mathematics

MDPI

Article

Comparing Distributions of Sums of Random Variables by Deficiency: Discrete Case

Vladimir E. Bening [1,2] and Victor Y. Korolev [1,2,3,*]

[1] Faculty of Computational Mathematics and Cybernetics, Moscow State University, 119991 Moscow, Russia; bening@cs.msu.ru or bening@yandex.ru
[2] Moscow Center for Fundamental and Applied Mathematics, 119991 Moscow, Russia
[3] Federal Research Center "Computer Science and Control", Russian Academy of Sciences, 119333 Moscow, Russia
* Correspondence: vkorolev@cs.msu.ru

Abstract: In the paper, we consider a new approach to the comparison of the distributions of sums of random variables. Unlike preceding works, for this purpose we use the notion of deficiency that is well known in mathematical statistics. This approach is used, first, to determine the distribution of a separate random variable in the sum that provides the least possible number of summands guaranteeing the prescribed value of the $(1-\alpha)$-quantile of the normalized sum for a given $\alpha \in (0,1)$, and second, to determine the distribution of a separate random variable in the sum that provides the least possible number of summands guaranteeing the prescribed value of the probability for the normalized sum to fall into a given interval. Both problems are solved under the condition that possible distributions of random summands possess coinciding three first moments. In both settings the best distribution delivers the smallest number of summands. Along with distributions of a non-random number of summands, we consider the case of random summation and introduce an analog of deficiency which can be used to compare the distributions of sums with random and non-random number of summands. The main mathematical tools used in the paper are asymptotic expansions for the distributions of $\mathbb{R}$-valued functions of random vectors, in particular, normalized sums of independent identically distributed r.v.s and their quantiles. Along with the general case, main attention is paid to the situation where the summarized random variables are independent and identically distributed. The approach under consideration is applied to determination of the distribution of insurance payments providing the least insurance portfolio size under prescribed Value-at-Risk or non-ruin probability.

Keywords: limit theorem; sum of independent random variables; random sum; asymptotic expansion; asymptotic deficiency; kurtosis

Citation: Bening, V.E.; Korolev, V.Y. Comparing Distributions of Sums of Random Variables by Deficiency: Discrete Case. *Mathematics* **2022**, *10*, 454. https://doi.org/10.3390/math10030454

Academic Editors: Anatoliy Swishchuk and Antonella Basso

Received: 27 December 2021
Accepted: 28 January 2022
Published: 30 January 2022

1. Introduction

1.1. The Problem under Consideration and the Structure of the Paper

The problem considered in the paper is very close to the problem of stochastic ordering and even may be considered as a a version of this problem. In probability theory and statistics, a stochastic order quantifies the concept of one random variable being "bigger" or "smaller" than another. Many different orders exist, which have different applications, see, e.g., the book [1]. Here we propose an approach to establishing stochastic order for the distributions of sums of independent random variables (r.v.s) based on the notion of deficiency that is well known in asymptotic statistics, see, e.g., [2] and later publications [3–5]. Roughly speaking, in statistics the deficiency of a statistical procedure with respect to an 'optimal' procedure is the number of additional observations required to attain the same quality of inference as is guaranteed by the 'optimal' procedure.

In this paper we deal with the case where the deficiency is measured in natural-valued discrete units (number of 'additional' summands) and therefore here we deal with discrete

case. The notion of deficiency can be extended to the case of the continuous parameter, say, time. This case will be considered in another work.

Along with the general case, in the paper main attention is paid to the situation where the r.v.s being summed are assumed to be independent and identically distributed.

The first problem to be considered below consists in determination of the distribution of a separate random variable in the sum that provides the least possible number of summands guaranteeing the prescribed value of the $(1-\alpha)$-quantile of the normalized sum for a given $\alpha \in (0,1)$. The second problem considered in the paper consists in determination of the distribution of a separate random variable in the sum that provides the least possible number of summands guaranteeing the prescribed value of the probability for the normalized sum to fall into a given interval. Actually, in both problems we deal with 'fine tuning' of the distribution of a separate summand since we assume that different possible distributions of random summands possess coinciding three first moments, so that they can differ only by their kurtosis. In both settings the best distribution delivers the smallest number of summands.

We also consider the problem where some additional randomization is introduced so that the number of summands in the sum can be random itself. This randomization may not be artificially induced, but also may occur when the exact number of summands is a priori unknown and only some its 'expected' value can be available as the parameter of the problem. For this case we introduce an analog of deficiency which can be used to compare the distributions of sums with random and non-random number of summands.

Both problems are closely related with the problem of quantification of the accuracy of approximations provided by limit theorems of probability theory. The main mathematical tools used in the paper are asymptotic expansions for the distributions of normalized sums of independent identically distributed r.v.s and their quantiles.

The formal settings mentioned above can be applied to solving practical problems where the models of the observed statistical regularities have the form of distributions of sums of r.v.s and the number of summands plays a substantial role. For example, consider an insurance company whose portfolio consists of a finite number of insurance contracts. Formally, the portfolio is assumed to be a finite set of r.v.s each of which characterizes the income of the company related to a separate contract. Instead of *income* we can speak of *loss* assuming that income is a negative loss or that loss is a negative income.

In these terms, the first setting concerns the problem of determination of the distribution of a possible loss within a separate insurance contract (say, the distribution of an insurance payment) providing the least possible portfolio size and guaranteeing the prescribed Value-at-Risk for the average losses. The approach considered in the paper can be used when the distributions of the summands (possible losses) are known only up to their three first moments and the exact Value-at-Risk is not known for sure. In the second setting the latter requirement is replaced by that of guaranteeing the prescribed 'non-ruin' probability. Within the framework of this example in both settings the problem consists in the description of the best strategy of the insurance company, if by a strategy we mean the choice of the terms of a contract (e.g., the amount of insurance payment related to each possible insurance event), that is, of the distribution of possible loss within a separate contract. Briefly, the problem is to choose an optimal distribution of a separate loss among the distributions that have the same first three moments so that the portfolio size is least possible.

The paper is organized as follows. Section 1.2 contains a short overview of the properties of statistical deficiency. In Section 2 we outline some results concerning the asymptotic expansions for the distributions of $\mathbb{R}$-valued measurable functions of r.v.s and, in particular, for the distributions of normalized sums of r.v.s, as well as for their quantiles. In Section 3 the problem of comparison of the distributions of two sums of independent r.v.s by their deficiency is considered. The notion of asymptotic deficiency is introduced and some formulas for the calculation of asymptotic deficiency are presented. Section 3.1 contains the solution of this problem for these distributions providing a prescribed value

of the $(1-\alpha)$-quantile for a given $\alpha \in (0,1)$. In Section 3.2 this problem is considered for the distributions of sums of independent r.v.s guaranteeing a prescribed probability for an $\mathbb{R}$-valued measurable function of r.v.s, in particular, for a normalized sum of r.v.s, to fall into a given interval. Section 4 contains an example of extension of the results of Section 3 to the case of a random number of summands in the sum (random portfolio size, in terms of the example dealing with an insurance company). In Section 4.1 asymptotic expansions for the asymptotic $(1-\alpha)$-quantile (called α-reserve here) under a random portfolio size are presented and an analog of deficiency of the sum of a random number of summands (or the strategy with a random portfolio size) with respect to the distribution of the sum of a non-random number of summands (or a strategy with a non-random portfolio size) is considered. In Section 4.2 the problem of comparison of these distributions by an analog of deficiency is considered in a special case of three-point distribution of portfolio size.

