THEORY OF PROBABILITY AND ITS APPLICATIONS

Volume XXIV Number 2

1979

THE USE OF ω^2 TESTS FOR TESTING PARAMETRIC HYPOTHESES

E. V. KHMALADZE

(*Translated by W. U. Sirk*)

1. Introduction

As we know, the ω^2 test for testing a simple hypothesis about an unknown continuous distribution function (d.f.) $F(x)$ of independent identically distributed random variables (r.v.'s) $X_1, \cdots, X_n$, is based on the statistic

$$\omega_n^2 = n \int [F_n(x) - F_0(x)]^2 \, dF_0(x),$$

where $F_n(x)$ denotes the empirical distribution function (e.d.f.) of the r.v.'s $X_1, X_2, \cdots, X_n$; according to the hypothesis, $F(x) = F_0(x)$. Information about the ω^2 test can be found, for example, in [1] or [2]. Carrying out the usual substitution $t = F_0(x)$, we obtain

$$\omega_n^2 = n \int [F_n(t) - t]^2 \, dt = \int v_n^2(t) \, dt,$$

where $F_n(t)$ denotes the e.d.f. of the independent uniformly distributed r.v.'s $T_1 = F_0(X_1), \cdots, T_n = F_0(X_n)$, and $v_n(t) = \sqrt{n}\,[F_n(t) - t]$. The sequence of distributions P_n of the processes $v_n(t)$ converges weakly in $L_2[0, 1]$ as $n \to \infty$ to the distribution P of the Brownian bridge $v(t)$, and consequently the distribution of the statistic ω^2 converges weakly to the distribution of the r.v.'s.

$$\omega^2 = \int v^2(t) \, dt.$$

Since P-a.s. has the expansion (see, for example, [1], p. 31, or [3], p. 229)

$$v(t) = \sum_{k=1}^{\infty} \frac{1}{\pi k} V_k \sin \pi k t,$$

where $V_k, k = 1, 2, \cdots$, are independent and have a standard normal distribution, then

$$\omega^2 = \sum_{k=1}^{\infty} \frac{1}{\pi^2 k^2} V_k^2.$$

The distribution of this quadratic form is called the distribution of ω^2.

In the case where a complex parametric hypothesis $F(x)\in\{F(x,\theta),\ \theta\in\Theta\}$ is being tested, we would like to consider the analogous statistic,

$$\hat{\omega}_n^2=n\int[F_n(x)-F(x,\hat{\theta})]^2\,dF(x,\hat{\theta}),$$

where $\hat{\theta}$ denotes the estimate of the unknown value of the parameter θ. Such a statistic has in fact been considered already (see, for example, [4], [5]); however, a serious difficulty becomes apparent at once. Let the hypothesis be valid, i.e., let there exist a value $\theta_0\in\Theta$ such that $F(x)=F(x,\theta_0)$. Making the substitution $t=F(x,\theta_0)$, we obtain

$$\hat{\omega}_n^2=n\int[F_n(t)-G(t,\hat{\theta})]^2\,dG(t,\hat{\theta}),$$

where $G(t,\theta)=F[F^{-1}(t,\theta_0),\theta]$. It is clear that $G(t,\theta_0)=t$. Under the usual regularity conditions, the sequence of distributions Q_n of the processes $u_n(t)=\sqrt{n}\,[F_n(t)-G(t,\hat{\theta})]$ converges weakly in $L_2[0,1]$ as $n\to\infty$ to the distribution of a Gaussian process $u(t)$ with the expansion

$$u(t)=\sum_{k=1}^{\infty}\mu_kU_ka_k(t).$$

The distribution of the statistic $\hat{\omega}_n^2$ converges weakly to the distribution of the quadratic form

$$\hat{\omega}_n^2=\sum_{k=1}^{\infty}\mu_k^2U_k^2$$

of the independent r.v.'s $U_1, U_2,\cdots$, with standard normal distribution. However (and here is the difficulty mentioned), the coefficients $\mu_1,\mu_2,\cdots$ (like the ortho-normalized system of functions $a_1(t), a_2(t),\cdots$) depend on the family $\{F(x,\theta),\ \theta\in\Theta\}$. Consequently, to obtain the limit distribution of the statistic $\hat{\omega}_n^2$ in testing each separate hypothesis, we have to recompute each time the coefficients $\mu_1,\mu_2,\cdots$, and the distribution of the corresponding quadratic form. This situation, highly inconvenient in itself, is aggravated by the fact that computing the coefficients μ_k is a fairly difficult task as can be seen in testing the normality hypothesis (for example, [5], [6], [7]).

On the other hand, we attempt to extend the class of quadratic functionals under consideration; specifically, we turn to integrals of the form

$$\iint K(t,s)\,du(t)\,du(s).$$

(Note that $\hat{\omega}^2=\iint\min(t,s)\,du(t)\,du(s)$.) These integrals are also represented as the quadratic form $\sum_{k=1}^{\infty}\lambda_k^2Z_k^2$ of the independent r.v.'s $Z_1, Z_2,\cdots$, with standard normal distribution and coefficients $\lambda_1^2,\lambda_2^2,\cdots$, depending on the kernel $K(t,s)$.

L. N. Bol'shev posed the question of finding a kernel $K(t,s)$ which probably depends on a hypothetical family $\{F(x,\theta),\ \theta\in\Theta\}$ for which the coefficients λ_k^2 do not depend on this family of d.f.'s, but coincide with previously specified numbers (for example, with $1/\pi^2k^2$, then the distribution of the corresponding double integral will be the distribution of ω^2). The present paper is an attempt to examine Bol'shev's question.

2. Description of the Limit Process for $u_n(t)$

Let the d.f. $F(x)$ of independent identically distributed r.v.'s $X_1, \cdots, X_n$ belong, by hypothesis, to the family $\mathscr{F}$ of functions of absolutely continuous distributions $F(x, \theta)$ depending on the parameter θ. We assume that θ is a scalar parameter which takes on values from a certain open set, say, from an interval $\Theta = (\theta_1, \theta_2)$. Extending the results to the case of a vector parameter presents no difficulty. About the family $\mathscr{F}\{F(x, \theta), \theta_1 < \theta < \theta_2\}$ we assume that:

1. The d.f. $F(x, \theta)$ is differentiable with respect to θ for all x and $\theta \in \Theta$; moreover, for each value $\theta \in \Theta$ there exists a value $\varepsilon > 0$ such that

$$\int \left[\left(\frac{\partial}{\partial \theta} F\right)(x, \theta')\right]^2 dF\,(x, \theta) < \infty, \qquad |\theta' - \theta| < \varepsilon,$$

and, in addition,

$$\int \left[\left(\frac{\partial}{\partial \theta} F\right)(x, \theta') - \left(\frac{\partial}{\partial \theta} F\right)(x, \theta)\right]^2 dF\,(x, \theta) = o(1) \quad \text{as } |\theta' - \theta| \to 0.$$

2. For all $\theta \in \Theta$ the Fisher information is

$$\varkappa = \int \left[\left(\frac{\partial}{\partial \theta} \log f\right)(x, \theta)\right]^2 dF\,(x, \theta) < \infty,$$

where $f(x, \theta)$ denotes the density of the distribution $F(x, \theta)$.

