STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING MATHEMATISCHE STATISTIEK

Report S 307

On Smirnov's problem of large deviations in the normal approximation of the binomial distribution

by

W. Molenaar

September 1962

1. Introduction

Consider a sequence of n independent Bernoulli trials with fixed probability p for success; we assume $0 < p < 1$.
Let $\underline{S}_n$ [1] be the number of successes in n trials, and let

$$\underline{S}_n^* = \frac{\underline{S}_n - n\,p}{\sqrt{n\,p\,q}} \qquad (q=1-p)$$

be the normalized number of successes. We denote by

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}u^2} du$$

the normal distribution function and by $\varphi(u)$ the corresponding frequency function.

Let x be a real number depending on n. If it is given that

$$(1.1) \qquad n \to \infty\,,\ x \to \infty\,,\ x^3 n^{-\frac{1}{2}} \to 0,$$

then a well-known theorem asserts that

$$(1.2) \qquad \frac{P\left\{\underline{S}_n^* > x\right\}}{1-\Phi(x)} \sim 1.$$

SMIRNOV (1934) has formulated the following theorem, which can also be found in FELLER (1957) [2]:

If for a given number $a > 0$ we have

$$(1.3) \qquad n \to \infty\,,\ x \to \infty\,,\ x^{3+a} n^{-\frac{1}{2}} \to 0,$$

then we can prove

$$(1.4) \qquad \frac{P\left\{\underline{S}_n^* > x\right\}}{1-\Phi(x)} = 1 + o(x^{-a}).$$

A summary of Smirnov's proof can be found in MOLENAAR, pages 106-113. After replacing $P\left\{\underline{S}_n^* > x\right\}$ by an incomplete B-integral, this proof uses a cunning device of repeated

----

1) Stochastic variables have been underlined.
2) VII problem 12, page 180. In (6.8) loc.cit. x^a should be replaced by x^{-a}.

partial integrations that yield a series of even powers of x^{-1}, which we can stop as soon as the exponent has become larger than a. This series, multiplied by $\frac{1}{x} e^{-\frac{1}{2} x^2}$, is then with the desired accuracy equal to $1-\Phi(x)$; see the Appendix of this Report. SMIRNOV's method is clever, but rather elaborate because of the numerous calculations.

A second method, derived from the way in which FELLER (VII.5) proves a simpler limit theorem, is given in section 2. It proceeds by dividing the event $\{\underline{S}_n^* > x\}$ in $\{x < \underline{S}_n^* \leq x_d\}$ and $\{\underline{S}_n^* > x_d\}$. One has to choose x_d in such a way, that the probability of the first component can be calculated according to the classical De Moivre-Laplace theorem, i.e.

$$x_d^3 \, n^{-\frac{1}{2}} \to 0,$$

while the probability of the second event must be small compared to that of the first one (the quotient must be smaller than x^{-a}).

In his University course 1961-'62 RUNNENBURG has given quite a different proof of FELLER's theorem VII.5. In section 3 his method will be reproduced, and an analogous result for the Poisson distribution will be derived in the same way. In section 4 we shall see the difficulties that will arise if we try to improve RUNNENBURG's method into a proof of (1.4). The Appendix (section 5) contains the proof of an auxiliary theorem.

2. A method based on FELLER's proof.

In this section, when we use n, x, $\mathcal{O}$ and σ, we tacitly assume (1.3) to be given.

First we want to apply STIRLING's formula to the k^{th} term of the binomial distribution. If we define x_k by $k=np+x_k\sqrt{npq}$, then $n-k=nq-x_k\sqrt{npq}$. Now let us assume that

$$x_k^3 \, n^{-\frac{1}{2}} \to 0; \text{ so } k=\mathcal{O}(n) \text{ and } n-k=\mathcal{O}(n);$$

then we have .

$$\binom{n}{k} p^k q^{n-k} = \frac{n^n e^{-n} \sqrt{2\pi n}\, e^{\mathcal{O}(n^{-1})}\, p^k q^{n-k} \ldots}{k^k e^{-k} \sqrt{2\pi k}\,(n-k)^{n-k} e^{-n+k} \sqrt{2\pi(n-k)}} =$$

$$= \frac{1}{\sqrt{2\pi \frac{k(n-k)}{n}}} \cdot \frac{1+\mathcal{O}(n^{-1})}{(\frac{k}{np})^k (\frac{n-k}{nq})^{n-k}} =$$

$$= \frac{1}{\sqrt{2\pi n(p+\frac{x_k\sqrt{pq}}{\sqrt{n}})(q-\frac{x_k\sqrt{pq}}{\sqrt{n}})}} \cdot \frac{1+\mathcal{O}(n^{-1})}{\left(1+\frac{x_k\sqrt{q}}{\sqrt{np}}\right)^k \left(1-\frac{x_k\sqrt{p}}{\sqrt{nq}}\right)^{n-k}}$$

