JOURNAL OF RESEARCH of the National Bureau of Standards—B. Mathematics and Mathematical Physics
Vol. 68B, No. 2, April–June 1964

Fitting $y=\beta x$ When the Variance Depends on x

John Van Dyke[1]

(February 25, 1964)

This paper presents some results concerning the selection of a method for estimating the slope of a straight line through the origin. For fitting the line $y=\beta x$ when the variance of y is proportional to x^p, it is well known that the best estimate of β depends on p. In practice, however, only integer values of p would be convenient to work with. One of the estimators appropriate for $p=0,1,2$ would probably be used if the value of p were in fact fractional, or if it were only approximately known. This paper provides some guides for choosing the best among these estimators in a particular situation.

Formulas for the best estimators of β and their variances are given. Estimators of β appropriate for integer values of p are compared in the case when p is not integral, but is known, and in the case when p is only approximately known. Estimation of the variances of estimators of β is considered. Finally, some results are given on the effect of the spacing of the x values on the comparison of the estimators.

This paper presents some results concerning the selection of a method for estimating the slope of a straight line through the origin. For fitting the line $y=\beta x$ when the variance of y is proportional to x^p, it is well known that the best estimator of β depends on p. In practice, however, only integer values of p would be convenient to work with. One of the estimators appropriate for $p=0,1,2$ would probably be used if the value of p were in fact fractional, or if it were only approximately known. This paper provides some guides for choosing the best one in a particular situation.

The line $y=\beta x$ is to be fitted to observed points $(x_1, y_1), \ldots, (x_n, y_n)$, where $x_i>0$. The values of x are assumed to be known without error. For fixed x_i, the corresponding y_i, is an observed value of the random variable Y_i with mean βx_i and variance $V(Y_i)=\sigma^2 x_i^p, p\geqslant 0$.

Let B_p denote the "best," i.e., the minimum variance unbiased linear, estimator of β when $V(Y_i)=\sigma^2 x_i^p$, and let $V_p(B_p)$ be its variance. Denote by $V_p(B_r)$ the variance of the estimator B_r when $V(Y_i)=\sigma^2 x_i^p$ and $p\neq r$. When r is restricted to the integers it will be represented by k.

Section 1 contains formulas for estimators of β and for the variances of these estimators. Comparisons of B_k and B_{k+1} are considered in sections 2 and 3, with numerical calculations given for the case of equally spaced values of x. Section 2 is a comparison of B_k and B_{k+1} (k integral) when p is known and $k<p<k+1$. Values of $p^*=p_n(k, k+1)$ are given such that when $k\leqslant p<p^*$, B_k is better than B_{k+1}, and when $p^*<p\leqslant k+1$, then B_{k+1} is better. In section 3, B_k and B_{k+1} are compared when it is known that $k<p<k+1$, but the value of p is unknown. Comparison of the *relative efficiencies* of these estimators shows that B_k is better than B_{k+1} for $k<p<k+1$. In section 4, an estimator, t_r, of the variance of B_r is considered. It is an unbiased estimator of $V_p(B_r)$ whenever $r=p$ and for all p when $r=2$. It is positively biased when $p<r<2$, and negatively biased when $r<p<2$. Section 5 contains some results on the effect of the spacing of the x values on the comparison of estimators.

Methods for fitting a straight line when the variance depends on x are treated in a general way in, for example, Brownlee's chapter on linear regression.[2] The present paper provides further details relating to the choice of weights, for a particular case of the problem considered in Brownlee's section 11.14, "Weighted Regression through the Origin." The methods employed here could be adapted for consideration of other cases.

1. Formulas for B_P and $V_P(B_P)$

Let $Y_1, Y_2, \ldots, Y_n$ be mutually independent random variables corresponding to a set of exact positive values $x_1, x_2, \ldots, x_n$, with expectations

$$E(Y_i)=\beta x_i,$$

and let the variances $V(Y_i)$ be proportional to x_i^p, i.e., let

$$V(Y_i)=\sigma^2 x_i^p, \qquad p\geqslant 0.$$

Then the best estimator B_p of β is the value of B_p that minimizes the sum of weighted squared residuals,

$$\sum_{i=1}^{n}\frac{(y_i-B_p x_i)^2}{\sigma^2 x_i^p}.$$

[1] Present address: Bureau of Scientific Research, Food and Drug Administration, Washington, D.C.