Let θ_0 be the value (unknown) of the parameter for which $F(x) = F(x, \theta_0)$. About the r.v. $\hat{\theta} = \hat{\theta}(X_1, \cdots, X_n)$, an estimate of the value θ_0, we assume that:

3. For a certain function $l(x, \theta)$ satisfying the conditions

$$\int l(x, \theta)\, dF\,(x, \theta) = 0 \quad \text{and} \quad \int l^2(x, \theta)\, dF\,(x, \theta) < \infty,$$

the representation

$$\sqrt{n}(\hat{\theta} - \theta_0) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} l(X_i, \theta_0) + o_P(1) \quad \text{as } n \to \infty$$

is valid.

The assertions stated below are presented without the words "theorem", "lemma" or "corollary", and are numbered with Roman numerals.

I. *Let $r_n(x, \hat{\theta})$ be given by the equation*

$$\begin{aligned} \sqrt{n}[F_n(x) - F(x, \hat{\theta})] &= \sqrt{n}[F(x) - F(x, \theta_0)] \\ &\quad - \left(\frac{\partial}{\partial \theta} F\right)(x, \theta_0) \frac{1}{\sqrt{n}} \sum_{i=1}^{n} l(X_i, \theta_0) + r_n(x, \hat{\theta}). \end{aligned} \tag{1}$$

If Conditions 1 *and* 3 *are satisfied, then, as* $n \to \infty$,

$$\int [r_n(x, \hat{\theta})]^2 \, dF(x, \theta_0) = o_P(1).$$

PROOF. The result follows easily from the equation

$$r_n(x, \hat{\theta}) = \left[\left(\frac{\partial}{\partial \theta} F\right)(x, \theta') - \left(\frac{\partial}{\partial \theta} F\right)(x, \theta_0)\right]\sqrt{n}(\hat{\theta} - \theta_0)$$
$$+ \left(\frac{\partial}{\partial \theta} F\right)(x, \theta_0)\left[\sqrt{n}\,(\hat{\theta} - \theta_0) - \frac{1}{\sqrt{n}} \sum l(X_i, \theta_0)\right],$$

where θ' lies between $\hat{\theta}$ and θ_0.

We now carry out the substitution $t = F(x, \theta_0)$ in (1). In so doing, assume that

$$F^{-1}(t, \theta) = \sup\{x : F(x, \theta) = t\},$$

and also that

4. Any two distributions from the family $\mathcal{F}$ are mutually absolutely continuous.

Then, as is easy to verify, $F[F^{-1}(t, \theta_0), \theta] = F(x, \theta)$ if $t = F(x, \theta_0)$. Therefore in (1) we can replace x by $F^{-1}(t, \theta_0)$. Using the notation presented in the Introduction, we obtain

$$\begin{aligned} u_n(t) &= \sqrt{n}[F_n(t) - G(t, \hat{\theta})] \\ &= \sqrt{n}[F_n(t) - t] - g(t) \frac{1}{\sqrt{n}} \sum_{i=1}^{n} l(T_i) + r_n(t, \hat{\theta}), \end{aligned} \tag{2}$$

where $g(t) = (\partial G/\partial \theta)\,(t, \theta_0)$, $l(t) = l[F^{-1}(t, \theta_0)\theta_0]$ and, according to Assertion I,

$$\int r_n^2(t, \hat{\theta}) \, dt = o_P(1) \qquad \text{as } n \to \infty. \tag{3}$$

We note that, from Condition 3,

$$\int l(t) \, dt = 0 \quad \text{and} \quad \int l^2(t) \, dt < \infty, \tag{4}$$

and that, from Condition 2,

$$g'(t) = \left(\frac{\partial^2}{\partial t \, \partial \theta} G\right)(t, \theta_0) = \left(\frac{\partial}{\partial \theta} \log f\right)(x, \theta_0), \quad \text{i.e.} \int [g'(t)]^2 \, dt < \infty. \tag{5}$$

Now let P, as in the Introduction, denote a Gaussian distribution in $L_2[0, 1]$ (see, for example, [3], p. 349, or [8], p. 28) with mean 0 and correlation function

$$R_P(t, s) = \min(t, s) - ts;$$

let $v(t)$ denote a process with distribution P (the so-called Brownian bridge). Below, we consider the stochastic Wiener integral

$$\langle \psi, v \rangle = \int \psi'(t) \, dv\,(t), \tag{6}$$

defined, as we know (see, for example, [8], pp. 87, 188), for any function ψ from the class C_1 of functions whose first derivative belongs to $L_2[0, 1]$. This integral is a bilinear functional on $C_1 \times L_2[0, 1]$, where, for $\psi(t)$ from the class $C_2(C_2 \subset C_1)$ of functions whose second derivative belongs to $L_2[0, 1]$, the integral $\langle \psi, v\rangle$ is a continuous linear functional on $L_2[0, 1]$. On $C_1 \times C_1$ the functional (6) is a scalar product.

For finite n, the sums

$$\sum_{i=1}^{n} \psi'(T_i)/\sqrt{n}, \quad \text{where } \psi(t) \in C_1 \quad \text{and} \quad \int \psi'(t)\, dt = 0,$$

can be conveniently assumed to be linear functionals of the process $v_n(t)$, and can be denoted by $\langle \psi, v_n\rangle$. In this notation Condition 3 can be rewritten in the form

3′. For a given function $l(t)$ satisfying (4), the representation

$$\sqrt{n}\,(\hat{\theta} - \theta_0) = \langle e, v_n\rangle + o_P(1), \quad n \to \infty, \quad \text{where } e(t) = \int_0^t l(t)\, dt$$

is valid.

We now consider the process

$$u(t) = v(t) - g(t)\langle e, v\rangle. \tag{7}$$

As follows from (4) and (5), the functions $g(t)$ and $e(t)$ belong to C_1.

II. *The process $u(t)$ is the projection of the process $v(t)$ if and only if $\langle e, g\rangle = 1$. At the same time $u(t)$ is a projection parallel to the function $g(t)$ on a subspace that is orthogonal* (*relative to* (6)) *to the function $e(t)$.*

Indeed, the equation $\Pi_1\Pi_1 x = \Pi_1 x$, where $\Pi_1 x = x(t) - g(t)\langle e, x\rangle$, is an operator defined for P-a.s. functions $x(t) \in L_2[0, 1]$, which is equivalent to the equation $\langle e, g\rangle = 1$. If this last equation is satisfied then $\langle e, \Pi_1 x\rangle = 0$ and $\Pi_1 g = 0$. Thus the statement is proved.

If $e(t) = \langle g, g\rangle^{-1} g(t)$, i.e., $l(x, \theta) = \varkappa^{-1}(\partial \log f/\partial\theta)\,(x, \theta)$, then $u(t)$ is an orthoprojection of the process $v(t)$ relative to (6). We note a characteristic property of an orthoprojection:

III. *The r.v. $\langle \psi, u\rangle$ and the r.v. $\langle e, v\rangle$ are independent for any function $\psi(t) \in C_1$ if and only if $e(t) = \langle g, g\rangle^{-1} g(t)$.*

It is worth recalling that the Brownian bridge $v(t)$ is also an orthoprojection, relative to (6), of a standard Wiener process $w(t)$ parallel to the function $j(t) \equiv t$: $v(t) = w(t) - t\langle j, w\rangle$. Thus $u(t)$ is a projection of a Wiener process parallel to the functions $j(t)$ and $e(t)$:

$$u(t) = \Pi w(t) = w(t) - j(t)\langle j, w\rangle - g(t)\langle e, w\rangle.$$

The condition in Assertion II before the substitution of variables has the form

$$\langle e, g\rangle = \int l(x, \theta_0)\Big(\frac{\partial}{\partial\theta} \log f\Big)(x, \theta_0)\, dF(x, \theta_0) = 1. \tag{8}$$

Supplementing II, we can say that if this condition is not fulfilled, then the operator Π_1 is one-to-one.