Everywhere in what follows the set of real numbers is denoted by $\mathbb{R}$, the set of natural numbers is denoted by $\mathbb{N}$. The distribution function of the standard normal law will be denoted by $\Phi(x)$,

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

The distribution of a random vector $(X_1, \dots, X_n)$ will be denoted $\mathcal{L}(X_1, \dots, X_n)$.

1.2. Asymptotic Deficiency

Following the classical terminology of [6], consider two decision rules (say, two statistical procedures) D_n^* and D_n whose quality is characterized by the quantities π_n^* and π_n, respectively. Here n is the number of observations $X_1, \dots, X_n$ delivering the information underlying the decision rules. Assume that the rule D_n^* is in some sense optimal whereas the rule D_n is competing. For example, in the problem of estimation usually π_n^* and π_n are mean square deviations and $\pi_n^* \le \pi_n$. In the problem of testing hypotheses usually π_n^* and π_n are powers of tests so that $\pi_n^* \ge \pi_n$.

By $m(n)$ denote the number of observations required for the decision rule $D_{m(n)}$ based on $m(n)$ observations $X_1, \dots, X_{m(n)}$ to attain the same quality as the 'best' rule D_n^* based on n observations $X_1, \dots, X_n$. In what follows we will keep to the asymptotic approach assuming that $n \to \infty$. Following [7], by the asymptotic relative efficiency (a.r.e.) of the rule D_n with respect to the rule D_n^* we will mean the limit

$$e \equiv \lim_{n\to\infty} \frac{n}{m(n)}$$

(if it exists and does not depend on the sequence $m(n)$).

Instead of the ratio of the required number of observations, the difference $m(n) - n$ can be considered as well, vividly showing the additional number of observations required by the decision rule D_n. However, many authors considered the ratio $n/m(n)$, possibly, because the asymptotic analysis of its properties is simpler.

The systematic analysis of the asymptotic behavior of the difference $m(n) - n$ was first carried out by Hodges and Lehmann in 1970 [2]. They suggested to call the difference $m(n) - n$ *deficiency* of the competing decision rule D_n with respect to the rule D_n^* and introduced the notation

$$d_n = m(n) - n. \tag{1}$$

If the limit $\lim_{n\to\infty} d_n$ exists, then it is called *the asymptotic deficiency* of the competing decision rule D_n with respect to the rule D_n^* and is denoted d. The number d is often called the *deficiency* of D_n with respect to D_n^*. Note that if a.r.e. $e \ne 1$, then $d = \infty$, so that this case is not so interesting. In [2] it was also noticed that for some decision rules (statistical procedures) there typically appear cases $e = 1$ (see, e.g., the book [8]), that is, in these cases the a.r.e. cannot give an answer to the question, which rule is better, whereas the deficiency

can clarify the case, because, generally speaking, in this case the asymptotic deficiency can be arbitrary.

So, the deficiency of D_n with respect to D_n^* shows, how many additional observations (that is, how much extra information) is required to attain the desired quality, if the decision rule D_n is used instead of the 'optimal' decision rule D_n^*. Therefore, the notion of deficiency provides natural grounds for the asymptotic comparison of D_n and D_n^* in the case $e = 1$. The study of the asymptotic behavior of the deficiency d_n requires more sophisticated techniques than is used to find the limit e. As a rule, this techniques employ the construction of asymptotic expansions (a.e.s) for the corresponding functions characterizing the quality of decision rules (see, e.g., the books [7–9]).

Since the rules D_n^* and D_n have the quality characteristics π_n^* and π_n, respectively, then, by the definition of the deficiency $d_n = m(n) - n$, for every n we have

$$\pi_n^* \;=\; \pi_{m(n)}. \tag{2}$$

So solve Equation (2), the integer-valued quantity $m(n)$ should be treated as a variable taking arbitrary real values. For this purpose the function $\pi_{m(n)}$ can be defined for non-integer $m(n)$ by the formula

$$\pi_{m(n)} = \big(1 - m(n) + [m(n)]\big)\pi_{[m(n)]} + \big(m(n) - [m(n)]\big)\pi_{[m(n)]+1}$$

(see [2]).

The functions π_n^* and π_n are usually unknown, so, in practice, their approximations are used. Assume that the a.e.s

$$\pi_n^* = \frac{a}{n^r} + \frac{b}{n^{r+s}} + o\big(n^{-r-s}\big), \tag{3}$$

and

$$\pi_n = \frac{a}{n^r} + \frac{c}{n^{r+s}} + o\big(n^{-r-s}\big), \tag{4}$$

hold, where a, b and c are some numbers that do not depend on n, and $r > 0$, and $s > 0$ are constants determining the rate of decrease of these quality criteria in n. The first terms in these expansions coincide which means that the a.r.e. of the corresponding rules equals one. It can be easily obtained from relations (1)–(4) that

$$d_n = \frac{c - b}{ra} n^{1-s} + o\big(n^{1-s}\big) \tag{5}$$

(see [2] or [7]). Thus, the asymptotic deficiency has the form

$$d = \begin{cases} \pm\infty, & 0 < s < 1, \\ \dfrac{c-b}{ra}, & s = 1, \\ 0, & s > 1. \end{cases} \quad . \tag{6}$$

The asymptotic deficiency possesses the following obvious property of transitivity: if there is some third decision rule $\overline{D}_n$ with the quality characteristic $\overline{\pi}_n$ admitting an a.e. of the form (4), then the deficiency $\overline{d}_n$ of the rule $\overline{D}_n$ with respect the the rule D_n^* satisfies the equality

$$\overline{d}_n = \widetilde{d}_n + d_n,$$

where $\widetilde{d}_n$ is the deficiency of the rule $\overline{D}_n$ with respect to D_n and d_n is the deficiency of D_n with respect to D_n^*.

The case where $s = 1$ is most interesting, because in this case the asymptotic deficiency is finite. In the paper [2] some simple examples are given illustrating that this case is quite natural in mathematical statistics (also see the book [8]).

2. Asymptotic Expansions for the Distributions of Normalized Sums of Random Variables

We begin with most general case. Let $n \in \mathbb{N}$. Consider a finite set of r.v.s $X_1, \dots, X_n$. For the time being we do not assume that the r.v.s $X_1, \dots, X_n$ are independent and identically distributed. Let $L_n = L_n(X_1, \dots, X_n)$ be an $\mathbb{R}$-valued measurable function of $X_1, \dots, X_n$. (In what follows when dealing with the example of the portfolio of an insurance company we will call this function *generalized loss*). In particular, L_n may be of the form $L_n = \sqrt{n}T_n$ where T_n is the arithmetic mean,

$$T_n \equiv \frac{1}{n}\sum_{i=1}^{n} X_i. \tag{7}$$

As it has been already said, the problem consists in description of the distribution of r.v.s X_i providing the least possible number of summands n and guaranteeing the prescribed value of the $(1-\alpha)$-quantile of the function L_n for a given $\alpha \in (0,1)$.

Let $\alpha \in (0,1)$ be a small number. Consider the quantity $c_\alpha(n)$ defined by the asymptotic relation

$$\mathsf{P}\big(L_n \geq c_\alpha(n)\big) = \alpha + o(n^{-1}), \ \ n \to \infty. \tag{8}$$

The quantity $c_\alpha(n)$ is the asymptotic $(1-\alpha)$-quantile of L_n. If $L_n = \sqrt{n}T_n$, then $c_\alpha(n)$ can be interpreted as the threshold, the exceedance of which by L_n is undesirable and is assumed to have the prescribed small probability α. In terms of an insurance company, $c_\alpha(n)$ is the asymptotic Value-at-Risk.

