# Correlations and Spectra for Non-Stationary Random Functions

By J. Kampé de Fériet and François N. Frenkiel

**1. Introduction.** The definition of the correlation function and the spectrum of a stationary random function is now classical, but in many applications one feels the need to extend this definition to random functions, which, although non-stationary, are in some sense *nearly* stationary. We suggest, therefore, for the definition of the correlation of a random function whose covariance $\Gamma(t, s)$ is known, the limit

$$(1.1) \qquad R(h) = \lim_{T \to +\infty} \frac{1}{T} \int_{\frac{|h|}{2}}^{T - \frac{|h|}{2}} \Gamma\left(\xi - \frac{h}{2}, \xi + \frac{h}{2}\right) d\xi,$$

if this limit exists for every $h$. The spectrum $S(\lambda)$ can then be obtained from $R(h)$ in the classical way.

We are led to the above definition of the *correlation function* $R(h)$ by the following considerations: we determine the *sample-correlation* from a truncated sample of the random function; we then obtain a *sub-correlation*, $R_T(h)$, of the random function (defined as the correlation of the truncated random function) by averaging the sample correlations; finally, the correlation $R(h)$ is defined by (1.1) as the limit of $R_T(h)$, if this limit exists.

The function $R(h)$, so defined, has all the properties of a correlation function. If the random function is stationary (wide sense) [4, p. 95–96], our definition coincides with the classical definition. The estimation of the correlation of a stationary random function has been considered extensively in the literature, particularly by U. Grenander and M. Rosenblatt [6], R. B. Blackman and J. W. Tukey [1], and E. Parzen [13, 14]. In order to evaluate how good the estimate $R(h)$ is from the sample-correlations $\rho_T(h, \omega)$, which are the only experimental observables, we compute the variance of the random variables $\rho_T(h, \omega)$ about $R_T(h)$, and then we compute (for a fixed $h$) an upper bound of $R(h) - R_T(h)$ for large $T$.

This paper is especially concerned with the case in which the random function has a periodic covariance $\Gamma(t + \tau, s + \tau) = \Gamma(t, s)$. To appreciate the scope of the above condition, let us note that it is always satisfied when the random function is a sum of two uncorrelated random functions, one being a stationary (wide sense) random function and the other a periodic random function.

The last part of the paper is devoted to the estimate of $R(h)$ for a non-stationary random step-function $V(t, \omega)$, similar to the one introduced by N. Wiener,

$$(1.2) \qquad V(t, \omega) = X_n(\omega), \quad n - 1 \leq t < n, n = 1, 2, \cdots,$$

where $[X_1(\omega), \cdots X_n(\omega), \cdots]$ is a sequence of independent random variables taking only the values $-1$ and $+1$ with equal probability.

An experimental function will be constructed using a table of random numbers, and sample-correlations will be determined. Estimates for the sub-correlations are

then determined by taking averages over the experimental correlations. The accuracy of these estimates is characterized by giving the variance for the departure of the sample-correlations from the estimated sub-correlations. The experimental data are then compared with the theoretical results. (Cf. [5], [9].)

**2. Random Functions, Covariances and Correlation Functions.** We consider here a *random function* of a real variable $t$ as an *ensemble* of real functions $f(t, \omega)$, where $\omega$ is a parameter chosen at random in some set $\Omega$ according to a probability measure $\mu$ [7]. A *sample* of the random function $f(t, \omega)$ is simply the real function $f(t, \omega_0)$ corresponding to a particular choice of $\omega_0$ in the set $\Omega$. It is convenient for many applications to take for $\Omega$ a *function space*; each point $\omega$ is then a function $\omega(t)$ belonging to some prescribed class of functions (e.g., a continuous function on $[0, 1]$). One has thus for each sample $f(t, \omega) = \omega(t)$. When this particular choice is made for $\Omega$, one says that the random function is of "function space type" [4, p. 67].

The following general hypotheses shall be made with regard to the random functions considered in the present paper:

$\mathbf{H_1}$. $f(t, \omega)$ is measurable with respect to the product measure $m \times \mu$ (where $m$ is the Lebesgue measure on the real line $-\infty < t < +\infty$).

$\mathbf{H_2}$. For each $t$, $f(t, \omega) \in L^2(\Omega)$: $\overline{f(t, \omega)^2} < +\infty$. (If $F(\omega) \in L(\Omega)$ we denote its mean value by $\int_\Omega F(\omega)d\mu = \overline{F(\omega)}$.)

$\mathbf{H_3}$. $\mathbf{H_2}$ implies $f(t, \omega) \in L(\Omega)$; we suppose $\overline{f(t, \omega)} = 0$.

$\mathbf{H_4}$. It follows from $\mathbf{H_2}$ that the *covariance*

$$(2.1) \qquad \Gamma(t, s) = \overline{f(t, \omega)f(s, \omega)}$$

exists for all $t$'s and $s$'s. We assume that,

$$(2.2) \qquad \Gamma(t, t) \in L[a, b]$$

for every finite interval $a \leq t \leq b$. From (2.2), by the Fubini-Tonelli theorem, [8, Vol. 1, p. 609] it follows that

$$(2.3) \qquad f(t, \omega) \in L^2[a, b]$$

for almost all samples, in any finite interval $a \leq t \leq b$. This implies also that

$$(2.4) \qquad f(t, \omega) \in L[a, b] \qquad \text{for almost all samples.}$$

In our earlier paper [9] instead of (2.2), we assumed that

$$(2.2') \qquad F(t, s) \in L(\Delta^2)$$

for every finite rectangle $\Delta^2$ in the plane $(t, s)$. We are indebted to the referee for a simple counter example showing that $(2.2')$ does not always imply (2.3) and for the suggestion that we replace $(2.2')$ by (2.2); the proof that (2.2) implies (2.3) by the Fubini-Tonelli theorem is straightforward.

If the random function $f(t, \omega)$ is stationary

$$(2.5) \qquad \Gamma(t, s) = \rho(t - s)$$

where $\rho(h)$ is called a *correlation function*. Due to $\mathbf{H_1}$, $\rho(h)$ is uniformly continuous in any finite interval [3]. A real function $\rho(h)$ is a correlation function if, and only

if, it is symmetric and positive-definite

$$(2.6) \qquad \rho(-h) = \rho(h)$$

$$(2.7) \qquad \sum_j \sum_k X_j X_k \rho(h_j - h_k) \geqq 0$$

for any set $[h_1, \cdots h_n]$ with $n$ arbitrary.

**3. Truncated Samples of Random Functions.** In experiments concerning a random function $f(t, \omega)$ one can materialize, as a rule, the sample of the function only for a finite interval, that is, one knows only *truncated samples*. As far as finite intervals are considered, one often uses the notation $[-T, +T]$ for the interval in which the samples are known in the experiment. Rather than this *two-sided* (symmetric with respect to $t = 0$) *truncation*, we shall prefer here a *one-sided truncation* (starting at $t = 0$) and we will define a *truncated sample* by

$$(3.1) \qquad \begin{aligned} f_T(t, \omega_0) &= f(t, \omega_0), & 0 \leqq t \leqq T, \\ f_T(t, \omega_0) &= 0, & t < 0 \text{ or } t > T. \end{aligned}$$

This definition implies that the experiment starts at $t = 0$; we assume that it could be extended for an arbitrary time $T$ in the future, but not in the past (time prior to the beginning of the experiment). From the samples we will draw some inference with regard to the random function for $0 \leqq t < +\infty$, but completely ignore it for $t < 0$.

**4. Correlation and Spectrum of a Truncated Sample.** For a truncated sample, corresponding to a given $\omega_0$, we define a *sample correlation* as

$$(4.1) \qquad \rho_T(h, \omega_0) = \frac{1}{T} \int_{|h|/2}^{T-|h|/2} f\left(\xi - \frac{h}{2}, \omega_0\right) f\left(\xi + \frac{h}{2}, \omega_0\right) d\xi, \qquad \text{for } |h| \leqq T,$$

and

$$(4.2) \qquad \rho_T(h, \omega_0) = 0, \qquad \text{for } |h| \geqq T.$$

Let us remark that both (4.1) and (4.2) can be replaced by the formula

$$
\begin{aligned}
(4.3) \qquad \rho_T(h, \omega_0) &= \frac{1}{T} \int_{-\infty}^{+\infty} f_T\left(\xi - \frac{h}{2}, \omega_0\right) f_T\left(\xi + \frac{h}{2}, \omega_0\right) d\xi \\
&= \frac{1}{T} \int_{-\infty}^{+\infty} f_T(\xi, \omega_0) f_T(\xi + |h|, \omega_0) \, d\xi, \qquad \text{for all } h.
\end{aligned}
$$

From $\mathbf{H_2}$ it follows that the correlation $\rho_T(h, \omega)$ exists for almost all samples (i.e., with probability one). The great advantage of our definition is that *the correlation $\rho_T(h, \omega)$ is a positive-definite function of $h$, uniformly continuous in $h$ (for each $\omega$ for which it exists)*.

