



Article

Second Order Chebyshev– Edgeworth-Type Approximations for Statistics Based on Random Size Samples

Gerd Christoph and Vladimir V. Ulyanov

Special Issue

Limit Theorems of Probability Theory

Edited by

Prof. Dr. Alexander Tikhomirov and Prof. Dr. Vladimir Ulyanov



Article

Second Order Chebyshev–Edgeworth-Type Approximations for Statistics Based on Random Size Samples

Gerd Christoph ^{1,*}  and Vladimir V. Ulyanov ^{2,3,†} ¹ Department of Mathematics, Otto-von-Guericke University Magdeburg, 39016 Magdeburg, Germany² Faculty of Computer Science, HSE University, 101000 Moscow, Russia³ Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University, 119991 Moscow, Russia

* Correspondence: gerd.christoph@ovgu.de

† These authors contributed equally to this work.

Abstract: This article completes our studies on the formal construction of asymptotic approximations for statistics based on a random number of observations. Second order Chebyshev–Edgeworth expansions of asymptotically normally or chi-squared distributed statistics from samples with negative binomial or Pareto-like distributed random sample sizes are obtained. The results can have applications for a wide spectrum of asymptotically normally or chi-square distributed statistics. Random, non-random, and mixed scaling factors for each of the studied statistics produce three different limit distributions. In addition to the expected normal or chi-squared distributions, Student’s t -, Laplace, Fisher, gamma, and weighted sums of generalized gamma distributions also occur.

Keywords: second order Chebyshev–Edgeworth expansions; negative binomially distributed sample sizes; Pareto-like distributed sample sizes; asymptotically normally distributed statistics; asymptotically chi-square distributed statistics; scaled Student’s t -distribution; normal distribution; discrete Pareto distribution; generalized Laplace distribution; weighted sums of generalized gamma distributions

MSC: 62E17; 62H10; 60E05



Citation: Christoph, G.; Ulyanov, V.V. Second Order Chebyshev–Edgeworth-Type Approximations for Statistics Based on Random Size Samples. *Mathematics* **2023**, *11*, 1848. <https://doi.org/10.3390/math11081848>

Academic Editor: Steve Drekic

Received: 27 February 2023

Revised: 3 April 2023

Accepted: 11 April 2023

Published: 13 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/>).

1. Introduction

To improve the convergence properties of sums of independent identically distributed random variables in the Central Limit Theorem, asymptotic expansions of distribution functions of normalized sums were considered. The history of asymptotic expansions in nonparametric statistics is presented in detail in Wallace [1], Bickel [2], and Hall [3], among others. Chebyshev–Edgeworth expansions, with which we are concerned here, are presented in great detail in Bhattacharya and Rao [4] for random vectors and in Petrov [5] for one-dimensional random variables. For instance, in Pfanzagl [6] and Bentkus et al. [7], the authors emphasize that asymptotic expansions can provide more effective approximations for asymptotic studies in statistical theory. Second order approximations of distribution functions of sums of random variables are of great importance because they take into account the skewness and kurtosis of the random variable in addition to the expected value and the variance, as in the Central Limit Theorem. In Burnashev [8], second order expansions are proved for the asymptotically normally distributed sample median M_m on a sample of size m and its MSE. Based on this, for a Laplace population with density $e^{-|x|}/2$, the actual MSE with exact data is compared numerically with approximations data. For the normal approximation, the influence of the remaining term is below 10% only for $m > 250$, while for the approximation with the second order expansion, the influence of the remaining term is below 10% already from $m = 8$. For a Cauchy population with smooth and heavy tailed density $1/(\pi(1+x^2))$, for the normal approximation, the influence of the remaining term is below 10% for $m \geq 23$, while for the approximation

with the second order expansion, the influence of the remaining term is below 10% already from $m = 11$. Consequently, as Burnashev [8] pointed out, asymptotic expansions can significantly improve the exactness of statistical conclusions, even in the case of a small number of observations. The results in the abovementioned papers are based on non-random sample sizes or non-random number of observations.

When planning statistical studies, situations often arise where the sample sizes are unknown in advance and they are modeled as realizations of random variables. Many models from medicine, finance, risk theory, physics, and reliability lead to samples with random dimensions. For instance, in the papers by Nunes et al. [9,10,11], different models in medical research random size samples were investigated in order to prevent false conclusions. In Esquivel et al. [12], the authors give an informative overview of statistical inference with a random number of observations and some applications. Results for mean and variance for normally distributed samples, calculation of quantiles, and interval estimates with random sample size were also proved. Döbler [13] gives a detailed review of the literature on random sums as well as recent results on approximation in various metrics. In Schluter and Trede [14] (Theorem 1, Proposition 1), the authors show, using the convergence of a negative binomial random sum, that the growth rate of cities is Student t-distributed with 2 degrees of freedom. Their empirical investigations verify the result. The references in the above-cited papers provide further applications for random dimension sampling.

Bening et al. [15,16] proved convergence rates and asymptotic expansions for distributions of statistics T_{N_n} based on samples with random dimension $N_n \geq 1$. Here, T_m is a statistic based on a non-random number $m \geq 1$ of independent observations. The random variables size $N_n \geq 1$ form a sequence of integer random sample sizes that depends on a natural parameter n with $N_n \rightarrow \infty$ in probability for $n \rightarrow \infty$. Inequalities with a convergence rate are assumed for the approximations of the distribution functions of both the normalized statistics T_m and the normalized random sample sizes N_n . As examples, convergence rates and first order asymptotic expansions are derived for the statistics T_{N_n} , where T_m is an asymptotically normal statistic and the random sample size N_n is either negatively binomial or Pareto-like distributed.

In Christoph et al. [17], inequalities for the second order approximations of the distribution functions of normalized negative binomial and Pareto-like sample sizes were proved. Consequently, second order Chebyshev–Edgeworth approximations and the corresponding Cornish–Fisher expansions could be obtained for the distribution of the normalized arithmetic mean of a sample with normalized negative binomial or Pareto-like sample sizes where the remainders are of order $n^{-3/2}$.

The present work provides a supplement to our paper, Christoph and Ulyanov [18], where we have developed a formal second order design for asymptotic Chebyshev–Edgeworth approximations. We considered asymptotically normal statistics with sample size having negative binomial distribution as well as asymptotically chi-squared statistics with Pareto-like distributed sample sizes. In addition to the distributions of statistic T_m and random sample size N_n , three scaling factors for T_{N_n} are also introduced, leading to different expansions. It is the first paper to consider approximations for asymptotic chi-square statistics based on random sample sizes. Some more applications of random sample size sampling were also mentioned.

In the present paper, we provide similar results for asymptotically normal statistics of samples with Pareto-like distributed sample sizes and for asymptotically chi-squared statistics with sample size having negative binomial distribution.

For better reader convenience, we list in Section 2 some notations, conditions, and statements that were also used in Christoph and Ulyanov [18]. Section 3 states the necessary approximations for the statistics T_m and the sample sizes N_n . The dependence of the limit distributions of the scaled statistic T_{N_n} on the distributions of the statistic T_m and the sample size N_n , as well as the scaling factors, is discussed in Section 4. Section 5 then presents the main results. As examples, we consider the same statistic T_m as in

Christoph and Ulyanov [18] (Corollaries 1 and 2), but with changed sample sizes. Section 6 provides the proofs of the main results, leaving three auxiliary lemmas to Appendix A. Conclusions are presented in Section 7.

2. Notation and Preliminaries

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space on which all occurring random variables are given. Set positive numbers, real axis, integer part $[y]$ of real y , and indicator function as follows:

$$\mathbb{N}_+ = \{1, 2, \dots\}, \quad \mathbb{R} = (-\infty, \infty), \quad y - 1 < [y] \leq y \quad \text{and} \quad \mathbb{I}_A = \mathbb{I}_A(x) = \begin{cases} 1, & x \in A \subset \mathbb{R} \\ 0, & x \notin A \subset \mathbb{R} \end{cases}.$$

Let $X_1, X_2, X_3 \dots \in \mathbb{R}$ be independent identically distributed random variables. Define the statistic

$$T_m := T_m(X_1, \dots, X_m) \quad \text{with} \quad m \in \mathbb{N}_+,$$

based on the random sample $\{X_1, X_2, \dots, X_m\}$ with a non-random sample size $m \in \mathbb{N}_+$.

Consider the sequence of discrete random variables N_1, N_2, \dots , depending on an integer parameter $n \geq 1$. This integer $N_n \geq 1$ indicates the random dimension of the observations X_1, \dots, X_{N_n} . Let us assume that the sample size N_n does not depend on $X_1, X_2, X_3 \dots$, where $N_n \rightarrow \infty$ in probability when $n \rightarrow \infty$. Define for each $n \in \mathbb{N}_+$ the statistic T_{N_n} obtained from a random sample $\{X_1, X_2, \dots, X_{N_n}\}$ by

$$T_{N_n}(\omega) := T_{N_n(\omega)}(X_1(\omega), X_2(\omega), \dots, X_{N_n(\omega)}(\omega)) \quad \text{for each} \quad \omega \in \Omega. \quad (1)$$

It follows from Esquivel et al. [12] (Theorem 2.1.1) that the statistic T_{N_n} is well-defined in (1).

Since we want to prove second order approximations for the statistic T_{N_n} in form of inequalities, we need the corresponding assumptions for the statistic T_m and for the random sample size N_n as well.

For the statistic T_m with $\mathbb{E}T_m = 0$ and the random sample sizes $N_n \in \mathbb{N}_+$ we suppose conditions on the structure of the approximating functions as well as on the convergence rate:

Assumption 1. There are a distribution function $F(x)$, bounded functions $f_1(x)$, $f_2(x)$ which are differentiable for all $x \neq 0$, $\gamma \in \{-1, -1/2, 0, 1/2, 1\}$, $a > 1/2$ as well as $0 < C_1 < \infty$ such that

$$\sup_x \left| \mathbb{P}(m^\gamma T_m \leq x) - F(x) - m^{-1/2} f_1(x) - \mathbb{I}_{a>1}(a) m^{-1} f_2(x) \right| \leq C_1 m^{-a}, \quad m \leq 1. \quad (2)$$

Assumption 2. There exists a distribution function $H(y)$ with $H(0+) = 0$, a bounded variation function $h_2(y)$, a sequence of numbers $0 < g_n \uparrow \infty$, $b > 0$, and $0 < C_2 < \infty$ such that for $n \in \mathbb{N}_+$

$$\left. \begin{aligned} \sup_{y \geq 0} |\mathbb{P}(g_n^{-1} N_n \leq y) - H(y)| &\leq C_2 n^{-b}, & \text{for } 0 < b \leq 1, \\ \sup_{y \geq 0} |\mathbb{P}(g_n^{-1} N_n \leq y) - H(y) - n^{-1} h_2(y)| &\leq C_2 n^{-b}, & \text{for } b > 1. \end{aligned} \right\} \quad (3)$$

Remark 1. Assumptions 1 and 2 require inequalities for the approximations of T_m and N_n for all $m, n \in \mathbb{N}_+$, leading to inequalities for the approximations of T_{N_n} . See also Remark 5 below on Poisson and binomial random variables N_n . For these sample sizes, we are so far only aware of estimates of the remaining terms with small- o or large- \mathcal{O} convergence rates. About the differences between inequalities and \mathcal{O} order bounds, see, e.g., Fujikoshi and Ulyanov [19] (Chapter 1).

