

A STATISTICAL THEORY OF CALIBRATION¹

Dedicated to Jerzy Neyman

BY HENRY SCHEFFÉ

University of California, Berkeley

The kind of calibration problem considered may be roughly described as follows: There are two related quantities \mathcal{U} and \mathcal{V} such that \mathcal{U} is relatively easy to measure and \mathcal{V} relatively difficult, requiring more effort or expense; furthermore the error in a measurement of \mathcal{V} is negligible compared with that for \mathcal{U} . A distinguishing feature of the problem is, that from a single calibration experiment, where measurements are made on a number of pairs $(\mathcal{U}, \mathcal{V})$, we wish subsequently to estimate the unknown values of \mathcal{V} corresponding to a very large number of measurements of \mathcal{U} . The problem is solved by a procedure of interval estimation, whose operating characteristic is expressed in terms of a reformulation of the law of large numbers. Some idea of the contents of the article may be obtained from the table of contents.

TABLE OF CONTENTS

1. General idea of the solution of the calibration problem	1
2. Assumptions 1-4	3
3. The calibration chart	4
4. The estimation sets $I(u_i, \mathcal{E})$	5
5. Calculation of $\Pr\{\mathcal{E}_i \mathcal{E}\}$	6
6. Choice of $w(v, c)$	8
7. The case of known σ	10
8. Calculation of c_1 and c_2	11
9. Proof that the sets $I(u_i, \mathcal{E})$ are intervals; Assumption 5	26
10. Resumé of construction and use of calibration chart; use of tables	26
11. The use of linear calibration curves	29
12. Discussion of the assumptions	31
Appendix A. Extension to a case of unequal variance	35
Appendix B. Frequency interpretation	36

1. General idea of the solution of the calibration problem. (Many portions of this article are enclosed in parentheses, like this one. You should treat these with your attitude toward footnotes. The parentheses are boldface to facilitate skipping to the end.) If you are interested only in the mathematical derivation of the theory you may skip the material in parentheses and also Section 10. At the other extreme, if you have no interest in the mathematical derivation you may read the text up to Condition 1 in Section 4, then skip to Section 10, and perhaps omit Appendix B.

The kind of calibration problem treated in this article is characterized by a

Received February 16, 1971; revised September 30, 1971.

¹ This research was partially supported by the Office of Naval Research, Contract Nonr N00014-69-A-0200-1038.



statistical relation between two quantities \mathcal{U} and \mathcal{V} such that \mathcal{U} is relatively easy to measure and \mathcal{V} relatively difficult, requiring more effort or expense. (My treatment seems to have little connection with the previous work on the problem included in a review by E. J. Williams [9]. Not included there is an important article by Lieberman, Miller, and Hamilton [4]: My treatment shares with theirs the use of two specified probabilities, as in Wilks' tolerance intervals [10]. Their two methods are limited to the case of linear calibration curves, and the generally more efficient method of the two is based on the application of the Bonferroni inequality to a confidence band for a regression line and a confidence interval for a standard deviation. Their "Bonferroni intervals" are compared with mine in Section 10.) If a measurement u of \mathcal{U} is made when \mathcal{V} has the value v , it is assumed for v in an interval $[v^{(1)}, v^{(2)}]$ that u is a normal random variable with mean $m(v, \beta) = \sum_{j=1}^p \beta_j g_j(v)$ and variance σ^2 , where the $\{\beta_j\}$ are unknown parameters and the $\{g_j(v)\}$ are known functions. In a calibration experiment \mathcal{E} , n measurements U_1, \dots, U_n of \mathcal{U} are made at known values V_1, \dots, V_n of \mathcal{V} . (It may be argued that it suffices to know V_i with an error small compared with $\sigma/m_v(V_i, \beta)$, where $m_v(v, \beta) = \partial m(v, \beta)/\partial v$.) We will denote by \mathcal{E} also the outcome of the calibration experiment. After the calibration experiment, measurements u_i of \mathcal{U} are made at unknown values v_i of \mathcal{V} , for $i = 1, 2, \dots$. Finally, it is assumed that the random variables $U_1, \dots, U_n, u_1, u_2, \dots$ are independent. A more detailed statement of the assumptions may be found in Section 2. The problem is to construct and interpret interval estimates of the v_i .

("Measurements" of the values of \mathcal{V} need not necessarily be possible in the usual sense; for example, suppose \mathcal{V} is the concentration of a certain chemical in a "sample." In some cases one might get an accurate but difficult measurement of \mathcal{V} in a given sample by separation and weighing, but in other cases it might happen that, given a sample with an unknown value of \mathcal{V} , an accurate measurement is impossible or practically so. The V_i might then be known concentrations in samples specially prepared for the calibration experiment with known amounts of the pure chemical. We might then say that the values V_i are measurements on \mathcal{V} but the values v_i cannot be measured.)

First I shall explain the idea of the proposed statistical procedure in very general terms with no details to divert attention from the probabilistic structure of the inferences. For each measurement u_i , an interval estimate is made of the corresponding unknown v_i . This interval will depend on u_i and also on the outcome \mathcal{E} of the calibration experiment; we shall denote it by $I(u_i, \mathcal{E})$. The construction of the intervals $I(u_i, \mathcal{E})$ is described in Section 4. For each v_i , the statement \mathcal{S}_i is made that $v_i \in I(u_i, \mathcal{E})$. In terms of a relation called "in the long run greater than or equal," defined in the next paragraph, we seek a procedure for which for given α and δ , and for every possible sequence of constants $\{v_i\}$, the probability is $\geq 1 - \delta$ that the proportion of statements \mathcal{S}_i that are true is in the long run $\geq 1 - \alpha$.

If the probabilities of a sequence $\{\mathcal{E}_i\}$ of independent events satisfy $\Pr\{\mathcal{E}_i\} \geq \pi$ we shall say “the proportion of events \mathcal{E}_i that occur is in the long run $\geq \pi$.” The frequency interpretation of this statement is discussed in Appendix B.

Now let \mathcal{S}_i be the event that the above statement \mathcal{S}_i is true. In the conditional distribution, given \mathcal{E} , the events

$$\mathcal{S}_i: v_i \in I(u_i, \mathcal{E})$$

are independent, since the u_i are. Hence if for all sequences of constants $\{v_i\}$, $\Pr\{\mathcal{S}_i | \mathcal{E}\} \geq 1 - \alpha$, then in the conditional distribution, given \mathcal{E} , the proportion of correct statements \mathcal{S}_i is in the long run $\geq 1 - \alpha$. We note that $\Pr\{\mathcal{S}_i | \mathcal{E}\}$ is a function of \mathcal{E} , and also in general of v_i , β , and σ . Let $\mathcal{R}(\beta, \sigma)$ be a set in the \mathcal{E} -space where $\Pr\{\mathcal{S}_i | \mathcal{E}\} \geq 1 - \alpha$ for all v_i . If $\Pr\{\mathcal{E} \in \mathcal{R}(\beta, \sigma)\} \geq 1 - \delta$ for all β, σ , then our estimation procedure has the property italicized above.

2. Assumptions 1–4. We now proceed to a more detailed statement of the assumptions.

ASSUMPTION 1. The measurements $U_1, \dots, U_n, u_1, u_2, \dots$ on \mathcal{U} are statistically independent, and their distributions depend respectively on the values $V_1, \dots, V_n, v_1, v_2, \dots$ of \mathcal{V} .

Let $v^{(1)} = \min_i V_i, v^{(2)} = \max_i V_i$. The closed interval $T = [v^{(1)}, v^{(2)}]$ will be called the “calibration interval” of \mathcal{V} ; it is a known interval, assumed to be nondegenerate.

ASSUMPTION 2. U_i is $N(m(V_i, \beta), \sigma^2)$, and if $v_i \in T, u_i$ is $N(m(v_i, \beta), \sigma^2)$, where for $v \in T$,

$$(1) \quad m(v, \beta) = \sum_{j=1}^p \beta_j g_j(v),$$

$\beta = (\beta_1, \dots, \beta_p)$, the $\{\beta_j\}$ are unknown parameters, and the $\{g_j(v)\}$ are known functions with continuous derivatives. The $n \times p$ matrix whose i, j -element is $g_j(V_i)$ is of rank p .

The variance σ^2 is unknown, unless the contrary is stated. A case of unequal variance is treated in Appendix A. We shall make no assumption about the “fine structure” of the distribution of u_i for v_i outside the calibration interval T ; the following will suffice.

ASSUMPTION 3. For all constants $C^{(1)}, C^{(2)}$,

$$\Pr\{u_i \leq C^{(1)} \text{ when } v_i < v^{(1)}\} \geq \Pr\{u_i \leq C^{(1)} \text{ when } v_i = v^{(1)}\}$$

and

$$\Pr\{u_i \geq C^{(2)} \text{ when } v_i > v^{(2)}\} \geq \Pr\{u_i \geq C^{(2)} \text{ when } v_i = v^{(2)}\}.$$

The estimates of the $\{\beta_j\}$ are calculated from the calibration experiment by the method of least squares. The calculation is assumed to be made as usual by minimizing

$$(2) \quad \sum_{i=1}^n [U_i - \sum_{j=1}^p \hat{\beta}_j g_j(V_i)]^2$$

by varying the $\{\hat{\beta}_j\}$ freely over the entire Euclidean p -space of β , and may lead to $\hat{\beta}$ lying outside the parameter space to which later assumptions restrict β . Nevertheless, $\hat{\beta} = (\hat{\beta}_1, \dots, \hat{\beta}_p)$ will denote the estimates obtained by the unrestricted minimization process, and they will then be unique by the rank assumption on the matrix $(g_j(V_i))$, and jointly normal with means $(\beta_1, \dots, \beta_p)$ and covariance matrix $(\sigma^2 b_{ij})$, where the b_{ij} are known constants. We denote by $\hat{\sigma}^2$ an estimate of σ^2 satisfying

ASSUMPTION 4. $\hat{\sigma}^2$ is independent of $\hat{\beta}$ and distributed as $\sigma^2 \chi^2(\nu)/\nu$, where $\chi^2(\nu)$ denotes a chi-square variable with ν df.

The estimate $\hat{\sigma}^2$ may be pooled from other experiments besides the calibration experiment. In that case it will not be a statistic formed from \mathcal{E} ; we now change the notation so that henceforth \mathcal{E} will denote the $(p+1)$ -dimensional statistic $(\hat{\beta}, \hat{\sigma})$. (This is one of the details glossed over in the general description of the procedure in Section 1. I remark here also that in practice $\nu \geq n - p$, but in our derivations we will not restrict ν , because we shall see that this suggests approximations for certain constants we shall need to calculate.)

3. The calibration chart. I shall describe the construction of a "calibration chart" and explain how the intervals $I(u_i, \mathcal{E})$ are read from the chart; their analytical definition which this implies will be clear. *On the calibration chart we shall take u as abscissa and v as ordinate.* The regression curve of u on v has the equation $u = m(v, \beta)$, where $m(v, \beta)$ is given by (1); the fitted regression curve has the equation $u = m(v, \hat{\beta})$, where

$$(3) \quad m(v, \hat{\beta}) = \sum_{j=1}^p \hat{\beta}_j g_j(v).$$

The *calibration curve*, defined for $v \in T$, is identical with the fitted regression curve, but we think of it as determining the inverse of the fitted regression function $u = m(v, \hat{\beta})$, namely, v as a function of u and $\hat{\beta}$. (The analytic inversion may not be practicable, and is not needed, as the calibration curve is easily plotted from its abscissa $u = m(v, \hat{\beta})$. A similar remark applies to the upper and lower calibration curves defined below. As indicated in Fig. 1, for any u_i in the "calibration interval" $[u^{(1)}, u^{(2)}]$ of \mathcal{U} , where $u^{(k)} = m(v^{(k)}, \hat{\beta})$, we can read from the calibration curve a point estimate \hat{v}_i of the corresponding v_i , namely, \hat{v}_i is the ordinate at $u = u_i$. For this reason the user of the calibration chart is probably interested at least as much in the calibration curve as in our intervals $I(u_i, \mathcal{E})$. Although I do not treat the problem of point estimation of the v_i , I remark that the interval $I(u_i, \mathcal{E})$, no matter how small we made it by increasing α or δ , would always contain \hat{v}_i in its interior. It is obviously essential for practical reasons that \hat{v}_i be a single-valued function of u_i . Necessary for this is the condition that the derivative $m_v(v, \hat{\beta})$ be of constant sign, which we may as well assume positive, so $m_v(v, \hat{\beta}) \geq 0$; sufficient is the following condition.)

CONDITION 1. For all $v \in T$, $m_v(v, \hat{\beta}) > 0$, where $m_v(v, \hat{\beta}) = \partial m(v, \hat{\beta})/\partial v$.

