HYPOTHESIS TESTING AND ESTIMATION FOR LAPLACIAN FUNCTIONS

R. FORTET
UNIVERSITY OF PARIS

1. Introduction

During the last few years, several papers have been devoted to the study of random functions. And though a large amount of work remains to be done in this field, it presents some difficulties. If we try to apply in the field of random functions—or, more generally, of random elements—some of the basic notions of classical mathematical statistics such as sufficient statistics or maximum likelihood, we find that conditional probabilities or probability densities do not obviously exist, that sets and spaces are not compact or even locally compact, and so on. The existence of conditional probabilities is a particularly important point. I emphasize that, in this paper, by "conditional probability" I always mean a "regular conditional probability," that is, with the *complete additivity* property.

Concerning this existence of conditional probabilities, a very important advance has been made by M. Jiřina [4]. Among the more general results given in this paper, there is the following statement.

THEOREM 1.1. *Let $\mathfrak{X}$ be a metric, separable, complete space of elements x, let S be the smallest σ-algebra of subsets of $\mathfrak{X}$ containing the spheres (or the Borel sets) of $\mathfrak{X}$, let $m(e)$ be a probability measure on $(\mathfrak{X}, S)$, that is to say, a function of the set e, defined for $e \in S$, which is nonnegative and completely additive on S, with $m(\mathfrak{X}) = 1$, and let Σ be any σ-algebra $\subset S$. Then there exists, associated with $m(e)$, at least one conditional probability $\mu(x; e)$ on S, relative to Σ, having the following properties.*

(a) *It is a nonnegative function of $x \in \mathfrak{X}$ and of $e \in S$, which, for every fixed $e \in S$, is Σ-measurable as a function of x.*

(b) *For every fixed x, it is a probability measure on S as a function of e, including complete additivity and $\mu(x; \mathfrak{X}) = 1$.*

(c) *For every $A \in \Sigma$ and every $e \in S$,*

$$m(A \cap e) = \int_A \mu(x; e) m\,(dx). \tag{1.1}$$

Jiřina has completed the preceding results in [5]. It must be pointed out that in his paper the condition that $\mu(x; e)$ is Σ-measurable as a function of x, is weakened to $\mu(x; e)$ is $\bar{\Sigma}$-measurable, $\bar{\Sigma}$ being the completion of Σ.

A particular application of theorem 1.1 is

THEOREM 1.2. *Let $y = y(x)$ be an S-measurable real-valued function on $\mathfrak{X}$, that is to say, a measurable mapping of $\mathfrak{X}$ into the space R of the real numbers. If x is random with probability law $m(e)$, then $y(x)$ is a random variable with some definite distribution function $F(y)$. For every real number y, there is a conditional probability law $\mu(y; e)$ for x in $\mathfrak{X}$, which is the probability $P\{x \in e | y(x) = y\}$. If Λ is any Borel subset of R, and e any subset of $\mathfrak{X}$ belonging to S, we have*

$$P\{y(x) \in \Lambda, x \in e\} = \int_{\Lambda} \mu(y; e)\, dF(y). \tag{1.2}$$

Theorem 1.2 still holds, of course, if, more generally, $y = y(x)$ is a mapping into n-dimensional Euclidean space, that is to say, if y is an n-dimensional random variable.

Jiřina's method is a direct set-theoretic method, but we can think of another approach. Let $\mathfrak{F}$ be a convenient vector space of S-measurable mappings f of $\mathfrak{X}$ into the space R of real numbers. An a priori or a conditional probability measure p on $(\mathfrak{X}, S)$ can be defined as a linear functional on $\mathfrak{F}$, that is, by means of the mathematical expectations

$$\int_{\mathfrak{X}} f(x) p\,(dx), \tag{1.3}$$

and the proof of the existence of some conditional probability measure p can be effected by establishing the existence of the corresponding mathematical expectations (1.3); I do not think that such a method would be more powerful than Jiřina's method, but it may be easier to handle.

As an example of such a method, in R. Fortet and E. Mourier [2], theorem 1.2 has been proved under the assumption that $\mathfrak{X}$ is a real separable reflexive Banach space, S being the smallest σ-field containing the Borel sets of $\mathfrak{X}$. In this case, $\mathfrak{F}$ can be restricted to be the space $\mathfrak{X}^*$ of the strongly continuous linear functionals on $\mathfrak{X}$, in such a way that a probability measure p on $(\mathfrak{X}, S)$ is defined by its characteristic functional

$$\varphi(x^*) = \int_{\mathfrak{X}} e^{i\langle x^*, x\rangle} p\,(dx), \tag{1.4}$$

where $x^* \in \mathfrak{X}^*$ is an arbitrary strongly continuous linear functional on $\mathfrak{X}$.

In particular, with this second method, it may be easier to handle the unicity problem. In his papers Jiřina says nothing on this question. In a short note by Mourier [8], she gives a definition of the unicity property which is obviously inadmissible. A tentative definition of the unicity property can be the following.

Two *given* solutions of the conditional probability existence problem, that is to say, two *given* functions $\mu(x; e)$, $\mu'(x; e)$ having the above properties (a), (b), and (c), are not considered as distinct if the set of the $x (x \in \mathfrak{X})$ such that there exists an $e \in S$ for which $\mu(x; e) \neq \mu'(x; e)$ is of m-measure zero.

Now let $\mathcal{C}$ be the family of the functions $\mu(x; e)$ having the properties (a), (b), and (c). We say that there is unicity if no two functions of $\mathcal{C}$ are distinct.

In this sense, and in the case treated by Fortet and Mourier [2], we can easily prove unicity. With their notations the problem reduces to the following: Λ being any Borel subset of R, and x^* any element of $\mathfrak{X}^*$, also $\Phi(x^*; \Lambda)$ being a given function and $\nu\,(d\lambda)$ a given probability measure on R, we look at the functions $\varphi(x^*; \lambda)$ of $x^* \in \mathfrak{X}^*$ and $\lambda \in R$, which have the following properties.

(i) For any fixed λ we have $\varphi(x^*; \lambda)$ as a function of x^* is a characteristic functional.

(ii) For any fixed $x^* \in \mathfrak{X}^*$ and for any Borel subset Λ of R,

$$\Phi(x^*; \Lambda) = \int_\Lambda \varphi(x^*; \lambda)\nu\,(d\lambda). \tag{1.5}$$

As $\mathfrak{X}^*$ is separable, we choose a dense denumerable subset $\mathfrak{X}^{*\prime}$ of $\mathfrak{X}^*$ and let x_j^*, $j = 1, 2, 3, \cdots$, be the elements of $\mathfrak{X}^{*\prime}$.

