



Alexey Kudryavtsev ^{1,2,*} and Oleg Shestakov ^{1,2,3,*}

- ¹ Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, Moscow 119991, Russia
- ² Moscow Center for Fundamental and Applied Mathematics, Moscow 119991, Russia
- ³ Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences, Moscow 119333, Russia
- * Correspondence: aakudryavtsev@cs.msu.ru (A.K.); oshestakov@cs.msu.ru (O.S.)

Abstract: In this paper, we study a new type of distribution that generalizes distributions from the gamma and beta classes that are widely used in applications. The estimators for the parameters of the digamma distribution obtained by the method of logarithmic cumulants are considered. Based on the previously proved asymptotic normality of the estimators for the characteristic index and the shape and scale parameters of the digamma distribution constructed from a fixed-size sample, we obtain a statement about the convergence of these estimators to the scale mixtures of the normal law in the case of a random sample size. Using this result, asymptotic confidence intervals for the estimated parameters are constructed. A number of examples of the limit laws for sample sizes with special forms of negative binomial distributions are given. The results of this paper can be widely used in the study of probabilistic models based on continuous distributions with an unbounded non-negative support.

Keywords: parameter estimation; digamma distribution; mixed distributions; generalized gamma distribution; generalized beta distribution; method of moments; cumulants; transfer theorem

MSC: 62F12

1. Introduction

Distributions belonging to beta and gamma classes play an essential role in probability theory and mathematical statistics. Such distributions have proven themselves as convenient and efficient tools in modeling a large number of real processes and phenomena [1–6]. Special cases of the generalized beta distribution of the second kind and the generalized gamma distribution can have the properties of infinite divisibility and stability, which makes it possible to use them as asymptotic approximations in various limit theorems. Ref. [7] proposed a new probability distribution closely related to both beta and gamma classes.

Definition 1. We say that the random variable ζ has the digamma distribution $DiG(r, v, p, q, \delta)$ with a characteristic index $r \in \mathbb{R}$ and the parameters of shape $v \neq 0$, concentration p, q > 0, and scale $\delta > 0$, if its Mellin transform is

$$\mathcal{M}_{\zeta}(z) = \frac{\delta^{z} \Gamma(p + z/\nu) \Gamma(q - rz/\nu)}{\Gamma(p) \Gamma(q)}, \quad p + \frac{\operatorname{Re}(z)}{\nu} > 0, \quad q - \frac{r\operatorname{Re}(z)}{\nu} > 0, \quad (1)$$

where $\operatorname{Re}(z)$ is the real part of a complex number *z*, and $\Gamma(z)$ is Euler's gamma function.

Particular types of digamma distribution include the generalized gamma distribution (also known as the Amoroso distribution with zero shift) [8], the generalized beta



Citation: Kudryavtsev, A.; Shestakov, O. Limit Distributions for the Estimates of the Digamma Distribution Parameters Constructed from a Random Size Sample. *Mathematics* 2023, *11*, 1778. https://doi.org/ 10.3390/math11081778

Academic Editor: Christophe Chesneau

Received: 9 March 2023 Revised: 31 March 2023 Accepted: 6 April 2023 Published: 7 April 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). distribution of the second kind (also known as the McDonald distribution) [9], and the gamma-exponential distribution [10].

The digamma distribution (1) can be represented as a scale mixture of two generalized gamma-distributed random variables, i.e., for $\zeta \sim DiG(r, \nu, p, q, \delta)$ and independent random variables $\lambda \sim \Gamma(p, 1)$ and $\mu \sim \Gamma(q, 1)$ with gamma distributions

$$\zeta \stackrel{d}{=} \delta \left(\frac{\lambda}{\mu^r}\right)^{1/\nu}.$$
(2)

This representation makes it possible [11] to use the digamma distribution for an adequate description of the Bayesian balance models proposed in [12].

Assuming that the process is modeled using the digamma distribution, the problem of statistical estimation of its unknown parameters inevitably arises [5,13,14]. As shown in Ref. [7], the density of the digamma distribution is expressed in terms of the special Fox's H-function. This significantly complicates the application of the maximum likelihood method. The form of the Mellin transform (1) of the digamma distribution also indicates the infeasibility of using the direct method of moments. Refs. [15–17] originally proposed a modified method for estimating the parameters of the gamma-exponential distribution based on logarithmic moments and cumulants. Due to the fact that the digamma distribution and the gamma-exponential distribution have the Mellin transform of the same type (up to the range of the parameter r), all previously obtained conclusions about the form of estimates by the method of logarithmic cumulants for the gamma-exponential distribution automatically remain valid for the digamma distribution, taking into account the formal expansion of the characteristic index range from a unit interval to the entire real line.

In today's rapidly changing world, it is quite problematic to use the traditional statistical approach based on the analysis of fixed-size samples. Thus, in the context of the global crisis caused by the COVID-19 epidemic, it is necessary to have a mechanism to respond to negative impacts using only the currently available data. Since the accumulation of a sufficient fixed amount of statistics can often take an indefinite time, it makes sense to strive for the possession of methods that allow one to draw adequate conclusions based on an a priori indefinite number of observations. This approach inevitably leads to the consideration of models with randomized sample sizes and is usually found not only in medicine but also in other fields in situations where the accumulation of statistical data continues not up to a certain amount but, rather, over a given period of time. For example, a similar situation can be observed in insurance when a different number of insurance events (insurance payments and/or insurance contracts) occur during different reporting periods of the same length (say, months), etc. Due to these circumstances, it becomes quite natural to study the asymptotic behavior of distributions of fairly general statistics based on the random size samples. When replacing a non-random sample size with a random variable, the asymptotic properties of statistics can radically change. This fact was apparently first noted by B.V. Gnedenko in 1989 [18,19]. It was shown that if the sample size is a geometrically distributed random variable, then instead of the normal law expected in accordance with the classical theory, a Student distribution with two degrees of freedom arises as an asymptotic distribution for the sample median, whose tails are so heavy that it does not have second-order moments. The "heaviness" of the tails of asymptotic distributions is of critical importance, in particular, in problems of testing hypotheses.

The distributions from the gamma class and their derivatives have become very popular for modeling random non-negative parameters, and, when modeling a random number of events and studying an a priori unknown number of observations, their discrete analogs are widely used, which are mixed Poisson distributions with corresponding continuous structural distributions. The discrete analog of the gamma distribution $\Gamma(p, \delta)$ is the negative binomial distribution, whose partial probabilities for n = 0, 1, ... are

$$\mathsf{P}(N=n) = \int_{0}^{\infty} \frac{\lambda^{n+p-1} e^{-(1+1/\delta)\lambda}}{\delta^{p} \Gamma(p) n!} d\lambda = \frac{\Gamma(n+p)}{\Gamma(n+1)\Gamma(p)} \left(\frac{\delta}{\delta+1}\right)^{n} \left(\frac{1}{\delta+1}\right)^{p}.$$
 (3)

A natural generalization of the distribution (3) is the mixed Poisson distribution whose structure is given by the generalized gamma distribution $GG(\nu, p, \delta)$ with the density

$$f(x) = \frac{|\nu|x^{\nu p - 1} e^{-(x/\delta)^{\nu}}}{\delta^{\nu p} \Gamma(p)}, \ \nu \neq 0, \ p > 0, \ x > 0.$$
(4)

Such distributions are called generalized negative binomial distributions and are widely used in insurance, financial mathematics, physics, and other fields [20–24].

