The Annals of Probability
1983, Vol 11, No. 4, 1042–1047

A BERRY-ESSEEN THEOREM FOR ASSOCIATED RANDOM VARIABLES[1]

BY THOMAS E. WOOD

Louisiana State University

Uniform rates of convergence in the Central Limit Theorem for associated random variables are given. Applications to the Ising Model and Diffusion and Gaussian processes are discussed.

1. Introduction. A real function $f(x_1, \cdots, x_k)$ of k real variables will be called increasing if it is a nondecreasing function of each of the separate variables $x_1, \cdots, x_k$. Let $\{X(t): t \in T\}$, where T is a subset of the real numbers, be a collection of random variables. If for any k, any $\{t_i, \cdots, t_k\}$ contained in T, and each pair of Borel measurable increasing functions of k variables f and g the inequality

$$\operatorname{Cov}[f(X(t_1), \cdots, X(t_k)), g(X(t_1), \cdots, X(t_k))] \geq 0$$

holds, then the variables $\{X(t): t \in T\}$ are called associated (Esary, Proschan, and Walkup, 1967).

Before stating our main result, we establish some notation to be used throughout the remainder of the paper. For a sequence $\{X_n: n = 1, 2, \cdots\}$ of random variables we set $\bar{S}_n = (X_1 + \cdots + X_n)/n^{1/2}$. $F_n(x)$ denotes the distribution function of $\bar{S}_n$, $\sigma_n^2 = \mathrm{E}(\bar{S}_n^2)$, and $\rho_n = E|\bar{S}_n|^3$. We use $N_A(x)$ to denote the normal distribution function with zero mean and variance A^2.

THEOREM 1. *Suppose* $\{X_n: n = 1, 2, \cdots\}$ *is a sequence of associated random variables satisfying the following*:

(1) *Zero mean*: $EX_n = 0$, *for all n*
Finite Variance: $0 < EX_n^2 < \infty$ *for all n*
Finite third moment: $E|X_n|^3 < \infty$, *for all n*

(2) *Stationarity*: *for all m and for all j*, $k_1, \cdots, k_m$ *integers*, $(X(k_1), \cdots, X(k_m))$ *has the same distribution as* $(X(j + k_1), \cdots, X(j + k_m))$.

(3) *Finite susceptibility*: $A^2 \equiv EX_1^2 + 2\sum_{k=2}^{\infty} \operatorname{Cov}(X_1, X_k) < \infty$.

Then for $n = m \cdot k$

$$|F_n(x) - N_A(x)| \leq [(16\,\sigma_k^4 m(A^2 - \sigma_k^2))/(9\pi\rho_k^2)] + 3\rho_k/\sigma_k^3 m^{1/2}.$$

This result provides uniform rates of convergence in the Central Limit Theorem for associated random variables due to C. M. Newman (1980). The rates of convergence provided by our result depend on the rate of convergence of σ_k^2 to A^2 and the growth of ρ_k. We exhibit below examples from statistical mechanics and diffusion processes where ρ_k is bounded as k goes to infinity and the theorem gives its strongest estimates. In the Berry-Esseen Theorem for independent random variables the uniform rate of convergence of $F_n(x)$ to the normal distribution has order of magnitude $O(n^{-1/2})$. We give an example of Gaussian processes to show that in the generality of our results such a convergence rate is not to be expected. We have no general results on the growth of ρ_k for a stationary associated sequence.

Received August 1982.

[1] The results contained here are taken from the author's dissertation written at the University of Virginia under the direction of Professor Loren D. Pitt.

AMS 1980 *subject classifications.* Primary, 60F05; secondary, 60J60.

Key words and phrases. Association, Berry-Esseen theorem, Lebowitz inequality.

2. Proof of Theorem 1. The proof requires two lemmas. One is the smoothing lemma of A. C. Berry (Feller, 1971). The other is an inequality on characteristic functions of associated random variables due to C. M. Newman (1980).

SMOOTHING LEMMA. *Let F be a probability distribution with zero mean and characteristic function $\varphi(s) = \int e^{isx} dF(x)$. Suppose that $F - G$ vanishes at $\pm\infty$ and that G has a derivative g such that $|g| \le B$. Finally, suppose that g has a continuously differentiable Fourier transform γ such that $\gamma(0) = 1$ and $\gamma'(0) = 0$. Then*

$$|F(x) - G(x)| \le \pi^{-1} \int_{-T}^{T} |(\varphi(s) - \gamma(s))/s| \, ds + 24B/\pi T$$

holds for all x and $T > 0$.

