ANNALES DE L'INSTITUT FOURIER

R. M. DUDLEY

# Metric entropy and the central limit theorem in $C(S)$

*Annales de l'institut Fourier*, tome 24, nº 2 (1974), p. 49-60

<http://www.numdam.org/item?id=AIF_1974__24_2_49_0>

Ann. Inst. Fourier, Grenoble
**24**, 2 (1974), 49-60.

# METRIC ENTROPY AND THE CENTRAL LIMIT THEOREM IN C(S)

by Richard M. DUDLEY

A central limit theorem will be proved in the Banach space C(S) where S is a compact metric space. It will be assumed that the individual random variables $X_n$ in C(S) are independent and identically distributed and satisfy

$$(1) \quad |X_1(s) - X_1(t)| \leqslant M(\omega)e(s, t) \quad \text{for all} \quad s, t \in S,$$

where M is a random variable with $EM^p < \infty, p > 2$, and $e$ is a metric on S for which the $\varepsilon$-entropy H satisfies $\limsup_{\varepsilon \downarrow 0} \varepsilon^\alpha H(S, e, \varepsilon) < \infty$ for some $\alpha < 2p/(p+2)$.

The first theorem of this type apparently was that of Strassen (1969) for $p = \infty$, i.e. for M bounded. The main credit for the further extension to $p < \infty$ belongs to Evarist Giné (1974), who treated the case $p = 2$. His method, using truncation and Bernstein's inequality, will also be followed below. Theorem 1 below, described in the previous paragraph, can be considered as interpolating between Strassen's result for $p = \infty$ and Giné's for $p = 2$.

If $p = \infty$ or $p = 2$, the condition on H can be weakened to

$$\int_0^1 H^{\frac{1}{2}+p^{-1}}(S, e, t)\,dt < \infty.$$

For $2 < p < \infty$, Giné (1973) has found a different and often weaker sufficient condition, namely

$$H^{(p+3)/2p}(S, e, \varepsilon) = o(1/\varepsilon|\log \varepsilon|) \quad \text{as} \quad \varepsilon \downarrow 0.$$

Section 2 presents some counterexamples which indicate why hypotheses of a weaker type would not be enough.

DEFINITIONS. — Given a compact space S, C(S) denotes the set of all continuous real-valued functions on S, metrized as usual by the supremum norm.

Given $\varepsilon > 0$, let

$$N(S, e, \varepsilon) = \inf \left\{ N : \exists A_j,\ S = \bigcup_{j=1}^{N} A_j,\ \sup_{x, y \in A_j} e(x, y) \leqslant 2\varepsilon \right\},$$
$$H(S, e, \varepsilon) = \log N(S, e, \varepsilon).$$

Given a C(S)-valued random variable X, $EX = f$ means that for any $\varphi \in C(S)^*$, $E \int X\, d\varphi = \int f\, d\varphi$.

If $X_1, X_2, \ldots,$ are independent C(S)-valued random variables, we say the *central limit theorem holds* for the $X_j$ iff there is a Gaussian process Z on S with continuous sample functions such that $\mathscr{L}\left(n^{-\frac{1}{2}}(X_1 + \cdots + X_n)\right) \to \mathscr{L}(Z)$ in C(S) as $n \to \infty$, i.e. for every bounded (non-linear) real continuous functional F on C(S),

$$EF\left(n^{-\frac{1}{2}}(X_1 + \cdots + X_n)\right) \to EF(Z).$$

If the $X_j$ are identically distributed with law $\mu$, where $\mu$ is a Borel probability measure on C(S), then we say the central limit theorem holds for $\mu$ iff it holds for the $X_j$.

Since the Lipschitz condition on $X_1$ may seem to be a strong assumption, it should be noted that S may originally be given with some other metric $d$. Then, since S is compact and $X_1 \in C(S)$, there are some numbers $\delta_m \downarrow 0$ fast enough so that

$$\Pr \{\sup \{|X_1(s) - X_1(t)| : d(s, t) \leqslant \delta_m\} \geqslant m^{-2}\} \leqslant m^{-2}.$$

Then there is a modulus of continuity $g$, i.e. a continuous, subadditive, increasing function on $[0, \infty)$ with $g(0) = 0$, such that $g(\delta_m) \geqslant m^{-2}$. Letting $e = g \circ d$ we now have a metric $e$ such that 1) holds for some random variable M, although M may not have a pth moment. Thus the hypotheses limit the size of S as measured in terms of the modulus of continuity of $X_1$.

