APPLICATIONES MATHEMATICAE
35,1 (2008), pp. 33–47

A. Zaigraev and A. Podraza-Karakulska (Toruń)

ON ASYMPTOTICS OF THE MAXIMUM LIKELIHOOD SCALE INVARIANT ESTIMATOR OF THE SHAPE PARAMETER OF THE GAMMA DISTRIBUTION

Abstract. The maximum likelihood scale invariant estimator of the shape parameter of the gamma distribution, proposed by the authors [Statist. Probab. Lett. 78 (2008)], is considered. The asymptotics of the mean square error of this estimator, with respect to that of the usual maximum likelihood estimator, is established.

1. Introduction. Let a sample $x = (x_1, \ldots, x_n)$ 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

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

Consider the problem of estimation of α. One of the most popular estimators is the well-known maximum likelihood estimator (ML-estimator) (e.g. [4, Sections 9.3, 9.4], [6], [7], [8]). Let

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

be the corresponding likelihood function. The ML-estimators α_n^* and σ_n^* of α and σ, respectively, are determined by the equations

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

2000 *Mathematics Subject Classification*: Primary 62F12; Secondary 62F10, 33B15.

Key words and phrases: gamma distribution, maximum likelihood estimator, maximum likelihood scale invariant estimator, mean square error, psi function.

where $\Psi(\alpha) = \frac{d}{d\alpha} \ln \Gamma(\alpha) = (\ln \Gamma(\alpha))'$ is the so-called Euler psi (digamma) function. Namely, α_n^* is the root of the equation

$$g(\alpha) = T(x)$$

with respect to α, while

$$\sigma_n^* = \frac{\overline{x}}{\alpha_n^*},$$

where

$$g(\alpha) = \ln \alpha - \Psi(\alpha), \qquad T(x) = \ln \overline{x} - \frac{1}{n} \sum_{j=1}^{n} \ln x_j, \quad \overline{x} = \frac{1}{n} \sum_{k=1}^{n} x_k.$$

Observe that the function g is strictly decreasing and takes values in $(0, \infty)$ (e.g. Theorem 1 of [2]). Therefore, the estimator α_n^* is well-defined and unique.

Moreover,

$$\mathrm{E}T(x) = g_n(\alpha), \tag{1}$$

where

$$g_n(\alpha) = \Psi(n\alpha) - \Psi(\alpha) - \ln n = g(\alpha) - g(n\alpha).$$

It is well-known (e.g. [10, Section 3.1], [11, Section 6.4]) that the limit distribution, as $n \to \infty$, of the random vector $n^{1/2}(\alpha_n^* - \alpha, \sigma_n^* - \sigma)$ is normal $\mathcal{N}(0, I^{-1}(\alpha, \sigma))$, i.e. with zero mean vector and covariance matrix $I^{-1}(\alpha, \sigma)$, where $I(\alpha, \sigma)$ is the Fisher information matrix having the form

$$I(\alpha, \sigma) = \begin{pmatrix} \Psi'(\alpha) & 1/\sigma \\ 1/\sigma & \alpha/\sigma^2 \end{pmatrix}.$$

This implies that the limit distribution of the random variable $n^{1/2}(\alpha_n^* - \alpha)$ is $\mathcal{N}(0, \kappa^2(\alpha))$, where

$$\kappa^2(\alpha) = (\psi'(\alpha) - 1/\alpha)^{-1} = -1/g'(\alpha). \tag{2}$$

Note that the estimator α_n^* is scale invariant. The question arises: *does there exist a better scale invariant estimator of α?* The positive answer is given in [13]. Estimating the shape parameter α, one can consider σ as a *nuisance* parameter. Therefore, it is natural to apply the maximum likelihood principle to the measure defined on the σ-algebra of scale invariant sets generated by the underlying gamma distribution. It is known (e.g. [9, Subsection 3.2.2], [12, Section 8.3]) that the density corresponding to this measure, with respect to that generated by $\mathcal{N}(0,1)$ distribution, is given as follows:

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

where

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

The maximum likelihood scale invariant estimator (IML-estimator) α_n^{**} of α is defined as $\alpha_n^{**} \in \arg\max_{\alpha>0} \mathbf{q}(x;\alpha)$. By direct calculations one can find that α_n^{**} is the root of the equation

$$g_n(\alpha) = T(x) \tag{3}$$

with respect to α. Therefore, the IML-estimator α_n^{**} coincides with that based on the method of moments. Of course, the estimator α_n^{**} is scale invariant, well-defined and unique since the function g_n is strictly decreasing and takes values in $(0,\infty)$ (see Lemma 1 of [13]).

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

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

Observe that for any $\lambda > 0$,

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

whence

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

Thus,

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

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

In [13] it is shown that the IML-estimator is better than the ML-estimator in the sense that it has smaller bias and smaller variance. The main goal of this paper is to establish the asymptotics of the mean square error of α_n^{**} compared to that of α_n^*.

The paper is organized as follows. Section 2 deals with the asymptotic normality of the IML-estimator. The main result is established in Section 3 while all the auxiliary results are formulated and proved in the Appendix.

