Large and small sample properties of estimators for a linear functional relation

Martin Robert Dorff

Iowa State University
LARGE AND SMALL SAMPLE PROPERTIES OF ESTIMATORS
FOR A LINEAR FUNCTIONAL RELATION

by

Martin Robert Dorff

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major Subject: Statistics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Head of Major Department

Signature was redacted for privacy.

Dean of Graduate College

Iowa State University
Of Science and Technology
Ames, Iowa
1960
# TABLE OF CONTENTS

1. INTRODUCTION ................. 1
   1.1. Statement of Problem ........ 1
   1.2. Review of Literature .......... 2
   1.3. Scope of Investigation ........ 7

2. LARGE-SAMPLE RESULTS WHEN ERRORS ARE INDEPENDENT . 9
   2.1. Heuristic Considerations Suggesting Estimators of the Slope ........ 9
   2.2. Expectations of Various Statistics Used in Section 2 .......... 15
   2.3. Asymptotic Variances and Covariances of Estimators of \( \alpha \) and \( \beta \) ......... 23
   2.4. Special Cases of 2.3: Constant Replication ........ 35
   2.5. Other Estimators of the Slope ........ 38
   2.6. Discussion of Results ........ 41

3. LARGE-SAMPLE RESULTS WHEN ERRORS ARE CORRELATED. 50
   3.1. Notation, Assumptions, and Preliminary Remarks ........ 50
   3.2. Heuristic Considerations Suggesting Estimators of the Slope ........ 52
   3.3. Expectations of Various Statistics Used in Section 3 .......... 56
   3.4. Asymptotic Variances and Covariances of Estimators of \( \alpha \) and \( \beta \) ......... 70
   3.5. Discussion of Results ........ 77

4. SMALL-SAMPLE RESULTS FOR UNREPLICATED OBSERVATIONS 82
   4.1. Introduction and Assumptions ........ 82
   4.2. Expectations of Various Statistics Used in Section 4 .......... 86
   4.3. Ratio-of-Linear-Forms Estimator ......... 93
   4.4. Least-Squares-Type Estimator ........ 115
   4.5. Discussion of Results ........ 122

5. SMALL-SAMPLE RESULTS FOR REPLICATED OBSERVATIONS 129
   5.1. Introduction and Assumptions ........ 129
   5.2. Expectations of Various Statistics Used in Section 5 ........ 130
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.3. Bias and Mean Square Error of Slope Estimators</td>
<td>143</td>
</tr>
<tr>
<td>5.4. Comparison of Bias for Estimators of $\beta$</td>
<td>158</td>
</tr>
<tr>
<td>5.5. Comparison of Mean Square Error for Estimators of $\beta$</td>
<td>167</td>
</tr>
<tr>
<td>6. CONCLUSIONS</td>
<td>170</td>
</tr>
<tr>
<td>7. BIBLIOGRAPHY</td>
<td>177</td>
</tr>
<tr>
<td>8. ACKNOWLEDGEMENT</td>
<td>183</td>
</tr>
</tbody>
</table>
1. INTRODUCTION

1.1. Statement of Problem

This dissertation is concerned with estimating the parameters of a linear functional relation between two variables $X, Y$ when both variables are observed with error. To be specific, it deals with the following model:

\[ Y_i = \alpha + \beta X_i \quad i = 1, 2, \ldots, n \]
\[ x_i = X_i + e_i \]
\[ y_i = Y_i + f_i \]

where $X_i$ and $Y_i$ are sure variables; i.e., non-random variables; while $e_i$ and $f_i$ are random variables, representing the errors of observation. Only $x_i$ and $y_i$ are observable. We want to estimate $\alpha$ and $\beta$. The scope of this investigation will be set forth in section 1.3.

Such problems occur in many contexts, of course; calibration problems typically fall into this mold. As a matter of fact, in virtually all applications of fitting a functional relation, errors of observation occur in all variables; very frequently we can neglect the errors in the independent variables in comparison with the errors in the dependent variable, but clearly we cannot always do so.

The importance of the problem of estimating the parameters of a functional relation when both variables are observed with error is too obvious to require elaboration. If
our knowledge in this area is rather scanty, it is because it is hard to come by and not because we are indifferent to its import.

1.2. Review of Literature

Classical least-squares is an old subject relative to the body of statistical knowledge, so it is not surprising to find references dating back eighty years and more to the problem of fitting constants to a functional relationship between two variables when both variables are in error. Adcock (2) in 1878 proposed fitting a straight line by minimizing the sum of squares of deviations orthogonal to the line; Karl Pearson (46) also tried this approach. Two serious objections can be raised against this method: firstly, there is no sound reason for singling out the orthogonal deviations, and secondly, the straight line so obtained is not invariant under transformations of the coordinate system.

Other attempts to solve the problem have been numerous. Prior to Wald's paper (57) in 1940, such attempts invariably rested upon a priori assumptions concerning the variance of both error variables, or at least the ratio of these variances. As early as 1879 Kummell (35) had given the weighted least-squares solution in the case where the variance ratio is known, which solution is identical with that given by maximum likelihood if the errors are assumed to be normally
distributed. Kummell's solution is invariant with respect to choice of coordinates and is consistent; it is, in fact, the only consistent estimator of the parameters employing second moments, apart from a constant factor; see Lindley (37, p. 237).

Fresh ground was broken by Wald (57) in 1940, when he showed that under certain conditions consistent estimators of the parameters of the linear functional relationship can be obtained even though the variance ratio is not known explicitly. Wald's method requires that one divide the observations into two equal groups independently of the errors; this is possible if the errors associated with one of the variables are sufficiently small, so small in fact that there is never any question as to which group contains a given observation. Wald's method was subsequently refined by Nair and Shrivastava (41), Nair and Banerjee (40), Bartlett (6), Theil and van Yzeren (55), and Gibson and Jowett (23). The refinements involve dividing the observations into three groups rather than two; the middle group of observations is ignored, and so the question of deciding how many observations to put in each group arises. The results obtained by the aforementioned writers are summarized by Madansky (38).

Another way of obtaining estimators of the parameters in question is to use so-called "instrumental sets of variables". Instrumental variables are simply additional variables closely
related to the "investigational set" of variables. For example, if one were interested in studying the relationship between the deflection of a beam and the stress at a given point on the beam, the weight used to produce the deflection would be an instrumental variable. This method has been elaborated at length by Reiersøl (47) and summarized briefly by Madansky (38). The method of grouping is actually nothing more than a special case of the method of instrumental variables wherein the instrumental variable can take on only the values -1, 0, or 1 when the observations are partitioned into three groups and -1 or 1 when the observations are partitioned into two groups. Housner and Brennan (27) and Durbin (18) both suggest that if the order (according to magnitude) of the observed values of one of the variables is the same as the order of the true values, then the order number itself would make a good instrumental variable. As far as I can tell, no one has seen fit to exploit this suggestion or to scrutinize it in detail.

This still leaves open the problem of estimating the parameters when no information is available concerning the magnitude of the errors or their relative variance and when no instrumental variable is relevant. A very substantial contribution to the whole subject was made by Reiersøl (48), who showed in 1950 that if the errors are normally distributed, then the parameters $\alpha$ and $\beta$ are non-identifiable if and
only if the true value of each variable is a constant or if the true values are normally distributed. That is to say, in either of these two situations it is possible to find more than one set of values of \( a \) and \( \beta \) which gives rise to the same distribution of \( x \) and \( y \) unless one has additional information sufficient to render the parameters identifiable.

It follows from Reiersøl's theorem that, except for the trivial case when all the \( X_i \) are exactly the same, the parameters \( a \) and \( \beta \) of the linear functional relation are always identifiable when the errors are normally distributed. They can, in fact, be estimated by the method of maximum likelihood, and Kiefer and Wolfowitz (34) have shown that the maximum likelihood estimators of these parameters are strongly consistent. This apparently came as something of a surprise to most statisticians who were familiar with Lindley's 1947 paper (37). In that paper Lindley had applied the method of maximum likelihood in the situation where the errors are normally distributed with \( \lambda \) unknown; he showed that it led to estimators of the variances which were not consistent, and he left the matter at that point, without solving for the estimators of \( a \) and \( \beta \). Kiefer and Wolfowitz in the paper just cited explain why it is that the maximum likelihood estimators of \( a \) and \( \beta \) are consistent while those of the variances and covariances are not.

Wolfowitz has developed another technique, known as the
minimum distance method, for estimating the parameters of the linear functional relation in identifiable situations. The estimates are chosen so as to minimize the "distance" between the empirical distribution function of the observations and the true distribution function of the random variables corresponding to the observations when the parameters are replaced by their estimates in the latter distribution. This, of course, requires a more-or-less arbitrary definition of distance as well as knowledge of the distribution of errors. The technique has been described at length by Wolfowitz in a series of papers (59, 60, and 61).

When replicate observations are made, the whole problem becomes far simpler, and a variety of estimators has been proposed. These are discussed, with appropriate references, in sections 2.1 and 3.2. As far as I can determine, no one has compared these estimators with regard to bias; Madansky (38) has compared the asymptotic variance of two of them, but the approximations he has made have led him to conclusions which are erroneous in my opinion. The matter is discussed at length in section 3.5 of this dissertation. Tukey (56) has given statistics which purport to estimate the variances of several slope estimators, which variances are not given, however. According to Madansky (38, p. 194) the estimates "will not be too good in the usual cases of grouping, where r (the number of replicates) is small, or when one has an
instrumental variable." Thus there seems to be virtually nothing in the way of small-sample results concerning slope estimators.

1.3. Scope of Investigation

This investigation is confined to point estimation only, which curiously enough seems to be a more formidable problem than interval estimation in this area of statistics. And since the slope is generally of greater interest than the intercept, from the standpoint of applications at least, attention has been focused largely on the former. However, large-sample results are given for estimators of \( a \) also. People seem generally satisfied to estimate the intercept, once the slope has been estimated, by imposing the condition that the fitted line pass through \((\bar{x}, \bar{y})\). This gives, of course, the consistent estimator

\[
a = \bar{y} - b\bar{x},
\]

where \( b \) is the estimated slope and \( a \) is the estimated intercept. I have contented myself, for the present at least, with this estimator of \( a \) although I have not as yet investigated its behavior in the small-sample case.

There are very many aspects of point estimation that one might study, but in a dissertation of moderate scope one must, of course, make some choice among them. I have chosen to study the bias, the mean-square-error, and the asymptotic
variance, partly because of their innate interest to the statistically unsophisticated and the statistically sophisticated alike, and partly because it seemed wise to attack first the simpler aspects of what is a formidable problem.

The small-sample variance of any statistic equals its mean-square-error minus the square of its bias. In all of the estimators studied, the square of the bias is very small compared with the mean-square-error, so that the variance is essentially equal to the latter.

Finally, it has proved necessary in the small-sample case to restrict the error of the independent variable, and for certain results, of the dependent variable as well, to a certain finite range in order that the investigation could be carried out. I feel that most workers in the physical and engineering sciences will regard the assumptions I have made concerning the range of the error as indeed very mild and will be prepared to endorse them unhesitatingly in most of their researches. On the other hand I am quite willing to concede that these assumptions may be untenable in many situations in the biological and behavioral sciences, and sometimes in the physical sciences as well. But then it is an old story that assumptions sufficiently general to satisfy one investigator are much too unspecific to satisfy another working in a different area. In the final analysis each must decide for himself what assumptions are appropriate.
2. LARGE-SAMPLE RESULTS WHEN ERRORS ARE INDEPENDENT

2.1. Heuristic Considerations Suggesting Estimators of the Slope

Suppose that \( Y_i = \alpha + \beta X_i \), \( i = 1, 2, \ldots, n \), where \( \alpha \) and \( \beta \) are unknown constants, while the \( X_i \) and \( Y_i \) are sure variables. However, it is not possible to observe either \( X_i \) or \( Y_i \), but only \( x_{it} \) and \( y_{it} \), where

\[
\begin{align*}
  x_{it} &= x_i + e_{it} \\
  y_{it}' &= y_i + f_{it}' \\
  i &= 1, 2, \ldots, n \\
  t &= 1, 2, \ldots, r_i \\
  t' &= 1, 2, \ldots, s_i
\end{align*}
\]

\[
g = \frac{s_i}{r_i}
\]

is a constant, independent of \( i \).

In this situation we have replicated observations on each \( X_i \) and \( Y_i \). The problem is to estimate \( \beta \).

We introduce now the following notation and assumptions:

\[
\begin{align*}
  R &= \Sigma r_i \\
  S &= \Sigma s_i \\
  e_i &= \frac{1}{r_i} \Sigma e_{it} \\
  f_i &= \frac{1}{s_i} \Sigma f_{it}' \\
  e_\ldots &= \frac{1}{R} \Sigma r_i e_i \\
  f_\ldots &= \frac{1}{S} \Sigma s_i f_i
\end{align*}
\]

and similarly for \( x_i \), \( x_\ldots \), \( y_i \), and \( y_\ldots \);

\[
\begin{align*}
  \bar{X} &= \frac{1}{R} \Sigma r_i X_i \\
  \bar{Y} &= \frac{1}{S} \Sigma s_i Y_i
\end{align*}
\]
\[ E(e_{it}) = 0 \quad E(e_{it}^2) = \mu_e = \sigma_e^2 \]

\[ E(e_{it}^3) = 0 \quad E(e_{it}^4) = \mu_e, \]

\( e_{it} \) and \( e_{i't'} \) are independent unless \( i = i' \) and \( t = t' \).

\[ E(f_{it}) = 0 \quad E(f_{it}^2) = \nu_f = \sigma_f^2 \]

\[ E(f_{it}^3) = 0 \quad E(f_{it}^4) = \nu_f, \]

\( f_{it} \) and \( f_{i't'} \) are independent unless \( i = i' \) and \( t = t' \).

\( e_{it} \) and \( f_{i't'} \) are independent for all \( i, i', t, \) and \( t' \).

Consider the following mean-squares, suggested by familiar analysis-of-variance procedures:

\[ s_{XXB} = \frac{1}{n-1} \sum r_i (x_{it} - x..)^2 \]

\[ s_{YYB} = \frac{1}{n-1} \sum s_i (y_{it} - y..)^2 \]

\[ s_{XYB} = \frac{1}{n-1} \sum r_i (x_{it} - x..)(y_{it} - y..) \]

\[ s_{XXW} = \frac{1}{R-n} \sum (x_{it} - x_.)^2 \]

\[ s_{YYW} = \frac{1}{S-n} \sum (y_{it} - y_.)^2 . \]

If one examines the expected mean squares, which are shown in Table 1, the following estimators of \( \beta \) suggest themselves by the fact that each converges in probability to \( \beta \) as \( n \) approaches infinity, provided that the expectation of the respective denominator is not zero. (These expectations can vanish only in the trivial case where all the \( X_i \) are all the
Table 1. Expected mean squares of components

<table>
<thead>
<tr>
<th>Mean square</th>
<th>Expected mean square</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_{XXB}$</td>
<td>$\frac{1}{n-1} \sum r_i (X_i - \bar{X})^2 + \sigma_e^2$</td>
</tr>
<tr>
<td>$s_{YYB}$</td>
<td>$\frac{1}{n-1} \sum s_i (Y_i - \bar{Y})^2 + \sigma_f^2$</td>
</tr>
<tr>
<td>$s_{XYB}$</td>
<td>$\frac{1}{n-1} \sum r_i (X_i - \bar{X})(Y_i - \bar{Y})$</td>
</tr>
<tr>
<td>$s_{XXW}$</td>
<td>$\sigma_e^2$</td>
</tr>
<tr>
<td>$s_{YYW}$</td>
<td>$\sigma_f^2$</td>
</tr>
</tbody>
</table>

same, or in the case of $b_2$, when $\beta = 0$.

$$b_1 = \frac{s_{XYB}}{s_{XXB} - s_{XXW}}$$

$$b_2 = \frac{s_{YYB} - s_{YYW}}{g s_{XYB}}$$

$$b_3 = \left[ \frac{s_{YYB} - s_{YYW}}{g(s_{XXB} - s_{XXW})} \right]^{1/2} \text{sgn } \beta$$

There is, of course, the possibility that the denominator of $b_1$, $b_2$, or $b_3$ might vanish, in which case no estimate of $\beta$ would be defined by the estimator in question. This could happen, however, only with probability zero.

For some time now statisticians have been using the analysis of variance to estimate correlation coefficients;
see Smith (54), Hazel (25), and Alexander (3). Of course, for regression models this is closely related to the estimation of slope; however, the first explicit appearance of these slope estimators that I have been able to find is in a paper by Tukey (56).

Yet another estimator is suggested by maximum-likelihood in the situation where the errors are assumed to follow independent normal distributions, the variance ratio \( \sigma_f^2/\sigma_0^2 \) is assumed to be known, say \( \lambda \), and \( r_i = s_i = 1 \). The estimator of \( \beta \) is then

\[
\hat{\beta} = \phi' \pm \left[ (\phi')^2 + \lambda \right]^{1/2}
\]

where

\[
\phi' = \frac{\Sigma(y_1 - \overline{y})^2 - \lambda \Sigma(x_1 - \overline{x})^2}{2 \Sigma(x_1 - \overline{x})(y_1 - \overline{y})}.
\]

This estimator is obtained as the solution of the equation

\[
\hat{\beta}^2 \Sigma(x_1 - \overline{x})(y_1 - \overline{y}) + \hat{\beta}[\lambda \Sigma(x_1 - \overline{x})^2 - \Sigma(y_1 - \overline{y})^2] - \lambda \Sigma(x_1 - \overline{x})(y_1 - \overline{y}) = 0.
\]

The details involved in arriving at this equation are given by Lindley (37). The plus sign is to be taken when

\[
\Sigma(x_1 - \overline{x})(y_1 - \overline{y}) > 0;
\]

the minus sign is to be taken when

\[
\Sigma(x_1 - \overline{x})(y_1 - \overline{y}) < 0.
\]

If \( \Sigma(x_1 - \overline{x})(y_1 - \overline{y}) = 0 \), the solution, and hence the maximum-likelihood estimator, is \( \hat{\beta} = 0 \), provided

\[
\lambda \Sigma(x_1 - \overline{x})^2 - \Sigma(y_1 - \overline{y})^2 \neq 0.
\]
When this last expression is also zero, maximum likelihood fails to give a determinate estimator of \( \beta \). Of course this can happen only with probability zero. This solution has appeared many times in the literature; see (5), (15), (35), (37), and (62). In (35) and (37) the estimator is given as

\[
\hat{\beta} = \varphi' + (\varphi'^2 + \lambda)^{1/2}
\]

Madansky (38) has called attention to the need for the double sign preceding the radical and has given the condition cited above for the appropriate choice of sign.

When \( \lambda \) is unknown, it can be estimated from the data provided there is replication. To be specific, a consistent estimator of \( \lambda \) is

\[
\hat{\lambda} = \frac{s_{YYW}}{s_{XXW}}.
\]

Proceeding on heuristic grounds one can construct an estimator of \( \beta \) which makes use of \( \hat{\lambda} \). For example, if \( s_i = r_i \) for every \( i \), one might estimate \( \beta \) by

\[
b_i = \varphi \pm (\varphi^2 + \hat{\lambda})^{1/2}
\]

where \( \varphi = \frac{s_{YYB} - \hat{\lambda} s_{XXB}}{2s_{XYB}} \), provided \( s_{XYB} \neq 0 \). The positive sign is to be taken when \( s_{XYB} > 0 \); the negative sign is to be taken when \( s_{XYB} < 0 \). When \( s_{XYB} = 0 \), we take \( b_i = 0 \) provided that \( \hat{\lambda} s_{XXB} - s_{YYB} \neq 0 \). If this last expression is also zero, \( b_i \) is indeterminate.
Of course \( b_4 \) is not the maximum-likelihood estimator of \( \beta \) even for the case wherein the errors are normally distributed; it might be called a pseudo maximum-likelihood estimator. The complete set of likelihood equations is very complicated when \( \lambda \) is unknown, and their explicit solution has not at this time been obtained. However, \( b_4 \) is a consistent estimator of \( \beta \) even though the errors are not normally distributed provided they are subject to the mild restrictions set down at the beginning of this section.

If \( s_i = g r_i, \ g \neq 1 \), the pseudo maximum-likelihood estimator is

\[
b_4 = \frac{g}{g} + [(\frac{g}{g})^2 + \frac{\lambda}{g}]^{1/2}
\]

provided \( s_{XYB} \neq 0 \). The plus sign is to be chosen when \( s_{XYB} > 0 \); the negative sign, when \( s_{XYB} < 0 \). If \( s_{XYB} = 0 \), we take \( b_4 = 0 \) provided \( \lambda s_{XXB} - s_{YYB} \neq 0 \). If this last expression is also zero, \( b_4 \) is indeterminate.

There is one final estimator we should like to indicate at this point:

\[
b_L = \frac{\Sigma w_i y_i}{\Sigma w_i x_i}
\]

where \( \Sigma w_i = 0 \), \( \Sigma w_i x_i \neq 0 \), and the \( w_i \) are constants, chosen in such a way that they are independent of the errors. The estimators of Wald, Bartlett, and Housner and Brennan are all special cases of this estimator. The question of choosing
the $w_i$ will be considered in detail in sections 4.3 and 4.5.

It is clear that $b_L$ is a consistent estimator of $\beta$ whenever $\Sigma w_i(X_i - \bar{X}) \neq 0$, for $\Sigma w_i Y_i$ converges in probability to $\Sigma w_i(Y_i - \bar{Y}) = \beta \Sigma w_i(X_i - \bar{X})$, while $\Sigma w_i x_i$ converges in probability to $\Sigma w_i(X_i - \bar{X})$.

2.2. Expectations of Various Statistics Used in Section 2

Expectations of various functions of the $e_{it}$, $f_{it'}$, $X_i$, and $Y_i$ appear continually throughout section 2 and elsewhere in this investigation as well. For the sake of coherence, the expectations occurring most often are collected in section 2.2.

To simplify the typography the following convention is introduced.

When more than one of the indices in the triad $(i, j, k)$ occur in a summation, the sum is to be taken only over those terms for which the indices differ. When $i$, $i'$, and $i''$ (or $j$, $j'$, and $j''$, or $k$, $k'$, and $k''$) occur in a summation, all values are assumed by each index. For example

$$\sum_{i=1}^{3} \sum_{j=1}^{3} x_{i} x_{j} = x_{1} x_{2} + x_{1} x_{3} + x_{2} x_{1} + x_{2} x_{3} + x_{3} x_{1} + x_{3} x_{2}$$

$$\sum_{i=1}^{3} \sum_{i'=1}^{3} x_{i} x_{i'} = x_{1}^{2} + x_{1} x_{2} + x_{1} x_{3} + x_{2} x_{1} + x_{2}^{2} + x_{2} x_{3} + x_{3} x_{1} + x_{3} x_{2} + x_{3}^{2}.$$
Finally, let \( \frac{n}{r_H} = \sum \frac{1}{r_i} \).

### 2.2.1. \( \text{Var} \sum r_i(e_i - e) \)

\[
\sum r_i(e_i - e) = \sum r_i e_i^2 - 2\sum r_i e_i e + R e^2 = \sum r_i e_i^2 - R e^2.
\]

\[
[\sum r_i(e_i - e)]^2 = \sum r_i^2 e_i^2 - 2R\sum r_i e_i e + R^2 e^2 + \sum r_i^2 e^2.
\]

\[
\mathbb{E}\sum r_i r_i^2 e_i^2 e_i^2 = \mathbb{E}\sum r_i^2 e_i^2 + \mathbb{E}\sum r_i^2 e_i^2 e^2.
\]

\[
\mathbb{E}\sum r_i^2 e_i^2 = \mathbb{E}\frac{r_i^2 e_i^2}{R^2} = \sum \frac{r_i^2 e_i^2}{R^2}.
\]

\[
\mathbb{E}r_i^2 e_i^2 e = \mathbb{E}\frac{r_i e_i}{R^2} = \sum \frac{r_i e_i}{R^2}.
\]

Let \( \kappa_i(z) \) denote the \( i \)th cumulant of the random variable \( z \). Then

\[
\kappa_1(e) = \frac{1}{R^4} \sum \kappa_4(r_i e_i) = \frac{1}{R^4} \sum r_i(\mu_4 - 3\mu_2^2)
\]

\[
= \frac{\mu_4 - 3\mu_2^2}{R^3}.
\]
\[
E(e_i) = \frac{\mu_4}{R^3} + \left(\frac{3}{R^2} - \frac{3}{R^3}\right) \mu_2.
\]

\[
E[\Sigma r_i(e_i - e_{..})^2]^2 = \left(\frac{n}{R_H} - \frac{2n}{R} + \frac{1}{R}\right) \mu_4
\]

\[+ \left( n^2 - \frac{3n}{R_H} + \frac{6n}{R} - \frac{3}{R} - 1 \right) \mu_2^2. \quad [1]
\]

Since \( E[\Sigma r_i(e_i - e_{..})] = (n-1) \mu_2 \),

\[
\text{Var}[\Sigma r_i(e_i - e_{..})]^2 = \left(\frac{n}{R_H} - \frac{2n}{R} + \frac{1}{R}\right) \mu_4
\]

\[+ \left( 2n - 2 - \frac{3n}{R_H} + \frac{6n}{R} - \frac{3}{R} \right) \mu_2^2.
\]

2.2.2. \text{Var} \Sigma \Sigma(e_i - e_{i..})^2

\[
\text{Var} \Sigma \Sigma(e_i - e_{i..})^2 = E(\Sigma \Sigma e_i^2 - \Sigma r_i e_i^2)^2 - [E \Sigma \Sigma(e_i - e_{i..})^2]^2.
\]

Now \( (\Sigma \Sigma e_i^2 - \Sigma r_i e_i^2)^2 = (\Sigma \Sigma e_i^2)^2 - 2 \Sigma \Sigma e_i^2 \Sigma r_i e_i^2 + (\Sigma r_i e_i^2)^2. \)

\[
\text{Var} \Sigma \Sigma e_i^2 = \Sigma (\mu_4 - \mu_2^2) = R(\mu_4 - \mu_2^2).
\]

\[
E(\Sigma \Sigma e_i^2)^2 = R(\mu_4 - \mu_2^2) + R^2 \mu_2^2 = R \mu_4 + (R^2 - R) \mu_2^2.
\]

\[
\Sigma \Sigma r_i e_i^2 = \Sigma \Sigma r_i e_i^2 + \Sigma \Sigma e_i^2 r_j e_j.
\]

\[
= \Sigma \frac{1}{r_i} \left[ \mu_4 + (r_i - 1) \mu_2^2 \right] + (R-n) \mu_2^2
\]

\[= n \mu_4 + n(R-1) \mu_2^2.
\]
\[
\text{Var } \varepsilon_i e_i^2 = \varepsilon_i^2 \text{Var } e_i^2 = \varepsilon_i^2 \left[ \frac{\mu_4}{r_i^3} - \frac{3\mu_2}{r_i^3} + \frac{2\mu_2}{r_i^2} \right]
\]
\[
= \frac{n\mu_4}{r_H^2} - \frac{3n\mu_2}{r_H^2} + 2n\mu_2^2.
\]
\[
E(\varepsilon_i e_i^2)^2 = \frac{n\mu_4}{r_H^2} - \frac{3n\mu_2}{r_H^2} + 2n\mu_2^2 + n^2\mu_2^2
\]
\[
= \frac{n\mu_4}{r_H^2} + n(n + 2 - \frac{3}{r_H})\mu_2^2.
\]

Consequently,
\[
\text{Var } \Sigma (e_{it}-e_i)^2 = R\mu_4 + (R^2-R)\mu_2^2 - 2[n\mu_4 + n(R-1)\mu_2^2]
\]
\[
+ \frac{n\mu_4}{r_H^2} + n(n + 2 - \frac{3}{r_H})\mu_2^2 + (R-n)^2\mu_2^2
\]
\[
= (\frac{n}{r_H} + R - 2n)\mu_4 + (4n - R - \frac{3n}{r_H})\mu_2^2.
\]

2.2.3. \quad \text{Var } \varepsilon_i (X_i - \overline{X}) e_i.

\[
\text{Var } \varepsilon_i (X_i - \overline{X}) e_i = E[\varepsilon_i (X_i - \overline{X}) e_i]^2
\]
\[
= E\varepsilon_i^2 (X_i - \overline{X})^2 e_i^2
\]
\[
= \sigma_e^2 \varepsilon_i^2 (X_i - \overline{X})^2.
\]
2.2.4. \[ \text{ESE}_{r_i'(e_{i'-e})}^2 = \text{ESE}_{e_{i't-e_{i't}}'}^2 = E[\Sigma e_{i'_t-e_{i't}}']^2] \]

\[ \text{ESE}_{r_i'(e_{i'-e})}^2 = n\mu_4 + n(R-1)\mu_2^2. \]

\[ \text{ESE}_{r_i'(e_{i'-e})}^2 = \frac{n\mu_4}{R} + n(n + 2 - \frac{3}{R^2})\mu_2^2. \]

\[ \text{RESE}_{r_i'(e_{i'-e})}^2 = \frac{n\mu_4}{R} + nR + 2R - 3n\mu_2^2. \]

\[ \text{RESE}_{e_{i't-e_{i't}}'}^2 = \frac{R^2}{R^2} \text{ESE}_{e_{i't-e_{i't}}'}^2 = \frac{1}{R} \text{ESE}_{e_{i't-e_{i't}}'}^2 = \frac{\mu_2^2(R - \frac{\Sigma e_{i't-e_{i't}}'}{R}) + \mu_4 + (\Sigma e_{i't-e_{i't}}' - R)\mu_2^2}{R} \]

Combining terms we obtain

\[ \text{ESE}_{r_i'(e_{i'-e})}^2 = (n - 1 - \frac{n}{R^2} + \frac{n}{R})\mu_4 \]

\[ + (R-n-3)(n-1)\mu_2^2 + 3n[\frac{1}{R^2} - \frac{1}{R}]\mu_2^2. \]

2.2.5. \[ \text{Var}_{r_i'(e_{i'-e})}\text{(r}_{i'-f_{i'-f}) \]

\[ \text{Var}_{r_i'(e_{i'-e})}\text{(r}_{i'-f_{i'-f}) = E[\Sigma(r_{i'e_{i'f_{i'-f}}}-\Sigma_{r_{i'e_{i'f_{i'-f}}}})]^2 \]

\[ = E[\Sigma r_{i'e_{i'f_{i'-f}}}.f_{i}f_{i'}-\Sigma r_{i'e_{i'f_{i'-f}}}.e_{i'f_{i'}}] \]

\[ - \Sigma r_{i'e_{i'f_{i'-f}}}.e_{i'f_{i'}}+\Sigma r_{i'e_{i'f_{i'-f}}}.f_{i}f_{i'}] \]
\[ \frac{1}{g} (\sigma^2_\epsilon \sigma^2_f - 2 \sigma^2_\epsilon \sigma^2_f + \sigma^2_\epsilon \sigma^2_f) = \frac{n-1}{g} \sigma^2_\epsilon \sigma^2_f. \]

2.2.6. \textit{Var} \( s_{XYB} \)

Let \( T = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \); then
\[ \text{Var} \ s_{XYB} = \left( \frac{1}{n-1} \right)^2 \text{Var} \ T. \]

\[ T = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) + \sum_{i=1}^n (y_i - \bar{y})e_i + \sum_{i=1}^n (x_i - \bar{x})f_i + \sum_{i=1}^n (e_i - \bar{e})(f_i - \bar{f}). \]

\[ \text{Var} \ \sum_{i=1}^n (y_i - \bar{y})e_i = \sigma^2_\epsilon \sum_{i=1}^n (y_i - \bar{y})^2. \]

\[ \text{Var} \ \sum_{i=1}^n (x_i - \bar{x})f_i = \frac{\sigma^2_f}{g} \sum_{i=1}^n (x_i - \bar{x})^2. \]

Now \( \bar{y} = \beta (x_i - \bar{x}) \).

Combining these with the results of section 2.2.5 we get
\[ \text{Var} \ s_{XYB} = \left( \frac{1}{n-1} \right)^2 \left[ (\beta^2 \sigma^2_\epsilon + \frac{1}{g} \sigma^2_f) \sum_{i=1}^n (x_i - \bar{x})^2 + \frac{n-1}{g} \sigma^2_\epsilon \sigma^2_f \right]. \]

2.2.7. \textit{Var} \( s_{XXW} \)

\[ s_{XXW} = \frac{1}{R-n} \sum (x_{it} - \bar{x}_i)^2 = \frac{1}{R-n} \sum (e_{it} - e_i)^2. \]

\[ \text{Var} \ s_{XXW} = \left( \frac{1}{R-n} \right)^2 \text{Var} \ \sum (e_{it} - e_i)^2. \]

Using the result of section 2.2.2,
\[ \text{Var} \ s_{XXW} = \left( \frac{1}{R-n} \right)^2 \left[ (\frac{n}{r_H} + R - 2n)\mu^2 + (4n - R - \frac{3n}{r_H})\mu^2_2 \right]. \]
2.2.8. \( \text{Var } s_{XXB} \)

Let \( T = \Sigma r_i(x_i - \bar{x})^2 \). Then \( s_{XXB} = \frac{1}{n-1} T \) and

\[
\text{Var } s_{XXB} = \frac{1}{(n-1)^2} \text{Var } T.
\]

\[
T = \Sigma r_i(x_i - \bar{x})^2 + 2 \Sigma r_i(x_i - \bar{x}) e_i + \Sigma r_i(e_i - \bar{e})^2.
\]

\[
\text{Var} \{2 \Sigma r_i(x_i - \bar{x}) e_i \} = 4 \sigma^2 \Sigma r_i(x_i - \bar{x})^2.
\]

Combining this with the results of section 2.2.1 and recalling that

\[
\text{Cov} \{ \Sigma r_i(x_i - \bar{x}) e_i, \Sigma r_i(e_i - \bar{e}) \} = 0
\]

we find that

\[
\text{Var } s_{XXB} = \frac{4 \sigma^2 \Sigma r_i(x_i - \bar{x})^2}{(n-1)^2} + \left( \frac{n}{r_H} - \frac{2n}{R} + \frac{1}{R} \right) \frac{\mu_4}{(n-1)^2}
\]

\[
+ \left( 2n - 2 - \frac{3n}{r_H} + \frac{6n}{R} - \frac{3}{R} \right) \frac{\mu_2^2}{(n-1)^2}.
\]

2.2.9. \( \text{Cov}(s_{XXB}, s_{XXW}) \)

\[
\text{Cov}(s_{XXB}, s_{XXW}) = \frac{1}{(n-1)(R-n)} \text{E} \Sigma r_i(x_i - \bar{x} + e_i - \bar{e})^2 \Sigma (e_i - \bar{e})^2
\]

\[- \text{E}(s_{XXB}) \text{E}(s_{XXW}).
\]

\[
\text{E} \Sigma r_i(x_i - \bar{x} + e_i - \bar{e})^2 \Sigma (e_i - \bar{e})^2 = \text{E} \Sigma r_i(x_i - \bar{x})^2 \Sigma (e_i - \bar{e})^2 + \text{E} r_i(e_i - \bar{e})^2 \Sigma (e_i - \bar{e})^2.
\]
Combining this with the result of section 2.2.4 we have

\[
\text{Cov}(s_{XXB}, s_{XXW}) = \left( \frac{1}{R-n} \right) \left[ 1 - \frac{n}{n-1} \left( \frac{1}{R} - \frac{1}{n} \right) \right] [\mu_1 - 3\mu_2^2].
\]

2.2.10. \text{Cov}(s_{XYB}, s_{XXB})

\[
\text{Cov}(s_{XYB}, s_{XXB}) = \left( \frac{1}{n-1} \right)^2 \text{E} \left[ \Sigma r_1(X_1 - \bar{X})^2 + 2 \Sigma r_1(X_1 - \bar{X})e_i \right. \\
+ \left. \Sigma r_1(e_i^t - e_i^t)^2 \right] [\Sigma r_1(X_1 - \bar{X})(Y_1 - \bar{Y}) \\
+ \Sigma r_1(e_i^t - e_i^t)f_i + \Sigma r_1(X_1 - \bar{X})f_i \\
+ \Sigma r_1(Y_1 - \bar{Y})e_i - \mu_1] - \text{E}(s_{XYB})\text{E}(s_{XXB}) .
\]

\[
\text{E}x_{r_1}(X_1 - \bar{X})^2x_{r_1}(X_1 - \bar{X})(Y_1 - \bar{Y}) = \Sigma r_1(X_1 - \bar{X})^2x_{r_1}(X_1 - \bar{X})(Y_1 - \bar{Y}).
\]

\[
\text{E}x_{r_1}(X_1 - \bar{X})e_i x_{r_1}(Y_1 - \bar{Y})e_i = \sigma_1^2 x_{r_1}(X_1 - \bar{X})(Y_1 - \bar{Y}).
\]

\[
\text{E}x_{r_1}(e_i^t - e_i^t)^2x_{r_1}(X_1 - \bar{X})(Y_1 - \bar{Y}) = (n-1)\sigma_1^2 x_{r_1}(X_1 - \bar{X})(Y_1 - \bar{Y}).
\]

All the other terms in the expansion of the bracket expressions have zero expectation. Hence

\[
\text{Cov}(s_{XYB}, s_{XXB}) = \frac{2\sigma_1^2}{(n-1)^2} \Sigma r_1(X_1 - \bar{X})(Y_1 - \bar{Y}).
\]
2.2.11. \( \text{Cov}(s_{XYB}, s_{XXW}) \)

\[
\text{Cov}(s_{XYB}, s_{XXW}) = \frac{1}{(n-1)(R-n)} \mathbb{E} \left[ \sum (e_{it} - e_i)^2 \left\{ \sum_i (x_i - \bar{x})(y_i - \bar{y}) + \sum_i (e_i - e_t) f_{i1} + \sum_i (x_i - \bar{x}) f_{i1} ight. \\
+ \left. \sum_i (y_i - \bar{y}) e_{i1} \right\} - \mathbb{E}(s_{XYB}) \mathbb{E}(s_{XXW}) \right].
\]

\[
\mathbb{E} \left[ \sum (e_{it} - e_i)^2 \sum_i (x_i - \bar{x})(y_i - \bar{y}) = (R-n) \sigma_e^2 \sum_i (x_i - \bar{x})(y_i - \bar{y}).
\]

The other terms in the expansion of the bracketed expressions have zero expectation.

\[
\text{Cov}(s_{XYB}, s_{XXW}) = \frac{\sigma_e^2}{n-1} \sum_i (x_i - \bar{x})(y_i - \bar{y}) - \frac{\sigma_e^2}{n-1} \sum_i (x_i - \bar{x})(y_i - \bar{y}) = 0.
\]

2.3. Asymptotic Variances and Covariances of Estimators of \( \alpha \) and \( \beta \)

The asymptotic variance of the various estimators suggested in section 2.1 can be obtained from their Taylor-series expansions. As \( b_1, b_2, \) and \( b_L \) are rational functions, they can all be treated as follows:

Let \( \mathbb{E}(u) = U \)

\[ \mathbb{E}(v) = V. \]

Assume that \( v \neq 0 \) and \( V \neq 0. \) We now expand \( \frac{u}{v} \) in the neighborhood of \( (U, V). \)

\[
\frac{u}{v} - \frac{U}{V} = \frac{1}{V} (u - U) + \frac{U}{V^2} (v - V) + \text{terms of higher order}
\]

in \( u \) and \( v \).
The asymptotic variance of $\frac{u}{v}$, $\text{Var}_A \frac{u}{v}$, is therefore

$$\text{Var}_A \frac{u}{v} = \frac{\text{Var} u}{v^2} + \frac{u^2 \text{Var} v}{v^4} - \frac{2u \text{Cov}(u, v)}{v^3}.$$ \hspace{1cm} [1]

The asymptotic variance of $b_3$ and $b_4$ can be obtained in a similar way and will be treated in subsequent sections.