NEWMAN'S INEQUALITY. Suppose $X_1, \cdots, X_n$ are associated random variables with finite variances; then for any real $\lambda_1, \cdots, \lambda_n$

$$|E \exp(i \textstyle\sum_{k=1}^{n} \lambda_k X_k) - \prod_{k=1}^{n} E \exp(i\lambda_k X_k)| \le \sum_{k=1, j>k}^{n} |\lambda_k| |\lambda_j| \operatorname{Cov}(X_k, X_j).$$

Applying the triangle inequality to the smoothing lemma we get the following

LEMMA 1. *Suppose $\{X_n: n = 1, 2, \cdots\}$ satisfies the conditions of Theorem 1. Let $\varphi_n(s)$ be the characteristic function of $\bar{S}_n$ and $B = \sup\{|N'_A(x)|: x \in \mathbb{R}\}$. Then*

$$|F_{mk}(x) - N_A(x)| \le \pi^{-1}(I_1 + I_2 + I_3) + 24B/\tau T$$

where

$$I_1 = \int_{-T}^{T} |(\varphi_{mk}(s) - \varphi_k^m(s/\sqrt{m}))/s| \, ds;$$

$$I_2 = \int_{-T}^{T} |(\varphi_k^m(s/\sqrt{m}) - \exp(-\sigma_k^2 s^2/2))/s| \, ds;$$

$$I_3 = \int_{-T}^{T} |(\exp(-\sigma_k^2 s^2/2) - \exp(-A^2 s^2/2))/s| \, ds.$$

The proof of Theorem 1 now follows by estimating the three integrals I_1, I_2, and I_3. We set

$$Y_j^k = (X((j-1)k+1) + \cdots + X(jk))/k^{1/2}$$

and $T = 4\sigma_k^2 m^{1/2}/3\rho_k$. Using Newman's inequality we get

$$\begin{aligned} I_1 &\le \int_{-T}^{T} (\textstyle\sum_{l=1, j>l}^{m} s^2 \operatorname{Cov}(Y_j^k, Y_l^k))/m \, |s| \, ds \\ &= (T^2 \textstyle\sum_{l=1, j>l}^{m} \operatorname{Cov}(Y_j^k, Y_l^k))/m = (T^2/2)(\operatorname{Var}(\bar{S}_{mk}) - \operatorname{Var}(\bar{S}_k)) \\ &= (T^2/2)(\sigma_{mk}^2 - \sigma_k^2) \le (T^2/2)(A^2 - \sigma_k^2) = (16\, \sigma_k^4 m (A^2 - \sigma_k^2))/18\rho_k^2. \end{aligned}$$

Next, by Taylor's Theorem we have

$$I_3 \le 2 \int_0^T [(A^2 - \sigma_k^2)s^2]/2s \, ds = T^2(A^2 - \sigma_k^2)/2 = (16\, \sigma_k^4 m (A^2 - \sigma_k^2))/18\rho_k^2.$$

The estimates we use next mimic those which prove the Berry-Esseen Theorem for independent sequences. First note the inequality for complex numbers α, β with

$$|\alpha| \le \lambda \text{ and } |\beta| \le \gamma: |\alpha^n - \beta^n| \le n|\alpha - \beta| \gamma^{n-1}.$$

From the stationarity of $\{X_n\}$ we have that $EY_j^k = 0$, $E(Y_j^k)^2 = \sigma_k^2$ and

$$\varphi_k(s) = E\exp(\text{is } Y_j^k) = E\exp(\text{is } \bar{S}_k).$$

Also

$$|\varphi_k(s) - 1 + \sigma_k^2 s^2/2| = \left| \int_{-\infty}^{\infty} (e^{isx} - 1 - isx + s^2x^2/2)\, dF_k(x) \right|$$

$$\leq \int_{-\infty}^{\infty} |s^3x^3/6|\, dF_k(x) = \rho_k |s|^3/6.$$

Hence, if $\sigma_k^2 s^2/2 \leq 1$, $|\varphi_k(s)| \leq 1 - \sigma_k^2 s^2/2 + \rho_k|s|^3/6$. For $|s| \leq T$ this gives

$$|\varphi_k(s/\sigma_k\sqrt{m})| \leq 1 - s^2/2m + \rho_k|s|^3/6\sigma_k^3 m^{3/2}$$

$$\leq 1 - 5s^2/18m \leq \exp(-5s^2/18m).$$

We note that since σ_k^3 is always less than or equal to ρ_k the theorem holds trivially when $m \leq 9$ and so we assume without loss of generality that $m \geq 10$. Now

$$I_2 = \int_{-\sigma_k T}^{\sigma_k T} |(\varphi_k^m(s/\sigma_k\sqrt{m}) - \exp(-s^2/2))/s|\, ds.$$

Setting $\gamma^{m-1} = \exp(-s^2/4)$ and using $e^{-x} - 1 + x \leq x^2/2$ for $x > 0$, it follows finally that

$$I_2/\pi + 24B/\pi T \leq \pi^{-1} \int_{-\sigma_k T}^{\sigma_k T} ((2s^2/9 + |s|^3/18)\exp(-s^2/4))/T\, ds + 24B/\pi T$$

$$\leq 3\rho_k/\sigma_k^3 m^{1/2}.$$

In the last inequality $\pi < 9/5$ was used to approximate B. The proof of Theorem 1 is complete.

