STUDIA UNIV. "BABEŞ–BOLYAI", MATHEMATICA, Volume **LII**, Number 4, December 2007

THE ORTHOGONAL PRINCIPLE AND CONDITIONAL DENSITIES

ION MIHOC AND CRISTINA IOANA FĂTU

Dedicated to Professor D. D. Stancu on his 80^{th} birthday

Abstract. Let $X, Y \in L^2(\Omega, K, P)$ be a pair of random variables, where $L^2(\Omega, K, P)$ is the space of random variables with finite second moments. If we suppose that X is an observable random variable but Y is not, than we wish to estimate the unobservable component Y from the knowledge of observations of X. Thus, if $g = g(x)$ is a Borel function and if the random variable $g(X)$ is an estimator of Y, then $e = E\{[Y - g(X)]^2\}$ is the mean -square error of this estimator. Also, if $\widehat{g}(X)$ is an optimal estimator (in the mean-square sense) of Y, then we have the following relation $e_{\min} = e(Y, \widehat{g}(X)) = E\{[Y - \widehat{g}(X)]^2\} = \inf_g E\{[Y - g(X)]^2\}$, where inf is taken over all Borel functions $g = g(x)$. In this paper we shall present some results relative to the mean-square estimation, conditional expectations and conditional densities.

1. **Convergence in the mean-square**

Let (Ω, K, P) be a probability space and $\mathcal{F}(\Omega, K, P)$ the family of all random variables defined on (Ω, K, P). Let

$$L^p = L^p(\Omega, K, P) = \{X \in \mathcal{F}(\Omega, K, P) \mid E(|X|^p) < \infty\}, p \in \mathbb{N}^* \tag{1.1}$$

be the set of random variables with finite moments of order p, that is

$$\beta_p = E(|X|^p) = \int_{\mathbb{R}} |x|^p \, dF(x) < \infty, p \in \mathbb{N}^*, \tag{1.2}$$

Received by the editors: 13.03.2007.

2000 *Mathematics Subject Classification.* 62H10, 62H12.

Key words and phrases. estimation, mean-square estimation, conditional means, orthogonality principle, conditional densities.

where

$$F(x) = P(X < x), x \in \mathbb{R} \tag{1.3}$$

is the distribution function of the random variable X.

This set $L^p(\Omega, K, P)$ represent a linear space. An important role among the spaces $L^p = L^p(\Omega, K, P), p \geq 1$, is played by the space $L^2 = L^2(\Omega, K, P)$, the space of random variables with finite second moments.

Definition 1.1. If $X, Y \in L^2(\Omega, K, P)$, then the distance in mean square between X and Y, denoted by $d_2(X, Y)$, is defined by the equality

$$d_2(X, Y) = \|X - Y\| = [E(|X - Y|^2)]^{1/2}. \tag{1.4}$$

Remark 1.1. It is easy to verify that $d_2(X, Y)$ represents a semi-metric on the linear space L^2.

Definition 1.2. If $(X, X_n, n \geq 1) \subset L^2(\Omega, K, P)$, then about the sequence $(X_n)_{n \in \mathbb{N}^*}$ is said to converge to X in mean square (converge in L^2) if

$$\begin{aligned} \underset{n\to\infty}{lim} d_2(X_n, X) = \underset{n\to\infty}{lim} E(|X_n - X|^2)^{1/2} = \\ = \underset{n\to\infty}{lim} E(|X_n - X|^2) = 0. \end{aligned} \tag{1.5}$$

We write $l.i.m.X_n = X$ *or* $X_n \overset{m.p.}{\longrightarrow} X, n \to \infty$, and call X the limit in the mean (or mean square limit) of X_n.

Remark 1.2. If $X \in L^2(\Omega, K, P)$, then

$$Var(X) = E[(X - m)^2] = E[|X - m|^2] = \|X - m\|^2 = d_2^2(X, m),$$

where $m = E(X)$.

Consider two random variables X and Y. Suppose that only X can be observed. If X and Y are correlated, we may expect that knowing the value of X allows us to make some inference about the value of the unobserved variable Y. In this case an interesting problem, namely that of estimating one random variable with another or one random vector with another. If we consider any function $\widehat{X} = g(X)$ on X, then that is called an estimator for Y.

