# Confidence Intervals on the Join Point in Segmented Regression

By WALTER W. PIEGORSCH

Biometrics Unit, Cornell University, Ithaca, NY, USA

SUMMARY

A modified maximum likelihood estimator of the join point abscissa in a two-phase, continuous, segmented regression model is developed. Its sampling distribution is considered, including its moments and variance. With the estimator, construction of confidence intervals is undertaken, using an approximate method based on a stable variance estimate and an exact method utilizing the join estimator's distribution function. Comparisons are made to a previously derived method based on Fieller's theorem, using a simulated and an observed data set.

Keywords: TWO-PHASE REGRESSION; PIECEWISE REGRESSION; CONFIDENCE INTERVAL ESTIMATION; JOIN POINT

## 1. INTRODUCTION

IN many experimental situations, the regression relationship under consideration is a two-phase segmented model. When the segments are constrained to intersect at some join point, J, this model is known as a bilinear spline. The problem of estimating J has been considered in great detail (Blischke, 1961; Hinkley, 1971), including extensions to more complicated segmentations (Hudson, 1966; Esterby and El-Shaarawi, 1981). The results have shown that the nonlinear nature of the problem leads to very complicated equations, some of which can only be solved iteratively. To ease these problems some recent works have approached the problem from a decision-theoretic viewpoint (Smith and Cook, 1980; Chin Choy and Brömeling, 1980), although the current trend has been to suggest algorithms and alternatives which help speed the estimation process (Ginsberg et al., 1980; Lerman, 1980; Tishler and Zang, 1981a,b).

One topic of interest which has not seen much exposure is confidence region construction; only a few papers considering it have appeared (Kastenbaum, 1959; Dathe and Müller, 1980). In general, the nonlinear nature of the problem makes these constructions difficult, and care must be taken in developing such. We will present procedures based on a modified maximum likelihood estimator (MLE), and consider comparisons with some previously derived procedures.

## 2. MODIFIED ESTIMATOR

We model the relationship from pairs of observations $(x_i, Y_i)$, where x is an independent, fixed variable, as

$$Y_i = \begin{cases} \alpha_1 + \beta_1 x_i + \epsilon_i & (i = 1, \cdots, \tau) \\ \alpha_2 + \beta_2 x_i + \epsilon_i & (i = \tau+1, \cdots, n) \end{cases} \quad . \tag{2.1}$$

(The $x_i$'s are assumed ordered such that there is some $j < \tau$ and $k > \tau + 1$ with $x_j < x_\tau < x_{\tau+1} < x_k$ .) The random variation in $Y_i$ is described in terms of $\epsilon_i$,

$$\epsilon_i \sim \text{iid } N(0, \sigma^2) \quad . \tag{2.2}$$

For such a model, Blischke (1961) derived the MLE by critically distinguishing the case of τ known. He showed that the estimator is simply the intersection of the two sample regressions, a ratio of the form

$$\hat{J} = (\hat{\alpha}_1 - \hat{\alpha}_2)/(\hat{\beta}_2 - \hat{\beta}_1) \quad , \tag{2.3}$$

where $\hat{\alpha}_1$ and $\hat{\beta}_1$ are the intercept and slope MLE's, respectively, for the first $\tau$ data pairs, and $\hat{\alpha}_2$ and $\hat{\beta}_2$ are those for the next $n - \tau$ data pairs. When $\hat{J}$ falls outside of the interval $[x_\tau, x_{\tau+1}]$, the endpoint which maximizes the likelihood function is taken as the join abscissa. [In the case $\tau$ unknown one can either apply the procedure over all $\tau$ and choose that $\hat{J}(\tau)$ which maximizes the likelihood as $\hat{J}_{MLE}$, or simply estimate $\tau$ first (e.g. as in Ferreira, 1975) and then calculate $\hat{J}$ .] We consider, throughout, the simpler (though not unusually restrictive) case of $\tau$ known.