[2] K. A. Brownlee, Statistical Theory and Methodology in Science and Engineering, New York, John Wiley & Sons, Inc., 1960 (Chapter 11).

The estimator is

$$B_p = \frac{\sum_{i=1}^{n} y_i x_i^{1-p}}{\sum_{i=1}^{n} x_i^{2-p}}. \tag{1–1}$$

Its variance is given by

$$V_p(B_p) = \frac{\sigma^2}{\sum_{i=1}^{n} x_i^{2-p}}. \tag{1–2}$$

Table 1 gives formulas for B_p and $V_p(B_p)$ for $p=0$, 1 and 2, when there are n points: (x_1, y_1), (x_2, y_2), . . . , (x_n, y_n).

TABLE 1.—*Formulas for* B_p *and* $V_p(B_p)$

p	B_p	$V_p(B_p)$
0	$\frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$	$\frac{\sigma^2}{\sum_{i=1}^{n} x_i^2}$
1	$\frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i} = \frac{\bar{y}}{\bar{x}}$	$\frac{\sigma^2}{\sum_{i=1}^{n} x_i}$
2	$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i}{x_i}\right) = \overline{\left(\frac{y}{x}\right)}$	$\frac{\sigma^2}{n}$

If there are m values of y for each x, then y_i in the above formulas is replaced by y_{ij} with $j=1, 2, . . . , m$, and the denominator of each estimate and variance has the additional factor m; e.g., for $p=1$,

$$B_1 = \frac{\sum_{j=1}^{m} \sum_{i=1}^{n} y_{ij}}{m \sum_{i=1}^{n} x_i}$$

and

$$V_1(B_1) = \frac{\sigma^2}{m \sum_{i=1}^{n} x_i}$$

It has been assumed for convenience that all x_i are positive. In general, one might have both positive and negative values of x. The above formulas are easily modified for use when the variance depends on the absolute value of x,

$$V(Y_i) = \sigma^2 |x_i|^p.$$

Alternatively, the formulas in table 1 may be used directly provided every pair (x_i, y_i) having $x_i < 0$ has been replaced by $(-x_i, -y_i)$.

2. Comparison of B_k and B_{k+1} (k Integral) When $k<p<k+1$

Since numerical calculation of B_p and its variance is not easy when p is not an integer, the question arises as to which estimator, B_k or B_{k+1}, is the "better" estimator when p is known but lies between the integers k and $k+1$. Regardless of the true value of p, B_k is an unbiased estimator of β. Thus the question of which of the two is the "better" reduces to the question of which one has the smaller variance. This can be answered if we can find a value p^* such that $k < p^* < k+1$ and

$$V_p(B_k) < V_p(B_{k+1}) \ , \ k \leq p < p^*$$

$$V_p(B_k) > V_p(B_{k+1}) \ , \ p^* < p \leq k+1.$$

Here $V_p(B_k)$ denotes the variance of B_k calculated under the assumption that $V(Y_i) = \sigma^2 x_i^p$. In general,

$$V_p(B_r) = \frac{\sigma^2 \sum_{i=1}^{n} x_i^{2-2r+p}}{\left[\sum_{i=1}^{n} x_i^{2-r}\right]^2}. \tag{2–1}$$

The value of p^* may be found by setting $V_p(B_k)$ equal to $V_p(B_{k+1})$, and then solving for p.

Values of p^* have been found for the case of equally spaced positive values of x; that is, $x_i = ic$, $i=1$, 2, . . . , n, and c is any positive constant. The line $y=\beta x$ is then fitted to the points (c, y_1), $(2c, y_2)$, . . . , (nc, y_n). Let $p_n(k, k+1)$ denote the solution of the equation $V_p(B_k) = V_p(B_{k+1})$ in the interval $k < p < k+1$. Values of $p_n(k, k+1)$ were found for $k=0$, 1, and for various values of n. For small n $(2 \leq n \leq 10)$, $p_n(k, k+1)$ was found by graphical methods. The exact values of $p_n(k, k+1)$ for $k=0$, $n=2$ and $k=1$, $n=2$, 3, were also obtained directly by solving the equation $V_p(B_k) = V_p(B_{k+1})$. These results agree well enough with the graphical results so that it is believed that the results obtained graphically are correct to within one unit in the second decimal place. The value for $n=100$ was obtained by approximate methods. For $n \to \infty$, $p_\infty(k, k+1)$ was found by letting n tend to infinity in the equation $V_p(B_k) = V_p(B_{k+1})$ and then solving the resulting equation for p. The values of $p_n(k, k+1)$ are given in table 2. Since the range of $p_n(k, k+1)$ is comparatively small over all values of n, we state a convenient rule:

When	*Use*
$0 < p < 0.6$	B_0
$0.6 < p < 1.6$	B_1
$1.6 < p < 2+?$	B_2

Table 3 contains formulas for the estimators B_k ($k=0, 1, 2$) for equally spaced x and gives their variances $V_p(B_k)$ for $p=0, 1, 2$.

Figure 1 and figure 2 give, for $n=3$ and $n=10$ respectively, values of $V_p(B_p)$ and $V_p(B_k)$ for p rang-

TABLE 2. *Values of* $p_n(k, k+1)$ *for equally spaced* x

n	$k=0$	$k=1$
2	0.5406	1.5146
3	.56	1.5350
4	.57	1.55
5	.57	1.56
6	.58	1.57
7	.58	1.58
8	.59	1.58
9	.59	1.59
10	.59	1.59
100	.599+	1.64
∞	.600 – – –	1.66 – – –

TABLE 3. *Variances* $V_p(B_k)$ *and slopes* B_k *for equally spaced* x

p	$k=0$	$k=1$	$k=2$
0	$\frac{6\sigma^2}{n(n+1)(2n+1)c^2}$	$\frac{4\sigma^2}{n(n+1)^2c^2}$	$\frac{\sigma^2 \sum_{i=1}^{n} i^{-2}}{n^2c^2}$
1	$\frac{9\sigma^2}{c(2n+1)^2}$	$\frac{2\sigma^2}{cn(n+1)}$	$\frac{\sigma^2 \sum_{i=1}^{n} i^{-1}}{cn^2}$
2	$\frac{6(3n^2+3n-1)\sigma^2}{5n(n+1)(2n+1)}$	$\frac{2(2n+1)\sigma^2}{3n(n+1)}$	$\frac{\sigma^2}{n}$
B_k	$\frac{6\sum_{i=1}^{n} iy_i}{cn(n+1)(2n+1)}$	$\frac{2\sum_{i=1}^{n} y_i}{cn(n+1)}$	$\frac{\sum_{i=1}^{n} i^{-1}y_i}{cn}$

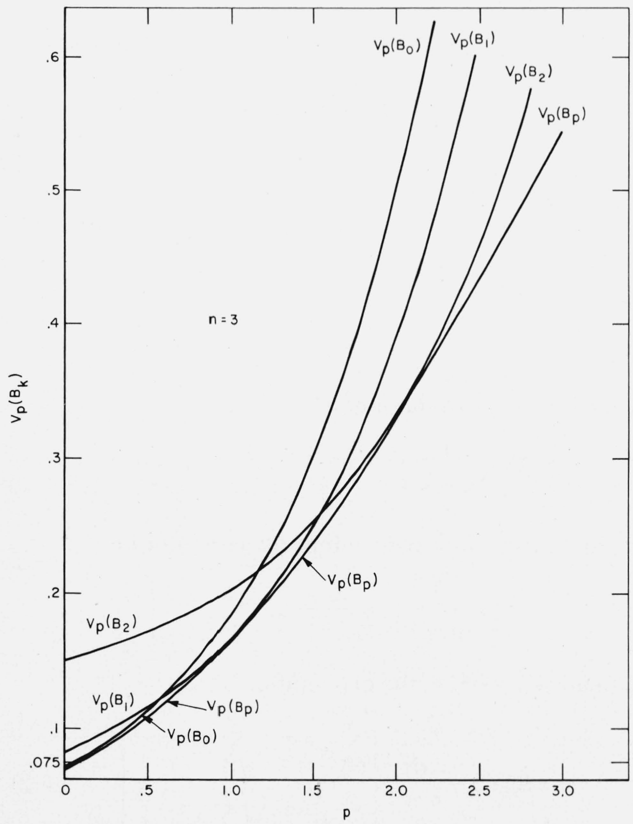

FIGURE 1. *Variances* $V_p(B_k)$, $V_p(B_p)$ *for* $c=1$, $\sigma^2=1$, $n=3$.

ing from 0.0 to at most 3.0 and $k=0$, 1, and 2. Figures such as these (on a larger scale) were used to obtain the values of $p_n(k, k+1)$, $2 \leq n \leq 10$. For convenience, the values of c and σ^2 were chosen equal to one. It is evident from the two figures, and it also can be shown mathematically, that $V_p(B_p)$ and $V_p(B_k)$ are increasing functions of p when all values of x are at least equal to one.

