The Annals of Probability
1975, Vol. 3, No. 4, 726–731

A GENERAL POISSON APPROXIMATION THEOREM[1]

BY R. J. SERFLING

The Florida State University

A sum of nonnegative integer-valued random variables may be treated as a Poisson variable if the summands have sufficiently high probabilities of taking 0 value and sufficiently weak mutual dependence. This paper presents simple exact upper bounds for the error of such an approximation. An application is made to obtain a new extension for dependent events of the divergent part of the Borel–Cantelli lemma. The bounds are illustrated for the case of Markov-dependent Bernoulli trials. The method of the paper is to reduce the general problem to the special case of independent 0–1 summands and then make use of known bounds for this special case.

1. Introduction and theorem. A natural measure of disparity between the distributions of two nonnegative integer-valued random variables X and Y is

$$d(X, Y) = \sup_A |P(X \in A) - P(Y \in A)| ,$$

where the sup is taken over all steps A of nonnegative integers. The alternative expression

$$d(X, Y) = \tfrac{1}{2} \sum_{k=0}^{\infty} |P(X = k) - P(Y = k)|$$

is easily verified. Another conventional disparity measure is

$$d_0(X, Y) = \sup_{k \geqq 0} |P(X \leqq k) - P(Y \leqq k)| .$$

Clearly $d_0(X, Y) \leqq d(X, Y)$.

The purpose of this paper is to provide simple exact upper bounds for $d_0(X, Y)$ and for $d(X, Y)$, for the case that X is a sum $\sum_1^n X_i$ of (possibly dependent) nonnegative integer-valued random variables and Y is a Poisson variable suitably chosen to approximate X in distribution.

Previous such approximation theorems have been restricted to independent summands, or to 0–1 valued summands or both. As a frame of reference, we state two such results which also will serve as lemmas in the present development. The first is due to Le Cam (1960), the second due to Franken (1964).

LEMMA 1. *Let $X_1, \cdots, X_n$ be independent Bernoulli variables with respective success probabilities $p_1, \cdots, p_n$ and let Y be Poisson with mean $\sum_1^n p_i$. Then*

$$d(\textstyle\sum_1^n X_i, Y) \leqq \sum_1^n p_i^2 . \tag{1.1}$$

LEMMA 2. *Let $X_1, \cdots, X_n$ be independent nonnegative integer-valued random*

Received July 29, 1974; revised December 2, 1974.

[1] Research supported by the Army, Navy and Air Force under Office of Naval Research Contract No. N00014-67-A-0235-0006.

AMS 1970 *subject classification*. Primary 60F05.

Key words and phrases. Poisson approximation, dependent summands, Borel–Cantelli lemma, Markov-dependent Bernoulli trials.

variables and let Y be Poisson with mean $\sum_1^n E(X_i)$. *Then*

$$d_0(\sum_1^n X_i, Y) \leqq \frac{2}{\pi} \sum_1^n [E^2(X_i) + EX_i(X_i - 1)] . \tag{1.2}$$

In each of the cited papers may be found additional results giving alternative bounds or dealing with alternative disparity measures. A result for *possibly dependent* 0–1 valued summands is provided by Freedman (1974), but in the case of independence it is cruder than (1.1).

The following theorem improves the bound stated in Lemma 2 and extends both lemmas to the general case of possibly dependent nonnegative integer-valued summands.

THEOREM 1. *Let* $X_1, \cdots, X_n$ *be (possibly depndent) nonnegative integer-valued random variables and put* $p_1 = P(X_1 = 1)$ *and*

$$p_i = P(X_i = 1 \mid \mathscr{F}_{i-1}) , \qquad 2 \leqq i \leqq n ,$$

where $\mathscr{F}_i$ *denotes the σ-field generated by* $X_1, \cdots, X_i$. *Let Y be Poisson with mean* $\sum_1^n E(p_i)$. *Then*

$$d(\sum_1^n X_i, Y) \leqq \sum_1^n E^2(p_i) + \sum_1^n E|p_i - E(p_i)| + \sum_1^n P(X_i \geqq 2) \tag{1.3}$$

and

$$d_0(\sum_1^n X_i, Y) \leqq \frac{2}{\pi} \sum_1^n E^2(p_i) + \sum_1^n E|p_i - E(p_i)| + \sum_1^n P(X_i \geqq 2) . \tag{1.4}$$

The result is proved by first reducing the general problem to the special case of independent 0–1 summands and then utilizing the implications of Lemmas 1 and 2 for this special case. The appropriate reduction lemma, which is of independent interest, and its application to obtain Theorem 1 are presented in Section 2.