2. Asymptotic normality of the IML-estimator. As already noted, the limit distribution of $n^{1/2}(\alpha_n^* - \alpha)$, as $n \to \infty$, is $\mathcal{N}(0, \kappa^2(\alpha))$, where $\kappa^2(\alpha)$ is defined by (2). Therefore, by Theorem 1.5 of [11, Section 5], the limit distribution of $n^{1/2}(g(\alpha_n^*) - g(\alpha))$ is $\mathcal{N}(0, \kappa^{-2}(\alpha))$. Since $g(\alpha_n^*) = g_n(\alpha_n^{**})$, the limit distribution of

$$n^{1/2}(g_n(\alpha_n^{**}) - g(\alpha)) = n^{1/2}(g(\alpha_n^{**}) - g(\alpha)) - n^{1/2}g(n\alpha_n^{**})$$

is also $\mathcal{N}(0, \kappa^{-2}(\alpha))$.

Now observe that the well-known asymptotic formula

$$\Psi(u) = \ln u - \frac{1}{2u} - \sum_{k=1}^{\infty} \frac{B_{2k}}{2ku^{2k}}, \qquad u \to \infty$$

(e.g. (6.3.18) of [1]), where $\{B_k\}$ are the so-called Bernoulli numbers, yields

$$g(u) = \frac{1}{2u} + \sum_{k=1}^{\infty} \frac{B_{2k}}{2ku^{2k}}, \qquad u \to \infty. \tag{4}$$

From (4) it follows that $n^{1/2}g(n\alpha) \to 0$ as $n \to \infty$. Therefore, applying the Slutsky theorem we infer that the limit distribution of

$$n^{1/2}(g(\alpha_n^{**}) - g(\alpha)) - n^{1/2}(g(n\alpha_n^{**}) - g(n\alpha))$$

is again $\mathcal{N}(0, \kappa^{-2}(\alpha))$.

From (4) one can also obtain that for any given α and all sufficiently large n we have $(n^{1/2}g(n\alpha))' \le 1$. Let $\varepsilon > 0$ be given. From Lemma 2 in the Appendix we get, as $n \to \infty$,

$$\begin{aligned} P(|n^{1/2}g(n\alpha_n^{**}) - n^{1/2}g(n\alpha)| \ge \varepsilon) &\le P(|\alpha_n^{**} - \alpha| \ge \varepsilon) \\ &= P(n^{1/2}|\alpha_n^{**} - \alpha| \ge n^{1/2}\varepsilon) \le ce^{-n\varepsilon^2(\Psi'(\alpha)-1/\alpha)/8} \to 0. \end{aligned}$$

Thus, as $n \to \infty$,

$$n^{1/2}(g(n\alpha_n^{**}) - g(n\alpha)) \to 0$$

in probability. Again, the Slutsky theorem implies that the limit distribution of

$$n^{1/2}(g(\alpha_n^{**}) - g(\alpha))$$

is $\mathcal{N}(0, \kappa^{-2}(\alpha))$. Finally, applying Theorem 1.5 of [11, Section 5] leads to the following result: the limit distribution of $n^{1/2}(\alpha_n^{**} - \alpha)$ is the same as that of $n^{1/2}(\alpha_n^* - \alpha)$, i.e. $\mathcal{N}(0, \kappa^2(\alpha))$.

3. Main result. Define

$$R_n^* = \mathrm{E}(\alpha_n^* - \alpha)^2, \qquad R_n^{**} = \mathrm{E}(\alpha_n^{**} - \alpha)^2.$$

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

$$n^2(R_n^* - R_n^{**}) = D(\alpha) + o(1), \qquad n \to \infty,$$

where

$$D(\alpha) = -\frac{3[g'(\alpha)/\alpha + 2g''(\alpha)]}{4\alpha(g'(\alpha))^3} > 0.$$

Proof. Take a number $1/3 < \delta < 1/2$ and divide the sample space $(0, \infty)^n$ into $X_n = X_{n,\delta}$ and $X_n^c = (0, \infty)^n \setminus X_n$, where

$$X_n = \{x : |\alpha_n^* - \alpha| < n^{-\delta},\ |\alpha_n^{**} - \alpha| < n^{-\delta}\}. \tag{5}$$

By the Cauchy–Schwarz inequality,

$$n^2\mathrm{E}((\alpha_n^* - \alpha)^2 \mathbf{1}_{X_n^c}(x)) \le (\mathrm{E}n^2(\alpha_n^* - \alpha)^4)^{1/2}(n^2 P(x \in X_n^c))^{1/2}.$$

In view of Lemmas 1 and 2 in the Appendix we obtain

$$\begin{aligned} n^2 P(x \in X_n^c) &\le n^2 P(|\alpha_n^* - \alpha| \ge n^{-\delta}) + n^2 P(|\alpha_n^{**} - \alpha| \ge n^{-\delta}) \\ &= n^2 P(n^{1/2}|\alpha_n^* - \alpha| \ge n^{1/2-\delta}) \\ &\quad + n^2 P(n^{1/2}|\alpha_n^{**} - \alpha| \ge n^{1/2-\delta}) \to 0, \qquad n \to \infty. \end{aligned}$$

Write for a moment $\xi_n = n^{1/2}(\alpha_n^* - \alpha)$. By Lemma 1 and integration by parts, for any $\gamma < \Psi'(\alpha)/8$ we have

$$\begin{aligned} \mathrm{E}e^{\gamma\xi_n^2} &= -\int_0^\infty e^{\gamma u^2}\, dP(|\xi_n| \ge u) = 1 + 2\gamma \int_0^\infty u e^{\gamma u^2} P(|\xi_n| \ge u)\, du \\ &\le 1 + 4\gamma \int_0^\infty u e^{(\gamma - \Psi'(\alpha)/8)u^2}\, du < \infty. \end{aligned}$$