The central limit theorem poses more difficulties in C(S) than in some other Banach spaces. For example in $L^r$ for $2 \leqslant r < \infty$, $EX_1 = 0$ and $E\|X_1\|^2 < \infty$ imply the central limit theorem (Fortet and Mourier [1955]). The counterexamples given in sec. 2 below confirm the known fact that in $C(S)$ stronger conditions are needed.

The reader familiar with Gaussian processes may also note that to ask whether a Gaussian process has sample functions in $L^p$ is usually a much deeper question for $p = \infty$ than for $p < \infty$.

Since every separable Banach space is isometric to a linear subspace of a space $C(S)$, our central limit theorem gives as a corollary a general central limit theorem for separable Banach spaces. This corollary will, however, be far from best possible for many Banach spaces, such as $L^2$, where metric entropy hypotheses are not really relevant.

THEOREM 1. — *Suppose* $(S, e)$ *is a compact metric space and* $\mu$ *is a probability measure on* $C(S)$ *such that for some* $p > 2$ *there is a random variable* $M \in \mathscr{L}^p(\mu)$ *such that*

*a)* $|f(x) - f(y)| \leqslant M(f)e(x, y)$ *for all* $x, y \in S$ *and* $\mu$*-almost all* $f \in C(S)$,

*b) For all* $x \in S$, $E_\mu f(x) = 0$ *and* $E_\mu f(x)^2 < \infty$, *and*

*c)* $H(S, e, \varepsilon) = O(\varepsilon^{-\alpha})$ *as* $\varepsilon \downarrow 0$ *for some* $\alpha < 2p/(p+2)$.

*Then the central limit theorem holds for* $\mu$.

COROLLARY. — *Let* $Y$ *and* $X$ *be separable Banach spaces and let* $T$ *be a bounded linear transformation from* $Y$ *into* $X$. *Let* $Y_n$ *be independent and identically distributed in* $Y$ *with* $EY_1 = 0$ *and* $E\|Y_1\|^p < \infty$ *for some* $p > 2$. *Let* $S$ *be the unit ball in the dual space* $X^*$, *with weak-* topology. Let* $e$ *be the usual norm metric on* $Y^*$. *Suppose* $H(T^*(S), e, \varepsilon) = O(\varepsilon^{-\alpha})$ *as* $\varepsilon \downarrow 0$ *for some* $\alpha < 2p/(p+2)$. *Then the central limit theorem holds in* $X$ *for the variables* $X_j = T(Y_j)$.

In the situation of Theorem 1 and its Corollary, the central limit theorem may fail if $\alpha > 2$, no matter how large $p$ is. This is not too surprising, since the limiting Gaussian process may fail to have continuous sample functions if the exponent of entropy is greater than 2, although that is for a possibly different metric $E^{\frac{1}{2}}(X_1(s) - X_1(t))^2$. I do not know whether $2p/(p+2)$ is a best possible bound for $\alpha$ if $p < \infty$; it is if $p = \infty$ (Strassen and Dudley [1969], section 2).

*Proof of Theorem* 1. — In *a*) we can assume

$$M(f) = \sup \{|f(x) - f(y)|/e(x, y) : x \neq y\},$$

and we can take $E_\mu M(f)^p \leqslant 1$ and $\alpha > 1$.

Let $Y_1, Y_2, \ldots,$ be independent C(S)-valued random variables with distribution $\mu$. By *a*) and *b*),

$$E_\mu(f(x) - f(y))^2 \leqslant E_\mu M(f)^2 e(x, y)^2 \leqslant e(x, y)^2.$$

Hence by *c*), since $\alpha < 2$, the Strassen-Sudakov theorem implies that the limiting Gaussian process Z has continuous sample functions. It remains to show that the distributions $\mathscr{L}(n^{-\frac{1}{2}}(Y_1 + \cdots + Y_n))$ are uniformly tight on C(S).