Under Assumption 2 and our definition of the $\{\hat{\beta}_j\}$ as unrestricted minimizing values, there is a positive probability of Condition 1 being violated for all β, σ ; however, it will be a consequence of the later Assumption 5 that this probability is negligible. This negligible probability must then be subtracted from the nominal probability $1 - \hat{\delta}$ for the procedure. (We might think of “negligible” meaning “small compared with $\hat{\delta}$,” perhaps of the order of $\hat{\delta}^2$. I remark that a negligible probability of violation of Condition 1 implies that the regression function has the property $m_v(v, \beta) > 0$ for all $v \in T$, since $m_v(v, \hat{\beta})$ is normal with mean $m_v(v, \beta)$: Suppose $m_v(v_0, \beta) \leq 0$ for a particular $v_0 \in T$. Then $\Pr\{m_v(v, \hat{\beta}) \leq 0 \text{ for some } v \in T\} \geq \Pr\{m_v(v_0, \hat{\beta}) \leq 0\} \geq \frac{1}{2}$, which is not negligible. The condition $m_v(v, \beta) > 0$ for all $v \in T$ restricts β to the convex set which is the intersection of the half-spaces $\{\beta : \sum_{j=1}^p \beta_j g_j'(v) > 0\}$ for $v \in T$. Some discussion of the Assumption 5 implying all this is offered in Section 12.)

The intervals $I(u_i, \mathcal{E})$ may be read from the upper and lower calibration curves, which are constructed as follows: They are to the left and right of the calibration curve at a horizontal distance $\hat{\sigma}w(v, c)$, where $w(v, c)$ is a positive function, to be specified later, and depending on one or more constants c , which will be determined to give the desired α and $\hat{\delta}$. The upper calibration curve has the equation

$$(4) \quad u = m^{(1)}(v, \mathcal{E}, c) \equiv m(v, \hat{\beta}) - \hat{\sigma}w(v, c),$$

the lower calibration curve,

$$(5) \quad u = m^{(2)}(v, \mathcal{E}, c) \equiv m(v, \hat{\beta}) + \hat{\sigma}w(v, c),$$

and they are to be constructed only between the horizontal lines $v = v^{(1)}$ and $v = v^{(2)}$. The intercepts with these lines might be denoted by

$$u^{(hk)} = m^{(h)}(v^{(k)}, \mathcal{E}, c) \quad (h, k = 1, 2),$$

but it will be more convenient to write

$$(6) \quad u^{11} = u^{(21)}, \quad u^{12} = u^{(12)}, \quad u^{01} = u^{(11)}, \quad u^{02} = u^{(22)},$$

as indicated in Fig. 1. The intervals $[u^{11}, u^{12}]$ and $[u^{01}, u^{02}]$ may be called respectively the “inner and outer calibration intervals” of \mathcal{U} .

4. The estimation sets $I(u_i, \mathcal{E})$. We shall define the sets $I(u_i, \mathcal{E})$, and calculate $\Pr\{\mathcal{E}_i | \mathcal{E}\} = \Pr\{v_i \in I(u_i, \mathcal{E}) | \mathcal{E}\}$, without assuming they are intervals. (Sets which are not necessarily intervals might be acceptable in some other problems, as in Fieller’s problem of estimating a ratio, or in estimating the slope of an unknown line when the fitted line is nearly vertical, but never in a calibration problem. The construction of the $I(u_i, \mathcal{E})$ will give intervals if the $m^{(h)}(v, \mathcal{E}, c)$ defined by (4) and (5) are strictly increasing functions of v , and a sufficient condition for this will be the following, since under our eventual choice of $w(v, c)$ the $m^{(h)}(v, \mathcal{E}, c)$ will have continuous derivatives with respect to v .)

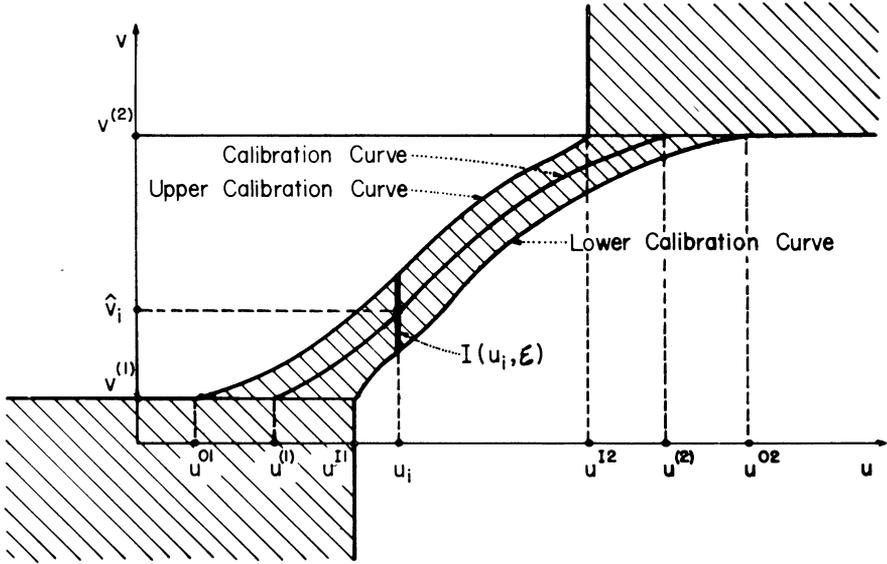


FIG. 1. Schematic diagram of calibration chart and the set $\mathcal{A}(\mathcal{E}, c)$, shaded.

CONDITION 2. For all $v \in T$, and for $h = 1, 2$, $m_v^{(h)}(v, \mathcal{E}, c) > 0$.

Again we can only require that this condition be violated with negligible probability, and this will follow from our later Assumption 5. We note, by adding for $h = 1, 2$ in Condition 2, that it implies Condition 1.

After the sets $I(u_i, \mathcal{E})$ have been shown to be intervals, they will have the following structure in terms of the calibration chart: If u_i is in the inner calibration interval $[u^{I1}, u^{I2}]$, the statement \mathcal{S}_i is that the corresponding v_i lies between the ordinates of the lower and upper calibration curves at $u = u_i$, that is, $I(u_i, \mathcal{E})$ is the interval between these curves on the line $u = u_i$, as shown for this case in Fig. 1. In every case the interval $I(u_i, \mathcal{E})$ will be closed. If u_i does not lie inside the inner calibration interval, the treatment of the edge effects is somewhat finicky: There are four cases, and the corresponding statements \mathcal{S}_i are given in Section 10; in every case the interval is the vertical section at $u = u_i$ of the shaded set $\mathcal{A}(\mathcal{E}, c)$ in Fig. 1.

Now without assuming they are intervals I will define the one-dimensional sets $I(u_i, \mathcal{E})$ from the following two-dimensional set $\mathcal{A} = \mathcal{A}(\mathcal{E}, c)$ in the u, v -plane. For the set $\mathcal{A}(\mathcal{E}, c)$, shaded in Fig. 1, the upper-or-left boundary is the upper calibration curve extended by the half of the line $u = u^{I2}$ above $v = v^{(2)}$ and the half of the line $v = v^{(1)}$ to the left of $u = u^{01}$; the lower-or-right boundary is the lower calibration curve extended by the half of the line $u = u^{I1}$ below $v = v^{(1)}$ and the half of the line $v = v^{(2)}$ to the right of $u = u^{02}$. Both boundaries are included in \mathcal{A} . The sets $I(u_i, \mathcal{E})$ are the vertical sections of \mathcal{A} by the lines $u = u_i$.

5. Calculation of $\Pr\{\mathcal{E}_i | \mathcal{E}\}$. Since for any v_i , v_i is in the vertical section of

\mathcal{A} by $u = u_i$ if and only if u_i is in the horizontal section of \mathcal{A} by $v = v_i$, we may calculate $\Pr\{\mathcal{E}_i | \mathcal{E}\}$ as the conditional probability, given \mathcal{E} , that u_i is in the horizontal section of $\mathcal{A}(\mathcal{E}, c)$ by $v = v_i$, which horizontal section will always be an interval.

If $v_i \in T$,

$$(7) \quad \Pr\{\mathcal{E}_i | \mathcal{E}\} = \Pr\{m^{(1)}(v_i, \mathcal{E}, c) \leq u_i \leq m^{(2)}(v_i, \mathcal{E}, c) | \mathcal{E}\},$$

where u_i is $N(m(v_i, \beta), \sigma^2)$ by Assumption 2. If in the double inequality we subtract $m(v_i, \beta)$ and divide by σ , we get by (4), (5), (1), (3),

$$\Pr\{\mathcal{E}_i | \mathcal{E}\} = \Pr\{m(v_i, \tilde{\beta}) - \tilde{\sigma}w(v_i, c) \leq z(v_i) \leq m(v_i, \tilde{\beta}) + \tilde{\sigma}w(v_i, c) | \mathcal{E}\},$$

where

$$\tilde{\beta} = (\tilde{\beta}_1, \dots, \tilde{\beta}_p), \quad \tilde{\beta}_j = (\hat{\beta}_j - \beta_j)/\sigma, \quad \tilde{\sigma} = \hat{\sigma}/\sigma,$$

and

$$z(v_i) = [u_i - m(v_i, \beta)]/\sigma$$

is $N(0, 1)$. Hence sufficient conditions for $\Pr\{\mathcal{E}_i | \mathcal{E}\} \geq 1 - \alpha$ are that

$$(8) \quad m(v_i, \tilde{\beta}) - \tilde{\sigma}w(v_i, c) \leq -z_\alpha,$$

$$(9) \quad m(v_i, \tilde{\beta}) + \tilde{\sigma}w(v_i, c) \geq z_\alpha,$$

where z_α is the two-tailed α -point of $N(0, 1)$.

If $v_i > v^{(2)}$,

$$\Pr\{\mathcal{E}_i | \mathcal{E}\} = \Pr\{u_i \geq u^{I2} | \mathcal{E}\}.$$

Now if \mathcal{E} is given, u^{I2} is fixed, and we may use Assumption 3 to conclude

$$\Pr\{\mathcal{E}_i | \mathcal{E}\} \geq \Pr\{u_i \geq u^{I2} \text{ when } v_i = v^{(2)} | \mathcal{E}\}.$$

Using the definition (6) of u^{I2} , and calculating as in the preceding paragraph, we then find

$$\Pr\{\mathcal{E}_i | \mathcal{E}\} \geq \Pr\{z(v^{(2)}) \geq m(v^{(2)}, \tilde{\beta}) - \tilde{\sigma}w(v^{(2)}, c) | \mathcal{E}\},$$

and hence that $\Pr\{\mathcal{E}_i | \mathcal{E}\} \geq 1 - \frac{1}{2}\alpha$ if condition (8) is satisfied for $v_i = v^{(2)}$. Similarly we find for $v_i < v^{(1)}$ that $\Pr\{\mathcal{E}_i | \mathcal{E}\} \geq 1 - \frac{1}{2}\alpha$ if condition (9) is satisfied for $v_i = v^{(1)}$. (It would be correct later to make the following stronger assertion than that italicized at the end of the third paragraph of Section 1: Divide the statements \mathcal{S}_i into two classes, the first, those made when u_i lies inside the inner calibration interval $[u^{I1}, u^{I2}]$, and the second, when u_i lies outside. Then the probability is $\geq 1 - \delta$ that the proportion of true statements in the first class is in the long run $\geq 1 - \alpha$, while that in the second class is in the long run $\geq 1 - \frac{1}{2}\alpha$. However, it hardly seems worthwhile to complicate in this way the statement of the operating characteristic of the method, especially since in practice relatively few of the u_i should be outside the inner calibration interval.)

We have now established that for all v_i , $\Pr\{\mathcal{E}_i | \mathcal{E}\} \geq 1 - \alpha$ if for all $v \in T$

$$m(v, \tilde{\beta}) - \tilde{\sigma}w(v, c) \leq -z_\alpha \quad \text{and} \quad m(v, \tilde{\beta}) + \tilde{\sigma}w(v, c) \geq z_\alpha,$$

or

$$(10) \quad |m(v, \tilde{\beta})| \leq \bar{\sigma}w(v, c) - z_\alpha \quad \text{for all } v \in T.$$

(At this point the following remarks may be helpful: The random variables involved in the inequality (10) are $(\hat{\beta}, \hat{\sigma}) = \mathcal{E}$, as is seen by writing it $|\sum_j(\hat{\beta}_j - \beta_j)g_j(v)| \leq \hat{\sigma}w(v, c) - \sigma z_\alpha$. The preceding probability calculations were conditional for fixed \mathcal{E} , but the succeeding ones are based on the distribution of the random variable \mathcal{E} .)

6. Choice of $w(v, c)$. In our remaining calculations we shall encounter the function $S(v)$ defined as

$$(11) \quad S(v) = \text{SD}(m(v, \tilde{\beta}))/\sigma \quad \text{for } v \in T,$$

where “SD” denotes “the standard deviation of.” A formula for $S(v)$ is given below by (13).