Remark 2. In Bening et al. [16], these conditions are formulated more generally. Assumption 1 requires the existence of f_1, \dots, f_l with $a > l/2$ and Assumption 2 that of h_1, \dots, h_k with $b > k/2$.

We restrict ourselves here, as in Christoph and Ulyanov [18], to the required approximation functions.

Assumptions 1 and 2 lead to the approximations for the distribution functions of statistics T_{N_n} :

Proposition 1. (Christoph and Ulyanov [18], Proposition 1) Let $\gamma \in \{-1, -1/2, 0, 1/2, 1\}$. The statistic T_m and the sample size N_n are supposed to satisfy Assumptions 1 and 2, respectively. Then,

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P}(g_n^\gamma T_{N_n} \leq x) - G_n(x, 1/g_n) \right| \leq C_1 \mathbb{E}(N_n^{-a}) + (C_3 D_n + C_4) n^{-b}, \quad (4)$$

where $a > 0, b > 0$ are the convergence rates in (2) and (3),

$$G_n(x, 1/g_n) = \int_{1/g_n}^{\infty} \left(F(xy^\gamma) + \frac{f_1(xy^\gamma)}{\sqrt{g_n y}} + \frac{f_2(xy^\gamma)}{g_n y} \right) d \left(H(y) + \frac{h_2(y)}{n} \right), \quad (5)$$

$$D_n = \sup_x \int_{1/g_n}^{\infty} \left| \frac{\partial}{\partial y} \left(F(xy^\gamma) + \frac{f_1(xy^\gamma)}{\sqrt{g_n y}} + \frac{f_2(xy^\gamma)}{y g_n} \right) \right| dy, \quad (6)$$

and $f_1(z), f_2(z), h_2(y)$ are given in (2) and (3). The constants C_1, C_3, C_4 do not depend on n .

Bening et al. [16] proved general transfer theorems under the conditions indicated in Remark 2 only for case $\gamma \geq 0$. Therefore, the proof is repeated in Christoph and Ulyanov [20] (Appendix A.1).

3. Second Order Estimates for Both the Statistics T_m and the Sample Sizes N_n

First we consider the following statistics T_m with non-random sample size m and $\mathbb{E}T_m = 0$ with the corresponding second order approximations. Let the asymptotically normal statistic T_m satisfy the following inequality:

$$\left| \mathbb{P}(\sqrt{m}T_m \leq x) - \Phi(x) - (m^{-1/2}(p_0 + p_2 x^2) + m^{-1}(p_1 x + p_3 x^3 + p_5 x^5) \mathbb{I}_{a>1}(a)) \varphi(x) \right| \leq C m^{-a} \quad (7)$$

with $a > 0$ and $\Phi(x)$ refers to the standard normal distribution function with density function $\varphi(y)$:

$$\Phi(x) = \int_{-\infty}^x \varphi(y) dy, \quad x \in \mathbb{R}, \quad \text{and} \quad \varphi(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}, \quad y \in \mathbb{R}.$$

Asymptotically chi-squared distributed statistics T_m satisfy the following inequality:

$$\left| \mathbb{P}(mT_m \leq x) - G_d(x) - m^{-1}(q_1 x + q_2 x^2) g_d(x) \right| \leq C m^{-2}, \quad (8)$$

where $G_d(x)$, $d \in \mathbb{N}_+$, denotes the chi-squared distribution function with d degrees of freedom and the density function $g_d(y)$:

$$g_d(y) = \frac{1}{2^{d/2} \Gamma(d/2)} y^{(d-2)/2} e^{-y/2}, \quad y > 0, \quad \text{and} \quad G_d(x) = \mathbb{P}(\chi_d^2 \leq x) = \int_0^x g_d(y) dy, \quad x > 0.$$

In Christoph and Ulyanov [18] (Sections 3.1 and 3.2), some examples of such statistics T_m are given that satisfy (7) or (8) and consequently, Assumption 1.

As already announced, we consider the following random sample sizes N_n with the corresponding second order approximations.

The Pareto-like random sample sizes $N_n(s)$ are defined as follows:

Let $Y_j(s) \in \mathbb{N}_+$, $j = 1, 2, \dots$ be independent discrete Pareto II random variables with parameter $s > 0$, which are discretized from continuous Lomax (Pareto II) random variables on \mathbb{N}_+ , for a review, see, e.g., Buddana and Kozubowski [21]. For $s > 0$, there are defined

$$\mathbb{P}(Y_j(s) \leq k) = \frac{k}{s+k}, \quad N_n(s) = \max_{1 \leq j \leq n} Y_j(s) \quad \text{and} \quad \mathbb{P}(N_n(s) \leq k) = \left(\frac{k}{s+k} \right)^n, \quad n, k \in \mathbb{N}_+. \quad (9)$$

Proposition 2. (Christoph and Ulyanov [18], Proposition 4) *Let $N_n(s)$ be the discrete Pareto-like random variable whose distribution function is given in (9); then, for all integers $n \geq 1$ and fixed positive $s > 0$, we have*

$$\sup_{y>0} \left| \mathbb{P} \left(\frac{N_n(s)}{n} \leq y \right) - W_s(y) - \frac{h_{2,s}(y)}{n} \right| \leq \frac{C_2(s)}{n^2} \quad (10)$$

$$W_s(y) = e^{-s/y} \quad y > 0, \quad h_{2,s}(y) = \frac{s e^{-s/y}}{2y^2} (s-1 + 2Q_1(ny)), \quad y > 0, \quad (11)$$

with jump correcting function $Q_1(y) = 1/2 - (y - [y])$ and $C_2(s) > 0$ does not depend on n . Furthermore,

$$\mathbb{E}(N_n(s))^{-a} \leq C(a, s) n^{-\min\{a, 2\}}, \quad (12)$$

with optimal bound in (12) for $0 < a \leq 2$, where a is the convergence rate in (7).

Remark 3. The inverse exponential random variable $W(s)$ with distribution function $H_s(y) = \mathbb{P}(W(s) \leq y) = e^{-s/y} \mathbb{I}_{(0, \infty)}(y)$ and rate parameter $s > 0$ is “heavy tailed” with shape parameter 1 as is $\mathbb{P}(N_n(s) \leq y)$. Thus, the expected values of these two random variables do not exist.

Suppose the positive integer $N_n(r)$ has a (shifted by 1) negative binomial distribution with probability of success $1/n$, $n \in \mathbb{N}_+$, parameter $r > 0$, probabilities

$$\mathbb{P}(N_n(r) = j) = \frac{\Gamma(j+r-1)}{\Gamma(j)\Gamma(r)} \left(\frac{1}{n} \right)^r \left(1 - \frac{1}{n} \right)^{j-1}, \quad j \in \mathbb{N}_+ \quad \text{and} \quad g_n = \mathbb{E}(N_n(r)) = r(n-1) + 1. \quad (13)$$

In statistical studies, for counting models, the negative binomial and Poisson distributions are the two most important ones. In Schluter and Trede [14] (Section 2.1), the authors emphasize that the negative binomial distribution with its two parameters can typically observe over-dispersion in count data, while this is not the case with the one-parameter Poisson distribution. They proved in a more general framework

$$\lim_{n \rightarrow \infty} \sup_y |\mathbb{P}(N_n(r)/g_n \leq y) - G_{r,r}(y)| = 0, \quad (14)$$

while $G_{r,r}(y)$ denotes the gamma distribution that has identical scale and shape parameters $r > 0$, whose density is

$$g_{r,r}(y) = \frac{r^r}{\Gamma(r)} y^{r-1} e^{-ry} \mathbb{I}_{(0, \infty)}(y), \quad y \in \mathbb{R}.$$

In Bening and Korolev [22] (Lemma 2.2), the result (14) was also obtained.

Proposition 3. (Christoph and Ulyanov [18], Proposition 3) *Let $r > 0$. The discrete random variable $N_n(r)$ has probabilities and expected value g_n given in (13). Then, for all $n \in \mathbb{N}_+$:*

$$\sup_{y \geq 0} \left| \mathbb{P} \left(\frac{N_n(r)}{g_n} \leq y \right) - G_{r,r}(y) - \frac{h_{2,r}(y)}{n} \right| \leq C_2(r) n^{-\min\{r, 2\}}, \quad (15)$$

where $C_2(r) > 0$ does not depend on n and with the jump correcting function $Q_1(y) = 1/2 - (y - [y])$,

$$h_{2,r}(y) = \begin{cases} 0, & \text{for } r \leq 1, \\ \frac{g_{r,r}(y)}{2r} ((y-1)(2-r) + 2Q_1(g_n y)), & \text{for } r > 1. \end{cases} \quad (16)$$

Moreover, negative moments $\mathbb{E}(N_n(r))^{-\alpha}$ satisfy the estimation for all $r > 0, \alpha > 0$

$$\mathbb{E}(N_n(r))^{-\alpha} \leq C(r) \begin{cases} n^{-\min\{r, \alpha\}}, & r \neq \alpha \\ \ln(n) n^{-\alpha}, & r = \alpha \end{cases} \quad (17)$$

and the convergence rate in case $r = \alpha$ cannot be improved.

Remark 4. Second order Chebyshev–Edgeworth expansions (10) and (15) with $r > 1$ were first proved in Christoph et al. [17] (Theorems 4 and 1). Approximations in (10) and (15) with remainder estimations C_s/n or $C_r n^{-\min\{r, 1\}}$ are given, e.g., in Bening et al. [16] and Gavrilenko et al. [23]. In Christoph et al. [24] (Corollaries 5.4 and 6.5), leading terms for the negative moments of $N_n(r)$ and $N_n(s)$ are derived that lead to (17) and (12).

Remark 5. The negative binomial distribution belongs to the class of Panjer distributions, which also includes the Poisson and binomial distributions. Samples with binomial or Poisson distributed sample sizes were studied among others in the above-cited papers [9–12]. Convergence rate bounds for statistics based on such samples are given in Döbler [13], Korolev [25], Bulinski and Slepov [26]. Döbler [13], Korolev and Shevtsova [27], Sunklodas [28] obtained Berry–Esseen bounds for sums based on samples with binomial and Poisson sample sizes. To the best of the authors' knowledge, Chebyshev–Edgeworth expansions for these lattice distributed random variables have only been proven so far with bounds of small- o or large- O rates, see, e.g., Petrov [29] (Chapter 6, Theorem 6) or Kolassa and McCullagh [30]. Therefore, inequality (3) in Assumption 2 is not fulfilled.