The process $u(t)$, as a linear transformation of a Gaussian process, is obviously a Gaussian process. Moreover:

IV. *If $\langle e, g\rangle = 1$, then the Gaussian distribution Q of the process $u(t)$ in $L_2[0, 1]$ has mean* 0 *and correlation function*

$$R_Q(t, s) = \min(t, s) - ts - g(t)e(s) - e(t)g(s) + \langle e, e\rangle g(t)g(s).$$

If $e(t) = \langle g, g\rangle^{-1}g(t)$, the correlation function $R_Q(t, s)$ assumes the form

$$R_Q(t, s) = \min(t, s) - ts - \langle g, g\rangle^{-1}g(t)g(s). \tag{9}$$

For the proof of Assertion IV we have to compute the mean and the correlation of the continuous linear functionals

$$\int \varphi_i(t)u(t)\,dt = -\langle \psi_i, u\rangle, \quad \text{where } \varphi_i(t) = \psi_i''(t) \in L_2[0, 1], \quad i = 1, 2.$$

But

$$\langle \psi_i, u\rangle = \langle \psi, \Pi w\rangle = \langle \Pi^*\psi_i, w\rangle,$$

where Π^* denotes the projector

$$\Pi^*\psi(t) = \psi(t) - j(t)\langle \psi, j\rangle - e(t)\langle \psi, g\rangle.$$

Now it is sufficient to recall (see, for example, [8], p. 88) that the mean of the r.v.'s $\langle \xi_i, w\rangle$, $\xi_i \in C_1$, $i = 1, 2$, is 0, while their correlation is $\langle \xi_1, \xi_2\rangle$, so that the mean of the r.v.'s $\langle \psi_i, u\rangle$, $i = 1, 2$, is 0, while the correlation equals $\langle \Pi^*\psi_1, \Pi^*\psi_2\rangle$. By direct computation it is easy to show that

$$\iint \varphi_1(t)\varphi_2(t)R_Q(t, s)\,dt\,ds = \langle \Pi^*\psi_1, \Pi^*\psi_2\rangle,$$

and, consequently, $R_Q(t, s)$ is indeed the correlation function of the distribution Q.

An expression of type (9) for the correlation function $R_Q(t, s)$ appeared in the very first papers devoted to the given category of problems, for the case of both scalar and vector parameters (see, for example, [5], or references in [1] and [2]), and the expression "Gaussian process" was applied, moreover, to the process $u(t)$. The exact formulation and proof of weak convergence of the processes $u_n(t)$ to $u(t)$ were given by J. Durbin considerably later, (in 1973; see [9]), for the space $C[0, 1]$ of continuous functions and for the space $D[0, 1]$ of functions without discontinuities of the second kind. The proof of weak convergence in $L_2[0, 1]$, which is necessary to us, looks very simple.

V. *If the relations* (3)–(5) *are fulfilled, then, as $n \to \infty$, the sequence of distributions Q_n of the processes $u_n(t)$ converges weakly in $L_2[0, 1]$ to the distribution Q of the process $u(t)$.*

In view of relation (3) it is sufficient to prove weak convergence for a sequence of processes $\Pi_1 v_n(t) = v_n(t) - g(t)\langle e, v_n\rangle$ (see, for example, [10], p. 25). Since the conditions of Assertion VI presented in Section 3 below are fulfilled, the sequence of the distributions of the processes $\Pi_1 v_n(t)$ is weakly compact. It remains for us to prove the weak convergence of the distributions of the continuous linear functionals $\langle\psi, \Pi_1 v_n\rangle$, $\psi(t)\in C_2$ to the normal distribution with mean 0 and variance $\langle\Pi^*\psi, \Pi^*\psi\rangle$. But it is easy to see that

$$\langle\psi, \Pi_1 v_n\rangle = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}[\psi'(T_i) - l(T_i)\langle\psi, g\rangle - \langle\psi, j\rangle]. \tag{10}$$

Since $T_1, \cdots, T_n$ are independent and identically (uniformly) distributed, expression (10) is a sum of independent identically distributed r.v.'s with mean 0 and variance $\langle\Pi^*\psi, \Pi^*\psi\rangle$; convergence to a normal distribution follows from the classical central limit theorem. Thus the statement is proved.

It is clear that the estimates themselves which satisfy Condition 3 converge weakly to the corresponding stochastic integrals

$$\sqrt{n}\,(\hat{\theta}-\theta) = \langle e, v_n\rangle + o_P(1) \Rightarrow \langle e, v\rangle, \qquad n\to\infty. \tag{11}$$

Thus Assertions II and VI show that substitution for the unknown value of the parameter of the estimate $\hat{\theta}$ which satisfies Conditions 3 and (8), is asymptotically equivalent to the projection of the process $v(t)$. Estimates satisfying Conditions 3 and (8) will be called *projecting* estimates. A particular case is presented by asymptotically efficient estimates (for example, estimates of maximum likelihood) for which the function

$$l(x, \theta) = \frac{1}{\varkappa}\frac{\partial \log f(x, \theta)}{\partial\theta},$$

leads, as already mentioned, to an orthogonal projection of the process $v(t)$. We must point out that an interesting discussion of such an orthogonal projection has been given by Yu. N. Tyurin (see [11], and also [9], [12]). Other estimates which are "not at all efficient" can also lead to projections, but not orthogonal ones. For example, for the estimate $\tilde{\theta}$ obtained by the method of moments, i.e., from the equation

$$nm_k(\theta) = \sum_{i=1}^{n}(X_i)^k, \quad \text{where } m_k(\theta) = \int x^k f(x, \theta)\,dx,$$

we have, under the usual regularity conditions,

$$\sqrt{n}\,(\tilde{\theta}-\theta_0) = \frac{1}{\sqrt{n}\,m_k'(\theta)}\sum_{i=1}^{n}[X_i^k - m_k(\theta)] + o_P(1), \qquad n\to\infty,$$

where $m_k'(\theta)$ is the derivative of the function $m_k(\theta)$ with respect to θ, so that

$$\frac{1}{m_k'(\theta)}\int x^k\left(\frac{\partial}{\partial\theta}f\right)(x, \theta)\,dx - \frac{m_k(\theta)}{m_k'(\theta)}\int\left(\frac{\partial}{\partial\theta}f\right)(x, \theta)\,dx = 1,$$

i.e., Condition (8) is fulfilled. Another example of projecting estimates is furnished by the so called M-estimates of Huber (see, for example, [13]); these are determined, in the case of a shift parameter, from the equation

$$\sum_{i=1}^{n}\xi(X_i - \theta^*) = 0,$$

where the function $\xi(x)$ satisfies the equation $\int \xi(x)\,dF(x)=0$, and is otherwise selected as a compromise between efficiency and stability (robustness) of the estimates θ^*. (We recall that the stability requirement usually leads to a decrease of efficiency of estimates.) For M-estimates, under the usual regularity conditions, the representation

$$\sqrt{n}\,(\theta^*-\theta_0)=\frac{1}{\sqrt{n}}\left[\int \xi(x)f'(x)\,dx\right]^{-1}\sum_{i=1}^{n}\xi(X_i-\theta_0)+o_P(1), \qquad n\to\infty,$$

holds, so that Condition (8) again is fulfilled and, consequently, the M-estimates are projecting estimates.

We note, in connection with (11), that stable estimates can be conceived as those which converge to continuous functionals of $v(t)$, i.e., for which $e(t)\in C_2$; asymptotically efficient estimates usually converge to measurable but not continuous linear functionals of $v(t)$, for which $e(t)\in C_1\backslash C_2$.

