A PROBLEM IN LEAST SQUARES

By Jan K. Wiśniewski

§1. We are dealing with two variables, the observed values of which are denoted x and y respectively. The pairs of observations are divided into r groups, numbering $n_1, n_2, \cdots n_r$ pairs. Suppose in each group we determine a regression equation of the following shape:

$$y_i = a_i + b_i x + \cdots m_i x^s \tag{1}$$

where y_i denotes the value of the "dependent" variable obtained from the regression equation, while y without any subscript denotes its observed value. The r regression equations of type (1) are not assumed independent; on the contrary, we postulate that

$$\sum_1^r y_i = a_0 + b_0 x + \cdots m_0 x^s \tag{2}$$

be fulfilled identically in x; $a_0, b_0, \cdots m_0$ being predetermined numbers. This leads to the following conditions:

$$\sum_1^r a_i = a_0 \qquad \sum_1^r b_i = b_0 \cdots \sum_1^r m_i = m_0. \tag{3}$$

The magnitude to be minimized under the theory of least squares is now

$$Z = \sum_1^{r-1} \sum_i [y - (a_i + b_i x + \cdots m_i x^s)]^2 + \sum_r \left\{ y - \left[\left(a_0 - \sum_1^{r-1} a_i \right) + \left(b_0 - \sum_1^{r-1} b_i \right) x + \cdots \left(m_0 - \sum_1^{r-1} m_i \right) x^s \right] \right\}^2. \tag{4}$$

The normal equations derived from (4) are of the following shape:

. . . .

$$n_j a_j + n_r \sum_1^{r-1} a_i + b_j \sum_j x + \left(\sum_1^{r-1} b_i \right) \left(\sum_r x \right) + \cdots m_j \sum_j x^s + \left(\sum_1^{r-1} m_i \right) \left(\sum_r x^s \right) = \sum_j y - \sum_r y + n_r a_0 + b_0 \sum_r x + \cdots m_0 \sum_r x^s \tag{5}$$

. . . .

. . . .