Regression with Non-independent Values of the Dependent Variable

December, 1957 Robert G. D. Steel * BU-89-M

Consider the model

$$Y_i = \mu + \beta x_i + \epsilon_i$$

where $x_i = X_i - \bar{x}$, $i=1,\cdots,n$, and ϵ_i is normally distributed with zero mean. The complete set of observations may be written as

$$Y = X\binom{\mu}{\beta} + R$$

where Y and R are n x 1 matrices and X is an n x 2 matrix. Let the covariance matrix of the ϵ's, i.e. RR', be M.

In the usual regression, we assume $M = I\sigma^2$, i.e. the ϵ's are independent and have a common variance. A fairly common regression problem arises when

$$M = \begin{pmatrix} w_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & & & w_n \end{pmatrix} \sigma^2 ,$$

i.e. when the ϵ's are assumed to be independent but are multiples of a common variance; a regression of means, based on different numbers of observations, on an independent variable exemplifies this case. The least common case, where M is a non-singular, covariance matrix, arises when the assumptions of independence are not justified.

When normality is assumed, the joint distribution of the Y_i's requires the quadratic form

$$Q = \frac{1}{2} \sum_{i,j} \sigma^{ij}(Y_i-\mu-\beta x_i)(Y_j-\mu-\beta x_j)$$

*Cornell University, Ithaca, New York

where σ^{ij} is an element of M^{-1}. It is required to minimize this expression by appropriately estimating μ and β.

The partial derivatives are:

$$\frac{\delta Q}{\delta \mu} = - \sum_i (Y_i - \mu - x_i\beta) \sum_j \sigma^{ij}$$

$$\frac{\delta Q}{\delta \beta} = - \sum_i (Y_i - \mu - x_i\beta) \sum_j \sigma^{ij} x_j$$

Equating the derivatives to zero, we obtain

$$(\sum_{i,j} \sigma^{ij})\hat{\mu} + (\sum_i x_i \sum_j \sigma^{ij})\hat{\beta} = \sum_i Y_i \sum_j \sigma^{ij}$$

$$(\sum_i \sum_j \sigma^{ij} x_j)\hat{\mu} + (\sum_i x_i \sum_j \sigma^{ij} x_j)\hat{\beta} = \sum_i Y_i \sum_j \sigma^{ij} x_j$$

Solution of these equations is given by

$$\hat{\mu} = \frac{\begin{vmatrix} \sum_i Y_i \sum_j \sigma^{ij} & \sum_i x_i \sum_j \sigma^{ij} \\ \sum_i Y_i \sum_j \sigma^{ij} x_j & \sum_i x_i \sum_j \sigma^{ij} x_j \end{vmatrix}}{\begin{vmatrix} \sum_{i,j} \sigma^{ij} & \sum_i x_i \sum_j \sigma^{ij} \\ \sum_{i,j} \sigma^{ij} x_j & \sum_i x_i \sum_j \sigma^{ij} x_j \end{vmatrix}} \qquad (1)$$

$$\hat{\beta} = \frac{\begin{vmatrix} \sum_{i,j} \sigma^{ij} & \sum_i Y_i \sum_j \sigma^{ij} \\ \sum_{i,j} \sigma^{ij} x_j & \sum_i Y_i \sum_j \sigma^{ij} x_j \end{vmatrix}}{|\text{(as for } \hat{\mu})|} \qquad (2)$$

If we define $\bar{x}$ as

$$(\sum_i X_i \sum_j \sigma^{ij}) / \sum_{i,j} \sigma^{ij} \quad ,$$

then $\quad (\sum_i x_i \sum_j \sigma^{ij}) / \sum_{i,j} \sigma^{ij} = 0$

and $\quad \hat{\mu} = \sum_i Y_i \sum_j \sigma^{ij} / \sum_{i,j} \sigma^{ij}$.

When the ϵ's are independent and their variances are multiples of a common variance, equations (1) and (2) reduce to

$$\hat{\mu} = \frac{\begin{vmatrix} \Sigma w_i Y_i & \Sigma w_i x_i \\ \Sigma w_i x_i Y_i & \Sigma w_i x_i^2 \end{vmatrix}}{\begin{vmatrix} \Sigma w_i & \Sigma w_i x_i \\ \Sigma w_i x_i & \Sigma w_i x_i^2 \end{vmatrix}}$$

$$\hat{\beta} = \frac{\begin{vmatrix} \Sigma w_i & \Sigma w_i Y_i \\ \Sigma w_i x_i & \Sigma w_i x_i Y_i \end{vmatrix}}{\left| \text{(as for } \hat{\mu}) \right|}$$

since the σ^2's now cancel. In this case, it is customary to define $\bar{x}$ as $(\Sigma w_i X_i)/\Sigma w_i$. Then $\Sigma w_i x_i = 0$ and $\hat{\mu} = (\Sigma w_i Y_i)/\Sigma w_i$.

