ANNALES
UNIVERSITATIS MARIAE CURIE-SKŁODOWSKA
LUBLIN—POLONIA
VOL. XI, 2 SECTIO A 1957

Z Zakładu Statystyki Matematycznej Wydziału Rolnego Wyższej Szkoły Rolniczej w Lublinie
Kierownik: z-ca prof. mgr Wiktor Oktaba

WIKTOR OKTABA

On the Linear Hypothesis in the Theory of Normal Regression

O hipotezie liniowej w teorii normalnej regresji

О линейной гипотезе в теории нормальной регрессии

1. Problem

Some problems of the testing of hypotheses in the theory of normal regression have been already discussed by "Student" [22] and A. A. Markoff [9] at the beginning of the present century. R. A. Fisher [1, 2], J. Neyman and E. S. Pearson [11, 12], J. Neyman [13], St. Kołodziejczyk [7], C. R. Rao [21] and many others have also dealt with these problems.

The assumptions connected with the problem of the testing of linear hypothesis in the theory of normal regression are the following.

Uncorrelated random variables

$$y_\alpha = \mu_\alpha + e_\alpha \tag{1}$$

with expected values

$$\mu_\alpha = E(y_\alpha) = x_{\alpha 1}\beta_1 + x_{\alpha 2}\beta_2 + \cdots + x_{\alpha p}\beta_p \tag{2}$$

($\alpha = 1, 2, \ldots, n$) depending on p ($p < n$) parameters $\beta_1, \beta_2, \ldots, \beta_p$ are normally distributed with the common variance σ^2. The symbols e_α denote residuals, and $x_{\alpha j}$ ($j = 1, 2, \ldots, p$) denote elements of given matrix X with n rows and p columns, which is marked by $X = \underset{np}{X}$. The g independent linear restrictions denoted by matrix relation

$$G\beta = \eta, \tag{3}$$

where $G = \underset{gp}{G}$ is the given matrix and η is the given column vector with components $\eta_1, \eta_2, \ldots, \eta_g$, are imposed on the unknown parameters $\beta_1, \beta_2, \ldots, \beta_p$ being the components of column vector β.

The problem consists in testing of linear hypothesis

$$H\beta = \nu \tag{4}$$

that h of given and independent parameter functions $H\beta$ which are linearly independent of (3) (where matrix $H = \underset{hp}{H}$, $h + g < p$, $h > 0$, $g \geqslant 0$) have certain values $\nu_1, \nu_2, \ldots, \nu_h$ that are components of vector ν.

This problem, with the omission of the restriction (3) and on the assumption that the rank r of matrix X is equal to p, has been dealt with and solved by St. Kołodziejczyk [7] by means of the test of significance (cf. sec. 3 in [7]) which he obtained as a likelihood ratio test λ, following a procedure due to J. Neyman and E. S. Pearson ([14], [15], [16] and [23]).

The general solution for $r \leqslant p$ is due to C. R. Rao. In the papers [20] and [21] he proved the theorem that under conditions (1), (2), (3) and (4) random variable $\sigma^{-2} \cdot Min \sum_{i=1}^{n} e_i^2$ (cf. (1)) where minimalization is performed with respect to parameters $\beta_1, \beta_2, \ldots, \beta_p$, is distributed as χ^2 (chi-square) with an appropriate number of degrees of freedom. Giving two proofs of this theorem C. R. Rao under the conditions (1)-(4) deduced an expression for the random variable F of the form similar to that in the theorem 1 of the present paper (cf. [19]). Since this random variable may be used for testing the linear hypothesis (4) in the model defined by the relations (1)-(3), when $r \leqslant p$, it provides a general solution for the discussed problem.

The other above mentionned authors mainly give different ways of testing linear hypothesis.

R. A. Fisher assigns for the model expressed by relations (1) and (2) the Student's t test for testing hypothesis that regression coefficient in population, β_i, assumes the given value β_i^0 $(i = 1, 2, \ldots, p)$.

As far as the further course of research in normal multiple regression is concerned, the corresponding sections in textbooks by R. A. Fisher [3], A. M. Mood [10], S. S. Wilks [25], H. B. Mann [8], C. R. Rao [19] and O. Kempthorne [6] deserve attention.

The survey of the methods of proofs applied in the above publications permits us to see certain gaps in the theory of normal multiple regression. To fill these gaps at least partly is the task of the present paper.

This paper deals with some problems connected with testing linear hypothesis in the theory of normal regression in the case of C. R. Rao's model [cf. (1), (2), (3) and (4)] restricted to the condition $r(X) = p$.

The solution of problem as defined on p. 18 in a general form and expressed in the present paper in the theorem 1 has been reached by generalizing H. B. Mann's lemma 4.1 and also by generalizing the theorem 4.1 of the same author (cf. [8]). Though C. R. Rao's results includes the theorem 1, it is given here another proof. This is justified by the fact that theorem 1 presents a general method of determining the form of random variable F (on the assumption that the linear hypothesis is true) for any linear regression model with restrictions given in (3).

Apart from the general problem we are primarily concerned with eliminating the minimalization marked in the formula for F in the theorem 1 and with presenting in the explicit form the expressions for the random variables which are to be deduced from this formula in connection with different types of multiple regression models under different variants of the linear hypothesis. The expressions so obtained are given in theorems 3, 4 and 5. To find these expressions [cf. (56), (57) and (58)] we express the known theorem [25] (cf. the theorem 2 of the present paper) in the matrix form [cf. (47) and (48)], while proving this theorem by the matrix calculus. It should be noted that the mentionned minimalization is unavoidable, if the theorem 1 is applied directly.

The random variables F given in the theorems 2, 3, 4 and 5 may be used for testing corresponding linear hypothesis (cf. sec. 6 and in particular (120)).

To show that the tests of significance based on random variables appearing in theorems 2, 3, 4 and 5 are right-tailed, a number of expected values for corresponding quadratic forms have been worked out. Incidentally other matrix relations have been found, such as some identities with conditional and unconditional estimates of parameters, related to multiple regression model, covariance matrix for linear combinations of parameter estimates, and other relations presented in sections 3 and 5.

Because of the simplicity of formulae and ease of operations with matrices we use matrix calculus and deal in particular with submatrices [4].

All throughout this paper the equality of variances for all random variables y_i is assumed. Some types of hypotheses met in the problems of regression analysis for any number of samples in the case of nonhomogeneous variances have been considered in W e l c h's investigations [24]. In this paper however this case is not considered.

We divide vectors into subvectors in the following way:

$$\text{(6)}\qquad \underset{p1}{\beta}=\begin{bmatrix}\beta_1\\ \beta_2\\ \vdots\\ \beta_q\\ \hdashline \beta_{q+1}\\ \beta_{q+2}\\ \vdots\\ \beta_p\end{bmatrix}=\begin{bmatrix}\underset{q1}{\gamma}\\ \hdashline \underset{p-q,1}{\delta}\end{bmatrix},\ \text{where } \gamma=\begin{bmatrix}\beta_1\\ \beta_2\\ \vdots\\ \beta_q\end{bmatrix} \text{ and } \delta=\begin{bmatrix}\beta_{q+1}\\ \beta_{q+2}\\ \vdots\\ \beta_p\end{bmatrix}.$$

For the sake of brevity we introduce the following notation:

$$\text{(7)}\qquad \underset{qq}{A}=X_1^* X_1,\qquad \underset{q,p-q}{B}=X_1^* X_2,\qquad \underset{p-q,p-q}{D}=X_2^* X_2.$$

The reciprocal matrix S^{-1} of the matrix $\underset{pp}{S}=X^* X$ is devided into four submatrices: $S^{11}=\underset{qq}{S^{11}}$, $S^{12}=\underset{q,p-q}{S^{12}}$, $S^{21}=\underset{p-q,q}{S^{21}}$ and $S^{22}=\underset{p-q,p-q}{S^{22}}$, and is represented in the relation:

$$\text{(8)}\qquad S^{-1}=\begin{bmatrix}S^{11} & S^{12}\\ S^{21} & S^{22}\end{bmatrix}$$

The unconditional estimates of parameters $\beta, \gamma, \delta, \ldots$ are denoted $\hat{\beta}, \hat{\gamma}, \hat{\delta}, \ldots$ respectively (strictly speaking: the estimates $\hat{\beta}$, $\hat{\gamma}$, $\hat{\delta}, \ldots$ are vectors, whose components are estimates of the corresponding components of vectors $\beta, \gamma, \delta, \ldots$). To convey that we estimate the parameters $\beta_1, \beta_2, \ldots, \beta_p$ forming the vector β, we say briefly that we estimate parameter β. The symbol $\tilde{\gamma}$ is used to denote the parameter estimate γ assuming that $\delta=\delta_0$, where δ_0 is the known vector; sismilarly $\tilde{\delta}$ denotes conditional estimate of parameter δ when $\gamma=\gamma_0$, where γ_0 is the known vector.

The sample variance for n observations $x_1, x_2, \ldots, x_n$ is denoted by

$$S_x^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

and the sample covariance for $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_n$ by

$$S_{xy}=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}),$$

2. Notation

In the present paper we use in general large italic letters G, H, M, P, R, ... for matrices and small italic letters x, y, z, u, v, δ, ϑ, ϱ, ... for column vectors with exceptions which will be evident from the text. We admit the following definitions and notations. Matrix $P=\{p_{ij}\}$, $i=1,2,\ldots,n$; $j=1,2,\ldots,m$; with n rows and m columns is called matrix of order $n\cdot m$ and is marked with symbol $P=\underset{nm}{P}$. Let the transpose of of a matrix $\underset{nm}{P}$ have corresponding symbol with a star i. e. P^*. The transpose of the column vector $\underset{n1}{y}=\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$ is called the row vector $\underset{1n}{y^*}=$ $=[y_1, y_2, \ldots, y_n]$, and conversely; the small italic letters with indices $y_1, y_2, \ldots, y_n$ denote elements of the vector y. The submatrices are separated by broken line and are written in square brackets i. e.

$$\underset{nm}{H}=\left[\underset{np}{W} \;\vdots\; \underset{n,\,m-p}{G}\right] \quad \text{or} \quad \underset{nm}{H}=\begin{bmatrix} \underset{qm}{P} \\ \hdashline \underset{n-q,\,m}{R} \end{bmatrix}.$$

Symbol E is reserved for expected values and $e=\begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$ for residual vector.

The reciprocal matrix of the matrix $\underset{gg}{P}$ is written, as usual, $P^{-1}=\{p^{ij}\}$; $i,j=1,2,\ldots,g$. The unit matrix and the zero matrix are marked by symbols I and O respectively. Let $r(P)$ denote the rank of matrix P. The determinant of the matrix P is marked by $|P|$. Definite matrices and vectors preserve throughout the paper consistent symbols. We admit convention

$$\mu=E(y)=X\beta,$$

where

$$\underset{n1}{y}=\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \mu=\begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix}, \quad \underset{np}{X}=\left[\begin{array}{cccc:cccc} x_{11} & x_{12} & \ldots & x_{1q} & x_{1,q+1} & x_{1,q+2} & \ldots & x_{1p} \\ x_{21} & x_{22} & \ldots & x_{2q} & x_{2,q+1} & x_{2,q+2} & \ldots & x_{2p} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{n1} & x_{n2} & \ldots & x_{nq} & x_{n,q+1} & x_{n,q+2} & \ldots & x_{np} \end{array}\right]=\left[\underset{nq}{X_1} \;\vdots\; \underset{n,\,p-q}{X_2}\right],$$

X_1 and X_2 are corresponding submatrices of the matrix X, where $r(X_1)=q$, $r(X_2)=p-q$ and $r(X)=p$.

where
$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i \quad \text{and} \quad \bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i;$$

thus the sum of squared deviations from the mean and sum of product deviations are defined by symbols nS_x^2 and nS_{xy} respectively.

Population covariance is written $E(y_i-\mu_i)(y_j-\mu_j)$ or $\operatorname{cov}(y_i, y_j)$, where $\mu_i=E(y_i)$ and $E(y_j)=\mu_j$.

Mean of squares of $x_1, x_2, \ldots, x_n$ is marked by

$$\overline{x^2}=\frac{1}{n}\sum_{i=1}^{n}x_i^2.$$

In the text we shall use the symbols $Q_a = Min\,(e^*e)$ and $Q_r = Min\, e^*e$; in the first case the minimalization is to be performed with respect to $\beta_1, \beta_2, \ldots, \beta_p$ under the restrictions imposed on these parameters, and in the second case the minimalization is to be performed under the same restrictions plus restrictions expressed in the null hypothesis. The minimalization with respect to the parameters $\beta_1, \beta_2, \ldots, \beta_p$ may be replaced by the minimalization with respect to the expected values $\mu_1, \mu_2, \ldots, \mu_n$ (cf. 5°-7°).

The meaning of other symbols is explained in the text.

3. Matrix relations

We deduce some relations which are used in the proofs of theorems presented in the following section. We are concerned with simple multiple regression model:

$$\mu = E(y) = X\beta = X_1\gamma + X_2\delta,$$

where $r(X)=p$, $r(X_1)=q$, $r(X_2)=p-q$ and $X=[X_1 \,\vdots\, X_2]$, μ, β, γ and δ are defined in the precending section.

1°. Let us write matrix $S=X^*X$ in the form including matrices A, B and D (cf. sec. 2). Considering the partition of the matrix X into two submatrices X_1 and X_2 we obtain:

$$S = X^*X = [X_1 \,\vdots\, X_2]^*[X_1 \,\vdots\, X_2] = \begin{bmatrix} X_1^* \\ X_2^* \end{bmatrix}[X_1 \,\vdots\, X_2] = \begin{bmatrix} X_1^*X_1 & X_1^*X_2 \\ X_2^*X_1 & X_2^*X_2 \end{bmatrix}$$

and using the notation introduced in (7) we obtain

$$S = \begin{bmatrix} A & B \\ B^* & D \end{bmatrix}. \tag{9}$$

On account of the identity $SS^{-1} = I$ and of the partition (8) of the matrix S^{-1} into four submatrices S^{11}, S^{12}, S^{21} and S^{22} we have

$$SS^{-1} = \begin{bmatrix} A & B \\ B^* & D \end{bmatrix} \begin{bmatrix} S^{11} & S^{12} \\ S^{21} & S^{22} \end{bmatrix} = \begin{bmatrix} AS^{11} + BS^{21} & AS^{12} + BS^{22} \\ B^* S^{11} + DS^{21} & B^* S^{12} + DS^{22} \end{bmatrix} = \begin{bmatrix} \underset{qq}{I} & \underset{q,\, p-q}{O} \\ \underset{p-q,\, q}{O} & \underset{p-q,\, p-q}{I} \end{bmatrix}$$

and hence from indirect comparison of submatrices we obtain the following four identities:

(10) $$AS^{11} + BS^{21} = I,$$

(11) $$B^* S^{11} + DS^{21} = O,$$

(12) $$AS^{12} + BS^{22} = O,$$

(13) $$B^* S^{12} + DS^{22} = I.$$

2°. The following relations of S^{11}, S^{12}, S^{21}, S^{22}, A, B and D hold:

(14) $$S^{11} = (A - BD^{-1} B^*)^{-1},$$

(15) $$S^{12} = - A^{-1} BS^{22} = - S^{11} BD^{-1},$$

(16) $$S^{21} = - D^{-1} B^* S^{11} = - S^{22} B^* A^{-1},$$

(17) $$S^{22} = (D - B^* A^{-1} B)^{-1}.$$

Proof: From (11) we receive $S^{21} = - D^{-1} B^* S^{11}$. Substituting this into (10) we obtain $AS^{11} - BD^{-1} B^* S^{11} = I$ and hence $(A - BD^{-1} B^*)S^{11} = I$, and S^{11} obtains in the form (14). Similarly using the identity (12) and (13) we prove the relation (17).

Now we shall deduce the formula (15). From identity (12) we obtain $AS^{12} = - BS^{22}$, hence immediately

(18) $$S^{12} = - A^{-1} BS^{22}.$$

Similarly from (11) we find

(19) $$S^{21} = - D^{-1} B^* S^{11}.$$

As the matrix $S = X^*X$ is symmetric, the reciprocal S^{-1} is also symmetric. From this we infer that $S^{12} = (S^{21})^* = (-D^{-1} B^* S^{11})^* = -S^{11} BD^{-1}$, which proves the relation (15). Taking into consideration (18) we obtain $S^{21} = (S^{12})^* = (-A^{-1} BS^{22})^* = -S^{22} B^* A^{-1}$, which jointly with (19) gives the result (16). The proof is concluded.

3°. The parameter estimates γ and δ in the model $E(y) = X_1\gamma + X_2\delta$, when $r(X) = p$, are marked with symbols $\hat{\gamma}$ and $\hat{\delta}$. We express them by submatrices X_1 and X_2 and by y.

It is known that the estimate of parameter β in the model $E(y) = X\beta$, where $r(X) = p$, is $\hat{\beta} = S^{-1} X^* y$ (cf. [6]). Dividing the vector β according to (6) into two subvectors γ and δ we obtain:

$$\hat{\beta} = \begin{bmatrix} \hat{\gamma} \\ \hat{\delta} \end{bmatrix} = S^{-1} X^* y = \begin{bmatrix} S^{11} & S^{12} \\ S^{21} & S^{22} \end{bmatrix} [X_1 \ \vdots \ X_2]^* \cdot y = \begin{bmatrix} S^{11} & S^{12} \\ S^{21} & S^{22} \end{bmatrix} \begin{bmatrix} X_1^* y \\ X_2^* y \end{bmatrix} = \begin{bmatrix} S^{11} X_1^* y + S^{12} X_2^* y \\ S^{21} X_1^* y + S^{22} X_2^* y \end{bmatrix},$$

hence

$$\hat{\gamma} = S^{11} X_1^* y + S^{12} X_2^* y \tag{20}$$

$$\hat{\delta} = S^{21} X_1^* y + S^{22} X_2^* y. \tag{21}$$

Using the relations (14), (15), (16) and (17) the estimates $\hat{\gamma}$ and $\hat{\delta}$ are given in the form:

$$\hat{\gamma} = (A - BD^{-1} B^*)^{-1} (X_1^* y - BD^{-1} X_2^* y), \tag{22}$$

$$\hat{\delta} = (D - B^* A^{-1} B)^{-1} (X_2^* y - B^* A^{-1} X_1^* y). \tag{23}$$

Hence it is evident, that in the case of $B = X_1^* X_2 = O$, i. e. when the parameters γ and δ are orthogonal, the independent estimates of simple form are obtained:

$$\hat{\gamma} = A^{-1} X_1^* y = (X_1^* X_1)^{-1} X_1^* y, \tag{22'}$$

$$\hat{\delta} = D^{-1} X_2^* y = (X_2^* X_2)^{-1} X_2^* y. \tag{23'}$$

These forms are similar to $\hat{\beta} = S^{-1} X^* y = (X^* X)^{-1} X^* y$.

4°. For the model $E(y) = X\beta = X_1\gamma + X_2\delta$ and any parameters γ_0 and δ_0 we have the identity:

$$\begin{aligned} (y - X_1\gamma_0 - X_2\delta_0)^* (y - X_1\gamma_0 - X_2\delta_0) = \\ = (y - X_1\hat{\gamma} - X_2\hat{\delta})^* (y - X_1\hat{\gamma} - X_2\hat{\delta}) + (\hat{\delta} - \delta_0)^* (S^{22})^{-1} (\hat{\delta} - \delta_0) + \\ + [\hat{\gamma} - \gamma_0 + A^{-1} B(\hat{\delta} - \delta_0)]^* A [\hat{\gamma} - \gamma_0 + A^{-1} B(\hat{\delta} - \delta_0)] \end{aligned} \tag{24}$$

where $\hat{\gamma} + A^{-1} B(\hat{\delta} - \delta_0) = \tilde{\gamma}$ is the estimate of the parameter γ when $\delta = \delta_0$, while as usual $\hat{\beta} = \begin{bmatrix} \hat{\gamma} \\ \hat{\delta} \end{bmatrix} = S^{-1} X^* y$.

