ON THE NORMAL APPROXIMATION TO THE BINOMIAL DISTRIBUTION

By W. Feller

Cornell University

1. Although the problem of an efficient estimation of the error in the normal approximation to the binomial distribution is classical, the many papers which are still being written on the subject show that not all pertinent questions have found a satisfactory solution. Let for a fixed n and $0 < p < 1$, $q = 1 - p$,

$$(1) \qquad T_k = \binom{n}{k} p^k q^{n-k}, \qquad P_{\lambda,\nu} = \sum_{k=\lambda}^{\nu} T_k .$$

For reasons of tradition (and, apparently, only for such reasons) one sets

$$(2) \qquad z_k = (k - np)\sigma^{-1}, \qquad \sigma = (npq)^{1/2},$$

and compares (1) with

$$(3) \qquad N_k = (2\pi)^{-1/2}\sigma^{-1} e^{-z_k^2/2} \quad \text{and} \quad \Pi_{\lambda,\nu} = \Phi\left(z_\nu + \frac{1}{2\sigma}\right) - \Phi\left(z_\lambda - \frac{1}{2\sigma}\right)$$

respectively,[1] where $\Phi(z)$ stands for the normalized error function. Many estimates are available for the maximum of the difference $|P_{\lambda,\nu} - \Pi_{\lambda,\nu}|$ for all λ, ν. Now this error is $O(\sigma^{-1})$ and even a precise appraisal will break down in the two most interesting cases: if σ is small, or if λ and ν are large as compared to σ. Indeed, even for moderately large values of k (such as are usually considered) the contribution of T_k to the sum in (1) will be considerably smaller than σ^{-1} so that any estimate of the form $O(\sigma^{-1})$ leaves us without guidance. With some modifications this remains true also for more refined estimates like Uspensky's remarkable result[2]

$$(4) \qquad P_{\lambda,\nu} = \Pi_{\lambda,\nu} + \frac{q-p}{6\sigma(2\pi)^{1/2}}\left[(1 - z^2)e^{-z^2/2}\right]\Big|_{z_\lambda - 1/2\sigma}^{z_\nu + 1/2\sigma} + \omega$$

with

$$|\omega| < \{.13 + .18\,|p - q|\}\sigma^{-2} + e^{-3\sigma/2}$$

provided $\sigma \geq 5$. What is really needed in many applications is an estimate of the relative error, but this seems difficult to obtain.

It should also be noticed that the accuracy of the normal approximation to the binomial is by no means quite as good as many texts would make appear. Exam-

[1] Very often the limits z_λ and z_ν instead of $z_\nu + \frac{1}{2\sigma}$ and $z_\lambda - \frac{1}{2\sigma}$ are used. This naturally results in an unnecessary systematic undervaluation.

[2] Uspensky [3], p. 129. A two-term development of T_r with an error of $O(\sigma^{-2})$ valid for $|x| < 2$, $\sigma > 3$ has been given by Mirimanoff and Dovaz [1927].

ples using $p = \frac{1}{2}$ and intervals which are symmetric with respect to np are hardly conclusive, since there the main error term drops out and systematic positive and negative errors cancel. Again, in practice comparatively small σ and comparatively large ν are frequently used. It works well to compare a $P_{\lambda,\nu}$ of a numerical value, say, .93 with a corresponding value $\Pi_{\lambda,\nu}$ of, say, .95. In classroom discussions the error may seem insignificant. However, in most actual applications one would consider the complementary probabilities, and the very same figures mean an approximation .05 to the correct value .07. If a confidence limit is set to the five per cent level, the normal approximation would in our example mean that two out of seven critical cases are missed. Consider next the *example* $p = \frac{1}{10}$, $n = 10{,}000$. For values of k around 1120 the relative error of N_k is about .30; it increases rapidly with increasing k. Around $k = 1150$ the relative error exceeds 2/3, around 1180 it is nearly 1.4. And yet this example is conservative in comparison with many cases where the normal approximation is used in practice.

