TECHNICAL REPORT NO 5/87

STATISTICAL CALIBRATION WHEN BOTH VARIABLES ARE SUBJECT TO ERROR

David R Fox
Department of Mathematics & Statistics
Curtin University of Technology

March 1987

Statistical Calibration when both variables are subject to error

David R Fox
Department of Mathematics and Statistics

ABSTRACT

In this paper we provide a methodology for calibration problems of the form Y = f(X) where measurements on Y contain two sources of variability, namely instrument error in the determination of Y and uncertainty in X.
The procedures presented are sufficiently general to cater for non-linear functions f, as well as correlated errors in X and Y.
We commence with a review of previous work associated with the so-called classical and inverse methods of regression.
A procedure due to Mandel (1984) has been adopted which removes the ambiguity associated with the two regressions. The successful application of this approach however, requires prior knowledge of the error covariance structure associated with X and Y. Given certain distributional assumptions and some simplifying approximations we show how estimates of the error variance components may be extracted from the calibration data.
A practical application associated with the determination of vehicle speed by airborne observation is given.

KEYWORDS : Calibration, inverse regression, classical regression, components of variance.

1. The Calibration Problem

In calibrating some instrument, we take readings Y on some physical process X and use the empirical relationship between Y and X to 'predict' the value of X given some future reading y_0.

In most instances it is assumed that the relationship between Y and X is linear, and thus estimation of the parameters α and β in the regression of y on x is readily achieved via OLS, to give :

$$\hat{Y}_i = \hat{\alpha} + \hat{\beta}\, x_i \tag{1.1}$$

In contrast to the 'normal' use of equation (1.1) where some new value of Y is to be predicted for a given value of $X = x_0$, the requirement here is to predict x_0 , having observed $Y = y_0$. In the so-called 'classical' calibration method we obtain the estimate $\hat{x}_0$ by a simple re-arrangement of the terms in equation (1.1) viz :

$$\hat{x}_0 = \frac{y_0 - \hat{\alpha}}{\hat{\beta}} \tag{1.2}$$

An alternative, and equally appealling approach is to treat X as dependent and regress X on Y . This procedure is known as the 'inverse' calibration method for which we obtain $\hat{\alpha}^*$ and $\hat{\beta}^*$ as our parameter estimates in the regression

$$X_i = \alpha^* + \beta^* y_i \tag{1.3}$$

The problem of deciding between these two methods is not new. Eisenhart (1939) suggests that both methods were in common use up to the time of his paper, although firmly rejects procedures based on equation (1.3) arguing that the least-squares line should be fitted to the variable which is observed with error. Krutchkoff (1967,1969), on the other hand, advocates the use of inverse calibration and presents the results of simulation studies in which the relative merits of each method were assessed.

A number of papers critical of Krutchkoff's work have since appeared, although general agreement on the 'best' approach has not been reached. The problem, it seems, stems from the fact that there is no universally accepted properties of an optimal estimator in the calibration problem.

Williams (1969) pointed out that in the case of normally distributed errors, the classical estimator has an undefined expectation and infinite variance and as such any comparison based on mean squared error (MSE) is rendered meaningless.

Berkson (1969) advocates the concept of Pitman closeness as a means for comparison although notes that estimators obtained by the inverse method are not consistent nor asymptotically unbiased. This lack of consistency was also observed by Madansky (1959).

Martinelle (1970) shows that the MSE for the inverse estimator is less than that of the classical estimator provided

$$(x_0 - \bar{x})^2 < s_x^2 \left[2 + \frac{1}{\theta^2 s_x^2}\right]$$

where $\bar{x}$ and s_x^2 are the sample mean and (biased) sample variance respectively and $\theta = \beta/\sigma$. Furthermore, Martinelle suggests that when $\theta^2 s_x^2$ is large, then there is little advantage in using the inverse method.

Lwin and Martiz (1980) proposed the 'non-linear' predictor of x_0 which was shown to have desirable properties not shared by other methods. In a later development, Lwin and Maritz (1982) considered a more general class of estimators and in particular demonstrated that

$$x^*(Y) = (1 + \theta^2 s_x^2)^{-1}\,\bar{x} + \theta^2 s_x^2\,(1 + \theta^2 s_x^2)^{-1}\,[(Y - \alpha)/\beta]$$

is optimal in the sense that it results in smallest MSE when applied to previous y_i's. This estimator makes use of the current observation y as well as previous x_i values of the calibration experiment.

Much work has also been devoted to other aspects of the calibration problem including extensions to the multivariate setting. Anders (1973), for example, considers the problem of finding simultaneous confidence intervals in the inverse regression. Spiegelman (1984) has explored the use of calibration curves in quality-control situations while Brown (1982), Wood (1982), Spezzaferri (1985), Oman & Wax (1984) and others have examined the use of multivariate calibration techniques.

Oman (1984) has derived a statistic, similar in nature to Cook's distance to measure the influence of a particular observation on future estimates from the calibration curve. Spiegleman (1984) has similarly considered the role of regression diagnostics in the calibration problem.

While much discussion continues on the relative merits of the inverse and classical approaches, a further complexity is introduced when one considers situations in which both X and Y are measured with error. Berkson (1950) formally raised the question of the existence of two separate regression lines in such cases. Small-sample properties of $\hat{\beta}$ were investigated by Halperin and Gurian (1971) and under certain prescribed conditions, results for $E[\hat{\beta}]$ and MSE $[\hat{\beta}]$ were derived. Clutton-Brock (1967) argues that there is "no paradox of two regression lines" and suggests that in the case of errors in both X and Y, the maximum likelihood estimates lie between the two separate regressions of Y on X and X on Y.

In the remainder of this paper we exploit a method due to Mandel (1984) in which the problems associated with regression in the case of both variables subject to error are largely removed. Furthermore, we show how the procedure is readily adapted to the calibration problem and how the ambiguities of the inverse and classical methods are avoided.

We now describe in more detail the method due to Mandel.

2. Fitting straight lines when both variables are subject to error

We commence with the model

$$y_i = \alpha + \beta x_i \qquad i = 1, \ldots, N \qquad (2.1)$$

The assumptions underlying the usual fitting procedure are :

(i) the x values have no error

(ii) the y values are subject to error - in particular ε_i is the error associated with y_i. These ε_i represent a random sample from a population with mean zero and variance σ^2_ε

(iii) Cov $[\varepsilon_i, \varepsilon_j] = 0 \;\; \forall \;\; i, j, \;\; i \neq j$.

