*IOSR Journal of Mathematics (IOSR-JM)*
*e-ISSN: 2278-5728.Volume 6, Issue 1 (Mar. - Apr. 2013), PP 53-58*
*www.iosrjournals.org*

# Optimal Estimating Sequence for a Hilbert Space Valued Parameter

P.Anilkumar[1] , N.V.Samiyya[2]
*Department of statistics farook college Kerala India*

***Abstract**: Some* optimality criteria used in estimation of parameters in finite dimensional space has been extended to a separable Hilbert space. Different optimality criteria and their equivalence are established for estimating sequence rather than estimator. An illustrious example is provided with the estimation of the mean of a Gaussian process
***Key words:** sieves, Gaussian process, reproducing kernel Hilbert space , MVUE , estimating sequence*

## I. Introduction

The classical theory of statistical inference deals with the sample spaces and parameter spaces as finite dimensional Euclidian spaces or subsets of it. The observations are assumed to be n independent and identical realizations of a random vector. Much of the inference methodologies can be said to be likelihood based. A first step in solving such a problem is the calculation of the likelihood function. The assumption that the samples are from $R^k$ is not true in general. Sometimes the data may be curves or in general, elements from some abstract spaces. Since most of the methods in classical inference are likelihood based, any attempts to extend classical method of interest should start with the determination of the likelihood function. A major difficulty is that in the spaces of interest, there is no natural invariant Lebesgue measure which plays an important role in classical inference. However, fortunately in the abstract spaces, Radon-Nikodym derivative of one hypothetical measure with respect to another plays the role of the likelihood. Likelihood based inferences in abstract spaces have been discussed in the literature with considerable depth. "Statistical inference for stochastic process" by Basava and Prakasa Rao (1980) contains good account of this exposition.

There are instance, for example, in inference for stochastic process where both sample space and parameter space are abstract. Here also the first step involved is the calculation of the likelihood. But even after obtaining the likelihood, the estimation is not straight forward. In most of the cases likelihood function will be unbounded. [See for example, Grenander (1981, P.395), Karr (1987), Beder (1987, 88)]. One possible approach to overcome this difficulty is to apply the method of sieves systematical exposition by Grenander (1981). Adopting Grenander's approach some authors have already studied estimation for stochastic processes. Works of Mckeague (1986), Beder (1987, 88), Karr (1987), Lesbow and Rozanski (1989) can be mentioned in this context.

In this paper we will define an optimality criterion for a Hilbert space valued parameter through the notations of estimating sequence. The criterion so developed is then applied to obtain an optimal estimating sequence for the mean of a Gaussian process'

## II. An optimality criterion for a sequence of parameters

We begin with certain optimality results available in simultaneous estimation of several parameters. Let $x_1,x_2,\ldots,x_n$ be random sample of size n from a population characterized by the probability density function f(x, θ) $\theta \in \Omega\ CR^m$ where $R^m$ being m dimensional Euclidian space. Suppose we are interested in k functions $\Psi(\theta) = (\Psi_1(\theta),\Psi_2(\theta),\ldots,\Psi_k(\theta))^T$. A vector of statistics $T = (T_1,T_2,\cdots T_k)^T$ is unbiased for Ψ(θ). If $E_\theta(T_i) = \Psi_i(\theta)$ for all $\theta \in \Omega$ for every i = 1,2,…,k. We would also say T is marginally unbiased for Ψ(θ) . For any two estimators T and S define

$M_T = E(T - E(T))(T - E(T))^T$

And $M_{TS} = E(T - E(T))(S - E(S))^T$

Let $U_\psi$ denote the class of all unbiased estimators of ψ.

**Definition 2.1**

The estimator $T^* \in U_\psi$ is M- optimal for Ψ if $M_T - M_{T^*}$ is nonnegative definite for all $\theta \in \Omega$.

A necessary and sufficient condition that $T^* \in U_\psi$ is M- optimal is that $M_{T^*U} = M_{UT^*} = 0$ for all $\theta \in \Omega$, $u \in U_0^{(k)}$ where $U_0^{(k)}$ is the class of all k-dimentional statistics that are unbiased for $0 = (0,0,\ldots,0)^T$. This is known as klebenov- Linnik- Rukhin theorem.