If we had used as correlation of the truncated sample, as is very often done, the function

$$(4.4) \qquad \bar{\rho}_T(h, \omega) = \frac{\rho_T(h, \omega)}{1 - \dfrac{|h|}{T}}$$

then we would have completely missed these important properties of $\rho_T(h, \omega)$. Indeed, $\check{\rho}_T(h, \omega)$ would have been, in general, discontinuous at $|h| = T$ and no longer a positive-definite function. Thus $\check{\rho}_T(h, \omega)$ would not have been the Fourier transform of a function $\psi_T(\lambda, \omega)$.

The only advantage gained by using $\check{\rho}_T(h, \omega)$ is that if $f$ is constant, $\check{\rho}_T$ is also a constant while $\rho_T$ is not. This fact was probably the reason which led the statisticians to use this definition for the correlation of a truncated function. However, the nonexistence of a spectrum $\psi_T(\lambda, \omega)$ which may have to be reintroduced later by various artifices, may lead to serious complications in the estimation of the correlation functions, particularly when numerical methods are used for that purpose.

The spectrum $\psi_T(\lambda, \omega)$ is very simply connected to the complex Fourier transform of the sample

$$(4.5) \qquad \alpha_T(\lambda, \omega) = \frac{1}{T} \int_0^T e^{-i\lambda t} f(t, \omega)\, dt = \frac{1}{T} \int_{-\infty}^{+\infty} e^{-i\lambda t} f_T(t, \omega)\, dt.$$

Due to (2.3), the Fourier transform exists for almost all samples and, by Plancherel's theorem, [8, Vol. 2, p. 742], $\alpha_T(\lambda, \omega) \in L^2[-\infty, +\infty]$ (but not, in general, to $L[-\infty, +\infty]$). From (4.5) we have:

$$(4.6) \qquad |\alpha_T(\lambda, \omega)|^2 = \frac{1}{T^2} \iint_{-\infty}^{+\infty} e^{i\lambda(s-t)} f_T(t, \omega)\, f_T(s, \omega)\, dt\, ds.$$

Let us consider the $(t, s)$ plane and make the change of variables

$$t = \xi - \frac{h}{2}, \qquad s = \xi + \frac{h}{2},$$

$$(4.7)$$

$$\xi = \frac{t + s}{2}, \qquad h = s - t.$$

For any $F(t, s) \in L(R^2)$

$$(4.8) \qquad \iint_{-\infty}^{+\infty} F(t, s)\, dt\, ds = \iint_{-\infty}^{+\infty} F\left(\xi - \frac{h}{2}, \xi + \frac{h}{2}\right) d\xi\, dh$$

$$= \int_{-\infty}^{+\infty} \left[ \int_{-\infty}^{+\infty} F\left(\xi - \frac{h}{2}, \xi + \frac{h}{2}\right) d\xi \right] dh$$

the last formula being a consequence of Fubini's theorem [8, Vol. 1, p. 631].

Using this transformation, we can write (4.6) in the new form

$$(4.10) \qquad |\alpha_T(\lambda, \omega)|^2 = \frac{1}{T} \int_{-\infty}^{+\infty} e^{i\lambda h} \rho_T(h, \omega)\, dh.$$

Using the fact that $\rho_T(-h, \omega) = \rho_T(h, \omega)$ we have

$$(4.11) \qquad |\alpha_T(\lambda, \omega)|^2 = \frac{2}{T} \int_0^T \rho_T(h, \omega) \cos \lambda h\, dh.$$

Thus we obtain immediately

$$(4.12) \qquad \psi_T(\lambda, \omega) = \frac{2}{\pi} \int_0^T \rho_T(h, \omega) \cos \lambda h\, db = \frac{T}{\pi} |\alpha_T(\lambda, \omega)|^2$$

which gives the expression of the spectrum in terms of the Fourier transform of the truncated sample.

Let us note the following properties of this function

$$(4.13) \qquad \psi_T(\lambda, \omega) \geqq 0,$$

$$(4.14) \qquad \psi_T(-\lambda, \omega) = \psi_T(\lambda, \omega),$$

$$(4.15) \qquad \psi_T(\lambda, \omega) \in L(0, +\infty).$$

Due to this last property we can invert the Fourier transform (4.12) and we obtain the reciprocal formula

$$(4.16) \qquad \rho_T(h, \omega) = \int_0^{+\infty} \psi_T(\lambda, \omega) \cos \lambda h \, d\lambda.$$

**5. Sub-Correlation and Sub-Spectrum of a Random Function.** Let us now define the *sub-correlation* $R_T(h)$ of the random function $f_T(t, \omega)$ as the average of the sample-correlations $\rho_T(h, \omega)$, i.e.,

$$(5.1) \qquad R_T(h) = \overline{\rho_T(h, \omega)}.$$

Obviously, we have

$$(5.2) \qquad R_T(h) = 0, \qquad\qquad\qquad |h| > T.$$

and for all other values of $h$

$$(5.3) \qquad R_T(h) = \int_\Omega \left[ \frac{1}{T} \int_{|h|/2}^{T-|h|/2} f\left(\xi - \frac{h}{2}, \omega\right) f\left(\xi + \frac{h}{2}, \omega\right) d\xi \right] d\mu, \qquad |h| \leq T.$$

Inverting the double integral in accordance with Fubini's theorem we find

$$(5.4) \qquad R_T(h) = \frac{1}{T} \int_{|h|/2}^{T-|h|/2} \Gamma\left(\xi - \frac{h}{2}, \xi + \frac{h}{2}\right) d\xi, \qquad 0 \leq h \leq T.$$

Let us refer to the change of variables (4.7) and let us consider the $(t, s)$ plane (Figure 1). We introduce the following notation

$$(5.5) \qquad \Delta_T = \{(t, s) : 0 \leqq t \leqq T, \quad 0 \leqq s \leqq T\}.$$

Then obviously

$$(5.6) \quad R_T(h) = \frac{1}{T} \times \text{Integral of } \Gamma(t, s) \text{ along the segment } \overline{AB} \text{ of } \delta(h) \text{ contained}$$

in $\Delta_T$.

If we define

$$(5.7) \qquad \begin{aligned} \Gamma_T(t, s) &= \Gamma(t, s), \qquad && (t, s) \in \Delta_T \\ \Gamma_T(t, s) &= 0, \qquad && (t, s) \notin \Delta_T \end{aligned}$$

we can also write (5.4) in the following way

$$(5.8) \qquad R_T(h) = \frac{1}{T} \int_{-\infty}^{+\infty} \Gamma_T\left(\xi - \frac{h}{2}, \xi + \frac{h}{2}\right) d\xi,$$

which is true for all values of $h$, giving $R_T(h) = 0$ for $h < T$ or $h > T$. Obviously,

FIG. 1.—Change of variables.

one has also

(5.9)
$$R_T(h) = \frac{1}{T} \int_{-\infty}^{+\infty} \Gamma_T(\xi, \xi + |h|) \, d\xi$$
$$= \frac{1}{T} \int_{0}^{T-|h|} \Gamma(\xi, \xi + |h|) \, d\xi.$$

Thus, *to compute $R_T(h)$ for all $h$ it is sufficient to know the covariance $\Gamma(t, s)$ in the square $\Delta_T$ .*

Let us define the *sub-spectrum of the random function $f(t, \omega)$* as the average over the spectra of the truncated random function

(5.10)                       $$\varphi_T(\lambda) = \overline{\psi_T(\lambda, \omega)}.$$

Then we have the two reciprocal formulas

(5.11)                   $$R_T(h) = \int_0^{+\infty} \varphi_T(\lambda) \cos \lambda h \, d\lambda$$

(5.12)                   $$\varphi_T(\lambda) = \frac{2}{\pi} \int_0^T R_T(h) \cos \lambda h \, dh.$$

From (4.13) it follows that $\varphi_T(\lambda) \geqq 0$; *thus, by S. Bochner's theorem* [2] *$R_T(h)$ is a continuous correlation function.*

**6. Correlation Function of a Random Function.** We define the *correlation function $R(h)$* of the random function $f(t, \omega)$ by

(6.1)       $$R(h) = \lim_{T \to +\infty} R_T(h) = \lim_{T \to +\infty} \frac{1}{T} \int_{|h|/2}^{T-|h|/2} \Gamma\left(\xi - \frac{h}{2}, \xi + \frac{h}{2}\right) d\xi,$$

*if this limit exists for every real $h$.*

If the random function is stationary (wide sense), then by definition

(6.2)            $$\Gamma(t, s) = \rho(s - t), \qquad \Gamma\left(\xi - \frac{h}{2}, \xi + \frac{h}{2}\right) = \rho(h),$$

and

$$(6.3) \qquad R_T(h) = \left(1 - \frac{|h|}{T}\right)\rho(h), \qquad\qquad |h| \leqq T,$$

$$R_T(h) = 0, \qquad\qquad |h| \geqq T.$$

Thus the limit of $R_T(h)$ exists and

$$(6.4) \qquad R(h) = \lim_{T\to\infty}\left(1 - \frac{|h|}{T}\right)\rho(h) = \rho(h).$$

Hence, for a stationary random function our definition gives the classical result, but we can also apply (6.1) to non-stationary random functions.

