decreasing and strictly convex.*

Proof. The integral representation

$$\Psi'(u) = \int_0^\infty \frac{t\mathrm{e}^{-ut}}{1-\mathrm{e}^{-t}}\,\mathrm{d}t, \quad u>0 \tag{3}$$

(e.g. Abramovitz and Stegun, 1965, formula 6.4.1) yields

$$g_n'(u) = n\Psi'(nu) - \Psi'(u) = \int_0^\infty \frac{t}{1-\mathrm{e}^{-t}}(n\mathrm{e}^{-nut} - \mathrm{e}^{-ut})\,\mathrm{d}t, \quad u>0. \tag{4}$$

Denote

$$h(t) := \frac{t}{1-\mathrm{e}^{-t}}, \quad \Delta_{n,u}(t) := n\mathrm{e}^{-nut} - \mathrm{e}^{-ut}, \quad t>0.$$

The function h is strictly increasing since

$$h'(t) = \frac{1-(1+t)\mathrm{e}^{-t}}{(1-\mathrm{e}^{-t})^2} > 0, \quad t>0$$

due to the inequality $\mathrm{e}^t > 1+t$, $t>0$. The function $\Delta_{n,u}$ is such that

$$\Delta_{n,u}(t) > 0, \quad 0<t<t_n(u) = \frac{\ln n}{(n-1)u},$$

$$\Delta_{n,u}(t) = 0, \quad t = t_n(u),$$

$$\Delta_{n,u}(t) < 0, \quad t > t_n(u).$$

This implies that

$$\int_0^{t_n(u)} h(t)\Delta_{n,u}(t)\,\mathrm{d}t < h(t_n(u)) \int_0^{t_n(u)} \Delta_{n,u}(t)\,\mathrm{d}t$$

and

$$\int_{t_n(u)}^\infty h(t)\Delta_{n,u}(t)\,\mathrm{d}t < h(t_n(u)) \int_{t_n(u)}^\infty \Delta_{n,u}(t)\,\mathrm{d}t.$$

Therefore,

$$g_n'(u) < h(t_n(u)) \int_0^\infty \Delta_{n,u}(t)\,\mathrm{d}t = 0$$

since

$$\int_0^\infty \Delta_{n,u}(t)\,\mathrm{d}t = 0.$$

Further, by differentiation of both sides of (4) we obtain

$$g_n''(u) = \int_0^\infty \frac{t}{1-\mathrm{e}^{-t}}(t\mathrm{e}^{-ut} - n^2 t\mathrm{e}^{-nut})\,\mathrm{d}t, \quad u>0.$$

Denote

$$\overline{\Delta}_{n,u}(t) := t\mathrm{e}^{-ut} - n^2 t\mathrm{e}^{-nut}, \quad t > 0$$

and observe that

$$\overline{\Delta}_{n,u}(t) > 0, \quad t > \overline{t}_n(u) = \frac{2\ln n}{(n-1)u},$$

$$\overline{\Delta}_{n,u}(t) = 0, \quad t = \overline{t}_n(u),$$

$$\overline{\Delta}_{n,u}(t) < 0, \quad 0 < t < \overline{t}_n(u).$$

This implies that

$$\int_0^{\overline{t}_n(u)} h(t)\overline{\Delta}_{n,u}(t)\,\mathrm{d}t > h(\overline{t}_n(u)) \int_0^{\overline{t}_n(u)} \overline{\Delta}_{n,u}(t)\,\mathrm{d}t$$

and

$$\int_{\overline{t}_n(u)}^{\infty} h(t)\overline{\Delta}_{n,u}(t)\,\mathrm{d}t > h(\overline{t}_n(u)) \int_{\overline{t}_n(u)}^{\infty} \overline{\Delta}_{n,u}(t)\,\mathrm{d}t.$$

Therefore,

$$g_n''(u) > h(\overline{t}_n(u)) \int_0^{\infty} \overline{\Delta}_{n,u}(t)\,\mathrm{d}t = 0$$

since

$$\int_0^{\infty} \overline{\Delta}_{n,u}(t)\,\mathrm{d}t = 0. \qquad \square$$

In the next lemma a property of the psi function is established. This result will be helpful in proving the last lemma.