Let $\varphi(x^*; \lambda)$ and $\varphi'(x^*; \lambda)$ be two given arbitrary solutions of the preceding problem. By (1.5) and by the Radon-Nikodym theorem, and excepting a set of values of λ of ν-measure 0, $\varphi(x^*; \lambda)$ and $\varphi'(x^*; \lambda)$ are equal for all $x^* \in \mathfrak{X}^{*\prime}$. Since φ and φ' are continuous functions of x^* on $\mathfrak{X}^*$ they are identical on $\mathfrak{X}^*$, and this proves the unicity.

2. Discrimination between two laws

Let X be a real random variable, and let us consider the two following hypotheses: hypothesis H_1, X obeys the probability law $\mathcal{L}_1$, with density $f_1(x)$; hypothesis H_2, X obeys the probability law $\mathcal{L}_2$, with density $f_2(x)$. The discrimination problem between these two hypotheses is a classical one and contains the two essential features listed below.

(a) The likelihood ratio $y(x) = f_1(x)/f_2(x)$ is a sufficient statistic. In other words, the probability law of X, conditional on the value y of the ratio $f_1(x)/f_2(x)$, is the same with $\mathcal{L}_1$ as with $\mathcal{L}_2$.

(b) Starting from this likelihood ratio, methods of testing can be constructed, following, for instance, Neyman's point of view.

Now let us assume that $\mathfrak{X}$ is an arbitrary space of elements x and that m_1 and m_2 are two probability measures defined on $\mathfrak{X}$, more precisely defined on some σ-algebra $\mathcal{B}$ of subsets of $\mathfrak{X}$. A random element X with values in $\mathfrak{X}$ obeys either the law m_1 or the law m_2, and the question is to discriminate between these two hypotheses by one observation or by several independent observations of X.

There is no difficulty in introducing the analogue of the ratio $y = f_1(x)/f_2(x)$. Several authors, such as U. Grenander [3] and B. Adhikari [1], have remarked that the measures m_1 and m_2 are absolutely continuous with respect to the measure $m = m_1 + m_2$, by the Radon-Nikodym theorem, with respect to m, m_1 and m_2 have densities $f_1(x)$ and $f_2(x)$ respectively, and the ratio $y = f_1(x)/f_2(x)$

will be the likelihood ratio. Hence methods and results like those recalled in (b) can be easily developed (see Adhikari [1]). Here we are faced only with the problem of actually determining y as a function of x.

Concerning point (a), the sufficiency can be handled only if, with each of $\mathfrak{L}_1$ and $\mathfrak{L}_2$, the variable X has conditional probability laws with respect to the random variable $Y = f_1(X)/f_2(X)$. Let us assume for instance, that $\mathfrak{X}$ is a metric complete separable space, and that $\mathcal{B}$ is the σ-algebra of the Borel sets of $\mathfrak{X}$. Assuming also that the set $\{x: f_1(x) \text{ or } f_2(x) = 0\}$ is of m-measure 0 (it would not be difficult to remove this assumption), let ν_1 and ν_2 be the two probability laws for Y that correspond respectively to $\mathfrak{L}_1$ and $\mathfrak{L}_2$, so that ν_1 and ν_2 are absolutely continuous each with respect to the other. If ω is any B-measurable subset of the y-axis and ω^{-1} the set $\{x: y(x) \in \omega\}$ we have

$$
\begin{aligned}
(2.1) \qquad \nu_1(\omega) &= \int_{\omega^{-1}} m_1\,(dx) = \int_{\omega^{-1}} f_1(x) m\,(dx) = \int_{\omega^{-1}} \frac{f_1(x)}{f_2(x)} f_2(x) m\,(dx) \\
&= \int_{\omega^{-1}} \frac{f_1(x)}{f_2(x)} m_2\,(dx) = \int_{\omega} y \nu_2\,(dy).
\end{aligned}
$$

Let us assume that under $\mathfrak{L}_1$ the measure $\mu_1(y; e)$ is a probability law for X, conditional on Y, and that under $\mathfrak{L}_2$ the measure $\mu_2(y; e)$ is a probability law for X, conditional on Y. Such conditional probability laws exist by theorem 1.2 and we have

$$
(2.2) \qquad \int_{\omega} \mu_2(y; e) \nu_2\,(dy) = m_2(e \cap \omega)
$$

for every $e \in \mathcal{B}$. Consequently

$$
\begin{aligned}
(2.3) \qquad \int_{\omega} \mu_2(y; e) \nu_1\,(dy) &= \int_{\omega} \mu_2(y; e) y \nu_2\,(dy) = \int_{\omega} y m_2(e \cap dy) \\
&= \int_{\omega} y \left[\int_{e\,dy} f_2(x) m\,(dx) \right] = \int_{e \cap \omega} f_1(x) m\,(dx) \\
&= m_1(e \cap \omega),
\end{aligned}
$$

so that $\mu_2(y; e)$ is also a conditional probability law for X with respect to Y under $\mathfrak{L}_1$.

We may admit that this result constitutes sufficiency. It will, however, have a deeper meaning if the unicity of the conditional probability laws of X with respect to Y under $\mathfrak{L}_1$ has been proved. In these circumstances it becomes possible to assert that the two functions $\mu_1(Y; e)$ and $\mu_2(Y; e)$ as functions of e are almost surely not distinct. This is obtained in the case treated in Fortet and Mourier [2].

From a concrete point of view, we may remark that a quantity like Y cannot be measured with absolute precision. The only conditional probabilities having a concrete meaning are of the following kind.

$$
(2.4) \qquad P_1 = \frac{1}{\nu_1(\omega)} \int_{\omega} \mu_1(y; e) \nu_1\,(dy),
$$

$$P_2 = \frac{1}{\nu_2(\omega)} \int_\omega \mu_2(y; e)\nu_2\,(dy), \tag{2.5}$$

where ω may be, for instance, the interval $(y, y + \Delta y)$. In this case it is easy to verify that, with two convenient numbers θ and θ' between 0 and 1, we have

$$P_1 = \frac{y + \theta\,\Delta y}{y + \theta'\,\Delta y} P_2 \tag{2.6}$$

and hence, if $y \neq 0$ and if $\Delta y \to 0$, the ratio P_1/P_2 tends toward 1 uniformly in e. For this, we do not need the unicity of the conditional probability law.