Because the MLE has no closed-form expression, its sampling distribution is difficult to derive. However, by modifying it slightly, the results become more tractible. Our suggestion is to simply truncate $\hat{J}$ at the endpoints of the known interval and define the new estimator, $\tilde{J}$, by

$$\tilde{J} = \begin{cases} x_\tau & \text{if} \quad \hat{J} < x_\tau \\ \hat{J} & \text{if} \quad x_\tau \le \hat{J} \le x_{\tau+1} \\ x_{\tau+1} & \text{if} \quad \hat{J} > x_{\tau+1} \end{cases} . \tag{2.4}$$

Development of the sampling behavior of $\tilde{J}$ follows quickly from that of $\hat{J}$ . The latter is a ratio of correlated Normal variables, the distribution function of which has been developed by D.V. Hinkley (1969). In this setting it becomes

$$\begin{aligned} F_{\hat{J}}(t) = P[\hat{J} \le t] &= L\left(\frac{\alpha_1-\alpha_2-(\beta_2-\beta_1)t}{v_1 v_2 a}, \frac{\beta_1-\beta_2}{v_2}; \frac{v_2 t - v_1\rho}{v_1 v_2 a}\right) \\ &\quad + L\left(\frac{(\beta_2-\beta_1)t-\alpha_1+\alpha_2}{v_1 v_2 a}, \frac{\beta_2-\beta_1}{v_2}; \frac{v_2 t - v_1\rho}{v_1 v_2 a}\right) , \end{aligned} \tag{2.5}$$

where

$$v_1^2 = \mathrm{var}(\hat{\alpha}_1 - \hat{\alpha}_2) = \sigma^2\left[\sum_{i=1}^{\tau} x_i^2/(\tau c_1) + \sum_{i=\tau+1}^{n} x_i^2/\{(n-\tau)c_2\}\right] , \tag{2.6}$$

$$v_2^2 = \mathrm{var}(\hat{\beta}_2 - \hat{\beta}_1) = \sigma^2\left(c_1^{-1} + c_2^{-1}\right) , \tag{2.7}$$

$$\rho = \frac{\text{cov}(\hat{\alpha}_1-\hat{\alpha}_2, \hat{\beta}_2-\hat{\beta}_1)}{v_1 v_2} \tag{2.8}$$

$$= \left(\frac{\bar{x}_1}{c_1} - \frac{\bar{x}_2}{c_2}\right) / \left(\left\{c_1^{-1} + c_2^{-1}\right\}\left\{\sum_{i=1}^{\tau} x_i^2/\tau c_1 + \sum_{i=\tau+1}^{n} x_i^2/(n-\tau)c_2\right\}\right)^{\frac{1}{2}} ,$$

$$\bar{x}_1 = \sum_{i=1}^{\tau} x_i/\tau , \qquad \bar{x}_2 = \sum_{i=\tau+1}^{n} x_i/(n-\tau) ,$$

$$c_1 = \sum_{i=1}^{\tau} (x_i - \bar{x}_1)^2 , \qquad c_2 = \sum_{i=\tau+1}^{n} (x_i - \bar{x}_2)^2 ,$$

$$a^2 \equiv a^2(t) = \frac{t^2}{v_1^2} - \frac{2\rho t}{v_1 v_2} + v_2^{-2} \tag{2.9}$$

and L is the standard bivariate Normal integral, defined as

$$L(h,k;g) = \frac{1}{2\pi\sqrt{(1-g^2)}} \int_h^{\infty}\int_k^{\infty} \exp\left\{-\frac{x^2-2gxy+y^2}{2(1-g^2)}\right\} dxdy . \tag{2.10}$$

From (2.4) and (2.5), one can show that

$$F_{\tilde{J}}(t) = P[\tilde{J} \le t] = \begin{cases} 0 & \text{if} \quad t < x_\tau \\ F_{\hat{J}}(x_\tau) & \text{if} \quad t = x_\tau \\ F_{\hat{J}}(t) & \text{if} \quad x_\tau < t < x_{\tau+1} \\ 1 & \text{if} \quad t \ge x_{\tau+1} \end{cases} . \tag{2.11}$$

Therefore, $F_{\tilde{J}}$ depends upon L in (2.10), and since values of L are available in tabular form (National Bureau of Standards, 1959), calculation of $F_{\tilde{J}}(t)$ is relatively simple.

