DOCUMENT RESUME

ED 067 405 TM 001 800

AUTHOR	Aitkin, Murray A.
TITLE	Fixed-Width Confidence Intervals in Linear Regression with Applications to the Johnson-Neyman Technique.
INSTITUTION	Educational Testing Service, Princeton, N.J.
REPORT NO	RB-72-18
PUB DATE	Apr 72
NOTE	17p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Analysis of Covariance; *Mathematical Applications; *Mathematical Models; *Multiple Regression Analysis; *Statistical Studies
IDENTIFIERS	Gafarian (A V); Johnson (P O); Neyman (J)

ABSTRACT

Fixed-width confidence intervals for a population regression line over a finite interval of x have recently been derived by Gafarian. The method is extended to provide fixed-width confidence intervals for the difference between two population regression lines, resulting in a simple procedure analogous to the Johnson-Neyman technique. (Author)

RESEARCH BULLETIN

RB-72-18

FIXED-WIDTH CONFIDENCE INTERVALS IN LINEAR REGRESSION WITH APPLICATIONS TO THE JOHNSON-NEYMAN TECHNIQUE

Murray A. Aitkin

This Bulletin is a draft for interoffice circulation. Corrections and suggestions for revision are solicited. The Bulletin should not be cited as a reference without the specific permission of the author. It is automatically superseded upon formal publication of the material.

Educational Testing Service
Princeton, New Jersey
April 1972

FIXED-WIDTH CONFIDENCE INTERVALS IN LINEAR REGRESSION WITH APPLICATIONS TO THE JOHNSON-NEYMAN TECHNIQUE

Abstract

Fixed-width confidence intervals for a population regression line over a finite interval of x have recently been derived by Gafarian. The method is extended to provide fixed-width confidence intervals for the difference between two population regression lines, resulting in a simple procedure analogous to the Johnson-Neyman technique.

FIXED-WIDTH CONFIDENCE INTERVALS IN LINEAR REGRESSION WITH APPLICATIONS TO THE JOHNSON-NEYMAN TECHNIQUE

1. SUMMARY AND INTRODUCTION

In the simple linear regression model, the length of the confidence interval for a predicted y at a given x depends on x, increasing as x departs from $\bar{x}$. Simultaneous confidence intervals for the predicted y for all x are given by the well-known Working-Hotelling [1929] hyperbolic band (Miller, 1966, p. 111). Gafarian [1964] has recently derived simultaneous fixed-width confidence intervals for the predicted y, for all x in a finite range centered at $\bar{x}$, and has provided tables for the calculation of these intervals.

An important application of the Working-Hotelling confidence band is in two-group analysis of covariance with a single covariate, when the two population regression lines are not parallel. The Johnson-Neyman [1936] technique, as modified by Potthoff [1964], for locating the values of x for which a significant difference can be asserted between the population regression lines, is a simple application of the Working-Hotelling procedure.

The fixed-width confidence interval procedure of Gafarian may also be applied to the above analysis of covariance model, providing a different and simple procedure for locating the values of x for which a significant difference can be asserted. Gafarian's tables, with slight modifications, may be used in this procedure.

Conclusions are drawn about the relative merits of the two procedures, and an example is given.

2. THE WORKING-HOTELLING AND GAFARIAN BANDS

For the linear regression model

$$\underline{y} = X\underline{\beta} + \underline{\epsilon}$$

where

$$X' = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{bmatrix} \quad , \quad \underline{\beta}' = (\beta_0, \beta_1) \quad , \quad \underline{y}' = (y_1, \ldots, y_n)$$

and $\underline{\epsilon} \sim N(\underline{0}, \sigma^2 I_n)$, the (unbiased) maximum likelihood estimates of $\underline{\beta}$ and σ^2 are

$$\hat{\underline{\beta}} = (X'X)^{-1}X'\underline{y} = S^{-1}X'\underline{y} \quad ,$$

$$\hat{\sigma}^2 = (\underline{y}'\underline{y} - \underline{y}'X\hat{\underline{\beta}})/(n - 2) \quad ,$$

and the covariance matrix of $\hat{\beta}_0, \hat{\beta}_1$ is

$$\sigma^2 S^{-1} = \sigma^2 \begin{bmatrix} S^{00} & S^{01} \\ S^{10} & S^{11} \end{bmatrix} = \sigma^2 \begin{bmatrix} \frac{1}{n} + \frac{\bar{x}^2}{\Sigma(x_i - \bar{x})^2} & -\frac{\bar{x}}{\Sigma(x_i - \bar{x})^2} \\ -\frac{\bar{x}}{\Sigma(x_i - \bar{x})^2} & \frac{1}{\Sigma(x_i - \bar{x})^2} \end{bmatrix} .$$