On account of the symmetry of the parameters γ and δ in the model we have also the identity:

$$(25)\qquad (y-X_1\gamma_0-X_2\delta_0)^*(y-X_1\gamma_0-X_2\delta_0)=$$
$$=(y-X_1\hat{\gamma}-X_2\hat{\delta})^*(y-X_1\hat{\gamma}-X_2\hat{\delta})+(\hat{\gamma}-\gamma_0)^*(S^{11})^{-1}(\hat{\gamma}-\gamma_0)+$$
$$+[\hat{\delta}-\delta_0+D^{-1}B^*(\hat{\gamma}-\gamma_0)]^*D[\hat{\delta}-\delta_0+D^{-1}B^*(\hat{\gamma}-\gamma_0)]$$

where $\hat{\delta}+D^{-1}B^*(\hat{\gamma}-\gamma_0)=\hat{\delta}$ is the estimate of the parameter δ when $\gamma=\gamma_0$.

Proof: To prove the relations (24) and (25) it is sufficient to prove the identities

$$(26)\qquad (\hat{\beta}-\beta_0)^*S(\hat{\beta}-\beta_0)=(\hat{\delta}-\delta_0)^*(S^{22})^{-1}(\hat{\delta}-\delta_0)+$$
$$+[\hat{\gamma}-\gamma_0+A^{-1}B(\hat{\delta}-\delta_0)]^*A[\hat{\gamma}-\gamma_0+A^{-1}B(\hat{\delta}-\delta_0)],$$

and

$$(27)\qquad (\hat{\beta}-\beta_0)^*S(\hat{\beta}-\beta_0)=(\hat{\gamma}-\gamma_0)^*(S^{11})^{-1}(\hat{\gamma}-\gamma_0)+$$
$$+[\hat{\delta}-\delta_0+D^{-1}B^*(\hat{\gamma}-\gamma_0)]^*D[\hat{\delta}-\delta_0+D^{-1}B^*(\hat{\gamma}-\gamma_0)],$$

respectively, both with aide of identity

$$(28)\qquad (y-X\beta_0)^*(y-X\beta_0)=(y-X\hat{\beta})^*(y-X\hat{\beta})+(\hat{\beta}-\beta_0)^*S(\hat{\beta}-\beta_0),$$

given in [6], where $\beta_0=\begin{bmatrix}\gamma_0\\ \delta_0\end{bmatrix}$ is any vector.

Let us first prove the identity (27). For the sake of brevity we write: $\varphi=\hat{\gamma}-\gamma_0$ and $\psi=\hat{\delta}-\delta_0$. Let us find the difference $R=(\hat{\beta}-\beta_0)^*S(\hat{\beta}-\beta_0)-(\hat{\gamma}-\gamma_0)^*(S^{11})^{-1}(\hat{\gamma}-\gamma_0)$. On account of (9) we have:

$$R=\begin{bmatrix}\hat{\gamma}-\gamma_0\\ \hat{\delta}-\delta_0\end{bmatrix}^*\begin{bmatrix}A & B\\ B^* & D\end{bmatrix}\begin{bmatrix}\hat{\gamma}-\gamma_0\\ \hat{\delta}-\delta_0\end{bmatrix}-(\hat{\gamma}-\gamma_0)^*(S^{11})^{-1}(\hat{\gamma}-\gamma_0)=$$
$$=[\varphi^*\ \psi^*]\begin{bmatrix}A & B\\ B^* & D\end{bmatrix}\begin{bmatrix}\varphi\\ \psi\end{bmatrix}-\varphi^*(S^{11})^{-1}\varphi=\varphi^*A\varphi+\psi^*B^*\varphi+\varphi^*B\psi+$$
$$+\psi^*D\psi-\varphi^*(S^{11})^{-1}\varphi=\varphi^*\{A-(S^{11})^{-1}\}\varphi+\psi^*B^*\varphi+\varphi^*B\psi+\psi^*D\psi.$$

Using (14) we have $(S^{11})^{-1}=A-BD^{-1}B^*$ and hence $A-(S^{11})^{-1}=BD^{-1}B^*$. Then $R=\varphi^*BD^{-1}B^*\varphi+\psi^*B^*\varphi+\varphi^*B\psi+\psi^*D\psi=(\psi+D^{-1}B^*\varphi)^*D(\psi+D^{-1}B^*\varphi)$. We have thus proved the identity (27) and because of (28) also the identity (25).

The identity (24) is proved in the same way using (26) instead of (27). To prove it it is sufficient to use the relation $D-(S^{22})^{-1}=B^*A^{-1}B$, which is obtained from the expression (17).

5°. It may be noted that the multiple regression model $\mu = E(y) = X\beta$, where $r(X) = p$, includes the implicit relation (cf. [8]):

$$\underset{n-p,n}{K} \cdot \underset{n1}{\mu} = 0, \tag{29}$$

where

$$\underset{n-p,n}{K} = \left[\underset{n-p,n-p}{I} \quad -L\text{Ł}^{-1} \right] \tag{30}$$

and

$$\underset{np}{X} = \begin{bmatrix} \underset{n-p,p}{L} \\ \hdashline \underset{pp}{\text{Ł}} \end{bmatrix}, \tag{31}$$

with matrix Ł including p independent rows of matrix X and matrix L — the remaining rows.

To prove the relation (29) let us divide the vector μ into two subvectors $\underset{n-p,1}{a}$ and $\underset{p1}{b}$, corresponding to the matrices L and Ł. Thus we obtain

$$\mu = \begin{bmatrix} a \\ \hdashline b \end{bmatrix} = X\beta = \begin{bmatrix} L \\ \hdashline \text{Ł} \end{bmatrix} \beta = \begin{bmatrix} L\beta \\ \hdashline \text{Ł}\beta \end{bmatrix}$$

and hence

$$b = \text{Ł}\beta \tag{32}$$

and $a = L\beta$. From expression (32) we obtain

$$\beta = \text{Ł}^{-1} b, \tag{33}$$

which exists since $r(X) = p$. It then follows that

$$a = L\text{Ł}^{-1} b \quad \text{and} \quad a - L\text{Ł}^{-1} b = 0 = [I \mid -L\text{Ł}^{-1}] \begin{bmatrix} a \\ \hdashline b \end{bmatrix} = K\mu,$$

which we were to prove. Thus the expected values $\mu_1, \mu_2, \ldots, \mu_n$ of random variables $y_1, y_2, \ldots, y_n$ are connected with $n-p$ linear relations given in (29), where the matrix K of rank $n-p$ is in the form (30).

6°. Consider the same multiple regression model as in 5° with the restrictions

$$\underset{gp}{G} \cdot \underset{p1}{\beta} = \underset{g1}{\eta}, \tag{34}$$

imposed on parameters β's, where the matrix G (of order $g \cdot p$) is of rank $g < p$ and η is the column vector with g components. We shall prove that both relations (34) and (29) connecting the expected values $\mu_1, \mu_2, \ldots \mu_n$ and the parameters $\beta_1, \beta_2, \ldots, \beta_p$ may be replaced by one relation of the form

$$M\mu = \vartheta, \tag{35}$$

where

(36) $$\underset{n-p+g,n}{M} = M = \left[\begin{array}{c:c} \underset{n-p,n-p}{I} & -\underset{n-p,p}{L}\cdot\underset{pp}{Ł} \\ \hdashline \underset{g,n-p}{O} & \underset{gp}{G}\cdot\underset{pp}{Ł^{-1}} \end{array}\right],$$

is the matrix of order $(n-p+g)\cdot n$ and of rank $n-p+g$, and

(37) $$\underset{n-p+g,1}{\vartheta} = \vartheta = \left[\begin{array}{c} \underset{n-p,1}{O} \\ \hdashline \underset{g1}{\eta} \end{array}\right].$$

In fact, let us write (34) as $O\cdot a + G\beta = \eta$. Using (33) we have $Oa + GŁ^{-1}b = \eta$ or $[O \vdots GŁ^{-1}]\left[\begin{array}{c} a \\ \hdashline b \end{array}\right] = \eta$. Hence considering $[I \vdots -LŁ^{-1}]\left[\begin{array}{c} a \\ \hdashline b \end{array}\right] = 0$, and ϑ in (37) we obtain (35). It is evident, that rank of matrix M equals $n-p+g$, since diagonal matrices I and $GŁ^{-1}$, being its submatrices, are of ranks $n-p$ and g respectively.

7°. The restrictions

$$H\beta = \nu,$$

may be replaced by the restrictions expressed by means of the relation

(38) $$P\mu = \nu$$

where

(39) $$P = [O \vdots HŁ^{-1}].$$

In fact, introducing β in the form (33) into (4) we obtain $H\beta = HŁ^{-1}b = \nu$ or $Oa + HŁ^{-1}b = \nu$, from which follows $[O \vdots HŁ^{-1}]\left[\begin{array}{c} a \\ \hdashline b \end{array}\right] = \nu$.

4. Lemma and Theorems

In this paper we discuss the model given by the relation (1), (2), (3) and (4) when $r(X) = p$.

It should be noted that some multiple regression models with matrix $X = \underset{np}{X}$ of rank $r < p$ may be transformed into models discused in the present paper i. e. into such models in which the numbers of independent parameters are equal to the ranks of matrices whose elements are the coefficients of these parameters (cf. for instance the model of the type a) in the "Applications").

In view of the great importance of the problem of testing linear hypothesis in the theory of experiment, and its extensive application in the analysis of data obtained from experiments carried out in different fields, such as agriculture, industry, biology, etc. (cf. paper by K. Iwaszkiewicz [5] and J. Neyman [13] and [17]), it is necessary to discuss it most thoroughly (cf. the two proofs of the same theorem given by C. R. Rao in [20] and [21]).

We present the lemma which generalizes H. B. Mann's lemma 4.1 (cf. [8]) for unhomogeneous relations.

Lemma. Let matrix relation

$$M\mu = \vartheta \tag{40}$$

where $M = \underset{kn}{M} = \{m_{ij}\}$, $i = 1, 2, ..., k$, $j = 1, 2, ..., n$, $\mu^* = [\mu_1, \mu_2, ..., \mu_n]$, $\vartheta^* = [\vartheta_1, \vartheta_2, ..., \vartheta_k]$, define k linearly independent linear restrictions imposed on n expected values $\mu_1, \mu_2, ..., \mu_n$. Then we may determine the system of restrictions

$$W\mu = \Theta \tag{41}$$

which is equivalent to (40), where $W = \underset{kn}{W} = \{w_{ij}\}$, $i = 1, 2, ..., k$, $j = 1, 2, ..., n$, $\Theta^* = [\Theta_1, \Theta_2, ..., \Theta_k]$, and where the matrix W is orthogonalized and normalized i. e. $WW^* = \underset{kk}{I}$.

Proof: The elements of matrix W are obtained according to Mann's recurrent method given in lemma 4.1. We use therefore the formulae:

$$\lambda_{lj} = \sum_{\alpha=1}^{n} w_{j\alpha} m_{l+1,\alpha}, \quad j = 1, 2, ..., l, \quad \text{where} \quad l < k, \tag{42}$$

$$w'_{l+1,\alpha} = m_{l+1,\alpha} - \lambda_{l1} w_{1\alpha} - \cdots - \lambda_{ll} w_{l\alpha}, \tag{43}$$

$$w_{l+1,\alpha} = \frac{w'_{l+1,\alpha}}{\sqrt{\sum_{\alpha=1}^{n} w'^2_{l+1,\alpha}}}, \quad \begin{matrix} \alpha = 1, 2, ..., n, \\ l = 0, 1, 2, ..., k-1, \end{matrix} \tag{44}$$

where for $l = 0$ we write $w'_{1\alpha} = m_{1\alpha}$.

Considering (40), (41), (42), (43) and (44), the components of vector Θ are obtained immediately:

$$(45)\quad \Theta_{l+1}=\sum_{\alpha=1}^{n} w_{l+1,\alpha}\mu_\alpha=\frac{\sum\limits_{\alpha=1}^{n} w'_{l+1,\alpha}\mu_\alpha}{\sqrt{\sum\limits_{\alpha=1}^{n} w'^2_{l+1,\alpha}}}=$$

$$=\frac{1}{\sqrt{\sum\limits_{\alpha=1}^{n} w'^2_{l+1,\alpha}}}\sum_{\alpha=1}^{n}\left(m_{l+1,\alpha}-\sum_{j=1}^{l}\lambda_{lj}w_{j\alpha}\right)\mu_\alpha=$$

$$=\frac{1}{\sqrt{\sum\limits_{\alpha=1}^{n} w'^2_{l+1,\alpha}}}\left[\sum_{\alpha=1}^{n} m_{l+1,\alpha}\mu_\alpha-\lambda_{l1}\sum_{\alpha=1}^{n} w_{1\alpha}\mu_\alpha-\dots-\lambda_{ll}\sum_{\alpha=1}^{n} w_{l\alpha}\mu_\alpha\right]=$$

$$=\frac{1}{\sqrt{\sum\limits_{\alpha=1}^{n} w'^2_{l+1,\alpha}}}(\vartheta_{l+1}-\lambda_{l1}\Theta_1-\lambda_{l2}\Theta_2-\dots-\lambda_{ll}\Theta_l),$$

where $l=0,1,2,\dots,k\text{-}1$, and for $l=0$ we admit $w'_{1\alpha}=m_{1\alpha}$ and $\lambda_{00}=0$. The expression (45) is a recurrent formula, which permits effectively to determine the coefficients appearing on the right sides of the relations (41). In particular we have

$$\Theta_1=\frac{\vartheta_1}{\sqrt{\sum\limits_{\alpha=1}^{n} m^2_{1\alpha}}}.$$

Theorem 1. Let the random variables y_α in the multiple regression model $y=X\beta+e=\mu+e$, where $E(y)=\mu=X\beta$, be normally and independently distributed with expected values μ_α, $\alpha=1,2,\dots,n$ and common variance σ^2. Let also matrix relation $K\mu=0$ (cf. (29)) determine $n-p$ independent linear restrictions for expected values $\mu_1,\mu_2,\dots,\mu_n$ derived from the model with p parameters $\beta_1,\beta_2,\dots,\beta_p$, which are subjected to g independent linear restrictions $G\beta=\eta$ (cf. (34)) where $g<p$. Let hypothesis be true that $h<p-g$ further linear restrictions $H\beta=\nu$ with respect to $\beta_1,\beta_2,\dots,\beta_p$ independent of preceding g restrictions and mutually independent, assume the values determined by the components of the vector ν (cf. (4)). Then the random variable

$$F=\frac{Q_r-Q_a}{h}:\frac{Q_a}{n-p+g}$$

has F distribution with h and $n-p+g$ degrees of freedom.

Proof: In the section 3 we proved, that both relations $K\mu=\theta$ and $G\beta=\eta$ may be replaced by one relation $M\mu=\vartheta$ (cf. (35), (36), (29) and (30)), and that the restrictions $H\beta=\nu$ may be expressed in the form $P\mu=\nu$ (cf. (4), (38) and (39)).

From the lemma it follows that the relations $M\mu=\vartheta$ and $P\mu=\nu$ may be transformed into their corresponding two equivalent sets of restrictions with such property that rows of two sets of these restrictions will belong to orthogonal and normalized matrices. Let the relations $M\mu=\vartheta$ and $P\mu=\nu$ be such transformed sets of restrictions i. e. let $MM^*=I$, $PP^*=I$, $MP^*=O$. Add orthogonal and normalized matrix

$$\begin{bmatrix} \underset{n-p+g,\, n}{M} \\ \hdashline \underset{h,\, n}{P} \end{bmatrix}$$

with $n-p+g+h$ rows and n columns to orthogonal and normalized square matrix

$$\underset{nn}{C}=\begin{bmatrix} \underset{n-p+g,\, n}{M} \\ \hdashline \underset{hn}{P} \\ \hdashline \underset{p-g-h,\, n}{\Pi} \end{bmatrix},$$

which is always possible to do. Then $CC^*=I$, $MM^*=I$, $PP^*=I$, $\Pi\Pi^*=I$, $MP^*=O$, $M\Pi^*=O$ and $P\Pi^*=O$. Consider orthogonal transformation $t=Cy$, where

$$t=\begin{bmatrix} z+\vartheta \\ \hdashline u+\nu \\ \hdashline w+\varrho \end{bmatrix},$$

and where the number of components of vectors z, u and w are equal: $n-p+g$, h and $p-g-h$ respectively. Since $t=Cy$ and $CC^*=I$ we have $z+\vartheta=My$, $u+\nu=Py$, $w+\varrho=\Pi y$ and hence

(46) $$z=My-\vartheta, \quad u=Py-\nu \quad \text{and} \quad w=\Pi y-\varrho.$$

Denoting $E(z)=m_z$, $E(u)=m_u$ and $E(w)=m_w$ we obtain: $m_z=M\mu-\vartheta$, $m_u=P\mu-\nu$ and $m_w=\Pi\mu-\varrho$.

Before finding Q_a note that $(y-\mu)^*(y-\mu)=(t-m_t)^*(t-m_t)$, where $E(t)=m_t$. In fact, since $t=Cy$ and $CC^*=I$ we have $y=C^{-1}t=C^*t$ and $\mu=E(y)=C^*m_t$. Hence $(y-\mu)^*(y-\mu)=(C^*t-C^*m_t)^*(C^*t-C^*m_t)=$ $=(t-m_t)^*CC^*(t-m_t)=(t-m_t)^*(t-m_t)$.

Now let us find Q_a, by minimalizing corresponding expression with respect to μ. We have $Q_a = \underset{M\mu=\vartheta}{Min}(y-\mu)^*(y-\mu) = \underset{M\mu=\vartheta}{Min}(t-m_t)^*(t-m_t) = Min[(z+\vartheta-m_z-\vartheta)^*(z+\vartheta-m_z-\vartheta)+(u+\nu-m_u-\nu)^*(u+\nu-m_u-\nu)+(w+\varrho-m_w-\varrho)^*(w+\varrho-m_w-\varrho)]$ and writing $m_z = M\mu-\vartheta$ we obtain $Q_a = z^*z + Min[(u-m_u)^*(u-m_u)+(w-m_w)^*(w-m_w)] = z^*z$.

Similarly we find Q_r by minimalizing the sum of squares of residuals $e^*e = (y-\mu)^*(y-\mu)$ with respect to μ, this time under two restrictions $M\mu = \vartheta$ and $P\mu = \nu$. We obtain in succession: $Q_r = Min\, e^*e = Min[(z-m_z)^*(z-m_z)+(u-m_u)^*(u-m_u)+(w-m_w)^*(w-m_w)] = z^*z+u^*u+Min(w-m_w)^*(w-m_w) = z^*z+u^*u$ and hence $Q_r - Q_a = u^*u$, where $u = \underset{h1}{u}$.

We shall prove that components of vectors $\underset{n-p+g,1}{z}$ and $\underset{h1}{u}$ are uncorrelated and normally distributed random variables with variance σ^2 and means equal to zero. For this purpose let us find covariance matrix of vector t:

$$E(t-m_t)(t-m_t)^* = E\begin{bmatrix} z-m_z \\ u-m_u \\ w-m_w \end{bmatrix}[(z-m_z)^* \; (u-m_u)^* \; (w-m_w)^*] =$$

$$= E\begin{bmatrix} (z-m_z)(z-m_z)^* & (z-m_z)(u-m_u)^* & (z-m_z)(w-m_w)^* \\ (u-m_u)(z-m_z)^* & (u-m_u)(u-m_u)^* & (u-m_u)(w-m_w)^* \\ (w-m_w)(z-m_z)^* & (w-m_w)(u-m_u)^* & (w-m_w)(w-m_w)^* \end{bmatrix} =$$

$$= \begin{bmatrix} \sigma^2 \cdot \underset{n-p+g,\, n-p+g}{I} & O & O \\ O & \sigma^2 \cdot \underset{hh}{I} & O \\ O & O & \sigma^2 \cdot \underset{p-g-h,\, p-g-h}{I} \end{bmatrix}$$

(since on account of relations (46) we obtain successively:

$$E(z-m_z)(z-m_z)^* = E[(z+\vartheta)-(m_z+\vartheta)][(z+\vartheta)-(m_z+\vartheta)]^* =$$
$$= (My-M\mu)(My-M\mu)^* = M(y-\mu)(y-\mu)^*M^* = MI\sigma^2M^* = \sigma^2 \cdot \underset{n-p+g,\, n-p+g}{I}$$

and further

$E(z-m_z)(u-m_u)^* = E(My-M\mu)(Py-P\mu)^* = MI\sigma^2P^* = \sigma^2 IMP^* = O$, etc.).