Since for any $\gamma > 0$,

$$\gamma^2 z^4 < e^{\gamma z^2}, \qquad z > 0,$$

we deduce that for any $0 < \gamma < \Psi'(\alpha)/8$,

$$\mathrm{E}[n^{1/2}(\alpha_n^* - \alpha)]^4 = \mathrm{E}|\xi_n|^4 < \gamma^{-2}\mathrm{E}e^{\gamma\xi_n^2} < \infty. \tag{6}$$

Therefore,

$$n^2\mathrm{E}((\alpha_n^* - \alpha)^2 \mathbf{1}_{X_n^c}(x)) \to 0, \qquad n \to \infty.$$

A similar reasoning with an application of Lemma 2 leads to

$$n^2\mathrm{E}((\alpha_n^{**} - \alpha)^2 \mathbf{1}_{X_n^c}(x)) \to 0, \qquad n \to \infty.$$

Thus, in order to prove the theorem it is enough to consider only the case $x \in X_n$ as $n \to \infty$, which we assume until the end of the proof.

Applying the Taylor formula and formula (4) yields

$$
\begin{aligned}
g(\alpha_n^{**}) - g(\alpha_n^*) &= g(\alpha_n^{**}) - g_n(\alpha_n^{**}) = g(n\alpha_n^{**}) \\
&= g(n\alpha) + g'(n\alpha)n(\alpha_n^{**} - \alpha) + O(n^{-1-2\delta}) \\
&= \frac{1}{2n\alpha} - \frac{\alpha_n^{**} - \alpha}{2n\alpha^2} + O(n^{-1-2\delta}).
\end{aligned} \tag{7}
$$

On the other hand, using the Taylor formula one can obtain

$$
g(\alpha_n^{**}) - g(\alpha_n^*) = g'(\alpha_n^*)(\alpha_n^{**} - \alpha_n^*) + \frac{g''(\alpha_n^*)}{2}(\alpha_n^{**} - \alpha_n^*)^2 + O(|\alpha_n^{**} - \alpha_n^*|^3). \tag{8}
$$

Similarly,

$$
g'(\alpha_n^*) = g'(\alpha) + g''(\alpha)(\alpha_n^* - \alpha) + O(n^{-2\delta}), \tag{9}
$$

$$
g''(\alpha_n^*) = g''(\alpha) + O(n^{-\delta}). \tag{10}
$$

By substituting (9) and (10) into (8), we have

$$
\begin{aligned}
&g(\alpha_n^{**}) - g(\alpha_n^*) \\
&= g'(\alpha)(\alpha_n^{**} - \alpha_n^*)\left(1 + \frac{g''(\alpha)}{g'(\alpha)}(\alpha_n^* - \alpha) + \frac{g''(\alpha)}{2g'(\alpha)}(\alpha_n^{**} - \alpha_n^*) + O(n^{-2\delta})\right).
\end{aligned} \tag{11}
$$

Comparing (7) and (11), we get

$$
g'(\alpha)(\alpha_n^{**} - \alpha_n^*)(1 + L_n(\alpha) + O(n^{-2\delta})) = \frac{1}{2n\alpha} - \frac{\alpha_n^{**} - \alpha}{2n\alpha^2} + O(n^{-1-2\delta}), \tag{12}
$$

where

$$
L_n(\alpha) = \frac{g''(\alpha)}{2g'(\alpha)}[2(\alpha_n^* - \alpha) + \alpha_n^{**} - \alpha_n^*] = \frac{g''(\alpha)}{2g'(\alpha)}[\alpha_n^* - \alpha + \alpha_n^{**} - \alpha].
$$

From (12) it follows that

$$
\begin{aligned}
\alpha_n^{**} - \alpha_n^* &= \frac{\alpha - (\alpha_n^{**} - \alpha)}{2n\alpha^2 g'(\alpha)(1 + L_n(\alpha))} + O(n^{-1-2\delta}) \\
&= \frac{\alpha - (\alpha_n^{**} - \alpha)}{2n\alpha^2 g'(\alpha)}\left(1 - \frac{g''(\alpha)}{2g'(\alpha)}[\alpha_n^* - \alpha + \alpha_n^{**} - \alpha]\right) + O(n^{-1-2\delta}).
\end{aligned} \tag{13}
$$

This can be rewritten as

$$
\begin{aligned}
\alpha_n^{**} - \alpha = \alpha_n^* - \alpha &+ \frac{1}{2n\alpha g'(\alpha)} - \frac{\alpha_n^{**} - \alpha}{2n\alpha^2 g'(\alpha)} - \frac{g''(\alpha)(\alpha_n^{**} - \alpha)}{4n\alpha(g'(\alpha))^2} \\
&- \frac{g''(\alpha)(\alpha_n^* - \alpha)}{4n\alpha(g'(\alpha))^2} + O(n^{-1-2\delta}).
\end{aligned}
$$

Therefore, we obtain

$$
A_n(\alpha)(\alpha_n^{**} - \alpha) = B_n(\alpha)(\alpha_n^* - \alpha) + C_n(\alpha) + O(n^{-1-2\delta}), \tag{14}
$$

where

$$A_n(\alpha) = 1 + \frac{1}{2n\alpha^2 g'(\alpha)} + \frac{g''(\alpha)}{4n\alpha(g'(\alpha))^2},$$

$$B_n(\alpha) = 1 - \frac{g''(\alpha)}{4n\alpha(g'(\alpha))^2}, \qquad C_n(\alpha) = \frac{1}{2n\alpha g'(\alpha)}.$$