REMARKS. Since $A^2 - \sigma_k^2$ goes to zero as k goes to infinity it is always possible to regulate the growth of m with k in such a way that $\rho_k/m^{1/2}$ and $m^{1/2}(A^2 - \sigma_k^2)/\rho_k$ both go to zero.

We have done some work in replacing variances and absolute third moments with truncated moments (Wood, 1982).

3. Ferromagnets. Some important examples of stationary associated sequences of random variables arise in models for ferromagnets in mathematical physics. Typical among these models is the classical Ising model for which association of the spin variables is equivalent to their satisfying the FKG-inequalities (Simon, 1974).

Moreover, for the ferromagnetic Ising models, the Lebowitz Inequality

$$EX_iX_jX_kX_l \leq (EX_iX_j)(EX_kX_l) + (EX_iX_k)(EX_jX_l) + (EX_iX_l)(EX_jX_k)$$

holds. A recent derivation of the Lebowitz Inequality for Ising models and Φ^4 field theories can be found in the paper by Brydges, Frölich, and Spencer (1982).

THEOREM 2. *Let $\{X_n : n = 1, 2, \cdots\}$ be a stationary sequence of mean zero associated random variables with finite fourth moments. Suppose $\{X_n\}$ satisfies the finite susceptibility condition and the Lebowitz Inequality. Then*

$$\rho_n = E|(X_1 + \cdots + X_n)/n^{1/2}|^3 = E|\bar{S}_n\|^3$$

is bounded as n goes to infinity,

PROOF. First note that $\rho_n \leq (E\bar{S}_n^4)^{3/4}$. The Lebowitz Inequality and the association give

$$E\,\bar{S}_n^4 = (1/n^2)\sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n E\,X_iX_jX_kX_l$$
$$\le (3/n^2)\sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n (EX_iX_j)(EX_kX_l)$$
$$\le (3/n^2)\sum_{i=1}^n \sum_{k=1}^n \sum_{j=1}^n (EX_iX_j)\sum_{l=1}^\infty (EX_kX_l).$$

The last sum is bounded by the finite susceptibility assumption. This completes the proof.

In the one dimensional Ising model, at sufficiently high temperatures, $\sum_{n=1}^\infty EX_1X_n$ is a convergent geometric series (Thompson, 1972). Theorem 1 and Theorem 2 then combine to give constants C and D so that

$$|F_n(x) - N_A(x)| \le C/m^{1/2}$$

where $n = D\,m \log m$.

4. Diffusions in $\mathbb{R}^1$. Consider the one dimensional diffusion process $\{X_t: t \in \mathbb{R}\}$ with generator

$$G = d^2/dx^2 - b(x)\,d/dx. \tag{4}$$

We take $b(x)$ to be a bounded, smooth even function with $b(x) = \beta/|x|$ for $|x| \ge \delta > 0$ and $\beta > 1$. We will show that if $\beta > 9$ then

$$\rho_T = T^{-3/2}E\left|\int_0^T X_t\,dt\right|^3$$

remains bounded as T goes to infinity. This process generated by G is associated by a theorem of I. Herbst and L. D. Pitt (Pitt, 1983). Before stating this theorem, we note a definition due to T. E. Harris (1977).

DEFINITION. A semigroup P^t is called monotone if P^tf is a bounded, increasing function whenever f is a bounded, increasing function in the domain of P^t. If P^t is the transition semigroup of a Markov process X_t and P^t is monotone then X_t is also called monotone.

HERBST-PITT THEOREM. *Let G generate a unique diffusion on $\mathbb{R}^n$ where G is of the form*

$$Gf(x) = \sum_{i=1}^n \sum_{j=1}^n a_{ij}(x)\partial_i\partial_j f(x) + \sum_{i=1}^n b_i(x)\partial_i f(x).$$

Let $P^t(x, dy)$ be the corresponding Markov semigroup.

A. *The condition that $a_{ij} \ge 0$ for all i and all j is necessary and sufficient for*

$$G(fg) - fGg - gGf \ge 0 \tag{5}$$

to hold for all smooth increasing functions f and g in the domain of G.