By applying the Taylor formula it is not difficult to obtain the following result.

Lemma 1. *Assume that there exist distribution function $G(x)$ and functions $g_1(x)$ and $g_2(x)$ such that*

$$\sup_{x\in\mathbb{R}}\left|\mathsf{P}\big(L_n < x\big) - G(x) - \frac{1}{\sqrt{n}}g_1(x) - \frac{1}{n}g_2(x)\right| = o(n^{-1}),$$

where the functions $G(x)$, $g_1(x)$ and $g_2(x)$ are smooth enough. Then the asymptotic $(1-\alpha)$-quantile $c_\alpha(n)$ of L_n admits the a.e.

$$c_\alpha(n) = c_\alpha - \frac{g_1(c_\alpha)}{\sqrt{n}G'(c_\alpha)} - \frac{1}{n}\left[\frac{G''(c_\alpha)g_1^2(c_\alpha)}{2(G'(c_\alpha))^3} + \frac{G'(c_\alpha)g_2(c_\alpha) - g_1(c_\alpha)g_1{}'(c_\alpha)}{(G'(c_\alpha))^2}\right] + o(n^{-1}),$$

where c_α satisfies the equation $G(c_\alpha) = 1-\alpha$.

Consider the application of this lemma to the case where $X_1, X_2, \dots$ are independent identically distributed r.v.s such that

$$\mathsf{E}X_1 = 0,\ \mathsf{E}X_1^2 = 1,\ \mathsf{E}|X_1|^{k+\delta} < \infty,\ k \in \mathbb{N},\ k \geq 3,\ \delta > 0 \tag{9}$$

and the function L_n has the form $L_n = \sqrt{n}T_n$ with T_n defined by (7). Here the condition $\mathsf{E}X_1 = 0$ means that the separate losses are centered by their expectations. Assume that the characteristic function $f(t)$ of the r.v. X_1 satisfies the Cramér condition (C)

$$\limsup_{|t|\to\infty}|f(t)| < 1. \tag{10}$$

Under conditions (9) and (10), from Theorem 6.3.2 of [10] (also see [9]) it follows that there exist functions $Q_1(x), \dots, Q_{k-2}(x)$ and a $C_{k,\delta} \in (0,\infty)$ such that

$$\sup_x\left|\mathsf{P}\big(\sqrt{n}T_n < x\big) - \Phi(x) - \sum_{i=1}^{k-2} n^{-i/2}Q_i(x)\right| \leq \frac{C_{k,\delta}}{n^{(k-2+\delta)/2}}, \ \ n \in \mathbb{N}, \tag{11}$$

For the definition of the functions $Q_1(x), \ldots, Q_{k-2}(x)$ see the book [10]. In particular,

$$Q_1(x) = -(x^2-1)\varphi(x)\frac{\mathsf{E}X_1^3}{6},$$

$$Q_2(x) = -(x^3-3x)\varphi(x)\frac{\mathsf{E}X_1^4-3}{24} - (x^5-10x^3+15x)\varphi(x)\frac{(\mathsf{E}X_1^3)^2}{72}. \tag{12}$$

Relations (11) and (12) and Lemma 1 directly imply the a.e. for the asymptotic $(1-\alpha)$-quantile $c_n(\alpha)$ of L_n presented in the following lemma.

Lemma 2. *Let conditions* (9) *and* (10) *hold with* $k = 4$, $\delta > 0$. *Then the the asymptotic* $(1-\alpha)$*-quantile* $c_n(\alpha)$ *of* L_n *admits the a.e.*

$$c_\alpha(n) = u_\alpha + \frac{\mathsf{E}X_1^3}{6\sqrt{n}}(u_\alpha^2-1) + \frac{1}{12n}\Big[\frac{\mathsf{E}^2X_1^3}{3}(5u_\alpha - 2u_\alpha^3) + \frac{\mathsf{E}X_1^4-3}{2}(u_\alpha^3-3u_\alpha)\Big] + o(n^{-1}),$$

where u_α *is the* $(1-\alpha)$*-quantile of the standard normal distribution:* $\Phi(u_\alpha) = 1-\alpha$.

3. The Comparison of the Distributions of Two Normalized Sums of Random Variables

3.1. The Asymptotic Deficiency of the Distributions of Summands Providing a Given $(1-\alpha)$*-Quantile of the Normalized Sums*

In this section we will present an approach to the comparison of the distributions of two sums of r.v.s in terms of the number of summands. The distribution of the random vector $X_1, \ldots, X_n$ will be denoted $\mathcal{L}(X_1, \ldots, X_n)$. Consider an $\mathbb{R}$-valued measurable function of $X_1, \ldots, X_n$.

From Lemma 1 we can easily obtain the following result.

Lemma 3. *Consider a sequence* $\{\epsilon_n\}_{n\geq 1}$ *such that* $\epsilon_n \to 0$ *as* $n \to \infty$. *Under the conditions of Lemma 1 we have*

$$\sup_{x\in\mathbb{R}}\Big|\mathsf{P}\big(L_n(X_1,\ldots,X_n) < x+\epsilon_n\big) - \mathsf{P}\big(L_n(X_1,\ldots,X_n) < x\big) -$$

$$-\epsilon_n G'(x) - \frac{\epsilon_n^2}{2}G''(x) - \frac{\epsilon_n}{\sqrt{n}}g_1{}'(x)\Big| = o\Big(\max\Big\{\epsilon_n^2, \frac{\epsilon_n}{\sqrt{n}}, n^{-1}\Big\}\Big).$$

Along with the r.v.s $X_1, \ldots, X_n$ resulting in the value $L_n(X_1, \ldots, X_n)$ of the function L_n, consider another set of r.v.s $Y_1, \ldots, Y_n$, according to which the value of the function L_n is $L_n(Y_1, \ldots, Y_n)$. For example, $L_n(X_1, \ldots, X_n)$ may have the form $L_n(X_1, \ldots, X_n) = \sqrt{n}T_n$ with T_n defined by (7) and $L_n(Y_1, \ldots, Y_n)$ may have the form $L_n(Y_1, \ldots, Y_n) = \sqrt{n}U_n$ where

$$U_n = \frac{1}{n}\sum_{i=1}^n Y_i. \tag{13}$$

Let to the distribution $\mathcal{L}(Y_1, \ldots, Y_n)$ there correspond the asymptotic $(1-\alpha)$-quantile $\overline{c}_\alpha(n)$ of L_n:

$$\mathsf{P}\big(L_n(Y_1,\ldots,Y_n) \geq \overline{c}_\alpha(n)\big) = \alpha + o(n^{-1}), \quad n \to \infty. \tag{14}$$

