The Annals of Mathematical Statistics
1970, Vol. 41, No. 4, 1334–1338

ERROR ESTIMATION FOR A LIMIT THEOREM FOR DEPENDENT RANDOM VARIABLES[1]

BY H. W. BLOCK

University of Pittsburgh

1. Introduction. Let (X_{nk}), $k = 1, 2, \cdots k_n$; $n = 1, 2, \cdots$ be a system of random variables. We investigate a limit theorem of a type studied by Loève [3] and particularly an error estimate for this theorem. The method used in finding the error estimate in this case applies to finding estimates for several of the theorems of [3].

2. A convergence theorem for independent systems. Let each X_{nk} have mean μ_{nk} and variance σ_{nk}^2 which we shall assume exists. Let $S_n = \sum_{k=1}^{k_n} X_{nk}$ and $\sigma_n^{\,2} = \sum_{k=1}^{k_n} \sigma_{nk}^2$. If for each n, $X_{n1}, X_{n2}, \cdots, X_{nk_n}$ are independent we say that (X_{nk}) is an *independent system*. We write $\mathscr{L}(S_n) \to \mathscr{L}(X)$ if $F_n(x)$, the distribution function of S_n, converges to $F(x)$, the distribution function of X, at each continuity point of X. We write $\mathscr{L}(X) = \mathscr{L}(Y)$ when X and Y have the same distribution.

It is well known that if a random variable is infinitely divisible and has finite variance it can be represented by the formula of Kolmogorov [2] with unique real constant c and bounded nondecreasing function $K(u)$ which is right continuous and $K(-\infty) = 0$. (Henceforth, we shall call a function with these properties a *Kolmogorov function*). Also it is known [2] that if (X_{nk}) is an independent system of random variables having finite variances and such that $(X_{nk} - \mu_{nk})$ is infinitesimal, then $\mathscr{L}(S_n) \to \mathscr{L}(X)$ if there is a Kolmogorov function $K(u)$ and a constant c such that as $n \to \infty$

$$\sum_{k=1}^{k_n} \int_{-\infty}^{u} x^2 \, dF_{nk}(x + \mu_{nk}) \to K(u) \quad \text{at continuity points of} \quad K(u), \tag{2.1}$$

$$\sum_{k=1}^{k_n} \int_{-\infty}^{\infty} x^2 \, dF_{nk}(x + \mu_{nk}) \to K(\infty), \tag{2.2}$$

$$\sum_{k=1}^{k_n} \mu_{nk} \to c, \tag{2.3}$$

where $\mathscr{L}(X)$ is the infinitely divisible distribution determined by $K(u)$ and c.

3. A convergence theorem for dependent systems. The following notation is the same as that used in [3] and [4] with the exception that distributions will be used in the usual sense (i.e. $F(\infty) = P(-\infty < X < \infty) = 1$) rather than in the more generalized sense of Loève (i.e. $F(\infty) \leqq 1$). We recall that

$$F'_{nk}(x) = P(X_{nk} \leqq x / \textstyle\sum_{j=1}^{k-1} X_{nj})$$

$$E'(X_{nk}) = E(X_{nk} / \textstyle\sum_{j=1}^{k-1} X_{nj})$$

$$\sigma'^2_{nk} = E((X_{nk} - E(X_{nk}))^2 / \textstyle\sum_{j=1}^{k-1} X_{nj}).$$

Received September 18, 1969.

[1] This paper is part of a Ph.D. dissertation accepted by the Ohio State University December 1968.

We say S is a *constancy set* of the function $K(u)$ if $S = \bigcup_{v=1}^{r} I_v$ and I_v is an interval of constancy of $K(u)$ and has for its endpoints continuity points of $K(u)$.

The following is representative of a class of limit theorems which can be obtained using the Comparison Theorem of Loève.

