The Singularity of Gaussian Measures in Function Space
Author(s): L. A. Shepp
Source: *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 52, No. 2 (Aug. 15, 1964), pp. 430-433
Published by: National Academy of Sciences
Stable URL: http://www.jstor.org/stable/72458
Accessed: 19/11/2010 14:40

nation thus permitted is thought to be especially conducive to the process of directional selection and fixation of novel homozygous states (homoselection). Such states might serve as the basis for the "genetic revolutions" underlying species formation.[13]

Summary.—New data are presented which indicate that two Nearctic species, *Drosophila robusta* and *D. americana*, are structurally homozygous on their extreme southern margins in central Florida. The same collecting sites are the apparent northern margin of the Neotropical species, *D. acutilabella*, which is similarly structurally homozygous. Like other species tending toward reduced polymorphism in Florida (*D. euronotus*, *D. nigromelanica*, and *D. willistoni*), the more central areas of these species show high polymorphism. Structurally monomorphic populations, especially if small, marginal and inbred, are considered ideal sites for species formation.

We are grateful to Mrs. M. S. Carson and Mr. S. Prakash for aid in collecting specimens, and to Dr. H. D. Stalker for supplying wild-caught flies from his collections. Miss Mary Christine Steiner and Mrs. Jean S. Russell rendered valuable technical assistance.

* Research supported by grants G20107 and GB-1607 of the National Science Foundation.

[1] Carson, H. L., in *Genetics and Twentieth Century Darwinism*, Cold Spring Harbor Symposia on Quantitative Biology, vol. 24 (1959), p. 87.

[2] Carson, H. L., *Asilomar Symp.*, in press.

[3] Carson, H. L., *Advan. Genet.*, **9**, 1 (1958).

[4] Carson, H. L., in *Population Genetics: The Nature and Causes of Genetic Variability in Populations*, Cold Spring Harbor Symposia on Quantitative Biology, vol. 20 (1955), p. 276.

[5] Carson, H. L., and H. D. Stalker, *Evolution*, **1**, 113–133 (1947).

[6] Hsu, T. C., *Texas Univ. Publ.*, **5204**, 35 (1952).

[7] Carson, H. L., and W. C. Blight, *Genetics*, **37**, 572 (1952).

[8] Stalker, H. D., *Genetics*, **49**, 669 (1964).

[9] *Ibid.*, 883 (1964).

[10] Townsend, J. I., Jr., *Evolution*, **6**, 428 (1952).

[11] Heed, W. B., and J. S. Russell, *Proc. Intern. Congr. Genet. 11th*, **1**, 139 (1963).

[12] Heed, W. B., and N. B. Krishnamurthy, *Texas Univ. Publ.* **5914**, 155 (1959).

[13] Mayr, E., *Evolution as a Process* (London: Geo. Allen & Unwin, 1954), p. 157.

THE SINGULARITY OF GAUSSIAN MEASURES IN FUNCTION SPACE

BY L. A. SHEPP

BELL TELEPHONE LABORATORIES, INC., MURRAY HILL, NEW JERSEY

Communicated by John W. Tukey, June 29, 1964

A pair of Gaussian measures on the space of functions on an interval are either mutually singular or equivalent.[7] This result has led to numerous attempts (see ref. 7 for a rather complete listing of the lengthy literature) to obtain convenient criteria for deciding between the two possibilities. As has been remarked in reference 7, completely satisfactory necessary and sufficient conditions have not yet been given.

In this note we give a tractable necessary and sufficient condition for mutual singularity which is concerned only with the given covariance functions. The method of proof depends on a generalization, first obtained by Kraft,[4] of a theorem of Kakutani.[3] We give a somewhat more general version of Kraft's theorem together with a simpler proof.

The condition of mutual *singularity* for two measures is the more important one from a statistical-philosophical viewpoint. In this case, a single sample path distinguishes the measures with almost certainty. The singular case does not seem to arise in actual practice, however (see ref. 6).

This work, originally done independently, has considerable overlap with a recent paper of C. R. Rao and V. S. Varadarajan.[5] Both papers use similar methods and rely on Kraft's theorem in proving their main theorems. The present treatment applies to continuous parameter stochastic processes (rather than to the discrete parameter case, as in ref. 5) and uses martingale methods throughout.