To do this we will truncate the $Y_n$. Let $M_n = M(Y_n)$. Then the $M_n$ are independent identically distributed random variables with a pth moment. Fix a $\gamma$ such that $\alpha < 1/\gamma < 2p/(p+2)$, i.e. $\frac{1}{2} + p^{-1} < \gamma < 1/\alpha$. Let $\delta = \gamma - \frac{1}{2} - p^{-1}$. Let

$$U_n = Y_n \quad \text{if} \quad |M_n| \leqslant \frac{1}{2} n^{\delta+(1/p)},$$
$$0 \text{ otherwise.}$$

Then $\sum_{n=1}^{\infty} \Pr(U_n \neq Y_n) < \infty$, so it suffices to prove that the central limit theorem holds for the $U_n$.

We have $E\|U_n\| \leqslant E\|Y_n\|$ which is bounded uniformly in $n$ by *a*) and *b*). Also

$$|EU_n(x) - EU_n(y)| \leqslant E|U_n(x) - U_n(y)| \leqslant E|Y_n(x) - Y_n(y)| \leqslant (E_\mu M)e(x, y),$$

for all $x, y \in S$, so that $EU_n(x)$ is a Lipschitzian function of $x$. Now

$$|EU_n(x)| = |EY_n(x) + E(U_n - Y_n)(x)| = |E(Y_n - U_n)(x)| \leqslant (EY_n^2(x))^{\frac{1}{2}} \Pr(|M_n| > n^{\delta+(1/p)})^{\frac{1}{2}} \leqslant Cn^{-\frac{1}{2}-\frac{1}{2}p\delta}$$

for some constant C. Thus $\sum_{j=1}^{n} EU_j(x)/n^{\frac{1}{2}} \to 0$ as $n \to \infty$, uniformly in $x$. Hence we can center the $U_n$: let $X_n = U_n - EU_n$. We need only prove the central limit theorem for the $X_n$, which satisfy: $EX_n = 0$, $EX_n(t)^2 < \infty$ for all $t \in S$, $M_p \equiv \sup_n EM(X_n)^p < \infty$, and we can assume $M_p \leqslant 1$; finally $M(X_n) \leqslant n^{\delta+(1/p)}$ for all $n \geqslant n_0$ (where $n_0$ does not depend on $\omega$), so we can assume it holds for all $n$.

Let $S_n = X_1 + \cdots + X_n$. We have

$$(2) \quad n^{-\frac{1}{2}}|S_n(x) - S_n(y)| \leqslant n^{\gamma}e(x, y) \quad \text{for all} \quad x, y \in S.$$

Next we use an upper exponential bound, specifically Bernstein's inequality (cf. Bennett [1962], and for a correction to the proof and a similar application, Giné [1974]). We have for any $\varepsilon > 0$

$$(3) \quad \Pr\left\{n^{-\frac{1}{2}}|S_n(s) - S_n(t)| \geqslant \varepsilon\right\} \leqslant \exp\left(-\varepsilon^2/[2e(s, t)^2 + \varepsilon n^{\gamma-1}e(s, t)]\right)$$

for any $s, t \in S$.

Since $\sup_n E(n^{-1}S_n(x)^2) < \infty$ for any $x \in S$, the uniform tightness of $\mathscr{L}(n^{-\frac{1}{2}}S_n)$ will be proved if we can establish the following « probable equicontinuity » result: for some $\varepsilon_m \to 0$,

$$(4) \quad \sup_n \Pr\left\{\sup\left\{|S_n(x) - S_n(y)|/n^{\frac{1}{2}} : e(x, y) \leqslant 2^{-m}\right\} > \varepsilon_m\right\} \leqslant \varepsilon_m.$$

Take any $K$ such that $\gamma < K < 1/\alpha$. In proving that (2) and (3) imply (4) for a given $n$, we will use (2) for $e(s, t) \leqslant n^{-K}$ and (3) for $e(s, t) > n^{-K}$.

If we use (2) for $2^{-m} \leqslant n^{-K}$, we will have (4) in this case for $\varepsilon_m \geqslant 2^{-m}n^{\gamma}$. To do this for all $n$ we need

$$\varepsilon_m > 2^{-m} \sup\{n^{\gamma} : n^K \leqslant 2^m\},$$

for which it will suffice to take

$$\varepsilon_m \geqslant 2^{-m+m(\gamma/K)} \equiv \varepsilon'_m \to 0 \qquad \text{as} \qquad m \to \infty$$

since $\gamma < K$.

It remains to show that (3) implies (4) for suitable $\varepsilon_m \to 0$ if $2^{-m} > n^{-K}$, i.e. for $m < K \log_2 n$. (Here $\log_2$ denotes logarithm to the base 2.)