For the case of independent 0–1 summands the bounds (1.3) and (1.4) reduce to (1.1) and (1.2) respectively. The bound (1.4) strictly improves (1.2) in the case of independent summands with $P(X_i > 1) > 0$ for at least one X_i. Note that the approximating Poisson variables in Lemma 2 and Theorem 1 differ except in the case of 0–1 summands.

By a further application of the reduction lemma, a new extension of the Borel–Cantelli lemma (divergent part) for dependent events is derived in Section 3. The result contains the extension given by Iosifescu and Theodorescu (1969) for ϕ-mixing dependent events.

In Section 4 the bounds of Theorem 1 are exemplified for the case of Markov-dependent Bernoulli trials. Section 5 provides a simple proof of Lemma 1.

2. Reduction lemma. Note that

$$d(X, Y) \leqq P(X \neq Y) \tag{2.1}$$

when X and Y are defined on a common probability space. Also, for arbitrary nonnegative integer-valued X, Y and Z,

$$d(X, Y) \leqq d(X, Z) + d(Y, Z) \tag{2.2}$$

and

$$d_0(X, Y) \leqq d_0(X, Z) + d_0(Y, Z) \leqq d(X, Z) + d_0(Y, Z)\,. \tag{2.3}$$

We now establish

LEMMA 3. *Let $X_1, \cdots, X_n$ be (possibly dependent) nonnegative integer-valued random variables and put $p_1 = P(X_1 = 1)$ and*

$$p_i = P(X_i = 1 \mid \mathscr{F}_{i-1})\,, \qquad 2 \leqq i \leqq n\,,$$

where $\mathscr{F}_i$ denotes the σ-field generated by $X_1, \cdots, X_i$. Write $X_i' = I(X_i = 1)$, $1 \leqq i \leqq n$, and put $p_1' = P(X_1' = 1) = p_1$ and

$$p_i' = P(X_i' = 1 \mid \mathscr{F}_{i-1}')\,, \qquad 2 \leqq i \leqq n\,,$$

where $\mathscr{F}_i'$ denotes the σ-field generated by $X_1', \cdots, X_i'$. Finally, let $X_1^*, \cdots, X_n^*$ be independent Bernoulli variables with respective success probabilities $p_1^*, \cdots, p_n^*$. Then

$$d(\textstyle\sum_1^n X_i, \sum_1^n X_i') \leqq \sum_1^n P(X_i \geqq 2) \tag{2.4a}$$

and

$$d(\textstyle\sum_1^n X_i', \sum_1^n X_i^*) \leqq \sum_1^n E|p_i' - p_i^*| \leqq \sum_1^n E|p_i - p_i^*|\,. \tag{2.4b}$$

Hence also

$$d(\textstyle\sum_1^n X_i, \sum_1^n X_i^*) \leqq \sum_1^n E|p_i - p_i^*| + \sum_1^n P(X_i \geqq 2)\,. \tag{2.4c}$$

PROOF. (2.4c) follows from (2.2), (2.4a) and (2.4b). Using (2.1), we have (2.4a) by

$$d(\textstyle\sum_1^n X_i, \sum_1^n X_i') \leqq \sum_1^n P(X_i \neq X_i') = \sum_1^n P(X_i \geqq 2)\,.$$

The second inequality of (2.4b) follows immediately from the relations $p_1' = p_1$ and $p_i' = E(p_i \mid \mathscr{F}_{i-1}')$, $2 \leqq i \leqq n$. It remains to establish the first inequality of (2.4b).

We proceed by constructing X_i' and X_i^*, $1 \leqq i \leqq n$, on a common probability space. Explicitly, we must construct a sequence $X_1^*, \cdots, X_n^*$ of independent Bernoulli variables having the given respective success probabilities $p_1^*, \cdots, p_n^*$ and a sequence $X_1', \cdots, X_n'$ of possibly dependent Bernoulli variables having the joint probability distribution which is determined by the set of quantities $p_1' = P(X_1' = 1)$ and

$$p_i'(x_1', \cdots, x_{i-1}') = P(X_i' = 1 \mid X_1' = x_1', \cdots, X_{i-1}' = x_{i-1}')\,,$$

for $x_j' = 0$ or 1, $1 \leqq j \leqq i-1$, and for $2 \leqq i \leqq n$. Introduce a sequence $U_1, \cdots, U_n$ of independent random variables uniformly distributed on the interval $[0, 1]$. Set $X_i^* = I(U_i \leqq p_i^*)$, $1 \leqq i \leqq n$. Set $X_1' = I(U_1 \leqq p_1')$ and, for $2 \leqq i \leqq n$, $X_i' = I(U_i \leqq p_i'(X_1', \cdots, X_{i-1}'))$. It is readily seen that $\{X_i^*\}$ and $\{X_i'\}$

fulfill the requirements. Further,

$$P(X_i' \neq X_i^*) = E\{P(X_i' \neq X_i^* \mid \mathscr{F}'_{i-1})\} = E|p_i' - p_i^*| , \qquad 1 \leqq i \leqq n .$$