THEOREM 1. *Let (X_{nk}) be a system of random variables such that $(X_{nk}-\mu_{nk})$ is infinitesimal. Then $\mathscr{L}(S_n) \to \mathscr{L}(X)$ if there exists a Kolmogorov function $K(u)$ and a constant c such that as $n \to \infty$,* (2.1), (2.2) *and* (2.3) *hold,*

$$\sum_{k=1}^{k_n} E\left|E'(X_{nk})-E(X_{nk})\right| \to 0, \tag{3.1}$$

there is a constancy set S (S can depend on n and/or k) of $k(u)$, whose complement, S^c, is a bounded set, such that

$$\sum_{k=1}^{k_n} E\left(\int_{S^c} \left|d(F'_{nk}(x+\mu_{nk})-F_{nk}(x+\mu_{nk}))\right|\right) \to 0. \tag{3.2}$$

The random variable X has the infinitely divisible distribution determined by $K(u)$ and c.

PROOF. The proof follows from Loève's Comparison Theorem ([3] or [4]) since

$$\begin{aligned}\sum_{k=1}^{k_n} E&\left|\int_S x\, d(F'_{nk}(x+\mu_{nk})-F_{nk}(x+\mu_{nk}))\right| \\ &\leqq \sum_{k=1}^{k_n} E\left|E'(X_{nk})-E(X_{nk})\right| \\ &\quad +(\max_{x\in S^c}|x|)\cdot \sum_{k=1}^{k_n} E\left(\int_{S^c} (d(F'_{nk}(x+\mu_{nk})-F_{nk}(x+\mu_{nk})\right|\right)\end{aligned}$$

and

$$\sum_{k=1}^{k_n} E\left|\int_S |x|^2 \left|d(F'_{nk}(x+\mu_{nk})-F_{nk}(x+\mu_{nk}))\right|\right| \leqq 2\sum_{v=1}^{r}\sum_{k=1}^{k_n}\int_{I_v} x^2\, dF_{nk}(x+\mu_{nk}).$$

In this finite variance case a similar theorem has been obtained by the author for convergence to the Poisson distribution. Convergence theorems to the normal distribution are proven in [3] and [4].

4. Error estimates. We consider the behavior of $M_n = \sup_{-\infty<x<\infty}|F_n(x)-F(x)|$ where $F_n(x)$ is the distribution function of S_n and $F(x)$ is the distribution function of X in Theorem 1.

LEMMA 1. *Let $K_1(x)$ and $K_2(x)$ be functions of the form*

$$K_j(x) = \int_{-\infty}^{x} u^2\, dF_j(u+\mu_j)$$

where $F_j(x)$ is the distribution function of a random variable Y_j which has mean μ_j and variance σ_j^2 for $j = 1, 2$.
Then

$$\begin{aligned}&\left|\int_{-\infty}^{\infty} x^{-2}(e^{itx}-1-itx)\, d(K_1(x)-K_2(x))\right| \\ &\quad\leqq (5/4)|t|^3\,\delta(\sigma_1^2+\sigma_2^2)+\tfrac{1}{2}t^2\sum_{j=0}^{m}|K_1(x_j)-K_2(x_j)| \\ &\qquad +2|t|A^{-1}[K_1(\infty)-K_1(A)+K_2(\infty)-K_2(A)+K_1(-A)+K_2(-A)]\end{aligned}$$

for any $A>0$, $0<\delta\leqq 2A$ with $m=[2A/\delta]+1$ and any x_j, $j=0,1,2,\cdots,m$, satisfying $-A=x_0<x_1<\cdots<x_m=A$ and $\max_{1\leqq j\leqq m}|x_j-x_{j-1}|<\delta$. (The proof of this result is the same as the proof of the first part of Lemma 3 page 619 of [5].)