By relaxing condition (i), we now allow the x's to be affected by an error denoted by δ, where the δ_i also represent a random sample from some population having zero mean and variance σ^2_δ .

Mandel shows that provided the quantity

$$\phi = \frac{|\beta|}{\sigma_\varepsilon \; \sigma_\delta} \ll 1$$

and that $\rho = 0$

then the OLS conditions still apply and that α and β may be estimated in the usual manner.

A more general situtation in which the OLS conditions do not apply is summarized by relations 2.2 (a), 2.2 (b), and 2.2 (c).

$$E[Y_i] = \alpha + \beta \; E[X_\lambda] \qquad (2.2\text{ a})$$

$$\frac{\sigma^2_\varepsilon}{\sigma^2_\delta} = \lambda \qquad (2.2\text{ b})$$

$$\rho(\varepsilon, \delta) = \rho \qquad (2.2\text{ c})$$

Under OLS conditions we find that each experimental point is projected vertically onto the fitted line. In the case where both variables are subject to error the angle of projection, γ will depend on λ .
The method proposed by Mandel relies on the construction of two new variables u and v, related to x and y, but formed in such a way that the OLS conditions are at least approximately fulfilled for (u,v)
The (u,v) data are obtained as follows :

$$u_i = x_i + ky_i \qquad (2.3\ a)$$

$$v_i = y_i - bx_i \qquad (2.3\ b)$$

for some constants k and b.
It is shown by Mandel that, while k and b are theoretically unknown, they can be approximated by solving equations (2.4 a) and (2.4 b).

$$b = \frac{s_{xy} + ks_{yy}}{s_{xx} + ks_{xy}} \qquad (2.4\ a)$$

$$k = \frac{b - \theta}{\lambda - b\theta} \qquad (2.4\ b)$$

where θ is defined by

$$\theta = \rho\sqrt{\lambda} \qquad (2.4\ c)$$

and s_{xx}, s_{yy} and s_{xy} are the usual sums of squares and cross-products defined by

$$s_{xx} = \sum x_i^2 - \frac{(\sum_i x_i)^2}{n} \qquad (2.5\ a)$$

$$s_{yy} = \sum_i y_i^2 - \frac{(\sum_i y_i)^2}{n} \qquad (2.5\ b)$$

$$s_{xy} = \sum_i x_i y_i - \frac{(\Sigma x_i)(\Sigma y_i)}{n} \qquad (2.5\ c)$$

Equations (2.4 a) and (2.4 b) define a quadratic in b which may be readily solved using

$$b = \frac{(s_{yy} - \lambda s_{xx}) \pm \sqrt{(s_{yy} - \lambda s_{xx})^2 - 4(s_{xy} - \theta s_{xx})(\theta s_{yy} - \lambda s_{xy})}}{2(\theta s_{yy} - \lambda s_{xy})} \qquad (2.6)$$

In addition, it is shown that the estimate of β in equation (2.1) is equal to b as defined in (2.6) and that $\hat{\alpha}$ is found in the usual manner ie. $\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$.

In the context of statistical calibration, Mandel's procedure is most appealing since the choice between the inverse and classical regressions does not arise. Furthermore, the uncertainty associated with a predicted value is the same, regardless of whether we consider X as the independent variable and Y as the dependent variable or vice-versa.

In the development of his approach, Mandel has tacitly assumed that σ^2_ε , σ^2_δ and ρ are all known constants. This will rarely be the case in practice and thus these quantities need to be estimated. It is this problem of estimation with which we now concern ourselves.

3. A model for the error components

The method of variance component estimation in the calibration problem has been discussed in detail by Fox (1987).

A generalization of these methods is now presented which allows for the possibility of correlated errors (the two error components were previously assumed to be independent).

We define a general calibration problem having the following components :

X_i is the true state of nature (unknown)

x_i is the assumed state of nature

Y_i is a measured response corresponding to X_i

y_i is the true value of the response for the state of nature X_i

Furthermore, we have

$$X_i = x_i + U_i \quad \text{and} \quad Y_i = y_i + V_i$$

where U_i and V_i are random errors reflecting

(a) our less than perfect knowledge of the true state of nature, and

(b) our inability to make error-free measurement.

We assume in our calibration problem that the data obtained consist of pairs (x_i, Y_i) and that the underlying response-generating model is of the form

$$\begin{aligned} Y_i &= f(X_i) + V_i \\ &= f(x_i + U_i) + V_i \end{aligned} \qquad (3.1)$$

Using a first-order approximation we have

$$Y_i \simeq f(x_i) + U_i f'(x_i) + V_i \qquad (3.2)$$

In what follows we shall assume that U_i and V_i have the bivariate normal distribution

$$h_{U_i,V_i}(u_i, v_i) = \frac{1}{2\pi\sigma\ \sigma\ \sqrt{1-\rho^2}} \exp\{-\tfrac{1}{2}(\frac{1}{1-\rho^2})[\frac{u_i^2}{\sigma_u^2} - 2\ \frac{u_i v_i}{\sigma_u \sigma_v} + \frac{v_i^2}{\sigma_v^2}]\}$$

(3.3)

From (3.2) it is apparent that

$$\sigma_i^2 = \mathrm{Var}[Y_i] \simeq \sigma_v^2 + \sigma_u^2\,[f'(x_i)]^2 + 2f'(x_i)\mathrm{Cov}[U_i,V_i]$$

$$= \sigma_v^2 + \sigma_u^2\,[f'(x_i)]^2 + 2f'(x_i)\,\rho\,\sigma_u\sigma_v \qquad (3.4)$$

Our problem now is to estimate σ_u^2, σ_v^2 and ρ given values of s_i^2 at various values of x_i.

We note from equation (3.4) that the model is no longer linear due to the simultaneous presence of σ_u^2, σ_v^2, σ_u and σ_v .