4. Limit Distributions of Statistics with Random Size Samples using Different Scaling Factors

We now consider the statistics T_m and the sample sizes N_n , which are supposed to satisfy the inequalities (2) and (3) in Assumptions 1 and 2, respectively. Let us investigate the scaled statistics $g_n^\gamma N_n^{\gamma^* - \gamma} T_{N_n}$ with the sequence $g_n \uparrow \infty$ as $n \rightarrow \infty$. We analyze the two cases Φ and G_u as limiting distributions F in Assumption 1 with respect to the exponents γ^* and γ : If $F = \Phi$, then $\gamma^* = 1/2$ and $\gamma \in \{-1/2, 0, 1/2\}$, while if $F = G_u$, then $\gamma^* = 1$ and $\gamma \in \{-1, 0, 1\}$. Then, conditioning on N_n and using (2) and (3), we have

$$\begin{aligned} \mathbb{P}(g_n^\gamma N_n^{\gamma^* - \gamma} T_{N_n} \leq x) &= \mathbb{P}(N_n^{\gamma^*} T_{N_n} \leq x (N_n/g_n)^\gamma) = \sum_{m=1}^{\infty} \mathbb{P}(m^{\gamma^*} T_m \leq x(m/g_n)^\gamma) \mathbb{P}(N_n = m) \\ &\stackrel{(2)}{\approx} \mathbb{E}\left(F(x(N_n/g_n)^\gamma)\right) = \int_{1/g_n}^{\infty} F(xy^\gamma) d\mathbb{P}(N_n/g_n \leq y) \stackrel{(3)}{\approx} \int_{1/g_n}^{\infty} F(xy^\gamma) dH(y). \end{aligned} \quad (18)$$

Consequently, the limit distribution of the scaled statistic $g_n^\gamma N_n^{\gamma^* - \gamma} T_{N_n}$ is a scale mixture of underlying F with mixing distribution H : $\mathbb{P}(g_n^\gamma N_n^{\gamma^* - \gamma} T_{N_n} \leq x) \rightarrow \int_0^\infty F(xy^\gamma) dH(y)$, as $n \rightarrow \infty$. Refer to, e.g., Choy and Chan [31], Fujikoshi et al. [32] (Chapter 13), and Fujikoshi and Ulyanov [19] (Chapter 2) and the references therein.

The limiting distributions $\int_{1/g_n}^\infty F(xy^\gamma) dH(y)$ therefore only arise from the leading distributions $F(x)$ and $H(y)$ in the inequalities (2) and (3) and also depend on the parameter γ .

In Christoph and Ulyanov [18] (Sections 5 and 6), the cases $F(x) = \Phi(x)$ with $H(y) = G_{r,r}(y)$ as well as $F(x) = G_u(x)$ with $H(y) = W_s(y)$ were considered. Now, we interchange the distributions of random sample sizes N_n . We first study the limiting distributions of asymptotically normally distributed statistics with Pareto-like distributed sample sizes $N_n(s)$ and also asymptotically chi-squared distributed statistics with negative binomial distributed sample sizes $N_n(r)$. Since $W_s(1/n) = e^{-s/n}$ and $G_{r,r}(1/g_n) \leq \frac{r^{r-1}}{\Gamma(r)} g_n^{-r}$ hold, the integral range in the last integral in (18) can be extended from $(1/g_n, \infty)$ to $(0, \infty)$ for further investigations.

4.1. The Case $F(x) = \Phi(x)$ and $H(y) = W_s(y)$

In Christoph and Ulyanov [20,33], asymptotically normally distributed statistics T_m for samples of m -dimensional normally distributed vectors were considered: correlation coefficient as well as the three geometric features: the length of a vector, the distance, and the angle between two vectors. Inequalities for second order approximations for statistic T_m are derived when the dimension m is replaced by Pareto-like distributed random dimension $N_n(s)$. For the median of a sample with random sample size $N_n(s)$ analogous results are shown in Christoph et al. [24] (Section 6). All these asymptotically normally distributed statistics $T_{N_n(s)}$ with Pareto-like random dimensions or sample sizes have the same limiting distribution.

Let $\gamma \in \{1/2, 0, -1/2\}$. Since $\mathbb{E}N_n(s) = \infty$, we choose as $g_n = n$. Then, the limit laws for

$$\mathbb{P}\left(n^\gamma N_n(s)^{1/2-\gamma} T_{N_n(s)} \leq x\right) \text{ are } V_\gamma(x, s) = \int_0^\infty \Phi(xy^\gamma) dH_s(y) = \int_0^\infty \Phi(xy^\gamma) \frac{s}{y^2} e^{-sy} dy.$$

with corresponding densities

$$v_\gamma(x, s) = \frac{s}{\sqrt{2\pi}} \int_0^\infty y^{\gamma-2} e^{-(x^2 y^{2\gamma}/2 + sy)} dy = \begin{cases} l_{1/\sqrt{s}}(x) = \frac{\sqrt{2s}}{2} e^{-\sqrt{2s}|x|}, & \gamma = \frac{1}{2}, \\ \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, & \gamma = 0, \\ s_2^*(x; \sqrt{s}) = \frac{1}{2\sqrt{2s}} \left(1 + \frac{x^2}{2s}\right)^{-3/2}, & \gamma = -\frac{1}{2}, \end{cases} \quad (19)$$

Therefore, the limit distributions $V_\gamma(x, s)$ are the Laplace law $L_{1/\sqrt{s}}(x)$ with density $l_{1/\sqrt{s}}(x)$ and scale parameter $\lambda = 1/\sqrt{s}$ for $\gamma = 1/2$, the standard normal law $\Phi(x)$ and density $\varphi(x)$ for $\gamma = 0$ and for $\gamma = -1/2$ the scaled Student's t -distribution $S_2^*(x; \sqrt{s})$ with 2 degrees of freedom and density $s_2^*(x; \sqrt{s})$. These mixed scale distributions $V_\gamma(x, s)$ are discussed in more detail in Christoph and Ulyanov [20] (Section 4.2).

4.2. The Case $F(x) = G_d(x)$ and $H(y) = G_{r,r}(y)$

Asymptotically chi-squared distributed statistics of samples with random sample size were considered for the first time in Christoph and Ulyanov [18] in case of $H(y) = W_s(y) = e^{-s/y}$, $y > 0$.

Now, negatively binomial distributed sample sizes $N_n(r)$ are considered. With $\gamma \in \{1, 0, -1\}$ and $g_n = \mathbb{E}N_n(r) = r(n-1) + 1$, the limit distributions for

$$\mathbb{P}\left(g_n^\gamma N_n(r)^{1-\gamma} T_{N_n(r)} \leq x\right) \text{ are } V_\gamma(x; d, r) = \int_0^\infty G_d(xy^\gamma) dG_{r,r}(y) = \int_0^\infty G_d(xy^\gamma) \frac{r^r}{\Gamma(r)} y^{r-1} e^{-ry} dy.$$

The corresponding densities are

$$v_\gamma(x; d, r) = \frac{r^r x^{d/2-1}}{\Gamma(r) 2^{d/2} \Gamma(d/2)} \int_0^\infty y^{r+\gamma d/2-1} e^{-(xy^\gamma/2 + ry)} dy$$

$$= \begin{cases} f^*(x; d, 2r) = \frac{\Gamma(d/2+r)}{\Gamma(d/2)\Gamma(r)} \frac{x^{d/2-1}}{2^{d/2} r^{d/2}} \left(1 + \frac{x}{2r}\right)^{-(d+2r)/2}, & \gamma = 1, \\ g_d(x) = \frac{1}{2^{d/2} \Gamma(d/2)} x^{d/2-1} e^{-x/2}, & \gamma = 0, \\ w_{r-d/2}(x; d, r) = \frac{r}{\Gamma(r) \Gamma(d/2)} \left(\frac{xr}{2}\right)^{r/2+d/4-1} K_{r-d/2}(\sqrt{2rx}). & \gamma = -1. \end{cases} \quad (20)$$

We prove (20) for $\gamma = \pm 1$ in Section 6 in the proof of Theorem 2.

The scale mixtures $V_\gamma(x; d, r)$ are the (scaled by d) F -distribution $F^*(x; d, 2r) = F(x/d; d, 2r)$ with parameters $d \in \mathbb{N}_+$ and $r > 0$ and density $f^*(x; d, 2r) = \frac{1}{d} f(\frac{x}{d}; d, 2r)$ for $\gamma = 1$, the chi-squared distribution $G_d(x)$ with d degrees of freedom and density $g_d(x)$ for $\gamma = 0$ and a gamma distribution of generalized type $W_{r-d/2}(x; d, r)$ occurs with density $w_{r-d/2}(x; d, r)$ for $\gamma = -1$. The modified Bessel function of the third kind or Macdonald

functions $K_\lambda(u)$ also occurred in Christoph and Ulyanov [18,20] in generalized gamma and Laplace densities.

Remark 6. The Macdonald function satisfying order-reflection formula $K_{-\lambda}(u) = K_\lambda(u)$ and $K_\lambda(u)$ may be expressed for $\lambda = m + 1/2$ with integer m in closed forms. In Oldham et al. [34] (Formulas 51:4:1 and 26:13:3), the Macdonald functions $K_{-\lambda}(u) = K_\lambda(u)$ for $\lambda = 1/2, 3/2, 5/2, 7/2, 9/2$ are explicitly given. Using Prudnikov et al. [35] (Formulas 2.3.16.1-3), the densities $w_{r-d/2}(x; d, r) = w_{m+1/2}(x; d, r)$ can be calculated:

$$w_{m+1/2}(x; d, r) = \frac{r^r x^{d/2-1}}{\Gamma(r) 2^{d/2} \Gamma(d/2)} \begin{cases} (-1)^m \sqrt{\pi} \frac{\partial^m}{\partial r^m} (r^{-1/2} e^{-\sqrt{2rx}}), & m = 0, 1, 2, \dots, \\ (-2)^{-m} \sqrt{\frac{\pi}{r}} \frac{\partial^{-m}}{\partial x^{-m}} e^{-\sqrt{2rx}}, & m = 0, -1, -2, \dots \end{cases} \quad (21)$$

Example 1. Some densities $w_{m+1/2}(x; d, r)$ for $m = r - (d + 1)/2 = -2, -1, 0, 1, 2$:

$$\begin{aligned} m = -2 \quad d = 7, r = 2 \quad w_{-3/2}(x; 7, 2) &= \frac{4x}{15} (1 + \sqrt{4x}) e^{-\sqrt{4x}} \\ m = -1 \quad d = 4, r = 3/2 \quad w_{-1/2}(x; 4, 3/2) &= \frac{3}{4} \sqrt{3x} e^{-\sqrt{3x}} \\ m = 0 \quad d = 4, r = 5/2 \quad w_{1/2}(x; 4, 5/2) &= \frac{1}{12} \sqrt{25x} e^{-\sqrt{5x}} \\ m = 0 \quad d = 3, r = 2 \quad w_{1/2}(x; 3, 2) &= \sqrt{4x} e^{-\sqrt{4x}} \\ m = 1 \quad d = 3, r = 3 \quad w_{3/2}(x; 3, 3) &= \frac{3}{8} (6x + \sqrt{6x}) e^{-\sqrt{6x}} \\ m = 2 \quad d = 3, r = 4 \quad w_{5/2}(x; 3, 4) &= \frac{1}{12} \left((8x)^{3/2} + 24x + 3\sqrt{8x} \right) e^{-\sqrt{8x}}. \end{aligned}$$

Remark 7. If $m = r - (d + 1)/2$ is an integer, the distribution functions $W_{m+1/2}(x; d, r)$ of the densities $w_{m+1/2}(x; d, r)$ can also be calculated explicitly by substitution and partial integration.