We now derive an inequality for $|m(v, \tilde{\beta})|$ by matrix calculus: In this paragraph boldface type will indicate a vector or matrix. The p -dimensional column vector $\tilde{\beta}$ is $N(\mathbf{0}, \mathbf{B})$, where $\mathbf{B} = (b_{ij})$. Hence $x = (\tilde{\beta}^T \mathbf{B}^{-1} \tilde{\beta})^{\frac{1}{2}}$ is $\chi(p)$. Since \mathbf{B} is positive definite there exists a nonsingular \mathbf{A} such that $\mathbf{A}^T \mathbf{B} \mathbf{A} = \mathbf{I}$, the identity matrix; hence $\mathbf{B} = \mathbf{A}^{-1T} \mathbf{A}^{-1}$ and $\mathbf{B}^{-1} = \mathbf{A} \mathbf{A}^T$. Write $m(v, \tilde{\beta}) = \mathbf{g}^T \tilde{\beta}$ for all $v \in T$, where the components of $\mathbf{g} = \mathbf{g}(v)$ are $g_j(v)$. Then applying the Schwarz inequality to $m(v, \tilde{\beta}) = (\mathbf{A}^{-1} \mathbf{g})^T (\mathbf{A}^T \tilde{\beta})$, we get

$$(12) \quad |m(v, \tilde{\beta})| \leq \|\mathbf{A}^{-1} \mathbf{g}\| \cdot \|\mathbf{A}^T \tilde{\beta}\|,$$

where for a vector \mathbf{y} , $\|\mathbf{y}\| = (\mathbf{y}^T \mathbf{y})^{\frac{1}{2}}$. Now $S(v) = \text{SD}(\mathbf{g}^T \tilde{\beta}) = (\mathbf{g}^T \mathbf{B} \mathbf{g})^{\frac{1}{2}}$, which we record here in nonmatrix notation for later use,

$$(13) \quad S(v) = [\sum_{j,k=1}^p b_{jk} g_j(v) g_k(v)]^{\frac{1}{2}},$$

and then continue:

$$S(v) = (\mathbf{g}^T \mathbf{A}^{-1T} \mathbf{A}^{-1} \mathbf{g})^{\frac{1}{2}} = \|\mathbf{A}^{-1} \mathbf{g}\|.$$

Also,

$$x = (\tilde{\beta}^T \mathbf{B}^{-1} \tilde{\beta})^{\frac{1}{2}} = (\tilde{\beta}^T \mathbf{A} \mathbf{A}^T \tilde{\beta})^{\frac{1}{2}} = \|\mathbf{A}^T \tilde{\beta}\|.$$

Substitution in (12) gives the desired inequality

$$(14) \quad |m(v, \tilde{\beta})| \leq x S(v) \quad \text{for all } v \in T,$$

where $x = (\tilde{\beta}^T \mathbf{B}^{-1} \tilde{\beta})^{\frac{1}{2}}$ is $\chi(p)$, does not depend on v , and is independent of $\bar{\sigma}$ since $\tilde{\beta}$ is, while $\bar{\sigma}$ is $\nu^{-\frac{1}{2}} \chi(\nu)$.

From here on our probability calculations will be made not in the $(p+1)$ -dimensional probability space of $\mathcal{E} = (\hat{\beta}, \hat{\sigma})$ but in the induced 2-dimensional probability space of $(x, \bar{\sigma})$, where $(x, \bar{\sigma})$ has the distribution just stated. Because of (14), the condition (10) will be satisfied if

$$(15) \quad x \leq \frac{\bar{\sigma}w(v, c) - z_\alpha}{S(v)} \quad \text{for all } v \in T.$$

The distribution in the $(x, \bar{\sigma})$ -space does not depend on the unknown (β, σ) , but the probability of the set where (15) is satisfied still depends on v , so let us “inf it out.” Let

$$(16) \quad G(\bar{\sigma}, c, \alpha) = \inf_{v \in T} \frac{\bar{\sigma}w(v, c) - z_\alpha}{S(v)}.$$

The event (15) will happen for all $v \in T$ if

$$(17) \quad x \leq G(\bar{\sigma}, c, \alpha).$$

Denote by \mathcal{R} the set in the $(x, \bar{\sigma})$ -space where (17) is satisfied. We shall now choose $w(v, c)$ so that $\Pr\{(x, \bar{\sigma}) \in \mathcal{R}\} = 1 - \delta$. Then the probability will be $\geq 1 - \delta$ that

$$\Pr\{\mathcal{E}_i | \mathcal{E}\} \geq 1 - \alpha \quad \text{for all } v_i.$$

(On contemplating (16) it seems to me that the only hope of obtaining relatively simple and general results in calculating $\Pr\{(x, \bar{\sigma}) \in \mathcal{R}\}$ is to choose $w(v, c)$ to be a linear function of $S(v)$.)

Let us choose

$$(18) \quad w(v, c) = c_1 + c_2 S(v),$$

where c_1 and c_2 are constants, so

$$G(\bar{\sigma}, c, \alpha) = c_2 \bar{\sigma} + \inf_{v \in T} \frac{c_1 \bar{\sigma} - z_\alpha}{S(v)}.$$

Then $G(\bar{\sigma}, c, \alpha)$ is easily calculated in terms of the constants

$$S_1 = \inf_{v \in T} S(v), \quad S_2 = \sup_{v \in T} S(v);$$

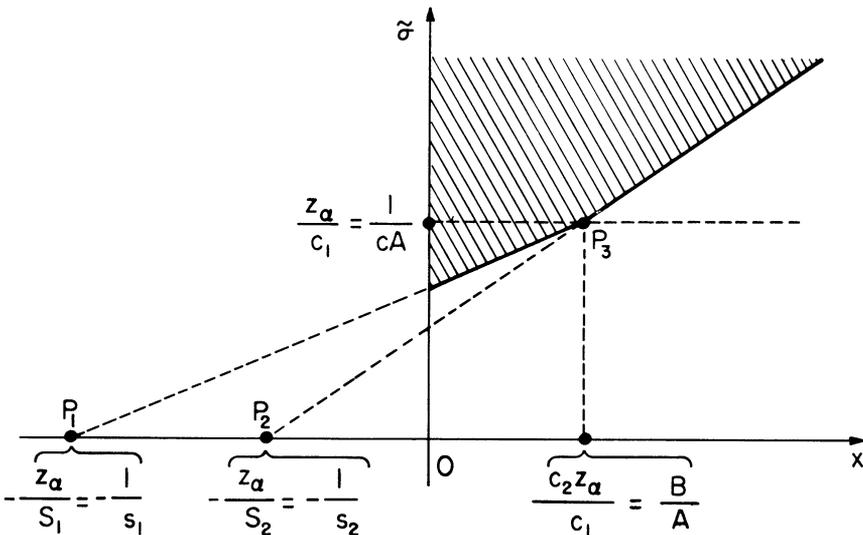


FIG. 2. The shaded set is \mathcal{R} , the domain of integration for $P(c)$.

it is given by the formulas which are the right members of (19) below. (The needed constants S_i are easily obtained graphically, as explained in Section 10.) if we substitute this for $G(\bar{\sigma}, c, \alpha)$ in (17), (17) becomes

$$(19) \quad \begin{aligned} x &\leq (c_2 + S_1^{-1}c_1)\bar{\sigma} - S_1^{-1}z_\alpha && \text{if } \bar{\sigma} \leq z_\alpha/c_1, \\ &\leq (c_2 + S_2^{-1}c_1)\bar{\sigma} - S_2^{-1}z_\alpha && \text{if } \bar{\sigma} \geq z_\alpha/c_1, \end{aligned}$$

and we now have the problem of calculating constants c_1 and c_2 for which the probability of the set \mathcal{R} defined by (19) is $1 - \delta$. \mathcal{R} is the shaded set in Fig. 2. It is completely determined by the three points P_1, P_2, P_3 whose coordinates are labeled on the axes in Fig. 2, where we use the first set of labels at this stage.

7. The case of known σ . Explicit formulas for the constants c_1 and c_2 which make the probability of (19) equal to $1 - \delta$ are easily found in the important case where σ is known. The case of known σ can be included in our treatment by taking $\nu = \infty$ and putting $\bar{\sigma} = \sigma$ in the equations (4) and (5) of the upper and lower calibration curves. All the probability in the $x, \bar{\sigma}$ -plane is then concentrated on the line $\bar{\sigma} = 1$, where x is $\chi(p)$. In the case $c_1 \leq z_\alpha$, the event (19) is defined by the upper inequality in (19) with $\bar{\sigma} = 1$; we then want

$$\Pr\{|x| \leq c_2 + S_1^{-1}(c_1 - z_\alpha)\} = 1 - \delta.$$

This will be true if $c_2 + S_1^{-1}(c_1 - z_\alpha) = \chi_\delta(p)$, the upper δ -point of $\chi(p)$. Similarly, we treat the case $c_1 \geq z_\alpha$. The result is

$$(20) \quad \begin{aligned} c_2 &= \chi_\delta(p) + S_1^{-1}(z_\alpha - c_1) && \text{if } c_1 \leq z_\alpha, \\ &= \chi_\delta(p) + S_2^{-1}(z_\alpha - c_1) && \text{if } c_1 \geq z_\alpha. \end{aligned}$$

To choose a pair (c_1, c_2) satisfying the condition (20) we shall minimize $w(v, c)$. (It is natural to try to minimize the vertical distance between the upper and lower calibration curves, which, for u_i in the inner calibration interval, is the length of the interval $I(u_i, \mathcal{E})$. Since this appears analytically very messy, I minimize instead the horizontal distance between the curves, namely $2\sigma w(v, c)$.) From (18) and (20),

$$\begin{aligned} w(v, c) &= c_1[1 - S_1^{-1}S(v)] + [\chi_\delta(p) + S_1^{-1}z_\alpha]S(v) && \text{if } c_1 \leq z_\alpha, \\ &= c_1[1 - S_2^{-1}S(v)] + [\chi_\delta(p) + S_2^{-1}z_\alpha]S(v) && \text{if } c_1 \geq z_\alpha. \end{aligned}$$

For any fixed v , the graph of $w(v, c)$ as a function of c_1 thus consists of two linear pieces with slope ≤ 0 for $c_1 \leq z_\alpha$, and slope ≥ 0 for $c_1 \geq z_\alpha$, and is continuous at $c_1 = z_\alpha$. It follows that $w(v, c)$ will be minimized, uniformly for $v \in T$, by taking $c_1 = z_\alpha$. Then $c_2 = \chi_\delta(p)$. We choose these values of c_1 and c_2 , so that

$$(21) \quad w(v, c) = z_\alpha + \chi_\delta(p)S(v)$$

in the case of known σ . (The following intuitive interpretation may be given to the two terms in $\sigma w(v, c) = \sigma z_\alpha + \sigma \chi_\delta(p)S(v)$: The horizontal distance $\sigma w(v, c)$

allows for the variation of a measurement u_i made at v , about the unknown regression curve of u on v ; the first term may be attributed to the variation it would still have if the regression curve were known, and the second term to the variation of the fitted regression curve about the unknown regression curve. This interpretation leads to a simplification in the case below where σ is unknown, by suggesting a relation to adopt between c_1 and c_2 which reduces their calculation to that of a single constant.)

8. Calculation of c_1 and c_2 . Returning now to the case where σ is unknown, we have to calculate values of the constants c_1 and c_2 which make the probability of the event (19) equal to $1 - \delta$; they will depend on S_1 and S_2 as well as on α, δ, p, ν . (The above strategy of minimizing $w(v, c)$ over all pairs (c_1, c_2) making the probability of (19) equal to $1 - \delta$ seems to me to lead to a calculation hopelessly complicated even for a computer, since the probability of (19) is given by a double integral. However, the values of c_1 and c_2 we shall adopt will agree with (21) in the limiting case for $\nu \rightarrow \infty$. We can get some new information by considering the limiting case as $S_2 \rightarrow 0$ for fixed ν . We might imagine achieving this by an appropriate sequence of calibration experiments for which $n \rightarrow \infty$, but for this hypothetical situation we keep the same estimate $\hat{\sigma}^2$ so that ν is not $\geq n - p$, but fixed.)