4. Estimation of the covariance structure for the components of error

As described in Technical report 3/87, we have at each x_i an estimate, S_i^2 , of σ_i^2

S_i^2 is assumed to have the pdf

$$g_{S_i^2}(s_i^2) = \frac{m^m}{\Gamma(m)}\,(s_i^2)^{m-1}\,\left(\frac{1}{\sigma_i^2}\right)^m e^{-ms_i^2/\sigma_i^2} \qquad (4.1)$$

where $m = \frac{(n-1)}{2}$ and $s_i^2 > 0$ and n is the number of replications at each x_i .

From equation (4.1) we obtain the likelihood function

$$L(\sigma_i^2, s_i^2) = K^N\{\exp[-m\sum_{i=1}^{N}\frac{s_i^2}{\sigma_i^2}]\}\prod_{i=1}^{N}(s_i^2)^{m-1}\left(\frac{1}{\sigma_i^2}\right)^m \qquad (4.2)$$

where $K = \frac{m^m}{\Gamma(m)}$

and the log-likelihood function

$$\ln L = N \ln K - m \sum_{i=1}^{N} \left(\frac{s_i^2}{\sigma_i^2}\right) - m \sum_{i=1}^{N} \ln \sigma_i^2 + (m - 1) \sum_{i=1}^{N} \ln (s_i^2) \tag{4.3}$$

Our aim is thus to estimate σ_i^2 such that equation (4.3) is maximized. Specifically

$$\max \quad F(\sigma_v^2;\sigma_u^2;\rho) = \sum_{i=1}^{N} (\ln w_i - s_i^2 w_i) \tag{4.4}$$

where $w_i = \frac{1}{\sigma_i^2}$ and

$$\sigma_i^2 = \sigma_v^2 + \sigma_u^2 \, [f'(x)]^2 + 2f'(x) \, \rho\sigma_u\sigma_v$$

subject to

$$\sigma_v^2 \geq 0$$

$$\sigma_u^2 \geq 0$$

$$-1 \leq \rho \leq +1 \, .$$

Instead of maximizing (4.4), we may choose to minimize

$$\Phi(\sigma_v^2;\sigma_u^2;\rho) = \sum_{i=1}^{N} (s_i^2 w_i - \ln w_i) \tag{4.5}$$

We have used the NAG routine EØ4LAF, a modified Newton algorithm, as described in the NAG reference manual (1977).

A brief description of the procedure follows.

4.1 Parameter estimation

Let $\underset{\sim}{\phi}$ denote the vector of parameters to be estimated. In this case

$$\underset{\sim}{\phi}^T = [\sigma_v^2, \sigma_u^2, \rho]$$

We wish to find values $\underset{\sim}{\phi}^*$ of $\underset{\sim}{\phi}$ for which $\Phi(\underset{\sim}{\phi})$ is minimized.

We commence with a given point $\underset{\sim}{\theta}_1$ and proceed to generate a sequence of points $\underset{\sim}{\theta}_2, \underset{\sim}{\theta}_3$ . . . which we hope converges to the point $\underset{\sim}{\theta}^*$ at which $\Phi(\underset{\sim}{\theta}^*)$ is minimum.

Let $\underset{\sim}{H}(\underset{\sim}{\theta})$ be the Hessian matrix of the function $\Phi(\underset{\sim}{\theta})$ and $\underset{\sim}{q}(\underset{\sim}{\theta}) = \frac{\partial\Phi}{\partial\theta}$ the gradient vector of $\Phi(\underset{\sim}{\theta})$.

Then, the i^{th} iteration of the Newton method is

$$\underset{\sim}{\theta}_{i+1} = \underset{\sim}{\theta}_i - \underset{\sim}{H}_i^{-1} \underset{\sim}{q}_i \tag{4.6}$$

Thus, for the present case we have :

$$\underset{\sim}{q}(\underset{\sim}{\theta}) = \begin{bmatrix} \dfrac{\partial\Phi}{\partial\sigma_v^2} \\[2ex] \dfrac{\partial\Phi}{\partial\sigma_u^2} \\[2ex] \dfrac{\partial\Phi}{\partial\rho} \end{bmatrix}$$

and

$$\underset{\sim}{H}(\underset{\sim}{\theta}) = \begin{bmatrix} \dfrac{\partial^2\Phi}{\partial^2\sigma_v^2} & \dfrac{\partial^2\Phi}{\partial\sigma_u^2\partial\sigma_v^2} & \dfrac{\partial^2\Phi}{\partial\rho\partial\sigma_v^2} \\[2ex] \dfrac{\partial^2\Phi}{\partial\sigma_u^2\partial\sigma_v^2} & \dfrac{\partial^2\Phi}{\partial\sigma_u^2} & \dfrac{\partial^2\Phi}{\partial\sigma_u^2\partial\rho} \\[2ex] \dfrac{\partial^2\Phi}{\partial\rho\partial\sigma_v^2} & \dfrac{\partial^2\Phi}{\partial\rho\partial\sigma_u^2} & \dfrac{\partial^2\Phi}{\partial^2\rho} \end{bmatrix}$$

We now demonstrate the procedures discussed in the preceding sections.

5. An example : Determination of vehicle speed by airborne observation

The problem of determining the speed of a vehicle by measuring from an overhead aircraft the time taken to travel some prescribed distance c is discussed in reports by Fox (1985, 1987).

In this case the calibration function takes the form

$$f(x_i) = \frac{c}{x_i}$$

where x_i is an assumed speed.

The components of the gradient vector and Hessian matrix are given in Appendix A while a listing of the computer program may be found in Appendix B.

We commence by validating the procedure using simulated data from a bivariate normal distribution for which σ_v^2, σ_u^2 and ρ are all known. In addition, by repeating the procedure a number of times some idea of the sampling variability of the parameter estimates may be obtained.

5.1 Model Validation

For the purpose of the exercise we used equation (3.1) to obtain at each x_i 30 values of Y_i . The error components u_i and v_i were generated from the pdf represented by equation (3.3) having $\sigma_u^2 = 4.5$, $\sigma_v^2 = 0.85$ and $\rho = 0.8$. Values of s_i^2 were then obtained at each x_i using the 30 observations in each case. The (x_i, s_i^2) data was then used as input to the modified Newton programme (Appendix B) and parameter estimates $\hat{\sigma}_u^2$, $\hat{\sigma}_v^2$ and $\hat{\rho}$ obtained.