Therefore, by (2.1) again, the first inequality of (2.4b) follows. □

PROOF OF THEOREM 1. By (2.2), write

$$d(\textstyle\sum_1^n X_i, Y) \leqq d(\sum_1^n X_i, \sum_1^n X_i^*) + d(\sum_1^n X_i^*, Y) , \tag{2.5}$$

where $X_1, \cdots, X_n, X_1^*, \cdots, X_n^*$ and Y are as given in Theorem 1 and Lemma 3. Choose $p_i^* = E(p_i)$, $1 \leqq i \leqq n$. Then, using (2.4c) and (1.1) respectively to bound the first and second terms on the righthand side of (2.5), we obtain (1.3). Similarly, the use of (2.3) with (2.4c) and (1.2) yields (1.4). □

REMARK. Note that the p_i^* may be chosen arbitrarily in Lemma 3. For example, if we seek to minimize $E|p_i - p_i^*|$, we may choose p_i^* to be a median of p_i instead of $E(p_i)$.

3. Extended Borel-Cantelli lemma. For arbitrary events $\{E_n\}$,

$$\textstyle\sum_1^\infty P(E_n) < \infty \Rightarrow P(E_n \text{ i.o.}) = 0 . \tag{3.1}$$

That is the "convergence part" of the Borel-Cantelli lemma (see, e.g., Chung (1974)). A converse, or "divergent part," is: if the events $\{E_n\}$ are *independent*, then

$$\textstyle\sum_1^\infty P(E_n) = \infty \Rightarrow P(E_n \text{ i.o.}) = 1 . \tag{3.2}$$

Both parts of the lemma are relevant in connection with studies such as the law of the iterated logarithm, and it thus becomes of interest to relax the independence assumption of the divergent part. One such extension is given by Chung (1974): if the events $\{E_n\}$ are *pairwise independent*, then (3.2) holds. Another dependent extension is given by Iosifescu and Theodorescu (1969), page 2. Let $\phi_1 = 0$ and

$$\phi_n = \sup_{F \in \mathscr{F}_{n-1}} |P(E_n \mid F) - P(E_n)| , \qquad n \geqq 2 ,$$

where $\mathscr{F}_n$ denotes the σ-field generated by $E_1, \cdots, E_n$. Their result is: if the events $\{E_n\}$ satisfy

$$\textstyle\sum_1^\infty \phi_n < \infty , \tag{3.3}$$

then (3.2) holds. We now show that (3.3) may be relaxed to

$$\textstyle\sum_1^\infty E|p_n - P(E_n)| < \infty , \tag{3.4}$$

where $p_1 = P(E_1)$ and, for $n \geqq 2$, $p_n = P(E_n \mid \mathscr{F}_{n-1})$. (Clearly (3.3) implies (3.4) since $|p_n - P(E_n)| \leqq \phi_n$ with probability 1.)

THEOREM 2. *If the events* $\{E_n\}$ *satisfy* (3.4), *then* (3.2) *holds.*

PROOF. It suffices to show that (3.4) and $\sum_1^\infty P(E_n) = \infty$ imply

$$P(\liminf_n E_n^c) = \lim_{m\to\infty} P(\textstyle\bigcap_{n=m}^\infty E_n^c) = 0 , \tag{3.5}$$

where E_n^c denotes the complement of E_n. Let $X_n = I(E_n) = 1$ or 0 according as E_n or E_n^c occurs. Then $\mathscr{F}_1, \mathscr{F}_2, \cdots$ and $p_1, p_2, \cdots$ as defined above also correspond to $X_1, X_2, \cdots$ as in Lemma 3. Hence, by (2.4c),

$$P(\bigcap_{n=m}^{M} E_n^c) = P(\sum_{n=m}^{M} X_n = 0) \leqq P(\sum_{n=m}^{M} X_n^* = 0) + \sum_{n=m}^{M} E|p_n - P(E_n)|,$$

where $X_1^*, X_2^*, \cdots$ are *independent* Bernoulli variables with respective success probabilities $P(E_1), P(E_2), \cdots$. Using the independence,

$$P(\sum_{n=m}^{M} X_n^* = 0) = \prod_{n=m}^{M} [1 - P(E_n)] \leqq \exp[-\sum_{n=m}^{M} P(E_n)].$$

