Conference on Applied Statistics in Agriculture 1991 - 3rd Annual Conference Proceedings

STRAIGHT LINE REGRESSION WHEN BOTH VARIABLES ARE SUBJECT TO ERROR

Norman R. Draper

STRAIGHT LINE REGRESSION WHEN BOTH VARIABLES ARE SUBJECT TO ERROR

Norman R. Draper
University of Wisconsin
Department of Statistics
1210 West Dayton Street
Madison, WI 53706

ABSTRACT

This expository note discusses the problem of fitting a straight line when both variables are subject to error. A brief review of the literature is undertaken, and one fitting method, the geometric mean functional realationship, is spotlighted and illustrated with two sets of example data. The emphasis is on providing practical advice. All methods have drawbacks, but the geometric mean functional relationship method appears to provide a sensible course of action in many practical problems, and could benefit from further investigation.

1. INTRODUCTION

Whenever we fit the model

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i, \quad i = 1, 2, \ldots, n \tag{1}$$

by least squares to a set of n data values (X_i, Y_i), we usually take it for granted that Y is subject to the error ϵ_i and X is not subject to error. If this is true, and if the vector of errors $\underset{\sim}{\epsilon} = (\epsilon_1, \epsilon_2, \ldots, \epsilon_n)'$ is distributed $N(0, I\sigma^2)$, maximum likelihood estimation and least squares estimation, namely

$$\underset{\beta_0, \beta_1}{\text{Minimize}} \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 X_i)^2$$

provide the same estimates (b_0, b_1) of (β_0, β_1).

What if both X and Y are subject to error? We can write

$$Y_i = \eta_i + \epsilon_i \tag{2}$$

$$X_i = \xi_i + \delta_i. \tag{3}$$

We assume that a straight line relationship

$$\eta_i = \beta_0 + \beta_1 \xi_i \tag{4}$$

holds between the true but unobserved values η_i and the n unknown parameters ξ_i. Substituting (4) into (2) and then substituting for ξ_i from (3) gives

$$Y_i = \beta_0 + \beta_1 X_i + (\epsilon_i - \beta_1 \delta_i). \tag{5}$$

Let us assume that $\epsilon_i \sim N(0, \sigma^2)$, with the ϵ_i uncorrelated, and $\delta_i \sim N(0, \sigma_\delta^2)$, with the δ_i uncorrelated, with ϵ_i and δ_i uncorrelated, and define

$$\sigma_\xi^2 = \sum_{i=1}^{n} (\xi_i - \bar{\xi})^2 / n, \tag{6}$$

$$\sigma_{\xi\delta} = \text{Covariance}\,(\xi, \delta), \tag{7}$$

$$\rho = \sigma_{\xi\delta} / (\sigma_\xi \sigma_\delta), \tag{8}$$

$$\tau = \sigma_\delta / \sigma_\xi. \tag{9}$$

In (7), $\sigma_{\xi\delta}$ would typically be zero; however, see case (2) below. If, mistakenly, we fit (1) by least squares, b_1 will be biased. In fact

$$E(b_1) = \beta_1 - \frac{\beta_1 \tau (\rho + \tau)}{1 + 2\rho\tau + \tau^2}. \tag{10}$$

The bias is negative if $\sigma_\xi^2 + \sigma_{\xi\delta} > 0$, this is, if $\rho + \tau > 0$. The bias arises from the fact that X_i is not independent of the error in (5), in general. In fact

$$\text{Covariance}\,[X_i, (\epsilon_i - \beta_1 \delta_i)] = -\beta_1 (\rho + \tau) \sigma_\xi \sigma_\delta. \tag{11}$$

We thus see that there are cases where fitting (1) by least squares will provide little or no bias. These are

1. If σ_δ^2 is small compared with σ_ξ^2, the errors in the X's are small compared with the spread in the ξ_i's (and so in the X's) and τ will be small. The bias in (10) is then small. This is what is often assumed in practice, when least squares is used.

2. If the X's are fixed and determined by the experimenter (see Berkson, 1950), then $\sigma_{\xi\delta} = \text{Covariance}(X_i - \delta, \delta) = -\sigma_\delta^2$, which means that $\sigma_{\xi\delta} + \sigma_\delta^2 = 0$, or $\rho + \tau = 0$, implying zero bias in (10).

3. We wish to fit $Y_i = \eta_i + \epsilon_i$ where $\eta_i = \beta_0 + \beta_1 X_i$ (the observed X_i, note) and not as in (4).