The purpose of this article is to study the asymptotic behavior of digamma distribution parameter estimates under conditions of an a priori unknown sample size.

The article has the following structure. Section 2 describes a method for obtaining digamma distribution parameter estimates; auxiliary relations are given. Section 3 contains the main statement of this paper on the asymptotic behavior of the digamma distribution parameter estimates constructed from random size samples. Section 4 discusses special cases of limit distributions. This paper also contains a section with our conclusions.

2. Auxiliary Relations

This section describes a method based on logarithmic cumulants for obtaining estimators for the parameters r, v, and δ of the digamma distribution (1) with fixed concentration parameters p and q and a sample of a non-random size n. Estimating the parameters p and q is a separate problem due to the analytical complexity of inverting the polygamma function.

The results and relations of this section were published in Ref. [17] and are provided as auxiliary statements.

To obtain an explicit form of theoretical logarithmic cumulants, consider the polygamma functions

$$\psi(z) = \frac{d}{dz} \ln \Gamma(z), \ \psi^{(m)}(z) = \frac{d^{m+1}}{dz^{m+1}} \ln \Gamma(z), \ m = 1, 2, \dots$$

The theoretical cumulants of the random variable $\ln \zeta$ for $\zeta \sim DiG(r, v, p, q, \delta)$ have the form

$$\kappa_{1}(r,\nu,\delta) = \mathsf{E}\ln\zeta = \frac{\nu\ln\delta + \psi(p) - r\psi(q)}{\nu};$$

$$\kappa_{m}(r,\nu) = (-i)^{m} \frac{d^{m}}{dy^{m}} \ln\mathsf{E}\zeta^{iy}\Big|_{y=0} = \frac{\psi^{(m-1)}(p) + (-r)^{m}\psi^{(m-1)}(q)}{\nu^{m}}, \ m > 1.$$

The moments of the random variable $\ln \zeta$ can be represented as [25]

$$\mu_m(r,\nu,\delta) \equiv \mathsf{E}\ln^m \zeta = B_m(\kappa_1(r,\nu,\delta),\kappa_2(r,\nu),\ldots,\kappa_m(r,\nu)),\tag{5}$$

where B_m is a complete (exponential) Bell polynomial that can be recurrently defined as

$$B_{m+1}(x_1,\ldots,x_{m+1}) = \sum_{k=0}^m C_m^k B_{m-k}(x_1,\ldots,x_{m-k}) x_{k+1}, \ B_0 = 1.$$

An explicit form of the necessary relations connecting moments and cumulants can be found in Ref. [25].

In addition, we will need the following moment characteristics of the logarithm of a random variable with a digamma distribution:

$$\sigma_m^2(r,\nu,\delta) \equiv \mathsf{D} \ln^m \zeta = \mu_{2m}(r,\nu,\delta) - \mu_m^2(r,\nu,\delta);$$

$$\sigma_{ml}(r,\nu,\delta) \equiv \mathsf{cov}(\ln^m \zeta, \ln^l \zeta) = \mu_{m+l}(r,\nu,\delta) - \mu_m(r,\nu,\delta)\mu_l(r,\nu,\delta).$$
(6)

To define the sample logarithmic cumulants, we introduce a notation for the sample logarithmic moments of the random variable ζ :

$$L_m(\mathbb{X}_n) = \frac{1}{n} \sum_{i=1}^n \ln^m X_i,$$

where $X_n = (X_1, ..., X_n)$ is a sample from the distribution ζ of non-random size n. Let us denote $l = (l_1, l_2, l_3, l_4)$. Consider the functions

$$K_{1}(l) \equiv K_{1}(l_{1}) = (\psi(q))^{-1}l_{1};$$

$$K_{2}(l) \equiv K_{2}(l_{1}, l_{2}) = (\psi'(q))^{-1}(l_{2} - l_{1}^{2});$$

$$K_{3}(l) \equiv K_{3}(l_{1}, l_{2}, l_{3}) = (\psi''(q))^{-1}(l_{3} - 3l_{2}l_{1} + 2l_{1}^{3});$$

$$K_{4}(l) \equiv K_{4}(l_{1}, l_{2}, l_{3}, l_{4}) = (\psi'''(q))^{-1}(l_{4} - 4l_{3}l_{1} - 3l_{2}^{2} + 12l_{2}l_{1}^{2} - 6l_{1}^{4}).$$

Consider the statistics

$$K_2(\mathbb{X}_n) \equiv K_2(L_1(\mathbb{X}_n), L_2(\mathbb{X}_n));$$

$$K_3(\mathbb{X}_n) \equiv K_3(L_1(\mathbb{X}_n), L_2(\mathbb{X}_n), L_3(\mathbb{X}_n));$$

$$K_4(\mathbb{X}_n) \equiv K_4(L_1(\mathbb{X}_n), L_2(\mathbb{X}_n), L_3(\mathbb{X}_n), L_4(\mathbb{X}_n)).$$
(8)

Note that the statistics $\psi^{(m-1)}(q)K_m(\mathbb{X}_n)$ are the *m*-th sample logarithmic cumulants of the digamma distribution.

 $K_1(\mathbb{X}_n) \equiv K_1(L_1(\mathbb{X}_n));$

The method for estimating the unknown parameters considered in this paper is based on solving the system for logarithmic cumulants:

$$\kappa_m(r,\nu,\delta) = \psi^{(m-1)}(q) K_m(\mathbb{X}_n), \ m = 1, 2, 3, 4.$$

To describe the solution of this system, we introduce a number of functions of sample logarithmic cumulants with the arguments $k = (k_1, k_2, k_3, k_4)$:

$$\phi_m = \frac{\psi^{(m)}(p)}{\psi^{(m)}(q)}; \ \tau(k) \equiv \tau(k_2, k_4) = \phi_1^2 k_4 + \phi_3 \left(k_4 - k_2^2\right); \tag{9}$$

$$R_{\pm}(k) \equiv R_{\pm}(k_{2},k_{4}) = \sqrt{\frac{\phi_{1}k_{4} \pm k_{2}\sqrt{\tau(k)}}{k_{2}^{2} - k_{4}}};$$
$$V_{\pm}(k) \equiv V_{\pm}(k_{2},k_{4}) = \sqrt{\frac{\phi_{1}k_{2} \pm \sqrt{\tau(k)}}{k_{2}^{2} - k_{4}}};$$
$$D_{\pm}(k) \equiv D_{\pm}(k_{1},k_{2},k_{4}) = \exp\left\{\psi(q)k_{1} + \frac{\psi(q)R_{\pm}(k) - \psi(p)}{V_{\pm}(k)}\right\}.$$
(10)

(7)