A $100(1-\alpha)\%$ confidence interval for $\beta_0 + \beta_1 x_0 = \underline{\beta}'\underline{x}_0$, where $\underline{x}_0' = (1 \quad x_0)$, when x_0 is specified, is

$$\underline{\beta}'\underline{x}_0 \in \hat{\underline{\beta}}'\underline{x}_0 \pm t_{n-2}^{\alpha/2}\,\hat{\sigma}(\underline{x}_0' S^{-1}\underline{x}_0)^{1/2} \quad .$$

If simultaneous confidence intervals for all x with confidence coefficient $100(1-\alpha)\%$ are required, then $t^{\alpha/2}$ is replaced by $\{2F_{2,n-2}^{\alpha}\}^{1/2}$, producing the Working-Hotelling confidence band for the entire regression line:

$$\underline{\beta}'\underline{x} \in \hat{\underline{\beta}}'\underline{x} \pm \{2F_{2,n-2}^{\alpha}\}\,\hat{\sigma}(\underline{x}' S^{-1}\underline{x})^{1/2}$$

(see Miller, 1966, p. 111, for further details).

In all practical cases, however, we will be interested in the regression equation over a finite range of x, say $a \le x \le b$. For such a range, the Working-Hotelling band will have a confidence coefficient greater than $100(1-\alpha)\%$, so this band is wastefully wide.

Gafarian [1964] derived fixed-width simultaneous confidence intervals over a finite range by the following argument. If the linear model is reparametrized as

$$\underline{y} = X^*\underline{\gamma} + \underline{\epsilon}$$

where

$$X^{*\prime} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 - \bar{x} & x_2 - \bar{x} & \cdots & x_n - \bar{x} \end{bmatrix}$$

then $\hat{\sigma}^2$ is as before, and

$$\hat{\gamma}' = (\bar{y}, \hat{\beta}_1) \quad ,$$

while the covariance matrix of $\hat{\gamma}_0$, $\hat{\gamma}_1$ is

$$\sigma^2 S_*^{-1} = \sigma^2 (X^{*\prime} X^*)^{-1} = \sigma^2 \begin{bmatrix} \frac{1}{n} & 0 \\ 0 & \frac{1}{\Sigma(x_i - \bar{x})^2} \end{bmatrix} = \sigma^2 \begin{bmatrix} s_*^{00} & s_*^{01} \\ s_*^{10} & s_*^{11} \end{bmatrix} = \sigma^2 \begin{bmatrix} s_0^2 & 0 \\ 0 & s_1^2 \end{bmatrix} .$$

Then

$$(\hat{\gamma}_0 - \gamma_0)/s_0 \sigma = \sqrt{n}\,(\hat{\gamma}_0 - \gamma_0)/\sigma$$

and

$$(\hat{\gamma}_1 - \gamma_1)/s_1 \sigma$$

are independent standard normal variables, and hence

$$T_0 = (\hat{\gamma}_0 - \gamma_0)/s_0 \hat{\sigma} \quad , \qquad T_1 = (\hat{\gamma}_1 - \gamma_1)/s_1 \hat{\sigma}$$

have a bivariate t -distribution with $n - 2$ degrees of freedom (see Dunnett & Sobel, 1954, Press 1972, p. 127). Then

$$P\{|(\hat{\underset{\sim}{\gamma}} - \underset{\sim}{\gamma})'\underset{\sim}{x}^*| \le \delta\hat{\sigma} \,,\ \forall\ x \in (a,b)\}$$

$$= P\{|s_0T_0 + s_1(x - \bar{x})T_1| \le \delta \,,\ \forall\ x \in (a,b)\}$$

$$= P\{|T_0 + s_1(x - \bar{x})T_1/s_0| \le \delta/s_0 \,,\ \forall\ x \in (a,b)\}$$

and this probability is given by the integral of the bivariate t -density over a parallelogram in the (t_0,t_1) plane. Gafarian tabulates these probabilities for the special case in which the interval (a,b) is symmetric about $\bar{x}$, i.e., $a = \bar{x} - h$, $b = \bar{x} + h$. In this case the probabilities reduce to a multiple of the integral of the density over the triangular region with vertices $(0,0)$, $(0,\delta/s_1h)$, $(\delta/s_0,0)$. The probabilities are tabulated as a function of n , $c = s_0/hs_1$ and $d = \delta/s_0$. Given n , c and the required confidence coefficient, d and hence δ may be obtained. The bivariate t -density has not been integrated over more general parallelograms.