Noting that $E(z) = E(My-\vartheta) = M\mu-\vartheta = 0$ and that by hypothesis: $P\mu = \nu$ we obtain $E(u) = E(Py-\nu) = P\mu-\nu = 0$. Thus it has been proved that Q_a is the sum of squares of $n-p+g$ normal and independent random variables with the means zero and common variance σ^2 i. e.

$$\frac{Q_a}{\sigma^2} = \frac{z^*z}{\sigma^2} = \chi^2_{n-p+g}$$

which means that Q_a/σ^2 is distributed as Chi-square with $(n-p+g)$ degrees of freedom. Similarly $Q_r - Q_a = u^* u$ is the sum of squares of h independent normally distributed random variables with means zero and common variance σ^2 i. e. $(Q_r - Q_a)/\sigma^2$ is distributed as Chi-square with h degrees of freedom. It has been also shown (cf. covariance matrix of vector t) that Q_a and $Q_r - Q_a$ are independent. Consequently we conclude that if the hypothesis $H\beta = \nu$ is true, the random variable

$$F = \frac{Q_r - Q_a}{h} : \frac{Q_a}{n-p+g}$$

has F distribution with h and $n-p+g$ degrees of freedom. Consequently it can be used to test the null hypothesis that $H\beta = \nu$.

Theorem 2. Consider the multiple regression model $y = X\beta + e = X_1\gamma + X_2\delta + e$ where according to the notation introduced on pages 20—21 we have $E(y) = X\beta = \mu$, and $\beta = \begin{bmatrix} \gamma \\ \delta \end{bmatrix}$. Let further the components of vector e be normally distributed with zero means, common variance σ^2 and zero covariances, and finally let rank of matrix X be equal to the number of parameters $\beta_1, \beta_2, \ldots, \beta_p$ i. e. let $r(X) = p$. Then on the null hypothesis: $\gamma = \gamma_0$ the random variable defined as the ratio

$$F = \frac{(\hat{\gamma} - \gamma_0)^* (S^{11})^{-1} (\hat{\gamma} - \gamma_0)}{q} : \frac{(y - X\hat{\beta})^* (y - X\hat{\beta})}{n - p} \tag{47}$$

has F distribution with q and $n-p$ degrees of freedom, and on the hypothesis: $\delta = \delta_0$ the random variable

$$F = \frac{(\hat{\delta} - \delta_0)^* (S^{22})^{-1} (\hat{\delta} - \delta_0)}{p - q} : \frac{(y - X\hat{\beta})^* (y - X\hat{\beta})}{n - p} \tag{48}$$

has F distribution with $p-q$ and $n-p$ degrees of freedom (the symbols here used are introduced in the sec. 2).

Proof: For the proof we shall apply the theorem 1. As it is known (cf. [6])

$$Q_a = Min\,(y - X\beta)^* (y - X\beta) = (y - X\hat{\beta})^* (y - X\hat{\beta}) = y^* y - \hat{\beta}^* X^* y \tag{49}$$

where $\hat{\beta} = S^{-1} X^* y$ and $S = X^* X$.

On the hypothesis that $\gamma = \gamma_0$ the model $y = X\beta + e = X_1\gamma + X_2\delta + e$ assumes the form $y = X_1\gamma_0 + X_2\delta + e$ or

$$E(y - X_1\gamma_0) = X_2\delta. \tag{50}$$

Now Q_r by (49) has for the model (50) the form:

$$Q_r = Min\,(y - X\beta)^* (y - X\beta) = Min\,(y - X_1\gamma_0 - X_2\delta)^* (y - X_1\gamma_0 - X_2\delta) =$$
$$= (y - X_1\gamma_0 - X_2\tilde{\delta})^* (y - X_1\gamma_0 - X_2\tilde{\delta}) = (y - X_1\gamma_0)^* (y - X_1\gamma_0) -$$
$$- \tilde{\delta}^* X_2^* (y - X_1\gamma_0) = y^* y - 2\,y^* X_1\gamma_0 + \gamma_0^* A\gamma_0 - \tilde{\delta}^* X_2^* (y - X_1\gamma_0),$$

where the conditional estimate

(51) $$\tilde{\delta} = D^{-1} X_2^* (y - X_1\gamma_0) = D^{-1} (X_2^* y - B^*\gamma_0)$$

is deduced from the normal equations

(52) $$D\tilde{\delta} = X_2^* (y - X_1\gamma_0)$$

where, as usual, $A = X_1^* X_1$, $B = X_1^* X_2$ and $D = X_2^* X_2$.

Since from the theory of normal regression it is clear that the estimete $\hat{\beta}$ is found from the normal equation $S\hat{\beta} = X^* y$, we obtain (cf. p. 21 and (9)):

$$S\hat{\beta} = \begin{bmatrix} A & B \\ B^* & D \end{bmatrix} \begin{bmatrix} \hat{\gamma} \\ \hat{\delta} \end{bmatrix} = \begin{bmatrix} A\hat{\gamma} + B\hat{\delta} \\ B^*\hat{\gamma} + D\hat{\delta} \end{bmatrix} = X^* y = \begin{bmatrix} X_1^* y \\ X_2^* y \end{bmatrix},$$

hence

(53) $$A\hat{\gamma} + B\hat{\delta} = X_1^* y$$

(54) $$B^*\hat{\gamma} + D\hat{\delta} = X_2^* y$$

Using the relation (54) we obtain

(55) $$\tilde{\delta} = D^{-1} (B^*\hat{\gamma} + D\hat{\delta} - B^*\gamma_0) = \hat{\delta} + D^{-1} B^* (\hat{\gamma} - \gamma_0)$$

instead of (51). By (49), (55), and the form of Q_r, as well as by the expression

$$\hat{\beta}^* X^* y = \begin{vmatrix} \hat{\gamma} \\ \hat{\delta} \end{vmatrix}^* |X_1 \; X_2|^* y = \hat{\gamma}^* X_1^* y + \hat{\delta}^* X_2^* y,$$

we have further

$$Q_r - Q_a = -2\,y^* X_1\gamma_0 + \gamma_0^* A\gamma_0 - [\hat{\delta} + D^{-1} B^* (\hat{\gamma} - \gamma_0)]^* (X_2^* y - B^*\gamma_0) + \hat{\beta}^* X^* y$$

and using (53) and (54) we find after the reduction:

$$Q_r - Q_a = (\hat{\gamma} - \gamma_0)^* [A - BD^{-1} B^*] (\hat{\gamma} - \gamma_0).$$

Comparing this with (14) we obtain

$$Q_r - Q_a = (\hat{\gamma} - \gamma_0)^* (S^{11})^{-1} (\hat{\gamma} - \gamma_0)$$

and finally by (49) it appears that random variable F mentionned in the theorem 1 assumes the form of the ratio (47). It may be used thus for testing the hypothesis $\gamma=\gamma_0$.

The alternative to the null hypothesis: $\gamma=\gamma_0$ is $\gamma\neq\gamma_0$, which means that at least one of the parameters $\beta_1,\beta_2,\ldots,\beta_q$ does not assume the value mentionned in the null hypothesis.

The formula (48) is deduced in the same manner. Here should be used the relation (17), from which it follows that $(S^{22})^{-1}=D-B^*A^{-1}B$. The proof of the theorem 2 is concluded.

Theorem 3. Consider the multiple regression model $\mu=E(y)=XB$, where the components of vector y are independent random variables normally distributed with means $\mu_1,\mu_2,\ldots,\mu_n$ and common variance σ^2. Besides, let the rank of matrix X be equal to the number p of parameters β and let the hypothesis that $\varphi=ŁB=\varphi_0$ be true, where $Ł=\{l_{ij}\}$, $(i=1,2,\ldots,q;\ j=1,2,\ldots,p)$ and $r(Ł)=q$. Then the random variable

$$(56)\qquad F=\frac{(Ł\hat\beta-\varphi_0)^*(ŁS^{-1}Ł^*)^{-1}(Ł\hat\beta-\varphi_0)}{q}:\frac{(y-X\hat\beta)^*(y-X\hat\beta)}{n-p}$$

has F distribution with q and $n-p$ degrees of freedom.

Proof: Since by assumption the rank of the matrix $Ł$ is q we can add to this matrix such a rectangular matrix M with $p-q$ rows and p columns, as to make the resulting square matrix $p\times p$ non-singular. Let such a matrix be $R=\begin{bmatrix}Ł\\ \hdashline M\end{bmatrix}$. Similarly let us add to vector $\varphi^*=[\varphi_1,\varphi_2,\ldots,\varphi_q]$ with q components a vector $\varrho^*=[\varrho_{q+1},\varrho_{q+2},\ldots,\varrho_p]$ with $p-q$ components, such that the relation $\underset{p1}{\psi}=\begin{bmatrix}\varphi\\ \hdashline \varrho\end{bmatrix}=R\beta$ is obtained. For this purpose it is sufficient to put $\varrho=M\beta$. Then $\beta=R^{-1}\psi$ and the multiple regression model $E(y)=X\beta$ becomes:

$$E(y)=XR^{-1}\psi=Z\psi=\underset{np}{X}\Big[\underset{pq}{R^{11}}\ \vdots\ \underset{p,p-q}{R^{12}}\Big]\begin{bmatrix}\varphi\\ \hdashline \varrho\end{bmatrix}=XR^{11}\varphi+XR^{12}\varrho=Z_1\varphi+Z_2\varrho,$$

where

$$\underset{pp}{R^{-1}}=\Big[\underset{pq}{R^{11}}\ \vdots\ \underset{p,p-q}{R^{12}}\Big]\quad\text{and}\quad \underset{np}{Z}=\Big[\underset{nq}{Z_1}\ \vdots\ \underset{n,p-q}{Z_2}\Big]=XR^{-1}\quad\text{or}\quad Z_1=XR^{11}$$

and $Z_2=XR^{12}$. Since the model $\mu=E(y)=X\beta$ has on the null hypothesis $\varphi=Ł\beta=\varphi_0$ been transformed into the model $E(y)=Z\psi=Z_1\varphi+Z_2\varrho$, the situation described in the theorem 2 is obtained. Consequently

$Q_a = (y - X\hat{\beta})^* (y - X\hat{\beta})$, where $\hat{\beta} = S^{-1} X^* y$ and, as usual, $S = X^* X$ and $Q_r - Q_a = (\hat{\varphi} - \varphi_0)^* (W^{11})^{-1} (\hat{\varphi} - \varphi_0)$, where W^{11} is the submatrix of matrix

$$W^{-1} = \begin{bmatrix} \underset{qq}{W^{11}} & \underset{q,\,p-q}{W^{12}} \\ \underset{p-q,\,q}{W^{21}} & \underset{p-q,\,p-q}{W^{22}} \end{bmatrix},$$

which is the reciprocal matrix of $\underset{pp}{W} = Z^* Z$. Let us find the matrix W^{11}. To do this let us observe that $W = Z^* Z = (R^{-1})^* X^* X R^{-1} = (R^{-1})^* S R^{-1}$ i.e.

$$W^{-1} = \begin{bmatrix} W^{11} & W^{12} \\ W^{21} & W^{22} \end{bmatrix} = [(R^{-1})^* S R^{-1}]^{-1} = R S^{-1} R^* =$$

$$= \begin{bmatrix} Ł \\ M \end{bmatrix} S^{-1} [Ł^* \; M^*] = \begin{bmatrix} ŁS^{-1}Ł^* & ŁS^{-1}M^* \\ MS^{-1}Ł^* & MS^{-1}M^* \end{bmatrix}.$$

Hence comparing the corresponding submatrices we have:

$$\underset{qq}{W^{11}} = ŁS^{-1}Ł^*, \quad \underset{q,\,p-q}{W^{12}} = ŁS^{-1}M^*, \quad \underset{p-q,\,q}{W^{21}} = MS^{-1}Ł^*, \quad \underset{p-q,\,p-q}{W^{22}} = MS^{-1}M^*.$$

Thus $Q_r - Q_a = (\hat{\varphi} - \varphi_0)^* (ŁS^{-1}Ł^*)^{-1} (\hat{\varphi} - \varphi_0)$. By the relation $\varphi = Ł\beta$ we obtain $\hat{\varphi} = Ł\hat{\beta} = ŁS^{-1}X^* y$. In consequence $Q_r - Q_a = (Ł\hat{\beta} - \varphi_0)^* (ŁS^{-1}Ł^*)^{-1} (Ł\hat{\beta} - \varphi_0)$. Considering further the form Q_a represented above, we obtain by the theorem 2 the random variable F defined in (56).

Similarly as in the theorems 1 and 2 this random variable may be used for testing the null hypothesis $Ł\beta = \varphi_0$. The alternative to the null hypothesis $Ł\beta = \varphi_0$ is: $Ł\beta$ is equal to any ϑ_0, different from φ_0. The alternative hypothesis states that q linear parametric functions expressed in the matrix form by $Ł\beta$ have values different from components of vector φ_0.

Theorem 4. Let g linearly independent restrictions defined by the matrix relation $G\beta = \eta$ (cf. (3)), where $G = \underset{gp}{G}$ and $\eta = \underset{g1}{\eta}$, $g < p$, be imposed on the parameters $\beta_1, \beta_2, \ldots, \beta_p$ of the multiple regression model $\mu = E(y) = X\beta$, where the components of vector y are independent random variables normally distributed with means $\mu_1, \mu_2, \ldots, \mu_n$ and common variance σ^2, and where $r(X) = p$. Let also the null hypothesis $\underset{hp}{H}\,\underset{p1}{\beta} = \underset{h1}{\nu}$ be true, where $r(H) = h$ and $g + h < p$, when all the $g + h$ linear parametric functions, expressed by $G\beta = \eta$ and $H\beta = \nu$, are mutually independent. Under these assumptions the random variable given by the ratio

$$(57)\quad F = \frac{(T\hat{\beta}-\tau)^*(TS^{-1}T^*)^{-1}(T\hat{\beta}-\tau)-(G\hat{\beta}-\eta)^*(GS^{-1}G^*)^{-1}(G\hat{\beta}-\eta)}{h} : \frac{(y-X\hat{\beta})^*(y-X\hat{\beta})+(G\hat{\beta}-\eta)^*(GS^{-1}G^*)^{-1}(G\hat{\beta}-\eta)}{n-p+g}$$

has F distribution with h and $n-p+g$ degrees of freedom, where

$$\underset{g+h,p}{T} = \begin{bmatrix} G \\ H \end{bmatrix} \quad \text{and} \quad \underset{g+h,1}{\tau} = \begin{bmatrix} \eta \\ \nu \end{bmatrix}.$$

Proof: Under the assumptions of the theorem 3, where $G\beta = \eta$ plays the rôle of the restrictions $Ł\beta = \varphi_0$, we obtain:

$$Q'_r = (G\hat{\beta}-\eta)^*(GS^{-1}G^*)^{-1}(G\hat{\beta}-\eta)+(y-X\hat{\beta})^*(y-X\hat{\beta}).$$

Under the conditions of theorem 4 Q_a has the latter form, so that:

$$Q_a = \underset{G\beta=\eta}{Min}(y-X\beta)^*(y-X\beta) =$$
$$= (y-X\hat{\beta})^*(y-X\hat{\beta})+(G\hat{\beta}-\eta)^*(GS^{-1}G^*)^{-1}(G\hat{\beta}-\eta)$$

with $n-p+g$ degrees of freedom, where as usual $\hat{\beta} = S^{-1}X^*y$ and $S = X^*X$.

Similarly we find Q_r after performing the minimalization under the restrictions $G\beta = \eta$ and $H\beta = \nu$: Since these restrictions are replaced by one relation $T\beta = \tau$, where the matrix T and vector τ are defined in the theorem 4, we have:

$$Q_r = Min(y-X\beta)^*(y-X\beta) =$$
$$= (y-X\hat{\beta})^*(y-X\hat{\beta})+(T\hat{\beta}-\tau)^*(TS^{-1}T^*)^{-1}(T\hat{\beta}-\tau)$$

with $n-p+g+h$ degrees of freedom. Hence

$$Q_r - Q_a = (T\hat{\beta}-\tau)^*(TS^{-1}T^*)^{-1}(T\hat{\beta}-\tau)-(G\hat{\beta}-\eta)^*(GS^{-1}G^*)^{-1}(G\hat{\beta}-\eta).$$

Applying the theorem 1 we may state that random variable expressed by the ratio (57) has F distribution with h and $n-p+g$ degrees of freedom.

Let us observe that this theorem may be also proved (at a greater length) using Lagrange's multipliers method.

Theorem 5. Let there be given the multiple regression model $\mu = E(y) = X\beta$, where the components of vector $y^* = |y_1, y_2, ..., y_n|$ are independent random variables normally distributed with means $\mu_1, \mu_2, ..., \mu_n$ and common variance σ^2, and the rank of the matrix X is equal to the number p of parameters β. Let the null hypothesis be true that the parameters $\beta_1, \beta_2, \ldots, \beta_p$ are represented by linear combinations of $p-q$ parameters $\vartheta_{q+1}, \vartheta_{q+2}, \ldots, \vartheta_p$, $q > 0$, expressed by the matrix relation $\beta = U\vartheta$, where the rank of matrix $U = \underset{p,\,p-q}{U}$ is $p-q$ and $\vartheta^* = \underset{p-q,\,1}{\vartheta^*} = [\vartheta_{q+1}, \vartheta_{q+2}, \ldots, \vartheta_p]$. Then the random variable expressed by the ratio

$$F = \frac{(S\hat{\beta})^* [S^{-1} - U(U^* SU)^{-1} U^*] S\hat{\beta}}{q} : \frac{(y - X\hat{\beta})^* (y - X\hat{\beta})}{n-p} \tag{58}$$

has F distribution with q and $n-p$ degrees of freedom.

Proof: Let matrix $P = \underset{pq}{P}$ complete the matrix $\underset{p,\,p-q}{U}$ to obtain the non-singular matrix $\underset{pp}{T} = [P \vdots U]$. In the same way the vector ϑ with $p-q$ components is completed by vector $\underset{q1}{\pi}$, to get the vector $\underset{p1}{\varrho} = \begin{bmatrix} \pi \\ \cdots \\ \vartheta \end{bmatrix}$ with p components. Thus we have the transformation of vector β into the vector ϱ, which is expressed by the matrix relation

$$\beta = T\varrho = [P \vdots U] \begin{bmatrix} \pi \\ \cdots \\ \vartheta \end{bmatrix} = P\pi + U\vartheta. \tag{59}$$

By (59) the hypothesis $\beta = U\vartheta$ may be written: $\pi = o$. Then the multiple regression model assumes the form:

$$E(y) = X\beta = XT\varrho = XP\pi + XU\vartheta = Z_1\pi + Z_2\vartheta = Z\varrho$$

where we put $\underset{nq}{Z_1} = \underset{np}{X} \cdot \underset{pq}{P}$, $\underset{n,\,p-q}{Z_2} = \underset{np}{X} \cdot \underset{p,\,p-q}{U}$ and $\underset{np}{Z} = [Z_1 \vdots Z_2] = XT$, and where $r(Z) = p$, since $r(X) = p$ and T is a non-singular matrix. Thus we come to the problem, which is dealt with in theorem 2: Considering the model $E(y) = Z_1\pi + Z_2\vartheta$ and assuming that the null hypothesis $\pi = o$ is true, we conclude that the random variable

$$F = \frac{\hat{\pi}^* (H^{11})^{-1} \hat{\pi}}{q} : \frac{(y - X\hat{\beta})^* (y - X\hat{\beta})}{n-p} \tag{60}$$

has F distribution with q and $n-p$ degrees of freedom, where $H = \underset{pp}{H} = Z^* Z$ and according to the symbols introduced in the sec. 2 the matrix H^{11} is the submatrix of

$$H^{-1} = \begin{bmatrix} \underset{qq}{H^{11}} & \vdots & \underset{q,\,p-q}{H^{12}} \\ \cdots & & \cdots \\ \underset{p-q,\,q}{H^{21}} & \vdots & \underset{p-q,\,p-q}{H^{22}} \end{bmatrix}$$

and $\hat{\pi} = H^{11} Z_1^* y + H^{12} Z_2^* y = (H^{11} P^* + H^{12} U^*) S \hat{\beta}$ (cf. (20)). In the applications it is more convenient to use the formula (58), which is deduced in the following way.