Definition 1.3. We say that a function $X^* = g^*(X)$ on X is best estimator in the mean-square sense if

$$E\{[Y - X^*]^2\} = E\{[Y - g^*(X)]^2\} = \inf_g E\{[Y - g(X)]^2\}. \tag{1.6}$$

If $X \in L^2(\Omega, K, P)$ then a very simple but basic problem consists in: find a constant a (i.e. the constant random variable $a, a \in L^2(\Omega, K, P)$) such that the mean-square error

$$e = e(X; a) = E[(X - a)^2] = \int_{\mathbb{R}} (x - a)^2 dF(x) =$$
$$= \|X - a\|^2 = d_2^2(X, a) \tag{1.7}$$

is minimum.

Evidently, the solution of a such problem is the following: if $a = E(X)$ then the mean-square error is minimum and we have

$$\min_{a \in \mathbb{R}} E[(X - a)^2] = Var(X).$$

Theorem 1.1. ([1]) (The orthogonality principle) Let X, Y be two random variables such that $E(X) = 0$, $E(Y) = 0$ and $\widehat{X}$ a new random variable, $\widehat{X} \in L^2(\Omega, K, P)$, defined as

$$\widehat{X} = g(X) = a_0 X, \ a_0 \in R. \tag{1.8}$$

The real constant a_0 that minimize the mean-square error

$$E[(Y - \widehat{X})^2] = E[(Y - a_0 X)^2] \tag{1.9}$$

is such that the random variable $Y - a_0 X$ is orthogonal to X; that is,

$$E[(Y - a_0 X)X] = 0 \tag{1.10}$$

and the minimum mean-square error is given by

$$e_{min}(Y, \widehat{X}) = e_{\min} = E[(Y - a_0 X)Y], \tag{1.11}$$

where

$$a_0 = \frac{E(XY)}{E(X^2)} = \frac{cov(X, Y)}{\sigma_1^2} \tag{1.12}$$

2. General mean-square estimation

Let us now remove the constraints of linear estimator and consider the more general problem of estimating Y with a (possibly nonlinear) function of X. For this, we recall the notion of inner (scalar) product.

Thus, if X and $Y \in L^2\ (\Omega, K, P)$, we put

$$(X, Y) = E(XY). \tag{2.1}$$

It is clear that if $X, Y, Z \in L^2\ (\Omega, K, P)$ then

$$\begin{cases} (aX + bY, Z) & = & a(X, Z) + b(Y, Z), \quad a, b \in \mathbb{R}, \\ (X, X) & \geq & 0, \\ (X, X) & = & 0 \Longleftrightarrow X = 0, \text{a.s.} \end{cases} \tag{2.2}$$

Consequently (X, Y) is a scalar product. The space $L^2\ (\Omega, K, P)$ is complete with respect to the norm

$$\| X \| = (X, X)^{1/2} \tag{2.3}$$

induced by this scalar product. In accordance with the terminology of functional analysis, a space with the scalar product (2.1) is a Hilbert space.

Hibert space methods are extensively used in the probability theory to study proprieties that depend only on the first two moments of random variables.

In the next, we want to estimate the random variable Y by a suitable function $g(X)$ of X so that the mean-square estimation error

$$e = e(Y, g(X)) = E\left\{[(Y - g(X)]^2\right\} = \iint\limits_{\mathbb{R}^2} [y - g(x)]^2 f(x, y) dx dy \tag{2.4}$$

is minimum.

Theorem 2.1. ([3]) Let $\widehat{X}$ be a random variable defined as a nonlinear function of X, namely

$$\widehat{X} = g(X) \tag{2.5}$$

where $g(x)$ represents the value of this random variable $g(X)$ in the point x, $x \in D_x = \{x \in R \mid f(x) > 0\}$.Then, the minimum value of the mean-square error, namely,

$$e_{\min} = e_{\min}(Y, \widehat{X}) = E\left\{[(Y - E(Y \mid X)]^2\right\} \tag{2.6}$$

is obtained if

$$g(X) = E(Y \mid X), \tag{2.7}$$

where

$$\begin{aligned} E(Y \mid X = x) = E(Y \mid x) = \\ = \int_{-\infty}^{\infty} y f(y \mid x) dy \end{aligned} \tag{2.8}$$

is the random variable defined by the conditional expectation of Y with respect to X.