Assume that the a.e. for the distribution function of $L_n(Y_1, \ldots, Y_n)$ has the form

$$\mathsf{P}\big(L_n(Y_1,\ldots,Y_n) < x\big) = G(x) + \frac{1}{\sqrt{n}}g_1(x) + \frac{1}{n}\overline{g}_2(x) + o(n^{-1}), \tag{15}$$

where the functions $G(x)$, $g_1(x)$ and $\overline{g}_2(x)$ are smooth enough. The a.e. (15) differs from the a.e. for the distribution function of $L_n(X_1, \ldots, X_n)$ established by Lemma 1 only by

the term of order n^{-1}, which means that the two distributions are rather close. Define the sequence of natural numbers $\{m(n)\}_{n\geq 1}$ by the equality

$$\mathsf{P}\big(L_{m(n)}(Y_1,\dots,Y_{m(n)}) \geq c_\alpha(m(n))\big) = \alpha + o(n^{-1}),\ \ n\to\infty. \tag{16}$$

If $m(n) - n = d + o(1)$, $d \in \mathbb{R}$, $n \to \infty$, then d is the asymptotic deficiency of the distribution $\mathcal{L}(Y_1,\dots,Y_1)$ with respect to the distribution $\mathcal{L}(X_1,\dots,X_n)$. In other words, d is the asymptotic number of 'additional' r.v.s be included in the set $Y_1,\dots,Y_1$ in order that the distribution $\mathcal{L}(Y_1,\dots,Y_{m(n)})$ provides the same quality as the distribution $\mathcal{L}(X_1,\dots,X_n)$.

Theorem 1. *Assume that the conditions of Lemma 1 and (15) hold and $G'(c_\alpha)c_\alpha \neq 0$. Then the asymptotic deficiency d of the distribution $\mathcal{L}(Y_1,\dots,Y_1)$ with respect to the distribution $\mathcal{L}(X_1,\dots,X_n)$ has the form*

$$d = \frac{2\big[g_2(c_\alpha) - \overline{g}_2(c_\alpha)\big]}{G'(c_\alpha)c_\alpha} + o(1).$$

Proof. From Lemma 1 and condition (15) it directly follows that

$$\overline{c}_\alpha(n) = c_\alpha - \frac{g_1(c_\alpha)}{\sqrt{n}G'(c_\alpha)} - \frac{1}{n}\Big[\frac{G''(c_\alpha)g_1^2(c_\alpha)}{2(G'(c_\alpha))^3} + \frac{G'(c_\alpha)\overline{g}_2(c_\alpha) - g_1(c_\alpha)g_1'(c_\alpha)}{(G'(c_\alpha))^2}\Big] + o(n^{-1}) \tag{17}$$

and therefore

$$\epsilon_n \equiv \sqrt{\frac{m(n)}{n}}\overline{c}_\alpha(m(n)) - c_\alpha(m(n)) = \frac{d}{2n}c_\alpha - \frac{1}{n}\frac{\big(g_2(c_\alpha) - \overline{g}_2(c_\alpha)\big)}{G'(c_\alpha)} + o(n^{-1}). \tag{18}$$

Further, with the account of the definitions of $m(n)$ (see (16)) and ϵ_n we have

$$\alpha + o(n^{-1}) = \mathsf{P}\big(L_{m(n)}(Y_1,\dots,Y_{m(n)}) \geq \overline{c}_\alpha(m(n))\big) =$$

$$= \mathsf{P}\Big(L_{m(n)}(Y_1,\dots,Y_{m(n)}) \geq \sqrt{\frac{n}{m(n)}}\big(c_\alpha(m(n)) + \epsilon_n\big)\Big)) \tag{19}$$

Applying Lemma 3 to the right-hand side of (19) we obtain

$$\alpha + o(n^{-1}) = \mathsf{P}\big(L_{m(n)}(Y_1,\dots,Y_{m(n)}) \geq c_\alpha(m(n))\big) - \epsilon_n G'(c_\alpha) + o(n^{-1}).$$

Now from (16) and (18) it follows that

$$d = \frac{2\big[g_2(c_\alpha) - \overline{g}_2(c_\alpha)\big]}{G'(c_\alpha)c_\alpha} + o(1).$$

The theorem is proved. □

Now consider an example of the application of Theorem 1 to the optimization of the portfolio size of an insurance company. Let the possible losses $X_1, X_2, \dots$ related with each insurance contract in the portfolio be independent identically distributed r.v.s satisfying conditions (9) and (10). Consider another distribution, under which the possible losses $Y_1, Y_2, \dots$ are assumed to be independent identically distributed r.v.s such that

$$\mathsf{E}Y_1 = 0,\ \mathsf{E}Y_1^2 = 1,\ \mathsf{E}|Y_1|^{4+\delta} < \infty,\ \delta > 0. \tag{20}$$

Assume that the characteristic function $p(t)$ of the r.v. Y_1 satisfies the Cramér (C) condition

$$\limsup_{|t|\to\infty} |p(t)| < 1. \tag{21}$$

For each n consider the average losses U_n defined by (13). Assume that

$$\mathsf{E}X_1^3 = \mathsf{E}Y_1^3, \tag{22}$$

(for example, the r.v.s X_i and Y_i are centered by their expectations and the distributions of these centered r.v.s are symmetric). From Lemma 2 and Theorem 1 we directly obtain the following statement.

Lemma 4. *Let conditions (9), (10) and (20)–(22) hold. Then the asymptotic (as $n \to \infty$) deficiency of the distribution $\mathcal{L}(Y_1, \dots, Y_n)$ with respect to the distribution $\mathcal{L}(X_1, \dots, X_n)$ (the 'additional number of contracts') d has the form*

$$d = \frac{\left(\mathsf{E}X_1^4 - \mathsf{E}Y_1^4\right)\left(3 - u_\alpha^2\right)}{12} + o(1).$$

Lemma 4 illustrates that if the distributions are close, then the deficiency is determined by the kurtosis.

3.2. The Asymptotic Deficiency of the Distributions of Summands Providing a Given Probability for the Normalized Sum to Fall into a Given Interval

To begin with, in this section we again consider the values of a measurable $\mathbb{R}$-valued function $L_n(X_1, \dots, X_n)$ and $L_n(Y_1, \dots, Y_n)$ on random vectors $(X_1, \dots, X_n)$ and $(Y_1, \dots, Y_n)$ with the the distributions $\mathcal{L}(X_1, \dots, X_n)$ and $\mathcal{L}(Y_1, \dots, Y_n)$, respectively. The goal is to provide that the value of L_n falls into the interval $[S_1, S_2)$ for some given numbers $S_1 < S_2$. As a quality characteristic consider the probabilities

$$\pi_n = \mathsf{P}\big(S_1 \le L_n(X_1, \dots, X_n) < S_2\big), \ \ \overline{\pi}_n = \mathsf{P}\big(S_1 \le L_n(Y_1, \dots, Y_n) < S_2\big). \tag{23}$$

If $L_n(X_1, \dots, X_n) = \sqrt{n}T_n$ (see (7)) and $L_n(Y_1, \dots, Y_n) = \sqrt{n}U_n$ (see (22)), that is, normalized sums of r.v.s are considered, then relation (23) means that π_n and $\overline{\pi}_n$ are probabilities of that the normalized sums of r.v.s are inside the interval $[S_1, S_2)$.

From the definition of π_n we directly obtain the following result.