The logarithm of the denominator of the second factor is equal to

$$(np+x_k\sqrt{npq}) \sum_{j=1}^{\infty} \frac{(-1)^{j-1}}{j} \left(\frac{x_k\sqrt{q}}{\sqrt{np}}\right)^j - (nq-x_k\sqrt{npq}) \sum_{j=1}^{\infty} \frac{1}{j} \left(\frac{x_k\sqrt{p}}{\sqrt{nq}}\right)^j =$$

$$= \tfrac{1}{2} x_k^2 + r(x_k,n).$$

Conclusion: $x_k^3\, n^{-\frac{1}{2}} \to 0$ and $k=np+x_k\sqrt{npq}$ imply

$$(2.1)\qquad \binom{n}{k} p^k q^{n-k} = \frac{1}{\sqrt{2\pi n\, p\, q}}\, e^{-\frac{1}{2} x_k^2} \left\{1+R(x_k,n)\right\} = \frac{\varphi(x_k)}{\sqrt{n\, p\, q}} \left\{1+R(x_k,n)\right\}.$$

Below, we will say more about the remainder term $R(x_k,n)$.

Given the quantity x, that increases with n in the way defined in (1.3), we can determine positive integers c and d such that

$$(2.2)\qquad \begin{gathered} c=np + x_c\sqrt{npq},\quad d=np + x_d\sqrt{npq}, \\ x_c < x \leqslant x_{c+1},\quad x_d \leqslant x+\log x < x_{d+1}. \end{gathered}$$

As $x \to \infty$, it follows that x_c, x_d, c and d increase with n and x.

Now we have $x_k^3\, n^{-\frac{1}{2}} \to 0$ for every k with $c < k \leq d$, because

$$0 < x_k^3\, n^{-\frac{1}{2}} \leq x_d^3\, n^{-\frac{1}{2}} \leq (x+\log x)^3\, n^{-\frac{1}{2}},$$

while we have $x^3\, n^{-\frac{1}{2}} = \sigma(x^{-a})$ according to (1.3). Thus we may apply (2.1):

$$P\left\{c < \underline{S}_n \leq d\right\} = \sum_{k=c+1}^{d} \binom{n}{k} p^k q^{n-k} =$$

$$= \frac{1}{\sqrt{npq}} \sum_{k=c+1}^{d} \varphi(x_k) \left\{1+R(x_k,n)\right\} .$$

From a more exact computation of $R(x_k,n)$ one can easily see that

$$\left|R(x_k,n)\right| \leq M\, x_d^3\, n^{-\frac{1}{2}} = \sigma(x^{-a}) \qquad (c < k \leq d),$$

where we can take for M a constant not depending on k,x or n, provided that x and n are sufficiently large and that $x_d n^{-\frac{1}{2}}$ is sufficiently small. Using this uniform estimate one finds that

$$P\left\{c < \underline{S}_n \leq d\right\} = \frac{1+\sigma(x^{-a})}{\sqrt{npq}} \sum_{k=c+1}^{d} \varphi(x_k).$$

The definition of x_k yields that

$$x_{k+\frac{1}{2}} - x_{k-\frac{1}{2}} = \frac{1}{\sqrt{npq}} \quad ;$$

hence the first mean value theorem of differential calculus implies the existence of numbers ξ_k that satisfy

$$\Phi(x_{k+\frac{1}{2}}) - \Phi(x_{k-\frac{1}{2}}) = \frac{1}{\sqrt{npq}}\varphi(\xi_k) \text{ and } x_{k-\frac{1}{2}} < \xi_k < x_{k+\frac{1}{2}} \ ;$$

from the last statement it follows that $\left|x_k - \xi_k\right| < \frac{1}{2\sqrt{npq}}$.

Now

$$\frac{1}{\sqrt{npq}}\varphi(x_k) = e^{\frac{1}{2}(\xi_k^2 - x_k^2)}\left\{\Phi(x_{k+\frac{1}{2}}) - \Phi(x_{k-\frac{1}{2}})\right\} ;$$

and

$$\tfrac{1}{2}\left|\xi_k^2-x_k^2\right| = \tfrac{1}{2}\left|\xi_k-x_k\right|(\xi_k+x_k) \leqslant \frac{1}{4\sqrt{npq}}\left\{2\ x_k+\frac{1}{2\sqrt{npq}}\right\} \leqslant$$

$$\leqslant \frac{1}{4\sqrt{npq}}\left\{2\ x_d+\frac{1}{2\sqrt{npq}}\right\} = \mathcal{O}(x\ n^{-\frac{1}{2}}) = \sigma(x^{-a}),$$

because we have $k \leqslant d$.