3. Testing for the presence of a signal

In information theory we encounter the following problem. Let $U(t)$ be a Laplacian random real function of t on $[0, T]$ with continuous covariance

$$\Gamma(t, \tau) = E[U(t)U(\tau)] \tag{3.1}$$

and with $E[U(t)] \equiv 0$. For instance, $U(t)$ may represent a noise. Then $U(t)$ may be considered as a random element U with values in the separable Hilbert space $\mathfrak{X}$ of the real functions of t on $[0, T]$, the square of which is L-integrable. Let m_1 be the probability measure of U on $\mathfrak{X}$ under these conditions. Let $\rho(t)$ be an arbitrary given nonrandom element of $\mathfrak{X}$ (a signal), let $V(t)$ be the Laplacian random function of t on $[0, T]$ defined by $V(t) = \rho(t) + U(t)$, and let m_2 be the probability measure of $V(t)$ considered as a random element V with values in $\mathfrak{X}$. Let $\mathfrak{B}$ be the smallest σ-algebra of subsets of $\mathfrak{X}$ containing the classical "cylindrical" sets defined by the strongly continuous linear functionals on $\mathfrak{X}$; it is also the smallest σ-algebra containing the spheres of $\mathfrak{X}$ (see Mourier [8]). Then $\Gamma(t, \tau)$ and $\rho(t)$ uniquely determine m_1 and m_2 on $\mathfrak{B}$.

Finally, we consider a random element X with values in $\mathfrak{X}$, so that actually X is a random function of t on $[0, T]$. There is an a priori probability p_1 that X obeys the law m_1 and an a priori probability p_2 that it obeys the law m_2, where $p_1 + p_2 = 1$. By one (and only one) complete observation of the values taken by $X(t)$ for all the values of $t \in [0, T]$, we have to discriminate between m_2 and m_1 (signal or no signal).

We shall make use of the preceding sufficient statistic $y(x) = f_1(x)/f_2(x)$. In fact, we shall see that under certain conditions m_2 is absolutely continuous with respect to m_1 (and reciprocally), and actually we shall make use, as a sufficient statistic, of the density $y(x)$ of m_2 with respect to m_1. At the same time we shall obtain an explicit determination of $y(x)$ as a function of x in $\mathfrak{X}$.

Let s_j, for $j = 1, 2, 3, \cdots$ be the eigenvalues, distinct or not, and $u_j(t)$ the corresponding eigenfunctions of the following linear operator in $\mathfrak{X}$:

$$b(t) = \int_0^T \Gamma(t, \tau)a(\tau)\,d\tau, \qquad a, b \in \mathfrak{X}, \tag{3.2}$$

or symbolically,

$$b = \Gamma(a). \tag{3.3}$$

As we know, Γ is a Hermitian, positive definite operator, the $u_j(t)$ are mutually orthogonal, and we assume that they are normed, that is,

$$\int_0^T |u_j(t)|^2\, dt = 1. \tag{3.4}$$

We have

$$s_j u_j = \Gamma(u_j), \tag{3.5}$$

where the s_j are real and positive.

Supplementing the u_j, if necessary, by some other unitary vectors of $\mathfrak{X}$, we may assume that the u_j constitute a complete orthonormal basis for $\mathfrak{X}$. Taking $s_j = 0$, (3.5) remains true for the supplementary u_j introduced in this way.

We assume that the s_j are numbered in nonincreasing order and we know that $\sum_j s_j < +\infty$. Let us put

$$U^j = \int_0^T U(t) u_j(t)\, dt, \tag{3.6}$$

so that the U^j are mutually independent Laplacian random variables with

$$E\{U^j\} = 0, \qquad E\{|U^j|^2\} = s_j, \tag{3.7}$$

and almost surely we have

$$U(t) = \sum_j U^j u_j(t) \tag{3.8}$$

(with strong convergence in $\mathfrak{X}$).

Any bounded linear functional x^* on $\mathfrak{X}$ can be defined by its components x_j^* corresponding to the basis $\{u_j\}$, namely

$$x_j^* = \langle x^*, u_j\rangle. \tag{3.9}$$

The characteristic functional $\varphi_1(x^*)$ corresponding to m_1 is easy to find (see Mourier [8]) and is given by

$$\varphi_1(x^*) = E[e^{i\langle x^*, U\rangle}] = \exp\left(-\frac{1}{2}\sum_j s_j x_j^{*2}\right), \tag{3.10}$$

because $\langle x^*, U\rangle = \sum_j x_j^* U^j$.

We call ρ^j the components of $\rho(t)$ on the basis u_j, x^j the components of an arbitrary element x of $\mathfrak{X}$, and we put

$$A = \sum_j \frac{(\rho^j)^2}{s_j}, \tag{3.11}$$

$$\Psi(x) = \sum_j \frac{\rho^j}{s_j}\left(x^j - \frac{\rho^j}{2}\right). \tag{3.12}$$

Under the (necessary and sufficient) assumption that

$$A = \sum_j \frac{(\rho^j)^2}{s_j} < +\infty, \tag{3.13}$$

the series (3.12) is convergent almost everywhere on $\mathfrak{X}$ with respect to the measure m_1 and the characteristic functional $\varphi_2(x^*)$ corresponding to m_2 is equal to

$$\varphi_2(x^*) = \int_{\mathfrak{X}} e^{i\langle x^*,x\rangle} e^{\Psi(x)} m_1(dx), \tag{3.14}$$

this result implying in particular that, with respect to m_1, the measure m_2 is absolutely continuous with density

$$y(x) = e^{\Psi(x)}. \tag{3.15}$$

Results (3.12) and (3.15) have already been obtained by Grenander [3]. In practice, the linear functional $\Psi(x) = \log y(x)$ can be used as a sufficient statistic for the discrimination problem. The two probability laws of $\Psi(X)$ with m_1 and with m_2 are Laplacian laws with different means $-A/2$ and $+A/2$, but with the same standard deviation $\sqrt{A}$. From these properties, it is easy to construct a discrimination procedure.