Let (X_{nk}) be a system of random variables and let (X_{nk}^*) be its independent version (i.e. $\mathscr{L}(X_{nk}) = \mathscr{L}(X_{nk}^*)$ and (X_{nk}^*) is an independent system). Also define

$$K_{nk}(x) = \int_{-\infty}^{x} u^2\, dF_{nk}(u+\mu_{nk}).$$

$$K'_{nk}(x) = \int_{-\infty}^{x} u^2\, dF'_{nk}(u+\mu_{nk}).$$

$$K_n(x) = \sum_{k=1}^{k_n} K_{nk}(x).$$

LEMMA 2. *Let (X_{nk}) be a system of random variables and let (X_{nk}^*) be its independent version. Then for any $A > 0, 0 \leqq \delta < 2A, m = [2A/\delta]+1$ and any $x_j, j = 0, 1, 2, \cdots, m$, satisfying $-A = x_0 < x < \cdots < x_m = A$ and $\max_{1<j<m}|x_j - x_{j-1}| < \delta$ we have*

$$|f_n(t) - f_n{}^*(t)| \leqq |t| \sum_{k=1}^{k_n} E|E'(X_{nk}) - E(X_{nk})| + (5/2)|t|^3 \delta\sigma_n{}^2$$
$$+ \tfrac{1}{2}t^2 \sum_{j=0}^{m} \sum_{k=1}^{k_n} E|K'_{nk}(x_j) - K_{nk}(x_j)| + 4|t|\sigma_n{}^2/A$$

where $\sigma_n{}^2 = \sum_{k=1}^{k_n} \sigma_{nk}^2$, $f_n(t)$ and $f_n{}^*(t)$ are the characteristic functions of S_n and $S_n{}^*$.

PROOF. Let $\phi'_{nk}(t) = E(e^{it}(X_{nk} - \mu_{nk}) \mid \sum_{l=1}^{k-1} X_{nl})$ and $\phi_{nk}(t) = E(e^{it}(X_{nk}^* - \mu_{nk}))$. Using the same proof as in Lemma B page 65 of [3] we have the following inequality $|f_n(t) - f_n{}^*(t)| \leqq \sum_{k=1}^{k_n} E|\phi'_{nk}(t) - \phi_{nk}(t)|$. By Lemma 1, at each value of the random variable $K'_{nk}(x)$ we have,

$$\left|\int_{-\infty}^{\infty} x^{-2}(e^{itx} - 1 - itx)\, d(K'_{nk}(x) - K_{nk}(x))\right|$$
$$\leqq (5/4)|t|^3 \delta\, \sigma'^2_{nk} + \sigma^2_{nk}) + \tfrac{1}{2}t^2 \sum_{j=0}^{m} |K'_{nk}(x_j) - K_{nk}(x_j)|$$
$$+ 2A^{-1}|t|\, |K'_{nk}(\infty) - K'_{nk}(A) + K_{nk}(\infty) - K_{nk}(A) + K'_{nk}(-A) + K_{nk}(-A)|.$$

Finally

$$|f_n(t) - f_n{}^*(t)| \leqq \sum_{k=1}^{k_n} E\left|\int_{-\infty}^{\infty} x^{-2}(e^{itx} - 1 - itx)\, d(K'_{nk}(x) - K_{nk}(x)) + it(E'(X_{nk}) - \mu_{nk})\right|$$
$$\leqq |t| \sum_{k=1}^{k_n} E|E'(X_{nk}) - E(X_{nk})| + (5/2)|t|^3 \delta\sigma_n{}^2$$
$$+ \tfrac{1}{2}t^2 \sum_{j=0}^{m} \sum_{k=1}^{k_n} E|K'_{nk}(x_j) - K_{nk}(x_j)| + 4t\sigma_n{}^2/A$$

which is the required inequality.

We now investigate the behavior of M_n for Theorem 1. A bound is first obtained for this quantity. Then under the conditions of Theorem 1, it is shown that the bound converges to 0 as $n \to \infty$.

