ANNALES
UNIVERSITATIS MARIAE CURIE-SKŁODOWSKA
LUBLIN — POLONIA
VOL. XVI, 7 SECTIO A 1962

Z Zakładu Statystyki Matematycznej Wydziału Rolniczego
Wyższej Szkoły Rolniczej w Lublinie
Kierownik: doc. dr Wiktor Oktaba

WIKTOR OKTABA

On the Testing of the Linear Hypothesis for the Model of Normal Regression

O weryfikowaniu liniowej hipotezy dla modelu normalnej regresji

О проверке линейной гипотезы для модели нормальной регрессии

1. Theorem

The problem of my previous work [1] is developed in the present paper, so we use the same matrix notation.

Theorem. *Let the components* $y_1, y_2, \ldots, y_n$ *of the column vector* $\underset{n1}{y}$ *of the multiple regression model* $\mu = E(y) = \underset{np}{X}\,\underset{p1}{\beta}$ *be independent random variables normally distributed with means* $\mu_1, \mu_2, \ldots, \mu_n$ *and common variance* σ^2. *Besides, let* $r(X)$ *i. e. the rank of matrix* X *be equal to the number* p *of parameters* β *and let the hypothesis that* $\underset{q1}{\varphi} = \underset{qm}{L}\,\underset{m1}{\gamma} = \underset{q1}{\varphi_0}$ *be true, where* $r(L) = q$ *and* $\gamma = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdots \\ \beta_m \end{bmatrix}$ *is the subvector of vector* $\underset{p1}{\beta} = \begin{bmatrix} \underset{m1}{\gamma} \\ \cdots \\ \underset{p-m,1}{\delta} \end{bmatrix}$. *Then the random variable*

$$(1)\qquad F = \frac{(L\hat{\gamma}-\varphi_0)^*(LS^{11}L^*)^{-1}(L\hat{\gamma}-\varphi_0)}{q} : \frac{(y-X\hat{\beta})^*(y-X\hat{\beta})}{n-p}$$

has F *distribution with* q *and* $n-p$ *degrees of freedom, where* $\underset{pp}{S} = X^*X$ *and* $\underset{mm}{S^{11}}$ *is submatrix of the matrix* $S^{-1} = \begin{bmatrix} S^{11} & S^{12} \\ S^{21} & S^{22} \end{bmatrix}$ *and* $\hat{\gamma}$ *is the estimate of the parameter* γ.

Proof. We use theorem 3 (cf. loc. cit.), which says that under the above mentionned assumptions, if the hypothesis that $\varphi = L\beta = \varphi_0$ is true (where $\underset{qp}{Ł} = Ł$ and $r(Ł) = q$), the random variable

$$(2)\qquad F = \frac{(Ł\hat{\beta}-\varphi_0)^*(ŁS^{-1}Ł^*)^{-1}(Ł\hat{\beta}-\varphi_0)}{q} : \frac{(y-X\hat{\beta})^*(y-X\hat{\beta})}{n-p}$$

has F distribution with q and $n-p$ degrees of freedom. As usual, $\hat{\beta} = S^{-1}X^*y$. In order to do this, let us consider $\underset{qp}{Ł} = [\underset{qm}{L} \vdots \underset{q,p-m}{0}]$. Then

$$\varphi = Ł\beta = [\underset{qm}{L} \vdots \underset{q,p-m}{0}]\begin{bmatrix}\underset{m1}{\gamma}\\ \underset{p-m,1}{\delta}\end{bmatrix} = L\gamma, \text{ so } Ł\hat{\beta} = L\hat{\gamma} \text{ and } ŁS^{-1}Ł^* = [L \vdots 0]\times$$

$$\times\begin{bmatrix}S^{11} & S^{12}\\ S^{21} & S^{22}\end{bmatrix}[L \vdots 0]^* = [LS^{11} \vdots LS^{12}]\begin{bmatrix}L^*\\ 0^*\end{bmatrix} = LS^{11}L^*.$$

The proof of the theorem is concluded if we substitute $Ł\hat{\beta} = L\hat{\gamma}$ and $ŁS^{-1}Ł^* = LS^{11}L^*$ into the expression (2).