Combination of these results leads to

(2.3) $$P\left\{c \leqslant \underline{S}_n \leqslant d\right\} = \left\{\Phi(x_{d+\frac{1}{2}}) - \Phi(x_{c+\frac{1}{2}})\right\}\left\{1+\sigma(x^{-a})\right\}.$$

Now we will use FELLER's result proved in section 5: for every integer $b \geqslant 0$ we have for $y \to \infty$

$$1-\Phi(y) \backsim \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2}\left\{\frac{1}{y}-\frac{1}{y^3}+\frac{1.3}{y^5}-\ldots+(-1)^b\frac{1.3.5..(2b-1)}{y^{2b+1}}\right\}.$$

From the choice of c and d in (2.2) we see that

$$\frac{x_{c+\frac{1}{2}}}{x} = 1+\sigma(n^{-\frac{1}{2}}); \quad \frac{x_{d+\frac{1}{2}}}{x+\log x} = 1+\sigma(n^{-\frac{1}{2}}).$$

So $1-\Phi(x_{c+\frac{1}{2}}) = \left\{1-\Phi(x)\right\}\left\{1+\sigma(x^{-a})\right\}$, and

$$\frac{1-\Phi(x_{d+\frac{1}{2}})}{1-\Phi(x)} = \frac{e^{-\frac{1}{2}x^2_{d+\frac{1}{2}}}\left\{\frac{1}{x_{d+\frac{1}{2}}}-\frac{1}{x^3_{d+\frac{1}{2}}}+\ldots\right\}}{e^{-\frac{1}{2}x^2}\left\{\frac{1}{x}-\frac{1}{x^3}+\ldots\right\}} =$$

$$= e^{-\frac{1}{2}(x+\log x)^2+\frac{1}{2}x^2}\cdot\frac{1-\frac{\log x}{x}+\frac{\log^2 x-1}{x^2}+\mathcal{O}(\frac{\log^3 x}{x^3})}{1-\frac{1}{x^2}+\mathcal{O}(\frac{1}{x^4})}\left\{1+\sigma(n^{-\frac{1}{2}})\right\} =$$

$$= x^{-x}e^{-\frac{1}{2}\log^2 x}\left\{1-\frac{\log x}{x}+\mathcal{O}(\frac{\log^2 x}{x^2})\right\}\left\{1+\sigma(n^{-\frac{1}{2}})\right\} = \sigma(x^{-a}),$$

for in the long run x will be larger than the fixed positive number a. We use the last inferences in (2.3):

$$P\left\{c < \underline{S}_n \leqslant d\right\} = \left\{\left[1-\Phi(x_{c+\frac{1}{2}})\right] - \left[1-\Phi(x_{d+\frac{1}{2}})\right]\right\}\left\{1+\sigma(x^{-a})\right\} =$$

$$(2.4) \quad = \left\{1-\Phi(x)\right\}\left\{1-\frac{1-\Phi(x_{d+\frac{1}{2}})}{1-\Phi(x)}\right\}\left\{1+\sigma(x^{-a})\right\} =$$

$$= \left\{1-\Phi(x)\right\}\left\{1+\sigma(x^{-a})\right\}.$$

Finally, we will show that $P\left\{\underline{S}_n > d\right\}$ becomes negligibly small [1]. It is trivial that

$$\frac{\binom{n}{k} p^k q^{n-k}}{\binom{n}{k-1} p^{k-1} q^{n-k+1}} = \frac{(n-k+1)p}{kq},$$

and thus

$$\frac{P\left\{\underline{S}_n > d\right\}}{P\left\{\underline{S}_n = d\right\}} = \sum_{\nu=1}^{n-d} \frac{\binom{n}{d+\nu} p^{d+\nu} q^{n-d-\nu}}{\binom{n}{d} p^d q^{n-d}} \leqslant \sum_{\nu=1}^{n-d} \left\{\frac{(n-d)p}{dq}\right\}^{\nu};$$

for the ν^{th} term of the sum is the product of ν quotients of subsequent binomial terms, and (2.5) states that every factor of this product is not larger than $\frac{(n-d)p}{dq}$. Now we majorize the finite geometric series by summing from $\nu=0$ to oo:

$$\frac{P\left\{\underline{S}_n > d\right\}}{P\left\{\underline{S}_n = d\right\}} \leqslant \frac{1}{1-\frac{(n-d)p}{dq}} = \frac{dq}{dq-np+dp} = \frac{npq+x_d q\sqrt{npq}}{x_d\sqrt{npq}} =$$

$$= \frac{\sqrt{npq}}{x_d}\left\{1+O\left(\frac{x_d}{\sqrt{n}}\right)\right\} = \frac{\sqrt{npq}}{x_d}\left\{1+\sigma(x^{-a})\right\}.$$