3. APPLICATION TO ANALYSIS OF COVARIANCE

Consider the two-sample analysis of covariance model with a single covariate, when the regression lines are not parallel. Let

$$\underset{\sim}{y} = X\underset{\sim}{\beta} + \underset{\sim}{\epsilon}$$

where

$$X' = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_{n_1} & x_{n_1+1} & \cdots & x_n \\ 0 & 0 & \cdots & 0 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 & x_{n_1+1} & \cdots & x_n \end{bmatrix} , \qquad \underset{\sim}{\beta}' = (\beta_0,\beta_1,\beta_2,\beta_3) ,$$

and

$$\underset{\sim}{\epsilon} \sim N(\underset{\sim}{0}, \sigma^2 I_n) \quad .$$

This model represents the two regressions

$$y_i = \beta_0 + \beta_1 x_i \qquad i = 1, \ldots, n_1$$

$$y_i = \beta_0 + \beta_2 + (\beta_1 + \beta_3) x_i \qquad i = n_1 + 1, \ldots, n \quad .$$

Unbiased maximum likelihood estimates are again

$$\hat{\underset{\sim}{\beta}} = S^{-1} X' \underset{\sim}{y} \quad ,$$

$$\hat{\sigma}^2 = (\underset{\sim}{y}' \underset{\sim}{y} - \underset{\sim}{y}' X \hat{\underset{\sim}{\beta}})/(n - 2) \quad ,$$

and it is easily verified that the covariance matrix of $\hat{\underset{\sim}{\beta}}$ is

$$\sigma^2 S^{-1} = \sigma^2 (S^{ij}) = \sigma^2 \begin{bmatrix} S_1^{-1} & -S_1^{-1} \\ -S_1^{-1} & S_1^{-1} + S_2^{-1} \end{bmatrix} \quad , \quad i, j = 0, \ldots, 3$$

where

$$S_1^{-1} = \begin{bmatrix} \dfrac{1}{n_1} + \dfrac{\bar{x}_1^2}{\sum_{i=1}^{n_1} (x_i - \bar{x}_1)^2} & - \dfrac{\bar{x}_1}{\sum_{i=1}^{n_1} (x_i - \bar{x}_1)^2} \\ & \dfrac{1}{\sum_{i=1}^{n_1} (x_i - \bar{x}_1)^2} \end{bmatrix} \quad ,$$

$$S_2^{-1} = \begin{bmatrix} \dfrac{1}{n_2} + \dfrac{\bar{x}_2^2}{\sum_{i=n_1+1}^{n} (x_i - \bar{x}_2)^2} & -\dfrac{\bar{x}_2}{\sum_{i=n_1+1}^{n} (x_i - \bar{x}_2)^2} \\ \cdot & \dfrac{1}{\sum_{i=n_1+1}^{n} (x_i - \bar{x}_2)^2} \end{bmatrix} ,$$

so that, for example,

$$S^{33} = \frac{1}{\sum_{i=1}^{n_1} (x_i - \bar{x}_1)^2} + \frac{1}{\sum_{i=n_1+1}^{n} (x_i - \bar{x}_2)^2} ,$$

where $\bar{x}_j$ is the mean of x for the j-th group ($j = 1,2$). The distance between the population regression lines at $x = x_0$ is $\beta_2 + \beta_3 x_0$, and a $100(1 - \alpha)\%$ confidence interval for this distance, when x_0 is specified, is

$$\beta_2 + \beta_3 x_0 \in \hat{\beta}_2 + \hat{\beta}_3 x_0 \pm t_{n-4}^{\alpha/2} \hat{\sigma}\{S^{22} + 2x_0 S^{23} + x_0^2 S^{33}\}^{1/2} .$$

If simultaneous confidence intervals for all x are desired with confidence coefficient $100(1 - \alpha)\%$, then $t_{n-4}^{\alpha/2}$ is replaced by $\{2F_{2,n-4}^{\alpha}\}^{1/2}$ (see Potthoff, 1964), and we obtain

$$\beta_2 + \beta_3 x \in \hat{\beta}_2 + \hat{\beta}_3 x \pm \{2F_{2,n-4}^{\alpha}\}^{1/2} \hat{\sigma}\{S^{22} + 2xS^{23} + x^2 S^{33}\}^{1/2} .$$