Set $b_n(\alpha) = \mathrm{E}(\alpha_n^* - \alpha)$. Applying the Taylor formula, we obtain

$$g(\alpha_n^*) - g(\alpha) = g'(\alpha)(\alpha_n^* - \alpha) + \frac{g''(\alpha)}{2}(\alpha_n^* - \alpha)^2 + O(n^{-3\delta}).$$

Further,

$$\mathrm{E}(g(\alpha_n^*) - g(\alpha)) = g'(\alpha)b_n(\alpha) + \frac{g''(\alpha)}{2} R_n^* + O(n^{-3\delta}). \tag{15}$$

On the other hand, by (3) and (1) we have

$$\mathrm{E}(g(\alpha_n^*) - g(\alpha)) = -g(n\alpha). \tag{16}$$

Comparing (15) and (16) and making use of (4) we get

$$b_n(\alpha) = -\frac{1}{2n\alpha g'(\alpha)} - \frac{g''(\alpha)}{2g'(\alpha)} R_n^* + O(n^{-3\delta}). \tag{17}$$

Observe that from (14) it follows that

$$A_n^2(\alpha)R_n^{**} = B_n^2(\alpha)R_n^* + C_n^2(\alpha) + 2B_n(\alpha)C_n(\alpha)b_n(\alpha) + O(n^{-2-2\delta}),$$

or, by substituting (17),

$$\begin{aligned} A_n^2(\alpha)R_n^{**} = B_n^2(\alpha)R_n^* &+ C_n^2(\alpha) - 2B_n(\alpha)C_n^2(\alpha) \\ &- B_n(\alpha)C_n(\alpha)\frac{g''(\alpha)}{g'(\alpha)}R_n^* + O(n^{-1-3\delta}). \end{aligned}$$

Therefore,

$$\begin{aligned} n^2(R_n^{**} - R_n^*) = n^2 R_n^* &\left(\frac{B_n^2(\alpha) - B_n(\alpha)C_n(\alpha)g''(\alpha)/g'(\alpha)}{A_n^2(\alpha)} - 1\right) \\ &+ n^2\,\frac{C_n^2(\alpha) - 2B_n(\alpha)C_n^2(\alpha)}{A_n^2(\alpha)} + o(1). \end{aligned} \tag{18}$$

Since the limit distribution of $n^{1/2}(\alpha_n^* - \alpha)$, as $n \to \infty$, is $\mathcal{N}(0, \kappa^2(\alpha))$, and

$$\mathrm{E}[n^{1/2}(\alpha_n^* - \alpha)]^4 < \infty$$

(see (6)), by the moment continuity theorem (e.g. Theorem 4 of [5, Section 1, §6]) we obtain

$$\mathrm{E}[n^{1/2}(\alpha_n^* - \alpha)]^2 \to \kappa^2(\alpha) = -1/g'(\alpha), \qquad n \to \infty.$$

Thus,

$$nR_n^* = -1/g'(\alpha) + o(1), \qquad n \to \infty. \tag{19}$$

The proof is finished by substituting (19) into (18) and taking into account Lemma 3 below, formula (4) and the relations

$$n\left(\frac{B_n^2(\alpha)-B_n(\alpha)C_n(\alpha)g''(\alpha)/g'(\alpha)}{A_n^2(\alpha)}-1\right)=-\frac{3g''(\alpha)}{2\alpha(g'(\alpha))^2}-\frac{1}{\alpha^2g'(\alpha)}+o(1),$$

$$n^2\frac{C_n^2(\alpha)-2B_n(\alpha)C_n^2(\alpha)}{A_n^2(\alpha)}=-\frac{1}{4\alpha^2(g'(\alpha))^2}+o(1).$$

Appendix

LEMMA 1. *If a sample $x=(x_1,\dots,x_n)$ is drawn from a $\Gamma(\alpha,\sigma)$ distribution, then for any $z>0$ and any $c>1$ there exists $N=N(z,c)$ such that for all $n\ge N$,*

$$P(n^{1/2}|\alpha_n^*-\alpha|\ge z)\le ce^{-\Psi'(\alpha)z^2/8}.$$

Proof. Consider the random function (cf. [5, §23])

$$Z_n(\beta)=\frac{\mathbf{p}(x;\alpha+\beta,1)}{\mathbf{p}(x;\alpha,1)},\qquad \beta>-\alpha.$$

Fix $z>0$. Since

$$\{|\alpha_n^*-\alpha|\ge z\}=\{\sup_{|\beta|\ge z}Z_n(\beta)\ge\sup_{|\beta|\le z}Z_n(\beta)\}\subset\{\sup_{|\beta|\ge z}Z_n(\beta)\ge Z_n(0)=1\},$$

we have

$$P(n^{1/2}|\alpha_n^*-\alpha|\ge z)\le P(\sup_{\beta\in B}Z_n(\beta)\ge 1),$$

where

$$B=(-\alpha,-z/\sqrt{n}]\cup[z/\sqrt{n},\infty).$$

From the Markov inequality, we obtain

$$P(\sup_{\beta\in B}Z_n(\beta)\ge 1)\le \mathrm{E}(\sup_{\beta\in B}Z_n(\beta))^{1/2}=\mathrm{E}\sup_{\beta\in B}Z_n^{1/2}(\beta).$$