The moments of $\tilde{J}$ are also as attainable. Integration by parts can be employed to show

$$E(\tilde{J}) = x_{\tau+1} - \int_{x_\tau}^{x_{\tau=1}} F_{\hat{J}}(t)dt \quad , \tag{2.12}$$

and

$$\mathrm{var}(\tilde{J}) = 2\int_{x_\tau}^{x_{\tau+1}} (x_{\tau+1} - t)F_{\hat{J}}(t)dt - \left(\int_{x_\tau}^{x_{\tau+1}} F_{\hat{J}}(t)dt\right)^2 \tag{2.13}$$

Again, since $F_{\hat{J}}$ depends upon L, values of the integrands in (2.12) and (2.13) can be calculated, and the expressions then evaluated using a numerical quadrature procedure. Simulations were carried out to examine the sampling behavior of $\tilde{J}$, and they suggested that this modified estimator performed adequately. Values of $E(\tilde{J})$ were commonly close to the true value of J, and the values of $\mathrm{var}(\tilde{J})$ were particularly small (Piegorsch, 1982). Specific exemplifications of the use of these equations will be presented in Section 4.

## 3. CONFIDENCE INTERVALS ON J

With an available, closed-form distribution function in (2.11), we can consider construction of confidence intervals on J . Mood, Graybill, and Boes (1974, Ch. 8) present a procedure for confidence interval construction which they call 'The Statistical Method'. It involves calculating two functions of an unknown parameter, $\theta$, from the distribution function of some estimator of $\theta$ — call it T — then solving for the confidence limits by using the observed value of the estimator. Specifically, if the estimator produces an estimate of $t_0$ then the lower limit is found by solving for $\theta$ in

$$p_1 = \int_{-\infty}^{t_o} f_T(t|\theta)dt \quad , \tag{3.1}$$

and the upper limit is found by solving for $\theta$ in

$$p_2 = \int_{t_o}^{\infty} f_T(t|\theta)dt \quad . \tag{3.2}$$

The resulting interval will have a confidence coefficient of $1-p_1-p_2$ (note that $p_i>0$, $i=1,2$, and $p_1+p_2<1$) .

In our setting this will be rather tricky. We do not have a simple one-parameter distribution function. Instead, a pair of parameters, $\mu_1=\alpha_1-\alpha_2$ and $\mu_2=\beta_2-\beta_1$, which contribute to the parameter of interest, J, are involved. To get around this we will make probability statements about each parameter, then use the Bonferroni inequality (actually just a simplified version: $P[A\cap B]\geq P[A]+P[B]-1$) to simultaneously combine the statements into one interval (as exemplified in Lieberman et al., 1967). We will use the fact that $\hat{\mu}_2\sim N(\mu_2,v_2^2)$, then combine confidence statements based on this with statements on $\mu_1$ derived using The Statistical Method . In each case we will need statements with probability $1-(\alpha/2)$ so as to combine them into one statement with confidence coefficient $1-\alpha$ . The result will be limits, $c_\ell$ and $c_u$, such that $P[c_\ell<J<c_u]\geq 1-\alpha$ . We will need to break things up into three cases: $x_\tau<\tilde{J}<x_{\tau+1}$, $\tilde{J}=x_\tau$, and $\tilde{J}=x_{\tau+1}$ .

CASE I: $x_\tau<\tilde{J}<x_{\tau+1}$ . Since $\hat{\mu}_2\sim N(\mu_2,v_2^2)$, it is well known that

$$P\left(\hat{\mu}_2-v_2\Phi^{-1}\left\{1-\frac{\alpha}{4}\right\}<\mu_2<\hat{\mu}_2+v_2\Phi^{-1}\left\{1-\frac{\alpha}{4}\right\}\right)=1-\frac{\alpha}{2}\quad, \tag{3.3}$$

so

$$P\left(\frac{1}{\mu_2}<u_o\cap\frac{1}{\mu_2}>\ell_o\right)=1-\frac{\alpha}{2}\quad, \tag{3.4}$$

where, for notation's sake,

$$u_o=\left(\hat{\mu}_2-v_2\Phi^{-1}\left\{1-\frac{\alpha}{4}\right\}\right)^{-1}\qquad\text{and}\qquad \ell_o=\left(\hat{\mu}_2+v_2\Phi\left\{1-\frac{\alpha}{4}\right\}\right)^{-1}\quad. \tag{3.5}$$

