A PROPERTY OF THE LOG-LIKELIHOOD-RATIO PROCESS* FOR GAUSSIAN PROCESSES

B. L. S. PRAKASA RAO AND HERMAN RUBIN

(Received Apr. 18, 1967)

The aim of this paper is to give a necessary and sufficient condition for a Gaussian process to be the log-likelihood ratio process of Gaussian processes. We assume for simplicity that all the measures encountered hereafter are mutually absolutely continuous.

Let X be multivariate normally distributed with unknown mean θ and known dispersion matrix Σ. We assume that Σ is non-singular without loss of generality. Let $p_\theta(x)$ denote the density of X with respect to the Lebesgue measure on R^n. Hereafter $E[Y\mid\psi]$ denotes the expectation of Y when ψ is the true parameter. Define

$$L(X,\theta,\varphi)=\log p_\theta(X)-\log p_\varphi(X)\,. \tag{1}$$

Since

$$p_\theta(X)=(2\pi)^{-n/2}\,|\,\Sigma\,|^{-1}\exp\left\{-\frac{1}{2}(X-\theta)'\Sigma^{-1}(X-\theta)\right\} \tag{2}$$

where $|\,\Sigma\,|$ denotes the determinant of Σ and X' denotes the transpose of the vector X, it is easily seen that

$$L(X,\theta,\varphi)=-\frac{1}{2}[(2X-\theta-\varphi)'\Sigma^{-1}(\varphi-\theta)]\,. \tag{3}$$

Now

$$\begin{aligned} E[L(X,\theta,\varphi)\mid\psi]&=-\frac{1}{2}[(2\psi-\theta-\varphi)'\Sigma^{-1}(\varphi-\theta)]\\ &=\frac{1}{2}(\psi-\varphi)'\Sigma^{-1}(\psi-\varphi)\\ &\quad-\frac{1}{2}(\psi-\theta)'\Sigma^{-1}(\psi-\theta)\,, \end{aligned} \tag{4}$$

* Prepared under Contract No. NONR-2587(02) while B.L.S. P. Rao was at Michigan State University. Reproduction in whole or in part is permitted for any purpose of the U.S. Government.

and

$$(5)\quad \begin{aligned} &\operatorname{Cov}[L(X, \theta_1, \varphi_1), L(X, \theta_2, \varphi_2) \mid \psi] \\ &= \frac{1}{4} \operatorname{Cov}[(2X-\theta_1-\varphi_1)'\Sigma^{-1}(\varphi_1-\theta_1), (2X-\theta_2-\varphi_2)'\Sigma^{-1}(\varphi_2-\theta_2) \mid \psi] \\ &= \frac{1}{4} \operatorname{Cov}[2X\Sigma^{-1}(\varphi_1-\theta_1), 2X\Sigma^{-1}(\varphi_2-\theta_2) \mid \psi] \\ &= (\varphi_1-\theta_1)'\Sigma^{-1}(\varphi_2-\theta_2) . \end{aligned}$$

Let

$$(6)\quad d(\theta, \varphi) = \frac{1}{2}(\varphi-\theta)'\Sigma^{-1}(\varphi-\theta) .$$

After some algebraic manipulation, it can be shown that

$$(7)\quad d(\theta_1, \varphi_2) + d(\theta_2, \varphi_1) - d(\theta_1, \theta_2) - d(\varphi_1, \varphi_2) = (\varphi_1-\theta_1)'\Sigma^{-1}(\varphi_2-\theta_2) .$$

The following result is obtained from (4), (5) and (7). For a multivariate normal distribution with unknown mean ψ and known dispersion matrix Σ, the log-likelihood-ratio process $\{L(X, \theta, \varphi)\}$ is Gaussian with

$$(8)\quad E[L(X, \theta, \varphi) \mid \psi] = d(\varphi, \psi) - d(\theta, \psi) ,$$

$$(9)\quad \begin{aligned} &\operatorname{Cov}[L(X, \theta_1, \varphi_1), L(X_2, \theta_2, \varphi_2) \mid \psi] \\ &= d(\theta_1, \varphi_2) + d(\theta_2, \varphi_1) - d(\theta_1, \theta_2) - d(\varphi_1, \varphi_2) \end{aligned}$$

where

$$d(\theta, \varphi) = \frac{1}{2}(\varphi-\theta)'\Sigma^{-1}(\varphi-\theta) .$$

We note that $d^{1/2}$ is an Euclidean metric on R^n when X is n-variate normally distributed.

