



# Article On Approximation of the Tails of the Binomial Distribution with These of the Poisson Law

Sergei Nagaev <sup>1,†</sup> and Vladimir Chebotarev <sup>2,\*,†</sup>

- <sup>1</sup> Sobolev Institute of Mathematics, 630090 Novosibirsk, Russia; nagaev@math.nsc.ru
- <sup>2</sup> Computing Center, Far Eastern Branch of the Russian Academy of Sciences, 680000 Khabarovsk, Russia
- \* Correspondence: vladimir.ch@ccfebras.ru
- + These authors contributed equally to this work.

**Abstract:** A subject of this study is the behavior of the tail of the binomial distribution in the case of the Poisson approximation. The deviation from unit of the ratio of the tail of the binomial distribution and that of the Poisson distribution, multiplied by the correction factor, is estimated. A new type of approximation is introduced when the parameter of the approximating Poisson law depends on the point at which the approximation is performed. Then the transition to the approximated binomial law is carried out. In both cases error estimates are obtained. A number of conjectures are made about the refinement of the known estimates for the Kolmogorov distance between binomial and Poisson distributions.

Keywords: binomial distribution; poisson approximation; esscher transformation



**Citation:** Nagaev, S.; Chebotarev, V. On Approximation of the Tails of the Binomial Distribution with These of the Poisson Law. *Mathematics* **2021**, *9*, 845. https://doi.org/10.3390/ math9080845

Academic Editor: Irina Shevtsova

Received: 11 March 2021 Accepted: 9 April 2021 Published: 13 April 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 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 and Main Results

The subject of this study is upper and lower bounds for probabilities of the type  $\mathbf{P}(\sum_{i=1}^{n} X_i \ge nx)$ , where  $X_1, \ldots, X_n$  are independent equally distributed Bernoulli random variables. In other words, we estimate tail probabilities for the binomial distribution. To this end we use the Poisson approximation.

It should be noted that although the binomial distribution is very special from the formal point of view it is of great concern in applications. Moreover, due to simplicity, more exact bounds are attainable for the binomial distribution than in the general case.

Let us start with the known Hoeffding inequality. Assuming that the independent random variables  $X_1, \ldots, X_n$  satisfy the condition  $0 \le X_i \le 1, i = 1, \ldots, n$ , W. Hoeffding [1] deduced the inequality

$$\mathbf{P}\Big(\sum_{i=1}^{n} X_i \ge n(\mu+t)\Big) \le \Big(\frac{\mu}{\mu+t}\Big)^{n(\mu+t)} \Big(\frac{1-\mu}{1-\mu-t}\Big)^{n(1-\mu-t)},\tag{1}$$

where  $\mu = \frac{1}{n} \sum_{i=1}^{n} \mathbf{E}X_i$ ,  $0 < t < 1 - \mu$ . In the case of identically distributed random variables  $X_j$  we have  $\mu = \mathbf{E}X_1$ , and the inequality (1) remains the same. Making in (1) the change of variable  $n(\mu + t) = y$  we get

$$\mathbf{P}\Big(\sum_{i=1}^{n} X_i \ge y\Big) \le \Big(\frac{n\mu}{y}\Big)^y \Big(\frac{1-\mu}{1-y/n}\Big)^{n(1-y/n)}$$

In turn this inequality can be written in the following form,

$$\mathbf{P}\Big(\sum_{i=1}^{n} X_i \ge y\Big) \le e^{-nH(y/n,\mu)},\tag{2}$$

where

$$H(t, p) = t \ln \frac{t}{p} + (1 - t) \ln \frac{1 - t}{1 - p}$$

is the so-called relative entropy or Kullback–Leibler distance between two two-point distributions (t, 1 - t) and (p, 1 - p) concentrated at the same pair of points.

Apparently, I. Sanov [2] was the first who stated probability inequalities in terms of the function of the type  $\sum_{j=1}^{m} t_j \ln \frac{p_j}{t_j}$ , where  $\sum_{j=1}^{m} p_j = \sum_{j=1}^{m} t_j = 1$ ,  $p_j$ ,  $t_j > 0$ , j = 1, ..., m.

The starting point in proving (1) and many other probability inequalities for independent random variables is the following bound.

Let there exist  $H_0 > 0$  such that

$$\int_{-\infty}^{\infty} e^{H_0 u} \, dV_j(u) < \infty, \quad j = 1, \dots n, \tag{3}$$

where  $V_j$  are the distribution functions of  $X_j$ , j = 1, ..., n. Then for every  $0 < h \le H_0$ , we have

$$\mathbf{P}\Big(\sum_{j=1}^n X_j \ge y\Big) \le e^{-hy} \prod_{j=1}^n R(h; V_j),$$

where

$$R(h; V_j) := \int_{-\infty}^{\infty} e^{hu} \, dV_j(u), \quad j = 1, \dots n.$$

$$\tag{4}$$

Thus,

$$\mathbf{P}\Big(\sum_{j=1}^{n} X_j \ge y\Big) \le \min_{h>0} e^{-hy} \prod_{j=1}^{n} R(h; V_j).$$
(5)

In the case of i. i. d. random variables inequality (5) can be written in the following form,

$$1 - G_n(y) \le \min_{h>0} e^{-hy} R^n(h; G), \tag{6}$$

where  $G_n(y) = \mathbf{P}(\sum_{j=1}^n X_j < y)$ , *G* is the distribution of  $X_1$ . On the other hand, for each  $0 < h \le H_0$  the following identity holds,

$$1 - G_n(y) = R^n(h;G) \int_y^\infty e^{-hu} \, dG_n^{(h)}(u), \tag{7}$$

where

$$G_n^{(h)}(y) = R^{-n}(h;G) \int_{-\infty}^{y} e^{hu} \, dG_n(u)$$
(8)

is the Esscher transformation of the distribution function  $G_n(y)$  (see [3]). Note that starting with the classic work of Cramér [4], Esscher's transform has been repeatedly used in the theory of large deviations.

Let  $h_0$  be such that

$$h_0 y - n \ln R(h_0; G) = \max_{h>0} \left( hy - n \ln R(h; G) \right), \tag{9}$$

and denote

$$\omega_n(y;G) = \int_y^\infty e^{-h_0 u} \, dG_n^{(h_0)}(u). \tag{10}$$

It follows from (7) and (10) that

$$1 - G_n(y) = R^n(h_0; G) \,\omega_n(y; G).$$
(11)

Notice, that although the method used in this work essentially coincides with the method of our previous article on estimates of large deviations in the case of normal

approximation [5], function  $\omega_n(y; G)$  differs from function  $\omega_n(y; G)$  from [5] by the absence of the factor  $e^{h_0 y}$ . The nuance is that in this work we are dealing with one-way distributions, and direct copying of the previous approach could make the reasoning unnecessarily cumbersome.

Taking into account (9) it is easily seen that  $h_0$  satisfies the equality  $m(h_0; G) = \frac{y}{n}$ , where

$$m(h;G) = \int_{-\infty}^{\infty} u \, dG^{(h)}(u) \equiv R^{-1}(h;G) \int_{-\infty}^{\infty} u e^{hu} \, dG(u).$$

For any nondegenerate random variable  $\xi$  we have  $\mathbf{P}(\xi < \mathbf{E}\xi) > 0$ . Therefore,  $\omega_n(y;G) < 1$ . Estimating  $\omega_n(h_0;G)$ , we can sharpen the Hoeffding inequality.

Note that in the case  $\mathbf{E}X_1 = 0$  the asymptotics of  $\omega_n(y; G)$  is found in [4] (p. 172) under condition  $\int_{-\infty}^{\infty} e^{hy} dG(y) < \infty, 0 < h \le H_0$ , namely,

$$\omega_n(y;G) \sim e^{\frac{y^2}{2n\sigma^2}} \left[1 - \Phi\left(\frac{y}{\sigma\sqrt{n}}\right)\right] \equiv M\left(\frac{y}{\sigma\sqrt{n}}\right),$$

where  $\sigma^2 = \mathbf{E}X_1^2$ , the restriction y = o(n) being imposed (see also [6]), and

$$M(t) = \frac{1 - \Phi(t)}{\varphi(t)} \tag{12}$$

is the so called Mills ratio ( $\Phi(t)$  and  $\varphi(t)$  are the distribution function and density function, respectively, of the standard normal law).

Let  $\lambda$  be an arbitrary positive number,  $\Pi_{\lambda}(y)$  the distribution function of the Poisson law with the mean  $\lambda$ . We will also use the notation  $\pi_{\lambda}(j) = \frac{\lambda^j}{j!} e^{-\lambda}$ . Note that we consider distribution functions to be continious from left.

In connection with (12), note that in the present work we define and use the following analogue of the Mills ratio for the Poisson distribution with an arbitrary parameter  $\lambda$ : for every integer  $k \ge 0$ 

$$M(k;\lambda) := \frac{1 - \Pi_{\lambda}(k)}{\pi_{\lambda}(k)} \equiv \frac{1}{\pi_{\lambda}(k)} \sum_{j=k}^{\infty} \pi_{\lambda}(j) = 1 + \sum_{m=1}^{\infty} \frac{\lambda^{m} k!}{(m+k)!}$$
$$= 1 + \sum_{m=1}^{\infty} \lambda^{m} \prod_{j=1}^{m} \frac{1}{k+j}.$$
(13)

M. Talagrand [7] sharpened the Hoeffding inequalities for  $y < \frac{n\sigma^2}{Kb}$ , where *K* is a constant, regarding which is known only that it exists. The bounds obtained in [7] are stated in terms of *K* as well.