Then

$$\mathrm{E}\sup_{\beta\in B}Z_n^{1/2}(\beta)\le\Big(\sup_{\beta\in B}\int_0^\infty p^{1/2}(u;\alpha,1)p^{1/2}(u;\alpha+\beta,1)\,du\Big)^n.$$

Simple calculation yields

$$\int_0^\infty p^{1/2}(u;\alpha,1)p^{1/2}(u;\alpha+\beta,1)\,du=\frac{\Gamma(\alpha+\beta/2)}{(\Gamma(\alpha)\Gamma(\alpha+\beta))^{1/2}}.$$

Let us investigate the properties of the function

$$\varrho(u)=\frac{\Gamma(\alpha+u/2)}{(\Gamma(\alpha)\Gamma(\alpha+u))^{1/2}},\qquad u>-\alpha.$$

Clearly,

$$(\ln \varrho(u))' = \frac{1}{2}\left[\Psi\left(\alpha + \frac{u}{2}\right) - \Psi(\alpha + u)\right]. \tag{20}$$

Since $\Psi(u)$ is increasing, it follows that $\varrho(u)$ is increasing in $(-\alpha, 0)$ and decreasing in $(0, \infty)$ with maximum equal to 1 at $u = 0$. Furthermore,

$$\ln \varrho(-u) < \ln \varrho(u), \qquad 0 < u < \alpha.$$

Indeed, the function $\widetilde{\varrho}(u) = \ln \varrho(u) - \ln \varrho(-u)$ has $\widetilde{\varrho}(0) = 0$ and

$$\widetilde{\varrho}'(u) = \frac{1}{2}\left[\Psi\left(\alpha + \frac{u}{2}\right) + \Psi\left(\alpha - \frac{u}{2}\right) - \Psi(\alpha + u) - \Psi(\alpha - u)\right] > 0$$

since the function $\Psi(\alpha + z) + \Psi(\alpha - z)$ is decreasing because $(\Psi(\alpha + z) + \Psi(\alpha - z))' < 0$ in view of $\Psi'(\alpha + z) < \Psi'(\alpha - z)$ for $z > 0$ ($\Psi'(u)$ is decreasing).

Hence,

$$\sup_{\beta \in B} \varrho(\beta) = \varrho\left(\frac{z}{\sqrt{n}}\right).$$

By the Taylor formula, for all sufficiently large n we obtain

$$\varrho\left(\frac{z}{\sqrt{n}}\right) = 1 + \frac{\varrho''(0)}{2} \cdot \frac{z^2}{n} + o(n^{-1}). \tag{21}$$

But from (20) it follows that

$$\varrho'(u) = \frac{\varrho(u)}{2}\left[\Psi\left(\alpha + \frac{u}{2}\right) - \Psi(\alpha + u)\right].$$

Therefore,

$$\varrho''(u) = \frac{\varrho'(u)}{2}\left[\Psi\left(\alpha + \frac{u}{2}\right) - \Psi(\alpha + u)\right] + \frac{\varrho(u)}{2}\left[\frac{1}{2}\Psi'\left(\alpha + \frac{u}{2}\right) - \Psi'(\alpha + u)\right]$$

and

$$\varrho''(0) = -\frac{1}{4}\Psi'(\alpha) < 0.$$

Substitution into (21) yields

$$\varrho\left(\frac{z}{\sqrt{n}}\right) = 1 - \frac{\Psi'(\alpha) z^2}{8n} + o(n^{-1}).$$

Thus, for any $c > 1$ and all sufficiently large n we obtain

$$\Big(\sup_{\beta \in B} \int_0^\infty p^{1/2}(u; \alpha, 1) p^{1/2}(u; \alpha + \beta, 1)\, du\Big)^n = \left(1 - \frac{\Psi'(\alpha) z^2}{8n} + o(n^{-1})\right)^n \le c e^{-\Psi'(\alpha) z^2/8}.$$

The proof is complete.

LEMMA 2. *If a sample $x = (x_1, \dots, x_n)$ is drawn from a $\Gamma(\alpha, \sigma)$ distribution, then for any $z > 0$ and any $c > 1$ there exists $N = N(z, c)$ such that for all $n \geq N$,*

$$P(n^{1/2}|\alpha_n^{**} - \alpha| \geq z) \leq ce^{-(\Psi'(\alpha)-1/\alpha)z^2/8}.$$

Proof. The method of establishing this result is similar to that of proving Lemma 1. Consider the random function

$$Z_n(\beta) = \frac{\mathbf{q}(x; \alpha+\beta)}{\mathbf{q}(x; \alpha)} = \frac{\Gamma(n(\alpha+\beta))(\prod_{i=1}^n x_i)^{\alpha+\beta-1}}{(\Gamma(\alpha+\beta))^n(\sum_{i=1}^n x_i)^{n(\alpha+\beta)}} \cdot \frac{(\Gamma(\alpha))^n(\sum_{i=1}^n x_i)^{n\alpha}}{\Gamma(n\alpha)(\prod_{i=1}^n x_i)^{\alpha-1}}$$
$$= \frac{\Gamma(n(\alpha+\beta))(\Gamma(\alpha))^n}{(\Gamma(\alpha+\beta))^n\Gamma(n\alpha)} \cdot \frac{(\prod_{i=1}^n x_i)^{\beta}}{(\sum_{i=1}^n x_i)^{n\beta}}.$$

