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

Research Report

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

Title of the Report: **Chebyshev's inequality for Hilbert-space valued random elements with estimated mean and covariance**

Research Report No.: **RR2015-13**

Date: **December 9, 2015**

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

Chebyshev's inequality for Hilbert-space valued random elements with estimated mean and covariance

B.L.S. Prakasa Rao
CR Rao Advanced Institute of Mathematics, Statistics
and Computer Science, Hyderabad 500046, India

Abstract : We obtain a generalization of the Chebyshev's inequality for random elements taking values in a separable Hilbert space with estimated mean and covariance.

Keywords and phrases: Chebyshev's inequality; Random element; Estimated mean; Estimated covariance; Hilbert space.

MSC 2010: 60B11, 60E15.

1 Introduction

Chebyshev' inequality gives a bound on the closeness of a real-valued random variable X to its mean μ in terms of its variance σ^2 whenever it is finite. In fact, for any $\epsilon > 0$,

$$P(|X-\mu| \geq \epsilon) \leq \min(1, \frac{\sigma^2}{\epsilon^2}). \tag{1. 1}$$

Although this inequality is an important inequality to explicitly compute distribution-free probability bounds based on the mean μ and variance σ^2 of the random variable X, it can not be used if the mean and variance are not known. One method that is suggested is to estimate the mean and variance from the sample and to substitute them in the inequality given above. This method might not give reliable bounds for the probability on the left side of equation (1.1) in case the estimators for μ and σ^2 are not good. Saw et al. (1984) studied this problem and obtained a empirical version of the Chebyshev's inequality. Let $X_1, \ldots X_n$ be independent and identically distributed (i.i.d.) random variables and

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

denote the sample mean and

$$s_n^2 = \frac{1}{n-1}\sum_{i=1}^{n-1}(X_i - \bar{X}_n)^2$$

denote the sample variance. Let X_{n+1} be a random variable independent of the sample $X_1, \ldots, X_n$ but identically distributed as the random variable X_1. The following result is due to Saw et al. (1984).

Theorem 1.1: Suppose $X_1, \ldots, X_{n+1}$ are i.i.d. random variables. Then, for any $\epsilon > 0$,

$$P(|X_{n+1} - \bar{X}_n| \geq \epsilon s_n) \leq \min(1, \frac{(n^2 - 1 + n\epsilon^2)}{n^2\epsilon^2}).$$

This result is a slight variant of the result in Saw et al. (1984) and is as given in Stellato et al. (2016).

Stellato et al. (2016) have recently obtained a multivariate version of the Chebyshev's inequality with estimated mean and variance. They obtained a generalization of the result of Saw et al. (1984) under the assumption that the observed sample consists of independent and identically distributed random vectors. Their result is as follows.

Theorem 1.2: Suppose $\mathbf{X}_1, \ldots, \mathbf{X}_{n+1}$ are i.i.d. d-dimensional random vectors with $n \geq d$. Let

$$\mu_n = \frac{1}{n}\sum_{i=1}^{n} \mathbf{X}_i$$

and

$$\Sigma_n = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{X}_i - \mu_n)(\mathbf{X}_i - \mu_n)'.$$

Suppose the matrix Σ_n is non-singular. Then, for any $\epsilon > 0$,

$$P((\mathbf{X}_{n+1} - \mu_n)'\Sigma_n^{-1}(\mathbf{X}_{n+1} - \mu_n) \geq \epsilon^2) \leq \min(1, \frac{d(n^2 - 1 + n\epsilon^2)}{n^2\epsilon^2}).$$

It can be seen that the upper bound in the inequality given above tends to $\min(1, \frac{d}{\epsilon^2})$ as $n \to \infty$.

2 Main Result

It is now known that results similar to those discussed Section 1 are also of interest and important for random variables which are function-valued as they will have possible applications in functional data analysis (cf. Ramsay and Silverman (1997)) which deals with modeling and analysis of observations which are function-valued. In many cases, the function space is the L^2-space of square integrable functions on the real line which is a Hilbert space.

