AN APPROXIMATION THEOREM FOR THE POISSON BINOMIAL DISTRIBUTION

LUCIEN LE CAM

1. Introduction. Let x_j; $j = 1, 2, \cdots$ be independent random variables such that $\text{Prob}(X_j = 1) = 1 - \text{Prob}(X_j = 0) = p_j$. Let $Q = \mathscr{L}(\Sigma X_j)$ be the distribution of their sum. This kind of distribution is often referred to as a Poisson binomial distribution. For any finite measure μ on the real line let $\|\mu\|$ be the norm defined by

$$\|\mu\| = \sup_f \left\{\left|\int f d\mu\right|\right\}.$$

the supremum being taken over all measurable functions f such that $|f| \leqq 1$. Let $\lambda = \Sigma p_j$, let $\Sigma p_j^2 = \lambda\varpi$ and let $\alpha = \sup_j p_j$. Finally let P be the Poisson distribution whose expectation is equal to λ.

The purpose of the present paper is to show that there exist absolute constants D_1 and D_2 such that $\|Q - P\| \leqq D_1\alpha$ for all values of the p_j's and $\|Q - P\| \leqq D_2\varpi$ if $4\alpha \leqq 1$.

The constant D_1 is not larger than 9 and the constant D_2 is not larger than 16.

Such a result can be considered a generalization of a theorem of Yu. V. Prohorov [9] according to which such constants exist when all the probabilities p_j are equal.

The norm $\|Q - P\|$ is always larger than the maximum distance $\rho(P, Q)$ between the cumulative distributions. For this distance ρ a very general theorem of A. N. Kolmogorov [6] implies that $\rho(P, Q)$ is at most of order $\alpha^{1/5}$. The improvement obtained here is made possible by the smaller scope of our assumptions.

The method of proof used in the present paper is not quite elementary, since it uses both operator theoretic methods and characteristic functions. The relevant concepts are described in § 2.

A completely elementary approach, described in [4] leads to bounds of the order of $3\alpha^{1/3}$ for the distance ρ. Unfortunately, the elementary method does not seem to be able to provide the more precise result of the present paper.

The developments given here were prompted by discussions with J. H. Hodges, Jr. in connection with the writing of [4].

2. Measures as operators. Let $\{\mathfrak{H}, \mathfrak{A}\}$ be a measurable Abelian group, that is, an Abelian group on which a σ-field $\mathfrak{A}$ has been selected

Received December 9, 1959. The author is a research fellow of the A.P. Sloan Foundation.

in such a way that the map $(x, y) \to x + y$ from $\mathfrak{X} \times \mathfrak{X}$ to $\mathfrak{X}$ is measurable for the σ-fields $\mathfrak{A} \times \mathfrak{A}$ and $\mathfrak{A}$.

Let $\mathscr{B}$ denote the set of bounded measurable numerical functions on $\{\mathfrak{X}, \mathfrak{A}\}$. A finite signed measure μ on $\mathfrak{A}$ defines an operator, also denoted μ, from $\mathscr{B}$ to itself. To the function $f \in \mathscr{B}$ the operator μ makes correspond the element μf whose value at the point x is $(\mu f)(x) = \int f(x + \xi)\mu(d\xi)$. Linear combinations of two operators are defined by the equality

$$(\alpha\mu + \beta\nu)f = \alpha(\mu f) + \beta(\nu f) .$$

The product of two operators will be defined by composition: $(\mu\nu)f = \mu(\nu f)$. In other words,

$$[(\mu\nu)f](x) = \int \mu(dy) \int f(x + \xi + y)\nu(d\xi) .$$

It follows from Fubini's theorem that $\mu\nu = \nu\mu$. The product $\mu\nu$ corresponds to the convolution of the two measures.

For any element f of $\mathscr{B}$ let $|f|$ be the norm $|f| = \sup |f(x)|$. Define the operator norm $\|\mu\|$ by

$$\|\mu\| = \sup \{|\mu f|; \ |f| \leqq 1\} .$$

The norm $\|\mu\|$ is equal to the total mass of μ considered as a measure. It is an immediate consequence of the operator representation of $\mu\nu$ that $\|\mu\nu\| \leqq \|\mu\| \, \|\nu\|$.

Let $\mathfrak{M}$ be the system of operators obtained from all the finite signed measures. What precedes can be summarized by saying that $\mathfrak{M}$ is a normed commutative algebra having for identity the operator I which is the probability measure whose mass is entirely concentrated at the point $x = 0$. It is not difficult to show that $\mathfrak{M}$ is complete for the norm, so that $\mathfrak{M}$ is in fact a real commutative Banach algebra.

Let φ be a complex-valued function of a complex variable z. Suppose that for $|z| < a$, the function φ has a convergent power series expansion. It is then possible to define $\varphi(A)$ for every $A \in \mathfrak{M}$ such that $\|A\| < a$ by simple formal substitution in the power series expansion of φ.

The entity $\varphi(A)$ is then of the form $\varphi(A) = B + iC$ where both B and C belong to $\mathfrak{M}$. Other possible definitions can be found in [3], [2], [8]. If $\hat{\mu}$ is the Fourier transform $\hat{\mu}(t) = \int e^{itx}\mu(dx)$ of the measure μ then $\varphi(\mu)$ is the measure where the Fourier transform is $\varphi(\hat{\mu})$.

In most cases of statistical interest, the space $\mathfrak{X}$ is either the real line, or the additive group of integers, or the circle, or a Euclidean space. In those circumstances, as well as in the case where $\mathfrak{X}$ is an arbitrary Abelian locally compact group, we may replace $\mathscr{B}$ by the space of continuous functions which tend to zero at infinity without affecting any of the above properties.