Let us consider as an example the random function

$$f(t, \omega) = \frac{W(t)}{\sqrt{t}}$$

where $W(t)$ is the classical Wiener-Lévy function, giving the abscissa at time $t$ of a particle, starting from the origin at time $t = 0$, and subjected to one-dimensional Brownian motion. This function is certainly not stationary; it has the covariance

$$\Gamma(t, s) = \sqrt{\frac{t}{s}}, \qquad\qquad 0 < t \leqq s.$$

$$\Gamma(t, s) = \sqrt{\frac{s}{t}}, \qquad\qquad 0 < s \leqq t.$$

According to our definition this nonstationary random function has the correlation

$$R(h) = \lim_{T\to+\infty}\frac{1}{T}\int_0^{T-|h|}\sqrt{\frac{\xi}{\xi + |h|}}\, d\xi = 1, \qquad \text{for all} \quad h.$$

**7. Spectrum of a Non-Stationary Random Function.** As far as the spectrum is concerned, $\varphi_T(\lambda)$ does not, in general, tend toward a limit when $T \to +\infty$, even if the correlation $R(h)$ exists, but exactly as in the stationary case [10, Vol. 2, p. 164–166] it can be shown, using Paul Lévy's continuity theorem [12, p. 195], that the integrated spectrum

$$(7.1) \qquad S_T(\lambda) = \int_0^\lambda \varphi_T(\eta)\, d\eta$$

does in fact tend toward a limit

$$(7.2) \qquad S(\lambda) = \lim_{T\to+\infty} S_T(\lambda)$$

if the correlation $R(h)$ defined by (6.1) exists and is continuous. (This is not necessarily true; $R(h)$ being the limit of a sequence of continuous functions can be discontinuous.) Thus,

$$(7.3) \qquad R(h) = \int_0^{+\infty} \cos \lambda h\, dS(\lambda),$$

this being a Fourier-Stieltjes integral and the spectrum $S(\lambda)$ being a monotonic

non-decreasing function such that $S(0) = 0$, $S(+\infty) < +\infty$. In general, $S(\lambda)$ is discontinuous and the spectrum has a countable number of *lines* corresponding to a finite amount of energy.

The spectrum $S(\lambda)$ can be computed from the correlation $R(h)$ by applying the Paul Lévy inversion formula [12, p. 166]

$$(7.4) \qquad S(\lambda_2) - S(\lambda_1) = \lim_{A \to +\infty} \frac{2}{\pi} \int_0^A \frac{\sin \lambda_2 h - \sin \lambda_1 h}{h} R(h) \, dh,$$

which is valid if $\lambda_1$ and $\lambda_2$ are continuity points of $S(\lambda)$. When $R(h)$ is known the spectrum is thus defined for every $\lambda > 0$ with the exception of, at most, a countable number of discontinuity points.

**8. Estimation of Correlations for a Non-Stationary Function.** For a non-stationary random function, even if we have not only truncated samples of the function, but also its covariance in $\Delta_T$, this does not give us sufficient information to determine $R(h)$. It is obvious, from (6.1), that large values of $\xi$ are most important in determining $R(h)$ (even at small values of $h$). The knowledge of $\Gamma(t, s)$ in the square $\Delta_T$ only does not give us any information about its values for large $\xi$ on $\delta(h)$ (See Fig. 1).

We shall consider here one class of random functions which is not stationary, but on which information is given, which enables us to make an estimate of $R(h)$.

*This class is defined by the condition that* $\Gamma\left(\xi - \dfrac{h}{2}, \xi + \dfrac{h}{2}\right)$ *is periodic in* $\xi$. This condition means that the covariance is invariant under a translation $\tau$ parallel to the first bissectrix (Fig. 1)

$$(8.1) \qquad \Gamma(t + \tau, s + \tau) = \Gamma(t, s).$$

The scope of the implications of this hypothesis for applications is better understood if one points out that (8.1) is satisfied when the phenomenon represented by the function $f(t, \omega)$ is the result of the superposition of two phenomena, one stationary and the other periodic. Thus

$$(8.2) \qquad f(t, \omega) = f_1(t, \omega) + f_2(t, \omega)$$

where $f_1(t, \omega)$ is a stationary (wide sense) random function:

$$(8.3) \qquad \overline{f_1(t, \omega)f_1(s, \omega)} = \rho_1(t - s)$$

and $f_2(t, \omega)$ is periodic

$$(8.4) \qquad f_2(t + \tau, \omega) = f_2(t, \omega)$$

$$(8.5) \qquad \begin{aligned} \overline{f_2(s, \omega)f_2(t, \omega)} &= \Gamma_2(t, s) \\ \Gamma_2(t + \tau, s) = \Gamma_2(t, s + \tau) &= \Gamma_2(t, s), \end{aligned}$$

the two random functions $f_1$ and $f_2$ being uncorrelated

$$(8.6) \qquad \overline{f_1(t, \omega)f_2(s, \omega)} = 0.$$

Thus one has

$$(8.7) \qquad \Gamma(t, s) = \rho_1(t - s) + \Gamma_2(t, s)$$

and $\Gamma(t, s)$ satisfies (8.1).

In this case the limit (6.1) obviously exists and the correlation is given by

$$(8.8) \qquad R(h) = \frac{1}{\tau} \int_0^\tau \Gamma(\xi, \xi + |h|) \, d\xi.$$

The sub-correlation for $T = N\tau$ (where $N$ is an integer), due to (5.9), is equal to

$$(8.9) \qquad R_{N\tau}(h) = \frac{1}{N\tau} \int_0^{N\tau - |h|} \Gamma(\xi, \xi + |h|) \, d\xi.$$

As a consequence of the periodicity

$$R(h) - R_{N\tau}(h) = \frac{1}{N\tau} \int_0^{|h|} \Gamma(\xi, \xi + |h|) \, d\xi.$$

We thus have the upper bound

$$(8.10) \qquad |R(h) - R_{N\tau}(h)| \leq \frac{1}{N\tau} \int_0^{|h|} |\Gamma(\xi, \xi + |h|)| \, d\xi$$

for the error obtained by using the sub-correlation $R_{N\tau}(h)$ instead of the correlation $R(h)$. The upper bound (8.10) is a function of $h$; however, at a given $h$, this bound tends to 0 as $1/N$.

Let us consider the case when $(k - 1)\tau \leq |h| \leq k\tau$ ($k$ integer). Using the periodicity of $\Gamma$ and Schwarz inequality we have

$$(8.11) \qquad |R(h) - R_{N\tau}(h)| \leq \frac{k}{N} \frac{1}{\tau} \int_0^\tau \Gamma(\xi, \xi) \, d\xi.$$

When the covariance $\Gamma(t, s)$ is known for one period over the diagonal $t = s$ we can compute

$$(8.12) \qquad A = \frac{1}{\tau} \int_0^\tau \Gamma(\xi, \xi) \, d\xi.$$

Finally, the upper bound for the approximation is given by

$$(8.13) \qquad \lceil R(h) - R_{N\tau}(h)| \leq A \frac{k}{N}, \qquad (k - 1)\tau \leq h \leq k\tau.$$

Let us now consider $R_T(h)$ when $N\tau < T < (N + 1)\tau$. We have

$$R_T(h) - \frac{N\tau}{T} R_{N\tau}(h) = \frac{1}{T} \int_{N\tau - |h|}^{T - |h|} \Gamma(\xi, \xi + |h|) \, d\xi$$

$$= \frac{1}{T} \int_{N\tau}^{T} \Gamma(\xi, \xi + |h|) \, d\xi.$$

Thus, due to the periodicity of $\Gamma$,

$$(8.14) \qquad \left. R_T(h) - \frac{N\tau}{T} R_{N\tau}(h) \right| \leq A \frac{\tau}{T} < \frac{A}{N}.$$

From (8.13) and (8.14) we obtain finally

(8.15)
$$\left| R_T(h) - \frac{N\tau}{T} R(h) \right| \leqq A \frac{1+k}{N}, \qquad N\tau < T < (N+1)\tau,$$
$$(k-1)\tau \leqq h < k\tau.$$

This relation shows that *the approximation is good if $k/N$ is small.*

**9. Accuracy of Estimates of Correlations.** In estimating the correlation $R(h)$ of a random function from truncated samples there are two steps:

(a) From the correlations $\rho_T(h, \omega_j)$ of the truncated samples we estimate the sub-correlation $R_T(h)$,

(b) From $R_T(h)$ we compute $R(h)$.

