The Annals of Probability
1981, Vol. 9, No. 3, 453–459

AN EXTENDED DICHOTOMY THEOREM FOR SEQUENCES OF PAIRS OF GAUSSIAN MEASURES

By G. K. Eagleson

University of Cambridge

A dichotomy for sequences of pairs of Gaussian measures is proved. This result is then used to give a simple proof of the famous equivalence/singularity dichotomy for Gaussian processes. The proof uses tightness arguments and can be directly applied to the theory of hypothesis testing to show that two sequences of simple hypotheses which specify Gaussian measures are either contiguous or entirely separable.

1. Introduction. In this note we prove a dichotomy for sequences of pairs of Gaussian measures. Roughly speaking, we show that the sequence of Radon-Nikodym derivatives of the pairs is either tight or there exists a subsequence along which all the mass is dissipated. The proof of this result is extremely simple and, apart from its own interest, it can be used to obtain the equivalance/singularity dichotomy for Gaussian processes and to extend that result to sequences of probability measures which are not necessarily defined on the same sample space. Thus we are able to apply our basic dichotomy to the theory of hypothesis testing to show that two sequences of simple hypotheses which specify Gaussian measures are either contiguous or entirely separable.

2. Background information. Suppose that $\{X_n, n = 1, 2, \cdots\}$ is a sequence of Gaussian random variables with $X_n \sim N(\mu_n, \sigma_n^2)$, say. It is obvious that the distributions of the $\{X_n\}$ are tight if and only if both the $\{\mu_n\}$ and $\{\sigma_n^2\}$ lie in compact sets. If this is not the case, there will exist a subsequence $\{n'\}$ such that for all $A > 0$

$$\lim_{n'\to\infty} P(|X_{n'}| \leq A) = 0. \tag{1}$$

That is, either the distributions are tight or there exists a subsequence along which all the mass is dissipated.

The proof of our theorem is based on the simple observation that a similar dichotomy also holds for the distributions of quadratic forms in Gaussian random variables.

Suppose that $\{X_j^{(n)}; j = 1, \cdots, n; n = 1, 2, \cdots\}$ is a triangular array of Gaussian random variables and that $A^{(n)} = (a_{ij}^{(n)})$ is, for each n, an $n \times n$ symmetric matrix. We consider the sequence of quadratic forms

$$Z_n = \sum_{i,j} X_i^{(n)} a_{ij}^{(n)} X_j^{(n)} \tag{2}$$

and ask when are their distributions tight. By simultaneously reducing the variance-covariance matrix of $\{X_j^{(n)}, j = 1, \cdots, n\}$ to the unit matrix and diagonalizing $A^{(n)}$, it is clear that the distribution of Z_n is the same as that of

$$Z_n' = \sum_{r=1}^{n} \lambda_r^{(n)} (Y_r + m_r^{(n)})^2, \tag{3}$$

where the $\{Y_r\}$ are independent, identically distributed $N(0, 1)$ random variables, the $\{m_r^{(n)}\}$ are constants and $\{\lambda_r^{(n)}\}$ are the eigenvalues of $A^{(n)}$.

Proposition 1. *A sequence of distributions of quadratic forms, Z_n, in Gaussian*

Received November 3, 1979; revised March 10, 1980.
AMS 1970 *subject classifications.* Primary, 60G30; secondary, 62F03.
Key words and phrases. Absolute continuity, Gaussian processes, contiguity.

random variables is either tight or there exists a subsequence $\{n'\}$ such that for all $A > 0$

$$\lim_{n'\to\infty} P(|Z_{n'}| \le A) = 0. \tag{4}$$

PROOF. Considering the $\{Z_n'\}$ introduced above, we have

$$EZ_n' = \mu_n = \sum_{r=1}^n \lambda_r^{(n)} (1 + m_r^{(n)2})$$

and

$$\operatorname{Var} Z_n' = \sigma_n^2 = 2 \sum_{r=1}^n \lambda_r^{(n)2}(1 + 2m_r^{(n)2}).$$

If $\sup_n \sigma_n^2 < \infty$, then Chebyshev's inequality shows that the Z_n' are tight if $\sup_n |\mu_n| < \infty$ and that the total mass is dissipated along a subsequence if $\sup_n |\mu_n| = \infty$.