The Statistical Method suggests that we can solve for $\mu_1$ in

$$\text{(a)}\ \frac{\alpha}{4}=F_{\tilde{J}}(\tilde{J})\qquad\text{and}\qquad\text{(b)}\ 1-\frac{\alpha}{4}=F_{\tilde{J}}(\tilde{J}) \tag{3.6}$$

and then take

$$P\{(\mu_1^b/\mu_2) < J < (\mu_1^a/\mu_2)\} \geq 1 - \frac{\alpha}{2} \quad . \tag{3.7}$$

The simultaneous combination of (3.4) and (3.7) gives us a (minimum) $1-\alpha$ interval on J, but it must be performed carefully:

(i) If $\mu_1^a \geq 0$ then $[J < \mu_1^a/\mu_2 \cap 1/\mu_2 < u_0] \Rightarrow J < \mu_1^a u_0$ .
Now, if $\mu_1^b \geq 0$ then $[J > \mu_1^b/\mu_2 \cap 1/\mu_2 > \ell_0] \Rightarrow J > \mu_1^b \ell_0$ . But, if $\mu_1^b < 0$ then $[J > \mu_1^b/\mu_2 \cap 1/\mu_2 < u_0] \Rightarrow J > \mu_1^b u_0$ .

(ii) If $\mu_1^a < 0$ then $[J < \mu_1^a/\mu_2 \cap 1/\mu_2 > \ell_0] \Rightarrow J < \mu_1^a \ell_0$ .
Also, since $\mu_1^a > \mu_1^b$ by construction, $\mu_1^b < 0$ so $[J > \mu_1^b/\mu_2 \cap 1/\mu_2 < u_0]$ $\Rightarrow J > \mu_1^b u_0$ .

CASE II: $\tilde{J} = x_\tau$ . Obviously here $c_\ell = x_\tau$, and we only need consider derivation of $c_u$, i.e., our only interest is in the value of $\mu_1^a$ combined simultaneously with a bound on $\mu_2$ . The former value satisfies $F_{\hat{J}}(x_\tau) = 1 - (\alpha/4)$ while the latter will follow from Normal distribution theory. Again, the simultaneous combination can be tricky:

(i) $\mu_1^a \geq 0$ . Since $P\{\mu_2 > \hat{\mu}_2 - v_2\Phi^{-1}[1-(\alpha/2)]\} = 1-(\alpha/2)$, combination with $P[x_\tau < J < \mu_1^a/\mu_2] = 1-(\alpha/2)$ produces $c_u = \mu_1^a/\{\hat{\mu}_2 - v_2\Phi^{-1}[1-(\alpha/2)]\}$ .

(ii) $\mu_1^a < 0$ . Similar to the above, $P[\mu_2 < \hat{\mu}_2 - v_2\Phi^{-1}(\alpha/2)] = 1-(\alpha/2)$, so we get $c_u = \mu_1^a/[\hat{\mu}_2 - v_2\Phi^{-1}(\alpha/2)]$ . Note that $\Phi^{-1}(\alpha/2) = -\Phi^{-1}[1-(\alpha/2)]$ .

CASE III: $\tilde{J} = x_{\tau+1}$ . In this case $c_u = x_{\tau+1}$ and our interest is in values of $\mu_1^b$ such that $F_{\hat{J}}(x_{\tau+1}) = 1-(\alpha/4)$ in simultaneity with a lower or upper bound on $\mu_2$ . The results are similar to those in Case II:

(i) $\mu_1^b < 0$ gives $c_\ell = \mu_1^b/\{\hat{\mu}_2 - v_2\Phi^{-1}[1-(\alpha/2)]\}$ .

(ii) $\mu_1^b \geq 0$ gives $c_\ell = \mu_1^b/[\hat{\mu}_2 - v_2\Phi^{-1}(\alpha/2)]$ .

All of these results are summarized in Table 1. Note that we have taken $\sigma^2$ as known. If this were not the case, and there were no previously establish estimate to use instead, the simultaneous combination could be extended to the 3 parameters, $\mu_1$, $\mu_2$, and $\sigma^2$ .