These formats will not fit all practical cases. One case that occurred at the University of Wisconsin in connection with a study on wild birds, required the observation of X_i = "the distance the bird was from a path". The student pointed out that, as she approached a bird, it flew away before she got close enough to see precisely where it had perched. Thus error in recording X was unavoidable.

In Section 2, we summarize some of the published work on this topic. In Section 3, we highlight the geometric mean functional relationship. The latter is applied to two data sets in Section 4.

2. SELECTED PRIOR WORK

If we attempt to obtain maximum likelihood estimates of β_0 and β_1 under the distributional assumptions made in connection with (5), we find that there is an identifiability problem. The estimation cannot be carried through without some additional information being added, for example, knowledge of the ratio $\lambda = \sigma^2/\sigma_\delta^2$. (Barnett, 1967; Wong, 1989). This is Case III of Sprent and Dolby (1980), discussed below. Various authors have suggested alternative analyses.

Geary (1942) proposed a method dependent on fourth order mixed cumulants of X and Y. However, the two estimates obtained sometimes lie outside the "regression limits" defined by the two least squares lines of Y against X and X against Y.

Sprent and Dolby (1980) distinguish four cases:

I. (X, Y) are bivariate normal variables and $E(Y|X) = \beta_0 + \beta_1 X$.

II. $Y \sim N(\beta_0 + \beta_1 X, \sigma^2)$. The observed X-values are fixed on realizations of a random variable with any (reasonable) distribution.

In both I and II, estimates via maximum likelihood are the usual least squares estimates $b_1 = S_{XY}/S_{XX}$ and $b_0 = \overline{Y} - b_1\overline{X}$, where $S_{XY} = \sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})$ and $S_{XX} = \sum_{i=1}^{n}(X_i - \overline{X})^2$.

III. The case of Section 1. If λ were known, maximum likelihood leads to estimates

$$\hat{\beta}_1 = [S_{YY} - \lambda S_{XX} + \{(S_{YY} - \lambda S_{XX})^2 + 4\lambda S_{XY}^2\}^{1/2}]/(2S_{XY})$$

$$\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1\overline{X}, \tag{12}$$

where $S_{YY} = \sum_{i=1}^{n}(Y_i - \overline{Y})^2$. Note that, if $\lambda = S_{YY}/S_{XX}$, $\hat{\beta}_1 = (S_{YY}/S_{XX})^{1/2}$ which is the geometric mean functional relationship, after attachment of the sign of S_{XY}. This is often called the functional relationship model. Note also that, when $\lambda = 1$, the solution (12) defines the line which minimizes the sum of squares of perpendicular deviations from the line. Many people find such a solution intuitively satisfying, but it is appropriate only when $\sigma^2 = \sigma_\delta^2$, that is, when $\lambda = 1$. This solution was first given by Adcock (1878).

IV. Similar to III but with ξ_i a normal random variable, independent of δ, so that because of (4), (ξ_i, η_i) follow a joint degenerate bivariate normal distribution. This is the so-called structural relationship model and again, the case III solution applies if λ is known.

Sprent and Dolby "do not recommend ad hoc use of the geometric mean functional relationship when there are errors in both variables," arguing that other ad hoc estimates could equally be used. The paper by Barker, Soh and Evans (1988) provides an excellent justification for the geometric mean functional relationship, however. These authors show that the estimator minimizes the sum of the triangular areas formed by drawing horizontal and vertical lines to the fitted lines from the observed points; see Figure 1. (This had previously been pointed out by Teissier (1948), but his paper was accessible only to those who read French, and so was not widely known. A clear restatement and diagram are given by Harvey and Mace (1982, p. 349). The geometric mean functional relationship has been condemned as being inconsistent, that is, the estimates do not tend to their true values as n tends to infinity. However, other estimators are biased, and what happens for large n is often not of concern to those with practical problems and small data sets.

Patefield (1981) looks at the multi-X case and extends the following single -X results: When λ is specified, and both X and Y distributions are normal, the maximum likelihood estimate of β_1 takes the same form for both structural and functional relationships and is bounded by the slopes from the two (Y on X) and (X on Y) regressions. (In the latter case, we transpose the fitted equation to a Y on X form to get the bound.) Some asymptotic comparisons are also made.

Reilly and Patino-Leal (1981) provide general methods for producing the posterior probability density function for the parameters. The error covariance matrix is assumed to be known. A virtue of the development is that both linear and nonlinear models can be handled using this technique.

Brown (1982) assumes λ known, discusses deficiencies in the maximum likelihood estimator, and offers a robust alternative.