It is surprising that the classical norming (2) is generally accepted although there does not seem to exist any deeper reason for it. The use of moments, though usually very convenient, does not necessarily lead to best results. For example, the density function

$$(5) \qquad f_n(x) = \frac{1}{n!} x^n e^{-x}$$

is the $(n + 1)$-fold convolution of $f_0(x)$ with itself and therefore, for large n, of nearly normal "type." The conventional norming would approximate $f_n(x)$ by $\{2\pi(n + 1)\}^{-1/2} e^{-\{x-(n+1)\}^2/2(n+1)}$, while the use of the norming factor n instead of $(n + 1)$ seems clearly indicated.

Actually, as will be seen, it is natural (at least for small values of $k - np$) to replace (2) by

$$(6) \qquad x_k = \{k + \tfrac{1}{2} - (n + 1)p\}\sigma^{-1},$$

and accordingly to approximate $P_{\lambda,\nu}$ by the error integral taken between the limits

$$(7) \qquad \{\lambda - (n + 1)p\}\sigma^{-1} \quad \text{and} \quad \{\nu + 1 - (n + 1)p\}\sigma^{-1}.$$

For example, let $p = \frac{1}{10}$, $n = 500$, $\lambda = 50$, $\nu = 55$. The correct value is $P_{50,55} \approx .317573$; the norming (2) leads to $\Pi_{50,55} \approx .32357$, while the more natural limits (6) lead to an approximation .31989. More important are the quite unexpected simplifications which the norming (6) permits when one studies the error for large x_k or small σ.

We are now led to reformulate the problem: *instead of starting with arbitrary limits for the error integral and to estimate the resulting error, we shall try to determine the limits so as to minimize the error.* Theoretically, for any given λ, ν these limits could be determined so as to give an exact value for $P_{\lambda,\nu}$. However, such limits would depend in the most intricate way on λ and ν. For practical purposes one would restrict the considerations to certain simple functions such as polynomials.

We shall here consider only the case where the limits are at most quadratic polynomials. Essentially our problem seems that treated by Serge Bernstein (and, apparently, only by him). In a series of papers since 1924, S. Bernstein has considered the accuracy of the normal approximation. Quite recently[3] he has, by a considerable computational effort, extended the range of validity from $npq \geq 365$ to $npq \geq 62.5$ and proved the following

THEOREM (*S. Bernstein*): *Let*

$$npq \geq 62.5 \tag{8}$$

and let α_x, β_x *be the solutions of the quadratic equations*

$$\begin{aligned} x - \tfrac{3}{2} - np &= \alpha_x(npq)^{1/2} + \frac{q-p}{6}\,\alpha_x^2 \\ x + \tfrac{1}{2} - np &= \beta_x(npq)^{1/2} + \frac{q-p}{6}\,\beta_x^2 . \end{aligned} \tag{9}$$

If

$$\alpha \geq 0, \qquad \beta < 2^{1/2}(npq)^{1/6} \tag{10}$$

then

$$\Phi(\beta_\nu) - \Phi(\beta_\lambda) \leqq P_{\lambda,\nu} \leq \Phi(\alpha_\nu) - \Phi(\alpha_\lambda). \tag{11}$$

The conditions (10) are practically equivalent to

$$\lambda \geq np + \tfrac{3}{2}, \qquad \nu \leq np + 2^{1/2}\sigma^{1/3}. \tag{12}$$

The remarkable feature of this excellent result is that the error remains $O(\sigma^{-1})$ throughout an interval which increases with σ (instead of the conventional uniformly bounded intervals).

In the sequel it will be shown that startling simplifications can be obtained if the norming (6) is used from the beginning instead of (2). Our main result is an improvement of S. Bernstein's theorem. The condition (8) will be replaced by $(n+1)pq \geq 9$. The first condition in (10) will be relaxed to $k \geq (n+1)p$, that is to say, our theorem will hold for all k exceeding the central value (for those less than the central value an analogous theorem holds); in the other condition (10), the numerical value $2^{1/2}$ will be replaced by an arbitrary constant. Instead of quadratic equations, we shall consider quadratic polynomials. And finally, the gap between the two sets of limits will be reduced.