Let M be an arbitrary finite positive measure on $\mathfrak{X}$. Then $\exp(M) = e^M = I + M + \cdots + (1/k)!\, M^k + \cdots$. It follows that $\exp[M - \|M\| I] = \exp[-\|M\|]\exp(M)$ is always a probability measure.

If a random variable X is equal to the origin of $\mathfrak{X}$ with probability $(1-p)$ the distribution $\mathscr{L}(X)$ can be written $\mathscr{L}(X) = I + p(M - 1)$ where M is a probability measure.

The following theorem, essentially due to Khintchin [5] and Doeblin [1] is concerned with the distribution Q of a sum ΣX_j of independent variables having distributions $G_j = I + p_j(M_j - I)$ where M_j is a probability measure. The product $\prod_j G_j$ is always convergent when $\lambda = \sum_j p_j$ is finite. Conversely finiteness of λ is necessary to the convergence of $\prod_j G_j$ when $\mathfrak{X}$ is the additive group of integers. More generally, suppose that $\mathfrak{X}$ is the real line and that there exists an $\varepsilon > 0$ such that $\lambda_\varepsilon = \Sigma p_j M_j\{[-\varepsilon, \varepsilon]^c\} = \infty$. Then $\prod_j G_j$ cannot be convergent. This follows for instance from a result of Paul Lévy [7] according to which any interval containing the sum ΣX_j with probability $\alpha > 0$ must have a length of the order of $\varepsilon\sqrt{\lambda_\varepsilon}$.

A refinement of Paul Lévy's theorem can be found in [6], Lemma 1. However, the finiteness of λ is not generally necessary to the convergence of $\prod_j G_j$. This is quite obvious if $\mathfrak{X}$ is the circle and G_1 is the Haar measure of the circle, but the condition is not even necessary on the line.

THEOREM 1. *Let X_j; $j = 1, 2, \cdots$ be independent random variables taking their values in the measurable Abelian group $\mathfrak{X}$. Assume that $\mathscr{L}(X_j) = I + p_j(M_j - I)$ where M_j is a probability measure and assume that $\lambda = \Sigma p_j < \infty$. Let $p_j = \lambda c_j$, let $\varpi = \Sigma c_j p_j$ and finally let $M = \Sigma c_j M_j$. Then*

$$\|Q - P\| \leqq 2\lambda\varpi$$

for $P = \exp[\lambda(M - I)]$.

Proof. The proof is essentially the same as the proof of Theorem 1 in [4], given there in terms of random variables. In terms of operators one can proceed as follows.

Let $F_j = \exp p_j(M_j - I)$ and let $R_1 = \prod_{j \geqq 2} G_j$. For $k > 1$ let $R_k = (\prod_{j \leqq k-1} F_j)(\prod_{j \geqq k+1} G_j)$. Then $R_k F_k = R_{k+1} G_{k+1}$ so that

$$\prod_j G_j - \prod_j F_j = \sum_j R_j(G_j - F_j)\ .$$

Since R_j is a probability measure, this implies

$$\| \prod_j G_j - \prod_j F_j \| \geqq \sum_j \| G_j - F_j \| .$$

The difference $F_j - G_j$ can be written

$$F_j - G_j = [e^{-p_j} - (1 - p_j)]I + p_j(e^{-p_j} - 1)M_j + \sum_{k=2}^{\infty} \frac{e^{-p_j}}{k!} p_j^k M_j^k .$$

Hence $\| F_j - G_j \| \leqq 2p_j(1 - e^{-p_j}) \leqq 2p_j^2$.

Noting that $\prod_j F_j = \exp[\lambda(M - I)]$, this proves the desired result.

REMARK. The literature does not seem to contain any reference to the fact that Theorem 1 can be proved as in [4] and coupled with Lindeberg's proof of the normal approximation theorem to obtain a completely elementary proof of the general Central Limit theorem.

3. Sums of indicator variables and binomial distributions. In all the subsequent sections of this paper $\mathfrak{X}$ will be the additive group of integers and $\{X_j; 1, 2, \cdots\}$ will be a family of independent random variables such that $\text{Prob}(X_j = 1) = 1 - \text{Prob}(X_j = 0) = P_j$. The distribution $\mathscr{L}(X_j)$ can then be written either as $I + p_j\Delta$ or $(1 - p_j)I + p_jH$ where Δ is the difference operator $\Delta = H - I$ and H is the probability measure whose mass is entirely concentrated at the point $x = 1$. The Poisson distribution whose expectation is λ can be written $P = \exp(\lambda\Delta)$.

Letting $\lambda c_j = p_j$ and $\varpi = \Sigma c_j p_j$, Theorem 1 implies that if $Q = \mathscr{L}(\Sigma X_j)$ then the following inequality holds.

PROPOSITION 1. $\| Q - \exp(\lambda\Delta) \| \leqq 2\lambda\varpi$.

From now on we shall assume that $\lambda < \infty$ *and that* $\alpha = \sup p_j$ *does not exceed* 1/4.