In the preceding section for the case of a random function with a periodic covariance we have solved problem (b) at large values of $T$. Now, let us look at the problem (a), namely, how to evaluate the approximation with which one determines $R_T(h)$ from the average of the correlations $\rho_T(h, \omega_j)$ of a number $q$ of samples

(9.1)
$$\hat{R}_{T,q}(h) = \frac{1}{q} \sum_{j=1}^{j=q} \rho_T(h, \omega_j).$$

It appears that the best way to make such an evaluation is to determine the variance of the random variables $\rho_T(h, \omega_j)$ about their mean value $R_T(h)$; if this variance is small enough we can expect that for a reasonably large number $q$ of samples the estimate $\hat{R}_T(h)$ will be fairly good.

In order to compute this variance, in addition to Hypotheses $\mathbf{H}_1$ to $\mathbf{H}_4$ of Section 2, we shall assume that

$\mathbf{H}_5 . f(t, \omega) \in L^4(\Omega)$ for all $t$'s.

This insures the existence of the fourth-order moment

(9.2)
$$\mathfrak{M}(t_1, t_2, t_3, t_4) = \int_\Omega f(t_1, \omega)f(t_2, \omega)f(t_3, \omega)f(t_4, \omega) \, d\mu$$

for all $[t_1, t_2, t_3, t_4]$.

We shall, moreover, assume that

$\mathbf{H}_6 . \mathfrak{M}(t_1, t_2, t_3, t_4) \in L(\Delta)$ for every finite parallelepiped of the four-dimensional space $(t_1, t_2, t_3, t_4)$.

Let us observe that the fourth-order moment exists in the important case of normal random functions, i.e., when the $n$ random variables, $f(t_1, \omega), \cdots f(t_n, \omega)$ $[t_1, \cdots, t_n]$ arbitrary follow an $n$-variate normal (Gaussian) law.

From (4.3) and (5.8) we find that the departure of a correlation for a particular truncated sample from the mean value taken over the correlations for all samples is given by

$$\rho_T(h, \omega) - R_T(h) = \frac{1}{T} \int_{-\infty}^{+\infty} \left[ \Gamma_T \left( \xi - \frac{h}{2}, \xi + \frac{h}{2} \right) \right.$$
$$\left. - f_T \left( \xi - \frac{h}{2}, \omega \right) f_T \left( \xi + \frac{h}{2}, \omega \right) \right] d\xi.$$

After taking a square and averaging, we obtain

$$(9.3) \qquad \sigma_T(h)^2 = \overline{[\rho_T(h,\omega) - R_T(h)]^2} = \frac{1}{T^2} \iint_{-\infty}^{+\infty} F_T(\xi, \eta, h) \, d\xi \, d\eta$$

with

$$
(9.4) \qquad
\begin{aligned}
F_T(\xi, \eta, h) = {}& \mathfrak{M}_T\left(\xi - \frac{h}{2}, \xi + \frac{h}{2}, \eta - \frac{h}{2}, \eta + \frac{h}{2}\right) \\
& - \Gamma_T\left(\xi - \frac{h}{2}, \xi + \frac{h}{2}\right) \Gamma_T\left(\eta - \frac{h}{2}, \eta + \frac{h}{2}\right).
\end{aligned}
$$

In the above equation $\Gamma_T(t, s)$ is defined by (5.7) and $\mathfrak{M}_T$ by the two relations

$$
(9.5) \qquad
\begin{aligned}
\mathfrak{M}_T(t_1, t_2, t_3, t_4) = {}& \mathfrak{M}(t_1, t_2, t_3, t_4) \\
& \text{when} \quad (t_1, t_2) \in \Delta_T \quad \text{and} \, (t_3, t_4) \in \Delta_T
\end{aligned}
$$

and

$$(9.6) \quad \mathfrak{M}_T(t_1, t_2, t_3, t_4) = 0 \quad \text{when} \quad (t_1, t_2) \notin \Delta_T \quad \text{or} \quad (t_3, t_4) \notin \Delta_T.$$

Equation (9.4) shows that whenever the covariance $\Gamma$ and the fourth-order moment $\mathfrak{M}$ are known, we can evaluate the variance, $\sigma_T(h)^2$, of the random variables $\rho_T(h, \omega)$ about their mean value $R_T(h)$. In particular, for $h = 0$, we have

$$(9.7) \qquad \sigma_T(0)^2 = \frac{1}{T^2} \iint_{\Delta_T} [\mathfrak{M}(\xi, \xi, \eta, \eta,) - \Gamma(\xi, \xi)\Gamma(\eta, \eta)] \, d\xi \, d\eta.$$

**10. Random Step- Function with Periodic Covariance.** As an example let us take the random function defined as follows*

$$(10.1) \qquad\qquad V(t, \omega) = X_n(\omega) \quad n - 1 \leq t < n, n = 1, 2, \cdots$$

where

$$(10.2) \qquad\qquad X_1(\omega), \cdots, X_n(\omega), \cdots$$

is a sequence of independent random variables, taking only the values $-1$ and $+1$ with equal probability

$$(10.3) \qquad\qquad \text{Prob}\,[X_n = -1] = \text{Prob}\,[X_n = +1] = \tfrac{1}{2}.$$

The random function $V(t, \omega)$ is essentially the same as a function considered by Norbert Wiener in his pioneering work on correlation and spectrum, [15, p. 151] except for the fact that his function is defined for $-\infty < t < +\infty$ and ours only in $[0, +\infty]$.

From (10.3) it follows that

$$(10.4) \qquad\qquad \bar{X}_n = 0$$

$$(10.5) \qquad\qquad \overline{X_m X_n} = \delta_{m,n},$$

* As measure space $\Omega$, one can take the interval $[0, 1]$, $0 \leq \omega < 1$, with $X_n(\omega) = 2a_n - 1$, where $a_n$ is the $n$th coefficient in the binary development $\omega = a_1/2 + \cdots + a_n/2^n + \cdots$. The measure $\mu$ on $[0, 1]$ is the Lebesgue measure $\mu\{\omega: a_n = 1\} = \mu\{\omega: a_n = 0\} = \tfrac{1}{2}$.

where $\delta_{m,n}$ is the Kronecker symbol. If one defines

$$\Delta_{m,n} = \{(t, s) : m - 1 \leqq t < m, n - 1 \leqq s < n\}$$

then

$$\Gamma(t, s) = 1, \qquad (t, s) \in \bigcup_{n=1}^{n=+\infty} \Delta_{n,n},$$

$$\Gamma(t, s) = 0, \qquad (t, s) \notin \bigcup_{n=1}^{n=+\infty} \Delta_{n,n}.$$

Obviously, the random function $V(t, \omega)$ is not stationary, but its covariance, when put in the form $\Gamma\left(\xi - \dfrac{h}{2}, \xi + \dfrac{h}{2}\right)$ is periodic, with period 1 in $\xi$ (see Fig. 2). Thus the results of Section 8 apply, and the non-stationary random function $V(t, \omega)$ has, according to (8.8), a correlation

$$\text{(10.6)} \qquad \begin{aligned} R(h) &= (1 - |h|), & |h| &\leqq 1, \\ R(h) &= 0, & |h| &\geqq 1. \end{aligned}$$