For $m = 1, 2, \ldots$, let $F_m$ be a finite set of minimal cardinality such that for every $x \in S$ there is some $x_m \in F_m$

with $e(x, x_m) \leqslant 2^{-m-3}$. We choose such an $x_m$ for each $x$. For some constant $C \geqslant 1$ we have by $c)$:

$$(5) \qquad \operatorname{card}(F_m) \leqslant \exp(C2^{m\alpha}).$$

For any positive integers $k$ and $n$ we shall choose numbers $\varepsilon_{kn} > 0$ such that

$$(6) \qquad \lim_{m\to\infty} \sup_n \sum_{k=m}^{[K \log_2 n]} P_{kn} = 0$$

where

$$P_{kn} = \Pr\{\exists x, y \in F_k \cup F_{k+1} : \\ e(x, y) \leqslant 2^{1-k},\ n^{-\frac{1}{2}}|S_n(x) - S_n(y)| \geqslant \varepsilon_{kn}\},$$

and such that $\lim_{m\to\infty} \beta_m = 0$ where $\beta_m = \sup\{\beta_{mn} : n > 2^{m-K}\}$

and
$$\beta_{mn} = \sum_{k=m}^{[K \log_2 n]} \varepsilon_{kn}.$$

Then to obtain (6) it will suffice to make $P_{kn} \leqslant \varepsilon_{kn}$.

If we can find such $\varepsilon_{kn}$, then given $m$ and $n$ and any $x, y$ with $e(x, y) \leqslant 2^{-m} > n^{-K}$, let $r = [K \log_2 n] + 1$, $T_n(x, y) = n^{-\frac{1}{2}}|S_n(x) - S_n(y)|$. Then since

$$e(x_m, y_m) \leqslant 2 \cdot 2^{-m-3} + 2^{-m} < 2^{1-m},$$

we have, except on a set of probability at most $\sum_{j=m}^{r-1} P_{jn}$, the inequality

$$\begin{aligned} T_n(x, y) &\leqslant T_n(x, x_r) + T_n(y, y_r) + T_n(x_m, y_m) \\ &\quad + \sum_{j=m}^{r-1} [T_n(x_j, x_{j+1}) + T_n(y_j, y_{j+1})] \\ &\leqslant 2n^{\gamma-K} + \sum_{j=m}^{r-1} 3\varepsilon_{jn} \leqslant 2 \cdot 2^{m(\gamma-K)/K} + 3\beta_m. \end{aligned}$$

Then we could take $\varepsilon_m = \max(\varepsilon'_m, 2 \cdot 2^{m(\gamma-K)/K} + 3\beta_m)$ and obtain (4) as desired.

We must still find $\varepsilon_{kn}$ to satisfy (6) and $\beta_m \to 0$. By (3) and (5) we have

$$P_{kn} \leqslant 4 \exp(8C2^{k\alpha} - \varepsilon_{kn}^2/[4^{2-k} + \varepsilon_{kn} n^{\gamma-1} 2^{1-k}]).$$

Thus $P_{kn} \leqslant \varepsilon_{kn}$ will follow from

$$8C2^{k\alpha} - \varepsilon_{kn}^2/[4^{2-k} + \varepsilon_{kn}n^{\gamma-1}2^{1-k}] \leqslant \log \varepsilon_{kn} - \log 4,$$

or from

$$\varepsilon_{kn}^2 > [9C2^{k\alpha} + |\log \varepsilon_{kn}|](4^{2-k} + 2^{1-k}n^{\gamma-1}\varepsilon_{kn}). \tag{7}$$

(7) will follow from the three inequalities

$$|\log \varepsilon_{kn}| \leqslant 7C2^{k\alpha}, \tag{A}$$
$$\varepsilon_{kn}^2 > 32C2^{k\alpha-2k+4}, \quad \text{and} \tag{B}$$
$$\varepsilon_{kn} > 32C2^{k\alpha-k+1}n^{\gamma-1}. \tag{C}$$

(B) and (C) will both be satisfied when we set, for a sufficiently large constant $N > 1$,

$$\varepsilon_{kn} = N \max\left(2^{\frac{1}{2}k(\alpha-2)}, n^{\gamma-1}2^{k(\alpha-1)}\right). \tag{8}$$