The approach throughout section 2 and section 3 of this dissertation is due to John Gurland, who considered the leading term of $\text{Var}_A$ for $b_1$, $b_2$, $b_3$, $b_4$, and $b_L$. His results were presented at the Chicago meetings of the Institute of Mathematical Statistics in 1956 and at the national meetings in Seattle that same year but have never been published.

2.3.1. $\text{Var}_A(b_1)$

In this case,

$$u = s_{XYB},$$

$$v = s_{XXB} - s_{XXW},$$

$$U = \frac{1}{n-1} \sum (X_i - \overline{X})(Y_i - \overline{Y}),$$

$$V = \frac{1}{n-1} \sum (X_i - \overline{X})^2 .$$

$$\text{Var} v = \text{Var} s_{XXB} + \text{Var} s_{XXW} - 2 \text{Cov}(s_{XXB}, s_{XXW}).$$

From sections 2.2.7, 2.2.8, and 2.2.9
\[
\text{Var } v = \frac{4\sigma_e^2 \Sigma r_1(X_i - \bar{X})^2}{(n-1)^2} + Q_1(r_{H,R})\sigma_e^4 + Q_2(r_{H,R})\mu_4
\]

where

\[
Q_1(r_{H,R}) = \frac{1}{(n-1)^2} \left( 2n - 2 - \frac{3n}{r_H} + \frac{6n}{R} - \frac{3}{R} \right) + \frac{1}{(R-n)^2} \left( 4n - R - \frac{3n}{r_H} \right)
\]

\[
+ \frac{6}{R-n} - \frac{6n}{(R-n)(n-1)} \left( \frac{1}{r_H} - \frac{1}{R} \right)
\]

\[
Q_2(r_{H,R}) = \frac{1}{(n-1)^2} \left( \frac{n}{r_H} - \frac{2n}{R} + \frac{1}{R} \right) + \frac{1}{(R-n)^2} \left( \frac{n}{r_H} + R - 2n - \frac{2}{R-n} \right)
\]

\[
+ \frac{2n}{(R-n)(n-1)} \left( \frac{1}{r_H} - \frac{1}{R} \right)
\]

\[
\text{Cov}(u, v) = \text{Cov}(s_{XYB}, s_{XXB}) - \text{Cov}(s_{XYB}, s_{XXW})
\]

From sections 2.2.10 and 2.2.11

\[
\text{Cov}(u, v) = \frac{2\sigma_e^2}{(n-1)^2} \Sigma r_1(X_i - \bar{X})(Y_i - \bar{Y})
\]

Consequently,

\[
\text{Var}_A(b_1) = \frac{\beta^2 \sigma_e^2 + \frac{1}{g} \sigma_f^2}{\Sigma r_1(X_i - \bar{X})^2} + \frac{(n-1)\sigma_e^2 \sigma_f^2}{g[\Sigma r_1(X_i - \bar{X})^2]^2}
\]

\[
+ \frac{\beta^2 [Q_1(r_{H,R})\sigma_e^4 + Q_2(r_{H,R})\mu_4](n-1)^2}{[\Sigma r_1(X_i - \bar{X})^2]^2}
\]

Some explanation is in order here concerning our use of the term "asymptotic variance" and the symbol \( \text{Var}_A \). When we employ the usual methods such as moments, maximum likelihood, or minimum chi-square to estimate the parameters of a
distribution which does not involve a covariate, the asymptotic variance is ordinarily $O\left(\frac{1}{N}\right)$. However, when the distribution involves a covariate, it is not possible to make such a simple statement; the order depends upon the covariate. For example, in ordinary regression the asymptotic variance of the slope estimator is $\frac{\sigma^2}{\sum(X_i - \bar{X})^2}$, whose order clearly depends upon the spacing of the $X_i$.

The situation with respect to $b_1$ is similar. Suppose, for example, that the $X_i$ are uniformly spaced. Then $\sum r_i (X_i - \bar{X})^2 = O(n^3)$, provided the constants $r_i$ are independent of $n$. Consequently the first term of $\text{Var}_A(b_1)$ is $O(n^{-3})$, the second term is $O(n^{-5})$, and the third term is $O(n^{-5})$. As a second example suppose that our $X_i$ are spaced according to the pattern

$$\ldots, -3\omega, -2\omega, -1\omega, 0, 1\omega, 2\omega, 3\omega, \ldots$$

where $\omega > 0$.

Since $\int_0^1 x^{2\omega} dx = \frac{n^{2\omega+1}}{2\omega+1}$, we have $\sum r_i (X_i - \bar{X})^2 = O(n^{2\omega+1})$. Consequently, the first term of $\text{Var}_A(b_1)$ is $O\left(\frac{1}{n^{2\omega+1}}\right)$ while the second and third terms are $O\left(\frac{1}{n^{4\omega+1}}\right)$. Thus the second and third terms are of higher order than the first, but if $\omega$ is small, the orders do not differ much. We shall here rule out negative values of $\omega$, for they would imply an
infinite number of the $X_i$ squeezed into a finite interval, in which case it does not seem reasonable to talk of the asymptotic variance at all because the first term is no longer the dominating term.

In view of the anomalous connotation of "order" in connection with the asymptotic variance of the slope estimators, it seems appropriate to retain the terms containing $[\Sigma r_i (X_i - \bar{X})^2]^{-2}$ so that we can make a finer comparison of our estimators in section 2.6. When we speak of the "asymptotic variance" of $b_1$ or use the symbol $\text{Var}_A(b_1)$, it should be understood that we are including these terms. The same comments apply to $b_2$, $b_3$, and $b_4$ considered subsequently.

2.3.2. $\text{Var}_A(b_2)$

Take $u = s_{YYB} - s_{YYW}$

$$v = g s_{XYB}$$

$$U = \frac{1}{n-1} \sum a_i (y_i - \bar{Y})^2$$

$$V = \frac{g}{n-1} \Sigma r_i (X_i - \bar{X}) (Y_i - \bar{Y}) .$$

$$\text{Var} u = \text{Var} s_{YYB} + \text{Var} s_{YYW} - 2 \text{Cov}(s_{YYB}, s_{YYW}) .$$

These variances and covariances can be easily obtained by permuting the symbols in the corresponding expressions involving the $X$'s. Thus,
\[ \text{Var } u = \frac{4\sigma_f^2 \Sigma_1 (y_1 - \bar{y})^2}{(n-1)^2} + Q_1(s_H, S) \sigma_f^4 + Q_2(s_H, S) \nu \cdot \]

\[ \text{Cov}(u, v) = \frac{2\sigma_f^2}{(n-1)^2} \Sigma_1 (x_1 - \bar{x})(y_1 - \bar{y}) \cdot \]

\[ \text{Var}_A(b_2) = \frac{\beta \sigma_e^2 + \frac{1}{g} \sigma_f^2 + \frac{(n-1)\sigma_e^2 \sigma_f^2}{\Sigma_1 (x_1 - \bar{x})^2} + \frac{\beta^2 [Q_1(s_H, S) \sigma_f^4 + Q_2(s_H, S) \nu]}{\Sigma_1 (x_1 - \bar{x})^2} (n-1)^2}{(n-1)^2} \cdot \]

2.3.3. \text{Var}_A(b_3)

\[ u = s_{YYB} - s_{YYW} \]

\[ v = g(s_{XXB} - s_{XXW}) \]

\[ U = \frac{1}{n-1} \Sigma_1 (y_1 - \bar{y})^2 \]

\[ V = \frac{g}{n-1} \Sigma_1 (x_1 - \bar{x})^2 \cdot \]

Now \[ b_3 - \beta = \frac{u - U}{2U^{1/2}V^{1/2}} - \frac{(v - V)U^{1/2}}{2V^{3/2}} \]

\[ + \text{terms of higher order in } u \text{ and } v \cdot \]

Since \( u \) and \( v \) are independent in this case,

\[ \text{Var}_A(b_3) = \frac{\text{Var } u}{4UV} + \frac{U \text{ Var } v}{4V^3} \cdot \]

Consequently,
\[ \text{Var}_A(b_3) = \frac{\beta^2 \sigma_e^2 + \frac{1}{g^2} \sigma_f^2}{\Sigma_1(X_i - \bar{X})^2} + \frac{\beta^2 [Q_1(rH,R) \sigma_f^2 + Q_2(rH,R) \mu_4](n-1)^2}{4[\Sigma_1(X_i - \bar{X})^2]^2} \]
\[ + \frac{[Q_1(s_H,S) \sigma_f + Q_2(s_H,S) \mu_4](n-1)^2}{4\beta^2 \sigma^2 [\Sigma_1(X_i - \bar{X})^2]^2}. \]  

2.3.4. \text{Var}_A(b_4)

When \( s_{XYB} \neq 0 \),
\[ b_4 = \frac{g}{\beta} + \left[ \left( \frac{g}{\beta} \right)^2 + \frac{\lambda}{g} \right]^{1/2} \]
where \( \lambda = \frac{s_{YYW}}{s_{XXW}} \)
\[ g = \frac{s_{YYB} - \hat{\lambda} s_{XXB}}{2s_{XYB}} = \frac{s_{YYB} s_{XXW} - s_{YYW} s_{XXB}}{2s_{XXW} s_{XYB}} \]

When \( s_{XYB} = 0 \), \( b_4 = 0 \) provided \( \hat{\lambda} s_{XXB} - s_{YYB} \neq 0 \). If this latter expression is also zero, we do not define \( b_4 \).

Now \( \frac{g}{\beta} \) converges in probability to
\[ \frac{\sigma_e^2 \left[ \frac{1}{n-1} \Sigma_1(Y_i - Y)^2 + \sigma_f^2 \right] - \sigma_f^2 \left[ \frac{1}{n-1} \Sigma_1(X_i - \bar{X})^2 + \sigma_e^2 \right]}{2 \beta \sigma_e^2 \Sigma_1(X_i - \bar{X})^2} = \frac{g \beta^2 \sigma_e^2 - \sigma_f^2}{2 \beta \sigma_e^2} \]
and \( \left( \frac{g}{\beta} \right)^2 + \frac{\lambda}{g} \) converges in probability to
\[ \frac{g^2 \beta^2 \sigma_e^2 + 2 \beta \sigma_e^2 \sigma_f^2 + \sigma_f^2}{4 \beta^2 \sigma_e^2 \sigma_f^2} + \frac{\sigma_f^2}{g \sigma_e^2} = \frac{(g \beta^2 \sigma_e^2 + \sigma_f^2)^2}{4 \beta^2 \sigma_e^2 \sigma_f^2}. \]
Therefore $b_4$ converges in probability to
\[
\frac{g\beta^2\sigma_e^2 - \sigma_f^2}{2\beta g\sigma_e^2} + \frac{g\beta^2\sigma_e^2 + \sigma_f^2}{2\beta g\sigma_e^2} = \beta.
\]

Consider $b_4 = F(s_{YW}, s_{XYB}, s_{XXB}, s_{XXW}, s_{XYB})$. Denote $E(s_{YW})$
by $S_{YYB}$ and the other four mean squares analogously. We now
expand in a Taylor series about $S_{YYB}, S_{YW}, S_{XXB}, S_{XXW},$ and $S_{XYB}$.

\[
b_4 - \beta = \frac{n-1}{(g\beta^2\sigma_e^2 + \sigma_f^2)\Sigma_i(X_i - \overline{X})^2} [\beta^2 \sigma_e^2 (s_{YYB} - s_{YYB}) - \beta^2 \sigma_e^2 (s_{YW} - s_{YW})]
\]

\[+ \beta^2 \sigma_e^4 \text{Var } s_{YYB} + \beta^2 \sigma_f^4 \text{Var } s_{YW} + \beta^2 \sigma_f^4 \text{Var } s_{XXB} + \beta^2 \sigma_f^4 \text{Var } s_{XXW}
\]

\[+ (\sigma_f^2 - g\beta^2 \sigma_e^2)(s_{XYB} - s_{XYB})] + \text{terms of higher order.}
\]

\[
\text{Var}_A(b_4) = \frac{(n-1)^2}{(g\beta^2\sigma_e^2 + \sigma_f^2)^2 [\Sigma_i(X_i - \overline{X})^2]^2} [\beta^2 \sigma_e^4 \text{Var } s_{YYB}
\]

\[+ \beta^2 \sigma_e^4 \text{Var } s_{YW} + \beta^2 \sigma_f^4 \text{Var } s_{XXB} + \beta^2 \sigma_f^4 \text{Var } s_{XXW}
\]

\[+ (\sigma_f^2 - g\beta^2 \sigma_e^2)^2 \text{Var } s_{XYB} - 2\beta^2 \sigma_e^2 \text{Cov}(s_{YYB}, s_{YW})
\]

\[+ 2\beta^2 \sigma_f^2 \text{Cov}(s_{XXB}, s_{XXW})].
\]

\[
\text{Var}_A(b_4) = \frac{\beta^2 \sigma_e^2}{\Sigma_i(X_i - \overline{X})^2} + \frac{1}{\Sigma_i(X_i - \overline{X})^2} \frac{(g\beta^2\sigma_e^2 - \sigma_f^2)^2}{(g\beta^2\sigma_e^2 + \sigma_f^2)^2} \frac{(n-1)^2 \sigma_f^2}{\Sigma_i(X_i - \overline{X})^2]^2}
\]

\[+ \frac{\sigma_f^4}{(g\beta^2\sigma_e^2 + \sigma_f^2)^2} \frac{\beta^2 [Q_1(r_H, R) \sigma_e^4 + Q_2(s_H, R) \mu_4]}{[\Sigma_i(X_i - \overline{X})^2]^2} (n-1)^2
\]

\[+ \frac{\sigma_e^4}{(g\beta^2\sigma_e^2 + \sigma_f^2)^2} \frac{\beta^2 [Q_1(s_H, S) \Sigma_i + Q_2(s_H, S) \mu_4]}{[\Sigma_i(X_i - \overline{X})^2]^2} (n-1)^2.
\]
2.3.5. \( \text{Var}_A(b_L) \)

We take \( u = \sum w_i y_i \).

\[ v = \sum w_i x_i \]

\[ U = \sum w_i Y_i \]

\[ V = \sum w_i X_i \].

Assume \( v \neq 0 \) and \( V \neq 0 \).

Then \( \text{Var} u = \sigma_f^2 \sum \frac{w_i^2}{s_i} \)

\[ \text{Var} v = \sigma_e^2 \sum \frac{w_i^2}{r_i} \].

Consequently,

\[ \text{Var}_A(b_L) = \frac{\sigma_f^2 \sum \frac{w_i^2}{s_i} + \beta^2 \sigma_e^2 \sum \frac{w_i^2}{r_i}}{(\sum w_i X_i)^2} \]

\[ = \left( \frac{\sigma_f^2}{\sigma_e^2} \right) \sum \frac{w_i^2}{r_i} \left( \frac{1}{(\sum w_i X_i)^2} \right) \].

It is interesting to consider the choice of the \( w_i \) that minimizes \( \text{Var}_A(b_L) \). The \( w_i \) can be obtained as follows:

\[ \sum \frac{w_i^2}{r_i} \]

Let \( F = \frac{\sum w_i}{(\sum w_i X_i)^2} \).

Since \( \sum w_i = 0 \), we can also write
Setting the $i^{th}$ partial derivative equal to zero we have
\[ \frac{w_i}{r_i} \sum w_i (x_i - \overline{x}) - (x_i - \overline{x}) \sum \frac{w_i^2}{r_i} = 0. \]

Solving this set of $n$ equations we obtain
\[ w_i = c r_i (x_i - \overline{x}), \]
where $c$ is any arbitrary constant except zero. Of course, we have found the unrestricted maximum of $F$, but since the restriction is satisfied by the $w_i$, this is also the maximum subject to the restriction $\sum w_i = 0$. Denoting the estimator which employs these weights by $b_0$, we have
\[ \text{Var}_A(b_0) = \frac{\beta \sigma_e^2 + \frac{1}{\varepsilon} \sigma_f^2}{\sum r_i (x_i - \overline{x})^2}. \]

2.3.6. **Comment on leading term of Var$_A$**

We should like to point out at this time that the leading term of Var$_A$ is precisely the same for $b_1$, $b_2$, $b_3$, $b_4$, and $b_0$. This fact was emphasized by Gurland at Chicago and at Seattle but has never appeared in print, so far as we are aware.

When the leading term is of appreciably lower order
than the remaining terms of $\text{Var}_A$, it certainly seems reasonable to neglect the higher order terms and to conclude that all the estimators have the same asymptotic variance. As we have pointed out in section 2.3.1, some spacings of the $X_i$ will cause the second and third terms of $\text{Var}_A$ to have an order of magnitude which does not greatly exceed that of the first term. It therefore seems worthwhile to retain all the terms in order to arrive at a more delicate comparison of the various estimators.

2.3.7. Asymptotic variance of intercept estimators

The estimator of $\alpha$ is almost invariably taken to be

$$a = \bar{y} - bx$$

inasmuch as it is a consistent estimator of $\alpha$. We shall expand $a$ in a Taylor series about $(\bar{X}, \bar{Y}, \beta)$.

$$a - \alpha = (\bar{y} - \bar{Y}) - \beta(x - \bar{X}) - \bar{X}(b - \beta)$$

$$+ \text{terms of higher order in } x, y, \text{ and } b.$$ 

$$\text{Var}_A(a) = \frac{\sigma_f^2}{R} + \beta^2 \frac{\sigma_e^2}{R} + \bar{X}^2 \text{Var}_A(b)$$

$$- 2\bar{X} \text{ Cov}_A(b, \bar{y}) + 2\beta \bar{X} \text{ Cov}_A(b, x).$$

Now $b$ depends only on $s_{XXB}, s_{XXW}, s_{XYB}, s_{YYB},$ and $s_{YYB}$. Consequently we write $b$ as a function of these:

$$b = F(s_{XXB}, s_{XXW}, s_{XYB}, s_{YYB}, s_{YYW}).$$
Denote \( \frac{2F}{\partial s_{XXB}} \), evaluated at the point
\[
(s_{XXB}, s_{XXW}, s_{XYB}, s_{YYB}, s_{YYW})
\]
by \( \left( \frac{\partial F}{\partial s_{XXB}} \right)_0 \), and similarly for the other derivatives.

\[
b - \beta = \left( \frac{\partial F}{\partial s_{XXB}} \right)_0 (s_{XXB} - s_{XXW}) + \left( \frac{\partial F}{\partial s_{XXW}} \right)_0 (s_{XXW} - s_{XXB})
\]
\[
+ \left( \frac{\partial F}{\partial s_{XYB}} \right)_0 (s_{XYB} - s_{XYW}) + \left( \frac{\partial F}{\partial s_{XYW}} \right)_0 (s_{XYW} - s_{XYB})
\]
\[
+ \left( \frac{\partial F}{\partial s_{YYW}} \right)_0 (s_{YYW} - s_{YYB}) + \text{terms of higher order}
\]
in \( s_{XXB}, s_{XXW}, s_{XYB}, s_{YYB}, s_{YYW} \).

\[
\text{Cov}_A(b, x_{..}) = \left( \frac{\partial F}{\partial s_{XYB}} \right)_0 \text{Cov}(s_{XYB}, x_{..}).
\]

\[
\text{Cov}(s_{XYB}, x_{..}) = \text{Cov}(s_{XYB}, e_{..})
\]
\[
= \frac{r}{n-1} \text{Ee} \cdot \Sigma [(X_i - \overline{X}) + (e_i - e_{..})][(Y_i - \overline{Y}) + (f_i - f_{..})].
\]

\[
\text{Ee} \cdot \Sigma (X_i - \overline{X})(Y_i - \overline{Y}) = 0.
\]

\[
\text{Ee} \cdot \Sigma (X_i - \overline{X})(f_i - f_{..}) = 0.
\]

\[
\text{Ee} \cdot \Sigma (Y_i - \overline{Y})(e_i - e_{..}) = \Sigma (Y_i - \overline{Y})\text{Ee} \cdot (e_i - e_{..}) = 0.
\]

\[
\text{Ee} \cdot \Sigma (e_i - e_{..})(f_i - f_{..}) = 0.
\]

Consequently, \( \text{Cov}_A(b, x_{..}) = 0 \). Similarly, \( \text{Cov}_A(b, y_{..}) = 0 \).
Therefore,

\[ \text{Var}_A(a) = \frac{\sigma_f^2}{\bar{e} R} + \beta^2 \frac{\sigma_e^2}{R^2} + \bar{x}^2 \text{Var}_A(b). \]

It is a simple matter then to write down the asymptotic variable for \( a \) once that of \( b \) has been obtained.

2.3.8. Asymptotic covariance of slope and intercept estimators

\[ (a - a)(b - \beta) = (y - \bar{y})(b - \beta) - \beta(x - \bar{x})(b - \beta) - \bar{x}(b - \beta)^2 \]

\[ + \text{terms of higher order in } x, y, \text{ and } b. \]

Therefore,

\[ \text{Cov}_A(a, b) = -\bar{x} \text{Var}_A(b). \]

2.4. Special Cases of 2.3: Constant Replication

The formulas of section 2.3 become much simpler when \( r_i = r \) and \( s_i = s \), where \( r \) and \( s \) are constants independent of \( i \). The asymptotic variances may be obtained from the previous results by substituting \( r_n \) for \( R \), \( s_n \) for \( S \), \( r \) for \( R_H \) and \( s \) for \( S_H \). Under these conditions

\[ Q_2(r_H, R) = Q_2(s_H, S) = 0. \]

\[ Q_1(r_H, R) = \frac{2(nr-1)}{n(n-1)(r-1)}. \]

\[ Q_2(s_H, S) = \frac{2(ns-1)}{n(n-1)(s-1)}. \]

Consequently,
\[
\text{Var}_A(b_1) = \frac{\beta^2 \sigma^2_e + \frac{1}{g^2} \sigma^2_f}{r \Sigma(X_1 - \bar{X})^2} + \frac{(n-1) \sigma^2_e \sigma^2_f}{g^2 \Sigma(X_1 - \bar{X})^2} + \frac{2 \beta^2 (n-1)(nr-1) \sigma^4_e}{nr^2 (r-1) \Sigma(X_1 - \bar{X})^2}
\]

\[
\text{Var}_A(b_2) = \frac{\beta^2 \sigma^2_e + \frac{1}{g^2} \sigma^2_f}{r \Sigma(X_1 - \bar{X})^2} + \frac{(n-1) \sigma^2_e \sigma^2_f}{g^2 \Sigma(X_1 - \bar{X})^2} + \frac{2(n-1)(ns-1) \sigma^4_f}{ns^2 (s-1) \beta^2 \Sigma(X_1 - \bar{X})^2}
\]

\[
\text{Var}_A(b_3) = \frac{\beta^2 \sigma^2_e + \frac{1}{g^2} \sigma^2_f}{r \Sigma(X_1 - \bar{X})^2} + \frac{\beta^2 (n-1)(nr-1) \sigma^4_e}{2nr^2 (r-1) \Sigma(X_1 - \bar{X})^2} + \frac{(n-1)(ns-1) \sigma^4_f}{2ns^2 (s-1) \beta^2 \Sigma(X_1 - \bar{X})^2}
\]

\[
\text{Var}_A(b_4) = \frac{\beta^2 \sigma^2_e + \frac{1}{g^2} \sigma^2_f}{r \Sigma(X_1 - \bar{X})^2} + \frac{(g\beta^2 \sigma^2_e - \sigma^2_f)^2}{(g\beta^2 \sigma^2_e + \sigma^2_f)^2} + \frac{(n-1) \sigma^2_e \sigma^2_f}{g^2 \Sigma(X_1 - \bar{X})^2} + \frac{2 \beta^2 (n-1)(nr-1) \sigma^4_e}{nr^2 (r-1) \Sigma(X_1 - \bar{X})^2}
\]

\[
\text{Var}_A(b_L) = \frac{(\frac{\sigma^2_f}{g^2} + \beta^2 \sigma^2_e) \Sigma w^2}{r \Sigma w_i X_i}
\]

\[
\text{Var}_A(b_0) = \frac{\sigma^2_f}{r \Sigma(X_1 - \bar{X})^2} + \frac{\beta^2 \sigma^2_e}{r \Sigma(X_1 - \bar{X})^2}
\]

If we further assume that \( r = s \), these expressions become
\[ \text{Var}_A(b_1) = \frac{\beta^2 \sigma_e^2 + \sigma_f^2}{r^2(X_1 - \bar{X})^2} + \frac{(n-1)\sigma_e^2 \sigma_f^2}{r^2(\Sigma(X_1 - \bar{X})^2)^2} + \frac{2\beta^2(n-1)(nr-1)\sigma_e^4}{nr^2(r-1)[\Sigma(X_1 - \bar{X})^2]^2}. \]

\[ \text{Var}_A(b_2) = \frac{\beta^2 \sigma_e^2 + \sigma_f^2}{r^2(X_1 - \bar{X})^2} + \frac{(n-1)\sigma_e^2 \sigma_f^2}{r^2(\Sigma(X_1 - \bar{X})^2)^2} + \frac{2(n-1)(nr-1)\sigma_f^4}{nr^2(r-1)\beta^2[\Sigma(X_1 - \bar{X})^2]^2}. \]

\[ \text{Var}_A(b_3) = \frac{\beta^2 \sigma_e^2 + \sigma_f^2}{r^2(X_1 - \bar{X})^2} + \frac{\beta^2(n-1)(nr-1)\sigma_e^4}{2nr^2(r-1)[\Sigma(X_1 - \bar{X})^2]^2} \]

\[ + \frac{(n-1)(nr-1)\sigma_f^4}{2nr^2(r-1)\beta^2[\Sigma(X_1 - \bar{X})^2]^2}. \]

\[ \text{Var}_A(b_4) = \frac{\beta^2 \sigma_e^2 + \sigma_f^2}{r^2(X_1 - \bar{X})^2} + \frac{(\beta^2 \sigma_e^2 - \sigma_f^2)^2}{(\beta^2 \sigma_e^2 + \sigma_f^2)^2} \frac{(n-1)\sigma_e^2 \sigma_f^2}{r^2(\Sigma(X_1 - \bar{X})^2)^2} \]

\[ + \frac{\sigma_f^4}{(\beta^2 \sigma_e^2 + \sigma_f^2)^2} \frac{2\beta^2(n-1)(nr-1)\sigma_e^4}{nr^2(r-1)[\Sigma(X_1 - \bar{X})^2]^2} \]

\[ + \frac{\sigma_e^4}{(\beta^2 \sigma_e^2 + \sigma_f^2)^2} \frac{2(n-1)(nr-1)\sigma_f^4}{nr^2(r-1)\beta^2[\Sigma(X_1 - \bar{X})^2]^2}. \]

\[ \text{Var}_A(b_L) = \frac{\beta^2 \sigma_e^2 + \sigma_f^2}{r} \frac{\Sigma w_1^2}{(\Sigma w_1 X_1)^2}. \]

\[ \text{Var}_A(b_0) = \frac{\beta^2 \sigma_e^2 + \sigma_f^2}{r^2(X_1 - \bar{X})^2}. \]
2.5. Other Estimators of the Slope

In addition to the estimators already suggested, there are other simple possibilities. In particular,

\[
\frac{\sum(x_i - \bar{x})(y_i - \bar{y})^2}{\sum(x_i - \bar{x})^2(y_i - \bar{y})}, \quad \frac{\sum(y_i - \bar{y})(x_i - \bar{x})^2}{\sum(x_i - \bar{x})^3}, \quad \text{and} \quad \frac{\sum(y_i - \bar{y})^3}{\sum(x_i - \bar{x})(y_i - \bar{y})^2}
\]

are all consistent estimators of \( \beta \) provided that the respective denominators do not have expectation zero. Furthermore, none of them require replication!

We shall show now that the first of these is consistent; the proof that the other two are consistent is quite similar.

\[
E\sum(x_i - \bar{x})(y_i - \bar{y})^2 = \sum E(x_i - \bar{x})E(y_i - \bar{y})^2
\]

\[
= \Sigma(X_i - \bar{X})[(Y_i - \bar{Y})^2 + \frac{n-1}{n} \sigma^2_i]
\]

\[
= \beta^2 \Sigma(X_i - \bar{X})^3
\]

\[
E\sum(x_i - \bar{x})^2(y_i - \bar{y}) = \beta \Sigma(X_i - \bar{X})^3.
\]

Consequently \( \frac{\sum(x_i - \bar{x})(y_i - \bar{y})^2}{\sum(x_i - \bar{x})^2(y_i - \bar{y})} \) converges in probability to \( \beta \).

In view of the fact that these estimators do not require replication, it is of interest to examine their asymptotic variance. We shall obtain the leading term of \( \text{Var}_A(b_c) \) where
\[ b_c = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})^2}{\Sigma(x_i - \bar{x})^2(y_i - \bar{y})} \]

by using equation [2.3 - 1]. (Henceforth equations will be designated by the section number first, followed by the number of the equation within that section, all enclosed within brackets.) We take

\[ u = \Sigma(x_i - \bar{x})(y_i - \bar{y})^2 \]
\[ v = \Sigma(x_i - \bar{x})^2(y_i - \bar{y}) \]
\[ U = \beta^2 \Sigma(x_i - \bar{x})^3 \]
\[ V = \beta \Sigma(x_i - \bar{x})^3 . \]

Now \( \text{Var } u = \Sigma \Sigma(x_i - \bar{x})^2(y_i - \bar{y})^4 \)
\[ + \Sigma \Sigma(x_i - \bar{x})(x_j - \bar{x})(y_i - \bar{y})^2(y_j - \bar{y})^2 - U^2 . \]
\[ \Sigma \Sigma(x_i - \bar{x})^2(y_i - \bar{y})^4 = \beta^4 \Sigma(x_i - \bar{x})^6 + \frac{n-1}{n} (\sigma_x^2 \beta^4 + 6 \sigma_x^2 \beta^2 \Sigma(x_i - \bar{x}))^4 \]
\[ + d_1 \Sigma(x_i - \bar{x})^2 + d_2 , \]

where \( d_1 = 6 \left( \frac{n-1}{n} \right)^2 \sigma_x^2 \sigma_y^2 \beta^2 + E(f_i - \bar{f})^4 \)
and \( d_2 = (n-1) \sigma_x^2 E(f_i - \bar{f})^4 . \)

\[ \Sigma \Sigma(x_i - \bar{x})(x_j - \bar{x})(y_i - \bar{y})^2(y_j - \bar{y})^2 = \beta^4 \Sigma(x_i - \bar{x})^3(x_j - \bar{x})^3 \]
\[ - \frac{2(n-1)}{n} \sigma_x^2 \beta^2 \Sigma(x_i - \bar{x})^4 - \frac{1}{n} (4 \sigma_x^2 \beta^2 + \sigma_y^2 \beta^4) \Sigma(x_i - \bar{x})^2(x_j - \bar{x})^2 \]
\[ + d_3 \Sigma(x_i - \bar{x})^2 + d_4 , \]
where \( d_3 \) and \( d_4 \) depend upon \( \sigma_e, \sigma_f, \beta, \) and the \( f_i \) but not upon the \( X_i \). (All the \( d_i \) used subsequently exhibit the property of depending upon \( \sigma_e, \sigma_f, \beta, \) the \( f_i \) and the \( e_i \) but not the \( X_i \).) Consequently,

\[
\text{Var } u = \frac{1}{n} \left( \sigma_e^2 \sigma_f^4 + 4 \sigma_f^2 \beta^2 \right) \left( (n-1) \Sigma (X_i - \bar{X})^4 - \Sigma (X_i - \bar{X})^2 (X_j - \bar{X})^2 \right) \\
+ d_5 \Sigma (X_i - \bar{X})^2 + d_6.
\]

It may similarly be shown that

\[
\text{Var } v = \frac{1}{n} \left( 4 \sigma_e^2 \beta^2 + \sigma_f^2 \right) \left( (n-1) \Sigma (X_i - \bar{X})^4 - \Sigma (X_i - \bar{X})^2 (X_j - \bar{X})^2 \right) \\
+ d_7 \Sigma (X_i - \bar{X})^2 + d_8.
\]

Now \( \text{Cov}(u, v) = E\Sigma (x_i - \bar{x})^3 (y_j - \bar{y})^3 + E\Sigma (x_i - \bar{x}) (x_j - \bar{x})^2 (y_j - \bar{y})^2 (y_j - \bar{y}) \)

\[= \text{UV}.
\]

\[E\Sigma (x_i - \bar{x})^3 (y_j - \bar{y})^3 = \beta^3 \Sigma (X_i - \bar{X})^6
\]

and

\[E\Sigma (x_i - \bar{x}) (x_j - \bar{x})^2 (y_j - \bar{y})^2 = \beta^3 \Sigma (X_i - \bar{X})^3 (X_j - \bar{X})^3
\]

\[- \frac{n-1}{n} (\sigma_e^2 \beta^3 + \sigma_f^2 \beta) \Sigma (X_i - \bar{X})^4 - \frac{2}{n} (\sigma_e^2 \beta^3 + \sigma_f^2 \beta) \Sigma (X_i - \bar{X})^2 (X_j - \bar{X})^2
\]

\[+ d_9 \Sigma (X_i - \bar{X})^2 + d_{10}.
\]

Consequently,

\[\text{Cov}(u, v) = - \frac{1}{n} (\sigma_e^2 \beta^3 + \sigma_f^2 \beta) \left( (n-1) \Sigma (X_i - \bar{X})^4 + 2 \Sigma (X_i - \bar{X})^2 (X_j - \bar{X})^2 \right)
\]

\[+ d_{11} \Sigma (X_i - \bar{X})^2 + d_{12}.
\]

From equation [2.3 - 1] we obtain for the leading term of
The factor
\[ \frac{1}{n}[\Sigma(X_i - \bar{X})^3]^2 [7(n-1)\Sigma(X_i - \bar{X})^4 - \Sigma\Sigma(X_i - \bar{X})^2(X_j - \bar{X})^2] \]

is very much larger than the factor \([\Sigma(X_i - \bar{X})^2]^{-1}\), which plays the corresponding role in the estimators discussed in section 2.3, unless the \(X_i\) are very skew indeed. For this reason I feel that \(b_c\) is not likely to prove of much practical interest.