Thus

$$P(\bigcap_{n=m}^{M} E_n^c) \leqq \exp[-\sum_{n=m}^{M} P(E_n)] + \sum_{n=m}^{M} E|p_n - P(E_n)|.$$

Letting $M \to \infty$, we have

$$P(\bigcap_{n=m}^{\infty} E_n^c) \leqq \sum_{n=m}^{\infty} E|p_n - P(E_n)|$$

since $\sum_1^\infty P(E_n) = \infty$. Now letting $m \to \infty$ we obtain (3.5) by virtue of (3.4). □

4. Example: Markov-dependent Bernoulli trials. Consider a sequence $X_1, X_2, \cdots$ of Markov-dependent Bernoulli variables with transition probabilities

$$p_i(1) = P(X_i = 1 \mid X_{i-1} = 1) = \alpha, \qquad p_i(0) = P(X_i = 1 \mid X_{i-1} = 0) = \beta$$

for $i \geqq 2$. Assume $0 < \alpha < 1$, $0 < \beta < 1$ and put

$$\delta = \alpha - \beta, \qquad p = \frac{\beta}{1 - \delta}.$$

Assume for convenience that $P(X_1 = 1) = p$, so that the X_i are identically distributed with success probability p. In this case

$$E|p_i - E(p_i)| = E|p_i - p| = 2|\delta|p(1 - p), \qquad i \geqq 2.$$

Hence, for approximation of $\sum_1^n X_i$ as a Poisson variable Y_n with mean np, Theorem 1 provides the error bound

$$d(\sum_1^n X_i, Y_n) \leqq np^2 + 2(n - 1)|\delta|p(1 - p).$$

The approximation is effective if p and $|\delta|$ are both small relative to np.

Extension to the case $P(X_1 = 1) \neq p$ is straightforward.

5. Simple proof of Lemma 1. Let Y_i be Poisson with mean p_i, let $Z_i = 0$ with probability $(1 - p_i)\exp(p_i)$ and 1 otherwise, and let Z_i and Y_i be independent. Define

$$X_i = I(Y_i \geqq 1) + I(Y_i = 0)I(Z_i = 1).$$

It is quickly checked that X_i is a Bernoulli variable with success probability p_i. Also

$$P(X_i \neq Y_i) = P(Y_i \geqq 2) + P(Y_i = 0, X_i = 1) = p_i(1 - e^{-p_i}) \leqq p_i^2.$$

Using (2.1) we thus have

$$d(\sum_1^n X_i, \sum_1^n Y_i) \leqq \sum_1^n P(X_i \neq Y_i) \leqq \sum_1^n p_i^2.$$

Introducing the assumption that $Y_1, \cdots, Y_n$ are mutually independent, Lemma 1 follows. □ (A similar type of proof was given by Hodges and Le Cam (1960) for a weaker version of Lemma 1.)

Acknowledgment. The author thanks the referee for helpful remarks.

REFERENCES

CHUNG, KAI LAI (1974). *A Course in Probability Theory* 2nd ed. Academic Press, New York.

FRANKEN, P. (1964). Approximation der Verteilungen von Summen unabhängiger nichtnegativer ganzzahler Zufallsgrossen durch Poissonsche Verteilungen. *Math. Nachr.* **27** 303–340.

FREEDMAN, DAVID (1974). The Poisson approximation for dependent events. *Ann. Probability* **2** 256–269.

HODGES, J. L. and LE CAM, L. (1960). The Poisson approximation to the Poisson binomial distribution. *Ann. Math. Statist.* **31** 737–740.

IOSIFESCU, M. and THEODORESCU, R. (1969). *Random Processes and Learning.* Springer-Verlag, New York.

LE CAM, L. (1960). An approximation theorem for the Poisson binomial distribution. *Pacific J. Math.* **10** 1181–1197.

DEPARTMENT OF STATISTICS
THE FLORIDA STATE UNIVERSITY
TALLAHASEE, FLORIDA 32306

Note added in proof. Theorem 2 may also be derived from Corollary 68 of Lévy, P. (1937). *Théorie de l'addition des variables aléatoires*, Paris. I am indebted to D. L. McLeish for bringing to my attention this type of approach.