On the other hand, if $f_n(t)$ denotes the characteristic function of Z_n', then

$$f_n(t) = [\textstyle\prod_{r=1}^n (1 - 2i\lambda_r^{(n)}t)^{-1/2}]\exp[\sum_{r=1}^n \tfrac{1}{2}m_r^{(n)2}((1 - 2i\lambda_r^{(n)}t)^{-1} - 1)]$$

and

$$|f_n(t)|^2 \le [4t^2 \textstyle\sum_{r=1}^n \lambda_r^{(n)2}]^{-1/2} \exp[-4t^2 \sum_{r=1}^n m_r^{(n)2}\lambda_r^{(n)2}/(1 + 4\lambda_r^{(n)2}\, t^2)].$$

When $\sup_n \sigma_n^2 = \infty$, either there exists a subsequence $\{n'\}$ along which $\sum_{r=1}^{n'} \lambda_r^{(n')2} \to \infty$, in which case $|f_{n'}(t)| \to 0$, or the $\{\lambda_r^{(n)2}\}$ are uniformly bounded in n and r. If

$$\lambda_r^{(n)2} \le C,\ (1 + 4\lambda_r^{(n)2}t^2) \le (1 + 4Ct^2)$$

and

$$\textstyle\sum_{r=1}^n m_r^{(n)2}\lambda_r^{(n)2}/(1 + 4\lambda_r^{(n)2}t^2) \ge (1 + 4Ct^2)^{-1} \sum_{r=1}^n m_r^{(n)2}\lambda_r^{(n)2}.$$

As, in this second case, there exists a subsequence $\{n'\}$ along which $\sum_{r=1}^{n'} m_r^{(n')2}\lambda_r^{(n')2} \to \infty$, we have, once again that $|f_{n'}(t)| \to 0$. But $|f_{n'}(t)| \to 0$ implies (4) (see, for example, Loève 12.4.4° page 197).

3. A dichotomy for sequences of Gaussian measures. If two Gaussian distributions are equivalent, their likelihood ratio is the exponential of a quadratic or a linear form. This simple observation, together with Proposition 1, gives the following Theorem.

THEOREM. *Suppose that for each $n = 1, 2, \cdots, (\mathbb{R}^n, \mathscr{B}^n)$ is endowed with two Gaussian probability measures, P_n and Q_n. We assume that $Q_n \ll P_n$ for all n and write*

$$L_n = dQ_n/dP_n.$$

Either $\{L_n\}$ is tight (Q_n) or there exists a subsequence $\{n'\}$ and constants $A_n \uparrow \infty$ such that

$$\lim_{n'\to\infty} Q_{n'}(L_{n'} \le A_{n'}) = 0. \tag{5}$$

PROOF. For each n, $\log L_n$ is (after maybe completing a square) either a quadratic form plus a constant or a linear form in random variables which are Gaussian under Q_n (as well as under P_n). If the $\{\log L_n\}$ are all essentially quadratic forms then, either they (and hence $\{L_n\}$) are tight (Q_n), or there exists a subsequence $\{n'\}$ such that for all $A > 0$

$$\lim_{n'\to\infty} Q_{n'}(|\log L_{n'}| \le A) = 0.$$

Hence there exist constants A_n increasing to ∞ and such that

$$\lim_{n'\to\infty} Q_{n'}(|\log L_{n'}| \le A_{n'}) = 0$$

(see Chung, Section 7.2, Lemma 1). Together with the estimate,

$$Q_n(L_n < A^{-1}) = \int_{[L_n<A^{-1}]} L_n \, dP_n < A^{-1},$$

this proves that (5) holds.

If the $\{\log L_n\}$ are all linear forms it is even more straightforward to see that they are either tight or (5) holds. It is also clear that the dichotomy of the theorem will continue to hold when the $\{\log L_n\}$ are sometimes linear and sometimes quadratic forms.