Let us consider what happens in Fig. 2 when S_2 and hence $S_1 \rightarrow 0$. Then the point P_1 stays to the left of P_2 while P_2 recedes indefinitely along the negative x -axis. The two linear pieces of the right boundary of the shaded region thus rotate about the fixed point P_3 into a limiting position which is the half of the horizontal line $\bar{\sigma} = z_\alpha/c_1$ to the right of the $\bar{\sigma}$ -axis. The limit of the probability of the set (19) is thus $\Pr\{\bar{\sigma} \geq z_\alpha/c_1\}$. This will equal $1 - \delta$ if z_α/c_1 is the lower δ -point of $\bar{\sigma}$, or $c_1 = z_\alpha A$, where

$$(22) \quad A = A(\delta, \nu) = \nu^{1/2}/\chi_{1-\delta}(\nu),$$

and $\chi_{1-\delta}(\nu)$ denotes the lower δ -point of $\chi(\nu)$. Thus the limit of c_1 as $S_2 \rightarrow 0$ is $z_\alpha A$. (In the case of small S_2 the first term of $\hat{\sigma}w(v, c) = \hat{\sigma}c_1 + \hat{\sigma}c_2S(v)$ has an intuitive interpretation similar to that below (21), but $\hat{\sigma}c_1 \rightarrow z_\alpha A\hat{\sigma}$ now also allows for the fact that σ is unknown, $A\hat{\sigma}$ being an upper $(1 - \delta)$ -confidence limit for σ . If we try to give an interpretation similar to the above to the second term, it suggests directly, or by analogy with my S -method of multiple comparison [6, Section 3.5], at least for small S_2 , the approximation of c_2 by the upper δ -point of $x/\bar{\sigma}$, namely $B = [pF_\delta(p, \nu)]^{1/2}$. I considered at first defining $c_1 = z_\alpha A$ and then calculating c_2 to make the probability of (19) equal to $1 - \delta$, hoping that B could be used as a first approximation in the computer calculation of c_2 , but I decided that this calculation would become very unstable as $S_2 \rightarrow 0$, since c_2 then multiplies a quantity in (18) that approaches zero, while c_1 remains fixed, so that large changes in c_2 would cause little change in $w(v, c)$, the quantity of interest. This led me to the following formulation employing a single constant c whose calculation should not be subject to this instability.

The preceding intuitive interpretations also led me to hope that c would not vary much from unity as S_1 and S_2 are varied.)

We shall now determine a constant c so that the probability of the event (19) is $1 - \delta$ if

$$c_1 = cz_\alpha A \quad \text{and} \quad c_2 = cB,$$

where A is given by (22), and B by

$$(23) \quad B = B(\delta, p, \nu) = [pF_\delta(p, \nu)]^{\frac{1}{2}},$$

where $F_\delta(p, \nu)$ denotes the upper δ -point of the F distribution with p and ν df. The form of $w(v, c)$ is then

$$(24) \quad w(v, c) = c[Az_\alpha + BS(v)],$$

the event (19) becomes

$$(25) \quad \begin{aligned} x \leq c(B + S_1^{-1}z_\alpha A)\bar{\sigma} - S_1^{-1}z_\alpha & \quad \text{if } \bar{\sigma} \leq (cA)^{-1}, \\ \leq c(B + S_2^{-1}z_\alpha A)\bar{\sigma} - S_2^{-1}z_\alpha & \quad \text{if } \bar{\sigma} \geq (cA)^{-1}, \end{aligned}$$

and its probability, which we will denote by $P(c)$, depends also on the six quantities $\alpha, \delta, p, \nu, S_1, S_2$. We can suppress the explicit dependence on α by introducing, for the purposes of computation and tabulation, the quantities

$$(26) \quad s_i = S_i/z_\alpha \quad (i = 1, 2).$$

Then $P(c)$ depends on the five quantities δ, p, ν, s_1, s_2 , besides c , as indicated by the second labeling in Fig. 2 of the coordinates of the points P_1, P_2, P_3 , with A depending on δ and ν through (22), B on δ, p , and ν through (23), while the probability density of $\bar{\sigma}$ depends on ν , and that of x , on p .

As c increases from 0 to ∞ , P_1 and P_2 remain fixed in Fig. 2 while P_3 descends the vertical line $x = B/A$, its ordinate decreasing from ∞ to 0, and hence the probability $P(c)$ in \mathcal{R} increases from 0 to 1. The value of c which makes $P(c) = 1 - \delta$ will be a function of δ, p, ν, s_1, s_2 , say $c = \varphi(\delta, p, \nu, s_1, s_2)$. Knowledge of the following limiting behavior of $\varphi(\delta, p, \nu, s_1, s_2)$ is useful for the tabulation of φ : (i) $\varphi \rightarrow 1$ as $\nu \rightarrow \infty$, for all δ, p, s_1, s_2 ; (ii) $\varphi \rightarrow 1$ as s_2 and hence $s_1 \rightarrow \infty$, for all δ, p, ν ; (iii) $\varphi \rightarrow 1$ as s_2 and hence $s_1 \rightarrow 0$, for all δ, p, ν . (i) follows from using in (24) the limits 1 for A and $\chi_\delta(p)$ for B as $\nu \rightarrow \infty$, so $w(v, c) \rightarrow c[z_\alpha + \chi_\delta(p)S(v)]$, whence $c \rightarrow 1$ from (21). (ii) follows from using in $c = c_1/(z_\alpha A)$ the limit $z_\alpha A$ for c_1 , established below (22), as $s_2 \rightarrow 0$. (iii) may be deduced from Fig. 2, where, as $s_1 \rightarrow \infty$, the right boundary of \mathcal{R} approaches the ray from 0 through P_3 , on which $x/\bar{\sigma} = cB$, so $P(c) \rightarrow \Pr\{x/\bar{\sigma} \leq cB\} = \Pr\{[pF(p, \nu)]^{\frac{1}{2}} \leq c[pF_\nu(p, \nu)]^{\frac{1}{2}}\} \rightarrow 1 - \delta$ if $c \rightarrow 1$. While the values of s_1 and s_2 in a calibration problem always satisfy $0 < s_1 < s_2$, it is convenient for tabulating purposes to extend the definition of $\varphi(\delta, p, \nu, s_1, s_2)$ to all s_1, s_2 satisfying $0 \leq s_1 \leq s_2$, by taking limits.

The probability $P(c) = P(c, \delta, p, \nu, s_1, s_2)$ of the event (25) may be calculated with the aid of Fig. 2 to be

$$P(c) = Q(c, \delta, p, \nu, s_1) + R(c, \delta, p, \nu, s_2),$$

where

$$(27) \quad Q(c, \delta, p, \nu, s_1) = \int_{[c(A+B s_1)]^{-1}}^{(cA)^{-1}} g(\bar{\sigma}) \{ \int_{\bar{\sigma}}^{L_1(\bar{\sigma})} f(x) dx \} d\bar{\sigma} ,$$

$$(28) \quad R(c, \delta, p, \nu, s_2) = \int_{(cA)^{-1}}^{\infty} g(\bar{\sigma}) \{ \int_{\bar{\sigma}}^{L_2(\bar{\sigma})} f(x) dx \} d\bar{\sigma} ,$$

and the limits of integration $L_i(\bar{\sigma})$ are

$$L_i(\bar{\sigma}) = c(B + s_i^{-1}A)\bar{\sigma} - s_i^{-1} .$$

The probability densities $f(x)$ and $g(\bar{\sigma})$ are

$$f(x) = [2^{(p-2)/2} \Gamma(p/2)]^{-1} x^{p-1} \exp(-x^2/2) ,$$

$$g(\bar{\sigma}) = 2(\nu/2)^{\nu/2} [\Gamma(\nu/2)]^{-1} \bar{\sigma}^{\nu-1} \exp(-\nu\bar{\sigma}^2/2) ,$$

and the constants A and B are given by (22) and (23). The formulas (27) and (28) are obtained from (25) for $0 < s_1 < s_2$, and we extend them to $0 < s_1 \leq s_2$. The value of $P(c)$ for $s_1 = 0$, defined as $\lim P(c)$ as $s_1 \rightarrow 0$, is needed for tabulation. It is calculated by setting $Q = 0$ instead of (27), and retaining the definition of R by (28); this is justified by the limiting behavior of the set \mathcal{R} as $s_1 \rightarrow 0$ in Fig. 2.

The value of c which makes $P(c) = 1 - \delta$ in the case $p = 2$ is given in Table 1, and corrections to this for other values of p , in Table 2. Use of the tables is discussed at the end of Section 10. (I decided, in view of how c would be used in the construction of the calibration chart, to tabulate it to two decimal places. The printing of the table was simplified by tabulating instead of c the integer c^* , where $c = 1 + (.01)c^*$, and c is to two decimals. Next I decided that instead of the solution of $P(c) = 1 - \delta$, I would tabulate the minimum integer c^* for which $P_1(c^*) \geq 1 - \delta$, where $P_1(c^*) = P(1 + (.01)c^*)$. This redefinition, while of some theoretical advantage in ensuring $P(c) \geq 1 - \delta$ for a two-decimal c , is of no practical importance, except that it eliminates the inverse interpolation problem of determining the solution of $P(c) = 1 - \delta$ from calculated values of $P(c)$: The CDC 6400 computer was programmed to move through integer values of c^* until it found two successive integer values for which $P_1(c^*)$ bracketed $1 - \delta$. The result is the same as though the inverse interpolation problem were solved and the resulting c then rounded up, instead of rounded off, to two decimals. Finally, I decided on another slight modification, which happened to make the table of c^* look a little smoother in some spots: The programmer believed that the computed value of $P_1(c^*)$ should be correct within $\pm .0001$, although the complicated derivation of an error bound was not made. Let $P_2(c^*)$ denote the computed value of $P_1(c^*)$ rounded off to 4 decimals. Then the tabled c^* is the minimum integer for which $P_2(c^*) \geq 1 - \delta$. It should agree within ± 1 with the desired minimum integer c^* for which $P_1(c^*) \geq 1 - \delta$. Besides c^* (actually, c) the computer also printed out $P_2(c^*)$ and $P_2(c^* - 1)$, respectively $\geq 1 - \delta$ and $< 1 - \delta$.)

TABLE 1 continued

		$\delta = .10$								$\delta = .25$																					
		0	.05	.10	.15	.20	.30	.40	.60	.80	1.0	1.5	2.0	3.0	4.0	∞	0	.05	.10	.15	.20	.30	.40	.60	.80	1.0	1.5	2.0	3.0	4.0	∞
s_2	ν																														
0	4	0															0														
	8	0															0														
	15	0															0														
	30	0															0														
	60	0															0														
	120	0															0														
	∞	0															0														
.05	4	1	I														1	0													
	8	1	I														1	0													
	15	1	I														1	I													
	30	1	I														1	I													
	60	1	I														1	I													
	120	1	I														1	I													
	∞	0	0														0	0													
.10	4	2	0	I													2	1	I												
	8	2	0	I													2	1	I												
	15	1	I	I													2	0	I												
	30	1	I	I													2	0	I												
	60	1	I	I													2	0	I												
	120	1	I	I													2	0	I												
	∞	0	0	0													0	0	0												
.15	4	3	1	I													3	2	0	I											
	8	2	0	I													3	1	0	I											
	15	2	0	I													3	1	0	I											
	30	2	0	I													3	1	0	I											
	60	2	0	I													3	1	0	I											
	120	2	0	I													2	1	0	I											
	∞	0	0	0													0	0	0												
.20	4	3	1	0	I												4	2	1	0	I										
	8	3	1	0	I												4	2	1	0	I										
	15	2	0	I													4	2	1	0	I										
	30	2	0	I													3	2	0	I											
	60	2	0	I													3	1	0	I											
	120	2	0	I													2	1	0	I											
	∞	0	0	0													0	0	0												
	4	3	1	0	I												4	2	1	0	I										
	8	3	1	0	I												4	2	1	0	I										
	15	2	0	I													3	2	0	I											
	30	2	0	I													3	1	0	I											
	60	2	0	I													3	1	0	I											
	120	2	0	I													3	1	0	I											
	∞	0	0	0													0	0	0												
$s_1 =$	0	.05	.10	.15	.20	.30	.40	.60	.80	1.0	1.5	2.0	3.0	4.0	∞	0	.05	.10	.15	.20	.30	.40	.60	.80	1.0	1.5	2.0	3.0	4.0	∞	