It will be seen that the computations leading to this improvement are almost negligible in comparison with S. Bernstein's deeper method; with slightly more sophisticated arguments and numerical evaluations, our results can be considerably improved. Our consideration will be based on a new expression for T_k, in which only exponential terms appear but the usual square root is missing.

[3] S. Bernstein [1], the first paper of the series appears to have appeared in Učenye Zapiski, Kiev, 1924.

In passing from approximations to T_k to approximations to $P_{\lambda,\nu}$ one has to replace sums by integrals. This procedure is cumbersome if an estimate of the relative error is desired. Euler's formula and other standard formulas are of little use. We shall therefore start with a lemma which, it is hoped, may be useful in this connection; it will therefore be proved in a slightly more general form than actually required for the present paper.

2. LEMMA[4] 1. *For* $0 < h < \frac{1}{3}$ *and* $|xh| \leq 1$

$$\int_{x-h/2}^{x+h/2} e^{-u^2/2}\,du = he^{-x^2/2+(x^2-1)h^2/24+\omega h^4}, \tag{13}$$

with

$$-\frac{x^4}{880} \leq \omega \leq \frac{1}{285}. \tag{14}$$

PROOF. Denote the integral in (13) by J. Then

$$h^{-1}e^{x^2/2}J = h^{-1}\int_{-h/2}^{h/2} e^{-xt-t^2/2}\,dt = 2h^{-1}\int_0^{h/2} chxte^{-t^2/2}\,dt. \tag{15}$$

We begin by showing that for $0 \leq \alpha \leq \frac{1}{2}$

$$e^{\alpha^2/2-\alpha^4/11} \leq ch\alpha \leq e^{\alpha^2/2-\alpha^4/16}. \tag{16}$$

In fact

$$e^{\alpha^4/11}ch\alpha \geq \left(1+\frac{\alpha^2}{2}+\frac{\alpha^4}{24}\right)\left(1+\frac{\alpha^4}{11}\right) \geq 1+\frac{\alpha^2}{2}+\frac{\alpha^4}{8}\sum_0^\infty\left(\frac{1}{24}\right)^\rho \geq e^{\alpha^2/2}, \tag{17}$$

and

$$e^{\alpha^2/2-\alpha^4/16} \geq \left(1+\frac{\alpha^2}{2}+\frac{\alpha^4}{8}\right)\left(1-\frac{\alpha^4}{16}\right) \geq 1+\frac{\alpha^2}{2}+\frac{\alpha^4}{24}\cdot\frac{1}{1-\frac{1}{120}} \geq ch\alpha. \tag{18}$$

It follows from (15) and (16) that

$$\begin{aligned} h^{-1}e^{x^2/2}J &\geq 2h^{-1}\int_0^{h/2} e^{(x^2-1)t^2/6-x^4t^4/55}\cdot e^{(x^2-1)t^2/3-4x^4t^4/55}\,dt \\ &\geq 2h^{-1}\int_0^{h/2} e^{(x^2-1)t^2/6-x^4t^4/55}\left\{1+\frac{x^2-1}{3}t^2-\frac{4x^4t^4}{55}\right\}dt \\ &= 2h^{-1}\left[te^{(x^2-1)t^2/6-x^4t^4/55}\right]_0^{h/2} \end{aligned} \tag{19}$$

which proves one part of the lemma.

To obtain an upper estimate we make use of the inequalities

$$e^{(x^2-1)t^2/3} \leq \left(1+\frac{x^2t^2}{3}\right)e^{-t^2/3+x^4t^4/18}$$

[4] The fraction $\frac{1}{3}$ is chosen quite arbitrarily; if h be restricted to $0 < h < 1$ the first member of (14) remains unchanged, while the fraction $\frac{1}{285}$ on the right side has to be replaced by $\frac{1}{264}$.

$$\leq \left(1 + \frac{x^2 t^2}{3}\right)\left(1 - \frac{t^2}{3}\right) e^{x^4 t^4/18} \cdot \frac{1 - \frac{t^2}{3} + \frac{t^4}{18}}{1 - \frac{t^2}{3}} \tag{20}$$

$$\leq \left(1 + \frac{x^2 - 1}{3} t^2\right) e^{x^4 t^4/18 + h^4/285}$$

Using (16) and (20), the proof of the second part of the lemma follows from a computation analogous to (19).