Definition 2.1. We say that the estimator (the nonlinear function)

$$\widehat{X} = g(X) = E(Y \mid X) \tag{2.9}$$

is best (optimal) in the mean-square sense for the unknown random variable Y if

$$\begin{aligned} e_{\min}(Y, \widehat{X}) = \min_{g(X)} E\left\{[(Y - g(X)]^2\right\} = \\ = E\left\{[(Y - E(Y \mid X)]^2\right\}. \end{aligned} \tag{2.10}$$

Lemma 2.1. ([1]) If X and Y are two independent random variable, then

$$E(Y \mid X) = E(Y). \tag{2.11}$$

Corollary 2.1. If X, Y are two independent random variables then the best mean-square estimator of Y in terms of X is $E(Y)$. Thus knowledge of X does not help in the estimation of Y.

3. **Conditional expectation and conditional densities**

We assume that the random vector (X, Y) have the bivariate normal distribution with the probability density function

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}e^{-\frac{1}{2(1-r^2)}\left[\left(\frac{x-m_1}{\sigma_1}\right)^2-\frac{2r(x-m_1)(y-m_2)}{\sigma_1\sigma_2}+\left(\frac{y-m_2}{\sigma_2}\right)^2\right]}, \tag{3.1}$$

where:

$$m_1=E(X)\in\mathbb{R}, m_2=E(Y)\in\mathbb{R}, \sigma_1^2=Var(X)>0, \sigma_2^2=Var(Y)>0, \tag{3.1a}$$

$$r=r(X,Y)=\frac{cov(X,Y)}{\sigma_1\sigma_2},\quad r\in(-1,1), \tag{3.2}$$

r being the correlation coefficient between X and Y.

First, we will recall some very important definitions and proprieties for a such normal distribution.

Lemma 3.1. If two jointly normal random variable X and Y are uncorrelated, that is, $cov(X, Y) = 0 = r(x, y)$, then they are independent and we have

$$f(x,y)=f(x;m_1,\sigma_1^2)f(y;m_2,\sigma_2), \tag{3.3}$$

where

$$f(x;m_1,\sigma_1^2)=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{1}{2}\left(\frac{x-m_1}{\sigma_1}\right)^2},\ f(y;m_2,\sigma_2^2)=\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{1}{2}\left(\frac{y-m_2}{\sigma_2}\right)^2} \tag{3.3a}$$

are the marginal probability density functions for the components X and Y of the normal random vector (X, Y).

Lemma 3.2. If (X, Y) is a random vector with the bivariate normal probability density function (3.1), then for the conditional random variable $(Y \mid X)$, for example, the probability density function, denoted by $f(y \mid x)$, has the form

$$f(y\mid x)=\frac{1}{\sqrt{2\pi(1-r^2)}\sigma_2}e^{-\frac{1}{2\sigma_2^2(1-r^2)}\left[y-\left(m_2+r\frac{\sigma_2}{\sigma_1}(x-m_1)\right)\right]^2}, \tag{3.4}$$

This conditional probability density function (3.4) may be obtained using the well-bred method which have in view the following relations

$$f(y\mid x)=\frac{f(x,y)}{f(x)}, f(x)>0,\ f(x)=f(x;m_1,\sigma_1^2)=\int\limits_{-\infty}^{\infty}f(x,y)dy. \tag{3.5}$$

In the next, we shall recover this conditional probability density function using the orthogonality principle.