3. Weak Convergence in the Case of Alternatives

We now consider the limit behavior of the process $u_n(t)$ in the case of alternatives. Namely, we consider a certain sequence of d.f.'s $B_n(x)$, $n=1,2,\cdots$, and let, for each n, $X_{1n},\cdots,X_{nn}$ be independent r.v.'s with d.f.'s $B_n(x)$. It is again convenient for us to use the r.v.'s $T_{in}=F(X_{in},\theta_0)$, where the value θ_0 is determined by Condition 5 (see below). The d.f. of the r.v. T_{in} is $A_n(t)=B_n[F^{-1}(t,\theta_0)]$. Let $F_n^A(t)$ denote the d.f.'s of the r.v.'s $T_{1n},\cdots,T_{nn}$, and let

$$v_n^A(t)=\sqrt{n}\,[F_n^A(t)-t].$$

About the estimate $\hat{\theta}(X_{1n},\cdots,X_{nn})$ and the sequence of d.f.'s $B_n(x)$ we assume that a condition analogous to 3 is fulfilled:

5. There exists a value of the parameter θ_0 and a function $l(t)$ satisfying the condition in (4), such that, as $n\to\infty$,

$$\sqrt{n}\,(\hat{\theta}-\theta_0)=\langle e, v_n^A\rangle+o_P(1), \quad \text{where } e(t)=\int_0^t l(t)\,dt.$$

We now consider the process

$$u_n^A(t)=\sqrt{n}\,[F_n^A(t)-G(t,\hat{\theta})]=v_n^A(t)-g(t)\langle e, v_n^A\rangle+r_n(t,\hat{\theta}\rangle. \tag{12}$$

Below, we note that the remainder $r_n(t,\hat{\theta})$ can be neglected as before. Consequently, when considering weak convergence of the distributions of processes $u_n^A(t)$, we have to establish weak compactness of the sequence of distributions of the processes

$$\Pi_1 v_n^A(t)=v_n^A(t)-g(t)\langle e, v_n^A\rangle,$$

and the convergence of linear functionals of these processes.

We consider first the process

$$\Pi_1 v_{n,A}(t)=v_{n,A}(t)-g(t)\langle e, v_{n,A}\rangle, \quad \text{where } v_{n,A}(t)=\sqrt{n}\,[F_n^A(t)-A_n(t)].$$

Relative to weak compactness of a family of distributions of centered processes $\Pi_1 v_{n,A}$, the following assertion is valid:

VI. *Let the family A_l of* d.f.'*s on the segment* $[0, 1]$ *and the function* $l(t) \in L_2[0, 1]$ *be such that* $\int l^2(t)\, dA\,(t) \leqq c < \infty$ *for all* $A(t) \in A_l$. *Then, for any sequence* $A_n(t) \in A_l$, *the sequence of distributions of the processes*

$$\Pi_1 v_{n,A} = v_{n,A}(t) - g(t)\langle e, v_{n,A}\rangle$$

is weakly compact in $L_2[0, 1]$.

Let us write out the trace of the correlation operator R_A of the process $\Pi_1 v_{n,A}$ for some base $b_k(t)$, $k = 1, 2, \cdots$:

$$\operatorname{Tr} R_A = \sum_{k=1}^{\infty} \mathbf{E}\Big(\int b_k(t)\Pi_1 v_{n,A}(t)\, dt\Big)^2.$$

For the proof of weak compactness (see, for example, [3], Chap. V, §2) it is sufficient to prove that this series converges uniformly with respect to A and n. We have

$$\Big[\sum_{k=N}^{\infty} \mathbf{E}\Big(\int b_k \Pi_1 v_{n,A}\Big)^2\Big]^{1/2} \leqq \Big[\sum_{k=N}^{\infty} \mathbf{E}\Big(\int b_k v_{n,A}\Big)^2\Big]^{1/2} + \Big[\sum_{k=N}^{\infty}\Big(\int b_k g\Big)^2\Big]^{1/2} [\mathbf{E}\langle e, v_{n,A}\rangle^2]^{1/2}.$$

Since

$$\mathbf{E}\Big(\int b_k v_{n,A}\Big)^2 = \int c_k^2\, dA - \Big(\int c_k\, dA\Big)^2 \leqq \int c_k^2\, dA, \quad \text{where } c_k(t) = \int_0^t b_k(t)\, dt,$$

choosing, say, $\sin \pi k t$, as $b_k(t)$, we find that

$$\mathbf{E}\Big(\int b_k v_{n,A}\Big)^2 \leqq 1/\pi^2 k^2.$$

This implies that as $N \to \infty$, the remainder $\sum_{k=N}^{\infty} \mathbf{E}(\int b_k v_{n,A})^2$ converges to 0, uniformly with respect to A and n. Since, in addition,

$$\mathbf{E}\langle e, v_{n,A}\rangle^2 = \int l^2\, dA - \Big(\int l\, dA\Big)^2 \leqq c,$$

the second remainder in the right side of the inequality also converges to 0 uniformly with respect to A and n. Consequently, the series $\operatorname{Tr} R_A$ also converges uniformly with respect to A and n. The statement is proved.

From the uniform convergence of the series $\sum_{k=1}^{\infty} \mathbf{E}(\int b_k v_{n,A})^2$, we note the weak compactness of the sequence of distributions of the processes $v_{n,A}(t)$ for any sequence of d.f.'s $A_n(t)$. We note, in addition, that the correlation operators of the processes $v_{n,A}(t)$ and $\Pi_1 v_{n,A}(t)$ do not depend on the value of the index n but only on the d.f.'s $A_n(t)$ and coincide with the correlation operators of the processes $v(A_n(t))$ and $\Pi_1 v(A_n(t))$, respectively, where $v(t)$ is the Brownian bridge. Consequently, Assertion VI is equivalent to the following: *if* $(\int l^2(t)\, dA\,(t)| \leqq c < \infty$ *for* $A(t) \in A_l$, *then the family of distributions of the processes* $\Pi_1 v(A(t))$, $A(t) \in A_l$, *is weakly compact*; *the family of distributions of the processes* $v(A(t))$, $A(t) \in A$, *where* A *is the set of all* d.f.'*s on the segment* $[0, 1]$, *is weakly compact.*

Since $v_n^A(t) = v_{n,A}(t) + \sqrt{n}\,[A_n(t) - t]$, for the weak compactness of the sequence of distributions of the processes $v_n^A(t)$ it is necessary and sufficient that the sequence of functions $\sqrt{n}\,[A_n(t) - t]$ be compact in $L_2[0, 1]$, while for weak convergence of the distributions, as Assertion VI shows, it is necessary and sufficient that this sequence of functions converge in $L_2[0, 1]$:

$$\int [\sqrt{n}\,(A_n(t) - t) - a(t)]^2\,dt \to 0, \qquad n \to \infty. \tag{13}$$

As D. M. Chibisov informed the author, this Assertion is a consequence of the more general result of [14].

In order to reach weak convergence of distributions of the processes $\Pi_1 v_n^A(t)$, we consider conditions under which the sequence of distributions of the r.v. $\langle e, v_n^A \rangle$ weakly converges to a normal distribution. In contrast to the case of measurable linear functionals of the processes $v_n(t)$, the convergence of measurable linear functionals of the processes $v_n^A(t)$ does not follow from weak convergence of the processes. Additional assumptions are necessary for such convergence.