For our purposes it is convenient to use Stirling's formula in a form which is not quite the usual one.

LEMMA 2. (*Stirling's formulas*). For $n \geq 4$,

$$n! = (2\pi)^{\frac{1}{2}}(n + \tfrac{1}{2})^{n+\frac{1}{2}} e^{-(n+\frac{1}{2}) - 1/24(n+\frac{1}{2}) + (7/2880)(1+\vartheta_1)/(n+\frac{1}{2})^3}, \tag{21}$$

or

$$n! = (2\pi)^{\frac{1}{2}} n^{n+\frac{1}{2}} e^{-n+1/12n-(1+\vartheta_2)/360n^3} \tag{22}$$

where

$$| \vartheta_i | < \tfrac{1}{6}, \qquad \vartheta \to 0 \quad \text{as} \quad n \to \infty. \tag{23}$$

Formula (21) can be derived from the gamma function or in any other way that leads to the standard form (22).[5]

3. From now on we shall put

$$\sigma^2 = (n + 1)pq \tag{24}$$

$$x_k = \{k + \tfrac{1}{2} - (n + 1)p\}\sigma^{-1}; \tag{25}$$

the subscript k will be omitted whenever no confusion is to be feared. To transform T_k we shall use (21) for the factorials in the denominator, but (22) for $(n + 1)!$ in the numerator.

[5] A simple proof runs as follows. Put $B_n = n!(n + \frac{1}{2})^{-(n+\frac{1}{2})} e^{n+\frac{1}{2}+1/24(n+\frac{1}{2})}$. Then

$$\log \frac{B_{\rho-1}}{B_\rho} - \sum_{\nu=2}^{\infty} \left\{\frac{1}{6} - \frac{1}{2\nu(2\nu + 1)}\right\} \frac{1}{(2\rho)^{2\nu}} = \frac{7}{60} \frac{1 + \delta_1}{(2\rho)^4}$$

with $0 < \delta_1 < \frac{1}{70}$ if $\rho \geq 5$. From here (21) follows using the fact that

$$\sum_{\rho=n+1}^{\infty} \log \frac{B_{\rho-1}}{B_\rho} = \log B_n - \tfrac{1}{2} \log (2\pi)$$

and that for $n \geq 4$

$$\frac{1 - \delta}{3(n + \frac{1}{2})^3} < \sum_{n+1}^{\infty} \frac{1}{\rho^4} < \frac{1}{3(n + \frac{1}{2})^3}$$

with $0 < \delta < \frac{3}{25}$. In this way the estimate (23) can be considerably improved.

Then

$$
(26)\quad \begin{aligned}
\log ((2\pi)^{\frac{1}{2}} \sigma T_k) &= (n + 1) \log (n + 1) - (k + \tfrac{1}{2}) \log \frac{k + \frac{1}{2}}{p} \\
&\quad - (n - k + \tfrac{1}{2}) \log \frac{n - k + \frac{1}{2}}{q} + \frac{1}{12(n + 1)} + \frac{1}{24(k + \frac{1}{2})} \\
&\quad + \frac{1}{24(n - k + \frac{1}{2})} - \rho \\
&= -\frac{\sigma^2}{q}\left(1 + \frac{qx}{\sigma}\right)\log\left(1 + \frac{qx}{\sigma}\right) - \frac{\sigma^2}{p}\left(1 - \frac{px}{\sigma}\right)\log\left(1 - \frac{px}{\sigma}\right) \\
&\quad + \frac{pq}{12\sigma^2} + \frac{q}{24\sigma^2\left(1 + \frac{qx}{\sigma}\right)} + \frac{p}{24\sigma^2\left(1 - \frac{px}{\sigma}\right)} - \rho
\end{aligned}
$$

