DOI 10.1007/s10958-016-3213-3
Journal of Mathematical Sciences, Vol. 220, No. 6, February, 2017

THE RATE OF CONVERGENCE OF THE DISTRIBUTIONS OF REGULAR STATISTICS CONSTRUCTED FROM SAMPLES WITH NEGATIVELY BINOMIALLY DISTRIBUTED RANDOM SIZES TO THE STUDENT DISTRIBUTION

S. V. Gavrilenko[1] , V. N. Zubov[1] , and V. Yu. Korolev[1,2]

New estimates are obtained for the rate of convergence of the negative binomial distribution with parameters (r, p) to the gamma-distribution with parameters (r, r) when $p \to 0$. The main result is a considerable improvement of convergence rate estimates for regular statistics constructed from samples with negatively binomially distributed random sizes to the Student distribution.

1. Introduction

The main goal of this paper is to refine the information on the convergence of asymptotically normal statistics with random indexes of the form T_{N_n}, where N_n is a random variable with negative binomial distribution, to the Student distribution. Let us recall that a statistic T_n is called asymptotically normal if

$$P(\sigma\sqrt{n}(T_n - \mu) \leqslant x) \overset{n\to\infty}{\Longrightarrow} \Phi(x).$$

Additionally assume that

$$\sup_x |P(\sqrt{n}\sigma(T_n - \mu) \leqslant x) - \Phi(x)| = O\left(\frac{1}{\sqrt{n}}\right), \tag{1}$$

and that the index distribution has the form

$$P(N_n = k) = \binom{k+r-1}{r-1} p^r (1-p)^k,\ p = \frac{r}{n},\ k = 0, 1, \ldots .$$

Then, as is shown in [1], it is possible to estimate the rate of convergence of $P(\sqrt{n-r}\sigma(T_{N_n} - \mu) \leqslant x)$ to the Student distribution function $F_{2r}(x)$ with $2r$ degrees of freedom. In the same paper the authors obtained an auxiliary result, for which authors of the present paper wish to improve. Namely, it is proved that

$$\sup_x \left| P\left(\frac{N_{p,r}}{\mathbf{E}N_{p,r}} \leqslant x\right) - G_{r,r}(x) \right| = O\left(p^{\frac{r}{r+1}}\right) \tag{2}$$

for random variable $N_{p,r}$ with negative binomial distribution and $G_{r,r}(x)$ — the function of gamma-distribution with equal parameters. On the basis of this result the following statement is proved.

Theorem 2.1. *Let T_n be asymptotically normal in the above sense, and let* (1) *hold. Then the following estimate is valid*:

$$\sup_x |P(\sqrt{n-r}\sigma(T_{N_n} - \mu) \leqslant x) - F_{2r}(x)| = O\left(n^{-\frac{r}{r+1}}\right).$$

In the present paper we considerably improve the estimate (2) and, as a consequence, refine the main theorem. Its new formulation is given in the final part of the paper.

[1] Lomonosov Moscow State University, Faculty of Computational Mathematics and Cybernetics, Moscow, Russia, e-mail: vkorolev@cs.msu.su
[2] Institute of Informatics Problems of RAS, Moscow, Russia

Translated from *Statisticheskie Metody Otsenivaniya i Proverki Gipotez*, Vol. 19, pp. 118–134, 2006
1072-3374/17/2206-0701

2. Convergence rate in terms of smoothed uniform metric for $0 < r < 1$

Consider $N_{p,r}$, a random variable having negative binomial distribution with parameters $(r,p), r \in (0,1)$, i.e.,

$$P\left(N_{p,r}=k\right)=\frac{\Gamma(r+k)}{k!\,\Gamma(r)}p^{r}(1-p)^{k}, k=0,1,2,\ldots,$$

and its normalization

$$N^{*}_{p,r}=\frac{N_{p,r}}{\mathbf{E}N_{p,r}}$$

as $p \to 0$. In [1] it was shown that

$$\Delta_{p,r}\equiv\sup_{x\geqslant 0}|P(N^{*}_{p,r}\leqslant x)-G_{r,r}(x)|\to 0,\ \ p\to 0,$$

where $G_{r,r}(x)=\frac{r^{r}}{\Gamma(r)}\int_0^x e^{-ry}y^{r-1}dy,\ x>0.$

We are interested in the rate of the weak convergence

$$N^{*}_{p,r}\overset{p\to 0}{\Rightarrow}\xi\sim\Gamma(r,r) \tag{3}$$

for all $r > 0$.

Let us begin with the case $0 < r < 1$.

In most limit theorems of probability theory and mathematical statistics, the rate of convergence is dealt with in terms of the uniform metric

$$\rho(F,G)=\sup_{x}|F(x)-G(x)|.$$

For given $F(x) = P(N_{p,r} < x)$ and $G(x)=\int\limits_{o}^{x}\frac{r^{r}}{\Gamma(r)}e^{-rt}t^{r-1}dt$ the following estimate was obtained in [1]:

$$\sup_{x\geqslant 0}|F(x)-G(x)|=O(p^{\frac{r}{r+1}}),$$

which, in the best case, considering that $r \in (0,1)$, gives $O(\sqrt{p})$.

However, in some cases, the uniform metric does not allow one to estimate the real closeness of distribution functions. As an example consider

$$F_n(x)=\begin{cases}1, & x>\frac{1}{n},\\ 0, & x\leqslant\frac{1}{n},\end{cases}\qquad G(x)=\begin{cases}1, & x>0,\\ 0, & x\leqslant 0,\end{cases}$$

$$\sup_{x\geqslant 0}|F_n(x)-G(x)|=1,\qquad \text{but}\quad F_n(x)\xrightarrow{n\to\infty}G(x)\ \text{for}\ \forall x.$$

Let us consider an alternative approach to the problem with the use of the smoothed uniform metric $\rho(F,G)=\sup\limits_{x}|(F*H)(x)-(G*H)(x)|$ for some fixed distribution H. We assume that the distribution function H is absolutely continuous ($h(u)$ is the distribution density).