B. *Necessary and sufficient conditions that P^t be monotone are*

a) $\partial_i b_j \ge 0$ if $i \ne j$

b) $\partial_k a_{ij} = 0$ if $k \notin \{i, j\}$.

The following formal calculation shows how one uses this theorem to prove the positive covariances necessary for association. We assume P^t is monotone and f and g are increasing. Set $h(s) = P^{t-s}[P^sf(x)P^sg(x)]$ so that $h(0) = P^t(fg)$ and $h(t) = (P^tf)(P^tg)$. Then for $s \in (0, t)$

$$h'(s) = -P^{t-s}[G(P^sfP^sg) - (GP^sf)P^sg - P^sf(GP^sg)]$$

which cannot be positive by (5). Therefore

$$\mathrm{Cov}(f,(X_t), g(X_t)) = h(0) - h(t) \ge 0.$$

We now return to showing ρ_T is bounded as T goes to inifinity for $\beta > 9$. We use Feller's representation $G = D_m D_s$ for the generator in terms of the speed measure $m(dx)$ and natural scale function $s(x)$. Elementary calculations show that in our case for $|x| \geq \delta$ we may take $m'(x) = |x|^{-\beta}$ and $s'(x) = |x|^{\beta}$. In particular we observe that $m(dx) = m'(x)\,dx$ is a finite measure since $\beta > 1$. Furthermore, P^t is self-adjoint on $L^2(\mathbb{R}, dm)$, or what is the same, X_t has time reversed symmetry if we take $m(dx)$ (appropriately normalized) as our initial distribution.

Let $\varphi(x) = x$ be the identity function on $\mathbb{R}$ and $0 \leq i < j < k < l$. Then

$$EX_iX_jX_kX_l = EP^{j-i}\varphi(X_j)X_jX_kP^{l-k}\varphi(X_k).$$

Observe that P^t monotone and $b(x)$ even imply that P^t preserves odd increasing functions. Hence, $0 \leq xP^t\varphi(x)$ as the product of two odd increasing functions. Denote by R_0 the 0th order resolvent operator for X_t:

$$R_0 f(x) = E_x \int_0^\infty f(X_t)\,dt.$$

Thus

$$\begin{aligned} \left| E \int_A X_qX_rX_sX_t\,dqdrdsdt \right| &\leq \int_B EP^{r-q}\varphi(X_r)X_rX_sP^{t-s}\varphi(X_s)\,dqdrdsdt \\ &\leq \int_C ER_0\varphi(X_r)X_rX_sR_0\varphi(X_s)\,drds \\ &\leq (T^2/2) \int_{-\infty}^{\infty} |xR_0\varphi(x)|^2\,dm(x) \end{aligned}$$

where $A = \{(q, r, s, t): 0 < q < r < s < t < T\}$, $B = \{(q, r, s, t): 0 < r < s < T \text{ and } -\infty < q < r < s < t < \infty\}$, and $C = \{(r, s): 0 < r < s < T\}$.

Assuming for the moment that $xR_0\varphi(x)$ is in $L^2(\mathbb{R}, dm)$ we now have the estimates

$$\rho_T \leq (1/T^2)\left(E\left(\int_0^T X_t\,dt\right)^4\right)^{3/4} \leq C\|xR_0\varphi(x)\|_2^2$$

for some constant $C > 0$.

We now show that $xR_0\varphi(x)$ is in $L^2(\mathbb{R}, dm)$ provided that $\beta > 9$. We make no claim that this is necessary for ρ_T to be bounded, but only that it is sufficient.

For $x > \delta$

$$R_0\varphi(x) = \int_0^x s'(y)\left\{\int_y^\infty \varphi(z)m'(z)\,dz\right\}dy = x^3/3(\beta - 2).$$

Hence $\int_{-\infty}^{\infty} |xR_0\varphi(x)|^2 dm(x) < \infty$ if and only if $\int_\delta^\infty \{x^4/3(\beta - 2)\}^2/x^\beta\,dx < \infty$. This is equivalent to $\beta > 9$.

REMARKS. The diffusion processes we have discussed provide examples in which we have rapid convergence in Theorem 1, but they do not satisfy the usual strong mixing conditions used for proving the Central Limit Theorem for Markov processes (Rosenblatt, 1978).