$$
(27)\quad \begin{aligned}
0 \le \rho &\le \frac{7}{6}\left\{\frac{1}{360(n + 1)^3} + \frac{7}{2880}\left[\frac{1}{(k + \frac{1}{2})^3} + \frac{1}{(n - k + \frac{1}{2})^3}\right]\right\} \\
&\le \frac{7}{6}\cdot\frac{1}{360\sigma^6}\left\{p^3q^3 + \frac{7}{8}(p^3 + q^3)\right\},
\end{aligned}
$$

provided only that $k \ge 4$, $(n - k) \ge 4$. Asymptotically ρ is equivalent to the right-hand member without factor $\frac{7}{6}$ (which, by the way, could be replaced by $1 + \frac{1}{40}$). Obviously

$$
(28)\qquad 0 < \rho < \frac{1}{300\sigma^6},
$$

if $k \ge 4$, $n - k \ge 4$. We shall consider later on the case $\sigma \ge 3$, $|x| \le \frac{2}{3}\sigma$; then clearly $k \ge 4$, $n - k \ge 4$, so that the use of (28) will be justified. Expanding (26) into a power series we obtain

THEOREM. *If* $k \ge 4$, $n - k \ge 4$,

$$
(29)\quad \begin{aligned}
T_k = (2\pi)^{-\frac{1}{2}}\sigma^{1}\exp\Bigg\{&-\sum_{2}^{\infty}\frac{p^{\nu-1} - (-q)^{\nu-1}}{\nu(\nu - 1)}\frac{x^{\nu}}{\sigma^{\nu-2}} \\
&+ \frac{1}{24\sigma^2}\sum_{3}^{\infty}\{p^{\nu-1} - (-q)^{\nu-1}\}\left(\frac{x}{\sigma}\right)^{\nu-2} + \frac{1 + 2pq}{24\sigma^2} - \rho\Bigg\}
\end{aligned}
$$

where ρ *satisfies* (28) (*and* (27)); x *and* σ *are defined by* (25) *and* (24), *respectively.*

Each term of the second series will usually be small as compared to the corresponding term of the first series; the second series can therefore, if desired, be absorbed in the error term. If x is small the first term of the first series will be preponderant. However, as x increases, more and more terms will make themselves noticeable; if $x \sim \sigma^{1/2}$, three terms will be essential, and so on.

Formula (29) permits us to approximate $P_{\lambda,\nu}$ by means of integrals. The tangent rule would suggest to compare $P_{\lambda,\nu}$ to

$$
(30)\qquad \Phi\left(x_{\nu} + \frac{1}{2\sigma}\right) - \Phi\left(x_{\mu} - \frac{1}{2\sigma}\right),
$$

and (29) together with lemma (1) permits easily to estimate the *relative error* in the practically most important cases. It is also seen that the limits in (30) are essentially the only limits depending linearly on λ and ν which will render the relative error $O(\sigma^{-1})$ for $x = O(1)$. Instead of elaborating on these simple questions we proceed to the more intricate problem of limits which are quadratic polynomials in λ and ν.

4. For brevity we shall from now on put

$$\frac{p-q}{6} = a. \tag{31}$$

The estimate $|a| \leqq \frac{1}{6}$ will be used constantly. It obviously suffices to consider values of $\lambda < \nu$ which exceed the central value $[(n+1)p]$.

THEOREM. *Suppose that*

$$\sigma > 3 \tag{32}$$

and

$$\lambda \geq (n+1)p \qquad \nu + \tfrac{1}{2} \leq (n+1)p + \tfrac{2}{3}\sigma^2. \tag{33}$$

Then

$$P_{\lambda,\nu} \leq e^{5(1-pq)/36\sigma^2}\{\Phi(\eta_{\nu+1}) - \Phi(\eta_\lambda)\}, \tag{34}$$

if

$$\eta_k = \frac{k-(n+1)p}{\sigma} + \frac{a}{\sigma}\left\{\frac{k-(n+1)p}{\sigma}\right\}^2 + \frac{2a}{\sigma} - \frac{1}{2\sigma^2}, \tag{35}$$

while the inequality in (34) *is reversed if*

$$\eta_k = \frac{k-(n+1)p}{\sigma} + \frac{a}{\sigma}\left\{\frac{k-(n+1)p}{\sigma}\right\}^2 + \frac{2a}{\sigma} + \frac{M}{6\sigma} + \frac{1}{7\sigma}, \tag{36}$$

where

$$M = \frac{x_\nu^3}{\sigma} = \frac{\{\nu + \frac{1}{2} - (n+1)p\}^3}{\sigma^4}. \tag{37}$$