Theorem 3.1. Let (X, Y) be a normal random vector which is characterized by the relations (3.1), (3.1a) and (3.2). If

$$\overset{o}{X} = X - m_1, \overset{o}{Y} = Y - m_2, \tag{3.6}$$

are the deviation random variables and U is a new random variable which is defined as

$$U = \overset{o}{Y} - c_0 \overset{o}{X}, \ where\ c_0 \in \mathbb{R} - \{0\}, \tag{3.7}$$

then the orthogonality principle implies the conditional density function (3.4), which corresponds to the conditional random variable $(\overset{o}{Y} \mid \overset{o}{X})$, and more we have the following relation

$$f(y \mid x) = f(u), \tag{3.8}$$

where $f(u)$ is the probability density function that corresponds to U.

Proof. Indeed, because

$$\begin{cases} E(\overset{o}{X}) = m_{\overset{0}{x}} &= 0, \quad Var(\overset{o}{X}) = \sigma^2_{\overset{0}{x}} &= Var(X) = \sigma_1^2, \\ E(\overset{o}{Y}) = m_{\overset{0}{y}} &= 0, \quad Var(\overset{o}{Y}) = \sigma^2_{\overset{0}{y}} &= Var(Y) = \sigma_2^2, \end{cases} \tag{3.9}$$

and

$$cov(\overset{o}{X}, \overset{o}{Y}) = E(\overset{o}{X}\overset{o}{Y}) = E[(X - m_1)(Y - m_2)] = cov(X, Y) = r\sigma_1\sigma_2, \tag{3.10}$$

then we obtain

$$E(U) = m_U = 0. \tag{3.11}$$

Also, for the variance of the random variable U, we obtain

$$\begin{aligned} Var(U) = \sigma_U^2 = E\left\{[U - E(U)]^2\right\} = E(U^2) = \\ = E\{[\overset{o}{Y} - c_0\overset{o}{X}]^2\} = \\ = \sigma_2^2 - 2c_0 cov(X, Y) + c_0^2\sigma_1^2 = \\ = \sigma_2^2 - 2c_0 r\sigma_1\sigma_2 + c_0^2\sigma_1^2, \end{aligned}$$

The value of the constant c_0 will be determined using the orthogonality principle, namely: the random variables U and $\overset{\circ}{X}$ to be orthogonal. This condition implies the following relation

$$E(U\overset{o}{X}) = E\left[(\overset{o}{Y} - c_0\overset{o}{X}) \mid \overset{o}{X})\right] = 0, \tag{3.12}$$

and, more, the constant c_0 must to minimize the mean-square error

$$e = E[(\overset{o}{Y} - c_0\overset{o}{X})^2], \tag{3.13}$$

that is,

$$e_{\min} = E[(\overset{o}{Y} - c_0\overset{o}{X})\overset{o}{Y}]. \tag{3.14}$$

Indeed, using (1.12) we obtain the following value

$$c_0 = \frac{E(\overset{o}{X}\overset{o}{Y})}{E(\overset{o}{X^2})} = r\frac{\sigma_2}{\sigma_1}, \tag{3.15}$$

if we have in view the relations (3.9) and (3.10).

Also, from (3.12), we obtain

$$cov(U, \overset{o}{X}) = E(U\overset{o}{X}) = 0, \rho(U, \overset{o}{X}) = 0, \tag{3.16}$$

where $\rho(U, \overset{o}{X})$ represents the correlation coefficient between the random variables U and $\overset{o}{X}$.

Because the random variables U and $\overset{o}{X}$ are normal distributed with $\rho(U, \overset{o}{X}) = 0$ then, using the Lemma 3.1, it follows that these random variables are independent and their joint probability density function, denoted by $f(\overset{0}{x}, u)$, has the form

$$f(\overset{o}{x}, u) = f(\overset{o}{x})f(u), \tag{3.17}$$

where $f(\overset{o}{x})$ is the probability density function for the random variable $\overset{o}{X}$, that is,

$$f(\overset{o}{x}) = \frac{1}{\sqrt{2\pi}\sigma_{\overset{o}{x}}} e^{-\frac{1}{2}\left[\frac{\overset{o}{x} - m_{\overset{o}{x}}}{\sigma_{\overset{0}{x}}}\right]^2} = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{2}\left(\frac{x-m_1}{\sigma_1}\right)^2} =$$
$$= f(x; m_1\sigma_1^2), \ x \in \mathbb{R}, \tag{3.18}$$

if we have in view the relations (3.6) and (3.9).