Chan (1982) offers a method of estimating β_1 when the ξ_i arise from a uniform distribution over a specified range. He seeks to find consistent estimators of the parameters by using a local maximum, rather than a global maximum, of the likelihood function that results. He concludes via simulations that his new method is better for larger n, and that both his method and the geometric mean functional relationship have "too large mean squared errors to be of practical use" under the uniform distribution assumption.

Wolter and Fuller (1982) provide formulas for estimating a quadratic model in one X. They provide (normal) asymptotic distribution results for the estimates and perform some "small-sample" ($n = 33$ and 66) simulation results.

Ketellaper (1983) concludes that a "corrected least squares estimator", $b_{CLS} = S_{XY}/(S_{XX} - \sigma_\delta^2)$, suggested by Madansky (1959), is better, at least for $n \geq 20$, than S_{XY}/S_{XX}, the usual least squares estimator.

Mandel (1984) gives a series of steps to get a straight line fit for the (X, Y)

situation. He also suggest a way of checking if the ordinary least squares Y on X solution is acceptable by evaluating a particular number (p. 10). The evaluation requires knowledge of λ and $\sigma_{\xi\delta}$.

Lakshminarayanan and Gunst (1984) examine maximum likelihood estimation when λ is known. They conclude that "effective use of asymptotic properties of the ... estimator ... requires a large sample size and accurate selection of ... λ."

Schnute (1984) proposed several estimation criteria based on minimization of various functions of sample moments of the data.

Stefanski (1985) uses an M-estimator for parameter estimation in a very general errors-in-variables formulation, assesses asymptotic bias and discusses the construction of an estimator with smaller bias.

Gleser and Hwang (1987) show that, for errors-in-variables regression models (and for other specific models), "it is impossible to construct confidence intervals for key parameters which have both positive confidence and finite expected length (p. 1351)." Their work "casts doubt upon the usefulness of large sample approximations in such models, at least when used for the purpose of forming confidence sets or assessing the accuracy of point estimators."

Burr (1988) suggest an ad hoc modification to the maximum likelihood solution in the "Berkson case ... under which the values of the predictor variables are set by the experimenter but not achieved exactly." She concluded that the modification was not worthwhile if $3\lambda < 1, 2\beta_1 < 1$, or $n < 60$. Some modifications suggested by Whittemore and Keller (1988) require knowledge of some of the parameters and "are most useful when applied to large data sets ... " (p. 1065).

Miller (1989) in a general multiresponse regression setting concludes that "if someone is comfortable with using a particular large sample test [on residuals] in the usual regression setting, than they should also feel comfortable with the same test when applied to errors in variables residuals." His conclusion applies to residuals obtained by any method of parameter estimation whose bias is of order $n^{-1/2}$ in probability.

Wong (1989) considers maximum likelihood estimation and slope-testing methods when λ is known (and assumed to be equal to 1).

Whittemore (1989) suggests a method where the unobserved variables ξ_i are estimated from the X_i via a James Stein estimation procedure, followed by M-estimation of the model parameters.

Jeffreys (1990) applies several robust estimation methods to astronomical data by adapting least squares software, and emphasizes the value of these procedures when outliers are present. See also Zamar (1989).

Naidu (1990) suggests an adjusted linear estimator (ALE) which depends on an unknown matrix $\mathbf{L}$, an estimate of which is obtained by a ridge-

regression-like method called the Extended Ridge Method. Some simulations indicate that the ALE improves on ordinary least squares (Y on X) in certain circumstances.

Riggs, Guarnieri, and Addelman (1978) study, partially through simulations, a variety of 34 different methods of fitting (X, Y) data. While they favor (12), they warn that a reasonably accurate estimate of λ is desirable. They also point out that the geometric mean functional realtionship occupies a "central position" in compromises between the two least squares solutions, Y on X and X on Y, an appealing characteristic (see their Figure 8, page 1338).

3. PRACTICAL ADVICE

Although many of us try to avoid the issue of errors in both X and Y by advising "Take data where the X-range is large compared with the X-error," this cannot always be done, and one must often suggest something specific. If λ is known (or can reasonably be estimated) use of the maximum likelihood solution (12) is probably best.