Looking at the above example, note that

$$(F_n*G)(x)=\int\limits_0^x F_n(x-u)h(u)du=\int\limits_{\substack{x-u>\frac{1}{n}\\ u\in[0,x]}}h(u)du=\int\limits_0^{x-\frac{1}{n}}h(u)du,$$

$$(G*H)(x)=\int\limits_{o}^{x}h(u)du,$$

$$\sup_x|(F_n*H)(x)-(G*H)(x)|=\sup_x|\int\limits_{x-\frac{1}{n}}^{x}h(u)du|\xrightarrow{n\to\infty}0,$$

at least for the class of all bounded densities h. Thus, the smoothed uniform metric is more adequate here.

The smoothed uniform metric is also convenient because, as is well known, the uniform metric is regular, i.e., if a random variable $Z\sim H(x)$ is independent of random variables $X\sim F(x)$ and $Y\sim G(x)$, then

$$\rho_H(F,G)=\rho(F*H,G*H)=\rho(X+Z,Y+Z)\leqslant\rho(X,Y)=\rho(F,G).$$

This means that, in view of the regularity, the results for the rate of convergence of $\Delta_{p,r}$ in terms of the smoothed metric obtained below may be combined with the results of D. O. Selivanova (see [3]): for natural r

$$\Delta_{p,r}\equiv\sup_{x\geqslant 0}|P(N^*_{p,r}\leqslant x)-G_{r,r}(x)|\leqslant\frac{rp}{1-p}=O(p).$$

Let us return to the initial problem of estimation of the rate of convergence (3). Using the uniform metric we will not be able to obtain good estimates, since we cannot use Esseen's estimates [1] because the derivative of $G_{r,r}(x)$ at zero is unbounded. Therefore we shall try to obtain good estimates of convergence rate in terms of the smoothed uniform metric. We shall take $H(x)=G_{r,1}(x)$ (i.e., exponential distribution with the density $h(x)=re^{-rx}$) as a smoothing distribution. Note that

$$(G*H)(x)=(G_{r,r}*G_{r,1}(x))=G_{r,r+1}(x).$$

This means that the derivative $(G*H)'(x)=\frac{r^{r+1}}{\Gamma(r+1)}e^{-rx}x^r$ is bounded for $0<r<1$ and, hence, we may apply the Esseen inequality to the convolutions $(F*H)(x)$ and $(G*H)(x)$

$$\sup_{x\geqslant 0}|(F*H)(x)-(G*H)(x)|\leqslant\frac{1}{\pi}\int\limits_{|t|\leqslant T}\frac{|f_h(t)-g_h(t)|}{|t|}\,dt+\frac{24A}{\pi T},\tag{4}$$

where $A=\sup\limits_x|(G*H)'(x)|$ and $f_h(t)$ and $g_h(t)$ are the characteristic functions of $(F*H)(x)$ and $(G*H)(x)$, respectively. This inequality is valid for all $T>0$.

Obviously, there exist a probability space, a random variable $Z\sim H$, and random variables X,Y independent of Z defined on the same space, such that $X+Z\sim F*H$ and $Y+Z\sim G*H$. Then for the characteristic functions of convolutions we have the inequality

$$|f_h(t)-g_h(t)|=|Ee^{it(Z+X)}-Ee^{it(Z+Y)}|=$$

$$=|Ee^{itZ}(Ee^{itX}-Ee^{itY})|\leqslant|f(t)-g(t)|\cdot E|e^{itZ}|=|f(t)-g(t)|.$$

Therefore, from the Esseen inequality, it follows that

$$\Delta_{p,r}=\sup_{x\geqslant 0}|(F*H)(x)-(G*H)(x)|\leqslant\frac{1}{\pi}\int\limits_{|t|\leqslant T}\frac{|f_h(t)-g_h(t)|}{|t|}dt+\frac{24A}{\pi T}\leqslant$$

$$\leqslant\frac{1}{\pi}\int\limits_{|t|\leqslant T}\frac{|f(t)-g(t)|}{|t|}dt+\frac{24A}{\pi T},\ \ A=\sup_{x\geqslant 0}|(G*H)'(x)|.$$

The latter inequality is valid for all T. Set $T(p) = cp^{\alpha}$, where $0 < \alpha < 1$. It is clear that the upper estimate includes two terms, the second of which has the order $O\left(\frac{1}{T}\right)$. The first term is $O(pT)$, which follows from the reasoning in [1], i.e., $\Delta_{p,r} = O(T^{-1}) + O(pT)$. It easy to see that the maximum convergence rate can be found by equating the terms' orders for $\alpha = \frac{1}{2}$.

Hence, when $p \to 0$, the value $\Delta_{p,r}$ has the order $O(\sqrt{p})$ irrespective of r from the interval $(0, 1)$.

3. Estimates in the uniform metric

3.1. Statement of the problem

It is necessary to estimate the expression

$$\sup_{x \geqslant 0} \left| P\left(\frac{N_{p,r}}{EN_{p,r}} \leqslant x \right) - G_{r,r}(x) \right|, \; p \to 0, \; r > 0,$$

where $N_{p,r}$ is a random variable, having negative binomial distribution with parameters r and p, and $G_{r,r}(x)$ is a gamma-distribution function with equal parameters. The first distribution function is piecewise constant, and the second one is continuous

$$P\left(\frac{N_{p,r}}{EN_{p,r}} \leqslant x \right) = \sum_{k=0}^{\left\lfloor \frac{rqx}{p} \right\rfloor} q^k p^r \frac{\Gamma(k+r)}{k! \cdot \Gamma(r)},$$

$$G_{r,r}(x) = \int_0^x \frac{r^r}{\Gamma(r)} e^{-rt} t^{r-1} dt. \qquad (*)$$

Because of their monotonicity, it is sufficient to study the behavior of $\Delta_{p,r}$ at the jump points of the negative binomial distribution function, i.e., at the points $\frac{np}{rq}$, $n \geqslant 0$, $n \in \mathbb{Z}$.