In what follows, we will need the derivatives of functions (10), expressed in terms of the functions ϕ_m and τ , defined in (9). Note that

$$\begin{split} R_{k_{2},\pm}(k) &\equiv \frac{\partial R_{\pm}}{\partial k_{2}}(k_{2},k_{4}) = \mp \frac{k_{4} \Big(\phi_{1}^{2}k_{2}^{2} + \tau(k) \pm 2\phi_{1}k_{2}\sqrt{\tau(k)} \Big)}{2 \big(k_{2}^{2} - k_{4} \big)^{3/2} \sqrt{\tau(k)} \sqrt{\phi_{1}k_{4} \pm k_{2}\sqrt{\tau(k)}} \big)}; \\ R_{k_{4},\pm}(k) &\equiv \frac{\partial R_{\pm}}{\partial k_{4}}(k_{2},k_{4}) = \pm \frac{k_{2} \Big(\phi_{1}^{2}k_{2}^{2} + \tau(k) \pm 2\phi_{1}k_{2}\sqrt{\tau(k)} \Big)}{4 (k_{2}^{2} - k_{4})^{3/2} \sqrt{\tau(k)} \sqrt{\phi_{1}k_{4} \pm k_{2}\sqrt{\tau(k)}}}; \\ V_{k_{2},\pm}(k) &\equiv \frac{\partial V_{\pm}}{\partial k_{2}}(k_{2},k_{4}) = \mp \frac{k_{2} \big(\phi_{1}^{2}k_{4} + \tau(k) \big) \pm \phi_{1}(k_{2}^{2} + k_{4})\sqrt{\tau(k)}}{2 (k_{2}^{2} - k_{4})^{3/2} \sqrt{\tau(k)} \sqrt{\phi_{1}k_{2} \pm \sqrt{\tau(k)}}}; \\ V_{k_{4},\pm}(k) &\equiv \frac{\partial V_{\pm}}{\partial k_{4}}(k_{2},k_{4}) = \pm \frac{\phi_{1}^{2}k_{2}^{2} + \tau(k) \pm 2\phi_{1}k_{2}\sqrt{\tau(k)}}{4 (k_{2}^{2} - k_{4})^{3/2} \sqrt{\tau(k)} \sqrt{\phi_{1}k_{2} \pm \sqrt{\tau(k)}}}; \\ D_{k_{1},\pm}(k) &\equiv \frac{\partial D_{\pm}}{\partial k_{4}}(k_{1},k_{2},k_{4}) = \psi(q) \exp \Big\{ \psi(q)k_{1} + \frac{\psi(q)R_{\pm}(k) - \psi(p)}{V_{\pm}(k)} \Big\}; \\ D_{k_{2},\pm}(k) &\equiv \frac{\partial D_{\pm}}{\partial k_{2}}(k_{1},k_{2},k_{4}) = \exp \Big\{ \psi(q)k_{1} + \frac{\psi(q)R_{\pm}(k) - \psi(p)}{V_{\pm}(k)} \Big\} \times \\ \times \frac{\psi(p)V_{k_{2},\pm}(k) + \psi(q)R_{k_{2},\pm}(k)V_{\pm}(k) - \psi(q)R_{\pm}(k)V_{k_{2},\pm}(k)}{V_{\pm}^{2}(k)}; \\ D_{k_{4},\pm}(k) &\equiv \frac{\partial D_{\pm}}{\partial k_{4}}(k_{1},k_{2},k_{4}) = \exp \Big\{ \psi(q)k_{1} + \frac{\psi(q)R_{\pm}(k) - \psi(p)}{V_{\pm}(k)} \Big\} \times \\ \times \frac{\psi(p)V_{k_{4},\pm}(k) + \psi(q)R_{k_{4},\pm}(k)V_{\pm}(k) - \psi(q)R_{\pm}(k)V_{k_{4},\pm}(k)}{V_{\pm}(k)} - \psi(q)R_{\pm}(k)V_{k_{4},\pm}(k)}. \end{split}$$

Using the formula for the derivative of a composite function, we obtain

$$\begin{split} \frac{\partial R_{\pm}}{\partial l_{1}}(l) &= -\frac{2l_{1}}{\psi'(q)} R_{k_{2},\pm}(K_{2}(l),K_{4}(l)) - \frac{4l_{3} - 24l_{2}l_{1} + 24l_{1}^{3}}{\psi'''(q)} R_{k_{4},\pm}(K_{2}(l),K_{4}(l)); \\ &\frac{\partial R_{\pm}}{\partial l_{2}}(l) = \frac{1}{\psi'(q)} R_{k_{2},\pm}(K_{2}(l),K_{4}(l)) - \frac{6l_{2} - 12l_{1}^{2}}{\psi'''(q)} R_{k_{4},\pm}(K_{2}(l),K_{4}(l)); \\ &\frac{\partial R_{\pm}}{\partial l_{3}}(l) = -\frac{4l_{1}}{\psi'''(q)} R_{k_{4},\pm}(K_{2}(l),K_{4}(l)); \\ &\frac{\partial R_{\pm}}{\partial l_{4}}(l) = \frac{1}{\psi'''(q)} R_{k_{4},\pm}(K_{2}(l),K_{4}(l)); \end{split}$$

To formulate the statement about the asymptotic normality of estimators for the parameters r, v, and δ with fixed concentration parameters p and q for a fixed sample size n, we introduce some notations. Let

$$\Sigma = \begin{pmatrix} \sigma_{1}^{2}(r,\nu,\delta) & \sigma_{12}(r,\nu,\delta) & \sigma_{13}(r,\nu,\delta) & \sigma_{14}(r,\nu,\delta) \\ \sigma_{12}(r,\nu,\delta) & \sigma_{2}^{2}(r,\nu,\delta) & \sigma_{23}(r,\nu,\delta) & \sigma_{24}(r,\nu,\delta) \\ \sigma_{13}(r,\nu,\delta) & \sigma_{23}(r,\nu,\delta) & \sigma_{3}^{2}(r,\nu,\delta) & \sigma_{34}(r,\nu,\delta) \\ \sigma_{14}(r,\nu,\delta) & \sigma_{24}(r,\nu,\delta) & \sigma_{34}(r,\nu,\delta) & \sigma_{4}^{2}(r,\nu,\delta) \end{pmatrix};$$

$$d_{R_{\pm}} = \left(\frac{\partial R_{\pm}}{\partial l_{1}}(l)\Big|_{l=\mu}, \frac{\partial R_{\pm}}{\partial l_{2}}(l)\Big|_{l=\mu}, \frac{\partial R_{\pm}}{\partial l_{3}}(l)\Big|_{l=\mu}, \frac{\partial R_{\pm}}{\partial l_{4}}(l)\Big|_{l=\mu}\right);$$

$$d_{V_{\pm}} = \left(\frac{\partial V_{\pm}}{\partial l_{1}}(l)\Big|_{l=\mu}, \frac{\partial V_{\pm}}{\partial l_{2}}(l)\Big|_{l=\mu}, \frac{\partial V_{\pm}}{\partial l_{3}}(l)\Big|_{l=\mu}, \frac{\partial V_{\pm}}{\partial l_{4}}(l)\Big|_{l=\mu}\right);$$

$$d_{D_{\pm}} = \left(\frac{\partial D_{\pm}}{\partial l_{1}}(l)\Big|_{l=\mu}, \frac{\partial D_{\pm}}{\partial l_{2}}(l)\Big|_{l=\mu}, \frac{\partial D_{\pm}}{\partial l_{3}}(l)\Big|_{l=\mu}, \frac{\partial D_{\pm}}{\partial l_{4}}(l)\Big|_{l=\mu}\right), \quad (12)$$

where the variances $\sigma_m^2(r, \nu, \delta)$ and the covariances $\sigma_{ml}(r, \nu, \delta)$ are defined in the relations (6), the partial derivatives $\partial R_{\pm}/\partial l_k(l)$, $\partial V_{\pm}/\partial l_k(l)$, and $\partial D_{\pm}/\partial l_k(l)$ are defined in (11), and $\mu = (\mu_1, \mu_2, \mu_3)$ is the vector of moments (5).

