AN ABSTRACT OF THE THESIS OF

LEONARD GEORGE SWANSON for the DOCTOR OF PHILOSOPHY
(Name) (Degree)

in MATHEMATICS presented on July 22, 1970
(Major) (Date)

Title: EQUIVALENCE OF FINITE MEASURES

Abstract approved: Redacted for Privacy
T. R. Chow

Necessary and sufficient conditions for the equivalence of finite measures μ and ν are given in terms of maps between function spaces and in terms of the Radon-Nikodym derivative $\frac{d\mu}{d\lambda}$, where $\lambda = \mu + \nu$. The derivative $\frac{d\mu}{d\lambda}$ is represented as the unique vector guaranteed by the Riesz representation theorem for a continuous linear form on $L_2(\lambda)$, and Fourier analysis is used to represent $\frac{d\mu}{d\nu}$ when μ is equivalent to ν. If μ and ν are equivalent and $\frac{d\mu}{d\nu}$ is known, then a complete orthonormal system is constructed for $L_2(\nu)$ in terms of a complete orthonormal system of $L_2(\mu)$. By applying Kakutani's theorem and our results on $\frac{d\mu}{d\lambda}$, necessary and sufficient conditions for the equivalence of countable Gaussian processes are given in terms of their one-dimensional distributions.

Equivalence of Finite Measures

by

Leonard George Swanson

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

June 1971

APPROVED:

Redacted for Privacy

Assistant Professor of Mathematics

in charge of major

Redacted for Privacy

Acting Chairman of Department of Mathematics

Redacted for Privacy

Dean of Graduate School

Date thesis is presented July 22, 1970

Typed by Clover Redfern for Leonard George Swanson

ACKNOWLEDGMENT

I wish to thank Professor Thomas R. Chow for his suggestion of this area of study and also thank him for his assistance in the preparation of this thesis.

TABLE OF CONTENTS

EQUIVALENCE OF FINITE MEASURES

I. INTRODUCTION

This paper is motivated by the problem of finding necessary and sufficient conditions for the equivalence of Gaussian stochastic processes. If $X = \{x_t : t \in T\}$ and $Y = \{y_t : t \in T\}$ are real-valued Gaussian processes, the Kolmogorov theorem (5, p. 32-37) allows one to consider both X and Y as processes defined on the measurable space $(R^T, \mathcal{C})$. R represents the real numbers, $\mathcal{C}$ represents the cylinder sets of R^T, and $(R^T, \mathcal{C})$ is usually referred to as the path space of the processes. If μ and ν are the probabilities generated respectively by X and Y on $(R^T, \mathcal{C})$ via the Kolmogorov theorem, then X is equivalent to Y if and only if μ is equivalent to ν. Here of course, two measures are equivalent provided they have exactly the same sets of measure zero. The processes X and Y are perpendicular or mutually singular if and only if the associated measures μ and ν are perpendicular.

Extensive research has been done on the problem of finding necessary and sufficient conditions for the equivalence of Gaussian processes, and on the problem of finding necessary and sufficient conditions for the perpendicularity of Gaussian processes. These investigations have taken many forms and have used a wide variety of methods. In the following, the major results of these investigations are given and outlines of important methods are presented.

In 1947, R. H. Cameron and W. T. Martin (3) were able to show that if one considers the measures induced on the path space by Wiener processes on the unit interval, then the measures are perpendicular if the processes have different variances. A wider sufficient condition for the perpendicularity of measures induced on the path space by continuous Gaussian processes on the unit interval was obtained in 1956 by G. Baxter (2).

In 1950, by using the idea of spectral measures, U. Grenander (10) was able to give necessary and sufficient conditions for the equivalence of stationary Gaussian processes in the case where the index set T represents the real numbers. His method was to consider functions $A(s,t)$ and $B(s,t)$ defined on $T \times T$ such that $A(s,t) = a(s-t)$ and $B(s,t) = b(s-t)$, where both a and b are continuous positive definite functions on T. It then follows from Bochner's theorem (5, p. 126-128) that there exist finite regular Borel measures ρ and σ on T such that

$$a(t) = \int_{-\infty}^{+\infty} e^{it\omega} d\rho(\omega) \quad \text{and} \quad b(t) = \int_{-\infty}^{+\infty} e^{it\omega} d\sigma(\omega).$$

The measures ρ and σ are called spectral measures. If $m(t)$ and $n(t)$ are two functions defined on T, it then follows (6, p. 72;

8, p. 2.2) that there exist unique stationary Gaussian processes $X = \{x_t : t \in T\}$ and $Y = \{y_t : t \in T\}$ such that X has mean $m(t)$ and covariance $A(s,t)$ while Y has mean $n(t)$ and covariance $B(s,t)$. Using this notation, Grenander's result may then be stated:

THEOREM 1.1. X is equivalent to Y if and only if

(i) ρ and σ have identical non-atomic parts;

(ii) ρ and σ have exactly the same set of atoms Λ and if the masses at $i \in \Lambda$ are respectively α_i and β_i, then $\sum_{i \in \Lambda} (1 - \frac{\alpha_i}{\beta_i})^2 < \infty$.

Jacob Feldman has used spectral measures to study the question of equivalence of Gaussian processes restricted to finite intervals of the real line and has obtained results (9) for a special class of cases. Many others have done work on finding conditions for equivalence of Gaussian processes in terms of their spectral measures in various special cases. A summary of many of these results is to be found in (19).

In 1969, T. R. Chow (4) generalized Grenander's result to the case where the index set T is a separable locally compact group. This generalization was accomplished by using techniques of von Neuman algebras and direct integral representations. Chow was able to show that any stationary Gaussian process on a locally compact group admits a direct integral representation. This result together with the following theorem gives the generalization. The notation is that of (4).

THEOREM 1.2. Let G be a separable locally compact group. Two stationary processes $\{X(t), t \in G\}$ and $\{Y(t), t \in G\}$ with means zero, covariances ρ and σ respectively are equivalent if and only if they admit the direct integral representations

$$x(t) = \int_{\Lambda}^{\oplus} x_{\lambda}(t)d\mu \qquad y(t) = \int_{\Lambda}^{\oplus} y_{\lambda}(t)d\nu$$

where μ and ν are the central Radon measures of ρ and σ such that

(i) μ and ν have the identical non-atomic parts,

(ii) they have the same set of atoms which is countable such that $\sum_{a \in A} d(a)(1-(\mu(a)/\nu(a)))^2 < \infty$ where $d(a)$ is the dimension of $H_{\rho}(a)$ if $H_{\rho}(a)$ is finite dimensional and ∞ otherwise, and A is the set of all atoms.

In 1958, a rather surprising result was obtained independently by J. Hájek (11) and Jacob Feldman (7). They showed that if X and Y are Gaussian processes with the same index set, then either X and Y are equivalent or X and Y are perpendicular. The two methods of proof are quite different and both give necessary and sufficient conditions for equivalence.

Hájek's approach is based on the idea of entropy distance between measures. Let μ and ν be the measures generated respectively by X and Y on the path space $(R^T, \mathcal{C})$. Let μ_o and ν_o denote respectively the measures obtained by restricting μ and ν to the σ-algebra $\mathcal{C}_o$ generated by the cylinders with conditions on only a finite number of values of the index set T. The entropy distance between μ_o and ν_o is defined by:

$$J(\mu_o, \nu_o) = \int \log \frac{d\mu_o}{d\nu_o} d\mu_o + \int \log \frac{d\nu_o}{d\mu_o} d\nu_o .$$

The entropy distance between μ and ν is defined by $J(\mu, \nu) = \sup J(\mu_o, \nu_o)$, where the supremum is taken over all possible finite dimensional restrictions of μ and ν. Hájek's result is then stated as:

THEOREM 1.3. X is equivalent to Y if and only if $J(\mu, \nu) < \infty$. Further, if $J(\mu, \nu) = \infty$, then X and Y are perpendicular.

Feldman's approach makes use of Kakutani's theorem (14) and equivalence operators between Hilbert spaces. If H and K are Hilbert spaces, an operator $A: H \to K$ is an equivalence operator if A is a linear homeomorphism and $I - A^*A$ is Hilbert-Schmidt. Let μ and ν be the probabilities generated respectively by X

and Y on the path space $(R^T, \mathcal{C})$. Assume that $\int x_t^2 d\mu < \infty$ for all $x_t \in X$ and $\int y_t^2 d\nu < \infty$ for all $y_t \in Y$. Let H be the subspace of $L_2(\mu)$ generated by X and let K be the subspace of $L_2(\nu)$ generated by Y. Feldman's necessary and sufficient conditions for equivalence are given by the following theorem. The version given here appears in (8).

THEOREM 1.4. If X and Y are Gaussian processes with means zero, then X is equivalent to Y if and only if there is an equivalence operator from H onto K, sending the equivalence class of x_t to that of y_t.

Yet another approach to the problem is to use the theory of reproducing kernel Hilbert spaces (1). In a paper (17) appearing in 1963, E. Parzen was able to give necessary and sufficient conditions for the equivalence of Gaussian processes in terms of their covariance functions by using reproducing kernel Hilbert spaces.

Let X and Y be Gaussian processes with covariance functions $A(s,t)$ and $B(s,t)$, respectively. Let $H_{A \otimes B}$ be the reproducing kernel Hilbert space generated by
$(A \times B)((s,t),(u,v)) = A(s,u)B(t,v)$. Parzen proved the following:

THEOREM 1.5. X is equivalent to Y if and only if $A - B \in H_{A \otimes B}$.

A proof of this last result also appears in (8), where an example case in which X represents Brownian motion on $T = [0, +\infty)$ is worked out. The same result for this special case was obtained by L. A. Shepp (21) using methods very much different from those of Parzen.

The applications of this subject to statistics, time series analysis, and related fields are numerous. A collection of applications is to be found in (18). Indeed, one of the major efforts in this area is the translation of the general results listed above to special cases so as to make them easily applicable.

A companion question to the question of finding necessary and sufficient conditions for the equivalence of measures deals with finding representations of Radon-Nikodym derivatives. That is, if μ and ν are equivalent measures, just what does $\frac{d\mu}{d\nu}$ look like? Results on this representation question have been obtained by Feldman (8) and Parzen (17). Feldman's results are obtained by using the theory of Hilbert-Schmidt and trace operators. Parzen derives his results by using reproducing kernel Hilbert space theory.

In this paper, necessary and sufficient conditions are found for the equivalence of finite measures in terms of maps between function spaces and in terms of a Radon-Nikodym derivative.

The companion question of finding a representation for a Radon-Nikodym derivative is also examined. If μ and ν are finite measures on the same measurable space and $\lambda = \mu + \nu$, then the function $f(\cdot) = <\cdot, 1>_\mu$ is shown to be a continuous linear functional on $L_2(\lambda)$. By the Riesz representation theorem, there exists a unique vector $a \epsilon L_2(\lambda)$ such that $f(x) = <x, a>_\lambda$ for all $x \epsilon L_2(\lambda)$. The uniqueness of $\frac{d\mu}{d\lambda}$ allows us to show $a = \frac{d\mu}{d\lambda}$ λ-a.e. We may thus represent $\frac{d\mu}{d\lambda}$ as a Fourier series using any complete orthonormal system of $L_2(\lambda)$ since $\frac{d\mu}{d\lambda} \epsilon L_2(\lambda)$. An immediate consequence of this representation of $\frac{d\mu}{d\lambda}$ is a representation of $\frac{d\mu}{d\nu}$ when μ and ν are equivalent.

The problem of finding complete orthonormal systems for L_2-spaces is examined in the case where μ and ν are equivalent and $\frac{d\mu}{d\nu}$ is known. If a complete orthonormal system is known for $L_2(\mu)$, we give a method for finding a complete orthonormal system for $L_2(\nu)$. This result is then applied to examples involving classical complete orthonormal systems.

Finally, we given necessary and sufficient conditions for the equivalence of Gaussian processes with countable index sets. These conditions are given in terms of the finite dimensional distributions of the processes by a direct application of Kakutani's theorem and our result on the Radon-Nikodym derivative.

II. TERMINOLOGY AND NOTATION

Measures and Related Topics

This section is devoted to the definitions and notations of measure theory needed in this paper.