It may be expected that Q would be approximable by a binomial distribution much more closely than by a Poisson distribution. Letting $\lambda = \nu\varpi$, a binomial distribution with ν trials and probability of success ϖ can be written

$$B = (I + \varpi\Delta)^\nu = (1 - \varpi)^\nu(I + \rho H)^\nu$$

with $\rho = \varpi/1 - \varpi$, at least when ν is an integer. If ν is not an integer the expression

$$B = (1 - \varpi)^\nu\left\{I + \binom{\nu}{1}\rho H + \cdots + \binom{\nu}{k}\rho^k H^k + \cdots\right\}$$

where

$$\binom{\nu}{k} = \frac{1}{k!}\nu(\nu - 1)\cdots(\nu - k + 1) = \frac{\Gamma(\nu + 1)}{k!\,\Gamma(\nu - k + 1)}$$

still possesses a precise meaning as long as $\rho < 1$. However, B is not a probability measure even though $\int 1dB = 1$. Let n be the integer such that $(n-1) < \nu \leqq n$. The coefficients $\binom{\nu}{k}$ of order $k = (n+1), (n+2)\cdots$ are alternately positive and negative.

Let $S = (1-\varpi)^{\nu} \sum_{k=n+1}^{\infty} \binom{\nu}{k} \rho^k H^k$. The norm of S is equal to

$$|| S || = (1-\varpi)^{\nu} \sum_{k=n+1}^{\infty} \left| \binom{\nu}{k} \right| \rho^k = (1-\varpi) \left| \sum_{k=n+1}^{\infty} \binom{\nu}{k} (-\rho)^k \right| .$$

The term inside the absolute value symbol is simply the remainder of the expansion of $(1-\rho)^{\nu}$. By Taylor's formula $|| S ||$ is equal to the absolute value of

$$\frac{1}{n!} \nu(\nu-1)\cdots(\nu-n)(1-\varpi)^{\nu}(1-\rho)^{\nu} \int_0^{\rho/1-\rho} (-1)^n t^n (1+t)^{\nu-n-1} dt .$$

Therefore, since $n-1 < \nu < n$

$$\begin{aligned} || S || &\leqq (1-\varpi)^{\nu}(1-\rho)^{\nu} \int_0^{\rho/1-\rho} t^n (1+t)^{-1} dt \\ &\leqq \frac{1}{n+1} (1-\varpi)^{\nu}(1-\rho)^{\nu} \left(\frac{\rho}{1-\rho} \right)^{n+1} \\ &= \frac{\varpi^{n+1}}{n+1} (1-2\varpi)^{\nu-n-1} \leqq \frac{4}{\nu+1} \varpi^{\nu+1} . \end{aligned}$$

In the cases considered here $\nu = (\Sigma p_j)^2 (\Sigma p_j^2)^{-1}$ is always larger than or equal to unity. In all cases where ν is large and ϖ is small $|| S ||$ will be rather negligible.

Note that $\lambda = \nu\varpi = \int x dB$ and $\nu\varpi(1-\varpi) = \int (x-\lambda)^2 dB$. However, this last quantity may not be treated as a variance, since B possesses negative terms.

In spite of this it will be convenient to bound the remainder term

$$S(m) = (1-\varpi)^{\nu} \sum_{k=m+1}^{\infty} \binom{\nu}{k} \rho^k H^k$$

for large values of m, by Chebyshev's inequality. Assuming $\lambda < m \leqq n$ the terms $(1-\varpi)^{\nu} \binom{\nu}{k} \rho^k$ are smaller than $(1-\varpi)^{\nu-n}$ $(1-\varpi)^n \binom{n}{k} \rho^k$. Therefore

$$|| S(m) || \leqq \frac{4\varpi^{\nu+1}}{\nu+1} + (1-\varpi)^{\nu-n} \sum_{k=m+1}^{n} (1-\varpi)^n \binom{n}{k} \rho^k .$$

Finally, by Chebyshev's inequality applied to the binomial $[1+\varpi\Delta]^n$, one obtains

$$\| S(m) \| \leqq \frac{4\varpi^{\nu+1}}{\nu+1} + (1-\varpi)^{\nu-n} \frac{n\varpi(1-\varpi)}{[m+1-n\varpi]^2} .$$

In particular, if $m \leqq 2n\varpi < m+1$

$$\| S(m) \| \leqq \frac{4\varpi^{\nu+1}}{\nu+1} + \frac{(1-\varpi)^{1-(n-\nu)}}{n\varpi}$$

$$\leqq [4\varpi^{\nu+2}+1]\frac{1}{\lambda} .$$

To show that Q can be approximated by the Poisson distribution P in the cases where λ is too large for Proposition 1 to have any significance, we shall first show that Q can be approximated by B and then show that B is very close to P. The argument will be divided into three parts according to the values of λ and λa^2 for $a^2 = \Sigma c_j(p_j - \varpi)^2$. If λ is large but λa^2 is small, bounds will be obtained through operator theoretic methods. If λ is so large that λa^2 becomes large, bounds will be obtained through computations on characteristic functions.

4. Approximations by binomial distributions. *In this section, it will be assumed throughout that* $\lambda \geqq 3$ *and that* $\alpha \leqq 1/4$.

For the distributions Q and B defined in the preceding section we can write

$$\begin{aligned}\log Q - \log B &= \sum_j \log(I + p_j\Delta) - \nu \log(I + \varpi\Delta)\\ &= \lambda \sum_j c_j\left\{\frac{1}{p_j}\log(I + p_j\Delta) - \frac{1}{\varpi}\log(I+\varpi\Delta)\right\}\\ &= \lambda \sum_{k=2}^{\infty} \frac{(-1)^k}{k+1}\beta_k \Delta^{k+1} = \lambda\Delta M ,\end{aligned}$$

with

$$M = \sum_{k=2}^{\infty} \frac{(-1)^k}{k+1}\beta_k\Delta^k$$

and $\beta_k = \sum_j c_j p_j^k - \varpi^k \geqq 0$.

