CRRAO Advanced Institute of Mathematics, Statistics and Computer Science (AIMSCS)

Research Report

Author (s): **B. L. S. Prakasa Rao**

Title of the Report: **Characterization of Gaussian Distribution on a Hilbert Space from Samples of Random Size**

Research Report No.: **RR2014-16**

Date: **June 11, 2014**

Prof. C R Rao Road, University of Hyderabad Campus,
Gachibowli, Hyderabad-500046, INDIA.
www.crraoaimscs.org

Characterization of Gaussian Distribution on a Hilbert Space from Samples of Random Size

B. L. S. PRAKASA RAO [1]

CR Rao Advanced Institute of Mathematics, Statistics and Computer Science, Hyderabad 500046, India

Abstract: We obtain two characterizations of the Gaussian distribution on a Hilbert space from samples of random size.

Key words : Characterization; Gaussian distribution; Samples of random size; Hilbert space.

1 Introduction

Several characterizations of the univariate and the multivariate normal distribution are known (cf. Kagan et al. [4], Prakasa Rao [9]). Most of these results involve statistics based on fixed sample sizes. However there are situations, such as in study of population growth using branching processes, the size of a generation depends on the size of the previous generation which itself is random. For the breeding habit and study growth of an organism in one generation, one needs to study distributions of statistics based on population sizes of the previous generation which in turn are random. In such cases, it is necessary to characterize the underlying distribution based on samples of random size. Cook [1] obtained a characterization of correlated normal random vectors. Kagan and Shalaevski [4]) obtained characterization of normal distribution by a property of the non-central chi-square distribution. Kotlarski and Cook [5] extended the results in Cook [1] and Kagan and Shalaevski [4] and obtained two characterizations of the multivariate normal distribution based on samples of random size. Prakasa Rao [8] obtained similar results in an unpublished report. In view of the recent development of methods of functional data analysis, it would be of interest to investigate whether the results on characterizations of Gaussian distribution obtained in the case of Euclidean spaces R and R^k continue to hold when the observation space is a function space such as a Hilbert space. Our aim is to characterize the Gaussian distribution on a real separable Hilbert space H from samples of random size. Example of such a space H is the

[1] E-mail: blsprao@gmail.com

space of square integrable functions f on the real line associated with the norm

$$||f|| = [\int_R |f(x)|^2 dx]^{1/2}.$$

2 Preliminaries

Let $(\Omega, \mathcal{F}, \mu)$ be a probability space and H be a real separable Hilbert space. Let $\mathcal{B}$ be the Borel-σ-algebra generated by the norm topology on the space H. A mapping $X : \Omega \to H$ is said to be a *random element* taking values in a Hilbert space H if $X^{-1}(B) \in \mathcal{F}$ for every $B \in \mathcal{B}$. Define

$$\mu_X(B) = \mu(X^{-1}(B)), B \in \mathcal{B}.$$

It is easy to check that μ_X is a probability measure on the measurable space $(H, \mathcal{B})$. Let $\mathcal{M}(H)$ denote the class of all probability measures on $(H, \mathcal{B})$. Let $< x, y >$ denote the inner product and $||x||$ the norm defined on the Hilbert space H. Let $\nu \in \mathcal{M}(H)$ be such that

$$\int_H ||x||^2 \nu(dx) < \infty.$$

Then the *covariance operator* S of ν is the Hermitian operator determined uniquely by the quadratic form

$$< Sy, y >= \int_H < x, y >^2 \nu(dx).$$

A positive definite Hermitian operator S on the Hilbert space H is called an S-operator if it has finite trace, that is, for some orthonormal basis $\{e_i, i \geq 1\}$, of the Hilbert space H, the sum $\sum_{i=1}^{\infty} < Se_i, e_i >< \infty$. In such a case, the inequality holds for every orthonormal basis of the Hilbert space H.