By repeating the procedure ten times we were able to assess the sampling variability in $\hat{\sigma}_u^2$, $\hat{\sigma}_v^2$ and $\hat{\rho}$.

The results of these simulations are given in the following table.

TABLE 5.1 Sample Variances at each x_i for each of 10 replications

Data generated using $\sigma_v = 0.85$ $\sigma_u = 4.5$ $\rho = 0.8$

x_i	1	2	3	4	5	6	7	8	9	10	Mean	S.D.
80	0.5461	0.5929	0.6037	0.8226	0.4706	0.5055	0.6529	0.8761	0.6989	1.0000		
85	0.4858	0.6304	0.5041	0.4007	0.3215	0.4147	0.5715	0.3697	0.5373	0.4928		
90	0.3493	0.2560	0.2490	0.3329	0.5716	0.4900	0.3387	0.4530	0.3341	0.3091		
95	0.3697	0.4083	0.2323	0.4083	0.2683	0.2694	0.3047	0.2070	0.2683	0.4045		
100	0.2830	0.2266	0.2172	0.2440	0.2852	0.2480	0.2948	0.4020	0.3709	0.3493		
105	0.2421	0.1989	0.3493	0.2756	0.2323	0.3564	0.3564	0.2256	0.2381	0.2052		
110	0.2381	0.3238	0.3364	0.1978	0.1529	0.2352	0.3260	0.1936	0.1980	0.1823		
115	0.1892	0.2025	0.4135	0.3434	0.2200	0.1197	0.3215	0.2905	0.2905	0.1624		
120	0.1849	0.2381	0.1706	0.2809	0.2052	0.2228	0.2070	0.1260	0.2381	0.1568		
125	0.3552	0.3318	0.2275	0.1747	0.3648	0.1616	0.3226	0.3376	0.2611	0.3215		
130	0.3612	0.2735	0.3058	0.2500	0.2652	0.3894	0.4007	0.1584	0.3376	0.3306		
135	0.4251	0.4775	0.2981	0.3069	0.1900	0.3919	0.2862	0.3215	0.3758	0.1498		
140	0.2798	0.2852	0.2798	0.2884	0.4409	0.2632	0.4147	0.3283	0.4529	0.2981		
$\hat{\sigma}_v$	0.9248	0.9772	0.7460	0.8123	0.8543	0.8230	0.8946	0.8763	1.0229	0.9224	0.8854	0.0817
$\hat{\sigma}_u$	5.0064	5.3932	3.7315	4.6493	4.6475	4.5126	4.6761	5.1612	5.6528	5.7595	4.9190	0.6071
ρ	0.8423	0.8624	0.7190	0.7934	0.8232	0.8020	0.7963	0.8384	0.8743	0.8707	0.8222	0.0470

With reference to Table 5.1, we see that in most cases the estimators are in close agreement with the true parameter values, although we note some evidence of bias.

All three parameters have, on average, been over estimated by about half a standard deviation. This is not totally unexpected since our expression for σ_i^2 (equation 3.4) is, after all, only a first-order approximation. Presumably the degree of bias will be very much dependent on the exact nature of f(x) and how well it is represented by a first-order approximation.

It is difficult to draw any specific conclusions from the above results. Further experimentation would be required to assess the effects of different values of n and different parameter values, σ_u^2 , σ_v^2 , and ρ. Nevertheless, as a general methodology designed to extract the components σ_u^2 ,σ_v^2 and ρ , the above results give us no reason to modify the procedure.

We now apply this method to actual experimental data in which a vehicle was timed by an observer in an overhead aircraft. A detailed discussion of the experimental procedure is given in Fox (1985).

6. Application to air survellance method

The procedure described in §5 is now applied to the following test data.

Assumed speed $\tilde{x}_i$	Sample variance s_i^2
85	0.0441126
90	0.0483296
95	0.0216649
100	0.0162971
105	0.0057608
110	0.0063680
115	0.0183088
120	0.0007868
125	0.0080210
130	0.0028164
135	0.0016314
140	0.0027468

We find convergence of the iterative procedure at the point

$\sigma_v^2 = 0.0072$; $\sigma_u^2 = 1.3250$; $\rho = 0.8981$.

In technical report 3/87 in which ρ was assumed zero we obtained $\hat{\sigma}_v^2 = -0.0035$ and $\hat{\sigma}_u^2 = 0.6614$. Clearly then, the spurious result for $\hat{\sigma}_v^2$ is an artifact of incomplete model specification.

The present results suggest that the variability (as measured by the standard deviation) in maintaining constant vehicle speed is about 1 km/hr; the variability associated with making time measurements is of the order of 1/10 th of a second and that the two errors are strongly (and positively) correlated. These results are not surprising and tally with what one would intuitively expect. With these estimates, we are now in a position to 'calibrate' the method according to the procedure outlined in § 2.

7. Calibration of the air surveillance method

We first require an estimate of λ which is defined to be the ratio of the error variance for Y to the error variance for X.
From equations (3.1) and (3.2) we have

$$Y_i = f(X_i) + V_i$$
$$= X_i' + V_i$$

where

$$X_i' = f(X_i)$$

Thus

$$Var[X_i'] = Var[f(X_i)]$$
$$Var[f(x_i) + U_i\ f'(x_i)$$
$$= [f'(x_i)]^2\sigma_u^2 \qquad (7.1)$$

Now, $Var[Y_i] = \sigma_i^2$ is defined by equation (3.4). We observe that $\lambda = Var[Y_i]/Var[X_i]$ is not constant, although as we will show later, this has little effect on the final outcome.
We now apply Mandel's regression procedure to the experimental data in order to estimate α and β in the model :

$$Y_i = \alpha + \beta X_i' \qquad (7.2)$$

where Y_i is the measured time at speed X_i and $X_i' = \frac{c}{X_i}$ represents the 'true' time for the assumed speed x_i.
At each x_i we compute λ and θ (equation 2.4 c). Equation (2.6) is then solved to obtain b which is our estimate $\hat{\beta}$ of β in equation (7.2).
For the experimental data we obtain :

$$Sx'x' = 686.1667$$
$$Sx'y = 679.7297$$
$$Syy = 674.8906 \qquad n = 78.$$

Our calculations are summarized in Table 7.1 below :

x_i	λ	θ	b
80	0.597	0.6939	1.0061
85	0.556	0.6697	1.0060
90	0.514	0.6439	1.0059
95	0.472	0.6170	1.0058
100	0.431	0.5896	1.0057
105	0.394	0.5637	1.0057
110	0.355	0.5351	1.0057
115	0.322	0.5096	1.0057
120	0.290	0.4386	1.0058
125	0.261	0.4588	1.0058
130	0.233	0.4335	1.0059
135	0.217	0.4184	1.0059
140	0.205	0.4066	1.0060

Table 7.1: Estimation of $\hat{\beta} = b$

As can be seen from Table 7.1, even though λ is not constant, its variation is not sufficient to substantially alter b. Thus to three decimal places we obtain $\hat{\beta} = 1.006$.