Lemma 2. *Let the function* θ *be defined by the equality*

$$g'(\theta(u)) = g_n'(u), \quad u > 0.$$

Then

$$g(\theta(u)) > g_n(u), \quad u > 0.$$

Proof. First of all, observe that the function θ is well defined since both the functions g' and g_n' take values in $(-\infty, 0)$, are negative and strictly increasing (see Alzer, 1997, Theorem 1 and Lemma 1 above).

Now we prove that

$$g'\left(\sqrt{n/(n-1)}u\right) > g_n'(u), \quad u > 0. \tag{5}$$

Consider the function

$$\rho(u) := g_n'(u) - g'\left(\sqrt{n/(n-1)}u\right) = g'(u) - ng'(nu) - g'\left(\sqrt{n/(n-1)}u\right), \quad u > 0.$$

Making use of (3) one can write the representation (Binet's formula):

$$g(u) = \int_0^{\infty} \varphi(t)\mathrm{e}^{-ut}\,\mathrm{d}t, \quad u > 0,$$

where

$$\varphi(t) := \frac{1}{1-\mathrm{e}^{-t}} - \frac{1}{t}, \quad t > 0.$$

Therefore,

$$\rho(u) = \int_0^\infty \varphi(t)(nt\mathrm{e}^{-nut} + t\mathrm{e}^{-\sqrt{n/(n-1)}ut} - t\mathrm{e}^{-ut})\,\mathrm{d}t, \quad u > 0.$$

The function φ is strictly increasing and

$$\lim_{t\to 0} \varphi(t) = \frac{1}{2}, \quad \lim_{t\to\infty} \varphi(t) = 1$$

(see Alzer, 1997, Proof of Theorem 1). Denote

$$\begin{aligned} \Delta_n(z) &:= nz\mathrm{e}^{-nz} + z\mathrm{e}^{-\sqrt{n/(n-1)}z} - z\mathrm{e}^{-z} \\ &= z\mathrm{e}^{-z}(n\mathrm{e}^{-(n-1)z} + \mathrm{e}^{-(\sqrt{n/(n-1)}-1)z} - 1), \quad z > 0. \end{aligned}$$

Then

$$\rho(u) = \frac{1}{u^2}\int_0^\infty \varphi\left(\frac{z}{u}\right)\Delta_n(z)\,\mathrm{d}z, \quad u > 0.$$

Observe that there exists $z_0 > 0$ such that

$$\Delta_n(z) > 0, \quad 0 < z < z_0,$$

$$\Delta_n(z) = 0, \quad z = z_0,$$

$$\Delta_n(z) < 0, \quad z > z_0.$$

The standard reasoning leads to the inequalities

$$\int_0^{z_0} \varphi\left(\frac{z}{u}\right)\Delta_n(z)\,\mathrm{d}z < \varphi\left(\frac{z_0}{u}\right)\int_0^{z_0} \Delta_n(z)\,\mathrm{d}z,$$

$$\int_{z_0}^\infty \varphi\left(\frac{z}{u}\right)\Delta_n(z)\,\mathrm{d}z < \varphi\left(\frac{z_0}{u}\right)\int_{z_0}^\infty \Delta_n(z)\,\mathrm{d}z.$$

Therefore,

$$u^2\rho(u) = \int_0^\infty \varphi\left(\frac{z}{u}\right)\Delta_n(z)\,\mathrm{d}z < \varphi\left(\frac{z_0}{u}\right)\int_0^\infty \Delta_n(z)\,\mathrm{d}z = 0$$

since

$$\int_0^\infty \Delta_n(z)\,\mathrm{d}z = 0,$$

and (5) holds.

Since the function g' is strictly increasing, from (5) it follows that

$$\theta(u) < \sqrt{n/(n-1)}u, \quad u > 0.$$

On the other hand,

$$\theta(u) > u, \quad u > 0$$

since

$$g_n'(u) - g'(u) = -ng'(nu) > 0, \quad u > 0.$$

Denote

$$\lambda(u) := g(\theta(u)) - g_n(u), \quad u > 0.$$

Since $\theta'(u) > 1$ and $g'(u) < 0$, we obtain

$$\lambda'(u) = g'(\theta(u))\theta'(u) - g_n'(u) = g'(\theta(u))(\theta'(u) - 1) < 0, \quad u > 0.$$