Previously, in Ref. [17], the following result was obtained for the gammaexponential distribution.

Theorem 1. Let $0 \le r < 1$ and v > 0. Assume that the concentration parameters p and q of the digamma distribution $DiG(r, v, p, q, \delta)$ are fixed. Then, for $r > \sqrt{\phi_3/\phi_1}$, the estimators $\hat{r}(\mathbb{X}_n) = R_+(K_2(\mathbb{X}_n), K_4(\mathbb{X}_n))$ for the unknown characteristic index r, $\hat{v}(\mathbb{X}_n) = V_+(K_2(\mathbb{X}_n), K_4(\mathbb{X}_n))$ for the unknown characteristic index r, $\hat{v}(\mathbb{X}_n) = V_+(K_2(\mathbb{X}_n), K_4(\mathbb{X}_n))$ for the unknown scale parameter v and $\hat{\delta}(\mathbb{X}_n) = D_+(K_1(\mathbb{X}_n), K_2(\mathbb{X}_n), K_4(\mathbb{X}_n))$ for the unknown scale parameter δ have the property of asymptotic normality when $n \to \infty$:

$$\sqrt{n} \frac{\hat{r}(\mathbb{X}_n) - r}{\sqrt{d_{R_+} \Sigma d_{R_+}^T}} \Longrightarrow N(0, 1), \quad \sqrt{n} \frac{\hat{v}(\mathbb{X}_n) - \nu}{\sqrt{d_{V_+} \Sigma d_{V_+}^T}} \Longrightarrow N(0, 1); \quad \sqrt{n} \frac{\hat{\delta}(\mathbb{X}_n) - \delta}{\sqrt{d_{D_+} \Sigma d_{D_+}^T}} \Longrightarrow N(0, 1).$$
(13)

Remark 1. *In addition to the property of asymptotic normality, the estimators listed in Theorem 1 have the property of strong consistency* [16].

Remark 2. In Theorem 1, if $0 \le r < \sqrt{\phi_3/\phi_1}$, then one should choose the statistics $\hat{r}(\mathbb{X}_n) = R_-(K_2(\mathbb{X}_n), K_4(\mathbb{X}_n)), \hat{v}(\mathbb{X}_n) = V_-(K_2(\mathbb{X}_n), K_4(\mathbb{X}_n)), \text{ and } \hat{\delta}(\mathbb{X}_n) = D_-(K_1(\mathbb{X}_n), K_2(\mathbb{X}_n), K_4(\mathbb{X}_n))$ with a corresponding modification of the normalizing constants in (13) [17].

Remark 3. In Theorem 1, if v < 0, then one should choose as an estimator for the unknown parameter v the statistics $\hat{v}(\mathbb{X}_n) = -V_+(K_2(\mathbb{X}_n), K_4(\mathbb{X}_n))$ if $r > \sqrt{\phi_3/\phi_1}$, and $\hat{v}(\mathbb{X}_n) = -V_-(K_2(\mathbb{X}_n), K_4(\mathbb{X}_n))$ if $0 \le r < \sqrt{\phi_3/\phi_1}$.

Remark 4. Since the gamma-exponential distribution and the digamma distribution have the Mellin transform of the same type (1), the results of Theorem 1 and Remark 1 remain valid for all $r \ge 0$. In the case when r < 0, one should consider as an estimator for the parameter r the statistics $\hat{r}(\mathbb{X}_n) = -R_+(K_2(\mathbb{X}_n), K_4(\mathbb{X}_n))$ for $r < -\sqrt{\phi_3/\phi_1}$ and $\hat{r}(\mathbb{X}_n) = -R_-(K_2(\mathbb{X}_n), K_4(\mathbb{X}_n))$ for $-\sqrt{\phi_3/\phi_1} < r \le 0$.

Remark 5. When processing real data, one should first choose one of the statistics $\pm R_{\pm}(K_2(\mathbb{X}_n), K_4(\mathbb{X}_n))$, and $\pm V_{\pm}(K_2(\mathbb{X}_n), K_4(\mathbb{X}_n))$ as the estimators $\hat{r}(\mathbb{X}_n)$ and $\hat{v}(\mathbb{X}_n)$, using the algorithm for eliminating unnecessary solutions described in Ref. [17]. The estimator for the unknown parameter δ is always defined by the formula

$$\hat{\delta}(\mathbb{X}_n) = \exp\left\{\psi(q)K_1(\mathbb{X}_n) + \frac{\psi(q)\hat{r}(\mathbb{X}_n) - \psi(p)}{\hat{v}(\mathbb{X}_n)}
ight\}.$$

3. Main Result

Everywhere below we will assume that the sample size is random. To obtain asymptotic approximations, it is reasonable to consider a situation in which the random size of the sample increases in some sense. We will consider a sequence N_n such that $N_n \rightarrow \infty$ in probability as $n \rightarrow \infty$.

Let the non-random size sample $\mathbb{X}_n = (X_1, ..., X_n)$ and the random size sample $\mathbb{X}_{N_n} = (X_1, ..., X_{N_n})$ be from the digamma distribution $DiG(r, \nu, p, q, \delta)$ with the known concentration parameters p and q.

Using the Functions (7) and (8), we construct the statistics

$$K_2(\mathbb{X}_{N_n}) \equiv K_2(L_1(\mathbb{X}_{N_n}), L_2(\mathbb{X}_{N_n}));$$
$$(\mathbb{X}_{N_n}) \equiv K_4(L_1(\mathbb{X}_{N_n}), L_2(\mathbb{X}_{N_n}), L_3(\mathbb{X}_{N_n}), L_4(\mathbb{X}_{N_n})),$$

 $K_4(\mathbb{X}_{N_n}) \equiv K_4(L_1(\mathbb{X}_N))$ based on sample logarithmic moments

$$L_m(\mathbb{X}_{N_n}) = \frac{1}{N_n} \sum_{i=1}^{N_n} \ln^m X_i$$

Let N_n be a sequence of natural-valued random variables independent of $X_1, X_2, ...,$ for each n, and let N_n tend toward infinity in probability as $n \to \infty$.

The following statement holds.