TABLE 1 continued

		$\delta = .10$										$\delta = .25$																									
s_2	ν	0	.05	.10	.15	.20	.30	.40	.60	.80	1.0	1.5	2.0	3.0	4.0	∞	0	.05	.10	.15	.20	.30	.40	.60	.80	1.0	1.5	2.0	3.0	4.0	∞						
.30	4	4	2	1	0	0	1	3	5	7	10	15	20	30	40	∞	5	4	2	1	0	0	1	2	2	2	2	2	2	2	∞						
	8	4	1	0	0	0	0	0	0	0	0	0	0	0	0	0	5	3	2	1	0	0	1	2	2	2	2	2	2	2	∞						
	15	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	2	1	0	0	0	1	2	2	2	2	2	2	2	2	∞					
	30	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	2	1	0	0	0	1	2	2	2	2	2	2	2	2	2	∞				
	60	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	2	1	0	0	0	1	2	2	2	2	2	2	2	2	2	∞				
.40	∞	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	∞				
	4	5	3	2	1	0	0	0	0	0	0	0	0	0	0	0	6	5	3	2	1	0	0	1	1	1	1	1	1	1	1	1	∞				
	8	4	2	1	0	0	0	0	0	0	0	0	0	0	0	0	6	4	2	1	0	0	0	1	1	1	1	1	1	1	1	1	1	∞			
	15	4	2	1	0	0	0	0	0	0	0	0	0	0	0	0	5	4	2	1	0	0	0	1	1	1	1	1	1	1	1	1	1	∞			
	30	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	5	3	2	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	∞		
.60	60	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	4	3	2	1	0	0	0	1	1	1	1	1	1	1	1	1	1	∞			
	120	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	4	3	2	1	0	0	0	1	1	1	1	1	1	1	1	1	1	∞			
	∞	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	∞			
	4	7	5	3	2	1	0	0	0	0	0	0	0	0	0	0	8	6	5	4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	∞		
	8	6	4	2	1	0	0	0	0	0	0	0	0	0	0	0	7	5	4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	∞		
.80	15	5	3	1	0	0	0	0	0	0	0	0	0	0	0	0	6	5	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	∞		
	30	4	2	1	0	0	0	0	0	0	0	0	0	0	0	0	6	4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	∞		
	60	4	2	1	0	0	0	0	0	0	0	0	0	0	0	0	6	4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	∞		
	120	4	2	1	0	0	0	0	0	0	0	0	0	0	0	0	6	4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	∞		
	∞	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	∞		
1.0	4	8	6	4	3	2	1	0	0	0	0	0	0	0	0	0	9	8	7	6	5	4	3	2	1	0	0	0	0	0	0	0	0	0	∞		
	8	6	4	3	2	1	0	0	0	0	0	0	0	0	0	0	8	7	6	5	4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	∞	
	15	5	3	2	1	0	0	0	0	0	0	0	0	0	0	0	7	6	5	4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	∞	
	30	4	2	1	0	0	0	0	0	0	0	0	0	0	0	0	7	5	4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	∞	
	60	4	2	1	0	0	0	0	0	0	0	0	0	0	0	0	6	5	4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	∞	
1.0	120	4	2	1	0	0	0	0	0	0	0	0	0	0	0	0	5	4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	∞	
	∞	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	∞	
	4	8	7	5	4	3	2	1	0	0	0	0	0	0	0	0	10	8	7	6	5	4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	∞
	8	7	5	4	2	1	0	0	0	0	0	0	0	0	0	0	9	7	6	5	4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	∞
	15	6	5	3	2	1	0	0	0	0	0	0	0	0	0	0	8	6	5	4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0

$\delta = .25$

$\delta = .10$

TABLE 1 continued

		$\delta = .01$										$\delta = .05$														
		0	.05	.10	.15	.20	.30	.40	.60	.80	1.0	0	.05	.10	.15	.20	.30	.40	.60	.80	1.0	1.5	2.0	3.0	4.0	∞
$s_1 =$	ν																									
1.5	4	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	8	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	15	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	30	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	60	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
2.0	4	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	8	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	15	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	30	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	60	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
3.0	4	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	8	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	15	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	30	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	60	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
4.0	4	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	8	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	15	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	30	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	60	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
∞	4	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	8	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	15	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	30	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	60	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
∞	4	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	8	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	15	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	30	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	60	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3

$\delta = .05$

$\delta = .01$

TABLE 2
 Corrections to c^* of Table 1 for $p = 1, 4, 6$. The first (second, third) digit
 of the three-figure number is the correction for $p = 1$ (4, 6)

		$\delta = .01$										$\delta = .05$																			
		0	.05	.10	.15	.20	.30	.40	.60	.80	1.0	1.5	2.0	3.0	4.0	∞	0	.05	.10	.15	.20	.30	.40	.60	.80	1.0	1.5	2.0	3.0	4.0	∞
s_2	s_1																														
0	4	000																													
	8	000																													
	15	000																													
	30	000																													
	60	000																													
	120	000																													
	∞	000																													
.05	4	000	100																												
	8	000	101																												
	15	000	111																												
	30	000	011																												
	60	000	011																												
	120	000	001																												
	∞	000	000																												
.10	4	100	011	101																											
	8	000	011	112																											
	15	000	011	112																											
	30	000	101	112																											
	60	000	011	112																											
	120	000	011	012																											
	∞	000	000	000																											
.15	4	000	101	111	211																										
	8	011	101	122	212																										
	15	000	101	112	212																										
	30	000	111	111	112																										
	60	000	101	012	112																										
	120	000	100	111	112																										
	∞	000	000	000	000																										
.20	4	100	011	111	112	212																									
	8	000	111	112	123	223																									
	15	011	111	112	223	223																									
	30	000	011	112	123	212																									
	60	000	101	111	112	112																									
	120	000	101	011	101	012																									
	∞	000	000	000	000	000																									
	4	000	101	011	111	112																									
	8	011	101	111	211	123																									
	15	000	011	112	112	212																									
	30	000	101	011	112	122																									
	60	000	011	101	101	111																									
	120	000	000	000	000	000																									
	∞	000	000	000	000	000																									
s_2	s_1																														
		$\delta = .01$										$\delta = .05$																			
		0	.05	.10	.15	.20	.30	.40	.60	.80	1.0	1.5	2.0	3.0	4.0	∞	0	.05	.10	.15	.20	.30	.40	.60	.80	1.0	1.5	2.0	3.0	4.0	∞

TABLE 2 continued

		$\delta = .01$										$\delta = .05$																				
$s_1 =$	s_2	0	.05	.10	.15	.20	.30	.40	.60	.80	1.0	1.5	2.0	3.0	4.0	∞	0	.05	.10	.15	.20	.30	.40	.60	.80	1.0	1.5	2.0	3.0	4.0	∞	
.30	4	011	011	101	112	212	223										000	100	011	111	112	222										
	8	011	101	121	223	224											000	101	111	122	122	212	222									
	15	000	101	122	212	223	234										000	001	111	111	112	223										
	30	100	101	111	123	123	212										000	001	011	012	112	122	122									
	60	000	001	101	111	112	212										000	000	011	012	111	112										
	120	000	011	111	012	111	111										000	000	100	101	012	111										
	∞	000	000	000	000	000	000										000	000	000	000	000	000	000									
.40	4	000	111	112	112	212	223	323									100	100	011	111	111	222	222									
	8	000	011	112	212	212	223	324									000	011	101	112	122	212	223	223								
	15	011	111	112	212	123	223	234									000	100	101	112	122	222	223	223								
	30	000	001	112	112	123	123	223									000	001	011	111	112	112	112	123								
	60	000	101	112	112	112	112	112									000	100	101	111	012	111	112									
	120	000	101	101	111	101	012	012									000	011	011	011	111	101	012									
	∞	000	000	000	000	000	000	100									000	000	000	000	000	000	000									
.60	4	000	011	101	201	112	212	223	323								000	000	100	011	111	211	123	323								
	8	000	011	112	112	222	223	234	334								000	100	111	101	211	212	223	323								
	15	000	111	112	212	212	212	223	223								000	100	101	012	211	122	123	312								
	30	011	011	112	112	122	212	123	212								011	000	011	111	112	112	212									
	60	000	100	111	112	111	112	112	112								000	011	101	111	012	111	012	112								
	120	000	011	011	111	101	012	011	011								000	000	100	011	101	011	011	111								
	∞	000	000	000	000	000	000	000	000								000	000	000	000	000	000	000	000								
.80	4	000	010	100	200	112	212	312	323	323							001	000	100	100	011	111	211	222	323							
	8	100	111	112	112	212	223	323	234	323							001	011	100	101	201	112	222	322	323							
	15	100	101	111	211	212	212	223	223	223							000	011	100	012	112	112	123	212	223							
	30	000	101	111	112	112	212	222	212								000	011	101	101	111	112	211	112	212							
	60	100	011	101	111	101	112	111	111								011	000	011	011	111	012	111	111	111							
	120	000	000	101	011	011	111	111	111								000	011	011	100	011	111	101	101	100							
	∞	000	000	000	000	000	000	000	000								000	000	000	000	000	000	000	000	000							
1.0	4	001	000	100	111	111	211	211	322	323	334						011	001	111	000	011	101	211	212	322	323						
	8	000	100	111	211	212	223	223	223	323	234						000	000	100	011	111	211	212	312	223	223						
	15	000	011	112	112	122	123	212	223	222	212						000	000	011	111	111	112	112	123	212	212						
	30	000	011	012	112	112	112	112	122	122	122						100	000	011	011	101	111	112	111	122	112						
	60	000	011	011	112	111	111	111	101	012	012						000	000	100	101	011	111	101	012	111	111						
	120	000	011	011	101	101	101	011	011	011	011						000	000	100	001	100	011	011	011	111	101						
	∞	000	000	000	000	000	000	000	000	000	000						000	000	000	000	000	000	000	000	000	000						

TABLE 2 continued

		$\delta = .10$										$\delta = .25$																					
s_2	ν	0	.05	.10	.15	.20	.30	.40	.60	.80	1.0	1.5	2.0	3.0	4.0	∞	0	.05	.10	.15	.20	.30	.40	.60	.80	1.0	1.5	2.0	3.0	4.0	∞		
.30	4	000	011	101	101	101	112										000	011	000	000	011	101											
	8	100	011	101	012	112	212										001	000	000	000	100	011	101										
	15	000	000	011	111	111	122										100	001	000	000	011	101											
	30	000	011	100	012	101	122										001	100	000	000	001	101											
	60	000	000	011	011	111	111										001	000	000	000	000	011											
	120	000	000	100	101	011	101										100	011	000	000	000	011											
.40	∞	000	000	000	000	000	000									000	000	000	000	000	000												
	4	000	000	100	100	101	112	212								001	011	100	000	000	000	012											
	8	000	011	100	011	112	212	123								011	000	001	000	000	011	111											
	15	000	100	101	101	111	112	123								100	011	000	000	000	101	101											
	30	000	000	011	011	012	112	211								001	000	000	000	000	000	100	101										
	60	000	000	001	011	011	101	112								100	011	000	000	000	100	011											
.60	120	000	000	001	011	101	011	011							000	000	000	000	100	000	100												
	∞	000	000	000	000	000	000	000							000	000	000	000	000	000	000												
	4	011	001	000	000	111	112	112	213						011	101	111	011	001	100	001	012											
	8	011	101	000	100	101	112	212	312						111	100	101	000	000	000	011	111											
	15	000	000	000	101	111	122	212	123						100	011	100	100	100	100	000	011	111										
	30	000	000	011	011	011	112	111	112						011	101	011	000	000	000	100	101	111										
.80	60	001	000	000	001	011	101	111						100	100	000	000	011	100	001	011												
	120	000	000	100	001	100	011	101	101					011	011	100	000	000	100	011	100												
	∞	000	000	000	000	000	000	000	000					000	000	000	000	000	000	000	000												
	4	011	001	000	000	000	100	112	212	312				112	022	111	111	101	011	100	100	112											
	8	100	100	000	100	100	111	112	212	222				111	012	111	101	101	000	011	111	111											
	15	011	100	000	100	101	112	211	123	212				101	011	101	100	100	000	000	101	111											
1.0	30	000	011	000	001	011	111	101	111	211				011	011	100	001	011	011	100	101	111											
	60	000	000	011	100	101	011	111	012	111				011	011	011	001	000	000	100	101	100											
	120	011	000	000	011	001	100	011	111	111				001	000	000	000	000	000	100	011	001											
	∞	000	000	000	000	000	000	000	000	000				000	000	000	000	000	000	000	000	000											
	4	101	011	100	000	000	100	111	211	223	312			112	212	112	112	111	011	100	100	011	112										
	8	001	000	001	011	011	111	111	211	212	223			112	111	111	111	111	100	000	100	101	111										
15	101	000	000	011	011	111	112	112	212	122			111	111	111	111	101	011	000	011	101	111											
30	100	000	011	011	100	101	111	112	112	111			111	101	111	001	000	000	000	011	101	011											
60	011	000	000	000	011	101	011	111	101	101			101	101	001	001	100	100	000	011	001	100											
120	000	000	011	000	011	001	100	100	100	011			101	100	000	100	000	000	000	100	011	001											
∞	000	000	000	000	000	000	000	000	000	000			000	000	000	000	000	000	000	000	000	000											

$\delta = .25$

$\delta = .10$

9. Proof that the sets $I(u_i, \mathcal{E})$ are intervals. **Assumption 5.** The only remaining question in the mathematical derivation of the procedure is whether the sets $I(u_i, \mathcal{E})$ are intervals. From (4), (5), and (24),

$$(29) \quad m^{(h)}(v, \mathcal{E}, c) = m(v, \hat{\beta}) + (-1)^h \hat{\sigma} c [Az_\alpha + BS(v)] \quad (h = 1, 2),$$

where $S(v)$ is given by (13). Thus the $m^{(h)}(v, \mathcal{E}, c)$ have continuous derivatives with respect to $v \in T$ since the $g_j(v)$ do by Assumption 2, and our task is reduced to formulating an Assumption 5 which ensures that the probability is negligible that Condition 2 is violated. From (29),

$$(30) \quad m_v^{(h)}(v, \mathcal{E}, c) = m_v(v, \hat{\beta}) + (-1)^h \hat{\sigma} c BS'(v).$$

(The question is whether (30) is positive. On contemplating (30) at first it seemed to me that even though the slope $m_v(v, \hat{\beta})$ of the fitted regression curve were comfortably positive, $S(v)$ might vary so wildly that the contribution of $S'(v)$ to (30) could still drag it to negative values. This possibility is ruled out by the following Lemma, which relates $S'(v)$ to the distribution of $m_v(v, \hat{\beta})$.)