The gap between the limits (35) and (36) is $O(\sigma^{-1})$ if $x_\nu^3 = O(\sigma)$. In S. Bernstein's case (12), $M \leq \sqrt{2}$ and the gap is about $2/(5\sigma)$. It will be seen from the proof that it requires only routine computations to improve the correction term $\left\{\frac{M}{6} + \frac{1}{7}\right\}\sigma^{-1}$ in (36).

PROOF. Put

$$\xi_k = x_k + \frac{a}{\sigma}x_k^2, \tag{38}$$

again suppressing the subscripts wherever convenient. As a consequence of (33), we shall be concerned only with values x_k satisfying

$$\frac{1}{2\sigma} < x < \frac{2}{3}\sigma. \tag{39}$$

Consider first the main series in (29) and write

$$\sum_{2}^{\infty} \frac{p^{\nu-1} - (-q)^{\nu-1}}{\nu(\nu-1)} \frac{x^\nu}{\sigma^{\nu-2}} = \tfrac{1}{2}\xi^2 + A, \tag{40}$$

where

$$A = \left(\frac{p^3+q^3}{12} - \frac{a^2}{2}\right)\frac{x^4}{\sigma^4} + \sum_{5}^{\infty} \frac{p^{\nu-1} - (-q)^{\nu-1}}{\nu(\nu-1)} \frac{x^\nu}{\sigma^{\nu-2}}. \tag{41}$$

We shall require some estimates of A. First consider the case $a > 0$. Then all terms of the series are positive, while the expression within parentheses assumes its minimum $\frac{1}{48}$ for $p = \frac{1}{2}$. By (39) $\xi < \frac{10}{9} x$, whence

$$A > \frac{1}{74}\frac{\xi^4}{\sigma^2} \quad \text{if} \quad a > 0. \tag{42}$$

If $a < 0$ the signs in the series (41) alternate, each negative term being smaller in absolute value than the preceding positive term. Therefore, using (39),

$$A \geq \left\{\frac{p^3+q^3}{12} - \frac{a^2}{2} - \frac{q^4-p^4}{30}\right\}\frac{x^4}{\sigma^2}. \tag{43}$$

The expression within braces is a cubic in p which assumes its minimum for $p = (1 + \sqrt{793})/72 = .405\ldots$ It follows that

$$A \geq \frac{1}{60}\frac{x^4}{\sigma^4} \geq \frac{1}{60}\frac{\xi^4}{\sigma^4} \quad \text{if} \quad a < 0 \tag{44}$$

(half of this estimate would actually suffice for our purposes). On the other hand, it is evident from (41) that the ratio A/x^4 attains its maximum for $p = 1$. Therefore, using (39)

$$A < \frac{2}{15}\frac{x^4}{\sigma^2}. \tag{45}$$

Next we write

$$\frac{1}{24\sigma^2}\sum_{3}^{\infty} \{p^{\nu-1} - (-q)^{\nu-1}\}\left(\frac{x}{\sigma}\right)^{\nu-2} = \frac{a}{4\sigma^3}\xi + B, \tag{46}$$

whence

$$B = \tfrac{1}{2}\left[\frac{p^3+q^3}{12} - \frac{a^2}{2}\right]\frac{x^2}{\sigma^4} + \frac{1}{24\sigma^2}\sum_{5}^{\infty} \{p^{\nu-1} - (-q)^{\nu-1}\}\left(\frac{x}{\sigma}\right)^{\nu-2}. \tag{47}$$