Next we use the estimate (2.1) for $P\left\{\underline{S}_n = d\right\} = \binom{n}{d} p^d q^{n-d}$:

---

1) This needs a separate proof, because we can not yet state that $P\left\{\underline{S}_n > d\right\} = \left\{1-\Phi(x_d)\right\}\left\{1+\sigma(x^{-a})\right\}$.

$$P\left\{\underline{S}_n > d\right\} \leqslant \frac{1}{\sqrt{2\pi npq}}\ e^{-\frac{1}{2}x_d^2}\ .\ \frac{\sqrt{npq}}{x_d}\ \left\{1+\sigma(x^{-a})\right\} =$$

$$= \frac{1}{x_d\sqrt{2\pi}}\ e^{-\frac{1}{2}x_d^2}\ \left\{1+\sigma(x^{-a})\right\} .$$

$$\frac{P\left\{\underline{S}_n > d\right\}}{1-\Phi(x)} \leqslant \frac{\frac{1}{x_d\sqrt{2\pi}}\ e^{-\frac{1}{2}x_d^2}\ \left\{1+\sigma(x^{-a})\right\}}{\frac{1}{\sqrt{2\pi}}\ e^{-\frac{1}{2}x^2}\left\{\frac{1}{x} - \frac{1}{x^3} + \frac{1.3}{x^5} - \dots\right\}} =$$

$$= e^{-\frac{1}{2}(x+\log x)^2+\frac{1}{2}x^2}\ \left\{1-\frac{\log x}{x} + O\left(\frac{\log^2 x}{x^2}\right)\right\}\left\{1+\sigma(x^{-a})\right\} =$$

$$= x^{-x}\ e^{-\frac{1}{2}\log^2 x}\ \left\{1+\sigma(x^{-a})\right\} = \sigma(x^{-a})\ ;$$

here we have used that $\frac{x_d}{x+\log x} = 1+\sigma(n^{-\frac{1}{2}})=1+\sigma(x^{-a})$, cf. (2.2) and (1.3). Now we have proved

$$\frac{P\left\{\underline{S}_n > d\right\}}{1-\Phi(x)} = \sigma(x^{-a})\ ; \qquad (2.6)$$

combining (2.4) and (2.6) we see that

$$P\left\{\underline{S}_n^* > x\right\} = P\left\{c < \underline{S}_n \leqslant d\right\} + P\left\{\underline{S}_n > d\right\} =$$

$$= \left\{1-\Phi(x)\right\}\ .\left\{1+\sigma(x^{-a})\right\} .$$

This completes our proof of Smirnov's result (1.4).

3. RUNNENBURG's method.

Again we have n independent Bernoulli trials with probability p for success, $0<p<1$. If (1.1) is given, RUNNENBURG proves (1.2) in the following way.

First of all he estimates one term of the binomial distribution, in the way we did it in section 2. He makes this estimation for $r=np+x\sqrt{npq}$; without loss of generality we may take r to be an integer. Because (1.1) is given, we find the analogon of (2.1):

$$(3.1) \qquad \binom{n}{r} p^r q^{n-r} \backsim \frac{1}{\sqrt{2\pi npq}} \; e^{-\frac{1}{2}x^2}.$$

Next we may put

$$(3.2) \qquad P\left\{\underline{S}_n^* > x\right\} = \sum_{\nu=1}^{n-r} \frac{\binom{n}{r+\nu} p^{r+\nu} q^{n-r-\nu}}{\binom{n}{r} p^r q^{n-r}} \cdot \binom{n}{r} p^r q^{n-r} .$$

Because of (2.5) we have

$$(3.3) \qquad \begin{aligned} \sum_{\nu=1}^{n-r} \frac{\binom{n}{r+\nu} p^{r+\nu} q^{n-r-\nu}}{\binom{n}{r} p^r q^{n-r}} &= \sum_{\nu=1}^{n-r} \prod_{j=1}^{\nu} \frac{(n-r-j+1)p}{(r+j)q} = \\ &= \sum_{\nu=1}^{n-r} \frac{(n-r)^{!\nu} p^{\nu}}{(r+\nu)^{!\nu} q^{\nu}} = Y, \end{aligned}$$

i.e. we denote the last mentioned sum by Y, and we use the notation $a^{!\nu} = a(a-1)(a-2)\ldots(a-\nu+1)$.

