APPLICATIONES MATHEMATICAE
39,2 (2012), pp. 129–135

Piotr Nowak (Wrocław)

STOCHASTIC COMPARISONS OF MOMENT ESTIMATORS OF GAMMA DISTRIBUTION PARAMETERS

Abstract. Recently the order preserving property of estimators has been intensively studied, e.g. by Gan and Balakrishnan and collaborators. In this paper we prove the stochastic monotonicity of moment estimators of gamma distribution parameters using the standard coupling method and majorization theory. We also give some properties of the moment estimator of the shape parameter and derive an approximate confidence interval for this parameter.

1. Introduction and preliminaries. Suppose that $\mathbf{X} = (X_1, \ldots, X_n)$ is a sample from the gamma distribution with density

$$f(x;\alpha,\lambda) = \frac{1}{\Gamma(\alpha)\lambda^\alpha} x^{\alpha-1}\exp(-x/\lambda), \qquad x>0,\ \lambda>0,\ \alpha>0.$$

The gamma distribution is one of the most popular distributions that arise in reliability and statistics and the problem of estimating its parameters is important. In this case the likelihood equation does not have an explicit solution and numerical methods have to be used to compute maximum likelihood estimators. On the other hand the method of moments gives very simple estimators.

It is well known that $E(X_1) = \alpha\lambda$ and $\mathrm{Var}(X_1) = \alpha\lambda^2$. Moment estimators of the parameters α and λ are

$$\hat{\alpha} = \frac{\bar{X}^2}{S^2} \quad \text{and} \quad \hat{\lambda} = \frac{S^2}{\bar{X}}, \tag{1.1}$$

where $\bar{X}$ and S^2 are the sample mean and the sample biased variance respectively.

2010 *Mathematics Subject Classification*: 60E15, 62E99, 62N05.
Key words and phrases: stochastic ordering, gamma distribution, moment estimator, confidence interval.

DOI: 10.4064/am39-2-1

For completeness we recall some definitions of stochastic orders (see Shaked and Shanthikumar [10]).

Let X and Y be two random variables, and F and G their respective probability distribution functions. We say that X is *stochastically smaller* than Y ($X \le_{\mathrm{st}} Y$) if $F(x) \ge G(x)$ for every x. We say that the family $\{X(\theta) : \theta \in \Theta \subset \mathbb{R}\}$ of random variables is *stochastically increasing* in θ if $X(\theta_1) \le_{\mathrm{st}} X(\theta_2)$ whenever $\theta_1 < \theta_2$. The estimator $\hat{\gamma}$ of the function $\gamma(\theta)$, $\theta \in \Theta$, is said to be stochastically increasing in θ if the family of distributions of $\hat{\gamma}$ is stochastically increasing in θ. Let F^{-1} and G^{-1} be quantile functions of F and G respectively. We say that X is *less dispersed* than Y (denoted $X \le_{\mathrm{disp}} Y$) if $G^{-1}(\alpha) - F^{-1}(\alpha)$ is an increasing function of $\alpha \in (0,1)$. It is well known that if $X \le_{\mathrm{disp}} Y$ then $\mathrm{Var}(X) \le \mathrm{Var}(Y)$ provided that the variances exist. Also if X and Y are such that their supports have a common finite left endpoint, then $X \le_{\mathrm{disp}} Y$ implies $X \le_{\mathrm{st}} Y$.

Analogously, we say that the family $\{X(\theta) : \theta \in \Theta \subset \mathbb{R}\}$ of random variables is *dispersively increasing* in θ if $X(\theta_1) \le_{\mathrm{disp}} X(\theta_2)$ whenever $\theta_1 < \theta_2$, and the estimator $\hat{\gamma}$ is dispersively increasing in θ if the family of distributions of this estimator is dispersively increasing in θ.

Now suppose that X and Y are nonnegative random variables. We say that X is *smaller* than Y *in the convex transform order* (denoted $X \le_{\mathrm{c}} Y$) if $G^{-1}F$ is convex on the support of F.