We have not pursued the investigation of the other two estimators mentioned above because of the conviction that they would also have large asymptotic variances.

One can construct similar estimators of \(\beta\) which require no replication by using forms of any odd order in numerator and denominator, but we feel that they would exhibit even more unstability than those just discussed. We have not, however, actually investigated such estimators as they do not seem promising.

### 2.6. Discussion of Results

We turn now to a comparison of the asymptotic variances of the various slope estimators when \(r_i = s_i = r\). The esti-
mators $b_1$ and $b_2$ are easy to compare, for

$$\text{Var}_A(b_1) - \text{Var}_A(b_2) = \frac{2(n-1)(nr-1)}{nr^2(r-1)[\Sigma(x_i^2 - \bar{X})^2]^2} \left( \beta^2 \frac{\sigma_f^2}{\sigma_e^2} - \frac{\sigma_f^2}{\beta^2} \right).$$

Consequently,

$$\text{Var}_A(b_1) - \text{Var}_A(b_2) > 0 \quad \text{when} \quad \beta^2 > \lambda$$

$$\text{Var}_A(b_1) - \text{Var}_A(b_2) < 0 \quad \text{when} \quad \beta^2 < \lambda.$$

In order to facilitate the comparison of $b_3$ with $b_1$ and $b_2$ we shall make use of the fact that

$$\lim_{n \to \infty} \frac{nr-1}{n(r-1)} = \frac{r}{r-1}$$

to justify approximating $\frac{nr-1}{n(r-1)}$ by $\frac{r}{r-1}$. Making this approximation we have

$$\text{Var}_A(b_1) - \text{Var}_A(b_3) = \frac{n-1}{r^2[\Sigma(x_i^2 - \bar{X})^2]^2} \left( \sigma_e^2 \sigma_f^2 + \frac{2r}{r-1} \beta^2 \sigma_f^2 \right)$$

$$- \frac{r}{r-1} \frac{\sigma_f^2}{\sigma_e^2} - \frac{r}{r-1} \frac{\sigma_f^2}{\beta^2}.$$

Thus, $\text{Var}_A(b_1) - \text{Var}_A(b_3) > 0$ when

$$\sigma_e^2 \sigma_f^2 + \frac{3}{2} \frac{r}{r-1} \beta^2 \sigma_f^2 \sigma_e^2 - \frac{r}{r-1} \frac{\sigma_f^2}{2\beta^2} > 0,$$

that is, when $\left( \frac{\lambda}{\beta^2} \right)^2 < \frac{r-1}{r} + \sqrt{\left( \frac{r-1}{r} \right)^2 + 3}$. This criterion is evaluated for a few typical values of $r$ in Table 2. The import of this table is that the asymptotic variance of $b_3$ is less than that of $b_1$ only when
Table 2. Evaluation of \( \left( \frac{\lambda}{\beta^2} \right)_0 \) = \( \left[ \frac{r-1}{r} + \sqrt{\left(\frac{r-1}{r}\right)^2 + 3} \right]^{1/2} \)

<table>
<thead>
<tr>
<th>( r )</th>
<th>( \left( \frac{\lambda}{\beta^2} \right)_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.52</td>
</tr>
<tr>
<td>3</td>
<td>1.59</td>
</tr>
<tr>
<td>5</td>
<td>1.65</td>
</tr>
<tr>
<td>( \infty )</td>
<td>1.73</td>
</tr>
</tbody>
</table>

\[
\beta^2 > \frac{\lambda}{\left( \frac{\lambda}{\beta^2} \right)_0}.
\]

In very much the same way we can compare \( b_2 \) with \( b_3 \);

\[
\text{Var}_A(b_2) - \text{Var}_A(b_3) > 0 \text{ when }
\]

\[
\sigma^2_e \sigma^2_f + \frac{3}{2} \frac{r}{r-1} \frac{\sigma^4_f}{2\beta^2} - \frac{1}{2} \frac{r}{r-1} \beta^2 \sigma^4_e > 0 ;
\]

that is, when \( \left( \frac{\lambda}{\beta^2} \right)^{-2} < \frac{r-1}{r} + \sqrt{\left(\frac{r-1}{r}\right)^2 + 3} \). Thus, the asymptotic variance of \( b_3 \) is less than that of \( b_2 \) only when

\[
\beta^2 < \left( \frac{\lambda}{\beta^2} \right)_0 \lambda .\]

We might summarize the situation diagrammatically as follows:
In this region $\text{Var}_A(b_2) < \text{Var}_A(b_3)$
\[ \lambda < \frac{1}{\beta^2} \left( \frac{\lambda}{\beta^2} \right)_0 \]

In this region $\text{Var}_A(b_3) < \text{Var}_A(b_1)$
\[ \frac{1}{\left( \frac{\lambda}{\beta^2} \right)_0} < \frac{\lambda}{\beta^2} < \left( \frac{\lambda}{\beta^2} \right)_0 \]

In this region $\text{Var}_A(b_1) < \text{Var}_A(b_2)$
\[ \frac{\lambda}{\beta^2} < \frac{1}{\left( \frac{\lambda}{\beta^2} \right)_0} \]

Comparisons with $b_4$ are somewhat more involved because $\beta$ itself enters as well as $\beta^2/\lambda$. Again approximating $\frac{n(r-1)}{n(r-1)}$ by $\frac{r}{r-1}$, we have

\[ \text{Var}_A(b_1) - \text{Var}_A(b_4) = \frac{n-1}{r^2 \left[ \Sigma (x_i - \bar{x})^2 \right]^2} \left[ \sigma_e^2 \frac{\lambda}{\sigma_r^2} + 2 \beta^2 \sigma_r^4 \right] \frac{r}{r-1} \]

\[ - \left( \frac{\beta^2 - \lambda}{\beta^2 + \lambda} \right)^2 \frac{\sigma_e^2}{\sigma_r^2} - \frac{2 \lambda^2}{(\beta^2 + \lambda)^2} \left( \frac{\lambda^4}{\sigma_r^4} + \frac{\sigma_e^4}{\beta^2} \right) \frac{r}{r-1} \]

The factor involving $\sigma_e^4$, $\sigma_r^4$, $\lambda$, and $\beta$ can be written

\[ \left[ \frac{\lambda + 2 \beta^2}{r-1} - \left( \frac{\beta^2 - \lambda}{\beta^2 + \lambda} \right)^2 \right] \frac{\lambda^4}{\sigma_e^4} = G, \text{ say.} \]

We are interested in the sign of $G$ for various values of $\lambda$, $r$, and $\beta$; in particular, we shall investigate the sign of $G$ when $\frac{\lambda}{\beta^2} > 1$, for it is then that $\text{Var}_A(b_1)$ is less than $\text{Var}_A(b_2)$. For given values of $\frac{\lambda}{\beta^2}$ and $r$ it is possible to find $\beta_0$ such that $G > 0$ when $|\beta| > \beta_0$. In order to do this
one need simply solve the equation \( G = 0 \). The results for representative values of \( \frac{\lambda}{\beta^2} \) and \( r \) are shown in Table 3. The import of this table is that the asymptotic variance of \( b_1 \) is less than that of \( b_4 \) whenever \( |\beta| < \beta_0 \).

**Table 3. Values of \( \beta_0 \)**

<table>
<thead>
<tr>
<th>( \frac{\lambda}{\beta^2} )</th>
<th>( r )</th>
<th>2</th>
<th>3</th>
<th>5</th>
<th>( \infty )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.707</td>
<td>.700</td>
<td>.683</td>
<td>.653</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.816</td>
<td>.790</td>
<td>.770</td>
<td>.738</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>.913</td>
<td>.865</td>
<td>.836</td>
<td>.800</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>.952</td>
<td>.896</td>
<td>.861</td>
<td>.820</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>.975</td>
<td>.914</td>
<td>.875</td>
<td>.831</td>
<td></td>
</tr>
<tr>
<td>( \infty )</td>
<td>1.000</td>
<td>.930</td>
<td>.890</td>
<td>.840</td>
<td></td>
</tr>
</tbody>
</table>

The comparison of \( b_2 \) with \( b_4 \) is simpler, for \( \text{Var}_A(b_2) \) always exceeds \( \text{Var}_A(b_4) \). This may be seen as follows:

\[
\text{Var}_A(b_2) - \text{Var}_A(b_4) = \frac{n-1}{r^2[\Sigma(X_i-\bar{X})^2]^2} \left[ \frac{\lambda + \frac{2\lambda^2}{\beta^2}}{r-1} \right] \sigma_e^4
\]

\[
- \frac{(\beta^2-\lambda)^2}{(\beta^2+\lambda)^2} \frac{\lambda}{\beta^2} - \frac{2\lambda^2}{(\beta^2+\lambda)^2} \left( \beta^2 + \frac{1}{\beta^2} \right) \frac{r}{r-1}
\]

Now \( \lambda - \frac{(\beta^2-\lambda)^2}{(\beta^2+\lambda)^2} \frac{\lambda}{\beta^2} - \frac{2\lambda^2}{(\beta^2+\lambda)^2} \frac{r}{r-1} \beta^2 \) is zero when \( r = 2 \) and
is positive for all other values of \( r \); \( \frac{2\lambda^2}{\beta^2} - \frac{2\lambda^2}{(\beta^2+\lambda)^2} \frac{1}{\beta^2} \) is always positive. Consequently \( \text{Var}_A(b_2) - \text{Var}_A(b_4) > 0 \) always.

From the diagram on page 44, one would not want to use \( b_3 \) unless he had reason to believe that \( \frac{\lambda}{\beta^2} \) was close to unity. Consequently we shall compare \( b_3 \) with \( b_4 \) only for \( \lambda = \beta^2 \).

Under these circumstances
\[
\text{Var}_A(b_3) - \text{Var}_A(b_4) = \frac{(n-1)(nr-1)}{nr^2[\Sigma(x_1-x)^2]^2} \left[ \beta^2 - \frac{1}{4}(\beta^2 + \frac{1}{\beta^2}) \right] \sigma^4_e.
\]

\[
\text{Var}_A(b_3) - \text{Var}_A(b_4) > 0 \quad \text{when} \quad |\beta| > .759.
\]

It is clear that one must have some information concerning the magnitude of \( \lambda \) before he can choose rationally among \( b_1 \), \( b_2 \), \( b_3 \), and \( b_4 \). This is not to say that the precise value of \( \lambda \) must be known; a rough approximation or perhaps an upper or lower bound might well be adequate. For example if one knew that \( \lambda \) was not greater than ten while \( |\beta| \) was roughly twenty, he could say categorically that \( b_4 \) would provide the best estimate of \( \beta \) insofar as the asymptotic variance was concerned. The information concerning \( \lambda \) might perhaps be obtained from the experimenter's knowledge of his technique, or failing that, one could take \( \hat{\lambda} \) as an estimator of \( \lambda \).

Inasmuch as one almost never knows the precise value of \( \lambda \), I feel that the estimator \( b_3 \) can scarcely ever be justified.
After all if one does not know $\lambda$, how can he say that $\frac{\lambda}{\beta^2}$ is close to unity; on the other hand, if he does know $\lambda$, he should be using the maximum likelihood estimator. The situation is now quite clear and may be summarized briefly as follows:

1. If $\beta^2 > \lambda$, use $b_\lambda$.
2. If $\beta^2 < \lambda$, use $b_\lambda$ when $|\beta| > \beta_0$; use $b_1$ when $|\beta| < \beta_0$. (See Table 3.)

If one does not know whether $|\beta| < \beta_0$ or $|\beta| > \beta_0$ and cannot make a reasonable inference concerning this from the data, then he is simply in no position to choose intelligently between $b_1$ and $b_\lambda$.

In order to handle the situation when $g \neq 1$, ($g = \frac{s}{r}$; see section 2.1) we make the approximations

$$
\frac{nr-1}{n(r-1)} \approx \frac{r}{r-1} \\
\frac{ns-1}{n(s-1)} \approx \frac{s}{s-1}.
$$

Proceeding just as before we find that

$$
\text{Var}_A(b_1) - \text{Var}_A(b_2) > 0 \quad \text{when} \quad \beta^2 > \lambda \frac{\sqrt{r(r-1)}}{\sqrt{s(s-1)}}.
$$

$$
\text{Var}_A(b_1) - \text{Var}_A(b_2) < 0 \quad \text{when} \quad \beta^2 < \lambda \frac{\sqrt{r(r-1)}}{\sqrt{s(s-1)}}.
$$
\[ \text{Var}_A(b_1) - \text{Var}_A(b_3) < 0 \quad \text{when} \quad \left( \frac{\lambda}{g \beta^2} \right)^2 < \frac{s-1}{s} + \frac{\sqrt{(s-1)^2 + 3 \frac{r(s-1)}{s(r-1)}}}{\frac{r-1}{r}}. \]

\[ \text{Var}_A(b_2) - \text{Var}_A(b_3) < 0 \quad \text{when} \quad \left( \frac{\lambda}{g \beta^2} \right)^2 < \frac{r-1}{r} + \frac{\sqrt{(r-1)^2 + 3s(r-1)}}{3r(s-1)}. \]

In order to compare \( b_4 \) with \( b_1 \) it is necessary to obtain values of \( \beta_0 \) corresponding to specified values of \( \frac{\lambda}{g \beta^2} \), \( r \), and \( s \); these can be obtained by solving the equation

\[ \frac{\lambda}{g} + 2\beta^2 \frac{r}{r-1} - \frac{(\beta^2 - \frac{\lambda}{g})^2}{(\beta^2 + \frac{\lambda}{g})^2} - \frac{(\lambda)^2}{(\beta^2 + \frac{\lambda}{g})^2} \left( \frac{r}{r-1} \beta^2 + \frac{s}{s-1} \beta^{-2} \right) = 0 \]

for \( \beta \). Results have not been tabulated here as they would require a triple-entry table.

It may be shown in exactly the same way as for the case \( g = 1 \) that \( \text{Var}_A(b_2) - \text{Var}_A(b_4) > 0 \) always.

Finally, if \( \lambda = \beta^2 \), \( \text{Var}_A(b_2) - \text{Var}_A(b_4) > 0 \) whenever

\[ \left( \frac{r}{r-1} + \frac{s}{s-1} \right) \beta^2 > \frac{1}{2} \left( \frac{r}{r-1} \beta^2 + \frac{s}{s-1} \frac{1}{\beta^2} \right). \]

It should be emphasized that the recommendations made in this section are based upon consideration of the asymptotic variance only. It is possible, of course, that different conclusions
might result from consideration of the bias, say, or the mean square error. We shall return to this point in section 4 of the dissertation.
3. LARGE-SAMPLE RESULTS WHEN ERRORS ARE CORRELATED

3.1. Notation, Assumptions, and Preliminary Remarks

We now modify the assumptions of section 2.1 by allowing the errors to be correlated. To be specific we shall consider the following model:

\[ x_{it} = X_i + e_{it} \]
\[ y_{it} = Y_i + f_{it} \]

where \( e_{it} \) and \( f_{it} \) are correlated. It will be observed that we consider only the situation wherein the number of replicates is a constant, \( r \), for each \( X_i \) and \( Y_i \). One might of course envisage situations wherein the number of replicates varies, just as in section 2.1, but in the case of correlated errors the algebra is very heavy indeed. Furthermore, it seems likely that the expressions would be so complicated as to defy interpretation, thus rendering the algebra empty.

With respect to the model above we introduce the following notation and assumptions:

\[ e_i^* = \frac{1}{r} \sum_{t=1}^{r} e_{it} \]
\[ f_i^* = \frac{1}{r} \sum_{t=1}^{r} f_{it} \]
\[ e_{..} = \frac{1}{n} \sum e_i^* \]
\[ f_{..} = \frac{1}{n} \sum f_i^* \]

and similarly for \( x_{i..} \), \( x_{..} \), \( y_{i..} \), and \( y_{..} \).
\[ \bar{X} = \frac{1}{n} \sum X_i \quad \bar{Y} = \frac{1}{n} \sum Y_i. \]

\[
E(e_{it}) = 0 \quad E(e_{it}^2) = \mu_{20} \quad E(e_{it}^4) = \mu_{40};
\]

\[ e_{it} \text{ and } e_{i't'} \text{ are independent unless } i = i' \text{ and } t = t'. \]

\[
E(f_{it}) = 0 \quad E(f_{it}^2) = \mu_{02} \quad E(f_{it}^4) = \mu_{04};
\]

\[ f_{it} \text{ and } f_{i't'} \text{ are independent unless } i = i' \text{ and } t = t'. \]

\[ E(e_{it} f_{it}) = \mu_{11} \]

\[ E(e_{it}^3) = E(e_{it}^2 f_{it}) = E(e_{it} f_{it}^2) = E(f_{it}^3) = 0 \]

\[
E(e_{it}^3 f_{it}) = \mu_{31} \quad E(e_{it}^2 f_{it}^2) = \mu_{22} \quad E(e_{it} f_{it}^3) = \mu_{13};
\]

\[ e_{it} \text{ and } f_{i't'} \text{ are independent unless } i = i' \text{ and } t = t'. \]

Denote \( E(e_{i}^3 f_{i}.) \) by \( \bar{\mu}_{31} \) and similar expectations correspondingly. Denote the cumulant corresponding to \( \bar{\mu}_{31} \) by \( \bar{K}_{31} \) and similarly for the other cumulants. The assumption that all third moments, pure and mixed alike, are zero is rather innocuous; such is the case for any bivariate distribution the contours of which are symmetric with respect to the origin.

The algebra involved in obtaining the expectations of the various statistics used in section 2.5 is greatly facilitated by the use of bivariate k-statistics, which are explained in detail by Kendall (32), with further elaboration by Cook (12). I have adhered scrupulously to the notation of these two writers and have also made specific references to
formulas found in their work, so that anyone desiring to verify my results can retrace the path I have followed.

Two independent checks have been continually applied to the results in this section. The first involves setting $\mu_{11} = \mu_{31} = \mu_{13} = 0$ and $\mu_{22} = \mu_{20}\mu_{02}$; under these conditions one should recover the results of sections 2.2 and 2.3 wherein it is assumed that the errors are independent. The second check involves putting $f_{it} = e_{it}$ for all $i$ and $t$; one should then recover familiar univariate results. For example, if $f_{it} = e_{it}$, then $s_{XYW}$ must reduce to $s_{XXW}$.

3.2. Heuristic Considerations Suggesting Estimators of the Slope

Paralleling the development in section 2.1 we consider the mean squares defined there as well as

$$s_{XYW} = \frac{1}{n(r-1)} \sum (x_{it}-x_{i.})(y_{it}-y_{i.}).$$

The expectations of these mean squares are exhibited in Table 4.

The following estimators of $\beta$ suggest themselves:

$$b_1 = \frac{s_{XYB} - s_{XYW}}{s_{XXB} - s_{XXW}}$$

$$b_2 = \frac{s_{YYB} - s_{YYW}}{s_{XXB} - s_{XXW}}$$

$$b_3 = \left[ \frac{s_{YYB} - s_{YYW}}{s_{XXB} - s_{XXW}} \right]^{-1/2} \text{sgn } \beta.$$
Table 4. Expected mean squares of components

<table>
<thead>
<tr>
<th>Mean square</th>
<th>Expected mean square</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_{XXB}$</td>
<td>$\frac{r}{n-1} \Sigma (x_i - \bar{x})^2 + \mu_{20}$</td>
</tr>
<tr>
<td>$s_{YYB}$</td>
<td>$\frac{r}{n-1} \Sigma (y_i - \bar{y})^2 + \mu_{02}$</td>
</tr>
<tr>
<td>$s_{XYB}$</td>
<td>$\frac{r}{n-1} \Sigma (x_i - \bar{x})(y_i - \bar{y}) + \mu_{11}$</td>
</tr>
<tr>
<td>$s_{XXW}$</td>
<td>$\mu_{20}$</td>
</tr>
<tr>
<td>$s_{YYW}$</td>
<td>$\mu_{02}$</td>
</tr>
<tr>
<td>$s_{XYW}$</td>
<td>$\mu_{11}$</td>
</tr>
</tbody>
</table>

Each of these is consistent provided that the expectation of its denominator is not zero (see section 2.1). Still a fourth consistent estimator of $\beta$ is suggested by maximum-likelihood when the errors are assumed to follow a bivariate normal distribution with known $R = \frac{\mu_{11}}{\mu_{20}}$ and known $\lambda = \frac{\mu_{02}}{\mu_{20}}$. The solution of the likelihood equations is given by Acton (1, page 135). The estimator of $\beta$ is the appropriate root of the equation

$$\hat{\beta}^2(s_{XYB} - Rs_{XXB}) - \hat{\beta}(s_{YYB} - \lambda s_{XXB}) - (\lambda s_{XYB} - Rs_{YYB}) = 0 . \quad [1]$$

This equation corresponds to equation [10], page 135, in Acton. Acton's notation differs slightly from that employed here; he denotes $s_{XYB}$, $s_{XXB}$, $s_{YYB}$ by $S_{x_1,y_1}$, $S_{x_1}$, and $S_{y_1}$, respectively, $\lambda$ by $\frac{R^2}{\rho^2}$ and omits the caret over $\beta$. 
Actually, Acton's equation [10] contains an error in the coefficient of \( \beta \); he gives it as \( \frac{S_{y_1,y_1}}{\rho^2} - \frac{R}{\rho^2} S_{x_1,x_1} \) whereas it should read \( \frac{S_{y_1,y_1}}{\rho^2} - \frac{R^2}{\rho^2} S_{x_1,x_1} \). Acton's expression for \( \beta R \), just preceding his equation [10] is also in error; it should read

\[
\beta R = \frac{\beta - \frac{R}{\rho^2}}{\beta - R} R \theta.
\]

Equation [3.2 - 1] has the solution

\[
\hat{\beta} = \varphi' + \left[ (\varphi')^2 + L' \right]^{1/2}
\]

where

\[
\varphi' = \frac{s_{YYB} - \lambda s_{XXB}}{2(s_{XYB} - R s_{XXB})},
\]

\[
L' = \frac{\lambda s_{XYB} - R s_{YYB}}{s_{XYB} - R s_{XXB}}.
\]

The positive sign is to be chosen when \( s_{XYB} - R s_{XXB} > 0 \); the negative sign is to be chosen when \( s_{XYB} - R s_{XXB} < 0 \). If \( s_{XYB} - R s_{XXB} = 0 \), one obtains from equation [3.2 - 1] above

\[
\frac{\lambda s_{XYB} - R s_{YYB}}{s_{YYB} - \lambda s_{XXB}}
\]

as an estimator of \( \beta \) provided \( s_{YYB} - \lambda s_{XXB} \neq 0 \). If \( s_{YYB} - \lambda s_{XXB} \) is also zero, no finite value \( \hat{\beta} \) can satisfy [3.2 - 1] except when \( \lambda s_{XYB} - R s_{YYB} \) is zero, in which case the solution of the likelihood equation is indeterminate.

If \( R \) and \( \lambda \) are unknown, they can be estimated with the
aid of \( s_{XYW}, s_{XXW}, \) and \( s_{YYW} \). Thus,
\[
\hat{R} = \frac{s_{XYW}}{s_{XXW}}
\]
\[
\hat{\lambda} = \frac{s_{YYW}}{s_{XXW}}
\]
This suggests a sort of pseudo maximum-likelihood estimator of \( \beta \), analogous to that of section 2.1:
\[
b_{4} = \phi + (\phi^{2} + L)^{1/2}
\]
where \( \phi = \frac{s_{YYB} - \hat{\lambda}s_{XXB}}{2(s_{XYB} - \hat{R}s_{XXB})} \)
\[
L = \frac{\hat{\lambda}s_{XYB} - \hat{R}s_{YYB}}{s_{XYB} - \hat{R}s_{XXB}}
\]
provided \( s_{XYB} - \hat{R}s_{XXB} \neq 0 \);
\[
b_{4} = -\frac{\hat{\lambda}s_{XYB} - \hat{R}s_{YYB}}{s_{YYB} - \hat{\lambda}s_{XXB}}
\]
when \( s_{XYB} - \hat{R}s_{XXB} = 0 \), provided that \( s_{YYB} - \hat{\lambda}s_{XXB} \neq 0 \). If
\( s_{XYB} - \hat{R}s_{XXB} = 0 \) and \( s_{YYB} - \hat{\lambda}s_{XXB} = 0 \) (which could happen only
with probability zero, of course), the method described gives
no determinate estimator of \( \beta \).

It is not suggested that \( b_{4} \) is actually the maximum-
likelihood estimator of \( \beta \), but it is a consistent estimator
of \( \beta \) if the assumptions of 2.5.1 are satisfied. This is shown
in section 3.4.4.
3.3 Expectations of Various Statistics Used in Section 3

3.3.1. \( \text{Var} s_{\text{XYB}} \)

Let

\[
T = \Sigma(x_i - \overline{x})(y_i - \overline{y})
\]

\[
T = \Sigma(x_i - \overline{x})(y_i - \overline{y}) + \Sigma(x_i - \overline{x})f_i. + \Sigma(y_i - \overline{y})e_i.
\]

\[
+ \Sigma(e_i - e..)(f_i - f..).
\]

Put \( m_{11} = \frac{1}{n} \Sigma(e_i - e..)(f_i - f..) \); then \( k_{11} = \frac{n}{n-1} m_{11} \).

\[
\text{Var } k_{11} = \kappa\left(\begin{array}{cc} 1 & 1 \\ 1 & 1 \end{array}\right) = \frac{1}{n} \kappa_{22} + \frac{1}{n-1} \kappa_{00} \kappa_{22} + \frac{1}{n-1} \kappa_{11}^2.
\]

(12, p. 187).

Now \( \kappa_{20} = \frac{\mu_{20}}{r} \); \( \kappa_{02} = \frac{\mu_{02}}{r} \); \( \kappa_{11} = \frac{\mu_{11}}{r} \); and

\[
\kappa_{22} = \bar{\mu}_{22} - \bar{\mu}_{20}\bar{\mu}_{02} - 2\mu_{11}^2 \quad (12, \text{p. } 183).
\]

\[
\bar{\mu}_{22} = E\left(\begin{array}{c} e_i^2 \\ f_i^2 \end{array}\right) = \frac{1}{r^4} E(\Sigma e_{it}^2 + \Sigma f_{it}^2)(\Sigma f_{it}^2) + \Sigma f_{it}^2 f_{iu}^2).
\]

\[
= \frac{1}{r^3} [\mu_{22} + (r-1)\mu_{20}\mu_{02} + 2(r-1)\mu_{11}^2].
\]

Consequently

\[
\kappa_{22} = \frac{1}{r^3} [\mu_{22} + (r-1)\mu_{20}\mu_{02} + 2(r-1)\mu_{11}^2] - \frac{\mu_{20}\mu_{02}}{r^2} - \frac{2\mu_{11}^2}{r^2}
\]

\[
= \frac{1}{r^3} [\mu_{22} - \mu_{20}\mu_{02} - 2\mu_{11}^2]. \quad [1]
\]
\[ \text{Var } k_{11} = \frac{1}{nr^3} \left[ \mu_{22} - \mu_{20} \mu_{02} - 2\mu_{11}^2 \right] + \frac{\mu_{20} \mu_{02} + \mu_{11}^2}{(n-1)r^2}. \]

\[ \text{Var } \Sigma(e_i - e..)(f_i - f..) = n^2 \text{Var } m_{11} = (n-1)^2 \text{ Var } k_{11} \]

\[ = \frac{(n-1)^2}{nr^3} (\mu_{22} - \mu_{20} \mu_{02} - 2\mu_{11}^2) + \frac{n-1}{r^2} (\mu_{20} \mu_{02} + \mu_{11}^2). \]

\[ \text{Var } \Sigma(X_i - \overline{X})f_i = \frac{\mu_{02}}{r} \Sigma(X_i - \overline{X})^2. \]

\[ \text{Var } \Sigma(Y_i - \overline{Y})e_i = \frac{\mu_{20}}{r} \Sigma(Y_i - \overline{Y})^2 = \frac{\beta^2 \mu_{20}}{r} \Sigma(X_i - \overline{X})^2. \]

\[ \text{Cov}[\Sigma(X_i - \overline{X})f_i, \Sigma(Y_i - \overline{Y})e_i] = \frac{\mu_{11}}{r} \Sigma(X_i - \overline{X})(Y_i - \overline{Y}) = \frac{\beta \mu_{11}}{r} \Sigma(X_i - \overline{X})^2. \]

\[ \text{Var } T = \frac{\beta^2 \mu_{20} + 2\beta \mu_{11} + \mu_{02}}{r} \Sigma(X_i - \overline{X})^2 \]

\[ + \frac{(n-1)^2}{nr^3} (\mu_{22} - \mu_{20} \mu_{02} - 2\mu_{11}^2) + \frac{n-1}{r^2} (\mu_{20} \mu_{02} + \mu_{11}^2). \]

Since \( \text{Var } s_\text{XYB} = \frac{r^2}{(n-1)^2} \text{ Var } T, \)

\[ \text{Var } s_\text{XYB} = \frac{r^2}{(n-1)^2} \frac{\beta^2 \mu_{20} + 2\beta \mu_{11} + \mu_{02}}{r} \Sigma(X_i - \overline{X})^2 \]

\[ + \frac{(n-1)^2}{nr^3} (\mu_{22} - \mu_{20} \mu_{02} - 2\mu_{11}^2) + \frac{n-1}{r^2} (\mu_{20} \mu_{02} + \mu_{11}^2). \]
3.3.2. Var $s_{XYW}$

Let $T = \sum (e_{it} - e_{i.})(f_{it} - f_{i.})$.

Put $m_{11} = \frac{1}{r} \sum (e_{it} - e_{i.})(f_{it} - f_{i.})$.

Then $\text{Var} T = r^2 n \text{Var} m_{11}$.

Now $k_{11} = \frac{r}{r-1} m_{11}$.

As in 3.3.1, we have

$$\text{Var} k_{11} = \frac{1}{r} \kappa_{22} + \frac{1}{r-1} \kappa_{20} \kappa_{02} + \frac{1}{r-1} \kappa_{11}^2$$

$$= \frac{1}{r} (\mu_{22} - \mu_{20}\mu_{02} - 2\mu_{11}^2) + \frac{\mu_{20}\mu_{02} + \mu_{11}^2}{r-1}.$$

$$\text{Var} T = (r-1)^2 n \text{Var} k_{11}$$

$$= (r-1)^2 n \frac{1}{r}(\mu_{22} - \mu_{20}\mu_{02} - 2\mu_{11}^2) + \frac{\mu_{20}\mu_{02} + \mu_{11}^2}{r-1}.$$ 

Hence,

$$\text{Var} s_{XYW} = \frac{1}{nr} (\mu_{22} - \mu_{20}\mu_{02} - 2\mu_{11}^2) + \frac{\mu_{20}\mu_{02} + \mu_{11}^2}{n(r-1)}.$$

3.3.3. Cov($s_{XYB}$, $s_{XYW}$)

Let $T = [\Sigma (X_{i1} - \overline{X})(Y_{i1} - \overline{Y}) + \Sigma (X_{i1} - \overline{X})f_{i1.} + \Sigma (Y_{i1} - \overline{Y})e_{i.}$

$$+ \Sigma (e_{i1.} - e_{...})f_{i1.}] \Sigma (e_{i1't} - e_{...})f_{i1't}.$$

$$= \sum (e_{i1.} - e_{...})f_{i1.} \sum (e_{i1't} - e_{...})f_{i1't} = \sum \sum (e_{i1.}e_{i1't}f_{i1.}f_{i1't}$

$$- e_{i1.}f_{i1.}f_{i1't} - e_{...}e_{i1't}f_{i1.}f_{i1't} + e_{...}e_{i1'.}f_{i1.}f_{i1't}).$$
\[
\begin{align*}
E\Sigma\Sigma\Sigma_{i. i'. t} f_{i. f_{i'. t}} &= \frac{n}{r} [\mu_{22} + (r-1)\mu_{11}^2] + n(n-1)\mu_{11}^2 \\
&= \frac{n}{r} [\mu_{22} + (nr-1)\mu_{11}^2]. \\
E\Sigma\Sigma\Sigma_{i. i'. f_{i. f_{i'. t}}} &= rE\Sigma\Sigma\Sigma_{i. e_{i'. f_{i. f_{i'. t}}}} = rE\Sigma\Sigma\Sigma e_{i. f_{i.}} \\
&\quad + rE\Sigma\Sigma\Sigma_{i. e_{i'. f_{j.}}} \\
&= nr\mu_{22} + n(n-1)r \frac{\mu_{11}^2}{r^2} \\
&= \frac{nr}{r^3} [\mu_{22} + (r-1)(\mu_{20}\mu_{02} + 2\mu_{11}^2)] \\
&\quad + n(n-1)r \frac{\mu_{11}^2}{r^2} \\
&= \frac{n}{r^2} [\mu_{22} + (r-1)(\mu_{20}\mu_{02} + 2\mu_{11}^2) \\
&\quad + \frac{n(n-1)\mu_{11}^2}{r}] \\
&= \frac{n}{r^2} [\mu_{22} + (r-1)\mu_{20}\mu_{02} + (rn+r-2)\mu_{11}^2].
\end{align*}
\]

\[
\begin{align*}
E\Sigma\Sigma\Sigma_{i. i'. t} e_{i. f_{i. f_{i'. t}}} &= \frac{1}{n} E\Sigma\Sigma\Sigma_{i. e_{i'. f_{i. f_{i'. t}}}}. \\
E\Sigma\Sigma\Sigma e_{i. i'. f_{i. f_{i'. t}}} &= \frac{1}{n} E\Sigma\Sigma\Sigma_{i. e_{i'. f_{i. f_{i'. t}}}}.
\end{align*}
\]

Consequently,

\[
E\Sigma(e_{i. e_{i'. t}}) f_{i. e_{i. t} - e_{i'. t}} f_{i. f_{i'. t}} = \frac{(n-1)(r-1)}{r^2} [\mu_{22} - \mu_{20}\mu_{02} \\
&\quad + (nr-2)\mu_{11}^2].
\]
\[ E(X_i - \bar{X})(Y_i - \bar{Y}) \sum (e_i - e)v \cdot f_{i'} t = \beta \mu_{11} n (r-1) E(X_i - \bar{X})^2. \]

Therefore,
\[ E(T) = \beta \mu_{11} n (r-1) E(X_i - \bar{X})^2 + \frac{(n-1)(r-1)}{r^2} [\mu_{22} - \mu_{20} \mu_{02} + (nr-2) \mu_{11}^2]. \]

Since \( \text{Cov}(s_{XYB}, s_{XYW}) = \frac{r}{n(n-1)(r-1)} E(T) - E(s_{XYB})E(s_{XYW}), \)
\[ \text{Cov}(s_{XYB}, s_{XYW}) = \frac{\mu_{22} - \mu_{20} \mu_{02} - 2\mu_{11}^2}{nr}. \]

3.3.4. \( \text{Cov}(s_{XYB}, s_{XXB}) \)
\[ \text{Cov}(s_{XYB}, s_{XXB}) = \frac{r^2}{(n-1)^2} E[\sum (X_i - \bar{X})^2 + 2\sum (X_i - \bar{X}) e_i \cdot \]
\[ + \sum (e_i - e \cdot \cdot \cdot)^2] \times [\sum (X_i - \bar{X})(Y_i - \bar{Y}) + \sum (X_i - \bar{X}) f_{i'} \cdot \]
\[ + \sum (Y_i - \bar{Y}) e_i \cdot \cdot \cdot + \sum (e_i - e \cdot \cdot \cdot) f_{i'} \cdot] - E(s_{XYB})E(s_{XXB}). \]