Let us first consider the case $T = N$, where $N$ is an integer. We find that

$$\rho_N(h, \omega) = (k - |h|)Y_N(k - 1, \omega) + (1 - k + |h|)Y_N(k, \omega),$$

$$\text{(10.7)} \qquad k - 1 \leqq |h| \leqq k, \qquad 1 \leqq k \leqq N - 1,$$

$$\rho_N(h, \omega) = 0, \qquad |h| \geqq N,$$

where $Y_N(0, \omega), Y_N(1, \omega) \cdots Y_N(N - 1, \omega)$ represent the random variables

$$\text{(10.8)} \qquad Y_N(k, \omega) = \frac{1}{N} \sum_{j=1}^{j=N-k} X_j(\omega)X_{j+k}(\omega), \quad k = 0, 1, \cdots (N - 1)$$
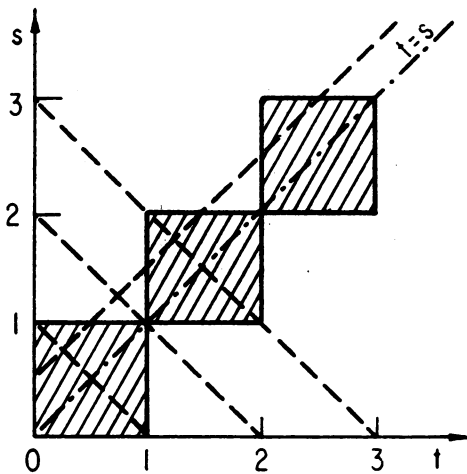


Fig. 2.—The $(t, s)$ diagram for a random step-function. $\Gamma(\xi - \frac{h}{2}, \xi + \frac{h}{2})$ is periodic in $\xi$ and for a translation parallel to $t = s$, $\Gamma(t + 1, s + 1) = \Gamma(t, s)$.

in particular

(10.9) $$Y_N(0, \omega) = 1.$$

For each sample $\omega$, the sample-correlation $\rho_N(h, \omega)$ is represented by a polygonal line (Fig. 3). The ordinates $Y_N(1, \omega)$, $Y_N(2, \omega) \cdots Y_N(N - 1, \omega)$ corresponding to $h = 1, 2, \cdots (N - 1)$, are the random variables defined by (10.8). We see immediately that $Y_N(k, \omega)$ can only take the values

$$-\frac{N - k}{N}, \qquad -\frac{N - k - 2}{N}, \qquad \cdots +\frac{N - k - 2}{N}, \qquad +\frac{N - k}{N}$$

following the binomial law. As a consequence

(10.10) $$| \rho_N(k, \omega) | \leqq 1 - \frac{k}{N}, \qquad k = 1, 2, \cdots, (N - 1).$$

We obtain for the mean, variances, and covariances of the ordinates $Y_N(k, \omega)$ respectively,

(10.11) $$\overline{Y_N(k, \omega)} = 0 \qquad k = 1, 2, \cdots, (N - 1)$$

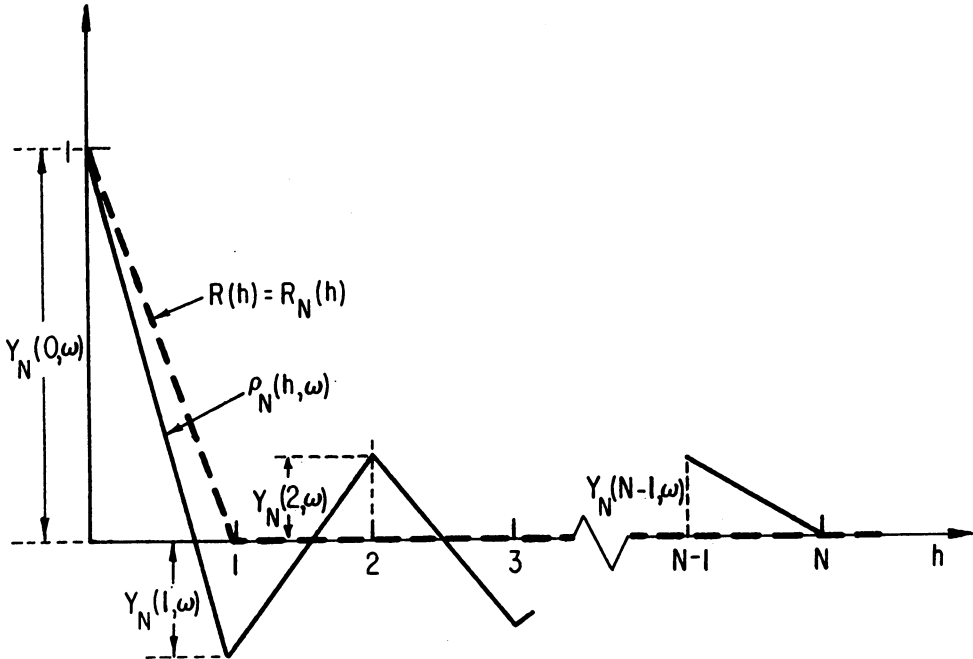(10.12) $$\overline{Y_N(k, \omega)^2} = \frac{N - k}{N^2},$$



FIG. 3.—The polygonal line $\rho_N(h, \omega)$ represents a sample-correlation of the step-function. $R_N(h) = \overline{\rho_N(h, \omega)}$ is the sub-correlation, which for the step-function, is equal to the correlation function $R(h)$.

and

(10.13) $$\overline{Y_N(k, \omega)Y_N(l, \omega)} = 0, \qquad\qquad k \neq l.$$

From (10.7), (10.9) and (10.11) we deduce

(10.14) $$\overline{\rho_N(h, \omega)} = 1 - |h|, \qquad\qquad |h| \leqq 1$$
$$\overline{\rho_N(h, \omega)} = 0, \qquad\qquad |h| \geqq 1.$$

Thus

(10.15) $$R_N(h) = \overline{\rho_N(h, \omega)} = R(h).$$

Let us next suppose $N \leqq T < N + 1$. We find

(10.16) $$\rho_T(h, \omega) = \frac{N}{T}\rho_N(h, \omega) + \frac{k - N + h}{T}X_k X_{N+1} + \frac{T - k - h}{T}X_{k+1}X_N,$$
$$N - k < h < T - k, \qquad k = 1, 2, \cdots N.$$

(10.17) $$\rho_T(h, \omega) = \frac{N}{T}\rho_N(h, \omega) + \frac{T - N}{T}X_k X_{N+1},$$
$$T - k < h < N + 1 - k, \qquad k = 1, 2, \cdots N.$$

Thus if $N < T < N + 1$,

(10.18) $$|\rho_T(h, \omega) - \rho_N(h, \omega)| \leqq \left(1 - \frac{N}{T}\right)|\rho_N(h, \omega)| + \left(1 - \frac{N}{T}\right).$$

As a consequence of (10.10), we have

(10.19) $$|\rho_T(h, \omega) - \rho_N(h, \omega)| \leqq \frac{2}{N + 1} \quad \text{for} \quad N \leqq T < N + 1.$$

For large values of $N$ the right-hand side of this inequality is as small as we want. Thus we can always, in studying $\rho_T(h, \omega)$, suppose that $T$ has an integer value $N$.

Let us now compute the variance $\sigma_N(h)^2$ of the sample-correlations about their mean value. We will make the computation only for $T = N$. Due to (10.14) one has

$$\sigma_N(h)^2 = \overline{\rho_N(h, \omega)^2}.$$

From (10.7), (10.12) and (10.13) we obtain

(10.20) $$\overline{\rho_N(h, \omega)^2} = (k - |h|)^2 \frac{N - k + 1}{N^2}$$
$$+ (1 - k + |h|)^2 \frac{N - k}{N^2}, \quad k - 1 \leqq |h| \leqq k, \ k = 1, 2, \cdots N - 1.$$

In particular for $h = k$

(10.21) $$\sigma_N(k)^2 = \overline{\rho_N(k, \omega)^2} = \frac{(N - k)}{N^2}.$$

The proposition that the *random variables $\rho_N(h, \omega)$ tend toward their limit $R(h)$ for almost all samples,*

$$\text{Prob}\left[\lim_{N \to +\infty} \rho_N(h, \omega) = R(h)\right] = 1,$$

will now be established. First from (10.8) we compute

(10.21) $$\overline{Y_N(k, \omega)^4} = \frac{(N - k)(2N - 2k + 1)}{N^4}.$$

Let us now observe that for a fixed $N$, when $k$ takes the values $1, 2, \cdots (N - 1)$ one has $\overline{Y_N(k, \omega)^4} < \dfrac{3}{N^2}$. Thus, $k$ now being fixed, we have

$$\sum_{N=1}^{N=+\infty} \overline{Y_N(k, \omega)^4} \leqq 3 \sum_{N=1}^{N=+\infty} \frac{1}{N^2} < + \infty$$

and, therefore

$$\text{Prob} \left[ \sum_{N=1}^{N=+\infty} Y_N(k, \omega)^4 < + \infty \right] = 1.$$

Here, we use the following criterion for the almost sure absolute convergence for a series of random variables. If $\sum_1^{+\infty} \overline{|X_n|} < + \infty$, then Prob $[\sum_1^{+\infty} |X_n| < + \infty]$ $= 1$. This criterion applies even when the random variables $X_1, \cdots, X_2, \cdots$ are not independent. However, the convergence of the series implies that $Y_N(k, \omega) \rightarrow 0$. We have thus proved that

$$\text{Prob} \left[ \lim_{N \rightarrow +\infty} Y_N(k, \omega) = 0 \right] = 1$$

for each fixed $k \geqq 1$.