Theorem 2. Let $r > \sqrt{\phi_3/\phi_1}$ and $\nu > 0$. Suppose that the concentration parameters p and q of the digamma distribution $DiG(r, \nu, p, q, \delta)$ are fixed. Assume that there exists a numerical sequence $\{b_n > 0\}$ and a random variable U such that

$$\frac{N_n}{b_n} \Longrightarrow U \tag{14}$$

when $n \to \infty$. Then, the estimators $\hat{r}(\mathbb{X}_{N_n}) = R_+(K_2(\mathbb{X}_{N_n}), K_4(\mathbb{X}_{N_n}))$ for the unknown characteristic index $r, \hat{v}(\mathbb{X}_{N_n}) = V_+(K_2(\mathbb{X}_{N_n}), K_4(\mathbb{X}_{N_n}))$ for the unknown shape parameter v, and $\hat{\delta}(\mathbb{X}_{N_n}) = D_+(K_1(\mathbb{X}_{N_n}), K_2(\mathbb{X}_{N_n}), K_4(\mathbb{X}_{N_n}))$ for the unknown scale parameter δ converge in distribution when $n \to \infty$:

$$\sqrt{b_n} \frac{\hat{r}(\mathbb{X}_{N_n}) - r}{\sqrt{d_{R_+} \Sigma d_{R_+}^T}} \Longrightarrow \frac{Y}{\sqrt{U}}, \quad \sqrt{b_n} \frac{\hat{\nu}(\mathbb{X}_{N_n}) - \nu}{\sqrt{d_{V_+} \Sigma d_{V_+}^T}} \Longrightarrow \frac{Y}{\sqrt{U}}, \quad \sqrt{b_n} \frac{\hat{\delta}(\mathbb{X}_{N_n}) - \delta}{\sqrt{d_{D_+} \Sigma d_{D_+}^T}} \Longrightarrow \frac{Y}{\sqrt{U}}, \tag{15}$$

where Y has a standard normal distribution, and U can be considered independent of Y.

Proof of Theorem 2. We consider the statement of the theorem for estimating the characteristic index *r*. The argument is based on the method proposed in Ref. [26]. Denote

 $a_n = rac{\sqrt{d_{R_+}\Sigma d_{R_+}^T}}{\sqrt{b_n}}, \ c_n = rac{\sqrt{d_{R_+}\Sigma d_{R_+}^T}}{\sqrt{n}}.$

Let $h_n(t)$ be the characteristic function of a random variable

$$Y_n \equiv \sqrt{n} \frac{\hat{r}(\mathbb{X}_n) - r}{\sqrt{d_{R_+} \Sigma d_{R_+}^T}} \equiv \frac{\hat{r}(\mathbb{X}_n) - r}{c_n};$$

and $f_n(t)$ be the characteristic function of

$$Z_n \equiv \sqrt{b_n} \frac{\hat{r}(\mathbb{X}_{N_n}) - r}{\sqrt{d_{R_+} \Sigma d_{R_+}^T}} \equiv \frac{\hat{r}(\mathbb{X}_{N_n}) - r}{a_n}.$$

Theorem 1 implies that when $n \to \infty$

$$Y_n \Longrightarrow Y \sim N(0,1).$$

Denote by h(t) the characteristic function of a standard normal random variable *Y*. Define the random variables

$$U_n\equiv\frac{c_{N_n}}{a_n}$$

Let

$$g_n(t) = \mathsf{E}h(tU_n).$$

Let us show that for any $t \in \mathbb{R}$

$$\lim_{n\to\infty}|f_n(t)-g_n(t)|=0.$$

For some positive number γ and positive integer *m*, we define

$$K_{1,n} \equiv K_{1,n}(\gamma) = \{ m \mid c_m \le \gamma a_n \}, \ K_{2,n} \equiv K_{2,n}(\gamma) = \{ m \mid c_m > \gamma a_n \}$$

For t = 0, the statement is obvious. Fix an arbitrary $t \neq 0$. Then,

$$|f_n(t) - g_n(t)| = |\mathsf{E}\exp\{itZ_n\} - \mathsf{E}h(tU_n)| =$$

$$= \left| \sum_{m=1}^{\infty} \mathsf{P}(N_n = m) \left[\mathsf{E} \exp\left\{ it \frac{\hat{r}(\mathbb{X}_m) - r}{a_n} \right\} - h\left(t \frac{c_m}{a_n}\right) \right] \right| =$$
$$= \left| \sum_{m=1}^{\infty} \mathsf{P}(N_n = m) \left[\mathsf{E} \exp\left\{ it \frac{c_m}{a_n} \cdot \frac{\hat{r}(\mathbb{X}_m) - r}{c_m} \right\} - h\left(t \frac{c_m}{a_n}\right) \right] \right| =$$
$$= \left| \sum_{m=1}^{\infty} \mathsf{P}(N_n = m) \left[h_m\left(t \frac{c_m}{a_n}\right) - h\left(t \frac{c_m}{a_n}\right) \right] \right| \le \sum_{m \in K_{1,n}} \mathsf{P}(N_n = m) \left| h_m\left(t \frac{c_m}{a_n}\right) - h\left(t \frac{c_m}{a_n}\right) \right| +$$

$$+\sum_{m\in K_{2,n}}\mathsf{P}(N_n=m)\left|h_m\left(t\frac{c_m}{a_n}\right)-h\left(t\frac{c_m}{a_n}\right)\right|=I_1+I_2.$$

Fix an arbitrary $\epsilon > 0$. Consider I_2 .

$$I_2 = \sum_{m \in K_{2,n}} \mathsf{P}(N_n = m) \left| h_m \left(t \frac{c_m}{a_n} \right) - h \left(t \frac{c_m}{a_n} \right) \right| \le 2 \sum_{m \in K_{2,n}} \mathsf{P}(N_n = m) = 2\mathsf{P}(U_n > \gamma) < \epsilon/2$$

for all $\gamma > \gamma_2(\epsilon)$, due to the convergence $U_n \Longrightarrow 1/\sqrt{U}$. Now, consider I_1 . Let $\gamma > \gamma_2(\epsilon)$. Since $|tc_m/a_n| \le |t|\gamma$,

$$I_{1} = \sum_{m \in K_{1,n}} \mathsf{P}(N_{n} = m) \left| h_{m} \left(t \frac{c_{m}}{a_{n}} \right) - h \left(t \frac{c_{m}}{a_{n}} \right) \right| \leq$$

$$\leq \sum_{m=1}^{\infty} \mathsf{P}(N_{n} = m) \sup_{|\tau| \leq \gamma|t|} |h_{m}(\tau) - h(\tau)| = \mathsf{E} \sup_{|\tau| \leq \gamma|t|} |h_{N_{n}}(\tau) - h(\tau)|.$$

Due to the uniform convergence of the sequence of characteristic functions $h_n(t)$ to h(t) on any finite interval and the convergence $N_n \longrightarrow \infty$ in probability,

$$\mathsf{E}\sup_{|\tau|\leq \gamma|t|}|h_{N_n}(\tau)-h(\tau)|<\epsilon/2$$

starting from some *n*.

Since $I_1 + I_2 < \epsilon$ starting from some *n*, we conclude that for any *t*

$$\lim_{n\to\infty}|f_n(t)-g_n(t)|=0.$$

Note that the function

$$\phi_t(x) = h(tx)$$

is bounded and continuous. Therefore, the weak convergence condition $U_n \implies 1/\sqrt{U}$ implies

$$\lim_{n\to\infty} \mathsf{E}\phi_t(U_n) = \mathsf{E}\phi_t(1/\sqrt{U}) = \mathsf{E}h(t/\sqrt{U}).$$

By the Fubini theorem, the right-hand side of the last equality is the characteristic function of the random variable Y/\sqrt{U} for a copy of the standard normal random variable Y independent of U.