A simple alternative initially suggested by Wald (1940), using two groups, and amended by Bartlett (1949) to three groups is the following: Divide the data into three equal (or as equal as possible) groups with: (1) the smaller, or most negative, X-values; let $P_1 \equiv (\overline{X}_1, \overline{Y}_1)$ be the center of gravity of these. (2) The larger, or least negative, X-values; let $P_3 \equiv (\overline{X}_3, \overline{Y}_3)$ be their center of gravity. (3) The remainder, which are used only in estimating the overall center of gravity, $(\overline{X}, \overline{Y})$. . Use the line passing through $(\overline{X}, \overline{Y})$ with slope $(\overline{Y}_3 - \overline{Y}_1)/(\overline{X}_3 - \overline{X}_1)$, that is, parallel to P_1P_3. For reasoning, see Wald (1940), and Bartlett (1949). Later studies by Gibson and Jowett (1957) indicate that maximum efficency is achieved by a division of observations closer to the ratio $1:2:1$, but the exact split is not crucial.

My own preference is to suggest the geometric mean function relationship for which the estimators are

$$\hat{\beta}_1 = (S_{YY}/S_{XX})^{1/2}, \qquad \hat{\beta}_0 = \overline{Y} - \hat{\beta}_1 \overline{X}. \tag{13}$$

The estimator $\hat{\beta}_1$ is the geometric mean of the quantities

$$b_1 = S_{XY}/S_{XX}, \qquad a_1^{-1} = (S_{XY}/S_{YY})^{-1},$$

where b_1 and a_1 are, respectively, the slopes in least squares fits of Y versus X $(\hat{Y} = b_0 + b_1 X)$ and of X versus Y $(\hat{X} = a_0 + a_1 Y)$. Inverting the latter relationship leads to

$$Y = -a_0/a_1 + a_1^{-1}\hat{X}$$

so the geometric mean $\hat{\beta}_1 = (b_1 a_1^{-1})^{1/2}$ is a compromise lying in between the two "Y on X equation" slopes. Note that, if the roles of X and Y are reversed, exactly the same line emerges, that is, the fitted line

$$Y = \hat{\beta}_0 + \hat{\beta}_1 X \tag{14}$$

is uniquely defined. This natural symmetry is most appealing. The attractiveness of the geometrical mean functional relationship has been greatly enhanced by the independent discoveries of Teissier (1948) and Barker, Soh, and Evans (1988) that this solution is an optimum solution to a specific problem. (See, also Harvey and Mace, 1982.) That is, the geometric mean functional relationship minimizes the sum of the areas obtained by drawing horizontal (parallel to the X-axis) and vertical (parallel to the Y-axis) lines from each data point (see Figure 1). The symmetry of the solution is again obvious; interchange of the X and Y axes leaves the areas unchanged. One disadvantage of the geometric mean functional relationship is that no easy calculations are available for conducting tests on the parameters or constructing confidence intervals for them. For the complications involved, see, for example, Creasy (1956). A referee remarked that applying PROC NLIN in SAS and defining the LOSS function as the sum of the areas might offer some help here; I have not evaluated this possibility. (While it is true that maximum likelihood methods can make appeal to asymptotic results at this point, such results do not seem to apply too well when n is small, judging by the comments of various authors.)

We now apply the geometric mean functional relationship solution to some published sets of data.

4. EXAMPLES

Example 1. The data in Table 1 were used by Jeffreys (1990) and taken from Dressler (1984). "They consist of the integrated V magnitudes V_{26}, and log of the central velocity dispersion, log σ, of a sample of 53 galaxies from two galaxy clusters, the Coma and Virgo clusters." (Jeffreys, 1990, p. 602). The model

$$\log \sigma = \beta_0 + \beta_1 V_{26}$$

is deemed appropriate with a common β_1 and a different β_0 for each cluster. Four outliers are present (asterisked in Table 1) which we ignore. (This bypasses some of the points made by Jefferys which are not our concern here.) We adopt a dummy or indicator variable z; $z = 1$ for the Coma sample and $z = 0$ for the Virgo sample. Two least squares fits using the models

$$\log \sigma = \beta_0 + \beta_1 V_{26} + \beta_2 z + \epsilon$$

and

$$V_{26} = \alpha_0 + \alpha_1 \log \sigma + \alpha_2 z + \epsilon$$

provide, respectively, fitted equations

$$\hat{\log \sigma} = 3.4795 - 0.116334 V_{26} + 0.44097 z$$

and

$$\hat{V}_{26} = 25.329 - 6.5641 \log \sigma + 3.7685 z.$$

The slope of the geometric mean functional relationship is thus $\{-0.116334/(-6.5641)\}^{1/2} = -0.133127$. Putting parallel straight lines with this slope through the individual centers of gravity of the two sets of data provides fitted equations

$$\hat{\log \sigma} = 4.159 - 0.133 V_{26} \text{ (Coma sample)}$$

and

$$\hat{\log \sigma} = 3.656 - 0.133 V_{26} \text{ (Virgo sample).}$$

These are very close to the reference solution of Jeffreys (1990) which was "an errors-in-variables least squares fit" to the same data. (The method is not further explained.) They are virtually identical to Dressler's (1984) values obtained via a sensible ad hoc procedure. (Jeffrey's: 4.14, 3.65, -0.132; Dressler's: 4.156, 3.656, -0.1333).