4. Applications. There exist in the literature a number of different proofs of the equivalence/singularity dichotomy for Gaussian processes; Hájek (1958) used information theoretic arguments, Feldman (1958) Hilbert space techniques, Brody (1971) properties of Gaussian distributions, Le Page and Mandrekar (1972) a 0-1 law for Gaussian processes, Kabanov, Liptser and Shiraev (1977) the fact that a series of squared Gaussian random variables converges with probability zero or one, and Chatterji and Mandrekar (1978) reproducing kernel Hilbert spaces. Here we show that this classical result is a straightforward consequence of our Theorem.

1. *Gaussian sequences.*

COROLLARY 1. *Suppose that $\{X_n, n = 1, 2, \cdots\}$ is a canonical sequence of random variables, defined on $(\mathbb{R}^\infty, \mathscr{B}^\infty)$, which is Gaussian under both the measures P and Q. Then either $P \sim Q$ or $P \perp Q$.*

PROOF. Let $\mathscr{F}_n$ denote the σ-field generated by $(X_1, \cdots, X_n)$ and set $P_n = P_{|\mathscr{F}_n}$, $Q_n = Q_{|\mathscr{F}_n}$. As P_n and Q_n are finite dimensional Gaussian distributions, either $P_n \sim Q_n$ or $P_n \perp Q_n$. As $P_n \perp Q_n$ for one n implies that $P \perp Q$, we may assume that $P_n \sim Q_n$ for all n. Then $\{L_n\}$ is an $(\mathscr{F}_n, P)$-martingale and it is well-known that $Q \ll P$ if and only if $\{L_n\}$ is uniformly integrable (P). But

$$\int_{[L_n>A]} L_n \, dP = Q(L_n > A)$$

so that $\{L_n\}$ is uniformly integrable (P) if and only if $\{L_n\}$ is tight (Q). Thus the Theorem gives immediately that either $Q \ll P$ or (5) holds. In the later case, as

$$P_{n'}(L_{n'} > A_{n'}) \le 1/A_{n'} \quad \text{for all} \quad n',$$

the measures P and Q must be mutually singular. By symmetry, either $P \ll Q$ or $P \perp Q$ and the dichotomy is proved.

2. *Gaussian processes.*

COROLLARY 2. *Suppose that $\{X_t, t \in [0, T]\}$ is a stochastic process on $(\mathbb{R}^T, \mathscr{B}^T)$ which is Gaussian under both P and Q. Then either $P \sim Q$ or $P \perp Q$.*

PROOF. The proof is exactly the same as that of Corollary 1, with $\mathscr{F}_n$ here denoting the σ-fields generated by nested finite subsets of $\{X_t, t \in [0, T]\}$.

3. *Testing Gaussian hypotheses.* Suppose that we are interested in testing two simple hypotheses. As we shall be interested in asymptotic size and power, we consider a sequence of probability spaces which we may, without loss of generality, take to be $(\mathbb{R}^n, \mathscr{B}^n, \mu_n)$. (Typically, n is the sample size.) For each n, we suppose that the null and alternative hypotheses are given by

$$H_0^{(n)} : \mu_n = P_n,$$

$$H_1^{(n)} : \mu_n = Q_n, \quad \text{say.}$$

Here the analogue of absolute continuity is Le Cam's idea of contiguity: the sequence of measures $\{Q_n\}$ is said to be *contiguous* to $\{P_n\}$ if whenever $P_n(A_n) \to_{n\to\infty} 0$ as $n \to \infty$ for some sequence of events $A_n \in \mathscr{B}^n$ then $Q_n(A_n) \to_{n\to\infty} 0$ also. If $\{Q_n\}$ is contiguous to $\{P_n\}$ then, at least in "nice" cases, the asymptotic distribution of statistics under Q_n can be derived from their asymptotic distribution under P_n, allowing the asymptotic power of tests based on these statistics to be calculated. (For a recent, elegant paper on contiguity, see Hall and Loynes (1977).)