Lemma 5. *Assume that for some $r > 0$ and $s > 0$ there exist a distribution function $H(x)$ and functions $h_1(x)$, $h_2(x)$ and $\overline{h}_2(x)$ such that*

$$\sup_{x \in \mathbb{R}} \left| \mathsf{P}\big(L_n(X_1, \dots, X_n) < x\big) - H(x) - \frac{1}{n^r} h_1(x) - \frac{1}{n^{r+s}} h_2(x) \right| = o(n^{-r-s}),$$

$$\sup_{x \in \mathbb{R}} \left| \mathsf{P}\big(L_n(Y_1, \dots, Y_n) < x\big) - H(x) - \frac{1}{n^r} h_1(x) - \frac{1}{n^{r+s}} \overline{h}_2(x) \right| = o(n^{-r-s}),$$

and, moreover, the functions $h_1(x)$, $h_2(x)$ and $\overline{h}_2(x)$ are measurable. Then π_n and $\overline{\pi}_n$ admit a.e.s

$$\pi_n = H(S_2) - H(S_1) + \frac{h_1(S_2) - h_1(S_1)}{n^r} + \frac{h_2(S_2) - h_2(S_1)}{n^{r+s}} + o(n^{-r-s}),$$

$$\overline{\pi}_n = H(S_2) - H(S_1) + \frac{h_1(S_2) - h_1(S_1)}{n^r} + \frac{\overline{h}_2(S_2) - \overline{h}_2(S_1)}{n^{r+s}} + o(n^{-r-s}).$$

Corollary 1. *Let $\epsilon_n \downarrow 0$ as $n \to \infty$ and $S_2 = S_1 + \epsilon_n$. Assume that the functions $H(x)$, $h_1(x)$, $h_2(x)$ and $\overline{h}_2(x)$ are smooth enough and $h_1(S_2) \neq h_1(S_1)$. Then*

$$\epsilon_n^{-1} \pi_n = H'(S_1) + \frac{\epsilon_n}{2} H''(S_1) + \frac{\epsilon_n^2}{6} H'''(S_1) + o(\epsilon_n^2) +$$

$$+ \frac{1}{n^r} h_1'(S_1) + \frac{1}{2n^r} h_1''(S_1)\epsilon_n + o(\epsilon_n n^{-r}) + \frac{1}{n^{r+s}} h_2'(S_1) + o(n^{-r-s}\epsilon_n^{-1}),$$

$$\epsilon_n^{-1}\overline{\pi}_n = H'(S_1) + \frac{\epsilon_n}{2}H''(S_1) + \frac{\epsilon_n^2}{6}H'''(S_1) + o(\epsilon_n^2) +$$
$$+\frac{1}{n^r}h_1'(S_1) + \frac{1}{2n^r}h_1''(S_1)\epsilon_n + o(\epsilon_n n^{-r}) + \frac{1}{n^{r+s}}\overline{h}_2'(S_1) + o(n^{-r-s}\epsilon_n^{-1}).$$

Lemma 5, Corollary 1 and formula (6) directly imply the expression for the asymptotic deficiency with quality characteristics (23).

Theorem 2. *Let conditions of Lemma 5 hold with $s = 1$. Then the deficiency d_n of the distribution $\mathcal{L}(Y_1, \ldots, Y_n)$ with the quality characteristic $\overline{\pi}_n$ with respect to the distribution $\mathcal{L}(X_1, \ldots, X_n)$ with the quality characteristic π_n has the form*

$$d_n = \frac{\overline{h}_2(S_2) - h_2(S_2) + \overline{h}_2(S_1) - \overline{h}_2(S_1)}{r(h_1(S_2) - h_1(S_1))} + o(1). \tag{24}$$

If $S_2 = S_1 + \epsilon_n$ with $\epsilon_n \downarrow 0$ as $n \to \infty$ and $h_1'(S_1) \neq 0$, then the formal passage to the limit in (3.13) yields the formula

$$d_n = \frac{\overline{h}_2'(S_1) - h_2'(S_1)}{rh_1'(S_1)} + o(1).$$

Consider an example of the application of Theorem 2 to the optimization of the portfolio size of an insurance company. Let the possible losses $X_1, X_2, \ldots$ related with each insurance contract in the portfolio be independent identically distributed r.v.s satisfying conditions (9) and (10). Consider another distribution, under which the possible losses $Y_1, Y_2, \ldots$ are assumed to be independent identically distributed r.v.s satisfying conditions (20) and (21). Assume that in (9) and (20) $k = 3$. We are interested in the asymptotic behavior of the average losses T_n (see (7)) and U_n (see (13)). With the account of Lemma 5 we obtain the following statement.

Lemma 6. *Let conditions* (9), (10), (19) *and* (20) *hold with $k = 3$. Then*

$$\mathsf{P}\big(\sqrt{n}T_n < x\big) = \Phi(x) + \frac{Q_1(x)}{\sqrt{n}} + \frac{Q_2(x)}{n} + o(n^{-1}),$$

$$\mathsf{P}\big(\sqrt{n}U_n < x\big) = \Phi(x) + \frac{\overline{Q}_1(x)}{\sqrt{n}} + \frac{\overline{Q}_2(x)}{n} + o(n^{-1}),$$

uniformly in $x \in \mathbb{R}$,

$$\pi_n = \Phi(S_2) - \Phi(S_1) + \frac{Q_1(S_2) - Q_1(S_1)}{\sqrt{n}} + \frac{Q_2(S_2) - Q_2(S_1)}{n} + o(n^{-1}),$$

$$\overline{\pi}_n = \Phi(S_2) - \Phi(S_1) + \frac{\overline{Q}_1(S_2) - \overline{Q}_1(S_1)}{\sqrt{n}} + \frac{\overline{Q}_2(S_2) - \overline{Q}_2(S_1)}{n} + o(n^{-1}),$$

where the functions $Q_1(x)$ and $Q_2(x)$ are defined in (12),

$$\overline{Q}_1(x) = -(x^2 - 1)\varphi(x)\frac{\mathsf{E}Y_1^3}{6},$$

$$\overline{Q}_2(x) = -(x^3 - 3x)\varphi(x)\frac{\mathsf{E}Y_1^4 - 3}{24} - (x^5 - 10x^3 + 15x)\varphi(x)\frac{(\mathsf{E}Y_1^3)^2}{72}.$$

Corollary 2. *Let $\epsilon_n \downarrow 0$ as $n \to \infty$ and $S_2 = S_1 + \epsilon_n$. Assume that conditions of Lemma 6 hold. Then*

$$\epsilon_n^{-1}\pi_n = \varphi(S_1) + \frac{\epsilon_n}{2}\varphi'(S_1) + \frac{\epsilon_n^2}{6}\varphi''(S_1) + o(\epsilon_n^2) +$$

$$+\frac{1}{\sqrt{n}}Q_1'(S_1)+\frac{\epsilon_n}{2\sqrt{n}}Q_1''(S_1)+o(\epsilon_n n^{-1/2})\frac{1}{n}Q_2'(S_1)+o(n^{-1}\epsilon_n^{-1}),$$

$$\epsilon_n^{-1}\overline{\pi}_n=\varphi(S_1)+\frac{\epsilon_n}{2}\varphi'(S_1)+\frac{\epsilon_n^2}{6}\varphi''(S_1)+o(\epsilon_n^2)+$$

$$+\frac{1}{\sqrt{n}}\overline{Q}_1'(S_1)+\frac{\epsilon_n}{2\sqrt{n}}\overline{Q}_1''(S_1)+o(\epsilon_n n^{-1/2})+\frac{1}{n}\overline{Q}_2'(S_1)+o(n^{-1}\epsilon_n^{-1}).$$

Theorem 2, Lemma 5 and formula (5) directly imply the following statement.