3. Relative Efficiency of B_k, B_{k+1} When $k<p<k+1$

Suppose it is only known that the value of p is in the interval $(k, k+1)$. Which estimator of β, B_k, or B_{k+1}, is then the "better" of the two? One way to answer the question is to use the idea of *relative efficiency*. The efficiency of an estimator T_1 of θ relative to an estimator T_2 of θ is defined to be

$$\text{R.E.}\,(T_1|T_2) = 100\,\frac{E(T_2-\theta)^2}{E(T_1-\theta)^2},$$

expressed as a percentage. If R.E. $(T_1|T_2) > 100$ percent, T_1 may be considered the better estimator of θ. If $E(T_1)=E(T_2)=\theta$, the relative efficiency is the ratio of the variances,

$$\text{R.E.}\,(T_1|T_2) = 100\,\frac{V(T_2)}{V(T_1)}.$$

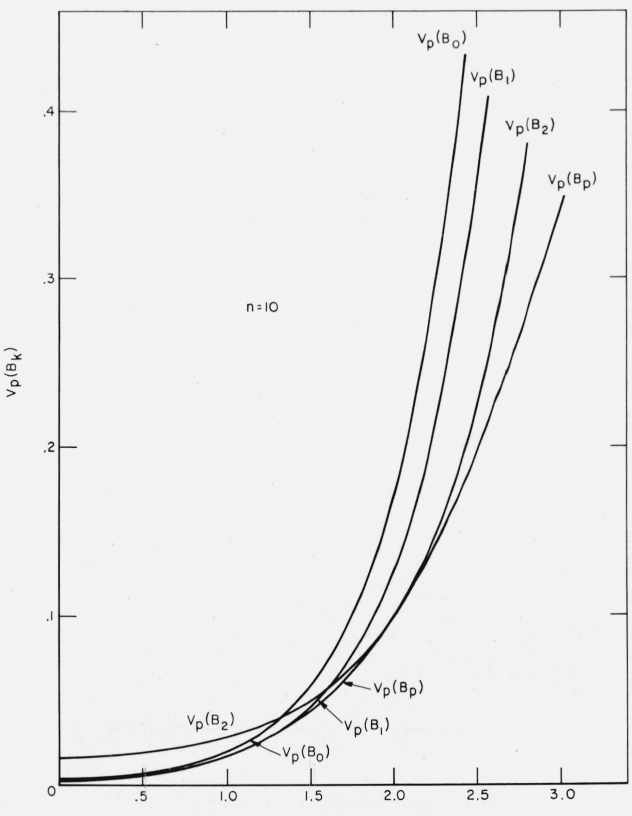

FIGURE 2. *Variances* $V_p(B_k)$, $V_p(B_p)$ *for* $c = 1$, $\sigma^2 = 1$, $n = 10$.

To compare B_k with B_{k+1}, we consider the least advantageous situation for each: (1) let B_k be the estimator when in fact $p=k+1$, and (2) let B_{k+1} be the estimator when $p=k$. Under these conditions the best estimator can be found by comparing two particular relative efficiencies:

(1) the efficiency of B_k relative to B_{k+1}, assuming $p=k+1$,

$$\text{R.E. } (B_k|B_{k+1})=\frac{V_{k+1}(B_{k+1})}{V_{k+1}(B_k)}, \tag{3-1}$$

(2) the efficiency of B_{k+1} relative to B_k, assuming $p=k$,

$$\text{R.E. } (B_{k+1}|B_k)=\frac{V_k(B_k)}{V_k(B_{k+1})}. \tag{3-2}$$

Both of these efficiencies are less than 100 percent, because the numerator is the variance of the estimator whose variance is minimum. However, if one of these relative efficiencies is larger than the other, then the estimator whose variance is the denominator of this larger ratio may be considered the better of the two estimators in the sense that it is less inefficient under the worst possible circumstances. A comparison between such pairs of relative efficiencies was made for the case of equally spaced x when $k=0$ and $k=1$. This comparison of the relative efficiencies showed that for both values of k,

$$\text{R.E. } (B_k|B_{k+1}) > \text{R.E. } (B_{k+1}|B_k)$$

whenever $n \geqslant 2$. Thus, B_0 is better than B_1 when $0 \leqslant p \leqslant 1$ and B_1 is better than B_2 when $1 \leqslant p \leqslant 2$.