The analogue of singularity is the entire separation of the two sequences of measures $\{P_n\}$ and $\{Q_n\}$ (see Le Cam (1977)). For two measures P and Q, let $\| P \wedge Q \|$ denote the infimum of the sum of error probabilities over all tests of P against Q. The sequences $\{P_n\}$ and $\{Q_n\}$ are said to *separate entirely* if

$$\liminf_{n\to\infty} \| P_n \wedge Q_n \| = 0.$$

That is, $\{P_n\}$ and $\{Q_n\}$ separate entirely if, and only if, there is a sequence of test functions $\{\phi_n\}$ and a subsequence, along which the ϕ_n give a consistent sequence of tests of $H_0^{(n)}$ against $H_1^{(n)}$ of asymptotic size zero.

It may seem strange to require only the existence of a consistent subsequence of tests, but this is forced upon us by the fact that the measures may be defined on different probability spaces. In practice, the subsequence will coincide with the whole sequence. That entire separation is the natural analogue of singularity is shown by the following Proposition.

PROPOSITION 2. *The following statements are equivalent*:
(i) *the sequences of measures $\{P_n\}$ and $\{Q_n\}$ separate entirely*;
(ii) *there exists a subsequence n' and sets $A_{n'} \in \mathscr{B}^{n'}$ such that*

$$P_{n'}(A_{n'}) \to 0 \qquad \text{while} \qquad Q_{n'}(A_{n'}) \to 1.$$

PROOF. Suppose that $\{P_n\}$ and $\{Q_n\}$ are separated entirely by the test functions $\{\phi_n\}$ along the subsequence $\{n'\}$. Set

$$A_n = [\phi_n \geq \tfrac{1}{2}].$$

Then

$$P_{n'}(A_{n'}) \leq 2E_{P_n'}(\phi_{n'}(\mathbf{X})) \to 0$$

and

$$Q_{n'}(A_{n'}^c) \leq 2E_{Q_n'}(1 - \phi_{n'}(\mathbf{X})) \to 0.$$

Conversely, if (ii) holds, then $\phi_{n'} = I_{A_n'}$ will provide a consistent subsequence of tests.

It follows that when the P_n and Q_n are the finite-dimensional distributions of two measures P and Q, contiguity is simply absolute continuity and entire separability, mutual singularity. In this case, Kraft (1955) showed that singularity is equivalent to the existence of a consistent sequence of tests of $H_0^{(n)}$ against $H_1^{(n)}$.

We can now apply our theorem to see that if the hypotheses specify Gaussian measures then only the two extreme situations of continguity and entire separability are possible.

COROLLARY 3. *If the sequence of null and alternative hypotheses $H_0^{(n)}$ and $H_1^{(n)}$ specify Gaussian measures P_n and Q_n respectively, then $\{P_n\}$ and $\{Q_n\}$ are either mutually contiguous or entirely separable.*

PROOF. As the P_n and Q_n are Gaussian measures, for each n either $P_n \sim Q_n$ or $P_n \perp Q_n$. If there exists infinitely many n for which $P_n \perp Q_n$, then $\{P_n\}$ and $\{Q_n\}$ are clearly entirely separable. On the other hand, if there are only finitely many such n, for n larger than some N_0 we may assume that $P_n \sim Q_n$ and that the likelihood ratio $L_n = dQ_n/dP_n$ exists. It then follows (Hall and Loynes (1977), Proposition 4) that $\{Q_n\}$ is contiguous to $\{P_n\}$ if and only if $\{L_n, n \geq N_0\}$ is tight (Q_n). The result follows immediately from the theorem above on noting that

$$P_n(L_n > A) \leq A^{-1} \qquad \text{for all} \quad n.$$

EXAMPLE. Suppose that $H_0^{(n)}$ specifies that the random variables $(X_{n1}, \cdots, X_{nn})$ are independent $N(0, 1)$ and $H_1^{(n)}$ that they are independent with $X_{ni} \sim N(a_{ni}, 1)$, $i = 1, \cdots, n$. Here $\log L_n$ is the linear form

$$\textstyle\sum_{i=1}^n a_{ni}x_{ni} - \frac{1}{2}\sum_{i=1}^n a_{ni}^2$$

which is distributed under $H_1^{(n)}$ as $N(\frac{1}{2}\sum a_{ni}^2, \sum a_{ni}^2)$. It follows immediately that $\{Q_n\}$ is contiguous to $\{P_n\}$ if and only if