DEFINITION 2.1. Let Ω be a set and let $\mathcal{A}$ be a σ-algebra of subsets of Ω. The ordered pair $(\Omega, \mathcal{A})$ is called a measurable space, and the elements of $\mathcal{A}$ are called $\mathcal{A}$-measurable sets or just measurable sets.

DEFINITION 2.2. By a measure μ on a measurable space $(\Omega, \mathcal{A})$ we mean a nonnegative extended real-valued set function defined on $\mathcal{A}$ satisfying $\mu(\emptyset) = 0$ and $\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$ for any sequence $\{A_i\}$ of disjoint measurable sets. A measurable space $(\Omega, \mathcal{A})$ together with a measure μ is called a measure space and is denoted by $(\Omega, \mathcal{A}, \mu)$. A measure μ on $(\Omega, \mathcal{A})$ is called a finite measure if $\mu(\Omega) < \infty$. A measure μ on $(\Omega, \mathcal{A})$ is called a probability measure if $\mu(\Omega) = 1$. A measure space $(\Omega, \mathcal{A}, \mu)$ is called a complete measure space, and μ is called a complete measure, provided $\mu(A) = 0$ and $B \subseteq A$ imply $B \in \mathcal{A}$. That is, every subset of a set of measure zero is a measurable set.

DEFINITION 2.3. Let μ and ν be two measures defined on the same measurable space $(\Omega, \mathcal{A})$. The measure ν is said to be absolutely continuous with respect to μ if $A \in \mathcal{A}$ and $\mu(A) = 0$ imply $\nu(A) = 0$. If ν is absolutely continuous with respect to μ, we write $\nu \ll \mu$. If $\mu \ll \nu$ and $\nu \ll \mu$, then μ and ν are said to be equivalent and we write $\mu \sim \nu$. Thus, μ and ν are equivalent if and only if they have exactly the same sets of measure zero. If there exists a set $A \in \mathcal{A}$ such that $\mu(A) = 0 = \nu(A^c)$, μ and ν are said to be mutually singular or perpendicular and we write $\mu \perp \nu$. We shall write $\mu = \nu$ if $\mu(A) = \nu(A)$ for all $A \in \mathcal{A}$ and $\mu \leq \nu$ if $\mu(A) \leq \nu(A)$ for all $A \in \mathcal{A}$. Note that if $\mu \leq \nu$, then $\mu \ll \nu$. We define the measure $\mu + \nu$ by $(\mu+\nu)(A) = \mu(A) + \nu(A)$ for all $A \in \mathcal{A}$.

The idea of a measurable function plays an important role in later results and is introduced in the following definition.

DEFINITION 2.4. Let $(\Omega_1, \mathcal{A}_1, \mu_1)$ and $(\Omega_2, \mathcal{A}_2, \mu_2)$ be any two measure spaces. Let x be a function from $(\Omega_1, \mathcal{A}_1, \mu_1)$ to $(\Omega_2, \mathcal{A}_2, \mu_2)$. We say that x is an $\mathcal{A}_1$-measurable function, or just a measurable function, if $x^{-1}(A) \in \mathcal{A}_1$ for all $A \in \mathcal{A}_2$. Of particular interest is the case where Ω_2 is the real numbers and $\mathcal{A}_2$ the Borel sets.

The next definition introduces an equivalence relation on the set of real-valued functions defined on a measure space. It plays a basic role in later results.

DEFINITION 2.5. If x and y are real-valued functions defined on $(\Omega, \mathcal{A}, \mu)$, we say x is equal to y μ-almost everywhere provided they differ only on a set of μ-measure zero. This will be written as $x = y$ μ-a.e. The equivalence classes will be denoted by $[x]_\mu = \{y : y = x \quad \mu\text{-a.e.}\}$.

Several of the later results depend on the concept of a function space. The following definition includes the notation to be used.

DEFINITION 2.6. Let p be any real number greater than or equal to one and let $(\Omega, \mathcal{A}, \mu)$ be any measure space. If x represents a measurable real-valued function defined on $(\Omega, \mathcal{A}, \mu)$, then:

$$\mathcal{L}_p(\Omega, \mathcal{A}, \mu) = \{x : \int |x|^p d\mu < \infty\}$$

$$\mathcal{L}_p^+(\Omega, \mathcal{A}, \mu) = \{x : x \in \mathcal{L}_p(\Omega, \mathcal{A}, \mu) \text{ and } x \geq 0\}.$$

If $[x]_\mu$ is the equivalence class introduced in DEFINITION 2.5, then:

$$L_p(\Omega, \mathcal{A}, \mu) = \{[x]_\mu : x \in \mathcal{L}_p(\Omega, \mathcal{A}, \mu)\}.$$

When there is no question as to what measurable space μ is defined on, we write $\mathcal{L}_p(\mu)$ instead of $\mathcal{L}_p(\Omega, \mathcal{A}, \mu)$ and $L_p(\mu)$ instead of $L_p(\Omega, \mathcal{A}, \mu)$.

Recall that the spaces $L_p(\Omega, \mathcal{A}, \mu)$ are real, normed linear spaces for $1 \leq p < \infty$ with a norm defined by:

$$\| [x]_\mu \|_\mu = \left[\int |x|^p d\mu \right]^{1/p}.$$

It should also be recalled that the space $L_2(\Omega, \mathcal{A}, \mu)$ is a Hilbert space with an inner product defined by:

$$<[x]_\mu, [y]_\mu>_\mu = \int xy d\mu.$$

Radon-Nikodym Theorem

Since the Radon-Nikodym derivative plays such an extensive role in this paper, we state the Radon-Nikodym theorem here in the form which is most useful to us for applications.

THEOREM 2.7. Let $(\Omega, \mathcal{A})$ be a measurable space and let μ and ν be finite measures on $(\Omega, \mathcal{A})$ such that $\nu \ll \mu$. Then there exists a function $x_o \in \mathcal{L}_1^+(\Omega, \mathcal{A}, \mu)$ such that $\nu(A) = \int_A x_o d\mu$ for all $A \in \mathcal{A}$. Also, if $x \in \mathcal{L}_1(\Omega, \mathcal{A}, \nu)$, then $xx_o \in \mathcal{L}_1(\Omega, \mathcal{A}, \mu)$ and $\int x d\nu = \int xx_o d\mu$. Moreover, x_o is unique in the sense that if y_o is any real-valued, $\mathcal{A}$-measurable function such that $\nu(A) = \int_A y_o d\mu$ for all $A \in \mathcal{A}$, then $x_o = y_o$ μ-a.e.

Proof:

See (12, p. 315).

Because of the uniqueness of x_o in the theorem above, we write $x_o = \frac{d\nu}{d\mu}$ μ-a.e. and call $\frac{d\nu}{d\mu}$ the Radon-Nikodym derivative of ν with respect to μ.

We now give two lemmas dealing with some elementary properties of Radon-Nikodym derivatives which are useful in later results.

LEMMA 2.8. Let μ, ν, and λ be finite measures on the same measurable space.

(a) If $\mu \ll \lambda$ and $\nu \ll \lambda$, then $\frac{d(\mu+\nu)}{d\lambda} = \frac{d\mu}{d\lambda} + \frac{d\nu}{d\lambda}$ λ-a.e.

(b) If $\nu \ll \mu \ll \lambda$, then $\frac{d\nu}{d\lambda} = \frac{d\nu}{d\mu}\frac{d\mu}{d\lambda}$ λ-a.e.

(c) If $\mu \sim \nu$, then $\frac{d\nu}{d\mu} = [\frac{d\mu}{d\nu}]^{-1}$ μ-a.e.

Proof:

See (12, p. 328).

LEMMA 2.9. Let μ and ν be finite measures on the same measurable space and let $\lambda = \mu + \nu$.

(a) $\frac{d\mu}{d\lambda} + \frac{d\nu}{d\lambda} = 1$ λ-a.e.

(b) $0 \leq \frac{d\mu}{d\lambda} \leq 1$ λ-a.e. and $0 \leq \frac{d\nu}{d\lambda} \leq 1$ λ-a.e.

(c) $0 < \frac{d\mu}{d\lambda} < 1$ λ-a.e. if and only if $0 < \frac{d\nu}{d\lambda} < 1$ λ-a.e.

Proof:

Part (a) follows at once from LEMMA 2.8(a) since

$$1 = \frac{d\lambda}{d\lambda} = \frac{d(\mu+\nu)}{d\lambda} = \frac{d\mu}{d\lambda} + \frac{d\nu}{d\lambda} \quad \lambda\text{-a.e.}$$

Recalling that $\frac{d\mu}{d\lambda}, \frac{d\nu}{d\lambda} \in \mathcal{L}_1^+(\lambda)$, parts (b) and (c) follow at once from part (a) of this lemma. Q.E.D.

III. MAPS BETWEEN FUNCTION SPACES

Maps between $L_p(\mu)$ and $L_p(\nu)$

This chapter contains two theorems which give necessary and sufficient conditions for the equivalence of finite measures in terms of maps between function spaces. Throughout this chapter μ and ν will represent finite measures on the same measurable space $(\Omega, \mathcal{A})$ and λ will be the finite measure on $(\Omega, \mathcal{A})$ defined by $\lambda = \mu + \nu$.

The next theorem is suggested by an approach due to Neveu (16, p. 112).

THEOREM 3.1. Let p be a real number such that $1 \leq p < \infty$. A necessary and sufficient condition for the equivalence of μ and ν is that the following hold:

(a) There exists an isometric isomorphism u_p of $L_p(\mu)$ onto $L_p(\nu)$.

(b) $\frac{d\mu}{d\lambda} \leq |u_p(1)|^p \frac{d\nu}{d\lambda}$ or $\frac{d\mu}{d\lambda} \geq |u_p(1)|^p \frac{d\nu}{d\lambda}$ λ-a.e.

(c) $\frac{d\nu}{d\lambda} \leq |u_p^{-1}(1)|^p \frac{d\mu}{d\lambda}$ or $\frac{d\nu}{d\lambda} \geq |u_p^{-1}(1)|^p \frac{d\mu}{d\lambda}$ λ-a.e.

Proof:

To show the necessity, we assume that $\mu \sim \nu$. It then follows at once that $\mu \sim \lambda \sim \nu$. To save writing and simplify the notation in this part of the proof, we shall write x instead of $[x]_\mu$. All equalities are therefore to be understood as μ-a.e. (and hence also as ν-a.e. and λ-a.e.).

Since $\mu \sim \nu$, $\frac{d\mu}{d\nu}$ exists, and for all $x \in L_p(\mu)$ we define $u_p(x) = x(\frac{d\mu}{d\nu})^{1/p}$.

Now $x \in L_p(\mu)$ implies $|x|^p \in L_1(\mu)$, so it follows from Theorem 2.7 that

$$\int |u_p(x)|^p d\nu = \int |x(\frac{d\mu}{d\nu})^{1/p}|^p d\nu = \int |x|^p \frac{d\mu}{d\nu} d\nu = \int |x|^p d\mu < \infty,$$

and therefore, u_p is a map from $L_p(\mu)$ into $L_p(\nu)$.

u_p is clearly a linear map. That is, u_p is a vector space homomorphism.

To see that u_p is one-to-one, assume $u_p(x_1) = u_p(x_2)$ for any $x_1, x_2 \in L_p(\mu)$. We thus have $x_1(\frac{d\mu}{d\nu})^{1/p} = x_2(\frac{d\mu}{d\nu})^{1/p}$. Now since $\frac{d\nu}{d\mu}$ exists and by LEMMA 2.8(c) we have $\frac{d\nu}{d\mu} = (\frac{d\mu}{d\nu})^{-1}$, we conclude that $x_1(\frac{d\mu}{d\nu})^{1/p}(\frac{d\mu}{d\nu})^{-1/p} = x_2(\frac{d\mu}{d\nu})^{1/p}(\frac{d\mu}{d\nu})^{-1/p}$. Therefore, $x_1 = x_2$ and u_p is one-to-one.