Example 2. The data in Table 2, from Kelly (1984), were taken from Miller (1980). Kelly uses the data to illustrate points she is making about (i) estimating the variance of the classical estimators of (12) and (ii) detecting influential observations. We analyze them using the geometric mean structural relationship estimator.

The two fits to all the data (X = heelstick, Y = catheter) are $\hat{Y} = 2.786 + 0.8805X$, and $\hat{X} = 4.210 + 0.7870Y$, which we can invert to the form $Y = -5.349 + 1.2706\hat{X}$. The geometric mean functional relationship is thus $Y = -0.91 + 1.058X$. Both individual regressions indicate that the second observation is influential, however, and a plot of the data indicates we might consider dropping it. The two fits to the remaining 19 observations give

$$\hat{Y} = -1.628 + 1.1147X,$$

and

$$\hat{X} = 5.482 + 0.70462Y,$$

which we can invert to the form

$$Y = -7.780 + 1.4192\hat{X}.$$

Then $\hat{\beta} = (1.1147/0.70462)^{1/2} = 1.258$ and the geometric mean functional relationship is

$$\hat{Y} = -4.52 + 1.258X.$$

(If the second observation is not deleted, the parallel result would be $\hat{Y} = -0.91 + 1.058X$, where the 1.058 is the geometric mean of the slopes 0.8805 and 1.2706.) Kelly (1984) obtains two 95% confidence intervals for the slope using all the data, getting (0.76, 1.38) via a bootstrap method, and (0.76, 1.52) via a method based on normal assumptions, given by Kendall and Stuart (1961, pages 388-390). She concludes that these support the hypothesis that $\beta_0 = 0, \beta_1 = 1$, which implies that the methods of measurement which gave rise to Table 2 are equivalent. She then points out that removal of the second observation takes the estimated point for (β_0, β_1) "to approximately the edge of a 60% confidence region around" her original estimates based on a maximum likelihood analysis assuming $\lambda = 1$. I interpret that to mean that the hypothesis $\beta_0 = 0, \beta_1 = 1$ is no longer supported.

The geometric mean functional relationship does not provide confidence intervals, but we can get a rough feel for the situation by looking at the estimates when all equations are written in Y on X form. When observation 2 is included, the two slopes are 0.8805 and 1.2706 and their geometric mean is 1.0577; the two intercept values are 2.786 and -5.349 and the intercept of the geometric mean functional relationship is -0.91. One feels that the hypothesis intercept = 0, slope = 1 is not unreasonable. Now remove the second observation. The slopes are now 1.1147 and 1.4192 with a geometric mean of 1.258 (all > 1) and the two intercepts are -1.628 and -7.780 (both $<$ 0) with an intercept of -4.52 from the geometric mean functional relationship. The impression we get is that the hypothesis is not valid. Thus the situation turns on the one influential data point. Can we regard the two lines that lead to the geometric mean functional relationship as confidence limits of some sort? No properties of them are known, it seems, but using them appears to be common sense. Comments are welcomed.

SUMMARY

Practical advice on what line to fit is often sought by researchers whose (X, Y) data have errors in both variables. The extensive literature available is hard to consult quickly. This expository note provides a selective summary of some of the methods available, and suggests use of the geometric mean functional relationship as a sensible way to proceed. This method is applied to two sets of published data for illustration.

ACKNOWLEDGEMENTS

I am grateful to Penny Reynolds for supplying some of these references and to the Wisconsin Alumni Research Foundation through the University of Wisconsin Graduate School for partial support. I also thank the referee for a number of helpful comments.

SELECTED BIBLIOGRAPHY

Adcock, R.J. (1878). A problem in least squares. *Analyst*, 5, 53-00.

Amemiya, Y., and Fuller, W.A. (1984). Estimation for the multivariate errors-in-variables model with estimated error covariance matrix. *Ann. Statist.* 12, 497-509.