**Definition 2.2**

An estimator $T^* = (T_1^*, T_2^* \ldots, T_k^*)$ is marginally optimal for ψ provided $T_i^*$ is MVUE for $\psi_i$, i = 1,2,…,k.

By an application of klebenov- Linnik- Rukhin theorem one conclude that an estimator is M – optimal if and only if it is marginally optimal. Some other equivalent optimality criteria are D-optimality, T- optimality and E-optimality. See B.K. Kale (1999) for a fairly good discussion.

A natural question that can arise is, can we consider a more general abstract parameter space and derive a similar optimality criteria. As an immediate generalization suppose that the parameter space Ω is a real separable Hilbert space. Then there exists a countable orthonormal basis $\{e_i\}$ such that any θ ∈ Ω can be written as $\Theta = \Sigma\psi_i e_i$ , where $\psi_i = (\theta, e_i)$ where (.,.) is the inner product associated with the Hilbert space Ω. Thus estimation of θ reduces to that of estimating $\psi_i$ , 1,2,…. It is also well known that

$\Sigma\psi_i^2 < \infty$ ………..(2.1)

Notice that practically one cannot estimate all $\psi_i$'s simultaneously. But the condition (2.1) tells us that $\psi_i \to 0$ as $i \to \infty$. Thus one take $\psi_i = 0$ after a stage say I d and estimate θ. This is precisely method of sieves suggested by Grenander (1981). Instead of resorting to method of sieves we will consider an estimating sequence and formulate an optimality criteria for an estimating sequence $\{T_i\}$.

**Definition 2.3**

A sequence of estimators T={Ti}is called an unbiased estimating sequence for $\psi=\{\psi_i\}$if $E_\theta(T_i) = \psi_i$ for all θ ∈ Ω, I =1,2,…….. Now we define three optimality criteria and establish their equivalence.

**Definition 2.4**

An unbiased estimating sequence $T=\{T_i\}$is sequential M-optimal for $\Psi=\{\psi_i\}$ if $(T_1,T_2.....T_k)'$ is M-opti∈∈mal for $(\psi_1, \psi_2,..... \psi_k)'$for every K=1,2,…

**Definition 2.5**

An unbiased estimating sequence $\{T_i\}$is marginally optimal if $T_i$ is MVUE for $\psi_i$, for all i = 1,2,3….

For an estimating sequence $\{T_i\}$, consider the linear functional $\sum_{i=1}^{\infty} l_i\psi_i$ , such that $\sum_{i=1}^{\infty} l_i^2 < \infty$.

**Definition 2.6**

An unbiased estimating sequence $\{T_i\}$ is functionally optimal if $\sum_{i=1}^{\infty} l_i T_i$ is MVUE for $\sum_{i=1}^{\infty} l_i\psi_i$ for all $\{l_i\}$ satisfying $\sum_{i=1}^{\infty} l_i^2 < \infty$.

**Theorem 2.1**

Let $U_0$ be the set of all unbiased estimators of 0 ,that is $U_0 = \{u, E_\theta(u) = 0$, for all θ ∈ Ω}, then an estimating sequence $T^* = \{T_i^*\}$ is functionally optimal if $E_\theta((\Sigma l_i T_i^*) u) = 0$ for all u ∈ $U_0$ and all θ ∈ Ω

Proof

Suppose the condition $E_\theta\ ((\Sigma\ l_i\ T_i^*)u) = 0$ for all u ∈ $U_0$ consider the estimator $\Sigma l_i T_i^* - \Sigma l_i T_i$ where $\{T_i\}$is any other unbiased estimating sequence. Then,