Remark that Talagrand, like Hoeffding, considers the case of non-identically distributed random variables.

In the present work we estimate  $\omega_n(y; G)$  in the case of Bernoulli trials with explicit values for constants, not laying any restriction on *y*.

In what follows we use the next notations: *F* the distribution function of the Bernoulli random variable with parameter p,  $0 , <math>F_{n,p} = F^{*n}$  the *n*-fold convolution of *F*. Obviously,

$$R(h;F) = r(h) := q + pe^{h}.$$
 (14)

In what follows we will assume *x* to satisfy the following condition,

$$0$$

It is not hard to verify that  $h_0$  satisfying (9) in the case G = F and y = nx has the following form,

$$h_0 = \ln \frac{qx}{p(1-x)}.$$
 (16)

Notice that  $h_0 > 0$  under condition (15). In what follows  $h = h_0$ . We get from (14) and (16) that

$$r(h) = \frac{q}{1-x}.$$
(17)

Hence, by (11),

$$1 - F_{n,p}(nx) = r^n(h)\,\omega_n(nx;F) = \left(\frac{q}{1-x}\right)^n \omega_n(nx;F),\tag{18}$$

where

$$\omega_n(nx;F) = \int_{nx}^{\infty} e^{-hu} \, dF_{n,p}^{(h)}(u). \tag{19}$$

Denote by  $\Pi_{\lambda}(t)$  the distribution function of Poisson law with a parameter  $\lambda > 0$ . If the variable *x* from (18) approaches 0, it is natural to take  $\Pi_{\lambda}$  with  $\lambda = np$  as the approximating distribution for  $F_{n,p}$ . Just this distribution is used in Theorem 2. However, first we need another approximating Poisson distribution with the mean

$$\lambda_1 = \lambda_1(n, p, x) = \frac{np(1-x)}{q},$$
(20)

depending not only on the parameters n and p, but on the variable x from formula (15). We shall call this distribution by *the variable Poisson distribution*.

Let us formulate the first statement about the connection between the behaviors of tails  $1 - F_{n,p}(nx)$  and  $1 - \prod_{\lambda_1}(nx)$ . First introduce the function

$$A(x,n,p) = \left(\frac{1-x}{q}\right)^{-n} e^{-\frac{n(x-p)}{q}}.$$
(21)

We have

$$A(x,n,p) = \left(1 - \frac{x-p}{q}\right)^{-n} e^{-\frac{n(x-p)}{q}} = e^{-n\left[\ln\left(1 - \frac{x-p}{q}\right) + \frac{x-p}{q}\right]} = e^{-n\left[\ln(1-u) + u\right]},$$
 (22)

where  $u = \frac{x-p}{q}$ . Function  $\left[-\ln(1-u) - u\right]$  is presented as a series:

$$-\ln(1-u) - u = \sum_{k=2}^{\infty} \frac{u^k}{k} =: \Lambda_2(u).$$

Thus,

$$A(x,n,p) = e^{n\Lambda_2(u)}.$$
(23)

Note that the series  $\Lambda_2(u)$  converges since by condition (15), we have  $0 < \frac{x-p}{q} < 1$ .

**Proposition 1.** If condition (15) is fulfilled, then

$$1 - F_{n,p}(nx) = \left[1 - \Pi_{\lambda_1}(nx)\right] A(x,n,p) + R_1 = \left[1 - \Pi_{\lambda_1}(nx)\right] e^{n\Lambda_2(u)} + R_1,$$
(24)

where

$$|R_1| \le 2e^{-nH(p,x)} \max_{y \ge nx} \left| F_{n,x}(y) - \Pi_{nx}(y) \right| \le 2xe^{-nH(x,p)}.$$
(25)

The following theorem gives one more form of the dependence of the tails of the binomial distribution on the tails  $1 - \prod_{\lambda_1}(nx)$  of the variable Poisson distribution. It is a consequence of Proposition 1, but by no means trivial, and requires the proof of a number of additional statements, which are given in Section 3.

**Theorem 1.** If condition (15) is fulfilled, then

$$1 - F_{n,p}(nx) = [1 - \Pi_{\lambda_1}(nx)] A(x, n, p) (1 + r_1)$$
  
=  $[1 - \Pi_{\lambda_1}(nx)] e^{n\Lambda_2(u)} (1 + r_1(x)),$  (26)

where  $u = \frac{x-p}{q}$ ,

$$|r_1(x)| \le c_1 \sqrt{nx^3}, \quad c_1 = 2e^{1/12} \sqrt{2\pi} = 5.4489 \dots$$
 (27)

**Example 1.** Let n = 500, p = 0.002, x = kp,  $k = \overline{2, 5}$ . Table 1 shows the corresponding values of the function  $c_1\sqrt{nx^3}$ .

**Table 1.** Values of the function  $c_1\sqrt{nx^3}$  for n = 500, p = 0.002, x = kp,  $k = \overline{2,5}$ .

<i>x</i>	0.004	0.006	0.008	0.01
nx	2	3	4	5
$c_1\sqrt{nx^3}$	0.0308	0.0566	0.0871	0.1218

Table 1, in accordance with Theorem 1, shows that with increasing x the approximation deteriorates.

**Remark 1.** It is known that the binomial distribution with parameters n, p is well approximated by the Poisson one with the parameter np if p is small enough [8]. The Poisson distribution from the equalities (24) and (26) has another parameter. However, we have  $\lambda_1 = \frac{np(1-x)}{q} \approx np$ , when x is close to 0 and p < x. In the next claims we consider the Poisson approximation with the parameter np. Note also that the Poisson distribution with parameter  $\lambda_1$  degenerates when x is close to 1. See also Table 2.

**Remark 2.** A necessary condition for good approximation in (26) is the smallness of x, namely,  $x < \theta n^{-1/3}$ . This agrees with the result by Yu.V. Prokhorov [9], according to which in the case  $x < \theta n^{-1/3}$  ( $\theta = 0.637$ ) Poisson approximation to the binomial distribution is more precise with respect to the normal approximation. However, as x is close to 0,  $\frac{\lambda_1}{\lambda} \approx 1$  ( $\lambda = np$ ). In this case,  $\lambda$  can be both large and small. This also applies to the values of nx. Note that  $\frac{d}{d\lambda}(1 - \Pi_{\lambda}(k)) > 0$  for any  $k \ge 1$ . Indeed, it is easy to see that  $\frac{d}{d\lambda}\Pi_{\lambda}(k) = \Pi_{\lambda}(k-1) - \Pi_{\lambda}(k) = -\pi_{\lambda}(k-1) < 0$ . Therefore,  $\frac{d}{d\lambda}(1 - \Pi_{\lambda}(k)) > 0$ . This means that  $1 - \Pi_{\lambda_1}(k) < 1 - \Pi_{\lambda}(k)$  for all  $k \ge 1$ .

**Theorem 2.** If condition (15) is fulfilled, then the following equality holds,

$$\frac{1 - F_{n,p}(nx)}{1 - \prod_{np}(nx)} = \frac{M(nx;\lambda_1)}{M(nx;np)} e^{-nq\Lambda_3\left(\frac{x-p}{q}\right)} (1 + r_1(x)),$$
(28)

where

$$\Lambda_3(u) = \sum_{k=2}^{\infty} \frac{u^k}{k(k-1)},$$

 $r_1(x)$  is the function from Theorem 1.

**Remark 3.** It follows from Remark 2 that if in the representation (26) the difference  $1 - \prod_{\lambda_1}(nx)$  is replaced by  $1 - \prod_{\lambda}(nx)$ , where  $\lambda = np$ , then instead of the function A(x, n, p), it will be

necessary to insert another correction factor, which will be less than A(x, n, p). The form of this factor is indicated in Theorem 2. In this connection, we note that the exponential function on the right-hand side of (28) has a negative exponent, in contrast to the exponential function in (26).

The following table gives an idea of the relationship between tails of the approximating distributions under consideration:  $\Pi_{\lambda_1}$  and  $\Pi_{\lambda}$ .

x	0.2	0.3	0.4	0.5	0.6	0.7	0.8
nx	2	3	4	5	6	7	8
$1 - \Pi_{\lambda_1}(nx)$	0.223	0.044	0.004	0.0002	$7.32 \times 10^{-6}$	$6.78 \times 10^{-8}$	$1.2 \times 10^{-10}$
$1 - \Pi_{\lambda}(nx)$	0.264	0.08	0.018	0.003	0.0005	0.00008	0.00001
$\frac{1-\Pi_{\lambda_1}(nx)}{1-\Pi_{\lambda}(nx)}$	0.845	0.551	0.255	0.076	0.0123	0.0008	$1.18 \times 10^{-5}$

**Table 2.** Values of the ratio  $\frac{1-\Pi_{\lambda_1}(nx)}{1-\Pi_{\lambda}(nx)}$  for n = 10, p = 0.1, x = kp,  $k = \overline{2,8}$ .