THEOREM 2. *Let X be an infinitely divisible random variable with distribution function $F(x)$, mean μ, variance σ^2, characteristic function $f(t)$, and Kolmogorov function $K(x)$. Let (X_{nk}) be a system of random variables with $\sigma_{nk}^2 \leqq 1$ for each X_{nk} and assume $dF(x)/dx \leqq B$ for all x where B is some positive constant. Then it follows that for any number $a > 1$ we have*

$$M_n = \sup_{-\infty<x<\infty} |F_n(x) - F(x)| \leqq k(a, B)g(n, m(A, \delta))$$

where $k(a, B)$ is a constant depending only on a and B,

$$g(n, m(A,\delta)) = [(5/16)n^2 \max_{1\leq k\leq k_n} \sigma_{nk}^2]^{1/5} + [(5/6)\,\delta(3\sigma_n{}^2+\sigma^2)]^{\frac{1}{3}}$$
$$+[\tfrac{1}{2}\textstyle\sum_{i=0}^{m}\sum_{k=1}^{k_n} E\,|K'_{nk}(x_i)-K_{nk}(x_i)| + \tfrac{1}{2}\sum_{i=0}^{m}|K_n(x_i)-K(x_i)|]^{\frac{1}{3}}$$
$$+[2\textstyle\sum_{k=1}^{k_n} E\,|E'(X_{nk})-E(X_{nk})| + 2\,|\mu_n-\mu|$$
$$+4A^{-1}(2\sigma_n{}^2+K_n(\infty)-K_n(A)$$
$$+K(\infty)-K(A)+K_n(-A)+K(-A))]^{\frac{1}{2}},$$

and $\sigma_n{}^2$, δ, A, $m(A, \delta) = m$, and $\{x_i\}$ are as defined in Lemma 2. (The numbers A and δ depend on n).

PROOF. We let (X_{nk}^*) be the independent version of (X_{nk}). We then have

$$|f_n(t)-f(t)| \leq |f_n(t)-f_n{}^*(t)| + |f_n{}^*(t)-f(t)|$$

where Lemma 2 may be applied to the first term on the right-hand side of the inequality. The second term may be bounded by $h(t, n, m(A, \delta))$ for $|t| \leq T_n$ where

$$h(t,n,m(A,\delta)) = [(5/8)t^4\sigma_n{}^2 \max_{1\leq k\leq k_n}\sigma_{nk}^2 + |t|\,|\mu_n-\mu| + (5/4)\,|t|^3\,\delta(\sigma_n{}^2+\sigma^2)$$
$$+\tfrac{1}{2}t^2\textstyle\sum_{i=0}^{m}|K_n(x_i)-K(x_i)|$$
$$+2\,|t|\,A^{-1}\{K_n(\infty)-K_n(A)+K(\infty)-K(A)+K_n(-A)+K(-A)\}],$$

and

$$T_n = [(1/3)\sigma_n{}^2 \max_{1\leq k\leq k_n}\sigma_{nk}^2)^{1/5} + ((5/6)\,\delta(\sigma_n{}^2+\sigma^2))^{\frac{1}{3}} + (\tfrac{1}{2}\textstyle\sum_{i=0}^{m}|K_n(x_i)-K(x_i)|)^{\frac{1}{3}}$$
$$+(4A^{-1}\{K_n(\infty)-K_n(A)-K(\infty)-K(A)+K_n(-A)+K(-A)\}+2\,|\mu_n-\mu|)^{\frac{1}{2}}]^{-1}.$$

This follows from the proof of Theorem 3 of [5]. Letting $T_n' = (g(n, m(A, \delta))^{-1}$ it is observed that $T_n' \geq T_n$. It is then easily seen that

$$\textstyle\int_{-T_n'}^{T_n'} |(f_n(t)-f(t))/t|\,dt \leq g(n,m(A,\delta)).$$

Then applying a result of Esseen [1] we obtain that for every $a > 1$, there is a finite positive number $c(a)$ depending only on a, such that

$$M_n \leq (2\pi)^{-1}a\textstyle\int_{-T_n'}^{T_n'}|(f_n(t)-f(t))/t|\,dt + c(a)\cdot B\cdot(T_n')^{-1}$$
$$\leq (2\pi)^{-1}ag(n,m(A,\delta)) + c(a)\cdot B\cdot g(n,m(A,\delta))$$
$$= k(a,B)\cdot g(n,m(A,\delta))$$

where $k(a, B) = (2\pi)^{-1}\cdot a + c(a)\cdot B$ which depends only on a and B. This completes the proof.