We now give a mathematical interpretation of condition (3.13). If Γ is a positive definite Hermitian linear operator, let $\Gamma^{1/2}$ be its positive definite Hermitian square root, so that

$$\sqrt{s_j}\, u_j = \Gamma^{1/2}(u_j) \qquad j = 1, 2, \cdots. \tag{3.16}$$

Let us denote by $\Gamma^{1/2}(\mathfrak{X})$ the set of the $b \in \mathfrak{X}$ such that there exists at least one $a \in \mathfrak{X}$ such that $b = \Gamma^{1/2}(a)$. It is easily seen that condition (3.13) means that

$$\rho \in \Gamma^{1/2}(\mathfrak{X}). \tag{3.17}$$

With the analogous notation $\Gamma(\mathfrak{X})$, if we assume that $\rho \in \Gamma(\mathfrak{X})$, it follows that

$$\sum_j \frac{|\rho^j|^2}{s_j^2} < +\infty, \tag{3.18}$$

and A can also be interpreted as the Hermitian product in $\mathfrak{X}$ of ρ by any $a \in \mathfrak{X}$ such that $\Gamma(a) = \rho$.

4. The case of a stationary Laplacian noise

The preceding results are given with more detail in Fortet and Mourier [2]. They are applied by Bethoux in his thesis, in the following way. Let $U'(t)$ be a Laplacian stationary random function over $(-\infty, +\infty)$ with a correlation function $r(h)$, a null spectrum outside the bandwidth $(-\Omega, +\Omega)$ and a spectral density $f(\omega)$ inside the bandwidth $(-\Omega, +\Omega)$. Then we have

$$r(h) = \int_{-\Omega}^{+\Omega} e^{i\omega h} f(\omega)\, d\omega. \tag{4.1}$$

The assumption of a spectrum limited to a finite bandwidth $(-\Omega, +\Omega)$ is rather unrealistic, but is usual in communication theory. We shall add the hypothesis that $f(\omega)$ has an upper bound, and also a *positive* lower bound (on this last point, the hypothesis can be weakened). We now define $U(t)$ by

$$U(t) = U'(t) \qquad \text{for} \quad 0 \leqq t \leqq T. \tag{4.2}$$

Again, let $\tilde{\rho}(t)$ be a real function of t on $(-\infty, +\infty)$, also with a null spectrum outside the bandwidth $(-\Omega, +\Omega)$, and of the form

$$\tilde{\rho}(t) = \int_{-\Omega}^{+\Omega} e^{i\omega t}\varphi(\omega)\,d\omega, \qquad -\infty < t < +\infty, \tag{4.3}$$

with

$$\int_{-\Omega}^{+\Omega} |\varphi(\omega)|^2\,d\omega < +\infty. \tag{4.4}$$

We define $\rho(t)$ by

$$\rho(t) = \tilde{\rho}(t) \qquad \text{for} \quad 0 \leqq t \leqq T. \tag{4.5}$$

Putting

$$\rho'(t) = \int_{-\Omega}^{+\Omega} e^{i\omega t}\frac{\varphi(\omega)}{[f(\omega)]^{1/2}}\,d\omega, \qquad -\infty < t < +\infty, \tag{4.6}$$

we assume that

$$\int_{-\infty}^{+\infty} |\rho'(t)|^2\,dt = 2\pi \int_{-\Omega}^{+\Omega} \frac{|\varphi(\omega)|^2}{f(\omega)}\,d\omega \leqq \frac{2\pi\Omega\epsilon}{N}, \tag{4.7}$$

where

$$N = E(|U(t)|^2) = \int_{-\Omega}^{+\Omega} f(\omega)\,d\omega \tag{4.8}$$

and ϵ is a given positive number. That is to say, we assume that $\rho'(t)$ has a finite total energy, which is equivalent to the fact that $\tilde{\rho}(t)$ has a finite total energy, because $f(\omega)$ has a positive lower bound.

Now we shall apply section 3 to the $U(t)$ and the $\rho(t)$ defined in the preceding way, and under the stated assumptions. Here $b = \Gamma(a)$ is given by

$$b(t) = \int_0^T r(t-\tau)a(\tau)\,d\tau, \tag{4.9}$$

or, putting

$$g(\omega) = \int_0^T e^{-i\omega\tau}a(\tau)\,d\tau, \tag{4.10}$$

by the expression

$$b(t) = \int_{-\Omega}^{+\Omega} e^{i\omega t}f(\omega)g(\omega)\,d\omega. \tag{4.11}$$

From this, the following lemma can be obtained.

LEMMA 4.1. *The u_j corresponding to positive values of the s_j constitute a complete basis.*

We do not need any supplementary vectors of $\mathfrak{X}$, and Γ and $\Gamma^{1/2}$ are one-to-one mappings of $\mathfrak{X}$, respectively on $\Gamma(\mathfrak{X})$ and $\Gamma^{1/2}(\mathfrak{X})$. Further results can be deduced, such as those indicated below.

First, let us put

$$g_j(\omega) = \int_0^T e^{-i\omega\tau}u_j(\tau)\,d\tau, \tag{4.12}$$

and

$$u_j'(t) = \int_{-\Omega}^{+\Omega} e^{i\omega t} g_j(\omega)[f(\omega)]^{1/2}\, d\omega, \qquad -\infty < t < +\infty, \tag{4.13}$$

that is to say, $u_j'(t)$ is the Fourier transform over $-\infty < t < +\infty$ of the function of ω, which is equal to 0 for $|\omega| > \Omega$ and equal to $g_j(\omega)[f(\omega)]^{1/2}$ for $|\omega| \leqq \Omega$. It follows that

$$\int_{-8}^{+\infty} u_j'(t)\overline{u_k'}(t)\, dt = 2\pi s_j \int_0^T u_j(\tau)\, \overline{u_k}(\tau)\, d\tau = \begin{cases} 0 & \text{if } k \neq j; \\ 2\pi s_j & \text{if } k = j. \end{cases} \tag{4.14}$$

In other words, the $u_j'(t)$ constitute a (not necessarily complete) orthogonal basis over $(-\infty, +\infty)$.

Second, putting

$$2\pi s_j \lambda^j = \int_{-\infty}^{+\infty} \rho'(t)\, \overline{u_j'}(t)\, dt, \tag{4.15}$$

$\rho'(t)$ is necessarily of the form

$$\rho'(t) = \sum_j \lambda^j u_j'(t) + H(t), \qquad -\infty < t < +\infty, \tag{4.16}$$

$H(t)$ being some function such that

$$\int_{-\infty}^{+\infty} |H(t)|^2\, dt < +\infty, \qquad \int_{-\infty}^{+\infty} H(t)\overline{u_j'}(t)\, dt = 0 \tag{4.17}$$

for every j. It appears that

$$\lambda^j = \frac{\rho^j}{s_j}, \tag{4.18}$$

and that

$$H(t) \equiv 0. \tag{4.19}$$

Consequently, assumption (4.7) can be written

$$\sum_j \frac{|\rho^j|^2}{s_j} \leqq \frac{\Omega\epsilon}{N}, \tag{4.20}$$

which implies that condition (3.13) is actually satisfied by the present signal ρ, and we have just to apply section 3.