We have $Q_a = Min (y - X\beta)^* (y - X\beta) = (y - X\hat{\beta})^* (y - X\hat{\beta})$, where $\hat{\beta} = S^{-1} X^* y$. Further when the null hypothesis $\beta = U\vartheta$ is true $Q_r = Min (y - X\beta)^* (y - X\beta)$ where the minimalization is to be performed with respect to parameters $\beta_1, \beta_2, \ldots, \beta_p$. Let us find an explicit expression for $Q_r - Q_a$. On the hypothesis: $\beta = U\vartheta$ the multiple regression model has the form

(61) $$E(y) = X\beta = XU\vartheta = Z_2 \vartheta$$

where $\underset{n,\, p-q}{Z_2} = XU$ and $r(Z_2) = p - q$, since under the assumption $r(X) = p$ and $r(U) = p - q$.

The estimate of parameter ϑ in the form

$$\tilde{\vartheta} = (Z_2^* Z_2)^{-1} Z_2^* y = (U^* SU)^{-1} U^* X^* y$$

is obtained from the normal equations $(Z_2^* Z_2) \tilde{\vartheta} = Z_2^* y$, where $Z_2^* Z_2 = U^* SU$. Hence

(62) $$Q_r - Q_a = (y - Z_2 \tilde{\vartheta})^* (y - Z_2 \tilde{\vartheta}) - (y - X\hat{\beta})^* (y - X\hat{\beta})$$

and by

$$(y - X\hat{\beta})^* (y - X\hat{\beta}) = y^* y - \hat{\beta}^* S \hat{\beta} = y^* y - y^* X S^{-1} X^* y$$

we have

(63) $$\begin{aligned} Q_r - Q_a &= y^* y - y^* Z_2 (U^* SU)^{-1} Z_2^* y - y^* y + y^* X S^{-1} X^* y = \\ &= y^* X S^{-1} X^* y - y^* XU (U^* SU)^{-1} U^* X^* y = \\ &= y^* X [S^{-1} - U(U^* SU)^{-1} U^*] X^* y = (S\hat{\beta})^* [S^{-1} - U(U^* SU)^{-1} U^*] S\hat{\beta}. \end{aligned}$$

Using the theorem 1 we obtain the random variable F in the form (58) with q and $n-p$ degrees of freedom. This random variable may be used for testing the null hypothesis $\beta = U\vartheta$ in the case of model (61). The same is true of the theorem 4.

5. Expected values of quadratic forms, covariance matrices, and other matrix relations

The problem of testing a null hypothesis against an alternative hypothesis is connected with the problem of determining the type of test i. e. whether the test is to be one-tailed or two-tailed. Since the random variables F discussed in this paper are the ratios of quadratic forms, the determination of the type of test based on these random variables requires finding expected values of corresponding quadratic forms. To do this it is necessary to deduce a number of matrix relations. Incidentally some covariance matrices and other matrix expressions which can be applied in the multiple regression will be deduced.

1°. Let $y_1, y_2, \ldots, y_n$ be random variables with covariances $cov(y_i, y_j) = E(y_i - \mu_i)(y_j - \mu_j)$, $i, j = 1, 2, \ldots, n$, and means $\mu_1, \mu_2, \ldots, \mu_n$. Let further $y^* = [y_1, y_2, \ldots, y_n]$ and $\mu^* = [\mu_1, \mu_2, \ldots, \mu_n] = E(y^*)$, and let $\underset{nn}{P} = \{p_{ij}\}$ be the matrix of quadratic form $y^* Py$. We shall prove that under these assumptions

(64) $$E(y^* Py) = \mu^* P\mu + \sum_i \sum_j p_{ij}\, cov(y_i, y_j).$$

Proof:

$$E(y^* Py) = E(y^* Py) - E\mu^* Py + E\mu^* Py = E(y-\mu)^* Py + \mu^* P\mu =$$

$$= E(y-\mu)^* P(y-\mu) + E(y-\mu)^* P\mu + \mu^* P\mu = E(y-\mu)^* P(y-\mu) + \mu^* P\mu =$$

$$= E\sum_i \sum_j (y_i - \mu_i) p_{ij} (y_j - \mu_j) + \mu^* P\mu = \mu^* P\mu + \sum_i \sum_j p_{ij}\, cov(y_i, y_j).$$

1°*a*. When the random variables $y_1, y_2, \ldots, y_n$ are uncorrelated and have variances equal to $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$ respectively, then $cov(y_i, y_j) = 0$ for $i \neq j$, $i, j = 1, 2, \ldots, n$, and $cov(y_i, y_i) = \sigma_i^2$ for $i = 1, 2, \ldots, n$. Then the identity (64) assumes the form

(64 *a*) $$E(y^* Py) = \mu^* P\mu + \sum_i p_{ii}\sigma_i^2.$$

2°. The sum of diagonal elements of matrix

(65) $$\underset{nn}{L} = \underset{nm}{A} \cdot \underset{m1}{b} \cdot (\underset{nm}{A} \cdot \underset{m1}{c})^* = Abc^* A^* \quad \text{is equal to} \quad (Ab)^* Ac$$

where $A = \underset{nm}{A}$, $b = \underset{m1}{b}$, and $c = \underset{m1}{c}$.

Proof: Let $A^* = [[a_1] \vdots [a_2] \vdots \cdots \vdots [a_n]]$ i.e. let

$$A = \begin{bmatrix} [a_1]^* \\ [a_2]^* \\ \vdots \\ [a_n]^* \end{bmatrix}$$

where $[a_i]^*$, $i=1,2,\ldots,n$, is the vector with the components constituting i-th row of matrix A. Then

(66) $$A^* A = \sum_i [a_i][a_i]^*$$

and matrix L has the form:

$$L = Abc^* A^* = \begin{bmatrix} [a_1]^* bc^* [a_1], & [a_1]^* bc^* [a_2], \ldots, [a_1]^* bc^* [a_n] \\ [a_2]^* bc^* [a_1], & [a_2]^* bc^* [a_2], \ldots, [a_2]^* bc^* [a_n] \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ [a_n]^* bc^* [a_1], & [a_n]^* bc^* [a_2], \ldots, [a_n]^* bc^* [a_n] \end{bmatrix}$$

Thus the sum of diagonal elements of matrix L equals:

$$\sum_i [a_i]^* bc^* [a_i] = \sum_i b^* [a_i][a_i]^* c = b^* \left(\sum_i [a_i][a_i]^* \right) c = b^* A^* Ac = (Ab)^* Ac.$$

3°. For the multiple regression model $y = X\beta + e$ we have

(67) $$Ee^* XS^{-1} X^* e = p\sigma^2$$

where, as usual, $S = X^*X$ and σ^2 is the variance of each of n independent random variables $y_1, y_2, \ldots, y_n$ that are the components of vector y.

Proof: By the identity (64a) we obtain

$$Ee^* XS^{-1} X^* e = \sigma^2 \sum_i^n f_{ii},$$

where f_{ii} are diagonal elements of matrix $F = XS^{-1}X^*$. To show this it is sufficient to prove that the sum of diagonal elements of matrix F equals p. This is proved immediately by writing explicitely the elements of this matrix. Thus, if by L we denote the matrix $\underset{np}{L} = XS^{-1} = \{l_{jk}\}$, then

$$l_{jk} = \sum_i^p x_{ij} s^{ik}; \qquad j = 1, 2, \ldots, n, \qquad k = 1, 2, \ldots, p.$$

Hence

$$f_{gh} = \sum_m^p l_{gm} x_{mh} = \sum_m^p \sum_i^p x_{ig} s^{im} x_{mh}; \qquad g, h = 1, 2, \ldots, n.$$

When $g = h$ then

$$f_{gg} = \sum_{m=1}^p \sum_{i=1}^p x_{ig} s^{im} x_{mg}.$$

Hence

$$\sum_{g}^{n} f_{gg} = \sum_{g}^{n} \sum_{m}^{p} \sum_{i}^{p} x_{ig}\, x_{mg}\, s^{im} = \sum_{m}^{p} \sum_{i}^{p} \left(\sum_{g}^{n} x_{ig}\, x_{mg} \right) s^{im} = \sum_{m}^{p} \left(\sum_{i}^{p} s_{im}\, s^{im} \right) = p,$$

since on account of $S = X^* X$ we obtain

$$s_{im} = \sum_{g}^{n} x_{ig}\, x_{mg}\,.$$

3°a. Note that

$$E\,(y - X\hat{\beta})^*\,(y - X\hat{\beta}) = (n-p)\,\sigma^2. \tag{68}$$

In fact, by $E\,(y - X\hat{\beta})^*\,(y - X\hat{\beta}) = Ee^*\,e - Ee^*\,XS^{-1}\,X^*\,e$ and by (64a) we have $Ee^*\,e = n\sigma^2$, and hence the expression (68) follows from (67).

4°. For any β_0 —

$$E\,(y - X\beta_0)^*\,(y - X\beta_0) = (\beta - \beta_0)^*\,S\,(\beta - \beta_0) + n\sigma^2. \tag{69}$$

Proof: The result (69) is obtained immediately with the help of the relation (68), of the identity (cf. [6])

$$(y - X\beta_0)^*\,(y - X\beta_0) = (y - X\hat{\beta})^*\,(y - X\hat{\beta}) + (\hat{\beta} - \beta_0)^*\,S\,(\hat{\beta} - \beta_0)$$

and of the matrix expression (cf. loc. cit.)

$$E\,(\hat{\beta} - \beta_0)^*\,S\,(\hat{\beta} - \beta_0) = (\beta - \beta_0)^*\,S\,(\beta - \beta_0) + p\sigma^2.$$

5°. The matrices of covariances between the unconditional estimates $\hat{\gamma}$ and the $\hat{\delta}$'s for parameters γ and δ appearing in the multiple regression model of the form $E\,(y) = X\beta = X_1\,\gamma + X_2\,\delta$

$$\left(\text{where, as usual, } \underset{np}{X} = \left[\underset{nq}{X_1} \,\vdots\, \underset{n,\,p-q}{X_2} \right] \text{ and } \underset{p1}{\beta} = \left[\begin{array}{c} \gamma \\ \hdashline \delta \end{array} \right] \right) \text{ are:}$$

$$E\,(\hat{\gamma} - \gamma)\,(\hat{\gamma} - \gamma)^* = \sigma^2 \cdot S^{11}, \tag{70}$$

$$E\,(\hat{\gamma} - \gamma)\,(\hat{\delta} - \delta)^* = \sigma^2 \cdot S^{12}, \tag{71}$$

$$E\,(\hat{\delta} - \delta)\,(\hat{\gamma} - \gamma)^* = \sigma^2 \cdot S^{21}, \tag{72}$$

$$E\,(\hat{\delta} - \delta)\,(\hat{\delta} - \delta)^* = \sigma^2 \cdot S^{22}, \tag{73}$$

where according to sec. 2: the symbol σ^2 denotes common variance of each of n components $y_1, y_2, \ldots, y_n$ of vector y, and S^{11}, S^{12}, S^{21} and S^{22} are submatrices of S^{-1} (cf. (8)).

Proof: It is known (cf. [6]) that covariance matrix of vector $\hat{\beta}$ is equal to $M = E(\hat{\beta}-\beta)(\hat{\beta}-\beta)^* = \sigma^2 S^{-1}$. Utilizing the fact that $\hat{\beta} = \begin{bmatrix} \hat{\gamma} \\ \hat{\delta} \end{bmatrix}$ we have on the one hand

$$(74)\qquad E\begin{bmatrix} \hat{\gamma}-\gamma \\ \hat{\delta}-\delta \end{bmatrix}\begin{bmatrix} \hat{\gamma}-\gamma \\ \hat{\delta}-\delta \end{bmatrix}^* = E\begin{bmatrix} \hat{\gamma}-\gamma \\ \hat{\delta}-\delta \end{bmatrix}[(\hat{\gamma}-\gamma)^* \,\vdots\, (\hat{\delta}-\delta)] =$$

$$= \begin{bmatrix} E(\hat{\gamma}-\gamma)(\hat{\gamma}-\gamma)^* & E(\hat{\gamma}-\gamma)(\hat{\delta}-\delta)^* \\ E(\hat{\delta}-\delta)(\hat{\gamma}-\gamma)^* & E(\hat{\delta}-\delta)(\hat{\delta}-\delta)^* \end{bmatrix}$$

and on the other hand

$$(75)\qquad M = \sigma^2 \cdot S^{-1} = \sigma^2 \begin{bmatrix} S^{11} & S^{12} \\ S^{21} & S^{22} \end{bmatrix} = \begin{bmatrix} \sigma^2 S^{11} & \sigma^2 S^{12} \\ \sigma^2 S^{21} & \sigma^2 S^{22} \end{bmatrix}.$$

The comparison of corresponding submatrices given in (74) and (75) leads to the covariance matrices in the form (70)-(73).

Remark 1. In particular case when the sets of parameters represented by vectors γ and δ are orthogonal i. e. when the matrix $B = X_1^* X_2 = O$, we deduce using the relations (14)—(17):

$$(76)\qquad S^{11} = A^{-1} = (X_1^* X_1)^{-1}, \qquad S^{12} = S^{21} = O, \qquad S^{22} = D^{-1} = (X_2^* X_2)^{-1}.$$

Consequently the covariance matrices (70)—(73) are equal to:

$$(77)\qquad E(\hat{\gamma}-\gamma)(\hat{\gamma}-\gamma)^* = \sigma^2 A^{-1},$$

$$(78)\qquad E(\hat{\gamma}-\gamma)(\hat{\delta}-\delta)^* = E(\hat{\delta}-\delta)(\hat{\gamma}-\gamma)^* = O,$$

$$(79)\qquad E(\hat{\delta}-\delta)(\hat{\delta}-\delta)^* = \sigma^2 D^{-1}.$$

6° Under the assumptions given in 5° for any parameter $\beta_0 = \begin{bmatrix} \gamma_0 \\ \delta_0 \end{bmatrix}$ we have:

$$(80)\qquad E(\hat{\gamma}-\gamma_0)(\hat{\gamma}-\gamma_0)^* = \sigma^2 \cdot S^{11} + (\gamma-\gamma_0)(\gamma-\gamma_0)^*,$$

$$(81)\qquad E(\hat{\gamma}-\gamma_0)(\hat{\delta}-\delta_0)^* = \sigma^2 \cdot S^{12} + (\gamma-\gamma_0)(\delta-\delta_0)^*,$$

$$(82)\qquad E(\hat{\delta}-\delta_0)(\hat{\gamma}-\gamma_0)^* = \sigma^2 \cdot S^{21} + (\delta-\delta_0)(\gamma-\gamma_0)^*,$$

$$(83)\qquad E(\hat{\delta}-\delta_0)(\hat{\delta}-\delta_0)^* = \sigma^2 \cdot S^{22} + (\delta-\delta_0)(\delta-\delta_0)^*,$$

where $\hat{\beta} = \begin{bmatrix} \hat{\gamma} \\ \hat{\delta} \end{bmatrix}$.

Proof: Using the relation (71) we obtain successively:

$$E(\hat{\gamma}-\gamma_0)(\hat{\delta}-\delta_0)^* = E\hat{\gamma}\hat{\delta}^* - E\hat{\gamma}\delta_0^* - E\gamma_0\hat{\delta}^* + \gamma_0\delta_0^* =$$

$$= E\hat{\gamma}\hat{\delta}^* - \gamma\delta_0^* - \gamma_0\delta^* + \gamma_0\delta_0^* = \sigma^2 S^{12} + \gamma\delta^* - \gamma\delta_0^* - \gamma_0\delta^* + \gamma_0\delta_0^* =$$

$$= \sigma^2 S^{12} + (\gamma-\gamma_0)(\delta-\delta_0)^*.$$

Thus we have proved the relation (81). The three remaining relations: (80), (82) and (83) may be proved in the same way; it is sufficient only to consider the relations (70), (72) and (73).

7°. Under the assumptions given in 5° we have for any parameter (any vector) γ_0 and any parameter (any vector) δ_0, respectively:

$$(84)\qquad E(\hat{\gamma}-\gamma_0)^*(S^{11})^{-1}(\hat{\gamma}-\gamma_0) = (\gamma-\gamma_0)^*(S^{11})^{-1}(\gamma-\gamma_0) + q\sigma^2,$$

$$(85)\qquad E(\hat{\delta}-\delta_0)^*(S^{22})^{-1}(\hat{\delta}-\delta_0) = (\delta-\delta_0)^*(S^{22})^{-1}(\delta-\delta_0) + (p-q)\sigma^2,$$

where γ_0 is any chosen set of β_i for $i=1, 2, ..., q$ and δ_0 is any chosen set of β_i for $i=q+1,\ q+2, ..., p$.

Proof: To prove the relation (84), let us put for brevity $\hat{\psi}=\hat{\gamma}-\gamma_0$. Let l^{ij} be the elements of matrix $(S^{11})^{-1}$, $i, j = 1, 2, ..., q$. Considering $E(\hat{\gamma})=\gamma$ and the relation (64) we obtain:

$$E(\hat{\gamma}-\gamma_0)^*(S^{11})^{-1}(\hat{\gamma}-\gamma_0) = (\gamma-\gamma_0)^*(S^{11})^{-1}(\gamma-\gamma_0) + \sum_i^q\sum_j^q l^{ij}\,cov(\hat{\psi}_i, \hat{\psi}_j).$$

Let us determine now the covariance matrix for vector $\hat{\psi}$. Using $\hat{\psi}-E(\hat{\psi})=\hat{\gamma}-\gamma$ and (70), we see that it is equal to

$$E(\hat{\psi}-E(\hat{\psi}))(\hat{\psi}-E(\hat{\psi}))^* = E(\hat{\gamma}-\gamma)(\hat{\gamma}-\gamma)^* = \sigma^2 S^{11}.$$

Thus we have $cov(\hat{\psi}_i, \hat{\psi}_j) = \sigma^2 l_{ij}$, so that

$$E(\hat{\gamma}-\gamma_0)^*(S^{11})^{-1}(\hat{\gamma}-\gamma_0) = (\gamma-\gamma_0)^*(S^{11})^{-1}(\gamma-\gamma_0) + \sigma^2\sum_i^q\sum_j^q l^{ij} l_{ij} =$$

$$= (\gamma-\gamma_0)^*(S^{11})^{-1}(\gamma-\gamma_0) + q\sigma^2.$$

The relation (85) is proved in the same manner.

7°*a*. From the relation (84) it is evident that, when the null hypothesis $\gamma=\gamma_0$ is true then $E(\hat{\gamma}-\gamma_0)^*(S^{11})^{-1}(\hat{\gamma}-\gamma_0)=q\sigma^2$. Using this we shall now show that test based on random variable (47) is right-tailed. In fact, it is sufficient to prove that when $\gamma\neq\gamma_0$ the quadratic form $(\gamma-\gamma_0)^*(S^{11})^{-1}(\gamma-\gamma_0)$ is positive definite.