Theorem 3. *Let, in addition to the conditions of Lemma 5.,* $\mathsf{E}X_1^3=\mathsf{E}Y_1^3$. *Then the deficiency* d_n *of the distribution* $\mathcal{L}(Y_1,\dots,Y_n)$ *with the quality characteristic* $\overline{\pi}_n$ *with respect to the strategy* $\mathcal{L}(X_1,\dots,X_n)$ *with the quality characteristic* π_n *(the 'additional number of contracts') has the form*

$$d_n=2\frac{\overline{Q}_2(S_2)-Q_2(S_2)+Q_2(S_1)-\overline{Q}_2(S_1)}{Q_1(S_2)-Q_1(S_1)}n^{1/2}+o(n^{1/2}).$$

Consider an example where the asymptotic deficiency is finite.

Corollary 3. *Let* $\epsilon_n=\frac{1}{n}$ *and* $S_2=S_1+\frac{1}{n}$, $\mathsf{E}X_1^3=\mathsf{E}Y_1^3=0$. *Then under the conditions of Lemma 5 we have*

$$\pi_n=\frac{\varphi(S_1)}{n}+\frac{\varphi'(S_1)+2Q_2'(S_1)}{n^2}+o(n^{-2}),$$

$$\pi_n=\frac{\varphi(S_1)}{n}+\frac{\varphi'(S_1)+2\overline{Q}_2'(S_1)}{n^2}+o(n^{-2}).$$

Moreover, the deficiency d_n *has the form*

$$d_n=\frac{2(\overline{Q}_2'(S_1)-Q_2'(S_1))}{\varphi(S_1)}+o(1)=\frac{S_1^4-6S_1^2+3}{12}(\mathsf{E}Y_1^4-\mathsf{E}X_1^4)+o(1).$$

4. Random Number of Summands

4.1. Asymptotic Expansions for the Asymptotic $(1-\alpha)$*-Quantile of* $\mathbb{R}$*-Valued Measurable Functions of a Random Number of Random Variables*

In this section we consider the case where an additional randomization can be introduced into the problem. In this case the number of summands in the sum can be considered as random. This randomization may not be artificially induced, but also may occur when the exact portfolio size can be unknown beforehand and only some 'expected' number of summands can be available as the parameter of the problem.

Let natural-valued r.v.s $N_1,N_2,\dots$ and r.v.s $X_1,X_2,\dots$ be defined on one and the same probability space $(\Omega,\mathcal{A},\mathsf{P})$. In what follows we will assume that n is the expected value of N_n,

$$\mathsf{E}N_n=n. \tag{25}$$

Assume that for each $n\geq 1$ the r.v. N_n is independent of the sequence $X_1,X_2,\dots$. As above, for each $n\geq 1$, consider the value of an $\mathbb{R}$-valued measurable function $L_n=L_n(X_1,\dots,X_n)$. For each $n\geq 1$ consider the r.v. L_{N_n} defined as

$$L_{N_n}(\omega)\equiv L_{N_n(\omega)}(X_1(\omega),\dots,X_{N_n(\omega)}(\omega)),\ \ \omega\in\Omega.$$

Below we will assume that the following condition holds.

Condition A. *There exist* $k\in\mathbb{N}\backslash\{1\}$, $\alpha_{i,n}\in\mathbb{R}$, $i=1,\dots,k$, $\beta_n>0$, $C_k>0$, *a differentiable distribution function* $G(x)$ *and measurable functions* $g_j(x)$, $j=1,\dots,k$ *such that*

$$\beta_n\to 0,\ \ \max_{1\leq i\leq k}|\alpha_{i,n}|\to 0$$

as $n \to \infty$ *and*

$$\sup_x \left| \mathsf{P}\left(L_n < x\right) - G(x) - \sum_{i=1}^{k} \alpha_{i,n} g_i(x) \right| \le C_k \beta_n, \ \ n \in \mathbb{N}.$$

Lemma 7. *Let the function* $L_n = L_n(X_1, \ldots, X_n)$ *satisfy Condition A. Then*

$$\sup_x \left| \mathsf{P}\left(L_{N_n} < x\right) \ - G(x) - \sum_{i=1}^{k} g_i(x) \mathsf{E}\alpha_{i,N_n} \right| \le C_k \mathsf{E}\beta_{N_n}.$$

The elementary proof of this lemma directly follows by the formula of total probability.

Consider an example of application of Lemma 7. Let $X_1, X_2, \ldots$ be independent identically distributed r.v.s satisfying conditions (9) and (10). Assume that the function L_n is the normalized arithmetic mean (or, which is the same, the normalized sum) $L_n = \sqrt{n} T_n$ with T_n defined in (7). Then, in accordance with what has been said in Section 2, relation (11) holds implying the validity of Condition A. From (11) playing the role of Condition A and Lemma 7 we obtain the following statement.

Lemma 8. *Assume that* $L_n = \sqrt{n} T_n$ *with* T_n *defined in (7) and conditions (9) and (10) hold. Then*

$$\sup_x \left| \mathsf{P}\left(\sqrt{N_n} T_{N_n} < x\right) - \Phi(x) - \sum_{i=1}^{k-2} Q_i(x) \mathsf{E} N_n^{-i/2} \right| \le C_{k,\delta} \mathsf{E} N_n^{-(k-2+\delta)/2},$$

where the functions $Q_i(x)$ *are defined in Theorem 6.3.2 of [10].*

Relation (11) and Lemma 8 imply the following statement.

Lemma 9. *Let conditions (9) and (10) hold with* $k = 4$ *and* $\delta > 0$. *Assume that condition (25) holds and*

$$\mathsf{E} N_n^{-1/2} = \frac{1}{\sqrt{n}} + \frac{a}{n} + o(n^{-1}), \ \ a \in \mathbb{R},$$

$$\mathsf{E} N_n^{-1} = \frac{b}{n} + o(n^{-1}), \ \ \mathsf{E} N_n^{-(2+\delta)/2} = o(n^{-1}), \ \ b \in \mathbb{R}.$$

Then

$$\sup_x \left| \mathsf{P}\left(\sqrt{n} T_n < x\right) - \Phi(x) - \frac{Q_1(x)}{\sqrt{n}} - \frac{Q_2(x)}{n} \right| = o(n^{-1})$$

and

$$\sup_x \left| \mathsf{P}\left(\sqrt{N_n} T_{N_n} < x\right) - \Phi(x) - \frac{Q_1(x)}{\sqrt{n}} - \frac{bQ_2(x) + aQ_1(x)}{n} \right| = o(n^{-1}).$$

We will use Lemma 9 in order to determine the asymptotic $(1-\alpha)$-quantile of L_n and calculate the asymptotic deficiency.

Recall that, for $\alpha \in (0,1)$, the asymptotic $(1-\alpha)$-quantile of L_n is the quantity $c_\alpha(n)$ satisfying the asymptotic equality

$$\mathsf{P}\left(L_n \ge c_\alpha(n)\right) = \alpha + o(n^{-1}), n \to \infty. \tag{26}$$

Correspondingly, we define the the asymptotic $(1-\alpha)$-quantile $\tilde{c}_\alpha(n)$ of L_{N_n} by the equation

$$\mathsf{P}\left(L_{N_n} \ge \tilde{c}_\alpha(n)\right) = \alpha + o(n^{-1}), n \to \infty. \tag{27}$$

From Lemmas 1 and 9 we directly obtain the a.e.s for these asymptotic $(1-\alpha)$-quantiles.