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let H be a real separable Hilbert space and let $(\mathbf{x}, \mathbf{y})$ denote the inner product between $\mathbf{x}$ and $\mathbf{y}$ for $\mathbf{x}, \mathbf{y} \in H$. Let $||\mathbf{x}||$ denote the norm of $\mathbf{x}$ for $x \in H$. Let $\mathbf{X}$ be a random element defined on $(\Omega, \mathcal{F}, P)$ and taking values in the Hilbert space H. Suppose μ is the probability measure of the random element $\mathbf{X}$. For properties of probability measures on a Hilbert space H, see Parthasarathy (1967). Suppose that

$$\int_H ||\mathbf{x}||^2 d\mu(\mathbf{x}) < \infty. \tag{2. 1}$$

The covariance operator S of μ is the Hermitian operator determined uniquely by the quadratic form

$$(S\mathbf{y}, \mathbf{y}) = \int_H (\mathbf{x}, \mathbf{y})^2 d\mu(\mathbf{x}). \tag{2. 2}$$

A positive semi-definite Hermitian operator S is called an S-operator if it has finite trace, that is, for some orthonormal basis $\{e_i, i \geq 1\}$,

$$\sum_{i=1}^{\infty} (Se_i, e_i) < \infty. \tag{2. 3}$$

If the above inequality holds for some orthonormal basis of H, then it will hold for every orthonormal basis of H.

It is obvious that

$$\int_H ||\mathbf{x}|| d\mu(\mathbf{x}) < \infty \tag{2. 4}$$

under the condition stated in (2.2). Then

$$\int_H (\mathbf{x}, \mathbf{y}) d\mu(\mathbf{x})$$

is defined for each $\mathbf{y} \in H$ and

$$\left|\int_H (\mathbf{x}, \mathbf{y}) d\mu(\mathbf{x})\right| \leq ||\mathbf{y}|| \int_H ||\mathbf{x}|| d\mu(\mathbf{x}) < \infty. \tag{2. 5}$$

Therefore the functional $\int_H (\mathbf{x}, \mathbf{y}) d\mu(\mathbf{x})$ is a bounded linear functional with the norm bounded by $\int_H ||\mathbf{x}|| \mu(d\mathbf{x})$. Hence there exists an element $\mathbf{x}_0 \in H$ such that

$$(\mathbf{x}_0, \mathbf{y}) = \int_H (\mathbf{x}, \mathbf{y}) d\mu(\mathbf{x}), \mathbf{y} \in H$$

by the Riesz representation theorem (cf. Dunford and Schwartz (1958); Kolmogorov and Fomin (1975), p. 188) The element $\mathbf{x}_0$ is called the expectation of the random element X and is denoted by $\int_H \mathbf{x}\ \mu(d\mathbf{x})$. Without loss of generality, suppose that $\mathbf{x}_0 = \mathbf{0}$ where $\mathbf{0}$ is the identity element in H.

The following theorem is proved in Prakasa Rao (2010) for random elements taking values in a real separable Hilbert space H.

Theorem 2.1: Suppose $\mathbf{X}$ is a random element taking values in a real sepearable Hilbert space H with expectation zero, covariance operator S, and probability distribution η such that

$$\int_H ||\mathbf{x}||^2 d\eta(\mathbf{x}) < \infty.$$

Then, for every $\epsilon > 0$,

$$P[(S\mathbf{X}, \mathbf{X}) > \epsilon] \leq \frac{[\int_H ||\mathbf{x}||^2 d\eta(\mathbf{x})]^2}{\epsilon}.$$

A more general version of Theorem 2.1 can be obtained in the following way. Let γ be a probability measure on the Borel σ-algebra generated by the norm topolgy on the space H. Further suppose that

(2. 6) $$\int_H ||\mathbf{z}||^2 d\gamma(\mathbf{z}) < \infty.$$

It is known that the bilinear form

$$B(\mathbf{x}, \mathbf{y}) = \int_H (\mathbf{x}, \mathbf{z})(\mathbf{y}, \mathbf{z}) d\gamma(\mathbf{z}), \mathbf{x}, \mathbf{y} \in H$$

determines a unique compact linear operator S_γ from H into H, (cf. Dunford and Schwartz (1958)) such that for all $\mathbf{x}, \mathbf{y} \in H$,

$$(S_\gamma \mathbf{x}, \mathbf{y}) = B(\mathbf{x}, \mathbf{y})$$

and the following result can be obtained extending Theorem 2.1.