Put \( m_{20} = \frac{\sum (e_i - e \cdot \cdot \cdot)^2}{n} \quad m_{11} = \frac{\sum (e_i - e \cdot \cdot \cdot) f_{i'}}{n} \)
\[ k_{20} = \frac{n}{n-1} m_{20} \quad k_{11} = \frac{n}{n-1} m_{11}. \]
\[ \text{Cov}(k_{20}, k_{11}) = \kappa(2 \ 1) = \frac{1}{n} \mu_{31} + \frac{2}{n-1} \mu_{20} \mu_{11}. \] (12, p. 187)

Now \( \mu_{31} = \mu_{31} - 3\mu_{20} \mu_{11}, \) (11, p. 183).
\[ \bar{\mu}_{31} = \mathbb{E}(e_{i_1}^3 f_{i_1}) = \frac{1}{r^4} \mathbb{E}(\sum_{t \in I} e_{it}^3 + 3 \sum_{t \in I} e_{it} e_{iu}) \]

\[ + \sum_{t \in I} e_{it} e_{iu} e_{iv} (\sum_{t \in I} f_{it}) = \frac{\mu_{31} + 3(r-1)\mu_{20}\mu_{11}}{r^3} . \]

Therefore \[ \bar{\kappa}_{31} = \frac{\mu_{31} - 3\mu_{20}\mu_{11}}{r^3} . \]

\[ \text{Cov}(k_{20}, k_{11}) = \frac{\mu_{31} - 3\mu_{20}\mu_{11}}{nr^3} + \frac{2\mu_{20}\mu_{11}}{(n-1)r^2} . \]

\[ \text{Cov}[\Sigma(e_{i_1}, e_{i_2})^2, \Sigma(e_{i_1}, e_{i_2})f_{i_1,f_{i_2}}] = n^2 \text{Cov}(m_{20}, m_{11}) \]

\[ = (n-1)^2 \text{Cov}(k_{20}, k_{11}) = \frac{(n-1)^2}{nr^3} (\mu_{31} - 3\mu_{20}\mu_{11}) \]

\[ + \frac{n-1}{r^2} (2\mu_{20}\mu_{11}) . \]

\[ \mathbb{E}[\Sigma(e_{i_1}, e_{i_2})^2 \Sigma(e_{i_1}, e_{i_2})f_{i_1,f_{i_2}}] = \frac{(n-1)^2}{nr^3} (\mu_{31} - 3\mu_{20}\mu_{11}) \]

\[ + \frac{n-1}{r^2} (2\mu_{20}\mu_{11}) + \frac{n-1}{r^2} \mu_{20}\mu_{11} . \]

\[ \mathbb{E}[\Sigma(X_{i_1} - \bar{X})^2 \Sigma(X_{i_1} - \bar{X})(Y_{i_1} - \bar{Y})] = \Sigma(X_{i_1} - \bar{X})^2 \Sigma(X_{i_1} - \bar{X})(Y_{i_1} - \bar{Y}) . \]

\[ \mathbb{E}[\Sigma(X_{i_1} - \bar{X})^2 \Sigma(e_{i_1}, e_{i_2})f_{i_1,f_{i_2}}] = \frac{n-1}{r} \mu_{11} \Sigma(X_{i_1} - \bar{X})^2 . \]

\[ \mathbb{E}[\Sigma(X_{i_1} - \bar{X}) e_{i_1} \Sigma(X_{i_1} - \bar{X})f_{i_1,f_{i_2}}] = \frac{\mu_{11}}{r} \Sigma(X_{i_1} - \bar{X})^2 . \]

\[ \mathbb{E}[\Sigma(X_{i_1} - \bar{X}) e_{i_1} \Sigma(Y_{i_1} - \bar{Y})e_{i_1,f_{i_2}}] = \frac{\mu_{20}}{r} \Sigma(X_{i_1} - \bar{X})^2 . \]

\[ \mathbb{E}[\Sigma(e_{i_1}, e_{i_2})^2 \Sigma(X_{i_1} - \bar{X})(Y_{i_1} - \bar{Y})] = \frac{n-1}{r} \mu_{20} \Sigma(X_{i_1} - \bar{X})(Y_{i_1} - \bar{Y}) . \]

Therefore,
\[
\text{Cov}(s_{XYB}, s_{XXB}) = \frac{2r(\beta \mu_{20} + \mu_{11})}{(n-1)^2} \Sigma(X_i - \bar{X})^2 + \frac{1}{nr} \left( \mu_{31} - 3 \mu_{20} \mu_{11} \right) \\
+ \frac{2 \mu_{20} \mu_{11}}{n-1}.
\]

3.3.5. \text{Cov}(s_{XYW}, s_{XXB})

\[
\text{Cov}(s_{XYW}, s_{XXB}) = \frac{r}{n(n-1)(r-1)} E[\Sigma(X_i - \bar{X})^2 + 2 \Sigma(X_i - \bar{X})e_i] \\
+ \Sigma(e_i - e.)^2 \Sigma(e_i't - e_i')f_i't \\
- E(s_{XYW})E(s_{XXB}).
\]

\[
E\Sigma(e_i - e.)^2 \Sigma(e_i't - e_i')f_i't = E(\Sigma e_i'^2 \cdot e_i't\cdot f_i't - ne^2 \cdot \Sigma e_i'^2 \cdot f_i't \\
- r \Sigma e_i'^2 \cdot e_i', f_i' + nre^2 \cdot e_i', f_i').
\]

\[
E\Sigma e_i'^2 \cdot e_i't\cdot f_i't = E\Sigma e_i'^2 \cdot e_i't\cdot f_i't + E\Sigma e_i'^2 \cdot e_j't\cdot f_j't \\
= \frac{n}{r} [\mu_{31}(r-1)\mu_{20}\mu_{11}] + n(n-1)\mu_{20}\mu_{11} \\
= \frac{n}{r} [\mu_{31}(rn-1)\mu_{20}\mu_{11}].
\]

\[
Ee^2 \Sigma e_i't\cdot f_i't = \frac{1}{n} E\Sigma e_i'^2 \cdot e_i't\cdot f_i't = \frac{1}{n} E\Sigma e_i'^2 \cdot e_i't\cdot f_i't.
\]
\[
\begin{align*}
\mathbb{E}r\sum_{i} e_{i}^{2} f_{i} & = r\mathbb{E}e_{i}^{3} f_{i} + r\mathbb{E}e_{i}^{2} f_{j} \\
& = nr\mu_{31} + nr(n-1) \frac{\mu_{20}\mu_{11}}{r^2} \\
& = nr[\mu_{31} + 3(n-1)\mu_{20}\mu_{11}] + nr(n-1)\mu_{20}\mu_{11} \\
& = \frac{n}{r^2} [\mu_{31} + (nr+2r-3)\mu_{20}\mu_{11}] .
\end{align*}
\]

\[
\begin{align*}
\mathbb{E}r\sum_{i} e_{i} f_{i} & = \frac{1}{n} r\mathbb{E}e_{i}^{2} f_{i} = \frac{1}{r^2} [\mu_{31} + (nr+2r-3)\mu_{20}\mu_{11}] .
\end{align*}
\]

Therefore,
\[
\begin{align*}
\mathbb{E}X^{2}\sum_{i} (e_{i} - e_{\cdot})^2 \sum_{i} (e_{i} - e_{\cdot}) f_{i} & = \frac{(n-1)(r-1)}{r^2} \mu_{31} \\
& + \frac{n-1}{r^2} (r^2 n-3r-nr+3)\mu_{20}\mu_{11} .
\end{align*}
\]

\[
\begin{align*}
\mathbb{E}X(X_{i} - \bar{X})^{2}\sum_{i} (e_{i} - e_{\cdot}) f_{i} & = n(r-1)\mu_{11} X(X_{i} - \bar{X})^2 .
\end{align*}
\]

Consequently,
\[
\text{Cov}(s_{XYW}, s_{XXW}) = \frac{\mu_{31} - 3\mu_{20}\mu_{11}}{nr} .
\]

3.3.6. \text{Cov}(s_{XYB}, s_{XXW})
\[
\begin{align*}
\text{Cov}(s_{XYB}, s_{XXW}) & = \frac{r}{n(n-1)(r-1)} \mathbb{E} \left[ \sum_{i} (e_{i} - e_{\cdot}) \sum_{i} (e_{i} - e_{\cdot})^2 \right] \\
& = \frac{r}{n(n-1)(r-1)} \left[ \Sigma(X_{i} - \bar{X})(Y_{i} - \bar{Y}) + \Sigma(X_{i} - \bar{X})f_{i} + \Sigma(Y_{i} - \bar{Y})e_{i} \\
& + \Sigma(e_{i} - e_{\cdot})f_{i} \right] .
\end{align*}
\]
$$\sum(e_i' t - e_i^*')^2 \sum(e_i' - e_i^*') f_i = E[\sum e_i' t e_i' f_i - r \sum e_i' e_i' f_i]$$

$$- e_0 \sum e_i' t f_i + re_0 \sum e_i' f_i].$$

$$\sum e_i' t e_i' f_i = \sum e_i' t e_i' f_i + \sum e_i' t e_j' f_j.$$

$$= \frac{n}{r} [\mu_{31} + (r-1)\mu_{20}\mu_{11}] + n(n-1)\mu_{20}\mu_{11}$$

$$= \frac{n}{r} [\mu_{31} + (nr-1)\mu_{20}\mu_{11}] .$$

$$E \sum e_i' t f_i = \frac{n}{r^2} [\mu_{31} + (nr+2r-3)\mu_{20}\mu_{11}]$$

from [3.3.5 - 1].

$$E e_0 \sum e_i' f_i = \frac{1}{n} \sum e_i' t e_i' f_i = \frac{1}{r} [\mu_{31} + (nr-1)\mu_{20}\mu_{11}] .$$

$$E e_0 \sum e_i' f_i = \frac{r}{n} \sum e_i' t e_i' f_i .$$

$$= \frac{1}{r^2} [\mu_{31} + (nr+2r-3)\mu_{20}\mu_{11}] .$$

Therefore,

$$\sum(e_i' t - e_i^*')^2 \sum(e_i' - e_i^*') f_i = \frac{(n-1)(r-1)}{r^2} \mu_{31}$$

$$+ \frac{n-1}{r^2} (r^2 n^2 - r - nr-3) .$$

Now

$$\sum(X_i - \bar{X})(Y_i - \bar{Y}) \sum(e_i' t - e_i^*')^2 = n(r-1)\mu_{20} \sum(X_i - \bar{X})(Y_i - \bar{Y}) .$$

Consequently,

$$\text{Cov}(s_{XYB}, s_{XXW}) = \frac{\mu_{31} - 3\mu_{20}\mu_{11}}{nr} .$$
3.3.7. $\text{Cov}(s_{XYW}, s_{XXW})$

Following the notation of section 3.3.2 put

$$m_{20} = \frac{1}{r} \Sigma (e_{it} - \bar{e}_i)^2.$$ 

Then $k_{20} = \frac{r}{r-1} m_{20}$.

$$\text{Cov}(k_{20}, k_{11}) = \kappa(0 1) = \frac{1}{r} \kappa_{31} + \frac{2}{r-1} \kappa_{20} \kappa_{11}$$

$$= \frac{1}{r} (\mu_{31} - 3\mu_{20}\mu_{11}) + \frac{2}{r-1} \mu_{20}\mu_{11}.$$ 

$$\text{Cov}\left[\Sigma(e_{it} - \bar{e}_i)^2, \Sigma(e_{i't} - \bar{e}_{i'})f_{i't}\right] = nr^2 \text{Cov}(m_{20}, m_{11})$$
$$= n(r-1)^2 \text{Cov}(k_{20}, k_{11}).$$

$$\text{Cov}(s_{XYW}, s_{XXW}) = \frac{1}{n} \text{Cov}(k_{20}, k_{11})$$

$$= \frac{1}{nr} (\mu_{31} - 3\mu_{20}\mu_{11}) + \frac{2}{n(r-1)} \mu_{20}\mu_{11}.$$ 

3.3.8. $\text{Cov}(s_{XYB} - s_{XYW}, s_{XXB} - s_{XXW})$

$$\text{Cov}(s_{XYB} - s_{XYW}, s_{XXB} - s_{XXW}) = \text{Cov}(s_{XYB}, s_{XXB}) - \text{Cov}(s_{XYB}, s_{XXW})$$
$$- \text{Cov}(s_{XYW}, s_{XXB}) + \text{Cov}(s_{XYW}, s_{XXW}).$$

Combining results from sections 3.3.4, 3.3.5, 3.3.6 and 3.3.7 one obtains
\[
\text{Cov}(s_{XYB}, s_{XXB}) = \frac{r^2}{(n-1)^2} \left[ \Sigma(Y_1 - \bar{Y})^2 + 2\Sigma(Y_1 - \bar{Y})f_i \right] \\
+ \Sigma(f_i - f_{..})^2 \left[ \Sigma(X_1 - \bar{X})^2 + 2\Sigma(X_1 - \bar{X})e_i \right] \\
+ \Sigma(e_{i} - e_{..})^2 - \text{E}(s_{XXB})\text{E}(s_{YYB}) \cdot
\]

Put \( m_{20} = \frac{\Sigma(e_{i} - e_{..})^2}{n} \) \( m_{02} = \frac{\Sigma(f_i - f_{..})^2}{n} \)

\( k_{20} = \frac{n}{n-1} m_{20} \) \( k_{02} = \frac{n}{n-1} m_{02} \)

\[
\text{Cov}(k_{20}, k_{02}) = \mathcal{K}(2, 0) = \frac{1}{n} \mathcal{K}_{22} + \frac{2}{n-1} \mathcal{K}_{11}^2 .
\]

Making use of [2.5.2 - 1] we have

\[
\text{Cov}(k_{20}, k_{02}) = \frac{\mu_{22} - \mu_{20}\mu_{02} - 2\mu_{11}^2}{nr^3} + \frac{2\mu_{11}^2}{(n-1)r^2}.
\]

\[
\text{Cov}[\Sigma(e_{i} - e_{..})^2, \Sigma(f_{i} - f_{..})^2] = n^2 \text{Cov}(m_{20}, m_{02})
\]

\[
= (n-1)^2 \text{Cov}(k_{20}, k_{02})
\]

\[
= \frac{(n-1)^2}{nr^3} (\mu_{22} - \mu_{20}\mu_{02} - 2\mu_{11}^2) + \frac{n-1}{r^2} (2\mu_{11}^2) .
\]
\[ \mathbb{E}[(e_i - \mu)^2 \Sigma(f_i^1 - f) + \mu_{11} + \mu_{22} - 2\mu_{02} - 2\mu_{11} + \frac{2(n-1)^2 \mu_{11}^2}{n^2} + (\frac{n-1}{n})^2 \mu_{11}^2] \]

\[ \mathbb{E}[(X_i - \bar{X})^2 \Sigma(Y_i - \bar{Y})^2 = \mathbb{E}[(X_i - \bar{X})^2 \Sigma(Y_i - \bar{Y})^2] . \]

\[ \mathbb{E}[(Y_i - \bar{Y})^2 \Sigma(e_i - \mu)^2 = \frac{n-1}{n} \mu_{20}^2 \Sigma(Y_i - \bar{Y})^2 . \]

\[ \mathbb{E}[(Y_i - \bar{Y}) f_i \Sigma(X_i - \bar{X}) e_i = \frac{\mu_{11}}{n} \Sigma(X_i - \bar{X})(Y_i - \bar{Y}) . \]

\[ \mathbb{E}[(X_i - \bar{X})^2 \Sigma(f_i - f) + \mu_{11} + \mu_{22} - 2\mu_{02} - 2\mu_{11} + \frac{2(n-1)^2 \mu_{11}^2}{n^2} + (\frac{n-1}{n})^2 \mu_{11}^2] \]

Therefore,

\[ \text{Cov}(s_{YYB} \cdot s_{XXB}) = \frac{4r \mu_{11}}{(n-1)^2} \Sigma(X_i - \bar{X})^2 + \frac{1}{n^2} (\mu_{22}^2 - 2\mu_{02}^2 - 2\mu_{11}^2) + \frac{2\mu_{11}^2}{n-1} . \]

3.3.10. Cov(s_{YYB} \cdot s_{XXW})

\[ \text{Cov}(s_{YYB} \cdot s_{XXW}) = \frac{r}{n(n-1)(n-2)} \mathbb{E}[\Sigma(Y_i - \bar{Y})^2] + \Sigma(Y_i - \bar{Y}) f_i . \]

\[ + \Sigma(f_i - f) \mathbb{E}[\sum(e_i - \mu)^2 \Sigma(e_i - \mu) e_i] - \mathbb{E}(s_{YYB}) \mathbb{E}(s_{XXW}) . \]
\[
\begin{align*}
\Sigma(f_1 - f), f_1 \Sigma(e_{it} - e_{it}) e_{it} &= \Sigma(\Sigma \Sigma \Sigma \Sigma e_{it}^2 f_{i}^2 - nr \Sigma \Sigma \Sigma e_{it}^2) \\
- r \Sigma \Sigma \Sigma e_{it}^2 f_{i}^2 + nr \Sigma \Sigma \Sigma e_{it}^2).
\end{align*}
\]

\[
\begin{align*}
\Sigma \Sigma \Sigma e_{it}^2 f_{i}^2 &= e \Sigma \Sigma \Sigma e_{it}^2 f_{i}^2 + E \Sigma \Sigma \Sigma e_{it}^2 f_{i}^2 \\
&= \frac{n}{r}[\mu_{22} + (r-1)\mu_{11}] + n(n-1)\mu_{11}^2
\end{align*}
\]

\[
\begin{align*}
E \Sigma \Sigma \Sigma e_{it}^2 f_{i}^2 &= \frac{1}{n} E \Sigma \Sigma \Sigma e_{it}^2 f_{i}^2.
\end{align*}
\]

\[
\begin{align*}
E \Sigma \Sigma \Sigma e_{it}^2 f_{i}^2 &= r \Sigma \Sigma \Sigma e_{it}^2 f_{i}^2 + r \Sigma \Sigma \Sigma e_{it}^2 f_{i}^2 \\
&= nr \mu_{22} + n(n-1)r \frac{\mu_{11}^2}{r^2}
\end{align*}
\]

\[
\begin{align*}
&= \frac{nr[\mu_{22} + (r-1)\mu_{20}^2 + 2(r-1)\mu_{11}^2]}{r^3} + \frac{n(n-1)\mu_{11}^2}{r^2}.
\end{align*}
\]

\[
\begin{align*}
E \Sigma \Sigma \Sigma e_{it}^2 f_{i}^2 &= \frac{r}{n} E \Sigma \Sigma \Sigma e_{it}^2 f_{i}^2.
\end{align*}
\]

\[
\begin{align*}
E \Sigma(y_i - \bar{y})^2 \Sigma(e_{it} - e_{it})^2 &= n(r-1)\mu_{20}^2 \Sigma(y_i - \bar{y})^2.
\end{align*}
\]

Consequently,

\[
\text{Cov}(s_{YYB}, s_{XXW}) = \frac{\mu_{22} - \mu_{20}^2 - 2\mu_{11}^2}{nr}.
\]
3.3.11. \( \text{Cov}(s_{YYW}, s_{XXB}) \)

This expectation can be obtained from \( \text{Cov}(s_{YYB}, s_{XXW}) \) by simply permuting the symbols. The result is

\[
\text{Cov}(s_{YYW}, s_{XXB}) = \frac{\mu_{22} - \mu_{20}\mu_{02} - 2\mu_{11}}{nr}.
\]

3.3.12. \( \text{Cov}(s_{YYW}, s_{XXW}) \)

Put \( m_{20} = \frac{1}{r} \sum_t (e_{it} - e_i)^2 \)

\[
m_{02} = \frac{1}{r} \sum_t (f_{it} - f_i)^2.
\]

Then \( k_{20} = \frac{r}{r-1} m_{20} \) and \( k_{02} = \frac{r}{r-1} m_{02} \).

\[
\text{Cov}(k_{20}, k_{02}) = K(2, 0) = \frac{1}{r} K_{22} + \frac{2}{r-1} K_{11}
\]

\[
= \frac{1}{r} (\mu_{22} - \mu_{20}\mu_{02} - 2\mu_{11}) + \frac{2\mu_{11}^2}{r-1}.
\]

\[
\text{Cov}[(\Sigma e_{it} - e_i)^2, (\Sigma f_{it} - f_i)^2] = n(r-1)\text{Cov}(k_{20}, k_{02}).
\]

\[
\text{Cov}(s_{YYW}, s_{XXW}) = \frac{1}{n} \text{Cov}(k_{20}, k_{02})
\]

\[
= \frac{1}{nr} (\mu_{22} - \mu_{20}\mu_{02} - 2\mu_{11}) + \frac{2}{n(r-1)} \mu_{11}^2.
\]
3.3.13. \[ \text{Cov}(s_{YYB}, s_{YYW}, s_{XXB}, s_{XXW}) \]

\[
\text{Cov}(s_{YYB}, s_{YYW}, s_{XXB}, s_{XXW}) = \text{Cov}(s_{YYB}, s_{XXB}) - \text{Cov}(s_{YYB}, s_{XXW})
- \text{Cov}(s_{YYW}, s_{XXB}) + \text{Cov}(s_{YYW}, s_{XXW}).
\]

Combining results from sections 3.3.9, 3.3.10, 3.3.11 and 3.3.12, one obtains

\[
\text{Cov}(s_{YYB}, s_{YYW}, s_{XXB}, s_{XXW}) = \frac{4r\sigma^2 \mu_{11}}{(n-1)^2} \sum (x_i - \bar{x})^2 + \frac{2(nr-1)}{n(n-1)(r-1)} \mu_{11}^2.
\]

3.4. Asymptotic Variances and Covariances of Estimators of \( \alpha \) and \( \beta \)

This section parallels section 2.3 closely. In particular the discussion in section 2.3.1 concerning the order of magnitude of the terms in \( \text{Var}_A \) is equally applicable in this section.

It will be shown in this section that the leading term of \( \text{Var}_A \) is the same for \( b_1, b_2, b_3, \) and \( b_4 \), just as in the case of independent errors. For the reasons given in that case (see section 2.3.6) we have retained additional terms in the expressions for \( \text{Var}_A \).

3.4.1. \( \text{Var}_A(b_1) \)

The asymptotic variance of \( b_1 \) can be found exactly as in section 2.3.1. We now take
\[ u = s_{XYB} - s_{XYW} \]
\[ v = s_{XXB} - s_{XXW} \]
\[ U = \frac{r}{n-1} \Sigma (X_i - \overline{X})(Y_i - \overline{Y}) \]
\[ V = \frac{r}{n-1} \Sigma (X_i - \overline{X})^2. \]

\[ \text{Var } u = \text{Var } s_{XYB} + \text{Var } s_{XYW} - 2\text{Cov}(s_{XYB}, s_{XYW}) \]
\[ = \frac{r^2}{(n-1)^2} \beta^2 \mu_{20} + \frac{2\beta \mu_{11} + \mu_{02}}{r} \Sigma (X_i - \overline{X})^2 \]
\[ + \frac{n r - 1}{n(n-1)(r-1)} (\mu_{20}^2 + \mu_{11}^2). \]

From section 2.3.1, by putting \( r_H = r \) and \( R = r_n \), we obtain

\[ \text{Var } v = \frac{4 r}{(n-1)^2} \frac{\mu_{20}}{r} \Sigma (X_i - \overline{X})^2 + \frac{2(n r - 1) \mu_{20}^2}{n(n-1)(r-1)}. \]

\[ \text{Cov}(u, v) \text{ is given in section 2.5.9.} \]

Making use of [2.3 - 1] we obtain

\[ \text{Var}_A(b_1) = \frac{\beta^2 \mu_{20} - 2\beta \mu_{11} + \mu_{02}}{r} \frac{1}{\Sigma (X_i - \overline{X})^2} \]
\[ + \frac{(n-1)(nr-1)}{nr^2(r-1)} \frac{2\beta^2 \mu_{20}^2 - 4\beta \mu_{20} \mu_{11}}{[\Sigma (X_i - \overline{X})^2]^2} \]
\[ + \mu_{20}^2 \mu_{02} + \mu_{11}^2 \frac{1}{[\Sigma (X_i - \overline{X})^2]^2}. \]

When \( \mu_{11} = 0 \), this becomes
Var_A(b_1) = \frac{\beta^2 \mu_{20} + \mu_{02} + \frac{1}{r} \frac{(n-1)(nr-1)}{2 \Sigma(X_i - \bar{X})^2} (2 \beta^2 \mu_{20} + \mu_{20} \mu_{02})}{n \Sigma(\Sigma(X_i - \bar{X})^2)^2}.

This exceeds the asymptotic variance of the estimator

\frac{s_{XYB}}{s_{XXB} - s_{XXW}}

by \frac{n-1}{r^2} \frac{nr-1}{n(r-1)} \left[ \frac{1}{\Sigma(X_i - \bar{X})^2} \right]^2 \mu_{20} \mu_{02}. \quad \text{For large } n \text{ this excess is approximately}

\frac{n-1}{r^2} \frac{r}{r-1} \left[ \frac{1}{\Sigma(X_i - \bar{X})^2} \right]^2 \mu_{20} \mu_{02}.

One would expect the asymptotic variance of \frac{s_{XYB}}{s_{XXB} - s_{XXW}} to be greater than that of \frac{s_{XYB}}{s_{XXB} - s_{XXW}} because of the additional variation introduced by \( s_{XYW} \). In effect one pays a penalty for lack of knowledge concerning the independence of \( x_{it} \) and \( y_{it} \).

3.4.2. \text{Var}_A(b_2)

The asymptotic variance of \( b_2 \) can be obtained from the results of sections 3.3.1 - 3.3.9 by simply permuting the symbols. It is found that
When $\mu_{11} = 0$, this becomes

$$\text{Var}_A(b_2) = \frac{\beta^2\mu_{20} + \mu_{02}}{r} \frac{1}{\Sigma(X_i - \bar{X})^2}$$

$$+ \frac{(n-1)(nr-1)}{nr^2(r-1)} \left( \frac{2\mu_{02}}{\beta^2} - \frac{4\mu_{02}\mu_{11}}{\beta} + \mu_{20}\mu_{02} + \mu_{11}^2 \right) \frac{1}{[\Sigma(X_i - \bar{X})^2]^2}.$$ 

This exceeds the asymptotic variance of the estimator

$$\frac{S_{YYB} - S_{YYW}}{S_{XYB}}$$

by $$\frac{n-1}{r^2} \frac{nr-1}{n(r-1)} \left[ \frac{1}{\Sigma(X_i - \bar{X})^2} \right]^2 \mu_{20}\mu_{02}.$$ 

3.4.3. $\text{Var}_A(b_3)$

Take $u = S_{YYB} - S_{YYW}$

$$v = S_{XXB} - S_{XXW}$$
U = \frac{r}{n-1} \Sigma (Y_i - \bar{Y})^2

V = \frac{r}{n-1} \Sigma (X_i - \bar{X})^2.

Since b_3 = \left( \frac{u}{v} \right)^{1/2}, we have from section 2.3.3

\text{Var}_A(b_3) = \frac{\text{Var}_A(u)}{4uv} + \frac{U \text{Var}_A(v)}{4v^3} - \frac{\text{Cov}(u,v)}{2v^2}.

Making use of the results of the preceding sections, we obtain

\text{Var}_A(b_3) = \frac{\beta^2 \mu_{20} - 2\beta \mu_{11} + \mu_{02}}{r} \frac{1}{\Sigma (X_i - \bar{X})^2}

+ \frac{(n-1)(nr-1)}{2nr^2(r-1)[\Sigma (X_i - \bar{X})^2]^2} (\beta^2 \mu_{20}^2 - 2\mu_{11}^2 + \frac{\mu_{02}^2}{\beta^2}).

If we take \mu_{11} = 0, we get precisely the same expression as in the case of uncorrelated errors.

3.4.4. \text{Var}_A(b_4)

b_4 = \varphi \pm \left[ \beta^2 + L \right]^{1/2}

where \varphi = \frac{s_{XYB} - \hat{\lambda}s_{XXB}}{2(s_{XYB} - \hat{\lambda}s_{XXB})}

L = \frac{s_{XYB} - \hat{\lambda}s_{YYB}}{s_{XYB} - \hat{\lambda}s_{XXB}},

provided that s_{XYB} - \hat{\lambda}s_{XXB} \neq 0. (See section 3.1.)
Since \( \hat{s}_{XYB} - \hat{s}_{YYB} \) converges in probability to

\[
\frac{r}{n-1} (\beta \lambda - R \bar{\beta}^2) \Sigma (X_1 - \bar{X})^2, \quad s_{XYB} - \hat{s}_{XXB} \text{ to } \frac{r}{n-1} (\beta - R) \Sigma (X_1 - \bar{X})^2, \text{ and}
\]

\[
s_{YYB} - \hat{s}_{XXB} \text{ to } \frac{r}{n-1} (\beta^2 - \lambda) \Sigma (X_1 - \bar{X})^2, \quad \phi \text{ converges in probability to } \frac{\beta^2 - \lambda}{2(\beta - R)}, \text{ while } L \text{ converges to } \frac{\beta \lambda - R \beta^2}{\beta - R}. \quad \text{Consequently}
\]

\[
\phi^2 + L \text{ converges to } \frac{(\beta^2 + \lambda - 2R\beta)^2}{[2(\beta - R)]^2}. \quad \text{Therefore } b_4 \text{ converges to}
\]

\[
\frac{\beta^2 - \lambda}{2(\beta - R)} + \frac{\beta^2 + \lambda - 2R\beta}{2(\beta - R)} = \beta. \quad \text{Thus, } b_4 \text{ is a consistent estimator of } \beta.
\]

We can now expand \( b_4 \) in a Taylor series, just as we did in section 2.3.4, obtaining

\[
b_4 - \beta = \frac{\sum_{i=1}^{n-1}}{r(\beta^2 + \lambda - 2R\beta) \Sigma (X_1 - \bar{X})^2} [ (\beta - R)(s_{YYB} - S_{YYB}) - (\beta - R)(s_{YYW} - S_{YYW}) + \beta (R\beta - \lambda)(s_{XXB} - S_{XXB}) - \beta (R\beta - \lambda)(s_{XXW} - S_{XXW}) + (\lambda - \beta^2)(s_{XYB} - S_{XYB}) - (\lambda - \beta^2)(s_{XYW} - S_{XYW}) ]
\]

+ terms of higher order in \( s_{YYB}, s_{YYW}, s_{XXB}, s_{XXW} \)

and \( s_{XYB}, s_{XYW} \).

Therefore,
\[
\text{Var}_A(b \mid \lambda) = \frac{(n-1)^2}{r^2(\beta^2+\lambda-2\beta\beta)^2\left[\Sigma(X_1-\bar{X})^2\right]^2} \left[2p+2\beta(R-\lambda)\right] \text{Var}(s_{YYB}-s_{YYW})
\]

\[+ \beta^2(R^2-R^2\lambda)^2 \text{Var}(s_{XXB}-s_{XXW}) + (\lambda-2\beta)^2 \text{Var}(s_{XYB}-s_{XYW})
\]

\[- 2\beta(R-\lambda)(R^2-R^2\lambda) \text{Cov}[(s_{YYB}-s_{YYW}),(s_{XXB}-s_{XXW})]
\]

\[- 2\beta(R^2-R^2\lambda)(\lambda-2\beta) \text{Cov}[(s_{YYB}-s_{YYW}),(s_{XYB}-s_{XYW})]
\]

\[- 2\beta(R-\lambda)(\lambda-2\beta)^2 \text{Cov}[(s_{XXB}-s_{XXW}),(s_{XYB}-s_{XYW})].
\]

Thus,

\[\text{Var}_A(b \mid \lambda) = \frac{\beta^2\mu_{20} - 2\beta\mu_{11} + \mu_{02}}{r\Sigma(X_1-\bar{X})^2}
\]

\[+ \frac{(n-1)(n\lambda-1)}{nr^2(\beta^2+\lambda-2\beta\beta)^2\left[\Sigma(X_1-\bar{X})^2\right]^2} \left[2p+2\beta(R-\lambda)\right]^2 \mu_{02}^2
\]

\[+ 2(\beta^2-R^2\lambda)^2 \mu_{20}^2 + (\lambda-2\beta)^2(\mu_{20}\mu_{02}+\mu_{11})^2
\]

\[- 4\beta(R-\lambda)(\beta^2-R^2)\mu_{11}^2 - 4\beta(R-\lambda)(\lambda-2\beta)\mu_{02}\mu_{11}
\]

\[- 4\beta(\beta-\lambda)(\lambda-2\beta)\mu_{20}\mu_{11}].
\]

3.4.6. Asymptotic Variance of Intercept Estimators

For the reasons discussed in section 2.3.6 for the case of uncorrelated errors, the intercept \(a\) is estimated by

\[a = y - bx.
\]

when the errors are correlated. The asymptotic variance of \(a\) can be obtained for this situation just as it was obtained
in section 2.3.6. Thus
\[ a - a = (y - \bar{y}) - \beta(x - \bar{x}) - \bar{x}(b - \beta) \]
\[ + \text{terms of higher order in } x, y, \text{ and } b. \]
Therefore
\[
\text{Var}_A(a) = \frac{\mu_{02}}{rn} + \beta^2 \frac{\mu_{20}}{rn} + \bar{x}^2 \text{Var}_A(b) - \frac{2\beta\mu_{11}}{rn}
\]
\[ - 2\bar{x} \text{Cov}_A(b, y) + 2\beta \bar{x} \text{Cov}_A(b, x). \]
It can be shown, exactly as was done in section 2.3.6, that
\[ \text{Cov}_A(b, y) = \text{Cov}_A(b, x) = 0. \]
Therefore
\[
\text{Var}_A(a) = \frac{\beta^2\mu_{20} - 2\beta\mu_{11} + \mu_{02}}{rn} + \bar{x}^2 \text{Var}_A(b). \]

3.4.7. Asymptotic covariance of slope and intercept estimators
\[ (a - a)(b - \beta) = (y - \bar{y})(b - \beta) - \beta(x - \bar{x})(b - \beta) - \bar{X}(b - \beta)^2 \]
\[ + \text{terms of higher order in } x, y, \text{ and } b. \]
Therefore
\[ \text{Cov}_A(a, b) = - \bar{X} \text{Var}_A(b). \]

3.5. Discussion of Results

We should like now to compare the various estimators as was done in section 2.6. To begin with the simpler comparisons, we have
\[ \text{Var}_A(b_1) - \text{Var}_A(b_2) = \frac{(n-1)(nr-1)}{nr^2(r-1)\left[\Sigma(X_i - \bar{X})^2\right]^2} \left(2\beta^2 - 4\beta R \right. \\
\left. + \frac{4}{\beta} R\lambda - \frac{2}{\beta^2} \lambda^2 \right) \mu_{20}^2. \]

Since \( R^2 = \rho^2\lambda \), we can write
\[ 2\beta^2 - 4\beta R + \frac{4}{\beta} R\lambda - \frac{2}{\beta^2} \lambda^2 = 2\beta^2 - 4\beta \rho \sqrt{\lambda} + \frac{4}{\beta} \lambda \rho \sqrt{\lambda} - \frac{2}{\beta^2} \lambda^2. \]

Let \( h = \frac{\sqrt{\lambda}}{\beta} \). Then the right-hand side of \([3.5 - 1]\) becomes
\[ \frac{2}{h^2} - \frac{4\rho}{h} + 4\rho h - 2h^2, \]
which equals \( \frac{1}{h^2} (1-h^2)(h^2-2\rho h+1) \). Since \(-1 \leq \rho \leq 1\), the only real roots of the equation
\[ (1-h^2)(h^2-2\rho h+1) = 0 \]
are \( h = \pm 1 \); furthermore, \( h^2 - 2\rho h + 1 \geq 0 \) for all \( h \) and \( \rho \).
Clearly \((1-h^2)(h^2-2\rho h+1)\) exceeds zero when \( h^2 < 1 \), and is less than zero when \( h^2 > 1 \). We thus arrive at the conclusion
\[ \text{Var}_A(b_1) - \text{Var}_A(b_2) > 0 \text{ when } \beta^2 > \lambda \]
\[ \text{Var}_A(b_1) - \text{Var}_A(b_2) < 0 \text{ when } \beta^2 < \lambda, \]
which is exactly the same conclusion as the one arrived at in the case of independent errors!