Let us now take a fixed interval $|h| \leqq M$, where $M$ is an integer. In this interval, there are exactly $M$ points, $h = 1, h = 2, \cdots h = M$, which completely determine the polygonal line. Each of the $M$ ordinates $Y_N(1, \omega), Y_N(2, \omega) \cdots Y_N(M, \omega)$ tends toward 0 with probability one. Their number being finite, this evidently implies

$$\text{Prob} \left[ \lim_{N \rightarrow +\infty} Y_N(k, \omega) = 0 \quad \text{for} \quad k = 1, 2, \cdots M \right] = 1.$$

We have thus proved that

(10.22) $$\text{Prob} \left[ \lim_{N \rightarrow +\infty} \rho_N(h, \omega) = R(h) \right] = 1$$

in the finite interval $|h| \leqq M$ and the proof is complete because $M$ could be taken arbitrarily large.

## 11. Correlation Estimates for a Continuous Step-Function Constructed Using Sequences of Random Numbers.

A continuous step-function has been constructed using a table of random numbers which are listed in 100 groups of 1000 digits each [11]. For our analysis we selected the 200 first digits of each of the 100 groups and divided them into five consecutive segments of 40 digits. All the even digits were then replaced by $+1$ and the odd digits by $-1$. We thus obtain one set of 500 sequences of 40 digits $(+1, \text{ or } -1)$. Other sets were constituted by combining two or more consecutive segments of 40 digits. As a result we have obtained the following five sets of experimental functions without overlapping sequences within each set: 500 sequences of 40 digits; 200 sequences of 80 digits; 100 sequences of 120, 160 and 200 digits, respectively.

Let us consider a set of $q$ sequences with $N$ digits in each sequence. A sample-correlation for the sequence is determined for $T = N$ by (10.8). The average taken over the $q$ sample-correlations of the set determines the estimate for the sub-correlation, cf. equation (9.1).

$$(11.1) \qquad \hat{R}_{N,q}(k) = \frac{1}{q} \sum_{\omega_i=1}^{q} \rho_N(k, \omega_i).$$

In a sequence of $N$ digits there are $N - k$ products $X_j(\omega)X_{j+k}(\omega)$ equal to either $+1$ or $-1$. If $n_N(k, \omega)$ is the number of products equal to $+1$ then

$$(11.2) \qquad \rho_N(k, \omega) = \frac{1}{N}[N - k - 2n_N(k, \omega)]$$

and

$$(11.3) \qquad \hat{R}_{N,q}(k) = \frac{N - k}{k} - \frac{2}{Nq} \sum_{\omega_i=1}^{q} n_N(k, \omega_i).$$

As it has been shown in Section 10 a sample-correlation $\rho_N(h, \omega)$ is represented by a polygonal line. The vertices of this polygonal line correspond to $h = k = 0, 1, 2, \cdots, (N - 1)$ and are given by (11.2). Similarly, the estimate $\hat{R}_{N,q}(k)$, for the sub-correlation is represented by a polygonal line which is determined by (11.3). It is, therefore, sufficient to determine the values of $\rho_N(k, \omega)$ and $\hat{R}_{N,q}(k)$ at the vertices of the polygonal lines to have the corresponding sample-correlations $\rho_N(h, \omega)$ and the estimate for the sub-correlation $\hat{R}_{N,q}(h)$ for the continuous step-function. Figure 4 illustrates several examples of sample-correlation $\rho_N(h, \omega)$ for individual sequences of digits. Numerical data for $\hat{R}_{N,q}(k)$, for the five sets of sequences, are listed in Table I and a few of them are illustrated on Figure 4. The
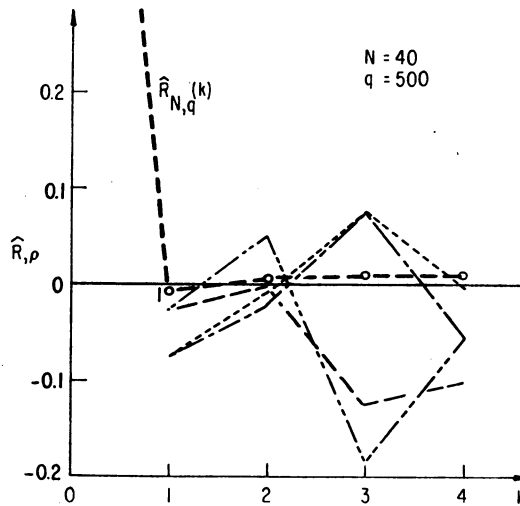


FIG. 4.—Experimental examples of sample-correlations $\rho_N(k, \omega)$ obtained for a random sequence of digits $-1, +1$, are represented using light interrupted lines. The heavy polygonal line represents the estimate of the sub-correlation $R_{N,q}(k)$ obtained by averaging over $q = 500$ sample-correlations.

theoretical sub-correlation $R_N(h)$ is in the present case equal to the correlation $R(h)$ (see (10.15)) and, therefore, the experimental results for $\hat{R}_{N,q}(k)$ should be compared to the theoretical value

$$(11.4) \qquad R_N(k) = R(k) = 0, \qquad\qquad k \geqq 1.$$

The departure of the value for a sample-correlation from the estimated sub-correlation, as determined from $q$ samples, is given by

$$\rho_N(k, \omega) - \hat{R}_{N,q}(k).$$

The variance $\sigma_{N,q}(k)^2$ for such departures is given by the relation

$$\sigma_{N,q}(k)^2 = \frac{1}{q}\sum_1^q [\rho_N(k, \omega) - \hat{R}_{N,q}(k)]^2$$

(11.5)

$$= \frac{4}{N^2q}\sum_1^q \left[ -n_N(k, \omega) + \frac{1}{q}\sum_1^q n_N(k, \omega) \right]^2.$$

Numerical results for the experimental variance $\sigma_{N,q}(k)^2$ are listed in Table II and compared with the theoretical values, $\sigma_N(k)^2$, computed by using the relation (10.21).

TABLE I

*Estimates for the Sub-Correlations of Sequences of Random Numbers*

| $N$ | 40 | 80 | 120 | 160 | 200 |
|---|---|---|---|---|---|
| $q$ | 500 | 200 | 100 | 100 | 100 |
| $\hat{R}_{N,q}(1)$ | −0.0041 | −0.0026 | −0.0047 | −0.0023 | −0.0053 |
| $\hat{R}_{N,q}(2)$ | +0.0053 | +0.0040 | +0.0143 | +0.0033 | +0.0054 |
| $\hat{R}_{N,q}(3)$ | +.0066 | +0.0071 | +0.0087 | +0.0081 | +0.0073 |
| $\hat{R}_{N,q}(4)$ | +0.0083 | +0.0036 | +0.0095 | +0.0061 | +0.0072 |

TABLE II

*Comparison between Experimental Variances $\sigma_{N,q}(k)$ and Theoretical*
*Variances $\sigma_N(k)$*

| $N$ | 40 | 80 | 120 | 160 | 200 |
|---|---|---|---|---|---|
| $q$ | 500 | 200 | 100 | 100 | 100 |
| $\sigma_{N,q}(1)^2$ | 0.0253 | 0.0125 | 0.0094 | 0.0074 | 0.0066 |
| $\sigma_N(1)^2$ | 0.0244 | 0.0123 | 0.0083 | 0.0062 | 0.0050 |
| $\sigma_{N,q}(2)^2$ | 0.0231 | 0.0125 | 0.0076 | 0.0061 | 0.0048 |
| $\sigma_N(2)^2$ | 0.0238 | 0.0122 | 0.0082 | 0.0062 | 0.0050 |
| $\sigma_{N,q}(3)^2$ | 0.0232 | 0.0121 | 0.0074 | 0.0059 | 0.0043 |
| $\sigma_N(3)^2$ | 0.0231 | 0.0120 | 0.0081 | 0.0061 | 0.0049 |
| $\sigma_{N,q}(4)^2$ | 0.0198 | 0.0089 | 0.0065 | 0.0038 | 0.0035 |
| $\sigma_N(4)^2$ | 0.0225 | 0.0119 | 0.0081 | 0.0061 | 0.0049 |