Now the ν^{th} term of Y is smaller than $\left\{\frac{(n-r)p}{rq}\right\}^{\nu}$; therefore Y is majorized by a finite geometric series, and a fortiori by the infinite one. Thus we find that

$$(3.4) \qquad Y \leqslant \sum_{\nu=0}^{\infty} \left\{\frac{(n-r)p}{rq}\right\}^{\nu} = \frac{rq}{rq-(n-r)p} = \frac{npq+xq\sqrt{npq}}{x\sqrt{npq}} \backsim \frac{\sqrt{npq}}{x} .$$

In order to find a lower bound for Y, we break off the series after N terms, with $N \leq n-r$. In that case we have for each term

$$\frac{(n-r)!\,p^{\nu}}{(r+\nu)!\,q^{\nu}} \geq \left\{\frac{(n-r-N)p}{(r+N)q}\right\}^{\nu} \qquad ;$$

so

$$\text{(3.5)} \qquad Y \geq \sum_{\nu=1}^{N} \left\{\frac{(n-r-N)p}{(r+N)q}\right\}^{\nu} = \frac{\frac{(n-r-N)p}{(r+N)q}\left[1-\left\{\frac{(n-r-N)p}{(r+N)q}\right\}^{N}\right]}{1-\frac{(n-r-N)p}{(r+N)q}} .$$

We choose $N=c\sqrt{npq}$, where c is a positive constant. Then for sufficiently large n we have $n-r=nq-x\sqrt{npq} \geq N$, and on the other hand

$$\log\left\{\frac{(n-r-N)p}{(r+N)q}\right\}^{N} = c\sqrt{npq}\ \log\frac{\{nq-(x+c)\sqrt{npq}\}p}{\{np+(x+c)\sqrt{npq}\}q} =$$

$$=c\sqrt{npq}\left\{\log\left(1-\frac{p(x+c)}{\sqrt{npq}}\right)-\log\left(1+\frac{q(x+c)}{\sqrt{npq}}\right)\right\} = -c(x+c)+o(1).$$

But $\lim\left\{-c(x+c)\right\} = -\infty$, and thus

$$1-\left\{\frac{(n-r-N)p}{(r+N)q}\right\}^{N} \longrightarrow 1 \quad ;$$

so the right hand member of (3.5) is asymptotically equal to

$$\frac{\frac{(n-r-N)p}{(r+N)q}}{1-\frac{(n-r-N)p}{(r+N)q}} = \frac{(n-r-N)p}{r+N-np} =$$

(3.6)

$$= \frac{npq-p(x+c)\sqrt{npq}}{(x+c)\sqrt{npq}} \sim \frac{\sqrt{npq}}{x+c} \sim \frac{\sqrt{npq}}{x} .$$

Now we have upper and lower bounds for Y that are both asymptotically equal to $\frac{\sqrt{npq}}{x}$, so we may infer that

$$\text{(3.7)} \qquad Y \sim \frac{\sqrt{npq}}{x} .$$

Finally we recall FELLER's result for the normal distribution function $\Phi(y)$ mentioned in section 2 and proved in section 5:

$$1-\Phi(y) \backsim \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} \left\{ \frac{1}{y} - \frac{1}{y^3} + \frac{1.3}{y^5} - \ldots + (-1)^b \frac{1.3.5\ldots(2b-1)}{y^{2b+1}} \right\}.$$

We substitute b=0:

$$(3.8) \qquad 1-\Phi(y) \backsim \frac{1}{y\sqrt{2\pi}} e^{-\frac{1}{2}y^2} .$$

Now we combine (3.1), (3.2), (3.3), (3.7) and (3.8):

$$(3.9) \quad P\left\{ \underline{S}_n^* > x \right\} = Y\binom{n}{r} p^r q^{n-r} \backsim \frac{\sqrt{npq}}{x} \cdot \frac{1}{\sqrt{2\pi npq}} e^{-\frac{1}{2}x^2} \backsim 1-\Phi(x).$$

This proof by RUNNENBURG can also be adapted to prove the asymptotic normality of other discrete distributions.
In this way we obtained the following result for the Poisson distribution. If

$$(3.10) \qquad \lambda \to \infty, \quad x \to \infty, \quad x^3 \lambda^{-\frac{1}{2}} \to 0,$$

then

$$(3.11) \qquad P\left\{ \underline{k} > \lambda + x\sqrt{\lambda} \right\} \backsim 1-\Phi(x),$$

where

$$P\left\{ \underline{k} = k \right\} = p(k;\lambda) = e^{-\lambda} \frac{\lambda^k}{k!} .$$

Without loss of generality we may assume that $r = \lambda + x\sqrt{\lambda}$ is an integer. Then we have

$$P\{\underline{k}\geq\lambda+x\sqrt{\lambda}\}=\sum_{\nu=1}^{\infty}\frac{p(r+\nu;\lambda)}{p(r;\lambda)}.p(r;\lambda)=Y'.p(r;\lambda);$$

$$p(r;\lambda)=\frac{e^{-\lambda}\lambda^r}{r!}\backsim\frac{e^{-\lambda}\lambda^r}{r^r e^{-r}\sqrt{2\pi r}}=\frac{e^{x\sqrt{\lambda}}}{(1+\frac{x}{\sqrt{\lambda}})^r\sqrt{2\pi r}};$$

$$\log p(r;\lambda)\backsim x\sqrt{\lambda}-\tfrac{1}{2}\log 2\pi-\tfrac{1}{2}\log\lambda-(\lambda+x\sqrt{\lambda}+\tfrac{1}{2})\log(1+\frac{x}{\sqrt{\lambda}})=$$

$$=-\tfrac{1}{2}\log 2\pi\lambda-\tfrac{1}{2}x^2+\mathcal{O}(\frac{x^3}{\sqrt{\lambda}});$$