Again, in practical cases, our interest will be limited to a finite range of x. We now extend Gafarian's results to this case. First reparametrize the model as

$$\underline{y} = X^*\underline{\gamma} + \underline{\epsilon}$$

where

$$X^{*\prime} = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 & \cdots & 1 \\ x_1 - k & x_2 - k & \cdots & x_{n_1} - k & x_{n_1+1} - k & \cdots & x_n - k \\ 0 & 0 & \cdots & 0 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 & x_{n_1+1} - k & \cdots & x_n - k \end{bmatrix}$$

where k is to be determined so that $\hat{\gamma}_2$ and $\hat{\gamma}_3$ are independently distributed. The covariance matrix of the $\hat{\gamma}_i$ is now

$$\sigma^2 S_*^{-1} = \sigma^2 \begin{bmatrix} S_{1*}^{-1} & -S_{1*}^{-1} \\ -S_{1*}^{-1} & S_{1*}^{-1} + S_{2*}^{-1} \end{bmatrix}$$

where

$$S_{1*}^{-1} = \begin{bmatrix} \dfrac{1}{n_1} + \dfrac{(\bar{x}_1 - k)^2}{\sum\limits_{i=1}^{n_1} (x_i - \bar{x}_1)^2} & -\dfrac{(\bar{x}_1 - k)}{\sum\limits_{i=1}^{n_1} (x_i - \bar{x}_1)^2} \\ & \dfrac{1}{\sum\limits_{i=1}^{n_1} (x_i - \bar{x}_1)^2} \end{bmatrix}$$

and similarly for S_{2*}^{-1}. Thus

$$\operatorname{cov}(\hat{\gamma}_2, \hat{\gamma}_3) = -\sigma^2 \left(\frac{\bar{x}_1 - k}{\sum\limits_{i=1}^{n_1} (x_i - \bar{x}_1)^2} + \frac{\bar{x}_2 - k}{\sum\limits_{i=n_1+1}^{n} (x_i - \bar{x}_2)^2} \right)$$

and this is zero when

$$k = \{\bar{x}_1 \sum_{i=n_1+1}^{n} (x_i - \bar{x}_2)^2 + \bar{x}_2 \sum_{i=1}^{n_1} (x_i - \bar{x}_1)^2\}/\{ \sum_{i=1}^{n_1} (x_i - \bar{x}_1)^2 + \sum_{i=n_1+1}^{n} (x_i - \bar{x}_2)^2\} \quad .$$

For this value of k, we have

$$\operatorname{var}(\hat{\gamma}_2) = \sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2} + \frac{(\bar{x}_1 - \bar{x}_2)^2}{\sum_1^{n_1} (x_i - \bar{x}_1)^2 + \sum_{n_1+1}^{n} (x_i - \bar{x}_2)^2} \right) = s_2^2\sigma^2$$

$$\operatorname{var}(\hat{\gamma}_3) = \sigma^2\left(\frac{1}{\sum_1^{n_1} (x_i - \bar{x}_1)^2} + \frac{1}{\sum_{n_1+1}^{n} (x_i - \bar{x}_2)^2} \right) = s_3^2\sigma^2 \quad .$$

Then as in Section 2,

$$(\hat{\gamma}_2 - \gamma_2)/s_2\sigma \quad \text{and} \quad (\hat{\gamma}_3 - \gamma_3)/s_3\sigma$$

are independent standard normal variates, and hence

$$T_2 = (\hat{\gamma}_2 - \gamma_2)/s_2\hat{\sigma} \quad \text{and} \quad T_3 = (\hat{\gamma}_3 - \gamma_3)/s_3\hat{\sigma}$$

have a bivariate t-distribution with $n - 4$ degrees of freedom. Hence

$$P\{|\hat{\gamma}_2 - \gamma_2 + (\hat{\gamma}_3 - \gamma_3)(x - k)| \leq \delta\hat{\sigma} \ , \ \forall\ x \in (a,b)\}$$

$$= P\{|T_2 + s_3(x - k)T_3/s_2| \leq \delta/s_2 \ , \ \forall\ x \in (a,b)\} \quad .$$

Provided that (a,b) is symmetric about k, i.e., $a = k - h$, $b = k + h$, these probabilities may be read directly from Gafarian's tables, with the sample size minus 2 used for Gafarian's n, $c = s_3 h/s_2$, and $d = \delta/s_2$. Again given n, c and the required confidence coefficient, d and hence δ may be obtained. The confidence band is then

$$\gamma_2 + \gamma_3(x - k) = \beta_2 + \beta_3 x \in \hat{\beta}_2 + \hat{\beta}_3 x \pm \delta\hat{\sigma} \quad , \qquad x \in (k - h, k + h) .$$

The inconvenient requirement of symmetry about k results from the limited tables available.