Example 2. Distribution functions $W_\lambda(x; d, r)$ for given densities $w_\lambda(x; d, r)$ with $\lambda = \pm 1/2$:

$$w_{-1/2}(x; 4, \frac{3}{2}) = \frac{3}{4} \sqrt{3x} e^{-\sqrt{3x}} \text{ and } W_{-1/2}(x; 4, \frac{3}{2}) = 1 - \frac{1}{2} (2\sqrt{3x} + 3x + 2) e^{-\sqrt{3x}} \quad (22)$$

$$w_{1/2}(x; 4, \frac{5}{2}) = \frac{25x}{12} e^{-\sqrt{5x}} \text{ and } W_{1/2}(x; 4, \frac{5}{2}) = 1 - \left(\frac{(5x)^{3/2}}{6} + \frac{5x}{2} + \frac{\sqrt{5x}}{6} + 1 \right) e^{-\sqrt{5x}} \quad (23)$$

$$w_{1/2}(x; 3, 2) = \sqrt{4x} e^{-\sqrt{4x}} \text{ and } W_{1/2}(x; 3, 2) = 1 - (2x + 2\sqrt{x} + 1) e^{-\sqrt{4x}}. \quad (24)$$

Remark 8. The generalized gamma distribution $G^*(x; \beta, \alpha, \lambda)$ has two shape parameters α and β , a scale parameter λ , and the density

$$g^*(x; \beta, \alpha, \lambda) = \frac{|\alpha| \lambda^\beta}{\Gamma(\beta)} x^{\alpha\beta-1} e^{-\lambda x^\alpha}, \quad x \geq 0, \quad |\alpha| > 0, \quad \beta > 0, \quad \lambda > 0. \quad (25)$$

The density (25) is given in Korolev and Zeifman [36] and Korolev and Gorshenin [37] and summarizes many known densities. Generalized gamma distributions are defined in many different ways, but they do not correspond to the ones that occur above.

Remark 9. The densities $w_{m+1/2}(x; d, r)$ with integer $m = r - (d + 1)/2$ are generalized gamma densities $g^*(x; \beta, \alpha, \lambda)$ given in formula (25) or may be represented as linear combinations of such densities. The parameters $\alpha = 1/2$ and $\lambda = \sqrt{2r}$ apply in all densities $g^*(x; \beta, \alpha, \lambda)$. The parameter β also depends on the number of derivatives $m = r - (d + 1)/2$ in the densities (21).

Example 3. Some linear combinations of generalized gamma densities:

$$\begin{aligned}
w_{1/2}(x; 3, 2) &= g^*(x; 3, 1/2, \sqrt{4}) \\
w_{3/2}(x; 3, 3) &= \frac{3}{4} g^*(x; 4, 1/2, \sqrt{6}) + \frac{1}{4} g^*(x; 3, 1/2, \sqrt{6}) \\
w_{5/2}(x; 3, 4) &= \frac{1}{2} g^*(x; 5, 1/2, \sqrt{8}) + \frac{3}{8} g^*(x; 4, 1/2, \sqrt{8}) + \frac{1}{8} g^*(x; 3, 1/2, \sqrt{8}).
\end{aligned}$$

5. Main Results

Inequalities for approximations to scaled statistics $\mathbb{P}(g_n^\gamma N_n^{\gamma^* - \gamma} T_{N_n} \leq x)$ for $\gamma \in \{0, \pm 1/2, \pm 1\}$ will be presented. Here, $\gamma^* = 1/2$ and $\gamma \in \{0, \pm 1/2\}$ when the statistic T_m is asymptotically normally distributed, or $\gamma^* = 1$ and $\gamma \in \{0, \pm 1\}$ when normalized T_m has chi-squared limit distribution.

5.1. Asymptotically Normal Statistics T_m and Pareto-like Sample Sizes $N_n(s)$

Let asymptotically normal statistic T_m satisfy inequality (7) with coefficients p_k and the rate of convergence $a > 0$. The Pareto-like sample size $N_n = N_n(s)$, $s > 0$, is given in (9), which fulfills the inequality (10). For the scaling factors, select $\gamma^* = 1/2$ and $\gamma \in \{0, \pm 1/2\}$ in formula (18).

Theorem 1. Under the conditions given above, the following approximations apply:

i: Let $\gamma = 1/2$. The non-random scaling factor \sqrt{n} for the statistic $T_{N_n(s)}$ leads to approximations by the Laplace distribution $L_{1/\sqrt{s}}(x)$ with the density $l_{1/\sqrt{s}}(x)$ stated in (19) for $\gamma = 1/2$:

$$\sup_x \left| \mathbb{P}(\sqrt{n} T_{N_n(s)} \leq x) - L_{1/\sqrt{s}, n}(x) \right| \leq C_s n^{-\min\{a, 2\}}$$

where $a > 0$ is the rate of convergence in (7) and

$$\begin{aligned}
L_{1/\sqrt{s}, n}(x) &= L_{1/\sqrt{s}}(x) + l_{1/\sqrt{s}}(x) \left(\frac{\mathbb{I}_{\{a > 1/2\}}(a)}{\sqrt{n}} \left[p_2 x^2 + p_0 \left(\frac{|x|}{\sqrt{2s}} + \frac{1}{2s} \right) \right] \right. \\
&\quad \left. + \frac{\mathbb{I}_{\{a > 1\}}(a)}{n} \left[p_5 x^3 |x| \sqrt{2s} + p_3 x^3 + \left(p_1 + \frac{s-1}{4} \right) x \left(\frac{|x|}{\sqrt{2s}} + \frac{1}{2s} \right) \right] \right).
\end{aligned}$$

ii: Let $\gamma = 0$. The random scaling factor $\sqrt{N_n(s)}$ with $T_{N_n(s)}$ leads to the normal approximation $\Phi(x)$:

$$\sup_x \left| \mathbb{P}(\sqrt{N_n(s)} T_{N_n(s)} \leq x) - \Phi(x) - \varphi_{n,2}(x) \right| \leq C_s n^{-\min\{a, 2\}},$$

where $a > 0$ is the rate of convergence in (7) and

$$\varphi_{n,2}(x) = \varphi(x) \left(\frac{\sqrt{\pi}(p_0 + p_2 x^2)}{2\sqrt{s}n} \mathbb{I}_{\{a > 1/2\}}(a) + \frac{p_1 x + p_3 x^3 + p_5 x^5}{s n} \mathbb{I}_{\{a > 1\}}(a) \right).$$

iii: Let $\gamma = -1/2$. The mixed scaling factor $n^{-1/2} N_n(s)$ at $T_{N_n(s)}$ results in Scaled Student's t -distribution $S_2^*(x; \sqrt{s})$ with density $s_2^*(x; \sqrt{s})$ given in (19) for $\gamma = -1/2$:

$$\sup_x \left| \mathbb{P}(n^{-1/2} N_n(s) T_{N_n(s)} \leq x) - S_{n,2}^*(x) \right| \leq C_s n^{-\min\{a, 2\}},$$

where $a > 0$ is the rate of convergence in (7) and

$$S_{n;2}^*(x; \sqrt{s}) = S_2^*(x; \sqrt{s}) + s_2^*(x; \sqrt{s}) \left(\frac{\mathbb{I}_{\{a>1/2\}}(a)}{\sqrt{n}} \left[p_0 + \frac{3p_2 x^2 \{\ominus > \mathbb{K}\}(a)}{(x^2 + 2s)} \right] \right. \\ \left. + \frac{\mathbb{I}_{\{a>1\}}(a)}{n} \left[\frac{3p_1 x}{x^2 + 2s} + \frac{15p_3 x^3}{(x^2 + 2s)^2} + \frac{105p_5 x^5}{(x^2 + 2s)^3} + \frac{3(s-1)x}{4(x^2 + 2s)} \right] \right).$$

As applications of the Theorem 1, we now examine the Student t -distribution, the Student t -test statistic, and the sample mean as asymptotically normal statistics T_m considered in Christoph and Ulyanov [18] (Section 3.1 and Corollary 1) for the case of negative binomial sample sizes $N_n = N_n(r)$.