To see that u_p is onto, choose any $x \in L_p(\nu)$ and consider $x(\frac{d\nu}{d\mu})^{1/p}$. Since:

$$\int |x(\frac{d\nu}{d\mu})^{1/p}|^p d\mu = \int |x|^p \frac{d\nu}{d\mu} d\mu = \int |x|^p d\nu < \infty,$$

we have $x(\frac{d\nu}{d\mu})^{1/p} \in L_p(\mu)$. By applying u_p to this element we get:

$$u_p(x(\frac{d\nu}{d\mu})^{1/p}) = x(\frac{d\nu}{d\mu})^{1/p}(\frac{d\mu}{d\nu})^{1/p} = x(\frac{d\nu}{d\mu})^{1/p}(\frac{d\nu}{d\mu})^{-1/p} = x.$$

Hence, u_p is onto and it follows that u_p is an isomorphism.

The map u_p is norm preserving since:

$$\|x\|_\mu^p = \int |x|^p d\mu = \int |x|^p \frac{d\mu}{d\nu} d\nu = \int |x(\frac{d\mu}{d\nu})^{1/p}|^p d\nu$$

$$= \int |u_p(x)|^p d\nu = \|u_p(x)\|_\nu^p$$

for all $x \in L_p(\mu)$.

Since u_p is linear and norm preserving, it follows that u_p is an isometry.

Properties (b) and (c) are obvious since $\frac{d\mu}{d\lambda} = \frac{d\mu}{d\nu}\frac{d\nu}{d\lambda}$ and $\frac{d\nu}{d\lambda} = \frac{d\nu}{d\mu}\frac{d\mu}{d\lambda}$ by LEMMA 2.8(b), while $|u_p(1)|^p = \frac{d\mu}{d\nu}$ and $|u_p^{-1}(1)|^p = \frac{d\nu}{d\mu}$. This proves necessity.

To prove the converse, assume conditions (a), (b), and (c) hold. Because u_p is norm preserving we have:

$$\|1\|_\mu^p = \|u_p(1)\|_\nu^p,$$

$$\int d\mu = \int |u_p(1)|^p d\nu .$$

Applying THEOREM 2.7, since $\mu \leq \lambda$ and $\nu \leq \lambda$, we get:

$$\int \frac{d\mu}{d\lambda}\, d\lambda = \int |u_p(1)|^p \frac{d\nu}{d\lambda}\, d\lambda .$$

It then follows that:

$$\int \left(\frac{d\mu}{d\lambda} - |u_p(1)|^p \frac{d\nu}{d\lambda}\right) d\lambda = 0 \quad \text{and} \quad \int \left(|u_p(1)|^p \frac{d\nu}{d\lambda} - \frac{d\mu}{d\lambda}\right) d\lambda = 0.$$

This means that if either inequality of (b) holds, we have:

$$\frac{d\mu}{d\lambda} - |u_p(1)|^p \frac{d\nu}{d\lambda} = 0 \quad \lambda\text{-a.e.},$$

$$\frac{d\mu}{d\lambda} = |u_p(1)|^p \frac{d\nu}{d\lambda} \quad \lambda\text{-a.e.}$$

If we let 1_A represent the characteristic function of $A \in \mathcal{A}$, then we have:

$$1_A \frac{d\mu}{d\lambda} = 1_A |u_p(1)|^p \frac{d\nu}{d\lambda} \quad \lambda\text{-a.e.} \quad \text{for all } A \in \mathcal{A},$$

$$\int_A \frac{d\mu}{d\lambda}\, d\lambda = \int_A |u_p(1)|^p \frac{d\nu}{d\lambda}\, d\lambda \quad \text{for all } A \in \mathcal{A},$$

$$\mu(A) = \int_A |u_p(1)|^p d\nu \quad \text{for all } A \in \mathcal{A}.$$

Therefore, $\mu \ll \nu$.

Similarly, $\|u_p^{-1}(1)\|_\mu^p = \|1\|_\nu^p$ implies

$$\int |u_p^{-1}(1)|^p \frac{d\mu}{d\lambda}\, d\lambda = \int \frac{d\nu}{d\lambda}\, d\lambda ,$$

and it then follows from condition (c) that:

$$\frac{d\nu}{d\lambda} = |u_p^{-1}(1)|^p \frac{d\mu}{d\lambda} \quad \lambda\text{-a.e.}$$

$$1_A \frac{d\nu}{d\lambda} = 1_A |u_p^{-1}(1)|^p \frac{d\mu}{d\lambda} \quad \lambda\text{-a.e.} \quad \text{for all} \quad A \in \mathcal{Q},$$

$$\int_A \frac{d\nu}{d\lambda} d\lambda = \int_A |u_p^{-1}(1)|^p \frac{d\mu}{d\lambda} d\lambda \quad \text{for all} \quad A \in \mathcal{Q},$$

$$\nu(A) = \int_A |u_p^{-1}(1)|^p d\mu \quad \text{for all} \quad A \in \mathcal{Q}.$$

Therefore, $\nu \ll \mu$ and $\mu \sim \nu$.

We note that because of the uniqueness of the Radon-Nikodym derivative we must have $|u_p(1)|^p = \frac{d\mu}{d\nu}$ ν-a.e. and $|u_p^{-1}(1)|^p = \frac{d\nu}{d\mu}$ μ-a.e. Q.E.D.

The following observations should be made from the proof of THEOREM 3.1. In the proof of the necessity of conditions (a), (b), and (c), the conditions (b) and (c) required no real proof since $|u_p(1)|^p = \frac{d\mu}{d\nu}$ λ-a.e. and $|u_p^{-1}(1)|^p = \frac{d\nu}{d\mu}$ λ-a.e. Thus, conditions (b) and (c) were really special cases of LEMMA 2.8(b).

In the proof that conditions (a), (b), and (c) are sufficient, the full strength of (a) was not used. We only needed the fact that the map u_p is one-to-one and norm preserving.

Finally, the constant function 1 is not as significant as it might appear. The entire theorem may be restated as follows:

THEOREM 3.1'. μ and ν are equivalent if and only if there exists an $x_o \in L_p(\mu)$ with $||x_o||_\mu = 1$, and the following hold:

(a') There exists an isometric isomorphism u_p of $L_p(\mu)$ onto $L_p(\nu)$.

(b') $\frac{d\mu}{d\lambda} \leq |u_p(x_o)|^p \frac{d\nu}{d\lambda} \mu(\Omega)$ or $\frac{d\mu}{d\lambda} \geq |u_p(x_o)|^p \frac{d\nu}{d\lambda} \mu(\Omega)$ λ-a.e.

(c') $\frac{d\nu}{d\lambda} \leq |x_o|^p \frac{d\mu}{d\lambda} \nu(\Omega)$ or $\frac{d\nu}{d\lambda} \geq |x_o|^p \frac{d\mu}{d\lambda} \nu(\Omega)$ λ-a.e.

<u>Maps from $L_p(\lambda)$ into $L_p(\mu)$ and $L_p(\nu)$</u>

The approach used in this section is suggested by Feldman's result on the equivalence of Gaussian processes (8). Several lemmas are needed for our next result and we state them now.

LEMMA 3.2. Let μ and δ be measures on $(\Omega, \mathcal{A})$ such that $\mu \leq \delta$. If $p \geq 1$ and $x \in \mathcal{L}_p(\Omega, \mathcal{A}, \delta)$, it follows that $x \in \mathcal{L}_p(\Omega, \mathcal{A}, \mu)$ and $\int |x|^p d\mu \leq \int |x|^p d\delta$.

Proof:

See (12, p. 313).

LEMMA 3.3. Let μ and ν be finite measures on the measurable space $(\Omega, \mathcal{A})$ and define $\lambda = \mu + \nu$. For $1 \leq p < \infty$, there is a well-defined map $\theta_p : L_p(\lambda) \to L_p(\mu)$ given by

$\theta_p([x]_\lambda) = [x]_\mu$.

Proof:

Since $\lambda = \mu + \nu$, it follows that $\mu \leq \lambda$.

If $[x]_\lambda \in L_p(\lambda)$, it follows from LEMMA 3.2 that $\int |x|^p d\mu \leq \int |x|^p d\lambda$ and hence $\theta_p([x]_\lambda) \in L_p(\mu)$. Thus, θ_p is a map from $L_p(\lambda)$ into $L_p(\mu)$.

To see that θ_p is well-defined, suppose $[x_1]_\lambda$ and $[x_2]_\lambda$ are elements of $L_p(\lambda)$ and assume $[x_1]_\mu \neq [x_2]_\mu$. Then there exists a set $A \in \mathcal{A}$ such that $\mu(A) > 0$ and $x_1(\omega) \neq x_2(\omega)$ for $\omega \in A$. But, $\mu(A) > 0$ implies $\lambda(A) > 0$ and hence $[x_1]_\lambda \neq [x_2]_\lambda$. Therefore, θ_p is a well-defined map. Q.E.D.

The map $\theta_p : [x]_\lambda \mapsto [x]_\mu$ is often referred to as the induced identity map.

The next lemma provides the major part of the proof of the theorem of this section.

LEMMA 3.4. A necessary and sufficient condition for $\mu \sim \lambda$ is that the map θ_p defined in LEMMA 3.3 be one-to-one.

Proof:

Assume $\mu \sim \lambda$ and $[x_1]_\lambda \neq [x_2]_\lambda$ for two elements $[x_1]_\lambda$, $[x_2]_\lambda$ of $L_p(\lambda)$. However, $\mu \sim \lambda$ means that μ and λ have exactly the same sets of measure zero, so that $[x_1]_\mu = [x_1]_\lambda$ and $[x_2]_\mu = [x_2]_\lambda$. Therefore, $[x_1]_\mu \neq [x_2]_\mu$, and θ_p is one-to-one.

Now assume θ_p is one-to-one. Since $\mu \leq \lambda$ implies $\mu \ll \lambda$, it suffices to show that $A \in \mathcal{A}$ and $\mu(A) = 0$ imply $\lambda(A) = 0$. If $A \in \mathcal{A}$, then $[1_A]_\lambda \in L_p(\lambda)$ since $|1_A|^p = 1_A$ and $\int |1_A|^p d\lambda = \int 1_A d\lambda = \lambda(A) < \infty$. If $\mu(A) = 0$, then $[1_A]_\mu = [0]_\mu$. However, $\theta_p([1_A]_\lambda) = [1_A]_\mu = [0]_\mu$ and since θ_p is one-to-one and $\theta_p([0]_\lambda) = [0]_\mu$, we must have $[1_A]_\lambda = [0]_\lambda$. Therefore, $\lambda(A) = 0$ and $\lambda \ll \mu$. Hence, $\mu \sim \lambda$. Q.E.D.

Since $\lambda = \mu + \nu$, a completely analogous procedure, or a direct application of the preceding lemmas, allows us to define a map $\phi_p : L_p(\lambda) \to L_p(\nu)$ by $\phi_p([x]_\lambda) = [x]_\nu$ and conclude:

LEMMA 3.5. A necessary and sufficient condition for $\nu \sim \lambda$ is that the map ϕ_p be one-to-one.

Combining the last three lemmas we obtain necessary and sufficient conditions for the equivalence of μ and ν in the next theorem.

THEOREM 3.6. Let μ and ν be finite measures on the same measurable space $(\Omega, \mathcal{A})$ and let $\lambda = \mu + \nu$. The measures μ and ν are equivalent if and only if both the maps θ_p and ϕ_p defined above are one-to-one.

Proof:

$\mu \sim \nu$ if and only if $\mu \sim \lambda$ and $\nu \sim \lambda$. But, $\mu \sim \lambda$ and $\nu \sim \lambda$ if and only if θ_p and ϕ_p are one-to-one by LEMMA 3.4 and LEMMA 3.5 respectively. Q.E.D.

It should be observed that θ_p and ϕ_p are continuous linear maps, but they are not in general norm preserving.

An examination of the proofs of this section also shows that THEOREM 3.6 is still valid if $L_p(\Omega, \mathcal{A}, \mu)$, $L_p(\Omega, \mathcal{A}, \nu)$, and $L_p(\Omega, \mathcal{A}, \lambda)$ are replaced by the vector spaces of equivalence classes of bounded measurable functions on $(\Omega, \mathcal{A}, \mu)$, $(\Omega, \mathcal{A}, \nu)$ and $(\Omega, \mathcal{A}, \lambda)$ respectively.