(3.12)
$$p(r;\lambda)\backsim\frac{1}{\sqrt{2\pi\lambda}}e^{-\frac{1}{2}x^2}.$$

$$\frac{p(k;\lambda)}{p(k-1;\lambda)}=\frac{\lambda}{k},\text{ so }\frac{p(r+\nu;\lambda)}{p(r;\lambda)}=\frac{\lambda^\nu}{(r+\nu)^{!\nu}};$$

(3.13)
$$Y'=\sum_{\nu=1}^{\infty}\frac{\lambda^\nu}{(r+\nu)^{!\nu}}\leq\sum_{\nu=0}^{\infty}\left(\frac{\lambda}{r}\right)^\nu=\frac{r}{r-\lambda}=\frac{\lambda+x\sqrt{\lambda}}{x\sqrt{\lambda}}\backsim\frac{\sqrt{\lambda}}{x}.$$

Take $N=c\sqrt{\lambda}$, where c is a positive constant; then we have

$$Y'\geq\sum_{\nu=1}^{N}\frac{\lambda^\nu}{(r+N)^\nu}=\frac{\frac{\lambda}{r+N}\left\{1-(\frac{\lambda}{r+N})^N\right\}}{1-\frac{\lambda}{r+N}}.$$

$$\log\left(\frac{\lambda}{r+N}\right)^N=-c\sqrt{\lambda}\log\left(1+\frac{x+c}{\sqrt{\lambda}}\right)=-c(x+c)+o(1)\longrightarrow-\infty.$$

So $(\frac{\lambda}{r+N})^N$ becomes negligible compared to 1, and we see that asymptotically

(3.14)
$$Y'\geq\frac{\frac{\lambda}{r+N}}{1-\frac{\lambda}{r+N}}=\frac{\lambda}{r+N-\lambda}=\frac{\sqrt{\lambda}}{x+c}\backsim\frac{\sqrt{\lambda}}{x}.$$

From (3.13) and (3.14) we deduce: $Y'\backsim\frac{\sqrt{\lambda}}{x}$; combining this with (3.12), we infer that

$$P\{\underline{k}>\lambda+x\sqrt{\lambda}\}\backsim\frac{\sqrt{\lambda}}{x}.\frac{1}{\sqrt{2\pi\lambda}}e^{-\frac{1}{2}x^2}\backsim 1-\Phi(x).$$

4. Applicability of this method in our case.

We will now investigate whether we can improve RUNNENBURG's method into a proof of (1.4) from (1.3), i.e. a solution of SMIRNOV's problem. The proof of section 3 is not sufficient here, not even if we make a different choice for N in (3.5). For the expressions of the form "$A \sim B$" mentioned in (3.6), (3.7) and (3.8) cannot be just replaced by "$A=B\left\{1+\mathcal{O}(x^{-a})\right\}$".

As can be seen from (2.1), the sharpened version of (3.1) is valid:

$$(4.1) \qquad \binom{n}{r} p^r q^{n-r} = \frac{1}{\sqrt{2\pi n\, p\, q}}\, e^{-\frac{1}{2}x^2}\left\{1+\mathcal{O}(x^{-a})\right\} .$$

Next, FELLER's result (5.1) and Remark 1 at the end of section 5 give us that

$$(4.2) \qquad 1-\Phi(x) = \frac{1}{x\sqrt{2\pi}}\, e^{-\frac{1}{2}x^2}\, S_b(x)\left\{1+\mathcal{O}(x^{-a})\right\} ,$$

if we put

$$(4.3) \qquad b=\left[\frac{a}{2}\right] \text{ and } S_b(x)=1-\frac{1}{x^2}+\frac{1.3}{x^4}-\ldots+(-1)^b\frac{1.3.5\ldots(2b-1)}{x^{2b}}.$$

In order to reach the desired result

$$P\left\{\underline{S}_n^* > x\right\} = \left\{1-\Phi(x)\right\}\left\{1+\mathcal{O}(x^{-a})\right\}$$

we will apparently have to prove for the sum Y defined in (3.3) that

$$(4.4) \qquad Y = \frac{\sqrt{npq}}{x}\, S_b(x)\left\{1+\mathcal{O}(x^{-a})\right\} .$$