That these processes do not satisfy the strong mixing condition is an easy consequence of the fact that G is self-adjoint on $L^2(\mathbb{R}, dm)$. In fact,

$$\sup\{Ef(X_0)g(X_t)\} = e^{-\Lambda t}$$

where the supremum is over all pairs (f, g) of functions with $Ef(X_0) = Eg(X_t) = 0$ and $Ef(X_0)^2 = 1$ and

$$\Lambda = \inf\left\{-\int f(x) \cdot Gf(x)m'(x)\,dx\right\}$$

over the set of all smooth functions with compact support satisfying $\int f\, dm = 0$ and $\int |f|^2\, dm = 1$. By choosing $f_n(x)$ to be smooth approximations of the odd function $g(x)$ with

$$g(x) = \begin{cases} 0 & \text{for } 0 \le x \le \lambda \\ x - \lambda & \text{for } \lambda \le x \le \lambda + 1 \\ 1 & \text{for } \lambda + 1 \le x \end{cases}$$

it is easy to see that $\Lambda = 0$ and hence X_t cannot be strongly mixing.

4. Gaussian processes. With the following example we intend to demonstrate that the convergence rate in Theorem 1 can be arbitrarily slow. This differs from the Berry-Esseen Theorem for independent random variables which provides the order of magnitude $n^{-1/2}$ for $|F_n(x) - N_1(x)|$.

As the Fourier transform of a positive L^1 function

$$r(s, t) = r(t - s) = [1 + (t - s)^2]^{-\nu};\ \tfrac{1}{2} < \nu < \tfrac{3}{4}$$

(Magnus, et. al., 1966) is the covariance function, $r(s, t) = EX_sX_t$, of a stationary, mean zero, Gaussian process $\{X_t : t \in \mathbb{R}\}$. By a theorem of L. D. Pitt (1982) the process X_t is associated.

Since $\nu > \frac{1}{2}$ we have the finite susceptibility condition, $\int_0^\infty r(t)\, dt < \infty$, holds. The process X_t satisfies the conditions of Theorem 1 and we can get estimates on the convergence rate of $|F_n(x) - N_A(x)|$ independently of the theorem because all the distributions involved are normal.

For two mean zero normal distributions $N_A(x)$ and $N_B(x)$ the maximum of $|N_B(x) - N_A(x)|$ occurs at $\pm x_0$ where $x_0 = (2A^2B^2(\log A - \log B))^{1/2}/(A^2 - B^2)$. Note that x_0 converges to A as B goes to A. A Taylor's Theorem approximation shows that $|N_B(x_0) - N_A(x_0)|$ has order of magnitude $O(A - B)$ as B goes to A.

Applying the notation of Theorem 1 we have $F_n(x) = N_B(x)$ for

$$B = \sigma_n^2 = 1 + (2/n) \textstyle\sum_{i=1, j>i}^n EX_iX_j = 1 + (2/n) \sum_{i=1, j>i}^n r(j - i).$$

By varying the choice of ν we can make σ_n^2 converge arbitarily slowly to A.

REFERENCES

[1] BRYDGES, D., FRÖHLICH, J., and SPENCER, T. (1982). The random walk representation of classical spin systems and correlation inequalities. *Commun. Math. Phys.* **83** 123–150.
[2] ESARY, J. D., PROSCHAN, F. and WALKUP, D. W. (1967). Association of random variables with applications. *Ann. Math. Statist.* **38** 1466–1474.
[3] FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications*, vol. II, 2nd ed. John Wiley, New York.
[4] HARRIS, T. E. (1977). A correlation inequality for Markov processes in partially ordered state spaces. *Ann. Probab.* **5** 451–454.
[5] MAGNUS, W., OBERHETTINGER, F., and SONI, R. P. (1966). *Formulas and Theorems for the Special Functions of Mathematical Physics.* Springer-Verlag, New York.
[6] NEWMAN, C. M. (1980). Normal fluctuations and the FKG-inequalities. *Commun. Math. Phys.* **74** 119–128.
[7] PITT, L. D. (1982). Positively correlated normal variables are associated. *Ann. Probab.* **10** 496–499.
[8] PITT, L. D. (1983). Proceedings of the Symposium on Stochastic Inequalities at Lincoln, Nebraska, 1982. (To appear).
[9] ROSENBLATT, M. (1978). Studies in probability theory. MAA Studies in Math. **18.**
[10] SIMON, B. (1974). *The $P(\Phi)_2$ Euclidean (Quantum) Field Theory.* Princeton University Press.
[11] THOMPSON, C. J. (1972). *Mathematical Statistical Mechanics.* Macmillan, New York.
[12] WOOD, T. E. (1982). *Sequences of Associated Random Variables.* Dissertation, University of Virginia.

DEPARTMENT OF MATHEMATICS
LOUISIANA STATE UNIVERSITY
BATON ROUGE, LOUISIANA 70803