

An Application of Fieller's Theorem to Obtain an Interval Estimate  
of the Common X-intercept of Several Regression Lines

BU-238-M

D. S. Robson

March, 1967

Abstract

Fieller's Theorem applies to the problem  $Y_{ij} = \alpha_i + \beta_i X_{ij} + \epsilon_{ij}\sigma$ ,  $i=1, \dots, r$ ;  
 $j=1, \dots, n_i$ ;  $\epsilon_{ij} \sim N(0,1)$ ;  $\alpha_1/\beta_1 = \dots = \alpha_r/\beta_r = -M$  to give the interval estimate

$$\frac{m - k\bar{X} \pm k \sqrt{(\bar{X}-m)^2 + (1-k^2)(C_X - \bar{X}^2)}}{1-k^2}$$

where

$$m = - \frac{\sum_1^r (\bar{Y}_i - b_i \bar{X}_i)}{\sum_1^r b_i}$$

$$k^2 = \frac{t_{\alpha, \Sigma(n_i-2)}^2 \sum_1^r s_{b_i}^2}{\left( \sum_1^r b_i^2 \right)^2}$$

$$s_{b_i}^2 = \frac{s^2}{\sum_1^{n_i} (X_{ij} - \bar{X}_i)^2}$$

$$s^2 = \frac{\sum_1^r \sum_1^{n_i} [Y_{ij} - \bar{Y}_i - b_i (X_{ij} - \bar{X}_i)]^2}{\sum_1^r (n_i - 2)}$$

$$\bar{X} = \frac{\sum_{i=1}^r s_{b_i}^2 \bar{X}_i}{\sum_{i=1}^r s_{b_i}^2}$$

$$C_X = \frac{s^2 \sum_{i=1}^r \frac{1}{n_i} + \sum_{i=1}^r s_{b_i}^2 \bar{X}_i^2}{\sum_{i=1}^r s_{b_i}^2}$$

An Application of Fieller's Theorem to Obtain an Interval Estimate  
of the Common X-intercept of Several Regression Lines

BU-238-M

D. S. Robson

March, 1967

Fieller's Theorem (see Finney, 1952, p. 27) provides an interval estimate of the X-intercept

$$M = - \frac{\alpha}{\beta}$$

for a linear regression problem  $Y = \alpha + \beta X + \epsilon$  in which the errors  $\epsilon$  are independent  $N(0,1)$  and in which the least squares estimator of  $\beta$

$$b = \frac{\sum_{i=1}^n (X_i - \bar{X}) Y_i}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

is significantly different from 0; i.e., in which

$$\frac{b^2}{s_b^2} > t_{\alpha}^2$$

or

$$k^2 = \frac{t_{\alpha}^2 s_b^2}{b^2} < 1 \quad .$$

---

Biometrics Unit, Department of Plant Breeding and Biometry, Cornell University,  
Ithaca, New York.

The  $100(1-\alpha)\%$  interval estimate of  $M$  is then given by the two roots of

$$\frac{(a+bM)^2}{s^2 \left[ \frac{1}{n} + \frac{(\bar{X}-M)^2}{\Sigma(X-\bar{X})^2} \right]} = t_{\alpha}^2$$

where

$$s^2 = \frac{1}{n-2} \Sigma(Y_i - a - bX_i)^2 \quad \text{and} \quad a = \bar{Y} - b\bar{X} .$$

These two roots may be expressed as

$$\frac{m - k^2\bar{X} \pm k \sqrt{(\bar{X}-m)^2 + (1-k^2)\Sigma(X-\bar{X})^2/n}}{1-k^2}$$

where

$$m = -\frac{a}{b} .$$

This approach may be extended to provide an interval estimate of an hypothesized common X-intercept of  $r$  regression lines

$$Y_{ij} = \alpha_i + \beta_i X_{ij} + \epsilon_{ij} \sigma \quad i=1, \dots, r; \quad j=1, \dots, n_i$$

in which

$$\frac{\alpha_1}{\beta_1} = \frac{\alpha_2}{\beta_2} = \dots = \frac{\alpha_r}{\beta_r} = -M ,$$

provided that the  $\epsilon_{ij}$  are independent  $N(0,1)$  and the average slope

$$\bar{b} = \frac{b_1 + \dots + b_r}{r}$$

differs significantly from 0; i.e.,

$$\frac{(b_1 + \dots + b_r)^2}{s^2 \left[ \frac{1}{\Sigma(X_1 - \bar{X}_1)^2} + \dots + \frac{1}{\Sigma(X_r - \bar{X}_r)^2} \right]} > t_\alpha^2$$

or

$$k^2 = \frac{t_\alpha^2 (s_{b_1}^2 + \dots + s_{b_r}^2)}{(b_1 + \dots + b_r)^2} < 1 .$$