1. *Statement of Results.*—We shall be concerned with the function space Ω of all functions $X(t)$, $0 \leq t \leq T$, with the σ-field $\mathcal{B}$ generated by the coordinate functions (see ref. 1, p. 67). Given two real-valued, positive-definite functions ρ_0 and ρ_1 on $[0,T] \times [0,T]$, we form the Gaussian measures μ_0 and μ_1 induced on $\mathcal{B}$ by ρ_0 and ρ_1, respectively (ref. 1, p. 72).

If π is a partition, $0 \leq t_1 < \ldots < t_n \leq T$, of $[0,T]$, we form the matrices ρ_0^π, ρ_1^π by restricting ρ_0 and ρ_1 to $\pi \times \pi$, i.e., $\rho_i^\pi(j,k) = \rho_i(t_j;t_k)$ for $1 \leq j,\, k \leq n$, $i = 0, 1$. Fix $0 < \alpha < 1$ throughout the paper and form the matrix sum $\rho_\alpha^\pi = (1 - \alpha)\rho_0^\pi + \alpha\rho_1^\pi$. The determinant of a matrix ρ is denoted $|\rho|$. If there is a set $\Lambda \in \mathcal{B}$ for which $\mu_0(\Lambda) = \mu_1(\Omega - \Lambda) = 0$, we write $\mu_0 \perp \mu_1$. We let

$$r_\alpha^\pi = |\rho_0^\pi|^{1-\alpha}|\rho_1^\pi|^\alpha/|\rho_\alpha^\pi|. \tag{1}$$

THEOREM 1. (i) *The numbers* r_α^π *are between zero and one and decrease as the partition* π *refines, i.e.,* $r_\alpha^{\pi_1} \leq r_\alpha^{\pi_2}$ *if* $\pi_1 \supseteq \pi_2$.

(ii) *Let* $r_\alpha = \inf\limits_\pi r_\alpha^\pi$; $r_\alpha = 0$ *if and only if* $\mu_0 \perp \mu_1$.

(iii) *If* ρ_0 and ρ_1 *are continuous, then* $\mu_0 \perp \mu_1$ *if and only if* $\lim r_\alpha^{\pi_n} = 0$, *where* $\pi_n = \{\tau_1,\ldots,\tau_n\}$ *and* $\tau_1, \tau_2, \ldots$ *is any sequence of distinct points which is dense in* $[0,T]$.

In the more general case with mean value functions u_0 and u_1 on $[0,T]$ also given, the solution is somewhat more difficult to handle since the criterion to be obtained involves the inverse of a matrix. The Gaussian measures are introduced on $\mathcal{B}$ in the same way as in the special case $u_0 = u_1 = 0$, above. A solution to this problem in the special case $\rho_0 = \rho_1$ was obtained earlier. We denote by u^π the vector whose jth component is $u_1(t_j) - u_0(t_j)$, $j = 1,2,\ldots,n$, where π is the above partition.

THEOREM 1′. (i) *The numbers* $s_\alpha^\pi = r_\alpha^\pi \exp(-\alpha(1 - \alpha)(\rho_\alpha^\pi)^{-1}u^\pi,u^\pi)$ *are between zero and one and decrease as the partition* π *refines.*

(ii) *Let* $s_\alpha = \inf\limits_\pi s_\alpha^\pi$; $s_\alpha = 0$ *if and only if* $\mu_0 \perp \mu_1$.

(iii) *If* ρ_0 *and* ρ_1 *are continuous, then* $\mu_0 \perp \mu_1$ *if and only if* $\lim s_\alpha^{\pi_n} = 0$, *where* $\pi_n \subset \pi_{n+1}$, $n = 1,2,\ldots$, *is a sequence of expanding partitions which become dense in* $[0,T]$.

An immediate corollary of these statements is that $r_\alpha = 0$ implies $s_\alpha = 0$ since the exponential term is bounded by unity. This has a simple intuitive interpretation.

2. *Proofs of the Assertions.*—Let Ω be a set and $\mathcal{B}_\pi$ be a σ-field of subsets of Ω,

$\pi \epsilon P$, where P is a directed set. Suppose that $\mathcal{B}_{\pi_1} \subseteq \mathcal{B}_{\pi_2}$ whenever $\pi_1 \leq \pi_2$. Let $\mathcal{B}$ be the smallest σ-field containing each of $\mathcal{B}_\pi$, $\pi \epsilon P$. Suppose further that μ_0 and μ_1 are probability measures on $\mathcal{B}$ and define the restrictions of these measures to $\mathcal{B}_\pi$ by

$$\mu_0^\pi = \mu_0|_{\mathcal{B}_\pi}, \qquad \mu_1^\pi = \mu_1|_{\mathcal{B}_\pi}, \ \pi \epsilon P. \tag{2.1}$$