Also, for the probability density function $f(u)$, we obtain the following forms

$$\begin{aligned} f(u) &= \frac{1}{\sqrt{2\pi}\sigma_u} e^{-\frac{1}{2}\left[\frac{u-m_u}{\sigma_u}\right]^2} = \\ &= \frac{1}{\sigma_2\sqrt{2\pi(1-r^2)}} e^{-\frac{1}{2(1-r^2)}\left(\frac{u}{\sigma_2}\right)^2} = \\ &= \frac{1}{\sigma_2\sqrt{2\pi(1-r^2)}} e^{-\frac{1}{2\sigma_2^2(1-r^2)}\left\{y-\left[m_2+r\frac{\sigma_2}{\sigma_1}(x-m_1)\right]\right\}^2} \\ &= \frac{1}{\sigma_2\sqrt{2\pi(1-r^2)}} e^{-\frac{1}{2\sigma_2^2(1-r^2)}(y-m_{y|x})^2}, \end{aligned} \tag{3.19}$$

if we have in view the relations (3.6) and (3.11) as well as the fact that the values of the random variable $U = \overset{o}{Y} - c_0\overset{o}{X}$ can be express as

$$u = \overset{o}{y} - c_0\overset{o}{x} = y - [m_2 + r\frac{\sigma_2}{\sigma_1}(x - m_1)] = y - m_{Y|x}. \tag{3.20}$$

Therefore, the form (3.1a) of the probability density function $f(u)$, together with the relation (3.4), give us just the relation (3.8), that is, we obtain the following equality

$$f(u) = f(y \mid x) = \frac{1}{\sigma_2\sqrt{2\pi(1-r^2)}} e^{-\frac{1}{2\sigma_2^2(1-r^2)}(y-m_{y|x})^2}. \tag{3.21}$$

Utilizing the forms (3.18) and (3.21) of the probability density functions $f(\overset{o}{x})$ and $f(u)$, from the relation (3.17), we obtain the following expressions

$$f(\overset{o}{x}, u) = f(\overset{o}{x})f(u) = f(x; m_1\sigma_1^2)f(y \mid x) = \tag{3.22}$$

$$\begin{aligned} &= \left[\frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{2}\frac{(x-m_1)^2}{\sigma_1^2}}\right] \left[\frac{1}{\sigma_2\sqrt{2\pi(1-r^2)}} e^{-\frac{1}{2\sigma_2^2(1-r^2)}\left\{y-\left[m_2+r\frac{\sigma_2}{\sigma_1}(x-m_1)\right]\right\}^2}\right] = \\ &= \left[\frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{2}\frac{(x-m_1)^2}{\sigma_1^2}}\right] \left[\frac{1}{\sigma_2\sqrt{2\pi(1-r^2)}} e^{-\frac{1}{2(1-r^2)}\left[\frac{y-m_2}{\sigma_2}-r\frac{x-m_1}{\sigma_1}\right]^2}\right] = \\ &= \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{1}{2(1-r^2)}\left[\frac{(x-m_1)^2}{\sigma_1^2}-2r\frac{(x-m_1)(y-m_2)}{\sigma_1\sigma_2}+\frac{(y-m_2)^2}{\sigma_2^2}\right]^2} = f(x, y), \end{aligned}$$

and, hence, it follows the equality

$$f(x, y) = f(x)f(y \mid x). \tag{3.23}$$

In the next, we must to prove that the minimum of the mean-square error, specified in the relation (3.14), can be obtained if the constant c_0 has the value (3.15).

But, in the beginning, we recall some definitions and some properties of the conditional means.