This step turns out to be very complicated. We can write for Y

$$Y = \sum_{\nu=1}^{n-r} f_1 f_2 f_3 \ldots f_\nu, \text{ with } f_j = \frac{(n-r-j+1)p}{(r+j)q} .$$

An obvious first approximation for Y is

$$Y_1=\sum_{\nu=1}^{n-r} f_1^{\nu} = \frac{\frac{(n-r)p}{(r+1)q}\left\{1-\left(\frac{(n-r)p}{(r+1)q}\right)^{n-r}\right\}}{1-\frac{(n-r)p}{(r+1)q}} =$$

(4.6)
$$= \frac{(n-r)p}{r+q-np}\left\{1-\left(\frac{npq-xp\sqrt{npq}}{npq+xq\sqrt{npq}+q}\right)^{n-r}\right\} =$$

$$= \frac{npq-xp\sqrt{npq}}{x\sqrt{npq}+q}\left\{1-\frac{\left(1-\frac{xp}{\sqrt{npq}}\right)^{nq-x\sqrt{npq}}}{\left(1+\frac{xq}{\sqrt{npq}}+\frac{1}{np}\right)^{nq-x\sqrt{npq}}}\right\}.$$

We expand the logarithm of the last factor into a power series; substituting $p+q=1$, $p^2-q^2=p-q$, we see that it is equal to

$$-\frac{x\sqrt{nq}}{\sqrt{p}}+\frac{x^2}{2p}+\mathcal{O}\left(\frac{x^3}{\sqrt{n}}\right) .$$

So from (4.6) we get

$$Y_1=\frac{\sqrt{npq}}{x}\left\{1-\frac{x\sqrt{p}}{\sqrt{nq}}+\mathcal{O}\left(\frac{1}{x\sqrt{n}}\right)\right\}\left\{1-\exp\left(-\frac{x\sqrt{nq}}{\sqrt{p}}+\frac{x^2}{2p}+\mathcal{O}(x^{-a})\right)\right\}.$$

Now for x and n sufficiently large, we have:

$$\frac{x\sqrt{p}}{\sqrt{nq}}=\mathcal{O}(x^{-a}),\ \text{and}\ \frac{x\sqrt{nq}}{\sqrt{p}}-\frac{x^2}{2p}>a\log x,$$

which gives $\exp\left(-\frac{x\sqrt{nq}}{\sqrt{p}}+\frac{x^2}{2p}+\mathcal{O}(x^{-a})\right)<x^{-a}$.

Substituting this, we may conclude that

$$Y_1 = \frac{\sqrt{npq}}{x}\left\{1+\mathcal{O}(x^{-a})\right\} .$$

The next step should be an estimation of the difference

$$\Delta_1=Y_1-Y=f_1(f_1-f_2)+f_1(f_1^2-f_2f_3)+f_1(f_1^3-f_2f_3f_4) +$$
$$+ \dots + f_1(f_1^{n-r-1}-f_2f_3\dots f_{n-r}) .$$

We would like to prove that $\Delta_1 < \Gamma_1 = \frac{\sqrt{npq}}{x^3}$; next to estimate $\Delta_2 = \Gamma_1 - \Delta_1$ by $\frac{3\sqrt{npq}}{x^5}$, etc.
Thus we would approximate Y by

$$\frac{\sqrt{npq}}{x}\left\{1-\frac{1}{x^2}+\frac{1.3}{x^4}-\frac{1.3.5}{x^6}+\ldots\right\};$$

which would be (4.4) if we continued until the exponent has become 2b. But the difficulty of finding the desired estimations for Δ_1, Δ_2, etc. remains unsolved.
Some calculation yields that for $\frac{j}{n} = \sigma(1)$

$$f_j = 1-\frac{x}{\sqrt{npq}}+\frac{x^2}{np}-\frac{j-p}{npq}+\mathcal{O}\left(\frac{x^3}{n\sqrt{n}}\right)+\mathcal{O}\left(\frac{jx}{n\sqrt{n}}\right)+\mathcal{O}\left(\frac{j^2}{n^2}\right).$$

After some more computations we find that the k^{th} term of Δ_1 is approximately equal to

$$\left(\sum_{j=2}^{k+1} j-k\right)(npq)^{-1} = \frac{k(k+1)}{2\,npq},$$

but this is only valid if $k=\sigma(n)$, and besides, the series of these estimations diverges.