Figure 3 and figure 4 show for $n=3$ and $n=10$ respectively values of

$$\text{R.E. } (B_l|B_k)=100\left[\frac{V_p(B_k)}{V_p(B_l)}\right],$$

for all p in the interval 0.0 to 3.0 and k and l equal to some combinations of 0, 1, and 2. To verify the conclusion stated in the previous paragraph, notice for example that (in both figures) V_0/V_1 at $p=0$ is smaller than V_1/V_0 at $p=1$; that is, the minimum of

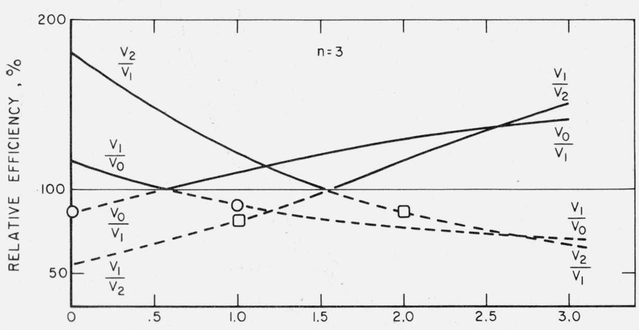

FIGURE 3. *Relative efficiency curves for* n=3.

$$\text{R.E. } (B_l|B_k)=100\left[\frac{V_p(B_k)}{V_p(B_l)}\right]=\frac{V_k}{V_l}$$

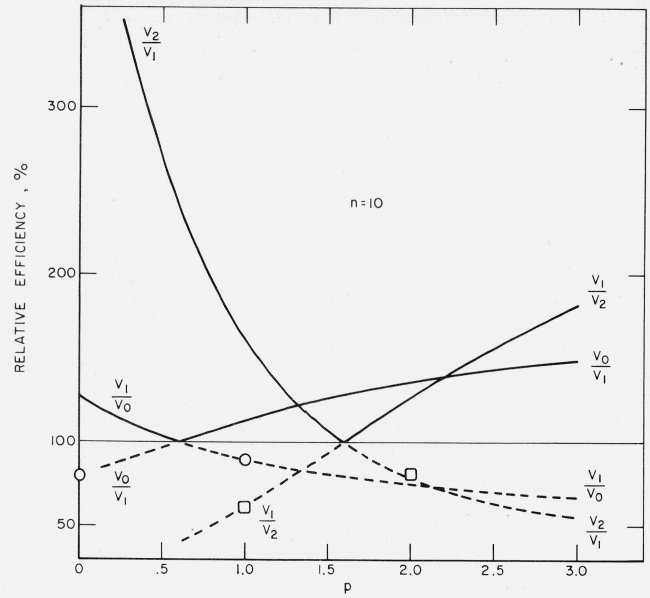

FIGURE 4. *Relative efficiency curves for* n=10.

$$\text{R.E. } (B_l|B_k)=100\left[\frac{V_p(B_k)}{V_p(B_l)}\right]=\frac{V_k}{V_l}$$

R.E. $(B_1|B_0)$ is less than the minimum of R.E. $(B_0|B_1)$ in the interval $0 \leqslant p \leqslant 1$.