4. RELATION TO THE JOHNSON-NEYMAN TECHNIQUE

The Johnson-Neyman technique is essentially the procedure for finding those values of x for which the Working-Hotelling simultaneous confidence intervals for the difference between the population regression lines do not include zero, so that a significant difference between the population means may be asserted for these values of x. The region of nonsignificance may be defined as consisting of those values of x for which zero is included in the confidence interval, i.e., for which

$$(\hat{\beta}_2 + \hat{\beta}_3 x)^2 < 2F^{\alpha}_{2,n-4}\hat{\sigma}^2\{S^{22} + 2xS^{23} + x^2 S^{33}\}$$

or

$$Q(x) = q_2 x^2 + 2q_1 x + q_0$$

$$= (\hat{\beta}_3^2 - \lambda\hat{\sigma}^2 S^{33})x^2 + 2(\hat{\beta}_2\hat{\beta}_3 - \lambda\hat{\sigma}^2 S^{23})x + (\hat{\beta}_2^2 - \lambda\hat{\sigma}^2 S^{22}) < 0$$

where $\lambda = 2F^{\alpha}_{2,n-4}$. If the roots of $Q(x) = 0$ are imaginary, then the region of nonsignificance is empty if $q_2 > 0$, and is the whole x-axis if $q_2 < 0$. If the roots $x_\theta < x_\phi$ are real, then the region of nonsignificance is $x_\Theta < x < x_\phi$ if $q_2 > 0$, and is $x > x_\phi$, $x < x_\Theta$ if $q_2 < 0$.

For the extension of the Gafarian technique, the region of nonsignificance consists of those x for which

$$\hat{\beta}_2 + \hat{\beta}_3 x - \delta\hat{\sigma} < 0 < \hat{\beta}_2 + \hat{\beta}_3 x + \delta\hat{\sigma} \quad ,$$

i.e.,

$$\frac{-\hat{\beta}_2 - \delta\hat{\sigma}}{\hat{\beta}_3} < x < \frac{-\hat{\beta}_2 + \delta\hat{\sigma}}{\hat{\beta}_3}$$

if $\hat{\beta}_3 > 0$, and

$$\frac{-\hat{\beta}_2 + \delta\hat{\sigma}}{\hat{\beta}_3} < x < \frac{-\hat{\beta}_2 - \delta\hat{\sigma}}{\hat{\beta}_3}$$

if $\hat{\beta}_3 < 0$. If both these values of x fall outside the interval $(k \pm h)$, on opposite sides of k , then the entire interval is a region of nonsignificance. If both fall outside, but on the same side of k , the region of nonsignificance is empty.

5. EXAMPLE AND DISCUSSION

We illustrate the two procedures with an example from Walker and Lev [1953, p. 402]. For two groups $n_1 = 8$ and $n_2 = 10$ slow readers, a pretest reading gain score x and a test gain score y were obtained. The first group received play therapy, the second did not. The data required are

Group 1 $\quad n_1 = 8 \quad \bar{x}_1 = 0.03125 \quad , \quad \sum_1^{n_1} (x_i - \bar{x}_1)^2 = 0.4197 \quad ,$

$$\hat{y} = 0.9675 + 1.4401x$$

Group 2 $n_2 = 10$ $\bar{x}_2 = -0.1900$, $\sum_{n_1+1}^{n} (x_i - \bar{x}_2)^2 = 0.6718$,

$$\hat{y} = 0.2063 - 0.0881x$$

$$\hat{\sigma}^2 = 0.1091 \qquad \hat{\sigma} = 0.3303 \quad .$$

The variances and covariance of $\hat{\beta}_2, \hat{\beta}_3$ may be obtained by substituting in $\sigma^2 S^{-1}$, but they are most easily obtained by running the four variable equation through a multiple regression program which outputs the covariance matrix of the regression coefficients. By either method,

$$\hat{\sigma}^2 S^{22} = 0.03065 \quad , \quad \hat{\sigma}^2 S^{23} = 0.02272 \quad , \quad \hat{\sigma}^2 S^{33} = 0.42219 \; .$$

For the Johnson-Neyman technique, the quadratic $Q(x) = 0$ becomes, taking $\lambda = 2F_{2,14}^{.05} = 7.48$,

$$-0.8226x^2 + 1.9866x + 0.3052 = 0$$

with roots -0.165 and 2.580. Since $q_2 < 0$, the region of nonsignificance is $x < -0.165$, $x > 2.580$ (this differs from the result in Walker and Lev, where the t critical value is used instead of F).