Then

$$Z_n^{1/2}(\beta) = \frac{(\Gamma(n(\alpha+\beta)))^{1/2}(\Gamma(\alpha))^{n/2}}{(\Gamma(\alpha+\beta))^{n/2}(\Gamma(n\alpha))^{1/2}} \cdot \frac{(\prod_{i=1}^n x_i)^{\beta/2}}{(\sum_{i=1}^n x_i)^{n\beta/2}}.$$

Therefore,

$$\mathrm{E}Z_n^{1/2}(\beta) = \frac{(\Gamma(n(\alpha+\beta)))^{1/2}}{(\Gamma(\alpha+\beta))^{n/2}(\Gamma(n\alpha))^{1/2}(\Gamma(\alpha))^{n/2}}$$
$$\cdot \int_0^\infty \dots \int_0^\infty \frac{(u_1 \dots u_n)^{\alpha+\beta/2-1}}{(u_1 + \dots + u_n)^{n\beta/2}} e^{-(u_1+\dots+u_n)}\, du_1 \dots du_n.$$

The change of variables $v_1 = u_1, \dots, v_{n-1} = u_{n-1}, v_n = u_1 + \dots + u_n$ yields

$$\mathrm{E}Z_n^{1/2}(\beta) = \frac{(\Gamma(n(\alpha+\beta)))^{1/2}}{(\Gamma(\alpha+\beta))^{n/2}(\Gamma(n\alpha))^{1/2}(\Gamma(\alpha))^{n/2}}$$
$$\cdot \int_0^\infty \dots \int_0^\infty \int_{v_1+\dots+v_{n-1}}^\infty \frac{[v_1 \dots v_{n-1}(v_n - v_1 - \dots - v_{n-1})]^{\alpha+\beta/2-1}}{v_n^{n\beta/2}}$$
$$\cdot e^{-v_n}\, dv_n \dots dv_2 dv_1.$$

The next change of variables $v_1 = v_n z_1, \dots, v_{n-1} = v_n z_{n-1}$ gives

$$\mathrm{E}Z_n^{1/2}(\beta) = \frac{(\Gamma(n(\alpha+\beta)))^{1/2}}{(\Gamma(\alpha+\beta))^{n/2}(\Gamma(n\alpha))^{1/2}(\Gamma(\alpha))^{n/2}} \int_0^\infty v_n^{n\alpha-1} e^{-v_n}\, dv_n$$
$$\cdot \int \dots \int_A [z_1 \dots z_{n-1}(1 - z_1 - \dots - z_{n-1})]^{\alpha+\beta/2-1}\, dz_1 \dots dz_{n-1}$$
$$= \frac{(\Gamma(n(\alpha+\beta)))^{1/2}(\Gamma(n\alpha))^{1/2}}{(\Gamma(\alpha+\beta))^{n/2}(\Gamma(\alpha))^{n/2}}$$
$$\cdot \int \dots \int_A [z_1 \dots z_{n-1}(1 - z_1 - \dots - z_{n-1})]^{\alpha+\beta/2-1}\, dz_1 \dots dz_{n-1}$$

$$= \frac{(\Gamma(n(\alpha+\beta)))^{1/2}(\Gamma(n\alpha))^{1/2}}{(\Gamma(\alpha+\beta))^{n/2}(\Gamma(\alpha))^{n/2}} \int_0^1 z_1^{\alpha+\beta/2-1} dz_1 \int_0^{1-z_1} z_2^{\alpha+\beta/2-1}\, dz_2 \ldots$$

$$\ldots \int_0^{1-z_1-\cdots-z_{n-2}} z_{n-1}^{\alpha+\beta/2-1}(1-z_1-\cdots-z_{n-2}-z_{n-1})^{\alpha+\beta/2-1}\, dz_{n-1}$$

$$= \frac{(\Gamma(n(\alpha+\beta)))^{1/2}(\Gamma(n\alpha))^{1/2}}{(\Gamma(\alpha+\beta))^{n/2}(\Gamma(\alpha))^{n/2}}\, I_n,$$

where

$$A = \{(z_1,\ldots,z_{n-1}) : z_1 > 0,\ldots,z_{n-1} > 0,\ z_1+\cdots+z_{n-1} < 1\}.$$

Now we calculate I_n. Observe that

$$\int_0^a u^y(a-u)^z\, du = a^{y+z+1}B(y+1,z+1), \qquad y,z > -1,$$

where $B(\cdot,\cdot)$ is the beta-function. Then

$$\int_0^{1-z_1-\cdots-z_{n-2}} z_{n-1}^{\alpha+\beta/2-1}(1-z_1-\cdots-z_{n-2}-z_{n-1})^{\alpha+\beta/2-1} dz_{n-1}$$
$$= (1-z_1-\cdots-z_{n-2})^{2\alpha+\beta-1}B(\alpha+\beta/2,\alpha+\beta/2).$$

Therefore,

$$I_n = B(\alpha+\beta/2,\alpha+\beta/2)\int_0^1 z_1^{\alpha+\beta/2-1}\, dz_1 \int_0^{1-z_1} z_2^{\alpha+\beta/2-1}\, dz_2 \ldots$$
$$\ldots \int_0^{1-z_1-\cdots-z_{n-3}} z_{n-2}^{\alpha+\beta/2-1}(1-z_1-\cdots-z_{n-3}-z_{n-2})^{2\alpha+\beta-1}\, dz_{n-2}.$$