Letting $\mu = \mu_0 + \mu_1$, we have $\mu_i < \mu$, $\mu_i^\pi < \mu$, $i = 0, 1$. We denote the Radon-Nikodym derivatives $d\mu_0/d\mu$ and $d\mu_1/d\mu$ by X_0 and X_1, respectively, and we let $X_i^\pi = E^{\mathcal{B}_\pi}X_i$, $i = 0,1$, denote the conditional expectation of X_i with respect to $\mathcal{B}_\pi$. We have $X_i^\pi = d\mu_i^\pi/d\mu, i = 0,1$, almost everywhere with respect to μ. We set, for $0 < \alpha < 1$,

$$r_\alpha^\pi = \left\{ \int_\Omega (X_0^\pi)^\alpha (X_1^\pi)^{1-\alpha} d\mu \right\}^2, \tag{2.2}$$

and note that μ could have been taken here as any measure dominating μ_0 and μ_1.

THEOREM (*Kraft*). (i) *The numbers r_α^π are between zero and one and decrease as π increases, i.e., $r_\alpha^{\pi_1} \leq r_\alpha^{\pi_2}$ if $\pi_1 \geq \pi_2$.*

(ii) *Let $r_\alpha = \inf_\pi r_\alpha^\pi$; $r_\alpha = 0$ if and only if $\mu_0 \perp \mu_1$.*

To prove the theorem we note that $\{X_i^\pi, \pi \epsilon P\}$ are martingale nets $i = 0,1$, and by a theorem of Helms[2] we obtain easily that X_i^π converges to X_i, $i = 0,1$, in $L^1(\mu)$. It follows from the fact that $d\mu_i^\pi/d\mu \leq 1$ almost everywhere (μ) $i = 0,1$, that

$$\sqrt{r_\alpha} = \int_\Omega (d\mu_0/d\mu)^\alpha (d\mu_1/d\mu)^{1-\alpha} d\mu \tag{2.3}$$

which is zero if and only if $(d\mu_0/d\mu)\cdot(d\mu_1/d\mu) = 0$ almost everywhere (μ). Since this occurs if and only if $\mu_0 \perp \mu_1$, we have proved (*ii*).

To prove (*i*) we note that $g(x_0,x_1) = x_0^\alpha x_1^{1-\alpha}$, $x_0 \geq 0$, $x_1 \geq 0$ is a concave function. By Jensen's inequality for conditional expectations we have, for $\pi_1 \geq \pi_2$,

$$E^{\mathcal{B}_1} g(X_0^{\pi_2}, X_1^{\pi_2}) \geq g(E^{\mathcal{B}_1}X_0^{\pi_2}, E^{\mathcal{B}_1}X_1^{\pi_2}) = g(X_0^{\pi_1}, X_1^{\pi_1}). \tag{2.4}$$

Taking expectations of both sides and squaring, we obtain

$$r_\alpha^{\pi_2} \geq r_\alpha^{\pi_1}. \tag{2.5}$$

It is an immediate consequence of Hölder's inequality that $r_\alpha^\pi \leq 1$. This completes the proof.

3. *Evaluation of r_α^π for Gaussian Measures.*—We now specialize P as the directed set of partitions of $[0,T]$, and suppose that $\mathcal{B}_\pi$ is the σ-field induced by the coordinate functions $X(t)$ for $t \epsilon \pi$. It is easy to see that we may write

$$\int_\Omega (X_0^\pi)^\alpha (X_1^\pi)^{1-\alpha} d\mu = \int_\Omega \left(\frac{d\mu_0^\pi}{d\mu_1^\pi}\right)^\alpha d\mu_1^\pi \tag{3.1}$$

if $\mu_0^\pi < \mu_1^\pi$ which is the case if ρ_1^π is positive-definite. Transferring this integration to Euclidean space, we obtain

$$\int_{-\infty}^{\infty} \cdots \int (p_0^\pi(x_1,\ldots,x_n))^\alpha (p_1^\pi(x_1,\ldots x_n))^{1-\alpha} dx_1 \ldots dx_n \tag{3.2}$$

$$p_i^\pi(x) = (2\pi)^{-n/2}|\rho_i^\pi|^{-1/2} \exp - \frac{1}{2}((\rho_i^\pi)^{-1}x,x) \quad i = 0,1,$$

where $x = (x_1,\ldots,x_n)$ and π has n elements. Standard integration formulas and some slight manipulations give

$$r_\alpha{}^\pi = |\rho_0{}^\pi|^{1-\alpha}|\rho_1{}^\pi|^\alpha/|\rho_\alpha{}^\pi|. \tag{3.3}$$