A trivial computation analogous to (43) shows that B > 0. Again, if $a < 0$, the signs in the series (47) alternate and in this case

$$0 \leq B \leq \tfrac{1}{2}\left[\frac{p^3+q^3}{12} - \frac{a^2}{2}\right]\frac{x^2}{\sigma^4} \leq \frac{5}{144}\frac{x^2}{\sigma^4} < \frac{1}{20}\frac{\xi^2}{\sigma^4}. \tag{48}$$

If $a > 0$ we can majorize (47) by a geometric series and obtain

$$0 \leq B \leq \frac{1}{8}\frac{x^2}{\sigma^4} \leq \frac{1}{8}\frac{\xi^2}{\sigma^4}. \tag{49}$$

Now put

$$\Delta\xi_k = \sigma^{-1}\left(1 + \frac{2a}{\sigma}x_k\right). \tag{50}$$

Then

$$\xi_k + \tfrac{1}{2}\Delta\xi_k = \xi_{k+1} - \tfrac{1}{2}\Delta\xi_{k+1} \tag{51}$$

so that the intervals with endpoints $\xi_k \pm \frac{1}{2}\Delta\xi_k$ are non-overlapping and contiguous. Clearly

$$\Delta\xi = \sigma^{-1}\left\{1 + \frac{4a}{\sigma}\xi\right\}^{1/2}. \tag{52}$$

Introducing (40), (46), and (52) into (29) we obtain

$$T_k = (2\pi)^{-1/2}\Delta\xi\cdot\exp\left\{-\frac{\xi^2}{\sigma} - A + B + \frac{a}{4\sigma^3}\xi - \tfrac{1}{2}\log\left(1 + \frac{4a\xi}{\sigma}\right) + \frac{1 + 2pq}{24\sigma^2} - \rho\right\}. \tag{53}$$

To appraise the logarithmic term we write

$$\tfrac{1}{2}\log\left(1 + \frac{4a\xi}{\sigma}\right) = \frac{2a\xi}{\sigma} - C. \tag{54}$$

$C\,\xi^{-2}$ attains its maximum value when $a = -\frac{1}{6}$, and it is readily seen that

$$\begin{aligned} 0 < C < \frac{4a^2\xi^2}{\sigma^2} \quad &\text{if} \quad a > 0 \\ 0 < C < \frac{6a^2\xi^2}{\sigma^2} \quad &\text{if} \quad a < 0. \end{aligned} \tag{55}$$

Finally we put, with a parameter u to be determined,

$$y = \xi + \frac{2a - u}{\sigma}, \qquad \Delta y = \Delta\xi. \tag{56}$$

If one puts

$$u = \frac{1}{2\sigma} - \frac{a}{4\sigma^2} \tag{57}$$

and η_k is defined by (35), then

$$y_k + \tfrac{1}{2}\Delta y_k = \eta_{k+1}, \qquad y_k - \tfrac{1}{2}\Delta y_k = \eta_k. \tag{58}$$

On the other hand, if

$$u = -\frac{M}{6} - \frac{1}{7} - \frac{a}{4\sigma^2} \tag{59}$$

and η_k is defined by (36), the identities (58) hold again. Accordingly, all we have to show is that, with u defined by (57),

$$T_k \leq \{\Phi(y_k + \tfrac{1}{2}\Delta y_k) - \Phi(y_k - \tfrac{1}{2}\Delta y_k)\}e^{5(1-pq)/36\sigma^2} \tag{60}$$

and that the inequality in (60) is reversed if u is defined by (59).

Elementary transformations lead from (53) to

$$T_k = (2\pi)^{-\frac{1}{2}} \Delta y \cdot \exp\left\{-\frac{y^2}{2} + \frac{(\Delta y)^2}{24}(y^2 - 1) + \frac{5(1 - pq)}{36\sigma^2} + E\right\}, \tag{61}$$

where

$$E = \frac{u^2 - 4au}{2\sigma^2} - \left(\frac{u}{\sigma} - \frac{5a}{12\sigma^3}\right)\xi - \frac{1}{24\sigma^2}\left(1 + \frac{4a\xi}{\sigma}\right)y^2 - A + B + C - \rho. \tag{62}$$