VII. *Assume that*

(a) $\int_{|l|>c} l^2(t)\,dA_n(t) \to 0$ *as* $c \to \infty$ *uniformly with respect to* n,

(b) $\int l^2(t)\,dA_n(t) \to \int l^2(t)\,dt$ *as* $n \to \infty$,

(c) $\sqrt{n} \int l(t)\,d[A_n(t) - t$ *converges as* $n \to \infty$.

Then $\langle e, v_n^A \rangle \Rightarrow \langle e, v \rangle + \langle e, a \rangle$, *where* $\langle e, a \rangle$ *denotes the limit of the sequence from Condition* (c).

Indeed $\langle e, v_n^A \rangle = \langle e, v_{n,A} \rangle + \sqrt{n}\langle e, A_n - j \rangle$, where according to Condition (c) the second term on the right has the limit $\langle e, a \rangle$. But the r.v. $\langle e, v_{n,A} \rangle$ is the normed sum

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(l(T_{in}) - \int l\,dA_n \right)$$

of independent r.v.'s with mean 0 and variance $\int l^2\,dA_n - (\int l\,dA_n)^2$. For the proof of convergence of distributions of these sums to the normal distribution we test whether the Lindeberg condition is satisfied. We have:

$$\left[\int_M \left(l - \int l\,dA_n \right)^2 dA_n \right]^{1/2} \leqq \left[\int_M l^2\,dA_n \right]^{1/2} + \left| \int l\,dA_n \right| \int_M dA_n,$$

where $M = \{t \colon |l(t) - \int l\,dA_n| > \varepsilon\sqrt{n}\}$. Here, for any $\varepsilon > 0$,

$$\int_M dA_n \leqq \int_M l^2\,dA_n \to 0 \quad \text{as } n \to \infty$$

from Condition (a); hence it follows from the inequality presented above that the Lindeberg condition is fulfilled. Therefore distributions of the r.v. $\langle e, v_{n,A} \rangle$ weakly converge to a normal distribution with mean 0 and variance, in view of Conditions (b) and (c), equal to $\int l^2\,dt - (\int l\,dt)^2$. This coincides with the variance of the r.v. $\langle e, v \rangle$, and the assertion is proved.

Relative to the set of functions $l(t)$ from $L_2[0, 1]$, for which the conditions of Assertion VII are fulfilled for a given sequence of alternatives $A_n(t)$, we can easily prove the following assertion:

VIII. *A set of functions $l(t)$ from $L_2[0, 1]$, for which the conditions of Assertion* VII *are fulfilled, forms a linear manifold. If the convergence* (13) *is valid, then this manifold contains the class C_1.*

If Conditions (a), (b) and (c) of Assertion VII are fulfilled for the function $l(t)$, then, obviously, they are fulfilled for a function $\gamma l(t)$ for any γ. If Conditions (b) and (c) are fulfilled for the functions $l_1(t)$ and $l_2(t)$, then it is also obvious that they are fulfilled for the function $l_1(t)+l_2(t)$. We consider Condition (a). We have:

$$\left[\int_{|l_1+l_2|>c} (l_1+l_2)^2\, dA_n\right]^{1/2} \leqq \left[\int_{|l_1+l_2|>c} l_1^2\, dA_n\right]^{1/2} + \left[\int_{|l_1+l_2|>c} l_2^2\, dA_n\right]^{1/2}.$$

In turn,

$$\begin{aligned}\int_{|l_1+l_2|>c} l_1^2\, dA_n &\leqq \int_{|l_1|>c/2} l_1^2\, dA_n + \int_{|l_1|<c/2<|l_2|} l_1^2\, dA_n \\ &\leqq \int_{|l_1|>c/2} l_1^2\, dA_n + \int_{|l_2|>c/2} l_2^2\, dA_n.\end{aligned}$$

Using an analogous estimate for $\int_{|l_1+l_2|>c} l_2^2\, dA_n$ we obtain:

$$\left[\int_{|l_1+l_2|>c} (l_1+l_2)^2\, dA_n\right]^{1/2} \leqq 2\left[\int_{|l_1|>c/2} l_1^2\, dA_n + \int_{|l_2|>c/2} l_2^2\, dA_n\right]^{1/2};$$

whence it follows that if Condition (a) is fulfilled for $l_1(t)$ and $l_2(t)$, then it is also fulfilled for $l_1(t)+l_2(t)$.

We shall show now that Conditions (a), (b) and (c) are fulfilled for any function $l(t)\in C_1$ if the convergence (13) holds. Since $l(t)$ is bounded on $[0, 1]$, Condition (a) is fulfilled for any sequence of d.f.'s $A_n(t)$. Since $l_2(t)\in C_1$ if $l(t)\in C_1$, then using integration by parts (see, for example, [15], p. 148) we find that (b) and (c) follow from (13). The assertion is proved.

IX. *If Conditions* 1, 2, 5, (13) *and the conditions of Assertion* VII *are fulfilled, then the sequence of distributions Q_n^A of the processes $u_n^A(t)$ weakly converges in $L_2[0, 1]$, as $n\to\infty$, to the Gaussian distribution Q^A of the process*

$$u^A(t) = u(t) + a(t) - g(t)\langle e, a\rangle.$$

Indeed, from Conditions 1 and 5 and the estimate $\langle e, v_n^A\rangle = O_P(1)$ (which follows from Assertion VII) we conclude that as $n\to\infty$ the remainder $r_n(t, \hat\theta)$ in formula (12) satisfies the relation

$$\int r_n^2(t, \hat\theta)\, dt = o_P(1), \qquad n\to\infty.$$

Therefore it is again sufficient to consider the processes $v_n^A(t) - g\langle e, v_n^A\rangle$. The convergence of distributions of continuous linear functionals of these processes:

$$\langle \psi, v_n^A\rangle - \langle \psi, g\rangle\langle e, v_n^A\rangle$$

to a Gaussian distribution with mean $\langle \psi, \Pi_1 a\rangle$ and variance $\langle \Pi^*\psi, \Pi^*\psi\rangle$ follows from Assertions VII and VIII. Weak compactness of the sequence of distributions of these processes follows from Assertion VI, Condition (13) and Condition (c) of Assertion VII. The proof is completed.

We return to the interpretation of stable estimates, mentioned at the end of Section 2, as those estimates, satisfying Condition 5, which converge in probability to continuous functionals $\langle e, v_n^A\rangle$ of the processes v_n^A. Then in connection with Assertion IX we may note that if the process $u_n^A(t)$ is obtained by means of a stable estimate, then weak convergence to the distribution Q^A takes place for any sequence of alternatives satisfying Condition (13).

4. Linear Transformations and Linear Functionals of the Process $u(t)$

As was mentioned in the Introduction, the process $u(t)$ admits the expansion

$$u(t)=\sum_{k=1}^{\infty} \mu_k U_k a_k(t),$$

where $a_k(t)$, $k=1,2,\cdots$, are normalized eigenfunctions, μ_k^2, $k=1,2,\cdots$, are eigenvalues of the correlation function $R_Q(t,s)$, and U_k, $k=1, 2, \cdots$, are independent r.v.'s having a standard normal distribution. Finding μ_k and $a_k(t)$ is, as was mentioned, a difficult task. Furthermore, μ_k and $a_k(t)$, $k=1, 2, \cdots$, depend on $g(t)$, so that they have to be determined anew for different hypothetical families $\mathcal{F}$ (if not for different subhypotheses $F(x,\theta)$ from the same family). Consequently, the distribution of the quadratic form

$$\int u^2(t)\,dt=\sum_{k=1}^{\infty} \mu_k^2 U_k^2$$

must also be computed anew for different hypotheses. Therefore it is difficult to use the statistic $\int u_n^2(t)\,dt$. On the other hand, Assertion II indicates how to construct the quadratic form

$$\iint K(t,s)\,du\,(t)\,du\,(s)=\sum_{k=1}^{\infty} \lambda_k^2 Z_k^2 \tag{14}$$

with given coefficients λ_k^2 of the independent r.v. Z_k, $k=1,2,\cdots$, with standard normal distribution. It is obvious that the construction of such a quadratic form is equivalent to the construction of a linear transformation

$$z(t)=\int H(t,s)\,du\,(s), \tag{15}$$

such that $\int z^2(t)\,dt=\sum \lambda_k^2 Z_k^2$.