I think that it is appropriate to mention at this point that Madansky (38) also has examined \( \text{Var}_A(b_1) - \text{Var}_A(b_2) \); he arrives at the erroneous conclusion that
\[ \text{Var}_A(b_1) - \text{Var}_A(b_2) > 0 \text{ when } \beta^2 > 1 \]
\[ \text{Var}_A(b_1) - \text{Var}_A(b_2) < 0 \text{ when } \beta^2 < 1. \]

This says in effect that one can choose rationally between \( b_1 \) and \( b_2 \) without knowing the relative variance of the errors in \( y \) and \( x \); this simply does not seem reasonable. In fact it seems very surprising to me that one can make such a choice without knowing the correlation coefficient! I think that the approximations made by Madansky are the source of the erroneous conclusions he has drawn.

Turning next to \( b_3 \) we find that
\[
\text{Var}_A(b_1) - \text{Var}_A(b_3) = \frac{(n-1)(nr-1)}{nr^2(r-1)[\Sigma(X_i - \bar{X})^2]^2} \left( \frac{3}{2} \beta^2 \mu_{20}^2 - 4\beta \mu_{20} \mu_{11} + \mu_{20}^2 - \frac{1}{2} \frac{\mu_{02}^2}{\beta^2} \right). 
\]

The factor containing \( \beta, \mu_{20}, \mu_{11} \), and \( \mu_{02} \) can be written

\[
\left[ \frac{3}{2} \beta^2 - 4\beta \rho + \lambda - \frac{1}{2} \frac{\lambda^2}{\beta^2} \right] \mu_{20}^2 = G, \text{ say.} \]

Putting \( h = \frac{\sqrt{\lambda}}{\beta} \) this becomes \( G = \frac{1}{2}(3-8 \rho h^2 + 2h^2 - h^4) \mu_{20}^2 \). The roots of the equation \( G = 0 \) for representative values of \( \rho \) are given in Table 5.

There is in each case one positive and one negative root; the remaining roots are complex. The import of this table is that
\[ \text{Var}_A(b_1) - \text{Var}_A(b_3) > 0 \text{ when } h_2 < \frac{\sqrt{\lambda}}{\beta} < h_1. \]
Table 5. Roots of equation $3 - 8 \rho h + 2h^2 - h^4 = 0$

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$h_1$</th>
<th>$h_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.42</td>
<td>-2.41</td>
</tr>
<tr>
<td>3/4</td>
<td>0.60</td>
<td>-2.28</td>
</tr>
<tr>
<td>1/2</td>
<td>1.00</td>
<td>-2.13</td>
</tr>
<tr>
<td>1/4</td>
<td>1.43</td>
<td>-1.95</td>
</tr>
<tr>
<td>0</td>
<td>1.73</td>
<td>-1.73</td>
</tr>
<tr>
<td>-1/4</td>
<td>1.95</td>
<td>-1.43</td>
</tr>
<tr>
<td>-1/2</td>
<td>2.13</td>
<td>-1.00</td>
</tr>
<tr>
<td>-3/4</td>
<td>2.28</td>
<td>-0.60</td>
</tr>
<tr>
<td>-1</td>
<td>2.41</td>
<td>-0.42</td>
</tr>
</tbody>
</table>

Similarly,

$$\text{Var}_1(b_2) - \text{Var}_2(b_3) = \frac{(n-1)(nr-1)}{nr^2(r-1)[\Sigma(X_i - \bar{X})^2]^2} \left( \frac{3}{2} \frac{\mu^2}{\beta^2} - \frac{4\mu_2\mu_1}{\beta} + \mu_{20} \mu_{02} - \frac{1}{2} \beta^2 \mu_{20}^2 \right).$$

The last factor can be written

$$\left(\frac{3}{2} \frac{\lambda^2}{\beta^2} - 4 \frac{\lambda \beta}{\beta} + \lambda - \frac{1}{2} \beta^2\right) \mu_{20}^2 = \bar{\sigma}, \text{ say.}$$

Putting $\bar{\sigma} = \frac{\beta}{\sqrt{\lambda}}$, we have

$$\bar{\sigma} = \frac{1}{2} \left(3 - 8 \rho \bar{h} + 2\bar{h}^2 - \bar{h}^4\right) \mu_{20}^2.$$

Thus the equation $\bar{\sigma} = 0$ has exactly the same roots as the
equation $G = 0$, and we can state without further ado that

$$\text{Var}_A(b_2) - \text{Var}_A(b_3) > 0 \text{ when } h_2 < \frac{\beta}{\sqrt{\lambda}} < h_1 \text{ where } h_1 \text{ and } h_2 \text{ are taken from Table 5}.$$  

With respect to the estimator $b_3$ then we arrive at substantially the same conclusions as for the case of independent errors: $b_3$ has lower asymptotic variance than $b_1$ or $b_2$ only for a narrow band of values of $\frac{\sqrt{\lambda}}{\beta}$; outside this band one would do better by making the appropriate choice between $b_1$ and $b_2$. Since one is scarcely ever in a position to decide when $b_3$ is more advantageous, its importance is dubious.

It has not proved feasible at this time to undertake a detailed study of $b_4$ because of the algebraic complexity. It appears that one would need to prepare a triple-entry table, giving values of $\beta_0$ for various values of $\frac{\lambda}{\beta^2}$, $R$, and $\lambda$. 
4. SMALL-SAMPLE RESULTS FOR UNREPLICATED OBSERVATIONS

4.1. Introduction and Assumptions

The remainder of this dissertation deals with two important small-sample characteristics of various slope estimators that have already been proposed; namely, the bias and the mean square error. Section 4 discusses the problem when there is no replication; section 5 takes up the situation when there is replication.

The problem of estimating the parameters of a linear functional relation when both variables are in error and there is no replication has long been regarded as a particularly ugly and intractable one. The procedure which seems to have greatest appeal for statisticians is to estimate $\beta$ by

$$b_L = \frac{\sum w_i y_i}{\sum w_i x_i}, \text{ where } \sum w_i = 0,$$

and $\alpha$ by

$$a = \bar{y} - b_L \bar{x}.$$

Here, and throughout the remainder of section 4, the subscript $i$ denotes the order number when the $x$'s are ordered according to magnitude. It is clear that we must require $\sum w_i = 0$, for otherwise $b_L$ would not be invariant with respect to translation of the coordinate axes. The question is, how does one choose the $w_i$ in an estimator of this type.
Wald, whose 1940 paper (57) first directed the serious attention of statisticians to this class of estimator suggested taking

\[ w_i = -1 \quad \text{for } i = 1, 2, \ldots, \frac{n}{2} \]

\[ w_i = +1 \quad \text{for } i = \frac{n}{2} + 1, \ldots, n \]

when \( n \) is even; that is, he suggested dividing the observations into two equal groups, to one of which is assigned weight -1 and to the other, weight +1. Wald showed that \( b_L \) is a consistent estimator of \( \beta \) provided that the partition of the observations into two groups can be carried out independently of the errors and provided that limit inferior

\[ \frac{(X_1 + \ldots + X_{n/2}) - (X_{n/2+1} + \ldots + X_n)}{n} \]

is positive. If the errors in the \( x_i \) are small enough so that partitioning the observations according to the magnitudes of the \( x_i \) yields the same two groups as obtained by partitioning the observations according to the magnitudes of the \( X_i \), the first condition is satisfied. The second condition merely guarantees the expectation of the denominator does not vanish.

Bartlett (6) subsequently suggested that when the number \( n \) of observations is divisible by three, a better choice of the \( w_i \) is

\[ w_i = -1 \quad \text{for } i = 1, 2, \ldots, \frac{n}{3} \]

\[ w_i = 0 \quad \text{for } i = \frac{n}{3} + 1, \ldots, \frac{2n}{3} \]
$w_i = +1$ for $i = \frac{2n}{3} + 1, \ldots, n$.

Other investigators have shown how Bartlett's partition can itself be improved by adjusting the size of the three groups (23), (40), (41), and (55). Again, in order to insure the consistency of the estimator, the partition into three groups must be carried out independently of the errors in the $x_i$, and the expected denominator must not vanish.

Housner and Brennan (27) and Durbin (18) both suggest that taking $w_i = 1 - \bar{X}$ will give an estimator that is often more efficient than any of the foregoing. It should be realized, however, that consistency of this estimator rests upon more stringent assumptions than those made by Wald and the writers mentioned in the paragraph which precedes this. One must now assume the entire ordering according to the magnitudes of the $x_i$ to be identical with the ordering according to the $X_i$. In some fields of research such an assumption might be decidedly objectionable and unrealistic, but in the physical sciences and the engineering sciences this assumption would ordinarily be regarded as mild. And the greater stringency with respect to the assumptions of Wald, Bartlett, and others is more apparent than real; in most cases it is just as reasonable to assume that one can correctly order all the points as to assume that one can partially order them. In any event occurrence of situations where this
assumption is reasonable seem sufficiently frequent to justify a detailed investigation.

We shall accordingly make the blanket assumption that all \( n \) observations can be ordered independently of the errors. If that is so, the \( w_i \) are simply constants, and \( b_L \) is the ratio of two linear forms in \( x_i \) and \( y_i \). For want of a generally accepted name, we shall henceforth refer to \( b_L \) as the "ratio-of-linear-forms" estimator. The remainder of section 4 deals with an investigation of the bias and the mean square error of \( b_L \) and with a comparison of these quantities with those obtained for \( b_Q \), the estimator of \( \beta \) one would obtain if he simply ignored the errors in the \( x_i \) and minimized the sum of squares of deviations in the vertical direction.

We introduce now the model and the assumptions which form the basis for section 4. Suppose that \( Y_i = \alpha + \beta X_i \), \( i = 1, 2, \ldots, n \), where \( \alpha \) and \( \beta \) are unknown constants, while the \( X_i \) and \( Y_i \) are sure variables. However, it is not possible to observe either \( X_i \) or \( Y_i \), but only \( x_i \) and \( y_i \), where

\[
x_i = X_i + e_i
\]
\[
y_i = Y_i + f_i
\]

\( e_i \) and \( f_i \) are random variables representing the errors of observation. The problem is to estimate \( \alpha \) and \( \beta \). Concerning the distribution of the \( e_i \) and the \( f_i \), it is assumed that

1) \( e_i \) and \( e_j \) are independent if \( i \neq j \)
2) $f_i$ and $f_j$ are independent if $i \neq j$

3) $e_i$ and $f_i$ are independent for all $i$ and $i'$

4) $E(f_i) = 0$ and $E(f_i^2) = \sigma_i^2$

5) Let $c = \min \mid X_{i+1} - X_i \mid$. Then $\text{Prob} \mid e_i \mid > c/2 = 0$. That is to say, the $e_i$ have finite range, extending from $-c/2$ to $+c/2$. This is the assumption which ensures that the ordering according to the $x_i$ is identical with the ordering according to the $X_i$.

6) $E(e_i^2) = \mu_2$, $E(e_i^4) = \mu_4$, $E(e_i^6) = \mu_6$

All odd moments of the $e_i$ are zero.

7) $\Sigma w_i x_i \neq 0$ and $\Sigma w_i X_i \neq 0$. It is easily verified that for each of the choices of the $w_i$ considered in section 4.3 this is the case.

4.2. Expectations of Various Statistics Used in Section 4

To preserve the continuity of the discussion in later parts of section 4, expectations of various statistics employed there are collected in section 4.2. We shall make use of the convention regarding indices introduced in section 2.2; we also introduce the following new symbols:

- $P = \Sigma w_i X_i$
- $E_{XX} = \Sigma(X_i - \bar{X})^2$
- $G_k = E[\Sigma(X_i - \bar{X})e_i]^{2k}$
4.2.1. \( \mathbb{E}(\Sigma w_i e_i)^2 \)

\[
\mathbb{E}(\Sigma w_i e_i)^2 = \Sigma w_1^2 e_i^2 + \Sigma \Sigma w_1 w_j e_i e_j = \mu_2 w_i^2 .
\]

4.2.2. \( \mathbb{E}(\Sigma w_i e_i)^4 \)

\[
\mathbb{E}(\Sigma w_i e_i)^4 = \Sigma w_1^4 e_i^4 + 3 \Sigma \Sigma w_1 w_j e_i e_j = \mu_4 w_i^4 + 3 \mu_2^2 \Sigma w_1 w_j^2 .
\]

4.2.3. \( \mathbb{E}(\Sigma w_i e_i)^6 \)

\[
\mathbb{E}(\Sigma w_i e_i)^6 = \Sigma w_1^6 e_i^6 + 15 \Sigma \Sigma w_1 w_j^2 e_i^2 e_j^2 + 15 \Sigma \Sigma \Sigma w_1 w_j w_k e_i e_j e_k
\]
\[
= \mu_6 w_i^6 + 15 \mu_4 \mu_2 \Sigma w_1 w_j^2 + 15 \mu_2^3 \Sigma w_1 w_j w_k .
\]

4.2.4. \( \mathbb{E}(\Sigma(e_i-e_i))^2 \)

This expectation may be extracted from \([2.2.1 - 1]\) simply by putting the \( r_i \) equal to 1. We obtain

\[
H_2 = \frac{(n-1)^2}{n} \mu_4 + (n^2-3n+5-\frac{3}{n}) \mu_2^2 .
\]
4.2.5. \((\text{GH})_{11} = \mathbb{E}[\Sigma(X_i - \bar{X})e_i]^2 \Sigma(e_i, -e_j)^2\)

\((\text{GH})_{11} = \mathbb{E}[\Sigma(X_i - \bar{X})e_i]^2[\Sigma e_i^2, -\frac{1}{n}(\Sigma e_i^2)]\) .

Now \(\mathbb{E}[\Sigma(X_i - \bar{X})e_i]^2 \Sigma e_i^2 = \mathbb{E} \Sigma \Sigma(X_i - \bar{X})^2 e_i^2 e_j^2 = [\mu_i + (n-1)\mu_2]^2 E_{XX} \), and \(\mathbb{E}[\Sigma(X_i - \bar{X})e_i]^2(\Sigma e_i^2)^2 = \mathbb{E}[\Sigma(X_i - \bar{X})e_i]^2 \Sigma e_i^2 + \Sigma(\Sigma e_i^2)^2) = [\mu_i + (n-1)\mu_2]^2 E_{XX} \)

Therefore

\((\text{GH})_{11} = \left[\frac{n-1}{n} \mu_i + \frac{n^2 - 2n + 3}{n} \mu_2^2\right] E_{XX} \).

4.2.6. \(G_2 = \mathbb{E}[\Sigma(X_i - \bar{X})e_i]^2\)

\(G_2 = \mathbb{E}[\Sigma(X_i - \bar{X})e_i]^2 + 3 \Sigma \Sigma(X_i - \bar{X})^2(X_j - \bar{X})^2 e_i^2 e_j^2\)

\(= \mu_i \Sigma(X_i - \bar{X})^4 + 3 \mu_2^2 \Sigma \Sigma(X_i - \bar{X})^2(X_j - \bar{X})^2 \).

4.2.7. \((\text{GH})_{12} = \mathbb{E}[\Sigma(X_i - \bar{X})e_i]^2 [\Sigma(e_i, -e_j)^2]^2\)

\((\text{GH})_{12} = \mathbb{E}[\Sigma(X_i - \bar{X})e_i]^2 [(\Sigma e_i^2)^2 - 2ne_i^2 e_j^2 + n^2 e_j^2] \).

Now $E[\Sigma(X_i - \overline{X})e_i]^2(\Sigma e_i^2)^2 = E[\Sigma(X_i - \overline{X})e_i]^2[\Sigma e_i^4 + \Sigma \Sigma e_i^2 e_j^2]$

\[= [\mu_6 + (n-1)\mu_4 \mu_2]E_{XX} + [2(n-1)\mu_4 \mu_2 + (n-1)(n-2)\mu_2^3]E_{XX}\]

\[= [\mu_6 + 3(n-1)\mu_4 \mu_2 + (n-1)(n-2)\mu_2^3]E_{XX},\]

and $E[\Sigma(X_i - \overline{X})e_i]^2(\Sigma e_i^2)^2(\Sigma e_i^4)^2 = E[\Sigma(X_i - \overline{X})e_i]^2[\Sigma e_i^2 + \Sigma \Sigma e_i^2 e_j^2]$

\[= [\mu_6 + 3(n-1)\mu_4 \mu_2 + (n-1)(n-2)\mu_2^3]E_{XX}\]

\[= [\mu_6 + 4(n-2)\mu_4 \mu_2 + 2(n-2)\mu_2^3]E_{XX},\]

and $E[\Sigma(X_i - \overline{X})e_i]^2(\Sigma e_i^2)^4 = E[\Sigma(X_i - \overline{X})e_i]^2[\Sigma e_i^4 + 4\Sigma \Sigma e_i^2 e_j^2 + 3\Sigma e_i^2 e_j^2 + \Sigma \Sigma \Sigma e_i^2 e_j^2 e_k^2 + \Sigma \Sigma \Sigma e_i^2 e_j^2 e_k e_m^2]$

\[= [\mu_6 + (n-1)\mu_4 \mu_2]E_{XX} + 4(-2\mu_4 \mu_2)E_{XX}\]

\[+ 3[2(n-1)\mu_4 \mu_2 + (n-1)(n-2)\mu_2^3]E_{XX}\]

\[+ 6[-2(n-2)\mu_2^3]E_{XX} + 0\]

\[= [\mu_6 + (7n-15)\mu_4 \mu_2 + 3(n-2)(n-5)\mu_2^3]E_{XX}.\]

Consequently,
\[(GH)_{12} = \{\mu_6 + 3(n-1)\mu_4\mu_2 + (n-1)(n-2)\mu_2^3\}E_{XX}
\]
\[-\frac{2}{n}\{\mu_6 + (3n-7)\mu_4\mu_2 + (n-2)(n-3)\mu_2^3\}E_{XX}
\]
\[+ \frac{1}{n^2}\{\mu_6 + (7n-15)\mu_4\mu_2 + 3(n-2)(n-5)\mu_2^3\}E_{XX}.
\]

4.2.8. \[(GH)_{21} = E[\Sigma(X_i - \overline{X})e_i]4\Sigma(e_i,e_j)^2
\]
\[(GH)_{21} = E[\Sigma(X_i - \overline{X})\Sigma(e_i,e_j)]4(0)
\]
\[+ 3\Sigma\Sigma(X_i - \overline{X})e_i^2 + 6\Sigma\Sigma(X_i - \overline{X})(X_j - \overline{X})e_i^2 e_j e_k
\]
\[+ \Sigma\Sigma\Sigma(X_i - \overline{X})(X_j - \overline{X})(X_k - \overline{X})e_i^2 e_j e_k e_m \Sigma e_i^2
\]
\[- \frac{1}{n}(2e_i,e_i)^2].
\]

\[(GH)_{21} = [\mu_6 + (n-1)\mu_4\mu_2] \Sigma(X_i - \overline{X})4 + 4(0)
\]
\[+ 3[2\mu_4\mu_2 + (n-2)\mu_2^3] \Sigma(X_i - \overline{X})e_i^2 \Sigma(X_j - \overline{X})^2 + 6(0) + 0
\]
\[- \frac{1}{n}\{\mu_6 + (n-1)\mu_4\mu_2\} \Sigma(X_i - \overline{X})4 + \frac{8}{n} \mu_4\mu_2 \Sigma(X_i - \overline{X})4
\]
\[- \frac{3}{n}[2\mu_4\mu_2 + (n-2)\mu_2^3] \Sigma(X_i - \overline{X})^2 \Sigma(X_j - \overline{X})^2
\]
\[+ \frac{12}{n} \mu_2^3 \Sigma(X_i - \overline{X})^2 \Sigma(X_j - \overline{X})^2 - \Sigma(X_i - \overline{X})4].
\]

Simplifying, we have
\[(GH)_{21} = [(1 - \frac{1}{n}) \mu_6 + (n-1 - \frac{n-9}{n}) \mu_4 \mu_2 + \frac{12}{n} \mu_2^3] \Sigma (x_i - \bar{x})^4 \]
\[+ [6(1 - \frac{1}{n}) \mu_4 \mu_2 + 3(n-2) \mu_2^3] \Sigma (x_i - \bar{x})^2 (x_j - \bar{x})^2. \]

4.2.9. \[G_3 = E[\Sigma (x_i - \bar{x}) e_i]^6 \]

\[G_3 = E[\Sigma (x_i - \bar{x}) e_i^6] + 15 \Sigma (x_i - \bar{x})^4 (x_j - \bar{x})^2 e_i^2 e_j^2 \]
\[+ 15 \Sigma (x_i - \bar{x})^2 (x_j - \bar{x})^2 (x_k - \bar{x})^2 e_i^2 e_j^2 e_k^2]. \]

Therefore \[G_3 = \mu_6 \Sigma (x_i - \bar{x})^6 + 15 \mu_4 \mu_2 \Sigma (x_i - \bar{x})^4 (x_j - \bar{x})^2 \]
\[+ 15 \mu_2^3 \Sigma (x_i - \bar{x})^2 (x_j - \bar{x})^2 (x_k - \bar{x})^2. \]

4.2.10. \[H_3 = E[\Sigma (e_i - e_j)^2]^3 \]

Following the notation of Kendall (32) we put \[\Sigma (e_i - e_j)^2 = nm_2. \] Since \(k_2 = \frac{n}{n-1} m_2\), we obtain

\[H_3 = (n-1)^3 E_k^3 = (n-1)^3 \left\{ \kappa(2^3) + 3 \kappa(2^2) \kappa(2^1) + [\kappa(2^1)]^3 \right\} \]
\[= (n-1)^3 \left[ \frac{\kappa_6}{n^2} + \frac{12 \kappa_4 \kappa_2}{n(n-1)} + \frac{4(n-2)}{n(n-1)^2} \kappa_3^2 + \frac{8}{(n-1)^2} \kappa_2^3 \right. \]
\[+ 3 \left( \frac{\kappa_4}{n} + \frac{2 \kappa_2^2}{n-1} \right) \kappa_2 + \kappa_2^3 \]
\[= (n-1)^3 \left[ \frac{\mu_6 - 15 \mu_4 \mu_2 + 30 \mu_2^3}{n^2} \right. \]
\[+ \left( \frac{12}{n(n-1)} + \frac{3}{n} \right) (\mu_4 - 3 \mu_2^2) \mu_2 + \left( \frac{8}{(n-1)^2} + \frac{6}{n-1} + 1 \right) \mu_2^3 \].
4.2.11. $(GH)_{13} = E[\Sigma(X_i - \bar{X})e_i]^2[\Sigma(e_i - e)^2]^3$

$$(GH)_{13} \leq \max_{e_i} E[\Sigma(X_i - \bar{X})e_i]^2[\Sigma(e_i - e)^2]^2.$$

Now $\max_{e_i} [\Sigma(e_i - e)^2] = n \frac{c^2}{4}$, for the maximum moment of inertia for a system of homogeneous particles is obtained when the particles are placed symmetrically and as far from the axis of symmetry as possible. Consequently,

$$(GH)_{13} \leq n \left(\frac{c^2}{2}\right)^2 (GH)_{12}.$$

4.2.12. $(GH)_{22} = E[\Sigma(X_i - \bar{X})e_i]^{4}[\Sigma(e_i - e)^2]^2$ 

$$(GH)_{22} \leq n \left(\frac{c^2}{2}\right)^2 (GH)_{21}.$$

4.2.13. $H_4 = E[\Sigma(e_i - e)^2]^4$

$$H_4 \leq n \left(\frac{c^2}{2}\right)^2 H_3.$$

4.2.14. $H_5 = E[\Sigma(e_i - e)^2]^5$

$$H_5 \leq n^2 \left(\frac{c^2}{2}\right)^4 H_3.$$
**4.2.15.** $H_6 = E[\Sigma (e_i - \bar{e})^2]^6$

$H_6 \leq n^3 \left( \frac{c}{2} \right)^6 H_3$.

**4.3. Ratio-of-Linear-Forms Estimator**

In this section we shall make a detailed study of the estimator $b_L$. In particular, we shall examine the bias and the mean square error for various choices of the $w_i$.

**4.3.1. Bias of $b_L$**

Consider the estimator $b_L = \frac{\Sigma w_i y_i}{\Sigma w_i x_i}$. Since the $y_i$ and the $x_i$ are independent,

$$E(b_L) = E(\Sigma w_i y_i) E\left( \frac{1}{\Sigma w_i x_i} \right).$$

Now $\frac{1}{\Sigma w_i x_i} = \frac{1}{P + \Sigma w_i e_i}$

$$= \frac{1}{P} \sum_{k=0}^{t} (-1)^{k+1}(\Sigma w_i e_i) \frac{k}{k+2} + (-1)^{t+2} \frac{(\Sigma w_i e_i)^{t+1}}{1 + \frac{\Sigma w_i e_i}{P}}.$$

Consequently,

$$E \left( \frac{1}{\Sigma w_i x_i} \right) = \frac{1}{P} \sum_{k=0}^{t} (-1)^{k+1}(\Sigma w_i e_i) \frac{k}{k+2} + D,$$

where $D = \frac{(-1)^t}{P + 2} \frac{(\Sigma w_i e_i)^{t+1}}{1 + \frac{\Sigma w_i e_i}{P}}.$
Except for a few situations, when \( n = 4 \), it has proved sufficiently accurate for the purposes of this investigation to take \( t = 5 \). Then

\[
E \left( \frac{1}{Z_{w_1} X_1} \right) = \frac{1}{p} + \frac{1}{p^3} E(S_{w_1} e_i)^2 + \frac{1}{p^7} E(S_{w_1} e_i)^4 + D
\]

where

\[
D = \frac{1}{p^7} E \left( \frac{(S_{w_1} e_i)^6}{1 + \frac{1}{p} \left( \frac{S_{w_1} e_i}{Z_{w_1} e_i} \right)} \right).
\]

In a few cases it has been necessary to take \( t = 7 \). We shall carry out the details of the analysis only for \( t = 5 \) as the modifications required when \( t = 7 \) are obvious.

Since \( E(S_{w_1} y_i) = \beta P \), we have

\[
E(b_L) = \beta \left[ 1 + \frac{1}{p^2} E(S_{w_1} e_i)^2 + \frac{1}{p^4} E(S_{w_1} e_i)^4 \right] + \beta PD.
\]

Making use of 4.2.1 and 4.2.2 we have

\[
E(b_L) = \beta \left[ 1 + \frac{1}{p^2} \mu_2 S_{w_1}^2 + \frac{1}{p^4} \left( \mu_4 S_{w_1}^4 + 3 \mu_2^2 S_{w_1}^2 \right) \right] + \beta PD.
\]

Let us denote the bias of \( b_L \) by \( B_L \). Then

\[
B_L = \beta \left[ \frac{1}{p^2} \mu_2 S_{w_1}^2 + \frac{1}{p^4} \left( \mu_4 S_{w_1}^4 + 3 \mu_2^2 S_{w_1}^2 \right) \right] + \beta PD.
\]

In all of the cases studied in section 4.3 it is easily verified that

\[
-1 < \frac{S_{w_1} e_i}{Z_{w_1} e_i} < 1.
\]

Consequently, for the cases we want to investigate it is possible to find simple bounds for \( D \). In fact,
Making use of 4.2.3, we find

\[ 0 < D < \frac{1}{2p^7} \left( \frac{1}{\sum |w_i|} + 1 \right) E(\Sigma w_i e_i)^6. \]

Thus we have the means for approximating the bias to any desired degree of accuracy, and furthermore, we have a way of assessing the accuracy of any approximation.

### 4.3.2. Mean square error of \( b_L \)

It is possible to obtain the mean square error, \( M_L \), of \( b_L \) in essentially the same way as \( B_L \). To begin with,

\[ E(b_L^2) = E(\Sigma w_i y_i)^2 E\left( \frac{1}{\Sigma w_i x_i} \right)^2. \]

Now

\[ \left( \frac{1}{\Sigma w_i x_i} \right)^2 = \frac{1}{p^2} \left[ 1 - \frac{2\Sigma w_i e_i}{p} + \frac{3(\Sigma w_i e_i)^2}{p^2} - \frac{4(\Sigma w_i e_i)^3}{p^3} \right. \]

\[ \left. + \frac{5(\Sigma w_i e_i)^4}{p^4} + \frac{1}{p^2} \right] \left( \frac{\Sigma w_i e_i}{p} \right)^6 + 6 \left( \frac{\Sigma w_i e_i}{p} \right)^7 \frac{1}{1 + \frac{\Sigma w_i e_i}{p}}. \]

Consequently,
\[
E \left( \frac{1}{\Sigma w_i x_i} \right)^2 = \frac{1}{p^2} [1 + \frac{3}{p^2} \mu_2 \Sigma w_i^2 + \frac{5}{p^4} (\mu_1^2 \Sigma w_i^4 + 3\mu_2^2 \Sigma \Sigma w_i^2 w_j^2)] + D',
\]

where \( D' = \frac{1}{p^2} E \frac{7(\Sigma w_i e_i)^6}{(1 + \frac{\Sigma w_i e_i}{p})^7}. \)

Since \( E(\Sigma w_i y_i)^2 = [E(\Sigma w_i y_i)]^2 + \text{Var}(\Sigma w_i y_i) = \beta^2 p^2 + \sigma^2_{\Sigma w_i} \), we obtain

\[
E(b_L^2) = \beta^2 [1 + \frac{3}{p^2} \mu_2 \Sigma w_i^2 + \frac{5}{p^4} (\mu_1^2 \Sigma w_i^4 + 3\mu_2^2 \Sigma \Sigma w_i^2 w_j^2)] + \frac{\sigma^2_{\Sigma w_i}^2}{p^2} [1 + \frac{3}{p^2} \mu_2 \Sigma w_i^2 + \frac{5}{p^4} (\mu_1^2 \Sigma w_i^4 + 3\mu_2^2 \Sigma \Sigma w_i^2 w_j^2)]
\]

\[
+ (\beta^2 p^2 + \sigma^2_{\Sigma w_i}^2)D'. \]

Now \( M_L = E(b_L^2) - 2\beta E(b_L) + \beta^2 \). Making use of this relation, and putting \( \lambda = \frac{\sigma^2_{\Sigma w_i}}{\mu_2} \), we have

\[
M_L = \beta^2 \left[ \frac{1}{p^2} \mu_2 \Sigma w_i^2 + \frac{3}{p^4} (\mu_1^2 \Sigma w_i^4 + 3\mu_2^2 \Sigma \Sigma w_i^2 w_j^2) \right] + \lambda \left[ \frac{1}{p^2} \mu_2 \Sigma w_i^2 \right.
\]

\[
+ \frac{3}{p^4} \mu_2 (\Sigma w_i^2)^2 + \frac{5}{p^6} \mu_2 \Sigma w_i^2 (\mu_1^2 \Sigma w_i^4 + 3\mu_2^2 \Sigma \Sigma w_i^2 w_j^2),
\]

\[
+ \triangle, \quad \text{where}
\]

\[
\triangle = (\beta^2 p^2 + \sigma^2_{\Sigma w_i}^2)D' - 2\beta^2 PD .
\]

In order to obtain simple bounds for \( \triangle \), we shall first show that when \( -1 < \frac{\Sigma w_i e_i}{p} < 1 \),
In order to do this, let \( z = \frac{\Sigma w_i e_i}{P} \) and let \( a = \max z \). Then

\[
E \frac{z^7}{(1+z)^2} = \int_{-a}^{a} \frac{z^7}{(1+z)^2} f(z)dz
\]

\[
< \int_{-a}^{0} \frac{z^7}{(1-a)^2} f(z)dz + \int_{0}^{a} z^7 f(z)dz .
\]

Since \( \int_{-a}^{0} z^7 f(z)dz = - \int_{0}^{a} z^7 f(z)dz \),

\[
E \frac{z^7}{(1+z)^2} < \left[ 1 - \left( \frac{1}{1-a} \right)^2 \right] \int_{0}^{a} z^7 f(z)dz ;
\]

that is, \( E \frac{z^7}{(1+z)^2} < 0 \). Consequently,

\[
0 < D' < \frac{7}{2p^2} \left[ \left( 1 - \frac{\Sigma |w_i|}{P} \right)^{-2} + 1 \right] E(\Sigma w_i e_i)^6 ,
\]

and

\[
0 < \Delta < \beta^2 \left[ \frac{7}{2} \left( 1 - \frac{\Sigma |w_i|}{P} \right)^{-2} + \frac{5}{2} - \left( 1 + \frac{\Sigma |w_i|}{P} \right)^{-1} \right] E(\Sigma w_i e_i)^6
\]

\[
+ \frac{\lambda \mu_2 \Sigma w_i^2}{p^2} \left[ \frac{7}{2} \left( 1 - \frac{\Sigma |w_i|}{P} \right)^{-2} + \frac{7}{2} \right] E(\Sigma w_i e_i)^6 ,
\]

where
Accordingly it is possible to approximate $M_L$ in a fairly simple way and to assess the accuracy of the approximation.

4.3.3. Optimal choice of the $w_i$

We should like now to consider two possible sets of $w_i$; namely,

1. The set which minimizes the bias $B_L$
2. The set which minimizes the mean square error $M_L$.

In general, of course, the two sets differ. The exact specification of either set does not appear possible in view of the complexity of the expressions for $B_L$ and $M_L$. On the other hand it is possible to obtain very good approximate specifications.

If one chooses the $w_i$ according to the scheme of Wald, or Bartlett, or Housner and Brennan, $\max \frac{\Sigma w_i e_i}{P}$ turns out to be considerably less than unity, even for values of $n$ as small as four; this is a consequence of assumption (5) of section 4.1. This leads one to conjecture that $\max \frac{\Sigma w_i e_i}{P}$ is small for any reasonable choice of the $w_i$ -- in particular, one would expect it to be small for good choices, like that which minimizes the bias or the mean square error. If this is the case, one would expect the bias and the mean square
error to be determined primarily by the term of order $P^{-2}$ in 
the expression for each; consequently, minimizing that term 
should approximately minimize the bias or the mean square 
error, as the case may be. Now this term in the expression 
for the bias is $\mu_2 \beta P^{-2} \Sigma w_i^2$, whereas the corresponding term in 
the expression for the mean square error is $\mu_2 (\beta^2 + \lambda) P^{-2} \Sigma w_i^2$. 
Obviously both of these terms are minimized by the same set of $w_i$, which has been obtained in section 2.3.5. We showed there 

$$w_i = (X_i - \bar{X})C$$  \[1\]

where $C$ is any arbitrary constant $\neq 0$.

Up to this point of section 4.3.3 we have proceeded 
heuristically; we should now like to show with the aid of sev­
eral examples that retention of terms in $P^{-4}$ makes only a 
trifling difference in the weights and far less difference in 
the bias itself.

As the first example, we consider four points: 

$X_1 = -\frac{3c}{2}$, $X_2 = -\frac{c}{2}$, $X_3 = \frac{c}{2}$, and $X_4 = \frac{3c}{2}$. We take 

$\mu_2 = \frac{1}{2} \left( \frac{c}{2} \right)^2$ and $\mu_4 = \frac{1}{5} \left( \frac{c}{2} \right)^4$ which happen to be the moments of 
the rectangular distribution on the interval $(-\frac{c}{2}, \frac{c}{2})$, whose 
range is the maximum possible consonant with the assumptions 
of section 4.1. We might equally well consider any other 
distribution on this finite range; the same technique would 
apply, and the results would be substantially the same.
Equation \([4.3.3 - 1]\) gives as the weights \(w_1 = -3\), \(w_2 = -1\), \(w_3 = +1\), \(w_4 = +3\) (or any multiple of these, of course).