LEMMA. $S'(v) = \rho(v) \text{SD}(m_v(v, \hat{\beta}))/\sigma$, where $\rho(v)$ is the correlation coefficient between $m(v, \hat{\beta})$ and $m_v(v, \hat{\beta})$.

To prove the Lemma we differentiate (13) to get

$$(31) \quad S'(v) = [S(v)]^{-1} \sum_{j,k=1}^p b_{jk} g_j(v) g_k'(v).$$

The double sum in (31) may be recognized as the covariance between $m(v, \hat{\beta})$ and $m_v(v, \hat{\beta})$, divided by σ^2 , and writing this as $\rho(v) \text{SD}(m(v, \hat{\beta})) \text{SD}(m_v(v, \hat{\beta}))/\sigma^2$, and using the definition (11) of $S(v)$, we obtain the Lemma.

Applying the Lemma in (30) we find that a necessary and sufficient condition for both the $m_v^{(h)}(v, \mathcal{E}, c)$ to be positive is that

$$m_v(v, \hat{\beta}) > cB|\rho(v)|\hat{\sigma} \text{SD}(m_v(v, \hat{\beta})),$$

and that hence the probability of Condition 2 being violated will be negligible under

ASSUMPTION 5. The probability is negligible that for some $v \in T$, $m_v(v, \hat{\beta}) \leq cB\hat{\sigma} \text{SD}(m_v(v, \hat{\beta}))$.

(We may notice that $\hat{\sigma} \text{SD}(m_v(v, \hat{\beta}))$ is a statistic, and is in fact the estimate of $\text{SD}(m_v(v, \hat{\beta}))$ formed by replacing σ by $\hat{\sigma}$ in the formula for $\text{SD}(m_v(v, \hat{\beta}))$. Thus Assumption 5 says that there is a negligible probability that for some $v \in T$, $m_v(v, \hat{\beta})$ be negative or fall within cB estimated standard deviations of zero. There is some further discussion of Assumption 5 in Section 12.)

10. Resumé of construction and use of calibration chart; use of tables. The numerical values of the least-squares estimates $\{\hat{\beta}_j\}$, the estimate $\hat{\sigma}^2$, and the coefficients $\{b_{jk}\}$ must be calculated, where b_{jk} is the covariance between $\hat{\beta}_j$ and $\hat{\beta}_k$ divided by σ^2 . The calibration chart is constructed by plotting, with a

horizontal u -axis and a vertical v -axis, the calibration curve $u = m(v, \hat{\beta}) = \sum_{j=1}^p \hat{\beta}_j g_j(v)$, the lower calibration curve $u = m(v, \hat{\beta}) + \hat{\sigma}[c_1 + c_2 S(v)]$, and the upper calibration curve $u = m(v, \hat{\beta}) - \hat{\sigma}[c_1 + c_2 S(v)]$, where the calculation of c_1 , c_2 , and $S(v)$ is reviewed in the next three paragraphs. The three curves are to be drawn only in the horizontal strip which intersects the v -axis in $[v^{(1)}, v^{(2)}]$, the calibration interval of \mathcal{V} , that is, the range of the values of \mathcal{V} used in the calibration experiment. A schematic diagram of the calibration chart is the horizontal strip in Fig. 2 between the lines $v = v^{(1)}$ and $v = v^{(2)}$; the diagram is schematic because in practice the three calibration curves would be closer together. It may be convenient to plot all three curves at the same time, plotting the three abscissas for each value of v . If any of the three curves should fail to be the graph of a single-valued everywhere-increasing function of u the calibration chart cannot be used: If no computational error can be found, nor a gross error in the calibration experiment, my method cannot be applied, presumably because of a violation of Assumption 5, which implies that there is a negligible probability of this disaster. (Assumption 5 might be violated because for some v the slope $m_v(v, \beta)$ of the true regression curve is negative or too close to zero, or because n is too small, causing too large an error in the estimated slope $m_v(v, \hat{\beta})$ or in its estimated standard deviation $\hat{\sigma} \text{SD}(m_v(v, \hat{\beta}))$ appearing in Assumption 5.)

The function $S(v)$ defined by (11) is calculated from (13). Its minimum and maximum values S_1 and S_2 in the interval $[v^{(1)}, v^{(2)}]$ must now be determined: This will generally not be possible analytically, but S_1 and S_2 can be obtained from a graph of $S(v)$; they need not be determined very precisely, because the values of c_1 and c_2 are not very sensitive to changes in S_1 and S_2 . The table of values of $S(v)$ calculated for this graph should be saved for plotting the upper and lower calibration curves.

At this point, if not before, the choice of probability levels α and δ has to be made. The choice of a physical scientist or engineer who determines physical or chemical constants in a form like 3280 ± 40 , where 40 is the "probable error," could be formulated as $\alpha = .50$. Experimenters accustomed to conventional statistical methodology might choose $\alpha = .10$ or $.05$. Next, δ must be chosen; my own choice would usually restrict $\delta < \alpha$. Tables 1 and 2, needed next for determining c allow the four choices $\delta = .01, .05, .10, .25$.

The value of c is found by entering Tables 1 and 2 as explained below, with $s_1 = S_1/z_\alpha$ and $s_2 = S_2/z_\alpha$, where z_α is the *two-tailed* α -point, that is, the upper $\alpha/2$ -point, of the standard normal distribution. Finally, c_1 and c_2 are calculated as

$$(32) \quad c_1 = cz_\alpha \nu^{\frac{1}{2}} [\chi_{1-\delta}^2(\nu)]^{-\frac{1}{2}}, \quad c_2 = c[pF_\delta(p, \nu)]^{\frac{1}{2}},$$

where $\chi_{1-\delta}^2(\nu)$ is the *lower* δ -point of the chi-square distribution with ν df, and $F_\delta(p, \nu)$ is the upper δ -point of the F -distribution with p and ν df. The three calibration curves can now be plotted.

In order to explain the use of the calibration chart we need to define three

nested intervals on the u -axis. The endpoints of the intervals are the abscissas of the six points of intersection of the three calibration curves with the two lines $v = v^{(1)}$ and $v = v^{(2)}$, and we denote them by the notation used in Fig. 1 for the labels on the u -axis: The calibration interval of \mathcal{U} is $[u^{(1)}, u^{(2)}]$, the inner calibration interval is $[u^{I1}, u^{I2}]$, and the outer calibration interval is $[u^{O1}, u^{O2}]$.

Corresponding to an observation u_i falling in the calibration interval of \mathcal{U} , a point estimate \hat{v}_i of v_i may be read as the ordinate of the calibration curve at $u = u_i$.

An interval estimate of v_i is provided by the following statement \mathcal{S}_i : If u_i falls in the inner calibration interval of \mathcal{U} , that is, $u^{I1} \leq u_i \leq u^{I2}$, the statement \mathcal{S}_i is that v_i lies in the closed interval between the ordinates read from the lower and upper calibration curves at $u = u_i$. If u_i falls outside the inner calibration interval—which should not occur often in practice—there are four possibilities, depending on how u_i relates to the outer calibration interval $[u^{O1}, u^{O2}]$. If $u^{O1} \leq u_i < u^{I1}$, \mathcal{S}_i states that v_i is \leq the ordinate of the upper calibration curve; if $u^{I2} < u_i \leq u^{O2}$, \mathcal{S}_i states that v_i is \geq the ordinate of the lower calibration curve; if $u_i < u^{O1}$, \mathcal{S}_i states that $v_i \leq v^{(1)}$; if $u_i > u^{O2}$, \mathcal{S}_i states that $v_i \geq v^{(2)}$. In every case the statement \mathcal{S}_i is that v_i lies in the vertical section by the line $u = u_i$ of the shaded set in Fig. 1.

Some discussion about what can be done if the calibration chart gives intervals for the v_i which are too wide to be useful may be found at the end of Section 12.

The rest of this section explains the use of Tables 1 and 2 to find the value of c . For compactness of the tables the value of the integer c^* is tabled instead of the value of c , where $c = 1 + (.01)c^*$ to two decimals. Steps in the tabled values of the variables s_1 , s_2 , ν with which the tables are entered are sufficiently small so that interpolation will generally not be necessary for the following reasons: The change of c^* between these tabled values of the variables is small considering the way changes of c^* affect the calibration chart, and the larger value of c^* at the two bracketing tabled values of the variable would be used. For example, an increase of 2 in c^* increases the width of the interval estimate of v_i by 2 percent, which would hardly be perceptible on the chart. With very few exceptions the changes $|\Delta c^*| > 2$ occur only in the steps $\nu = 120$ to ∞ , $s_1 = 0$ to $.05$, and $s_2 = 4$ to ∞ , intervals in which the values of ν , s_1 , or s_2 will very rarely lie.

Suppose first that ν has one of the tabled values 4, 8, 15, 30, 60, 120, ∞ . Interpolation on s_1 and s_2 may be avoided because it can be proved for all δ , p , ν that c^* is nonincreasing in s_1 and nondecreasing in s_2 . Hence a conservative value of c^* is obtained by entering with the largest tabled value of $s_1 \leq$ its actual value, and the smallest tabled value of $s_2 \geq$ its actual value. If $p = 2$, the c^* thus read from Table 1 is the value to be used. If $p = 1, 4$, or 6 , Table 1 is first entered in the same way as above, and then the correction to be added to the value of c^* found there is obtained by entering Table 2 in exactly the same

place that Table 1 was entered, and using respectively the first, second or third digit or the 3-digit number listed in Table 2 as the correction. For $p = 3$, use the larger of the values of c^* for $p = 2$ and 4; for $p = 5$, the larger of the values of c^* for $p = 4$ and 6. For $p = 7, 8, 9, 10$, add 1 to the value of c^* found for $p = 6$. (The last prescription is based on a computer calculation of c^* for $p = 10$ and all (s_1, s_2, ν, δ) with s_1 and $s_2 = 0, .1, .3, .8, 2, \infty$ ($s_1 \leq s_2$); $\nu = 2, 15, 120$; $\delta = .01, .05, .10, .25$. Nowhere did the value of c^* for $p = 10$ exceed that for $p = 6$ by more than 1.) Finally, for ν not one of the tabled values use the larger of the c^* found by the above method for the two bracketing tabled values of ν .

(If you nevertheless wish to interpolate on s_1 and s_2 the 4-point bivariate interpolation formula may be used, except when (s_1, s_2) is too close to the diagonal $s_1 = s_2$ to allow a fourth point, in which case the 3-point formula may be used. However, if $s_2 > 4$, a case that should arise very rarely, the 4-point formula fails because one of the tabled bracketing values of s_2 is ∞ ; in this case before using the 4-point formula first transform from s_2 to $1 - (4/s_2)$, since for large s_2 , c is almost linear in $1/s_2$. Graphical interpolation may be used on ν , or on $1/\nu$ if ν is large, and on p .)

11. The use of linear calibration curves. As an illustration I consider how my general method applies in the frequently assumed case of linear regression. It is assumed that the regression curve is a straight line not necessarily passing through the origin, so $p = 2$. All the following equations or inequalities containing the symbol v are for v in the calibration interval $[v^{(1)}, v^{(2)}]$. The form of the regression function is assumed to be

$$m(v, \beta) = a + bv,$$

where β denotes the pair (a, b) . In the least-squares estimates \hat{a}, \hat{b} , and their covariance matrix, we have to replace the x, y of the usual notation by V, U , respectively, and we can then calculate that

$$(33) \quad S(v) = [n^{-1} + k(v - \bar{V})^2]^{\frac{1}{2}},$$

where

$$k = 1/\sum_{i=1}^n (V_i - \bar{V})^2, \quad \bar{V} = \sum_{i=1}^n V_i/n.$$

This seems to be one of the relatively few cases where we can explicitly solve the equation of the calibration curve for v , namely $v = (u - \hat{a})/\hat{b}$, and where we can determine S_1 and S_2 without resorting to numerical or graphical methods. Indeed, by inspection of (33) we see that $S_1 = n^{-\frac{1}{2}}$ and $S_2 = [n^{-1} + kM^2]^{\frac{1}{2}}$, where M is the larger of $\bar{V} - v^{(1)}$ and $v^{(2)} - \bar{V}$. After c has been obtained from Table 1, c_1 and c_2 are given by (32) with $p = 2$, and the upper and lower calibration curves are found to be the two branches of hyperbolas

$$(34) \quad u = \hat{a} + \hat{b}v + (-1)^h \hat{\sigma} \{c_1 + c_2 [n^{-1} + k(v - \bar{V})^2]^{\frac{1}{2}}\}$$

for $h = 1$ and 2, respectively.