Theorem 2.2: Let $\mathbf{X}$ be a random element taking values in a real separable Hilbert space H with probability distribution μ on the Borel σ-algebra generated by the norm on the space H. Further suppose that the probability measures μ and γ satisfy the inequality of the type (2.6). Then, for every $\epsilon > 0$,

$$P((S_\gamma \mathbf{X}, \mathbf{X}) \geq \epsilon) \leq \frac{\int_H ||\mathbf{z}||^2 d\mu(\mathbf{z}) \int_H ||\mathbf{z}||^2 d\gamma(\mathbf{z})}{\epsilon}.$$

Proof : Since $(S_\gamma \mathbf{x}, \mathbf{x}) \geq 0$ for all $\mathbf{x} \in H$, by Markov's inequality, it follows that

$$P((S_\gamma \mathbf{X}, \mathbf{X}) \geq \epsilon) \quad \leq \quad \frac{E((S_\gamma \mathbf{X}, \mathbf{X}))}{\epsilon}$$

$$
\begin{aligned}
&= \frac{E(\int_H (\mathbf{X},\mathbf{z})^2 d\gamma(\mathbf{z}))}{\epsilon} \\
&\leq \frac{E(||\mathbf{X}||^2)\int_H ||\mathbf{z}||^2 d\gamma(\mathbf{z})}{\epsilon}.
\end{aligned}
$$

Remarks: If $\gamma = \mu$ is the probability distribution of the random element $\mathbf{X}$, then the result in Theorem 2.2 reduces to the result in Theorem 2.1. Furthermore, if H is infinite-dimensional, then $S_\mu = S_\gamma$ is compact and will not have a bounded inverse. We take this opportunity to point out that the inequality in the equation (2.11) in Theorem 2.2 in Prakasa Rao (2010) does not hold in the case of infinite-dimensional Hilbert space.

We now extend the result obtained by Stellato et al. (2016) for real separable Hilbert-space valued random elements with possible applications to functional data analysis. Note that the method used by Stellato et al. (2016) can not be generalized and in fact the upper bound as obtained in Theorem 1.3 (cf. Stellato et al. (2016)) tends to infinity as $d \to \infty$ and hence of no relevance for obtaining upper bounds for probability of an event.

Suppose $\{\mathbf{X_i}, 1 \leq i \leq n\}$ are independent and identically distributed random elements taking values in a separable Hilbert space H with mean x_0 and with the same covariance operator S. Suppose further that the mean x_0 and the covariance operator S are *unknown.* Let

$$\bar{\mathbf{X}}_n = n^{-1}(\mathbf{X}_1 \ldots + \mathbf{X}_n). \tag{2. 7}$$

Then $\bar{\mathbf{X}}_n$ is a random element taking values in H. We call $\bar{X}_n$ as the *empirical mean* of the random elements $\{\mathbf{X_i}, 1 \leq i \leq n\}$. Let

$$S_n = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{X}_i - \bar{\mathbf{X}}_n) \otimes (\mathbf{X}_i - \bar{\mathbf{X}}_n) \tag{2. 8}$$

where $\otimes$ denotes the tensor operator $\mathbf{u} \otimes \mathbf{v}(.) = (\mathbf{u}, .)\mathbf{v}$ for $\mathbf{u}, \mathbf{v} \in H$. Here $(\mathbf{x}, \mathbf{y})$ denotes the inner product between the elements $\mathbf{x}, \mathbf{y} \in H$ and the norm of element $\mathbf{x}$ in the Hilbert space H is denoted by $||\mathbf{x}||$ as indicated earlier. The operator S_n is called the *empirical covariance operator*. Let $\mathbf{X}_{n+1}$ be a random element independent of the i.i.d. sample $\{\mathbf{X}_i, 1 \leq i \leq n\}$ but identically distributed as $\mathbf{X}_1$. Let $\epsilon > 0$. We will now obtain an upper bound for the probability

$$P((S_n(\mathbf{X}_{n+1} - \bar{\mathbf{X}}_n), (\mathbf{X}_{n+1} - \bar{\mathbf{X}}_n)) \ \geq \epsilon)$$

for any $\epsilon > 0$. Note that, for any element $\mathbf{y} \in H$,

$$\begin{aligned}
(S_n\mathbf{y}, \mathbf{y}) &= \frac{1}{n}\sum_{i=1}^{n}((\mathbf{X}_i - \bar{\mathbf{X}}_n) \otimes (\mathbf{X}_i - \bar{\mathbf{X}}_n)\mathbf{y}, \mathbf{y}) \\
&= \frac{1}{n}\sum_{i=1}^{n}[(\mathbf{X}_i - \bar{\mathbf{X}}_n, \mathbf{y})]^2 \\
&= \int_H [(\mathbf{z} - \bar{\mathbf{X}}_n, \mathbf{y})]^2 \nu_n(d\mathbf{z})
\end{aligned}$$