Let us now consider a collection of Gaussian processes $\{X_t\}$ with $E[X_t \mid \psi] = \psi_t$ and $\operatorname{Cov}[X_t, X_u \mid \psi] = \sigma_{tu}$ where t and $u \in T = [0, 1]$. Let B be the class of all finite subsets F of T. B is a directed set by the inclusion relation ($F_1 \succ F_2$ if $F_1 \supseteqq F_2$). Let $L(F, \theta, \varphi)$ denote the log-likelihood-ratio of the densities of $(X_{t_1}, X_{t_2}, \cdots, X_{t_k})$ when θ and φ are the mean functions and $F = \{t_1, t_2, \cdots, t_k\}$. It is well known that $L(F, \theta, \varphi)$ is Gaussian and it follows from (8) and (9) that

$$(10)\quad E[L(F, \theta, \varphi) \mid \psi] = d_F(\varphi, \psi) - d_F(\theta, \psi)$$

and

$$(11)\quad \begin{aligned} &\operatorname{Cov}[L(F, \theta_1, \varphi_1), L(F, \theta_2, \varphi_2) \mid \psi] \\ &= d_F(\theta_1, \varphi_2) + d_F(\theta_2, \varphi_1) - d_F(\theta_1, \theta_2) - d_F(\varphi_1, \varphi_2) \end{aligned}$$

where $d_F^{1/2}$ is an Euclidean metric on R^k and k denotes the number of elements in F. Let ν_θ be the Gaussian measure induced by the process $\{X_t\}$ when θ is the true parameter through the usual Kolmogorov extension. Since the processes $\{X_t\}$ have the same covariance structure, ν_θ and ν_φ are mutually absolutely continuous for any θ and φ. Furthermore if $L(\theta, \varphi)=\log d\nu_\theta/d\nu_\varphi$, then it follows from theorems on Martingale convergence (See Doob [1]) or from the results of Feldman [2], that $L(F, \theta, \varphi)$ converges in quadratic mean to $L(\theta, \varphi)$. In fact $L(\theta, \varphi)$ is a Gaussian random variable such that

$$E[L(F, \theta, \varphi) \mid \psi] \to E[L(\theta, \varphi) \mid \psi]\,, \tag{12}$$

and

$$\mathrm{Cov}\,[L(F, \theta_1, \varphi_1), L(F, \theta_2, \varphi_2) \mid \psi] \to \mathrm{Cov}\,[L(\theta_1, \varphi_1), L(\theta_2, \varphi_2) \mid \psi]\,. \tag{13}$$

Therefore, we have from (10) and (11) that

$$E[L(\theta, \varphi) \mid \psi]=\lim_F\, [d_F(\varphi, \psi)-d_F(\theta, \psi)] \tag{14}$$

and

$$\begin{aligned}&\mathrm{Cov}\,[L(\theta_1, \varphi_1), L(\theta_2, \varphi_2) \mid \psi]\\ &\quad=\lim_F\, [d_F(\theta_1, \varphi_2)+d_F(\theta_2, \varphi_1)-d_F(\theta_1, \theta_2)-d_F(\varphi_1, \varphi_2)]\,.\end{aligned} \tag{15}$$

Substituting $\theta=\psi$ in (14), we note that $\lim_F d_F(\varphi, \psi)$ exists for any φ and ψ. Let us denote this limit by $d(\varphi, \psi)$. It is easy to see that $d^{1/2}(\varphi, \psi)$ is an Euclidean metric on R^T. From (14) and (15) we have

$$E[L(\theta, \varphi) \mid \psi]=d(\varphi, \psi)-d(\theta, \psi) \tag{16}$$

and

$$\mathrm{Cov}\,[L(\theta_1, \varphi_1), L(\theta_2, \varphi_2) \mid \psi]=d(\theta_1, \varphi_2)+d(\theta_2, \varphi_1)-d(\theta_1, \theta_2)-d(\varphi_1, \varphi_2)\,. \tag{17}$$

We shall now prove the following theorem.