Proof: First consider two known theorems:

Theorem I. If the quadratic form is positive definite, the reciprocal form is also positive definite. *Theorem* II. Any principal submatrix of a positive definite symmetric matrix is a positive definite matrix.

To prove that matrix $(S^{11})^{-1}$ is positive, it is sufficient to show (by theorem I) that its reciprocal matrix i. e. matrix S^{11} is positive definite. Since S^{11} is the principal submatrix of matrix $S = \underset{pp}{S} = X^*X$, then by theorem II it is sufficient to prove that matrix S is positive definite. We exclude here the trivial case $X = O$ when the matrix S^{-1} does not exist. Let $u = \underset{p1}{u} \neq 0$ be any vector with p components, not all of which are zeros. Then $u^*Su = u^*X^*Xu = (Xu)^*(Xu)$ is a quadratic form (sum of squares) which means that the matrix S is positive definite. It has been thus proved that the test based on random variable (47) is right-tailed.

8°. Under the assumptions given in 5° the covariance matrices of conditional estimates $\tilde{\gamma}$ and $\tilde{\delta}$ which are obtained when $\delta = \delta_0$ and $\gamma = \gamma_0$ respectively, are equal to

$$E\,[\tilde{\gamma} - E\,(\tilde{\gamma})]\,[\tilde{\gamma} - E\,(\tilde{\gamma})]^* = \sigma^2 \cdot A^{-1}, \tag{86}$$

$$E\,[\tilde{\delta} - E\,(\tilde{\delta})]\,[\tilde{\delta} - E\,(\tilde{\delta})]^* = \sigma^2 \cdot D^{-1}. \tag{87}$$

Proof: First we shall deduce the relation (86), whose left side we shall denote by C. Since $E\,(\hat{\gamma}) = \gamma$ and $E\,(\hat{\delta}) = \delta$, and the conditional estimate $\tilde{\gamma} = \hat{\gamma} + A^{-1}B\,(\hat{\delta} - \delta_0)$ (cf. formula (55)), we can write $\tilde{\gamma} - E\,(\tilde{\gamma}) = (\hat{\gamma} - \gamma) + A^{-1}B\,(\hat{\delta} - \delta)$ and hence

$$\begin{aligned} C &= E\,\{(\hat{\gamma} - \gamma) + A^{-1}B\,(\hat{\delta} - \delta)\}\,\{(\hat{\gamma} - \gamma) + A^{-1}B\,(\hat{\delta} - \delta)\}^* = \\ &= E\,(\hat{\gamma} - \gamma)(\hat{\gamma} - \gamma)^* + E\,(\hat{\gamma} - \gamma)(\hat{\delta} - \delta)^* B^* A^{-1} + A^{-1}B \cdot E\,(\hat{\delta} - \delta)(\hat{\gamma} - \gamma)^* + \\ &\qquad + A^{-1}B \cdot E\,(\hat{\delta} - \delta)(\hat{\delta} - \delta)^* B^* A^{-1}. \end{aligned}$$

Utilizing the expressions (70)—(73) and reducing the corresponding matrix relations we obtain:

$C = \sigma^2\,(S^{11} + S^{12}B^*A^{-1} + A^{-1}BS^{21} + A^{-1}BS^{22}B^*A^{-1})$. Note further that from the relations (10) and (12) follow the corresponding matrix relations: $S^{12}B^* = I - S^{11}A$ and $BS^{22} = -AS^{12}$. Hence $S^{12}B^*A^{-1} = A^{-1} - S^{11}$, $A^{-1}BS^{21} = A^{-1} - S^{11}$ and $A^{-1}BS^{22}B^*A^{-1} = -A^{-1}\,(AS^{12})\,B^*A^{-1} = -S^{12}B^*A^{-1} = -(I - S^{11}A)\,A^{-1} = S^{11} - A^{-1}$. Using these relations and performing appropriate reduction we obtain $C = E\,[\tilde{\gamma} - E\,(\tilde{\gamma})]\,[\tilde{\gamma} - E\,(\tilde{\gamma})]^* = \sigma^2 A^{-1}$. Thus we have proved the relation (86).

To prove the relation (87) we proceed similarly, but in place of the relations (10) and (12) we use the relations (11) and (13).

Remark 2. Note that the right sides of (86) and (77) are identical, and also the right sides at (87) and (79). This means that covariance matrix of unconditional estimate $\hat{\gamma}$, determined under the assumption that vectors γ and δ are orthogonal, is identical with covariance matrix of conditional estimate $\tilde{\gamma}$ determined without assumption of orthogonality. Mutatis mutandis the same is true of the vector $\tilde{\delta}$. The estimate $\tilde{\gamma}$ is, of course, determined under the condition that $\delta=\delta_0$ and estimate $\tilde{\delta}$ is determined when $\gamma=\gamma_0$.

9°. Under the asumptions given in 5° the following equalities hold:

$$(88)\quad E(\tilde{\gamma}-\gamma_0)^* A(\tilde{\gamma}-\gamma_0)= \\ =\{\gamma-\gamma_0+A^{-1}B(\delta-\delta_0)\}^* A\{\gamma-\gamma_0+A^{-1}B(\delta-\delta_0)\}+q\sigma^2$$

$$(89)\quad E(\tilde{\delta}-\delta_0)^* D(\tilde{\delta}-\delta_0)= \\ =\{\delta-\delta_0+D^{-1}B^*(\gamma-\gamma_0)\}^*\cdot D\{\delta-\delta_0+D^{-1}B^*(\gamma-\gamma_0)\}+(p-q)\sigma^2$$

where $\underset{q1}{\tilde{\gamma}}=\hat{\gamma}+A^{-1}B(\hat{\delta}-\delta_0)$ is the estimate of the parameter γ under assumption that $\delta=\delta_0$ while $\underset{p-q,1}{\tilde{\delta}}=\hat{\delta}+D^{-1}B^*(\gamma-\gamma_0)$ is the estimate of the parameter δ under assumption that $\gamma=\gamma_0$.

Proof: First we shall prove the relation (88). Let $\hat{\varphi}=\hat{\gamma}-\gamma_0$. Considering formula (64) which gives the expected value of quadratic form we have:

$$E(\tilde{\gamma}-\gamma_0)^* A(\tilde{\gamma}-\gamma_0)=E\tilde{\varphi}^* A\varphi=\{\gamma-\gamma_0+A^{-1}B(\delta-\delta_0)\}^*\cdot \\ \cdot A\{\gamma-\gamma_0+A^{-1}B(\delta-\delta_0)\}+\sum_i^q\sum_j^q a_{ij}\operatorname{cov}(\tilde{\varphi}_i,\tilde{\varphi}_j)$$

where a_{ij}; $i,j=1,2,\ldots,q$; are the elements of the matrix $\underset{qq}{A}=X_1^*X_1$. Since the covariance matrix of vector $\tilde{\varphi}$ (on the account of $\tilde{\varphi}-E(\tilde{\varphi})=\tilde{\gamma}-\gamma_0-E(\tilde{\gamma}-\gamma_0)=\tilde{\gamma}-E(\tilde{\gamma})$ and (86)) is equal to

$$E[\tilde{\varphi}-E(\tilde{\varphi})][\tilde{\varphi}-E(\tilde{\varphi})]^*=E[\tilde{\gamma}-E(\tilde{\gamma})][\tilde{\gamma}-E(\tilde{\gamma})]^*=\sigma^2A^{-1},$$

it follows that

$$\sum_i^q\sum_j^q a_{ij}\operatorname{cov}(\tilde{\varphi}_i,\tilde{\varphi}_j)=\sum_i^q\sum_j^q\sigma^2\cdot a_{ij}\cdot a^{ij}=q\sigma^2,$$

which completes the proof of the relation (88).

In the same way we prove the relation (89) taking into consideration the result (87).

$9^\circ a$. From the relations (88) and (89), where it has been put

(90) $$\gamma_0 = E(\tilde{\gamma}) = E\{\hat{\gamma} + A^{-1} B(\hat{\delta} - \delta_0)\} = \gamma + A^{-1} B(\delta - \delta_0)$$

and

(91) $$\delta_0 = E(\tilde{\delta}) = E\{\hat{\delta} + D^{-1} B^*(\hat{\gamma} - \gamma_0)\} = \delta + D^{-1} B^*(\gamma - \gamma_0)$$

respectively follow directly two expressions:

(92) $$E[\tilde{\gamma} - E(\tilde{\gamma})]^* A[\tilde{\gamma} - E(\tilde{\gamma})] = q\sigma^2$$

(93) $$E[\tilde{\delta} - E(\tilde{\delta})]^* D[\tilde{\delta} - E(\tilde{\delta})] = (p-q)\sigma^2.$$

$9^\circ b$. When the sets of parameters represented by the vectors γ and δ are orthogonal, we have

(94) $$E(\hat{\gamma} - \gamma_0)^* A(\hat{\gamma} - \gamma_0) = (\gamma - \gamma_0)^* A(\gamma - \gamma_0) + q\sigma^2$$

(95) $$E(\hat{\delta} - \delta_0)^* D(\hat{\delta} - \delta_0) = (\delta - \delta_0)^* D(\delta - \delta_0) + (p-q)\sigma^2.$$

The relations (94) and (95) are obtained directly from expressions (88) and (89) under the condition $B = X_1^* X_2 = O$ (which determine the orthogonality of parameter sets γ and δ). Under this condition the unconditional and conditional estimates are identical i. e. $\tilde{\gamma} = \hat{\gamma}$ and $\tilde{\delta} = \hat{\delta}$.

$9^\circ c$. When the sets of parameters determined by the vectors $[\psi_1], [\psi_2], \ldots, [\psi_r]$ (with $p_1, p_2, \ldots, p_r$ components respectively) are reciprocally orthogonal in the linear regression model

(96) $$\mu = E(y) = X\beta = \sum_i^r X_i[\psi_i]$$

where

$$\underset{np}{X} = \left[\underset{np_1}{X_1} \,\vdots\, \underset{np_2}{X_2} \,\vdots \cdots \vdots\, \underset{np_r}{X_r}\right] \quad \text{and} \quad \beta = \begin{bmatrix} [\psi_1] \\ [\psi_2] \\ \vdots \\ [\psi_r] \end{bmatrix} \quad \text{and} \quad p = \sum_i^r p_i,$$

then

(97) $$E([\tilde{\psi}_i] - [\psi_i^0])^* X_i^* X_i([\tilde{\psi}_i] - [\psi_i^0]) = ([\psi_i] - [\psi_i^0])^* X_i^* X_i([\psi_i] - [\psi_i^0]) + p_i\sigma^2$$

where $[\tilde{\psi}_i] = (X_i^* X_i^*)^{-1} X_i^* y$, $i = 1, 2, \ldots, r$, is the estimate of the parameter $[\psi_i]$ under the condition that the remaining parameters equal zero. As usual, σ^2 is the common variance of each of n independent and normally distributed random variables $y_1, y_2, \ldots, y_n$ that are the components of vector y whose expected value is μ.

To prove the relation (97) it is sufficient to substitute $X_i^* X_i$ for A and $[\tilde{\psi}_i] = [\hat{\psi}_i]$, and $[\psi_i^0]$ for γ and γ_0 respectively.

Remark 3. Note that the expression $([\widetilde{\psi}_i] - [\psi_i^0])^* X_i^* X_i ([\widetilde{\psi}_i] - [\psi_i^0])$ whose expected value is determined by the formula (97), is the denominator of the random variable (47) which may be used for testing the null hypothesis $[\psi_i] = [\psi_i^0]$, $i = 1, 2, ..., r$. In particular, when the null hypothesis $[\psi_i] = 0$ is true, this expression becomes equal to $[\widetilde{\psi}_i]^* X_i^* X_i [\widetilde{\psi}_i]$ and constitutes the sum of squares of regression y on X_i.

10°. Under the asumptions given in 5° consider q independent linear relations of parameters $\beta_1, \beta_2, ..., \beta_p$ defined by the matrix relation $Ł\beta = \varphi$, where the rank of matrix $Ł = \underset{qp}{Ł}$ is equal to q and vector $\varphi = \underset{q1}{\varphi}$. The following relation holds:

$$(98) \qquad E(Ł\hat{\beta} - \varphi_0)(Ł\hat{\beta} - \varphi_0)^* = ŁS^{-1}Ł^* \sigma^2 + (Ł\beta - \varphi_0)(Ł\beta - \varphi_0)^*$$

where φ_0 is any vector and $\hat{\beta}$ is, as usual, equal to $S^{-1} X^* y$.

Proof: Using the relation $E\hat{\beta}\hat{\beta}^* = \sigma^2 \cdot S^{-1} + \beta\beta^*$, which is obtained from the known form of covariance matrix of vector $\hat{\beta}$ i. e. from expression $E(\hat{\beta} - \beta)(\hat{\beta} - \beta)^* = \sigma^2 \cdot S^{-1} = E\hat{\beta}\hat{\beta}^* - \beta\beta^*$, we obtain successively $E(Ł\hat{\beta} - \varphi_0) \times (Ł\beta - \varphi_0)^* = ŁE\hat{\beta}\hat{\beta}^* Ł^* - EŁ\hat{\beta}\varphi_0^* - \varphi_0 E\hat{\beta}^* Ł^* + \varphi_0\varphi_0^* = Ł(\sigma^2 S^{-1} + \beta\beta^*)Ł^* - Ł\beta\varphi_0^* - \varphi_0\beta^* Ł^* + \varphi_0\varphi_0^* = ŁS^{-1}Ł^* \sigma^2 + (Ł\beta - \varphi_0)(Ł\beta - \varphi_0)^*$.

10°*a*. In particular case when φ_0 is the true value φ of the product $Ł\beta$ i. e. when $\varphi_0 = \varphi = Ł\beta = E(Ł\hat{\beta})$, then the matrix expression (98) defines the covariance matrix of linear parametric function $Ł\hat{\beta}$. It can be immediately seen that this matrix is equal to

$$(99) \qquad E(Ł\hat{\beta} - \varphi)(Ł\hat{\beta} - \varphi)^* = ŁS^{-1}Ł^* \cdot \sigma^2$$

where $\hat{\beta} = S^{-1} X^* y$.

Remark 4. The covariance matrix (99) is applied in the formula (56) to determine the random variable F in the theorem 3.

11°. Under the assumptions given in 10° we have

$$(100) \quad E(Ł\hat{\beta} - \varphi_0)^*(ŁS^{-1}Ł^*)^{-1}(Ł\hat{\beta} - \varphi_0) = (Ł\beta - \varphi_0)^*(ŁS^{-1}Ł^*)^{-1}(Ł\beta - \varphi_0) + q\sigma^2$$

Proof: Let $\hat{\psi} = Ł\hat{\beta} - \varphi_0$ be the vector and let b_{ij}, $i, j = 1, 2, ..., q$, be the elements of matrix $ŁS^{-1}Ł^*$. Then by the formula (64) and by $E(Ł\hat{\beta}) = Ł\beta$ we obtain:

$$E(Ł\hat{\beta} - \varphi_0)^*(ŁS^{-1}Ł^*)^{-1}(Ł\hat{\beta} - \varphi_0) =$$

$$= (Ł\beta - \varphi_0)^*(ŁS^{-1}Ł^*)^{-1}(Ł\beta - \varphi_0) - \sum_i^q \sum_j^q b^{ij} cov(\hat{\psi}_i, \hat{\psi}_j),$$

which is denoted by M. Let us find the covariance matrix of vector $\hat{\psi}$. Note that

$$E(\hat{\psi} - E(\hat{\psi}))(\hat{\psi} - E(\hat{\psi}))^* = E\hat{\psi}\hat{\psi}^* - (E\hat{\psi})(E\hat{\psi}^*) =$$
$$= E(Ł\hat{\beta} - \varphi_0)(Ł\hat{\beta} - \varphi_0)^* - E(Ł\hat{\beta} - \varphi_0) \cdot E(Ł\hat{\beta} - \varphi_0)^* = ŁS^{-1}Ł^* \cdot \sigma^2,$$

which follows immediately from the relation (98). Thus

$$cov(\hat{\psi}_i, \hat{\psi}_j) = b_{ij}\sigma^2 \quad \text{and} \quad \sum_i^q \sum_j^q b^{ij} cov(\hat{\psi}_i, \hat{\psi}_j) = \sum_i^q \sum_j^q \sigma^2 b^{ij} b_{ij} = q\sigma^2.$$

Hence we obtain

$$M = (Ł\beta - \varphi_0)^* (ŁS^{-1}Ł^*)^{-1} (Ł\beta - \varphi_0) + q\sigma^2.$$

11°*a*. In particular case when φ_0 is the true value φ of the parametric function $Ł\beta$ i.e. when $\varphi_0 = \varphi = Ł\beta = E(Ł\hat{\beta})$, we have

$$E(Ł\hat{\beta} - \varphi_0)^* (ŁS^{-1}Ł^*)^{-1} (Ł\hat{\beta} - \varphi_0) = q\sigma^2. \tag{101}$$

11°*b*. The test based on random variable F given in (56) is right-tailed.

Proof: The comparison of relation (100) with result (101) indicates that to prove that the test based on random variable (56) is right-tailed it is sufficient to show that the quadratic form

$$(Ł\beta - \varphi_0)^* (ŁS^{-1}Ł^*)^{-1} (Ł\beta - \varphi_0)$$

is positive definite, in other words it is sufficient to prove by the theorem I (cf. p. 44) that the reciprocal of the quadratic form of the latter (i.e. of the form $u^*ŁS^{-1}Ł^*u = W$, where $u = \underset{q1}{u}$ is any vector non-equal to zero), is positive definite. Since the quadratic form W may be given as $W = (Ł^*u)^* S^{-1} (Ł^*u) = m^*S^{-1}m$, where $m = \underset{p1}{m} = \underset{pq}{Ł^*}\underset{q1}{u}$, it suffices to prove that the matrix S^{-1} is potitive definite. The latter statement follows directly frcm the theorem I and from the fact that matrix S is positive definite (cf. proof on p. 44 in 7°α).

Remark 5. Note that the quadratic form, whose expected value is given in the relation (100), appears in the formula (56) that gives the random variable F in the theorem 3.

12°. Under the assumptions of the theorem 4 the following relation holds

$$E\{(T\hat{\beta} - \tau)^* (TS^{-1}T^*)^{-1} (T\hat{\beta} - \tau) - (G\hat{\beta} - \eta)^* (GS^{-1}G^*)^{-1} (G\hat{\beta} - \eta)\} =$$
$$= (T\beta - \tau)^* (TS^{-1}T^*)^{-1} (T\beta - \tau) + h\sigma^2. \tag{102}$$

Proof: Note that under the assumptions given in the theorem 4 $E(\underset{gp}{G}\hat{\beta})=G\cdot\beta=\underset{g1}{\eta}$, and for any vector $\underset{h1}{\nu}$ we have generally $E(\underset{hp}{H}\hat{\beta})=H\cdot\beta\neq\nu$. Hence from the fact that $T=\begin{bmatrix} G \\ \cdots \\ H \end{bmatrix}$ and $\tau=\begin{bmatrix} \eta \\ \cdot \\ \nu \end{bmatrix}$, it follows that generally we have $E(T\hat{\beta})=T\beta\neq\tau$, and the equality here holds only, when the null hypothesis $H\beta=\nu$ is true. Hence using the relation (100) and noting that $\underset{g+h,p}{r(T)}=g+h$ we find

$$E\{(T\hat{\beta}-\tau)^*(TS^{-1}T^*)^{-1}(T\hat{\beta}-\tau)\}=(T\beta-\tau)^*(TS^{-1}T^*)^{-1}(T\beta-\tau)+(g+h)\sigma^2.$$

Further, since $r(G)=g$, we obtain by relation (101) the value

$$E(G\hat{\beta}-\eta)^*(GS^{-1}G^*)^{-1}(G\hat{\beta}-\eta)=g\sigma^2.$$

Since the difference of the latter two expressions constitutes (102), the theorem is proved.