In Theorem 2 let $\sigma < \delta < 1$ be such that $\pm\delta$ are continuity points of $K(x)$ and choose $A = \delta^{-\frac{1}{2}}$. Let $m(\delta) = m(\delta^{-\frac{1}{2}}, \delta) = m(A, \delta)$.

We now show that the bound obtained in Theorem 2 converges to zero as $n \to \infty$ under the conditions of Theorem 1.

THEOREM 3. *Let (X_{nk}) be a system of random variables satisfying the conditions of Theorem* 1. *Let $F(x)$ be the limit distribution and assume that there is a positive constant B such that $dF(x)/dx < B$ for all x. Then there exists a sequence (δ_n) with $0 < \delta_n < 1$ such that $\delta_n \to 0$ as $n \to \infty$ and $g(n, m(\delta_n)) \to 0$ as $n \to \infty$.*

PROOF. We first show how the (δ_n) are chosen. Let S be a set satisfying (3.2). Then

$$\sum_{k=1}^{k_n} E\left|K'_{nk}(x_i) - K_{nk}(x_i)\right| \leqq 2\sum_{k=1}^{k_n} \int_S x^2\, d(F_{nk}(x+\mu_{nk}) + (\max_{x\in S^c} x^2)\sum_{k=1}^{k_n} E(\int_{S^c} \left|d(F'_{nk}(x+\mu_{nk}) - F_{nk}(x+\mu_{nk}))\right|)$$

which converges to 0 as $n \to \infty$ by (2.1) and (3.2). Thus for each $\delta < 0$

$$\lim_{n\to\infty} \sum_{i=0}^{m(\delta)} \sum_{k=1}^{k_n} E\left|K'_{nk}(x_i) - K_{nk}(x_i)\right| = 0. \tag{4.1}$$

We also have $\lim_{n\to\infty} K_n(x_i) = K(x_i)$ by (2.1) since (x_i) were chosen to be continuity points of $K(x)$. Thus for each $\delta < 0$,

$$\lim_{n\to\infty} \sum_{i=0}^{m(\delta)} \left|K_n(x_i) - K(x_i)\right| = 0. \tag{4.2}$$

It is then possible to choose (δ_n) with $0 < \delta_n < 1$ such that $\pm\delta_n^{-\frac{1}{2}}$ are continuity points of $K(x)$, $\lim_{n\to\infty} \delta_n = 0$, and (4.1) and (4.2) hold with δ replaced by δ_n.

We need only notice that under the conditions of Theorem 1 $\lim_{n\to\infty} \sigma_n^{\ 2} = \sigma^2$, where σ^2 is the variance of the limit distribution, and $\lim_{n\to\infty} \max_{1\leqq k\leqq k_n} \sigma_{nk}^2 = 0$. Thus the theorem follows.

Also notice that under the conditions of the previous theorem $\sigma_{nk}^2 \leqq 1$ for large n since $\lim_{n\to\infty} \max_{1\leqq k\leqq k_n} \sigma_{nk}^2 = 0$. This condition was needed in Theorem 2 to obtain the bound and is not a serious restriction.

Using similar techniques bounds can be obtained for the aforementioned convergence theorems to the normal and the Poisson law.

REFERENCES

[1] ESSEEN, C. G. (1945). Fourier analysis of distribution functions. *Acta Math.* **77** 1–125.

[2] GNEDENKO, B. V. and KOLMOGOROV, A. N. (1954). *Limit Distributions for Sums of Independent Random Variables*, tr. K. L. Chung. Addison-Wesley, Reading.

[3] LOÈVE, M. (1950). On sets of probability laws and their limit elements. *Univ. California Publ. Statist.* **1** 53–87.

[4] LOÈVE, M. (1963). *Probability Theory* (3rd ed.). Van Nostrand, Princeton.

[5] SHAPIRO, J. M. (1955). Error estimates for certain probability theorems. *Ann. Math. Statist.* **26** 617–630.