$$\limsup_{n\to\infty} \textstyle\sum_{i=1}^n a_{ni}^2 < \infty.$$

This condition should be compared with that obtained by Hájek and Šidák (VI.2.1) where they show that $\{Q_n\}$ is contiguous to $\{P_n\}$ (and that $\log L_n$ is asymptotically normal under P_n) if

$$\lim_{n\to\infty} \max_{1\leq i\leq n} a_{ni}^2 = 0$$

and

$$\lim_{n\to\infty} \textstyle\sum_{i=1}^n a_{ni}^2 = b^2, \qquad 0 < b^2 < \infty.$$

REMARK. While the existence of a dichotomy is interesting, one also needs to have straightforward necessary and sufficient conditions for one or other of the two possibilities to hold. Such conditions have been derived for Gaussian processes by Rao and Varadarajan (1963) and Shepp (1964). However, one can use the Theorem (as in the above example) to immediately give conditions for equivalence/singularity or contiguity/separability. In fact the proof of the Theorem shows that $\{L_n\}$ will be tight (Q_n) if, and only if, both

$$|E_{Q_n} \log L_n|$$

and

$$\mathrm{Var}_{Q_n} (\log L_n)$$

are bounded. By simultaneously reducing Q_n to $N(\mathbf{0}, I_n)$ and P_n to $N(\boldsymbol{\mu}_n, \mathrm{diag}(\sigma_{n1}^2, \cdots, \sigma_{nn}^2))$,

$$\begin{aligned}\log L_n &= \textstyle\sum_{j=1}^n \log \sigma_{nj} - \dfrac{1}{2}\sum_{j=1}^n x_j^2 + \dfrac{1}{2}\sum_{j=1}^n (x_j - \mu_{nj})^2/\sigma_{nj}^2 \\ &= \textstyle\sum_{j=1}^n \log \sigma_{nj} - \dfrac{1}{2}\sum_{j=1}^n (1 - \sigma_{nj}^{-2})\left[x_j + \dfrac{\mu_{nj}}{\sigma_{nj}^2 - 1}\right]^2 \\ &\quad + \dfrac{1}{2}\textstyle\sum_{j=1}^n \dfrac{\mu_{nj}^2}{\sigma_{nj}^2 - 1}.\end{aligned}$$

It follows that

$$E_{Q_n} \log L_n = -\frac{1}{2}\textstyle\sum_{j=1}^n \left(\log \dfrac{1}{\sigma_{nj}^2} + 1 - \dfrac{1}{\sigma_{nj}^2}\right) + \dfrac{1}{2}\sum_{j=1}^n \mu_{nj}^2/\sigma_{nj}^2$$

and that

$$\mathrm{Var}_{Q_n} \log L_n = \tfrac{1}{2}\textstyle\sum_{j=1}^n (1 - \sigma_{nj}^{-2})^2 + \sum_{j=1}^n \mu_{nj}^2/\sigma_{nj}^4.$$

In order that $\mathrm{Var}_{Q_n} \log L_n$ be bounded, it is necessary and sufficient that

(6) $$\sup_n \sum_{j=1}^n (1 - \sigma_{nj}^{-2})^2 < \infty$$

and that

(7) $$\sup_n \sum_{j=1}^n \mu_{nj}^2/\sigma_{nj}^4 < \infty.$$

As (6) implies that the σ_{nj}^2 are uniformly bounded away from zero, (7) may be replaced by

(7′) $$\sup_n \sum_{j=1}^n \mu_{nj}^2 < \infty.$$

On the other hand, as $\log x + 1 - x \le 0$, $E_{Q_n} \log L_n$ will be bounded if and only if

(8) $$\sup_n \sum_{j=1}^n (\log \sigma_{nj}^2 - 1 + \sigma_{nj}^{-2}) < \infty$$

and

(9) $$\sup_n \sum_{j=1}^n \mu_{nj}^2/\sigma_{nj}^2 < \infty.$$

As (8) implies that the σ_{nj}^2 are uniformly bounded, and as there exist positive constants c and d such that

$$y - cy^2 \le \log(1 + y) \le y - dy^2$$

if $-1 < a_1 < y < a_2 < \infty$ $(a_1 < 1 < a_2)$, condition (8) is equivalent to