$\boldsymbol{E_\theta\ (\Sigma\ l_i\ T^*_i - \Sigma\ l_i\ T_i) = 0}$ **and so** $\boldsymbol{E_\theta(\Sigma\ l_i\ T^*_i)\ (\Sigma\ l_i\ T^*_i - \Sigma\ l_i\ T_i) = 0}$ **which implies**

$$\boldsymbol{E_\theta(\Sigma\ l_i\ T^*_i)^2 = E_\theta(\Sigma\ l_i\ T^*_i)\ (\Sigma\ l_i\ T_i)}$$
$$\boldsymbol{\leq E_\theta^{1l2}(\Sigma\ l_i\ T^*_i)^2\ \ E_\theta^{1l2}(\Sigma\ l_i\ T_i)^2}$$

Hence

$$\boldsymbol{E_\theta\ (\Sigma\ l_i\ T^*_i)^2 \leq E_\theta\ (\Sigma\ l_i\ T_i)^2}$$

**Or** $\qquad \boldsymbol{Var_\theta\ (\Sigma\ l_i\ T^*_i) \leq Var_\theta(\Sigma\ l_i\ T_i)}$

Proving that $\sum_{i=1}^{\infty} l_i T_i^*$ is MVUE for $\sum_{i=1}^{\infty} li\,\psi i$. Conversely suppose $\sum_{i=1}^{\infty} l_i T_i^*$ is MVUE and if possible suppose $E_{\theta_0}\ ((\sum_{i=1}^{\infty} l_i T_i^*)u_0) \neq 0$ for some given choice of $\theta = \theta_0$ and $u_0 \in U_0$

Define

$$\boldsymbol{\lambda = \frac{-E_{\theta_0}\left((\Sigma li\ Ti^*)u_0\right)}{E_{\theta_0}(u_0{}^2)}}$$

Then

$$\boldsymbol{E_{\theta_0}(\sum l_i\ T^*_i + \lambda u_0)^2 = E_{\theta_0}\ (\sum l_i\ T^*_i)^2 + \lambda^2\ E_{\theta_0}(u_0^2) - 2E_{\theta_0}((\sum l_i\ T^*_i)u_0)\frac{E_\theta\ (\sum l_i\ T^*_i)u_0}{E_{\theta_0}(u_0^2)}}$$

$$\boldsymbol{= E_\theta(\sum l_i\ T^*_i)^2 - \frac{E_{\theta_0}{}^2((\sum li\ T^*_i)u_0}{E_{\theta_0}(u_0^2)}}$$

$$\boldsymbol{< E_{\theta_0}(\sum l_i\ T^*_i)^2}$$

Hence

$$\boldsymbol{var_{\theta_0}((\sum l_i\ T^*_i)\ < var_{\theta_0}(\sum l_i\ T^*_i + \lambda u_0)}$$

This is a contradiction to the assumption that $\sum l_i\ T_i^*$ is MVUE for $\sum_{i=1}^{\infty} l_i\psi_i$ this proves the theorem,

**Theorem 2.2**

An estimating sequence $\{T_i\}$ is marginally optimal if and only if it is functionally optimal, Proof

Suppose $\{T_i\}$ is functionally optimal. Then $\sum_{i=1}^{\infty} l_i T_i$ is MVUE for $\sum_{i=1}^{\infty} l_i \psi_i$. Choose $l_i = 0$, $i \neq k$ and $l_k = 1$, then $T_k$ is MVUE for $\Psi_k$ , k = 1,2,… Thus $\{T_i\}$ is marginally optimal. Conversely suppose $\{T_i\}$ is marginally optimal. Then

$$E_\theta(T_i\, u) = 0, \forall\, \theta \in \Omega, u \in U_0$$

Hence

$E_\theta((\sum l_i\, T_i)\, u) = \sum_1^{\infty} l_i\, E_\theta\ (T_i\, u) = 0$

Thus by theorem 2.1 $\{T_i\}$ is functionally optimal.

**Theorem 2.3**

The estimating sequence $\{T_i\}$ is sequentially M-optimal if and only if it is marginally optimal

Proof

Let $T^k = [T_i, \ldots\ldots T_{k_i}]^T$ and

$U^{(k)} = [u_i \ldots\ldots\ldots u_k]^T$, where $u_i \in U_0$.