By  $\theta$  we will denote quantities, maybe different in different places, satisfying the bound  $|\theta| \leq 1$ .

Rewrite (28) in the form:

$$1 - F_{n,p}(nx) = \left(1 - \Pi_{np}(nx)\right) \Omega_1(x, n, p) e^{-nq\Lambda_3\left(\frac{x-p}{q}\right)} (1 + r_1(x)),$$
(29)

where  $\Omega_1(x, n, p) = \frac{M(nx;\lambda_1)}{M(nx;np)}$ .

Let us give a table of values of the functions:  $M(nx; \lambda_1)$ , M(nx; np) and  $\Omega_1(x, n, p) \equiv \frac{M(nx;\lambda_1)}{M(nx;np)}$ . Let n = 10, p = 0.1,  $x = \frac{k}{n}$ ,  $k = \overline{2, 10}$ . Calculations arrive at the following table (Table 3).

Table 3.	Values of $M(nx; \lambda_1$	), M(nx; np)	and $\Omega_1(x, n, p)$ :	for $n = 10$ , $p = 0.1$ , $x =$	$\frac{k}{n}, k =$	2,9.
----------	-----------------------------	--------------	---------------------------	----------------------------------	--------------------	------

x	nx	$\lambda_1$	$M(nx;\lambda_1)$	M(nx;np)	$\Omega_1(x,n,p)$
0.2	2	0.88888	1.37583	1.43656	0.95772
0.3	3	0.77777	1.22909	1.30969	0.93846
0.4	4	0.66666	1.14969	1.23876	0.92809
0.5	5	0.55555	1.10048	1.19382	0.92181
0.6	6	0.44444	1.0672	1.16292	0.91769
0.7	7	0.33333	1.04326	1.14042	0.91481
0.8	8	0.22222	1.02525	1.12332	0.91269
0.9	9	0.11111	1.01167	1.10991	0.91148

Taking into account that  $\Omega_1(x, n, p)$  is not much different from 1 (see Table 3) write up  $\Omega_1(x, n, p) (1 + r_1(x)) = 1 + r_2(x)$ . We will use the elementary identity b(1 + a) - 1 = a - (1 - b)(1 + a). Putting  $a = r_1(x)$ ,  $b = \Omega_1(x, n, p)$ , we obtain

$$r_2(x) \equiv \Omega_1(x,n,p) \left(1 + r_1(x)\right) - 1 = r_1(x) - \left(1 - \Omega_1(x,n,p)\right) (1 + r_1(x)), \quad (30)$$

and

$$1 - F_{n,p}(nx) = \left(1 - \prod_{np}(nx)\right) e^{-nq\Lambda_3\left(\frac{x-p}{q}\right)} (1 + r_2(x)).$$
(31)

Note that Equality (31) is another form of Theorem 2.

The following inequalities hold:  $1 - \Omega_1(x, n, p) > 0$  and  $1 + r_1(x) > 0$ . Hence, by (30),  $|r_2(x)| \ge |r_1(x)|$  if  $r_1(x) < 0$ , and  $r_2(x) \le r_1(x)$  if  $r_1(x) > 0$ . In the next theorem, the estimate of  $r_2(x)$  is got.

**Theorem 3.** If condition (15) is fulfilled, then

$$\frac{1 - F_{n,p}(nx)}{1 - \prod_{np}(nx)} = e^{-nq\Lambda_3\left(\frac{x-p}{q}\right)} (1 + r_2(x)), \tag{32}$$

where

$$r_2(x) = \theta \Big( c_1 \sqrt{nx^3} + \frac{p}{x} \Big).$$

#### Remark 4.

1. The closeness of  $xn^{1/3}$  and  $\frac{p}{x}$  to 0 ensure the closeness of  $r_2(x)$  to zero. Moreover, as it was said in Remark 2, the closeness of  $xn^{1/3}$  to 0 agrees with [9].

2. Under the condition  $x = o(n^{-1/3})$  the quantity  $nx^2$  may not tend to zero.

**Remark 5.** Let us discuss the connection between x, n and p the function  $r(x) = c_1 \sqrt{nx^3} + \frac{p}{r}$ approaching zero. Obviously, r(x) can tend to zero only if  $x \to 0$  and  $\frac{p}{x} \to 0$ .

Let the parameters n and p be fixed. Find  $\min_{p < x < 1} r(x)$ . We write r(x) for brevity as

 $r(x) = ax^{3/2} + \frac{p}{x}$ . Obviously,  $\frac{dr(x)}{dx} = \frac{3a}{2}x^{1/2} - \frac{p}{x^2}$ . Therefore, the minimum of r(x) is attained at the point  $x_0 = \left(\frac{2p}{3a}\right)^{2/5}$ , and  $\frac{dr(x)}{dx} < 0$  for  $0 < x < x_0$ , and  $\frac{dr(x)}{dx} > 0$  for  $x > x_0$ . As a result of calculations, we make sure that  $p < x_0$  if  $p < \left(\frac{2}{3c_1}\right)^{2/3} n^{-1/3}$ . This condition

can be considered fulfilled. From here,

$$\min_{p < x < 1} r(x) = r(x_0) = a^{2/5} p^{3/5} \left( (2/3)^{3/5} + (2/3)^{-2/5} \right) = (np^3)^{1/5} c_2,$$

where  $c_2 = c_1^{2/5} (5/2) (2/3)^{3/5} = 3.8619 \dots$ . Thus,  $\min_{p < x < 1} r(x) \to 0$  if and only if  $np^3 \to 0$ .

Indeed, let f(x) and g(x) be the left and right branches of the function r(x) with respect to the line  $x = x_0$ . In this case, the domain of f(x) is  $(0, x_0]$ , and g(x) is  $[x_0, \infty)$ . On the other hand, for each  $\varepsilon > 0$ , you can specify an interval  $(x_1, x_2)$ , containing  $x_0$  such that for  $x \in (x_1, x_2)$  the inequality  $r(x) - r(x_0) < \varepsilon$  will hold.

These functions are strictly monotone and therefore have inverse ones:  $f^{-1}(\cdot)$  and  $g^{-1}(\cdot)$ , respectively. Then the required interval has the form  $(f^{-1}(\varepsilon), g^{-1}(\varepsilon))$ . Note that the domain of *these inverse functions is the same:*  $[r(x_0), \infty)$ *.* 

**Example 2.** Let n = 100,  $p = 4.75 \times 10^{-4}$ . Then  $a = c_1 \sqrt{n} = 54.4893$ ,  $x_0 = \left(\frac{2p}{3a}\right)^{2/5} =$ 0.00804853,  $r(x_0) = 0.0983617$ . The graph of the function  $r(x) = ax^{3/2} + \frac{p}{x}$  is shown in Figure 1.

Take  $\varepsilon = 0.05$  for example. Finding the roots of the equation  $r(x) - r(x_0) = \varepsilon$ , we get:  $x_1 = 0.0034603$ ,  $x_2 = 0.0169604$ ,  $x_2 - x_1 = 0.0135001$ . Note that  $\varepsilon$  can be chosen arbitrarily small only if  $np^3$  is sufficiently close to 0.

The following table (Table 4) shows the behavior of the interval  $(x_1, x_2) \equiv ((f^{-1}(\varepsilon), g^{-1}(\varepsilon)))$ with decreasing  $\varepsilon$ .

ε	1/2	1/4	1/8	1/16	1/32	1/1024
<i>x</i> <sub>1</sub>	0.0007	0.001	0.002	0.003	0.004	0.007
<i>x</i> <sub>2</sub>	0.048	0.033	0.024	0.018	0.014	0.009
$x_2 - x_1$	0.048	0.032	0.021	0.015	0.01	0.001

**Table 4.** Behavior of the interval  $(x_1, x_2)$  depending on  $\varepsilon$ .



**Figure 1.** Graph of the function r(x) from Example 2. Points  $x_1 = f^{-1}(1/8) = 0.00218$  and  $x_2 = g^{-1}(1/8) = 0.024$  are indicated.

*Figure* 1 *illustrates the column* " $\varepsilon = \frac{1}{8}$ " *from Table* 4.

Note that near the point  $x_0$  both functions  $ax^{3/2}$  and  $\frac{p}{x}$  that form r(x), make approximately the same contribution to r(x). For instance, r(0.008) = 0.098..., where  $ax^{3/2}|_{x=0.008} = 0.0389...$ ,  $\frac{p}{x}|_{x=0.008} = 0.059...$ 

**Corollary 1.** Let condition (15) be fulfilled and  $c_1\sqrt{nx^3} < 1$ . Then

$$\frac{1 - F_{n,p}(nx)}{1 - \Pi_{np}(nx)} = e^{-\frac{n(x-p)^2}{2}} \left[ 1 + \theta \left( 5.74\sqrt{nx^3} + \frac{p}{x} \right) \right].$$
(33)

**Remark 6.** Note that the behavior of the series  $nq\Lambda_3(\frac{x-p}{q})$  is defined by the first summand in contrast to the Cramér series in the case of Gaussian approximation [4].

## 2. Proof of Proposition 1

We will use one result from [8]. The latter is formulated as follows.