Then (8) also implies (A) for $k$ large enough, since

$$\sup_n |\log \varepsilon_{kn}| \leqslant \log N + k(2 - \alpha) < 7C2^{k\alpha}, \quad k \text{ large.}$$

To evaluate the maximum in the definition (8) let $\zeta = 2(1 - \gamma)/\alpha$. Then

$$\varepsilon_{kn} = \begin{cases} 2^{\frac{1}{2}k(\alpha-2)}N & \text{for} \quad k \leqslant \zeta \log_2 n, \\ n^{\gamma-1}2^{k(\alpha-1)}N & \text{for} \quad \zeta \log_2 n < k \leqslant r. \end{cases}$$

Hence

$$\beta_{mn} = \sum_{k=m}^{r-1} \varepsilon_{kn} \leqslant \sum_{k=m}^{[\zeta \log_2 n]} N2^{\frac{1}{2}k(\alpha-2)} + \sum_{k=[\zeta \log_2 n]}^{r-1} Nn^{\gamma-1}2^{k(\alpha-1)}.$$

The first sum is part of the tail of a convergent geometric series, since $\alpha < 2$, so it approaches $0$ as $m \to \infty$, uniformly in $n$. The second sum is at most

$$N(1 + K \log_2 n)n^{\gamma-1}2^{r(\alpha-1)},$$

since $\alpha > 1$. As $m \to \infty$, $n > 2^{m/K} \to \infty$ so for $m$ large,

$$N(1 + K \log_2 n)n^{\gamma-1} < n^{K-1},$$

so the second sum is smaller than $2^{m(\alpha-1+1-(1/K))} \to 0$ as $m \to \infty$ since $\alpha < 1/K$, and the proof is complete.

## Counterexamples.

The following examples seem to show that assumptions on moments of the norm of $X_1$ and of differences $X_1(s) - X_1(t)$ do not give good central limit theorems. The examples are based on the same scheme as those in Strassen and Dudley (1969), sec. 3. The idea of extending this scheme to find stronger counter-examples was suggested by A. de Araujo (1973), although his examples there do not go as far as the ones below.

Proposition. — *For any* $K < \frac{1}{2}$ *there is a process* $X(t)(\omega)$, $0 \leqslant t \leqslant 1$, *with continuous sample functions,* $|X(t)(\omega)| \leqslant 1$ *for all* $t$ *and* $\omega$, *and* $E(X(s) - X(t))^2 \leqslant |s - t|^K$ *for all* $s, t \in [0, 1]$, *such that the central limit theorem does not hold for (independent identically distributed variables in* $C([0, 1])$ *with) the distribution of* $X$.

*Proof.* — For each $n = 1, 2, \ldots$, we shall decompose $[0, 1]$ into a set $I_n$ of $N_n$ equal subintervals, where $N_n = \prod_{s=1}^{n} 6k_s$, $k_s$ integers. Thus each interval in $I_{n-1}$ is decomposed into $6k_n$ equal subintervals to form $I_n$, where $I_0 = \{[0, 1]\}$.

For each $n$ and each $j = 0, \ldots, k_n - 1$, we define a piecewise linear continuous function $g_{nj}$ as follows. Let

$$g_{nj}(x) = \begin{cases} 0 & \text{if } N_n x/3 \text{ is an integer,} \\ 1 & \text{if } 6i + 1 \leqslant N_n x \leqslant 6i + 2, \\ -1 & \text{if } 6i + 4 \leqslant N_n x \leqslant 6i + 5, \end{cases}$$

where

$$i = j + rk_n, \qquad r = 0, 1, \ldots, N_{n-1} - 1,$$

and let $g_{nj}$ be continuous and linear on those closed intervals for which it was previously defined only at the endpoints, namely $6i + u \leqslant N_n x \leqslant 6i + u + 1, u = 0, 2, 3, 5$.

Note that for each $j$, inside every interval in $I_{n-1}$ is an interval in $I_n$ on which $g_{nj} = 1$ and another on which $g_{nj} = -1$.

Let $p_n = cn^{-\beta}$ where $1 < \beta < 2$ and $c = 1\Big/\sum_{n=1}^{\infty} n^{-\beta}$.

In Strassen and Dudley (1969) we took $\beta = 5/4$ but here the choice is not important. To be definite, we take $\beta = 3/2$.