**15. Frequency Distributions of Sample Correlations.** Let us denote as

$$q\mathcal{F}_q[n_N(k, \omega) = u]$$

the number of sequences of a set for which $n_N(k, \omega)$ is equal to $u$. We find then, for the *frequency distribution of sample-correlations*

(12.1)               $$\mathcal{F}_q[\rho_N(k, \omega) = U] = \mathcal{F}_q[n_N(k, \omega) = u],$$

where

$$U = \frac{1}{N}(N - k - 2u).$$

Numerical values for the frequency distribution of sample-correlations can easily be obtained from Table III where $q\mathcal{F}_q[n_N(k, \omega) = u]$ is listed for the five sets of sequences. It may be noted here that (11.3) can also be written as

(12.2)               $$\hat{R}_{N,q}(k) = \frac{N - k}{N} - \frac{2}{N}\sum_{u=0}^{N-k} u\mathcal{F}_q[n_N(k, \omega) = u],$$

which is often more convenient for numerical computations.

TABLE III
*Frequency Distributions of Sample-Correlations $\mathcal{F}_q(\rho_N = U) = \mathcal{F}_q(n_N = u)$*

| $n_{40}$ | 500 $\mathcal{F}_{500}$ ($n_{40} = u$) | | | | $n_{80}$ | 200 $\mathcal{F}_{200}$ ($n_{80} = u$) | | | | $n_{120}$ | 100 $\mathcal{F}_{100}$ ($n_{120} = u$) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $k=1$ | $k=2$ | $k=3$ | $k=4$ | | $k=1$ | $k=2$ | $k=3$ | $k=4$ | | $k=1$ | $k=2$ | $k=3$ | $k=4$ |
| 9 | 0 | 0 | 0 | 1 | 24 | 0 | 0 | 1 | 0 | 45 | 0 | 1 | 0 | 0 |
| 10 | 0 | 0 | 0 | 1 | 25 | 0 | 0 | 0 | 1 | 46 | 1 | 0 | 2 | 2 |
| 11 | 4 | 3 | 2 | 3 | 26 | 0 | 0 | 0 | 0 | 47 | 0 | 1 | 2 | 0 |
| 12 | 3 | 3 | 2 | 12 | 27 | 0 | 1 | 1 | 0 | 48 | 1 | 2 | 1 | 1 |
| 13 | 8 | 16 | 10 | 19 | 28 | 0 | 1 | 2 | 0 | 49 | 2 | 2 | 1 | 0 |
| 14 | 14 | 13 | 18 | 23 | 29 | 2 | 1 | 1 | 0 | 50 | 2 | 0 | 3 | 1 |
| 15 | 23 | 33 | 27 | 49 | 30 | 0 | 3 | 3 | 1 | 51 | 1 | 3 | 0 | 2 |
| 16 | 30 | 45 | 26 | 44 | 31 | 4 | 3 | 2 | 3 | 52 | 3 | 4 | 3 | 7 |
| 17 | 49 | 53 | 46 | 61 | 32 | 4 | 8 | 7 | 8 | 53 | 6 | 5 | 4 | 7 |
| 18 | 50 | 51 | 55 | 74 | 33 | 9 | 7 | 12 | 11 | 54 | 2 | 3 | 4 | 7 |
| 19 | 58 | 66 | 69 | 72 | 34 | 5 | 12 | 14 | 13 | 55 | 7 | 7 | 10 | 9 |
| 20 | 56 | 70 | 68 | 65 | 35 | 13 | 14 | 12 | 19 | 56 | 5 | 13 | 8 | 8 |
| 21 | 68 | 57 | 59 | 38 | 36 | 18 | 15 | 13 | 18 | 57 | 6 | 7 | 9 | 11 |
| 22 | 51 | 29 | 46 | 18 | 37 | 11 | 13 | 15 | 18 | 58 | 7 | 7 | 9 | 10 |
| 23 | 38 | 32 | 28 | 6 | 38 | 16 | 15 | 19 | 18 | 59 | 5 | 7 | 9 | 9 |
| 24 | 22 | 9 | 20 | 8 | 39 | 16 | 20 | 20 | 29 | 60 | 7 | 6 | 8 | 4 |
| 25 | 11 | 12 | 13 | 5 | 40 | 18 | 10 | 20 | 23 | 61 | 6 | 8 | 2 | 4 |
| 26 | 7 | 4 | 8 | 0 | 41 | 14 | 11 | 15 | 6 | 62 | 5 | 7 | 5 | 3 |
| 27 | 2 | 3 | 0 | 1 | 42 | 19 | 21 | 12 | 14 | 63 | 4 | 1 | 6 | 4 |
| 28 | 6 | 1 | 3 | 0 | 43 | 16 | 12 | 8 | 5 | 64 | 9 | 4 | 4 | 4 |
| | | | | | 44 | 6 | 11 | 5 | 5 | 65 | 6 | 2 | 3 | 2 |
| | | | | | 45 | 8 | 12 | 7 | 0 | 66 | 6 | 3 | 5 | 0 |
| | | | | | 46 | 6 | 4 | 6 | 4 | 67 | 1 | 3 | 0 | 0 |
| | | | | | 47 | 8 | 3 | 3 | 1 | 68 | 0 | 2 | 0 | 0 |
| | | | | | 48 | 3 | 2 | 1 | 1 | 69 | 3 | 1 | 0 | 2 |
| | | | | | 49 | 2 | 0 | 0 | 0 | 70 | 2 | 0 | 0 | 2 |
| | | | | | 50 | 1 | 0 | 0 | 2 | 71 | 0 | 0 | 0 | 1 |
| | | | | | 51 | 0 | 1 | 1 | 0 | 72 | 2 | 0 | 1 | 0 |
| | | | | | 52 | 0 | 0 | 0 | 0 | 73 | 0 | 0 | 0 | 0 |
| | | | | | 53 | 0 | 0 | 0 | 0 | 74 | 0 | 1 | 1 | 0 |
| | | | | | 54 | 1 | 0 | 0 | 0 | 75 | 0 | 0 | 0 | 0 |
| | | | | | | | | | | 76 | 0 | 0 | 0 | 0 |
| | | | | | | | | | | 77 | 1 | 0 | 0 | 0 |

TABLE III (*Continued*)