Lemma 3.3. ([1]) The conditional mean $E(. \mid X)$ is a linear operator, that is,

$$E(cY + dZ \mid X) = cE(Y \mid X) + dE(Z \mid X), c, d \in R. \tag{3.24}$$

Definition 3.1. If (X, Y) is a bivariate random vector with the probability density function $f(x, y)$ and $Z = g(X, Y)$ is a new random variable which is a function of the random variables X and Y, then the conditional mean of the random variable $Z = g(X, Y)$, given $X = x$, is defined as

$$E[g(X, Y) \mid X = x] = \int_{-\infty}^{\infty} g(x, y) f(y \mid X = x) dy, \tag{3.25}$$

for any $x \in D_x = \{x \in R \mid f(x) > 0\}$.

Lemma 3.4. ([1]) If the random variable Z has the form

$$Z = g(X, Y) = g_1(X) g_2(Y), \tag{3.26}$$

then we have the following relation

$$E[g_1(X) g_2(Y) \mid X] = g_1(X) E[g_2(Y) \mid X]. \tag{3.27}$$

Lemma 3.5. ([1]) If X is a random variable and c is a real constant, then

$$E[c \mid X] = c. \tag{3.28}$$

Now, we can return to the our problem, namely to prove that the minimum of the mean-square error, specified in the relation (3.14), can be obtained if the constant c_0 has the value (3.15).

Thus, because the random variables $U = \overset{0}{Y} - c_0\overset{0}{X}$ and $\overset{0}{X}$ are independent, then from (3.12) and Lemma 2.1, we obtain

$$\begin{aligned} E(U\overset{0}{X}) &= E[(\overset{0}{Y} - c_0\overset{0}{X}) \mid \overset{0}{X}] = \\ &= E[(\overset{0}{Y} - c_0\overset{0}{X})] = \\ &= E(\overset{0}{Y}) - c_0 E(\overset{0}{X}) = 0, \end{aligned}$$

that is, we have the following equality

$$E(U\overset{o}{X}) = E(\overset{o}{Y}) - c_0 E(\overset{o}{X}) = 0. \tag{3.29}$$

On the other hand, in accordance with the Lemma 3.4, (respectively, in accordance with the relation (3.26)) and the Lemma 3.5, where $g_1(X) = c_0\overset{o}{X}$ and $g_2(Y) = 1$, we obtain

$$E[c_0\overset{o}{X} \mid \overset{o}{X} = \overset{o}{x}] = c_0\overset{o}{X}\underbrace{E[1 \mid \overset{o}{X} = \overset{o}{x}]}_{=1} = c_0\overset{o}{X}, \tag{3.30}$$

for any $\overset{o}{x} = x - m_1, x \in \mathbb{R}$

This last relation, together with the Lemma 3.3 give us the possibility to rewritten the conditional mean $E(U\overset{o}{X}) = E[(\overset{o}{Y} - c_0\overset{o}{X}) \mid \overset{o}{X}]$ in an useful form

$$\begin{aligned} E(U\overset{o}{X}) &= E[(\overset{o}{Y} - c_0\overset{o}{X}) \mid \overset{o}{X}] = \\ &= E(\overset{o}{Y} \mid \overset{o}{X}) - E(c_0\overset{o}{X} \mid \overset{o}{X}) = \\ &= E(\overset{o}{Y} \mid \overset{o}{X}) - c_0\overset{o}{X}, \end{aligned}$$

that is,

$$E(U\overset{o}{X}) = E(\overset{o}{Y} \mid \overset{o}{X}) - c_0\overset{o}{X}. \tag{3.31}$$

From (3.29) and (3.31), we obtain the random variable

$$E(\overset{o}{Y} \mid \overset{o}{X}) = c_0\overset{o}{X} = r\frac{\sigma_2}{\sigma_1}\overset{o}{X}, \tag{3.32}$$

which has the real values of the form

$$E(\overset{o}{Y} \mid \overset{o}{X} = \overset{o}{x}) = r\frac{\sigma_2}{\sigma_1}\overset{o}{x}, \text{ for any } \overset{o}{x} = x - m_1, x \in \mathbb{R}. \tag{3.32a}$$