When the ϵ's are independent and have a common variance, equations (1) and (2) become

$$\hat{\mu} = \frac{\begin{vmatrix} \Sigma Y_i & \Sigma x_i \\ \Sigma x_i Y_i & \Sigma x_i^2 \end{vmatrix}}{\begin{vmatrix} n & \Sigma x_i \\ \Sigma x_i & \Sigma x_i^2 \end{vmatrix}}$$

$$\hat{\beta} = \frac{\begin{vmatrix} n & \Sigma Y_i \\ \Sigma x_i & \Sigma x_i Y_i \end{vmatrix}}{|(\text{as for } \hat{\mu})|}$$

and the σ^2's again cancel. Here we define $\bar{x}$ as $(\Sigma X_i)/n$. Now $\Sigma x_i=0$ and $\hat{\mu}=\Sigma Y_i/n$.

In the general case, i.e. correlated ϵ's, we have assumed that M was at our disposal. Since $\hat{\mu}$ and $\hat{\beta}$ turn out to be linear combinations of the Y's, we are able to compute their variances and covariance. In the two special cases, the same is true. However if σ^2 is unknown, it is now possible and relatively easy to estimate σ^2 and, in turn, compute an estimate of the variances and covariance. Known sampling distributions apply to these two special cases.

Example

Consider the following set of data

i	x	Y
1	-1	114
2	0	124
3	+1	143

with

$$M = \begin{pmatrix} 7/24 & 1/6 & 1/12 \\ 1/6 & 4/6 & 2/6 \\ 1/12 & 2/6 & 7/6 \end{pmatrix}$$

and

$$M^{-1} = \begin{pmatrix} 4 & -1 & 0 \\ -1 & 2 & -1/2 \\ 0 & -1/2 & 1 \end{pmatrix} \quad ; \quad \begin{matrix} \Sigma\sigma^{1j} = 3 \\ \Sigma\sigma^{2j} = 1/2 \\ \Sigma\sigma^{3j} = 1/2 \end{matrix}$$

Applying equations (1) and (2), we find

$$\hat{\mu} = \frac{\begin{vmatrix} 3(114)+\frac{1}{2}(124)+\frac{1}{2}(143) & -1(3)+1(\frac{1}{2}) \\ -4(114)+(1-\frac{1}{2})(124)+1(143) & -1(-4)+1(1) \end{vmatrix}}{\begin{vmatrix} 4 & -2\frac{1}{2} \\ 3(-1)+1(\frac{1}{2}) & 5 \end{vmatrix}}$$

$$= 127.3$$

$$\hat{\beta} = \frac{\begin{vmatrix} 4 & 475.5 \\ -2.5 & -251 \end{vmatrix}}{55/4}$$

$$= 13.4$$

If there were zero covariances and the variances were 7/24, 4/6 and 7/6, then

$$M = \begin{pmatrix} 7/24 & 0 & 0 \\ 0 & 4/6 & 0 \\ 0 & 0 & 7/6 \end{pmatrix} \quad \text{and} \quad M^{-1} = \begin{pmatrix} 24/7 & 0 & 0 \\ 0 & 6/4 & 0 \\ 0 & 0 & 6/7 \end{pmatrix}$$

Now

$$\bar{x} = \frac{\Sigma w_i x_i}{\Sigma w_i}$$

$$= \frac{\frac{24}{7}(-1)+\frac{6}{4}(0)+\frac{6}{7}(1)}{\frac{24}{7}+\frac{6}{4}+\frac{6}{7}}$$

$$= -\frac{4}{9}$$

and $x_1' = -\frac{5}{9}$, $x_2' = \frac{4}{9}$, $x_3' = 1\frac{4}{9}$.

$$\hat{\mu} = \frac{\Sigma w_i Y_i}{\Sigma w_i}$$

$$= \frac{\frac{24}{7}(114) + \frac{6}{4}(124) + \frac{6}{7}(143)}{\frac{24}{7} + \frac{6}{4} + \frac{6}{7}}$$

$$= 126.8$$

$$\hat{\beta} = \frac{\Sigma w_i x_i' Y_i}{\Sigma w_i x_i'^2}$$

$$= \frac{\frac{24}{7}(-\frac{5}{9})114 + \frac{6}{4}(\frac{4}{9})124 + \frac{6}{7}(\frac{13}{9})143}{\frac{24}{7}(\frac{5}{9})^2 + \frac{6}{4}(\frac{4}{9})^2 + \frac{6}{7}(\frac{13}{9})^2}$$

$$= 13.5 \quad .$$

If the variances were equal, then we would find

$$\hat{\mu} = \frac{\Sigma Y}{3}$$

$$= 127$$

$$\hat{\beta} = \frac{\Sigma xY}{\Sigma x^2}$$

$$= 14.5$$

References

Rao, C. R., Advanced Statistical Methods in Biometric Research, Chap. 3, sec'n 3a.9, 1952. John Wiley and Sons.

Robson, D. R., B. U. Seminar, December 6, 1957.