2. Applications

1°. Let us put $q = 1$ and $\varphi_0 = 0$. Then $LS^{11}L^*$ and $L\hat{\gamma}$ are numbers so we obtain $(LS^{11}L^*)^{-1} = 1/LS^{11}L^*$ and $(L\hat{\gamma})^* = L\hat{\gamma}$. Since $\hat{\beta} = S^{-1}X^*y$ we obtain $\mathrm{Var}(\underset{1m}{L}\,\underset{m1}{\hat{\gamma}}) = \mathrm{Var}(\sum_{i=1}^{m} l_i\hat{\gamma}_i) = \sum_{i=1}^{m} l_i^2\mathrm{Var}(\hat{\gamma}_i) + \sum_{i\neq j} l_i l_j \mathrm{Cov}(\hat{\gamma}_i, \hat{\gamma}_j) = \sum_{i=1}^{m}\sum_{j=1}^{m} l_i l_j s^{ij}\sigma^2 = \sigma^2 LS^{11}L^*$. Thus the random variable (1) can be written as follows

$$F = \frac{\sigma^2(L\hat{\gamma})^2}{\mathrm{Var}(L\hat{\gamma})} : \frac{(y-X\hat{\beta})^*(y-X\hat{\beta})}{n-p}$$

with $\nu_1 = 1$ and $\nu_2 = n-p$ degrees of freedom.

Consider a model of one-way classification

$$y_{ij} = a_i + e_{ij}, \quad i = 1, 2, \ldots, I, \quad j = 1, 2, \ldots, n_i.$$

We want to test the hypothesis that $\underset{1m}{L}\,\underset{m1}{\gamma} = 0$ where

$$\underset{1m}{L} = \{l_k\},\ k = 1, 2, \ldots, m,\ m \leqslant I,\ \underset{m1}{\gamma} = \begin{bmatrix}\gamma_1\\ \gamma_2\\ \cdots\\ \gamma_m\end{bmatrix} \text{ and } \beta = \begin{bmatrix}\gamma\\ \delta\end{bmatrix} = \begin{bmatrix}a_1\\ a_2\\ \cdots\\ a_m\\ a_{m+1}\\ \cdots\\ a_I\end{bmatrix}$$

It is easy to verify that in this case $\hat{\gamma}_i = \bar{y}_{i.} = \dfrac{\sum_j y_{ij}}{n_i}$, and $\mathrm{Var}(L\hat{\gamma}) =$

$= \operatorname{Var}\Big(\sum_{k=1}^{m} l_k\hat{\gamma}_k\Big) = \sigma^2 \sum_{k=1}^{m} \frac{l_k^2}{n_k}$. The random variable

$$F = \frac{\Big(\sum_{k=1}^{m} l_k\hat{\gamma}_k\Big)^2}{\sum_{k=1}^{m} l_k^2/n_k} : \frac{\sum_{ij}(y_{ij}-\bar{y}_{i.})^2}{n-I}$$

with 1 and $n-I$ d. f. can be used to test the hypothesis mentioned above.

2° In the model of two-way classification

$$y_{ijl} = \mu + a_i + \beta_j + e_{ijl}, \qquad \begin{array}{l} i = 1, 2, \ldots, I \\ j = 1, 2, \ldots, \ldots, J \\ l = 1, 2, \ldots, n_{ij} \end{array}$$

we can use the theorem to verify that comparisons of some parameters a_i or of some parameters β_j are equal to zero. Before doing this it is necessary to reparametrize the model to obtain non-singular matrix $S = X^*X$.

The same situation holds in the case of the model

$$y_{ij} = \mu + a_i + e_{ij}, \quad i = 1, 2, \ldots, I; \quad j = 1, 2, \ldots, n_i$$

because then the determinant of the matrix S is equal to zero.

REFERENCES

[1] Oktaba, W., *On the linear hypothesis in the theory of normal regression*, Ann. Univ. Mariae Curie-Skłodowska, Sectio A, **11**, 2 (1957), p. 17-71.