Again,

$$\int_0^{1-z_1-\cdots-z_{n-3}} z_{n-2}^{\alpha+\beta/2-1}(1-z_1-\cdots-z_{n-3}-z_{n-2})^{2\alpha+\beta-1}\, dz_{n-2}$$
$$= (1-z_1-\cdots-z_{n-3})^{3\alpha+3\beta/2-1}B(\alpha+\beta/2,2(\alpha+\beta/2)),$$

and we obtain

$$I_n = B(\alpha+\beta/2,\alpha+\beta/2)B(\alpha+\beta/2,2(\alpha+\beta/2))$$
$$\cdot\int_0^1 z_1^{\alpha+\beta/2-1}\, dz_1 \int_0^{1-z_1} z_2^{\alpha+\beta/2-1}\, dz_2 \ldots$$
$$\ldots \int_0^{1-z_1-\cdots-z_{n-4}} z_{n-3}^{\alpha+\beta/2-1}(1-z_1-\cdots-z_{n-4}-z_{n-3})^{3\alpha+3\beta/2-1}\, dz_{n-3}.$$

Repeating this calculation, we get

$$\begin{aligned} I_n &= B(\alpha+\beta/2, \alpha+\beta/2) B(\alpha+\beta/2, 2(\alpha+\beta/2)) \\ &\quad \ldots B(\alpha+\beta/2, (n-1)(\alpha+\beta/2)) \\ &= \frac{(\Gamma(\alpha+\beta/2))^n}{\Gamma(n(\alpha+\beta/2))}. \end{aligned}$$

Thus,

$$\mathrm{E} Z_n^{1/2}(\beta) = \frac{(\Gamma(n(\alpha+\beta)))^{1/2} (\Gamma(\alpha+\beta/2))^n (\Gamma(n\alpha))^{1/2}}{(\Gamma(\alpha+\beta))^{n/2} \Gamma(n(\alpha+\beta/2)) (\Gamma(\alpha))^{n/2}} = (\Delta_n(\beta))^n.$$

Let us investigate the properties of the function

$$\Delta_n(u) = \frac{\Gamma(\alpha+u/2)(\Gamma(n(\alpha+u)))^{1/(2n)} (\Gamma(n\alpha))^{1/(2n)}}{(\Gamma(\alpha+u))^{1/2} (\Gamma(\alpha))^{1/2} (\Gamma(n(\alpha+u/2)))^{1/n}}, \qquad u > -\alpha.$$

Clearly,

$$\begin{aligned} \ln \Delta_n(u) &= \ln \Gamma(\alpha+u/2) - \frac{1}{2} \ln \Gamma(\alpha+u) - \frac{1}{2} \ln \Gamma(\alpha) \\ &\quad + \frac{1}{2n} \ln \Gamma(n(\alpha+u)) + \frac{1}{2n} \ln \Gamma(n\alpha) \\ &\quad - \frac{1}{n} \ln \Gamma(n(\alpha+u/2)), \end{aligned}$$

$$\begin{aligned} (\ln \Delta_n(u))' &= \frac{1}{2} \Psi(\alpha+u/2) - \frac{1}{2} \Psi(\alpha+u) + \frac{1}{2} \Psi(n(\alpha+u)) \\ &\quad - \frac{1}{2} \Psi(n(\alpha+u/2)) \\ &= \frac{1}{2} [g_n(\alpha+u) - g_n(\alpha+u/2)]. \end{aligned} \tag{22}$$

Since $g_n(u)$ is decreasing (see Lemma 1 of [13]), it follows that $\Delta_n(u)$ is increasing in $(-\alpha, 0)$ and decreasing in $(0, \infty)$ with maximum equal to 1 at $u = 0$. Furthermore,

$$\ln \Delta_n(-u) < \ln \Delta_n(u), \qquad 0 < u < \alpha.$$

Indeed, the function $\widetilde{\Delta}_n(u) = \ln \Delta_n(u) - \ln \Delta_n(-u)$ has $\widetilde{\Delta}_n(0) = 0$ and

$$\widetilde{\Delta}_n'(u) = \frac{1}{2} \left[g_n(\alpha+u) + g_n(\alpha-u) - g_n\left(\alpha + \frac{u}{2}\right) - g_n\left(\alpha - \frac{u}{2}\right) \right] > 0$$

since the function $\zeta(z) = g_n(\alpha+z) + g_n(\alpha-z)$ is increasing because $(g_n(\alpha+z) + g_n(\alpha-z))' > 0$ in view of $g_n'(\alpha+z) > g_n'(\alpha-z)$ for $z > 0$ ($g_n'(u)$ is increasing).

Hence,

$$\sup_{\beta \in B} \Delta_n(\beta) = \Delta_n\left(\frac{z}{\sqrt{n}}\right).$$

Now observe that from (22) it follows that

$$\Delta_n'(u) = \frac{\Delta_n(u)}{2}\left[g_n(\alpha+u) - g_n\left(\alpha+\frac{u}{2}\right)\right],$$

$$\Delta_n''(u) = \frac{\Delta_n'(u)}{2}\left[g_n(\alpha+u) - g_n\left(\alpha+\frac{u}{2}\right)\right] + \frac{\Delta_n(u)}{2}\left[g_n'(\alpha+u) - \frac{1}{2}\,g_n'\left(\alpha+\frac{u}{2}\right)\right],$$

$$\Delta_n'''(u) = \frac{\Delta_n''(u)}{2}\left[g_n(\alpha+u) - g_n\left(\alpha+\frac{u}{2}\right)\right] + \Delta_n'(u)\left[g_n'(\alpha+u) - \frac{1}{2}\,g_n'\left(\alpha+\frac{u}{2}\right)\right] + \frac{\Delta_n(u)}{2}\left[g_n''(\alpha+u) - \frac{1}{4}\,g_n''\left(\alpha+\frac{u}{2}\right)\right].$$