Finally, we make some observations about THEOREM 3.6 in comparison to THEOREM 3.1. The maps in THEOREM 3.6 will always exist, while the existence of the map in THEOREM 3.1 is not guaranteed in one direction. On the other hand, $L_p(\lambda)$ must be used in THEOREM 3.6, so we have maps between three spaces while THEOREM 3.1 requires maps between only two spaces. However, $L_1(\lambda)$ is used in THEOREM 3.1 since we have conditions on $\frac{d\mu}{d\lambda}$ and $\frac{d\nu}{d\lambda}$.

IV. A RADON-NIKODYM DERIVATIVE

Conditions on the Derivative

In this chapter we give necessary and sufficient conditions for the equivalence of finite measures on the same measurable space in terms of a Radon-Nikodym derivative. As before, μ and ν will represent finite measures on $(\Omega, \mathcal{A})$ and $\lambda = \mu + \nu$. Now, $\mu \leq \lambda$ and $\nu \leq \lambda$, so that $\mu \ll \lambda$ and $\nu \ll \lambda$. It therefore follows from THEOREM 2.7 that $\frac{d\mu}{d\lambda}$ and $\frac{d\nu}{d\lambda}$ exist in $\mathcal{L}_1^+(\lambda)$.

The first theorem of this section requires two lemmas. The next lemma is a modified form of THEOREM 2.7.

LEMMA 4.1. The measures μ and ν are equivalent if and only if there exist functions $x_o \in \mathcal{L}_1^+(\mu)$ and $y_o \in \mathcal{L}_1^+(\nu)$ such that $\frac{d\mu}{d\lambda} = y_o \frac{d\nu}{d\lambda}$ λ-a.e. and $\frac{d\nu}{d\lambda} = x_o \frac{d\mu}{d\lambda}$ λ-a.e.

Proof:

If $\mu \sim \nu$, then $\mu \sim \lambda \sim \nu$. It follows from THEOREM 2.7 that $\frac{d\mu}{d\nu}$ and $\frac{d\nu}{d\mu}$ both exist and by LEMMA 2.8(b) we have $\frac{d\mu}{d\lambda} = \frac{d\mu}{d\nu}\frac{d\nu}{d\lambda}$ λ-a.e. and $\frac{d\nu}{d\lambda} = \frac{d\nu}{d\mu}\frac{d\mu}{d\lambda}$ λ-a.e. Thus, we let $y_o = \frac{d\mu}{d\nu}$ ν-a.e. and $x_o = \frac{d\nu}{d\mu}$ μ-a.e.

To prove the converse, suppose x_o and y_o exist. Then we have:

$$\frac{d\mu}{d\lambda} = y_o \frac{d\nu}{d\lambda} \quad \lambda\text{-a.e.},$$

$$1_A \frac{d\mu}{d\lambda} = 1_A y_o \frac{d\nu}{d\lambda} \quad \lambda\text{-a.e.} \quad \text{for all} \quad A \in \mathcal{A},$$

$$\int_A \frac{d\mu}{d\lambda} d\lambda = \int_A y_o \frac{d\nu}{d\lambda} d\lambda \quad \text{for all} \quad A \in \mathcal{A},$$

$$\mu(A) = \int_A y_o d\nu \quad \text{for all} \quad A \in \mathcal{A}.$$

We thus have $\mu \ll \nu$.

A completely analogous argument starting from $\frac{d\nu}{d\lambda} = x_o \frac{d\mu}{d\lambda}$ λ-a.e. shows that $\nu \ll \mu$. Therefore, $\mu \sim \nu$. Q.E.D.

Our next lemma requires that μ and ν be complete measures. Note that if μ and ν are complete, then λ is also complete.

LEMMA 4.2. Suppose μ and ν are complete finite measures. If $0 < \frac{d\mu}{d\lambda} < 1$ λ-a.e., then there exists a $y_o \in \mathcal{L}_1^+(\nu)$ and an $x_o \in \mathcal{L}_1^+(\mu)$ such that $\frac{d\mu}{d\lambda} = y_o \frac{d\nu}{d\lambda}$ λ-a.e. and $\frac{d\nu}{d\lambda} = x_o \frac{d\mu}{d\lambda}$ λ-a.e.

Proof:

Define the real-valued function z on Ω by:

$$z(\omega) = \begin{cases} \frac{d\mu}{d\lambda}(\omega) & \text{if} \quad \frac{d\mu}{d\lambda}(\omega) \neq 0 \\ \frac{1}{2} & \text{if} \quad \frac{d\mu}{d\lambda}(\omega) = 0. \end{cases}$$

Since z is everywhere nonzero and since $\frac{d\mu}{d\lambda}$ is measurable, it follows (12, p. 151) that $\frac{1}{z} = z^{-1}$ is measurable. Because λ is complete and because $0 < \frac{d\mu}{d\lambda} < 1$ λ-a.e. gives us $z^{-1} = (\frac{d\mu}{d\lambda})^{-1}$ λ-a.e., it follows (12, p. 155) that $(\frac{d\mu}{d\lambda})^{-1}$ is measurable.

Since $(\frac{d\mu}{d\lambda})^{-1}$ and $\frac{d\nu}{d\lambda}$ are nonnegative measurable functions and the product of measurable functions is measurable (12, p. 153), it follows that $(\frac{d\nu}{d\lambda})(\frac{d\mu}{d\lambda})^{-1}$ is measurable and nonnegative.

Since $\frac{d\nu}{d\lambda} \; (\frac{d\mu}{d\lambda})^{-1} \; \frac{d\mu}{d\lambda} = \frac{d\nu}{d\lambda}$ λ-a.e., we have:

$$\int \frac{d\nu}{d\lambda} \; (\frac{d\mu}{d\lambda})^{-1} d\mu = \int \frac{d\nu}{d\lambda} \; (\frac{d\mu}{d\lambda})^{-1} \frac{d\mu}{d\lambda} d\lambda = \int \frac{d\nu}{d\lambda} d\lambda = \nu(\Omega) < \infty .$$

Therefore, $\frac{d\nu}{d\lambda} \; (\frac{d\mu}{d\lambda})^{-1} \in \mathcal{L}_1^+(\mu)$ and has the desired property. Thus, we take $x_o = \frac{d\nu}{d\lambda} \; (\frac{d\mu}{d\lambda})^{-1}$ λ-a.e.

It follows from LEMMA 2.9(c) that $0 < \frac{d\mu}{d\lambda} < 1$ λ-a.e. implies that $0 < \frac{d\nu}{d\lambda} < 1$ λ-a.e., so a similar argument shows y_o can be taken to be $\frac{d\mu}{d\lambda} \; (\frac{d\nu}{d\lambda})^{-1}$ λ-a.e. Q.E.D.

We now may prove the following theorem which gives a condition for equivalence in terms of $\frac{d\mu}{d\lambda}$.

THEOREM 4.3. If μ and ν are complete finite measures on $(\Omega, \mathcal{A})$ and $\lambda = \mu + \nu$, then μ and ν are equivalent if and only if $0 < \frac{d\mu}{d\lambda} < 1$ λ-a.e.

Proof:

Suppose $\mu \sim \nu$, so that $\mu \sim \lambda \sim \nu$. Since LEMMA 2.9(b) shows that $0 \leq \frac{d\mu}{d\lambda} \leq 1$ λ-a.e. in every case, we need only show that $\frac{d\mu}{d\lambda}$ assumes the values zero and one only on sets of λ-measure zero.

Suppose $\frac{d\mu}{d\lambda}(\omega) = 0$ for all $\omega \in A \in \mathcal{A}$. Then we have $\mu(A) = \int_A \frac{d\mu}{d\lambda} d\lambda = 0$. Now since $\mu \sim \lambda$, we have $\lambda(A) = 0$.

On the other hand, if $\frac{d\mu}{d\lambda}(\omega) = 1$ for all $\omega \in A$, we have $\mu(A) = \int_A \frac{d\mu}{d\lambda} d\lambda = \int_A d\lambda = \lambda(A) = \mu(A) + \nu(A)$ and hence $\nu(A) = 0$. But, $\nu \sim \lambda$ implies that $\lambda(A) = 0$. Therefore, $\mu \sim \nu$ implies $0 < \frac{d\mu}{d\lambda} < 1$ λ-a.e.

To prove the converse, assume $0 < \frac{d\mu}{d\lambda} < 1$ λ-a.e. By LEMMA 4.2, there must exist $y_o \in \mathcal{L}_1^+(\nu)$ and $x_o \in \mathcal{L}_1^+(\mu)$ such that $\frac{d\mu}{d\lambda} = y_o \frac{d\nu}{d\lambda}$ λ-a.e. and $\frac{d\nu}{d\lambda} = x_o \frac{d\mu}{d\lambda}$ λ-a.e. It then follows at once from LEMMA 4.1 that $\mu \sim \nu$. Q.E.D.

Completeness can be discarded in the last theorem. If μ is not equivalent to ν, then either μ is not absolutely continuous with respect to ν or ν is not absolutely continuous with respect to μ. Without loss of generality, we assume that μ is not absolutely continuous with respect to ν. Then there must exist a set $A \in \mathcal{A}$ such that $\nu(A) = 0$ and $\mu(A) > 0$.

Since $\nu \ll \lambda$, we have $0 = \nu(A) = \int_A \frac{d\nu}{d\lambda} d\lambda = \int 1_A \frac{d\nu}{d\lambda} d\lambda$. It follows that $1_A \frac{d\nu}{d\lambda} = 0$ λ-a.e. since $1_A \frac{d\nu}{d\lambda} \geq 0$ λ-a.e. Because $1_A = 1$ on A and $\lambda(A) = \mu(A) + \nu(A) = \mu(A) > 0$, it cannot be the case that $1_A = 0$ λ-a.e. Consequently, we must have $\frac{d\nu}{d\lambda} = 0$ on A. By LEMMA 2.9(a) we have $\frac{d\mu}{d\lambda} = 1 - \frac{d\nu}{d\lambda}$ λ-a.e., so that $\frac{d\mu}{d\lambda} = 1$ on A. Thus, we have $\frac{d\mu}{d\lambda} = 1$ on a set of positive λ-measure, so it cannot be the case that $0 < \frac{d\mu}{d\lambda} < 1$ λ-a.e.

Since the proof that $\mu \sim \nu$ implies $0 < \frac{d\mu}{d\lambda} < 1$ λ-a.e. did not require completeness, we have proved the following result:

THEOREM 4.4. If μ and ν are finite measures on $(\Omega, \mathcal{A})$ and $\lambda = \mu + \nu$, then μ is equivalent to ν if and only if $0 < \frac{d\mu}{d\lambda} < 1$ λ-a.e.

It should be pointed out that by LEMMA 2.9(c) we have $0 < \frac{d\mu}{d\lambda} < 1$ λ-a.e. if and only if $0 < \frac{d\nu}{d\lambda} < 1$ λ-a.e. This fact together with THEOREM 4.4 allows us to state the following result.

COROLLARY 4.5. If μ and ν are finite measures on the measurable space $(\Omega, \mathcal{A})$ and $\lambda = \mu + \nu$, then μ is equivalent to ν if and only if $0 < \frac{d\mu}{d\lambda} < 1$ λ-a.e. or $0 < \frac{d\nu}{d\lambda} < 1$ λ-a.e.

Special Cases

In this section we examine special cases in which the measurable space $(\Omega, \mathcal{A})$ is of such a nature that it is possible to define finite measures of the form $d\mu(\omega) = f(\omega)d\omega$ and $d\nu(\omega) = g(\omega)d\omega$. Of particular interest are the cases where Ω represents the real numbers R, finite subintervals of R, or n-dimensional Euclidean space R^n and $\mathcal{A}$ represents the respective Borel sets. Consequently, throughout this section $(\Omega, \mathcal{A})$ will be considered to be a space so that finite measures of the form $d\mu(\omega) = f(\omega)d\omega$ do exist on it.