The proof immediately follows from the following observations

$E_\theta(T_i\ U_j)$ for all j,θ, I =1,2,…. If and only if $E\ (T^{(K)} U^{(k)^T}) = 0$, for all $u^{(k)} \in U_0^k$ and for every θ if and only if

**$Cov(T^K, U^K)=0$,** For all $u^{(k)} \in U_0^K$ and all θ

**Remark**

The last two theorems tells us that all the optimality criteria are one and the same . later on we simply say the estimating sequence is optimal if it satisfy any one of the criteria discussed above.

We now proceed to illustrate the criteria in the estimation of mean of a Gaussian process

## III. Gaussian process

Let $\{x(t), t \in T\}$ be a stochastic process defined on (Ω, Ғ) where T is some general index set. Let V be the set of all finite linear combinations of the form $\sum_{i=1}^{n} c_i\, x(t_i)$. Let P be a probability measure attached to the measurable space (Ω, Ғ). Then under the probability measure P,$V_P$ – the set of all P – equivalent classes of elements of V becomes a vector space. We say that the process is Gaussian under P, if each element of $V_p$ is a normal random variable. In this case $V_P$ C $L_2$(Ω, Ғ, p) and its completion $H_p$ consist of only normal random variables. We denote the norm and inner product in $H_p$ by $\|\ .\ \|_p$ and $(\ .\ ,\ .)_p$ respectively

where $(x, y)_p = cov_p(x , y)$ for x , y ∈ $H_p$ the space $H_p$ is usually called Gaussian space. The process $\{x(t), t \in T\}$ is called a Gaussian process.

The function $m(t)=E_P(x(t))$ is called mean function of Gaussian process and R( s, t)= Cov(x(s),x(t)) is called covariance function . We now introduce the concept of reproducing kernel Hilbert space with kernel R, for every t ∈ T

**Definition 3.1**

(a) R (. , t) ∈ K (R,T) and

(b) For every f ∈ K(R,T)

< f , R(. , t) > = f ( t )

Where < . , . > denote inner product on K( R , T )

**Example**

Let T=[o,b], b <∞ and R (s,t)=min (s,t) for s , t ∈ [o,b]. define

**K (R,T)** $= \{ f :\ f\ (t) = \int_0^t f^I(u)du, f^I \in L_2(T)\}$

Where $L_2(T)$ is the set of all square integrable functions defined on T and $f^I$ is the derivative of f. if we define

$< f , g > = \int_0^b f^I\ (u)\ g^I\ (u)\ du$ , for f , g ∈ k (R,T),then k (R,T) is a Hilbert space

To see this observe that R(s,t)= Min (s,t) $= \int_0^s I_{[0,t]}(u)du$

Therefore R(.,t) ∈ K(R,T) as $I_{[0,t]} \in L_2(T)$

Also $< f, R\ (.\ , t) > = \int_0^b f'(s)\ I_{[0,t]}\ (s)\ ds = \int_0^t f'(t)\ dt = f(t)$

The following lemma from Becker (1987) gives a Fourier type expansion for the covariance function R

**Lemma 3.1**

Let $\{e_k, K \in A\}$ be complete orthonormal set in K (R,T),then we have $R(s,t) = \sum_{k \in A} e_k(s) e_k(t)$ For all (s, t)∈ T x T and that the set $A' = \{k \in A: e_k(s)\ e_k(t) \neq 0\}$ is atmost countable . The isomorphic isomorphism between $H_P$ and K(R, T) is established through the next lemma.

**Lemma 3.2**

Let $\{x(t), t \in T\}$ be a centered Gaussian process (that is E(x(t)=0) define on (Ω, Ғ,P)and R(s,t)be the covariance function of the process then the following holds:

(a) The R.K.H.S. k(R,T) is given by { f : f(t) = $(x(t), y_f)$ for a unique $y_f \in H_P$} with inner product

< f , g > = $(y_f, y_g)$ the map given by ʌ $(y_f)$ = f is an isomorphism of $H_P$ onto K(R,T).

(b) For each t, x(t)= $\sum_{k\in A} e_k(t)(U_k)$ where $U_k = \Lambda^{-1}(e_k)$, k ∈ A are iid N(0,1) variables . Further the series $\sum_{k\in A} e_k(t)e_k(U_K)$ converges almost surely. The map $\Lambda$ $(y_f)$=f is called Loeve map.