Therefore,

$$\Delta_n^{(k)}(0) = \frac{1}{2}\left(1-\frac{1}{2^{k-1}}\right)g_n^{(k-1)}(\alpha), \qquad k \ge 1.$$

Since for all sufficiently large n (see (4)),

$$g_n^{(k-1)}(\alpha) = g^{(k-1)}(\alpha) + O(n^{-1}),$$

by the Taylor formula we obtain, for all sufficiently large n,

$$\Delta_n\left(\frac{z}{\sqrt{n}}\right) = 1 + \frac{\Delta_n''(0)}{2}\cdot\frac{z^2}{n} + O(n^{-2}) = 1 + \frac{g'(\alpha)z^2}{8n} + O(n^{-2}).$$

Thus, for any $c > 1$ and all sufficiently large n,

$$\mathrm{E}\sup_{\beta\in B} Z_n^{1/2}(\beta) \le \left(1 - \frac{(\Psi'(\alpha)-1/\alpha)z^2}{8n} + o(n^{-1})\right)^n \le ce^{-(\Psi'(\alpha)-1/\alpha)z^2/8}.$$

The proof is complete.

LEMMA 3. *For $u > 0$, $g'(u)/u + 2g''(u) > 0$.*

Proof. We use the method utilized e.g. in the proof of Lemma 1 in [3]. Consider the function

$$q(u) = \frac{g'(u)}{u} + 2g''(u) = -\frac{1}{u^2} - \frac{\Psi'(u)}{u} - 2\Psi''(u), \qquad u > 0.$$

From the integral representations [1, formula (6.4.1)]

$$\Psi'(u) = \int_0^\infty \frac{te^{-ut}}{1-e^{-t}}\,dt, \qquad \Psi''(u) = -\int_0^\infty \frac{t^2e^{-ut}}{1-e^{-t}}\,dt,$$

and the evident relations

$$\frac{1}{u} = \int_0^\infty e^{-ut}\, dt, \qquad \frac{1}{u^2} = \int_0^\infty te^{-ut}\, dt,$$

by the convolution theorem for Laplace transforms we get

$$q(u) = \int_0^\infty e^{-ut} d(t)\, dt, \qquad u > 0,$$

where

$$d(t) = -t + \frac{2t^2}{1-e^{-t}} - \int_0^t \frac{v}{1-e^{-v}}\, dv, \qquad t > 0.$$

Differentiating yields

$$d'(t) = \frac{h(t)}{(e^t-1)^2}, \qquad t > 0,$$

where

$$h(t) = 3t(e^{2t}-e^t) - 2t^2e^t - (e^t-1)^2 = (3t-1)e^{2t} - (2t^2+3t-2)e^t - 1.$$

Making use of the series representation for e^t, we obtain

$$h(t) = \sum_{k=3}^\infty \frac{c_k t^k}{k!},$$

where

$$c_k = 2^{k-1}(3k-2) + 2 - k - 2k^2.$$

Since $c_k > 0$ for any $k \geq 3$, we conclude that $d'(t) > 0$, and therefore, for any $t > 0$,

$$d(t) > \lim_{s\to 0} d(s) = 0.$$

Thus, $q(u) > 0$ for any $u > 0$. The proof is complete.

References

[1] M. Abramovitz and I. A. Stegun (eds.), *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, Dover, New York, 1965.

[2] H. Alzer, *On some inequalities for the gamma and psi functions*, Math. Comp. 66 (1997), 373–389.

[3] —, *Inequalities for the gamma function*, Proc. Amer. Math. Soc. 128 (1999), 141–147.

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

[5] A. A. Borovkov, *Mathematical Statistics*, Nauka, Moscow, 1984 (in Russian).

[6] K. O. Bowman and L. R. Shenton, *Properties of Estimators for the Gamma Distribution*, Dekker, New York, 1988.

[7] B. R. Crain, *Exponential models, maximum likelihood estimation, and the Haar condition*, J. Amer. Statist. Assoc. 71 (1976), 737–740.

[8] H. Dang and G. Weerakkody, *Bounds for the maximum likelihood estimates in two-parameter gamma distribution*, J. Math. Anal. Appl. 245 (2000), 1–6.

[9] J. Hájek, Z. Šidák and P. K. Sen, *Theory of Rank Tests*, 2nd ed., Academic Press, San Diego, 1999.

[10] I. A. Ibragimov and R. Z. Has'minskii, *Statistical Estimation Asymptotic Theory*, Springer, New York, 1981.

[11] E. L. Lehmann, *Theory of Point Estimation*, Wiley, New York, 1987.

[12] A. Nagaev, *Limit theorems under testing hypotheses*, UMK, Toruń, 1996.

[13] A. Zaigraev and A. Podraza-Karakulska, *On estimation of the shape parameter of the gamma distribution*, Statist. Probab. Letters 78 (2008), 286–295.

Faculty of Mathematics and Computer Science
Nicolaus Copernicus University of Toruń
Chopina 12/18
87-100 Toruń, Poland
E-mail: alzaig@mat.uni.torun.pl
agap@mat.uni.torun.pl

Received on 10.9.2007;
revised version on 22.2.2008 (1886)