Now we define a probability measure $\mu$ on $C([0, 1])$ by setting $\mu(\{g_{nj}\}) = \mu(\{-g_{nj}\}) = p_n/2k_n$ for each $n = 1, 2, \ldots$ and each $j = 0, \ldots, k_n - 1$. Let $X$ be a random variable with distribution $\mu$. Then clearly $|X(t)| \leqslant 1$. Also for each $t$, $EX(t) = 0$ since $X$ is symmetric and bounded.

Now we prove that the central limit theorem never holds for $\mu$ with the given $p_n$, for any $k_n \geqslant 2$.

Let $\Omega$ be a probability space over which independent processes $X_1, X_2, \ldots,$ are defined, each with distribution $\mu$. For $\omega \in \Omega$ let $A_{mn} = A_{mn}(\omega) = \{r \leqslant m : (\exists j)\, X_r(\omega) = \pm g_{nj}\}$

Let $f_{mn} = \sum_{r \in A_{mn}} X_r$. Let $B_{mn} = \{\omega : (\exists t)\ f_{mn}(t) \neq 0\}$. If $\omega \in B_{mn}$, then there is a $j$ such that $f_{mn} \geqslant 1$ either on all intervals where $g_{nj} = 1$ or on all those where $g_{nj} = -1$. Thus for any $j_1, \ldots, j_n$ with $j_s = 0, 1, \ldots, k_s - 1$, $s = 1, \ldots, n$, and for any signs $\sigma_s = \pm 1$, there is an interval in $I_n$ on which $\sigma_s g_{sj_s} = 1$ for all $s = 1, \ldots, n$.

Let $Z_m = m^{-\frac{1}{2}}(X_1 + \cdots + X_m)$. Then $\max Z_m \geqslant m^{-\frac{1}{2}} J_m$ where $J_m = J_m(\omega)$ is the number of values of $n$ with $\omega \in B_{mn}$.

Let $M_{mn}$ be the number of elements of $A_{mn}$. Then

$$\Pr(M_{mn} = 0) = (1 - p_n)^m \leqslant \exp(-mp_n) \leqslant 1/e \text{ if } mp_n \geqslant 1.$$

This holds for $n = 1, 2, \ldots, [(cm)^{2/3}] \equiv n_m$ where $[x]$ denotes the greatest integer $\leqslant x$.

Let $K_m = [n_m - m^\gamma]$ where $\frac{1}{2} < \gamma < 2/3$. For definiteness et $\gamma = 5/8$. Then

$$\begin{aligned}
&\Pr\{M_{mn} = 0 \text{ for at least } K_m \text{ values of } n \leqslant n_m\}\\
&\quad \leqslant \exp(-K_m)\binom{n_m}{K_m} \leqslant \exp(-K_m + (m^\gamma + 1)\log(cm)^{2/3})\\
&\quad \leqslant \exp\left(-\frac{1}{2}(cm)^{2/3}\right) < \frac{1}{2}
\end{aligned}$$

for $m$ large enough.

The conditional probability of $B_{mn}$, given that $M_{mn} \geqslant 1$ and any conditions on the $X_r$ for $r \notin A_{mn}$, is at least $\frac{1}{2}$.

Thus by comparison to binomial probabilities, the conditional probability that at least $1/3$ of $m^\gamma$ such events occur is asymptotically at least $\frac{1}{2}$, and

$$\liminf_{m\to\infty} \Pr\{J_m \geqslant m^\gamma/3\} \geqslant \frac{1}{4}$$

by the weak law of large numbers. Hence

$$\liminf_{m\to\infty} \Pr\{\max Z_m \geqslant m^{1/8}/3\} \geqslant \frac{1}{4},$$

so the distributions of the $Z_m$ are not uniformly tight and cannot converge.

Now we estimate mean-square differences. Given $s, t \in [0, 1]$, take $n$ such that $1/N_{n+1} < |s - t| \leqslant 1/N_n$, where $N_0 = 1$. Note that $X(s) - X(t) = 0$ unless either $s$ or $t$ belongs to some interval on which $X = \pm g_{nj} \neq 0$. Thus

$$\begin{aligned} E(X(s) - X(t))^2 \leqslant & \sum_{m<n} (2p_m k_m^{-1}(6N_m|s-t|)^2) \\ & + 2p_n k_n^{-1} N_n^2 |s-t|^2 + 8\Big(\sum_{m>n} p_m/k_m\Big) \\ \leqslant & \ 72 k_n^{-2} + 2p_n k_n^{-1} N_n^2 |s-t|^2 + 8 \sum_{m>n} p_m/k_m \end{aligned}$$

since $N_m|s - t| \leqslant 1/k_n$ for $m < n$.