The interval is given by the two roots

$$\frac{m - k^2 \tilde{X} \pm k \sqrt{(\tilde{X} - m)^2 + (1 - k^2)(C_X - \tilde{X}^2)}}{1 - k^2}$$

where

$$m = - \frac{a_1 + \dots + a_r}{b_1 + \dots + b_r}$$

$$\tilde{X} = \frac{\frac{\bar{X}_1}{\Sigma(X_1 - \bar{X}_1)^2} + \dots + \frac{\bar{X}_r}{\Sigma(X_r - \bar{X}_r)^2}}{\frac{1}{\Sigma(X_1 - \bar{X}_1)^2} + \dots + \frac{1}{\Sigma(X_r - \bar{X}_r)^2}}$$

$$C_X = \frac{\left( \frac{1}{n_1} + \frac{\bar{X}_1^2}{\Sigma(X_{1i} - \bar{X}_1)^2} \right) + \dots + \left( \frac{1}{n_r} + \frac{\bar{X}_r^2}{\Sigma(X_{ri} - \bar{X}_r)^2} \right)}{\frac{1}{\Sigma(X_{1i} - \bar{X}_1)^2} + \dots + \frac{1}{\Sigma(X_{ri} - \bar{X}_r)^2}}$$

The point estimator  $m$  is evidently not the most efficient; a better point estimator may be derived by an iterative solution of the maximum likelihood equations. The latter approach, however, provides only asymptotically valid interval estimators while the above interval estimation procedure is exact for small samples.

This problem was posed by Dr. Alan Dobson of the Cornell Veterinary College; in the context of his problem there was no reason to suppose homogeneity of either the slopes ( $\beta_i$ ) or the Y-intercepts ( $\alpha_i$ ).

#### Reference

Finney, D. J. Statistical Method in Biological Assay. Charles Griffin and Co., London, 1952.

Appendix: Maximum likelihood point estimation.

The assumed model is  $Y_{ij} = \beta_i(X_{ij} - M) + \epsilon_{ij}\sigma$ ,  $i=1, \dots, r$ ;  $j=1, \dots, n$ ; with  $\epsilon_{ij}$  independent and identically distributed  $N(0,1)$ . The likelihood equations are:

$$\hat{M} = \frac{\sum_{i=1}^r n_i \hat{\beta}_i (\hat{\beta}_i \bar{X}_i - \bar{Y}_i)}{\sum_{i=1}^r n_i \hat{\beta}_i^2} = \hat{M}(\hat{\beta})$$

$$\hat{\beta}_i = \frac{\sum_{j=1}^{n_i} (X_{ij} - \hat{M}) Y_{ij}}{\sum_{j=1}^{n_i} (X_{ij} - \hat{M})^2} = \hat{\beta}_i(\hat{M})$$

These solutions may be calculated iteratively, starting with

$$\hat{\beta}_{0i} = \frac{\sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)(Y_{ij} - \bar{Y}_i)}{\sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2}$$

$$\hat{M}_0 = \hat{M}(\hat{\beta}_0)$$

and letting

$$b_{0i} = \hat{\beta}_i(\hat{M}_0) \quad .$$

The revised estimates are then

$$\hat{M}_1 = \hat{M}_0 + \frac{\sum_i n_i \hat{\beta}_{0i} [b_{0i} (\bar{X}_i - \hat{M}_0) - \bar{Y}_i]}{\sum_i n_i \hat{\beta}_{0i}^2 \left[ \frac{j}{\sum_j (X_{ij} - \hat{M}_0)^2} \right]}$$

$$\hat{\beta}_{1i} = b_{0i} - \hat{\beta}_{0i} \left[ \frac{n_i (\bar{X}_i - \hat{M}_0) (\hat{M}_0 - \hat{M}_1)}{\sum_j (X_{ij} - \hat{M}_0)^2} \right] .$$

With  $b_{1i}$  defined as

$$b_{1i} = \hat{\beta}(\hat{M}_1)$$

these calculations may be repeated to obtain  $\hat{M}_2$  and  $\hat{\beta}_{2i}$ , and so on until the estimates stabilize at the maximum likelihood values  $\hat{M}$  and  $\hat{\beta}_i$ .

The variances and covariances of the maximum likelihood estimators are

$$\text{var}(\hat{M}) = \frac{\sigma^2}{\sum_i n_i \beta_i^2 \left[ \frac{j}{\sum_j (X_{ij} - \bar{X}_i)^2} \right]}$$

$$\text{var}(\hat{\beta}_i) = \left\{ n_i \beta_i^2 + \sum_{i' \neq i} n_{i'} \beta_{i'}^2 \left[ \frac{j}{\sum_j (X_{i',j} - \bar{X}_{i'})^2} \right] \right\} \frac{\text{var}(\hat{M})}{\sum_j (X_{i,j} - \bar{M})^2}$$



$$\text{cov}(\hat{M}, \hat{\beta}_i) = - n_i \beta_i \frac{\bar{X}_i - M}{\sum_j (X_{ij} - M)^2} \text{var}(\hat{M})$$

$$\text{cov}(\hat{\beta}_i, \hat{\beta}_j) = - \frac{\text{cov}(\hat{M}, \hat{\beta}_i) \text{cov}(\hat{M}, \hat{\beta}_j)}{\text{var}(\hat{M})} .$$

The variance  $\sigma^2$  may be estimated by

$$\hat{\sigma}^2 = \frac{1}{\sum_i (M_i - 1) - 1} \sum_i \sum_j [Y_{ij} - \hat{\beta}_i (X_{ij} - M)]^2 .$$