As we know (see, for example, [3], p. 237), the linear transformation (15) is defined P-a.s. for all Hilbert–Schmidt kernels $H(t,s)$ (i.e., such that $\iint H^2(t,s)\,dt\,ds<\infty$). Let $\mathcal{H}$ denote a class of Hilbert–Schmidt kernels $H(t,s)$ for which

$$\int H(t,s)g'(s)\,ds=\int H(t,s)\,ds=0.$$

When considering linear transformations of the process $u(t)$, it is sufficient to deal with kernels of the class $\mathcal{H}$. Indeed, as follows from Assertion II, for any

Hilbert–Schmidt kernel $H(t, s)$ we have

$$\int H(t, s)\, du\,(s) \tag{16}$$
$$= \int \left[H(t, s) - \int H(t, s)\, ds - \langle g, g\rangle^{-1} g'(s) \int H(t, s) g'(s)\, ds\right] du\,(s),$$

and the kernel in the right side belongs to $\mathscr{H}$.

Let now $h_k(t)$, $k = 1, 2, \cdots$, be a normalized system of functions from $L_2[0, 1]$, each of which is orthogonal to $j'(t) = 1$ and $g'(t)$, i.e.,

$$\int h_k(t) h_r(t)\, dt = \delta_{kr}, \quad \int h_k(t)\, dt = 0, \quad \int h_k(t) g'(t)\, dt = 0, \tag{17}$$

or, in the notation adopted above,

$$\langle e_k, e_r\rangle = \delta_{kr}, \quad \Pi^* e_k = e_k, \quad \text{where } e_k(t) = \int_0^t h_k(t)\, dt. \tag{18}$$

X. *If the functions $h_k(t)$, $k = 1, 2, \cdots$, satisfy conditions* (17), *then the* r.v.'*s*

$$Z_k = \int h_k(t)\, du\,(t), \qquad k = 1, 2, \cdots,$$

form a sequence of independent r.v.'s having standard normal distribution.

Indeed, since $u(t) = \Pi w(t)$,

$$\langle e_k, u\rangle = \langle \Pi^* e_k, w\rangle = \langle e_k, w\rangle$$

in view of (18), while, in view of the first condition in (17), any linear combination

$$\sum \lambda_k Z_k = \langle \sum \lambda_k e_k, w\rangle, \quad \text{where } \sum \lambda_k^2 < \infty,$$

has normal distribution with mean 0 and the variance $\sum \lambda_k^2$; this in fact proves X.

XI. *For any correlation function $R(t, s)$ there exists a kernel $H(t, s) \in \mathscr{H}$ such that the process $z(t) = \int H(t, s)\, du\,(s)$ will have the correlation function $R(t, s)$.*

Indeed, let $m_k(t)$ and λ_k^2, $k = 1, 2, \cdots$, be a normalized sequence of eigenfunctions and a sequence of eigenvalues of the correlation function $R(t, s)$, and let $h_k(t)$, $k = 1, 2, \cdots$, be any sequence of functions from $L_2[0, 1]$, which satisfies (17). We put

$$H(t, s) = \sum \lambda_k m_k(t) h_k(s).$$

Then the process

$$z(t) = \int H(t, s)\, du\,(s) = \sum \lambda_k Z_k m_k(t) \tag{19}$$

will have the correlation function $R(t, s)$.

It is clear that here

$$\int z^2(t)\, dt = \sum \lambda_k^2 Z_k^2. \tag{20}$$

Turning to linear transformations of the process $u^A(t)$, we assume that $a(t)\in C_1$. We recall that this condition and the condition $g(t)\in C_1$ together are necessary and sufficient for mutual absolute continuity of the distributions Q and Q^A. Here the r.v. is

$$\langle e_k, u^A\rangle = Z_k+\alpha_k, \quad \text{where } \alpha_k=\int h_k(t)a'(t)\,dt,$$

and the process is

$$z^A(t)=\int H(t,s)\,du^A(s)=z(t)+c(t),$$

where $c(t)=\int H(t,s)a'(s)\,ds=\sum \lambda_k\alpha_k m_k(t)$, so that

$$\int [z^A(t)]^2\,dt=\sum \lambda_k^2(Z_k+\alpha_k)^2. \tag{21}$$

For finite n, according to (12),

$$\langle e_k, u_n^A\rangle=\langle e_k, v_n^A\rangle+\langle e_k, r_n\rangle, \tag{22}$$

with

$$\int h_k(t)\,dv_n^A\,(t)=\frac{1}{\sqrt{n}}\sum_{i=1}^n h_k(X_{in},\theta_0). \tag{23}$$

Consequently, in order to establish the weak convergence

$$\langle e_k, u_n^A\rangle \Rightarrow Z_k+\alpha_k, \tag{24}$$

we need, in addition to the conditions of Assertion VII, conditions under which

$$\int h_k(t)\,dr_n\,(t,\hat\theta)=o_P(1), \qquad n\to\infty$$

This of course does not yet follow from the estimate $\int r_n^2(t,\theta)\,dt=O_p(1)$, $n\to\infty$. We strengthen Condition 1 in the following way:

6. For all $\theta\in\Theta$,

$$\int\left[\left(\frac{\partial}{\partial\theta}\log f\right)(x,\theta')-\left(\frac{\partial}{\partial\theta}\log f\right)(x,\theta)\right]^2 dF\,(x,\theta)=o(1) \quad \text{as } |\theta'-\theta|\to 0.$$

Then:

XII. *If for the functions $h_k(t)$, $l(t)$ and the sequence of* d.f.'s *$A_n(t)$, the conditions of Assertion* VII *and Condition* 6 *are fulfilled, then the weak convergence* (24) *holds. Consequently,*

$$z_n^A(t)\Rightarrow z^A(t) \quad \text{and} \quad \int [z_n^A(t)]^2\,dt\Rightarrow\sum\lambda_k^2(Z_k+\alpha_k)^2, \tag{25}$$

where $z_n^A(t)=\int H(t,s)\,du_n^A\,(s)$.

For the proof it is sufficient to note that if $l(t)$ satisfies the conditions of Assertion VII, then $\langle e, v_n^A\rangle = O_P(t)$, $n\to\infty$, which together with Condition 6 gives

$$\int [r_n'(t,\hat\theta)]^2\,dt = o_P(1), \qquad n\to\infty.$$

Consequently, (24) follows from (22).

In connection with XII we can say that for $h_k(t)\in C_1$ with $\alpha_k = -\int h_k'(t)a(t)\,dt$, the convergence (24) holds from Assertion IX even without the assumption that $a(t)\in C_1$. That is, when the limit distributions Q and Q^A are orthogonal, the tests based on continuous (linear) functionals at all levels less than unity also have an asymptotic power less than unity. For $h_k(t)\in L_2[0,1]\backslash C_1$ the condition $a(t)\in C_1$ is necessary but not sufficient for (24), so that tests based on linear measurable functionals can also have a power tending to unity when the limit distributions Q and Q^A are mutually absolutely continuous.