Suppose ν is a probability measure in $\mathcal{M}(H)$ such that

$$\int_H ||x|| \nu(dx) < \infty.$$

It can be shown that there exists an element x_0 in H such that

$$< x_0, y >= \int_H < x, y > \nu(dx), y \in H.$$

The element x_0 is called the mean of the probability measure ν or the expectation of the random element X if the distribution of the random element X is ν. We denote the mean or expectation x_0 by the notation

$$\int_H x \, \nu(dx).$$

For any probability measure ν on the measurable space $(H, \mathcal{B})$, the *characteristic functional* $\hat{\nu}(.)$ is a functional defined on H by the relation

$$\hat{\nu}(y) = \int_H e^{i<x,y>} \nu(dx), y \in H.$$

The characteristic functional $\phi_X(.)$ of the random element X is given by

$$\begin{aligned} \phi_X(y) &= \int_H e^{i<x,y>} \mu_X(dx), y \in H \\ &= \int_\Omega e^{i<X(\omega),y>} \mu(d\omega), y \in H. \end{aligned}$$

It is known that there is a one-to-one correspondence between the characteristic functionals and the probability measures on H. Furthermore the characteristic functional ϕ_X of a random element X satisfies the conditions

$$\phi_X(0) = 1; |\phi_X(y)| \leq 1, \phi_X(y) = \overline{\phi_X(-y)}, y \in H$$

where 0 denotes the identity element in H. Moreover the function $\phi_X(.)$ is continuous in the norm topology . In addition, if X and Y are independent random elements taking values in h, then $X + Y$ is also a random element taking values in H, and

$$\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t), t \in H.$$

For proofs of these results, see Parthasarathy [7] or Grenander [2].

A probability measure μ_X generated by a random element X on a Hilbert space H is said to be *Gaussian* if its characteristic functional $\phi_X(y)$ is of the form

$$\phi_X(y) = \exp\{i < x_0, y > -\frac{1}{2} < Sy, y >\}$$

where x_0 is a fixed element in H and S is an S-operator on H. It can be shown that x_0 is the mean and the operator S is the covariance operator for Gaussian measure with characteristic functional $\phi_X(y), y \in H$ (cf. Grenander [2], Theorem 6.3.1.). The following result is due to Grenander [2], p. 141.

Theorem 2.1: (i) Suppose X and Y are two independent random elements taking values in a Hilbert space H with Gaussian measures with the means m_X and m_Y and covariance operators S_X and S_Y respectively. Then $X + Y$ is a H-valued random element with Gaussian

measure with the mean $m_X + m_Y$ and the covariance operator $S_X + S_Y$. Conversely, if $Z = X + Y$ is a sum of independent random elements taking values in H with Gaussian measure, then the random elements X and Y must have Gaussian measures.
(ii) If X is a random element taking values in a Hilbert space H with Gaussian measure μ_X, Then X can be represented as

$$X = m + \sum_{i=1}^{\infty} \psi_i e_i$$

where $\{e_i, i \geq 1\}$ is an orthonormal basis on H and $\{\psi_i, i \geq 1\}$ are independent mean zero Gaussian random variables with $Var(\psi_i) = \sigma_i^2, i \geq 1$ and $\{\sigma_i^2, i \geq 1\}$ are the eigenvalues of the operator S. Furthermore the infinite series is convergent (strongly) with probability one.
(iii)If B is a bounded linear operator from H to H and X is a random element with Gaussian measure with mean m and covariance operator S, then the random element $Y = BX$ has a Gaussian measure with mean Bm and covariance operator $S = BSB^*$.

Let $\mathbf{X}_i, 1 \leq i \leq N$ and $\mathbf{Y}_j, 1 \leq j \leq N$ be two independent samples of independent identically distributed Hilbert space valued random elements with X_i distributed with probability measure μ_X and Y_j distributed with with probability measure μ_Y and N be a discrete integer valued random variable independent of $\mathbf{X}_i, 1 \leq i \leq N$ and $\mathbf{Y}_j, 1 \leq j \leq N$. Let

$$W = \sum_{j=1}^{N} [< S_1(\mathbf{X}_j - \mathbf{a}), (\mathbf{X}_j - \mathbf{a}) > + < \mathbf{S_2}(\mathbf{Y_j} - \mathbf{b}), (\mathbf{Y_j} - \mathbf{b}) >]$$

where S_1, S_2 are known positive definite Hermitian operators with finite traces and $\mathbf{a}$ and $\mathbf{b}$ are elements in H. Suppose that $E[e^{-\frac{1}{2}W}] = J(\mathbf{a}, \mathbf{b}) < \infty$. We prove that the function $J(\mathbf{a}, \mathbf{b})$ is a measurable function of the function

$$< S_1\mathbf{a}, \mathbf{a} > + < S_2\mathbf{b}, \mathbf{b} >$$

if and only if the probability measures μ_X and μ_Y are Gaussian with mean zero vector. This result generalizes a result characterizing the multivariate normal distribution by Kotlarski and Wood [6] (cf. Prakasa Rao [8]). Characterization problems of similar nature for identifiability in stochastic models are discussed in Prakasa Rao [9].