Let {x(t),t ∈ T} be a Gaussian process define on (Ω,π,P) with mean function m(t) = $E_p(x(t))$ and covarience function R(s,t) = $Cov_p$( x (s) , x (t)). Let $P_0$ be a another Gaussian measure set . $E_{p0}(x(x))=0$ and $cov_p$(x ( t ),x (s))= R (s,t).assume that the RKHS K(R,T) generated by R is complete so that an orthonormal basis {$e_k$}is always countable Chatterji,S.P. and Handrekar,V.(1978) has shown that $P_0$ and p are equivalent if and only if m ∈ k(R , T) assume that m ∈ k(R , T). since {$e_k$}is a complete orthonormal basis for k(R , T) and m ∈ k(R , T),m=$\sum_{k=1}^{\infty} a_k e_k$ where $a_k \in l_2$ ,the set of square summable sequence. From elementary Hilbert space theory {$a_k$}is unique to m. notice that

$$\boldsymbol{E_P^{(X(t))} = m(t) = (x(t)\,,y_m)_{p_0}}$$

Where Y = $\Lambda^{-1}$ (m). since $U_k = \Lambda^{-1}$ ($e_k$)

$$\boldsymbol{E_P\,(Uk) = (U_k\,,Y_m)_{p_0}}$$

$$= (\Lambda^{-1}(\boldsymbol{e_k}), \Lambda^{-1}(\boldsymbol{e_k}))_{\boldsymbol{p_0}}$$

$$= <\boldsymbol{e_k\,, m} > = \boldsymbol{a_k}$$

And

$$\boldsymbol{Cov_p\,(U_k, U_l)} = (\Lambda^{-1}(\boldsymbol{e_k}), \Lambda^{-1}(\boldsymbol{e_l}))_{\boldsymbol{p_0}}$$

$$= <\boldsymbol{e_k, e_l}>$$

$$= \boldsymbol{\delta_{kl}}$$

Thus $U_k$'s are independent N($a_k$,1) random variable under the probability measure P.
Finally the Radon Nikodym derivative of P with respect to $P_0$ is given by

$$\boldsymbol{\frac{dp}{dp_0} = exp\{\sum_{k=1}^{\infty}(a_k\,U_k - \frac{a_k^2}{2})\}} \qquad \text{(see Beder (1987)}$$

**Example 3.1**

Let= [0,b] and R is any continuous Covarience function defined on T X T. Define the integral operator with kernel R as

$$\boldsymbol{R(f(s)) = \int_0^b R(s,t)\,f(t)\,dt}$$

Then R is an operator on $L_2$ with countable system of eigen values {$\lambda_k$}and eigen functions {$\varphi_k$}. We use R for the operator as well as Kernel ) and we have

$$\boldsymbol{R(s,t) = \sum_k \lambda_k\,e_k(s)e_k(t)} \qquad \text{(C.F. Ash and Gardner(1975) p 37)}$$

Let k(R,T) be the space spanned by {$\varphi_k$}and define the inner product on K(R,T) as follows. For f= $\sum_k a_k \varphi_k$ and g = $\sum b_k \varphi_k$, define .

$$\boldsymbol{<f\,,g> = \sum\sum a_j\,b_j\,\frac{\varphi_j\,(t)\varphi_k(t)}{\surd\lambda_k\,\lambda_j}\,dt}$$

$$= \boldsymbol{\sum_k \frac{a_k\,b_k}{\lambda_k}}$$

Set $e_k=\surd\lambda_k\, e_k$, then {$e_k$}is a complete orthonormal basis for k(R,T)

$$\textbf{R(s,t)} = \boldsymbol{\sum_k e_k(s)e_k(t)} \qquad \& \qquad \boldsymbol{<R(\,.\,,t), f> = <\sum e_k(t)e_k\,, \sum a_j\,e_k>}$$

$$= \boldsymbol{\sum a_k\,e_k(t) = \mathbf{f}\,(\mathbf{t})}$$

Thus k(R,T) is a R.K.H.S. of x(t) is Gaussian process with mean function m ∈ k (R,T) and covariance function R then m admits the decomposition. M=$\sum a_k e_k$ and x(t)=$\sum_k U_k e_k$ (t) where $a_k$ = < m , $e_k$ > and

$$\boldsymbol{U_k = \frac{1}{\lambda_k}\int_0^b x\,(t)\,e_k(t)\,dt} \qquad \text{(C.F. Ash and Gardner(1975))}$$