12°*a*. Note that when the null hypothesis $H\beta=\nu$ is true, the relation (102) (on the account of $E(H\hat{\beta})=H\beta=\nu$, and what follows $E(T\hat{\beta}=\tau)$ assumes the form

$$(103)\quad E\{(T\hat{\beta}-\tau)^*(TS^{-1}T^*)^{-1}(T\hat{\beta}-\tau)-(G\hat{\beta}-\eta)^*(GS^{-1}G^*)^{-1}(G\hat{\beta}-\eta)\}=h\sigma^2$$

where h is the rank of matrix $\underset{hp}{H}$ appearing in the null hypothesis.

12°*b*. The test based on random variable (57) which appears in the theorem 4 is right-tailed. To prove this, it should be noted that from the relations (102) and (103) it follows that it is sufficient to prove that, when $T\beta\neq\tau$, the quadratic form $(T\beta-\tau)^*(TS^{-1}T^*)^{-1}(T\beta-\tau)$ is positive definite. The latter statement is proved similarly as in 11°*b*.

13. Let $\beta=Wa$, where $W=\underset{pm}{W}$ is any matrix of rank m, $1\leqslant m\leqslant p$, and $a=\underset{m1}{a}$ is any vector with m parameters $a_1, a_2, \ldots, a_m$. The relation $\beta=Wa$ expresses p parameters $\beta_1, \beta_2, \ldots, \beta_p$ in terms of m parameters $a_1, a_2, \ldots, a_m$. Under these assumptions and under those of theorem 5 the following relation holds:

$$(104)\quad E(S\hat{\beta})^*[S^{-1}-W(W^*SW)^{-1}W^*]S\hat{\beta}=$$
$$=(S\beta)^*[S^{-1}-W(W^*SW)^{-1}W^*]S\beta+(p-m)\sigma^2.$$

Proof: Let L denote the left side of the expression (104). Utilizing the fact that the covariance matrix of vector $\hat{\beta}$ is equal to $S^{-1}\cdot\sigma^2$ and using the formula (64), we obtain:

$$(105)\quad E\hat{\beta}^* S\hat{\beta} = \beta^* S\beta + \sum_i^p \sum_j^p s_{ij}\, cov(\hat{\beta}_i, \hat{\beta}_j) = $$

$$= \beta^* S\beta + \sum_i^p \sum_j^p s_{ij}\, s^{ij}\, \sigma^2 = \beta^* S\beta + p\sigma^2.$$

Putting $\underset{m1}{\hat{\varphi}} = W^* S\hat{\beta}$, we may state that covariance matrix of vector $\hat{\varphi}$ is equal to $\sigma^2 \cdot W^* SW$. In fact,

$$(106)\quad E(\hat{\varphi} - E(\hat{\varphi}))(\hat{\varphi} - E(\hat{\varphi}))^* = E(W^* S\hat{\beta} - W^* S\beta)(W^* S\hat{\beta} - W^* S\beta)^* =$$

$$= W^* S \cdot E\{(\hat{\beta} - \beta)(\hat{\beta} - \beta)^*\} \cdot SW = W^* SS^{-1} \sigma^2 \cdot SW = \sigma^2 W^* SW.$$

Applying expression (106) and formula (64), and putting $(W^* SW)^{-1} = \{h^{ij}\}$, we find

$$(107)\quad E(S\hat{\beta})^* W(W^* SW)^{-1} W^* S\hat{\beta} = E(W^* S\hat{\beta})^* (W^* SW)^{-1} W^* S\hat{\beta} =$$

$$= (W^* S\beta)^* (W^* SW)^{-1} W^* S\beta + \sum_i^m \sum_j^m h^{ij}\, cov(\hat{\varphi}_i, \hat{\varphi}_j) =$$

$$= (W^* S\beta)^* (W^* SW)^{-1} W^* S\beta + m\sigma^2.$$

Hence by the formulae (105) and (107) we get the result (104). In fact,

$$L = E\hat{\beta}^* S\hat{\beta} - E(S\hat{\beta})^* W(W^* SW)^{-1} W^* S\hat{\beta} =$$

$$= \beta^* S\beta + p\sigma^2 - (W^* S\beta)^* (W^* SW)^{-1} W^* S\beta - m \cdot \sigma^2 =$$

$$= (S\beta)^* [S^{-1} - W(W^* SW)^{-1} W^*] S\beta + (p - m)\sigma^2.$$

13°*a*. Let the null hypothesis $\beta = \underset{p,p-q}{U} \cdot \underset{p-q,1}{\vartheta}$ be true under the assumptions of the theorem 5. Then the relation (104) assumes the form

$$(108)\quad E(S\hat{\beta})^* [S^{-1} - U(U^* SU)^{-1} U^*] S\hat{\beta} = q\sigma^2.$$

To prove this we substitute in (104): $p-q$, U and $U\vartheta$ for m, W and β respectively; we thus obtain

$$E(S\hat{\beta})^* [S^{-1} - W(W^* SW)^{-1} W^*] S\hat{\beta} = (S\beta)^* S^{-1} S\beta -$$

$$- (S\beta)^* U(U^* SU)^{-1} U^* S\beta + q\sigma^2 = (U\vartheta)^* SU\vartheta + q\sigma^2 -$$

$$- (SU\vartheta)^* U(U^* SU)^{-1} U^* SU\vartheta = \vartheta^* U^* SU\vartheta - \vartheta^* U^* SU\vartheta + q\sigma^2 = q\sigma^2.$$

$13°b$. The test based on random variable (58) given in the theorem 5 is right-tailed [1]).

Proof: The random variable (60), which has the form of the random variable (47) in the theorem 2 is equivalent to the random variable (58). Since the test based on random variable (47) (cf. p. 44) is right-tailed we conclude that the test in question is also right-tailed. The proof is concluded.

Note. It should be noted that the null hypothesis $\beta = \underset{p,p-q}{U} \cdot \underset{p-q,1}{\vartheta}$ is tested against the alternative hypothesis that β equals any $\underset{p,p-q}{W} \cdot \underset{p-q,1}{\vartheta}$, where $W \neq U$, meaning that p parameters $\beta_1, \beta_2, \ldots, \beta_p$ of the given multiple regression model may be represented by the linear combinations of $p-q$ parameters, and that these combinations are different from those of the null hypothesis.

Note that the tests of significance based on random variables F given in the theorems 2, 3, 4 and 5 are right-tailed; this follows from the fact (cf. [8]) that these F's may be expressed as the decreasing functions of random variable lambda which by the maximum likelihood ratio criterion may be used as the left-tailed test. Using in this paper the theorems of matrix calculus we have presented the direct proofs that the tests based on random variables F (given here) are right-tailed.

6. Applications

In the present section we shall give some examples of the applications of random variables F given in the matrix notation in the theorems 2, 3, 4 and 5 for testing the linear hypotheses. Since the random variables have the matrix forms, their deduction in the case of any particular concrete form of linear hypothesis in the given linear regression model (design of experiment) is different from that presented in the papers published so far.

We shall consider three types of the multiple regression models:

a) the model, which may be transformed into a model with matrix whose rank is equal to the number of independent parameters,

b) the model of the one-way classification, and

c) the model with the regression lines y on x in several populations.

---

[1]) From $13°b$ it follows that the quadratic form $(S\beta)^* [S^{-1} - W(W^* SW)^{-1} W^*] S\beta$ (cf. also (104)) is non-negative. In particular, considering the linear regression model $y_\alpha = x_{1\alpha}\beta_1 + x_{2\alpha}\beta_2$; $\alpha = 1, 2, \ldots, n$; and treating the matrix W as the column vector $W = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, we obtain the Schwarz's inequality for the sums: $(\Sigma x_1^2)(\Sigma x_2^2) \geqslant (\Sigma x_1 x_2)^2$.

Ad a). If in the multiple regression model $y = X\beta$ the number of independent parameters among p parameters $\beta_1, \beta_2, \ldots, \beta_p$ is equal to the rank r of the matrix X, the number of restrictions imposed on the parameters β is equal to $p-r$. As a result of the elimination of the $p-r$ dependent parameters β we obtain the model with r parameters. If the matrix rank of this model on such reparametrization does not change and remains equal to r, we obtain the situation which is being discussed in this paper i. e. we have the multiple regression model whose number of parameters is equal to the rank of the matrix formed from the coefficients of these parameters.

To illustrate such model, consider a model of the two-way classification with one element in each cell, i. e. design of randomized blocks. If the effects for the treatments and for the blocks are marked by a_i; $i = 1, 2, \ldots, c$; and by b_j; $j = 1, 2, \ldots, k$; respectively, the model assumes the form:

$$(109) \quad E(y_a) = a_1 x_{1a} + a_2 x_{2a} + \cdots + a_{c-1} x_{c-1,\,a} + a_c x_{ca} + b_1 x_{c+1,\,a} + \\ + b_2 x_{c+2,\,a} + \cdots + b_{k-1} x_{c+k-1,\,a} + b_k x_{c+k,\,a} + \mu'$$

where μ' (ordinary number) is the mean in population embracing all the elements arranged in ck cells. Note that when $a = 1, 2, \ldots, n$, we have simultaneously $i = 1, 2, \ldots, c$, and $j = 1, 2, \ldots, k$. Two restrictions

$$(110) \qquad \sum_i^c a_i = 0, \qquad \sum_j^k b_j = 0$$

are imposed on the parameters of the model. The assumptions with regard to matrix X are as follows: a) $x_{c+k+1,\,a} = 1$ for $a = 1, 2, \ldots, n$; b) when $i = 1$, $j = 1$, then $x_1 = 1$ and $x_{c+1} = 1$, and the remaining $x's$ except $x_{c+k+1,\,a}$ are equal to zero; c) when $i = 1$, $j = 2$, then $x_1 = 1$, $x_{c+2} = 1$, and the remaining $x's$ except $x_{c+k+1,\,a}$ are equal to zero; and so on. We are concerned with the testing of the null hypothesis that no differences exist between the treatments i. e. that $a_1 = a_2 = \ldots = a_c = 0$. It is easy to verify that under these assumptions the rank of the matrix X of the order $n \times (c+k+1)$ is equal to $c+k-1$ i. e. is less by two than the number $p = c+k+1$ of parameters in the model. Thus we cannot make direct use of the F test that follows from the theorem 2. However if we determine a_c and b_c from the restrictions (110) and if thus obtained expressions are introduced into relation (109) the following model with the $c+k-1$ parameters will be obtained

$$(111) \quad E(y_a) = a_1 z_{1a} + a_2 z_{2a} + \cdots + a_{c-1} z_{c-1,\,a} + b_1 z_{c+1,\,a} + b_2 z_{c+2,\,a} + \\ + \cdots + b_{k-1} z_{c+k-1,\,a} + \mu'$$

where

$$z_{lm} = x_{la} - x_{ca}; \qquad l = 1, 2, \ldots, c-1; \qquad z_{ma} = x_{ma} - x_{c+k,a};$$

$$m = c+1, \quad c+2, \ldots, c+k-1.$$

The rank of matrix $Z = \underset{n,\, c+k-1}{Z}$ is equal to the number of the parameters i. e. is equal to $c+k-1$. Putting

$$\underset{1,\, c-1}{\gamma^*} = [a_1, a_2, \ldots, a_{c-1}], \quad \underset{1k}{\delta^*} = [b_1, b_2, \ldots, b_{k-1}, \mu'] \quad \text{and} \quad \beta^* = [\gamma^* \vdots \delta^*]$$

we obtain the model

$$E(y) = Z\beta = Z_1\gamma + Z_2\delta, \tag{112}$$

where

$$Z = \Big[\underset{n,\, c-1}{Z_1} \vdots \underset{nk}{Z_2}\Big] \quad \text{and} \quad Z = \{z_{ua}\}, \quad u = 1, 2, \ldots, c-1; \quad a = 1, 2, \ldots, n.$$

Since the null hypothesis is: $\gamma = 0$ we may test it by using the random variable F given in the theorem 2.

The estimates of the parameters $a_1, a_2, \ldots, a_{c-1}, b_1, b_2, \ldots, b_{k-1}, \mu'$ are found from the relation $\hat{\beta} = (Z^*Z)^{-1}Z^*y$ which holds since the number of parameters is equal to $r(Z^*Z) = c+k-1$. Since $Z_1^*Z_2 = O$, which can be easily verified, the parameter sets γ and δ are orthogonal. The matrices Z, Z_1 and Z_2 play here the rôle of the matrices X, X_1 and X_2 respectively in the model discussed in the sec. 3. Hence, on the account of $B = Z_1^*Z_2 = O$ and the relation (14) we find $S^{11} = A^{-1} = (Z_1^*Z_1)^{-1}$, and the random variable (47) assumes the form:

$$F = \frac{\hat{\gamma}^* A \hat{\gamma}}{c-1} : \frac{(y - Z\hat{\beta})^*(y - Z\hat{\beta})}{(c-1)(k-1)}$$

with $c-1$ and $(c-1)(k-1)$ degrees of freedom, where, as it can easily be verified,

$$A = Z_1^*Z_1 = \begin{bmatrix} 2k & k & \ldots & k \\ k & 2k & \ldots & k \\ \cdot & \cdot & \cdot & \cdot \\ k & k & \ldots & 2k \end{bmatrix}.$$

Using the formula (22′), which in the actual case has the form $\hat{\gamma} = (Z_1^*Z_1)^{-1}Z_1^*y$, and noting that the determinant of the matrix A has value $|A| = |Z_1^*Z_1| = ck^{c-1}$ (by application of the well known algebraic formulae (cf. [18]), and next by obtaining

$$A^{-1}=\begin{bmatrix} c-1 & -1 & -1 & \dots & -1 \\ -1 & c-1 & -1 & \dots & -1 \\ -1 & -1 & c-1 & \dots & -1 \\ \dots & \dots & \dots & \dots & \dots \\ -1 & -1 & -1 & \dots & c-1 \end{bmatrix}$$

we get the known formulae for the estimates of parameters a_i: $\hat{a}_i=\overline{y}_i-\overline{y}$, $i=1,2,\dots,c-1$. Since it follows that $\hat{\gamma}^*=[\overline{y}_1-\overline{y}, \overline{y}_2-\overline{y}, \dots, \overline{y}_{c-1}-\overline{y}]$, by performing some simple calculations we get

$$\hat{\gamma}^* A \hat{\gamma}=k\sum_i^c(\overline{y}_i-\overline{y})^2.$$

The numerator of the random variable F is thus found. Writing the matrix Z in explicit form we derive without difficulty the denominator of F,

$$(y-Z\hat{\beta})^*(y-Z\hat{\beta})=\sum_i^c\sum_j^k(y_{ij}-\overline{y}_i-\overline{y}_j+\overline{y})^2.$$

Ad *b)*. Consider some examples of the testing of linear hypothesis in the case of the model of one-way classification under the assumptions given in the theorem 3.

We take p random samples from p populations in such a way that the i'th sample with n_i observed values, $i=1,2,\dots,p$, is drawn from i'th population. The total number of the observations is equal to $n=\sum_i^p n_i$. The multiple regression model $\mu=E(y)=X\beta$ expressed in symbols used in this paper takes the form:

$$\mu_\alpha=E(y_\alpha)=x_{1\alpha}\beta_1+x_{2\alpha}\beta_2+\dots+x_{p\alpha}\beta_p \tag{113}$$

where we put: $x_{1\alpha}=1$ for $\alpha=1,2,\dots,n_1$ and $x_{1\alpha}=0$ for the remaining α's; $x_{2\alpha}=1$ for $\alpha=n_1+1, n_1+2,\dots,n_1+n_2$ and $x_{2\alpha}=0$ for the remaining α's, etc. Let $\beta_1,\beta_2,\dots,\beta_p$ define the means in the corresponding populations. Then $\mu_\alpha=\beta_1$ for $\alpha=1,2,\dots,n_1$; $\mu_\alpha=\beta_2$ for n_1+1, $n_1+2,\dots$, $n_1+n_2,\dots$; $\mu_\alpha=\beta_p$ for

$$\alpha=\sum_{l=1}^{p-1}n_l+1,\quad \alpha=\sum_{l=1}^{p-1}n_l+2,\dots,\alpha=n.$$

It is obvious that the matrix $S=X^*X$ has the form $S=\begin{bmatrix} n_1, & 0,\dots,0 \\ 0, & n_2,\dots,0 \\ \dots & \dots \\ 0, & 0,\dots,n_p \end{bmatrix}$

and the reciprocal matrix of S is equal to

$$S^{-1} = \begin{bmatrix} \frac{1}{n_1}, & 0, \ldots, & 0 \\ 0, & \frac{1}{n_2}, \ldots, & 0 \\ \ldots & \ldots & \ldots \\ 0, & 0, \ldots, & \frac{1}{n_p} \end{bmatrix}.$$

Example 1. Under the above assumptions of one-classification and under the usual assumption that random variables y_α are all normally distributed with the common unknown variance σ^2, we shall verify the null hypothesis that the linear combination of p parameters is equal to zero i. e. $\sum_i^p \ell_i \beta_i = 0$, where the coefficients $\ell_1, \ell_2, \ldots, \ell_p$ are known. To deduce the form of random variable F which may be used to test the mentioned hypothesis it is sufficient to apply theorem 3. In this connection we note that the matrix $Ł = \underset{qp}{Ł}$ is in our case the row vector: $Ł = \underset{1p}{Ł} = [\ell_1, \ell_2, \ldots, \ell_p]$.

We find directly that

$$(ŁS^{-1}Ł^*)^{-1} = \frac{1}{\sum_i^p \frac{\ell_i^2}{n_i}}, \qquad \hat{\beta} = S^{-1}X^*y = \begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \bar{y}_p \end{bmatrix},$$

$$(Ł\hat{\beta})^*(ŁS^{-1}Ł^*)^{-1}(Ł\hat{\beta}) = \frac{\left(\sum_i^p \ell_i \bar{y}_i\right)^2}{\sum_i^p \frac{\ell_i^2}{n_i}}$$

and finally that

$$(y - X\hat{\beta})^*(y - X\hat{\beta}) = \sum_i^p \sum_j^{n_i} (y_{ij} - \bar{y}_i)^2, \quad \text{where} \quad \bar{y}_i = \frac{1}{n_i}\sum_j^{n_i} y_{ij}$$

is the mean of the observations in the i'th sample. Using theorem 3 we obtain the following expression for the random variable F which may be used to test the null hypothesis that $\sum_i^p \ell_i \beta_i = 0$:

$$F = \frac{(n-p)\left(\sum_i^p \ell_i \bar{y}_i\right)^2}{\left(\sum_i^p \frac{\ell_i^2}{n_i}\right) \cdot \sum_i^p \sum_j^{n_i} (y_{ij} - \bar{y}_i)^2} \tag{114}$$

with 1 and $n—p$ degrees of freedom.

In particular case, when the samples have identical numbers of observations i. e. when $n_i = k$ for $i = 1, 2, \ldots, c$ the random variable F assumes the form

$$F = \frac{k(n-p)\left(\sum_i^p l_i \bar{y}_i\right)^2}{\left(\sum_i^p l_i^2\right) \cdot \sum_i^p \sum_j^{n_i} (y_{ij} - \bar{y}_i)^2}. \tag{115}$$

The test based on random variable (115) may be applied, for example, in the analysis of factorial experiments performed in the glasshouse. In fact, if in the pot experiment with two factors, we investigate the effect of the two fertilizers "N" and "W" on the yield of a wheat variety and if each of the factors is introduced at two levels 0 and 1, then we get the four following treatment-combinations: $n^0 w^0$, $n^1 w^0$, $n^0 w^1$ and $n^1 w^1$. When each of these treatment combinations is replicated k times, the experiment contains $n = 4k$ pots. Let the means of the four populations corresponding to the four treatment combinations be denoted by β_1, β_2, β_3 and β_4. The experimenter is generally interested not only in the main effects but also in the interaction "WN" of the investigated factors. The correspoding hypothesis that interaction is equal to zero we express as the parameter linear function: $\beta_1 - \beta_2 - \beta_3 + \beta_4 = 0$. The null hypotheses that the main effect of the factor "W" and that of the factor "N" are not significant are expressed by the relations $-\beta_1 - \beta_2 + \beta_3 + \beta_4 = 0$ and $-\beta_1 + \beta_2 - \beta_3 + \beta_4 = 0$ respectively. Each of these hypotheses may be separately tested by means of test based on random variable (115).