Incidentally, we have found a *physical* interpretation of condition (3.13), which in the context of the present section 4 is equivalent to the fact that $\rho'(t)$, or $\tilde{\rho}(t)$, has finite total energy.

Obviously, the discrimination will be performed in the best conditions if A is as large as possible; that is, if $\rho'(t)$ is such that

$$\int_{-\infty}^{+\infty} |\rho'(t)|^2\, dt = 2\pi \frac{\Omega\epsilon}{N}. \tag{4.21}$$

5. Capacity of a channel

We now consider the same noise $U(t)$, but a number $n+1$ of real signals $\rho_0(t), \rho_1(t), \cdots, \rho_n(t)$ over $(0, T)$, with the following assumptions.

(a) $\rho_k(t)$ is the part relative to $(0, T)$ of a real signal $\tilde{\rho}_k(t)$ defined over $(-\infty, +\infty)$, with a spectral density $\varphi_k(\omega)$ limited to the bandwidth $(-\Omega, +\Omega)$, that is,

$$\rho_k(t) = \int_{-\Omega}^{+\Omega} e^{i\omega t}\varphi_k(\omega)\, d\omega, \qquad -\infty < t < +\infty. \tag{5.1}$$

(b) If we put

$$\rho_k'(t) = \int_{-\Omega}^{+\Omega} e^{i\omega t} \frac{\varphi_k(\omega)}{[f(\omega)]^{1/2}}\, d\omega, \qquad -\infty < t < +\infty, \tag{5.2}$$

we have

$$\int_{-\infty}^{+\infty} |\rho_k'(t)|^2\, dt \leqq \frac{2\pi\Omega\epsilon}{N}. \tag{5.3}$$

Let m_k be the probability law of the random element $U(t) + \rho_k(t)$. By an observation over $(0, T)$ of the random element $V(t) = U(t) + \rho_k(t)$, where k is fixed but unknown, we have to discriminate between the m_k, that is, to decide on the value of k.

Putting

$$\rho_k^j = \int_0^T \rho_k(t)\overline{u_j}(t)\, dt, \tag{5.4}$$

$$V^j = \int_0^T V(t)\overline{u_j}(t)\, dt, \tag{5.5}$$

$$\Psi_\alpha = \sum_j \frac{(\rho_\alpha^j - \rho_0^j)}{s_j}\left(V^j - \frac{\rho_\alpha^j + \rho_0^j}{2}\right), \qquad \alpha = 1, 2, \cdots, n, \tag{5.6}$$

it can be seen by a generalization of (3.12) that $\Psi = \{\Psi_1, \Psi_2, \cdots, \Psi_n\}$ is a sufficient statistic for our problem, and it is an n-dimensional Laplacian random variable.

The functions $\rho(t) = \sum_j \rho^j u_j(t)$ for which

$$\sum_j \frac{|\rho^j|^2}{s_j} < +\infty \tag{5.7}$$

are the elements of $\Gamma^{1/2}(\mathfrak{X})$. This space can be considered as a separable Hilbert space where the scalar product is defined by

$$\rho' \cdot \rho'' = \sum_j \frac{\rho'^j \rho''^j}{s_j}. \tag{5.8}$$

The characterization of the Laplacian law of $\Psi = \{\Psi_1, \cdots, \Psi_n\}$, that is to say, the specification of the values of $E\{\Psi_\alpha\}$ and $E\{\Psi_\alpha\Psi_\beta\}$, depends only on the scalar products of the signals $\rho_0, \rho_1, \cdots, \rho_n$. Hence the conditions of the test remain unchanged if $\rho_0, \rho_1, \cdots, \rho_n$ are replaced by $n+1$ signals $\rho_0', \rho_1', \cdots, \rho_n'$ derived from the ρ_j by a unitary operator in $\Gamma^{1/2}(\mathfrak{X})$, and the condition $\sum_j |\rho^j|^2/s_j \leqq \Omega\epsilon/N$ will also remain unchanged by such a unitary operator.

Let $\mathfrak{H}_n$ be the subspace of $\Gamma^{1/2}(\mathfrak{X})$ consisting of all the $\rho(t)$ such that $\rho^j = 0$ for every $j > n + 1$, so that $\mathfrak{H}_n$ is of dimension $(n + 1)$. There is always a unitary operator $\mathfrak{z}$ in $\Gamma^{1/2}(\mathfrak{X})$ which transforms $\rho_0, \rho_1, \cdots, \rho_n$ into signals $\rho_0', \rho_1', \cdots, \rho_n'$ belonging to $\mathfrak{H}_n$. That is, we do not lose any generality in assuming that

$$\rho_k \in \mathfrak{H}_n \tag{5.9}$$

for $k = 0, 1, \cdots, n$, or that

$$\rho_k^j = 0 \tag{5.10}$$

for $k = 0, 1, 2, \cdots, n$ and for every $j > n + 1$. Under these conditions Ψ depends only on $V^1, V^2, \cdots, V^{n+1}$, that is, on $\{V^1, \cdots, V^{n+1}\}$ or, in other and better words, $Q = \{V^1 s_1^{-1/2}, V^2 s_2^{-1/2}, \cdots, V^{n+1} s_{n+1}^{-1/2}\}$ is a sufficient statistic. Clearly Q is a system of $n + 1$ independent Laplacian random variables, with standard deviation equal to 1. Finally, if we put

$$x_k^j = \frac{\rho_k^j}{\sqrt{s_j}}, \qquad x^j = \frac{U^j}{\sqrt{s_j}}, \tag{5.11}$$

our problem is the following: let E_{n+1} be a proper Euclidean $(n + 1)$-dimensional affine space referred to an orthonormal reference system with origin 0. In E_{n+1}, in the sphere Σ_n of center 0 and radius $\sqrt{A} = (\Omega\epsilon/N)^{1/2}$, the $n + 1$ points $M_0, M_1, \cdots, M_n$ are given, the coordinates of M_k being the x_k^j, with $j = 1, 2, \cdots, n + 1$. We make one observation of the random position of the random point M with coordinates $x_k^j + X^j$, where k is unknown, and we have to deduce k from the observation.