Let  $X_1, \ldots, X_n$  be independent Bernoulli random variables. We denote  $S_n = \sum_{j=1}^n X_j$ ,

 $F_{S_n}$  the distribution of the sum  $S_n$ ,  $p_j = \mathbf{P}(X_j = 1)$ ,  $\lambda = \sum_{j=1}^n p_j$ ,  $\Pi_{\lambda}$  Poisson distribution with parameter  $\lambda$ . In the paper [8], the following estimate for the total variation distance  $d_{TV}(F_{S_n}, \Pi_{\lambda})$  between  $F_{S_n}$  and  $\Pi_{\lambda}$  is obtained,

$$d_{TV}(F_{S_n}, \Pi_{\lambda}) \le (1 - e^{-\lambda}) \frac{1}{\lambda} \sum_{j=1}^{n} p_j^2.$$
 (34)

In the particular case when

$$p_1 = p_2 = \ldots = p_n = p,$$
 (35)

we have  $\lambda = np$ . Then it follows from (34) that

$$d_{TV}(F_{S_n}, \Pi_{\lambda}) \le (1 - e^{-\lambda}) p, \tag{36}$$

whence

$$d_{TV}(F_{S_n}, \Pi_{\lambda}) \le p. \tag{37}$$

In the case (35) we will use the notation

$$d_{n,p} = d_K(F_{S_n}, \Pi_\lambda),$$

where  $d_K(F_{S_n}, \Pi_{\lambda})$  is the Kolmogorov distance between the distributions  $F_{S_n}$  and  $\Pi_{\lambda}$ . Since

$$d_{n,p} \le d_{TV}(F_{S_n}, \Pi_\lambda), \tag{38}$$

It follows from (37) that

$$d_{n,p} \le p. \tag{39}$$

In what follows, we will use the estimate (39), although there is reason to believe that on the right-hand side of (39) the coefficient 1 in front of *p* can be replaced by a smaller number, namely,  $e^{-1} - \frac{1}{4} \approx 0.236$  (see Section 6 Supplement).

**Proof of Proposition 1.** Note that  $F^{(h)}(t)$  is the Bernoulli distribution function with the parameter  $p_h = \frac{p e^h}{r(h)}$ , and  $F^{(h)}_{n,p}(t)$  the binomial distribution function with the parameters n and  $p_h$ . It is easily seen that by (16),

$$p_h = x, \quad q_h = 1 - p_h = 1 - x.$$
 (40)

Therefore, instead of  $F_{n,p}^{(h)}(t)$ , we can write  $F_{n,x}(t)$ .

1

The approximating Poisson distribution for  $F_{n,p}^{(h)}$  has the parameter nx. Taking into account this and (19), we have

$$\omega_n(nx,F) = \int_{nx}^{\infty} e^{-hy} \, dF_{n,p}^{(h)}(y) = I_1 + I_2,\tag{41}$$

where

$$I_{1} = \int_{nx}^{\infty} e^{-hy} \, d\Pi_{nx}(y), \quad I_{2} = \int_{nx}^{\infty} e^{-hy} \, d\Big(F_{n,x}(y) - \Pi_{nx}(y)\Big)$$

It is easily seen that

$$I_{1} = e^{-nx} \sum_{k \ge nx} e^{-hk} \frac{(nx)^{k}}{k!} = e^{-nx(1-e^{-h})} \sum_{k \ge nx} \frac{(nxe^{-h})^{k}}{k!} e^{-nxe^{-h}}$$

$$= e^{-nx(1-e^{-h})} \left(1 - \Pi_{\lambda(h)}(nx)\right),$$
(42)

where  $\lambda(h) = nxe^{-h}$ . It follows from (16) and (20) that

$$\lambda(h) = nx \frac{(1-x)p}{xq} = np \frac{1-x}{q} = \lambda_1.$$
(43)

Moreover, by (16),

$$nx(1-e^{-h}) = nx\left(1-\frac{(1-x)p}{xq}\right) = \frac{n(x-p)}{q}.$$

Then we get from (42) that

$$I_1 = e^{-\frac{n(x-p)}{q}} \left( 1 - \Pi_{\lambda_1}(nx) \right).$$
(44)

Further, integrating by parts, we get

$$|I_{2}| = \left| h \int_{nx}^{\infty} e^{-hy} \left( F_{n,x}(y) - \Pi_{nx}(y) \right) dy - e^{-hnx} \left( F_{n,x}(nx) - \Pi_{nx}(nx) \right) \right| \leq$$

$$\leq 2e^{-nhx} \max_{y \geq nx} \left| F_{n,x}(y) - \Pi_{nx}(y) \right|.$$

$$(45)$$

It follows from (41), (44) and (45) that

$$\omega_n(nx;F) = e^{-\frac{n(x-p)}{q}} \left(1 - \Pi_{\lambda_1}(nx)\right) + R,$$

$$|R| \le 2e^{-nhx} \max_{y \ge nx} \left|F_{n,x}(y) - \Pi_{nx}(y)\right|.$$
(46)

Using (16) and the definition of H(t, p), we get

$$e^{-nhx} \left(\frac{q}{1-x}\right)^n = \left(\frac{qx}{p(1-x)}\right)^{-nx} \left(\frac{q}{1-x}\right)^n = \\ = \left(\left(\frac{x}{p}\right)^x \left(\frac{1-x}{q}\right)^{1-x}\right)^{-n} = e^{-nH(x,p)}.$$
(47)

It follows from (18), (21) and (46) that

$$1 - F_{n,p}(nx) = \left(1 - \Pi_{\lambda_1}(nx)\right) A(x, n, p) + \left(\frac{q}{1-x}\right)^n R,$$
(48)

where R satisfies the inequality in (46).

Now applying (39) and (46)–(48), one after the other, we obtain the statement of Proposition 1.  $\Box$ 

#### 3. Proof of Theorems 1 and 2

It is assumed in Section 3 condition (15) to be fulfilled.

Lemma 1. If the condition (15) is satisfied and nx is an integer, then

$$\frac{|I_2|}{I_1} \le \frac{2\sqrt{2\pi nx^3} e^{1/(12nx)}}{M(nx;\lambda_1)} \le \frac{c_1\sqrt{nx^3}}{M(nx;\lambda_1)}.$$
(49)

**Proof.** First of all, we write (44) as

$$I_1 = M(nx;\lambda_1) \exp\left\{-\frac{n(x-p)}{q} - \lambda_1\right\} \frac{\lambda_1^{nx}}{(nx)!}.$$

In view of (43),

$$\frac{n(x-p)}{q} + \lambda_1 = \frac{nx - nxp}{q} = nx$$

Consequently,

$$I_1 = M(nx;\lambda_1) e^{-nx} \frac{\lambda_1^{nx}}{(nx)!}.$$
 (50)

$$\frac{|I_2|}{I_1} \le \frac{2x(nx)! \, e^{nx - nhx}}{M(nx;\lambda_1) \, \lambda_1^{nx}}.$$
(51)

In accordance to 
$$(16)$$
,

$$e^{-nhx} = \left(\frac{p(1-x)}{qx}\right)^{nx}.$$
(52)

Further, by the Stirling formula,

$$(nx)! = \sqrt{2\pi nx} (nx)^{nx} e^{-nx+\theta}, \tag{53}$$

where  $\frac{1}{12nx+1} < \theta < \frac{1}{12nx}$ . Finally, according to (43),

$$\lambda_1^{nx} = (nx)^{nx} \left(\frac{p(1-x)}{qx}\right)^{nx}.$$
(54)

Substituting the expressions (52)–(54) into the right-hand side of (51), we obtain (49).  $\Box$ 

**Proof of Theorem 1.** It follows from (18), (41), and (49) that

$$1 - F_{n,p}(nx) = I_1 \cdot \left(\frac{1-x}{q}\right)^{-n} (1+r_1),$$
(55)

where  $|r_1| \le c_1 \sqrt{nx^3}$ . Note that by (55), we have  $1 + r_1 > 0$ . Taking into account (21) and (44), we get

$$I_1 \cdot \left(\frac{1-x}{q}\right)^{-n} = A(x,n,p) \left(1 - \Pi_{\lambda_1}(nx)\right).$$
(56)

Theorem 1 follows from (23), (55) and (56).  $\Box$ 

Proof of Theorem 2. Using (55) and (56), we obtain

$$1 - F_{n,p}(nx) = \left(1 - \Pi_{\lambda_1}(nx)\right) A(x, n, p)(1 + r_1) = \left(1 - \Pi_{np}(nx)\right) \frac{M(nx; \lambda_1) \pi_{\lambda_1}(nx)}{M(nx; np) \pi_{np}(nx)} A(x, n, p)(1 + r_1).$$
(57)

We have

$$\frac{\pi_{\lambda_1}(nx)}{\pi_{np}(nx)} = (1-u)^{nx} e^{npu} = e^{npu+nx\ln(1-u)},$$

where  $u = \frac{x-p}{q}$ . Hence, taking into account (22), we get

$$\frac{\pi_{\lambda_1}(nx)}{\pi_{np}(nx)} A(x,n,p) = e^{-nq\left(\frac{1-x}{q}\ln(1-u)+u\right)} = e^{-nq\left((1-u)\ln(1-u)+u\right)}.$$