Since

$$|f_n(t) - \mathsf{E}h(t/\sqrt{U})| \le |f_n(t) - g_n(t)| + |g_n(t) - \mathsf{E}h(t/\sqrt{U})| < 2\epsilon$$

for all $\epsilon > 0$ starting from some *n*,

$$\lim_{n\to\infty}f_n(t)=\mathsf{E}h(t/\sqrt{U}),$$

which completes the proof of the theorem for the estimator of the characteristic index r.

The statements of the theorem for the estimators of the form parameter v and the scale parameter δ are proved in a completely similar way. The theorem is proved. \Box

Remark 6. Similarly to Remarks 2–5, the statement of Theorem 2 remains valid in the cases $r < -\sqrt{\phi_3/\phi_1}, -\sqrt{\phi_3/\phi_1} < r \le 0, 0 \le r < \sqrt{\phi_3/\phi_1}$ and $\nu < 0$ for the estimators $\hat{r}(\mathbb{X}_{N_n}) = \pm R_{\pm}(K_2(\mathbb{X}_{N_n}), K_4(\mathbb{X}_{N_n}))$ and $\hat{\nu}(\mathbb{X}_{N_n}) = \pm V_{\pm}(K_2(\mathbb{X}_{N_n}), K_4(\mathbb{X}_{N_n}))$ with the corresponding modification of the normalizing constants in (15). The choice of the "correct" signs of the estimators is carried out using the algorithm for eliminating unnecessary solutions from Ref. [17]. The estimator for the unknown parameter δ is always defined by the formula

$$\hat{\delta}(\mathbb{X}_{N_n}) = \exp\left\{\psi(q)K_1(\mathbb{X}_{N_n}) + \frac{\psi(q)\hat{r}(\mathbb{X}_{N_n}) - \psi(p)}{\hat{\nu}(\mathbb{X}_{N_n})}\right\}$$

Let us introduce additional notation

$$s_{mm}(\mathbb{X}_{N_n}) \equiv \sigma_m^2(\hat{r}(\mathbb{X}_{N_n}), \hat{v}(\mathbb{X}_{N_n}), \hat{\delta}(\mathbb{X}_{N_n}));$$

$$s_{ml}(\mathbb{X}_{N_n}) = s_{lm}(\mathbb{X}_{N_n}) \equiv \sigma_{ml}(\hat{r}(\mathbb{X}_{N_n}), \hat{v}(\mathbb{X}_{N_n}), \hat{\delta}(\mathbb{X}_{N_n}));$$

$$d_r^{[m]}(\mathbb{X}_{N_n}) \equiv \frac{\partial \hat{r}(\mathbb{X}_{N_n})}{\partial l_m}; \ d_v^{[m]}(\mathbb{X}_{N_n}) \equiv \frac{\partial \hat{v}(\mathbb{X}_{N_n})}{\partial l_m}; \ d_{\delta}^{[m]}(\mathbb{X}_{N_n}) \equiv \frac{\partial \hat{\delta}(\mathbb{X}_{N_n})}{\partial l_m}, \tag{16}$$

where $\sigma_m^2(r, \nu, \delta)$ and $\sigma_{ml}(r, \nu, \delta)$ are defined in (6) and $\hat{r}(\mathbb{X}_{N_n})$, $\hat{\nu}(\mathbb{X}_{N_n})$, and $\hat{\delta}(\mathbb{X}_{N_n})$ satisfy the conditions of Theorem 2.

Theorem 2 implies a statement about the form of the asymptotic confidence intervals for unknown parameters of the digamma distribution. Denote by u_{γ} the $(1 + \gamma)/2$ -quantile of the limiting random variable γ/\sqrt{U} .

Corollary 1. Suppose that the conditions of Theorem 2 are met; then the asymptotic confidence intervals with a confidence level γ based on the estimators $\hat{r}(\mathbb{X}_{N_n})$, $\hat{v}(\mathbb{X}_{N_n})$, and $\hat{\delta}(\mathbb{X}_{N_n})$ for the unknown parameters r, v, and δ have the form

$$(A_r(\mathbb{X}_{N_n}), B_r(\mathbb{X}_{N_n})) = \left(\hat{r}(\mathbb{X}_{N_n}) - \frac{u_{\gamma}}{\sqrt{n}}C_r(\mathbb{X}_{N_n}), \hat{r}(\mathbb{X}_{N_n}) + \frac{u_{\gamma}}{\sqrt{n}}C_r(\mathbb{X}_{N_n})\right);$$

$$(A_{\nu}(\mathbb{X}_{N_n}), B_{\nu}(\mathbb{X}_{N_n})) = \left(\hat{\nu}(\mathbb{X}_{N_n}) - \frac{u_{\gamma}}{\sqrt{n}}C_{\nu}(\mathbb{X}_{N_n}), \hat{\nu}(\mathbb{X}_{N_n}) + \frac{u_{\gamma}}{\sqrt{n}}C_{\nu}(\mathbb{X}_{N_n})\right);$$

$$(A_{\delta}(\mathbb{X}_{N_n}), B_{\delta}(\mathbb{X}_{N_n})) = \left(\hat{\delta}(\mathbb{X}_{N_n}) - \frac{u_{\gamma}}{\sqrt{n}}C_{\delta}(\mathbb{X}_{N_n}), \hat{\delta}(\mathbb{X}_{N_n}) + \frac{u_{\gamma}}{\sqrt{n}}C_{\delta}(\mathbb{X}_{N_n})\right),$$

where

$$C_{r}(\mathbb{X}_{N_{n}}) = \sqrt{\sum_{m=1}^{4} \sum_{l=1}^{4} d_{r}^{[m]}(\mathbb{X}_{N_{n}}) s_{ml}(\mathbb{X}_{N_{n}}) d_{r}^{[l]}(\mathbb{X}_{N_{n}})};$$

$$C_{\nu}(\mathbb{X}_{N_{n}}) = \sqrt{\sum_{m=1}^{4} \sum_{l=1}^{4} d_{\nu}^{[m]}(\mathbb{X}_{N_{n}}) s_{ml}(\mathbb{X}_{N_{n}}) d_{\nu}^{[l]}(\mathbb{X}_{N_{n}})};$$

$$C_{\delta}(\mathbb{X}_{N_{n}}) = \sqrt{\sum_{m=1}^{4} \sum_{l=1}^{4} d_{\delta}^{[m]}(\mathbb{X}_{N_{n}}) s_{ml}(\mathbb{X}_{N_{n}}) d_{\delta}^{[l]}(\mathbb{X}_{N_{n}})};$$

and $s_{ml}(\mathbb{X}_{N_n})$, $d_r^{[m]}(\mathbb{X}_{N_n})$, $d_{\nu}^{[m]}(\mathbb{X}_{N_n})$, $d_{\delta}^{[m]}(\mathbb{X}_{N_n})$ are defined in (16).

The proof is completely analogous to the proof of Corollary 2 from Ref. [17].

4. Examples of Limit Distributions

Let us give a number of examples of possible limit distributions in Theorem 2.

As noted in Section 1, special forms of the negative binomial distribution have gained great popularity in modeling a random number of events. Since the negative binomial distribution is concentrated on non-negative integers, it cannot be directly used as a random sample size. We will consider such distributions with a shift by one, which will ensure the natural value of the sample size. According to the generalized Slutsky theorem, all conclusions concerning the asymptotic behavior of "shifted" distributions are equivalent to the statements about the asymptotics for sequences of random variables that have a classical negative binomial distribution or a mixed Poisson distribution with a structural gamma distribution.