If μ and ν are finite measures such that $d\mu(\omega) = f(\omega)d\omega$ and $d\nu(\omega) = g(\omega)d\omega$ and $\lambda = \mu + \nu$, then $d\lambda(\omega) = d\mu(\omega) + d\nu(\omega) = (f(\omega)+g(\omega))d\omega$. Since

$$\mu(A) = \int_A d\mu(\omega) = \int_A f(\omega)d\omega = \int_A \frac{f(\omega)}{f(\omega)+g(\omega)} (f(\omega)+g(\omega))d\omega$$

for all $A \in \mathcal{A}$, it follows that:

$$\frac{d\mu}{d\lambda}(\omega) = \frac{f(\omega)}{f(\omega)+g(\omega)} \quad \lambda\text{-a.e.}$$

Similarly:

$$\frac{d\nu}{d\lambda}(\omega) = \frac{g(\omega)}{f(\omega)+g(\omega)} \quad \lambda\text{-a.e.}$$

Note that $0 < \frac{f(\omega)}{f(\omega)+g(\omega)} < 1$ λ-a.e. may be restated as $0 < \frac{f(\omega)}{f(\omega)+g(\omega)} < 1$ wherever $\frac{f(\omega)}{f(\omega)+g(\omega)}$ is defined.

With the notation introduced above, it is now possible to restate

COROLLARY 4.5 in the context of this section.

COROLLARY 4.6. If μ and ν are finite measures on $(\Omega, \mathcal{A})$ having the form $d\mu(\omega) = f(\omega)d\omega$ and $d\nu(\omega) = g(\omega)d\omega$ and $\lambda = \mu + \nu$, then $\mu \sim \nu$ if and only if $0 < \frac{f(\omega)}{f(\omega)+g(\omega)} < 1$ wherever $\frac{f(\omega)}{f(\omega)+g(\omega)}$ is defined or $0 < \frac{g(\omega)}{f(\omega)+g(\omega)} < 1$ wherever $\frac{g(\omega)}{f(\omega)+g(\omega)}$ is defined.

If $d\mu(\omega) = f(\omega)d\omega$ and $d\nu(\omega) = g(\omega)d\omega$ and it is the case that $\mu \sim \nu$, then it is possible to use LEMMA 2.8 to find $\frac{d\mu}{d\nu}$. Thus:

$$\frac{d\mu}{d\nu} = \frac{d\mu}{d\lambda}\,\frac{d\lambda}{d\nu} = \frac{d\mu}{d\lambda}\left(\frac{d\nu}{d\lambda}\right)^{-1} = \frac{f}{f+g}\;\frac{f+g}{g} = \frac{f}{g} \quad \nu\text{-a.e.}$$

Therefore, $\frac{d\mu}{d\nu} = \frac{f}{g}$ wherever $\frac{f}{g}$ is defined.

We conclude this chapter by giving two specific examples.

EXAMPLE 4.7. Let $\Omega = [-1, 1]$ and let $\mathcal{B}$ be the Borel sets of $[-1, 1]$. On $(\Omega, \mathcal{B})$, let μ be Lebesgue measure so that $d\mu(\omega) = 1d\omega$ and let ν be the measure defined by $d\nu(\omega) = e^{\omega}d\omega$. Since $e^{\omega} > 0$ for all $\omega \in \Omega$, we have $0 < \frac{1}{1+e^{\omega}} < 1$ everywhere on Ω. It therefore follows from COROLLARY 4.6 that $\mu \sim \nu$. We also have $\frac{d\mu}{d\nu}(\omega) = e^{-\omega}$.

Our final example makes use of Gaussian density functions.

EXAMPLE 4.8. Let $\Omega = R$ and let $\mathcal{B}$ be the Borel sets of R. Let $m, \sigma \in R$ such that $\sigma > 0$ and suppose they are fixed.

For any $\omega \in R$, define the function $N(\sigma, m, \omega)$ by:

$$N(\sigma, m, \omega) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left[-\frac{1}{2}\left(\frac{\omega-m}{\sigma}\right)^2\right].$$

This is the density function of a nondegenerate Gaussian distribution with mean m and variance σ^2.

Define the measures μ and ν on $(\Omega, \mathcal{B})$ by $d\mu(\omega) = N(\sigma_1, m_1, \omega)d\omega$ and $d\nu(\omega) = N(\sigma_2, m_2, \omega)d\omega$ where $\sigma_1 > 0$ and $\sigma_2 > 0$. Since $\exp(\omega) > 0$ for all $\omega \in R$, it follows that:

$$0 < \frac{N(\sigma_1, m_1, \omega)}{N(\sigma_2, m_2, \omega) + N(\sigma_1, m_1, \omega)} < 1$$

for all $\omega \in R$. Therefore, $\mu \sim \nu$ and:

$$\frac{d\mu}{d\nu}(\omega) = \frac{N(\sigma_1, m_1, \omega)}{N(\sigma_2, m_2, \omega)}.$$

V. REPRESENTATIONS OF RADON-NIKODYM DERIVATIVES

The Riesz Theorem

The first representation of a Radon-Nikodym derivative given in this chapter makes extensive use of the Riesz representation theorem. For completeness and later reference we give the Riesz theorem and its proof here as found in Horvāth (13, p. 42).

THEOREM 5.1. Let E be a Hilbert space and g a continuous linear form on E. Then there exists a unique element $a \in E$ such that $g(x) = \langle x, a \rangle$ for all $x \in E$.

Proof:

If g is identically zero, choose $a = 0$. If g is not identically zero, then $(\ker g)$ is distinct from E and there must exist a nonzero vector $b \in E$ which is orthogonal to $(\ker g)$. Let:

$$(*) \qquad a = \frac{bg(b)}{\|b\|^2} .$$

Then $g(a) = \|a\|^2 = |g(b)|^2/\|b\|^2$. Since $(\ker g)$ is a hyperplane, every vector $x \in E$ can be written in the form $y + \alpha a$ where $y \in (\ker g)$ and α is a real number. Thus:

$$g(x) = g(y+\alpha a) = \alpha g(a) = \alpha\|a\|^2 = \langle \alpha a, a\rangle = \langle y+\alpha a, a\rangle = \langle x, a\rangle .$$

The uniqueness of a is clear. Q.E.D.

The next proposition introduces a continuous linear form on an L_2-space.

PROPOSITION 5.2. Let μ and λ be finite measures defined on the same measurable space $(\Omega, \mathcal{A})$ such that $\mu \leq \lambda$. Then the map g defined by $g([x]_\lambda) = <[x]_\lambda, [1]_\lambda>_\mu = \int x d\mu$ is a continuous linear form on $L_2(\Omega, \mathcal{A}, \lambda)$.

Proof:

Let $[x]_\lambda \in L_2(\lambda)$. Then $x \in \mathcal{L}_2(\lambda)$, and since $\mu \leq \lambda$, it follows from LEMMA 3.2 that $x \in \mathcal{L}_2(\mu)$. By the Schwarz inequality we have:

$$\left|\int x d\mu\right| \leq \int |x| d\mu \leq \left(\int |x|^2 d\mu\right)^{1/2} \left(\int 1 d\mu\right)^{1/2}$$

$$= \left(\int |x|^2 d\mu\right)^{1/2} (\mu(\Omega))^{1/2} < \infty.$$

The map g is thus defined and real-valued on $L_2(\lambda)$.

To see that g is well-defined, let $[x_1]_\lambda$ and $[x_2]_\lambda$ be elements of $L_2(\lambda)$ such that $g([x_1]_\lambda) \neq g([x_2]_\lambda)$. This means that $\int x_1 d\mu \neq \int x_2 d\mu$, so there must exist a set $A \in \mathcal{A}$ such that $\mu(A) > 0$ and $x_1(\omega) \neq x_2(\omega)$ for all $\omega \in A$. However, $\mu \leq \lambda$ and $\mu(A) > 0$ imply $\lambda(A) > 0$. Therefore, x_1 and x_2 differ on a set of positive λ-measure. Thus, $[x_1]_\lambda \neq [x_2]_\lambda$ and g is well-defined.

The fact that g is linear on $L_2(\lambda)$ follows from the fact that g is linear on $L_2(\mu)$ and LEMMA 3.2 which shows that if $x \in \mathcal{L}_2(\lambda)$, then $x \in \mathcal{L}_2(\mu)$.

To see that g is continuous on $L_2(\lambda)$, choose $[x]_\lambda \in L_2(\lambda)$ and use the Schwarz inequality to get:

$$|g([x]_\lambda)| = |\int x d\mu| \leq M \left(\int |x|^2 d\mu\right)^{1/2}$$

where $M = (\mu(\Omega))^{1/2}$.

By LEMMA 3.2 we have $\int |x|^2 d\mu \leq \int |x|^2 d\lambda$, so:

$$|g([x]_\lambda)| \leq M \left(\int |x|^2 d\lambda\right)^{1/2} = M\|[x]_\lambda\|_\lambda .$$

Thus, g is continuous on $L_2(\lambda)$. Q.E.D.

Since $L_2(\lambda)$ is a real Hilbert space and the function g defined in PROPOSITION 5.2 is a continuous linear form on $L_2(\lambda)$, THEOREM 5.1 applies directly to this case. We restate it here in this context.

THEOREM 5.3. If μ and λ are finite measures on the same measurable space $(\Omega, \mathcal{A})$ such that $\mu \leq \lambda$ and g is the continuous linear form defined on $L_2(\lambda)$ by $g([x]_\lambda) = \langle [x]_\lambda, [1]_\lambda \rangle_\mu$, then there exists a unique element $[a]_\lambda \in L_2(\lambda)$ such that $g([x]_\lambda) = \langle [x]_\lambda, [a]_\lambda \rangle_\lambda$ for every $[x]_\lambda \in L_2(\lambda)$.

We are now in a position to state and prove the theorem which relates the Radon-Nikodym derivative $\frac{d\mu}{d\lambda}$ and the Riesz vector a.

THEOREM 5.4. Let μ and ν be finite measures defined on the same measurable space $(\Omega, \mathcal{A})$ and let $\lambda = \mu + \nu$. If $[a]_\lambda$ is the unique element of $L_2(\lambda)$ guaranteed by THEOREM 5.3, then $\frac{d\mu}{d\lambda} = a$ λ-a.e.

Proof:

It follows from THEOREM 5.3 that $\int x d\mu = \int x a d\lambda$ for all $x \in \mathcal{L}_2(\lambda)$. Since $1_A \in \mathcal{L}_2(\lambda)$ for all $A \in \mathcal{A}$, we have $\int 1_A d\mu = \int 1_A a d\lambda$ for all $A \in \mathcal{A}$. Therefore, $\mu(A) = \int_A a d\lambda$ for all $A \in \mathcal{A}$.

But, by THEOREM 2.7 we must have $\mu(A) = \int_A \frac{d\mu}{d\lambda} d\lambda$ for all $A \in \mathcal{A}$. It therefore follows by the uniqueness of $\frac{d\mu}{d\lambda}$ that $\frac{d\mu}{d\lambda} = a$ λ-a.e. Q.E.D.

By using formula (*) of the proof of THEOREM 5.1 to express the Riesz vector a, we may restate THEOREM 5.4 in the following form:

COROLLARY 5.5. Let μ and ν be finite measures on the same measurable space $(\Omega, \mathcal{A})$ and let $\lambda = \mu + \nu$. Then

$$\frac{d\mu}{d\lambda} = \frac{b \int b d\mu}{\int b^2 d\lambda} \quad \lambda\text{-a.e.},$$

where $[b]_\lambda$ is any nonzero element of $(\ker g)^\perp \subsetneq L_2(\lambda)$ and g is the continuous linear functional on $L_2(\lambda)$ defined by $g([x]_\lambda) = \langle [x]_\lambda, [1]_\lambda \rangle_\mu$.

The last result together with THEOREM 4.4 allows us to give necessary and sufficient conditions for the equivalence of finite measures μ and ν in terms of the element b.

We know that $\mu \sim \nu$ if and only if $0 < \frac{d\mu}{d\lambda} < 1$ λ-a.e. by THEOREM 4.4. Using the representation of $\frac{d\mu}{d\lambda}$ given by COROLLARY 5.5, we have μ equivalent to ν if and only if

$$0 < \frac{b \int b d\mu}{\int b^2 d\lambda} < 1 \quad \lambda\text{-a.e.}$$

This reduces at once to $\mu \sim \nu$ if and only if $0 < b \int b d\mu < \int b^2 d\lambda$ λ-a.e. We may therefore state the following theorem.