**Example 3.2 (wiener process)**

Let x(t) be a wiener process defined on [0,b] with mean m (.) and covariance function R(s,t)=Min (s,t).

since R(s,t)= Min (s,t) the integral equation $\boldsymbol{\int_0^b R(s,t)\,\varphi(t)\,dt = \lambda\,\varphi(\mathbf{s})}$ becomes

$$\boldsymbol{\int_0^s t\,\varphi(t)\,dt + \int_s^b s\,\varphi(t)dt = \lambda\,\varphi(s)} \quad \rightarrow \mathbf{(3.1)}$$

On differentiating (3.1) with respect to S we have

$$\boldsymbol{S\,\varphi(s) + \int_s^b \varphi(t)dt - s\,\varphi(s) = \lambda\,\varphi^I(s)}$$

$$\boldsymbol{i.e.\ \int_s^b \varphi(t)dt = \lambda\,\varphi^I\,(s)} \quad \rightarrow \mathbf{(3.2)}$$

Differentiating again we get

$$\boldsymbol{-\varphi(s) = \lambda\,\varphi^{II}(s)}$$

If $\lambda = 0$ then $\varphi$ (s)=0 and so $\lambda = 0$ is not an eigen value. Putting s = 0 in (3.1) we get

$$\lambda\, \varphi(0) = 0 \quad \rightarrow \varphi(0) = 0$$

Putting s= b in (3.2) we get $\mathbf{0 = \lambda\, \varphi^{I}(b) \rightarrow \varphi^{I}(b) = 0}$

Thus $\{\varphi, \lambda\}$ satisfies

$$\varphi^{II} + \lambda^{-1}\varphi = 0 \ \text{ with } \varphi(0) = \varphi^{I}(b) = 0$$

solving the above differential equation after putting $\lambda^{-1} = \beta^2$ we get

$$\varphi_k(t) = \left(\frac{2}{b}\right)^{\frac{1}{2}} Sin\, \beta_k(t), \quad k \in z^+$$

with eigen values $\lambda_k = \beta_k^{-2}$, where $\beta_k = (k - \frac{1}{2})\, \pi/b$

thus the function $e_k(t) = \sqrt{\lambda_k}\, \varphi_k(t) = \beta_k^{-1}\, (2/b)^{1/2} \sin \beta_k(t)$ is a completely orthonormal basis for k (R , T ) and

$$U_k = \beta_k\, (2/b)^{1/2} \int_0^b x(t) \sin \beta_k\, t\, dt$$

**Estimation problem**

Let $\{x_i(t), t \in T\}$, i = 1,2,……..n consist of n iid sample of observations of the process x(t) defined on the probability space ($\Omega$, Ғ,p) with mean function m(t) = $E_p$ (x(t)) and covariance function R(s,t) = $cov_p(x(s),x(t))$ as before take $P_0$ as Gaussian measure so that $E_{p_0}$x(t)) = 0 and $cov_{p_0}$ (x(s), x(t)) = R (s , t) assume that m $\in$ k(R,T) then the Radon Nikodym derivative of P with respect to $P_0$ in the product sample space is

$$\frac{dp^{n\Theta}}{dp_0^{n\Theta}} = \prod_{i=1}^{n} exp\ \{\sum_{k=1}^{\infty} (a_k U_{k_i} - \frac{1}{2}\, a_k^{\ 2})\}$$