Expanding  $\ln(1 - u)$  to the power of *u*, we have

$$u + (1 - u)\ln(1 - u) = u - (1 - u)\sum_{k=1}^{\infty} \frac{u^k}{k} = \Lambda_3(u).$$
(58)

Thus,

$$\frac{\pi_{\lambda_1}(nx)}{\pi_{np}(nx)}A(x,n,p) = e^{-nq\Lambda_3(u)}.$$
(59)

The equality (28) follows from (57) and (59).  $\Box$ 

## 4. Proof of Theorem 3 and Corollary 1

**Lemma 2.** Let nx be an integer. Then

$$M(nx;\lambda_1) \leq \frac{1}{1-p/x}.$$

**Proof.** By formula (13),

$$M(nx;\lambda_1) \leq 1 + \sum_{m=1}^{\infty} \left(\frac{\lambda_1}{nx}\right)^m.$$

It follows from (15) that

$$\lambda_1 \leq np.$$

These inequalities imply Lemma 2.  $\Box$ 

**Lemma 3.** The following inequality holds,

$$\frac{1 - \Pi_{\lambda_1}(nx)}{1 - \Pi_{np}(nx)} > \frac{\pi_{\lambda_1}(nx)}{\pi_{np}(nx)} (1 - x^{-1}p).$$

Proof. Obviously,

$$\frac{1 - \Pi_{\lambda_1}(nx)}{1 - \Pi_{np}(nx)} > \frac{\pi_{\lambda_1}(nx)}{1 - \Pi_{np}(nx)}.$$
(60)

Since for every  $k \ge 0$ ,  $\lambda > 0$ ,

$$rac{\pi_\lambda(k+1)}{\pi_\lambda(k)} = rac{\lambda}{k+1},$$

then for  $k \ge nx$ 

$$\frac{\pi_{np}(k+1)}{\pi_{np}(k)} = \frac{np}{k+1} < \frac{p}{x}$$

and, consequently,

$$1 - \Pi_{np}(nx) = \pi_{np}(nx) \sum_{k=0}^{\infty} \frac{\pi_{np}(nx+k)}{\pi_{np}(nx)} < \pi_{np}(nx) \sum_{k=0}^{\infty} \frac{p^k}{x^k} = \pi_{np}(nx) \frac{1}{1 - x^{-1}p}.$$
 (61)

The statement of the lemma follows from (60) and (61).  $\Box$ 

Lemma 4. The following inequality holds,

$$\frac{1-\Pi_{\lambda_1}(nx)}{1-\Pi_{np}(nx)} \leq \frac{\pi_{\lambda_1}(nx)}{\pi_{np}(nx)}.$$

**Proof.** We have

$$\frac{1-\Pi_{\lambda_1}(nx)}{1-\Pi_{np}(nx)} \leq \max_{k \geq nx} \frac{\pi_{\lambda_1}(k)}{\pi_{np}(k)} = \frac{\pi_{\lambda_1}(nx)}{\pi_{np}(nx)}$$

The last equality is true because of

$$\frac{\pi_{\lambda_1}(k)}{\pi_{np}(k)} = \frac{\lambda_1^k}{(np)^k} e^{-\lambda_1 + np}$$
(62)

and, by (43),

$$\frac{\lambda_1}{np} = \frac{1-x}{q} < 1. \tag{63}$$

Lemma 5. The following equality holds,

$$\frac{1 - \Pi_{\lambda_1}(nx)}{1 - \Pi_{np}(nx)} = \frac{\pi_{\lambda_1}(nx)}{\pi_{np}(nx)} (1 - r_3),$$

13 of 21

where

$$r_3 = \theta_1 \frac{p}{x}, \quad 0 < \theta_1 < 1$$

Proof. By Lemmas 3 and 4,

$$\frac{\pi_{\lambda_1}(nx)}{\pi_{np}(nx)} (1 - x^{-1}p) < \frac{1 - \Pi_{\lambda_1}(nx)}{1 - \Pi_{np}(nx)} \le \frac{\pi_{\lambda_1}(nx)}{\pi_{np}(nx)}$$

This implies the statement of Lemma 5.  $\Box$ 

Proof of Theorem 3. Using (55) and (56) and Lemma 5, we have

$$1 - F_{n,p}(nx) = \left(1 - \Pi_{np}(nx)\right) e^{-n\left(u + \ln(1-u)\right)} \frac{\pi_{\lambda_1}(nx)}{\pi_{np}(nx)} (1 + r_1)(1 - r_3).$$
(64)

By (62) and (63),

$$\frac{\pi_{\lambda_1}(nx)}{\pi_{np}(nx)} = (1-u)^{nx} e^{npu} = e^{npu+nx\ln(1-u)}.$$
(65)

It follows from (64), (65) that

$$1 - F_{n,p}(nx) = \left(1 - \Pi_{np}(nx)\right) e^{-n\left(uq + (1-x)\ln(1-u)\right)} (1+r_1)(1-r_3).$$
(66)

Taking into account that 1 - x = q(1 - u) and expanding  $\ln(1 - u)$  to the power of u, we have

$$uq + (1-x)\ln(1-u) = q \left[ u + (1-u)\ln(1-u) \right]$$
  
=  $q \left[ u - (1-u) \sum_{k=1}^{\infty} \frac{u^k}{k} \right] = q \Lambda_3(u).$  (67)

The statement of the theorem follows from (66), (67) and the inequality

$$|r_1 - r_3 - r_1 r_3| = |r_1(1 - r_3) - r_3| \le |r_1|(1 - r_3) + r_3 < |r_1| + r_3$$

which holds because  $0 < r_3 < 1$ .  $\Box$ 

**Proof of Corollary 1.** As before, we use the notation  $u = \frac{x-p}{q}$ . Write up

$$nq\Lambda_3(u) = nq\Big(\frac{u^2}{2} + \sum_{k=3}^{\infty} \frac{u^k}{k(k-1)}\Big) = A + B,$$
(68)

where  $A = n \frac{(x-p)^2}{2q}$ ,  $B = nq \sum_{k=3}^{\infty} \frac{u^k}{k(k-1)}$ . Firstly,

$$A = \frac{n(x-p)^2}{2} + \frac{n(x-p)^2}{2} \left(\frac{1}{q} - 1\right) = \frac{n(x-p)^2}{2} + \theta_1 \frac{nx^3}{2(1-x)},$$
(69)

where  $0 < \theta_1 < 1$ . Now let us estimate *B*. We have

$$B = nq \left(\frac{x-p}{q}\right)^3 \sum_{k=3}^{\infty} \frac{u^{k-3}}{k(k-1)}.$$
(70)

Using (58) we can write up

$$\sum_{k=3}^{\infty} \frac{u^{k-3}}{k(k-1)} = \frac{1}{u^3} \left( \Lambda_3(u) - \frac{u^2}{2} \right) = \frac{1}{u^3} \left( (1-u) \ln(1-u) + u - \frac{u^2}{2} \right) =: g(u).$$
(71)

The condition  $c_1 \sqrt{nx^3} < 1$  implies that  $x < n^{-1/3} c_1^{-2/3} \le c_1^{-2/3}$ . Hence,

$$u < \frac{x}{1-x} < \frac{x}{1-x} \Big|_{x=c_1^{-2/3}} < 0.477.$$
(72)

Then

$$\sum_{k=3}^{\infty} \frac{u^{k-3}}{k(k-1)} < g(0.477) < 0.224.$$
(73)

By (70) and (73),

$$B < nq \left(\frac{x-p}{q}\right)^3 0.224 < \frac{0.224 \, nx^3}{(1-x)^2}.$$
$$B = \frac{\theta_2 \, 0.224 \, nx^3}{(1-x)^2},\tag{74}$$

where  $0 < \theta_2 < 1$ .

Thus,

Collecting (68), (69) and (74), we arrive at

$$nq\Lambda_{3}\left(\frac{x-p}{q}\right) = \frac{n(x-p)^{2}}{2} + \frac{\theta_{1}nx^{3}}{2(1-x)} + \frac{\theta_{2} \ 0.224 \ nx^{3}}{(1-x)^{2}} = \frac{n(x-p)^{2}}{2} + \frac{\theta_{3} \ 0.724 \ nx^{3}}{(1-x)^{2}},$$
(75)

where  $0 < \theta_3 < 1$ . Hence,

$$e^{-nq\Lambda_3\left(\frac{x-p}{q}\right)} = e^{-n\frac{(x-p)^2}{2} - \theta_3 0.724 \frac{nx^3}{(1-x)^2}} = e^{-n\frac{(x-p)^2}{2}}(1-r_4),$$
(76)

where  $r_4 = \theta_4 0.724 \frac{nx^3}{(1-x)^2}$ ,  $0 < \theta_4 < 1$ . Using the condition  $c_1 \sqrt{nx^3} < 1$  again, we obtain

$$\frac{\sqrt{nx^3}}{(1-x)^2} < \frac{1}{c_1(1-c_1^{-2/3})^2} < 0.4004.$$

Since  $0.724 \cdot 0.4004 < 0.29 \equiv c_3$ , we have

$$r_4 = \theta_5 \cdot c_3 \sqrt{nx^3},\tag{77}$$

where  $0 < \theta_5 < 1$ .