A nongraphical application of my method is possible in this case. First I

remark that Condition 2, ensuring that the estimation sets $I(u_i, \mathcal{E})$ be intervals, may be shown to be equivalent to $\hat{b}/\hat{\sigma} > c_2 k M / S_2$. Next we need to solve for v the equations of the three calibration curves. The calibration curve satisfies

$$(35) \quad v = (u - \hat{a})/\hat{b};$$

the upper (lower) calibration curve is given for $h = 1(2)$ by

$$(36) \quad v = \bar{V} + C^{-1}[\hat{b}D_h - (-1)^h \hat{\sigma} c_2 (n^{-1}C + kD_h^2)^{\frac{1}{2}}],$$

where

$$C = \hat{b}^2 - (\hat{\sigma} c_2)^2 k,$$

$$D_h = D_h(u) = u - \hat{a} - \hat{b}\bar{V} - (-1)^h \hat{\sigma} c_1.$$

To use equations (35) and (36) we need the endpoints of the three calibration intervals of \mathcal{U} : For $k = 1, 2$, $u^{(k)} = \hat{a} + \hat{b}v^{(k)}$; $u^{11}(u^{12})$ is found by putting $v = v^{(1)}(v^{(2)})$ in (34) with $h = 2(1)$; $u^{01}(u^{02})$, by putting $v = v^{(1)}(v^{(2)})$ in (34) with $h = 1(2)$. Then for $u^{(1)} \leq u_i \leq u^{(2)}$, the point estimate \hat{v}_i of v_i is given by putting $u = u_i$ in (35). For $u^{01} \leq u_i \leq u^{02}$ ($u^{11} \leq u_i \leq u^{12}$), the upper (lower) endpoint of the interval estimate for v_i is given by putting $u = u_i$ in (36) with $h = 1(2)$.

(For $u_i \in [u^{(1)}, u^{(2)}]$ the "Bonferroni method" of Lieberman, Miller, and Hamilton [4] may be shown to be equivalent to using the vertical intervals at $u = u_i$ between two curves (34) if c_1 and c_2 are defined by (32) with p equated to 2, c replaced by 1, and δ by $\delta/2$; these intervals are generally longer—certainly if $c \leq 1$.)

The rest of this section consists of some practical considerations peculiar to the case of linear regression; the considerations of Section 12 for the general case are of course also pertinent.

The use of the calibration chart with the preceding construction is valid in the case when departures of the regression function from linearity are small compared with σ , under our other underlying assumptions. The hypothesis formulating this case could be tested with sufficiently high power if it were feasible in the development stage of the calibration method to obtain enough suitable known values V_i in a calibration-type experiment. It is also conceivable that in some calibration situations there might exist a structural theory of the relation of \mathcal{U} and \mathcal{V} implying linearity of the regression. Examples of instruments for which linear calibration curves are not used are thermometers and ammeters. A few examples of calibration problems in which the authors evidently consider linear regression adequate are: In [8, Section 2.2] \mathcal{U} is the scale reading of a flame photometer designed to measure sodium concentration, and $\{V_i\}$ are the concentrations in n samples made up with preassigned sodium concentrations. In [2, Fig. 4] \mathcal{V} is the log concentration of a certain kind of serum immunoglobulin in a sample, \mathcal{U} is the ring-diameter of the precipitate formed when a sample is introduced and incubated, with a specified technique, in a plate of antibody-agar, while the $\{V_i\}$ are the log concentrations in specially prepared

samples with known amounts of the purified immunoglobulin. In [7, Fig. 1] the problem is the measurement of body density in obese adolescent girls, \mathcal{V} is the body density measured by underwater weighing of the girl, while \mathcal{U} is the log thickness of a skinfold measured with a sliding and spreading caliper device applied to the triceps. (I am indebted to Professor Lila Elveback for finding [2] for me, and Professor Ruth L. Huenemann for [7].)

The upper and lower calibration curves (34) will be closest to the calibration line if k has its minimum value, which is attained for even n if half the observations in the calibration experiment are taken at the value $v^{(1)}$ of \mathcal{V} and half at $v^{(2)}$. I should almost never recommend this design, but rather a more or less uniform distribution of the V_i on the calibration interval $[v^{(1)}, v^{(2)}]$, with perhaps some of the V_i being duplicated if possible. A plot of the n points (U_i, V_i) might then indicate some suspicious departures from the calibration line. Also, if the regression actually is nonlinear, the resulting calibration line will generally tend to be a better approximation to the regression curve than the above two-point determination of the line—consider for example the case where the regression function is convex.

12. Discussion of the assumptions. Some of the assumptions made in deriving statistical inferences are adopted in order to facilitate the mathematical derivation, and may not be realistic; in the analysis of the effects of such assumptions on the validity of these inferences we cannot in general expect a degree of rigor and completeness comparable to that possible in the mathematical derivation.

Concerning Assumption 1, I have nothing to add to the usual warning that lack of independence can be disastrous to the validity of the statistical inferences, and concerning Assumption 2, I shall consider only the effects of nonnormality and of unequal variance, offering nothing on the more difficult question of the adequacy of the assumed functional form of the regression function.

If one regards the second term in $\hat{\sigma}w(v, c) = c_1\hat{\sigma} + c_2\hat{\sigma}S(v)$ as due to the variation of the abscissa u of the fitted regression curve about that of the regression curve, one might expect the effect of nonnormality of the $\{U_i\}$ on the correct value of c_2 to be not very serious, for the following reason: If $c = 1$, $c_2\hat{\sigma}S(v) = [pF_\delta(p, \nu)]^{\frac{1}{2}}\hat{\sigma}S(v)$ is the half-width of the interval estimate for the abscissa $u = m(v, \beta) = \sum_1^p \beta_j g_j(v)$ given by the S -method, and this has the same robustness against nonnormality as the corresponding F -test [6, Section 10.6]. But the effect of nonnormality of the $\{u_i\}$ on the correct value of c_1 can be extremely serious: The effect enters our probability calculations when we proceed from (7) to (8) and (9), where z_α would have to be replaced by the two-tailed α -point of $\sigma^{-1}[u_i - E(u_i)]$, which might differ greatly from z_α , and also in general might depend on σ ; by "two-tailed α -point" I mean here the upper α -point of $\sigma^{-1}|u_i - E(u_i)|$. If $\alpha = .05$ or $.10$, instead of $.50$, say, the experimental determination of this replacement for the constant z_α would require a very large number of observations. However if these were made, we might then consider that we have the relatively

simple case of known σ , and this would have two further advantages: First, it would replace by unity the constant $A(\delta, \nu)$ defined by (22), which enters c_1 , and whose product with $\hat{\sigma}$ may be regarded as an upper $(1 - \delta)$ -confidence limit for σ , a quantity known to be very sensitive to nonnormality. Second, it would eliminate the need of Assumption 4. I am thinking here of a very large number of observations, perhaps on different instruments of the same type as that to be calibrated, made to determine σ , before a calibration experiment with a relatively small number n of observations, so that (21) would be used with a nonzero $S(v)$ and the replacement for z_α ; however, the use in (21) of the coefficient $\chi_\delta(p)$ calculated from the normality assumption would be justified by the remark at the beginning of this paragraph. (In the light of the preceding interpretation the constant c_2 benefits from a sort of generalized central-limit-theorem effect because the distance $m(v, \hat{\beta}) - E[m(v, \hat{\beta})]$ with which it is associated is linear in the observations $\{U_i\}$, whereas c_1 enjoys no such benefit, being associated with the distance $u_i - E(u_i)$ for a *single* observation.)

Concerning violation of the assumption of equal variance, if you believe that $SD(u)$, the standard deviation of a measurement u of \mathcal{U} made at a value v of \mathcal{V} , depends on v , you might try to estimate it where you believe it is largest, and use this estimate for $\hat{\sigma}$ in my method, since you must guard against the case where most of the v_i might lie in that neighborhood. I think a better strategy would be this: Try to get a rough idea of how $SD(u)$ varies with v , from prior knowledge of the calibration situation, or from the scatter of the n points (V_i, U_i) about the fitted regression curve. Express this idea in the form $SD(u) = \sigma F(v)$, where $F(v)$ is an appropriately chosen function, and σ is a constant. If you now pretend that $SD(u)$ really is of this form, with σ a parameter to be estimated from the data, you can apply the generalization of my method given in Appendix A.

Obviously if Assumption 3 is violated in certain ways the inferences would be grossly invalidated; for example, if a v_i can occur outside the calibration interval $T = [v^{(1)}, v^{(2)}]$ such that the distribution of the corresponding u_i assigns little probability to the half of the u -axis where the interval estimate covers v_i . Assumption 3 is in a very convenient form for the mathematical derivation of my procedure, since, as we have seen, it reduces the proof to the case $v \in T$. It says that the graph of the cdf of u_i for $v_i < v^{(1)}$ ($v_i > v^{(2)}$) is nowhere to the right (left) of that for $v_i = v^{(1)}$ ($v_i = v^{(2)}$). (Perhaps you can find an intuitive justification for it directly, or a heuristic path less circuitous or more persuasive than the following: Let us consider the situation, given \mathcal{E} , so that the interval $[u^{(1)}, u^{(2)}]$ is fixed. If $v_i \notin T$ we do not want a "large" probability of an "on-scale" reading, that is, of $u_i \in [u^{(1)}, u^{(2)}]$, for we then infer $v_i \in T$. But we want more than this: Suppose $v_i < v^{(1)}$; then not only do we not want a "large" probability of an "on-scale" reading, but we do not want a large probability of an "above-scale" reading, for in either case our interval estimate does not cover v_i , that is, we do not want a "large" probability that $u_i \geq u^{(1)}$. Let us specify this to

mean that we do not want the probability of the event $u_i \geq u^{T1}$ when $v_i < v^{(1)}$ to be larger than that when $v_i = v^{(1)}$. But under the normality assumption on the $\{U_i\}$, u^{T1} is a random variable with possible value anywhere on the real line, and identifying its value with $C^{(1)}$, we are led to make the first part of Assumption 3; similarly for the second part.)

Assumption 4 is of a different character from the other assumptions; it was formulated in order to get into the problem more quickly. If $\hat{\sigma}^2$ is the residual mean-square of the calibration experiment, that is, the minimized value of (2) divided by $n - p$, then Assumption 4 is a consequence of Assumptions 1 and 2. More generally, if $\hat{\sigma}^2$ is pooled from the residual mean-squares of several independent experiments of the type of a "calibration experiment," then Assumption 4 will be satisfied if each of the experiments satisfies Assumptions 1 and 2 with the same σ^2 but possibly different T , n , and β . (In principle the known $\{g_j(v)\}$ and p could also be different, but this would be hard to imagine in practice.) Of course, if it is possible to replicate observations with the same value v of \mathcal{V} , then if $\hat{\sigma}^2$ is pooled from the mean-squares within replicates, it will satisfy Assumption 4 even if Assumption 2 is violated by assuming a wrong functional form of the regression function, unlike a $\hat{\sigma}^2$ pooled from residual mean-squares. We have already remarked that Assumption 4 is unnecessary in the case where σ^2 is known.

Finally, we consider Assumption 5, which I find the most troublesome. I believe that there is no possibility of a convincing empirical validation of this assumption—only of an invalidation. We could define a "negligible" probability to be one not exceeding a specified ϵ , and then calculate what part of the parameter space is excluded by Assumption 5, but I see no profit in this exercise—all the questions about accepting the assumption would remain. A robust high-power test of Assumption 5 does not seem practicable: It would require a tremendous number of observations to establish the "negligible" probability, and much of the test's probability of rejection would no doubt come from parts of the critical region where the various components of Assumption 2 would be most suspect.