Since $(-1)^k\Delta^k = \sum_{s=0}^{\infty}\binom{k}{s}(-1)^s H^s$, the measure M assigns negative masses to the odd positive integers and positive masses to the even nonnegative integers.

The norm of M is precisely equal to

$$\| M \| = \sum_{k=2}^{\infty}\frac{\beta_k 2^k}{k+1} = -\sum_j c_j\left\{\frac{1}{2p_j}\log(1-2p_j) - \frac{1}{2\varpi}\log(1-2\varpi)\right\} .$$

Letting $u = 2\varpi$ and $v_j = 2(p_j - \varpi)$ this can also be written

$$\| M \| = \int_0^1 \Sigma c_j \left\{ \left[\frac{1}{1 - t(u + v_j)} - \frac{1}{1 - tu} \right] \right\} dt .$$

Since $\Sigma c_j v_j = 0$ and $\Sigma c_j v_j^2 = 4a^2$ while

$$[1 - t(u + v_j)]^{-1} - (1 - tu)^{-1} = (1 - tu)^2 \{1 + (tv_j)[1 - t(u + v_j)]^{-1}\} tv_j$$

one can write

$$\begin{aligned} \| M \| &= \int_0^1 \left\{ \Sigma c_j \frac{v_j^2}{[1 - t(u + v_j)]} \right\} \frac{t^2}{(1 - tu)^2} dt \\ &\leqq \frac{4a^2}{1 - 2\alpha} \int_0^1 \frac{t^2}{(1 - tu)^2} dt \\ &\leqq \frac{4a^2}{3(1 - 2\alpha)} \left\{ 1 + \frac{3\varpi}{(1 - 2\varpi)^2} \right\} . \end{aligned}$$

Hence $\| M \| = ha^2$ with

$$h \leqq \frac{4}{3(1-2\alpha)} \left\{ 1 + \frac{3\varpi}{(1-2\varpi)^2} \right\} .$$

One can also write $M = \Delta M_1 = \Delta^2 M_2$ with $\| M \| = 2 \| M_1 \| = 4 \| M_2 \|$.

It results from these equalities that

$$Q = B \exp [\lambda \Delta M] .$$

For every measure μ, Taylor's formula gives

$$e^{\mu} = I + \mu \int_0^1 e^{\xi \mu} d\xi .$$

Hence

$$\begin{aligned} Q - B &= \lambda \Delta B M \int_0^1 e^{\xi \lambda \Delta M} d\xi \\ &= \lambda \Delta^2 B M_1 \int_0^1 e^{\xi \lambda \Delta M} d\xi . \end{aligned}$$

Finally

$$\| Q - B \| \leqq \lambda \| M \| \; \| \Delta B \| e^{2h\lambda a^2}$$

and

$$\| Q - B \| \leqq \frac{1}{2} \lambda \| M \| \; \| \Delta^2 B \| e^{2h\lambda a^2} .$$

One can also note that there exist probability measures F and G such that if $\varepsilon = \| M \|$ then

$$Q \exp [\lambda \varepsilon (F - I)] = B \exp [\lambda \varepsilon (G - I)] .$$

According to the foregoing expressions, to obtain bounds on $\| Q - B \|$ it will be sufficient to evaluate $\| \Delta B \|$ and $\| \Delta^2 B \|$.

Let $f(x) = \binom{\nu}{x} \varpi^x (1 - \varpi)^{\nu - x}$ and consider only values x such that $x \leqq n - 1$. In this range f achieves its maximum at a value x such that $\lambda + \varpi - 1 < x \leqq \lambda + \varpi$. It follows that $(\Delta f)(x')$ is positive for $x' \leqq x$ and negative for $x' > x$. Finally

$$\| \Delta B \| \leqq 2f(x) + \| S \| .$$

Let $x = \nu\xi$. An application of Stirling's formula leads to the inequality

$$\log f(x) \leqq -\frac{1}{2} \log [2\pi\nu\xi(1 - \xi)]$$

$$f(x) \leqq \frac{\theta}{\sqrt{\lambda}}$$

with

$$\theta = \frac{1}{\sqrt{2\pi}} \left[\frac{\xi}{\varpi} (1 - \xi) \right]^{-1/2} .$$

Since $\varpi(1 + 1/\nu) - 1/\nu < \xi \leqq \varpi(1 + 1/\nu)$ the quantity $\xi/\varpi(1 - \xi)$ is larger than

$$\left[1 + \frac{1}{\nu} - \frac{1}{\lambda}\right]\left[1 - \varpi\left(1 + \frac{1}{\nu}\right)\right] = \left[1 - \frac{1 - \varpi}{\lambda}\right]\left[1 - \varpi\left(1 + \frac{1}{\nu}\right)\right]$$

$$\geqq \left(1 - \frac{1}{\lambda}\right)\left[1 - \varpi\left(1 + \frac{\varpi}{\lambda}\right)\right] \geqq \frac{2}{3}\left(1 - \frac{13}{48}\right) .$$

Consequently,

$$\theta \leqq \left(\frac{72}{70\pi}\right)^{1/2}$$

and

$$\| \Delta B \| \leqq \frac{2\theta}{\sqrt{\lambda}} + \frac{4\varpi^{\nu+2}}{\lambda} .$$

Thus, we have shown the validity of the following proposition.