Lemma 10. *Under the conditions of Lemma 8, we have*

$$c_\alpha(n) = u_\alpha + \frac{\mathsf{E}X_1^3}{6\sqrt{n}}(u_\alpha^2 - 1) + \frac{1}{12n}\Big[\frac{\mathsf{E}^2X_1^3}{3}(5u_\alpha - 2u_\alpha^3) + \frac{\mathsf{E}X_1^4 - 3}{2}(u_\alpha^3 - 3u_\alpha)\Big] + o(n^{-1}),$$

$$\tilde{c}_\alpha(n) = u_\alpha + \frac{\mathsf{E}X_1^3}{6\sqrt{n}}(u_\alpha^2 - 1) +$$

$$+\frac{1}{12n}\Big[\frac{\mathsf{E}^2X_1^3}{3}(5u_\alpha - 2u_\alpha^3) + \frac{b(\mathsf{E}X_1^4 - 3)}{2}(u_\alpha^3 - 3u_\alpha) + 2a\mathsf{E}X_1^3(u_\alpha^2 - 1)\Big] + o(n^{-1}),$$

where u_α satisfies the equation $\Phi(u_\alpha) = 1 - \alpha$.

Now define the sequence $m(n)$ of natural numbers by the relation

$$\mathsf{P}\big(\sqrt{n}L_{N_{m(n)}} \geq \sqrt{m(n)}c_\alpha(m(n))\big) = \alpha + o(n^{-1}), n \to \infty. \tag{28}$$

If

$$m(n) = n + d + o(1), \tag{29}$$

$n = 1, 2, \ldots$, then d can have the meaning of the expected additional number of summands to be included in the sum in order that the function L_{N_n} exceeds $c_\alpha(n)$ for the loss under a non-random number n of summands. The quantity d will be called the *asymptotic deficiency*.

In the same way that Theorem 1 was proved, we can establish the following statement.

Theorem 4. *Assume that*

$$\mathsf{E}N_n = n, \ \ \mathsf{E}N_n^{-1/2} = \frac{1}{\sqrt{n}} + \frac{a}{n} + o(n^{-1}), \ \ a \in \mathbb{R},$$

$$\mathsf{E}N_n^{-1} = \frac{b}{n} + o(n^{-1}), \ \ \mathsf{E}N_n^{-(2+\delta)/2} = o(n^{-1}), \ \ b \in \mathbb{R},$$

and there exist $\delta > 0$, a differentiable distribution function $G(x)$ and measurable functions $g_1(x)$ and $g_2(x)$ such that

$$\sup_x \Big|\mathsf{P}\big(L_n < x\big) - G(x) - \frac{g_1(x)}{\sqrt{n}} - \frac{g_2(x)}{n}\Big| \leq \frac{C}{n^{(2+\delta)/2}}$$

and $G'(c_\alpha)c_\alpha \neq 0$. Then the expected number d of additional summands (see (28) and (29)) in the normalized random sum L_{N_n} with respect to the normalized sum L_n has the form

$$d = \frac{2\big[g_2(c_\alpha)(1-b) - ag_1(c_\alpha)\big]}{G'(c_\alpha)c_\alpha} + o(1),$$

where c_α satisfies the equation $G(c_\alpha) = 1 - \alpha$.

Theorem 4 implies the following statement.

Corollary 4. *Under the conditions of Lemma 8 the expected additional number of summands d (see (28) and (29)) corresponding to the normalized sum $\sqrt{N_n}T_{N_n}$ with a random number of summands with respect to the normalized sum $\sqrt{n}T_n$ has the form*

$$d = \frac{2\big((1-b)Q_2(u_\alpha) - aQ_1(u_\alpha)\big)}{\varphi(u_\alpha)u_\alpha} + o(1).$$

If additionally $\mathsf{E}X_1^3 = 0$, *then*

$$d = \frac{(1-b)(3-u_\alpha^2)(\mathsf{E}X_1^4 - 3)}{12} + o(1).$$

4.2. An Example of Three-Point Distribution of the Number of Summands

In this section, keeping to the terminology of the example related to optimization of the portfolio size of an insurance company, we will use Corollary 4 to obtain a.e.s for the asymptotic Value-at-Risk (asymptotic $(1-\alpha)$-quantile of the normalized average loss, or asymptotic normalized α-reserve) in the case where the portfolio size N_n has a special distribution concentrated in three points so that is symmetric around the central point.

Assume that the portfolio size N_n has the distribution of the form

$$\mathsf{P}(N_n = n - h_n) = \mathsf{P}(N_n = n) = \mathsf{P}(N_n = n + h_n) = \tfrac{1}{3}, \tag{30}$$

where $h_n \in \mathbb{N}$, $h_n < n$, $n = 1, 2, \ldots$, and

$$\lim_{n\to\infty} \frac{h_n}{n} = 0. \tag{31}$$

Lemma 11. *Let the random portfolio size* N_n *have distribution (30) and let condition (31) hold. Then* $\mathsf{E}N_n = n$ *and, as* $n \to \infty$,

$$\mathsf{E}N_n^{-1/2} = \frac{1}{\sqrt{n}} - \frac{1}{4\sqrt{n}}\Big(\frac{h_n}{n}\Big)^2 + O\Big(\frac{1}{\sqrt{n}}\Big(\frac{h_n}{n}\Big)^3\Big),$$

$$\mathsf{E}N_n^{-1} = \frac{1}{n} + \frac{2}{3n}\Big(\frac{h_n}{n}\Big)^2 + O\Big(\frac{1}{n}\Big(\frac{h_n}{n}\Big)^4\Big), \quad \mathsf{E}N_n^{-3/2} = \frac{1}{n^{3/2}} + O\Big(\frac{1}{n^{3/2}}\Big(\frac{h_n}{n}\Big)^2\Big).$$

Proof. The desired statements follow from the relations

$$\mathsf{E}N_n^{-1} = \frac{3n^2 - h_n^2}{3n(n^2 - h_n^2)} = \frac{1}{n}\Big(1 - \frac{h_n^2}{3n}\Big)\Big(1 + \frac{h_n^2}{n^2} + O\Big(\frac{h_n^4}{n^4}\Big)\Big) = \frac{1}{n} + \frac{2}{3n}\Big(\frac{h_n}{n}\Big)^2 + O\Big(\frac{1}{n}\Big(\frac{h_n}{n}\Big)^4\Big),$$

$$\mathsf{E}N_n^{-3/2} = \frac{1}{3n^{3/2}}\Big(\frac{1}{(1-h_n/n)^{3/2}} + 1 + \frac{1}{(1+h_n/n)^{3/2}}\Big) = \frac{1}{n^{3/2}} + O\Big(\frac{1}{n^{3/2}}\Big(\frac{h_n}{n}\Big)^2\Big).$$

The formula for $\mathsf{E}N_n^{-1/2}$ is established in a similar way. □

Lemmas 10 and 11 imply the following statement.