Corollary 1. Let the conditions of Theorem 1 be satisfied:

- i: Let $\gamma = 1/2$. In case of the Student's t -statistic $T_m = Z / \sqrt{\chi_m^2}$ with m degrees of freedom estimated in [18] (Formula (18)), inequality (7) is valid with $p_0 = p_2 = p_5 = 0$, $p_1 = p_3 = 1/4$ and $a = 2$. The non-random scaling factor \sqrt{n} and Pareto-like $N_n(s)$ sample sizes lead to:

$$\sup_x \left| \mathbb{P} \left(\frac{\sqrt{n} Z}{\sqrt{\chi_{N_n(s)}^2}} \leq x \right) - L_{1/\sqrt{s}}(x) - \frac{l_{1/\sqrt{s}}(x)}{8n} (2x^3 + x(1 + |x|\sqrt{2s})) \right| \leq C_s n^{-2}$$

- ii: Let $\gamma = 0$. Let $T_m = (\bar{X}_m - \mu) / \hat{\sigma}_m$ be the Student's t -statistic with sample mean \bar{X}_m and sample variance $\hat{\sigma}_m$, which was considered in [18] (Formulas (21) and (20)). The first order approximation (7) with $p_0 = \lambda_3/6$, $p_2 = \lambda_3/3$, $a = 1$, the Pareto-like random sample sizes $N_n(s)$ and the random scaling factor $\sqrt{N_n(s)}$ result in:

$$\sup_x \left| \mathbb{P} \left(\sqrt{N_n(s)} T_{N_n(s)} \leq x \right) - \Phi(x) - \varphi(x) \frac{\sqrt{\pi}(\lambda_3 + 2\lambda_3 x^2)}{12\sqrt{s}n} \right| \leq C_s n^{-1},$$

- iii: Let $\gamma = -1/2$. Considering sample mean $T_m = \bar{X}_m$ estimated in [18] (Formulas (15) and (16)), one has (7) with $p_0 = -p_2 = \lambda_3/6$, $p_1 = \lambda_4/8 - 5\lambda_3^2/24$, $p_3 = -\lambda_4/24 + 5\lambda_3^2/36$, $p_5 = -\lambda_3^2/72$, $a = 3/2$, Pareto-like random sample sizes $N_n(s)$ and mixed scaling factor $n^{-1/2} N_n(s)$, then

$$\sup_x \left| \mathbb{P} \left(n^{-1/2} N_n(s) T_{N_n(s)} \leq x \right) - S_2^*(x; \sqrt{s}) - s_{n;2}^*(x; \sqrt{s}) \right| \leq C_s n^{-3/2},$$

with

$$s_{n;2}^*(x; \sqrt{s}) = s_2^*(x; \sqrt{s}) \left(\frac{1}{\sqrt{n}} \left(\frac{\lambda_3}{6} - \frac{\lambda_3 x^2}{2(x^2 + 2s)} \right) \right. \\ \left. + \frac{1}{n} \left(\frac{(3\lambda_4 - 5\lambda_3^2)x}{8(x^2 + 2s)} - \frac{5(3\lambda_4 - 10\lambda_5)x^3}{24(x^2 + 2s)^2} - \frac{35\lambda_3^2 x^5}{24(x^2 + 2s)^3} + \frac{3(s-1)x}{4(x^2 + 2s)} \right) \right).$$

5.2. Asymptotically Chi-Squared Distributed T_m with Negative Binomially Distributed Sample Sizes $N_n(r)$

Let the asymptotically chi-squared distributed statistics T_m satisfy inequality (8) with coefficients q_1, q_2 and the rate of convergence $a = 2$. The negative binomially distributed sample sizes $N_n = N_n(r)$ with parameter $r > 0$ and success probability $1/n$ are given in (13) and fulfill the inequality (15). For the scaling factors, choose $\gamma^* = 1$ and $\gamma \in \{0, \pm 1\}$ in formula (18).

Theorem 2. Under the conditions given above, the following approximations apply.

- i: Let $\gamma = 1$. The non-random scaling factor $g_n = \mathbb{E}N_n(r) = r(n-1) + 1$ at statistics $T_{N_n(r)}$ leads to approximations by the scaled F-distribution $F^*(x; d, 2r) = F(x/d; d, 2r)$ having parameters $d \in \mathbb{N}_+$ and $r > 0$ and density $f^*(x; d, 2r) = \frac{1}{d} f(\frac{x}{d}; d, 2r)$ given in (20) with $\gamma = 1$:

$$\sup_x \left| \mathbb{P}(g_n T_{N_n(r)} \leq x) - F^*(x; d, 2r) - f_n^*(x; d, 2r) \right| \leq C_r \begin{cases} n^{-\min\{r, 2\}}, & r \neq 2, \\ n^{-2} \ln n, & r = 2, \end{cases}$$

where

$$f_n^*(x; d, 2r) = \frac{f^*(x; d, 2r)}{g_n} \mathbb{I}_{\{r > 1\}}(r) \left(\left(q_1 - \frac{2-r}{2} \right) \frac{x(2r+x)}{2r+d-2} + q_2 x^2 + \frac{x(2-r)}{2} \right). \quad (26)$$

- ii: For $\gamma = 0$ and random scaling factor $N_n(r)$ at $T_{N_n(r)}$, the approximation $G_d(x)$ does not change:

$$\sup_x \left| \mathbb{P}(N_n(r) T_{N_n(r)} \leq x) - G_d(x; n) \right| \leq C_r \begin{cases} n^{-\min\{r, 2\}}, & r \neq 2, \\ n^{-2} \ln n, & r = 2, \end{cases}$$

where

$$G_d(x; n) = G_d(x) + \frac{g_d(x)}{g_n} \mathbb{I}_{\{r > 1\}}(r) (q_1 x + q_2 x^2) \frac{r}{r-1}.$$

- iii: Let $\gamma = -1$ and $r \geq 2$. The mixed scaling factor $g_n^{-1} N_n^2(r)$ at $T_{N_n(r)}$ results in a gamma distribution of generalized type $W_{r-d/2}(x; d, r)$ with density $w_{r-d/2}(x; d, r)$ given in (20) for $\gamma = -1$:

$$\sup_x \left| \mathbb{P}\left(\frac{N_n^2(r)}{g_n} T_{N_n(r)} \leq x\right) - W_{r-d/2; n}(x; d, r) \right| \leq C_r \begin{cases} n^{-2}, & r > 2, \\ n^{-2} \ln n, & r = 2, \end{cases}$$

where

$$\begin{aligned} W_{r-d/2; n}(x; d, r) &= W_{r-d/2}(x; d, r) + \frac{w_{r-d/2}(x; d, r)}{g_n} \mathbb{I}_{\{r > 1\}}(r) \left(2q_2 r x + \frac{(r-2)x}{2} \right. \\ &\quad \left. + \frac{\sqrt{2rx}}{2} (2q_1 + 2q_2(d+2-2r) + 2-r) \frac{K_{r-d/2-1}(\sqrt{2rx})}{K_{r-d/2}(\sqrt{2rx})} \right). \end{aligned}$$

The restriction $r \geq 2$ in Theorem 2(iii) has a purely proof-technical character. In Proposition 4, a result is shown with $r = 3/2$.

Remark 10. The function $R(u; d, r) = \frac{K_{\lambda-1}(u)}{K_{\lambda}(u)}$ can be calculated explicitly for $\lambda = m + 1/2$ with integer $m = r - (d+1)/2$. Then, for example, $R(\sqrt{3x}; 4, 3/2) = 1 + \frac{1}{\sqrt{3x}}$ and $R(\sqrt{4x}; 3, 2) = 1$.

Example 4. Let $\gamma = -1$ in (20), $r = 2$ and $d = 3$. Then, for an asymptotically chi-squared distributed test variable T_m satisfying (8), with scale factor $\frac{N_n^2(2)}{2n-1}$, the estimation holds:

$$\sup_{x>0} \left| \mathbb{P}\left(\frac{N_n^2(2)}{2n-1} T_{N_n(2)} \leq x\right) - W_{1/2}(x; 3, 2) + \frac{w_{1/2}(x; 3, 2)}{4(2n-1)} (\sqrt{4x} (q_2 \sqrt{4x} + q_1 + q_2)) \right| \leq C_2 \frac{\ln n}{n^2},$$

where $W_{1/2}(x; 3, 2)$ and $w_{1/2}(x; 3, 2)$ are specified in (24).

As applications to Theorem 2, we now examine Hotelling's T_0^2 distribution and normalized quotients of two independent chi-square distributions as asymptotic chi-square

distributions, considered in Christoph and Ulyanov [18] (Section 3.2 and Corollary 2) where the sample sizes $N_n = N_n(s)$ had Pareto-like distribution.

Corollary 2. *The conditions of the Theorem 2 shall be fulfilled:*

- i: Let $\gamma = 1$. Consider Hotelling's generalized T_0^2 -statistic $T_0^2 = T_m = \text{tr}(\mathbf{S}_q \mathbf{S}_m^{-1})$ with independently distributed random matrices \mathbf{S}_q and \mathbf{S}_m having Wishart distributions $W_p(q, \mathbf{I}_p)$ and $W_p(m, \mathbf{I}_p)$, respectively. Then, inequality (8) holds with limit distribution $G_d(x)$, $d = pq$, $q_1 = (p + 1 - q)/2$ and $q_2 = (p + 1 + q)/(2d + 4)$. The non-random scaling factor $g_n = \mathbb{E}N_n(r)$ by $T_{N_n(r)}$ leads to

$$\sup_x \left| \mathbb{P}(g_n T_{N_n(r)} \leq x) - F^*(x; pq, 2r) - f_n^*(x; pq, 2r) \right| \leq C_r \begin{cases} n^{-\min\{r, 2\}}, & r \neq 2, \\ n^{-2} \ln n, & r = 2, \end{cases} \quad (27)$$

where the scaled F-distribution $F^*(x; pq, 2r)$ with density $f^*(x; pq, 2r)$ is given in (20) for $\gamma = 1$

$$f_n^*(x; pq, 2r) = \frac{f^*(x; pq, 2r)}{g_n} \mathbb{I}_{\{r > 1\}}(r) \left(\left(\frac{p+1-q}{2} - \frac{2-r}{2} \right) \frac{x(2r+x)}{2r+pq-2} + \frac{(p+1+q)x^2}{(2pq+4)} + \frac{x(2-r)}{2} \right). \quad (28)$$

- ii: Let $\gamma = 0$, χ_d^2 and χ_m^2 be independent and $T_m = \chi_d^2 / \chi_m^2$ be scale mixtures satisfying inequality (8) with coefficients $q_1 = (d - 2)/2$ and $q_2 = -1/2$. Random degrees of freedom $N_n(r)$ instead of m and random scaling factor $N_n(r)$ lead to

$$\sup_{x > 0} \left| \mathbb{P}(N_n(r) T_{N_n(r)} \leq x) - G_d(x; n) \right| \leq C_r \begin{cases} n^{-\min\{r, 2\}}, & r \neq 2, \\ n^{-2} \ln n, & r = 2, \end{cases}$$

where

$$G_d(x; n) = G_d(x) + \frac{g_d(x)}{2g_n} \mathbb{I}_{\{r > 1\}}(r) ((d-2)x - x^2) \frac{r}{r-1}.$$

- iii: Let $\gamma = -1$. The statistics $T_m = \chi_4^2 / \chi_m^2$ satisfy the inequality (8) with the limiting distribution $G_4(x)$ and the coefficients $q_1 = 1$ and $q_2 = -1/2$. The mixed scaling factor $g_n^{-1} N_n^2(r)$ at $T_{N_n(r)}$ results in a limiting gamma distribution of generalized type $W_{r-d/2}(x; d, r)$. Only if $r - (d + 1)/2 = m$ is an integer, the involved Macdonald functions $K_{r-d/2}(\sqrt{2rx})$ may be explicitly calculated. Since $d = 4$, we choose $r = 5/2$ and find $r - (d + 1)/2 = 0$. Then, uniformly in $x > 0$:

$$\left| \mathbb{P} \left(\frac{N_n^2(5/2)}{(5n-3)/2} \frac{\chi_4^2}{\chi_{N_n(5/2)}^2} \leq x \right) - W_{1/2}(x; 4, 5/2) + \frac{w_{1/2}(x; 4, 5/2)}{2(5n-3)} (9x - \sqrt{5}x) \right| \leq \frac{C_{3/2}}{n^{3/2}},$$

where $W_{1/2}(x; 4, 5/2)$ and $w_{1/2}(x; 4, 5/2)$ are specified in (23).