PROPOSITION 2. *Let $\lambda \geqq 3$ and $\alpha \leqq 1/4$, then*

$$\| Q - B \| \leqq 2ha^2\sqrt{\lambda} \exp(2h\lambda a^2)\left\{\theta + \frac{4\varpi^{\nu+2}}{\sqrt{\lambda}}\right\}$$

with

$$h \leqq \frac{4}{3}\left(\frac{1}{1-2\alpha}\right)\left[1+\frac{3\varpi}{(1-2\varpi)^2}\right] \leqq \frac{32}{3}$$

and

$$\theta \leqq \left(\frac{36}{35\pi}\right)^{1/2} \leqq \frac{1}{\sqrt{3}} .$$

A computation using the fact that $\Delta M = \Delta^2 M_1$ and the bounds for $\| \Delta^2 B \|$ can be carried out as follows.

Let $u = x + 1 - \nu\varpi$ and let $f(u)$ be the probability of $x = \nu\varpi + u - 1$ for the binomial B. Let $\delta^{-1} = \nu\varpi(1-\varpi)$ and let $\beta = \varpi\delta$ and $\gamma = (1-\varpi)\delta$. Then

$$\frac{f(u+1)}{f(u)} = \frac{1-\beta(u-1)}{1+\gamma u} .$$

The second differences of the function f for $x \leqq n$ are equal to some positive quantity multiplied by

$$g(u) = u^2 - (2\varpi - 1)u - (\nu + 2)\varpi(1-\varpi) .$$

Let r_1 and r_2, $r_1 < r_2$ be the roots of this polynomial. The second differences $(\Delta^2 f)(u)$ are negative for $u \in (r_1, r_2)$ and positive otherwise. Letting $\varphi(u) = (\Delta f)(u)$ it follows that

$$\| \Delta^2 B \| \leqq \varphi(u_1) + |\varphi(u_2) - \varphi(u_1 - 1)| + \varphi(n - \lambda + 1) - \varphi(u_2 - 1) + \frac{8}{\nu+1}\varpi^{\nu+1} .$$

The values u_i are determined by the condition that the corresponding x values, say x_1 and x_2, are respectively the largest integer not exceeding $r_1 + \lambda$ and the smallest integer as large as $r_2 + \lambda$. The roots r_1 and r_2 are given by the expression

$$r = (\varpi - 1/2) \pm \left[(\nu+1)\omega(1-\omega) + \frac{1}{4}\right]^{1/2} .$$

If $\lambda \geqq 3$ the value u_1 is negative while $u_2 - 1$ is positive.

In this case

$$\begin{aligned} \varphi(u_1) &\leqq f(u_1+1)\left[1 - \frac{1+\gamma u_1}{1-\beta(u_1-1)}\right] \\ &\leqq f(u_1+1)\delta[|u_1| + \varpi] \\ &\leqq \frac{\theta}{\sqrt{\lambda}} \| u_1 | + \varpi | \frac{1}{\lambda(1-\varpi)} . \end{aligned}$$

Similarly,

$$| \varphi(u_2) | \leqq f(u_2)\left[\frac{1 - \beta(u_2 - 1)}{1 + \gamma u_2} - 1\right]$$

$$\leqq \frac{\theta}{\lambda\sqrt{\lambda}} [| u_2 | + \varpi]\left(\frac{1}{1 - \varpi}\right).$$

Note that $| u_1 - 1 | \leqq 1 + 1/2 + \sqrt{\nu\varpi(1 - \varpi)} + 1/6 \leqq 5/3 + \sqrt{\lambda(1 - \varpi)}$. Hence

$$\varphi(u_1 - 1) \leqq \frac{\theta}{\lambda}\left\{\frac{1}{(1 - \varpi)\sqrt{\lambda}}\left[\frac{5}{\lambda} + \sqrt{\lambda(1 - \varpi)}\right] + \varpi\right\}$$

$$\leqq \frac{9\theta}{4\lambda}.$$

The other terms can be bounded in a similar manner giving

$$|| \Delta^2 B || \leqq 9\frac{\theta}{\lambda} + \frac{16}{\lambda}\varpi^{\nu+2} \leqq \frac{5.4}{\lambda}.$$

Finally the following result holds.

PROPOSITION 3. *If* $\lambda \geqq 3$ *and* $\alpha \leqq 1/4$ *then*

$$|| Q - B || \leqq (2.7)h \exp [2h\lambda a^2]a^2$$

with $h \leqq 32/3$.

It is possible to obtain bounds on the third difference $|| \Delta^3 B ||$ by similar procedures. The algebra becomes somewhat more cumbersome. Nevertheless, it is not difficult to see that bounds of the type

$$|| Q - B || \leqq C\frac{\log \lambda}{\sqrt{\lambda}} \exp [2\lambda a^2 h]a^2$$

can be obtained in this manner.

The bounds given in Propositions 2 and 3 will be of value if λa^2 is small. When λ is so large that λa^2 is large, better inequalities than the preceding may be obtained through the use of Fourier transforms.

Let $\hat{\mu}$ be the Fourier transform of the measure μ. For instance $\hat{Q}(t) = \int e^{itx} Q(dx)$. Note the following inequalities.

First

$$| 1 + p(e^{it} - 1) |^2 = 1 - 2p(1 - p)(1 - \cos t).$$

Hence, if $| t | \geqq \pi/2$

$$|1+p(e^{it}-1)|^2 \leqq 1-2p(1-p)\frac{2}{\pi}|t|\,.$$

If $|t| \leqq \pi/2$ then

$$1-\cos t = \frac{t^2}{2}\left[1-\frac{t^4}{12}\cos\xi t\right]$$

with $|\xi| \leqq 1$.