$\hat{\alpha}$ is estimated in the usual manner, ie.

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} = -0.3238 .$$

Our estimated regression model is thus

$$\hat{y} = 1.006x - 0.3238 \tag{7.3}$$

In comparison, we obtain the following models using

(a) 'classical' regression

and

(b) inverse regression :

$$\hat{y}^* = 0.9906x \tag{7.4 a}$$

$$\hat{x}^* = 1.0072y + 0.1042 \tag{7.4 b}$$

With respect to equation (7.3) we note that $\hat{\alpha}$ represents a constant bias of approximately 0.32 seconds. This is very close to the value of 0.36 reported previously [Fox (1985), §4.2.1]. Secondly, it is observed that $\hat{\beta}$ in equation (7.3) falls between the corresponding estimates in equations (7.4 a) and (7.4 b).

For equations (7.3) and (7.4 a) we obtain the following calibration equations :

$$\hat{x} = \frac{y + 0.3238}{1.006} \tag{7.5 a}$$

$$\hat{x}^* = \frac{y + 0.0655}{0.9906} \tag{7.5 b}$$

$\hat{x}^*$, obtained via the inverse calibration method is obtained from equation (7.4 b) .

We now compare all three estimators.

True Speed	Speed as determined by equation * (7.5 a)	(7.5 b)	(7.4 b)
80	79.34	79.02	79.06
85	84.22	83.94	83.98
90	89.10	88.86	88.90
95	93.96	93.78	93.81
100	98.82	98.70	98.72
105	103.67	103.62	103.62
110	108.51	108.53	108.53
115	113.35	113.44	113.43
120	118.17	118.36	118.33
125	122.98	123.26	123.22
130	127.79	128.17	128.11
135	132.59	133.08	133.00
140	137.38	137.98	137.89

* The three equations yield computed <u>times</u>. These are then converted to a speed over a 500m distance.

In this case the differences between the three methods are negligable. This is due to the fact that the relationship between independent and dependent variables is almost perfect and the regression line has slope very close to 1.

In general, however, such close agreement cannot be expected.

APPENDIX A : Gradient Vector and Hessian Matrix

A.1

We give here computing formulae required to determine the elements of the gradient vector and Hessian matrix for the components of variance estimation associated with the example of § 5.

We have :

$$\Phi(\sigma_v^2;\sigma_u^2;\rho) = \sum_{i=1}^{N} (s_i^2 w_i - \ln w_i)$$

where $$w_i = \frac{1}{\sigma_i^2}$$

and $$\sigma_i^2 = \sigma_v^2 + \frac{c^2}{x_i^4}\sigma_u^2 - \frac{2c}{x_i^2}\rho\sigma_u\sigma_v$$

A.1 Elements of the gradient vector

$$\underset{\sim}{q}(\underset{\sim}{\theta})^T = \left[\frac{\partial\Phi}{\partial\sigma_v^2} \quad \frac{\partial\Phi}{\partial\sigma_u^2} \quad \frac{\partial\Phi}{\partial\rho} \right]$$

$$\frac{\partial\Phi}{\partial\sigma_v^2} = \sum_{i=1}^{N} (w_i - w_i^2 s_i^2)\frac{\partial\sigma_i^2}{\partial\sigma_v^2}$$

$$\frac{\partial\Phi}{\partial\sigma_u^2} = \sum_{i=1}^{N} (w_i - w_i^2 s_i^2)\frac{\partial\sigma_i^2}{\partial\sigma_u^2}$$

$$\frac{\partial\Phi}{\partial\rho} = \sum_{i=1}^{N} (w_i - w_i^2 s_i^2)\frac{\partial\sigma_i^2}{\partial\rho}$$

and

$$\frac{\partial\sigma_i^2}{\partial\sigma_v^2} = 1 - \frac{c\rho\sigma_u}{x_i^2\sigma_v}$$

$$\frac{\partial \sigma_i^2}{\partial \sigma_u^2} = \frac{c^2}{x_i^4} - \frac{c\rho\sigma_v}{x_i^2\sigma_u}$$

$$\frac{\partial \sigma_i^2}{\partial \rho} = \frac{-2c\,\sigma_u\sigma_v}{x_i^2}$$

A.2 Elements of the Hessian matrix

$\underset{\sim}{H}(\underset{\sim}{\theta})$ is given in § 4.1

$$\frac{\partial^2\Phi}{\partial^2\sigma_v^2} = \sum_{i=1}^{N}\{[w_i^2 - 2s_i^2w_i^3][1 - \frac{c\rho\sigma_u}{x_i^2\sigma_v}]^2 + \tfrac{1}{2}[s_i^2w_i^2 - w_i][\frac{c\rho\sigma_u}{x_i^2\sigma_v^3}]\}$$

$$\frac{\partial^2\Phi}{\partial\sigma_u^2\partial\sigma_v^2} = \sum_{i=1}^{N}\{[w_i^2 - 2s_i^2w_i^3][\frac{c^2}{x_i^4} - \frac{c\rho\sigma_v}{x_i^2\sigma_u}][1 - \frac{c\rho\sigma_u}{\sigma_v x_i^2}] - \tfrac{1}{2}[S_i^2w_i^2-w_i][\frac{c\rho}{\sigma_v\sigma_u x_i^2}]\}$$