3.2. Estimation at the point

First let us try to obtain the estimate at a fixed point x for $r > 1$. Decompose the integral (*) into two parts

$$G_{r,r}(x) = \int_0^x \frac{r^r}{\Gamma(r)} e^{-rt} t^{r-1} dt = \int_0^{\frac{p}{rq}} \frac{r^r}{\Gamma(r)} e^{-rt} t^{r-1} dt + \int_{\frac{p}{rq}}^{\frac{np}{rq}} \frac{r^r}{\Gamma(r)} e^{-rt} t^{r-1} dt = I_1 + I_2.$$

Using the generalized mean value formula proved in the Appendix, we can state that

$$I_1 = \frac{r^r}{\Gamma(r)} e^{-r\frac{\theta p}{rq}} \int_0^{\frac{p}{rq}} t^{r-1} dt = \frac{p^r}{r\Gamma(r)} \frac{e^{-\frac{\theta p}{q}}}{q^r}, \quad 0 \leqslant \theta \leqslant 1.$$

Thus, for sufficiently small p

$$|I_1| \leqslant \frac{(2p)^r}{r\Gamma(r)} = C_1(r) p^r,$$

and the problem reduces to obtaining a similar estimate for the second term I_2. The main idea is to apply the formula of rectangles on the segment $[\frac{p}{rq}, \frac{np}{rq}]$ with the constant step $\frac{p}{rq}$, and to take the values of the integrand at the right boundaries of the partitioning segments

$$I_2 = \frac{p}{rq} \sum_{k=2}^{n} \frac{r^r}{\Gamma(r)} e^{-\frac{pk}{q}} \left(\frac{kp}{rq} \right)^{r-1} + R(r, p, x).$$

In what follows we denote the difference $G_{r,r}(x) - R(r,p,x)$ by $\widehat{G}_{r,r}(x)$. On the other hand,

$$\sum_{k=0}^{n} q^k p^r \frac{\Gamma(k+r)}{\Gamma(r)\cdot k!} = p^r + qp^r r + \sum_{k=2}^{n} q^k p^r \frac{\Gamma(k+r)}{\Gamma(r)\cdot k!}.$$

Denoting by Δ the expression $P(\frac{N_{p,r}}{\mathbf{E}N_{p,r}} \leqslant x) - G_{r,r}(x)$, we have

$$\Delta = \underbrace{p^r + qp^r r - \frac{p^r}{\Gamma(r)\cdot r}\frac{e^{-\frac{\theta p}{q}}}{q^r}}_{\Delta_1} +$$

$$+\sum_{k=2}^{n} \underbrace{\left(q^k p^r \frac{\Gamma(k+r)}{\Gamma(r)\cdot k!} - \frac{r^r}{\Gamma(r)} e^{-\frac{pk}{q}} \left(\frac{kp}{rq}\right)^{r-1} \cdot \frac{p}{rq}\right)}_{\Delta_k} + R(r,p,x). \tag{5}$$

The term Δ_1 is $O(p^r)$ with the constant depending only on r. Therefore, the key step is to obtain the upper estimate for Δ_k and their sum Δ. In Appendix A it is shown that

$$\frac{\Gamma(r+k)}{k!} = k^{r-1}\left(1 + \frac{r(r-1)}{2k} + O\left(\frac{1}{k^2}\right)\right),$$

where the constant depends only on r. Note that from the Taylor expansion with the Lagrange remainder term it follows that

$$(1-p)^{-r} = 1 + pr + \frac{r(r+1)p^2}{(1-\beta p)^{r+2}},$$

$$e^{-\frac{-pk}{q}} = e^{-pk(1-p)^{-1}} = e^{-pk(1+p+\frac{p^2}{(1-\gamma p)^3})} = e^{-pk-p^2k-\frac{p^3k}{(1-\gamma p)^3}},$$

$$(1-p)^k = (1-p)^{-\frac{1}{p}\cdot(-pk)} = e^{-pk-\frac{p^2k}{2}-\frac{p^3k}{3(1-\alpha p)^3}}.$$

Here α,β,γ are numbers from the segment $[0,1]$. Substituting this expressions into Δ_k and taking common factors out, we obtain

$$\Delta_k = \frac{p^r}{\Gamma(r)} e^{-pk} k^{r-1} \cdot \left(e^{-pk\left(\frac{p}{2}+\frac{p^2}{3(1-\alpha p)^3}\right)} \cdot \left(1 + \frac{r(r-1)}{2k} + O\left(\frac{1}{k^2}\right)\right) - \right.$$

$$\left. - \left(1 + pr + \frac{r(r+1)p^2}{(1-\beta p)^{r+2}}\right) e^{-pk\left(p+\frac{p^2}{(1-\gamma p)^3}\right)}\right). \tag{6}$$

Let us slightly simplify this expression. It is obvious that

$$e^{-p^2k-k\cdot O(p^3)} = 1 - p^2k - k\cdot O(p^3) + \frac{e^{\mu}}{2}\cdot(-p^2k - k\cdot O(p^3))^2 = 1 - p^2k + O(p^2). \tag{7}$$

However, it is less obvious that the constant in $O(p^3)$ may be taken not depending on k. Indeed, in the remainder we have the terms of the form $p^{2+t}k$, $t \geqslant 1$. We can write

$$p^{2+t}k = p^t\cdot(p^2k) \leqslant p^t\cdot p^2 n = p^t\cdot p^2\cdot\frac{xrq}{p} = p^t xqrp = O(p^2). \tag{8}$$