Consequently, for $|t| \leqq \pi/2$

$$|1+p(e^{it}-1)|^2 \leqq 1-2p(1-p)\frac{t^2}{2}\left(\frac{48-\pi^2}{48}\right)$$

and for $|t| \leqq \pi/4$

$$|1+p(e^{it}-1)|^2 \leqq 1-2p(1-p)\frac{t^2}{2}\left(\frac{192-\pi^2}{192}\right)\,.$$

It follows that $|\hat{B}(t)| \leqq 1$ and

(1) For $\pi/2 \leqq |t| \leqq \pi$
$\max\{|\hat{B}(t)|, |\hat{Q}(t)|\} \leqq \exp -\{\lambda(1-\varpi)(2/\pi)|t|\}$.

(2) For $\pi/4 \leqq |t| \leqq \pi/2$
$\max\{|\hat{B}(t)|, |\hat{Q}(t)|\} \leqq \exp[-(b^2/2)\lambda t^2]$
with $b^2 = (1-\varpi)-\pi^2/48$.

(3) For $|t| \leqq \pi/4$
$\max\{|\hat{B}(t)|, |\hat{Q}(t)|\} \leqq \exp[-(\beta^2/2)\lambda t^2]$
with $\beta^2 = (1-\varpi)(1-\pi^2/192)$.

In addition, for $|t| \leqq \pi/4$ and for $z = e^{it}-1$ one can write

$$\begin{aligned}\log\hat{Q}-\log\hat{B} &= \lambda\Sigma c_j\left[\frac{1}{p_j}\log(1+p_jz)-\frac{1}{\varpi}\log(1+\varpi z)\right]\\ &= -\lambda z^3\int_0^1\frac{\xi^2}{(1+\xi\varpi z)^2}\left[\sum\nolimits_j\frac{c_j\delta_j^2}{1+\xi p_jz}\right]d\xi\end{aligned}$$

with $c_j = p_j/\lambda$ and $\delta_j = p_j-\varpi$.

This gives

$$|\log\hat{Q}-\log\hat{B}| \leqq \frac{1}{3}\lambda\alpha^2|z|^3\psi(z)$$

where

$$\psi(z) = \sup_{|t|\leqq\pi/4}\sup_j\left|\int_0^1\frac{3\xi^2}{[1+\xi\omega z]^2}\,\frac{1}{(1+\xi p_jz)}\right|d\xi\,.$$

Since
$$|1+\xi\varpi z|^2 = |(1-\xi\varpi)+\xi\varpi e^{it}|^2$$
$$= 1 - 2\xi\varpi(1-\xi\varpi)(1-\cos t)$$

one has

$$|1+\xi\varpi z|^2 \geqq 1-(2-\sqrt{2})\frac{\pi}{4}\varpi .$$

Finally

$$\psi(z) \leqq \frac{1}{\sqrt{1-\frac{\alpha}{2}}}\frac{1}{\left(1-\frac{\varpi}{2}\right)} .$$

Hence

$$|\log \hat{Q} - \log \hat{B}| \leqq K^2\lambda a^2 |t|^3$$

with

$$K^2 \leqq \frac{1}{3}\left(1-\frac{\varpi}{2}\right)^{-1}\left(1-\frac{\alpha}{2}\right)^{-1/2} .$$

It follows that, for $|t| \leqq \pi/4$ one can write

$$|\hat{Q}(t)-\hat{B}(t)| \leqq |\hat{B}(t)|\lambda a^2K^2|t|^3 \exp[\lambda a K^2|t|^3]$$
$$\leqq \lambda a^2K^2|t|^3 \exp\left[-\frac{1}{2}\lambda\gamma^2t^2\right]$$

with $\gamma^2 = \beta^2 - a^2K^2\pi/4 \geqq 0$.

Let $V=(Q-B)$. The individual terms of V are given by the formula

$$V(k) = \frac{1}{2\pi}\int_{-\pi}^{+\pi} e^{-ikt}\hat{V}(t)dt .$$

Applying to this formula the above inequalities one obtains:

$$2\pi|V(k)| \leqq 2\lambda a^2K^2\int_0^\infty t^3 \sup\left[-\frac{1}{2}\lambda\gamma^2t^2\right]dt$$
$$+4\int_{\pi/4}^\infty \exp\left[-\lambda b^2\frac{t^2}{2}\right]dt$$
$$+4\int_{\pi/2}^\infty \exp\left[-\lambda(1-\varpi)\frac{2}{\pi}\right]dt .$$

Therefore,

$$2\pi|V(k)| \leqq \frac{4K^2a^2}{\lambda\gamma^4} + \frac{16}{\lambda\pi b^2}\exp\left[-\frac{\lambda b^2\pi^2}{32}\right] + \frac{2\pi}{\lambda(1-\varpi)}\exp[-(1-\varpi)\lambda] .$$

Noting that $xe^{-x} \leqq e^{-1}$ for $x \geqq 0$, this gives

$$2\pi\lambda \,|\, V(k) \,| \leqq \frac{4K^2a^2}{\gamma^4} + \left\{\frac{16\times 32}{\pi^3 cb^4} + \frac{2\pi}{(1-\varpi)^2 e}\right\}\frac{1}{\lambda}\,.$$