THEOREM. *A necessary and sufficient condition for a Gaussian process* $\{L(\theta, \varphi), \theta \in R^T, \varphi \in R^T\}$ *to be the log-likelihood-ratio process of Gaussian processes* $\{X_t\}$ *with the same covariance structure is that there exists an Euclidean metric* d *on* R^T *such that*

(i) $E[L(\theta, \varphi) \mid \psi]=d^2(\varphi, \psi)-d^2(\theta, \psi)$

and

(ii) $\mathrm{Cov}\,[L(\theta_1, \varphi_1), L(\theta_2, \varphi_2) \mid \psi]=d^2(\theta_1, \varphi_2)+d^2(\theta_2, \varphi_1)-d^2(\theta_1, \theta_2)-d^2(\varphi_1, \varphi_2)$

PROOF. The necessity of the conditions has been proved in (16) and (17).

Sufficiency. Let μ_ϕ denote the Gaussian measures corresponding to the processes $\{L(\theta, \varphi)\}$ when ϕ is the true mean function. We shall first prove that $L(\phi, \eta)$ is sufficient for the measures μ_ϕ and μ_η when the conditions (i) and (ii) of the theorem are satisfied. Since the process $\{L(\theta, \varphi)\}$ is Gaussian, it is enough to prove that the conditional distribution of $L(\theta, \varphi)$ for any θ and φ given $L(\phi, \eta)$ is the same under both ϕ and η.

Define

$$Y(\theta, \varphi; \phi, \eta) = L(\theta, \varphi) - \frac{\mathrm{Cov}\,[L(\varphi, \theta), L(\phi, \eta)]}{\mathrm{Var}\,[L(\phi, \eta)]} L(\phi, \eta) .$$

From elementary results, it is known that $Y(\theta, \varphi; \phi, \eta)$ is Gaussian under both ϕ and η. Further,

$$\begin{aligned} (18) \quad & E[Y(\theta, \varphi; \phi, \eta) \mid \phi] \\ &= E[L(\theta, \varphi) \mid \phi] - \frac{\mathrm{Cov}\,[L(\varphi, \theta), L(\phi, \eta) \mid \phi]}{\mathrm{Var}\,[L(\phi, \eta) \mid \phi]} E[L(\phi, \eta) \mid \phi] \\ &= \frac{1}{2}[d(\varphi, \phi) - d(\theta, \phi) - d(\theta, \eta) + d(\varphi, \eta)] \\ &= E[Y(\theta, \varphi; \phi, \eta) \mid \eta] , \end{aligned}$$

and

$$\begin{aligned} (19) \quad & \mathrm{Var}\,[Y(\theta, \varphi; \phi, \eta) \mid \phi] \\ &= \mathrm{Var}\,[L(\theta, \varphi) \mid \phi] - \frac{\mathrm{Cov}^2\,[L(\theta, \varphi), L(\phi, \eta) \mid \phi]}{\mathrm{Var}\,[L(\phi, \eta) \mid \phi]} \\ &= 2d(\theta, \varphi) - \frac{\{d(\theta, \eta) + d(\varphi, \phi) - d(\theta, \phi) - d(\varphi, \eta)\}^2}{2d(\phi, \eta)} \\ &= \mathrm{Var}\,[Y(\theta, \varphi; \phi, \eta) \mid \eta] . \end{aligned}$$

(18) and (19), together with the remark made earlier, prove that $Y(\theta, \varphi; \phi, \eta)$ has the same Gaussian distribution. In other words, $L(\phi, \eta)$ is sufficient for the Gaussian measures μ_ϕ and μ_η. Hence the process $\{L(\theta, \varphi)\}$ is the log-likelihood-ratio process of the class of Gaussian measures $\{\mu_\phi\}$ which are generated by Gaussian processes $\{X_t\}$ with the same covariance structure.

UNIVERSITY OF CALIFORNIA, BERKELEY
MICHIGAN STATE UNIVERSITY

REFERENCES

[1] J. L. Doob, *Stochastic Processes*, John Wiley, New York, 1953.
[2] J. Feldman, Unpublished lecture notes, University of California, Berkeley, 1966.