3 Characterizations

We now state and prove the main results.

Theorem 3.1: Suppose that the function $J(\mathbf{a},\mathbf{b})) = E[e^{-\frac{1}{2}W}] < \infty$ for all elements $\mathbf{a}$ and $\mathbf{b}$ in H. Then the function $J(\mathbf{a},\mathbf{b}))$ is a measurable function of the function $< S_1\mathbf{a},\mathbf{a} > + < S_2\mathbf{b},\mathbf{b} >$ for $\mathbf{a},\mathbf{b} \in H$ if and only if the distributions μ_X and μ_Y are Gaussian with mean zero vector.

Proof : It is clear that

$$\begin{aligned} E[e^{-\frac{1}{2}W}] &= \sum_{n=1}^{\infty} E[e^{-\frac{1}{2}W}|N=n]P(N=n) \\ &= \sum_{n=1}^{\infty}(E[\exp(-\frac{1}{2} < S_1(\mathbf{a}-\mathbf{X}_j),(\mathbf{a}-\mathbf{X}_j) >]E[\exp(-\frac{1}{2} < S_2(\mathbf{b}-\mathbf{Y}_j),(\mathbf{b}-\mathbf{Y}_j) >])^n P(N=n). \end{aligned}$$

The last equality follows from the assumption that $\mathbf{X}_i, 1 \le i \le N$ and $\mathbf{Y}_j, 1 \le j \le N$ are two independent samples of independent identically distributed k-dimensional random elements independent of the random variable N. Let

$$\alpha(\mathbf{a}) = E[\exp(-\frac{1}{2} < S_1(\mathbf{a}-\mathbf{X}_j),(\mathbf{a}-\mathbf{X}_j) >]$$

and

$$\beta(\mathbf{b}) = E[\exp(-\frac{1}{2} < S_2(\mathbf{b}-\mathbf{Y}_j),(\mathbf{b}-\mathbf{Y}_j) >].$$

Then, it follows that,

$$E[e^{-\frac{1}{2}W}] = Q(\alpha(\mathbf{a})\beta(\mathbf{b}))$$

where

$$Q(x) = \sum_{n=1}^{\infty} x^n P(N=n), 0 \le x \le 1.$$

Note that the function $Q(.)$ is a strictly increasing continuous function on the interval $[0,1]$. Hence its inverse is well defined. Suppose that the function $E[e^{-\frac{1}{2}W}]$ is a measurable function of the function $< S_1\mathbf{a},\mathbf{a} > + < S_2\mathbf{b},\mathbf{b} >$. Then there exists a measurable real-valued function $\psi(.)$ such that

(3. 1) $$\psi(< S_1\mathbf{a},\mathbf{a} > + < S_2\mathbf{b},\mathbf{b} >) = Q(\alpha(\mathbf{a})\beta(\mathbf{b}))$$

or equivalently

(3. 2) $$\alpha(\mathbf{a})\beta(\mathbf{b}) = \gamma(< S_1\mathbf{a},\mathbf{a} > + < S_2\mathbf{b},\mathbf{b} >)$$

where $\gamma = Q^{-1}\ o\ \psi$ for all $\mathbf{a},\mathbf{b} \in R^k$. It is easy to see that $\alpha(\mathbf{0}) \neq 0$ and $\beta(\mathbf{0}) \neq 0$ for $\mathbf{a} = \mathbf{0}$ and $\mathbf{b} = \mathbf{0}$. Substituting $\mathbf{a} = \mathbf{0}$ and $\mathbf{b} = \mathbf{0}$ alternately, we obtain that

(3. 3) $$\gamma(< S_1\mathbf{a},\mathbf{a} >)\gamma(< S_2\mathbf{b},\mathbf{b} >) = \alpha(\mathbf{0})\beta(\mathbf{0})\gamma(< S_1\mathbf{a},\mathbf{a} > + < S_2\mathbf{b},\mathbf{b} >)$$

for all $\mathbf{a}, \mathbf{b} \in H$. Let

$$\theta(t) = \frac{\gamma(t)}{\alpha(\mathbf{0})\beta(\mathbf{0})}, t \geq 0.$$

Note that the function $\theta(.)$ is measurable and the equation (3.3) implies that

$$\theta(< S_1\mathbf{a}, \mathbf{a} >)\theta(< S_2\mathbf{b}, \mathbf{b} >) = \theta(< S_1\mathbf{a}, \mathbf{a} > + < S_2\mathbf{b}, \mathbf{b} >) \tag{3. 4}$$

for all $\mathbf{a}, \mathbf{b} \in H$. Hence the function $\theta(.)$ is a measurable function such that

$$\theta(t)\theta(s) = \theta(t+s) \tag{3. 5}$$

for all $t, s \geq 0$ since S_1 and S_2 are positive definite operators. Therefore

$$\theta(t) = e^{c\,t}, t \geq 0 \tag{3. 6}$$

for some constant c. Hence

$$\gamma(t) = e^{c\,t}\alpha(\mathbf{0})\beta(\mathbf{0}), t \geq 0. \tag{3. 7}$$