Barker, F., Soh, Y.C., and Evans, R.J. (1988). Properties of the geometric mean functional relationship. *Biometrics*, 44, 279-281.

Barnett, V.D. (1967). A note on linear structural relationships when both residual variances are known. *Biometrika*, 54, 670-672.

Barnett, V.D. (1969). Simultaneous pairwise linear structural relationships. *Biometrics*, 25, 129-142.

Bartlett, M.S. (1949). Fitting a straight line when both variables are subject to error. *Biometrics*, 5, 207-212.

Berkson, J. (1950). Are there two regression? *J. Am. Statist. Assoc.*, 45, 164-180.

Brown, M.L. (1982). Robust line estimation with errors in both variables. *J. Am. Statist. Assoc.*, 77, 71-79.

Burr, D. (1988). On errors-in-variables in binary regression - Berkson case. *J. Am. Statist. Assoc.*, 83, 739-743.

Carlson, F.D., Sobel, E., and Watson, G.S. (1966). Linear relationship between variables affected by errors. *Biometrics*, 22, 252-267.

Carter, R.L., and Fuller, W.A. (1980). Instrumental variable estimation of the simple errors-in-variables model. *J. Am. Statist. Assoc.*, 75, 687-692.

Chan, N.N. (1982). Linear structural relationships with unknown error variances. *Biometrika*, 69, 277-279.

Chang, T. (1989). Spherical regression with errors in variables. *The Annals of Statistics*, 17, 293-306.

Clarke, M.R.B. (1980). The reduced major axis of a bivariate sample. *Biometrika*, 67, 441-446.

Creasy, M.A. (1956). Confidence limits for the gradient in the linear functional relationship. *J. Roy. Statist. Soc. B*, 18, 65-69.

Dressler A. (1984). International kinematics of galaxies in cluster, I. Velocity dispersions for elliptical galaxies in Coma and Virgo. *Astrophysics J.*, 281, 512-524.

Eisenhart, C. (1939). The interpretation of certain regression methods and their use in biological and industrial research. *An. Math. Statist.*, 10, 162-186.

Feldstein, M. (1974). Errors in variables: A consistent estimator with smaller mean square error in finite samples. *J. Am. Statist. Assoc.*, 69, 990-996.

Fuller, W.A. (1987). Measurement Error Models. New York: Wiley.

Ganse, R.A., Amemiya, Y., and Fuller, W.A. (1983). Prediction when both variables are subject to error, with application to earthquake magnitudes. *J. Am. Statist. Assoc.*, 78, 761-765.

Ghose, B.K. (1970). Regression analysis in paleobiometrics - a reappraisal. *J. Geology*, 78, 545-557.

Gibson, W.M. and Jowett, G.H. (1957). Three-group regression analysis, Part I. *Applied Statistics*, 6, 114-122.

Gleser, L.J. (1981). Estimation in a multivariate errors-in-variables regression
model: Large sample results. *Ann. Statist.*, 9, 24-44.

Gleser, L.J., and Hwang, J.T. (1987). The nonexistence of 100(1 - α)% confidence sets of finite expected diameter in errors-in-variables and related models. *Ann. Statist.*, 15, 1351-1362.

Halperin, M. (1970). On inverse estimation in linear regression. *Technometrics*, 12(4), 727-736.

Halperin, M., and Gurian, J. (1971). A note on estimation in straight line regression when both variables are subject to error. *J. Am. Statist. Assoc.*, 66, 587-589.

Harvey, P.H. and Mace, G.M. (1982). Comparisons between taxa and adaptive trends: problems of methodology. In *Current Problems in Sociobiology*. Cambridge University Press.

Jefferys, W.H. (1990). Robust estimation when more than one variable per equation of condition has error. *Biometrika*, 77, 597-607.

Jewell, N.P., and Shiboski, S.C. (1990). Statistical analysis of HIV infectivity based on partner studies. *Biometrics*, 46, 1133-1150.

Jensen, A.L. (1986). Functional regression and correlation analysis. *Can. J. Fish. Aquat. Sci.*, 43, 1742-1745.

Jolicoeur, P. (1975). Linear regressions in fishery research: some comments. *J. Fish. Res. Board Canada*, 1491-1494.

Jolicoeur, P. (1978). Interval estimation of the slope of the major axis of a bivariate normal distribution in the case of a small sample. *Biometrics*, 24, 679-682.

Jolicoeur, P. and Heusner, A.A. (1971). The allometry equation in the analysis of the standard oxygen consumption and body weight of the white rat. *Biometrics*, 27, 841-855.