It is evident that these remarks also apply in the case of testing the null hypothesis that the interactions of any order are equal to zero, if in the factorial experiments the treatment combinations constitute the one-way classification. In the experiments of this type one may verify by means of the test based on random variable (115) the significance of the regression components and in particular case the significance of the linear, quadratic and cubic components of the main effects and the combinations between these components of the two- and more-factor interactions.

For example, consider the 4×4 factorial experiment, where each of the two factors "W" and "N" occurs at 4 levels denoted successively by 0, 1, 2 and 3; let $\beta_1, \beta_2, \ldots, \beta_{16}$ be the means of the populations corresponding to the combinations: $n^0 w^0$, $n^0 w^1$, $n^0 w^2$, ..., $n^3 w^3$. Since for four levels of a factor the orthogonal coefficients determining the linear and quadratic effects are equal to: —3, —1, 1, 3 and 1, —1, —1, 1 respectively the null hypothesis ,that the regression linear $\times$ quadratic component $N_l W_q$ of the interaction "NW" is equal to zero, assumes the following form of the linear combination of the 16 parameters:

$$-3\beta_1+3\beta_2+3\beta_3-3\beta_4-\beta_5+\beta_6+\beta_7-\beta_8+\beta_9-\beta_{10}-\beta_{11}+\beta_{12}+$$
$$+3\beta_{13}-3\beta_{14}-3\beta_{15}+3\beta_{16}=0.$$

The significance of this function is tested by the test based on random variable (115).

Now let us present one more example of the application of theorem 3. Consider the third type of the multiple regression model, which contains c parameter groups i. e. vectors:

$$[\gamma_1]=\begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix},\quad [\gamma_2]=\begin{bmatrix}\beta_3\\ \beta_4\end{bmatrix},\ \ldots,\ [\gamma_c]=\begin{bmatrix}\beta_{p-1}\\ \beta_p\end{bmatrix},$$

forming jointly the vector $\beta^*=\left[[\gamma_1]^* \vdots [\gamma_2]^* \vdots \cdots \vdots [\gamma_c]^*\right]$. This model takes the form:

$$(116)\quad E(y_\alpha)=(x_{1\alpha}\beta_1+x_{2\alpha}\beta_2)+(x_{3\alpha}\beta_3+x_{4\alpha}\beta_4)+\cdots+(x_{p-1,\alpha}\beta_{p-1}+x_{p\alpha}\beta_p)$$

where $p=2c$, $\alpha=1,2,\ldots,n$, $n=\sum_i^c n_i$; $x_{1\alpha}=1$ for $\alpha=1,2,\ldots,n_1$ and $x_{1\alpha}=0$ for the remaining α; $x_{3\alpha}=1$ for $\alpha=n_1+1,\ n_1+2,\ldots,n_1+n_2$ and $x_{3\alpha}=0$ for the remaining α, etc.; suppose also that each of the $x_{2\alpha}$, $x_{4\alpha}, \ldots, x_{2c,\alpha}$ takes at least two different values. As usual we suppose that y_α's are independently and normally distributed with unknown common variance σ^2. This model may be interpreted in the following way. In each of c partial populations P_i, $i=1,2,\ldots,c$, we observe the linear regression y on x. The sample drawn from the population P_i contains n_i observations. Under these conditions the parameters with the even indices denote the unknown regression coefficients y on x, while the remaining parameters β with odd indices determine the points of the intersection of the successive linear regressions with the coordinate axis OY.

Note that $r(X)=p$. In fact, to prove this it is sufficient to take from each of the c pairs of columns two such rows of matrix $X=\underset{np}{X}$ that the elements of the even columns should be different. Under the above assumptions this is always possible to do. In this way we obtain the $p=2c$ independent rows, which means that the rank of the matrix X is equal to p i. e. to the number of the parameters of the model.

Having the matrix X we easily find the form of the matrix $S = X^*X$:

$$S = \underset{pp}{S} = \begin{bmatrix} n_1, & \sum^{n_1} x_2, & 0, & 0, \ldots, & 0, & 0 \\ \sum^{n_1} x_2, & \sum^{n_1} x_2^2, & 0, & 0, \ldots, & 0, & 0 \\ 0, & 0, & n_2 & \sum^{n_2} x_4, \ldots, & 0, & 0 \\ 0, & 0, & \sum^{n_2} x_4, & \sum^{n_2} x_4^2, \ldots, & 0, & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0, & 0, & 0, & 0, \ldots, & n_c, & \sum^{n_c} x_{2c} \\ 0, & 0, & 0, & 0, \ldots, & \sum^{n_c} x_{2c}, & \sum^{n_c} x_{2c}^2 \end{bmatrix}.$$

Since the value of the determinant of this matrix is equal to $|S| = n_1 n_2 \ldots n_c (n_1 S_2^2)(n_2 S_4^2) \ldots (n_c S_{2c}^2)$, where according to the notation in the sec. 2, the expression $n_i S_{2i}^2$ denotes the sum of squared deviations of observations x belonging to the i'th sample ($i = 1, 2, \ldots, c$), the matrix S^{-1} assumes the form

$$S^{-1} = \begin{bmatrix} \frac{\bar{x}_2^2}{n_1 S_2^2}, & \frac{-\bar{x}_2}{n_1 S_2^2}, & 0, & 0, \ldots, & 0, & 0 \\ \frac{-\bar{x}_2}{n_1 S_2^2}, & \frac{1}{n_1 S_2^2}, & 0, & 0, \ldots, & 0, & 0 \\ 0, & 0, & \frac{\bar{x}_4^2}{n_2 S_4^2}, & \frac{-\bar{x}_4}{n_2 S_4^2}, \ldots, & 0, & 0 \\ 0, & 0, & \frac{-\bar{x}_4}{n_2 S_4^2}, & \frac{1}{n_2 S_4^2}, \ldots, & 0, & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0, & 0, & 0, & 0, \ldots, & \frac{\bar{x}_{2c}^2}{n_c S_{2c}^2}, & \frac{-\bar{x}_{2c}}{n_c S_{2c}^2} \\ 0, & 0, & 0, & 0, \ldots, & \frac{-\bar{x}_{2c}}{n_c S_{2c}^2}, & \frac{1}{n_c S_{2c}^2} \end{bmatrix},$$

The symbols $\overline{x}_{2i}$ and $\overline{x^2_{2i}}$; $i = 1, 2, \dots, c$; denote the mean and the mean square of x's appearing in the i'th sample (cf. sec. 2). Using the known matrix relation $\hat{\beta} = S^{-1} X^* y$, we find the following estimates of the parameters:

$$\hat{\beta}_i = \overline{y}_{i+1} - \overline{x}_{i+1} \frac{S_{i+1, i+1}}{S^2_{i+1}} \quad \text{for } i = 1, 3, \dots, 2c-1$$

and $\hat{\beta}_j = S_{jj}/S_j^2$ for $j = 2, 4, \dots, 2c$, where $\overline{y}_{i+1}$ denotes the mean of y's in the $(i+1)/2$ — th sample, and $S_{i+1, i+1}$ — the covariance in this sample. The symbol S_{jj} stands for the covariance in the $j/2$ — th sample, while S^2_{i+1} and S_j^2 represent corresponding variances for x's. Developing the expression $(y - X\hat{\beta})^* (y - X\hat{\beta}) = y^* y - y^* X S^{-1} X^* y$, we easily find that it constitutes the sum of squares of deviations of y's from the regression "within the samples" i. e.

$$(117) \qquad (y - X\hat{\beta})^* (y - X\hat{\beta}) = \sum_i^c \left[n_i S^2_{y_{2i}} - \frac{(n_i S_{2i, 2i})^2}{n_i S^2_{2i}} \right]$$

where $S^2_{y_{2i}}$ is the variance of y's in the i'th sample, $i = 1, 2, \dots, c$.

For the model (116) we may verify by means of the test based on random variable in theorem 3, the null hypothesis that the c linear regressions intersect the coordinate axis OY at the same point; we may also test the null hypothesis that the regression coefficients y on x are all equal in the c populations. The latter problem raised by K. Iwaszkiewicz ([5]) has been solved by St. Kołodziejczyk ([7]). This paper offers the following solution of this problem:

We derive the corresponding random variable from the form (56) i. e. from the formula:

$$F = \frac{(Ł\hat{\beta} - \varphi_0)^* (Ł S^{-1} Ł^*)^{-1} (Ł\hat{\beta} - \varphi_0)}{q} : \frac{(y - X\hat{\beta})^* (y - X\hat{\beta})}{n - p}$$

with q and $n - p$ degrees of freedom, where the null hypothesis is: $\underset{q1}{\varphi} = Ł\beta = \varphi_0$. In the actual problem this hypothesis assumes the form:

$$\beta_2 - \beta_p = \varphi_1^0 = 0, \quad \beta_4 - \beta_p = \varphi_2^0 = 0, \dots, \beta_{p-2} - \beta_p = \varphi_{c-1}^0 = 0.$$

Hence $q = r(Ł) = c - 1 = (p/2) - 1$ and

$$Ł = \underset{c-1, p}{Ł} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0, \dots, & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0, \dots, & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1, \dots, & 0 & 0 & -1 \\ . & . & . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & 0 & 0 & 0 & 0, \dots, & 1 & 0 & -1 \end{bmatrix}.$$

To find the expression for the random variable F, it is sufficient to calculate its denominator. We note that

$$\underset{c-1,c-1}{K} = ŁS^{-1}Ł^{*} = \begin{bmatrix} \frac{1}{n_c S_{2c}^2} + \frac{1}{n_1 S_2^2}, & \frac{1}{n_c S_{2c}^2}, \ldots, & \frac{1}{n_c S_{2c}^2} \\ \frac{1}{n_c S_{2c}^2}, & \frac{1}{n_c S_{2c}^2} + \frac{1}{n_2 S_4^2}, \ldots, & \frac{1}{n_c S_{2c}^2} \\ \ldots & \ldots & \ldots \\ \frac{1}{n_c S_{2c}^2}, & \frac{1}{n_c S_{2c}^2}, \ldots, & \frac{1}{n_c S_{2c}^2} + \frac{1}{n_{c-1} S_{2c-2}^2} \end{bmatrix}.$$

Since the determinant of the matrix K has the value:

$$|K| = |ŁS^{-1}Ł^{*}| = \frac{\sum_i^c n_i S_{2i}^2}{\prod_i^c n_i S_{2i}^2},$$

the reciprocal of the matrix K is the matrix

$$K^{-1} = \frac{1}{W} \begin{bmatrix} (W - m_1)\, m_1, & -m_1 m_2, \ldots, & -m_1 m_{c-1} \\ -m_2 m_1, & (W - m_2)\, m_2, \ldots, & -m_2 m_{c-1} \\ \ldots & \ldots & \ldots \\ -m_{c-1} m_1, & -m_{c-1} m_2, \ldots, & (W - m_{c-1})\, m_{c-1} \end{bmatrix}$$

where for brevity we put $m_j = n_j S_{2j}^2$; $j = 1, 2, \ldots, c-1$; and $W = \sum_i^c n_i S_{2i}^2$.

Knowing the form of the matrix $Ł$, we have directly: $(Ł\hat{\beta})^{*} = [\hat{\beta}_2 - \hat{\beta}_p, \hat{\beta}_4 - \hat{\beta}_p, \ldots, \hat{\beta}_{p-2} - \hat{\beta}_p]$, and taking into account the form of the matrix K^{-1}, we obtain the following expression for the numerator of the random variable F

$$(Ł\hat{\beta})^{*} (ŁS^{-1}Ł^{*})^{-1} (Ł\hat{\beta}) = \left\{ \sum_{i=1}^{c-1} \sum_{j=i+1}^{c} (n_i S_{2i}^2)(n_j S_{2j}^2)(\hat{\beta}_{2i} - \hat{\beta}_{2j})^2 \right\} : \sum_{i=1}^{c} n_i S_{2i}^2 .$$

As a result, the random variable which may be used to test the null hypothesis that the regression coefficients y on x in the $c = p/2$ samples are identical has the form:

$$F = \frac{(n-2c)\sum_{i=1}^{c-1}\sum_{j=i+1}^{c}(n_i S_{2i}^2)(n_j S_{2j}^2)(\hat{\beta}_{2i}-\hat{\beta}_{2j})^2}{(c-1)\left(\sum_{i=1}^{c} n_i S_{2i}^2\right)\cdot\sum_{i=1}^{c}\left[n_i S_{y_{2i}}^2 - \frac{(n_i S_{2i,2i})^2}{n_i S_{2i}^2}\right]}$$

with c—*1* and n—2c degrees of freedom.

Next we will deal with the problems, whose solutions are reached by the application of the theorem 4. We shall illustrate this by the following example.

In the model of one-way classification with different numbers of observations in the subclasses (defined in the present section under *b*) we assume that the known linear combination of the parameters $\beta_1, \beta_2, \ldots, \beta_p$ has the given value η i. e.

$$\sum_{i}^{p} l_i \beta_i = \eta. \tag{118}$$

Let us find the form of random variable under the null hypothesis

$$\sum_{i}^{p} t_i \beta_i = 0. \tag{119}$$

The comparison of the present model with the assumptions of the theorem 4 shows that the matrices $G = \underset{gp}{G}$ and $H = \underset{hp}{H}$ are vectors i. e. $G = \underset{1p}{G} = [l_1, l_2, \ldots, l_p]$ and $H = \underset{1p}{H} = [t_1, t_2, \ldots, t_p]$, and that

$$T = \underset{g+h,\,p}{T} = \left[\frac{G}{H}\right] = \left[\frac{l_1, l_2, \ldots, l_p}{t_1, t_2, \ldots, t_p}\right] \text{ and } \tau = \left[\frac{\eta}{\nu}\right] = \left[\frac{\eta}{0}\right].$$

As usual, we assume that the functions (118) and (119) are independent. In order to apply the random variable given in (57) we calculate easily:

$$TS^{-1}T^* = \begin{bmatrix} \sum_{i}^{p}\frac{l_i^2}{n_i}, & \sum_{i}^{p}\frac{l_i t_i}{n_i} \\ \sum_{i}^{p}\frac{l_i t_i}{n_i}, & \sum_{i}^{p}\frac{t_i^2}{n_i} \end{bmatrix} \text{ and } (TS^{-1}T^*)^{-1} = \frac{1}{W}\begin{bmatrix} \sum_{i}^{p}\frac{t_i^2}{n_i}, & -\sum_{i}^{p}\frac{l_i t_i}{n_i} \\ -\sum_{i}^{p}\frac{l_i t_i}{n_i}, & \sum_{i}^{p}\frac{l_i^2}{n_i} \end{bmatrix}$$

where

$$W = |TS^{-1}T^*| = \left(\sum_{i}^{p}\frac{l_i^2}{n_i}\right)\left(\sum_{i}^{p}\frac{t_i^2}{n_i}\right) - \left(\sum_{i}^{p}\frac{l_i t_i}{n_i}\right)^2.$$

Considering further

$$T\hat{\beta}-\tau=\begin{bmatrix}\sum_i^p l_i\bar{y}_i-\eta\\ \sum_i^p t_i\bar{y}_i\end{bmatrix}\text{ where }\bar{y}_i=\hat{\beta}_i=\frac{1}{n_i}\sum_j^{n_i}y_{ij};\ i=1,2,\ldots,p;$$

we obtain:

$$(T\hat{\beta}-\tau)^*(TS^{-1}T^*)^{-1}(T\hat{\beta}-\tau)=$$

$$=\frac{1}{W}\left\{\left(\sum_i^p l_i\bar{y}_i-\eta\right)^2\sum_i^p\frac{t_i^2}{n_i}-2\left(\sum_i^p l_i\bar{y}_i-\eta\right)\left(\sum_i^p t_i\bar{y}_i\right)\left(\sum_i^p\frac{l_it_i}{n_i}\right)+\right.$$

$$\left.+\left(\sum_i^p t_i\bar{y}_i\right)^2\left(\sum_i^p\frac{l_i^2}{n_i}\right)\right\}$$

and finally

$$(G\hat{\beta}-\eta)^*(GS^{-1}G^*)^{-1}(G\hat{\beta}-\eta)=\frac{\left(\sum_i^p l_i\bar{y}_i-\eta\right)^2}{\sum_i^p\frac{l_i^2}{n_i}}.$$

Introducing the obtained expressions in the formula (57) we obtain the following expression for the random variable F:

$$F=\frac{(n-p+1)\,|a^2b^2-2(c-\eta)abd+(c-\eta)^2d^2|}{\left(b\sum_{i=1}^p\frac{t_i^2}{n_i}-d^2\right)\left[b\sum_{i=1}^p\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_i)^2+(c-\eta)^2\right]}\tag{120}$$

with 1 and $n-p+1$ degrees of freedom, where

$$a=\sum_{i=1}^p t_i\bar{y}_i,\quad b=\sum_{i=1}^p\frac{l_i^2}{n_i},\quad c=\sum_{i=1}^p l_i\bar{y}_i,\quad d=\sum_{i=1}^p\frac{l_it_i}{n_i}\text{ and }\eta=\sum_{i=1}^p l_i\beta_i$$

(t_i, l_i and η are given; $i=1,2,\ldots,p$).

The random variable (121) may be used by the experimenter, who takes an interest in the problem connected with the factorial experiments performed in a glasshouse. Suppose that in the pot experiment one investigates the effect of the factor "N" (for example the fertilizer) occuring at four levels 0, 1, 2 and 3, on the yield of some corn variety. Let β_1,β_2,β_3 and β_4 be the means of the populations corresponding to the given levels. Let the difference, η, between the yields obtained at the first and second level be known from the former experiments and be expressed by the relation $-\beta_1+\beta_2=\eta$. Then one is frequently interested in the testing of the null hypothesis that the linear regression component N_l is equal to zero i.e. $-3\beta_1-\beta_2+\beta_3+3\beta_4=0$. This hypothesis is verified by test based on random variable (121). In the case of rejection of the hypothesis we have the basis to conclude that the regression of the yields on the levels of investigated factor is significant.

Now we present a second example of applying test based on random variable (121). Observe the experiment described on p. 56. Using the random variable (121), one may verify the null hypothesis that the main efect of one of the factors is equal to zero, when it is known that the interaction between the factors also equals zero; this is equivalent to the testing of the hypothesis that $\beta_1-\beta_2+\beta_3-\beta_4=0$ when $\beta_1-\beta_2-\beta_3+\beta_4=0$.

The significance of the main effects in the field factorial experiments performed according to the model of randomized blocks may be also verified on the basis of random variable (57) (cf. theor. 4). According to the reparametrization of the model described on p. 52—53 we transform it into the model of the form (111). Let for illustration $c=4$ denote the number of combinations n^0w^0, n^0w^1, n^1w^0 and n^1w^1 obtained by combining two levels 0 and 1 of each of the two factors "N" and "W". Then the treatment effects a_1, a_2, a_3 and a_4 are identical with the effects of the respective combinations. Let the corresponding means of four populations represented by the four combinations be denoted by v_0, v_1, v_2 and v_3. Then, it is easy to see that

$$v_i=a_{i+1}+\mu', \qquad i=0,1,2,3. \tag{122}$$

The null hypothesis that the mean effect, "N", is equal to zero, under the assumption that the interaction WN between the factor "W" and "N" equals zero, is expressed by the relation $-v_0-v_1+v_2+v_3=0$, while the assumption about the interaction WN is expressed by the relation $v_0-v_1-v_2+v_3=0$. Using (122) these may be written as $-a_1-a_2+a_3+a_4=0$ and $a_1-a_2-a_3+a_4=0$ respectively. Since the restriction $\sum_i^4 a_i=0$ is imposed on the parameters a_i $(i=1,2,3,4)$, we obtain, instead of the latter three relations, only two relations: $a_1+a_2=0$ and $a_2+a_3=0$. The first of these constitutes the null hypothesis, and the second — the restriction for the parameters of the model (111), where we put $c=4$. Thus we obtain for the actual problem the model with the restriction and with the null hypothesis ,that satisfies the assumption of the theorem 4. It then follows that the test based on random variable (57) may be applied here.