Let us assume that the different possible values of k, $k = 0, 1, 2, \cdots, n$, have definite prior probabilities $p_0, p_1, \cdots, p_n$, where $\sum_k p_k = 1$. We divide E_{n+1} into $n + 1$ disjoint subsets R_k, with $k = 0, 1, \cdots, n$, the union of which is E_{n+1}. We denote by R such a partition, and we take $k = h$ if $M \in R_h$.

With such a procedure, the probability of accepting a false value of k depends on R, on the disposition $\mathfrak{D}$ of the M_k inside Σ_n, on n, and on $\Omega\epsilon/N$. We denote it by

$$P\left(R; \mathfrak{D}; n; \frac{\Omega\epsilon}{N}\right). \tag{5.12}$$

There is one and only one partition R_0 such that

$$P\left(R_0; \mathfrak{D}; n; \frac{\Omega\epsilon}{N}\right) = P_m\left(\mathfrak{D}; n, \frac{\Omega\epsilon}{N}\right) = \min_R P\left(R; \mathfrak{D}; n; \frac{\Omega\epsilon}{N}\right), \tag{5.13}$$

and it is not difficult to find R_0, which has a relatively simple geometrical form.

Now there exists at least one disposition $\mathfrak{D}_0$ such that

$$P_m\left(\mathfrak{D}_0; n; \frac{\Omega\epsilon}{N}\right) = \Pi\left(n; \frac{\Omega\epsilon}{N}\right) = \min_{\mathfrak{D}} P_m\left(\mathfrak{D}; n; \frac{\Omega\epsilon}{N}\right); \tag{5.14}$$

in the case where $p_0 = p_1 = \cdots = p_n = 1/(n+1)$, intuitively such a $\mathfrak{D}_0$ seems to be unique and to be obtained when $M_0, M_1, \cdots, M_{n+1}$ are the vertices of a regular polyhedron $\mathcal{P}_n$ inscribed in Σ_n but Bethoux gives no rigorous proof of this fact. Let η be a given arbitrary positive number, and let $n(\Omega\epsilon/N, \eta)$ be the largest integer n satisfying the inequality $\Pi(n; \Omega\epsilon/N) \leqq \eta$. We assume that $\epsilon = PT$, where P is a given constant. We interpret the ρ_k as signals emitted by a sender, and $V(t) = U(t) + \rho_k(t)$ as the corresponding responses received by the receiver after transmission by a noisy channel.

The limit

$$C = \lim_{\eta \to +0} \lim_{T \to +\infty} \frac{\log n\left(\frac{\Omega PT}{N}; \eta\right)}{T} \tag{5.15}$$

if it exists, can be interpreted as the *theoretical capacity of the channel*, at least in the case $p_0 = p_1 = \cdots = p_n = 1/(n+1)$.

Classically (Shannon and Weaver [11] and several other authors), it is stated that, at least in the case of a "white noise," the capacity of a channel has the value C', where

$$C' = \Omega \log (1 + P/N). \tag{5.16}$$

Obviously this C' is different from our C, and it appears from (5.15) that C cannot depend on the spectrum of the noise. More precisely, if we put

$$Z = \frac{\Omega PT}{N}, \qquad c(\eta) = \lim_{Z \to +\infty} \frac{N(Z; \eta)}{Z}, \qquad c = \lim_{\eta \to +0} c(\eta), \tag{5.17}$$

we find

$$C = c \frac{\Omega P}{N}. \tag{5.18}$$

I do not know any complete rigorous proof of (5.16), and whether or not the classical definition of the capacity is identical with the definition that I have adopted here is not clear to me.

G. Bethoux, in his current investigations, has not been able to compute $\Pi(n; \Omega\epsilon/N)$ and consequently he has not given a rigorous proof of the existence of the limit c, nor the exact value of c, but he has proved that

$$c \geqq \frac{1}{2}. \tag{5.19}$$

6. Sufficient statistic for an estimation

We make use of the notations of section 3. Let $\mathfrak{F}$ be a given family of elements ρ of $\mathfrak{X}$. We consider the random function

$$V(t) = \rho(t) + U(t) \tag{6.1}$$

on $[0, T]$, where $\rho(t)$ is a nonrandom (unknown but fixed) element of $\mathfrak{F}$.

The problems which may arise in these conditions have different aspects following the nature of $\mathfrak{F}$, and in sections 6 to 8 I shall always assume that $\mathfrak{F} \subset \Gamma^{1/2}(\mathfrak{X})$. The functional defined by (3.12) will be introduced in the form

$$\Psi(\rho; V) = \sum_j \frac{\rho^j}{s_j}\left(V^j - \frac{\rho^j}{2}\right), \tag{6.2}$$

and I also introduce

$$\Theta(\rho) = \Theta(\rho; U) = \sum_j \frac{\rho^j}{s_j} U^j. \tag{6.3}$$

We have to remark that the study of $\Psi(\rho; V)$ reduces to that of $\Theta(\rho; U)$ and that, almost surely, neither U nor V belong to $\Gamma^{1/2}(\mathfrak{X})$. For an arbitrarily given $\rho \in \Gamma^{1/2}(\mathfrak{X})$, almost surely $\Theta(\rho; U)$ exists, but we do not know whether almost surely $\Theta(\rho; U)$ exists for all the ρ belonging to $\mathfrak{F}$. The study of this point would be connected with the study of $\Theta(\rho; U)$ as a random function of $\rho \in \Gamma^{1/2}(\mathfrak{X})$. We recall that $\Gamma^{1/2}(\mathfrak{X})$ can be considered as a (separable) Hilbert space, in which the scalar product is defined by

$$\rho \cdot \rho' = \sum_j \frac{\rho^j \rho'^j}{s_j} \tag{6.4}$$

and the square of the norm $||\rho||$ is defined by

$$||\rho||^2 = \sum_j \frac{|\rho^j|^2}{s_j}; \tag{6.5}$$

the notation $\rho \times \rho'$ will be used for the scalar product in $\mathfrak{X}$.