Karni, E., and Weissman, I. (1974). A consistent estimator of the slope in a regression model with errors in the variables. *J. Am. Statist. Assoc.*, 69, 211-213, corrections, 840.

Keeping, E.S. (1962). Introduction to Statistical Inference. Princeton, NJ: Van Nostrand. (Pages 298-305.)

Kelly, G. (1984). The influence function in the errors in variables problem. *Ann. Statist.*, 12, 87-100.

Kendall, M.G. (1951). Regression, structure and functional relationship. Part
I. *Biometrika*, 38, 11-25.

Kendall, M.G. (1952). Regression, structure and functional relationship. Part
II. *Biometrika*, 39, 96-108.

Kendall, M.G., and Stuart, A. (1961). *The Advanced Theory of Statistics.* Vol. 2. New York: Hafner.

Kent, J.T. (1983). Information gain and a measure of correlation. *Biometrika*. 70, 163-173.

Kermack, K.A., and Haldane, J.B.S. (1950). Organic correlation and allometry. *Biometrika*, 37, 3-41.

Kerrich, J.E. (1966). Fitting the line $Y = \alpha X$ when errors of observation are present in both variables. *Am. Statist.*, 20, 24.

Ketellapper, R.H. (1983). On estimation of parameters in a simple linear errors-in-variables models. *Technometrics*, 25, 43-47.

Kuhry, B., and Marcus, L.F. (1977). Bivariate linear models in biometry. *Syst. Zool.*, 26, 201-209.

Lakshminarayanan, M.Y., and Gunst, R.F. (1984). Estimation of parameters in linear structural relationships: Sensitivity to the choice of the ratio of error variances. *Biometrika*, 71, 569-573.

Lindley, D.V. and El-Sayyad, G.M. (1968). The Bayesian estimation of a linear functional relationship. *J. Roy. Statist. Soc. B,* 30, 190-202.

Mandansky, A. (1959). The fitting of straight lines when both variables are subject to error. *J. Am. Statist. Assoc.,* 54, 173-205.

Mandel, J. (1964). The fitting of straight lines. Pages 272-311 in *The Statistical Analysis of Experimental Data.* New York: Wiley Interscience.

Mandel, J. (1984). Fitting straight lines when both variables are subject to error. *J. Quality Technol.,* 16, 1-14.

McArdle, B.G. (1988). The structural relationship: regression in biology. *Can. J. Zool.*, 66, 2329-2339.

Miller, R.G. (1980). Kanamycin levels in premature babies. Biostatistics Casebook, Vol. III, 127-142. (Technical Report No. 57, Division of Biostatistics, Stanford University.)

Miller, S.M. (1989). Empirical processes based upon residuals from errors-in-variables regressions. *Ann. Statist.,* 17, 282-292.

Moran, P.A.P. (1971). Estimating structural and functional relationships. *J. Multivar. Anal.* 1, 232-255.

Naidu, L.K. (1990). An adjusted linear estimator. *Computational Statistics and Data Analysis*, 10, 143-151.

Patefield, W.M. (1981). Multivariate linear relationships: Maximum likelihood estimation and regression bounds. *J. Roy. Statist. Soc. B*, 43, 342-352.

Rayner, J.M.V. (1985). Linear relations in biomechanics: the statistics of scaling functions. *J. Zool.* (Lond.), 206, 415-439.

Reilly, P.M., and Patineo-Leal, H. (1981). A Bayesian study of the errors-in-variables models. *Technometrics*, 23, 221-231.

Ricker, W.E. (1973). Linear regressions in fishery research. *J. Fish. Res. Board Can.*, 30, 409-434.

Ricker, W.E. (1975). A note concerning Professor Jolicouer's comments. *J. Fish. Res. Board Canada*, 32, 1494-1498.

Ricker, W.E. (1984). Computation and uses of central trend lines. *Can. J. Zool.*, 62, 1897-1905.

Riggs, D.S., Guarnieri, J.A., and Addelman, S. (1978). Fitting straight lines when both variables are subject to error. *Life Sciences*, 22, 1305-1360.

Rivest, L-P. (1989). Spherical regression for concentrated Fisher-von Mises distributions. *Ann. Statist.*, 17, 293-306.

Sampson, A.R. (1974). A tale of two regressions. *J. Am. Statist. Assoc.*, 69, 682-689.

Schnute, J. (1984). Linear mixtures: A new approach to bivariate trend lines. *J. Am. Statist. Assoc.*, 79, 1-8.