THEOREM 5.6. Let μ and ν be finite measures on the same measurable space $(\Omega, \mathcal{A})$ and let $\lambda = \mu + \nu$. Then $\mu \sim \nu$ if and only if $0 < b \int b d\mu < \int b^2 d\lambda$ λ-a.e., where $[b]_\lambda$ is any nonzero element of $(\ker g)^\perp$ and g is the continuous linear form defined on $L_2(\lambda)$ by $g([x]_\lambda) = <[x]_\lambda, [1]_\lambda>_\mu$.

The last result of this section gives a representation for $\frac{d\mu}{d\nu}$ in the case where $\mu \sim \nu$.

THEOREM 5.7. Let μ and ν be finite measures on the same measurable space $(\Omega, \mathcal{A})$ and let $\lambda = \mu + \nu$. If $\mu \sim \nu$, then

$$\frac{d\mu}{d\nu} = \frac{b\int bd\mu}{\int b^2 d\lambda - b\int bd\mu} \quad \lambda\text{-a.e.},$$

where $[b]_\lambda$ is any nonzero element of $(\ker g)^\perp$ and g is the continuous linear form defined on $L_2(\lambda)$ by $g([x]_\lambda) = <[x]_\lambda, [1]_\lambda>_\mu$.

Proof:

If $\mu \sim \nu$, then $\mu \sim \nu \sim \lambda$. It then follows from LEMMA 2.8 that:

$$\frac{d\mu}{d\nu} = \frac{d\mu}{d\lambda}\frac{d\lambda}{d\nu} = \frac{d\mu}{d\lambda}\left(\frac{d\nu}{d\lambda}\right)^{-1} \quad \lambda\text{-a.e.}$$

By LEMMA 2.9(a) we have $\frac{d\nu}{d\lambda} = 1 - \frac{d\mu}{d\lambda}$ λ-a.e., so that $\frac{d\mu}{d\nu} = \frac{d\mu}{d\lambda}\left(1 - \frac{d\mu}{d\lambda}\right)^{-1}$ λ-a.e.

Using the representation of $\frac{d\mu}{d\lambda}$ given by COROLLARY 5.5, we have:

$$\frac{d\mu}{d\nu} = \frac{b\int b\,d\mu}{\int b^2 d\lambda} \cdot \left[1 - \frac{b\int b\,d\mu}{\int b^2 d\lambda}\right]^{-1} = \frac{b\int b\,d\mu}{\int b^2 d\lambda} \cdot \frac{\int b^2 d\lambda}{\int b^2 d\lambda - b\int b\,d\mu}$$

$$= \frac{b\int b\,d\mu}{\int b^2 d\lambda - b\int b\,d\mu} \qquad \lambda\text{-a.e.}$$

Q. E. D.

Fourier Representations

In this section we give a representation of $\frac{d\mu}{d\lambda}$ in terms of any complete orthonormal system of $L_2(\lambda)$.

DEFINITION 5.8. Let E be a Hilbert space with inner product $<\cdot,\cdot>$. A family of elements $\{\varphi_j\}_{j\in J}$ of E is called a complete orthonormal system of E if

$$<\varphi_j, \varphi_k> = \begin{cases} 0 & \text{if } j \neq k \\ 1 & \text{if } j = k \end{cases}$$

and there does not exist any other nonzero element of E which is orthogonal to all the elements of $\{\varphi_j\}_{j\in J}$.

The following theorem gives a representation of every element of E in terms of any complete orthonormal system of E. This representation of $x \in E$ is usually called the Fourier representation of x.

THEOREM 5.9. Let $\{\varphi_j\}_{j \epsilon J}$ be a complete orthonormal system of the Hilbert space E. Then for every $x \epsilon E$ we have

$$x = \sum_{j \epsilon J} <x, \varphi_j> \varphi_j .$$

Proof:

See (13, p. 34).

Using the results of THEOREM 5.4, or for that matter LEMMA 2.9(b), we note that $[\frac{d\mu}{d\lambda}]_\lambda \epsilon L_2(\lambda)$. It follows that THEOREM 5.9 may be applied to find a Fourier representation for $[\frac{d\mu}{d\lambda}]_\lambda$. This is done in the next theorem.

THEOREM 5.10. Let μ and ν be finite measures on the same measurable space $(\Omega, \mathcal{A})$ and let $\lambda = \mu + \nu$. If $\{[\varphi_j]_\lambda\}_{j \epsilon J}$ is any complete orthonormal system of $L_2(\lambda)$, then

$$\frac{d\mu}{d\lambda} = \sum_{j \epsilon J} \left(\int \varphi_j d\mu \right) \varphi_j \quad \lambda\text{-a.e.}$$

Proof:

Since $[\frac{d\mu}{d\lambda}]_\lambda \epsilon L_2(\lambda)$, it follows from THEOREM 5.9 that:

$$[\frac{d\mu}{d\lambda}]_\lambda = \sum_{j \in J} < [\frac{d\mu}{d\lambda}]_\lambda, [\varphi_j]_\lambda >_\lambda [\varphi_j]_\lambda$$

$$= \sum_{j \in J} \left(\int \varphi_j \frac{d\mu}{d\lambda} d\lambda\right) [\varphi_j]_\lambda$$

$$= \sum_{j \in J} \left(\int \varphi_j d\mu\right) [\varphi_j]_\lambda$$

Therefore, $\frac{d\mu}{d\lambda} = \sum_{j \in J} \left(\int \varphi_j d\mu\right) \varphi_j \quad \lambda\text{-a.e.}$ Q.E.D.

The Gram-Schmidt orthonormalization process (12, p. 240) is available for the generation of complete orthonormal systems of $L_2(\lambda)$, provided linearly independent subsets which span $L_2(\lambda)$ can be found. The usefulness of THEOREM 5.10 depends much, therefore, on the structure of $L_2(\lambda)$.

Analogs of THEOREM 4.4 and THEOREM 5.7 may be stated by using the Fourier representation of $\frac{d\mu}{d\lambda}$. They are simply a matter of substitution and we shall not state them here.

VI. COMPLETE ORTHONORMAL SYSTEMS

Computing Complete Orthonormal Systems

Let μ and ν be finite measures on the same measurable space $(\Omega, \mathcal{A})$ such that μ is equivalent to ν. If a complete orthonormal system is known for $L_2(\mu)$ and $\frac{d\mu}{d\nu}$ is known, then a method is given in this section for finding a complete orthonormal system for $L_2(\nu)$. This method of computing complete orthonormal systems is given by the following theorem.

THEOREM 6.1. Let μ and ν be equivalent finite measures on the same measurable space $(\Omega, \mathcal{A})$. If $\{\varphi_j\}_{j \in J}$ is a complete orthonormal system for $L_2(\mu)$, then $\{\varphi_j(\frac{d\mu}{d\nu})^{1/2}\}_{j \in J}$ is a complete orthonormal system for $L_2(\nu)$.

Proof:

Since μ is equivalent to ν, and to simplify the notation, we shall write x in place of $[x]_\mu$.

To see that $\{\varphi_j(\frac{d\mu}{d\nu})^{1/2}\}_{j \in J}$ is an orthogonal system for $L_2(\nu)$, assume $j \neq k$. Then:

$$< \varphi_j(\frac{d\mu}{d\nu})^{1/2}, \varphi_k(\frac{d\mu}{d\nu})^{1/2} >_\nu = \int \varphi_j \varphi_k \frac{d\mu}{d\nu} d\nu = \int \varphi_j \varphi_k d\mu = 0,$$

since $\{\varphi_j\}_{j \in J}$ is an orthogonal system for $L_2(\mu)$.

To see that $\{\varphi_j(\frac{d\mu}{d\nu})^{1/2}\}_{j\in J}$ is a normal system for $L_2(\nu)$, we observe:

$$<\varphi_j(\frac{d\mu}{d\nu})^{1/2}, \varphi_j(\frac{d\mu}{d\nu})^{1/2}>_\nu = \int \varphi_j^2 \frac{d\mu}{d\nu} d\nu = \int \varphi_j^2 d\mu = 1,$$

since $\{\varphi_j\}_{j\in J}$ is a normal system for $L_2(\mu)$.

Since μ is equivalent to ν, it follows from THEOREM 3.1 that $L_2(\mu)$ is isometrically isomorphic to $L_2(\nu)$ via the map $u_2 : x \mapsto x(\frac{d\mu}{d\nu})^{1/2}$. It therefore follows that $\{\varphi_j(\frac{d\mu}{d\nu})^{1/2}\}_{j\in J} \subseteq L_2(\nu)$.

It remains to show that $\{\varphi_j(\frac{d\mu}{d\nu})^{1/2}\}_{j\in J}$ is a complete system of $L_2(\nu)$. Suppose that this is not the case. Then there exists a nonzero $y \in L_2(\nu)$ which is orthogonal to all the element of $\{\varphi_j(\frac{d\mu}{d\nu})^{1/2}\}_{j\in J}$. Since u_2 is one-to-one and onto, there exists a unique nonzero $x \in L_2(\mu)$ such that $y = u_2(x) = x(\frac{d\mu}{d\nu})^{1/2}$. Then by our assumption about y we have:

$$0 = <\varphi_j(\frac{d\mu}{d\nu})^{1/2}, y>_\nu = <\varphi_j(\frac{d\mu}{d\nu})^{1/2}, x(\frac{d\mu}{d\nu})^{1/2}>_\nu = \int \varphi_j x \frac{d\mu}{d\nu} d\nu$$

$$= \int \varphi_j x d\mu = <\varphi_j, x>_\mu,$$

for all φ_j, $j\in J$. We thus have a nonzero element $x \in L_2(\mu)$ which is orthogonal to $\{\varphi_j\}_{j\in J}$. But, this contradicts the fact that $\{\varphi_j\}_{j\in J}$ is a complete orthonormal system of $L_2(\mu)$. Thus, $\{\varphi_j(\frac{d\mu}{d\nu})^{1/2}\}_{j\in J}$ is a complete orthonormal system for $L_2(\nu)$. Q.E.D.

Examples

This section presents several applications of THEOREM 6.1. These applications make use of the classical complete orthonormal systems known as the trigonometric functions, the Legendre functions, the Hermite functions, and the Laguerre functions. These systems have many applications in applied mathematics and mathematical physics. They have been studied extensively, and for further information the reader is referred to the works by N. N. Lebedev (15), G. Sansone (20), and Gabor Szegö (22).

In all the examples which follow, we shall be in the situation discussed in Chapter IV under Special Cases. The finite measures μ and ν will be defined on a measurable space $(\Omega, \mathcal{A})$ so that they have the form $d\mu(\omega) = f(\omega)d\omega$ and $d\nu(\omega) = g(\omega)d\omega$. Recall that in this context, COROLLARY 4.6 showed that $\mu \sim \nu$ if and only if $0 < \frac{f(\omega)}{f(\omega)+g(\omega)} < 1$ wherever $\frac{f(\omega)}{f(\omega)+g(\omega)}$ is defined. Also recall that $\mu \sim \nu$ implies that $\frac{d\mu}{d\nu}(\omega) = \frac{f(\omega)}{g(\omega)}$ wherever the fraction is defined.

Since we will be dealing with equivalent measures in our examples, the functions are to be understood as being defined only almost everywhere and we write φ in place of $[\varphi]$.

In our first example, we make use of the trigonometric functions.

EXAMPLE 6.2. Let $\Omega = [0, 2\pi]$ and let $\mathcal{A}$ be the Lebesgue measurable sets of $[0, 2\pi]$. Let μ be Lebesgue measure on

$([0, 2\pi], \mathcal{A})$, so that $d\mu(x) = dx$. The trigonometric functions:

$$\{\frac{1}{\sqrt{2\pi}}, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin nx}{\sqrt{\pi}} : 0 \leq x \leq 2\pi, \; n = 1, 2, 3, \ldots\}$$

are a complete orthonormal system (20, p. 43-46) for $L_2([0, 2\pi], \mathcal{A}, \mu)$.