where ν_n is the counting measure assigning probability $\frac{1}{n}$ to each of the elements $\mathbf{X}_i, 1 \leq i \leq n$. Let $\epsilon > 0$ and η denote the probability measure of the random element $\mathbf{X}_1$ induced by the probability measure P on the space H associated with the Borel-σ algebra generated by the norm topology on the Hilbert space H. Let $H^{(n)}$ denote the tensor product of the space H over n copies of the space H. Let $\eta^{(n)}(.)$ denote the product measure generated by the random vector $(\mathbf{X}_1, \ldots, \mathbf{X_n})$ and

$$D_\epsilon = \{\mathbf{z} \in H : (S_n(\mathbf{z} - \bar{\mathbf{X}}_n), (\mathbf{z} - \bar{\mathbf{X}}_n)) \geq \epsilon\}, \epsilon > 0.$$

Then

$$\begin{aligned}
P(D_\epsilon) &= \int_{D_\epsilon \otimes H^n} \; d\eta(\mathbf{z}) d\eta^{(n)}(\mathbf{x}_1, \ldots, \mathbf{x}_n) \\
&\leq \frac{1}{\epsilon}\int_{D_\epsilon \otimes H^{(n)}} (S_n(\mathbf{z} - \bar{\mathbf{x}}_n), (\mathbf{z} - \bar{\mathbf{x}}_n)) \; d\eta(\mathbf{z}) d\eta^{(n)}(\mathbf{x}_1, \ldots, \mathbf{x}_n) \\
&\leq \frac{1}{\epsilon}\int_{H^{(n+1)}} (S_n(\mathbf{z} - \bar{\mathbf{x}}_n), (\mathbf{z} - \bar{\mathbf{x}}_n)) \; d\eta(\mathbf{z}) d\eta^{(n)}(\mathbf{x}_1, \ldots, \mathbf{x}_n) \\
&= \frac{1}{\epsilon}\int_H [\int_{H^{(n)}} \{\int_H (\mathbf{y} - \bar{\mathbf{x}}_n, \mathbf{z} - \bar{\mathbf{x}}_n)^2 \nu_n(dy)\} d\eta^{(n)}(\mathbf{x}_1, \ldots, \mathbf{x}_n)] d\eta(\mathbf{z}) \\
&\leq \frac{1}{\epsilon}\int_H [\int_{H^{(n)}} \{\int_H ||\mathbf{y} - \bar{\mathbf{x}}_n||^2 ||\mathbf{z} - \bar{\mathbf{x}}_n||^2 \nu_n(dy)\} d\eta^{(n)}(\mathbf{x}_1, \ldots, \mathbf{x}_n)] d\eta(\mathbf{z}) \\
&= \frac{1}{\epsilon}\int_H [\int_{H^{(n)}} ||\mathbf{z} - \bar{\mathbf{x}}_n||^2 \{\int_H ||\mathbf{y} - \bar{\mathbf{x}}_n||^2 \nu_n(dy)\} d\eta^{(n)}(\mathbf{x}_1, \ldots, \mathbf{x}_n)] d\eta(\mathbf{z}) \\
&= \frac{1}{\epsilon}\int_H [\int_{H^{(n)}} ||\mathbf{z} - \bar{\mathbf{x}}_n||^2 \{\frac{||\mathbf{x}_1 - \bar{\mathbf{x}}_n||^2 + \ldots + ||\mathbf{x}_n - \bar{\mathbf{x}}_n||^2}{n}\} d\eta^{(n)}(\mathbf{x}_1, \ldots, \mathbf{x}_n)] d\eta(\mathbf{z}) \\
&= \frac{1}{\epsilon}\int_H [\int_{H^{(n)}} ||\mathbf{z} - \bar{\mathbf{x}}_n||^2 ||\mathbf{x}_1 - \bar{\mathbf{x}}_n||^2 d\eta^{(n)}(\mathbf{x}_1, \ldots, \mathbf{x}_n)] d\eta(\mathbf{z}) \\
&\quad \text{since } \mathbf{X}_1, \ldots, \mathbf{X}_n \text{ are i.i.d. random elements.}
\end{aligned}$$