The conditional variance of the random variable $(\overset{o}{Y} \mid \overset{o}{X})$ can be express as

$$
\begin{aligned}
Var(\overset{o}{Y} \mid \overset{o}{X}) = \sigma^2_{\overset{o}{Y}\mid\overset{o}{X}} = \\
= E\{[\overset{o}{Y} - E(\overset{o}{Y} \mid \overset{o}{X})]]^2 \mid \overset{o}{X}\} = \\
= E[(\overset{o}{Y} - c_0\overset{o}{X})^2 \mid \overset{o}{X}],
\end{aligned} \tag{3.33}
$$

and, evidently, it is a random variable which has the real values of the form

$$
Var(\overset{o}{Y} \mid \overset{o}{X} = \overset{o}{x}) = E[(\overset{o}{Y} - c_0\overset{o}{X})^2 \mid \overset{o}{X} = \overset{o}{x}],\ for\ any \overset{o}{x} = x - m_1, x \in \mathbb{R}. \tag{3.33a}
$$

Because the random variables $U = \overset{o}{Y} - c_0\overset{o}{X}$ and $\overset{o}{X}$ are independent then, evidently, it follows that and the random variable $U^2 = (\overset{o}{Y} - c_0\overset{o}{X})^2$ and $\overset{o}{X}$ are independent. Then, from (3.36), we obtain

$$
\begin{aligned}
Var(\overset{o}{Y} \mid \overset{o}{X}) = E[(\overset{o}{Y} - c_0\overset{o}{X})^2 \mid \overset{o}{X}] = \qquad (3.34) \\
= E[(\overset{o}{Y} - c_0\overset{o}{X})^2] = \\
= E[(\overset{o}{Y} - c_0\overset{o}{X})\overset{o}{Y} + c_0(\overset{o}{Y} - c_0\overset{o}{X})\overset{o}{X}] = \\
= E[(\overset{o}{Y} - c_0\overset{o}{X})\overset{o}{Y}] + c_0\underbrace{E[(\overset{o}{Y} - c_0\overset{o}{X})\overset{o}{X}]}_{=0\text{ (see, (3.14))}} = \\
= \underbrace{E[(\overset{o}{Y} - c_0\overset{o}{X})\overset{o}{Y}]}_{\text{(see, (3.16))}} = \\
= e_{\min} = e_{\min}(\overset{o}{Y}, \overset{o}{X}) = \\
= E(\overset{o}{Y^2}) - 2c_0E(\overset{o}{Y}\overset{o}{X}) + c_0^2E(\overset{o}{X^2}) = \\
= \sigma_2^2 - r\frac{\sigma_2}{\sigma_1}r\sigma_1\sigma_2 = \sigma_2^2(1 - r^2). \qquad (3.34a)
\end{aligned}
$$

Therefore, the conditional variance of the deviation random variable $\overset{o}{Y}$, given $\overset{o}{X}$, represents just the minimum mean-square error.

References

[1] Fătu, C.I., *Metode optimale în teoria estimării statistice*, Editura RISOPRINT, Cluj-Napoca, 2005.

[2] Mihoc, I., Fătu, C.I., *Calculul probabilitaţilor şi statistică matematică,* Casa de Editură Transilvania Pres, Cluj-Napoca, 2003.

[3] Mihoc, I., Fătu, C.I., *Mean-square estimation and conditional densities,* in The 6th Romanian-German Seminar on Approximation Theory and its Applications (RoGer 2004), Mathematical Analysis and Approximation Theory, MediamiraScience Publisher, 2005, pp.147-159.

[4] Papoulis, A., *Probability, Random variables and Stochastic Processes*, McGraw-Hill Book Company, New York, London, Sydney, 1965.

[5] Rao, C.R., *Linear Statistical Inference and Its Applications*, John Wiley and Sons, Inc., New York, 1965.

[6] Rényi, A., *Probability Theory,* Akadémiai Kiado, Budapest, 1970.

[7] Shiryaev, A.N., *Probability,* Springer-Verlag, New York, Berlin, 1996.

[8] Wilks, S.S., *Mathematical Statistics,* Wiley, New York, 1962.

Faculty of Economics at Christian University "Dimitrie Cantemir",
Cluj-Napoca, Romania
E-mail address: `imihoc@cantemir.cluj.astral.ro`