$$\frac{\partial^2\Phi}{\partial\rho\partial\sigma_v^2} = \sum_{i=1}^{N}\{[w_i^2 - 2s_i^2w_i^3][\frac{-2c\sigma_u\sigma_v}{x_i^2}][1 - \frac{c\rho\sigma_u}{x_i^2\sigma_v}] - [s_i^2w_i^2 - w_i][\frac{c\sigma_u}{\sigma_v x_i^2}]\}$$

$$\frac{\partial^2\Phi}{\partial\rho\partial\sigma_u^2} = \sum_{i=1}^{N}\{[w_i^2 - 2s_i^2w_i^3][\frac{-2c\sigma_u\sigma_v}{x_i^2}][\frac{c^2}{x_i^4} - \frac{c\rho\sigma_v}{x_i^2\sigma_u}] - [s_i^2w_i^2 - w_i][\frac{c\sigma_v}{\sigma_u x_i^2}]\}$$

$$\frac{\partial^2\Phi}{\partial^2\sigma_u^2} = \sum_{i=1}^{N}\{[w_i^2 - 2s_i^2w_i^3][\frac{c^2}{x_i^4} - \frac{c\rho\sigma_v}{x_i^2\sigma_u}]^2 + \tfrac{1}{2}[s_i^2w_i^2 - w_i][\frac{c\rho\sigma_v}{x_i^2\sigma_u^3}]\}$$

$$\frac{\partial^2\Phi}{\partial^2\rho} = \sum_{i=1}^{N}\{[w_i^2 - 2s_i^2w_i^3][\frac{-2c\sigma_u\sigma_v}{x_i^2}]^2\}$$

APPENDIX B : Program Listings

B.1 Mainline programme

1DUA10:[TFOXDR.WORK]VARCOMP.FOR;14

```
      double precision x(3),f,g(3),w(30),bl(3),bu(3),y(30),s(30)
        common k,y,s
      integer iw(5)
      open(unit=21,file='varcomp.dat',status='old',readonly)
        do 1 i=1,30
      read(21,210,end=98)y(i),s(i)
    1 write(6,699)y(i),s(i)
  699 format(2x,2(1x,f12.6))
   98 k=i-1
  210 format(f6.0,2x,f12.0)
      n=3
      liv=5
      lw=30
      ifail=1
      ibound=0
      bl(1)=1e-6
      bl(2)=1e-6
      bl(3)=-0.99
      bu(1)=1e6
      bu(2)=1e6
      bu(3)=0.99
        write(6,610)
  610   format(//2x,'Enter initial estimates :')
      write(6,650)
  650   format(//2x,'Sigma-squared V=',$)
      read(5,500)x(1)
      write(6,651)
  651   format(//2x,'Sigma-squared U=',$)
      read(5,500)x(2)
      write(6,652)
  652 format(//2x,'Rho=',$)
      read(5,500)x(3)
  500             format(f12.0)
      call e04laf(n,ibound,bl,bu,x,f,g,iw,liv,w,lw,ifail)
      if(ifail.ne.0)write(6,600)ifail
      if(ifail.eq.1)go to 99
  600 format(//2x,'Error exit type ',i3,' see NAG documentation')
      write(6,601)f
      write(6,602)(x(j),j=1,n)
      write(6,603)(g(j),j=1,n)
  601 format(//2x,'Function value on exit is ',f12.6)
  602             format(//2x,'at the point ',3f9.4)
  603 format(//2x,'The corresponding gradient is' /15x,3f12.4)
   99 stop
      end
```

1DUA10:[TFOXDR.WORK]VARCOMP.FOR;14

PROGRAM SECTIONS

Name	Bytes	Attributes
0 $CODE	607	PIC CON REL LCL SHR EXE RD NOWRT LONG
1 $PDATA	284	PIC CON REL LCL SHR NOEXE RD NOWRT LONG
2 $LOCAL	476	PIC CON REL LCL NOSHR NOEXE RD WRT QUAD
3 $BLANK	484	PIC OVR REL GBL SHR NOEXE RD WRT LONG
Total Space Allocated	1851	

ENTRY POINTS

Address	Type	Name
0-00000000		VARCOMP$MAIN

VARIABLES

Address	Type	Name	Address	Type	Name	Address	Type	Name	Address	Type	Name
2-00000168	R*8	F	**	I*4	I	2-00000180	I*4	IBOUND	2-0000017C	I*4	IFAIL
**	I*4	J	3-00000000	I*4	K	2-00000174	I*4	LIW	2-00000178	I*4	LW
2-00000170	I*4	N									

RRAYS

Address	Type	Name	Bytes	Dimensions
2-00000120	R*8	BL	24	(3)
2-00000138	R*8	BU	24	(3)
2-00000018	R*8	G	24	(3)
2-00000150	I*4	IW	20	(5)
3-000000F4	R*8	S	240	(30)
2-00000030	R*8	W	240	(30)
2-00000000	R*8	X	24	(3)
3-00000004	R*8	Y	240	(30)

BELS

Address	Label	Address	Label	Address	Label	Address	Label	Address	Label	Address	Label
**	1	0-00000081	98	0-00000258	99	1-00000019	210'	1-0000007E	500'	1-00000082	600'
1-000000B3	601'	1-000000D7	602'	1-000000F0	603'	1-00000022	610'	1-00000042	650'	1-0000005A	651'
1-00000072	652'	1-0000000C	699'								

```
$1$DUA10:[TFOXDR.WORK]VARCOMP.FOR;14

FUNCTIONS AND SUBROUTINES REFERENCED

 Type  Name            Type  Name

       E04LAF                FOR$OPEN


COMMAND QUALIFIERS

 FOR VARCOMP/LIST

 /CHECK=(NOBOUNDS,OVERFLOW,NOUNDERFLOW)
 /DEBUG=(NOSYMBOLS,TRACEBACK)
 /STANDARD=(NOSYNTAX,NOSOURCE_FORM)
 /SHOW=(NOPREPROCESSOR,NOINCLUDE,MAP,NODICTIONARY,SINGLE)
 /WARNINGS=(GENERAL,NODECLARATIONS,NOULTRIX)
 /CONTINUATIONS=19  /NOCROSS_REFERENCE  /NOD_LINES  /NOEXTEND_SOURCE  /F77
 /NOG_FLOATING  /I4  /NOMACHINE_CODE  /OPTIMIZE


COMPILATION STATISTICS

 Run Time:          0.58 seconds
 Elapsed Time:      1.64 seconds
 Page Faults:       574
 Dynamic Memory:    355 pages
```