(10) $$\sup_n \max_{j \le n} \sigma_{nj}^2 < \infty$$

together with

(11) $$\sup_n \sum_{j=1}^n (1 - \sigma_{nj}^{-2})^2 < \infty.$$

Thus we have proved:

COROLLARY 4. *Suppose that for each $n = 1, 2, \cdots$, $(\mathbb{R}^n, \mathscr{B}^n)$ is endowed with two Gaussian probability measures, P_n and Q_n, with $P_n \sim N(\boldsymbol{\mu}_n, \mathrm{diag}(\sigma_{n1}^2, \cdots, \sigma_{nn}^2))$ and $Q_n \sim N(\mathbf{0}, I_n)$. Then $\{P_n\}$ and $\{Q_n\}$ are mutually contiguous or entirely separable. They will be mutually contiguous if, and only if,*

(i) $$\sup_n \max_{j \le n} \sigma_{nj}^2 < \infty,$$

(ii) $$\sup_n \sum_{j=1}^n (1 - \sigma_{nj}^{-2})^2 < \infty,$$

and

(iii) $$\sup_n \sum_{j=1}^n \mu_{nj}^2 < \infty.$$

These conditions are exactly the same as those obtained by Chatterji and Mandrekar (1978) (Lemma 4.1(c)) when the P_n and Q_n are finite-dimensional distributions of a Gaussian process.

Acknowledgments. The author is grateful to Dave Aldous for many useful discussions and to the referee for advice and some references.

REFERENCES

BRODY, E. J. (1971). An elementary proof of the Gaussian dichotomy theorem. *Z. Wahrscheinlichkeitstheorie und verw. Gebiete* **20** 217–226.

CHATTERJI, S. D. AND MANDREKAR, V. (1978). Equivalence and singularity of Gaussian measures and applications. In *Probabilistic Analysis and Related Topics I* (A. T. Bharucha-Reid, ed.). Academic, New York.

CHUNG, K. L. (1974). *A Course in Probability Theory* (2nd ed.) Academic, New York.

FELDMAN, J. (1958). Equivalence and perpendicularity of Gaussian measures. *Pacific J. Math.* **8** 699–708. (Correction, ibid, **9** (1960) 1295–1296.)

HÁJEK, J. (1958). On a property of the normal distribution of an arbitrary stochastic process (Russian). *Czechoslovak Math. J.* **8** 610–618.

HÁJEK, J. AND ŠIDÁK, Z. (1967). *Theory of rank tests.* Academia, Prague.

HALL, W. J. AND LOYNES, R. M. (1977). On the concept of contiguity. *Ann. Probability* **5** 278–282.

KABANOV, YU. M., LIPTSER, R. S. AND SHIRAEV, A. N. (1977). On the question of the absolute continuity and singularity of probability measures (Russian). Mat. Sb. **104** (146) 227–247.

KRAFT, C. (1955). Some conditions for consistency and uniform consistency of statistical procedures. Univ. Calif. Publ. Statist. **2** 125–142.

LE CAM, L. (1977). On the asymptotic normality of estimates. Proc. Symp. to Honour Jerzey Neyman (Warsaw 1974). 203–217. *Państw. Wydawn Nauk.* Warsaw.

LE PAGE, R. D. AND MANDREKAR, V. (1972). Equivalence-singularity dichotomies from zero-one laws. *Proc. Amer. Math. Soc.* **31** 251–254.

LOÈVE, M. (1963). *Probability Theory* (3rd ed.) Van Nostrand, New York.

RAO, C. R. AND VARADARAJAN, V. S. (1963). Discrimination of Gaussian processes. *Sankhyā* **25** 303–330.

SHEPP, L. A. (1964). The singularity of Gaussian measures in function space. *Proc. Nat. Acad. Sci. U.S.A.* **52** 430–433.

CSIRO
DIVISION OF MATHEMATICS AND STATISTICS
P.O. BOX 218
LINDFIELD, NEW SOUTH WALES 2070
AUSTRALIA