It may be helpful to consider the following set of four assumptions; they are listed in order of decreasing severity, for we know each implies the next: First, Assumption 5. Next, 5.1: The probability is negligible that Condition 2 is violated, that is, that for some $v \in T$, and for $h = 1$ or 2 , $m_v^{(h)}(v, \mathcal{E}, c) \leq 0$. Then, 5.2: The probability is negligible that Condition 1 is violated, that is, that for some $v \in T$, $m_v(v, \hat{\beta}) \leq 0$. Finally, 5.3: For all $v \in T$, $m_v(v, \beta) > 0$. Considering the assumptions in reverse order, I find 5.3 indispensable for justifying point estimation of the v_i from a calibration curve which is the graph of a strictly increasing function, and furthermore it seems to me that the stronger 5.2 must be assumed for this purpose, for reasons indicated in Section 3 in connection with Condition 1. I strongly suspect that for interval estimation of the $\{v_i\}$ some assumption resembling 5.1 must be made. And Assumption 5 speaks

to me more simply than 5.1. For 5.1 and 5.2, as well as 5, empirical validation does not seem practicable. (Nor does it for Assumption 1, whose violation can also be disastrous, but this is of a type to which we are more accustomed.) It would be formally possible to avoid Assumption 5 by redefining $I(u_i, \mathcal{E})$ as the smallest interval containing the set $I(u_i, \mathcal{E})$. When Condition 2 is violated this redefinition would cause the upper or lower calibration curve to have a vertical portion, which the practitioner would regard as queer, and it would probably produce the unhappy situation of the next paragraph.

I would guess that what would happen more frequently in practice than violation of Condition 2 ensuring that the upper and lower calibration curves are graphs of single-valued functions of u , is that for some u in $[u^{11}, u^{12}]$ the intervals $I(u, \mathcal{E})$ would be too wide to be useful, and the calibration chart would therefore not be used. An extreme case of this would happen if $u^{11} > u^{12}$: Then if Condition 2 is satisfied, we see from the defining set $\mathcal{A}(\mathcal{E}, c)$ in Section 4, that for u in the interval $[u^{12}, u^{11}]$, $I(u, \mathcal{E})$ would be the whole real line, giving $-\infty < v_i < +\infty$, while outside this interval it would be a half-line. But even when $u^{11} < u^{12}$, so that the intervals $I(u, \mathcal{E})$ are finite for $u^{11} \leq u \leq u^{12}$, they may still be too large to be useful. (One might consider conditioning on the event that the intervals $I(u, \mathcal{E})$ be of less than a specified length, but I expect this calculation would be difficult and the result unsatisfactory, in that the probability which we bounded by $1 - \delta$ would depend seriously on unknown parameters and could not be thus bounded. I try to console myself that this defect in my method is shared by all the commonly used confidence intervals whose width is a random variable not satisfying a preassigned bound.)

If the intervals $I(u, \mathcal{E})$ are too wide you may consider the remedy of using a larger number n of observations in a suitably designed calibration experiment. In this way you could make the second term in $w(v, c) = c_1 + c_2 S(v)$ as small as you please. For large n the first term would approach z_α , the horizontal distance $2\hat{\sigma}w(v, c)$ between the upper and lower calibration curves would approach $2\sigma z_\alpha$, and the vertical distance, or width of $I(u, \mathcal{E})$, would approach a limit approximately equal to $2\sigma z_\alpha / m_v(v, \beta)$. This could be estimated as $2z_\alpha G$, where

$$G = \hat{\sigma} / m_v(v, \hat{\beta}) = \hat{\sigma} [\sum_{j=1}^p \hat{\beta}_j g_j'(v)]^{-1},$$

and the standard deviation of $2z_\alpha G$ could be estimated as

$$2z_\alpha G [G^2 \sum_{j,k=1}^p b_{jk} g_j'(v) g_k'(v) + (2\nu)^{-1}]^{\frac{1}{2}}.$$

You would have to consider the estimated limiting width $2z_\alpha G$ as a function of v in the calibration interval $[v^{(1)}, v^{(2)}]$, and also its estimated standard deviation, and if you decided it was still too large, it would indicate that the calibration method being developed, of estimating the value of \mathcal{V} from a measurement u of \mathcal{U} , is not sufficiently precise for your purposes, because even if the calibration curve were to be determined with negligible error, the product of its slope $[m_v(v, \beta)]^{-1}$ by the standard deviation σ of a measurement u would be too large.

APPENDIX A

Generalization to a case of unequal variance. Let us generalize Assumption 2 to permit $SD(U_i) = \sigma F(V_i)$, and if $v_i \in T$, $SD(u_i) = \sigma F(v_i)$, where $F(v)$ is a known function with a continuous derivative for $v \in T$, and σ is an unknown parameter. Let us leave Assumptions 1, 3, 4 unchanged, and postpone generalizing Assumption 5 until we see what is needed.

The estimates $\{\hat{\beta}_j\}$ are calculated by the method of weighted least squares, where instead of (2) we minimize

$$(37) \quad \sum_{i=1}^n [F(V_i)]^{-2} [U_i - \sum_{j=1}^p \hat{\beta}_j g_j(V_i)]^2.$$

If the estimate $\hat{\sigma}^2$ of σ^2 is calculated only from the residual sum of squares in the calibration experiment, it is the minimized quantity (37) divided by $n - p$. The calibration chart is constructed with the lower and upper calibration curves defined by

$$(38) \quad u = m(v, \hat{\beta}) \pm \hat{\sigma}[c_1 F(v) + c_2 S(v)],$$

where $S(v)$ is still defined by (11), and c_1 and c_2 are calculated as follows: Let $S^*(v) = S(v)/F(v)$; for $i = 1, 2$ let S_i^* be the inf and sup of $S^*(v)$ for $v \in T$; find c by entering Table 1 with $s_i = S_i^*/z_\alpha$; c_i are then given by (32). The interval estimates for the v_i are given in terms of the lower and upper calibration curves by the same statements \mathcal{S}_i as in Section 10.

This generalization can be derived by transforming to “pseudo-observations” depending on the observations and the unknown v_i as follows: $U_i^* = U_i/F(V_i)$, $u_i^* = u_i/F(v_i)$ if $v_i \in T$, $u_i^* = u_i/F(v^{(1)})$ if $v_i < v^{(1)}$, $u_i^* = u_i/F(v^{(2)})$ if $v_i > v^{(2)}$. Under the transformation the quantities $v, \beta, \hat{\beta}, \sigma, \hat{\sigma}$ remain invariant, while for $v \in T$, the following quantities transform like $u^* = u/F(v)$: $g_j(v), m(v, \beta), m(v, \hat{\beta})$, and $S(v)$, that is, $g_j^*(v) = g_j(v)/F(v)$, etc. It may be verified that the “pseudo-observations” satisfy the Assumptions 1–4 made for the original method with $g_j(v)$ replaced by $g_j^*(v)$, and $m(v, \beta)$ by $m^*(v, \beta)$. The lower and upper calibration curves in the transformed problem have the equations

$$(39) \quad u^* = m^*(v, \hat{\beta}) \pm \hat{\sigma}[c_1 + c_2 S^*(v)],$$

and the sets $I^*(u_i^*, \mathcal{E})$ defined by these cover the v_i with the required probabilities; however they need not be intervals. If we did make the appropriate form of Assumption 5 in the transformed problem to ensure, with probability one minus a negligible amount, that the curves (39) have positive slope for all $v \in T$, this would not ensure that the corresponding curves (38) in the untransformed problem have positive slope. What we need is

CONDITION 2'. For $h = 1, 2$ and all $v \in T$,

$$m_i(v, \hat{\beta}) + (-1)^h \hat{\sigma}[c_1 F'(v) + c_2 S'(v)] > 0.$$

Although the lemma about $S'(v)$ in Section 9 is still valid, it no longer leads me to a neat analogue of Assumption 5, and so after some vain attempts to untie the Gordian knot I cut it with

ASSUMPTION 5'. The probability is negligible that Condition 2' is violated.

APPENDIX B

Frequency interpretation. Suppose the events of a sequence \mathcal{E}_i are independent, and let $\pi_i = \Pr(\mathcal{E}_i)$. Define $x_i = 1$ or 0 according as \mathcal{E}_i occurs or not, so that $p_m = \sum_1^m x_i/m$ is the proportion of the m events $\mathcal{E}_1, \dots, \mathcal{E}_m$ that occur. If all $\pi_i = \pi$ we may say that "the proportion of events \mathcal{E}_i that occur is in the long run equal to π ." A frequency interpretation of this case is usually based on the law of large numbers. The weak law says p_m converges to π in probability, i.e., for every $\varepsilon > 0$, $\Pr\{|p_m - \pi| < \varepsilon\} \rightarrow 1$ as $m \rightarrow \infty$. The strong law says p_m converges to π almost surely, i.e., $\Pr\{\lim_m p_m = \pi\} = 1$. Using the weak law, a frequency interpretation is given to the statement that for every $\varepsilon > 0$ the probability approaches 1 as $m \rightarrow \infty$ that p_m lies in the interval $(\pi - \varepsilon, \pi + \varepsilon)$; using the strong law, a frequency interpretation is given to the occurrence of events of probability one.

In the case that all $\pi_i \geq \pi$ we say that "the proportion of events \mathcal{E}_i that occur is in the long run $\geq \pi$." The weak law can then be reformulated thus: for every $\varepsilon > 0$, $\Pr\{p_m \geq \pi - \varepsilon\} \rightarrow 1$; and the strong law thus: $\Pr\{\liminf_m p_m \geq \pi\} = 1$. The weak result follows from applying to p_m the Chebyshev inequality and the formulas $E(p_m) \geq \pi$, $V(p_m) \leq (4m)^{-1}$. The strong result follows from applying to the sequence $y_i = x_i - \pi_i$ Kolmogorov's criterion [1, Theorem 3.4] and the formulas $E(y_i) = 0$, $V(y_i) \leq \frac{1}{4}$. Whatever method of obtaining a frequency interpretation is used in the previous case can be applied here.

(In case all $\pi_i \leq \pi$ we would say "the proportion of events \mathcal{E}_i that occur is in the long run $\leq \pi$," and the frequency interpretation would be based on the reformulated weak law that for every $\varepsilon > 0$, $\Pr\{p_m \leq \pi + \varepsilon\} \rightarrow 1$ as $m \rightarrow \infty$, or on the reformulated strong law, $\Pr\{\limsup_m p_m \leq \pi\} = 1$. All these reformulations could be made for any sequence of random variables of the structure $P_m = \sum_1^m X_i/m$, where the X_i are independent and $\sum_1^\infty V(X_i)/i^2 < \infty$. The frequency interpretations are as above, and are useful in many situations, besides that of calibration, where there is a routine application of the same statistical procedure to many successive cases, for example with attribute sampling inspection plans: Let $L(\pi)$ denote the probability of the plan accepting a lot with proportion defective π . Under the assumption that $L(\pi)$ is monotone nondecreasing, it follows that if a stream of lots characterized by an arbitrary sequence $\{\pi_i\}$ is submitted to the plan the proportion of bad (good) lots accepted (rejected) by the plan is in the long run \leq the consumer's (producer's) risk, and if it is an AOQL plan, the average proportion defective in the outgoing lots is in the long run \leq the AOQL, defined as $\inf_\pi [\pi L(\pi)]$. The usual frequency interpretations are made for the unrealistic model where all the π_i are equal.)

Acknowledgments. I shall feel forever indebted to Rupert G. Miller, Jr. for having found a basic error in an earlier version of the manuscript of this article,

which claimed for all p the present result for $p = 1$. Emil Spjøtvoll worked out a correct formulation of the operating characteristic for that version. Having decided nevertheless to attempt the present version, I asked Paul G. Hoel for references to work subsequent to his [3] on confidence bands for regression curves; he responded by sending me an unpublished manuscript which analyzed the relation between his method and C. R. Rao's [5]. David R. Brillinger suggested extending the weak-law treatment of the frequency interpretation in Appendix B to strong-law. David W. Hutchinson programmed the computer calculations.

REFERENCES

- [1] DOOB, J. L. (1953). *Stochastic Processes*. Wiley, New York.
- [2] FAHEY, J. L. and MCKELVEY, E. M. (1965). Quantitative determination of serum immunoglobulins in antibody-agar plates. *J. Immunology* **94** 84-90.
- [3] HOEL, PAUL G. (1954). Confidence bands for polynomial curves. *Ann. Math. Statist.* **25** 534-542.
- [4] LIEBERMAN, G. J., MILLER, R. G., JR., and HAMILTON, M. A. (1967). Unlimited simultaneous discrimination intervals in regression. *Biometrika* **54** 133-145. Corrections in *Biometrika* **58** 687.
- [5] RAO, C. RADHAKRISHNA (1959). Some problems involving linear hypotheses in multivariate analysis. *Biometrika* **46** 49-58.
- [6] SCHEFFÉ, HENRY (1959). *The Analysis of Variance*. Wiley, New York.
- [7] SELTZER, C. C., GOLDMAN, R. F., and MAYER, J. (1965). The triceps skinfold as a predictive measure of body density and body fat in obese adolescent girls. *Pediatrics* **36** 212-218.
- [8] WILLIAMS, E. J. (1959). *Regression Analysis*. Wiley, New York.
- [9] WILLIAMS, E. J. (1969). Regression methods in calibration problems. *Proc. 37-th Session, Bulletin of the International Statistical Institute* **43** Book 1, 17-28.
- [10] WILKS, S. S. (1941). Determination of sample sizes for setting tolerance limits. *Ann. Math. Statist.* **12** 91-96.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720