Minimization of \(B_L = \frac{1}{p^2} \mu_2 \Sigma w_1^2 + \frac{1}{p^4} (\mu_4 \Sigma w_1^4 + 3 \mu_2^2 \Sigma w_1^2 w_j^2)\)
gives for the \(w_i\)
\[
\begin{align*}
    w_1 &= -3.04\frac{1}{10}, \\
    w_2 &= -1, \\
    w_3 &= +1, \\
    w_4 &= +3.04\frac{1}{10}.
\end{align*}
\]
This introduces a relative change in the bias, \(\frac{\Delta B_L}{B_L}\), of
\[
9 \times 10^{-5}.
\]

As a second example, we consider the four points
\[
X_1 = -\frac{5c}{2}, \quad X_2 = -\frac{c}{2}, \quad X_3 = \frac{c}{2}, \quad \text{and} \quad X_4 = \frac{5c}{2},
\]
with \(\mu_2\) and \(\mu_4\) just as in the previous example. Equation \([4.3.3 - 1]\) gives
\[
\begin{align*}
    w_1 &= -5, \\
    w_2 &= -1, \\
    w_3 &= +1, \\
    w_4 &= +5.
\end{align*}
\]

Minimization of \(B_L = \frac{1}{p^2} \mu_2 \Sigma w_1^2 + \frac{1}{p^4} (\mu_4 \Sigma w_1^4 + 3 \mu_2^2 \Sigma w_1^2 w_j^2)\)
gives for the \(w_i\)
\[
\begin{align*}
    w_1 &= -5.03\frac{1}{10}, \\
    w_2 &= -1, \\
    w_3 &= +1, \\
    w_4 &= 5.03\frac{1}{10}.
\end{align*}
\]
This introduces a relative change in the bias \(\frac{\Delta B_L}{B_L}\) of
\[
3 \times 10^{-6}.
\]

It will be observed that the perturbations produced both in the weights and in the bias by the term in \(p^{-4}\) are smaller in the second example than in the first. That is exactly what one would expect, for although the spread of the \(X_i\) has increased in the second example, the range of the \(e_i\) has not
changed. This has the effect of decreasing $\max \frac{\sum w_i e_i}{p}$, thereby decreasing the importance of the term in $p^{-4}$ relative to that of the term in $p^{-2}$.

As a final example we consider the six points $X_1 = -\frac{5c}{2}$, $X_2 = -\frac{3c}{2}$, $X_3 = -\frac{c}{2}$, $X_4 = \frac{c}{2}$, $X_5 = \frac{3c}{2}$, and $X_6 = \frac{5c}{2}$, with $\mu_2$ and $\mu_4$ the same as in the preceding examples. For the weights [4.3.3 - 1] gives

$w_1 = -5$, $w_2 = -3$, $w_3 = -1$, $w_4 = 1$, $w_5 = 3$, $w_6 = 5$.

Minimization of $B_L = \frac{1}{p^2} \mu_2 \sum w_i^2 + \frac{1}{p^4} (\mu_4 \sum w_i^4 + 3 \mu_2^2 \sum w_i^2 w_j^2)$ is exceedingly laborious; the results are

$w_1 = -5.00007$, $w_2 = -3.00001$, $w_3 = -1$, $w_4 = +1$, $w_5 = +3.00001$, $w_6 = +5.00007$.

Clearly, the use of these weights instead of the weights given by [4.3.3 - 1] would make only a trifling change in the bias.

One might examine the mean square error in similar fashion but with increased labor due to the additional parameter $\lambda$. We have not actually undertaken such an investigation; we feel that the considerations set forth in the next paragraph obviate its necessity.

The remainder of section 4.3 compares the bias and mean
square error of various ratio-of-linear-forms estimators such as the Wald estimator and the Housner-Brennan estimator with the corresponding properties of the estimator \( \frac{\sum(X_i - \bar{X})y_i}{\sum(X_i - \bar{X})x_i} \) of section 2.3.5. The investigation embraces a wide range of values of \( n \) and many different spacings of the \( X_i \). Its import is unmistakable: the bias and the mean square error both have very flat minimums. Consequently the approximate minimization of the bias and the mean square error given by choosing \( w_i = X_i - \bar{X} \) is a very good approximation, and the estimator which employs these weights can, to all intents and purposes, be regarded as having minimum bias and minimum mean square error. It is interesting that the conclusions based on the small-sample approach employed here agree with those based on the large-sample approach in section 2. We shall refer to this choice of weights as optimal and to the estimator \( \frac{\sum(X_i - \bar{X})y_i}{\sum(X_i - \bar{X})x_i} \) as the optimal (ratio-of-linear-forms) estimator, which we shall denote by the symbol \( b_0 \).

4.3.4. Comparison of various well-known estimators with the optimal estimator

In general the spacing of the \( X_i \) is unknown, and the optimal estimator is therefore not obtainable. It is cus-
tomary to use the weights of Wald, of Bartlett, or of Housner and Brennan. The corresponding estimators will henceforth be denoted by $b_w^*$, $b_B^*$, $b_H^*$ respectively, and the corresponding biases and mean square errors will also carry these sub­scripts. For example, $b_B^*$ denotes the bias of Bartlett's estimator.

When the $X_i$ are uniformly spaced, $b_H^*$ and $b_0$ are identi­cal, which is to say that the Housner-Brennan estimator is optimal for this spacing. For any other spacing of the $X_i$ neither $b_H^*$, $b_B^*$, nor $b_w^*$ is optimal; the question is, how far do they depart from optimality.

To answer this question we have studied the bias and the mean square error of these estimators for $n = 4, 6, 8, 12, 16,$ and 20 and for various types of spacings of the $X_i$, which may be classified as follows:

1. Symmetric spacing: $\ldots, -5^\omega p, -3^\omega p, -1^\omega p, 1^\omega p, 3^\omega p, 5^\omega p, \ldots$ for $\omega \geq 0$

2. Symmetric spacing: $-1^\omega p, -2^\omega p, -3^\omega p, \ldots, 3^\omega p, 2^\omega p, 1^\omega p$ for $\omega < 0$

3. Asymmetric spacing: $0, 2(1^\omega p), 2(2^\omega p), 2(3^\omega p), \ldots$ for $\omega > 1$ and $0 < \omega < 1$

4. Asymmetric spacing: $-1^\omega p, -2^\omega p, -3^\omega p, \ldots$ for $\omega < 0$.
In addition we have taken

\[ \mu_2 = \frac{1}{3} \left( \frac{c}{2} \right)^2 \]

\[ \mu_4 = \frac{1}{7} \left( \frac{c}{2} \right)^4 \]

\[ \mu_6 = \frac{1}{7} \left( \frac{c}{2} \right)^6 \]

which coincide with the second, fourth, and sixth moments of a rectangular distribution having the maximum range consonant with the assumptions of 4.1. As indicated in section 4.3.3 there is no serious loss of generality in confining our attention to a rectangular distribution of errors.

Tables 6 to 12 present

1. The bias of the optimal estimator.
2. The ratio of the bias of \( b_w, b_B, \) and \( b_H \) to the bias of the optimal estimator.
3. The ratio of the mean square error of \( b_w, b_B, \) and \( b_H \) to that of the optimal estimator.

The values cited in the tables are correct to within one unit in the last place given there.

Investigation of the ratios \( \frac{M_H}{M_0}, \frac{M_W}{M_0}, \) and \( \frac{M_B}{M_0} \) is complicated only slightly by the presence of the additional parameter \( \lambda \). For example, \( \frac{M_H}{M_0} \) takes the form
\[ \frac{F_1 \beta^2 + F_2 \lambda}{F_3 \beta^2 + F_4 \lambda} = \frac{F_1 + F_2 \frac{\lambda}{\beta^2}}{F_3 + F_4 \frac{\lambda}{\beta^2}} \]

where \(F_1, F_2, F_3,\) and \(F_4\) are functions of the \(w_i,\) the \(X_i,\) and the \(\mu_i.\) If we consider \(\frac{M_H}{M_0}\) as a function of \(\frac{\lambda}{\beta^2},\) we know that it must take one of the two forms shown in Figure 1, depending upon the \(F_i.\)

It is clear that in the range \(0 < \frac{\lambda}{\beta^2} \leq \infty\) \(\frac{M_H}{M_0}\) is either monotonic increasing or monotonic decreasing. We have evaluated \(\frac{M_H}{M_0}\) for \(\frac{\lambda}{\beta^2} = 0\) and \(\frac{\lambda}{\beta^2} = \infty;\) where the values agree up to the third decimal place, we cite the common value, where they differ we cite them both.

Exactly the same sort of analysis applies to \(\frac{M_W}{M_0}\) and to \(\frac{M_B}{M_0}\).

As we have indicated in section 1.3, \(E_L^2\) is much smaller than \(M_L;\) consequently, the variance of \(b_L\) is essentially equal to \(M_L,\) and the ratios we have calculated for the mean-square-errors are substantially the same for the corresponding variances. The same comment applied to the other estimators studied in sections 4 and 5.
Figure 1. $\frac{M_H}{M_0}$ as a function of $\frac{\lambda}{\phi^2}$
Table 6. Maximum value of $B_0/\beta$

<table>
<thead>
<tr>
<th>$\omega n$</th>
<th>4</th>
<th>6</th>
<th>8</th>
<th>12</th>
<th>16</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Symmetric spacing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.0</td>
<td>0.00204</td>
<td>0.00024</td>
<td>0.00005</td>
<td>0.00001</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>1.5</td>
<td>0.00603</td>
<td>0.00109</td>
<td>0.00033</td>
<td>0.00006</td>
<td>0.0002</td>
<td>0.0002</td>
</tr>
<tr>
<td>1.2</td>
<td>0.0114</td>
<td>0.00268</td>
<td>0.00098</td>
<td>0.00024</td>
<td>0.0009</td>
<td>0.0004</td>
</tr>
<tr>
<td>1.0</td>
<td>0.0174</td>
<td>0.00482</td>
<td>0.00199</td>
<td>0.00058</td>
<td>0.00024</td>
<td>0.0012</td>
</tr>
<tr>
<td>0.75</td>
<td>0.0113</td>
<td>0.00282</td>
<td>0.00107</td>
<td>0.00030</td>
<td>0.0008</td>
<td>0.0006</td>
</tr>
<tr>
<td>0.5</td>
<td>0.00566</td>
<td>0.00121</td>
<td>0.00042</td>
<td>0.00012</td>
<td>0.0003</td>
<td>0.0002</td>
</tr>
<tr>
<td>-0.5</td>
<td>0.00239</td>
<td>0.00040</td>
<td>0.00011</td>
<td>0.00006</td>
<td>0.0002</td>
<td>0.0001</td>
</tr>
<tr>
<td>-1.0</td>
<td>0.00852</td>
<td>0.00085</td>
<td>0.00020</td>
<td>0.00003</td>
<td>0.0002</td>
<td>0.0001</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.0100</td>
<td>0.00075</td>
<td>0.00009</td>
<td>0.00001</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td><strong>Asymmetric spacing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.0</td>
<td>0.00171</td>
<td>0.00017</td>
<td>0.00004</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>1.5</td>
<td>0.00546</td>
<td>0.00090</td>
<td>0.00027</td>
<td>0.00005</td>
<td>0.0002</td>
<td>0.0001</td>
</tr>
<tr>
<td>1.2</td>
<td>0.0108</td>
<td>0.00245</td>
<td>0.00088</td>
<td>0.00022</td>
<td>0.00008</td>
<td>0.0004</td>
</tr>
<tr>
<td>0.75</td>
<td>0.0107</td>
<td>0.00296</td>
<td>0.00012</td>
<td>0.00038</td>
<td>0.00015</td>
<td>0.0008</td>
</tr>
<tr>
<td>0.5</td>
<td>0.00049</td>
<td>0.00139</td>
<td>0.00060</td>
<td>0.00019</td>
<td>0.00007</td>
<td>0.0003</td>
</tr>
<tr>
<td>-0.5</td>
<td>0.00325</td>
<td>0.00055</td>
<td>0.00013</td>
<td>0.00002</td>
<td>0.0001</td>
<td>--</td>
</tr>
<tr>
<td>-1.0</td>
<td>0.00171</td>
<td>0.00018</td>
<td>0.00004</td>
<td>0.00001</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>-2.0</td>
<td>0.00034</td>
<td>0.00002</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>
Table 7. $\frac{B_H}{B_0}$

<table>
<thead>
<tr>
<th>$\omega$</th>
<th>$n$</th>
<th>4</th>
<th>6</th>
<th>8</th>
<th>12</th>
<th>16</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Symmetric spacing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.0</td>
<td>1.046</td>
<td>1.057</td>
<td>1.061</td>
<td>1.064</td>
<td>1.065</td>
<td>1.066</td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td>1.018</td>
<td>1.019</td>
<td>1.020</td>
<td>1.020</td>
<td>1.020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.2</td>
<td>1.004</td>
<td>1.004</td>
<td>1.004</td>
<td>1.004</td>
<td>1.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.75</td>
<td>1.008</td>
<td>1.008</td>
<td>1.008</td>
<td>1.008</td>
<td>1.008</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>1.042</td>
<td>1.042</td>
<td>1.044</td>
<td>1.039</td>
<td>1.042</td>
<td>1.042</td>
<td></td>
</tr>
<tr>
<td>-0.5</td>
<td>1.092</td>
<td>1.083</td>
<td>1.075</td>
<td>1.063</td>
<td>1.063</td>
<td>1.065</td>
<td></td>
</tr>
<tr>
<td>-1</td>
<td>1.020</td>
<td>1.020</td>
<td>1.035</td>
<td>1.086</td>
<td>1.139</td>
<td>1.196</td>
<td></td>
</tr>
<tr>
<td>-2</td>
<td>1.006</td>
<td>1.095</td>
<td>1.212</td>
<td>1.473</td>
<td>1.745</td>
<td>2.021</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Asymmetric spacing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.0</td>
<td>1.089</td>
<td>1.085</td>
<td>1.082</td>
<td>1.077</td>
<td>1.075</td>
<td>1.073</td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td>1.032</td>
<td>1.030</td>
<td>1.027</td>
<td>1.025</td>
<td>1.024</td>
<td>1.023</td>
<td></td>
</tr>
<tr>
<td>1.2</td>
<td>1.006</td>
<td>1.006</td>
<td>1.006</td>
<td>1.005</td>
<td>1.005</td>
<td>1.005</td>
<td></td>
</tr>
<tr>
<td>0.75</td>
<td>1.016</td>
<td>1.016</td>
<td>1.015</td>
<td>1.013</td>
<td>1.012</td>
<td>1.011</td>
<td></td>
</tr>
<tr>
<td>0.50</td>
<td>1.084</td>
<td>1.105</td>
<td>1.095</td>
<td>1.087</td>
<td>1.079</td>
<td>1.073</td>
<td></td>
</tr>
<tr>
<td>-0.50</td>
<td>1.096</td>
<td>1.166</td>
<td>1.225</td>
<td>1.318</td>
<td>1.393</td>
<td>1.455</td>
<td></td>
</tr>
<tr>
<td>-1</td>
<td>1.116</td>
<td>1.297</td>
<td>1.418</td>
<td>1.633</td>
<td>1.822</td>
<td>1.997</td>
<td></td>
</tr>
<tr>
<td>-2</td>
<td>1.314</td>
<td>1.620</td>
<td>1.922</td>
<td>2.512</td>
<td>3.092</td>
<td>3.666</td>
<td></td>
</tr>
</tbody>
</table>
Table 8. $\frac{B_w}{B_0}$

<table>
<thead>
<tr>
<th>$\omega$</th>
<th>4</th>
<th>6</th>
<th>8</th>
<th>12</th>
<th>16</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symmetric spacing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1.65</td>
<td>1.73</td>
<td>1.76</td>
<td>1.78</td>
<td>1.79</td>
<td>1.79</td>
</tr>
<tr>
<td>1.5</td>
<td>1.47</td>
<td>1.52</td>
<td>1.54</td>
<td>1.55</td>
<td>1.56</td>
<td>1.56</td>
</tr>
<tr>
<td>1.2</td>
<td>1.34</td>
<td>1.39</td>
<td>1.41</td>
<td>1.42</td>
<td>1.42</td>
<td>1.42</td>
</tr>
<tr>
<td>1.0</td>
<td>1.27</td>
<td>1.30</td>
<td>1.32</td>
<td>1.32</td>
<td>1.33</td>
<td>1.33</td>
</tr>
<tr>
<td>0.75</td>
<td>1.16</td>
<td>1.20</td>
<td>1.20</td>
<td>1.21</td>
<td>1.22</td>
<td>1.22</td>
</tr>
<tr>
<td>0.50</td>
<td>1.08</td>
<td>1.09</td>
<td>1.10</td>
<td>1.11</td>
<td>1.12</td>
<td>1.12</td>
</tr>
<tr>
<td>-0.50</td>
<td>1.03</td>
<td>1.06</td>
<td>1.08</td>
<td>1.11</td>
<td>1.14</td>
<td>1.16</td>
</tr>
<tr>
<td>-0.75</td>
<td>1.06</td>
<td>1.11</td>
<td>1.16</td>
<td>1.26</td>
<td>1.34</td>
<td>1.41</td>
</tr>
<tr>
<td>-1</td>
<td>1.12</td>
<td>1.22</td>
<td>1.31</td>
<td>1.50</td>
<td>1.65</td>
<td>1.81</td>
</tr>
<tr>
<td>-2</td>
<td>1.38</td>
<td>1.74</td>
<td>2.13</td>
<td>2.93</td>
<td>3.71</td>
<td>4.51</td>
</tr>
<tr>
<td>Asymmetric spacing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1.36</td>
<td>1.41</td>
<td>1.42</td>
<td>1.43</td>
<td>1.43</td>
<td>1.43</td>
</tr>
<tr>
<td>1.5</td>
<td>1.28</td>
<td>1.31</td>
<td>1.33</td>
<td>1.34</td>
<td>1.34</td>
<td>1.34</td>
</tr>
<tr>
<td>1.2</td>
<td>1.25</td>
<td>1.29</td>
<td>1.31</td>
<td>1.32</td>
<td>1.32</td>
<td>1.32</td>
</tr>
<tr>
<td>0.50</td>
<td>1.49</td>
<td>1.58</td>
<td>1.56</td>
<td>1.56</td>
<td>1.54</td>
<td>1.53</td>
</tr>
<tr>
<td>-0.50</td>
<td>1.47</td>
<td>1.69</td>
<td>1.84</td>
<td>2.08</td>
<td>2.25</td>
<td>2.40</td>
</tr>
<tr>
<td>-1.00</td>
<td>1.62</td>
<td>1.99</td>
<td>2.30</td>
<td>2.83</td>
<td>3.29</td>
<td>3.72</td>
</tr>
<tr>
<td>-2.00</td>
<td>1.98</td>
<td>2.81</td>
<td>3.63</td>
<td>5.24</td>
<td>6.85</td>
<td>8.45</td>
</tr>
</tbody>
</table>

|  

Table 9. $\frac{B_B}{B_0}$

<table>
<thead>
<tr>
<th>$\omega$</th>
<th>$n$</th>
<th>6</th>
<th>12</th>
<th>6</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Symmetric spacing</td>
<td>Asymmetric spacing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.00</td>
<td></td>
<td>1.22</td>
<td>1.28</td>
<td>1.19</td>
<td>1.20</td>
</tr>
<tr>
<td>1.50</td>
<td></td>
<td>1.14</td>
<td>1.18</td>
<td>1.12</td>
<td>1.14</td>
</tr>
<tr>
<td>1.20</td>
<td></td>
<td>1.11</td>
<td>1.14</td>
<td>1.10</td>
<td>1.18</td>
</tr>
<tr>
<td>1.00</td>
<td></td>
<td>1.10</td>
<td>1.12</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>0.50</td>
<td></td>
<td>1.14</td>
<td>1.15</td>
<td>1.27</td>
<td>1.27</td>
</tr>
<tr>
<td>-0.50</td>
<td></td>
<td>1.25</td>
<td>1.26</td>
<td>1.34</td>
<td>1.61</td>
</tr>
<tr>
<td>-1.00</td>
<td></td>
<td>1.21</td>
<td>1.38</td>
<td>1.53</td>
<td>2.12</td>
</tr>
<tr>
<td>-2.00</td>
<td></td>
<td>1.38</td>
<td>2.13</td>
<td>2.03</td>
<td>3.66</td>
</tr>
<tr>
<td>( \omega )</td>
<td>( n )</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>2.0</td>
<td>1.046 to 1.047</td>
<td>1.057</td>
<td>1.061</td>
<td>1.064</td>
<td>1.065</td>
</tr>
<tr>
<td>1.5</td>
<td>1.019 to 1.021</td>
<td>1.020</td>
<td>1.020</td>
<td>1.020</td>
<td>1.020</td>
</tr>
<tr>
<td>1.2</td>
<td>1.003 to 1.004</td>
<td>1.004</td>
<td>1.004</td>
<td>1.004</td>
<td>1.004</td>
</tr>
<tr>
<td>1.0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0.75</td>
<td>1.005 to 1.009</td>
<td>1.008 to 1.009</td>
<td>1.008</td>
<td>1.008</td>
<td>1.008</td>
</tr>
<tr>
<td>0.5</td>
<td>1.039 to 1.043</td>
<td>1.042</td>
<td>1.044</td>
<td>1.039</td>
<td>1.042</td>
</tr>
<tr>
<td>-0.5</td>
<td>1.092</td>
<td>1.083</td>
<td>1.075</td>
<td>1.065</td>
<td>1.063</td>
</tr>
<tr>
<td>-1</td>
<td>1.018 to 1.021</td>
<td>1.020</td>
<td>1.035</td>
<td>1.086</td>
<td>1.139</td>
</tr>
<tr>
<td>-2</td>
<td>1.006 to 1.007</td>
<td>1.095 to 1.096</td>
<td>1.212</td>
<td>1.473</td>
<td>1.745</td>
</tr>
</tbody>
</table>

**Symmetric spacing**
Table 10. (Continued)

<table>
<thead>
<tr>
<th>$\omega$</th>
<th>$n$</th>
<th>4</th>
<th>6</th>
<th>8</th>
<th>12</th>
<th>16</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asymmetric spacing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.0</td>
<td>1.089</td>
<td>1.085</td>
<td>1.082</td>
<td>1.077</td>
<td>1.075</td>
<td>1.073</td>
<td></td>
</tr>
<tr>
<td></td>
<td>to 1.090</td>
<td>to 1.086</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td>1.032</td>
<td>1.030</td>
<td>1.027</td>
<td>1.025</td>
<td>1.024</td>
<td>1.023</td>
<td></td>
</tr>
<tr>
<td>1.2</td>
<td>1.005</td>
<td>1.006</td>
<td>1.005</td>
<td>1.005</td>
<td>1.005</td>
<td>1.005</td>
<td></td>
</tr>
<tr>
<td></td>
<td>to 1.007</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>0.75</td>
<td>1.011</td>
<td>1.016</td>
<td>1.015</td>
<td>1.013</td>
<td>1.012</td>
<td>1.011</td>
<td></td>
</tr>
<tr>
<td></td>
<td>to 1.016</td>
<td>to 1.017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>1.083</td>
<td>1.105</td>
<td>1.095</td>
<td>1.087</td>
<td>1.079</td>
<td>1.073</td>
<td></td>
</tr>
<tr>
<td></td>
<td>to 1.087</td>
<td>to 1.106</td>
<td>to 1.096</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.5</td>
<td>1.096</td>
<td>1.166</td>
<td>1.225</td>
<td>1.318</td>
<td>1.393</td>
<td>1.455</td>
<td></td>
</tr>
<tr>
<td></td>
<td>to 1.098</td>
<td>to 1.169</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-1</td>
<td>1.116</td>
<td>1.297</td>
<td>1.418</td>
<td>1.633</td>
<td>1.822</td>
<td>1.997</td>
<td></td>
</tr>
<tr>
<td>-2</td>
<td>1.314</td>
<td>1.621</td>
<td>1.922</td>
<td>2.512</td>
<td>3.092</td>
<td>3.666</td>
<td></td>
</tr>
<tr>
<td>$\omega$</td>
<td>$n$</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>12</td>
<td>16</td>
<td>20</td>
</tr>
<tr>
<td>--------</td>
<td>-----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>1.65 to 1.66</td>
<td>1.73</td>
<td>1.76</td>
<td>1.78</td>
<td>1.79</td>
<td>1.79</td>
</tr>
<tr>
<td>1.5</td>
<td></td>
<td>1.47 to 1.50</td>
<td>1.52</td>
<td>1.54</td>
<td>1.55</td>
<td>1.56</td>
<td>1.56</td>
</tr>
<tr>
<td>1.2</td>
<td></td>
<td>1.34 to 1.38</td>
<td>1.39</td>
<td>1.41</td>
<td>1.42</td>
<td>1.42</td>
<td>1.42</td>
</tr>
<tr>
<td>1.0</td>
<td></td>
<td>1.29 to 1.32</td>
<td>1.30</td>
<td>1.32</td>
<td>1.32</td>
<td>1.33</td>
<td>1.33</td>
</tr>
<tr>
<td>0.75</td>
<td></td>
<td>1.16 to 1.18</td>
<td>1.20</td>
<td>1.20</td>
<td>1.21</td>
<td>1.22</td>
<td>1.22</td>
</tr>
<tr>
<td>0.50</td>
<td></td>
<td>1.07 to 1.08</td>
<td>1.09</td>
<td>1.10</td>
<td>1.11</td>
<td>1.12</td>
<td>1.12</td>
</tr>
<tr>
<td>-0.50</td>
<td></td>
<td>1.03</td>
<td>1.06</td>
<td>1.08</td>
<td>1.11</td>
<td>1.14</td>
<td>1.16</td>
</tr>
<tr>
<td>-1</td>
<td></td>
<td>1.11 to 1.12</td>
<td>1.22</td>
<td>1.31</td>
<td>1.50</td>
<td>1.65</td>
<td>1.81</td>
</tr>
<tr>
<td>-2</td>
<td></td>
<td>1.38 to 1.41</td>
<td>1.74</td>
<td>2.13</td>
<td>2.93</td>
<td>3.71</td>
<td>4.51</td>
</tr>
</tbody>
</table>

**Symmetric spacing**
<table>
<thead>
<tr>
<th>( \omega )</th>
<th>( n )</th>
<th>4</th>
<th>6</th>
<th>8</th>
<th>12</th>
<th>16</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.50</td>
<td>1.27</td>
<td>1.31</td>
<td>1.33</td>
<td>1.34</td>
<td>1.34</td>
<td>1.34</td>
<td></td>
</tr>
<tr>
<td>1.20</td>
<td>1.25</td>
<td>1.29</td>
<td>1.31</td>
<td>1.32</td>
<td>1.32</td>
<td>1.32</td>
<td></td>
</tr>
<tr>
<td>0.50</td>
<td>1.49</td>
<td>1.58</td>
<td>1.58</td>
<td>1.56</td>
<td>1.54</td>
<td>1.53</td>
<td></td>
</tr>
<tr>
<td>-0.50</td>
<td>1.47</td>
<td>1.69</td>
<td>1.84</td>
<td>2.08</td>
<td>2.25</td>
<td>2.40</td>
<td></td>
</tr>
<tr>
<td>-1.00</td>
<td>1.62</td>
<td>1.99</td>
<td>2.30</td>
<td>2.83</td>
<td>3.29</td>
<td>3.72</td>
<td></td>
</tr>
<tr>
<td>-2.00</td>
<td>1.98</td>
<td>2.81</td>
<td>3.63</td>
<td>5.24</td>
<td>6.85</td>
<td>8.45</td>
<td></td>
</tr>
</tbody>
</table>

Asymmetric spacing
<table>
<thead>
<tr>
<th>$\omega\ n$</th>
<th>6</th>
<th>12</th>
<th>6</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Symmetric spacing</strong></td>
<td></td>
<td><strong>Asymmetric spacing</strong></td>
<td></td>
</tr>
<tr>
<td>2.00</td>
<td>1.22</td>
<td>1.28</td>
<td>1.19</td>
<td>1.20</td>
</tr>
<tr>
<td>1.50</td>
<td>1.14</td>
<td>1.18</td>
<td>1.12</td>
<td>1.14</td>
</tr>
<tr>
<td>1.00</td>
<td>1.10</td>
<td>1.12</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>0.50</td>
<td>1.14</td>
<td>1.15</td>
<td>1.27</td>
<td>1.27</td>
</tr>
<tr>
<td>-0.50</td>
<td>1.25</td>
<td>1.26</td>
<td>1.34</td>
<td>1.61</td>
</tr>
<tr>
<td>-2.00</td>
<td>1.38</td>
<td>2.13</td>
<td>2.03</td>
<td>3.66</td>
</tr>
</tbody>
</table>

4.4. Least-Squares-Type Estimator

Section 4.4 deals with $b_Q = \frac{\Sigma(x_i - \bar{x})y_i}{\Sigma(x_i - \bar{x})x_i}$, which is the estimator one would obtain if he minimized the sum of squares of deviations in the vertical direction, simply ignoring the fact that the $x_i$ are random variables. It is well-known that $b_Q$ is not a consistent estimator of $\beta$; however, it is conceivable that it might have desirable small-sample properties. We shall now investigate the bias and mean square error of $b_Q$. 

Table 12. $\frac{M_B}{M_0}$
4.4.1. Bias of $b_Q$

Denote the conditional expectation of $b_L$ when all the $x_i$ are fixed by $E(b_Q|x)$. Then

$$E(b_Q|x) = \beta \frac{\Sigma(X_i-\bar{X})x_i}{\Sigma(x_i-\bar{X})x_i}$$

$$= \beta \frac{E_{XX} + \Sigma(X_i-\bar{X})e_i}{E_{XX} + 2\Sigma(X_i-\bar{X})e_i + \Sigma(e_i-e)^2}$$

$$= \beta \left[1 + \frac{\Sigma(X_i-\bar{X})e_i}{E_{XX}}\right]^{-1} \left[1 + \frac{2\Sigma(X_i-\bar{X})e_i + \Sigma(e_i-e)^2}{E_{XX}}\right]^{-1}.$$

Now

$$\left[1 + \frac{2\Sigma(X_i-\bar{X})e_i + \Sigma(e_i-e)^2}{E_{XX}}\right]^{-1} = \sum_{k=0}^{5} (-1)^k E_{XX}^{k} \left[2\Sigma(X_i-\bar{X})e_i + \Sigma(e_i-e)^2\right]^k + R,$$

where

$$R = \left[1 + \frac{2\Sigma(X_i-\bar{X})e_i + \Sigma(e_i-e)^2}{E_{XX}}\right]^{-1} \left[2\Sigma(X_i-\bar{X})e_i + \Sigma(e_i-e)^2\right]^6 E_{XX}^{-6}.$$

Consequently,
\[
E(b^*_Q) = \beta \left[ 1 - \frac{(n-3)\mu^2}{E^{XX}} + \frac{H_2^2}{E^{XX}} - \frac{8(GH)_{11}}{E^{XX}} - \frac{H_3^3}{E^{XX}} + \frac{18(GH)_{12}}{E^{XX}} \right.
\]
\[
+ \frac{8G_2^2}{E^{XX}} + \frac{H_4^4}{E^{XX}} - \frac{48(GH)_{21}}{E^{XX}} - \frac{32(GH)_{13}}{E^{XX}} - \frac{H_5^5}{E^{XX}} - \frac{32G_3^3}{E^{XX}}
\]
\[
- \frac{80(GH)_{22}}{E^{XX}} - \frac{10(GH)_{14}}{E^{XX}} \right] + \beta E \left[ 1 + \frac{\Sigma(X_i - \bar{X})e_i}{E^{XX}} \right] R.
\]

Now \[
E \left[ 1 + \frac{\Sigma(X_i - \bar{X})e_i}{E^{XX}} \right] \left[ \frac{2\Sigma(X_i - \bar{X})e_i + \Sigma(e_i - \bar{e})^2}{E^{XX}} \right]^t >
\]
\[
\left[ 1 - \frac{\frac{1}{2}c\Sigma|X_i - \bar{X}|}{E^{XX}} \right] E \left[ \frac{2\Sigma(X_i - \bar{X})e_i + \Sigma(e_i - \bar{e})^2}{E^{XX}} \right]^t.
\]

The right side of the inequality is clearly positive, which implies that \(E(b^*_Q)\) is, before rearrangement in the form given above, an alternating series of monotonically-decreasing terms. Consequently,

\[
0 < \beta E \left[ 1 + \frac{\Sigma(X_i - \bar{X})e_i}{E^{XX}} \right] R < \beta E \left[ 1 + \frac{\Sigma(X_i - \bar{X})e_i}{E^{XX}} \right] \left[ 2\Sigma(X_i - \bar{X})e_i + \Sigma(e_i - \bar{e})^2 \right]^t E^{-6}.
\]

The bound given by the right side of the inequality is not excessively laborious to evaluate, but it can be simplified at only a small sacrifice in precision by using \(1 + \frac{\frac{1}{2}c\Sigma|X_i - \bar{X}|}{E^{XX}}\).
in place of $1 + \frac{\Sigma(X_i - \bar{X})e_i}{E_{XX}}$. We should then have

$$0 < \beta \left[ 1 + \frac{\Sigma(X_i - \bar{X})e_i}{E_{XX}} \right] R < \beta \left[ 1 + \frac{1/2 \Sigma |X_i - \bar{X}|}{E_{XX}} \right] \left( 2 \Sigma (X_i - \bar{X})e_i \right.

$$

$$\left. + \Sigma (e_i - e)^2 \right] ^6 \frac{E^{-6}}{E_{XX}},$$

where $E[2 \Sigma(X_i - \bar{X})e_i + \Sigma(e_i - e)^2] ^6 = 64G_3 + 24O(GH)_{22} + 60(GH)_{14} + H_6$.

Thus, when computing $E(b_Q)$, if we ignore the term

$$\beta \left[ 1 + \frac{\Sigma(X_i - \bar{X})e_i}{E_{XX}} \right] R,$$

we incur an error $\Delta$, where

$$0 < \Delta < \beta \left[ 1 + \frac{1/2 \Sigma |X_i - \bar{X}|}{E_{XX}} \right] (64G_3 + 24O(GH)_{22} + 60(GH)_{14} + H_6).$$

It was found that $\Delta$ was tolerably small for all values of $n$ except four. To investigate the bias for $n = 4$ would have entailed a very great increase in the amount of labor, an increase which did not seem justifiable.

Table 13 cites values of $\frac{B_Q}{B_0}$ correct to within one unit in the last place given there. It will be observed that $\frac{B_Q}{B_0}$ is approximately equal to $-(n-3)$. It is remarkable that the magnitude of this ratio increases greatly with the sample size.
4.4.2 Mean square error of $b_Q$

The mean square error $M_Q$ of $b_Q$ can be obtained in the same way as the bias. Thus

$$E(b_Q^2|x) = \beta^2 \left[ 1 + \frac{\Sigma(X_i - \bar{X})e_i}{E_{XX}} \right] \left[ 1 + \frac{2\Sigma(X_i - \bar{X})e_i + \Sigma(e_i - \bar{e})^2}{E_{XX}} \right]^{-2}$$

$$+ \sigma_F^2 \left[ 1 + \frac{2\Sigma(X_i - \bar{X})e_i + \Sigma(e_i - \bar{e})^2}{E_{XX}} \right]^{-1}.$$
\[ E(b^2_Q) = \beta^2 E \left[ 1 + \frac{\Sigma(x_1^2 - \overline{x})e_i}{E_{XX}} \right]^2 \left[ 1 + \frac{2\Sigma(x_1 - \overline{x})e_i + \Sigma(e_i - e)^2}{E_{XX}} \right]^{-2} \]

\[ + \sigma^2_{E_{XX}} \left[ 1 + \frac{2\Sigma(x_1 - \overline{x})e_i + \Sigma(e_i - e)^2}{E_{XX}} \right]^{-1} \]

\[ M_L = E(b^2_Q) - 2\beta E(b^2_Q) + \beta^2 \]

\[ = \beta^2 \left[ \frac{\mu_2}{E_{XX}} + \frac{H_2}{E_{XX}} - \frac{10(GH)}{E_{XX}} \frac{11}{1} - \frac{2H_3}{E_{XX}} + \frac{39(GH)}{E_{XX}} \frac{12}{1} + \frac{12G_2}{E_{XX}} + \frac{3H_4}{E_{XX}} \right. \]

\[ - \frac{112(GH)}{E_{XX}} \frac{21}{1} - \frac{100(GH)}{E_{XX}} \frac{13}{1} - \frac{4H_5}{E_{XX}} - \frac{240G_3}{E_{XX}} - \frac{680(GH)}{E_{XX}} \frac{22}{1} \]

\[ - \frac{95(GH)}{E_{XX}} \frac{14}{1} - \frac{480(GH)}{E_{XX}} \frac{31}{1} - \frac{240(GH)}{E_{XX}} \frac{23}{1} - \frac{6(GH)}{E_{XX}} \frac{15}{1} \]

\[ + \frac{\lambda\mu_2}{E_{XX}} \left[ 1 - \frac{(n-5)\mu_2}{E_{XX}} + \frac{H_2}{E_{XX}} - \frac{12(GH)}{E_{XX}} \frac{11}{1} - \frac{H_3}{E_{XX}} + \frac{24(GH)}{E_{XX}} \frac{12}{1} \right. \]

\[ + \frac{16G_2}{E_{XX}} + \frac{H_4}{E_{XX}} - \frac{80(GH)}{E_{XX}} \frac{21}{1} - \frac{40(GH)}{E_{XX}} \frac{13}{1} - \frac{H_5}{E_{XX}} \]

\[ + \frac{\lambda\mu_2}{E_{XX}} \right] + \Delta , \]

where

\[ 0 < \Delta < \left\{ 7\beta^2 \left[ 1 + \frac{\frac{1}{2}\Sigma|x_1^2 - \overline{x}|}{E_{XX}} \right]^2 - 2\beta^2 \frac{1 - \frac{1}{2}c\Sigma|x_1^2 - \overline{x}|}{E_{XX}} - 2\beta^2 \frac{c\Sigma|x_1^2 - \overline{x}| + nc^2}{E_{XX}} \right. \]

\[ + \frac{\lambda\mu_2}{E_{XX}} \left. \right\} \left[ 64G_3 + 240(GH)_{22} + 60(GH)_{14} + H_6 \right] . \]
We found $\Delta$ to be tolerably small for $n > 6$ except in the case where $n = 6$ and the $X_i$ are uniformly spaced. Rather than include additional terms in the expansion of $\frac{M_Q}{M_0}$, at the expense of very considerable additional labor, we have omitted this particular situation.