Assertions X, XI and XII, in the context of the problem under consideration, signify that if a system of functions $h_k(t)$, $k=1,2,\cdots$, satisfying (17) exists, we are essentially free to choose the normalized functions $m_k(t)$ and the numbers λ_k, provided only that $\sum\lambda_k^2<\infty$. For example, if we choose

$$m_k(t)=\sqrt{2}\sin\pi kt, \qquad \lambda_k = 1/\pi k,$$

then $z_n(t)$ will converge to the Brownian bridge, while if

$$m_k(t)=\sqrt{2}\sin\pi(k+\tfrac12)t, \qquad \lambda = 1/\pi(k+\tfrac12),$$

then $z_n(t)$ will converge to a Wiener process. For the construction of quadratic forms the form of the functions $m_k(t)$ is not important: the choice of the numbers λ_k^2, $k=1,2,\cdots$, completely determines the limit distribution of the statistic

$$\int [z_n(t)]^2\,dt = \sum\lambda_k^2\left[\frac{1}{\sqrt{n}}\sum h_k(T_i)\right]^2 + o_P(1), \qquad n\to\infty, \tag{26}$$

for the hypothesis case. That is, the distribution of the quadratic form (20), is such that, for example, for $\lambda_k = 1/\pi k$ we obtain the ω^2 distribution, while for $\lambda_k=1$, $k\leqq r$, and $\lambda_k = 0$, $k>r$, we obtain a chi-square distribution with r degrees of freedom, and so forth.

For the hypothesis case it is also not important exactly how the functions $h_k(t)$ satisfying (17) are chosen: As (24) shows, however, the power of the test based on the r.v.'s (23), (for example, on the quadratic statistic (26)), depends substantially on the choice of the functions $h_k(t)$, specifically, on how fast the Fourier coefficients α_k of the function

$$a'(t)-g'(t)\langle e, a\rangle$$

tend to zero as $k\to\infty$, for some sequence of alternatives. Recall that we encounter the same situation when we use the usual ω^2; test its power depends on how fast the Fourier coefficients

$$\sqrt{2}\int a'(t)\cos\pi kt\,dt$$

decrease.

Using the equation $t = F(x, \theta_0)$, Conditions (17) can be rewritten in the form

$$\int h_k(x,\theta)h_r(x,\theta)\,dF(x,\theta)=\delta_{kr},\qquad \int h_k(x,\theta)\,dF(x,\theta)=0,$$

(27)

$$\int h_k(x,\theta)\left(\frac{\partial}{\partial\theta}f\right)(x,\theta)\,dx=0,$$

where $h_k(x, \theta) = h_k(F(x, \theta))$, and hence $\theta = \theta_0$. We now turn our attention to the fact that the functions $h_k(x, \theta)$ depend on an unknown value of the parameter θ, and, consequently, the r.v.'s (23) cannot be constructed. However, it can be shown in a standard way that for the r.v.'s

$$\langle \hat{e}_k, u_n^A\rangle = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} h_k(X_{in},\hat{\theta}),\qquad k=1,2,\cdots, \tag{28}$$

the relation (24) again is valid. For example, let the following condition be fulfilled.

7. For each $\theta \in \Theta$ the function

$$\sup_{\theta-\varepsilon\leqq\theta'\leqq\theta+\varepsilon}\left|\frac{\partial}{\partial\theta}h_k(x,\theta')\right| = M_k(x,\theta,\varepsilon)$$

satisfies the relation

$$\int M_k(x,\theta,\varepsilon)\,dB_n(x) = o(1)\quad \text{as } \varepsilon\to\infty$$

uniformly with respect to n.

Then it is easy to see that:

XIII. *If Condition 7 is fulfilled and $\sqrt{n}(\hat{\theta}-\theta) = O_P(1)$ as $n \to \infty$, then*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} h_k(X_{in},\hat{\theta}) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} h_k(X_{in},\theta_0)+o_P(1)\quad \text{as } n\to\infty.$$

From XIII it follows that, for any finite r,

$$\sum_{k=1}^{r}\lambda_k^2\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n} h_k(X_{in},\hat{\theta})\right]\Rightarrow\sum_{k=1}^{r}\lambda_k^2 Z_k^2.$$

If in addition to XIII we can assert that $\sum_{i=1}^n h_k(X_{in}, \hat{\theta})/\sqrt{n} = O_P(1)$, $n \to \infty$, uniformly with respect to k, then an analogous relation will be true also for infinite sums, provided only that $\sum \lambda_k^2 < \infty$. Incidentally this is in fact not important: as we know (see [1]), the distribution of infinite quadratic form $\sum \lambda_k^2 Z_k^2$ can be very accurately approximated by a distribution of a finite quadratic form $\sum_{k=1}^r \lambda_k^2 Z_k^2 + \eta_r Z_{r+1}^2$, replacing the remainder $\sum_{k=r+1}^{\infty} \lambda_k^2 Z_k^2$ of the r.v. $\eta_r Z_{r+1}^2$ by some multiplier of η_r or even simply by a constant.

5. Some Remarks

We make a few concluding observations.

(1) The functions $h_k(x, \theta)$ satisfying conditions (27) can be obtained by means of the process of orthogonalization. If $h_k(t)$, $k = 1, 2, \cdots$, is a normalized

system in $L_2[0, 1]$, then the functions $h_k[F(x, \theta)]$ satisfy the first two conditions in (27). In order to satisfy the third, we carry out an orthogonalization process. It is clear that this process must be repeated for each hypothesis (or even for each simple subhypothesis), but this, as may be imagined, is considerably simpler than to compute the eigenvalues μ_k^2 and then the d.f.'s of the corresponding quadratic form.

We point out, (as mentioned in a similar connection in [16]), that to compute statistics of the form (28) it is not necessary to order the observations according to magnitude, as in the computation of the statistics $\int u_n^2(t)\, dt$ and $\int v_n^2(t)\, dt$.

(2) Statistics of the form (28) are of course quite analogous to the "components" of ω^2 statistics, i.e., the r.v.'s

$$\sqrt{2} \int v_n(t) \sin \pi kt\, dt \quad \text{and} \quad \int u_n(t) a_k(t)\, dt,$$

considered in [16] and [17]. When constructing the statistics in (28), only one additional circumstance is taken into account, namely, that it is not absolutely necessary to use the above mentioned "eigen" components and that to obtain the statistics in (28) with the functions $h_x(x, \theta)$ we have chosen is both much easier and more meaningful for the actual purposes of hypothesis verification. Specifically, it is easy to see that for any intentional choice of the functions $h_k(x, \theta)$ (i.e., such that the coefficients for the "expected" alternatives $a(t)$ are "not too small"), two or three, or even one statistic of the form (28), will lead to a test with "sufficiently good" power; this power as a rule, is greater than the power of the tests based on the statistics $\int u_n^2(t)\, dt$ and $\int v_n^2(t)\, dt$. Tables 3 and 4 in [16] and Tables 3, 4, 5 and 8 in [17] can serve as illustrations.

Thus the choice of the functions $h_k(x, \theta)$ and of the number of statistics (28) can be used to construct tests which are consistent for a fairly broad class of alternatives (although not for all), and which have a greater power with respect to the expected alternatives. Finally, the choice of the functions $h_k(x, \theta)$ can also be subordinated, to other goals, for example, to the condition of stability (robustness) of the test.