B.2 Subroutine FUNCT2 : Evaluation of function and gradient vector B.2

```
$1$DUA10:[TFOXDR.WORK]FUNCT2.FOR;13

        subroutine funct2(n,xc,fc,gc)
        common k,y,s
        double precision gc(n),xc(n),fc,y(30),s(30),w,z,x1,x2,x3,x4
c       write(6,600)xc(1),xc(2),xc(3)
        fc=0
        gc(1)=0
        gc(2)=0
        gc(3)=0
        do 1 i=1,k
        z=y(i)**2
        w=xc(1)+xc(2)*(1800/z)**2-(3600/z)*xc(3)*(xc(1)*xc(2))**0.5
        w=1/w
        fc=fc+s(i)*w-log(w)
        x1=1-xc(3)*1800*((xc(2)/xc(1))**0.5)/z
        x2=(1800/z)**2-xc(3)*1800*((xc(1)/xc(2))**0.5)/z
        x3=-(3600/z)*(xc(1)*xc(2))**0.5
        x4=w-s(i)*w**2
        gc(1)=gc(1)+x1*x4
        gc(2)=gc(2)+x2*x4
        gc(3)=gc(3)+x3*x4
c       write(6,600)z,w,y(i),s(i),fc,x1,x2,x3,x4,gc(1),gc(2),gc(3)
  600   format(/2x,3(3x,f12.6))
    1   continue
c       write(6,620)gc(1),gc(2),gc(3)
  620   format(2x,3f12.6)
          return
        end
```

PROGRAM SECTIONS

Name	Bytes	Attributes
0 $CODE	340	PIC CON REL LCL SHR EXE RD NOWRT LONG
2 $LOCAL	112	PIC CON REL LCL NOSHR NOEXE RD WRT QUAD
3 $BLANK	484	PIC OVR REL GBL SHR NOEXE RD WRT LONG
Total Space Allocated	936	

ENTRY POINTS

Address	Type	Name
0-00000000		FUNCT2

VARIABLES

Address	Type	Name	Address	Type	Name	Address	Type	Name	Address	Type	Name
AP-0000000C@	R*8	FC	2-00000018	I*4	I	3-00000000	I*4	K	AP-00000004@	I*4	N
2-00000000	R*8	W	2-00000010	R*8	X1	**	R*8	X2	**	R*8	X3
**	R*8	X4	2-00000008	R*8	Z						

```
_$DUA10:[TFOXDR.WORK]FUNCT2.FOR;13

ARRAYS

   Address  Type  Name              Bytes  Dimensions

AP-00000010@ R*8  GC                  **   (*)
 3-000000F4  R*8  S                  240   (30)
AP-00000008@ R*8  XC                  **   (*)
 3-00000004  R*8  Y                  240   (30)


LABELS

   Address   Label      Address   Label      Address   Label

     **      1            **      600'         **      620'


FUNCTIONS AND SUBROUTINES REFERENCED

 Type  Name

  R*8  MTH$DLOG


COMMAND QUALIFIERS

 FOR FUNCT2/LIST

 /CHECK=(NOBOUNDS,OVERFLOW,NOUNDERFLOW)
 /DEBUG=(NOSYMBOLS,TRACEBACK)
 /STANDARD=(NOSYNTAX,NOSOURCE_FORM)
 /SHOW=(NOPREPROCESSOR,NOINCLUDE,MAP,NODICTIONARY,SINGLE)
 /WARNINGS=(GENERAL,NODECLARATIONS,NOULTRIX)
 /CONTINUATIONS=19  /NOCROSS_REFERENCE  /NOD_LINES  /NOEXTEND_SOURCE  /F77
 /NOG_FLOATING  /I4  /NOMACHINE_CODE  /OPTIMIZE


COMPILATION STATISTICS

 Run Time:          0.49 seconds
 Elapsed Time:      1.28 seconds
 Page Faults:       554
 Dynamic Memory:    330 pages
```