Table 2
Numerical calculations of bounds for $\theta(u)$

u	$n=2$		$n=10$		$n=50$	
	$\theta(u)$	$\sqrt{n/(n-1)}u$	$\theta(u)$	$\sqrt{n/(n-1)}u$	$\theta(u)$	$\sqrt{n/(n-1)}u$
0.1	0.134913918	0.141421356	0.103577225	0.105409255	0.100565420	0.101015254
0.5	0.651452797	0.707106781	0.517553390	0.527046277	0.503191582	0.505076272
1	1.316748502	1.414213562	1.039115016	1.054092553	1.007252736	1.010152544
3	4.090593081	4.242640686	3.140261683	3.162277660	3.026250755	3.030457633
5	6.905201515	7.071067810	5.246643827	5.270462768	5.046219652	5.050762721
10	13.96570362	14.14213562	10.51571989	10.54092553	10.09672350	10.10152544
50	70.52586082	70.71067810	52.67831719	52.70462768	50.50261897	50.50762721
100	141.2355025	141.4213562	105.3828081	105.4092553	101.0102207	101.0152544

Hence, the function λ is strictly decreasing. Now we establish that $\lambda(u) > 0$ for all sufficiently large u. Due to the representations

$$g(u) = \frac{1}{2u}(1 + o(1)), \quad g_n(u) = \frac{1}{2u}\left(1 - \frac{1}{n}\right)(1 + o(1)), \quad u \to \infty$$

(e.g. Abramovitz and Stegun, 1965, formula 6.3.18), we get

$$\lambda(u) = \frac{1}{2}\left(\frac{1}{\theta(u)} - \frac{1}{u}\left(1 - \frac{1}{n}\right)\right)(1 + o(1)), \quad u \to \infty.$$

Simple calculations leads to the conclusion that $\lambda(u) > 0$, as $u \to \infty$, for $\theta(u) < \sqrt{n/(n-1)}u$. Thus, $\lambda(u) > 0$ for all $u > 0$. □

In the proof of Lemma 2 it is established that $u < \theta(u) < \sqrt{n/(n-1)}u$. The numerical calculations, made with the help of Maple for $n = 2, 10, 50$, and presented in Table 2, confirm those inequalities.

Lemma 3. *For any $n > 1$ the function $f(u) = g^{-1}(u) - g_n^{-1}(u)$, $u > 0$, is strictly decreasing.*

Proof. Clearly,

$$f'(u) = \frac{1}{g'(g^{-1}(u))} - \frac{1}{g_n'(g_n^{-1}(u))}, \quad u > 0.$$

For any $u > 0$ we have

$$g(g^{-1}(u)) = g_n(g_n^{-1}(u)) < g(\theta(g_n^{-1}(u))), \tag{6}$$

where the function θ is defined as in Lemma 2. Since the function g is strictly decreasing, from (6) it follows that

$$g^{-1}(u) > \theta(g_n^{-1}(u)), \quad u > 0$$

while since the function g' is strictly increasing, we get

$$g_n'(g_n^{-1}(u)) = g'(\theta(g_n^{-1}(u))) < g'(g^{-1}(u)), \quad u > 0.$$

Thus, $f'(u) < 0$ for all $u > 0$. □

Acknowledgement

The authors are grateful to the referee for very helpful comments and useful suggestions.

References

Abramovitz, M., Stegun, I.A. (Eds.), 1965. Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables. Dover, New York.

Alzer, H., 1997. On some inequalities for the gamma and psi functions. Math. Comp. 66, 373–389.

Barndorff-Nielsen, O., 1978. Information and Exponential Families in Statistical Theory. Wiley, New York.

Bowman, K.O., Shenton, L.R., 1988. Properties of Estimators for the Gamma Distribution. Marcel Dekker, New York.

Crain, B.R., 1976. Exponential models, maximum likelihood estimation, and the Haar condition. J. Amer. Statist. Assoc. 71, 737–740.

Dang, H., Weerakkody, G., 2000. Bounds for the maximum likelihood estimates in two-parameter gamma distribution. J. Math. Anal. Appl. 245, 1–6.

Hájek, J., Šidák, Z., Sen, P.K., 1999. Theory of Rank Tests, second ed. Academic Press, San Diego.

Nagaev, A.V., 1996. Limit Theorems under Testing Hypotheses. UMK, Torun.