From (76) we obtain

$$e^{-nq\Lambda_3\left(\frac{x-p}{q}\right)}(1+r) = e^{-n\frac{(x-p)^2}{2}}(1-r_4)(1+r).$$
(78)

Write up

$$(1 - r_4)(1 + r) = 1 + r_5,$$

where

$$r_5 = r - r_4 - rr_4.$$

Denote for brevity  $a_1 = \sqrt{nx^3}$ ,  $a_2 = \frac{p}{x}$ . Then  $r = \theta(c_1a_1 + a_2)$ ,  $|\theta| < 1$ ,  $r_4 = \theta_5c_3a_1$ ,  $0 < \theta_5 < 1$ .

Let us consider two cases:  $r \ge 0$  and r < 0. Consider the case  $r \ge 0$ . Then

$$|r_5| \leq \max\{r, r_4 + rr_4\}.$$

Note that r < 2. Since

$$r \leq c_1 a_1 + a_2$$
,  $r_4 + rr_4 < 3r_4 \leq 3c_3 a_1$ ,  $c_1 a_1 + a_2 > 3c_3 a_1$ ,

then

If r < 0, then

$$|r_5| \le c_1 a_1 + a_2.$$
  
 $|r_5| \le \max\{|r| + r_4, |r| \, r_4\}.$ 

Since

$$|r| + r_4 \le a_1(c_1 + c_3) + a_2$$
,  $|r| r_4 \le 2c_3a_1$ ,  $a_1(c_1 + c_3) + a_2 > 2c_3a_1$ ,

then

$$|r_5| \le (c_1 + c_3)a_1 + a_2.$$

Taking into account that  $c_1 + c_3 < 5.74$  we can conclude that Corollary 1 is proved.  $\Box$ 

## 5. Numerical Experiments

Further we will use the following Tables 5 and 6.

**Table 5.** Values of  $d_{n,p} = d_K(F_{S_n}, \Pi_{np})$  for  $p = 1/k, k = \overline{2, 10}, n = \overline{1, 20}$ .

n	$d_{n,1/2}$	$d_{n,1/3}$	$d_{n,1/4}$	$d_{n,1/5}$	$d_{n,1/6}$	$d_{n,1/7}$	$d_{n,1/8}$	$d_{n,1/9}$	$d_{n,1/10}$
1	0.10653	0.0498	0.0288	0.01873	0.0131	0.00973	0.00749	0.00595	0.0048
2	0.11787	0.06897	0.044	0.03022	0.022	0.01678	0.01317	0.01061	0.0087
3	0.09813	0.07158	0.050	0.03681	0.0278	0.0217	0.01736	0.01419	0.011
4	0.0935	0.06606	0.05147	0.03972	0.031	0.02494	0.02034	0.01688	0.0142
5	0.09979	0.05718	0.049	0.04019	0.0327	0.02687	0.02235	0.01882	0.016
6	0.08977	0.05483	0.045	0.03905	0.03298	0.02780	0.02357	0.02014	0.0173
7	0.09428	0.05986	0.040	0.03688	0.032	0.02796	0.02416	0.02096	0.0182
8	0.09357	0.05968	0.0389	0.03412	0.031	0.02754	0.02427	0.02136	0.0188
9	0.08838	0.05608	0.042	0.03108	0.029	0.02671	0.02399	0.02144	0.0191
10	0.09315	0.05363	0.043	0.03019	0.027	0.02559	0.02342	0.02124	0.0192
11	0.08841	0.057	0.0426	0.03244	0.025	0.02426	0.0226	0.02084	0.0191
12	0.0912	0.05698	0.04076	0.03356	0.02467	0.02282	0.0217	0.02028	0.0187
13	0.09024	0.05447	0.03808	0.03373	0.026	0.02131	0.02067	0.01959	0.0183
14	0.08873	0.05376	0.0397	0.03316	0.02727	0.02085	0.0195	0.01882	0.0178
15	0.09055	0.05578	0.04098	0.03202	0.02767	0.02208	0.0184	0.01798	0.0172
16	0.08617	0.05536	0.04099	0.03046	0.02760	0.02290	0.01807	0.01711	0.0165
17	0.09002	0.05309	0.04008	0.03011	0.02714	0.02335	0.0190	0.01621	0.0152
18	0.08783	0.05398	0.03827	0.03140	0.02636	0.02348	0.0197	0.01594	0.0159
19	0.08902	0.05497	0.03879	0.03200	0.025	0.02333	0.0201	0.01670	0.0144
20	0.08863	0.05411	0.0398	0.03201	0.024	0.02296	0.02034	0.01727	0.0137
$\infty$	0.08303	0.04888	0.03474	0.02696	0.02204	0.01864	0.0161	0.01424	0.0127

n	$2d_{n,1/2}$	$3d_{n,1/3}$	$4d_{n,1/4}$	$5d_{n,1/5}$	$6d_{n,1/6}$	$7d_{n,1/7}$	$8d_{n,1/8}$	$9d_{n,1/9}$	$10d_{n,1/10}$
1	0.21306	0.14959	0.11520	0.09365	0.07889	0.06814	0.05997	0.05355	0.0483
2	0.23575	0.20691	0.17612	0.1516	0.13252	0.11748	0.10540	0.09552	0.087
3	0.19626	0.21474	0.20196	0.18405	0.16696	0.1519	0.13893	0.12779	0.11
4	0.18701	0.19819	0.20589	0.19864	0.18698	0.17460	0.16279	0.15196	0.142
5	0.19959	0.17156	0.1968	0.20099	0.19632	0.18814	0.17882	0.16942	0.16
6	0.17954	0.16452	0.18060	0.19525	0.19788	0.19462	0.18857	0.18132	0.173
7	0.18856	0.17959	0.16116	0.18440	0.19392	0.19573	0.19332	0.18867	0.182
8	0.18714	0.17904	0.15570	0.17062	0.18617	0.19284	0.19416	0.19231	0.88
9	0.17677	0.16824	0.16883	0.15540	0.17594	0.18702	0.19195	0.1929	0.191
10	0.18630	0.16090	0.17308	0.15098	0.1642	0.17914	0.18743	0.19122	0.1920
11	0.17683	0.17102	0.17052	0.16223	0.15175	0.16988	0.18118	0.18760	0.1906
12	0.18241	0.17094	0.16306	0.16781	0.14803	0.15977	0.17370	0.18253	0.187
13	0.18049	0.16343	0.15234	0.16869	0.15779	0.14922	0.16537	0.1763	0.183
14	0.17746	0.16130	0.15887	0.16583	0.16362	0.14603	0.15651	0.16940	0.178
15	0.18111	0.16735	0.16392	0.16011	0.16605	0.15461	0.14736	0.16188	0.172
16	0.17235	0.16609	0.16396	0.15231	0.16562	0.16031	0.14456	0.15401	0.165
17	0.18005	0.15928	0.16003	0.15058	0.16284	0.16345	0.15221	0.14595	0.1529
18	0.17567	0.16195	0.15309	0.15702	0.15820	0.16436	0.15766	0.14346	0.1520
19	0.17804	0.16492	0.15517	0.16003	0.15213	0.16337	0.16111	0.15034	0.144
20	0.17726	0.16234	0.15947	0.16009	0.1450	0.16077	0.16279	0.1555	0.137
$\infty$	0.166	0.14666	0.13898	0.13484	0.13225	0.13048	0.12919	0.12821	0.1274

**Table 6.** Values of product  $k d_{n,1/k}$  for  $k = \overline{2, 10}, n = \overline{1, 20}$ .

Let us represent  $r_1(x)$  from the equality (26) in Theorem 1 in the form

$$r_1(x) = g(x, n, p) \sqrt{nx^3}.$$

In view of (27) the following estimate holds,

$$|g(x, n, p)| \le c_1 = 5.4489 \dots$$
 (79)

The question arises: how accurate is this estimate? In other words, we need to find

$$c_0 = \inf\{c : |g(x, n, p)| \le c, n \ge 1, 0$$

and check how much  $c_1$  differs from  $c_0$ . Let us rewrite (26) as

$$\frac{1 - F_{n,p}(nx)}{1 - \Pi_{\lambda_1}(nx)} \left( A(x, n, p) \right)^{-1} = 1 + g(x, n, p) \sqrt{nx^3},\tag{80}$$

where the parameter  $\lambda_1$  is defined by formula (43), and A(x, n, p) by formula (21). It follows from (80) that

$$g(x,n,p) = \frac{1}{\sqrt{nx^3}} \left[ \frac{1 - F_{n,p}(nx)}{1 - \prod_{\lambda_1}(nx)} \left( A(x,n,p) \right)^{-1} - 1 \right].$$

Let us carry out numerical experiments to get closer to understanding the answer to the question posed.

Note that if x = j/n, where j = 0, 1, ..., (n-1), then  $\lambda_1 = \frac{p}{a}(n-j)$ .

#### 5.1. Illustration to Theorem 1

Let n = 10, p = 0.1. Table 7 contains values of the function g(x, n, p), found for  $x = \frac{1}{n}$ ,  $j = \overline{2,9}$ , as well as intermediate results.