It appears that $\Theta(\rho)$ is a Laplacian random function of $\rho \in \Gamma^{1/2}(\mathfrak{X})$, with the following properties:

$$E\{\Theta(\rho)\} \equiv 0; \tag{6.6}$$

$$E\{\Theta(\rho)\Theta(\rho')\} = \rho \cdot \rho', \tag{6.7}$$

the scalar product in $\Gamma^{1/2}(\mathfrak{X})$;

$$E\{|\Theta(\rho) - \Theta(\rho')|^2\} = ||\rho - \rho'||^2, \tag{6.8}$$

the square of the norm in $\Gamma^{1/2}(\mathfrak{X})$. It is not the Laplacian random function of a variable in a Hilbert space considered by Lévy [6]. Some facts can be heuristically deduced, but a deeper study of $\Theta(\rho)$ on $\Gamma^{1/2}(\mathfrak{X})$ is not yet available.

Actually, however, we are interested in $\Theta(\rho)$, not in the whole space $\Gamma^{1/2}(\mathfrak{X})$, but only on $\mathfrak{F}$, and for some $\mathfrak{F}$ the situation is quite simple, as the following show.

EXAMPLE 6.1. $\mathfrak{F} \subset \Gamma(\mathfrak{X})$. This is possible because $\Gamma(\mathfrak{X}) \subset \Gamma^{1/2}(\mathfrak{X})$. In this case $\Theta(\rho; U)$ can be interpreted as $\bar{\rho} \times U$, where $\bar{\rho}$ is the element of $\mathfrak{X}$ with components ρ^j/s_j. Almost surely, $\Theta(\rho)$ exists for every $\rho \in \Gamma(\mathfrak{X})$.

EXAMPLE 6.2. *The parametric case.* Let τ be a variable element in a subset $\mathfrak{D}$ of some Euclidean space of dimension ν, and f a mapping of $\mathfrak{D}$ into $\Gamma^{1/2}(\mathfrak{X})$. I put $\rho_\tau = f(\tau)$ and if $\mathfrak{F}$ is the set of the ρ_τ, I say that we are in the ν-dimensional parametric case. In such a case $\Theta(\rho)$ reduces to an ordinary Laplacian random function.

7. Uniformly most powerful tests

Let ρ_0 be any given element of $\mathfrak{F}$; let H_0 be the hypothesis that $\rho = \rho_0$; let H_ρ be the hypothesis that $\rho = \rho'$, where ρ' is any element of $\mathfrak{F}$ different from ρ_0; and let us consider the test of H_0 against the whole set of hypotheses $H_{\rho'}$. It is easy to see that

(a) if each element ρ' of $\mathfrak{F}$ is of the form

$$\rho' = C\rho_0, \tag{7.1}$$

where C is any real constant independent of t, there exists a U.M.P. test;

(b) if $\mathfrak{F} \subset \Gamma(\mathfrak{X})$ then, for the existence of a U.M.P. test, condition (7.1) is necessary.

8. Maximum likelihood estimators

I now suppose that our problem is, from a single observation of V, that is, the observation of the values of $V(t)$ for all $t \in [0, T]$, to estimate ρ, knowing that $\rho \in \mathfrak{F}$.

Let $\hat{\rho}$ be some definite element of $\mathfrak{F}$, and let us assume that, for the observed V and for any element ρ' of $\mathfrak{F}$, the functional $\Psi(\rho'; V)$ exists and that

$$\Psi(\hat{\rho}; V) \geqq \Psi(\rho'; V). \tag{8.1}$$

Let $\hat{m}$ and m' be the two probability measures (on $\mathfrak{X}$) corresponding to the two random elements

$$\hat{V}(t) = \hat{\rho}(t) + U(t), \qquad V'(t) = \rho'(t) + U(t). \tag{8.2}$$

Then we know from section 3 that $\hat{m}$ is absolutely continuous with respect to m', and that its corresponding density is given by

$$\exp\{\Psi(\hat{\rho}; x) - \Psi(\rho'; x)\}, \qquad x \in \mathfrak{X}. \tag{8.3}$$

Consequently we can call $\hat{\rho}$ a maximum likelihood estimator of ρ.

Now let us suppose that there are two different maximum likelihood estimators $\hat{\rho}_1$ and $\hat{\rho}_2$. Then we must have $\Psi(\hat{\rho}_1; V) = \Psi(\hat{\rho}_2; V)$ almost surely and, if λ is the common value of $\Psi(\hat{\rho}_1; V)$ and $\Psi(\hat{\rho}_2; V)$, we can write

$$\lambda = \frac{1}{2}[\Psi(\hat{\rho}_1; V) + \Psi(\hat{\rho}_2; V)] = \sum_j \frac{\hat{\rho}_1^j + \hat{\rho}_2^j}{2}\frac{1}{s_j} V^j - \frac{1}{4}(||\hat{\rho}_1||^2 + ||\hat{\rho}_2||^2). \tag{8.4}$$

Putting $\rho' = 1/2(\hat{\rho}_1 + \hat{\rho}_2)$, we get

$$\begin{aligned}\lambda &= \Psi(\rho'; V) - \frac{1}{2}\frac{||\hat{\rho}_1||^2 + ||\hat{\rho}_2||^2}{2} - ||\rho'||^2 \\ &= \Psi(\rho'; V) - \frac{1}{8}||\hat{\rho}_1 - \hat{\rho}_2||^2,\end{aligned} \tag{8.5}$$

and hence

THEOREM 8.1. *If $\mathfrak{F}$ is such that $\Psi(\rho; V)$ exists almost surely for every $\rho \in \mathfrak{F}$, and is also such that it contains $(\rho_1 + \rho_2)/2$ whenever it contains ρ_1 and ρ_2, then there is at most one maximum likelihood estimator.*

Now let us assume that $\mathfrak{F} \subset \Gamma(\mathfrak{X})$. I call $\bar{\rho}$ the element of $\mathfrak{X}$ with components ρ^j/s_j; $|x|$ means the norm of $x \in \mathfrak{X}$ as an element of $\mathfrak{X}$; we have $|\bar{\rho}| \geqq ||\rho|| \geqq |\rho|$, and

$$\Psi(\rho; V) = \bar{\rho} \times V - \frac{1}{2}\bar{\rho} \times \rho \leqq |\bar{\rho}|\,|V| - \frac{1}{2}\bar{\rho} \times \rho; \tag{8.6}$$

it appears that $\Psi(\rho; V)$ is bounded when ρ varies in $\Gamma(\mathfrak{X})$, and a fortiori if ρ varies in $\mathfrak{F}$. With the use of (8.5) and a classical tool, we can prove

THEOREM 8.2. *If*

(a) *$\mathfrak{F}$ contains $(\rho_1 + \rho_2)/2$ whenever it contains ρ_1 and ρ_2;*

(b) *$\mathfrak{F}$ is closed in the sense of the strong topology in $\Gamma^{1/2}(\mathfrak{X})$;*

(c) *$\mathfrak{F} \subset \Gamma(\mathfrak{X})$,*

then there exists one and only one maximum likelihood estimator.