$\frac{M_Q}{M_0}$ is a monotonic function of $\frac{\lambda}{\beta^2}$, just as $\frac{M_H}{M_0}$ is, and its analysis has been carried out in the same way; cf. section 4.3.4. It will be observed from Table 14 that $\frac{M_Q}{M_0}$ is very nearly unity.

<table>
<thead>
<tr>
<th>$\omega$</th>
<th>$n$</th>
<th>6</th>
<th>8</th>
<th>12</th>
<th>16</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symmetric spacing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.99</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>a</td>
<td>.96</td>
<td>to</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>to</td>
<td>1.02</td>
<td>1.02</td>
<td>1.03</td>
<td></td>
</tr>
<tr>
<td>1/2</td>
<td>.99</td>
<td>.98</td>
<td>to</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>to</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>-1/2</td>
<td>.99</td>
<td>.99</td>
<td>to</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>to</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>-2</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

*We do not yet have a reliable value to cite here.*
Table 14. (Continued)

<table>
<thead>
<tr>
<th>$\omega$</th>
<th>n</th>
<th>6</th>
<th>8</th>
<th>12</th>
<th>16</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Asymmetric spacing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>1.2</td>
<td>to</td>
<td>.96</td>
<td>.98</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>to</td>
<td>.99</td>
<td>.99</td>
<td>1.00</td>
<td>1.01</td>
<td>1.01</td>
</tr>
<tr>
<td>1/2</td>
<td>to</td>
<td>.98</td>
<td>.99</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>to</td>
<td>.99</td>
<td>1.00</td>
<td>1.00</td>
<td>1.01</td>
<td>1.01</td>
</tr>
<tr>
<td>-1/2</td>
<td>to</td>
<td>.99</td>
<td>.99</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>-2</td>
<td></td>
<td>.99</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

4.5. Discussion of Results

4.5.1. Comparison of estimators

We should like now to compare the estimators discussed in section 4 with a view toward making some practical suggestions for workers who are actually concerned with fitting straight-line relations.

Generally speaking $b_B$ is clearly superior to $b_W$ but there are some situations when such is not the case; we found that $b_W$ had lower bias and lower mean square error than
b_B when the X_1 are symmetrically spaced, with \( \lambda = 0.5 \) or \( \lambda = -0.50 \). However, in almost every case investigated, \( b_H \) proved superior to \( b_W \) and to \( b_B \). This, of course, is hardly surprising. The use of \( b_H \) presupposes through the ability to order all the observations more information available to the experimenter than that of \( b_W \) or \( b_B \), and estimators based upon more complete information are typically more efficient than those based upon partial information. We did, nevertheless, find two instances where \( b_W \) proved superior to \( b_B \) and \( b_H \).

These were firstly \( n = 4, \omega = -0.50 \), symmetric spacing, and secondly, \( n = 6, \omega = -0.50 \), symmetric spacing. For asymmetric spacings corresponding to these values of \( n \) and \( \omega \) \( b_H \) was found decidedly superior to the others.

As the comparison of \( b_H \) and \( b_Q \) is slightly more involved, we should like to first compare \( b_0 \) with \( b_Q \). The conclusion seems inescapable that when \( b_0 \) can be obtained it is preferable to \( b_Q \) for two reasons:

1) \( b_0 \) has much smaller bias than \( b_Q \), while its variance is essentially no greater,

2) \( b_0 \) is a consistent estimator of \( \beta \), whereas \( b_Q \) is not.

The difficulty, of course, is that \( b_0 \) cannot usually be obtained, inasmuch as the spacing of the \( X_1 \) is unknown. (An example where \( b_0 \) can be obtained is given in section 4.5.2.)

Now, what is the lesson conveyed by Tables 7-14? Surely, it is that \( b_H \) is a reasonable substitute for \( b_0 \) whenever we know that the \( X_1 \) possess no marked skewness and are not
bunched excessively. When such is the case we see that
neither \( \frac{B_H}{B_0} \) nor \( \frac{M_H}{M_Q} \) greatly exceed unity; this means that \( \frac{M_H}{M_Q} \)
will not greatly exceed unity, while \( \frac{B_H}{B_Q} \) will be approxi-
mately \(- \frac{1}{n^2} \). Thus, when using \( b_H \) in preference to \( b_Q \), one
stands to do much better with respect to bias and very little
worse with respect to mean square error.

One could roughly assess the skewness or bunchiness of
his data by plotting the \( x_i \) versus \( i \) on log-log paper; the
slope of the line which best fits the points is then \( \omega \).
Ordinarily this should not be necessary; in most cases one
could decide whether there was evidence of skewness or
bunchiness by merely looking at the data.

4.5.2. An example

An interesting example which illustrates the application
of the ideas under discussion is furnished by an actual ex-
periment undertaken for the purpose of calibrating a Baldwin
SR-4 strain gage. The gage consists essentially of a wire
which changes its electrical resistance when it is put under
strain. We wanted to estimate the gage factor, which is the
ratio of unit change in resistance to unit strain.

The gage was placed at the fixed end of a cantilever
beam. The beam was then deflected through various distances
by means of weights placed on the free end; the deflection of the free end and the change in resistance were observed. The data is reproduced below.

<table>
<thead>
<tr>
<th>End deflection</th>
<th>Change of resistance</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 inches</td>
<td>0 ohms</td>
</tr>
<tr>
<td>.107</td>
<td>.10</td>
</tr>
<tr>
<td>.207</td>
<td>.20</td>
</tr>
<tr>
<td>.307</td>
<td>.23</td>
</tr>
<tr>
<td>.397</td>
<td>.39</td>
</tr>
</tbody>
</table>

There are theoretical reasons for believing the relation between the change of resistance and the end deflection to be linear, and a plot of the data, shown in Figure 2, supports this belief. We desired to estimate the slope of the linear relation. We obtained

$$b_H = 0.915$$

$$b_Q = 0.912$$

In this experiment we did not, of course, know the true deflections; however, the deflections were produced by equal load increments on the beam, and therefore the true deflections are equally spaced. In this case then $b_H$ is, in fact, the optimal estimator. Had the load increments been unequal but known, it would have been easy enough to compute the optimal estimator. Of course we were unusually fortunate in having a very good instrumental variable, but even when no such instrumental variable is available, it may well be
Figure 2. Calibration curve of SR-\(\frac{1}{4}\) strain gage
CHANGE OF RESISTANCE IN OHMS

DEFLECTION IN INCHES

b_H = 0.915
b_Q = 0.912
possible to improve upon $b_H$ if one has even a crude idea of the spacing of the $X_i$ obtained independently of the observations.

4.5.3. An extension of $b_H$

It sometimes happens that one can foresee a possible mis-ordering of two or more of the $X_i$ when they are very close together. For example, one might be calibrating two soil-bearing gages with the aid of seven different types of soil; he might know that two of the soils were so similar that in ordering the observations he might easily mis-order the two $X_i$ corresponding to these soils. For concreteness let us suppose that it is $X_3$ and $X_4$ that are causing the difficulty.

Under these circumstances a good choice of the weights would be achieved by treating $x_3$ and $x_4$ as a unit and assigning the Housner-Brennan weights to the six units. Each weight is then divided equally among the observations comprising the unit. Thus,

$$w_1 = -5, \quad w_2 = -3, \quad w_3 = -1/2, \quad w_4 = -1/2, \quad w_5 = 1, \quad w_6 = 3, \quad w_7 = 5.$$  

It is obvious how the scheme may be extended when there are any number of unorderable observations. In fact the estimators of Wald and Bartlett are simply special cases of this kind of estimator.
5. SMALL-SAMPLE RESULTS FOR REPLICATED OBSERVATIONS

5.1. Introduction and Assumptions

We shall conclude this dissertation with an investigation of the bias and the mean square error of the estimators proposed in section 2. Regrettably, we shall have to omit \( b_4 \); at the time of this writing we have been unable to find any suitably simple technique for assessing the bias or mean square error of this estimator, which is much more complicated than the others.

We shall employ the same notation as in section 2.1 but shall make the stronger distributional assumptions of section 4.1. To be specific we shall assume

1) \( e_{it} \) and \( e_{i't'} \) are independent unless \( i = i' \) and \( t = t' \).

2) \( f_{it} \) and \( f_{i't'} \) are independent unless \( i = i' \) and \( t = t' \).

3) \( e_{it} \) and \( f_{i't'} \) are independent for all \( i, i', t, \) and \( t' \).

4) \( E(f_{it}) = 0 \) and \( E(f_{it}^2) = \nu_2 = \sigma^2_f \).

5) Let \( c = \min_i |X_{i+1} - X_i| \). Then \( \text{Prob} \{ |e_{it}| \geq \frac{c}{2} \} = 0 \).

6) \( E(e_{it}^2) = \mu_2 \) \( E(e_{it}^4) = \mu_4 \) \( E(e_{it}^6) = \mu_6 \).

All odd moments of the \( e_{it} \) are zero.
7) \( s_i = r_i = r \), a constant independent of \( i \).

We introduce this last restriction in order to simplify the algebra.

8) The denominators of the statistics used have non-vanishing expectations.

5.2. Expectations of Various Statistics Used in Section 5

In order to discuss the bias and mean square error of the estimators proposed in section 2, various additional expectations must be considered. Some of these can be obtained from earlier results by making minor changes where necessary; others are unlike those already obtained and must be worked out in their entirety. We shall make use of the conventions and symbols already introduced, augmenting the latter as follows:

Put \( E_{XY} = \Sigma(X_i - \bar{X})(Y_i - \bar{Y}) \) and \( E_{YY} = \Sigma(Y_i - \bar{Y})^2 \)

Let \( \overline{G}_k = E[\Sigma(X_i - \bar{X})e_i]^{2k} \)
\( \overline{H}_k = E[\Sigma(e_i - e_..)^2]^k \)
\( \overline{I}_k = E[\Sigma(e_{it} - e_i..)^2]^k \)
\( \overline{J}_k = E[\Sigma(e_{it} - e_.)]^2]^k \)

\( \overline{GHIJ}_{k_1 k_2 k_3 k_4} = E[\Sigma(X_i - \bar{X})e_i]^{2k_1}[\Sigma(e_i - e_..)]^{2k_2}[\Sigma(e_{i2} - e_..)]^{2k_3}[\Sigma(e_{i3} - e_.)]^{2k_4} \).
Now $G_k$, $H_k$, and $(GH)_{kk'}$ can be easily obtained from $G_k$, $H_k$, and $(GH)_{kk'}$, respectively, by simply substituting $\mu_k$ for $\mu_k$ in the expressions for the latter given in section 4.2. Here, of course,

$$\mu_k = \mathbb{E}(e_i^k)$$

$$= r^{-k} \mathbb{E}(\sum_i e_i^k).$$

In particular,

$$\mu_2 = \frac{\mu_2}{r}$$

$$\mu_4 = r^{-4}[\mu_4 + 3(r-1)\mu_2^2]$$

$$\mu_6 = r^{-6}[\mu_6 + 15(r-1)\mu_4\mu_2 + 15(r-1)(r-2)\mu_2^3].$$

The $(GJ)_{kk'}$ can be obtained from the expressions for the $(GH)_{kk'}$ by a simple artifice, which is illustrated in section 5.2.2. The $(IJ)_{kk'}$ and $(GI)_{kk'}$ are the main source of difficulty and have been handled by whatever technique appeared easiest. Expectations which involve only the substitution of $\mu_k$ for $\mu_k$ in formulas already obtained will not be set down explicitly.

5.2.1. \[ I_2 = \mathbb{E}[(\sum_i e_i^k)^2] \]

From section 2.2.2, it follows directly that

$$I_2 = \frac{n(r-1)^2}{r} \mu_4 + \frac{n(r-1)(3-r)}{r} \mu_2^2 + \frac{n^2(r-1)^2}{r} \mu_2^2.$$
5.2.2. \((\bar{G}J)_{11} = E[\Sigma(X_1 - \bar{X})e_1]\Sigma^2(e_{i't' - e_1})^2\)

\[(\bar{G}J)_{11} = \frac{1}{r^2} E[\Sigma(X_{it} - \bar{X})e_{it}]^2 \Sigma^2(e_{i't' - e_1})^2\]

where \(X_{it} = X_i\) for every \(t\).

Now let us renumber the \(e_{it}\) in the following way:

\[e_{1,1} = e_1\]
\[\vdots\]
\[e_{1,r} = e_r\]
\[e_{2,1} = e_{r+1}\]
\[\vdots\]
\[e_{2,r} = e_{2r}\]
\[\vdots\]
\[e_{n,r} = e_{nr}\]

Further, let us put

\[X_{1,1} = Z_1\]
\[\vdots\]
\[X_{1,r} = Z_r\]
\[\vdots\]
\[X_{n,r} = Z_{nr}\]

Then \((\bar{G}J)_{11} = \frac{1}{r^2} E[\Sigma(Z_1 - \bar{Z})e_1]^2 \Sigma^2(e_{i't' - e_1})^2\)
where now \( i = 1, 2, \ldots, nr \)
\[ i' = 1, 2, \ldots, nr . \]

But this has already been worked out in section 4.2.5. We obtain
\[
(GJ)_{ll} = \left( \frac{nr-1}{nr} \mu_4 + \frac{n^2r^2 - 2nr + 3}{nr} \mu_2^2 \right) \frac{E_{XX}}{r} .
\]

5.2.3. \( (\bar{G}J)_{ll} = E[\Sigma(X_1 - \bar{X})e_{i1}]^2 \Sigma \Sigma(e_{i'1} - e_{i1})^2 \)

Since \( \Sigma \Sigma(e_{i'1} - e_{i1})^2 = \Sigma \Sigma(e_{i1} - e_{..})^2 - r \Sigma(e_{i1} - e_{..})^2 \),
\[
(\bar{G}I)_{ll} = (\bar{G}J)_{ll} - r(\bar{G}H)_{ll} .
\]

5.2.4. \( (\bar{H}I)_{ll} = E\Sigma(e_{i1} - e_{..})^2 \Sigma \Sigma(e_{i'1} - e_{i1})^2 \)

This expectation has already been obtained in section 2.2.4. We found that
\[
(\bar{H}I)_{ll} = \frac{(n-1)(r-1)}{r^2} \mu_4 + \frac{n-1}{r} (rn - 1 - \frac{rn+2r-3}{r} \mu_2^2 .
\]

5.2.5. \( I_3 = E[\Sigma e_{i1} - e_{i1})^2]^3 \)

The treatment here follows that of section 4.2.10 and employs the same notation. We first obtain the third cumulant of \( \Sigma e_{i1} - e_{i1})^2 = N \), say, which we denote by \( \kappa_3(N) \). Thus
\[ K_3(N) = n K_3(\Sigma (e_{it} - e_i)^2) \]

\[ = n(r-1)^3 K(3) \]

\[ = n(r-1)^3 \frac{K_6}{r^2} + \frac{12 K_4 K_2}{r(r-1)} + \frac{8}{(r-1)^2} K^3_2 \]

\[ = n(r-1)^3 \left[ \frac{\mu_6 - 15\mu_4\mu_2 + 30\mu_2^3}{r^2} + \frac{12}{r(r-1)} \left( \mu_4 - 3\mu_2^2 \right) \mu_2 + \frac{8}{(r-1)^2} \mu_2^3 \right] . \]

Since \( I_3 = K_3(N) + 3 K_2(N) K_1(N) + [K_1(N)]^3 \),

we obtain

\[ I_3 = n(r-1)^3 \left[ \frac{\mu_6 - 15\mu_4\mu_2 + 30\mu_2^3}{r^2} + \frac{12}{r(r-1)} \left( \mu_4 - 3\mu_2^2 \right) \right. \]

\[ + \frac{8}{(r-1)^2} \mu_2^3 \right] + 3n(r-1)\mu_2 \left[ \frac{n(r-1)^2}{r} \mu_4 \right. \]

\[ - n(r-1)(r-3) \frac{\mu_2^3}{r} \] \[ + n^3(r-1)^3 \mu_2^3 . \]

5.2.6. \( (\overline{II})_{21} = \mathbb{E} [\Sigma (e_{i1} - e_i)^2] \Sigma (e_{it} - e_i)^2 \)

\[ (\overline{II})_{21} = r\mathbb{E} [\Sigma e_{i1}^2 - ne_{i1}^2]^2 [2e_{i1}^2 - 2\Sigma e_{i1}^2 + e_{i1}^2] . \]
\[ E(\Sigma e^2_{i*})e^2_{it} = E\Sigma e^4_{i*}. \Sigma e^2_{it} + E\Sigma e^2_{i*}. e^2_j. \Sigma e^2_{it} \]

\[ = \frac{n}{r^4} \left[ \mu_6 + 7(r-1)\mu_4\mu_2 + e(r-1)(r-2)\mu_2^3 \right] \]

\[ + \frac{n(n-1)\mu_2^2}{r^2} + \frac{2n(n-1)}{r^3} [\mu_4 + (r-1)\mu_2^2]\mu_2 \]

\[ + \frac{n(n-1)(n-2)}{r^2} \mu_2^3. \]

\[ E(\Sigma e^2_{i*}). \Sigma e^2_{it} e_{i*} = \frac{1}{r} E(\Sigma e^2_{i*}). \Sigma e^2_{it} \]

\[ E(\Sigma e^2_{i*}). \Sigma e^2_{it} = \frac{1}{r} E(\Sigma e^2_{i*}). \Sigma e^2_{it} \]

\[ E e^2. \Sigma e^2_{i*}. \Sigma e^2_{it} = \frac{1}{n^2} E(\Sigma e^2_{i*}). \Sigma e^2_{it} \]

\[ E e^2. \Sigma e^2_{i*}. \Sigma e^2_{it} = \frac{1}{r} E e^2. \Sigma e^2_{i*}. \Sigma e^2_{it} \]

\[ E e^2. \Sigma e^2_{i*}. \Sigma e^2_{it} = \frac{1}{r} E e^2. \Sigma e^2_{i*}. \Sigma e^2_{it} \]

\[ E e^4. \Sigma e^2_{it} = \frac{1}{n^4} E e^4. \Sigma e^2_{it} + \frac{3}{n^4} E\Sigma e^2_{i*}. e^2_j. \Sigma e^2_{it} \]

\[ = \frac{1}{n^3 r^4} \left[ \mu_6 + 7(r-1)\mu_4\mu_2 + 3(r-1)(r-2)\mu_2^3 \right] + \frac{n-1}{n^3 r^3} \mu_2^2 \]

\[ + \frac{6(n-1)}{n^3 r^3} [\mu_4 + (r-1)\mu_2^2]\mu_2 + \frac{3(n-1)(n-2)}{n^3 r^2} \mu_2^3. \]

\[ E e^4. \Sigma e^2_{it} e_{i*} = \frac{1}{r} E e^4. \Sigma e^2_{it} \]
\[ Ee^{4} \Sigma e_{1.}^{2} = \frac{1}{r} Ee^{4} \sum_{i=1}^{2} e_{it}^{2} \]

Therefore,

\[
(\overline{\Pi})_{21} = \frac{(n-1)^{2}(r-1)}{r^{4}n} [\mu_{6} + 7(r-1)\mu_{4}\mu_{2} + 3(r-1)(r-2)\mu_{3}^2] \\
+ \frac{(n-1)^{3}(r-1)}{n} \mu_{2}m_{4}^{2} + \frac{2(n-1)^{3}(r-1)}{r^{3}n} [\mu_{4} + (r-1)\mu_{2}^2] \\
+ \frac{(n-1)^{3}(n-2)(r-1)}{r^{2}n} \mu_{2}^3 + \frac{4(n-1)(r-1)}{r^{3}n} [\mu_{4} + (r-1)\mu_{2}^2] \\
+ \frac{2(n-1)(n-2)(r-1)}{r^{2}n} \mu_{2}^3.
\]

5.2.7. \[(\overline{\Pi})_{12} = E \Sigma (e_{1..}^{2})^{2} \Sigma (e_{it-e_{i.}}^{2})^{2} \]

\[(\overline{\Pi})_{12} = E \Sigma (e_{1..}^{2})^{2} \Sigma (e_{1..}^{2})^{2} + \Sigma (e_{1..}^{2})^{2} \Sigma (e_{1..}^{2})^{2} \Sigma (e_{1..}^{2})^{2} \Sigma (e_{1..}^{2})^{2} \]

\[
E \Sigma (e_{1..}^{2})^{4} = E \Sigma \Sigma (e_{1..}^{2})^{4} + E \Sigma \Sigma (e_{1..}^{2})^{4} + E \Sigma \Sigma (e_{1..}^{2})^{4} + E \Sigma \Sigma (e_{1..}^{2})^{4}.
\]

Now \[E \Sigma \Sigma (e_{1..}^{2})^{4} = nrE e_{1..}^{2} (e_{1..}^{2})^{4} \]

\[
= nrE \sum (e_{1..}^{2})^{4} (e_{1..}^{2})^{3} + 6e_{1..}^{2} (e_{1..}^{2})^{4} + 4e_{1..}^{2} (e_{1..}^{2})^{4} + 6e_{1..}^{2} (e_{1..}^{2})^{4} \]

\[
= nr \left[ \frac{\mu_{6} + (r-1)\mu_{4}\mu_{2}}{r^{2}} - 4 \frac{\mu_{6} + 3(r-1)\mu_{4}\mu_{2}}{r^{3}} \\
+ 6 \frac{\mu_{6} + 7(r-1)\mu_{4}\mu_{2} + 3(r-1)(r-2)\mu_{3}^{3}}{r^{4}} \right].
\]
\[
\frac{4 \mu_6 + 6(r-1)\mu_4 \mu_2 + 15(r-1)(r-2)\mu_2^3}{r^5} + \bar{\mu}_6
\]

and
\[
E\Sigma_i \Sigma(e_{jt} - e_j)^4 = E\Sigma_i \Sigma e_{jt}^2(e_{jt} - e_j)^4
\]
\[
= n(n-1) r E\Sigma_i e_{jt}^2 \Sigma (e_{jt} - e_j)^4
\]
\[
= n(n-1) r E\Sigma_i 2 e_{jt}^2(e_{jt} - e_j) + 6e_{jt}^2 e_j^2 - 4e_{jt} e_j^3 + e_j^4
\]
\[
= n(n-1) \mu_2 \left[ \mu_4 - 4 \frac{\mu_4}{r} + 6 \frac{\mu_4 + (r-1)\mu_2^2}{r^2}
\right.
\]
\[
- 4 \frac{\mu_4 + 3(r-1)\mu_2^2}{r^3} + \bar{\mu}_4 \right]
\]

Also,
\[
E\Sigma_i \Sigma(e_{it} - e_i)^4 = \frac{1}{n} E\Sigma_i \Sigma(e_{it} - e_i)^4
\]

Turning now to the second term of [5.2.7 - 1], we consider first
\[
E\Sigma_i^2 \Sigma\Sigma(e_{it} - e_i)^2(e_{jt} - e_j)^2 = 2E\Sigma\Sigma\Sigma e_i^2(e_{it} - e_i)^2
\]
\[
(e_{jt} - e_j)^2 + E\Sigma\Sigma\Sigma e_i^2(e_{it} - e_i)^2(e_{jt} - e_j)^2
\]

Now \[
E\Sigma\Sigma\Sigma e_i^2(e_{it} - e_i)^2(e_{jt} - e_j)^2
\]
\[
= n(n-1) r(r-1) \mu_2 \left[ \frac{\mu_4 + (r-1)\mu_2^2}{r^2} - 2 \frac{\mu_4 + 3(r-1)\mu_2^2}{r^3} + \bar{\mu}_4 \right]
\]

Consequently,
\[ + \left[ \frac{\xi^\alpha}{\mathcal{Z}(1-\alpha)\xi + \mathcal{I}} \mathcal{I} + \frac{\xi^\alpha}{\mathcal{Z}(1-\alpha)\xi + \mathcal{I}} \mathcal{I} - \right. \\
\frac{\xi^\alpha}{\mathcal{Z}(1-\alpha)\xi + \mathcal{I}} \mathcal{I} + \frac{\xi^\alpha}{\mathcal{Z}(1-\alpha)\xi + \mathcal{I}} \mathcal{I} - \left. \right] \mathcal{I}(1-u) = \mathcal{Z}(I H) \]

After combining terms we have

\[ = \mathcal{Z}(I H) \]

\[ = \mathcal{Z}(I H) \]

Finally,

\[ \frac{\xi^\alpha}{\mathcal{Z}(1-\alpha)\xi + \mathcal{I}} \mathcal{I} + \left[ \frac{\xi^\alpha}{\mathcal{Z}(1-\alpha)\xi + \mathcal{I}} \mathcal{I} - \right. \\
\frac{\xi^\alpha}{\mathcal{Z}(1-\alpha)\xi + \mathcal{I}} \mathcal{I} - \left. \right] \mathcal{I}(1-u) = \mathcal{Z}(I H) \]

139
\[ 2(n-1)^2 r(n-1) \mu_2 \left[ \frac{\mu_4 + (r-1)\mu_2^2}{r^2} - 2 \frac{\mu_4 + e(r-1)\mu_2^2}{r^3} + \frac{\mu_4}{r} \right] \]

\[ + \frac{(n-1)^2(n-2)(r-1)^2 \mu_3}{r} \cdot \]

5.2.8. \( (\overline{GHI})_{111} = E[\Sigma(X_i - \overline{X})e_{i1}]^2 \Sigma(e_{i1}^2 - e_{i1}) \Sigma(e_{i1} e_{i1}) \]

\( \overline{GHI}_{111} = E[\Sigma(X_i - \overline{X})e_{i1}]^2 [\Sigma e_{i1}^2 - ne_{i1}^2] \Sigma(e_{i1} e_{i1}) \).

Now \( E[\Sigma(X_i - \overline{X})e_{i1}]^2 [\Sigma e_{i1}^2 - ne_{i1}^2] \Sigma(e_{i1} e_{i1})^2 = E\Sigma(X_i - \overline{X})^2 e_{i1}^4 \Sigma(e_{i1} e_{i1})^2 \]

\[ + E\Sigma(X_i - \overline{X})^2 e_{i1}^4 \Sigma(e_{j1} e_{j1})^2 + E\Sigma(X_i - \overline{X})^2 e_{i1}^4 \Sigma(e_{i1} e_{i1})^2 \Sigma e_{j1}^2 \]

\[ + E\Sigma(X_i - \overline{X})^2 e_{i1}^4 \Sigma e_{j1}^2 (e_{j1} e_{j1})^2 \]

\[ + E\Sigma(X_i - \overline{X})^2 e_{i1}^4 \Sigma e_{j1}^2 \Sigma e_{k1}^2 (e_{k1} e_{k1})^2 \cdot [1] \]

\( E\Sigma(X_i - \overline{X})^2 e_{i1}^4 \Sigma e_{i1}^2 = r^{-3} [\mu_4^6 + 7(\mu_4^3 + 3(r-1)(\mu_2^3)] \Sigma e_{i1}^2 \overline{X}_i \cdot \)

\( E\Sigma(X_i - \overline{X})^2 e_{i1}^4 (r e_{i1}^2) = r \mu_4 \Sigma e_{i1}^2 \cdot \)

\( E\Sigma(X_i - \overline{X})^2 e_{i1}^4 \Sigma(e_{j1} e_{j1})^2 = (n-1)(r-1) \mu_4 \Sigma e_{i1}^2 \overline{X}_i \cdot \)

\( E\Sigma(X_i - \overline{X})^2 e_{i1}^4 (e_{i1} e_{i1})^2 \Sigma e_{j1}^2 \cdot \)

\( (n-1) \mu_2 \left[ \frac{\mu_4 + (r-1)\mu_2^2}{r^2} - 2 \frac{\mu_4 + 3(\mu_2^2)}{r^3} + \frac{\mu_4}{r} \right] \Sigma e_{i1}^2 \overline{X}_i \cdot \)
This completes the evaluation of [5.2.8 - 1].

We consider now

$$E[\Sigma(X_i-X)\Sigma(X_j-X)e_i.e_j.]^2 \Sigma \Sigma(e_i'-e_i'')^2 =$$

$$\frac{1}{n} E[\Sigma(X_i-X)e_i.]^2 \Sigma \Sigma(e_i'-e_i'').^2 + \frac{1}{n} E[\Sigma(X_i-X)(X_j-X)e_i.e_j.] \Sigma \Sigma(e_i'-e_i'').^2$$

$$E[\Sigma(X_i-X)(X_j-X)e_i.e_j.] \Sigma \Sigma(e_i'-e_i'').^2 =$$

$$2E \Sigma \Sigma(X_i-X)(X_j-X)e_i.e_j. \Sigma \Sigma(e_i'-e_i'').^2 =$$

$$4E \Sigma \Sigma(X_i-X)(X_j-X)e_i.e_j. \Sigma(e_i'-e_i'').^2 + 2E \Sigma \Sigma(X_i-X)(X_j-X)e_i.e_j. \Sigma(e_i'-e_i'').^2$$

Now \( \Sigma(X_i-X)(X_j-X)e_i.e_j. \Sigma(e_i'-e_i')^2 = \)

$$- \left[ \frac{\mu_i + (r-1)\mu_2^2}{r^2} - 2 \frac{\mu_i + 3(r-1)\mu_2^2}{r^3} + \bar{\mu}_4 \right] \mu_2 E_{XX}$$

and
\[
\Sigma_{i} \left( \Sigma_{j} (X_{i} - \overline{X}) (X_{j} - \overline{X}) e_{i}^{2} e_{j}^{2} \Sigma (e_{i} - e_{j})^{2} \right) = - \frac{(n-2)(r-1)}{r^2} \mu_{2}^{3} \frac{E_{XX}}{r^2}
\]

Combining terms we obtain

\[
(GHI)_{111} = \frac{n-1}{nr^3} \left[ \mu_{6} + 7(r-1)\mu_{4}\mu_{2} + 3(r-1)(r-2)\mu_{2}^{3} \right] E_{XX}
\]

\[
+ \frac{r(n-1)}{n} \bar{\mu}_{6} E_{XX} + \frac{(n-1)^{2}(r-1)}{n} \bar{\mu}_{4}\mu_{2} E_{XX}
\]

\[
+ \frac{(n-1)^{2}}{n} \left[ \frac{\bar{\mu}_{4} + (r-1)\mu_{2}^{2}}{r^2} - 2 \frac{\mu_{4} + 3(r-1)\mu_{2}^{2}}{r^3} + \bar{\mu}_{4} \right].
\]

\[
\left[ \mu_{2} E_{XX} + rG[1] \right]
\]

\[
+ \frac{(n-1)^{2}(n-2)(r-1)}{nr} \mu_{2}^{2} \bar{G}_{1} + \frac{(n-2)(r-1)}{nr^2} \mu_{2}^{3} E_{XX}
\]

\[
+ \frac{1}{n} \left[ \frac{\bar{\mu}_{4} + (r-1)\mu_{2}^{2}}{r^2} - 2 \frac{\mu_{4} + 3(r-1)\mu_{2}^{2}}{r^3} + \bar{\mu}_{4} \right] \mu_{2} E_{XX}.
\]

5.2.9. \( (GJ)_{12} = E[\Sigma_{i} (X_{i} - \overline{X}) e_{i}]^{2} \Sigma (e_{i} - e_{j})^{2} ]^{2} \)

This may be obtained through the same artifice as employed in section 5.2.2. The result is

\[
(GJ)_{12} = [\mu_{6} + 3(nr-1)\mu_{4}\mu_{2} + (nr-1)(nr-2)\mu_{2}^{3}] \frac{E_{XX}}{r}
\]

\[
- \frac{2}{nr} \left[ \mu_{6} + (3nr-7)\mu_{4}\mu_{2} + (nr-2)(nr-3)\mu_{2}^{3} \right] \frac{E_{XX}}{r}
\]

\[
+ \frac{1}{n^2 r^2} \left[ \mu_{6} + (7nr-15)\mu_{4}\mu_{2} + 3(nr-2)(nr-5)\mu_{2}^{3} \right] \frac{E_{XX}}{r}.
\]
5.2.10. \((\bar{G}I)_{12} = E[\Sigma(X_i - \bar{X})e_i]_1^2[\Sigma(e_{i't'} - e_{i't})_1^2]\)

Since \(\Sigma(e_{i't'} - e_{i't})_1^2 = \Sigma(e_{i't'} - e_{i't})_1^2 - r\Sigma(e_{i't'} - e_{i't})_1^2\),

\[(\bar{G}I)_{12} = (\bar{G}J)_{12} - 2r(\bar{G}HI)_{111} - r^2(\bar{G}H)_{12} \cdot \]

5.2.11. \((\bar{G}J)_{21} = E[\Sigma(X_i - \bar{X})e_i]_1^4\Sigma(e_{i't'} - e_{i't})_1^2\)

Using the technique of section 5.2.2 we can write

\[(\bar{G}J) = \frac{1}{r^2} E[\Sigma(Z_i - \bar{Z})e_i]_1^4\Sigma(e_{i't'} - e_{i't})_1^2\]

where now \(i = 1, 2, \ldots, nr\)

\(i' = 1, 2, \ldots, nr\).

From the result of section 4.2.8 we can say immediately

\[(\bar{G}J)_{21} = \frac{1}{r} \left\{ (1 - \frac{1}{nr})\mu_6 + (nr - 1 - \frac{nr - 2}{nr})\mu_2^2 + \frac{12}{nr^2} \mu_2^3 \right\}^{\Sigma(X_i - \bar{X})}_1^4\]

\[
+ \left\{ 6(1 - \frac{1}{nr})\mu_2^2 + e(nr-2)\mu_2^3 - \frac{3(nr+2)}{nr} \mu_2^3 \right\}^{1} \cdot \\
\left\{ [\Sigma(X_i - \bar{X})^2]_1^2 - \frac{1}{r} \Sigma(X_i - \bar{X})^{\Sigma}_1^4 \right\} .
\]

5.2.12. \((\bar{G}I)_{21} = E[\Sigma(X_i - \bar{X})e_i]_1^4\Sigma(e_{i't'} - e_{i't})_1^2\)

Since \(\Sigma(e_{i't'} - e_{i't})_1^2 = \Sigma(e_{i't'} - e_{i't})_1^2 - r\Sigma(e_{i't'} - e_{i't})_1^2\),

\[(\bar{G}I)_{21} = (\bar{G}J)_{21} - r(\bar{G}H)_{21} . \]
5.2.13. Higher order expectations

Higher order expectations may be bounded by the same technique as employed in section 4.2.11. For example

\[ I_4 \leq nr \left( \frac{c}{2} \right)^2 I_3 \]

\[ (\bar{G}I)_{22} \leq nr \left( \frac{c}{2} \right)^2 (\bar{G}I)_{21} \]

and so on.