Let m be an integer such that $m \leqq 2n\varpi < m+1$ with $n-1 < \nu \leqq n$. The sum of the first m terms of $|V(k)|$ is inferior to

$$\frac{1}{\pi}\left\{\frac{4K^2a^2}{\gamma^2} + \frac{16\times 32}{\lambda\pi^3 eb^4} + \frac{2\pi}{\lambda(1-\varpi)^2 e}\right\}\left(1+\frac{1}{\nu}\right).$$

From this and Chebyshev's inequality it follows that

$$\|\, Q - B \,\| \leqq \frac{1}{\pi}\left(1+\frac{1}{\nu}\right)\left\{\frac{4K^2a^2}{\gamma^4} + \frac{16\times 32}{\lambda\pi^3 eb^4} + \frac{2\pi}{\lambda(1-\varpi)^2 e}\right\}$$
$$+ \frac{(1-\varpi)}{\lambda} + \frac{1}{\lambda}[1 + 4\varpi^{\nu+2}]\,.$$

As a summary, one can state the following.

PROPOSITION 4. *Assume* $\lambda \geqq 3$ *and* $\alpha \leqq 1/4$. *Then, there exist constants* C_1 *and* C_2 *such that*

$$\|\, Q - B \,\| \leqq C_1 a^2 + C_2\lambda^{-1}\,.$$

5. Approximation of the binomial by a Poisson distribution. A theorem of Yu. V. Prohorov [9] states that the binomial $B = [I + \varpi\Delta]^2$ and the Poisson $P = \exp(\lambda\Delta)$ differ little. Explicitly, there is a constant C_0 such that $\|P - B\| \leqq C_0\varpi$.

Prohorov's result is proved in [9] only for integer values of ν. For this reason we shall give here a complete proof which happens to be somewhat simpler than Prohorov's original argument. This proof leads to an evaluation of the constant C_0 which may not be the best available but will serve our purposes.

Let $R(x)$ be the ratio of the binomial probability $B[\{x\}]$ to the Poisson probability $P[\{x\}]$

$$R(x) = \nu(\nu-1)\cdots(\nu-x+1)\varpi^x(1-\varpi)^{\nu-x}e^{\lambda}\lambda^{-2}\,.$$

Let us restrict ourselves to the interval $0 \leqq x \leqq n$. Since

$$\frac{R(x+1)}{R(x)} = \frac{\nu - x}{\nu(1-\varpi)}$$

the ratio R achieves in this interval a maximum at the point x such that $x - 1 \leqq \lambda < x$.

For this particular value of x, Stirling's formula leads to the inequality

$$\log R(x) \leqq -\frac{1}{2}\log(1-\xi)$$

with

$$\varpi < \xi \leqq \varpi\left(1+\frac{1}{\lambda}\right).$$

Finally for $\lambda \geqq 3$ and $4\varpi \leqq 1$,

$$\begin{aligned} R(x) &\leqq \frac{1}{\sqrt{1-\xi}} \leqq 1 + \frac{\xi}{2\sqrt{1-\xi}} \\ &\leqq 1 + \frac{1}{2}\varpi\left(1+\frac{1}{\lambda}\right)\left[1-\varpi\left(1+\frac{1}{\lambda}\right)\right]^{-1/2} \\ &\leqq 1 + \left(\frac{2}{3}\right)^{1/2}\varpi . \end{aligned}$$

Let f be a nonnegative function such that $0 \leqq f \leqq 1$. The above inequalities imply that

$$\begin{aligned} \int f dB &\leqq \frac{4\varpi^{\nu+1}}{\nu+1} + \int_{x\leqq n} R(x)f(x)P(dx) \\ &\leqq \frac{4\varpi^{\nu+1}}{\nu+1} + \left(\frac{2}{3}\right)^{1/2}\varpi\int f(x)P(dx) + \int f(x)P(dx) \\ &\leqq \int f(x)P(dx) + \varpi\left\{\left(\frac{2}{3}\right)^{1/2}\frac{4\varpi^{\nu}}{\nu+1}\right\}. \end{aligned}$$

Similarly,

$$\int(1-f)dB = 1 - \int f dB \leqq \int (1-f)dP + \varpi\left[\left(\frac{2}{3}\right)^{1/2} + \frac{4\varpi^{\nu}}{\nu+1}\right].$$

Consequently:

PROPOSITION 5. *If $\lambda \geqq 3$ and $4\varpi \leqq 1$, then*

$$\begin{aligned} \| B - P \| &\leqq 2\varpi\left[\left(\frac{2}{3}\right)^{1/2} + \frac{4\varpi^{\nu}}{\nu+1}\right] \\ &\leqq [1.64]\varpi . \end{aligned}$$

Collecting the inequalities established in the preceding sections one obtains the following statement.

THEOREM 2. *Let $\{X_j; j = 1, 2, \cdots\}$ be a family of independent random variables. Assume that $\mathscr{L}(X_j) = I + p_j\Delta$ and that $\lambda = \Sigma p_i$ is* *finite. Let* $p_j = \lambda c_j$ *and* $\varpi = \Sigma c_j p_j$ *and* $\alpha = \sup_j p_j$. *Denote by* Q *the distribution* $Q = \mathscr{L}(\Sigma X_j)$ *and* P *the Poisson distribution* $P = \exp(\lambda \Delta)$.

There exist constants D_1 *and* D_2 *such that*

(1) *For all values of the* p_j *one has*

$$\| P - Q \| \leqq 2\lambda\varpi$$

and

$$\| P - Q \| \leqq D_1\alpha \ .$$

(2) *If* $4\alpha \leqq 1$ *then*

$$\| P - Q \| \leqq D_2\varpi \ .$$

The constant D_1 *is inferior to* 9 *and the constant* D_2 *is inferior to* 16.