Similarly to relation (7),

$$e^{\frac{-p^2k}{2}+kO(p^3)} = 1 - \frac{p^2k}{2} + O(p^2), \tag{9}$$

where the constant also depends only on n. Now, using exponent expansions, we transform expression (6) as

$$\Delta_k = \frac{p^r}{\Gamma(r)} e^{-pk} k^{r-1} \left(\left(1 - \frac{p^2 k}{2} + O(p^2)\right) \left(1 + \frac{r(r-1)}{2k} + O\left(\frac{1}{k^2}\right)\right) - \right.$$

$$\left. -(1 + pr + O(p^2))(1 - p^2 k + O(p^2)) \right) = \frac{p^r}{\Gamma(r)} e^{-pk} k^{r-1} \left(\left(1 + \frac{r(r-1)}{2k} - \right.\right.$$

$$\left.\left. - \frac{p^2 k}{2} + O(p^2) + O\left(\frac{1}{k^2}\right)\right) - \left(1 + pr - p^2 k + O(p^2)\right) \right) =$$

$$= \frac{p^r}{\Gamma(r)} e^{-pk} k^{r-1} \left(-pr + \frac{p^2 k}{2} + \frac{r(r-1)}{2k} + O(p^2) + O\left(\frac{1}{k^2}\right)\right). \qquad (10)$$

Now let us estimate $\sum\limits_{k=2}^{n} \Delta_k$:

$$\left|\sum_{k=2}^{n} \Delta_k\right| \leqslant \sum_{k=1}^{n} |\Delta_k| \leqslant \sum_{k=1}^{n} \frac{p^r}{\Gamma(r)} e^{-pk} k^{r-1} \cdot \left(pr + \frac{p^2 k}{2} + \frac{r(r-1)}{2k} + c_2 p^2 + \frac{c_3}{k^2}\right) = \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV} + \mathrm{V}.$$

We have five terms, each of which will be considered separately. First,

$$\mathrm{I} = \frac{p^r}{\Gamma(r)} \sum_{k=1}^{n} e^{-pk} k^{r-1} pr \leqslant \frac{p^{r+1} \cdot r}{\Gamma(r)} n^{r-1} \cdot n \leqslant p x^r C_{\mathrm{I}}(r).$$

Similarly

$$\mathrm{II} = \frac{p^r}{2\Gamma(r)} \sum_{k=1}^{n} e^{-pk} k^r p^2 \leqslant C_{\mathrm{II}}(r) x^{r+1} p,$$

$$\mathrm{IV} \leqslant \ [\text{by analogy with I}] \ \leqslant p^r \cdot n^r \cdot O(p^2) \leqslant C_{\mathrm{IV}}(r) x^{r+d} p^2,$$

d is an integer nonnegative number determined from estimate (7).

The estimate for V follows from III, since $\frac{1}{k^2} \leqslant \frac{1}{k}$. The estimate for III is less trivial than the others and is based on the representation of the term in the form of the integral sum for a proper integral (convergent improper integral) with small p. Also note that there exists $\lambda > 0$: $p \geqslant \frac{\lambda}{n}$, since $n = \frac{rqx}{p} \geqslant \frac{1}{2}\frac{rqx}{p}$. So,

$$\mathrm{III} = \sum_{k=1}^{n} \frac{p^r}{\Gamma(r)} e^{-pk} k^{r-2} \frac{r^2 - r}{2} \leqslant$$

$$\leqslant \sum_{k=1}^{n} \frac{p^r}{\Gamma(r)} e^{-\lambda \frac{k}{n}} \left(\frac{k}{n}\right)^{r-2} n^{r-2} \left(\frac{1}{n} \cdot n\right) \frac{r^2 - r}{2} = \left[x_k = \frac{k}{n}, \Delta x_k = \frac{1}{n} \right] =$$

$$= \frac{p^r}{\Gamma(r)} n^{r-1} \frac{r^2 - r}{2} \cdot \sum_{k=1}^{n} f(\xi_k) \Delta x_k, \text{ where } f(x) = e^{-\lambda x} x^{r-2}.$$

The integral sum obviously converges to the integral $\int_0^1 e^{-\lambda x} x^{r-2} dx$. This integral is proper for $r \geqslant 2$ and convergent improper for $1 < r < 2$. Therefore, for such r it may be estimated by a constant not depending on p. It follows from the above that

$$\mathrm{III} \leqslant C_{\mathrm{III}}(r) p x^{r-1}.$$

Now from the obtained five estimates we may conclude that

$$|\sum_{k=2}^{n} \Delta_k| \leqslant C_0(x)p.$$

It may be shown (see Appendix B) that in our version of the method of rectangles for $r \geqslant 1$ $R(r,p,x) = O(p)$ (the constant does not depend on the chosen point and even on the segment). Moreover, $O(p^r) + O(p) = O(p)$ for $r > 1$. Therefore $\Delta = O(p)$ for the points $x = \frac{np}{rq}$. Since

$$\forall x\, \exists x_0 = \frac{n_0 p}{rq} \colon \Delta(x) \leqslant \Delta(x_0),$$

the estimate $O(p)$ is obtained for all x.

If now we assume that $0 < r \leqslant 1$, then all the considerations including expansion (6) remain valid. The only thing to be amended are the estimates I–V.

Using a technique similar to the one we used to estimate III, for estimates of I, II, and IV under new assumptions we obtain the same rate of convergence. For example, for the estimate of I:

$$\mathrm{I} = \frac{p^r}{\Gamma(r)} \sum_{k=1}^{n} e^{-pk} k^{r-1} pr \leqslant \left[p \geqslant \frac{\lambda}{n}\right] \leqslant \frac{p^{r+1} r}{\Gamma(r)} n^r \sum_{k=1}^{n} e^{-\frac{k\lambda}{n}} \left(\frac{k}{n}\right)^{r-1} \frac{1}{n} \leqslant C \int_0^1 e^{-\lambda t} t^{r-1} dt \cdot x^r p = O(p).$$

For the estimates of I, IV everything is similar:

$$\mathrm{I} = O(p),\ \mathrm{IV} = O(p).$$

The estimation of V, as was already noted, reduces to the estimation of III. For III, $0 < r < 1$, we have

$$\sum_{k=1}^{n} \frac{p^r}{\Gamma(r)} e^{-pk} k^{r-2} \frac{r^2 - r}{2} < \frac{p^r}{\Gamma(r)} \sum_{k=1}^{n} \frac{1}{k^{2-r}} \leqslant C'_{\mathrm{III}} \cdot p^r.$$

For $r = 1$ this estimate, obviously, should be used for V only. For $r \in (0,1]$ there is an estimate of the remainder at the point (and on the finite segment) of the form $O(p^r)$. Noting that $O(p^r)+O(p^r) = O(p^r)$ for $r \in (0,1]$, we have $\Delta = O(p^r)$.