The Poisson distribution presents analogous difficulties. Here we have

$$p(r;\lambda)=\frac{1}{\sqrt{2\pi\lambda}}\, e^{-\frac{1}{2}x^2}\left\{1+\mathcal{O}\left(\frac{x^3}{\sqrt{\lambda}}\right)\right\},$$

cf. the derivation of (3.12). Now because of (4.2) there remains to be proved that

$$Y' = \sum_{\nu=1}^{\infty}\frac{\lambda^{\nu}}{(r+\nu)!_{\nu}} = \frac{\sqrt{\lambda}}{x}\; S_b(x)\left\{1+\sigma(x^{-a})\right\},$$

if it is given that $x^3\lambda^{-\frac{1}{2}} = \sigma(x^{-a})$. After a first approximation

$$Y'_1 = \sum_{\nu=1}^{\infty}\frac{\lambda^{\nu}}{(r+1)^{\nu}} = \frac{\sqrt{\lambda}}{x}\left\{1+\sigma(x^{-a})\right\},$$

the estimation of $\Delta_1' = Y_1' - Y'$ leaves us in almost the same situation as in the binomial case.

Another procedure that was investigated is the division of the series Y (or Y', for the Poisson distribution) into a large number of subseries, each of which can be approximated sufficiently closely by a geometric series. The fact that we cannot write the k^{th} term of Y, for somewhat large values of k, in a manageable form, appeared to be an obstacle for this method too.

Summary

Our proof given in section 2 is a short solution of SMIRNOV's problem of proving (1.4) from (1.3). SMIRNOV's proof is very ingenious, but somewhat elaborate; RUNNENBURG's method exposed in section 3, which seems so clear and elegant, does not appear to furnish a solution for this particular problem just now.

Appendix

5. Proof of FELLER's result.

The normal distribution function $\Phi(y)$ has for every non-negative integer b and for $y \to \infty$ the property

$$1-\Phi(y) \sim \frac{1}{y\sqrt{2\pi}} e^{-\frac{1}{2}y^2} S_b(y) =$$

(5.1)

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} \left\{ \frac{1}{y} - \frac{1}{y^3} + \frac{1.3}{y^5} - \ldots + (-1)^b \frac{1.3.5\ldots(2b-1)}{y^{2b+1}} \right\};$$

for each positive y, the right-hand side overestimates $1-\Phi(y)$ if b is even and underestimates it if b is odd.
This result is mentioned, but not proved, in FELLER (1957), VII.6 problem 1, p.179. For completeness' sake we will prove it here.

Differentiating the last member of (5.1) with respect to y, we get

$$-\frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}y^2} \left\{1+(-1)^b \frac{1.3.5\ldots(2b-1)(2b+1)}{y^{2b+2}}\right\} \quad ;$$

this is equal to

$$\frac{d}{dy} \int_y^\infty \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}t^2} \left\{1+(-1)^b \frac{1.3.5.\ldots.(2b+1)}{t^{2b+2}}\right\} dt.$$

Both functions differentiated here tend to zero for $y \to \infty$; so the functions are identically equal. The integrand of the last mentioned function is for even b larger, for odd b smaller than the normal frequency function $\varphi(x)$. Now if we assume that b is even, and put R(b,y) for the right-hand member of (5.1), then

(5.2) $$R(b+1,y) < 1-\Phi(y) < R(b,y),$$

and the quotient of R(b+1,y) and R(b,y) is equal to $1+O(y^{-2b-2})$. If b is odd, we have the same formula with ">" instead of "<". This proves (5.1) and also the assertion about over- and underestimation.

Remark 1: The series $S_b(y)$ has more or less the same character as STIRLING's series. For fixed y and $b \to \infty$ it diverges, but for fixed finite b we have for every $\varepsilon > 0$ a minimum-y, such that the outer members of (5.2) do not differ more than ε. Because $1-\Phi(y)$ is alternately overestimated and underestimated, the error is always smaller in absolute value than the first neglected term.
Therefore

$$1-\Phi(y) = \frac{1}{y\sqrt{2\pi}}\, e^{-\frac{1}{2}y^2}\, S_b(y) \left\{1+O(y^{-2b-2})\right\}.$$

For $b=\left[\frac{a}{2}\right]$ we have $2b+2 > a$; so (4.2) is correct.

Remark 2: FELLER's result is not valid for $y<0$ ($y \to -\infty$), and our proof fails in that case because of the divergence in $y=0$. Here we can prove, in the same way, that

$$\Phi(y) \backsim - \frac{1}{y\sqrt{2\pi}} e^{-\frac{1}{2}y^2} S_b(y) =$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} \left\{ - \frac{1}{y} + \frac{1}{y^3} - \frac{1.3}{y^5} + \ldots + (-1)^{b-1} \frac{1.3.5 \ldots (2b-1)}{y^{2b+1}} \right\}.$$

References.

W. FELLER, An Introduction to Probability Theory and its Applications, 2nd Impression, John Wiley and Sons, New York (1957).

W. MOLENAAR, Problems from FELLER's book, (in Dutch), typewritten internal report, 1962.

N. SMIRNOV, About the probabilities of large deviations (in Russian, German summary), Mat.Sbornik, Moscou, 40 (1933) 443-454.