Proof. The proof of Theorem 2 consists essentially of an evaluation of the constants involved in the bounds given by Propositions 2, 3 and 4. To these propositions one must add the following remarks.

The quantity $a^2 = \Sigma c_j(p_j - \varpi)^2$ can be written

$$a^2 = \Sigma c_j\left(p_j - \frac{\alpha}{2}\right)^2 - \left(\frac{\alpha}{2} - \varpi\right)^2 .$$

Hence

$$a^2 \leqq \alpha\varpi\left(1 - \frac{\varpi}{2}\right) \leqq \left(\frac{\alpha}{2}\right)^2 .$$

In particular $a^2 \leqq \alpha\varpi$ and $a \leqq \alpha/2 \leqq 1/8$ for $\alpha \leqq 1/4$. The bound $\| Q - P \| \leqq D_1\alpha$ is operative only when $D\alpha \leqq 2$. It is therefore sufficient to prove that $\| Q - P \| \leqq D_1\alpha$ for $\alpha \leqq 2D_1^{-1}$ and $2\lambda \geqq D_1$. A constant D_1 can then be obtained through application of Proposition 2 for $\lambda a^2 \leqq y^2$ and Proposition 4 for $\lambda a^2 \geqq y^2$, the quantity y^2 being adjusted to give the best value available.

Similarly, the second inequality can be proved by use of Propositions 3 and 4, assuming $2\lambda \geqq 16$ and $\varpi \leqq 1/8$.

Note that the constants 9 and 16 are certainly much too large. For very small values of α or ϖ one can obtain much better values of D_1 and D_2.

Statement 2 of Theorem 2 implies that the approximation by a Poisson distribution will be good even though a few of the probabilities P_j may be close to the bound $\alpha \leqq 1/4$. This will happen provided only that these large values contribute relatively little to the value of λ, the bulk of λ being due to very small values of the p_j.

6. Concluding remarks.

REMARK 1. It would be highly desirable for the applications to lower the values of the coefficients D_1 and D_2 to a more reasonable level. When α is fixed, this can be achieved for D_2 by restricting the range of values of ϖ to which the inequalities apply. For instance, taking $4\alpha = 1$ but $\varpi = 10^{-2}$, the coefficient D_2 can be taken approximately equal to 8. Such a value being still too large one may inquire whether there is a lower bound to the acceptable values of D_2.

In this connection the following remarks may be of interest. When λ becomes very large the distance $(1/\varpi) \| Q - B \|$ becomes rapidly negligible. This can be seen for instance by using the inequalities which led to Proposition 4 and the bounds in $\alpha^2 \log \lambda / \sqrt{\lambda}$ obtained through the use of third differences.

The main contribution to $(1/\varpi) \| - P \|$ is then attributable to the difference between the binomial B and the Poisson measure P.

Prohorov's theorem implies that $(1/\varpi) \| B - P \|$ cannot be much smaller than (.483). Therefore, one cannot expect to obtain a result of the type $\| Q - P \| \leqq D_2 \varpi$ where D_2 would be substantially smaller than 1/2.

REMARK 2. The result of Theorem 1 cannot be materially improved unless one is willing to restrict further the measures M_j or the group $\mathfrak{X}$.

A slight modification of the proof given here leads to the inequality

$$\| Q - P \| \leqq 2\Big[1 - \prod_j (1 - \beta_j)\Big],$$

where β_j is taken equal to $p_j(1 - e^{-p_j})$. The bound so obtained is actually reached for certain choices of the measures M_j. An example of this can be constructed when $\mathfrak{X}$ is the real line. It is sufficient to take M_j to be the probability measure giving all its mass to a point x_j and select the values $\{x_j; j = 1, 2, \cdots\}$ to be rationally independent. For any fixed $\varepsilon > 0$ one may find values $p_j < \varepsilon$ such that $2[1 - \prod(1 - \beta_j)] > 2 - \varepsilon$ and such that $\lambda = \sum_j p_j$ be finite.

REFERENCES

1. W. Doeblin, *Sur les sommes d'un grand nombre de variables aléatoires indépendantes*, Bull. des Sciences Mathématiques, **53**, Paris (1939), 23–32.

2. Nelson Dunford and Jacob T. Schwartz, *Linear operators, Part I. General theory*, Interscience Publishers, New York, 1958.

3. Einar Hille and Ralph S. Phillips, *Functional analysis and semi groups*, Amer. Math. Soc. Coll. Publ. 31, Providence, R. I. 1957.

4. J. L. Hodges, Jr. and Lucien Le Cam, *The Poisson approximation to the Poisson*

binomial distribution, to be published in Ann. Math. Stat.
5. A. Khintchine, *Asymptotische Gesetze der Wahrscheinlichkeitsrechnung*, Ergebnisse der Mathematik und ihrer grenzgebiete, Julius Springer, Berlin, 1933.
6. A. N. Kolmogorov, *Deux théorèmes asymptotiques pour les sommes de variables aléatoires* (Russian, French summary), Teoriia Veroiatnosteii, **1** (4), Moscow (1956), 426–436.
7. Paul Lévy, *Théorie de l'addition des variables aléatoires*, Gauthier-Villars, Paris, 1937.
8. M. A. Naimark, *Normed rings*, Moscow, 1956.
9. Yu. V. Prohorov, *Asymptotic behavior of the binomial distribution* (Russian), Uspekhii Matematicheskiikh Nauk, **8** (3), Moscow (1953), 135–142.