COROLLARY 8.1. *Under assumptions* (a) *and* (b) *of theorem* 8.2 *and the assumption that, almost surely, $\Psi(\rho; V)$ exists and is bounded for every $\rho \in \mathfrak{F}$, there is one and only one maximum likelihood estimator.*

Incidentally, from some heuristic geometrical considerations, it can be seen that a maximum likelihood estimator is not in general a sufficient estimator, unless $\mathfrak{F}$ belongs to some convenient space.

Application to the linear parametric case. For example, let us suppose that $\mathfrak{F}$ consists of a finite number ν of functions $g_1(t), g_2(t), \cdots, g_\nu(t)$ of t on $[0, T]$, belonging to $\Gamma^{1/2}(\mathfrak{X})$, together with all the linear combinations

$$\rho(t) = \sum_{k=1}^{\nu} \lambda^k g_k(t) \tag{8.7}$$

of the g_k. Without loss of generality, we assume that the g_k are linearly independent. This is an example of a parametric case of order ν; we shall call it the linear parametric case.

From corollary 8.1, or directly, putting

$$a_{hk} = \sum_j \frac{g_h^j g_k^j}{s_j} = g_h g_k, \qquad b_h = \sum_j \frac{g_h^j}{s_j} V^j, \tag{8.8}$$

where the g_h^j are the components of g_h on the basis $\{u_j\}$, we get

THEOREM 8.3. *In the linear parametric case, there is one and only one maximum likelihood estimator $\hat{\rho} = \{\hat{\lambda}^k\}$, which is also an unbiased sufficient estimator, and which is given by*

$$\sum_{k=1}^{\nu} a_{hk}\hat{\lambda}^k = b_h, \qquad h = 1, 2, \cdots, \nu. \tag{8.9}$$

The sufficiency results from the fact that

$$\Psi(\rho; V) = \Psi(\hat{\rho}; V) + [\Psi(\rho; V) - \Psi(\hat{\rho}; V)] \tag{8.10}$$

and that

$$\Psi(\rho; V) - \Psi(\hat{\rho}; V) = \frac{1}{2}\sum_j \frac{|\hat{\rho}^j|^2}{s_j} - \frac{1}{2}\sum_j \frac{|\rho^j|^2}{s_j} + \sum_{k=1}^{\nu} (\lambda^k - \hat{\lambda}^k) b_k, \tag{8.11}$$

which, from (8.9), *does not depend on V.*

Without loss of generality we can assume that the g_k are orthonormed, as elements of $\Gamma^{1/2}(\mathfrak{X})$; under these conditions it is immediately proved that

$$E(\|\hat{\rho} - \rho\|^2) = \nu. \tag{8.12}$$

Since $\hat{\rho}$ is linearly deduced from V, it follows that $\hat{\rho} \in \mathfrak{F}$ is a Laplacian random variable of dimension ν. Hence we are finally concerned with classical mathematical statistics in finite-dimensional space. In particular, taking into account the fact that $\hat{\rho}$ is a sufficient estimator, it is easy to prove that any other unbiased estimator $\tilde{\rho}$ will be such that $E(\|\tilde{\rho} - \rho\|^2) > \nu$, which proves

THEOREM 8.4. *In the linear parametric case, the maximum likelihood estimator is also a minimum variance estimator.*

These theoretical results can be applied effectively, for instance if we actually know the s_j and the u_j, which is not difficult in some cases, for example if $U(t)$ is a Wiener-Lévy process. Examples of this kind have already been investigated (see H. B. Mann [7] and C. T. Striebel [11]).

An interesting application concerns the case where the family $\mathfrak{F}$ consists of functions $\rho(t)$ of the form

$$\rho(t) = s(t - \lambda), \tag{8.13}$$

where $s(t)$, the signal, is a given function, and λ, the delay, an unknown parameter. In an observation by radar, the delay λ will be proportional to the distance of the target from the observing radar, and the estimation of λ gives an estimate of this distance.

A few results of sections 6 to 8 have been obtained by Bethoux and several others have been obtained by A. Hanen, in their as yet unpublished theses.

I should mention also that these sections are substantially identical with the paper presented by E. Parzen [10] at this Symposium.

REFERENCES

[1] B. ADHIKARI, "Tests d'hypotheses pour processus stochastiques," unpublished Ph.D. thesis, University of Paris, 1957.

[2] R. FORTET and E. MOURIER, "Les fonctions aléatoires comme éléments aléatoires dans un espace de Banach," *J. Math. Pures Appl.*, Vol. 38 (1959), pp. 347–364.

[3] U. GRENANDER, "Stochastic processes and statistical inference," *Ark. Mat.*, Vol. 1 (1950), pp. 195–277.

[4] M. JIŘINA, "Conditional probabilities on strictly separable algebras," *Czechoslovak Math. J.*, Vol. 4 (1954), pp. 372–380. (In Russian with English summary.)

[5] ———, "On regular conditional probabilities," *Czechoslovak Math. J.*, Vol. 9 (1959), pp. 445–450. (In English.)

[6] P. LÉVY, "Random functions: a Laplacian random function depending on a point of Hilbert space," Univ. California Publ. Statist., Vol. 2 (1956), pp. 195–206.

[7] H. B. MANN, "A theory of estimation for the fundamental random process and the Ornstein Uhlenbeck process," *Sankhyā*, Vol. 13 (1954), pp. 325–350.

[8] E. MOURIER, "Éléments aléatoires dans un espace de Banach," *Ann. Inst. H. Poincaré*, Vol. 13 (1953), pp. 161–244.

[9] ———, "Lois de probabilités conditionnelles; existence et détermination d'un résumé exhaustif pour la discrimination entre plusieurs lois de probabilités dans des espaces de Banach," *C. R. Acad. Sci. Paris*, Vol. 247 (1958), pp. 1552–1554.

[10] E. Parzen, "Regression analysis of continuous parameter time series," *Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1961, Vol. 1, pp. 469–489.

[11] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, Urbana, University of Illinois Press, 1949.

[12] C. T. Striebel, "Densities for stochastic processes," *Ann. Math. Statist.*, Vol. 30 (1959), pp. 559–567.