3.3. Estimation on the finite segment

Consider $x \in [0,a]$. As before, we are interested in points of the form $\frac{np}{rq}$. Let us use the considerations of Section 3.2. Obviously,

$$\forall a > 0\, \exists (n_1, n_2) \colon \frac{n_1 p}{rq} \leqslant a \leqslant \frac{n_2 p}{rq}.$$

Replacing n by $n_2 > n$ and x with $a \geqslant x$ in the estimates for I–V, we obtain

$$\left|\sum_{k=2}^{n} \Delta_k\right| \leqslant \begin{cases} C'_0(a)p, & r > 1, \\ C''_0(a)p^r, & 0 < r \leqslant 1. \end{cases}$$

In Appendix B it is shown that

$$\sup_{x\in[0,\,a]} R(r,p,x) = \begin{cases} O(p), & r \geqslant 1, \\ O(p^r), & r \in (0,1]. \end{cases}$$

Therefore Δ has the order $O(p)$ $(O(p^r))$ on the whole segment $[0,a]$.

3.4. Uniform estimation on the whole axis

Let us recall that $G_{r,r}(x) = \widehat{G}_{r,r}(x) + R(r,p,x)$. Now we shall prove that the following estimate is valid:

$$\sup_{x\geqslant 0} |P(N^*_{p,r} \leqslant x) - \widehat{G}(x)| = \begin{cases} O(p), & r \geqslant 1, \\ O(p^r), & 0 < r < 1. \end{cases} \tag{11}$$

To do this, we first prove that for some a depending only on r,

$$\sup_{x> a} |P(N^*_{p,r} \leqslant x) - \widehat{G}(x)| = \begin{cases} O(p), & r \geqslant 1, \\ O(p^r), & 0 < r < 1. \end{cases} \tag{12}$$

Consider the function $\Delta(x) = P(N^*_{p,r} \leqslant x) - \widehat{G}(x)$. Using (6), we derive that

$$\Delta(x) = \frac{1}{\Gamma(r)} p^r \sum_{k=2}^{n} k^{r-1} \Big[\Big(1 + \frac{r(r-1)}{2k} + O\Big(\frac{1}{k^2}\Big)\Big) e^{-pk(1+\frac{p}{2}+\frac{p^2}{3(1-\alpha p)^3})} -$$

$$-\Big(1 + pr + r(r+1)\frac{p^2}{(1-\beta p)^{r+2}}\Big) e^{-pk\left(1+p+\frac{p^2}{(1-\gamma p)^3}\right)}\Big] + A(p,r), \text{ where } \alpha,\beta,\gamma \in [0,1].$$

Let us find out for which x the function $\Delta(x)$ is monotonically nondecreasing. It is obvious that all points of increase of the function $\Delta(x)$ are the points of the form $\frac{p}{rq}k$,

$$0 \leqslant k \leqslant n = \Big[\frac{xrq}{p}\Big].$$

Hence, $\Delta(x)$ is monotonically nondecreasing if and only if

$$\Delta(\frac{p}{rq}n) - \Delta(\frac{p}{rq}(n-1)) \geqslant 0. \tag{13}$$

Let us solve this inequality denoting $\frac{np}{rq}$ by x. We have

$$\Delta(\frac{p}{rq}n) - \Delta(\frac{p}{rq}(n-1)) = \frac{1}{\Gamma(r)} p^r n^{r-1} e^{-pn} \Big[\Big(1 + \frac{r(r-1)}{2n} + O\Big(\frac{1}{n^2}\Big)\Big) e^{-C_1(r,p)x} -$$

$$-\Big(1 + pr + r(r+1)\frac{p^2}{(1-\beta p)^{r+2}}\Big) e^{-C_2(r,p)x}\Big], \tag{14}$$

where

$$C_1(r,p) = rq\Big(\frac{p}{2} + \frac{p^2}{3(1-\alpha p)^3}\Big), \quad C_2(r,p) = rq\Big(p + \frac{p^2}{(1-\gamma p)^3}\Big).$$

First, let $0 < r < 1$. By the definition of $O\Big(\frac{1}{n^2}\Big)$, for sufficiently small p

$$\Big(1 + \frac{r(r-1)}{2n} + O\Big(\frac{1}{n^2}\Big)\Big) \geqslant \Big(1 + \frac{p(r-1)}{2xq}\Big) \geqslant (1 - \frac{p}{2xq}) \geqslant \Big(1 - \frac{p}{q}\Big) \tag{15}$$

for arbitrary $x \geqslant 1$. Therefore, all solutions of the inequality

$$e^{-C_1(r,p)x}\Big(1 - \frac{p}{q}\Big) - e^{-C_2(r,p)x}\Big(1 + pr + r(r+1)\frac{p^2}{(1-\beta p)^{r+2}}\Big) \geqslant 0 \tag{16}$$

are solutions of inequality (13). Let us rewrite (16) in the equivalent form:

$$(C_2(r,p)-C_1(r,p))x \geqslant \ln\Big(1+pr+r(r+1)\frac{p^2}{(1-\beta p)^{r+2}}\Big)-\ln\left(1-\frac{p}{q}\right).$$

Expanding both logarithms in the neighborhood of zero, we obtain

$$(C_2(r,p)-C_1(r,p))x \geqslant pr+\frac{p}{q}+O(p^2),$$

$$x \geqslant \frac{p(r+\frac{1}{q}+O(p))}{p(\frac{1}{2}+C_3p)rq+C_4p^2}, \text{ where } |C_3|\leqslant 2, |C_4|\leqslant 2.$$

The last passage is valid, since the value $C_2(r,p)-C_1(r,p)=p(\frac{1}{2}+C_3p)rq+C_4p^2$, by which we divide both sides of the inequality, is positive for sufficiently small p. Thus, when $p\to 0$, we have

$$x \geqslant 2\Big(\frac{r+1}{r}\Big)+O(p).$$

Note that $O(p)$ in the last expression depends on r, but for each fixed r this value is arbitrarily small with the appropriate choice of p. Then we have that for $x\geqslant a=2(\frac{r+1}{r})+1$ the function $\Delta(x)$ is monotonically nondecreasing.