Hence

$$P(D_\epsilon) \leq \frac{1}{\epsilon} E[||\mathbf{X}_{n+1} - \bar{\mathbf{X}}_n||^2 ||\mathbf{X}_1 - \bar{\mathbf{X}}_n||^2].$$

It is easy to see that

$$\mathbf{X}_{n+1} - \bar{\mathbf{X}}_n = \frac{n+1}{n}(\mathbf{X}_{n+1} - \bar{\mathbf{X}}_{n+1}).$$

Hence

$$\begin{aligned} P(D_\epsilon) &\leq (\frac{n+1}{n})^2 \frac{1}{\epsilon} E[||\mathbf{X}_{n+1} - \bar{\mathbf{X}}_{n+1}||^2 ||\mathbf{X}_1 - \bar{\mathbf{X}}_n||^2] \\ &= (\frac{n+1}{n})^2 \frac{1}{\epsilon} E[||\mathbf{X}_{n+1} - \bar{\mathbf{X}}_{n+1}||^2 ||\mathbf{X}_n - \bar{\mathbf{X}}_n||^2] \\ &\quad \text{since } \mathbf{X}_1, \ldots, \mathbf{X}_n \text{ are i.i.d. random elements} \end{aligned}$$

and we obtain the following theorem.

Theorem 2.3: Suppose $\mathbf{X}_i, 1 \leq i \leq n+1$ are i.i.d. random elements taking values in a real separable Hilbert space such that

$$E[||\mathbf{X}_{n+1} - \bar{\mathbf{X}}_{n+1}||^2 ||\mathbf{X}_n - \bar{\mathbf{X}}_n||^2] < \infty.$$

Then, for every $\epsilon > 0$,

$$P((S_n(\mathbf{X}_{n+1} - \bar{\mathbf{X}}_n), (\mathbf{X}_{n+1} - \bar{\mathbf{X}}_n)) \geq \epsilon) \leq \min\{1, (\frac{n+1}{n})^2 \frac{1}{\epsilon} E[||\mathbf{X}_{n+1} - \bar{\mathbf{X}}_{n+1}||^2 ||\mathbf{X}_n - \bar{\mathbf{X}}_n||^2]\}.$$

Suppose, in addition, the random variable $||\mathbf{X}||$ has 4-th moment, that is,

$$\int_H ||\mathbf{x}||^4 \eta(d\mathbf{x}) < \infty.$$

Let

$$\gamma_n^2 = E[||\mathbf{X}_n - \bar{\mathbf{X}}_n||^4].$$

Then

$$\begin{aligned} \gamma_n^2 n^4 &= E[||\mathbf{X}_1 - \mathbf{X}_n + \ldots + \mathbf{X}_n - \mathbf{X}_n||^4] \\ &\leq E[(||\mathbf{X}_1 - \mathbf{X}_n|| + \ldots + ||\mathbf{X}_n - \mathbf{X}_n||)^4] \\ &\leq n^3 \sum_{i=1}^n E[||\mathbf{X}_i - \mathbf{X}_n||^4] \\ &\quad \text{(by the } c_r\text{-inequality, see Lin and Bai (2010), p.97)} \\ &= n^4 E[||\mathbf{X}_1 - \mathbf{X}_n||^4] \\ &\quad \text{(since } \mathbf{X}_1, \ldots, \mathbf{X}_n \text{ are i.i.d. random elements)} \end{aligned}$$

$$\leq \quad 8n^4(E||\mathbf{X}_1||^4 + E||\mathbf{X}_n||^4)$$

(by the c_r-inequality, see Lin and Bai (2010). p.97)

$$= \quad 16n^4 E[||\mathbf{X}_1||^4] < \infty$$

(since $\mathbf{X}_1, \mathbf{X}_2$ are identically distributed random elements).