Remark 11. In the paper Monahkov [38], an analogous to (27) estimation is shown, but with 11 approximation terms in corresponding formula (28). Instead of (8) with $q_1 = (p + 1 - q)/2$, $q_2 = (p + 1 + q)/(2d + 4)$ and $d = pq$, the following equivalent inequality is used; see Fujikoshi et al. [39] (Theorem 4.1(ii)):

$$\sup_x \left| \mathbb{P}(m \text{tr}(\mathbf{S}_q \mathbf{S}_m^{-1}) \leq x) - G_d(x) - \frac{d}{4m} (a_0 G_d(x) + a_1 G_{d+2}(x) + a_2 G_{d+4}(x)) \right| \leq \frac{C}{m^2}$$

where $a_0 = q - p - 1$, $a_1 = -2q$, $a_2 = q + p + 1$ with $a_0 + a_1 + a_2 = 0$ and $d = pq$.

Proposition 4. Let $\gamma = -1$. Consider the statistics $T_m = \chi_4^2 / \chi_m^2$, satisfying the inequality (8) with the limiting distribution $G_4(x)$, the coefficients $q_1 = 1$ and $q_2 = -1/2$ and the mixed scaling factor $g_n^{-1} N_n^2(r)$ at $T_{N_n(r)}$. If $r = 3/2$ and $d = 4$, then $r - (d + 1)/2 = -1$, $g_n = (3n - 1)/2$ and, uniformly in $x > 0$:

$$\left| \mathbb{P} \left(\frac{N_n^2(3/2)}{(3n-1)/2} \frac{\chi_4^2}{\chi_{N_n(3/2)}^2} \leq x \right) - W_{-1/2}(x; 4, 3/2) + \frac{w_{-1/2}(x; 4, 3/2)}{2(3n-1)} (7x + \sqrt{3x} + 1) \right| \leq \frac{C_{3/2}}{n^{3/2}},$$

where $W_{-1/2}(x; 4, 3/2)$ and $w_{-1/2}(x; 4, 3/2)$ are specified in (22).

6. Proofs

For the proofs of Theorems 1 and 2, we use Proposition 1. The statistics T_m and the sample size N_n are either asymptotically normally and discretely Pareto-like distributed (i.e., $F = \Phi$ and $H = W_s$) or asymptotically chi-squared and negatively binomially distributed (i.e., $F = G_d$ and $H = G_{r,r}$). In both cases, the size D_n defined in (6) is uniformly bounded for all $n \in N_+$, see Christoph and Ulyanov [18] (Lemma A1). Next, the bounds that are required in (4) for the negative moments of sample sizes $\mathbb{E}N_n(s)^{-a}$ and $\mathbb{E}N_n(r)^{-a}$ are provided by (12) and (17). Furthermore, it follows from Christoph and Ulyanov [18] (Proposition 2 and Lemma A2) that in both cases the domain of integration of the integrals in the function $G_n(x, 1/g_n)$ defined in (5) can be extended from $(1/g_n, \infty)$ to $(0, \infty)$:

$$\sup_x |G_n(x, 1/g_n) - G_{n,2}(x)| \leq C g_n^{-b},$$

where $b = 2$ if $F = \Phi$ and $H = W_s$ or $b = \min\{r, 2\}$ if $F = G_d$ and $H = G_{r,r}$, respectively, and

$$G_{n,2}(x) = \begin{cases} \int_0^\infty F(xy^\gamma) dH(y), & \text{for } 0 < b \leq 1/2, \\ \int_0^\infty \left(F(xy^\gamma) + \frac{f_1(xy^\gamma)}{\sqrt{g_n y}} \right) dH(y) =: G_{n,1}(x), & \text{for } 1/2 < b \leq 1, \\ G_{n,1}(x) + \int_0^\infty \frac{f_2(xy^\gamma)}{g_n y} dH(y) + \int_0^\infty \frac{F(xy^\gamma)}{n} dh_2(y), & \text{for } b > 1, \end{cases} \quad (29)$$

We still have to calculate the integrals in (29) that contain f_1 , f_2 , and h_2 , respectively.

Proof of Theorem 1. We now consider $F = \Phi$, $H = H_s$ and $\gamma \in \{0; \pm 1/2\}$. Here, $f_1(xy^\gamma) = (p_0 + p_2 x^2 y^{2\gamma}) \varphi(xy^\gamma)$, $f_2(xy^\gamma) = (p_1 x y^\gamma + p_3 x^3 y^{3\gamma} + p_5 x^5 y^{5\gamma}) \varphi(xy^\gamma)$ and we divide the function $h_2(y) = h_{2,s}(y)$ given in (11) into two parts: $h_{2,s}^*(y) = s(s-1)e^{-s/y}/(2y^2)$ and $h_{2,s}^{**}(y) = sQ_1(ny)y^{-2}e^{-s/y}$. The densities of the limit distributions $V_\gamma(x; d, r) = \int_0^\infty \Phi(xy^\gamma) dW_s(y)$ were given in (20). If $\gamma = 1/2$ to calculate the integrals in (29) involving $f_1(x\sqrt{y})$, $f_2(x\sqrt{y})$ and $h_{2,s}^*(y)$ we use Prudnikov et al. [35] (Formulas 2.3.16.2 and 2.3.16.3):

$$\int_0^\infty y^{-m-1/2} e^{-py-q/y} dy = \begin{cases} (-1)^{-m} \sqrt{\pi} \frac{\partial^{-m}}{\partial p^{-m}} (p^{-1/2} e^{-2\sqrt{pq}}), & m = 0, -1, -2, \dots \\ (-1)^m \frac{\sqrt{\pi}}{\sqrt{p}} \frac{\partial^m}{\partial q^m} (e^{-2\sqrt{pq}}), & m = 0, 1, 2, \dots \end{cases}, \quad p, q > 0, \quad (30)$$

for $p = x^2/2 > 0$, $q = s > 0$ and $m = 0, 1, 2$, respectively. The corresponding integral with $h_{2,s}^{**}(y)$ was estimated in Christoph et al. [17] (see Proof of Theorem 5) by $c(s) e^{-\sqrt{\pi s n}/2} \leq C(s) n^{-2}$.

In case of $\gamma = 0$, we obtain $\int_0^\infty \Phi(x) dh_2(y) = \Phi(x)(h_2(\infty) - \lim_{y \rightarrow 0} h_2(y)) = 0$. To calculate the integrals with $f_1(x)$ and $f_2(x)$ we use [35] (Formula 2.3.3.1) with $\alpha = 3/2, 2$ and $q = s$:

$$\int_0^\infty y^{-\alpha-1} e^{-q/y} dy \stackrel{1/y=z}{=} \int_0^\infty z^{\alpha-1} e^{-qz} dz = \Gamma(\alpha) q^{-\alpha}, \quad \alpha > 0, \quad q > 0. \quad (31)$$

If $\gamma = -1/2$, the integrals with $f_1(x/\sqrt{y})$, $f_2(x/\sqrt{y})$ and $h_{2,s}^*(x/\sqrt{y})$ are calculated using (31) with $\alpha = 3/2, 5/2, 7/2, 9/2$ and $q = s + x^2/2$. From Christoph and Ulyanov [20]

(see Proof of Theorem 8), it follows that holds: $n^{-1} \sup_x \left| \int_0^\infty \Phi(x/\sqrt{y}) dh_{2,s}^{**}(y) \right| \leq C(s)n^{-2}$ and Theorem 1 is proved. \square

Proof of Theorem 2. Now, we consider the case $F(x) = G_d(x)$, $H(y) = G_{r,r}(y)$ and $\gamma \in \{0; \pm 1\}$. This combination has not yet been studied in the literature. Only if $\gamma = 1$, there is a result by Monahkov [38]; see Remark 11 above. Then, $f_1(xy^\gamma) = 0$, $f_2(xy^\gamma) = (q_1xy^\gamma + q_2x^2y^{2\gamma})g_d(xy^\gamma)$ and we divide the function $h_2(y) = h_{2,r}(y)$ given in (16) into two parts: $h_{2,r}^*(y) = (2r)^{-1}g_{r,r}(y)(y-1)(2-r)$ and $h_{2,r}^{**}(y) = r^{-1}g_{r,r}(y)Q_1(g_ny)$.

For $\gamma = 1$, the density $v_1(x; d, r)$ in (20) and the integrals in (29) with $f_2(xy)$ and $h_{2,r}^*(y)$ are computed with (31) for $\alpha = r + d/2$, $r + d/2 - 1$. The integral with $h_{2,r}^{**}(y)$ is estimated in (A1) in Lemma A1. Together with the inequality $|1/g_n - 1/(rn)| \leq \max\{2, r\}(r-1)(rn)^{-2}$, we get (26).

In case of $\gamma = 0$, we obtain $\int_0^\infty G_d(x)dh_2(y) = G_d(x)(h_{2,r}(\infty) - \lim_{y \rightarrow 0} h_{2,r}(y)) = 0$. To calculate the integrals with $f_2(x)$, we use (31) with $\alpha = r - 1$ and $q = r$.