The case $r\geqslant 1$ is considered similarly. Here $a=3$, which follows from the inequalities

$$\Big(1+\frac{r(r-1)}{2n}+O\Big(\frac{1}{n^2}\Big)\Big)\geqslant 1,$$

$$\Downarrow$$

$$x \geqslant \frac{\ln\Big(1+pr+r(r+1)\frac{p^2}{(1-\beta p)^{r+2}}\Big)}{p(\frac{1}{2}+C_3p)rq+C_4p^2}, \text{ where } |C_3|\leqslant 2, |C_4|\leqslant 2,$$

$$\Downarrow$$

$$x \geqslant \frac{p(r+O(p))}{prq(\frac{1}{2}+O(p))} \Rightarrow x\geqslant 2+O(p).$$

Note that $P(N^*_{p,r}\leqslant x)-G(x)\xrightarrow[x\to+\infty]{} 0$. But

$$|\Delta(x)|=|P(N^*_{p,r}\leqslant x)-\widehat{G}(x)|\leqslant |P(N^*_{p,r}\leqslant x)-G(x)|+R(p),$$

i.e., $\Delta(x)$ is bounded and nondecreasing. Hence,

$$\Delta(x)\xrightarrow[x\to+\infty]{}\theta R(p), \text{ where } |\theta|\leqslant 1.$$

Therefore,

$$\sup_{x>a}|P(N^*_{p,r}\leqslant x)-\widehat{G}(x)|=\sup_{x>a}|\Delta(x)|=\max(|\Delta(a)|,|\theta R(p)|)\leqslant$$
$$\leqslant \max(\sup_{x\in[0,a]}|P(N^*_{p,r}\leqslant x)-\widehat{G}(x)|,R(p)).$$

Using the remainder estimate obtained in the Appendix and the estimate on the segment, we derive

$$\sup_{x>a}|P(N^*_{p,r}\leqslant x)-\widehat{G}(x)|=\begin{cases}O(p), & r\geqslant 1,\\ O(p^r), & 0<r<1,\end{cases}$$

i.e., (12) is proved, and (11) immediately follows from equality (12).

Using (11) and the fact that

$$\sup_{x\geqslant 0}|P(N^*_{p,r}\leqslant x)-G(x)|\leqslant \sup_{x\geqslant 0}|P(N^*_{p,r}\leqslant x)-\widehat{G}(x)|+R(p),$$

we finally obtain the uniform estimate

$$\sup_{x\geqslant 0}|P(N^*_{p,r}\leqslant x)-G(x)|=\begin{cases}O(p), & r\geqslant 1,\\ O(p^r), & 0<r<1.\end{cases}\tag{17}$$

4. Conclusion

Let us finally return to the paper [1] and Theorem 2.1 mentioned in the Introduction. Now we are able without significant modifications of the proof to obtain the following new result:

Theorem 2.1'. *Let T_n be asymptotically normal, and let* (1) *hold. Then the following relations are valid*:

$$\sup_x|P(\sqrt{n-r}\,\sigma(T_{N_n}-\mu)\leqslant x)-F_{2r}(x)|=\begin{cases}O\left(\dfrac{1}{\sqrt{n}}\right), & r\in[\frac12,1],\\ O\left(\dfrac{\ln n}{\sqrt{n}}\right), & r=\frac12,\\ O\left(\dfrac{1}{n^r}\right), & 0<r<\frac12.\end{cases}$$

Appendix

A. Asymptotic expansion for $\frac{\Gamma(k+r)}{k!}$

Let us use the Stirling formula in the following form:

$$\Gamma(\lambda+1)=\sqrt{2\pi\lambda}\left(\frac{\lambda}{e}\right)^{\lambda}\left(1+\frac{1}{12\lambda}+O\left(\frac{1}{\lambda^2}\right)\right),\ \lambda\to\infty.$$

Then

$$\frac{\Gamma(k+r)}{k!}=\frac{\sqrt{2\pi(k+r-1)}\left(\frac{k+r-1}{e}\right)^{k+r-1}\left(1+\frac{1}{12(k+r-1)}+O\left(\frac{1}{k^2}\right)\right)}{\sqrt{2\pi k}\left(\frac{k}{e}\right)^{k}\left(1+\frac{1}{12k}+O\left(\frac{1}{k^2}\right)\right)}.$$

Using the expansion

$$\left(1+\frac{1}{12k}+O\left(\frac{1}{k^2}\right)\right)^{-1}=1-\frac{1}{12k}+O\left(\frac{1}{k^2}\right),$$

we derive that

$$\frac{\Gamma(k+r)}{k!}=\left(\frac{k+r-1}{k}\right)^{k+\frac12}\left(\frac{k+r-1}{e}\right)^{r-1}\left(1+O\left(\frac{1}{k^2}\right)\right).$$

Next,

$$(k+r-1)^{r-1}=k^{r-1}\left(1+\frac{r-1}{k}\right)^{r-1}=k^{r-1}\left(1+\frac{(r-1)^2}{k}+O\left(\frac{1}{k^2}\right)\right),$$

$$\left(\frac{k+r-1}{k}\right)^{k+\frac12}=e^{\ln\left(1+\frac{r-1}{k}\right)\left(k+\frac12\right)}=\ =e^{r-1}\left(1+\frac{3r-2-r^2}{2k}+O\left(\frac{1}{k^2}\right)\right).$$

Thus,

$$\left(\frac{k+r-1}{k}\right)^{k+\frac12}\left(\frac{k+r-1}{e}\right)^{r-1}=k^{r-1}\left(1+\frac{(r-1)^2}{k}+O\left(\frac{1}{k^2}\right)\right)\times$$
$$\times\left(1+\frac{3r-2-r^2}{2k}+O\left(\frac{1}{k^2}\right)\right)=k^{r-1}\left(1+\frac{r(r-1)}{2k}+O\left(\frac{1}{k^2}\right)\right).$$