4. Estimation of the Variance of B_r

It has been shown that the variance of estimator B_r is

$$V_r(B_r)=\frac{\sigma^2}{\sum_{i=1}^{n} x_i^{2-r}}.$$

Let t_r denote the following estimator for the variance of B_r:

$$t_r=\frac{S_r^2}{\sum_{i=1}^{n} x_i^{2-r}}, \tag{4-1}$$

where S_r^2 is an estimator of σ^2,

$$S_r^2=\frac{1}{\mathrm{n}-1}\sum_{i=1}^{n}\frac{(y_i-B_r x_i)^2}{x_i^r}, \qquad r \geqslant 0,$$

or, in a form more convenient for computation

$$S_r^2=\frac{1}{n-1}\left[\sum_{i=1}^{n}\frac{y_i^2}{x_i^r}-B_r{}^2\sum_{i=1}^{n} x_i^{2-r}\right].$$

When $V(Y_i)=\sigma^2 x_i^p$, the expectation of t_r is

$$E_p(t_r)=\frac{\sigma^2}{n-1}\left[\frac{\sum_{i=1}^{n} x_i^{p-r}}{\sum_{i=1}^{n} x_i^{2-r}}-\frac{\sum_{i=1}^{n} x_i^{2-2r+p}}{\left(\sum_{i=1}^{n} x_i^{2-r}\right)^2}\right]. \tag{4-2}$$

When $p=r$, t_r is an unbiased estimator of $V_r(B_r)$. When $p \neq r$, one would like the expectation, $E_p(t_r)$, to be $V_p(B_r)$ as given in eq (2–1), but this is not usually the case. However, when $r=2$, t_r is an unbiased estimator; i.e., $E_p(t_2)=V_p(B_2)$ for all p. The bias of estimator t_r is the difference, $E_p(t_r)-V_p(B_r)$, and this quantity is positive or negative depending upon the values of p and r. In general, for positive values of x which are not all equal, the bias of estimator t_r is positive if $p < r < 2$, and it is negative if $r < p < 2$; i.e.,

$$E_p(t_r) > V_p(B_r), p < r < 2,$$

$$E_p(t_r) < V_p(B_r), r < p < 2.$$

A sketch of the proof of this result follows.

$$E_p(t_r)-V_p(B_r)=\frac{\sigma^2}{n-1}\,\frac{\left(\sum_{i=1}^{n} x_i^{p-r}\sum_{i=1}^{n} x_i^{2-r}-n\sum_{i=1}^{n} x_i^{2-2r+p}\right)}{\left(\sum_{i=1}^{n} x_i^{2-r}\right)^2}, \tag{4–3}$$

which can also be written as

$$\frac{\sigma^2}{\left(\sum_i x_i^{2-r}\right)^2}\,\frac{1}{n-1}\sum_{i<j}(x_i^{p-r}-x_j^{p-r})(x_j^{2-r}-x_i^{2-r}).$$

Consider the product $(x_i^{p-r}-x_j^{p-r})(x_j^{2-r}-x_i^{2-r})$; and take $x_1 \leq x_2 \leq \ldots \leq x_n$, but not all equal.

(1) If $p < r < 2$, the signs of the two factors are (+) (+) for all pairs x_i, x_j for which $x_i \neq x_j$. The sum of such terms is positive and hence the bias is positive.

(2) If $r < p < 2$, the signs are (−) (+) for all pairs x_i, x_j where $x_i \neq x_j$, and in this case the bias is negative.

Thus the estimator t_r is likely to overestimate the variance $V_p(B_r)$ if r is such that $p < r < 2$, and it is likely to underestimate $V_p(B_r)$ if $r < p < 2$.

For the case of equally spaced x, the effect of differences between p and r is shown in figure 5, where the *relative bias* of estimator t_r,

$$\text{R.B.}=100\left[\frac{E_p(t_r)-V_p(B_r)}{V_p(B_r)}\right], \tag{4–4}$$

is plotted against $p, 0 \leq p \leq 2$, for $r=k=0$, 1 and $n=3, 10$. This quantity is independent of σ^2 and of the constant c.

Figure 5 illustrates the possibility that the estimated variance of B_k may be substantially biased if k is not close to p. For example, suppose that $n=10$ and the value of p is unknown, but it is known to be in the interval (0, 1). If $k=0$, t_0 would most likely underestimate $V_p(B_0)$, for the relative bias ranges downward from 0.0 percent (at $p=0$) to about −33 percent (at $p=1$). On the other hand, if $k=1$, t_1 would be likely to overestimate $V_p(B_1)$, and the bias in this case could be as high as 68 percent (if $p=0$) of $V_p(B_1)$.