On the last row, the maximum value is shown in bold.

x	0.2	0.3	0.4	0.5
nx	2	3	4	5
$1 - F_{n,p}(nx)$	0.263	0.07	0.012	0.001
$1 - \Pi_{\lambda_1}(nx)$	0.223	0.044	0.004	0.0002
$\frac{1-F_{n,p}(nx)}{1-\Pi_{\lambda_1}(nx)}$	1.181	1.585	2.633	5.871
$\frac{1}{A(x,n,p)}$	0.935	0.747	0.486	0.238
$\frac{1 - F_{n,p}(nx)}{\left(1 - \Pi_{\lambda_1}(nx)\right) A(x,n,p)}$	1.104	1.185	1.28	1.400
$\frac{1 - F_{n,p}^{1}(nx)}{(1 - \Pi_{\lambda_{1}}(nx)) A(x,n,p)} - 1$	0.104	0.185	0.28	0.4002
$\sqrt{nx^3}$	0.282	0.519	0.8	1.118
g(x,n,p)	0.3704	0.356	0.3503	0.358
<i>x</i>	0.6	0.7	0.8	0.9
nx	6	7	8	9
$1 - F_{n,p}(nx)$	0.00014	$9.12 imes10^{-6}$	$3.7 \times 10^{-7}$	$9.1  imes 10^{-9}$
$1 - \Pi_{\lambda_1}(nx)$	$7.3 imes10^{-6}$	$6.78 imes10^{-8}$	$1.21 \times 10^{-10}$	$6.4 \times 10^{-15}$
$\frac{1-F_{n,p}(nx)}{1-\Pi_{\lambda_{n}}(nx)}$	20.05	134.5	3085.37	$1.4 imes10^6$
$\frac{1}{A(x,n,p)}$	0.0777	0.0133	0.0007	$2  imes 10^{-6}$
$\frac{1 - F_{n,p}(nx)}{\left(1 - \Pi_{\lambda_1}(nx)\right) A(x,n,p)}$	1.56	1.789	2.16	2.9
$\frac{1-F_{n,p}(nx)}{\left(1-\Pi_{\lambda_1}(nx)\right)A(x,n,p)}-1$	0.56	0.789	1.16	1.9
$\sqrt{nx^3}$	1.469	1.852	2.262	2.7
g(x, n, p)	0.38	0.426	0.51	0.718

**Table 7.** Values of the function g(x, n, p) for n = 10, p = 0.1 and  $x = \frac{j}{n}$ ,  $j = \overline{2, 9}$ .

The third row from bottom in Table 7 shows a real value of the remainder in (26) for various *x*. These values are acceptable for  $x \le 0.6$ , taking into account small value of the parameter *n*. The last line contains the values of the coefficient at  $\sqrt{nx^3}$  in  $r_1 = r_1(x)$ , for which the latter coincides with the real remainder for the given value of *x*. For example, for x = 0.2, the coefficient for  $\sqrt{nx^3}$  must be equal to 0.26764 (column "x = 0.2", last line), i.e.,  $r_1(0.2) = 0.26764\sqrt{0.08} = 0.0757$ . If for  $d_K(F_{10,1/5}, \Pi_2)$  to take the bound 0.150981*x*, where x = 0.2, which is found with the help of Table 6 (column " $5d_{n,1/5}$ ", line "n = 10"), we obtain  $0.822685 \equiv 0.150981c_1$  as the coefficient at  $\sqrt{nx^3}$  in  $r_1(0.2)$ . The latter is more than two times the number 0.370456 from Table 7. Accordingly, we obtain for  $r_1(0.2)$  is 0.104781 (see Table 7 or Table 8).

We observe that the product  $(1 - \Pi_{\lambda_1}(nx)) A(x, n, p)$  approximates  $1 - F_{n,p}$  by excess, which is compensated by multiplying by  $1 + r_1(0.2)$ .

5.2. Illustration to Theorems 2 and 3

Denote

$$V(x,n,p) = \frac{1 - F_{n,p}(nx)}{1 - \prod_{np}(nx)} e^{nq\Lambda_3\left(\frac{x-p}{q}\right)}, \quad \Omega_2(x,n,p) = \frac{M(nx;np)}{M(nx;\lambda_1)} \equiv \frac{1}{\Omega_1(x,n,p)}$$

Put n = 10, p = 0.1. Using Table 8 we can compare the remainders  $r_1(x)$  and  $r_2(x)$ .

**Table 8.** Values of the remainders  $r_2(x) = V(x, n, p) - 1$  (see (31)) and  $r_1(x) = \Omega_2(x, n, p) V(x, n, p) - 1$  (Theorems 1 and 2) for n = 10, p = 0.1 and  $x = \frac{j}{10}$ ,  $j = \overline{2,9}$ .

x	0.2	0.3	0.4	0.5
nx	2	3	4	5
$r_2(x) = V(x, n, p) - 1$	0.05807	0.11207	0.18822	0.29078
$r_1(x) = \Omega_2(x, n, p)V(x, n, p) - 1$	0.10478	0.185	0.28029	0.4002

Let us do more detailed calculations. We calculate for  $x = \frac{1}{n}$ ,  $j = \overline{2,9}$ :

$$V(x, n, p) - 1, \qquad \frac{V(x, n, p) - 1}{\sqrt{nx^3}}, \\ \Omega_2(x, n, p)V(x, n, p) - 1, \qquad \frac{\Omega_2(x, n, p)V(x, n, p) - 1}{\sqrt{nx^3}}.$$
(81)

Table 9 is similar to Table 7. The difference is that in Table 9 the tail  $1 - F_{n,p}(nx)$  is compared with two products:  $\frac{M(nx;\lambda_1)}{M(nx;np)} \left(1 - \prod_{np}(nx)\right) e^{-nq\Lambda_3\left(\frac{x-p}{q}\right)}$  and  $\left(1 - \prod_{np}(nx)\right) e^{-nq\Lambda_3\left(\frac{x-p}{q}\right)}$ . It turns out that the second of them more accurately approximates  $1 - F_{n,p}(nx)$  (compare the sixth and second rows from the bottom in Table 9; see also Table 8).

To compare the remainders  $r_2(x)$  (see (31)) and  $r_1(x)$  (Theorems 1 and 2), we can select in Table 9 the rows "V(x, n, p) - 1" and " $\Omega_2(x, n, p) V(x, n, p) - 1$ ".

**Table 9.** Values of the functions from (81) for n = 10, p = 0.1.

x	0.2	0.3	0.4	0.5
пх	2	3	4	5
$1 - F_{n,p}(nx)$	0.2639	0.07019	0.01279	0.0016
$1 - \Pi_{np}(nx)$	0.2642	0.0803	0.01898	0.0036
$\frac{1-F_{n,p}(nx)}{1-\Pi_{np}(nx)}$	0.9987	0.87409	0.6738	0.4467
$\frac{x-p}{q}$	0.1111	0.2222	0.3333	0.4444
$\Lambda_3(\frac{x-p}{q})$	0.0064	0.0267	0.063	0.1178
$nq\Lambda_3(\frac{x-p}{q})$	0.0577	0.24079	0.5672	1.061
$e^{nq\Lambda_3\left(rac{x-p}{q} ight)}$	1.0594	1.2722	1.7633	2.8894
V(x,n,p)	1.058	1.112	1.188	1.29
V(x,n,p)-1	0.058	0.112	0.188	0.29
$\frac{V(x,n,p)-1}{\sqrt{nx^3}}$	0.2053	0.2156	0.2352	0.26
$\Omega_2(x,n,p)$	1.0441	1.0655	1.0774	1.0848
$\Omega_2(x,n,p)V(x,n,p)$	1.1047	1.185	1.2802	1.4002
$\Omega_2(x,n,p)V(x,n,p)-1$	0.1047	0.185	0.2802	0.4002
$\frac{\Omega_2(x,n,p)V(x,n,p)-1}{\sqrt{nx^3}}$	0.37	0.356	0.350	0.358

x	0.6	0.7	0.8	0.9
пх	6	7	8	9
$1 - F_{n,p}(nx)$	0.0001469	$9.12 imes10^{-6}$	$3.7  imes 10^{-7}$	$9.1 imes10^{-9}$
$1-\Pi_{np}(nx)$	0.000594	$8 imes 10^{-5}$	$10^{-5}$	$1.12  imes 10^{-6}$
$\frac{1-F_{n,p}(nx)}{1-\Pi_{np}(nx)}$	0.24723	0.10958	0.03645	0.00808
$\frac{x-p}{a}$	0.5555	0.6666	0.7777	0.8888
$\Lambda_3(\frac{x-p}{q})$	0.1951	0.3004	0.4435	0.6447
$nq\Lambda_3(rac{\dot{x}-p}{q})$	1.7562	2.7041	3.9918	5.8027
$e^{nq\Lambda_3\left(rac{x-p}{q} ight)}$	5.7908	14.9418	54.1547	331.218
V(x, n, p)	1.43169	1.63733	1.97403	2.6787
V(x, n, p) - 1	0.43169	0.63733	0.97403	1.6787
$\frac{V(x,n,p-1)}{\sqrt{nx^3}}$	0.2937	0.3441	0.4304	0.6217
$\Omega_2(x, n, p)$	1.0896	1.0931	1.0956	1.0971
$\Omega_2(x,n,p)V(x,n,p)$	1.5601	1.7898	2.1628	2.9388
$\Omega_2(x,n,p)V(x,n,p)-1$	0.5601	0.7898	1.1628	1.9388
$\frac{\Omega_2(x,n,p)V(x,n,p)-1}{\sqrt{nx^3}}$	0.381	0.426	0.513	0.718

Table 9. Cont.