Let ν be another finite measure on $([0, 2\pi], \mathcal{A})$ such that $d\nu(x) = f(x)dx$ and assume that ν is equivalent to μ. Then we have:

$$\frac{d\mu}{d\nu}(x) = \frac{1}{f(x)}, \quad \text{so that} \quad (\frac{d\mu}{d\nu}(x))^{1/2} = \frac{1}{\sqrt{f(x)}}.$$

It then follows at once from THEOREM 6.1 that a complete orthonormal system for $L_2([0, 2\pi], \mathcal{A}, \nu)$ will be given by the set:

$$\{\frac{1}{\sqrt{2\pi f(x)}}, \frac{\cos nx}{\sqrt{\pi f(x)}}, \frac{\sin nx}{\sqrt{\pi f(x)}} : 0 \leq x \leq 2\pi, \; n = 1, 2\; 3, \ldots\}.$$

Our next example makes use of the Legendre functions.

EXAMPLE 6.3. Let $\Omega = [-1, 1]$ and let $\mathcal{A}$ represent the Lebesgue measurable sets of $[-1, 1]$. Let μ be Lebesgue measure on $([-1, 1], \mathcal{A})$ so that $d\mu(x) = dx$. If $P_n(x)$ represent the Legendre polynomials (15, p. 44) on $[-1, 1]$, then the Legendre functions are defined on $[-1, 1]$ by:

$$\varphi_n(x) = \sqrt{\frac{2n+1}{2}}\, P_n(x), \quad n = 0, 1, 2, \ldots .$$

The set of Legendre functions $\{\varphi_n(x) : n = 0, 1, 2, \ldots\}$ is a complete orthonormal system (20, p. 189-193) for the space $L_2([-1, 1], \mathcal{A}, \mu)$.

If ν is any other finite measure defined on $([-1, 1], \mathcal{A})$ which is equivalent to μ and has the form $d\nu(x) = f(x)dx$, so that $\frac{d\mu}{d\nu}(x) = \frac{1}{f(x)}$, then it follows from THEOREM 6.1 that:

$$\left\{\frac{\varphi_n(x)}{\sqrt{f(x)}} : n = 0, 1, 2, \ldots\right\} = \left\{\sqrt{\frac{2n+1}{2f(x)}}P_n(x) : n = 0, 1, 2, \ldots\right\}$$

is a complete orthonormal system for $L_2([-1, 1], \mathcal{A}, \nu)$.

The next example makes use of the Hermite functions.

EXAMPLE 6.4. Let $\Omega = R = (-\infty, +\infty)$ and let $\mathcal{A}$ be the Lebesgue measurable sets of R. Let μ be the measure on $(R, \mathcal{A})$ defined by $d\mu(x) = e^{-x^2}dx$. Since $\int_{-\infty}^{+\infty} e^{-x^2}dx = \sqrt{\pi}$, μ is a finite measure on $(R, \mathcal{A})$.

If $H_n(x)$ represent the Hermite polynomials (20, p. 306) on R, then the Hermite functions are defined on R by:

$$\varphi_n(x) = \frac{H_n(x)}{\sqrt{2^n n! \sqrt{\pi}}}, \quad n = 0, 1, 2, \ldots .$$

The set of Hermite functions $\{\varphi_n(x) : n = 0, 1, 2, \ldots\}$ is a complete orthonormal system (20, p. 351-355) for $L_2(R, \mathcal{A}, \mu)$.

Now, let ν be a finite measure on $(R, \mathcal{A})$ which has the form $d\nu(x) = f(x)dx$ and is equivalent to μ. Then:

$$\frac{d\mu}{d\nu}(x) = \frac{e^{-x^2}}{f(x)}, \quad \text{so that} \quad \left(\frac{d\mu}{d\nu}(x)\right)^{1/2} = \frac{1}{\sqrt{e^{x^2} f(x)}} .$$

It then follows at once from THEOREM 6.1 that a complete orthonormal system for $L_2(R, \mathcal{A}, \nu)$ is given by:

$$\left\{ \frac{\varphi_n(x)}{\sqrt{e^{x^2} f(x)}} : n = 0, 1, 2, \ldots \right\} .$$

Our final example makes use of the Laguerre functions.

EXAMPLE 6.5. Let $\Omega = [0, +\infty)$ and let $\mathcal{A}$ be the Lebesgue measurable sets of $[0, +\infty)$. Define the measures μ_α on $[0, +\infty)$ by $d\mu_\alpha(x) = e^{-x} x^\alpha dx$ for $\alpha > -1$. Since $\int_0^{+\infty} e^{-x} x^\alpha dx = \Gamma(\alpha+1)$, Γ being the Gamma function (20, p. 392), μ_α is a finite measure on $([0, +\infty), \mathcal{A})$ for each $\alpha > -1$.

If $L_n^{(\alpha)}(x)$ represent the Laguerre polynomials (20, p. 295) on $[0, +\infty)$, then the Laguerre functions are defined on $[0, +\infty)$ by:

$$\varphi_n^{(\alpha)}(x) = \sqrt{\frac{\Gamma(n+1)}{\Gamma(n+\alpha+1)}} \cdot L_n^{(\alpha)}(x), \quad n = 0, 1, 2, \ldots .$$

The collection $\{\varphi_n^{(\alpha)}(x) : n = 0, 1, 2, \ldots\}$ is a complete orthonormal system (20, p. 349-351) for $L_2([0, +\infty), \mathcal{A}, \mu_\alpha)$.

Let ν be a finite measure on $([0, +\infty), \mathcal{A})$ which has the form $d\nu(x) = f(x)dx$. Suppose ν is equivalent to μ. Then we have:

$$\frac{d\mu}{d\nu}(x) = \frac{e^{-x}x^{a}}{f(x)}, \quad \text{so that} \quad \left(\frac{d\mu}{d\nu}(x)\right)^{1/2} = \sqrt{\frac{x^{a}}{e^{x}f(x)}}$$

It then follows from THEOREM 6.1 that a complete orthonormal system for $L_2([0, +\infty), \mathcal{A}, \nu)$ is given by the set:

$$\left\{ \sqrt{\frac{x^{a}}{e^{x}f(x)}} \cdot \varphi_{n}^{(a)}(x) : n = 0, 1, 2, \ldots \right\}.$$

VII. GAUSSIAN PROCESSES

Background

In this chapter we give necessary and sufficient conditions for the equivalence of Gaussian processes, with countable index sets, in terms of the finite dimensional distributions of the processes. The proof of this result requires the application of Kakutani's theorem together with the results of Chapter IV.

We first introduce the necessary terminology of Gaussian processes. As for notation, R will denote the real numbers, $\mathcal{B}$ will denote the Borel sets of R, and P will be used to represent probability measures when there is no danger of confusion.

DEFINITION 7.1. Let $(\Omega, \mathcal{A}, \mu)$ be any probability space. If x is a measurable function from $(\Omega, \mathcal{A}, \mu)$ into $(R, \mathcal{B})$, then x is called a real random variable or simply a random variable.

A random variable x is called a Gaussian random variable if one of the following holds:

(a) x has a distribution with a density function of the form:

$$N(\sigma, m, \omega) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left[-\frac{1}{2}\left(\frac{\omega - m}{\sigma}\right)^2\right],$$

where $\sigma > 0$. That is, $P(x \leq r) = \int_{-\infty}^{r} N(\sigma, m, \omega)\, d\omega$. (This case will be called the nonsingular case.)

(b) x has a point mass at some point m. (This case will be called the singular case.)

DEFINITION 7.2. Let T be an arbitrary set. The collection $\{x_t : t \in T\}$ is called a stochastic process on the probability space $(\Omega, \mathcal{A}, \mu)$ provided each x_t is a random variable on $(\Omega, \mathcal{A}, \mu)$. If each finite linear combination $\Sigma\, a_t x_t$ is a Gaussian random variable, then $\{x_t : t \in T\}$ is called a Gaussian process.

If T is a finite set and $\{x_t : t \in T\}$ is a stochastic process, then there is a natural way (5, p. 32-33) to define a σ-algebra and measure on R^T. This leads to the following definition.

DEFINITION 7.3. If T is a finite set and $\{x_t : t \in T\}$ is a stochastic process, then its joint distribution is the measure on R^T which is determined by assigning $P\{x_t \in A_t : t \in T\}$ to $\prod_{t \in T} A_t$ where $A_t = (a_t, b_t)$, a_t and b_t being extended real numbers.

In particular, if $\{x_t : t \in T\}$ is a finite Gaussian process and each x_t is nonsingular and $A_t = (-\infty, a_t)$, then:

$$P\{x_t \in A_t : t \in T\} = P\{x_t \leq a_t : t \in T\}$$

$$= \int_{-\infty}^{a_1} \cdots \int_{-\infty}^{a_n} N_n(\omega_1, \ldots, \omega_n) d\omega_1 \ldots d\omega_n,$$

where N_n represents the n-dimensional Gaussian density function (5, p. 26).

DEFINITION 7.4. If T is an arbitrary set and $\{x_t : t \in T\}$ is a stochastic process, then for each finite subset $T_o \subset T$, we may obtain a joint distribution μ_{T_o} on R^{T_o} by considering $\{x_t : t \in T_o\}$ and using DEFINITION 7.3 to define $\mu_{T_o}(\prod_{t \in T_o} A_t) = P\{x_t \in A_t : t \in T_o\}$. The set $\{\mu_{T_o} : T_o \subset T, T_o \text{ finite}\}$ is called the family of finite joint distributions of $\{x_t : t \in T\}$ or the family of finite dimensional distributions of $\{x_t : t \in T\}$.

If two stochastic processes X and Y on the same index set have the same family of finite joint distributions, then X and Y are called isomorphic.

As a matter of notation, we shall let $\mathcal{B}_{T_o}$ be the σ-algebra of R^{T_o} on which the measure μ_{T_o} is defined. It should be noted that if $T_o \subset T_1$ are finite subsets of T, then there exists an embedding of $\mathcal{B}_{T_o}$ into $\mathcal{B}_{T_1}$ defined by the map $\phi : \prod_{t \in T_o} A_t \mapsto \prod_{t \in T_1} B_t$ where $B_t = A_t$ for $t \in T_o$ and $B_t = R$ for $t \in T_1 - T_o$. We thus may write $\mathcal{B}_{T_o} \subset \mathcal{B}_{T_1}$.

DEFINITION 7.5. If $T_o \subset T_1$ are finite subsets of T, then the finite joint distributions of the stochastic process $\{x_t : t \in T\}$ are said to satisfy the consistency condition if the restriction of μ_{T_1} to the σ-algebra $\mathcal{B}_{T_o}$ is equal to μ_{T_o}.

LEMMA 7.6. The finite joint distributions of a Gaussian process satisfy the consistency condition.

Proof:

Let $T_o \subset T_1$ be finite subsets of T such that $T_1 = \{t_1, \ldots, t_n\}$. Without loss of generality we assume $T_o = \{t_1, \ldots, t_m\}$, where $m < n$. Then we need only show that

$$\mu_{T_o}\left(\prod_{t \in T_o} A_t\right) = \mu_{T_1}\left(\prod_{t \in T_o} A_t\right)$$

where $A_t = (a_t, b_t)$. This follows since:

$$\mu_{T_o}\left(\prod_{t \in T_o} A_t\right) = \int_{a_{t_1}}^{b_{t_1}} \int_{a_{t_2}}^{b_{t_2}} \cdots \int_{a_{t_m}}^{b_{t_m}} N_m(\omega_1, \ldots, \omega_m) d\omega_1 \ldots d\omega_m$$

$$= \int_{a_{t_1}}^{b_{t_1}} \int_{a_{t_2}}^{b_{t_2}} \cdots \int_{a_{t_m}}^{b_{t_m}} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} N_n(\omega_1, \ldots, \omega_n) d\omega_1 \ldots d\omega_n$$

$$= \mu_{T_1}\left(\prod_{t \in T_o} A_t\right).$$

Q.E.D.

We are now in a position to state the Kolmogorov theorem.