It is easy to prove that the family of gamma distributions is stochastically increasing with respect to both parameters, α and λ. If X and Y have the gamma distributions $\mathcal{G}(1, \alpha_1)$ and $\mathcal{G}(1, \alpha_2)$ respectively, where $\alpha_1 < \alpha_2$, then $X \ge_{\mathrm{c}} Y$ (see van Zwet [11]) and also $X \le_{\mathrm{disp}} Y$ (see Shaked [9]).

A vector $\mathbf{x} = (x_1, \ldots, x_n)$ is said to be *smaller in the majorization order* than the vector $\mathbf{y} = (y_1, \ldots, y_n)$, or $\mathbf{x}$ is *majorized* by $\mathbf{y}$ (denoted $\mathbf{x} \prec \mathbf{y}$) if $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$ and if $\sum_{i=1}^j x_{(i)} \ge \sum_{i=1}^j y_{(i)}$ for $j = 1, \ldots, n-1$, where $x_{(i)}$ (resp. $y_{(i)}$) is the ith smallest element of $\mathbf{x}$ (resp. $\mathbf{y}$), $i = 1, \ldots, n$.

We recall a classical result of mathematical analysis.

THEOREM 1.1. $\mathbf{x}$ *is majorized by* $\mathbf{y}$ *if and only if*

$$\sum_{i=1}^n \phi(x_i) \le \sum_{i=1}^n \phi(y_i)$$

for all convex functions $\phi : \mathbb{R} \to \mathbb{R}$.

We refer to Marshall and Olkin [7] for comprehensive details on the theory of majorization.

One of the most useful and simplest methods of proving stochastic monotonicity is the coupling method (see Lindvall [6]), described in the following lemma.

LEMMA 1.2.

(a) *If* $X \le_{\rm st} Y$ *then there exists a random variable* $Z =_{\rm st} X$ *such that* $Z \le Y$ *pointwise.*

(b) *If* $X =_{\rm st} Z$ *and* $Z \le Y$ *pointwise, then* $X \le_{\rm st} Y$.

The following lemma may be easily proved using the coupling method.

LEMMA 1.3. *Let* X, Y, Z *be nonnegative random variables such that* $X \le_{\rm st} Y$ *and* Z *is independent of* X *and* Y. *Then* $X \cdot Z \le_{\rm st} Y \cdot Z$.

2. Results. In this section we prove the stochastic monotonicity of moment estimators given by (1.1). We also discuss the problem of dispersive ordering of the estimator $1/\hat\alpha$.

THEOREM 2.1. *The estimator* $\hat\lambda$ *is stochastically increasing in* λ.

Proof. Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a sample from the gamma distribution $\mathcal{G}(\lambda_1, \alpha)$ and similarly $\mathbf{Y} = (Y_1, \ldots, Y_n)$ be a sample from the gamma distribution $\mathcal{G}(\lambda_2, \alpha)$, where $\lambda_1 < \lambda_2$. Assume that the random vectors $\mathbf{X}$ and $\mathbf{Y}$ are independent. We show that $\hat\lambda_1 \le_{\rm st} \hat\lambda_2$, where

$$\hat\lambda_1 = \frac{S_X^2}{\bar X} \quad \text{and} \quad \hat\lambda_2 = \frac{S_Y^2}{\bar Y}.$$

For fixed α the sample mean $\bar X$ is a complete sufficient statistic for λ and $S_X^2/\bar X^2$ is an ancillary statistic for λ, hence by the Basu theorem (see e.g. Lehmann and Casella [5, p. 42]) they are independent random variables. From the assumptions it also follows that $\bar X \le_{\rm st} \bar Y$. Using Lemma 1.3 we have

$$\hat\lambda_1 = \frac{S_X^2}{\bar X} = \frac{S_X^2}{\bar X^2} \cdot \bar X \le_{\rm st} \frac{S_X^2}{\bar X^2} \cdot \bar Y =_{\rm st} \frac{S_Y^2}{\bar Y^2} \cdot \bar Y = \frac{S_Y^2}{\bar Y} = \hat\lambda_2.$$

The above stochastic equivalence holds since $S_X^2/\bar X^2$ is an ancillary statistic for λ, and $S_Y^2/\bar Y^2$ and $\bar Y$ are independent. ∎

Similarly we can formulate the next theorem about the moment estimator for α.