For the Gafarian method, we must first decide on the range of x over which the confidence interval is desired. The extreme values of x in the original data are -0.6 and +0.33, while

$$k = \{(0.03125)(0.6718) - (0.19)(0.4197)\}/\{0.4197 + 0.6718\}$$

$$= -0.05382 \quad .$$

We choose an interval symmetric about k which just includes the upper value of x, i.e., we take $h = 0.39$ (this will not cover the lower value of x; h would have to be increased to 0.55 to do this). Then

$$s_2^2 = \frac{1}{8} + \frac{1}{10} + \frac{(0.22125)^2}{1.0915} = 0.26985$$

$$s_2 = 0.51947$$

$$s_3^2 = \frac{1}{0.4197} + \frac{1}{0.6718} = 3.87119$$

$$s_3 = 1.96753$$

$$c = 1.4772 \quad .$$

From Gafarian's table, for $c = 1.5$, $n - 2 = 16$, $1 - \alpha = .95$, $d = 2.975$ by interpolation, while for $c = 1.4$, $n - 2 = 16$, $1 - \alpha = .95$, $d = 3.050$ by interpolation. Linear interpolation gives $d = 2.992$ for $c = 1.4772$. Hence $\delta = ds_2 = 1.554$, and the region of nonsignificance is given by

$$\frac{-.7612 - (1.554)(.3303)}{1.5282} < x < \frac{-.7612 + (1.554)(.3303)}{1.5282}$$

i.e.,

$$-0.834 < x < -0.162 \quad .$$

Note that the lower limit of x lies beyond $k - h$, so the region of nonsignificance is $-0.444 < x < -0.162$. The practical conclusions are identical from the two methods. For $x > -0.162$ (Gafarian) or -0.165 (Johnson-Neyman), the therapy group is superior to the nontherapy group.

The Gafarian method restricts such conclusions to the range (-0.444, 0.336), while the Johnson-Neyman method finds another nonsignificance region for $x > 2.580$. Since this is far beyond the range of the data, it is of no practical interest.

It should be noted that the Gafarian method is very sensitive to h, i.e., to the length of the interval over which the confidence band is to be constructed. If we wished to cover the entire range of x, then $h = 0.55$, whence $c = 1.0475$, $d = 3.454$, $\delta = 1.794$, and the region of nonsignificance is $(-0.886 < x < -0.110)$. Thus h should be chosen just large enough to cover the x-interval of practical interest.

It seems difficult to give a simple rule for the choice of the better procedure. It is not, of course, valid to calculate both intervals and choose the shorter, for the resulting interval will correspond to a confidence coefficient less than the required $100(1 - \alpha)\%$. Gafarian finds the area of the fixed-width confidence band to be less than that of the Working-Hotelling band restricted to the finite interval when $c > 1.5$, but greater when $c < 1.5$, for $\alpha = .05$. It does not immediately follow that the regions of significance will behave in the same way, but in the absence of further information such a rule might be considered. The above example lends some support to such a rule.

REFERENCES

Dunnett, C. W. and Sobel, M. A bivariate generalization of Student's t -distribution, with tables for certain special cases. *Biometrika*, 1954, *41*, 153-169.

Gafarian, A. V. Confidence bands in straight line regression. *J. Amer. statist. Ass.*, 1964, *59*, 182-213.

Johnson, P. O. and Neyman, J. Tests of certain linear hypotheses and their application to some educational problems. *Statist. Res. Mem.*, 1936, *1*, 57-93.

Miller, R. G. *Simultaneous statistical inference*. New York: McGraw-Hill, 1966.

Potthoff, R. F. On the Johnson-Neyman technique and some extensions thereof. *Psychometrika*, 1964, *29*, 241-256.

Press, S. J. *Applied multivariate analysis*. New York: Holt, Rinehart, and Winston, 1972.

Walker, H. M. and Lev, J. *Statistical inference*. New York: Holt, Rinehart, and Winston, 1953.

Working, H. and Hotelling, H. Application of the theory of error to the interpretation of trends. *J. Amer. statist. Ass.*, 1929, *24*, 73-85.