B.3 Subroutine HESS2 : Evaluation of Hessian Matrix


```
$1$DUA10:[TFOXDR.WORK]HESS2.FOR;5

      subroutine hess2(n,xc,heslc,lh,hesdc)
      common k,y,s
      double precision xc(n),heslc(lh),hesdc(n),x1,x2,x3,x4,x5
      double precision x6,x7,x8,x9,x10,x11,y(30),s(30),v,z
      heslc(1)=0
      heslc(2)=0
      heslc(3)=0
      hesdc(1)=0
      hesdc(2)=0
      hesdc(3)=0
      do 1 i=1,k
      z=y(i)**2
      v=xc(1)+xc(2)*(1800/z)**2-(3600/z)*xc(3)*(xc(1)*xc(2))**0.5
      v=1/v
      x1=2*(v**3)*s(i)-v**2
      x2=v-s(i)*v**2
      x3=(1800/z)**2-((xc(1)/xc(2))**0.5)*1800*xc(3)/z
      x4=1-(1800*xc(3)/z*(xc(2)/xc(1))**0.5)
      x5=-1800*xc(3)/((xc(1)*xc(2)**0.5)*z)
      x6=-(3600/z)*(xc(1)*xc(2))**0.5
      x8=-1800/z*(xc(2)/xc(1))**0.5
      x7=x8
      x9=-1800/z*(xc(1)/xc(2))**0.5
      heslc(1)=heslc(1)+(x1*x3*x4+0.5*x2*x5)
      heslc(2)=heslc(2)+(x1*x6*x4+x2*x7)
      heslc(3)=heslc(3)+(x1*x6*x3+x2*x9)
      x10=-(1800*xc(3)*xc(2)**0.5)/(z*xc(1)**1.5)
      x11=-(1800*xc(3)*xc(1)**0.5)/(z*xc(2)**1.5)
      hesdc(1)=hesdc(1)+(x1*x4**2-0.5*x2*x10)
      hesdc(2)=hesdc(2)+(x1*x3**2-0.5*x2*x11)
      hesdc(3)=hesdc(3)+(x1*x6**2)
    1 continue
      return
      end
```

```
$DUA10:[TFOXDR.WORK]HESS2.FOR;5

PROGRAM SECTIONS

  Name                                        Bytes   Attributes

 0 $CODE                           649   PIC CON REL LCL   SHR   EXE    RD NOWRT LONG
 2 $LOCAL                          176   PIC CON REL LCL NOSHR NOEXE    RD   WRT QUAD
 3 $BLANK                          484   PIC OVR REL GBL   SHR NOEXE    RD   WRT LONG

  Total Space Allocated           1309


NTRY POINTS

  Address   Type  Name

 0-00000000       HESS2


RIABLES

  Address   Type  Name             Address   Type  Name             Address    Type  Name             Address    Type  Name

 2-00000030  I*4  I              3-00000000  I*4  K             AP-00000010@ I*4  LH            AP-00000004@ I*4  N
     **      R*8  W              2-00000000  R*8  X1                 **      R*8  X10                **      R*8  X11
 2-00000008  R*8  X2             2-00000010  R*8  X3             2-00000018  R*8  X4                 **      R*8  X5
 2-00000020  R*8  X6                 **      R*8  X7                 **      R*8  X8                 **      R*8  X9
 2-00000028  R*8  Z


RRAYS

  Address     Type  Name               Bytes  Dimensions

AP-00000014@  R*8  HESDC                **   (*)
AP-0000000C@  R*8  HESLC                **   (*)
 3-000000F4   R*8  S                   240   (30)
AP-00000008@  R*8  XC                   **   (*)
 3-00000004   R*8  Y                   240   (30)


ABELS

  Address   Label

    **       1
```

```
1$DUA10:[TFOXDR.WORK]HESS2.FOR;5

OMMAND QUALIFIERS

 FOR HESS2/LIST

 /CHECK=(NOBOUNDS,OVERFLOW,NOUNDERFLOW)
 /DEBUG=(NOSYMBOLS,TRACEBACK)
 /STANDARD=(NOSYNTAX,NOSOURCE_FORM)
 /SHOW=(NOPREPROCESSOR,NOINCLUDE,MAP,NODICTIONARY,SINGLE)
 /WARNINGS=(GENERAL,NODECLARATIONS,NOULTRIX)
 /CONTINUATIONS=19  /NOCROSS_REFERENCE  /NOD_LINES  /NOEXTEND_SOURCE  /F77
 /NOG_FLOATING  /I4  /NOMACHINE_CODE  /OPTIMIZE


MPILATION STATISTICS

 Run Time:           0.69 seconds
 Elapsed Time:       1.96 seconds
 Page Faults:        566
 Dynamic Memory:     358 pages
```

REFERENCES

Bard, Y. (1974) *Nonlinear parameter estimation.* Academic Press, 1974.

Barker, D.R. and Diana, L. (1974) *Simple method for fitting data when both variables have uncertainties.* A.J.P. 42, 224-227.

Berkson, J. (1950) *Are there two regressions?* Am. Statistical Assoc. 1950, 164-180.

Berkson, J. (1969) *Estimation of a linear function for a calibration line consideration of a recent proposal.* Technometrics. 11, (4) 649-660.

Brown, J.P. (1982) *Multivariate calibration.* J.R.Statist. Soc. (B) 44, (3) 287-321.

Clutton-Brock, M. (1967) *Likelihood distributions for estimating functions when both variables are subject to error.* Technometrics. 9, (2) 261-269.

Fox, D.R. (1985) *A statistical appraisal of vehical speed determination from airborne observation.* Int. Report. Dept. Maths & Stats., Curtin Univ. 1985.

Fox, D.R. (1987) *Components of variance estimation in a calibration problem where both variables are subject to error.* Tech. Report 3/87, Dept. Maths & Stats., Curtin Univ.

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.* JASA. 54, (177) 173-205.

Krutchkoff, R.G. (1967) *Classical and Inverse regression methods of calibration.* Technometrics. 9, (3) 425-439.

Lechner, J.A. et. al. (1982) *An implementation of the Scheffe approach to calibration using spline functions, illustrated by a pressure-volume calibration.* Technometrics. 24, (3) 229-234.

Lwin, T. and Martiz, J.S. (1980) *A note on the problem of statistical calibration.* App. Statist. 29, (2) 135-141.

Lwin, T. and Maritz, J.S. (1982) *An analysis of the linear calibration controversy from the perspective of compound estimation.* Technometrics. 24, (3) 235-242.

Madansky, A. (1959) *The fitting of straight lines when both variables are subject to error.* JASA. 54, (177) 173-205.

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

NAG, (1984) *NAG programme documentation.* Numerical Algorithms Group, Oxford, U.K.

Naes, T. (1985) *Multivariate Calibration when the error covariance matrix is structured.* Technometrics. 27, (3) 301-311.

Oden, A. (1973) *Simultaneous confidence intervals in inverse linear regression.* Biometrika. 60, (2) 339-343.

Oman, S.D. (1982) *Analyzing residuals in calibration problems.* Technometrics. 26, (4) 347-353.

Oman, S.D. and Wax, Y. (1984). *Estimating fetal age by ultrasound measurements : an example of multivariate calibration.* Biometrics. 40, 947-960.

Shukla, G.K. (1972) *On the problem of calibration.* Technometrics. 14, (3) 547-553.

Sjostrom, M. and Wold, S. (1983) *A multivariate calibration problem in analytical chemistry solved by partial least-squares models in latent variables.* Analytica Chimica Acta. 150, 61-70.

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

Spiegelman, C.H. (1984) *An iterative calibration curve procedure.* J. of Research of Nat. Bureau of Standards. 89, (2) 187-192.

Spiegelman, C.H. (1984) *A new statistic for detecting influential observations in a Scheffe type calibration curve.* Austral. J. Statist. 26, (3) 290-297.

Spiegelman, C.H. and Studden, W. (1980) *Design aspects of Scheffe's calibration theory using linear splines.* J. of Research of Nat. Bureau of Standards. 85 (4) 295-304.

Wood, J.T. (1982) *Estimating the age of an animal : An application of multivariate calibration.* Proc. International Biometrics Conf. (1982). 117-121.