This proves (i) and (ii) of Theorem 1. To obtain the corresponding results for Theorem 1′, $p_i{}^\pi(x)$ becomes

$$(2\pi)^{-n/2}|\rho_i{}^\pi|^{-1/2} \exp - \frac{1}{2}((\rho_i{}^\pi)^{-1}(x - u^\pi),(x - u^\pi)). \tag{3.4}$$

We employ a straightforward calculation, completing the square, and the easily established identity

$$\alpha\rho_0{}^{-1} - \alpha\rho_0{}^{-1}(\alpha\rho_0{}^{-1} + (1 - \alpha)\rho_1{}^{-1})^{-1}\alpha\rho_0{}^{-1} = \alpha(1 - \alpha)((1 - \alpha)\rho_0 + \alpha\rho_1)^{-1}. \tag{3.5}$$

This reduces the square of the integral in (3.2) to the required form, i.e.,

$$s_\alpha{}^\pi = r_\alpha{}^\pi \exp - \alpha(1 - \alpha)((\alpha\rho_0{}^\pi + (1 - \alpha)\rho_1{}^\pi)^{-1}u^\pi,\ u^\pi). \tag{3.6}$$

We turn now to the proof of the last part of the theorems. Suppose that $t_1,t_2,\ldots$ is an infinite sequence of distinct points dense in $[0,T]$ and that ρ_0 and ρ_1 are continuous on $[0,T] \times [0,T]$. In applications one might take the t sequence as the binary points of the interval. We shall show that $\mathcal{B}$ may be replaced by the sub-σ-field $\mathcal{B}_0$ generated by the coordinate functions $X(t_n)$, $n = 1,2,\ldots$. If t is any point in $[0,T]$, there is a subsequence n' for which $t_{n'} \to t$. By the continuity of ρ_0 and ρ_1, $X(t_{n'}) \to X(t)$ in the mean topology relative to $\mu = \mu_0 + \mu_1$. Passing to a further subsequence n'', we obtain $X(t_{n''}) \to X(t)$ almost everywhere (μ). It follows that $X(t)$ is almost everywhere, with respect to both μ_0 and μ_1, equal to a function measurable $\mathcal{B}_0$. Further, to every set $B \epsilon \mathcal{B}$ there is a set $B_0 \epsilon \mathcal{B}_0$ for which $(\mu_0 + \mu_1)$ $(B_0 \triangle B) = 0$. This means that $\mu_0 \perp \mu_1$ relative to $\mathcal{B}_0$ is the same as $\mu_0 \perp \mu_1$ relative to $\mathcal{B}$. The proof is now complete since the sequence $r_\alpha{}^{\pi_n} \geq r_\alpha{}^{\pi_{n+1}}$, $n = 1,2,\ldots$ decreases to zero if and only if $\mu_0 \perp \mu_1$ on $\mathcal{B}_0$. This is merely the special case of Kraft's theorem where P is a linearly ordered set in which case one could apply the usual martingale theorem.

The author is indebted to David Slepian for help and encouragement. He has also benefited from discussions with S. P. Lloyd.

[1] Doob, J. L., *Stochastic Processes* (New York: Dover, 1953).

[2] Helms, L. L., "Mean convergence of martingales," *Trans. Am. Math. Soc.*, **87**, 439–446 (1958).

[3] Kakutani, S., "On equivalence of infinite product measures," *Ann. Math.*, **49**, 214–224 (1948).

[4] Kraft, C., "Some conditions for consistency and uniform consistency of statistical procedures," *Univ. California Publ. Statist.*, **2**, No. 6, 125–142 (1955).

[5] Rao, C. R., and V. S. Varadarajan, "Discrimination of Gaussian processes," *Sankhyā, Ser. A*, **25**, 303–330 (1963).

[6] Slepian, D., "Some comments on the detection of Gaussian signals in Gaussian noise," *IRE, Trans. Inform. Theory*, **4**, No. 2, 65–68 (1958).

[7] Yaglom, A. M., "On the equivalence and perpendicularity of two Gaussian probability measures in function space," in *Proceedings of the Symposium on Time Series Analysis, Brown University, 1962*, ed. Murray Rosenblatt (New York: John Wiley & Sons, 1963).