B. Estimates for the remainder $R(r,p,x)$**,** $r \geqslant 1$**,** $x \in \mathbf{R}$

Let us represent $R(r,p,x)$ in the form of the sum of remainders on each decomposition part of the segment $[0,x]$:

$$R(r,p,x) = \frac{r^r}{\Gamma(r)} \sum_{k=2}^{n} \left(\int\limits_{\frac{(k-1)p}{rq}}^{\frac{kp}{rq}} e^{-rt} t^{r-1} dt - \frac{p}{rq} \left(\frac{kp}{rq}\right)^{r-1} e^{-\frac{pk}{q}} \right).$$

Using the mean value theorem and the Lagrange theorem, we have

$$R_k(r,p,x) = e^{-\frac{p}{q}(k-1+\theta_k)} \left(\frac{p}{rq} \cdot (k-1+\theta_k)\right)^{r-1} \cdot \frac{p}{rq} - \frac{p}{rq} \cdot \left(\frac{kp}{rq}\right)^{r-1} e^{-\frac{pk}{q}} =$$

$$= \frac{p}{rq} \cdot \left(e^{-\frac{p(k-1+\theta_k)}{q}} \left(\frac{p(k-1+\theta_k)}{rq}\right)^{r-1} - \left(\frac{pk}{rq}\right)^{r-1} e^{-\frac{kp}{q}} \right) =$$

$$= [0 \leqslant \theta_k \leqslant 1] = \left(\frac{p}{rq}\right)^2 \cdot \underbrace{(-1+\theta_k)}_{\geqslant -1} (e^{-rt} t^{r-1})' \big|_{t=\xi_k}.$$

Here the point ξ_k lies between the points $\frac{kp}{rq}$ and $\frac{p(k-1+\theta_k)}{rq}$, i.e., also on the partial segment $[\frac{(k-1)p}{rq}, \frac{kp}{rq}]$. Note that the derivative of $e^{-rx}x^{r-1}$ equals

$$e^{-rx}(-rx^{r-1} + (r-1)x^{r-2}).$$

Therefore

$$|R(r,p,x)| \leqslant \frac{r^r}{\Gamma(r)} \sum_{k=1}^{n} \left(\frac{p}{rq}\right)^2 \left(e^{-r\xi_k} \left(r\xi_k^{r-1} + (r-1)\xi_k^{r-2} \right) \right).$$

Recalling that $n = \frac{xrq}{p}$, on the right-hand side of the inequality we can see the integral sum. Hence,

$$\frac{|R(r,p,x)|}{p} \leqslant \frac{r^r}{\Gamma(r)} \frac{p}{r^2 q^2} \sum_{k=1}^{n} e^{-r\xi_k} \left[r\xi_k^{r-1} + (r-1)\xi_k^{r-2} \right] \leqslant$$

$$\leqslant C \cdot \int\limits_0^x e^{-ry}(ry^{r-1} + (r-1)y^{r-2})\, dy \leqslant C \cdot \int\limits_0^\infty e^{-ry}(ry^{r-1} + (r-1)y^{r-2})\, dy.$$

The integral on the right-hand side converges for $r \geqslant 1$. Therefore in this case $|R(r,p,x)| \leqslant R(p) = O(p)$, $0 < r < 1$, $x \in [0,a]$, and

$$R_k(r,p,x) = \int\limits_{\frac{(k-1)p}{rq}}^{\frac{kp}{rq}} e^{-rt} t^{r-1} dt - \frac{p}{rq} \left(\frac{kp}{rq}\right)^{r-1} e^{-\frac{pk}{q}} = e^{-\frac{p}{q}(k-1+\theta_k)} \int\limits_{\frac{(k-1)p}{rq}}^{\frac{kp}{rq}} t^{r-1} dt - \frac{p}{rq} \left(\frac{kp}{rq}\right)^{r-1} e^{-\frac{pk}{q}}.$$

Since

$$e^{-\frac{pk}{q} + \frac{p}{q}(1-\theta_k)} = e^{-\frac{pk}{q}} e^{\frac{p}{q} - \frac{p}{q}\theta_k} \leqslant e^{-\frac{pk}{q}} e^{\frac{p}{q}} = e^{-\frac{pk}{q}} \left(1 + \frac{p}{q} + O(p^2)\right),$$

using the Lagrange theorem, we obtain

$$|R_k(r,p,x)| \leqslant e^{-\frac{kp}{q}} \left(\int\limits_{\frac{(k-1)p}{rq}}^{\frac{kp}{rq}} t^{r-1} dt \left(1 + \frac{p}{q} + C_1 p^2\right) - \left(\frac{p}{rq}\right)^r k^{r-1} \right) \leqslant$$

$$\leqslant (k^r-(k-1)^r)\cdot\left(\frac{p}{rq}\right)^r\frac{1}{r}\left(1+\frac{p}{q}+C_1p^2\right)-\left(\frac{p}{rq}\right)^r k^{r-1}\leqslant$$

$$\leqslant ((k-1)^{r-1}C_2p+(k-1)^{r-1}-k^{r-1})\cdot\left(\frac{p}{rq}\right)^r\leqslant$$

$$\leqslant [\text{ by Lagrange theorem }]\leqslant (C_2p(k-1)^{r-1}+(1-r)(k-1)^{r-2})\cdot\left(\frac{p}{rq}\right)^r\leqslant$$

$$\leqslant [(k-1)^{r-1}\leqslant 1]\leqslant C_3p^r(k-1)^{r-2}+C_4p^{r+1}.$$

Summing over k from 2 to $\frac{rqx}{p}$, we obtain for $0<r<1$, $x\geqslant a$

$$|R(r,p,x)|\leqslant C_3'p^r\sum_{k=2}^{n}\frac{1}{(k-1)^{2-r}}+C_4'(n-1)p^{r+1}\leqslant C(a)p^r.$$