Now we want to choose the $k_n$ to make $E(X(s) - X(t))^2$ as small as possible. Fix any $b > 0$ and let

$$k_n = [\exp((1 + b)^n)].$$

Then for $n$ large,

$$\begin{aligned} E(X(s) - X(t))^2 \leqslant & \ 8/k_{n+1} + 72 k_n^{-2} \\ & + \exp(-(1 + b)^n) N_n^2 |s - t|^2. \end{aligned}$$

Now by summation of a finite geometric series,

$$N_n^2 = \prod_{j=1}^{n} k_j^2 \leqslant \exp(2b^{-1}(1 + b)^{n+1}),$$

so

$$\begin{aligned} & E(X(s) - X(t))^2 \\ & \leqslant 24 \exp(-(1 + b)^{n+1}) + 216 \exp(-2(1 + b)^n) \\ & + |s - t|^2 \exp((2 + b)\, b^{-1}(1 + b)^n). \end{aligned}$$

Now again by summing a geometric series, we have for any $\varepsilon > 0$ and $n$ large enough

$$N_n = \prod_{j=1}^{n} [\exp(1+b)^j] \geqslant \exp(-n + b^{-1}((1+b)^{n+1} - 1 - b))$$
$$\geqslant \exp(b^{-1}(1+b)^{n+1}(1-\varepsilon)).$$

Thus for $b \geqslant 1$,

$$E(X(s) - X(t))^2$$
$$\leqslant 240\, N_{n+1}{}^{-2b/(1+b)^2} + |s-t|^{\varepsilon} N_n{}^{(2+b)/(1+b)(1-\varepsilon)}$$
$$\leqslant 240|s-t|^{2b/(1+b)^2} + |s-t|^{(1-\delta)b/(1+b)},$$

for some $\delta > 0$. To maximize the smaller of the two exponents of $|s-t|$ we let $\varepsilon \downarrow 0$, so that $\delta \downarrow 0$, and let $b = 1$, so we get the upper bound $241|s-t|^{K}$ for any $K < \frac{1}{2}\cdot$ Replacing X by X/13 we can get rid of the constant 241 and the proof is complete.

It is known that if $E(X(s) - X(t))^2 \leqslant C|s-t|^{1+\varepsilon}$ for some constants $C < \infty$ and $\varepsilon > 0$, $s, t \in [0, 1]$, then X has a version with continuous sample functions. (This was first proved by Kolmogorov; see Loève (1963), p. 519.) Since $n^{-\frac{1}{2}} S_n$ has the same second-moment structure as $X_1$, for all $n$, it is not hard to see that if also $EX(s)^2 < \infty$ for some (and hence all) $s \in [0, 1]$, then Kolmogorov's theorem works uniformly in $n$ to estimate the modulus of sample function continuity and boundedness, so that the central limit theorem must hold. This observation apparently was first made by A. de Araujo (1973).

## BIBLIOGRAPHY


Aloisio de ARAUJO, On the central limit theorem for $C(I^k)$ valued random variables (preprint, Statistics Dept., Univ. of Calif., Berkeley), 1973.

G. BENNETT, Probability inequalities for the sum of independent random variables, *Jour. Amer. Statist. Assoc.*, 57 (1962), 33-45.

R. M. DUDLEY, Sample functions of the Gaussian process, *Ann. Probability*, 1 (1973), 66-103.

R. FORTET and E. MOURIER, Les fonctions aléatoires comme éléments aléatoires dans les espaces de Banach, *Studia Math.*, 15 (1955), 62-79.

Evarist Giné, On the central limit theorem for sample continuous processes, to appear in *Annals of Probability*, 1974.

Evarist Giné, A note on the central limit theorem in C(S), (preprint), 1973.

M. Loève, (1963), *Probability Theory* (Princeton, Van Nostrand).

V. Strassen and R. Dudley, (1969), The central limit theorem and $\varepsilon$-entropy, *Lecture Notes in Math.*, 89, 224-231.

Richard M. Dudley,
Room 2-245,
Massachusetts Institute of Technology,
Cambridge, Mass. 02139 (USA).