Therefore, for any element $\mathbf{a} \in H$,

$$\gamma(< S_1\mathbf{a}, \mathbf{a} >) = e^{c\,<S_1\mathbf{a},\mathbf{a}>}\beta(\mathbf{0})\alpha(\mathbf{0}), \mathbf{a} \in H. \tag{3. 8}$$

Note that

$$\gamma(< S_1\mathbf{a}, \mathbf{a} >) = \alpha(\mathbf{a})\beta(\mathbf{0}), \mathbf{a} \in H \tag{3. 9}$$

from (3.2). Combining the equations (3.8) and (3.9) and noting that $\beta(\mathbf{0}) \neq 0$, it follows that

$$\begin{aligned} e^{c\,<S_1\mathbf{a},\mathbf{a}>}\alpha(\mathbf{0}) &= \alpha(\mathbf{a}) \\ &= \int_H \exp[-\frac{1}{2} < S_1(\mathbf{x}-\mathbf{a}), (\mathbf{x}-\mathbf{a} >]\mu_X(d\mathbf{x}). \end{aligned} \tag{3. 10}$$

The expression on the right side of the equation (3.10) is the convolution of a Gaussian density with the distribution μ_X within a constant. Hence the expression on the left side of the equation (3.10) also has to be a probability density function which implies that the constant $c < 0$ with a suitable normalizing constant $\alpha(\mathbf{0})$. The characteristic functions of the probability densities on both sides of the equation (3.10), then, should satisfy the relation

$$\exp[-\frac{1}{2} < S_1\mathbf{t}, \mathbf{t} > \sigma^2] = \exp[-\frac{1}{2} < S_1\mathbf{t}, \mathbf{t} >]\ \phi_X(\mathbf{t}), \mathbf{t} \in H \tag{3. 11}$$

where $\phi_\mathbf{X}$ is the characteristic function of the random element $\mathbf{X}$ for some $\sigma^2 > 0$. Hence

$$\phi_\mathbf{X}(\mathbf{t}) = \exp[-\frac{1}{2} < (\sigma^2 S_1 - S_1)\mathbf{t}, \mathbf{t} >], \mathbf{t} \in H. \tag{3. 12}$$

for some $\sigma^2 > 0$. Since $\phi_{\mathbf{X}}$ is the characteristic function of the random element $\mathbf{X}$, it follows that $\sigma^2 > 1$ and the random vector $\mathbf{X}$ has the Gaussian measure with the mean zero and the covariance operator $(\sigma^2 S_1 - S_1)$. Similar arguments prove that the random element $\mathbf{Y}$ is also Gaussian with mean zero and the covariance operator $(\sigma^2 S_2 - S_2)$ for some constant $\sigma^2 > 1$.

The converse part of the result stated in the theorem can be easily verified.

Suppose f and g are probability density functions on H. Let

$$Z = \Pi_{j=1}^{N} f(\mathbf{a} - \mathbf{X}_j) g(\mathbf{b} - \mathbf{Y}_j), \mathbf{a}, \mathbf{b} \in H.$$

Theorem 3.2: Suppose that the function $L(\mathbf{a}, \mathbf{b}) = E[Z] < \infty, \mathbf{a}, \mathbf{b} \in H$. Then the function $L(\mathbf{a}, \mathbf{b})$ is a measurable function of the function $(S_1\mathbf{a}, \mathbf{a}) + (S_2\mathbf{b}, \mathbf{b})$ if and only if the distributions μ_X and μ_Y are Gaussian with mean vectors x_0 and y_0 and covariance operators S_X and S_Y respectively and the probability density functions f and g are Gaussian measures with mean vectors μ_f and μ_g and the covariance matrices S_f and S_g respectively with

$$x_0 + \mu_f = y_0 + \mu_g = \mathbf{0}$$

and

$$S_X + S_f = \sigma^2 S_1; S_Y + S_g = \sigma^2 S_2$$

for some $\sigma^2 > 0$.