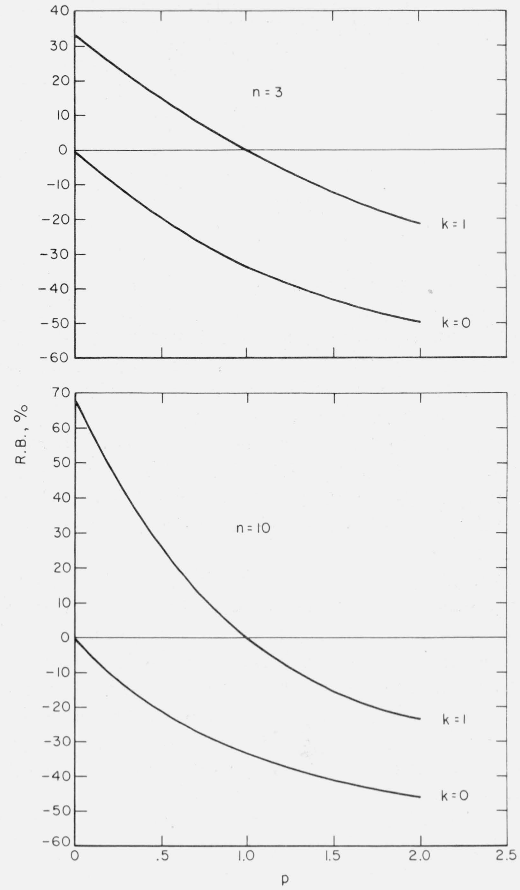

FIGURE 5. *Relative bias of estimator* t_r, *when* $r=k=0, 1$.

$\text{R.B.}=100\,[E_p(t_k)-V_p(B_k)]/V_p(B_k)$

5. Effect of the Spacing of the x Values on the Comparison of the Estimators

Up to this point emphasis has been on the case where the x values were n in number and were equally spaced on the interval $(0, nc)$; that is, $x_i=ic$, $i=1, 2, \ldots, n$; and to each x there corresponded a single y.

Now suppose that instead of n different x's, there are $N=n/m$ equally spaced x's on the interval $(0, nc)$, and m y's corresponding to each x; that is, the line $y=\beta x$ is fitted to the n points, $(x_i, y_{ij})=(mic, y_{ij})$, $i=1, 2, \ldots, N, j=1, 2, \ldots, m$.

Of course, if $m=1$, we are back to the case of n x's and a single y for each, so it is a special case of this more general form. For the more general case, the estimator B_p and variances $V_p(B_p)$ and $V_p(B_r)$ have the following forms:

$$B_p=\frac{\sum_{j=1}^{m}\sum_{i=1}^{N} y_{ij}i^{1-p}}{cm^2\sum_{i=1}^{N} i^{2-p}}, \tag{5–1}$$

$$V_p(B_p)=\frac{\sigma^2}{c^{2-p}m^{3-p}\sum_{i=1}^{N} i^{2-p}}, \tag{5–2}$$

$$V_p(B_r)=\frac{\sigma^2 \sum_{i=1}^{N} i^{2-2r+p}}{c^{2-p}m^{3-p}\left[\sum_{i=1}^{N} i^{2-r}\right]^2}. \tag{5–3}$$

In comparing estimators B_k and B_{k+1} in section 2, the equation, $V_p(B_k)=V_p(B_{k+1})$, was solved for p, and the solutions, $p_n(k, k+1)$, are given in table 2. The solutions of this equation when the variances are of the general form, eq (5–3), are also found in table 2. It is the number of x values, N, that determines the solution, and so the solutions are found by substituting the value of N for the value n of the table. For example, if $n=15$ and $m=3$, then $N=15/3=5$, and the solutions are those corresponding to $n=5$, i.e., 0.57 and 1.56.

In section 3, the relative efficiency of estimators was used to show that estimator B_k was better than estimator B_{k+1} under certain circumstances. The same result is obtained for the more general case. In addition, any combination of n and m for which N is constant gives the same relative efficiency (eq (3–1) or (3–2)). Thus, figure 4 is identical to those obtained for all pairs n and m such that $N=n/m=10$.

The estimator t_r (introduced in sec. 4) of the variance $V_r(B_r)$ can here be written in the form

$$t_r=\frac{S_r^2}{c^{2-r}m^{3-r}\sum_{i=1}^{N} i^{2-r}}, \tag{5–4}$$

where

$$S_r^2=\frac{1}{(n-1)c^r m^r}\left[\sum_{j=1}^{m}\sum_{i=1}^{N}\frac{y_{ij}^2}{i^r}-\frac{\left(\sum_{j=1}^{m}\sum_{i=1}^{N} y_{ij}i^{1-r}\right)^2}{m\sum_{i=1}^{N} i^{2-r}}\right].$$

The results given in section 4 on the sign of the bias of estimator t_r apply in the general case. However, the magnitude of the relative bias (R.B.) of t_r is affected by the choice of N and m for fixed $n=mN$.