\_

## 6. Supplement

In this section, we offer the reader some conjectures regarding the behavior of  $d_K(F_{n,p}, \Pi_{\lambda})$ .

Due to the cumbersomeness of the table, we did not place columns corresponding to  $11 \le k \le 20$ .

Nevertheless, we made sure that for each  $2 \le k \le 20$  the equality

$$\max_{1 \le n \le 20} d_{n,1/k} = d_{k,1/k}$$

holds. Our conjecture is as follows: for every  $2 \le k \le n$ ,

$$\max_{n \ge 1} d_{n,1/k} = d_{k,1/k}.$$
(82)

Remark that the sequence  $d_{k,1/k}$  decreases monotonically for  $2 \le k \le 20$ . This property is also true for sequences  $d_{k+j,1/k}$  for every fixed  $j \ge 1$  and  $d_{jk,1/k}$  for fixed  $j \ge 2$ . According to CLT,  $F_{n,p}(np + x\sqrt{n})$  converges in a uniform metric with the normal law  $\Phi(x/\sqrt{pq})$ . On the other hand,  $\prod_{np}(np + x\sqrt{n})$  approaches  $\Phi(x/\sqrt{p})$ . It means that

$$\lim_{n \to \infty} d_{n,p} = \max_{x > 0} \left( \Phi(x/\sqrt{q}) - \Phi(x) \right).$$
(83)

Hence,

$$\lim_{p\to 0}\lim_{n\to\infty}d_{n,p}=0.$$

Using formula (83), we get the elements of the last row of Table 5 and, hence, the elements of the last row of Table 6.

The next conjecture concerns existence and the value of the limit of  $d_{k,1/k}$ , when  $k \to \infty$ . Calculations for  $k = \overline{2,20}$  suggest that

$$d_{k,1/k} \equiv \max_{0 \le j \le k} |F_{k,1/k}(j+) - \Pi_1(j+)| = |F_{k,1/k}(0+) - \Pi_1(0+)| \equiv e^{-1} - (1 - 1/k)^k.$$
(84)

Indeed, according to Table 5,

$$d_{2,1/2} = 0.11879 \approx e^{-1} - \frac{1}{4}, \quad d_{3,1/3} = 0.07183 \approx e^{-1} - \left(\frac{2}{3}\right)^3, \ldots$$

In connection with this we remark that the behaviour of differences  $C_n^k p^k (1-p)^{n-k} - \pi_{np}(k)$  under condition  $np \le 2 - \sqrt{2}$  is investigated in [10].

Note that (84) is equivalent to the assumption

$$d_K(F_{n,1/n}, \Pi_1) = |F_{n,1/n}(0+) - \Pi_1(0+)|,$$

i.e.,  $\max_{0 \le k \le n} |F_{n,1/n}(k+) - \Pi_1(k+)$  is realized at k = 0. This fact is fairly easy to prove in the case  $\lambda \le 2 - \sqrt{2}$ , using the results of the paper [10], in which this case is considered. In the case of  $\lambda = 1$ , it is more difficult to find a proof, but it certainly exists.

After that we can assert the formula (84) is valid for all k and, moreover, there exists lim  $kd_{k,1/k}$ . Indeed, according to [10],

$$k\left(e^{-1} - \left(1 - \frac{1}{k}\right)^{k}\right) = k\left(e^{-1} - e^{k\ln(1-1/k)}\right) = e^{-1}k\left(1 - e^{-\frac{1}{2k} + O(k^{-2})}\right) \underset{k \to \infty}{\to} \frac{1}{2e'}$$

whence

$$\lim_{k \to \infty} k d_{k,1/k} = \frac{1}{2k}$$

Accordingly to Table 6 the constant  $c_0$  in the inequality

$$\sup_{n,p} \frac{d_{n,p}}{p} \le c_0 \tag{85}$$

cannot be less  $2(e^{-1} - \frac{1}{4}) \approx 0.235759$ .

If we impose the constraint  $p \leq \frac{1}{4}$ , then the lower bound for  $c_0$  is not less than  $4(e^{-1} - (\frac{3}{4})^4) \approx 0.205892$  (see Table 6). As for the upper bound for  $c_0$ , it is equal to  $2(e^{-1} - \frac{1}{4})$  if in (85) supremum with respect to p is taken over all p such that p = 1/k,  $k = \overline{2, \infty}$ .

If we adhere to the principle of incomplete induction, then the available information is sufficient to assert that  $c_0 = 2(e^{-1} - \frac{1}{4})$ .

Note that in the case  $p > \frac{1}{2}$ , it is sufficient to swap the roles of p and 1 - p. Table 6 demonstrates the following remarkable property:

$$\max_{1 \le n \le 20} \max_{2 \le k \le 9} k \, d_{n,1/k} = \max_{2 \le k \le 9} k \, d_{k,1/k} = 2 \, d_{2,1/2} = 2 \left( e^{-1} - \frac{1}{4} \right) = 0.235759.$$

Therefore, it is highly plausible that

$$\max_{1 \le n \le 20} \sup_{0$$

Moreover, the following equality is highly plausible,

$$\sup_{n \ge 1} \sup_{0 (86)$$

The equality (86) is another our conjecture. If this assumption is true, then instead of (39) we have a more precise estimate

$$d_{n,p} \le 2\left(e^{-1} - \frac{1}{4}\right)p < 0.23576p.$$
(87)

If the hypothetical estimate (87) is correct, main statements of the present work can be sharpen.

Since  $2 \cdot 0.23576 < 0.472$ , in the right-hand side of inequality (25) in Proposition 1 the product  $2xe^{-nH(x,p)}$  can be replaced by  $0.472xe^{-nH(x,p)}$ .

Taking into account the inequality  $0.23576 c_1 < 1.285$ , in all places the constant  $c_1 = 5.4489...$  can be replaced by  $\tilde{c}_1 = 1.285$ . In particular, in the formulations of Theorems 1–3,  $c_1$  can be replaced by  $\tilde{c}_1$ .

Taking into account that  $d_{k,1/k} = e^{-1} - (1 - 1/k)^k$ , and using Table 6, we arrive at the conclusion that in the case under  $k \ge 2$  and  $p = \frac{1}{k}$  (see the row "n = 2", the column " $2d_{n,1/2}$ "),

$$d_{K}(F_{n,1/k}, \Pi_{n/k}) \le k \left( e^{-1} - (1 - 1/k)^{k} \right) \Big|_{k=2} p < 0.236 \, p.$$
(88)

If *k* is growing, the coefficient at *p* in (88) decreases, but cannot be less than  $\frac{1}{2e} = 0.18...$ 

## 7. Conlusions

The main results of this paper are Theorems 1–3. In them, estimates are obtained for the errors arising when the tails of binomial distributions are replaced by the tails of the corresponding Poisson distributions. The constant  $c_1 \approx 5.5$  included in the error estimates is large enough for the estimates to be of practical use. However, further improvement of the estimates will allow apply our results, in particular, when constructing confidence intervals for the parameter *p*. The basis for the hope of obtaining such an improvement is the inequality (88).

**Author Contributions:** Investigation, S.N. and V.C.; Writing—original draft, S.N. and V.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: We are grateful to the reviewers for a number of useful comments.

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

#### References

- 1. Hoeffding, W. Probability inequalities for sums of bounded random variables. J. Am. Stat. Assoc. 1963, 58, 13–30. [CrossRef]
- 2. Sanov, I. On probability of large deviations of random variables. Mat. Sb. 1957, 42, 11-44.
- 3. Esscher, F. On the probability function in the collective theory of risk. Skand. Aktuarietidskr. 1932, 15, 175–195.
- 4. Cramér, H. On a new limit theorem of probability theory. *Russ. Math. Surv.* **1944**, *10*, 166–178.
- Nagaev, S.V.; Chebotarev, V.I. On large deviations for sums of i.i.d. Bernoulli random variables. J. Math. Sci. 2018, 234, 816–828. [CrossRef]
- 6. Petrov, V.V. A generalization of Cramer's limit theorem. Uspekhi Mat. Nauk 1954, 9, 195–202.
- 7. Talagrand, M. The missing factor in Hoeffding inequalities. Ann. Inst. Henri Poincaré 1995, 31, 689–702.
- 8. Barbour, A.D.; Hall, P. On the rate of Poisson convergence. Math. Proc. Camb. Philos. Soc. 1984, 95, 473–480. [CrossRef]
- 9. Prokhorov, Y.V. Asymptotic behaviour of binomial distribution. Russ. Math. Surv. 1953, 55, 135–142.
- 10. Kennedy, J.E.; Quine, M.P. The total variation distance between the binomial and Poisson distributions. *Ann. Probab.* **1989**, *17*, 396–400. [CrossRef]