Theorem 5. *Assume that the normalized average loss has the form* $L_n = \sqrt{n}T_n$ *with* T_n *defined in (7). Let the r.v.* N_n *be distributed according to (30) and condition (31) hold. Under the conditions of Lemma 9, for the asymptotic* α*-reserve* $\tilde{c}_\alpha(n)$ *corresponding to the normalized average loss* $\sqrt{N_n}T_{N_n}$ *there holds the relation*

$$\tilde{c}_\alpha(n) = c_\alpha(n) - \frac{\mathsf{E}X_1^3(u_\alpha^2 - 1)}{24\sqrt{n}}\Big(\frac{h_n}{n}\Big)^2 + o(n^{-1}), n \to \infty.$$

Remark 1. *In addition to the conditions of Theorem 5, let*

$$h_n = \gamma n^\beta + o(n^\beta), \ \ \gamma \geq 0, \ \ 0 \leq \beta < 1.$$

Then, as $n \to \infty$,

$$n^{5/2-2\beta}\big(c_\alpha(n) - \tilde{c}_\alpha(n)\big) \to \frac{\gamma^2}{24}\mathsf{E}X_1^3(u_\alpha^2 - 1).$$

Applying Lemma 9, by simple calculations we obtain the following statement.

Lemma 12. *Assume that conditions (9) and (10) hold with $k = 4$ and $0 < \delta \leq 1$. Let conditions (30) and (31) hold. Then*

$$\sup_x \left| \mathsf{P}\left(\sqrt{N_n} T_{N_n} < x\right) - \Phi(x) - \left(1 - \frac{h_n^2}{4n^2}\right) \frac{Q_1(x)}{\sqrt{n}} - \left(1 + \frac{2h_n^2}{3n^2}\right) \frac{Q_2(x)}{n} \right| = O\left(\frac{h_n^{(4+2\delta)/3}}{n^{7(2+\delta)/6}}\right).$$

Corollary 5. *Let conditions of Lemma 12 hold and $h_n = n^{3/4}$. Then*

$$\sup_{x \in \mathbb{R}} \left| \mathsf{P}\left(\sqrt{N_n} T_{N_n} < x\right) - \Phi(x) - \frac{1}{\sqrt{n}} Q_1(x) - \frac{1}{n}\left(Q_2(x) - \frac{1}{4} Q_1(x)\right) \right| = o(n^{-1}).$$

Relations (12), Lemmas 10 and 11 yield the following theorem.

Theorem 6. *Let the conditions of Corollary 5 hold. Then the asymptotic α-reserves $c_\alpha(n)$ and $\tilde{c}_\alpha(n)$ related to the normalized average losses $\sqrt{n} T_n$ and $\sqrt{N_n} T_{N_n}$ have the form*

$$c_\alpha(n) = u_\alpha + \frac{\mathsf{E} X_1^3}{6\sqrt{n}}(u_\alpha^2 - 1) + \frac{1}{12n}\left[\frac{\mathsf{E}^2 X_1^3}{3}(5u_\alpha - 2u_\alpha^3) + \frac{\mathsf{E} X_1^4 - 3}{2}(u_\alpha^3 - 3u_\alpha)\right] + o(n^{-1}),$$

$$\tilde{c}_\alpha(n) = u_\alpha + \frac{\mathsf{E} X_1^3}{6\sqrt{n}}(u_\alpha^2 - 1) +$$

$$+ \frac{1}{12n}\left[\frac{\mathsf{E}^2 X_1^3}{3}(5u_\alpha - 2u_\alpha^3) + \frac{\mathsf{E} X_1^4 - 3}{2}(u_\alpha^3 - 3u_\alpha) - \frac{1}{2}\mathsf{E} X_1^3 (u_\alpha^2 - 1)\right] + o(n^{-1}),$$

where u_α satisfies the equation $\Phi(u_\alpha) = 1 - \alpha$. The corresponding expected additional number d of contracts has the form

$$d = \frac{Q_1(u_\alpha)}{2\varphi(u_\alpha) u_\alpha} + o(1) = \frac{(1 - u_\alpha^2)\mathsf{E} X_1^3}{12 u_\alpha} + o(1).$$

5. Conclusions

The paper deals with an approach to the comparison of distributions of sums of a finite number of independent random variables by deficiency. The notion of asymptotic deficiency of the distribution of a measurable $\mathbb{R}$-valued function of a random vector with respect to the distribution of the same function of another random vector was introduced. Some formulas for the calculation of asymptotic deficiency were presented in the cases where the function has the form of a normalized sum of independent identically distributed r.v.s. The formulas for the asymptotic deficiency were obtained as the solution of two problems, one of which deals with the description of the distribution of a separate summand minimizing the number of summands and providing a prescribed value of the $(1 - \alpha)$-quantile of the normalized sum for a given $\alpha \in (0, 1)$. The second problem deals with minimization of the number of summands and guaranteeing a prescribed probability for a normalized sum of r.v.s to fall into a given interval. These results were extended to the case of a random number of summands in the sum (or random portfolio size, in terms of the example dealing with an insurance company). For this case, an analog of deficiency of the sum of a random number of summands with respect to the distribution of the sum of a non-random number of summands was introduced. The problem of comparison of these distributions by an analog of deficiency was considered in a special case of three-point distribution of portfolio size. The main mathematical tools used in the paper were asymptotic expansions for the distributions of average losses and their quantiles.

Author Contributions: Conceptualization, V.E.B. and V.Y.K.; Formal analysis, V.Y.K.; Funding acquisition, V.Y.K.; Investigation, V.E.B. and V.Y.K.; Writing – original draft, V.E.B. and V.Y.K. All authors have read and agreed to the published version of the manuscript.

Funding: The research was supported by the Ministry of Science and Higher Education of the Russian Federation, project No. 075-15-2020-799.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors thank the anonymous referees for their comments and suggestions that improved the paper. We also thank A. K. Gorshenin for his help in formatting the paper.

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

References

1. Müller, A.; Stoyan, D. *Comparison Methods for Stochastic Models and Risks*; J. Wiley & Sons: Chichester, UK, 2002; 352p, ISBN 978-0-471-49446-1.
2. Hodges, J.L.; Lehmann, E.L. Deficiency. *Ann. Math. Stat.* **1970**, *41*, 783–801._50. [CrossRef]
3. Torgersen, E. *Comparison of Statistical Experiments*; Printed online: May 2013; Cambridge University Press: Cambridge, UK, 1991; doi:10.1017/CBO9780511666353.007. [CrossRef]
4. Xiang, X. Deficiency of the sample quantile estimator with respect to kernel quantile estimators for censored data. *Ann. Stat.* **1995**, *23*, 836–854. [CrossRef]
5. Bening, V.E.; Korolev, V.Y.; Zeifman, A.I. Calculation of the deficiency of some statistical estimators constructed from samples with random sizes. *Colloq. Math.* **2019**, *157*, 157–171. [CrossRef]
6. Blackwell, D.; Girshick, M.A. *Theory of Games and Statistical Decisions. Wiley Publications in Statistics*; J. Wiley & Sons: New York, NY, USA; Chapman & Hall: London, UK, 1954; pp. XI, 355.
7. Lehmann, E.L.; Casella, G. *Theory of Point Estimation*; Springer: Berlin, Germany, 1998; 589p.
8. Bening, V.E. *Asymptotic Theory of Testing Statistical Hypotheses: Efficient Statistics, Optimality, Power Loss, and Deficiency*; Walter de Gruyter: Berlin, Germany, 2011; 277p, ISBN 978-3-11-093599-8.
9. Cramér, H. *Mathematical Methods of Statistics*; Princeton University Press: Princeton, NJ, USA, 1946; 647p.
10. Petrov, V.V. *Limit Theorems of Probability Theory: Sequences of Independent Random Variables*; Clarendon Press: Oxford, UK, 1985; 437p.