Proof : Let $\alpha(\mathbf{a}) = E[f(\mathbf{a} - \mathbf{X})]$ and $\beta(\mathbf{b}) = E[g(\mathbf{b} - \mathbf{Y})], \mathbf{a}, \mathbf{b} \in H$. It is easy to check that

$$(3.\ 13) \qquad \begin{aligned} E[Z] &= \sum_{n=1}^{\infty} [E(f(\mathbf{a} - \mathbf{X}))E(g(\mathbf{b} - \mathbf{Y}))]^n P(N = n) \\ &= Q(\alpha(\mathbf{a})\beta(\mathbf{b})) \quad \text{(say)}. \end{aligned}$$

Suppose that $E(Z) = \psi(< S_1\mathbf{a}, \mathbf{a} >> + < S_2\mathbf{b}, \mathbf{b} >)$ for some function $\psi(.)$ Then

$$Q(\alpha(\mathbf{a})\beta(\mathbf{b})) = \psi(< S_1\mathbf{a}, \mathbf{a} > + < S_2\mathbf{b}, \mathbf{b}) >, \mathbf{a}, \mathbf{b} \in H.$$

This relation is similar to that in equation (3.1). Arguments similar to those given earlier show that there exists a constant c such that

$$(3.\ 14) \qquad \alpha(\mathbf{0}) \exp[c \ < S_1\mathbf{a}, \mathbf{a} >] = \int_H f(\mathbf{a} - \mathbf{x}) \ \ \mu_X(d\mathbf{x}), \mathbf{a} \in H.$$

Note that the expression on the right side of the equation (3.14) is the convolution of the probability density function f with the distribution function μ_X . Hence the function on the left side of the equation (3.14) has to be a probability density function which implies that $c < 0$ and $\alpha(\mathbf{0})$ is a suitable normalizing constant for the corresponding Gaussian density function with the mean zero and the covariance operator $\sigma^2 S_1$ for some $\sigma^2 > 0$. An application of the Cramer's theorem for probability measures on a Hilbert space H stated above (Theorem 2.1) proves that f is a Gaussian probability density function and μ_X is a Gaussian probability measure such that

$$\mu_f + x_0 = \mathbf{0}$$

and

$$S_f + S_X = \sigma^2 S_1.$$

Similar arguments show that g and μ_Y are also Gaussian with

$$\mu_g + y_0 = \mathbf{0}$$

and

$$S_g + S_Y = \sigma^2 S_2.$$

The converse part of the result in Theorem 3.2 can be established easily.

Acknowledgement : This work was supported under the scheme "Ramanujan Chair Professor" by grants from the Ministry of Statistics and Programme Implementation, Government of India (M 12012/15(170)/2008-SSD dated 8/9/09), the Government of Andhra Pradesh, India (6292/Plg.XVIII dated 17/1/08) and the Department of Science and Technology, Government of India (SR/S4/MS:516/07 dated 21/4/08) at the CR Rao Advanced Institute for Mathematics, Statistics and Computer Science, Hyderabad, India.

References

1 Cook. L. 1973. A characterization of correlated normal random vectors, *Sankhya, Series A*, 35: 79-84.

2 Grenander, U. 1963. *Probabilities on Algebraic Structures*, Wiley, New York.

3 Kagan, A.M., Linnik, Yu.V. and Rao, C.R. 1973. *Characterization Problems in Mathematical Statistics*, Wiley, New York.

4 Kagan , A.M. and Shalaevski,O. 1967. Characterization of normal law by a property of the non-central chi-square distribution, *Lithuanian Journal of Math.*, 7. (In Russian).

5 Kotlarski, I.I. and Cook. 1977. Characterization of normal random vectors from samples of random size, *Sanhkya, Series B*, 39: 196-200.

6 Kotlarski, I.I. and Wood. 1976. Characterization of normality from samples of random size, Abstract, Institute of Mathematical Statistics, 5, 268.

7 Parthasarathy, K.R. 1968. *Probability Measures on Metric Spaces*, Academic Press, London.

8 Prakasa Rao, B.L.S. 1976. Characterization of multivariate normal distribution from samples of random size, Discussion paper Series, Indian Statistical Institute, New Delhi.

9 Prakasa Rao, B.L.S. 1992. *Identifiability in Stochastic Models: Characterization of Probability Distributions*, Academic Press, Boston.