If $\gamma = -1$ the density $v_{-1}(x; d, r)$ in (20) and the integrals with $f_2(x/y)$ and $h_{2,r}^*(y)$ are calculated using Prudnikov et al. [35] (Formula 2.3.16.1):

$$\int_0^\infty y^{\alpha-1} e^{-py-q/y} dy = 2(p/q)^{\alpha/2} K_\alpha(2\sqrt{pq}), \quad p, q > 0,$$

with $\alpha = r - d/2$, $r - d/2 - 1$, $r - d/2 - 2$, $p = r$ and $q = x/2$. We use the order-reflection formula $K_\alpha(u) = K_{-\alpha}(u)$ and the recursion formula; see Oldham et al. [34] (Chapter 51.5):

$$K_{r-d/2-2}(\sqrt{2rx}) = K_{d/2+2-r}(\sqrt{2rx}) = \frac{2(d/2-r+1)}{\sqrt{2rx}} K_{d/2-r+1}(\sqrt{2rx}) + K_{d/2-r}(\sqrt{2rx}).$$

The integral with $h_{2,r}^{**}(y)$ is estimated in (A4) in Lemma A2 and Theorem 2 is proved. \square

Proof of Proposition 4. We consider $\gamma = -1$, $r = 3/2$, $d = 4$ and $g_n = (3n-1)/2$. The integrals in (29) with $f_2(x/y)$ and $h_{2,r}^*(y)$ are calculated using (30) with $m = -1, -2, -3$, $p = r$ and $q = x/2$. The integral with $h_{2,r}^{**}$ is estimated in (A5) in Lemma A3 and Proposition 4 is proved. \square

7. Conclusions

The common goal of the present work and that of Christoph and Ulyanov [18] is to develop formal second order Chebyshev–Edgeworth expansions for sample statistics with random sample sizes. Corresponding expansions are assumed for the statistics with non-random sample sizes as well as for the random sample sizes. The statistics examined are asymptotically normally distributed and, for the first time in this setting, also asymptotically chi-squared distributed. The random sample sizes have negative binomial or Pareto-like distributions. The formal construction of the approximating functions allows the results to be used for a whole family of asymptotically normal or chi-squared distributed statistics. The Student t -distribution with m degrees of freedom, the one-sample Student t -test statistic, and the sample mean are considered as examples of asymptotic normal statistics. Hotelling's generalized T_0^2 statistic and scale mixture of a normalized quotient of two independent chi-squared random variables were studied as examples of the asymptotic chi-squared distributions. In addition, random, non-random, and mixed scaling factors for the statistics are considered, which have a significant influence on the limit distributions. The limit laws are scale mixtures of the normal with mixing gamma or chi-squared with mixing inverse exponential distributions. In addition to the normal distribution and the chi-square distribution, there are a variety of limit distributions: the Laplace, the scaled Student t -, the scaled Fisher, the generalized gamma, and linear combinations of generalized gamma distributions.

The remaining terms in the approximations of the scaled statistics are estimated by inequalities.

Author Contributions: Conceptualization, G.C. and V.V.U.; methodology, V.V.U. and G.C.; formal analysis, G.C. and V.V.U.; investigation, G.C. and V.V.U.; writing—original draft, G.C. and V.V.U.; writing—review and editing, V.V.U. and G.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding. It was carried out within the project “Analysis of the quality of approximations in the statistical analysis of multivariate observations” of the Magdeburg University, the program of the Moscow Center for Fundamental and Applied Mathematics, Lomonosov Moscow State University, and HSE University Basic Research Programs.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors thank the Editor for his support and the Reviewers for their appropriate comments which have improved the quality of this paper.

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

Appendix A. Auxiliary Lemmas

Lemma A1. Let $r > 1$ then

$$|J_1(x)| = \left| \int_0^\infty G_d(xy) dh_{2,r}^{**}(y) \right| \leq \frac{c(r, d)}{g_n^{r-1}} \quad \text{with} \quad h_{2,r}^{**}(y) = r^{-1} g_{r,r}(y) Q_1(g_n y). \quad (\text{A1})$$

Proof of Lemma A1. We use the Fourier series expansion of the jump correcting function $Q_1(y)$ at all non-integer points y ; see Prudnikov et al. [35] (Formula 5.4.2.9 for $a = 0$):

$$Q_1(y) = \frac{1}{2} - (y - [y]) = \sum_{k=1}^\infty \frac{\sin(2\pi k y)}{k\pi}, \quad y \neq [y], \quad (\text{A2})$$

and Prudnikov et al. [35] (Formula 2.5.31.4):

$$\int_0^\infty y^{\alpha-1} e^{-py} \sin(by) dy = \frac{\Gamma(\alpha)}{(b^2 + p^2)^{\alpha/2}} \sin(\alpha \arctan(b/p)) \quad \text{with} \quad \alpha > -1, \quad b, p > 0. \quad (\text{A3})$$

Integration by parts in the integral $J_1(x)$, using (A2), interchanging sum and integral and applying (A3) with $\alpha = r + d/2 - 1$, $p = (r + x/2)$ and $b = 2\pi k g_n$ leads to

$$\begin{aligned} J_1(x) &= -\frac{r^{r-1} x^{d/2}}{\Gamma(r) 2^{d/2} \Gamma(d/2)} \int_0^\infty y^{r+d/2-2} Q_1(g_n y) e^{-(r+x/2)y} dy \\ &= -\frac{r^{r-1} x^{d/2}}{\pi \Gamma(r) 2^{d/2} \Gamma(d/2)} \sum_{k=1}^\infty \frac{1}{k} \int_0^\infty y^{r+d/2-2} e^{-(r+x/2)y} \sin(2\pi k g_n y) dy \\ &= -\frac{r^{r-1} \Gamma(r + d/2 - 1)}{\pi \Gamma(r) 2^{d/2} \Gamma(d/2)} \sum_{k=1}^\infty \frac{a_k(x; n)}{k} \end{aligned}$$

with

$$a_k(x; n) = \frac{x^{d/2} \sin\left((r + d/2 - 1) \arctan(2\pi k g_n / (r + x/2))\right)}{\left((2\pi k g_n)^2 + (r + x/2)^2\right)^{(r+d/2-1)/2}}.$$

Now, we split the exponent $(r + d/2 - 1)/2 = (r - 1)/2 + d/4$ and obtain

$$|a_k(x; n)| \leq \frac{x^{d/2}}{(2\pi k g_n)^{r-1} (r + x/2)^{d/2}} \leq \frac{2^{d/2}}{(2\pi k g_n)^{r-1}}.$$

Since $r > 1$, we find uniform in $x \geq 0$

$$|J_1(x)| \leq \frac{c_1(r, d)}{g_n^{r-1}} \sum_{k=1}^{\infty} k^{-r} = \frac{c(r, d)}{g_n^{r-1}}$$

and Lemma A1 is proved. \square

Lemma A2. Let $r \geq 2$, then

$$|J_{-1}(x)| = \left| \int_0^{\infty} G_d(x/y) dh_{2,r}^{**}(y) \right| \leq \frac{c(r, d)}{g_n} \quad \text{with} \quad h_{2,r}^{**}(y) = r^{-1} g_{r,r}(y) Q_1(g_n y). \quad (\text{A4})$$

Proof of Lemma A2. Integration by parts in the integral $J_{-1}(x)$, using the Fourier series expansion (A2), interchanging sum and integral, we find

$$J_{-1}(x) = \frac{r^{r-1} x^{d/2}}{\Gamma(r) 2^{d/2} \Gamma(d/2)} \int_0^{\infty} y^{r-d/2-2} Q_1(g_n y) e^{-(ry+x/(2y))} dy = \frac{r^{r-1}}{\pi \Gamma(r) 2^{d/2} \Gamma(d/2)} \sum_{k=1}^{\infty} \frac{J_{k,n}(x)}{k}$$

with $J_{k,n}(x) = \int_0^{\infty} x^{d/2} y^{r-d/2-2} e^{-(ry+x/(2y))} \sin(2\pi k g_n y) dy$.

In the literature, we have only found integrals $J_{k,n}(x)$ with power functions $y^{-1/2}$ and $y^{-3/2}$. Therefore, we integrate by parts in the integral $J_{k,n}(x)$:

$$J_{k,n}(x) = \frac{-1}{2} \int_0^{\infty} \left((d-2r+4)f_1(x, y) + 2rf_2(x, y) - f_3(x, y) \right) e^{-(ry+x/(2y))} \frac{\cos(2\pi k g_n y)}{2\pi k g_n} dy,$$

where $f_1(x, y) = x^{d/2} y^{r-d/2-3}$, $f_2(x, y) = x^{d/2} y^{r-d/2-2}$ and $f_3(x, y) = x^{d/2+1} y^{r-d/2-4}$.

Since $r \geq 2$ and $d \geq 1$ we obtain $y^{r-2} e^{-ry/2} \leq c_r$ and $(x/y)^{(d-1)/2} e^{-x/(4y)} \leq c_d$. Using (30) with $m = 0, 1, 2$, $p = r/2$, and $q = x/4$ we find

$$\int_0^{\infty} f_1(x, y) dy \leq c_r c_d x^{1/2} \int_0^{\infty} y^{-3/2} e^{-(ry/2+x/(4y))} dy = c_r c_d 2\sqrt{\pi} e^{-\sqrt{rx}/2} \leq C_1(r, d),$$

$$\int_0^{\infty} f_2(x, y) dy \leq c_r c_d x^{1/2} \int_0^{\infty} \frac{y^{-1/2}}{e^{(ry/2+x/(4y))}} dy = c_r c_d \sqrt{2\pi x/r} e^{-\sqrt{2rx}/2} \leq C_2(r, d),$$

$$\int_0^{\infty} f_3(x, y) dy \leq c_r c_d x^{3/2} \int_0^{\infty} \frac{y^{-5/2}}{e^{(ry/2+x/(4y))}} dy = c_r c_d 2\sqrt{\pi} (\sqrt{2rx} + 2) e^{-\sqrt{rx}/2} \leq C_3(r, d)$$

and

$$|J_{k,n}| \leq \frac{1}{4\pi k g_n} (|d-2r+4| C_1(r, d) + 2r C_2(r, d) + C_3(r, d)) \leq \frac{C^*(r, d)}{k g_n}.$$

Hence,

$$|J_{-1}(x)| \leq \frac{r^{r-1}}{\pi \Gamma(r) 2^{d/2} \Gamma(d/2)} \frac{\pi^2}{6 g_n} C^*(r, d) \leq \frac{c(r, d)}{g_n}.$$

Lemma A2 is proved. \square

Lemma A3. Let $\gamma = -1$, $r = 3/2$, $d = 4$ and $g_n = (3n-1)/2$, then

$$|J_{-1}^*(x)| = \left| \int_0^{\infty} G_d(x/y) dh_{2,3/2}^{**}(y) \right| \leq \frac{c(3/2, 4)}{\sqrt{g_n}} \quad \text{with} \quad h_{2,3/2}^{**}(y) = (2/3) g_{3/2,3/2}(y) Q_1(g_n y). \quad (\text{A5})$$

Proof of Lemma A3. Integration by parts in the integral $J_{-1}^*(x)$, using the Fourier series expansion (A2), interchanging sum and integral, we find

$$J_{-1}^*(x) = \frac{\sqrt{3/2} x^2}{4 \Gamma(3/2)} \int_0^{\infty} y^{-5/2} Q_1(g_n y) e^{-(3y/2+x/(2y))} dy = \frac{\sqrt{3/2}}{\sqrt{\pi}} \sum_{k=1}^{\infty} \frac{J_{k,n}^*(x)}{k}$$

with

$$J_{k,n}^*(x) = x^2 \int_0^\infty y^{-5/2} e^{-(3y/2+x/(2y))} \sin(2\pi k g_n y) dy.$$