THEOREM 2.2. *The estimator* $\hat\alpha$ *is stochastically increasing in* α.

Proof. We use the coupling method. Assume that $\mathbf{X} = (X_1, \ldots, X_n)$ is a sample from the gamma distribution $\mathcal{G}(\lambda, \alpha_1)$ and similarly $\mathbf{Y} = (Y_1, \ldots, Y_n)$ is a sample from the gamma distribution $\mathcal{G}(\lambda, \alpha_2)$, where $\alpha_1 < \alpha_2$. Without loss of generality we can assume that $\lambda = 1$. Let F and G denote the distribution functions of $\mathcal{G}(1, \alpha_1)$ and $\mathcal{G}(1, \alpha_2)$ respectively. Equivalently we prove that

$$\frac{\sum_{i=1}^n X_i^2}{(\sum_{i=1}^n X_i)^2} \ge_{\rm st} \frac{\sum_{i=1}^n Y_i^2}{(\sum_{i=1}^n Y_i)^2}.$$

Observe that

$$\frac{\sum_{i=1}^n X_i^2}{(\sum_{i=1}^n X_i)^2} = \sum_{i=1}^n \left[\frac{X_{(i)}}{X_{(1)} + \cdots + X_{(n)}}\right]^2,$$

where $X_{(1)} \le \cdots \le X_{(n)}$ are the order statistics of $\mathbf{X}$. It is well known that

$$\mathbf{Y}^* = (Y_{(1)}, \ldots, Y_{(n)}) =_{\mathrm{st}} (G^{-1}F(X_{(1)}), \ldots, G^{-1}F(X_{(n)})),$$

where $\mathbf{Y}^*$ is the vector of the order statistics of $\mathbf{Y}$. Since the family of gamma distributions is stochastically increasing in α, i.e. $F \le_{\mathrm{st}} G$, the following inequalities hold pointwise:

$$X_{(1)} \le G^{-1}F(X_{(1)}), \ \ldots, \ X_{(n)} \le G^{-1}F(X_{(n)})$$

and

$$\frac{\sum_{i=1}^n Y_i^2}{(\sum_{i=1}^n Y_i)^2} =_{\mathrm{st}} \sum_{i=1}^n \left[\frac{Z_{(i)}}{\sum_{j=1}^n Z_{(j)}}\right]^2,$$

where $Z_{(i)} = G^{-1}F(X_{(i)})$, $i = 1, \ldots, n$.

As $F >_{\mathrm{c}} G$, we know that $G^{-1}F(x)/x$ is decreasing in $x > 0$. Hence

$$\frac{X_{(j)}}{X_{(i)}} \ge \frac{Z_{(j)}}{Z_{(i)}}, \qquad 1 \le i < j \le n.$$

Using the last inequalities it is easy to see that

$$\frac{\sum_{i=1}^k X_{(i)}}{\sum_{i=1}^n X_{(i)}} \le \frac{\sum_{i=1}^k Z_{(i)}}{\sum_{i=1}^n Z_{(i)}}, \qquad k = 1, \ldots, n-1.$$

So we have proved that

$$\left(\frac{X_{(1)}}{\sum_{i=1}^n X_{(i)}}, \ldots, \frac{X_{(n)}}{\sum_{i=1}^n X_{(i)}}\right) \succ \left(\frac{Z_{(1)}}{\sum_{i=1}^n Z_{(i)}}, \ldots, \frac{Z_{(n)}}{\sum_{i=1}^n Z_{(i)}}\right).$$

Applying Theorem 1.1 with $\phi(x) = x^2$ to the above majorization ends the proof. ∎

Remark 2.3. Deriving exact distributions of the moment estimators $\hat{\alpha}$ and $\hat{\lambda}$ is tedious for $n > 2$. One may prove that for $n = 2$ the statistic $1/\hat{\alpha}$ has beta distribution $\mathcal{B}(1/2, \alpha)$ and so $\hat{\alpha}$ is stochastically increasing in α.