Spezzaferri, F. (1985). A note on multivariate calibration experiments. *Biometrics,* 41, 267-272.

Sprent, P., and Dolby, G.R. (1980). The geometric mean functional relationship. *Biometrics*, 36, 547-550. (See, also, 38 pp. 859-860.)

Stefanski, L.A. (1985). The effect of measurement error parameter estimation. *Biometrika*, 72, 583-592.

Stroud, T.W.F. (1972). Comparing conditional means and variances in a regression model with measurement errors of known variances. *J. Am. Statist. Assoc.*, 67, 407-412, discussion 412-414, correction (1973) 68, 251.

Teissier, G. (1948). La relation d'allometrie sa signification statistique et biologique. *Biometrics*, 4, 14-48 (discussion, 48-53).

Tosteson, T.D., and Ware, J.H. (1990). Designing a logistic regression study using surrogate measures for exposure and outcome. *Biometrika*, 77, 11-21.

Wakkers, P.J.M., Hellendoorn, H.B.A., op der Weegh, G.J., and Heerspink, W. (1975). Applications of statistics in clinical chemistry: a critical evaluation of regression lines. *Clinica Chimica Acta*, 64, 173-184.

Wald, A. (1940). The fitting of straight lines if both variables are subject to error. *Ann. Math. Statist.*, 11, 284-300.

Ware, J.H. (1972). The fitting of straight lines when both variables are subject to error and the ranks of the means are known. *J. Am. Statist. Assoc.*, 67, 891-897.

Whittemore, A.S. (1989). Errors-in-variables regression using Stein estimates. *Am. Statistician*, 43, 226-228.

Whittemore, A.S., and Keller, J.B. (1988). Approximations for regression with covariate measurement error. *J. Am. Statist. Assoc.*, 83, 1057-1066.

Wolter, K.M., and Fuller, W.A. (1982). Estimation of the quadratic errors-in-variables model. *Biometrika*, 69, 175-182.

Wong, M.Y. (1989). Likelihood estimation of a simple linear regression model when both variables have error. *Biometrika*, 76, 141-148.

Zamar, R.H. (1989). Robust estimation in the errors-in-variables model. *Biometrika*, 76, 149-160.

Zar, J.H. (1967). The effect of changes in units of measurement on least squares regression lines. *Bioscience*, 1967, 818-819.

Table 1
Star data from Dressler (1984) and Jeffreys (1990). In the example, the four asterisked observations will be ignored.

Coma sample ($z = 1$)		Virgo sample ($z = 0$)	
V_{26}	$\log \sigma$	V_{26}	$\log \sigma$
12.60	2.449	11.39	2.242
13.12	2.394	12.53*	1.716*
14.23	2.285	9.98	2.412
14.86	2.166	9.37	2.480
15.88*	1.863*	12.24	2.059
13.92	2.286	9.20	2.355
15.45*	1.761*	12.17	2.009
14.36	2.209	12.01	1.949
15.07	2.113	12.50	2.079
14.07	2.301	8.56	2.474
14.53	2.243	10.28	2.268
15.60	2.169	11.28	2.170
12.27	2.383	8.79	2.528
14.36	2.311	12.02*	1.778*
14.50	2.339	11.92	2.021
15.52	2.251	9.95	2.391
13.46	2.361	11.30	2.185
11.85	2.584	9.88	2.338
15.31	2.007	9.82	2.303
13.98	2.180	8.90	2.514
15.28	2.099	10.97	2.262
14.26	2.275	9.28	2.276
14.11	2.320	11.37	2.027
14.87	2.191		
14.82	2.247		
15.37	2.059		
13.49	2.394		
15.04	2.154		
13.67	2.274		
12.88	2.383		

Table 2
Serum kanamycin levels in blood samples drawn simultaneously from an umbilical catheter and a heel venapuncture in twenty babies

Baby	Heelstick (X)	Catheter (Y)
1	23.0	25.2
2	33.2	26.0
3	16.6	16.3
4	26.3	27.2
5	20.0	23.2
6	20.0	18.1
7	20.6	22.2
8	18.9	17.2
9	17.8	18.8
10	20.0	16.4
11	26.4	24.8
12	21.8	26.8
13	14.9	15.4
14	17.4	14.9
15	20.0	18.1
16	13.2	16.3
17	28.4	31.3
18	25.9	31.2
19	18.9	18.0
20	13.8	15.6

Figure 1

The geometrical mean functional relationship line minimizes the sum of the shaded areas. (The dots are data points.)