Table 4 gives the relative bias for $n=10$, and shows that it is smallest for $N=1$ (which holds for all n). Table 4 can also be used to calculate the relative bias of t_r for other values of n.

From section 4, the R.B. for n observations is given by eq (4–4), using eqs (4–3) and (2–1). Let the formula obtained from eq (4–4) with $x_i=i$ be denoted by $G(n)$. For the general case, the R.B. for $N=n/m$ equally spaced x values is found to be

$$\text{R.B.}\ (N, n)=\left(\frac{n}{n-1}\right)\left(\frac{N-1}{N}\right)G(N).$$

Table 4 gives R.B. (N, n) for $n=10$, and also approximately R.B. (N, n) for other values of n, where the approximation should be fairly good if N and n are large. More precisely,

$$\text{R.B.}\ (N, hn)=\frac{h(n-1)}{hn-1}\text{R.B.}\ (N, n).$$

TABLE 4. *Values of* $V_p(B_k)$, $E_p(t_k)$, *and R.B.* (N, n) *for* $n=10, \sigma^2=1$, $c=1$, *and* $N=n/m$

p	N	$k=0$			$k=1$			$k=2$*
		$V_p(B_k)$	$E_p(t_k)$	R.B.	$V_p(B_k)$	$E_p(t_k)$	R.B.	$V_p(B_k)$
0	10	0.002597	0.002597	0	0.003306	0.005550	68	0.01550
	5	.002273	.002273	0	.002778	.003920	41	.00732
	2	.001600	.001600	0	.001778	.002025	14	.00250
	1	.001000	.001000	0	.001000	.00100	0	.00100
1	10	.02041	.01361	−33	.01818	.01818	0	.0293
	5	.01860	.01309	−30	.01667	.01667	0	.0228
	2	.01440	.01173	−19	.01333	.01333	0	.0150
	1	.01000	.01000	0	.01000	.01000	0	.0100
2	10	.1709	.0921	−46	.1273	.0970	−24	.100
	5	.1618	.0931	−42	.1222	.0975	−20	.100
	2	.1360	.0960	−29	.1111	.0988	−11	.100
	1	.1000	.1000	0	.1000	.1000	0	.100

*When $k=2$, $E_p(t_k)=V_p(B_k)$, and R.B. $=0$ for all p.

In comparing B_k, $k=0, 1, 2$, for fixed p, notice in table 4 that the differences among the three estimators are less pronounced when N is smaller. Furthermore, table 4 illustrates the way in which the choice of spacing for x-values affects the efficiency of the estimators B_k, and t_k for fixed p and k.

Thus, it is seen in table 4 that for given p and k, $V_p(B_k)$ is smallest when $N=1$. This result is not unexpected since it is well known for $p=0$ that the best estimate of the slope of a line through the origin is obtained with all observations made at one point as far away from the origin as possible. The same is true for any $p \leq 2$, as follows. When all x_i are equal to the largest, x_n, then from eq (2–1)

$$V_p(B_k)=\sigma^2/n\, x_n^{2-p}.$$

This choice of the x_i minimizes $V_p(B_k)$, since

$$\left(\sum_{i=1}^{n} x_i^{2-k}\right)^2 \leq \left(\sum_{i=1}^{n} x_i^{2-2k+p}\right)\left(\sum_{i=1}^{n} x_i^{2-p}\right)$$

and

$$\sum_{i=1}^{n} x_i^{2-p} \leq n\, x_n^{2-p} \text{ if } p \leq 2.$$

Minimizing the variance of the estimate of B_k by choosing $N=1$ is desirable only when the experimenter does not need observations at intermediate points to check on linearity. The dependence of the variance of y_i on x_i is irrelevant, of course, when all x_i are equal; for then all observations y_i have equal variances.

I thank Churchill Eisenhart for suggesting this study and Joan R. Rosenblatt for helpful advice and assistance.

(Paper 68B2–119)