—SHOW TABLE 1 HERE—

The number of calculations involved here is formidable, and one would probably need to turn to the computer. A question that naturally arises is whether or not a computationally simpler procedure could be developed, at the expense of some fixed level of confidence. Perhaps some approximate procedure, with less computational involvement, could be formulated? We considered intervals of the form $(c'_\ell, c'_u)$, where

$$c'_\ell = \max[x_\tau, \tilde{J} - k\sqrt{\hat{var}(\tilde{J})}] \quad (3.8)$$

and

$$c'_u = \min[\tilde{J} + k\sqrt{\hat{var}(\tilde{J})}, x_{\tau+1}] \quad , \quad (3.9)$$

and found that $var(\tilde{J})$ was very neatly estimated by simply replacing the parameters with their ML estimates in the expression for $F_{\hat{J}}(t)$ in (2.5). This produced an estimate, $\hat{F}_{\hat{J}}$, which could then replace its theoretical counterpart in (2.13), i.e. simply use

$$\hat{var}(\tilde{J}) = 2x_{\tau+1} \int_{x_\tau}^{x_{\tau+1}} \hat{F}_{\hat{J}}(t)dt - 2\int_{x_\tau}^{x_{\tau+1}} t\hat{F}_{\hat{J}}(t)dt - \left(\int_{x_\tau}^{x_{\tau+1}} \hat{F}_{\hat{J}}(t)dt\right)^2 \quad . \quad (3.10)$$

In considering a value for k, we found that $k \doteq 2$ proved empirically stable. Average interval lengths were not so large as to effectively duplicate the known information — i.e., an interval with $c'_\ell = x_\tau$ and $c'_u = x_{\tau+1}$ imparts little additional knowledge — while the empirical confidence coefficients worked out to about 0.9 (Piegorsch, 1982). The amount of necessary, on-line computing was also reduced. However, we reiterate the fact that these intervals are of an ad hoc nature and can only approximately provide the experimenter with a pre-set confidence level.

## 4. EXAMPLES

We exemplify these procedures using both a simulated and an observed set of data. Comparisons are made to a procedure suggested by M.A. Kastenbaum (1959), which involves applying Fieller's Theorem to the ratio in (2.3). This method requires solving a quadratic equation in J to calculate the lower and upper limits. The unfortunate possibility exists, therefore, of finding two complex roots, at which points the lower and upper limits should strictly be set to $-\infty$ and $\infty$, respectively. However, if such an occurrence were to happen, in this setting, we would simply set the lower limit to $x_\tau$, and the upper limit to $x_{\tau+1}$. We will, for comparison's sake, use a confidence level of $\gamma = .90$.

### Simulated Data

Data were artificially produced from the model

$$Y_i = \begin{cases} -7.9 + 4.5x_i + \epsilon_i & i = 1,\cdots,7 \\ 8.9 - 7.5x_i + \epsilon_i & i = 8,\cdots,17 \end{cases}, \qquad (4.1)$$

and $\epsilon_i \sim \text{iid } N(0,4)$. The data set is presented in Table 2. Note that the true join occurs at $x = 1.4$.

—SHOW TABLE 2 HERE—

To find $E(\tilde{J})$ and $\text{var}(\tilde{J})$ we need to evaluate the integrals in (2.12) and (2.13). As mentioned earlier, the availability of specified values for F(t), and therefore tF(t), makes numerical quadrature relatively easy. Applying the data values in Table 2 to (2.6), (2.7), and (2.8) yields $v_1^2 = 4.1659$, $v_2^2 = 1.6542$, and $\rho = 0.3989$ . Using these in (2.5) shows

$$F(t) = L\left(\frac{4.5712t-6.3997}{a(t)}, 9.3301; \frac{0.4899t-0.3101}{a(t)}\right) + L\left(\frac{6.3997-4.5712t}{a(t)}, -9.3301; \frac{0.4899t-0.3101}{a(t)}\right) , \qquad (4.2)$$

with $a(t) = \sqrt{(0.24t^2 - 0.3039t + 0.6045)}$, from (2.9).