Note that the gamma distribution belongs to the class of distributions with a scale parameter. It means that if $\Lambda \sim \Gamma(s, \theta)$, then $\hat{\Lambda} \stackrel{d}{=} \Lambda/\theta \sim \Gamma(s, 1)$. The following statements are based on the fact that for a standard Poisson process $N_1(t)$ independent of the random variable Λ ,

$$rac{N_1(\Lambda n)}{ heta n} \Longrightarrow \hat{\Lambda}, \ n o \infty.$$

Note also that if a random variable ξ has a generalized gamma distribution $GG(v, s, \theta)$ with the density (4), then

$$rac{1}{\sqrt{\xi}} \sim GGigg(-2v,s,rac{1}{\sqrt{ heta}}igg).$$

Denote by $\Pi(\Lambda)$ the mixed Poisson distribution whose structure is given by the random variable Λ . To specify particular cases of Theorem 2, we consider the distribution $D(\theta)$ degenerate at the point θ , the gamma distribution $\Gamma(s,\theta)$, the exponential distribution $E(\theta) \equiv \Gamma(1,\theta)$, and the scaled χ^2 -distribution $\chi^2(k,\theta) \equiv \Gamma(k/2,\theta)$, $k \in \mathbb{N}$ as the structural one. To determine the corresponding mixed Poisson distributions, consider the negative binomial distribution $NB(p, 1/(1 + \theta))$ whose partial probabilities are given by (3), and the geometric distribution $G(1/(1 + \theta)) \equiv NB(1, 1/(1 + \theta))$. To determine the limit distributions, consider the type VII Pearson distribution $P7(m, \alpha)$, $m \geq 1/2$, $\alpha > 0$, with the density

$$f_{P7}(x) = \frac{\alpha^{2m-1}}{B(m-1/2,1/2)} (\alpha^2 + x^2)^{-m};$$

the Student distribution $St(n) \equiv P7((n+1)/2, \sqrt{n})$; and the Cauchy distribution $K(\alpha) \equiv P7(1, \alpha)$.

For $b_n = \theta n$, let us list several examples of limit distributions of the random variable Y/\sqrt{U} from (15).

Let $N_n - 1 \stackrel{d}{=} N_1(\Lambda n) \sim \Pi(\Lambda n)$. Then, the limit random variable *U* in (14) coincides in distribution with $\hat{\Lambda}$, and the distributions of the random variable Y/\sqrt{U} have the form shown in Table 1.

Table 1. Special cases of the limit distribution.

$\Lambda \sim$	$\Pi(\Lambda n)$	$Y/\sqrt{u} \sim$			
D(heta)	$\Pi(\theta n)$	N(0,1)			
E(heta)	$G\left(\frac{1}{1+\theta n}\right)$	St(2)			
$\chi^2(1,\theta)$	$NB\left(\frac{1}{2},\frac{1}{1+\theta n}\right)$	$K\left(\sqrt{2}\right)$			
$\Gamma(s, heta)$	$NB\left(s, \frac{1}{1+\theta n}\right)$	$P7\left(s+rac{1}{2},\sqrt{2} ight)$			

Let us give some numerical examples of calculating the estimates of the parameters r, v, and δ of the digamma distribution $DiG(r, v, p, q, \delta)$ from the model samples. The concentration parameters p and q are fixed. The data given in Table 2 are obtained using the algorithm described in Ref. [17].

The pseudorandom sample size N_n for each n is generated for the distributions $\Pi(\Lambda n)$ from Table 1. The simulation of pseudorandom samples from the digamma distribution is based on Relation (2).

Table 2 lists the values of the estimates $\hat{r}(X_{N_n})$, $\hat{v}(X_{N_n})$, and $\hat{\delta}(X_{N_n})$ of the parameters r, v, and δ , obtained by simulating a sample from the digamma distribution DiG(0.5; 2.5; 2.4; 1.9; 1.0), and the corresponding boundaries of the confidence intervals. The distributions of the random sample size are taken from Table 1 with $\theta = 1$ and s = 2.

$N_n - 1 \sim$	$\hat{r}(\mathbb{X}_{N_n})$	$A_r(\mathbb{X}_{N_n})$	$B_r(\mathbb{X}_{N_n})$	$\hat{ u}(\mathbb{X}_{N_n})$	$A_\nu(\mathbb{X}_{N_n})$	$B_{\nu}(\mathbb{X}_{N_n})$	$\hat{\delta}(\mathbb{X}_{N_n})$	$A_{\delta}(\mathbb{X}_{N_n})$	$B_{\delta}(\mathbb{X}_{N_n})$
$\Pi(10^4)$	0.5754	0.0458	1.1051	2.5877	1.8475	3.3278	1.0159	0.8915	1.1403
$\Pi(10^{5})$	0.4693	0.3526	0.5859	2.4633	2.3205	2.6061	0.9912	0.9630	1.0195
$\Pi(10^{6})$	0.5032	0.4631	0.5433	2.5039	2.4525	2.5554	1.0005	0.9909	1.0101
$G\left(\frac{1}{1+10^4}\right)$	0.4073	-0.2401	1.0549	2.3678	1.6620	3.0735	0.9767	0.8180	1.1355
$G\left(\frac{1}{1+10^5}\right)$	0.5613	0.1982	0.9243	2.5793	2.077	3.0808	1.0140	0.9284	1.0995
$G\left(\frac{1}{1+10^6}\right)$	0.4942	0.4392	0.5493	2.4927	2.4229	2.5626	0.9985	0.9853	1.0118
$NB\left(\frac{1}{2}, \frac{1}{1+10^4}\right)$	0.4575	-2.9511	3.8662	2.4297	-1.6338	6.4933	0.9884	0.1551	1.8217
$NB\left(\frac{1}{2}, \frac{1}{1+10^5}\right)$	0.4137	-0.7791	1.6065	2.4056	1.0728	3.7384	0.9780	0.6893	1.2668
$NB\left(\frac{1}{2}, \frac{1}{1+10^6}\right)$	0.5107	0.2524	0.7690	2.5133	2.1781	2.8485	1.0024	0.9404	1.0644
$NB(2, \frac{1}{1+10^4})$	0.5700	0.2545	0.8855	2.5935	2.1529	3.0341	1.0148	0.9407	1.0888
$NB(2, \frac{1}{1+10^5})$	0.5317	0.3867	0.6767	2.5418	2.3483	2.7353	1.0079	0.9733	1.0424
$NB\left(2,\frac{1}{1+10^{6}}\right)$	0.5052	0.4848	0.5256	2.5056	2.4793	2.5319	1.0009	0.9960	1.0058

Table 2. Examples of parameter estimates and boundaries of confidence intervals for a model distribution for r = 0.5, $\nu = 2.5$, and $\delta = 1.0$.

5. Conclusions

This paper has considered the problem of estimating the parameters of the digamma distribution with a random sample size. The consideration of a random sample size is very important since the accumulation of a sufficient fixed amount of statistical data can often take an indefinite amount of time, and, sometimes, it is impossible, in principle. Therefore, it becomes natural to study the asymptotic behavior of statistics based on random size samples.