Now we examine the dispersive ordering of the estimator $1/\hat{\alpha}$.

Theorem 2.4. *The estimator* $1/\hat{\alpha}$ *is not dispersively monotone in* α.

Proof. Suppose for contradiction that $1/\hat{\alpha}$ is dispersively monotone. Then $\mathrm{Var}(1/\hat{\alpha})$ should be a monotone function in α. We show that is not true.

It is easy to calculate the moments of the sample mean:

$$E(\bar{X}^k) = \alpha\left(\alpha + \frac{1}{n}\right)\left(\alpha + \frac{2}{n}\right)\cdots\left(\alpha + \frac{k-1}{n}\right)\lambda^k, \qquad k = 1, 2, \ldots.$$

By the Basu theorem the statistics $S^4/\bar{X}^4$ and $\bar{X}^4$ are independent, as also are $S^2/\bar{X}^2$ and $\bar{X}^2$. Hence we obtain

$$\begin{aligned}\psi(\alpha) := \mathrm{Var}\left(\frac{S^2}{\bar{X}^2}\right) &= \frac{E(S^4)}{E(\bar{X}^4)} - \frac{E^2(S^2)}{E^2(\bar{X}^2)} \\ &= \frac{2\alpha(\alpha+1)}{(n-1)(\alpha+1/n)^2(\alpha+2/n)(\alpha+3/n)}.\end{aligned}$$

Since $\lim_{\alpha\to 0}\psi(\alpha) = \lim_{\alpha\to\infty}\psi(\alpha) = 0$, the function ψ is not monotone and therefore $1/\hat{\alpha}$ cannot be dispersively monotone. ■

3. Applications of stochastic comparison to confidence intervals. In this section we derive an asymptotic distribution of the estimator $\hat{\alpha}$ and give an application of Theorem 2.2 to comparison of approximate confidence intervals for α. This concept was considered by Balakrishnan et al. [2].

Using the delta method (see e.g. Serfling [8, p. 122]) we prove the following lemma.

Lemma 3.1. *The estimator $\hat{\alpha}$ is asymptotically normally distributed:*

$$\sqrt{n}(\hat{\alpha}-\alpha) \xrightarrow{\mathcal{L}} N(0, 2\alpha(1+\alpha)).$$

Proof. By the multivariate CLT it follows that

$$\sqrt{n}\begin{pmatrix}\bar{X}-\alpha\lambda \\ S^2-\alpha\lambda^2\end{pmatrix} \xrightarrow{\mathcal{L}} N(0,\Sigma),$$

where

$$\begin{aligned}\Sigma &= \begin{pmatrix}\mathrm{Var}(X_1) & E(X_1-E(X_1))^3 \\ E(X_1-E(X_1))^3 & E(X_1-E(X_1))^4-\mathrm{Var}^2(X_1)\end{pmatrix} \\ &= \begin{pmatrix}\alpha\lambda^2 & 2\alpha\lambda^3 \\ 2\alpha\lambda^3 & 2\alpha(3+\alpha)\lambda^4\end{pmatrix}.\end{aligned}$$

Thus by the delta method with $g(u,v) = u^2/v$, so that

$$\dot{g}(u,v) = (2u/v, -u^2/v^2) \quad \text{and} \quad \dot{g}(\alpha\lambda,\alpha\lambda^2) = (2/\lambda, -1/\lambda^2),$$

we have

$$\begin{aligned}\sqrt{n}(\hat{\alpha}-\alpha) &= \sqrt{n}(g(\bar{X},S^2) - g(\alpha\lambda,\alpha\lambda^2)) \\ &\xrightarrow{\mathcal{L}} N(0, \dot{g}(\alpha\lambda,\alpha\lambda^2)\Sigma\dot{g}(\alpha\lambda,\alpha\lambda^2)^T) = N(0,2\alpha(1+\alpha)). \ ■\end{aligned}$$

Similarly, applying the delta method to the function $g(v,u) = v/u$ we obtain the asymptotic distribution of the estimator $\hat{\lambda}$.