For a series of values of $t \in [1.2,1.6]$, $F(t)$, and $tF(t)$ were calculated. The quadrature procedure known as Simpson's Rule (Forsythe et al., 1977) was applied over a (constant) sub-interval length of $h_j = t_{j+1} - t_j = 0.00025$ . This produced approximate values with errors so small (on the order of $h^4$, or about $4 \times 10^{-15}$, per sub-interval; this is a total error on the order of $10^{-12}$) that the results will be considered exact. Specifically, it was found that

$$\int_{1.2}^{1.6} F(t)dt = 0.19757 \quad \text{and} \quad \int_{1.2}^{1.6} tF(t)dt = 0.28523 .$$

The resulting expectations were thus $E(\tilde{J}) = 1.40243$ and $\text{var}(\tilde{J}) = 0.02273$ . As a comparison, the asymptotic variance of the MLE (Hinkley, 1971) was calculated from the formula

$$\text{asymp. var}(\hat{J}_{MLE}) = \frac{\sigma^2}{\mu_2^2}\left\{\frac{1}{\tau} + \frac{1}{n-\tau} + \frac{(\bar{x}_1-J)^2}{c_1} + \frac{(\bar{x}_i-J)^2}{c_2}\right\} . \qquad (4.3)$$

This produced a value of 0.03109, almost 150% larger than $\text{var}(\tilde{J})$ .

The estimation process yielded a value of $\tilde{J} = 1.2403$ . Simpson's rule was again applied to estimate the variance using (3.10), and the resulting value was $\hat{\text{var}}(\tilde{J}) = 0.02111$ .

The resulting 90% confidence intervals from The Statistical Method, and from Kastenbaum's Fieller-based procedure were both found to reiterate the known information, i.e. (1.2,1.6) . The ad hoc procedure did slightly better, giving an interval of (1.2,1.5309) .

Liver Secretion Example

Chicken livers secrete lipids and proteins as a response to dietary fluctuations. For example, after imposing a change in cholesterol intake, triglyceride production over time follows an increasing trend which levels off sharply after some time period, $t = J$ (Behr, 1982). Model (2.1) is assumed to estimate this switchover time using the data in Table 3. From previous experience it is suggested that the variance about the regression model here is well-approximated by the value $\sigma^2 = 2.5$ . The join is expected to occur between four and six hours so that $\tau = 3$ .

—SHOW TABLE 3 HERE—

The estimation procedures in Sections 2 and 3 produce the estimates $\tilde{J} = 4.7387$ and $\hat{\text{var}}(\tilde{J}) = 0.1734$ . Confidence intervals can be calculated using the various methods of Section 3. They produce the following results, again at $\gamma = 0.9$:

| | | |
|---|---|---|
| Statistical method | (4.0 ,4.807 ) , |
| Kastenbaum method | (4.1008,5.4996) ; |
| *ad hoc* procedure | (4.0 ,5.5715) . |

The liver secretion results are quite pleasing, but with the simulated data we can see that the intervals produced are not adding any sufficiently new information to the experimental situation. One of the major reasons for this is that the known interval lengths are small relative to the measures of variation (e.g., $v_1^2$ or $\text{var}(\tilde{J})$) involved. Since these variances are critically dependent upon the choices of $x_\tau$, $x_{\tau+1}$, and the other design points, some element of care must be taken when selecting the values of the $x_i$'s . Unfortunately, only a few papers have appeared which consider experiment designs for segmented models. Extensions of works such as Agarwal and Studden (1978), or Park (1978) would most certainly add a great deal to the development of an overall strategy for the statistical analysis, including interval estimation, of segmented regressions.

## 5. EXTENSIONS TO UNKNOWN TAU

One critical distinction we have made is to intersect the derived confidence interval with the interval of known information, $[x_\tau, x_{\tau+1}]$. When $\tau$ is unknown, the situation becomes more complicated. Of the methods considered, the Kastenbaum procedure is still available, although the problem of infinite endpoints is still of concern.

As an alternative, Hinkley (1971) has suggested approximate approaches, based on likelihood ratios and the asymptotic Normality of MLE's. His intervals' small-sample behaviors seem empirically acceptable under the constraint $\beta_2 = 0$, but when the constraint is removed, increases in variability are observed. Further, the computations involved grow as numerous as those in the statistical method.