Hence

$$\gamma_n^2 \leq 16E[||\mathbf{X}_1||^4] < \infty \tag{2. 9}$$

for every $n \geq 1$. We will now compute a bound on the term

$$\zeta_n = E[||\mathbf{X}_{n+1} - \bar{\mathbf{X}}_{n+1}||^2||\mathbf{X}_n - \bar{\mathbf{X}}_n||^2].$$

Note that

$$\begin{aligned} \zeta_n &\leq (E[||\mathbf{X}_{n+1} - \bar{\mathbf{X}}_{n+1}||^4]\ E[||\mathbf{X}_n - \bar{\mathbf{X}}_n||^4])^{1/2} \\ &= [\gamma_{n+1}^2\gamma_n^2]^{1/2} \\ &\leq 16E[||\mathbf{X}_1||^4] \ \text{(by the equation (2.9))} \end{aligned}$$

Hence

$$P((S_n(\mathbf{X}_{n+1} - \bar{\mathbf{X}}_n), (\mathbf{X}_{n+1} - \bar{\mathbf{X}}_n)) \ \geq \epsilon) \leq \min\{1, (\frac{n+1}{n})^2 \frac{16}{\epsilon} E[||\mathbf{X}_1||^4]\} \tag{2. 10}$$

and we have the following result.

Theorem 2.4: Suppose $\mathbf{X}_i, 1 \leq i \leq n+1$ are i.i.d. random elements taking values in a separable Hilbert space such that $E[||\mathbf{X}_1||^4] < \infty$. Then, for every $\epsilon > 0$,

$$P((S_n(\mathbf{X}_{n+1} - \bar{\mathbf{X}}_n), (\mathbf{X}_{n+1} - \bar{\mathbf{X}}_n)) \ \geq \epsilon) \leq \min\{1, (\frac{n+1}{n})^2 \frac{16}{\epsilon} E[||\mathbf{X}_1||^4]\}. \tag{2. 11}$$

3 Application :

Let $\mathbf{X} = \{X(t), 0 \leq t \leq 1\}$ be a second order stochastic process with sample paths in the Hilbert space $L_2[0,1]$ with probability one. Suppose that $E(X(t)) = m(t)$ and $Cov(X(t), X(s)) = r(t,s)$ where the mean function $m(.)$ and the covariance function $r(t,s)$ are unknown and

$$E[(\int_0^1 X_1^2(t)dt)^2] < \infty.$$

Suppose we take repeated independent observations $\mathbf{X}_i, 1 \leq i \leq n$ of the process $\mathbf{X}$. Let

$$m_n(t) = \frac{1}{n}\sum_{i=1}^{n} X_i(t). \tag{3. 1}$$

Then the function $m_n(.)$ is an estimator of the function $m(.)$. As an application of Theorem 2.4, we get the following result: for any $\epsilon > 0$,

$$P(\frac{1}{n}\sum_{i=1}^{n}[\int_0^1 (X_i(t) - m_n(t))(X_{n+1} - m_n(t))dt]^2 \geq \epsilon) \leq (\frac{n+1}{n})^2 \frac{16}{\epsilon} E[(\int_0^1 X_1^2(t)dt)^2].$$

4 Remarks:

The upper bound obtained in Theorem 2.4 is not distribution free as it depends on the fourth moment of the distribution of the random elementX_1. It would be interesting to check whether an upper bound can be obtained which is distribution free as in the finite dimensional case obtained in Saw et al. (1984) and Stellato et al. (2016).

References :

Dunford, N. and and Schwartz, J.T. (1958)*Linear Operators, Part I: General Theory*, Interscience Publishers Inc, New York.

Kolmogorov, A.N and Fomin, S.V. (1975) *Introductory Real Analysis*, Translated and Edited by Richard A. Silverman, Dover Publications, Inc., New York.

Lin, Zhengyan and Bai, Zhidong (2010) *Probability Inequalities*, Science Press, Beijing and Springer, Berlin.

Parthasarathy, K. R. (1967) *Probability Measures on Metric Spaces*, Academic Press, London.

Prakasa Rao, B.L.S. (2010) Chebyshev's inequality for Hilbert-space-valued random elements, *Statist. Probab. Lett.*, **80**, 1039-1042.

Ramsay, J. and Silverman, B.W.(1997) *Functional Data Analysis*, Springer, Berlin.

Saw, J.G., Yang, M.C.K., and Mo, T.C. (1984) Chebyshev's inequality with estimated mean and variance, *The American Statistician*, **38**, 130-132.

Stellato, B., Van Parys, Bart P.G., and Goulart, Paul J. (2016) Multivariate Chebyshev inequality with estimated mean and variance, To appear *The American Statistician*, DOI:10.1080/00031305.2016.1186559