Lemma 3.2. *The estimator $\hat{\lambda}$ is asymptotically normally distributed:*

$$\sqrt{n}(\hat{\lambda}-\lambda) \xrightarrow{\mathcal{L}} N\left(0, \frac{3+2\alpha}{\alpha}\lambda^2\right).$$

Now we derive an approximate confidence interval for α. From asymptotic normality of $\hat{\alpha}$ it follows that for any $\alpha > 0$,

$$P\left(\left|\frac{\hat{\alpha}-\alpha}{\sqrt{2\alpha(1+\alpha)/n}}\right| \le z_{\gamma/2}\right) = 1-\gamma \quad \text{as } n \to \infty,$$

where $z_{\gamma/2}$ denotes the $(1-\gamma/2)$-quantile of the standard normal distribution. Solving the above inequality with respect to α we obtain an interval of the form

$$I = (L, U), \tag{3.1}$$

where

$$L = \frac{n\hat{\alpha} + z_{\gamma/2}^2 - z_{\gamma/2}\sqrt{2n\hat{\alpha}^2 + 2n\hat{\alpha} + z_{\gamma/2}^2}}{n - 2z_{\gamma/2}^2},$$

$$U = \frac{n\hat{\alpha} + z_{\gamma/2}^2 + z_{\gamma/2}\sqrt{2n\hat{\alpha}^2 + 2n\hat{\alpha} + z_{\gamma/2}^2}}{n - 2z_{\gamma/2}^2},$$

provided $n > 2z_{\gamma/2}^2$.

Theorem 2.2 immediately yields the following theorem.

THEOREM 3.3. *Let $I_1 = (L_1, U_1)$ and $I_2 = (L_2, U_2)$ be two random intervals of the form* (3.1) *for α_1 and α_2 respectively, where $\alpha_1 < \alpha_2$. Then $U_1 - L_1 \le_{\mathrm{st}} U_2 - L_2$, and I_1 is to the left of I_2 in probability, i.e. $L_1 \le_{\mathrm{st}} L_2$ and $U_1 \le_{\mathrm{st}} U_2$.*

References

[1] N. Balakrishnan and J. Mi, *Order-preserving property of maximum likelihood estimator*, J. Statist. Plann. Inference 98 (2001), 89–99.

[2] N. Balakrishnan, C. Brain and J. Mi, *Stochastic order and MLE of the mean of the exponential distribution*, Methodol. Comput. Appl. Probab. 4 (2002), 83–93.

[3] N. Balakrishnan and G. Iliopoulos, *Stochastic monotonicity of the MLE of exponential mean under different censoring schemes*, Ann. Inst. Statist. Math. 61 (2009), 753–772.

[4] G. Gan, *MLE are stochastically increasing if likelihoods are unimodal*, Statist. Probab. Lett. 40 (1998), 289–292.

[5] E. L. Lehmann and G. Casella, *Theory of Point Estimation*, 2nd ed., Springer, New York, 1998.

[6] T. Lindvall, *Lectures on the Coupling Method*, Wiley, New York, 1992.

[7] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, Academic Press, New York, 1979.

[8] R. J. Serfling, *Approximation Theorems of Mathematical Statistics*, Springer, New York, 1980.

[9] M. Shaked, *Dispersive ordering of distributions*, J. Appl. Probab. 19 (1982), 310–320.

[10] M. Shaked and J. G. Shanthikumar, *Stochastic Orders*, Springer, New York, 2007.

[11] W. van Zwet, *Convex Transforms of Random Variables*, Mat. Centrum, Amsterdam, 1964.

Piotr Nowak
Mathematical Institute
University of Wrocław
Pl. Grunwaldzki 2/4
50-384 Wrocław, Poland
E-mail: nowak@math.uni.wroc.pl

Received on 4.5.2011;
revised version on 25.7.2011 (2088)