What is needed is perhaps a conditional approach, first estimating $\tau$ with some associated level of confidence, and then applying the statistical method (or some other procedure). The opportunities for such improvement in join point estimation seem endless and although burdened by very complicated models and equations, the development of improved approaches in the estimation, testing, and design of segmented models is certainly within reach.

## ACKNOWLEDGEMENTS

This research is based on the author's M.S. thesis at Cornell University, Ithaca, NY, and was supported in part by a grant from Sigma Xi, The Scientific Research Society. The author wishes to thank Dr. George Casella for his many suggestions and gracious assistance, and Dr. D.S. Robson for a number of helpful comments.

REFERENCES

AGARWAL, G. F. and STUDDEN, W. J. (1978). Asymptotic design and estimation using linear splines. Commun. Statist., B7, 309-319.

BEHR, S. R. (1982). Personal communication.

BLISCHKE, W. R. (1961). Least squares estimators of two intersecting lines. Paper No. BU-135-M in the Biometrics Unit Mimeo Series, Cornell University, Ithaca, NY.

CHIN CHOY, J. H. and BROEMELING, L. (1980). Some Bayesian inferences for a changing linear model. Technometrics, 22, 71-78.

DATHE, E. M. and MÜLLER, P. (1980). A contribution to spline-regression. Biometrical J., 22, 259-270.

ESTERBY, S. R. and EL-SHAARAWI, A. H. (1981). Inference about the point of change in a regression model. Appl. Statist., 30, 277-285.

FERREIRRA, P. E. (1975). A Bayesian analysis of a switching regression model: Known number of regimes. J. Amer. Statist. Assoc., 70, 370-374.

FORSYTHE, G. E., MALCOLM, M. A., and MOLER, C. B. (1977). Computer Methods for Mathematical Computations. Englewood Cliffs, NJ: Prentice-Hall.

GINSBERG, V., TISHLER, A., and ZANG, I. (1980). Alternative estimation methods for two-regime models. Europ. Econ. Rev., 13, 207-228.

HINKLEY, D. V. (1969). On the ratio of two correlated normal random variables. Biometrika, 56, 635-639.

HINKLEY, D. V. (1971). Inference in two-phase regression. J. Amer. Statist. Assoc., 66, 736-743.

HUDSON, D. J. (1966). Fitting segmented curves whose join points have to be estimated. J. Amer. Statist. Assoc., 61, 1097-1129.

KASTENBAUM, M. A. (1959). A confidence interval on the point of intersection of two fitted linear regressions. Biometrics, 15, 323-324.

LERMAN, P. M. (1980). Fitting segmented regression models by grid search. Appl. Statist., **29**, 77-84.

LIEBERMAN, G. J., MILLER, R. G., and HAMILTON, M. A. (1967). Unlimited simultaneous discrimination intervals in regression. Biometrika, **54**, 133-145.

MOOD, A. M., GRAYBILL, F. A., and BOES, D. C. (1974). Introduction to the Theory of Statistics. New York: McGraw-Hill.

NATIONAL BUREAU OF STANDARDS (1959). Tables of the Bivariate Normal Distribution Function and Related Functions. Washington, D.C.: U.S. Government Printing Office.

PARK, S. H. (1978). Experiment designs for fitting segmented polynomial regression models. Technometrics, **20**, 151-154.

PIEGORSCH, W. W. (1982). A modification of the least squares join point estimator in bilinear segmented regression. M.S. Thesis, Cornell University, Ithaca, NY.

SMITH, A. F. M. and COOK, D. G. (1980). Straight lines with a change point: A Bayesian analysis of some renal transplant data. Appl. Statist., **29**, 180-189.

TISHLER, A. and ZANG, I. (1981a). A maximum likelihood method for piecewise regression models with a continuous dependent variable. Appl. Statist., **30**, 116-124.

TISHLER, A. and ZANG, I. (1981b). A new maximum likelihood algorithm for piecewise regression. J. Amer. Statist. Assoc., **76**, 980-987.