$$= \exp\{\sum_{i=1}^{n}\sum_{k=1}^{\infty} (a_k U_{k_i} - \frac{a_k^2}{2})\}$$

$$= exp\, \{n \sum_{k=1}^{\infty} (a_k\, \overline{U_k} - \frac{a_k^2}{2})\}$$

where $\overline{U_k}$ =1/n $\sum_{i=1}^{n} U_{k_i}$ . Since m(t) = $\sum_{k=1}^{\infty}(a_k e_k\ (t)$, estimation of m can be carried out by estimating $\{a_k\}$. For the estimation purpose the above Radom Nikodym derivative can be used as the likelihood function . Berder (1987) observed that the above likelihood function is unbounced in $\{a_k\}$ and a direct maximum likelihood method cannot be adopted. He introduced a sieve based on orthogonal projection to derive a consistant estimator Subramanyam, A and U. N. NaikNimbalkar (1990) has shown that $\sum_{k=1}^{\infty}(w_k n(\overline{U_k} - a_k)e_k$ in an optimal estimating function for estimating the mean function m . Solving the equation

$$\sum_{k=1}^{\infty}(w_k n\ (\overline{u_k} - a_k)e_k = 0$$

We get $\widehat{a_k} = \overline{U_k}$, k=1,2,…

**Theorem 3.1**

The estimating sequence $\{\widehat{a_k} = \overline{U_k}\}$ is an optimal estimating sequence

Proof

**Since E ($U_{k_i}$) = $a_k$**

**$E(\overline{U_k}) = \frac{1}{n}\sum^{n}_{i=1} U_{k_i} = a_k$**

thus $\{\overline{U_k}\}$ is an unbiased estimating sequence. Again since $U_{k1}$,………$U_{kn}$ and iid observations from N($a_k$,1) further ( $U_{k1}$,………$U_{kn}$) and $U_{l1}$,………$U_{ln}$ are independent. Therefore $\overline{U_k}$ is a complete sufficient statistic for $a_k$ and hence it is MVDE for $a_k$. Thus we consider that $\{\overline{U_k}\}$is marginally optimal for $\{a_k\}$. This proves the theorem.

## IV. Conclusion

A limitation for the estimating sequence is that usually it can not be directly used as an estimator of the infinite dimensional parameter (Anilkumar (1994), Subramanian & Naik Nimbalkar(1990). Some modifications is to be made on the estimator so that range of the modified estimator falls in the parameter space. Method of sieves is one such approach (Beder (1988, 1989). Sometimes a Bayes procedure is useful in picking up a suitable estimator (Anilkumar (1994)). But in such modification, one has to sacrifice the optimality property enjoyed by the estimating sequence and has to be satisfied by asymptotic properties. However in all such situations optimal estimating sequence is a right point to start.

## References


[1]. Anilkumar, p. (1994)on estimating the mean function of a Gaussian process. Statistics and prob. Letters 19 (1994) 77-84

[2]. Ash, R.B. and Gardner, M.F(1975) Topics in Stochastic process, Academic press , New York

[3]. Basavu, I.V, Prakasa Rao,B.L.V (1987) Statistical inference for stochastic process .Academic press, New York

[4]. Beder, J. H. (1987) A Sieve estimator for the mean of a Gaussian process . Ann. Statist 15, 59-78

[5]. Beder, J. H. (1988) A Sieve estimator for the covariance of a Gaussian process. Ann statist 16 No: 2,648-660

[6]. Chatterji,S.P. and Handrekar,V.(1978) equivalence and singularity of a Gaussian measure and applications. In probabilistic analysis and related topics. Vol.1 Ed. Bharucha – Reid A.T

[7]. Geman and Hwang, C.R (1982) Non – Parametric Maximum Likelihood Estimation by the method of sieves Annals. Statist. 10,401-414

[8]. Grenander,U (1981) Abstract inference John wiley and sons

[9]. Kale, B.K. (1999) A First Course on Parametric Inference Narosa Publishing House, Delhi

[10]. Leskow,J and Rozanski (1989) Histogram maximum likelihood estimation for the multiplicative intencity models. Stochastic process and their applications, 31, 151-159

[11]. Mckeague, I (1986) Estimation of a martingale regression using method of sieves Annals of Statistics. 14, 579-589.

[12]. Subramanyam, A and U. N. NaikNimbalkar (1990) Optimal unbiased estimating functions for the Hilbert space valued parameters