| $n_{160}$ | $100\,\mathfrak{F}_{100}(n_{160}=u)$ | | | | $n_{200}$ | $100\,\mathfrak{F}_{100}(n_{200}=u)$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $k=1$ | $k=2$ | $k=3$ | $k=4$ | | $k=1$ | $k=2$ | $k=3$ | $k=4$ |
| 60 | 0 | 0 | 1 | 0 | 81 | 0 | 0 | 1 | 0 |
| 61 | 0 | 0 | 1 | 0 | 82 | 0 | 0 | 1 | 0 |
| 62 | 0 | 0 | 0 | 0 | 83 | 1 | 1 | 1 | 1 |
| 63 | 1 | 0 | 0 | 0 | 84 | 2 | 1 | 0 | 1 |
| 64 | 0 | 0 | 0 | 0 | 85 | 1 | 1 | 1 | 2 |
| 65 | 0 | 0 | 0 | 0 | 86 | 1 | 2 | 2 | 1 |
| 66 | 2 | 2 | 0 | 2 | 87 | 0 | 3 | 3 | 0 |
| 67 | 0 | 4 | 2 | 0 | 88 | 1 | 1 | 1 | 0 |
| 68 | 2 | 1 | 1 | 1 | 89 | 3 | 4 | 2 | 4 |
| 69 | 2 | 1 | 1 | 2 | 90 | 4 | 2 | 3 | 3 |
| 70 | 0 | 3 | 5 | 3 | 91 | 2 | 4 | 1 | 5 |
| 71 | 1 | 4 | 3 | 1 | 92 | 1 | 2 | 1 | 4 |
| 72 | 5 | 4 | 5 | 3 | 93 | 5 | 2 | 4 | 6 |
| 73 | 6 | 1 | 3 | 8 | 94 | 8 | 5 | 9 | 3 |
| 74 | 5 | 3 | 6 | 4 | 95 | 4 | 2 | 4 | 8 |
| 75 | 3 | 7 | 8 | 8 | 96 | 1 | 4 | 3 | 5 |
| 76 | 7 | 4 | 6 | 14 | 97 | 5 | 7 | 10 | 12 |
| 77 | 6 | 6 | 6 | 9 | 98 | 5 | 8 | 8 | 8 |
| 78 | 4 | 4 | 9 | 7 | 99 | 6 | 5 | 5 | 4 |
| 79 | 6 | 7 | 5 | 8 | 100 | 2 | 5 | 8 | 2 |
| 80 | 9 | 12 | 8 | 6 | 101 | 5 | 9 | 3 | 4 |
| 81 | 4 | 7 | 1 | 3 | 102 | 7 | 4 | 5 | 5 |
| 82 | 6 | 5 | 5 | 6 | 103 | 3 | 7 | 7 | 6 |
| 83 | 1 | 4 | 6 | 4 | 104 | 5 | 4 | 2 | 7 |
| 84 | 8 | 3 | 5 | 3 | 105 | 5 | 2 | 2 | 3 |
| 85 | 4 | 4 | 4 | 2 | 106 | 1 | 3 | 3 | 0 |
| 86 | 3 | 2 | 0 | 2 | 107 | 4 | 3 | 1 | 0 |
| 87 | 2 | 3 | 2 | 0 | 108 | 3 | 1 | 5 | 1 |
| 88 | 3 | 4 | 2 | 1 | 109 | 2 | 3 | 1 | 4 |
| 89 | 1 | 1 | 2 | 0 | 110 | 2 | 1 | 1 | 0 |
| 90 | 3 | 0 | 1 | 1 | 111 | 1 | 0 | 1 | 0 |
| 91 | 0 | 2 | 1 | 2 | 112 | 3 | 2 | 1 | 0 |
| 92 | 1 | 1 | 0 | 0 | 113 | 1 | 0 | 0 | 1 |
| 93 | 0 | 0 | 0 | 0 | 114 | 0 | 1 | 0 | 0 |
| 94 | 2 | 0 | 1 | 0 | 115 | 2 | 1 | 0 | 0 |
| 95 | 1 | 1 | 0 | 0 | 116 | 0 | 0 | 0 | 0 |
| 96 | 1 | 0 | 0 | 0 | 117 | 2 | 0 | 0 | 0 |
| 97 | 0 | 0 | 0 | 0 | 118 | 1 | 0 | 0 | 0 |
| 98 | 1 | 0 | 0 | 0 | 119 | 1 | 0 | 0 | 0 |

The theoretical probability distribution of sample-correlations is given by the relation

(12.3)
$$\text{Prob}\,[\rho_N(k,\omega)=U]=\frac{1}{2^{N-k}}\binom{N-k}{u},\quad u=0,1,\cdots(N-k)$$

where

$$u=\frac{N-k}{2}-\frac{NU}{2}.$$

For large values of $(N-k)$, this binomial distribution can be approximated by the Gaussian distribution

(12.4)
$$\text{Prob}\,[\rho_N(k,\omega)=U]\approx\frac{2}{\sqrt{2\pi(N-k)}}$$
$$\cdot\exp\left[-\frac{(N-k-2u)^2}{2(N-k)}\right],\qquad u=0,1,\cdots N-k.$$
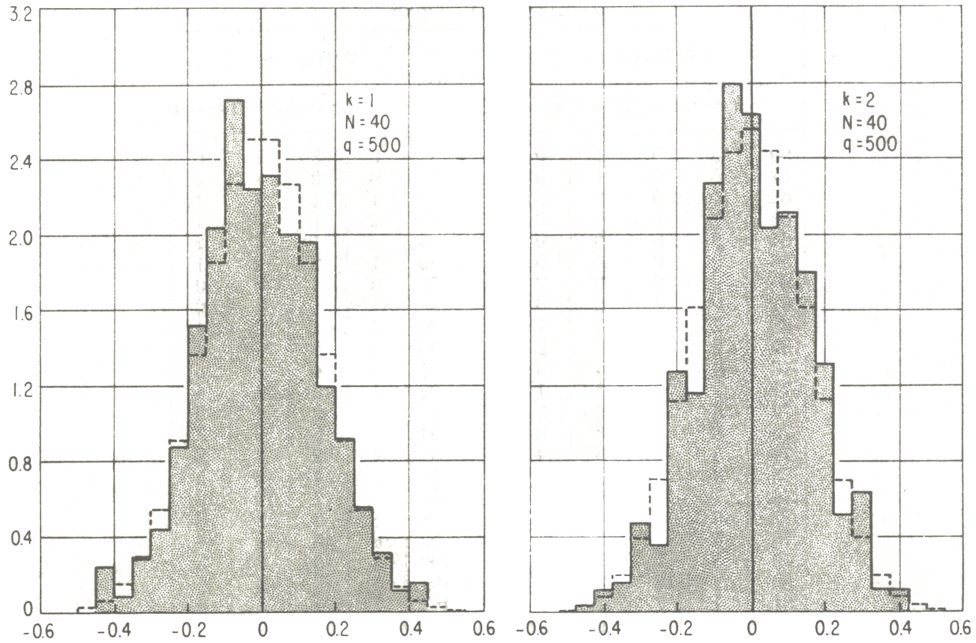
Fig. 5.—Experimental frequency distributions of sample-correlations, $\mathcal{F}_q[\rho_N(k, \omega) = U]$ (full line) are compared with the theoretical binomial frequency distribution (dashed lines).

In particular, the probability of $\rho_N(k, \omega)$ being equal to the sub-correlation

$$R_N(k) = \overline{\rho_N(k, \omega)} = 0, \qquad\qquad k \geqq 1$$

(which, in the present case, is also equal to the correlation $R(k)$) is given by

$$(12.5) \qquad \mathrm{Prob}\,[\rho_N(k, \omega) = 0] = \frac{1}{2^{N-k}} \binom{N-k}{\frac{1}{2}(N-k)}, \qquad k \geqq 1$$

with the approximated formula

$$(12.6) \qquad \mathrm{Prob}\,[\rho_N(k, \omega) = 0] \approx \frac{2}{\sqrt{2\pi(N-k)}}, \qquad k \geqq 1.$$

In Figure 5 we are comparing some of the experimental frequency distributions with these theoretical results.

Faculté des Sciences
University of Lille
Lille, France

Applied Mathematics Laboratory
David Taylor Model Basin
Washington 7, D. C.

1. R. B. BLACKMAN & J. W. TUKEY, *The Measurement of Power Spectra*, Dover Publications, Inc., New York, 1958.
2. S. BOCHNER, *Vorlesungen über Fouriersche Integrale*, Akad-Verlag, Leipzig, 1932. (Reprint: Chelsea, New York, 1948).
3. M. M. CRUM, "On positive-definitive functions," *Proc. London Math. Soc.*, v. 6, 1956, p. 548–560.

4. J. L. Doob, *Stochastic Processes*, John Wiley and Sons, New York, 1953.

5. F. N. Frenkiel & J. Kampé de Fériet, *Proceedings of the International Congress of Mathematicians*, Nordhoff, Groningen, v. 2, 1954, p. 291.

6. U. Grenander & M. Rosenblatt, *Statistical Analysis of Stationary Time Series*, John Wiley and Sons, New York, 1957.

7. P. R. Halmos, *Measure Theory*, D. Van Nostrand Company, New York, 1950.

8. E. W. Hobson, *The Theory of Functions of a Real Variable*, Dover Publications, New York, 1957, v. 1 and 2.

9. J. Kampé de Fériet & F. N. Frenkiel, "Estimation de la corrélation d'une fonction aléatoire non stationnaire," *Comptes Rendus, Acad. Sciences, Paris*, v. 249, 1959, p. 348–351.

10. J. Kampé de Fériet, "Introduction to the statistical theory of turbulence," *J. Soc. Indust. Appl. Math.*, v. 2, 1954, p. 1–9, 143–174, 244–271; v. 3, 1955, p. 90–117.

11. M. G. Kendall & B. B. Smith, *Random Sampling Numbers*, Cambridge University Press, 1939.

12. Paul Lévy, *Calcul des Probabilités*, Gauthier-Villars, Paris, 1925.

13. E. Parzen, "On consistent estimates of the spectrum of a stationary time series," *Ann. Math. Statist.*, v. 28, 1957, p. 329–348.

14. E. Parzen, "On choosing an estimate of the spectral density function of a stationary time series," *Ann. Math. Statist.*, v. 28, 1957, p. 921–932.

15. N. Wiener, *The Fourier Integral and Certain of its Applications*, Cambridge University Press, 1933.