TABLE 1

Statistical method confidence limits; $P[c_\ell < J < c_u] \geq 1 - \alpha$

| Case | $\mu_1^a$ satisfies | $\mu_1^b$ satisfies | $\mu_1^a < 0$ | $\mu_1^a \geq 0$ | $\mu_1^b < 0$ | $\mu_1^b \geq 0$ | $c_\ell$ | $c_u$ |
|---|---|---|---|---|---|---|---|---|
| I | $\frac{\alpha}{4} = F_{\tilde{J}}(\tilde{J})$ | $1 - \frac{\alpha}{4} = F_{\tilde{J}}(\tilde{J})$ | x | | x | | $\mu_1^b/[\hat{\mu}_2 - v_2\Phi^{-1}(1-\frac{\alpha}{4})]$ | $\mu_1^a/[\hat{\mu}_2 + v_2\Phi^{-1}(1-\frac{\alpha}{4})]$ |
| I | $\frac{\alpha}{4} = F_{\tilde{J}}(\tilde{J})$ | $1 - \frac{\alpha}{4} = F_{\tilde{J}}(\tilde{J})$ | | x | x | | $\mu_1^b/[\hat{\mu}_2 - v_2\Phi^{-1}(1-\frac{\alpha}{4})]$ | $\mu_1^a/[\hat{\mu}_2 - v_2\Phi^{-1}(1-\frac{\alpha}{4})]$ |
| I | $\frac{\alpha}{4} = F_{\tilde{J}}(\tilde{J})$ | $1 - \frac{\alpha}{4} = F_{\tilde{J}}(\tilde{J})$ | | x | | x | $\mu_1^b/[\hat{\mu}_2 + v_2\Phi^{-1}(1-\frac{\alpha}{4})]$ | $\mu_1^a/[\hat{\mu}_2 - v_2\Phi^{-1}(1-\frac{\alpha}{4})]$ |
| II | $F_{\hat{J}}(x_\tau) = \frac{\alpha}{4}$ | – | x | | | | $x_\tau$ | $\mu_1^a/[\hat{\mu}_2 - v_2\Phi^{-1}(\frac{\alpha}{2})]$ |
| II | $F_{\hat{J}}(x_\tau) = \frac{\alpha}{4}$ | – | | x | | | $x_\tau$ | $\mu_1^a/[\hat{\mu}_2 - v_2\Phi^{-1}(1-\frac{\alpha}{2})]$ |
| III | – | $1 - \frac{\alpha}{4} - F_{\hat{J}}(x_{\tau+1})$ | | | x | | $\mu_1^b/[\hat{\mu}_2 - v_2\Phi^{-1}(1-\frac{\alpha}{2})]$ | $x_{\tau+1}$ |
| III | – | $1 - \frac{\alpha}{4} = F_{\hat{J}}(x_{\tau+1})$ | | | | x | $\mu_1^b/[\hat{\mu}_2 - v_2\Phi^{-1}(\frac{\alpha}{2})]$ | $x_{\tau+1}$ |

TABLE 2

Simulated data set

| $x_i$ | $Y_i$ |
|---|---|
| -.5 | -10.1026 |
| -.4 | - 8.5996 |
| -.2 | - 9.7397 |
| .4 | - 2.6177 |
| .7 | - 3.6366 |
| .8 | - 3.462 |
| 1.2 | - 7.5573 |
| 1.6 | - 4.7696 |
| 3.2 | -10.5459 |
| 3.7 | -21.3723 |
| 3.8 | -22.1392 |
| 4.7 | -27.9223 |
| 6.1 | -38.5426 |
| 6.2 | -42.2086 |
| 6.9 | -43.7475 |
| 7.0 | -42.9145 |
| 7.2 | -40.8345 |

TABLE 3

Liver secretion data

| Hours $x_i$ | Triglyceride level $Y_i$ |
|---|---|
| 0 | 22.825 |
| 1 | 29.625 |
| 2 | 39.3 |
| 3 | 43.8 |
| 4 | 51.7 |
| 6 | 55.425 |
| 7 | 57.9 |
| 8 | 59.1 |
| 9 | 58.8 |
| 10 | 60.85 |
| 11 | 61.025 |
| 12 | 59.9625 |
| 13 | 60.0625 |
| 14 | 58.6 |
| 15 | 61.425 |
| 16 | 60.6 |