Streszczenie

W niniejszej pracy, będącej kontynuacją pracy [1] przedstawiam dowód następującego twierdzenia:

Niech składowe $y_1 y_2, \ldots, y_n$ *wektora kolumnowego* $\underset{n1}{y}$ *wielokrotnego modelu regresyjnego*

(1) $$\mu = E(y) = \underset{np}{X}\,\underset{p1}{\beta}$$

będą normalnymi niezależnymi zmiennymi losowymi o średnich $\mu_1, \mu_2, \ldots \mu_n$ *i wspólnej wariancji* σ^2. *Nadto, niech rząd macierzy* X *będzie równy liczbie* p *parametrów* β *i niech prawdziwą będzie hipoteza*

(2) $$\underset{q1}{\varphi} = \underset{qm}{L}\;\underset{m1}{\gamma} = \underset{q1}{\varphi_0}$$

gdzie rzędem macierzy L jest q i $\gamma = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdots \\ \beta_m \end{bmatrix}$ *jest podwektorem wektora* $\underset{p1}{\beta} = \begin{bmatrix} \underset{m1}{\gamma} \\ \cdots \\ \underset{p-m,1}{\delta} \end{bmatrix}$. *Wtedy zmienna losowa*

$$(3) \qquad F = \frac{(L\hat{\gamma}-\varphi_0)^*(LS^{11}L^*)^{-1}(L\hat{\gamma}-\varphi_0)}{q} : \frac{(y-X\hat{\beta})^*(y-X\hat{\beta})}{n-p}$$

ma rozkład F z q i n-p stopniami swobody, przy czym $\underset{pp}{S} = X^*X$ *i* $\underset{mm}{S^{11}}$ *jest podmacierzą macierzy* $S^{-1} = \begin{bmatrix} S^{11} & S^{12} \\ S^{21} & S^{22} \end{bmatrix}$ *i $\hat{\gamma}$ jest oceną parametru γ.*

Poza tym zamieszczam kilka przykładów, w których korzystając z formy zmiennej (3) można zweryfikować hipotezę (2) głoszącą, że q kombinacji liniowych m parametrów $\beta_1, \beta_2, \ldots, \beta_m$ przybiera q danych z góry wartości.

Резюме

В этой работе, составляющей продолжение работы [1], мы представляем доказательство следующей теоремы.

Пусть компоненты $y_1, y_2, \ldots, y_n$ столбцевого вектора $\underset{n1}{y}$ модели множественной регрессии

$$(1) \qquad \mu = E(y) = \underset{np}{X}\,\underset{p1}{\beta}$$

будут нормальными независимыми случайными величинами с математическими ожиданиями $\mu_1, \mu_2, \ldots, \mu_n$ и с общей дисперсией σ^2. Кроме того, пусть ранг матрицы X будет равен числу p параметров β и пусть гипотеза

$$(2) \qquad \underset{q1}{\varphi} = \underset{qm}{L}\,\underset{m1}{\gamma} = \underset{q1}{\varphi_0}$$

будет правильна; ранг матрицы L равен q, а $\gamma = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdots \\ \beta_m \end{bmatrix}$ *блочному вектору вектора* $\underset{p1}{\beta} = \begin{bmatrix} \underset{m1}{\gamma} \\ \underset{p-m,1}{\delta} \end{bmatrix}$. *Тогда случайная величина*

$$(3) \qquad F = \frac{(L\hat{\gamma}-\varphi_0)^*(LS^{11}L^*)^{-1}(L\hat{\gamma}-\varphi_0)}{q} : \frac{(y-X\hat{\beta})^*(y-X\hat{\beta})}{n-p}$$

следует закону F *о* q *и* $n-p$ *степеней свободы, где* $S = X^{*}X$, *а* S^{11}_{mm} *является блочной матрицей матрицы* $S^{-1} = \left[\begin{array}{c|c} S^{11} & S^{12} \\ \hline S^{21} & S^{22} \end{array}\right]$, $\hat{\gamma}$ *же является оценкой параметра* γ.

Наконец мы дали несколько примеров, в которых пользуясь величиной (3), можно проверить нулевую гипотезу (2), которая гласит что q линейных комбинаций m параметров $\beta_1, \beta_2, \ldots, \beta_m$ имеют определённые значения.