(3) In view of the third condition in (27), the asymptotic properties of the statistics in (28) (both in the case of a hypothesis and in the case of a sequence of alternatives), do not depend on whether the estimate $\hat{\theta}$, used in the construction of the process $u_n(t)$, is projecting or not, or on how "efficient" it is (more precisely, what its asymptotic variance is). In essence, it is only important that $\sqrt{n}\,(\hat{\theta} - \theta_0) = O_P(1)$ as $n \to \infty$.

For the linear statistic $\langle \psi, u_n \rangle$, where $\langle \psi, g \rangle \neq 0$, the use of some projecting estimate is important: when the conditions of Assertion VII are fulfilled, the asymptotic mean of this statistic in the case of alternatives is $\langle \psi, \Pi a \rangle$, while in the case of a hypothesis it is 0; the asymptotic variance in both cases is $\langle \Pi^* \psi, \Pi^* \psi \rangle$. Here it is easy to see that the use of a projecting estimate instead of an exact (say, unknown) value of the parameter θ_0 can lead to an increase of the power of the test. Thus, the asymptotic mean of the statistic $\langle \psi, v_n \rangle$ in the case of alternatives is $\langle \psi, a \rangle$, while in the case of a hypothesis it is 0, the asymptotic variance in both cases being $\langle \psi, \psi \rangle$ which, for example, is greater than $\langle \Pi^* \psi, \Pi^* \psi \rangle$ if one uses an orthogonally projecting estimate. Therefore, for alternatives for which $a = \Pi a$

the power of a test based on the statistic $\langle\psi, u_n\rangle$ is greater than that of a test based on the statistic $\langle\psi, v_n\rangle$.

(4) Although for a *specified* statistic of an empirical process the choice of projecting estimate to use is important, for the problem of "overall" verification of a given parametric hypothesis the choice in essence does not matter. In particular, for any given sequence of alternatives we can use an orthogonally projecting estimate, with, as was noted, the least asymptotic variance, to achieve the same asymptotic power as any other projecting estimate.

Indeed, let $u_{1n}(t)$ and $u_{2n}(t)$ be sequences of processes obtained by means of the projecting estimates $\hat{\theta}_1$ and $\hat{\theta}_2$, respectively, which weakly converge to the processes $\Pi_1 v(t)$ and $\Pi_2 v(t)$, where the operator Π_j projects onto a subspace that is orthogonal to the function $e_j(t)$, $j=1, 2$. Then

$$\hat{\Pi}_2 u_{1n}(t) = u_{1n}(t) - \hat{g}(t)\langle\hat{e}_2, u_{1n}\rangle \Rightarrow \Pi_2(t),$$

where $\hat{g}(t) = (\partial G/\partial\theta)(t, \hat{\theta})$ $l_2(t) = l_2[F^{-1}(t, \theta_0), \hat{\theta}]$, $\hat{e}_2(t) = \int_0^t l_2(t)\,dt$, and for $\hat{\theta}$ we can use either $\hat{\theta}_1$ or $\hat{\theta}_2$ (or any other consistent estimate). Consequently, a test based on a (P-a.s. continuous) statistic $T(u_{2n}(t))$ of the process $u_{2n}(t)$ is asymptotically equivalent to a test based on the statistic $T[\hat{\Pi}_2 u_{1n}(t)]$ of the process $u_{1n}(t)$.

The circumstance mentioned above may be noted also in a simpler context. Let $\mathbf{X} = (X_1, \cdots, X_n)^T$ be a random vector with independent coordinates having normal distribution and variance 1. We denote by the symbol $\mathbf{1}$ an n-dimensional vector $(1, \cdots, 1)^T$. We consider a hypothesis consisting of the fact that the means θ_i of the r.v. X_i, $i=1, \cdots, n$, are equal, i.e., that for a certain number $\hat{\theta}$ the vector $\mathbf{X} - \theta\mathbf{1}$ has the mean 0. Let the n-dimensional vector $\mathbf{1}$ be such that $\mathbf{I}^T\mathbf{1} = \mathbf{1}$. Then the r.v. $\mathbf{I}^T\mathbf{X}$ is the projecting estimate θ. For example, if $\mathbf{I} = (1/n)\mathbf{1}$, the estimate

$$\mathbf{1}^T\mathbf{X} = \sum_{i=1}^n X_i/n$$

is the nonbiased estimate with the least variance equal to $1/n$, while in the case $\mathbf{I} = (1, 0, \cdots, 0)^T$ the estimate $\mathbf{I}^T\mathbf{X} = X_1$ has variance 1. However, it is obvious that for all $\mathbf{I}$ the projection $\mathbf{Y} = \mathbf{X} - \mathbf{1}\mathbf{I}^T\mathbf{X}$ is the maximum invariant relative to shifts, and the tests based on $\mathbf{Y}$ for different $\mathbf{I}$ are equivalent.

Received by the editors
November 1, 1977

REFERENCES

[1] J. DURBIN, *Distribution theory for tests based on the sample distribution function*, Res. Conf. Ser. in Appl. Math., 9th issue, SIAM.

[2] G. V. MARTYNOV, *The Omega-Square Criterion*, Nauka, Moscow, 1978. (In Russian.)

[3] I. I. GIKHMAN AND A. V. SKOROKHOD, *The Theory of Stochastic Processes*, vol. I, Springer, New York, 1974.

[4] I. I. GIKHMAN, *On a problem of the theory of the ω^2-test*, Matem. Sbirn. Kiyvs'k. Derzh. Univ., 5 (1954), pp. 51–59. (In Ukrainian.)

[5] M. KAC, J. KIEFER AND J. WOLFOWITZ, *On a test of goodness-of-fit based on a distance method*, Ann. Math. Statist., 26, 2 (1955), pp. 189–211.

[6] G. V. MARTYNOV, *Computation of the limit distributions of statistics of normality tests of type ω^2*, Theory Prob. Appl., 21, 1 (1976), pp. 3–13.
[7] M. A. STEPHENS, *Asymptotic results for goodness-of-fit statistics with unknown parameters*, Ann. Statist., 4, 2 (1976), pp. 357–369.
[8] N. N. KUO, *Gaussian measures in Banach spaces*, Lect. Notes Math., 453, 1975.
[9] J. DURBIN, *Weak convergence of the sample distribution function when parameters are estimated*, Ann. Statist., 1, 2 (1973), pp. 279–290.
[10] P. BILLINGSLEY, *Convergence of Probability Measures*, Wiley, New York, 1968.
[11] YU. N. TYURIN, *A linear model of a multi-dimensional nonparametric statistic*, in *Multi-Dimensional Statist. Analysis in Soc.-Econ. Investigations*, Nauka, Moscow, 1974. (In Russian.)
[12] YU. N. TYURIN, *On parametric hypotheses testing with non-parametric tests*, Theory Prob. Appl., 15, 4 (1970), pp. 722–726.
[13] P. HUBER, *Robust estimation of location parameter*, Ann. Math. Statist., 35, 1 (1964), pp. 73–101.
[14] D. M. CHIBISOV, *An investigation of the asymptotic power of the test of fit*, Theory Prob. Appl., 10, 3 (1965), pp. 421–437.
[15] W. FELLER, *An Introduction to Probability Theory and Its Applications*, vol. II, Wiley, New York, 1966.
[16] J. DURBIN AND M. KNOTT, *Components of Cramér-von Mises statistics, I*, J. Royal Statist. Soc., B 34, 2 (1972), pp. 290–307.
[17] J. DURBIN, M. KNOTT AND C. C. TAYLOR, *Components of Cramér-von Mises Statistics, II*, J. Royal Statist. Soc., B 37, 2 (1975), pp. 216–237.