Let now u be defined by (57). In view of lemma 1 and (61), the inequality (60) will be proved if we show that

$$E_1 \equiv E + \frac{y^4(\Delta y)^4}{880} \leq 0. \tag{63}$$

Now clearly

$$\frac{y^2}{24\sigma^2}\left(1 + \frac{4a\xi}{\sigma}\right) = \frac{y^2(\Delta y)^2}{24} \geq \frac{y^4(\Delta y)^4}{880}. \tag{64}$$

Moreover, introducing the estimates (28), (32), (42), (44), (48), (49), and (55) into (62) it is seen that for $a > 0$

$$\sigma^2 E_1 < \frac{1}{24\sigma} - \frac{25}{54}\xi + \frac{1}{8}\xi^2 - \frac{1}{74}\xi^4, \tag{65}$$

and for $a < 0$

$$\sigma^2 E_1 < \frac{2}{9\sigma} - \tfrac{1}{2}\xi + \tfrac{1}{5}\xi^2 - \frac{1}{60}\xi^4. \tag{66}$$

The derivatives of the right-hand members in (65) and (66) are both negative for $\xi > 0$. Now we are interested only in values x satisfying (39). For such values $\xi \geq \frac{107}{216\sigma}$. For $\xi = \frac{107}{216\sigma}$ the right-hand members in (65) and (66) are negative, so that $E_1 < 0$ for $x > \frac{1}{2\sigma}$. This proves the first part of our theorem.

The proof that with (59) the inequality in (60) is reversed proceeds on similar lines. We have to show that

$$E_2 \equiv E - \frac{(\Delta y)^4}{285} \geq 0. \tag{67}$$

Suppose that $a < 0$, which is the less favorable case. Then, by (45), (37), and (39),

$$A \leq \frac{2M}{15}\frac{x}{\sigma} \leq \frac{3M}{20}\cdot\frac{\xi}{\sigma}. \tag{68}$$

Similarly

$$\frac{1}{24\sigma^2}\left(1 + \frac{4a\xi}{\sigma}\right)y^2 \leq \frac{1}{24\sigma^2}\left(\xi - \frac{u}{\sigma}\right)^2 \tag{69}$$

Using (62) we have therefore, neglecting the non-negative terms B and C,

$$\begin{aligned} E_2 &\geq \frac{u^2}{2\sigma^2} - \frac{u^2}{24\sigma^4} - \frac{u}{3\sigma^2} - \frac{1}{250\sigma^4} \\ &\quad + \xi\left\{-\frac{u}{\sigma} - \frac{5}{72\sigma^3} - \frac{3M}{20\sigma} + \frac{u}{12\sigma^3}\right\} - \frac{1}{24\sigma^2}\xi^2. \end{aligned} \tag{70}$$

The expression at the right side represents a parabola, and it suffices to show that it assumes positive values at the endpoints of our interval (39). Now

$$\frac{u^2}{2}\left(1 - \frac{1}{12\sigma^2}\right) - \frac{u}{3} \geq -\frac{1}{18}\cdot\frac{1}{1 - \dfrac{1}{12\sigma^2}} > -\frac{6}{107}, \tag{71}$$

and simple arithmetic shows that, with (59) the expression within the braces more than counterbalances the negative terms outside.[6] If $a > 0$ the situation is more favorable and the estimate (59) can then be further improved.

REFERENCES

[1] S. Bernstein, "Retour au problème de l'évaluation de l'approximation de la formule limite de Laplace," [In Russian]. *Izvestia Akademii Nauk SSR. Serija matematičeskaja*, *7* (1943), pp. 3–16.

[2] S. Mirimanoff and R. Dovaz, "Les épreuves répétées et la formule de Laplace," *C. R. Acad. Sci., Paris*, *185* (1927), pp. 827–829.

[3] J. V. Uspensky, *Introduction to Mathematical Probability*. New York, McGraw-Hill, 1937.

[6] A more careful computation shows that it suffices if we put $u = -\frac{M}{6} - \frac{1}{8} - \frac{a}{4\sigma^2}$ instead of $-\frac{M}{6} - \frac{1}{7} - \frac{a}{4\sigma^2}$.