We shall mention briefly one more problem, which requires the application of the theorem 4. Let the $c=p/2$ regression straight-lines y on x pass through such points on the Y axis, whose distances from each other are equal to the quantities $\eta_1^0, \eta_2^0, \ldots, \eta_{c-1}^0$. The remaining assumptions are presented on p. 57. Now we are interested in the test which under these conditions verifies the null hypothesis that the differences among the regression coefficients y on x in the c populations are equal to the given quantities $\nu_1^0, \nu_2^0, \ldots, \nu_{c-1}^0$. It may be easily observed that the ranks of both matrices $H = \underset{hp}{H}$ and $G = \underset{gp}{G}$ appearing in the theorem 4 are identical and equal to c—1. The application of this problem in the experimentation may be illustrated by means of the following example. Knowing differences between the yields of the c wheat varietes obtained at the zero level of investigated factor (for example: a fertilizer), one will be interested in verification the null hypothesis that the differences in the yield increment in successive levels have a priori given values, and in particular that these differences are identical.

We proceed to present one of the problems, which may be solved by means of theorem 5. Consider the multiple regression model determined by the relation (116). For this model (which represents the regression lines y on x in the c populations, and which contains the $p=2c$ parameters), we shall find a test of significance of the null hypothesis that all lines are identical. This hypothesis is expressed in the form: $\beta_i = \vartheta_1$ $(i=1, 3, 5, \ldots, 2c-1)$ and $\beta_j = \vartheta_2$ $(j=2, 4, \ldots, 2c)$ which in the matrix notation is written as

$$\beta = U\vartheta \tag{123}$$

where, according to the symbols given in the theorem 5, the rank of the matrix $U = \underset{p,\, p-q}{U}$ is in the actual case equal to $p-q=2$, since

$$U^* = \begin{bmatrix} 1 & 0 & 1 & 0 & \ldots & 1 & 0 \\ 0 & 1 & 0 & 1 & \ldots & 0 & 1 \end{bmatrix} \quad \text{while} \quad \vartheta = \underset{q-p,\, 1}{\vartheta} = \begin{bmatrix} \vartheta_1 \\ \vartheta_2 \end{bmatrix}.$$

Since the assumptions specified in the theorem 5 are here satisfied, and since the null hypothesis determines each of the parameters $\beta_1, \beta_2, \ldots, \beta_p$ as the linear combination of two parameters ϑ_1 and ϑ_2, the random variable (58) may be applied. Let us find the matrix expression for its numerator. Using the form of matrix S presented on p. 58, we obtain:

$$(U^* SU)^{-1} = \frac{1}{n^2 S^2} \begin{vmatrix} \Sigma x^2, & -\Sigma x \\ -\Sigma x\,, & n \end{vmatrix},$$

where summation in extended over all n observations and

$$S^2 = \frac{1}{n} \sum (x - \bar{x})^2$$

is the sample variance of x. Further we find

$$U(U^* SU)^{-1} U^* = Z = \underset{pp}{Z} = \frac{1}{n^2 S^2} \begin{bmatrix} MM \ldots M \\ MM \ldots M \\ \cdots\cdots\cdots \\ MM \ldots M \end{bmatrix}, \text{ where } M = \begin{bmatrix} \sum x^2, & -\sum x \\ -\sum x, & n \end{bmatrix}$$

and finally

$$y^* XU(U^* SU)^{-1} U^* X^* y = \frac{(nS_{xy})^2}{nS^2} + \frac{(\sum^n y)^2}{n} \tag{124}$$

where $nS_{xy} = \sum^n (x - \bar{x})(y - \bar{y})$. Utilizing the matrix S^{-1} given explicitely on p. 58, we have

$$y^* XS^{-1} X^* y = \sum_i^c \left[\frac{(n_i S_{2i,\,2i})^2}{n_i S_{2i}^2} + \frac{(\sum^{n_i} y_{2i})^2}{n_i} \right] \tag{125}$$

where y_{2i}, $n_i S_{2i}^2$ and $n_i S_{2i,\,2i}$ $(i = 1, 2, \ldots, c)$ denote successively the observations belonging to the i-th sample, the sum of squared deviations of x's in the i-th sample, and the sum of the product of mixed deviations in this sample. Considering the relation $(S\hat{\beta})^* [S^{-1} - U(U^* SU)^{-1} U^*] S\hat{\beta} = y^* XS^{-1} X^* y - y^* XU(U^* SU)^{-1} U^* X^* y$ and the expressions (124), (125) and (117), we may state that the random variable (58) in the actual model assumes the form

$$F = \frac{\sum_i^c \left[\frac{(n_i S_{2i,\,2i})^2}{n_i S_{2i}^2} + \frac{(\sum^{n_i} y_{2i})^2}{n_i} \right] - \frac{(nS_{xy})^2}{nS^2} - \frac{(\sum^n y)^2}{n}}{2c - 2} : \frac{\sum_i^c \left[n_i S_{y_{2i}}^2 - \frac{(n_i S_{2i,\,2i})^2}{n_i S_{2i}^2} \right]}{n - 2c} \tag{126}$$

with $2c - 2$ and $n - 2c$ degrees of freedom, where $n_i S_{y_{2i}}^2$ stands for the sum of squared deviations of y's in the i-th sample; $i = 1, 2, \ldots, c$. Thus the test based on random variable (126) verifies the null hypothesis that the c regression lines are identical.

REFERENCES

[1] F i s h e r R. A. — *The goodness of fit of regression formulae, and the distribution of regression coefficients.* — Jour. Roy. Stat. Soc., 85, Part IV, 597—612, 1922, and reprinted in Contributions to mathematical statistics. John Wiley and Sons., New York. 1950.

[2] F i s h e r R. A. — *Application of "Student's" distribution* — Metron V, No 3, 30—104, 1925.

[3] F i s h e r R. A. — *Statistical methods for research workers.* — Oliver and Boyd, Edinburgh. 1st ed., 1925; 11th ed., 1950.

[4] G a n t m a c h e r F. R. — *Tieoria matric* — GITTL, Moskwa, 1953.

[5] I w a s z k i e w i c z K. — *Sur la généralisation de la méthode de corrélation partielle pour le cas ou la variable éliminée n'est pas mesurable* (resumé in French). Kwartalnik Statystyczny, 9, Z. 3, Warszawa, 1932.

[6] K e m p t h o r n e O. — *The design and analysis of experiments* — John Wiley and Sons, New York, 1952.

[7] K o ł o d z i e j c z y k St. — *On an important class of statistical hypotheses* — Biometrika, 27, 161—190, 1935.

[8] M a n n H. B. — *Analysis and design of experiments.* — Dover Publications, Inc. New York, 1949.

[9] M a r k o f f A. A. — *Wahrscheinlichkeitsrechnung* — B. G. Teubner, Berlin und Leipzig, 1912.

[10] M o o d A. M. — *Introduction to the theory of statistics* — Mc Graw-Hill, New York, 1950.

[11] N e y m a n J. and P e a r s o n E. S. — *On the problem of two samples* — Bulletin de l'Académie Polonaise des Sciences et des Lettres, 1930 A, 73—96.

[12] N e y m a n J. and P e a r s o n E. S. — *On the problem of k samples* — Ibid. 1931 A, 460—482.

[13] N e y m a n J. — *O pewnych twierdzeniach z rachunku prawdopodobieństwa, które służą za podstawę do rozwiązania szeregu zagadnień doświadczalnictwa rolniczego* (résumé in English). Roczniki Nauk Rolniczych i Leśnych, 31, 224—276, Poznań, 1934.

[14] N e y m a n J. and P e a r s o n E. S. — *On the use and interpretation of certain test criteria for purposes of statistical inference* — Biometrika, 20 A, 175—240 and 264—294, 1928.

[15] N e y m a n J. and P e a r s o n E. S. — *On the problem of the most efficient tests of statistical hypotheses* — Phil. Trans. of the Royal Society, London, 231 A, 289—337, 1933.

[16] N e y m a n J. and P e a r s o n E. S. — *The testing of statistical hypoteses in relation to probabilities a priori* — Proceedings of the Cambridge Philosophical Society, 29, 492—510, 1933.

[17] N e y m a n J. — *O metodach opracowywania doświadczeń wielokrotnych* (résumé in English). Roczniki Nauk Rolniczych i Leśnych, 28, 154—210, Poznań, 1932.

[18] O k u n i e w L. J. — *Wysszaja algebra* — GITTL, Moskwa, Leningrad, 1949.

[19] R a o C. R. — *Advanced statistical methods in biometric research* — John Wiley and Sons, New York, 1952.

[20] Rao C. R. — *On the linear combination of observations and the general theory of least squares* — Sankhya, 7, 237, 1946.
[21] Rao C. R. — *A theorem in least squares* — Sankhya, 11, 9—12, 1951.
[22] „Student" — *The probable error of a mean* — Biometrika, 6, 1—25, 1908.
[23] Wald A. — *Test of statistical hypotheses concerning several parameters, when the number of observations is large* — Trans. Am. Math., Sec. 54, 426—482.
[24] Welch B. L. — *Some problems in the analysis of regression among k samples of two variables* — Biometrika, 27, 145—160, 1935.
[25] Wilks S. S. — *Mathematical Statistics* — University Press, Princeton, 1947.

Streszczenie

Zagadnienie sprawdzania hipotezy liniowej w teorii normalnej regresji, którym zajmujemy się w niniejszej pracy i którym interesowało się od niemal pół wieku wielu autorów, można zreferować w sposób następujący:

Niezależne zmienne losowe

$$y_k = m_k + e_k \tag{1}$$

o wartościach oczekiwanych

$$m_k = E(y_k) = x_{k1}\beta_1 + x_{k2}\beta_2 + \cdots + x_{kp}\beta_p \tag{2}$$

zależnych od p ($p < n$) parametrów $\beta_1, \beta_2, \ldots, \beta_p$; $k = 1, 2, \ldots, n$; mają rozkład normalny ze wspólną wariancją σ^2. Znak E określa wartość oczekiwaną, symbol $X = \{x_{kj}\}$, $k = 1, 2, \ldots, n$; $j = 1, 2, \ldots, p$; oznacza daną macierz a e_k — rezidua (por. [6]). Na parametry $\beta_1, \beta_2, \ldots, \beta_p$ modelu (2) nałożono g znanych liniowo niezależnych restrykcyj opisanych relacją macierzową

$$G\beta = \eta \tag{3}$$

gdzie β jest wektorem kolumnowym o p składowych $\beta_1, \beta_2, \ldots, \beta_p$ i gdzie macierz $G = \{g_{ql}\}$; $q = 1, 2, \ldots, g$; $l = 1, 2, \ldots, p$; jest znana podobnie jak wektor kolumnowy η o g składowych $\eta_1, \eta_2, \ldots, \eta_g$.

Ogólny problem polega na sprawdzeniu słuszności hipotezy liniowej

$$H\beta = \nu \tag{4}$$

że h znanych i liniowo niezależnych funkcyj parametrycznych określonych za pomocą relacji macierzowej (4) jako iloczyn macierzy $H = \{h_{ij}\}$; $i = 1, 2, \ldots, h$; $j = 1, 2, \ldots, p$; i wektora β mają określone wartości $\nu_1, \nu_2, \ldots, \nu_h$ będące składowymi wektora ν. Zakłada się przy tym, że funkcje podane w (3) i (4) są liniowo niezależne.

Rozwiązanie tego problemu można bezpośrednio uzyskać opierając się na twierdzeniu z teorii rozkładów, którego autorem jest C. R. Rao (por. [21]). Ze względu na podstawowe znaczenie tego twierdzenia w teorii normalnej regresji i w zastosowaniach podaję inny jego dowód w przypadku gdy rząd macierzy X wynosi p (por. twierdzenie 1 w niniejszej pracy oraz tw. 4.1 Manna [8]).

Poza wymienionym ogólnym zagadnieniem zajmuję się wyprowadzeniem wyraźnych form zmiennych losowych F w postaci macierzowej dla najczęściej spotykanych w zastosowaniach modeli regresji liniowej. Zmienne te (por. wzory (47), (48), (56), (57) i (58) w twierdzeniach 2, 3, 4 i 5) można traktować jako podstawy testów istotności F sprawdzających różne hipotezy liniowe w zależności od typu modelu regresji wielokrotnej.

Test F wynikający ze wzoru (56) może służyć do sprawdzenia hipotezy (4), gdy na parametry β modelu (2) nie nakładamy restrykcji (3), a test oparty na zmiennej losowej (57) może służyć do sprawdzenia tejże hipotezy, gdy restrykcje (3) są nałożone. Wreszcie test oparty na zmiennej losowej (58) (por. tw. 5) może służyć do sprawdzenia hipotezy zerowej, że parametry modelu wyrażają się jako określone z góry kombinacje liniowe mniejszej liczby innych parametrów.

Uzyskane formy testów ułatwiają w znacznej mierze wyznaczenie ich postaci dla konkretnych modeli eksperymentalnych i danych hipotez liniowych bez przeprowadzania minimalizacji względem parametrów, co byłoby nieodzowne w każdym przypadku, gdyby bezpośrednio stosowano twierdzenie 1. Należy zauważyć, że tę minimalizację trzebaby również zawsze stosować przy korzystaniu z twierdzenia C. R. Rao (loc. cit.). Tymczasem dla uzyskania testu istotności na podstawie twierdzeń 2, 3, 4 i 5 wystarczy jedynie wykonać kilka prostych operacyj na macierzach. Odpowiednie przykłady na zastosowanie twierdzeń 2, 3, 4 i 5, jak również przykłady zagadnień występujących w doświadczalnictwie polowym i szklarniowym, których rozwiązanie bądź naświetlenie wymaga zastosowania tego rodzaju testów przedstawiam w paragrafie 6.

Dowód twierdzenia 2 jest nowy, jakkolwiek inny dowód nie w formie macierzowej, znajdujemy u Manna (por. tw. 4.3 w [8]). Wszystkie dowody przedstawiam w rachunku macierzowym.

Korzystając z ogólnej postaci zmiennej losowej F wymienionej w twierdzeniu 4 znajduję wyraźną jej formę (wzór (120)) służącą do sprawdzania w modelu klasyfikacji pojedynczej (one-way), z różnymi liczebnościami obserwacyj w podklasach, hipotezy, że kombinacja liniowa parametrów jest zerem, gdy wiadomo, że pewna inna kombinacja liniowa tych parametrów ma określoną wartość.

Nadto podaję dowody (oparte na rachunku macierzowym) na to, że wszystkie testy F rozpatrywane w niniejszej pracy są prawostronne. W związku z tym wyznaczam szereg wartości oczekiwanych odpowiednich form kwadratowych. Poza tym wyprowadzam: tożsamości dla warunkowych i bezwarunkowych ocen parametrów (por. (24) i (25)), macierz kowariancji między kombinacjami liniowymi ocen parametrów oraz szereg innych macierzy kowariancji przedstawionych w paragrafach 3 i 5.

Резюме

Проблема проверки линейной гипотезы в теории нормальной регрессии, которой мы занимаемся в этой работе и которой интересовались почти от полусотни лет многие авторы, может быть изложена сокращённо следующим образом.

Независимые случайные величины

(1) $$y_k = m_k + e_k$$

с математическими ожиданиями

(2) $$m_k = E(y_k) = x_{k1}\beta_1 + x_{k2}\beta_2 + \cdots + x_{kp}\beta_p$$

зависящими от p ($p<n$) параметров следуют нормальному закону распределения с общей дисперсией σ^2. Символ E обозначает математическое ожидание, символ $X = \{x_{kj}\}$ ($k = 1, 2, \ldots, n$; $j = 1, 2, \ldots, p$) обозначает данную матрицу, а e_k — резидуумы (ср. [6]). На параметры $\beta_1, \beta_2, \ldots, \beta_p$ модели (2) наложено g известных линейно независимых ограничений, описанных матричным соотношением

(3) $$G\beta = \eta$$

где β есть столбцевой вектор с p компонентами $\beta_1, \beta_2, \ldots, \beta_p$, а матрица $G = \{g_{ql}\}$ ($q = 1, 2, \ldots, g$; $l = 1, 2, \ldots, p$) известна, равно как и столбцевой вектор η с g компонентами $\eta_1, \eta_2, \ldots, \eta_g$.

Общая проблема состоит в проверке правильности линейной гипотезы

(4) $$H\beta = \nu$$

что h известных и линейно независимых параметрических функций, определённых с помощью матричного соотношения (4), как произведение матрицы $H = \{h_{ij}\}$ ($i = 1, 2, \ldots, h$); $j = 1, 2, \ldots, p$) и вектора β, имеют определённые значения $\nu_1, \nu_2, \ldots, \nu_h$, являющиеся компонентами вектора ν. При этом полагаем, что данные в (3) и (4) функции линейно независимы.

Решение этой проблемы можно получить непосредственно, основываясь на теореме из теории распределений, которой автором был Ц. Р. Рао (ср. [21]).

Из-за фундаментального значения этой теоремы в теории нормальной регрессии и в применениях я провожу иное её доказательство для случая, когда порядок матрицы X есть p (ср. теорему 1 в этой работе и теорему 4.1 Манна [8]).

Сверх приведённого общего вопроса я занимаюсь выведением явных форм случайных величин F в матричном виде для чаще всего встречаемых в применениях моделей линейной регрессии. Эти переменные (ср. формулы (47), (48), (56), (57) и (58) в теоремах 2, 3, 4 и 5) можно рассматривать, как основания критериев значимости F, проверяющих различные линейные гипотезы в зависимости от типа модели множественной регрессии.

Критерий F, вытекающий из формулы (56), может служить для проверки гипотезы (4), когда на параметры β модели (2) не наложены ограничения (3), а критерий, опирающийся на случайной величине (57), может служить для проверки этой же гипотезы, когда наложены ограничения (3). Наконец, основанный на случайной величине (58) критерий (ср. теор. 5) может служить для проверки нулевой гипотезы, что параметры модели выражаются, как наперёд определённые комбинации меньшего числа иных параметров.

Полученные формы критериев облегчают в значительной степени определение их вида для конкретных экспериментальных моделей и данных линейных гипотез без проведения минимализации относительно параметров, что было бы всегда неизбежно, если бы применять непосредственно теорему 1. Следует заметить, что эту минимализацию следовало бы тоже всегда применять при пользовании теоремой Ц. Р. Рао (там же). Между тем, для получения критерия значимости на основании теорем 2, 3, 4 и 5 достаточно лишь выполнить несколько простых действий с матрицами. Подходящие примеры применения теорем 2, 3, 4 и 5, как и примеры проблем, выступающих в полевом и парниковом опытах, которых решение или разъяснение требует применения этого рода критериев, представлены в § 6.

Доказательство теоремы 2 ново, хотя иное доказательство не в матричном виде находим у Манна (ср. теор. 4.3 в [8]). Все доказательства я представляю в матричном исчислении.

Пользуясь общим видом случайной величины F, упомянутым в теореме 4, я нахожу явную его форму (формула (120)), служащую для проверки, в модели единичной классификации (one-way) с различными численностями наблюдений в классах, гипотезы, что линейная комбинация параметров равна нулю, когда известно, что некоторая другая линейная комбинация этих параметров имеет определённое значение.

Сверх того я даю доказательства (матричным исчислением) того, что всякие критерии F, рассматриваемые в этой работе, суть правосторонние. В связи с этим я нахожу ряд значений ожидаемых соответствующих квадратичных форм. Сверх того я вывожу: тождество для условных и безусловных оценок параметров (ср. (24) и (25)), матрицу ковариации между линейными комбинациями оценок параметров, а также ряд иных матриц ковариации, представленных в §§ 3 и 5.