THEOREM 7.7. Suppose that for each finite subset $T_o \subset T$ we are given the finite joint distribution μ_{T_o} of the stochastic process $\{x_t : t \in T\}$ and suppose that these finite joint distributions satisfy the consistency condition. Then there exists a measure μ on the σ-algebra generated by the cylinder sets of R^T making the coordinate functions into a stochastic process with the preassigned measures μ_{T_0} as its set of finite joint distributions.

Proof:

See (16, p. 79-83) or (5, p. 34-37).

Some general remarks about the proof of the Kolmogorov theorem are in order. The probability space we obtain is $(R^T, \mathcal{C}, \mu)$, where $\mathcal{C}$ is the infinite product σ-algebra generated by the cylinder sets (16, p. 80), and the measure μ is the infinite product measure generated by the finite joint distributions. The process defined on $(R^T, \mathcal{C}, \mu)$ is called the path space version of the process and is isomorphic to the original process.

DEFINITION 7.8. Let X and Y be Gaussian processes with the same index set, and let μ and ν be the measures generated on $(R^T, \mathcal{C})$ respectively by X and Y via the Kolmogorov theorem. We shall say X is equivalent to Y if and only if μ is equivalent to ν.

The final background result we need is Kakutani's theorem and we now state it:

THEOREM 7.9. Let J be a countable set. Let $\mathcal{A}_j$ be a σ-algebra of Ω_j and let $\mathcal{A}$ be the product σ-algebra generated by the $\mathcal{A}_j$ on the product $\Omega = \prod_{j \in J} \Omega_j$. Let μ_j and ν_j be probability measures on $(\Omega_j, \mathcal{A}_j)$ with $\mu_j \sim \nu_j$. Let $\mu = \prod_{j \in J} \mu_j$ and

$\nu = \prod_{j \in J} \nu_j$ be the product measures on $(\Omega, \mathcal{A})$. Then either $\mu \sim \nu$ or $\mu \perp \nu$. A necessary and sufficient condition for $\mu \sim \nu$ is that:

$$\prod_{j \in J} \int \left(\frac{d\mu_j}{d\nu_j}\right)^{1/2} d\nu > 0$$

Proof:

See (14) or (12, p. 453-455).

Conditions for Equivalence

Throughout this section we shall let $X = \{x_t : t \in T\}$ and $Y = \{y_t : t \in T\}$ be Gaussian processes with countable index set T. Let $\{\mu_{T_o} : T_o \subset T,\ T_o \text{ finite}\}$ be the finite joint distributions of X and let $\{\nu_{T_o} : T_o \subset T,\ T_o \text{ finite}\}$ be those of Y. We shall write μ_{t_o} in place of $\mu_{\{t_o\}}$ and ν_{t_o} in place of $\nu_{\{t_o\}}$ for $\{t_o\} \subset T$. It then follows that the measures μ and ν on $(R^T, \mathcal{C})$ generated by X and Y respectively via the Kolmogorov theorem are the infinite product measures $\mu = \prod_{t \in T} \mu_t$ and $\nu = \prod_{t \in T} \nu_t$.

Our result requires the following lemma.

LEMMA 7.10. If $\mu \sim \nu$, then $\mu_{T_o} \sim \nu_{T_o}$ for all finite $T_o \subset T$.

Proof:

We prove the contrapositive. Suppose there exists a finite $T_o \subset T$ such that μ_{T_o} is not equivalent to ν_{T_o}. Since μ and

ν are Gaussian measures, it follows that μ_{T_o} and ν_{T_o} are Gaussian, so that $\mu_{T_o} \perp \nu_{T_o}$ by THEOREM 1.3. Thus, there exists a set $A \in \mathcal{B}_{T_o}$, the σ-algebra of R^{T_o} on which μ_{T_o} and ν_{T_o} are defined, so that $\mu_{T_o}(A) = 0 = \nu_{T_o}(A^c)$. But, $\mathcal{B}_{T_o}$ is embedded in the σ-algebra $\mathcal{C}$ of R^T via a map ϕ, where $\phi(A)$ has restrictions only on the coordinates $t \in T_o$ and is unrestricted in the coordinates $t \in T - T_o$. Then $\mu(\phi(A)) = \mu_{T_o}(A) = 0$ and $\nu(\phi(A^c)) = \nu_{T_o}(A^c) = 0$. Thus, $\mu \perp \nu$, and the lemma follows.

Q.E.D.

With the aid of this lemma, THEOREM 7.9, and COROLLARY 4.6, we can now prove the following theorem.

THEOREM 7.11. Let X and Y be Gaussian processes with the same countable index set T. A necessary and sufficient condition for the equivalence of X and Y is that there exist disjoint sets U and V such that $T = U \cup V$ and the following hold:

(a) for each $t \in U$, x_t and y_t are nonsingular Gaussian random variables,

(b) for each $t \in V$, x_t and y_t are singular Gaussian random variables with point mass at the same point m_t,

(c) $\prod_{t \in U} \int \left(\frac{d\mu_t}{d\nu_t}\right)^{1/2} d\nu > 0.$

Proof:

Assume $X \sim Y$ so that $\mu \sim \nu$. It then follows by LEMMA 7.10 that $\mu_t \sim \nu_t$ for all $t \in T$. Then for each $t \in T$, we consider four cases: (i) μ_t and ν_t both have nonsingular distributions, (ii) μ_t is nonsingular while ν_t is singular, (iii) μ_t is singular while ν_t is nonsingular, and (iv) both μ_t and ν_t are singular.

In case (i), μ_t must have a density function of the form $N(\sigma_1, m_1)$, with $\sigma_1 > 0$, and ν_t will have $N(\sigma_2, m_2)$, with $\sigma_2 > 0$, for its density function. It then follows at once from EXAMPLE 4.8 that $\mu_t \sim \nu_t$.

In case (ii), μ_t has a density function of the form $N(\sigma, m)$, with $\sigma > 0$, while ν_t has point mass at some point m_t. Let $A = (a, b)$ such that $m_t \notin A$. Then $\mu_t(A) = \int_a^b N(\sigma, m, \omega)d\omega > 0$ and $\nu_t(A) = 0$. Thus, μ_t is not equivalent to ν_t and case (ii) must be ruled out.

In a similar manner, case (iii) must also be ruled out.

In case (iv), let μ_t have point mass at m_t and let ν_t have point mass at n_t. If $m_t \neq n_t$, we let $A = (a, +\infty)$ so that $n_t \in A$ while $m_t \notin A$. Then $\mu_t(A) = 0 = \nu_t(A^c)$, contradicting $\mu_t \sim \nu_t$. If, however, $m_t = n_t$, then μ_t and ν_t have exactly the same distribution function and we must have $\mu_t = \nu_t$ and hence $\mu_t \sim \nu_t$.

Consequently, if $\mu \sim \nu$ we must have conditions (a) and (b).

Condition (c) is also necessary by THEOREM 7.9. The product need be taken only over U since $\frac{d\mu_t}{d\nu_t} = 1$ for $t \in V$.

To prove the converse, assume (a), (b), and (c) hold.

In case (a), μ_t has a density of the form $N(\sigma_1, m_1)$ with $\sigma_1 > 0$ and ν_t has a density of the form $N(\sigma_2, m_2)$ with $\sigma_2 > 0$. If $\lambda_t = \mu_t + \nu_t$, then

$$\frac{d\mu_t}{d\lambda_t} = \frac{N(\sigma_1, m_1)}{N(\sigma_1, m_1)+N(\sigma_2, m_2)},$$

so that $0 < \frac{d\mu_t}{d\lambda_t} < 1$ every where. It then follows from COROLLARY 4.6 that $\mu_t \sim \nu_t$ for all $t \in U$.

In case (b), $\mu_t = \nu_t$ for all $t \in V$.

It therefore follows that $\mu_t \sim \nu_t$ for all $t \in T$. Since (c) also holds, it follows at once from THEOREM 7.9 that $\mu \sim \nu$. Therefore, $X \sim Y$. Q.E.D.

If T is finite, condition (c) of the last theorem may be omitted. To see this we need only show that conditions (a) and (b) imply (c) in this case.

If $t \in U$, it follows from EXAMPLE 4.8 that $\frac{d\mu_t}{d\nu_t} = \frac{N(\sigma_1, m_1)}{N(\sigma_2, m_2)}$, and since the exponential function is always positive, we have $\frac{d\mu_t}{d\nu_t} > 0$ for all $t \in U$.

If T is finite, then of course U is finite so $\prod_{t \in U} \int (\frac{d\mu_t}{d\nu_t})^{1/2} d\nu > 0$ is always true. Thus, (a) and (b) do imply (c) when T is finite and we have proved the following corollary.

COROLLARY 7.12. Let X and Y be Gaussian processes defined on the same finite index set T. A necessary and sufficient condition for equivalence of X and Y is the existence of disjoint sets U and V such that $T = U \cup V$ and the following hold:

(a) for each $t \in U$, x_t and y_t are nonsingular Gaussian random variables,

(b) for each $t \in V$, x_t and y_t are singular Gaussian random variables with point mass at the same point m_t.

BIBLIOGRAPHY

1. Aronszajn, N. Theory of reproducing kernels. Transactions of the American Mathematical Society 68:337-404. 1950.

2. Baxter, G. A strong limit theorem for Gaussian processes. Proceedings of the American Mathematical Society 7:522-527. 1956.

3. Cameron, R. H. and W. T. Martin. The behavior of measure and measurability under change of scale in Wiener space. Bulletin of the American Mathematical Society 53:130-137. 1947.

4. Chow, T. R. Equivalence of Gaussian stationary processes. Annals of Mathematical Statistics 40:197-202. 1969.

5. Cramér, Harald and M. R. Leadbetter. Stationary and related stochastic processes. New York, Wiley, 1967. 348 p.

6. Doob, J. L. Stochastic processes. New York, Wiley, 1953. 654 p.

7. Feldman, Jacob. Equivalence and perpendicularity of Gaussian processes. Pacific Journal of Mathematics 8:699-708. 1958.

8. Feldman, Jacob. Professor, University of California, Department of Mathematics. Lecture notes from Statistics 273. Equivalence of Gaussian processes. Berkeley, 1966.

9. Feldman, Jacob. Some classes of equivalent Gaussian processes on an interval. Pacific Journal of Mathematics 10:1211-1220. 1960.

10. Grenander, U. Stochastic processes and statistical inference. Arkiv für Mathematik 1:195-277. 1950.

11. Hájek, J. On a property of normal distribution of any stochastic process. Czechoslovak Mathematics Journal 8:610-618. 1958. (Translated in Institute of Mathematical Statistics. Selected Translations in Mathematical Statistics and Probability 1:245-252. 1961)

12. Hewitt, Edwin and Karl Stromberg. Real and abstract analysis. New York, Springer-Verlag New York, 1965. 476 p.

13. Horváth, John. Topological vector spaces and distributions. Vol. 1. Reading, Massachusetts, Addison-Wesley, 1966. 449 p.

14. Kakutani, S. On equivalence of infinite product measures. Annals of Mathematics 49:214-224. 1948.

15. Lebedev, N. N. Special functions and their applications. Englewood Cliffs, New Jersey, Prentice-Hall, 1965. 308 p.

16. Neveu, Jacques. Mathematical foundations of the calculus of probability. San Francisco, Holden-Day, 1965. 223 p.

17. Parzen, Emanuel. Probability density functionals and reproducing kernel Hilbert spaces. In: Symposium on time series analysis, Brown University, 1962, ed. by Murray Rosenblatt. New York, Wiley, 1963. p. 155-169.

18. Parzen, Emanuel. Time series analysis papers. San Francisco, Holden-Day, 1967. 565 p.

19. Rozanov, Yu. A. On probability measures in functional spaces corresponding to stationary Gaussian processes. Theory of Probability and its Applications 9:404-420. 1964.

20. Sansone, G. Orthogonal functions. New York, Interscience Publishers, 1959. 411 p.

21. Shepp, L. A. Radon-Nikodym derivatives of Gaussian measures. Annals of Mathematical Statistics 37:321-354. 1966.

22. Szegö, Gabor. Orthogonal polynomials. Rev. ed. New York, American Mathematical Society, 1959. 421 p. (Colloquium Publications. Vol. 23)