Let us split the sum, representing the remainder, into two parts:

$$R(r,p,x)=\frac{r^r}{\Gamma(r)}\sum_{k=2}^{\left\lfloor\frac{rqa}{p}\right\rfloor}\left(\int\limits_{\frac{(k-1)p}{rq}}^{\frac{kp}{rq}}e^{-rt}t^{r-1}dt-\frac{p}{rq}\left(\frac{kp}{rq}\right)^{r-1}e^{-\frac{pk}{q}}\right)+$$
$$+\frac{r^r}{\Gamma(r)}\sum_{k=\left\lfloor\frac{rqa}{p}\right\rfloor+1}^{\left\lfloor\frac{rqx}{p}\right\rfloor}\left(\int\limits_{\frac{(k-1)p}{rq}}^{\frac{kp}{rq}}e^{-rt}t^{r-1}dt-\frac{p}{rq}\left(\frac{kp}{rq}\right)^{r-1}e^{-\frac{pk}{q}}\right).$$

Obviously, the first part may be estimated similarly to the previous case, and the second one similarly to the case $r\geqslant 1$ (the upper estimate with the use of the integral $\int_a^\infty e^{-ry}(ry^{r-1}+(r-1)y^{r-2})\,dy$). Thus, we obtain the upper estimate of the form $C(a)p^r+Cp=O(p^r)$, where a is fixed in the proof of the uniform estimate on the axis.

C. Generalized mean value formula

Lemma (generalization of the theorem on continuous function, taking all intermediate values). *Let $f(x)\in C((a,b])$. Then for $\forall\mu$ strictly between infimum and supremum of the function on $(a,b]$ (finite or infinite) at some point $\xi\in(a,b]$ we have $\mu=f(\xi)$.*

Proof. Denote $M=\sup\limits_{(a,b]}f(x)$, $m=\inf\limits_{(a,b]}f(x)$. Consider, for example, the infimum. There may be two cases.

1. If $\exists[a_0,b]\subset(a,b]$: $\inf\limits_{(a,b]}f(x)=\inf\limits_{[a_0,b]}f(x)$, then we may state (using the second Weierstrass theorem) that $\exists\xi\in[a_0,b]:\ f(\xi)=\inf\limits_{[a_0,b]}f(x)=m$. Therefore for arbitrary sequence $x_n\to\xi,\ x_n\in[a_0,b],\ x_n\in(a,b],\ f(x_n)\to m$.

2. If for $\forall[a_0,b]\subset[a,b]$ $\inf\limits_{[a_0,b]}f(x)>\inf\limits_{(a,b]}f(x)$, then we take $a_0=a+\frac{1}{n}$ and denote $\inf\limits_{[a_0,b]}f(x)=m_n,\ m_n=m+\varepsilon_n$. By the properties of infimum for $\varepsilon_n'=\frac{\varepsilon_n}{2}$ $\exists x_n'$: $x_n'\in(a,b]$ and $f(x_n')<\frac{\varepsilon_n}{2}+m<m_n$. From the latter inequality it follows that $x_n'\in(a,a+\frac{1}{n})$, i.e., $x_n'\xrightarrow[n\to\infty]{}a$. On the other hand, $m\leqslant f(x_n')<m+\frac{\varepsilon_n}{2},\ \varepsilon_n\to 0$, since $m=\inf_n m_n$. Thus, in this case also $\exists x_n'\in(a,b]$: $f(x_n')\to m$.

Note that in the second case, as opposed to the first one, it as allowed that $m=-\infty$.

The case $M=\sup\limits_{(a,b]}f(x)$ is considered similarly. Summing up, for both infimum and supremum there exist sequences from the interval $(a,b]$ such that $f(x_n)\to m,\ f(x_n')\to M$. Therefore, for $\forall\mu$ strictly between m and M, $m \geqslant -\infty$, $M \leqslant +\infty$. there exists a segment whose boundaries are the elements of the above sequences and

$$\inf_{[y,z]} f(x) \leqslant \mu \leqslant \sup_{[y,z]} f(x).$$

But a function continuous on a *segment* takes all intermediate values on it, i.e., $\mu = f(\xi)$, $\xi \in [y, z]$. This concludes the proof of lemma.

Theorem. *Let a function $f(x)$ be nonconstant and continuous on a bounded interval $(a, b]$ and let the improper integral of the second kind $\int\limits_a^b f(x)\,dx$ converge. Then*

$$\exists \xi \in (a, b]\colon\ f(\xi)(b-a) = \int\limits_a^b f(x)\,dx.$$

Note that function $f(x)$ may be bounded or unbounded, and it may be unbounded in both directions (for example, $\sin\frac{1}{x}$, $x^{r-1}\sin\frac{1}{x}$, $0 < r < 1$, $(0, 1]$).

Denote by μ the value $\frac{1}{b-a} \cdot \int_a^b f(x)\,dx$. Obviously, it is between the infimum and supremum of the function $f(x)$ on $(a, b]$, and $m < \mu < M$. Otherwise $f \equiv const$. Therefore by the above lemma we have $\mu = f(\xi)$.

Acknowledgements

This research is supported by the Russian Scientific Fund, project 14–11–00364.

REFERENCES

1. V. E. Bening, V.Yu. Korolev, and U Da, "Estimates for speed of convergence of certain statistics distributions to student distribution," *Herald Russian University of Peoples' Friendship, "Applied Mathematics and Informatics" Series*, **12**, No. 1, 59–74 (2004).
2. H. Bevrani, V. E. Bening, and V.Yu. Korolev, "On the Accuracy of Approximation of the Negative Binomial Distribution by the Gamma Distribution and Convergence Rate of the Distributions of Some Statistics to the Student Distribution," *J. Math. Sci.*, **205**, No. 1, 34–44 (2015).
3. D. O. Selivanova, *Convergence Rate Estimates for Random Sums*, Ph.D. thesis in Phys. and Math. Sci. MSU, Moscow (1995).