The digamma distribution is a generalization of popular distributions from the gamma and beta classes, as well as the gamma-exponential distribution. This paper has discussed a method for estimating unknown parameters of the digamma distribution based on the logarithmic cumulants. Assuming that the sample size is random, the weak convergence of the studied estimators to the scale mixtures of the normal law is proved. This result allows for the construction of asymptotic confidence intervals for the estimated parameters. It is shown that the asymptotic properties of the statistics can change radically when passing from a fixed sample size to a random one. In particular, it leads to heavier tails of the limit distribution. For example, the type VII Pearson distribution may appear to be a limiting distribution whose representatives may not have a mathematical expectation.

The results proposed in this paper concern the estimation of the characteristic index and the shape and scale parameters of the digamma distribution assuming that the concentration parameters are known. Naturally, the question arises about the form of statistical estimates in the case in which all five parameters are unknown. The equations for constructing the estimates contain polygamma functions with arguments depending on the concentration parameters. Theoretical methods for inverting polygamma functions are being actively developed at the present time, but, apparently, there are currently no effective tools suitable for use in the method under consideration. At the same time, polygamma functions have nice properties that make their inversion easy using numerical methods. The authors plan to continue their studies in this direction.

Author Contributions: Conceptualization, A.K. and O.S.; methodology, A.K. and O.S.; formal analysis, A.K. and O.S.; investigation, A.K. and O.S.; writing—original draft preparation, A.K. and O.S.; writing—review and editing, A.K. and O.S.; supervision, A.K. and O.S.; funding acquisition, O.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Russian Science Foundation, project no. 22-11-00212.

Data Availability Statement: Data sharing not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Feng, M.; Qu, H.; Yi, Z.; Kurths, J. Subnormal Distribution Derived From Evolving Networks With Variable Elements *IEEE Trans. Cybern.* **2018**, *48*, 2556–2568.
- Iriarte, Y.A.; Varela, H.; Gómez, H.J.; Gómez, H.W. A Gamma-Type Distribution with Applications. Symmetry 2020, 12, 870. [CrossRef]
- 3. Feng, M.; Deng, L.-J.; Chen, F.; Perc, M.; Kurths, J. The accumulative law and its probability model: An extension of the Pareto distribution and the log-normal distribution. *Proc. R. Soc.* 2020, 476, 20200019. [CrossRef] [PubMed]
- 4. Barranco-Chamorro, I.; Iriarte, Y.A.; Gómez, Y.M.; Astorga, J.M.; Gómez, H.W. A Generalized Rayleigh Family of Distributions Based on the Modified Slash Model. *Symmetry* **2021**, *13*, 1226. [CrossRef]
- 5. López-Rodríguez, F.; García-Sanz-Calcedo, J.; Moral-García, F.J.; García-Conde, A.J. Statistical Study of Rainfall Control: The Dagum Distribution and Applicability to the Southwest of Spain. *Water* **2019**, *11*, 453. [CrossRef]
- Santoro, K.I.; Gómez, H.J.; Barranco-Chamorro, I.; Gómez, H.W. Extended Half-Power Exponential Distribution with Applications to COVID-19 Data. *Mathematics* 2022, 10, 942. [CrossRef]
- Kudryavtsev, A.A.; Nedolivko, Y.N.; Shestakov, O.V. Main Probabilistic Characteristics of the Digamma Distribution and the Method of Estimating Its Parameters. *Moscow Univ. Comput. Math. Cybern.* 2022, 46, 79–86. [CrossRef]
- 8. Amoroso, L. Ricerche intorno alla curva dei redditi. Ann. Mat. Pura Appl. 1925, 21, 123–159. [CrossRef]
- 9. McDonald, J.B. Some Generalized Functions for the Size Distribution of Income. Econometrica 1984, 52, 647–665. [CrossRef]
- 10. Kudryavtsev, A.A. On the representation of gamma-exponential and generalized negative binomial distributions. *Inform. Appl.* **2019**, *13*, 78–82.
- 11. Kudryavtsev, A.A.; Shestakov, O.V. Digamma Distribution as a Limit for the Integral Balance Index. *Moscow Univ. Comput. Math. Cybern.* **2022**, *46*, 133–139. [CrossRef]
- 12. Kudryavtsev, A.A. Bayesian balance models. Inform. Appl. 2018, 12, 18–27.
- 13. Combes, C.; Ng, H.K.T. On parameter estimation for Amoroso family of distributions. *Math. Comp. Sim.* **2021**, *191*, 309–327. [CrossRef]
- 14. Liu, S.; Gui, W. Estimating the Parameters of the Two-Parameter Rayleigh Distribution Based on Adaptive Type II Progressive Hybrid Censored Data with Competing Risks. *Mathematics* **2020**, *8*, 1783. [CrossRef]
- 15. Kudryavtsev, A.; Shestakov, O. Asymptotically Normal Estimators for the Parameters of the Gamma-Exponential Distribution. *Mathematics* **2021**, *9*, 273. [CrossRef]
- 16. Kudryavtsev, A.A.; Shestakov, O.V. A Method for Estimating Bent, Shape and Scale Parameters of the Gamma-Exponential Distribution. *Inform. Appl.* **2021**, *15*, 57–62.
- 17. Kudryavtsev, A.; Shestakov, O. The Estimators of the Bent, Shape and Scale Parameters of the Gamma-Exponential Distribution and their Asymptotic Normality. *Mathematics* 2022, *10*, 619. [CrossRef]
- 18. Korolev, V.Y. Product representations for random variables with Weibull distributions and their applications. *J. Math. Sci.* 2016, 218, 298–313. [CrossRef]
- 19. Gnedenko, B.V. On the estimation of unknown distribution parameters with a random number of independent observations. *Tr. Tbilis. Mat. Inst.* **1989**, *92*, 146–150.
- 20. Korolev, V.Y.; Zeifman, A.I. Generalized negative binomial distributions as mixed geometric laws and related limit theorems. *Lith. Math. J.* **2019**, *59*, 366–388. [CrossRef]
- 21. Wang, X.; Zhao, X.; Sun, J. A compound negative binomial distribution with mutative termination conditions based on a change point. *J. Comput. Appl. Math.* **2019**, *351*, 237–249. [CrossRef]
- 22. Bhati, D.; Ahmed, I.S. On uniform-negative binomial distribution including Gauss hypergeometric function and its application in count regression modeling. *Commun. Stat. Theory Methods* **2021**, *50*, 3106–3122. [CrossRef]
- 23. Zhang, J.; Wang, D.; Yang, K. A study of RCINAR(1) process with generalized negative binomial marginals. *Commun. Stat. B Simul. Comput.* **2020**, 49, 1487–1510. [CrossRef]
- 24. Mangiola, S.; Thomas, E.A.; Modrák, M.; Vehtari, A.; Papenfuss, A.T. Probabilistic outlier identification for RNA sequencing generalized linear models. *NAR Genom. Bioinform.* **2021**, *3*, lqab005. [CrossRef]
- 25. Kendall, M.G.; Stuart, A. *The Advanced Theory of Statistics*, 3rd ed.; Griffin: London, UK, 1969; Volume 1.
- 26. Korolev, V.Y.; Zeifman, A.I. On Convergence of the Distributions of Random Sequences with Independent Random Indexes to Variance–Mean Mixtures. *Stoch. Models.* **2016**, *32*, 414–432. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.