Using Prudnikov et al. [35] (Formula 2.5.37.3), with the real constants $p > 0$, $q > 0$ and $b > 0$, we obtain

$$\int_0^\infty y^{-3/2} e^{-p y - q/y} \sin(b y) dy = \frac{\sqrt{\pi}}{\sqrt{q}} e^{-2\sqrt{q} z_+} \sin(2\sqrt{q} z_-) \quad \text{and} \quad 2z_\pm^2 = \sqrt{p^2 + b^2} \pm p. \quad (\text{A6})$$

It was shown in Christoph et al. [17] (Proof of Theorem 5) that Leibniz's integral rule allows differentiation to q under the integral sign in (A6). Therefore,

$$\begin{aligned} \int_0^\infty y^{-5/2} e^{-p y - q/y} \sin(b y) dy &= (\sqrt{\pi}/2) e^{-2\sqrt{q} z_+} \left(q^{-3/2} \sin(2\sqrt{q} z_-) \right. \\ &\quad \left. + 2q^{-1} z_+ \sin(2\sqrt{q} z_-) - 2q^{-1} z_- \cos(2\sqrt{q} z_-) \right). \end{aligned}$$

Since $0 < z_- \leq z_+$, $p = 3/2$, $q = x/2$, $b = 2\pi k g_n$, $k \geq 1$ and $g_n \geq 1$ we find $z_+ \geq \sqrt{\pi k g_n}$,

$$|J_{k,n}^*(x)| \leq x^2 \frac{\sqrt{\pi}}{2} e^{-\sqrt{2} x z_+} \left(\frac{2\sqrt{2}}{x^{3/2}} + \frac{8}{x} z_+ \right) = \frac{\sqrt{\pi}}{z_+} e^{-\sqrt{2} x z_+} \left(\sqrt{2} x z_+ + 4 x z_+^2 \right) \leq \frac{e^{-1} + 8e^{-2}}{\sqrt{k n}}$$

and

$$|J_{-1}(x)| \leq \frac{\sqrt{3/2}}{\sqrt{\pi}} \sum_{k=1}^\infty \frac{e^{-1} + 8e^{-2}}{k^{3/2} \sqrt{g_n}}.$$

Lemma A3 is proved. \square

References

- Wallace, D.L. Asymptotic approximations to distributions. *Ann. Math. Statist.* **1958**, *29*, 635–654. [\[CrossRef\]](#)
- Bickel, P.J. Edgeworth expansions in nonparametric statistics. *Ann. Statist.* **1974**, *2*, 1–20. [\[CrossRef\]](#)
- Hall, P. *The Bootstrap and Edgeworth Expansion*; Springer Series in Statistics; Springer: New York, NY, USA, 1992.
- Bhattacharya, R.N.; Ranga Rao, R. *Normal Approximation and Asymptotic Expansions*; Wiley: New York, NY, USA, 1976.
- Petrov, V.V. *Limit Theorems of Probability Theory, Sequences of Independent Random Variables*; Clarendon Press: Oxford, UK, 1995.
- Pfanzagl, J. Asymptotic expansions related to minimum contrast estimators. *Ann. Statist.* **1973**, *1*, 993–1026. [\[CrossRef\]](#)
- Bentkus, V.; Götze, F.; van Zwet, W.R. An Edgeworth expansion for symmetric statistics. *Ann. Statist.* **1997**, *25*, 851–896. [\[CrossRef\]](#)
- Burnashev, M.V. Asymptotic expansions for median estimate of a parameter. *Theory Probab. Appl.* **1997**, *41*, 632–645. [\[CrossRef\]](#)
- Nunes, C.; Capistrano, G.; Ferreira, D.; Ferreira, S.S.; Mexia, J.T. Exact critical values for one-way fixed effects models with random sample sizes. *J. Comput. Appl. Math.* **2019**, *354*, 112–122. [\[CrossRef\]](#)
- Nunes, C.; Capistrano, G.; Ferreira, D.; Ferreira, S.S.; Mexia, J.T. Random sample sizes in orthogonal mixed models with stability. *Comp. Math. Methods* **2019**, *1*, e1050. [\[CrossRef\]](#)
- Nunes, C.; Mário, A.; Ferreira, D.; Moreira, E.M.; Ferreira, S.S.; Mexia, J.T. An algorithm for simulation in mixed models with crossed factors considering the sample sizes as random. *J. Comput. Appl. Math.* **2022**, *404*, 113463. [\[CrossRef\]](#)
- Esquivel, M.L.; Mota P.P.; Mexia J.T. On some statistical models with a random number of observations. *J. Stat. Theory Pract.* **2016**, *10*, 805–823. [\[CrossRef\]](#)
- Döbler, C. New Berry-Esseen and Wasserstein bounds in the CLT for non-randomly centered random sums by probabilistic methods. *ALEA Lat. Am. J. Probab. Math. Stat.* **2015**, *12*, 863–902.
- Schluter, C.; Trede, M. Weak convergence to the Student and Laplace distributions. *J. Appl. Probab.* **2016**, *53*, 121–129. [\[CrossRef\]](#)
- Bening, V.E.; Galieva, N.K.; Korolev, V.Y. On rate of convergence in distribution of asymptotically normal statistics based on samples of random size. *Ann. Math. Inform.* **2012**, *39*, 17–28.
- Bening, V.E.; Galieva, N.K.; Korolev, V.Y. Asymptotic expansions for the distribution functions of statistics constructed from samples with random sizes. *Inform. Appl.* **2013**, *7*, 75–83. (In Russian)
- Christoph, G.; Monakhov, M.M.; Ulyanov, V.V. Second-order Chebyshev-Edgeworth and Cornish-Fisher expansions for distributions of statistics constructed with respect to samples of random size. *J. Math. Sci.* **2020**, *244*, 811–839; Translated from *Zap. Nauchnykh Semin. POMI* **2017**, *466*, 167–207. [\[CrossRef\]](#)
- Christoph, G.; Ulyanov, V.V. Chebyshev-Edgeworth-type approximations for statistics based on samples with random sizes. *Mathematics* **2021**, *9*, 775. [\[CrossRef\]](#)
- Fujikoshi, Y.; Ulyanov, V.V. *Non-Asymptotic Analysis of Approximations for Multivariate Statistics*; Springer: Singapore, 2020.

20. Christoph, G.; Ulyanov, V.V. Second order expansions for high-dimension low-sample-size data statistics in random setting. *Mathematics* **2020**, *8*, 1151; Reprinted in Special Issue: *Stability Problems for Stochastic Models: Theory and Applications*; Zeifman, A., Korolev, V., Sipin, A., Eds.; MPDI: Basel, Switzerland, 2021; pp. 259–286. [\[CrossRef\]](#)
21. Buddana, A.; Kozubowski, T.J. Discrete Pareto distributions. *Econ. Qual. Control.* **2014**, *29*, 143–156. [\[CrossRef\]](#)
22. Bening, V.E.; Korolev, V.Y. On the use of Student’s distribution in problems of probability theory and mathematical statistics. *Theory Probab. Appl.* **2005**, *49*, 377–391. [\[CrossRef\]](#)
23. Gavrilenko, S.V.; Zubov, V.N.; Korolev, V.Y. The rate of convergence of the distributions of regular statistics constructed from samples with negatively binomially distributed random sizes to the Student distribution. *J. Math. Sci.* **2017**, *220*, 701–713. [\[CrossRef\]](#)
24. Christoph, G.; Ulyanov, V.V.; Bening, V.E. Second order expansions for sample median with random sample size. *ALEA Lat. Am. J. Probab. Math. Stat.* **2022**, *19*, 339–365. [\[CrossRef\]](#)
25. Korolev, V. Bounds for the rate of convergence in the generalized Rényi theorem. *Mathematics* **2022**, *10*, 4252. [\[CrossRef\]](#)
26. Bulinski, A.; Slepov, N. Sharp estimates for proximity of geometric and related sums distributions to limit laws. *Mathematics* **2022**, *10*, 4747. [\[CrossRef\]](#)
27. Korolev, V.; Shevtsova, I. An improvement of the Berry-Esseen inequality with applications to Poisson and mixed Poisson random sums. *Scand. Actuar. J.* **2012**, *2012*, 81–105. [\[CrossRef\]](#)
28. Sunklodas, J.K. On the normal approximation of a binomial random sum. *Lith. Math. J.* **2014**, *54*, 356–365. [\[CrossRef\]](#)
29. Petrov, V.V. *Sums of Independent Random Variables*; Akademie-Verlag: Berlin, Germany, 1975.
30. Kolassa, J.E.; McCullagh, P. Edgeworth series for lattice distributions. *Ann. Statist.* **1990**, *18*, 981–985. [\[CrossRef\]](#)
31. Choy, T.B.; Chan, J.E. Scale mixtures distributions in statistical modelling. *Aust. N. Z. J. Stat.* **2008**, *50*, 135–146. [\[CrossRef\]](#)
32. Fujikoshi, Y.; Ulyanov, V.V.; Shimizu, R. *Multivariate Statistics. High-Dimensional and Large-Sample Approximations*; Wiley Series in Probability and Statistics; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2010.
33. Christoph, G.; Ulyanov, V.V. Random dimension low sample size asymptotics. In *Recent Developments in Stochastic Methods and Applications*; Shiryayev, A.N., Samouylov, K.E., Kozyrev, D.V., Eds.; Springer Proceedings in Mathematics & Statistics; Springer International Publishing: Cham, Switzerland, 2021; Volume 371, pp. 215–228, ISBN 978-3-030-83266-7 and 978-3-030-83266-0. [\[CrossRef\]](#)
34. Oldham K.B.; Myland, J.C.; Spanier, J. *An Atlas of Functions*, 2nd ed.; Springer Science + Business Media: New York, NY, USA, 2009.
35. Prudnikov, A.P.; Brychkov, Y.A.; Marichev, O.I. *Integrals and Series, Vol. 1: Elementary Functions*, 3rd ed.; Gordon & Breach Science Publishers: New York, NY, USA, 1992.
36. 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. [\[CrossRef\]](#)
37. Korolev, V.Y.; Gorshenin, A. Probability models and statistical tests for extreme precipitation based on generalized negative binomial distributions. *Mathematics* **2020**, *8*, 604. [\[CrossRef\]](#)
38. Monakhov, M.M. Chebyshev–Edgeworth expansions for distributions of generalized Hotelling-type statistics based on random size samples. *Inform. Primen. [Informatics Its Appl.]* **2021**, *15*, 72–81. (In Russian) [\[CrossRef\]](#)
39. Fujikoshi, Y.; Ulyanov, V.V.; Shimizu, R. L_1 -norm error bounds for asymptotic expansions of multivariate scale mixtures and their applications to Hotelling’s generalized T_0^2 . *J. Multivar. Anal.* **2005**, *96*, 1–19. [\[CrossRef\]](#)

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.