5.3. Bias and Mean Square Error of Slope Estimators

5.3.1. The estimator \( b_1 \)

The investigation of \( b_1 \) is very similar to that of \( b_Q \) in section 4.4. We can write

\[ b_1 = \frac{s_{XYB}}{s_{XXB} - s_{XXW}} = \frac{\Sigma(X_i - \bar{X})(Y_i - \bar{Y}) + \Sigma(Y_i - \bar{Y})e_i + \Sigma(X_i - \bar{X})f_i + \Sigma(e_i - e_{..})f_i}{E_{XX} + 2\Sigma(X_i - \bar{X})e_i + \Sigma(e_i - e_{..})^2 - m\Sigma(e_{it} - e_{i..})^2} \]

where \( m = \frac{n-1}{nr(r-1)} \).

\[ E(b_1 | x) = \beta \frac{\Sigma(X_i - \bar{X})e_i}{1 + \frac{2\Sigma(X_i - \bar{X})e_i + \Sigma(e_i - e_{..})^2 - m\Sigma(e_{it} - e_{i..})^2}{E_{XX}}} \]

We can now write the denominator as a polynomial plus a remainder, exactly as in 4.4. We then take expectations,
term by term, obtaining

\[ E(b_1) = \beta \left[ 1 + \frac{2 \mu_2}{rE_{XX}} + \frac{H_2 + m^2 I_2}{E_{XX}^2} - \frac{2m(HI)_{11}}{E_{XX}} \right. \\
- 8 \frac{(GH)_{11} - m(GI)_{11}}{E_{XX}^3} - \frac{H_3 - m^3 I_3}{E_{XX}^3} \\
+ 3 \frac{m(HI)_{21} - m^2(HI)_{12}}{E_{XX}^4} + 18 \frac{(GH)_{12} + m^2(GI)_{12}}{E_{XX}^4} \\
+ 8 \frac{H_4 + m^4 I_4}{E_{XX}^4} - 36 \frac{m(GHI)_{111}}{E_{XX}^4} \\
- 4 \frac{m(HI)_{31} + m^3(HI)_{13}}{E_{XX}^4} + 6 \frac{m^2(HI)_{22}}{E_{XX}^4} \\
- 48 \frac{(GH)_{21} - m(GI)_{21}}{E_{XX}^5} - 32 \frac{(GH)_{13} - m^3(GI)_{13}}{E_{XX}^5} \\
- \frac{H_5 - m^5 I_5}{E_{XX}^5} \\
+ 96 \frac{m(GHI)_{121} - m^2(GHI)_{112}}{E_{XX}^5} + 5 \frac{m(HI)_{41} - m^4(HI)_{14}}{E_{XX}^5} \\
- 10 \frac{m^2(HI)_{32} - m^3(HI)_{23}}{E_{XX}^5} - 32 \frac{G_3}{E_{XX}^6} \]
\[
- 80 \frac{(\overline{G}H)_{22} + m^2(\overline{G}I)_{22}}{E_{XX}^6} - 10 \frac{(\overline{G}H)_{14} + m^4(\overline{G}I)_{14}}{E_{XX}^6} \\
+ 160 \frac{m(\overline{G}HI)_{211}}{E_{XX}^6} + 40 \frac{m(\overline{G}HI)_{131} + m^3(\overline{G}HI)_{113}}{E_{XX}^6} \\
- 60 \frac{m^2(\overline{G}HI)_{122}}{E_{XX}^6} + \beta \Delta, \text{ where}
\]

\[
|\Delta| < \left[ 1 + \frac{\frac{1}{2}c \Sigma |x_i - \overline{x}|}{E_{XX}} \right] \left[ 1 - \frac{c \Sigma |x_i - \overline{x}| + mn r \frac{c^2}{2}}{E_{XX}} \right]^{-1}
\]

\[
E_{XX}^6 \left[ 64 \overline{G}_3 + \overline{H}_6 + m^2 I_6 + 240(\overline{G}H)_{22} + 240m^2(\overline{G}I)_{22} \\
+ 60(\overline{G}H)_{14} + 60m^4(\overline{G}I)_{14} + 15m^2(\overline{H}I)_{42} \\
+ 15m^4(\overline{H}I)_{24} \right] = \Delta_0, \text{ say.}
\]

We obtain in the same way for the mean square error,

\[
M_1 = \beta^2 \left[ \frac{\mu_2}{mE_{XX}} + \frac{\overline{H}_2 + m^2 I_2}{E_{XX}^2} - 2 \frac{m(\overline{H}I)_{11} + m^4(\overline{G}I)_{11}}{E_{XX}^3} \\
- 2 \frac{\overline{H}_3 - m^3 I_3}{E_{XX}^3} + 6 \frac{m(\overline{H}I)_{21} - m^2(\overline{H}I)_{12}}{E_{XX}^3} \\
+ 39 \frac{(\overline{G}H)_{12} + m^2(\overline{G}I)_{12}}{E_{XX}^4} + 12 \frac{\overline{G}_2}{E_{XX}^4} + 3 \frac{\overline{H}_4 + m^4 I_4}{E_{XX}^4} \right]
\]
\[
\begin{align*}
&\ -72 \frac{m(GHI)_{111}}{E_{XX}^4} - 12 \frac{m(HI)_{31} + m^3(HI)_{13}}{E_{XX}^4} + 18 \frac{m^2(HI)_{22}}{E_{XX}^4} \\
&\ -112 \frac{(GH)_{21} - m(GI)_{21}}{E_{XX}^5} - 100 \frac{(GH)_{13} - m^3(GI)_{13}}{E_{XX}^5} \\
&\ -4 \frac{m^5(GI)_{15} + 300 m^2(GHI)_{122}}{E_{XX}^5} - 40 \frac{m^2(GHI)_{112}}{E_{XX}^5} \\
&\ + 20 \frac{m(HI)_{41} - m^4(HI)_{14}}{E_{XX}^5} - 40 \frac{m^2(HI)_{32} - m^3(HI)_{23}}{E_{XX}^5} \\
&\ -240 \frac{(GH)_{22} + m^2(GI)_{22}}{E_{XX}^6} \\
&\ -95 \frac{(GH)_{14} + m^4(GI)_{14}}{E_{XX}^6} + 1360 \frac{m(GHI)_{211}}{E_{XX}^6} \\
&\ + 380 \frac{m(GHI)_{131} + m^3(GHI)_{113}}{E_{XX}^6} - 570 \frac{m^2(GHI)_{122}}{E_{XX}^6} \\
&\ -480 \frac{(GH)_{31} - m(GI)_{31}}{E_{XX}^7} - 240 \frac{(GH)_{23} - m^3(GI)_{23}}{E_{XX}^7} \\
&\ -6 \frac{(GH)_{15} - m^5(GI)_{15}}{E_{XX}^7} + 30 \frac{m(GHI)_{141} - m^4(GHI)_{114}}{E_{XX}^7} \\
&\ -60 \frac{m^2(GHI)_{132} - m^3(GHI)_{123}}{E_{XX}^7}
\end{align*}
\]
\[
+ 720 \frac{m(\overline{GHI})_{211} - m^2(\overline{GHI})_{122}}{E_{XX}^7} + \frac{\sigma^2_f}{rE_{XX}^2} \left[ 1 + \frac{4\mu_2}{rE_{XX}} \right] \\
+ \frac{H_2 + m^2I_2}{E_{XX}^2} - 2 \frac{m(HI)_{111}}{E_{XX}^2} - 12 \frac{(\overline{GH})_{111} - m(\overline{GI})_{111}}{E_{XX}^3} \\
+ \frac{H_3 - m^3I_3}{E_{XX}^3} + 3 \frac{m(HI)_{211} - m^2(HI)_{122}}{E_{XX}^3} + 24 \frac{(GH)_{122} + m^2(GI)_{122}}{E_{XX}^4} \\
+ 16 \frac{G_2}{E_{XX}^4} + \frac{H_4 + m^4I_4}{E_{XX}^4} - 48 \frac{m(GHI)_{111}}{E_{XX}^4} \\
- 4 \frac{m(HI)_{311} + m^3(HI)_{113}}{E_{XX}^4} + 6 \frac{m^2(HI)_{122}}{E_{XX}^4} \\
- 80 \frac{(GH)_{211} - m(\overline{GI})_{211}}{E_{XX}^5} - 40 \frac{(GH)_{133} - m^3(\overline{GI})_{133}}{E_{XX}^5} \\
- \frac{1}{5} \frac{- m^5I_5}{E_{XX}^5} + 120 \frac{m(GHI)_{121} - m^2(\overline{GHI})_{112}}{E_{XX}^5} \\
+ 5 \frac{m(HI)_{411} - m^4(HI)_{114}}{E_{XX}^5} - 10 \frac{m^2(HI)_{322} - m^3(HI)_{233}}{E_{XX}^5} \\
+ \beta^2 \Delta' + \frac{\sigma^2_f}{rE_{XX}^2} \Delta''
\]

where \(|\Delta'| < \left[ 9 + \frac{7cE|\chi_i - \chi_j|}{E_{XX}} \right] \Delta_0 \)
5.3.2. The estimator $b_2$

The estimator $b_2 = \frac{S_{YXY} - S_{YYW}}{S_{XYB}}$ is more troublesome to investigate because of the presence of both the $e_i$ and $f_i$ in the denominator. Not only is the algebra more involved; there is the additional difficulty that the higher moments of the $f_i$ now affect the bias and the mean square error, whereas heretofore only the second moment of the $f_i$ has entered. To simplify the investigation we shall now make the following assumptions in addition to those of section 5.1:

8) $\text{Prob} \left| f_{it} \right| \geq \frac{1}{2} h \beta c = 0$.

That is to say, we now restrict $f_{it}$ to a finite range $(-\frac{3}{2} h \beta c, \frac{3}{2} h \beta c)$; the symbol $h$ is defined by this relation.

9) $\frac{\nu_2}{\mu_2} = h^2 \beta^2$; $\frac{\nu_4}{\mu_4} = h^4 \beta^4$; $\nu_3 = 0$.

These assumptions would be satisfied if the distribution of the $f_{it}$ differed from that of the $e_{it}$ only in its range.
Now \(b_2 = \frac{\Sigma(y_i - \bar{y})^2 - m\Sigma(y_{it} - y_{i\cdot})^2}{E_{XY}}\).

We can write

\[
1 + \frac{\Sigma(y_i - \bar{y})e_i}{E_{XY}} + \frac{\Sigma(X_i - \bar{X})f_i}{E_{XY}} + \frac{\Sigma(e_i - e_{\cdot \cdot})f_i}{E_{XY}}
\]

as a polynomial in \(\frac{\Sigma(y_i - \bar{y})e_i}{E_{XY}} + \frac{\Sigma(X_i - \bar{X})f_i}{E_{XY}} + \frac{\Sigma(e_i - e_{\cdot \cdot})f_i}{E_{XY}}\),

plus a remainder, just as in sections 4.4 and 5.3.1. \(E(b_2)\) can then, of course, be obtained by taking expectations, term by term.

Because of the labor involved we have not carried the polynomial to as high a degree as in our earlier investigations; however we have carried enough terms to ensure satisfactory accuracy for the comparisons we shall want to make later. We find upon taking expectations that

\[
E(b_2) = \beta \left[ 1 + \frac{(1-h^2)\mu_2}{rE_{XX}} + \frac{(n+3)h^2\mu_2^2}{r^2E_{XX}^2} + \frac{h^4(\bar{G}H)_{11}}{E_{XX}^3} - \frac{m^h(1-6r^{-1}\mu_2)(\bar{G}I)_{11}}{E_{XX}^3} + \frac{h^4\mu_2H_{2}}{rE_{XX}^3} - \frac{m^h\mu_2(\bar{H}I)_{11}}{rE_{XX}^3} + \frac{(1-h^4)\bar{G}_2}{E_{XX}^4} + 6(h^2+h^4)\frac{\mu_2(\bar{G}H)_{12}}{rE_{XX}^4} + \frac{h^6(\bar{G}H)_{21}}{E_{XX}^5} - \frac{m^6(\bar{G}I)_{21}}{E_{XX}^5} + \Delta \right].
\]
When \[ \frac{\sum(Y_i - \bar{Y})e_i + \sum(X_i - \bar{X})f_i + \sum(e_i - e_\cdot)f_i}{E_{XY}} \] is less than 1, the series of expectations is, before rearrangement, a convergent alternating series, and in this situation, which is the only one we attempt to deal with, it is easy to find bounds for \( \Delta \). In fact,

\[
0 < \Delta < \left[ 1 + \frac{\chi E|X_i - \bar{X}| + nh^2\left(\frac{e}{2}\right)^2}{E_{XX}} \right].
\]

The mean square error, \( M_2 \), can be investigated in the same way, but the algebra is very much more tedious and will not be reproduced here. Suffice it to say that the first four terms of \( M_2 \) are

\[
\frac{\beta^2\mu_2 + \sigma_f^2}{rE_{XX}} + \beta^2 \frac{\bar{H}_2 + m^2I_2}{E_{XX}^2} - 2m\beta^2 \frac{(\bar{H}I)_{11}}{E_{XX}^2} + \frac{4\mu_2\sigma_f^2}{r^2E_{XX}^2},
\]

which agrees exactly with the first four terms of \( M_1 \). Consequently, we feel that \( M_2 \) cannot differ greatly from \( M_1 \) for errors of the magnitude we are considering, and it does not seem worthwhile to make a detailed investigation of \( M_2 \) for
the sake of detecting what could be at most a small difference from \( \mu \).

5.3.3. The estimator \( b_3 \)

The estimator \( b_3 = [s_{YY} - s_{YW}]^{1/2}[s_{XX} - s_{WW}]^{-1/2} \text{sgn } \beta \)
is considerably easier to deal with than \( b_2 \) because we can exploit the independence of the numerator and the denominator. However, we do have to expand the numerator in a series; consequently, the higher moments of the \( f_1 \), enter the picture, just as they did with \( b_2 \). We shall therefore make assumptions 8 and 9 of section 5.3.2 and shall further assume that \( \nu_5 = 0 \) and \( \nu_6 = h^6 \rho^6 \).

We can expand \( [s_{XX} - s_{WW}]^{-1/2} \) in an infinite series of powers of

\[
\frac{2 \sum (x_i - \bar{x})e_i + \sum (e_i - \bar{e})^2 - m \sum e_i t - \bar{e})^2}{\sum e_{XX}}
\]

which series converges whenever this quantity is less than unity in absolute value. It is easily verified in all of the cases we consider that the series does actually converge. We can accordingly take expectations term by term and find bounds for the error just as we have done all along. We find
\[ E[s_{XXB} - s_{XXW}]^{-1/2} = \frac{1}{E_{XX}} \left\{ 1 + \frac{3\mu_2}{2rE_{XX}} + \frac{3}{8E_{XX}^2} [\bar{H}_2 + m^2 I_2]ight\} \]

- \[2m(\bar{H}I)_{11} - \frac{5}{16E_{XX}^3} [12(\bar{G}H)_{11} - 12m(\bar{G}I)_{11} + \bar{H}_3 - m^3 I_3] \]

- \[3m(\bar{H}I)_{21} + 3m^2(\bar{H}I)_{12}] + \frac{35}{128E_{XX}^4} [16\bar{G}_2 + \bar{H}_4 + m^4 I_4 \]

- \[4m(\bar{H}I)_{31} - 4m^3(\bar{H}I)_{13} + 24(\bar{G}H)_{12} + 24m^2(\bar{G}I)_{12} \]

+ \[6m^2(\bar{H}I)_{22} - 48m(\bar{G}HI)_{111} - \frac{63}{256E_{XX}^5} [\bar{H}_5 - m^5 I_5] \]

- \[80(\bar{G}H)_{21} - 80m(\bar{G}I)_{21} - 5m(\bar{H}I)_{41} + 5m^4(\bar{H}I)_{14} \]

+ \[40(\bar{G}H)_{13} - 40m^3(\bar{G}I)_{13} + 10m^2(\bar{H}I)_{32} - 10m^3(\bar{H}I)_{23} \]

- \[120(\bar{G}HI)_{121} + 120m^2(\bar{G}HI)_{112}] + E_{XX}^{-1/2} \Delta_1 \]

where

\[|\Delta_1| < \frac{693}{512} \left[ 1 - \frac{cE|x|-\bar{X}| + mnr(S)^2}{E_{XX}} \right]^{-3/2} \]

\[E_{XX}^{-6} [64\bar{G}_3 + \bar{H}_6 + m^2 I_6 + 240(\bar{G}H)_{22} + 240m^2(\bar{G}I)_{22} \]

+ \[60(\bar{G}H)_{14} + 60m^4(\bar{G}I)_{14} + 15m^2(\bar{H}I)_{42} + 15m^4(\bar{H}I)_{24}] \]

This expression for \(|\Delta_1|\) follows directly from Cauchy's formula for the remainder of a power-series expansion.

In the same way we can expand \([s_{YYB} - s_{YYW}]^{1/2}\) in an infinite series of powers of
\[
\frac{2\Sigma (Y_{i} - Y) f_{i} + E(f_{i} - f_{..})^{2} - m \Sigma (f_{i} - f_{..})^{2}}{E_{YY}}
\]

which series converges whenever this quantity is less than unity in absolute value. We confine ourselves to situations wherein the series does converge and take expectations term by term in the familiar way. We obtain

\[
E[s_{YYB} - s_{YYW}]^{1/2} = E_{YY}^{1/2} \left\{ 1 - \frac{h^{2} \nu_{2}}{2E_{XX}} - \frac{h^{4}}{8E_{XX}} [H_{2} + m^{2} I_{2}]
\right.
\]

\[\left. - 2m(HI)_{11} + \frac{1}{16E_{XX}} [12h^{4}(GH)_{11} - 12h^{4}m(GI)_{11} + h^{6}I_{3}
\right.
\]

\[\left. - h^{6}m^{2} I_{3} - 3h^{6}m(HI)_{21} + 3h^{6}m^{2}(HI)_{12} \right) - \frac{5}{128E_{XX}} [16h^{4}I_{2}
\right.
\]

\[\left. + h^{6}I_{4} + h^{8}m^{4} I_{4} - 4h^{8}m(HI)_{31} - 4h^{8}m^{3}(HI)_{13} + 24h^{6}(GH)_{12}
\right.
\]

\[\left. + 24h^{6}m^{2}(GI)_{12} + 6h^{6}m^{2}(HI)_{22} - 43h^{6}m(GHI)_{111} \right)
\]

\[\left. - \frac{7}{256E_{XX}} [h^{10}H_{5} - H^{10}m^{5} I_{5} + 80h^{6}(GH)_{21} - 80h^{6}m(GI)_{21}
\right.
\]

\[\left. - 5h^{10}m(HI)_{41} + 5h^{10}m^{4}(HI)_{14} + 40h^{8}(GH)_{13} - 40h^{8}m^{3}(GI)_{13}
\right.
\]

\[\left. + 10h^{10}m^{2}(HI)_{32} - 10h^{10}m^{3}(HI)_{23} - 120h^{8}m(GHI)_{121}
\right.
\]

\[+ 120h^{8}m^{2}(GHI)_{112} \right) + E_{YY}^{1/2} \Delta_{2},
\]

where
\[ |\Delta_2| < \frac{63}{12} \left[ 1 - \frac{\text{ho} \Sigma |X_1 - \bar{X}| + \text{mnr} \left( \frac{1}{2} \text{ch} \right)^2}{E_{XX}} \right]^{-1/2}. \]

\[ E_{XX} \{ 64h^6 \sigma_3 + h^{12} \delta_6 + h^{12} m^2 I_6 + 240h^8 (\delta H)_{22} \]
\[ + 240h^8 m^2 (\delta I)_{22} + 60h^{10} (\delta H)_{14} + 60h^{10} m^4 (\delta I)_{14} \]
\[ + 15h^{12} m^2 (\delta I)_{44} + 15h^{12} m^4 (\delta I)_{24} \} . \]

\( E(b_3) \) can now, of course, be found quite easily, since
\[ E(b_3) = E[s_{YYB} - s_{YYW}]^{1/2} [s_{XXB} - s_{XXW}]^{-1/2} . \]

We shall write down only the first term of the expression for the bias, for reference in section 5.4:
\[ B_3 = \beta \frac{(3-h^2)\mu_2}{2rE_{XX}} + \text{terms involving higher powers of } E_{XX}^{-1} . \]

It is easy to find \( M_3 \) now from the relation
\[ M_3 = E(b_3^2) - 2\beta E(b_3) + \beta^2 \]
where
\[ E(b_3^2) = \beta^2 E \left[ 1 + \frac{2 \Sigma (X_i - \bar{X})e_i \cdot + \Sigma (e_i \cdot - e_i \cdot) \cdot - m \Sigma (e_i \cdot - e_i \cdot)^2}{E_{XX}} \right]^{-1} . \]

We obtain, in the usual way,
\[ E(b_3^2) = \beta^2 \left\{ 1 + \frac{4\mu_2}{rE_{XX}} + \frac{1}{E_{XX}^2} [\bar{H}_2 + m^2 I_2 - 2m(\bar{H} I)_{11}] \right. \]
\[ - \frac{1}{E_{XX}^3} [12(\delta H)_{11} - 12m(\delta I)_{11} + H_3 - m^3 I_3 - 3m(\delta I)_{21}] \]
\[ - \frac{1}{E_{XX}} \left[ 16 \bar{G}_2 + \bar{H}_4 + m^2 I_4 - 4m(\bar{H}I)_31 - 4m^3(\bar{H}I)_{13} \right. \\
+ 24(\bar{G}H)_{12} + 24m^2(\bar{G}I)_{12} + 6m^2(\bar{H}I)_{22} - 48m(\bar{G}HI)_{111} \left. \\
- \frac{1}{E_{XX}} \left[ \bar{H}_5 - m^5 I_5 + 80(\bar{G}H)_{21} - 80m(\bar{G}I)_{21} - 5m(\bar{H}I)_{41} \right. \\
+ 5m(\bar{H}I)_{14} + 40(\bar{G}H)_{13} - 40m(\bar{G}I)_{13} + 10m^2(\bar{H}I)_{32} \\
- 10m^3(\bar{H}I)_{23} - 120m(\bar{G}HI)_{121} + 120m^2(\bar{G}HI)_{112} \right] + \beta^2 \Delta_3 \]

where

\[ |\Delta_3| < \left[ 1 - \frac{c2|x_i - x'| + mnr(Gh)}{E_{XX}} \right]^{-1} \]

\[ E_{XX} \left[ 64\bar{G}_3 + \bar{H}_6 + m^2 I_6 \right. \\
+ 240(\bar{G}H)_{22} + 240m^2(\bar{G}I)_{22} + 60(\bar{G}H)_{14} + 60m^2(\bar{G}I)_{14} \\
+ 15m^2(\bar{H}I)_{42} + 15m^4(\bar{H}I)_{24} \right] \]

5.3.4. The estimator \( b_L \)

The bias and mean square error of \( b_L = \frac{\Sigma w_i y_i}{\Sigma w_i x_i} \), where \( \Sigma w_i = 0 \), can be obtained from the corresponding characteristics of \( \frac{\Sigma w_i y_i}{\Sigma w_i x_i} \) by substituting \( \bar{\mu}_2, \bar{\mu}_4, \bar{\mu}_6, \) and \( \frac{1}{r} \sigma_r^2 \) for \( \mu_2, \mu_4, \mu_6, \) and \( \sigma_r^2 \), respectively. The results are

\[ B_L = \beta \left[ \frac{\mu_2}{r} \frac{\Sigma w_1^2}{p^2} + \frac{1}{p^4} \left( \bar{\mu}_4 \Sigma w_1^4 + 3 \frac{\mu_2^2}{r^2} \Sigma \Sigma w_{1w_1}^2 \right) \right. \\
+ \text{remainder}. \]
\[ M_L = \beta^2 \left[ \frac{\mu_2^2}{\mu} \frac{S_{W_1}^2}{p^2} + \frac{3}{p^4} (\bar{\mu}_4 S_{W_1}^4 + 3 \frac{\mu_6^2}{\mu} S_{W_1}^2 S_{W_1}^2) \right] + \lambda \left[ \frac{\mu_2^2}{\mu} \frac{S_{W_1}^2}{p^2} + \frac{3}{p^6} \frac{\mu_6^2}{\mu} (S_{W_1}^2)^2 + \frac{5}{6} \frac{\mu_2^2}{\mu} S_{W_1}^2 (\bar{\mu}_4 S_{W_1}^4) + 3 \bar{\mu}_2^2 S_{W_1}^2 S_{W_1}^2 \right] + \text{remainder.} \]

Bounds for the errors can be easily obtained using the formulas developed in the unreplicated case, with \( \mu_2, \mu_4, \mu_6, \) and \( \sigma_f^2 \) replaced by \( \bar{\mu}_2, \bar{\mu}_4, \bar{\mu}_6, \) and \( \frac{\sigma_f^2}{r} \) respectively.

5.3.5. The estimator \( b\_Q \).

The estimator \( b\_Q = \frac{\Sigma x_{it}(y_{it}-\bar{y}..)}{\Sigma x_{it}(x_{it}-\bar{x}..)} \) is what one would obtain if he employed the conventional least-squares procedure, simply ignoring the randomness of the \( x_{it} \) and the fact that there are \( r \) observations for each \( X_i \). We are not suggesting that this is an appropriate way of estimating \( \beta \); in fact, we shall show that quite the contrary is the case. We have nevertheless investigated \( b\_L \) for two reasons:

1. Many investigators use the conventional least-squares procedures uncritically, and it is not unthinkable that some might use \( b\_L \).

2. It is interesting to assess the price one must pay for ignoring randomness of the \( x_{it} \) and replication.

The examination of this estimator is rendered quite
simple by the same artifice we have already used. We write

\[ x_1,1 = z_1 \quad y_1,1 = v_1 \]
\[ x_1,2 = z_2 \quad y_1,2 = v_2 \]
\[ \vdots \quad \vdots \]
\[ x_1,r = z_r \quad y_1,r = v_r \]
\[ x_2,l = z_{r+1} \quad y_2,l = v_{r+1} \]
\[ \vdots \quad \vdots \]
\[ x_n,r = z_{nr} \quad y_n,r = v_{nr} \]

Then \( b_L = \frac{\Sigma v_1(z_1 - \overline{Z})}{\Sigma z_1(z_1 - \overline{Z})} \), where now \( i = 1, 2, \ldots, nr \). \( B_L \) and \( M_L \) can be obtained from the corresponding expressions of section 4.4 by substituting \( nr \) for \( n \), \( \Sigma(Z_1 - \overline{Z})^2 \) for \( E_{XX} \), \( \Sigma \Sigma(Z_1 - \overline{Z})^2(Z_j - \overline{Z})^2 \) for \( \Sigma \Sigma(X_1 - \overline{X})^2(X_j - \overline{X})^2 \), and so forth. We can then express quantities involving the \( Z_1 \) in terms of the \( X_1 \), as follows:

\[ \Sigma(Z_1 - \overline{Z})^2 = r \Sigma(X_1 - \overline{X})^2. \]

\[ \Sigma \Sigma(Z_1 - \overline{Z})^2(Z_j - \overline{Z})^2 = [\Sigma(Z_1 - \overline{Z})^2]^2 - \Sigma(Z_1 - \overline{Z})^4 \]
\[ = r^2 [\Sigma(X_1 - \overline{X})^2]^2 - r \Sigma(X_1 - \overline{X})^4. \]

\[ \Sigma \Sigma(Z_1 - \overline{Z})^4(Z_1 - \overline{Z})^2 = \Sigma(Z_1 - \overline{Z})^4 \Sigma(Z_1 - \overline{Z})^2 - \Sigma(Z_1 - \overline{Z})^6 \]
\[ = r^2 [\Sigma(X_1 - \overline{X})^4][\Sigma(X_1 - \overline{X})] - r \Sigma(X_1 - \overline{X})^6. \]
$$\sum (z_i - \bar{z})^2 (z_j - \bar{z})^2 (z_k - \bar{z})^2 = [\sum (z_i - \bar{z})^2]^3 - 3 \sum (z_i - \bar{z})^5 (z_j - \bar{z})^2$$

$$- \sum (z_i - \bar{z})^6$$

$$= r^3 [\sum (x_i - \bar{x})^2]^3 - 3 r^2 [\sum (x_i - \bar{x})^4] [\sum (x_i - \bar{x})^2]$$

$$+ 2 r^2 (x_i - \bar{x})^6.$$  

5.4. Comparison of Biases for Estimators of \( \beta \)

In order to compare the estimators discussed in section 5.3 we shall first take

$$\mu_2 = \frac{1}{2} \left( \frac{c}{2} \right)^2, \quad \mu_4 = \frac{1}{4} \left( \frac{c}{2} \right)^2, \quad \text{and} \quad \mu_6 = \frac{1}{7} \left( \frac{c}{2} \right)^2,$$

which are the moments of the rectangular distribution on the interval \((- \frac{c}{2}, \frac{c}{2})\), whose range is the maximum possible consonant with the assumptions of section 5.1. As we have already pointed out, we could consider any other finite-range distribution on this interval in the same way, and we should arrive at essentially the same conclusions.

It is easily verified that when \( n \geq 6 \) the series expansion employed for the denominator of \( b_3 \) is convergent and the bounds for the remainder of this series, as well as the bounds employed for \( b_1 \), \( b_L \), and \( b_8 \), are satisfactorily small. The situation is more complicated with respect to the numerator of \( b_3 \) and the denominator of \( b_2 \) because these involve the moments of the fit, which contain various power of \( h \) as
To identify the difficulty specifically, let us consider the expansion of the denominator of $b_2$. We pointed out in section 5.3.2 that in order to obtain satisfactory bounds for the remainder term we required that

$$\frac{\sum (X_i - \bar{X}) e_i + \sum (X_i - \bar{X}) f_i + \sum (e_i - e_{..}) f_i}{E_{XY}} < 1.$$ 

Now the maximum value of the left-hand side of this inequality is

$$\frac{\frac{c}{2}(1+h) \sum |X_i - \bar{X}| + hn(c/2)^2}{E_{XX}}.$$ 

It is clear that if we make $h$ too large [5.4 - 1] will exceed unity. In fact for $n = 6$, $\omega = 1$ we cannot get a satisfactory remainder term when $h > 1$. Of course for larger values of $n$ or other values of $\omega$ we could allow larger values of $h$.

It should, however, be remembered that we have insisted on allowing the $e_{it}$ to have the maximum possible range. If one is willing to decrease the range of the $e_{it}$, he can permit larger values of $h$. We shall pursue this point shortly, but we should like first to give some representative results when the $e_{it}$ are allowed their maximum range and $h$ is limited accordingly. The figures in Table 15 give the bias ratios when the $X_i$ are symmetrically spaced with $\omega = 1$, for $n = 6$, $r = 2$, and for $n = 8$, $r = 2$. The upper entry in
Table 15. Bias ratios of various slope estimators

<table>
<thead>
<tr>
<th></th>
<th>h = 1</th>
<th>h = 1/2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$8.81 &lt; \frac{B_0}{B_0} &lt; 8.82$</td>
<td>$2.04 &lt; \frac{E_1}{B_0} &lt; 2.11$</td>
<td>$2.01 &lt; \frac{B_1}{B_0} &lt; 2.02$</td>
</tr>
<tr>
<td>$0.0345 &lt; \frac{B_2}{B_0} &lt; 0.0395$</td>
<td>$0.756 &lt; \frac{B_2}{B_0} &lt; 0.758$</td>
<td>$B_2 = 0.0190 = 0.749$</td>
</tr>
<tr>
<td>$0.99 &lt; \frac{B_3}{B_0} &lt; 1.04$</td>
<td>$1.43 &lt; \frac{B_3}{B_0} &lt; 1.44$</td>
<td>$B_3 = 1.38$</td>
</tr>
<tr>
<td>$0.98 &lt; \frac{B_3}{B_1} &lt; 0.99$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
each cell corresponds to \( n = 6 \), the lower, to \( n = 8 \).

Had we approximated these bias ratios by using only the leading term of the expression for the bias of each estimator, we should have obtained

\[
\frac{B_0}{B_0} = 9 \text{ for } n = 6 ; \quad \frac{B_0}{B_0} = 13 \text{ for } n = 8
\]

\[
\frac{B_1}{B_0} = 2
\]

\[
\frac{B_2}{B_0} = 0 \text{ for } h = 1 ; \quad \frac{B_2}{B_0} = 3/4 \text{ for } h = 1/2
\]

\[
\frac{B_3}{B_0} = 1 \text{ for } h = 1 ; \quad \frac{B_3}{B_0} = 11/8 \text{ for } h = 1/2
\]

Bearing in mind that for asymmetric spacing, for other values of \( \omega \), for higher values of \( n \), and for lower values of \( h \) the convergence of the bias expressions is even more rapid than in the cases just examined, we feel that a satisfactory study of the bias ratios can be made using the leading terms only. We have then, approximately,

\[
\frac{B_0}{B_0} = - (nr - 3)
\]

\[
\frac{B_1}{B_0} = 2
\]

\[
\frac{B_2}{B_0} = 1 - h^2 \quad \text{provided } 0 \leq h \leq 1
\]
It will be recalled that the restriction on $h$ arises because of the need to ensure the convergence of series involving the $f_{it}$, whose range we have assumed to be $h\beta$ times the range of the $e_{it}$. If we are willing to limit the range of the $e_{it}$ (and hence of the $f_{it}$) sufficiently to ensure the convergence of the series, we can allow values of $h$ larger than unity. In fact it is clear that for any arbitrary $h$, $n$, and $X_i$ we can make

$$\frac{\sum (Y_i - \bar{Y})e_i}{E_{XY}} + \frac{\sum (X_i - \bar{X})f_i}{E_{XY}} + \frac{\sum (e_i - \bar{e})f_i}{E_{XY}} < 1$$

by simply restricting the range of the $e_{it}$ sufficiently. Precisely the same argument applies to the expansion of the numerator of $b_3$. Consequently we may say that for sufficiently small errors in the $e_{it}$

$$\frac{B_q}{B_0} = -nr-3$$

$$\frac{B_1}{B_0} = 2$$

$$\frac{B_2}{B_0} = 1 - h^2$$

$$\frac{B_3}{B_0} = \frac{3 - h^2}{2}$$
where now we do not insist that \( h \leq 1 \). \( \frac{B_Q}{B_0} \) increases very rapidly with \( n \); consequently \( b_Q \) is not suitable as an estimator of \( \beta \) if one is at all concerned about bias. The last three ratios are plotted in Figure 3, from which figure it is possible to draw the following conclusions when the errors are small:

1. When \( h < \sqrt{5/3} \), \( b_2 \) has the smallest absolute bias among the estimators considered. However, for larger values of \( h \) the absolute bias of \( b_2 \) increases very rapidly.

The value \( \sqrt{5/3} \) is the largest value of \( h \) such that the following inequalities are satisfied simultaneously:

\[
| 1 - h^2 | < 1
\]
\[
| \frac{1}{2} (3-h^2) | < 1
\]
\[
| 1 - h^2 | < \frac{1}{2} (3-h^2) |
\]

2. When \( h > \sqrt{5} \), \( b_0 \) has the smallest absolute bias.

The value \( \sqrt{5} \) is the smallest value of \( h \) such that the following inequalities are satisfied simultaneously:

\[
| 1 - h^2 | > 1
\]
\[
| \frac{1}{2} (3-h^2) | > 1
\]
Figure 3. Relative bias as a function of $\log_{10} h$
Relative Bias

\[ \frac{B^2}{b^0} \]

\[ \frac{B^3}{b^0} \]

\[ \frac{B^1}{b^0} \]

\[ \log_{10} \]

\[ \log_{10} \]

\[ b_0 \]
3. When \( h > \sqrt{3} \), \( b_1 \) has smaller absolute bias than \( b_2 \) since \( |1 - h^2| > 2 \).

When \( h > \sqrt{7} \), \( b_1 \) has smaller absolute bias than either \( b_2 \) or \( b_3 \) since the inequalities
\[
|1 - h^2| > 2
\]
and
\[
|\frac{1}{2} (3 - h^2)| > 2
\]
are simultaneously satisfied.

4. There is a small range of values of \( h \) for which \( b_3 \) has the smallest absolute bias; in fact when \( \sqrt{5/3} < h < \sqrt{5} \), this is the case. This range of values of \( h \) is obtained from the simultaneous solution of the inequalities
\[
|\frac{1}{2} (3 - h^2)| < 1
\]
and
\[
|\frac{1}{2} (3 - h^2)| < |1 - h^2|.
\]
However, we should ordinarily be unable to take advantage of this fact because of lack of knowledge about \( h \). (It will be remembered that \( h^2 = \frac{\lambda}{\beta^2} \).)

Moreover, for \( h > \sqrt{5} \) the absolute bias increases very rapidly. Consequently it seems doubtful that \( b_3^* \) would be of much practical use in providing an estimator with desirable bias characteristics. We have seen in section 2.6 that it would also be hard to justify using \( b_3 \) from the standpoint of asymptotic
It must be conceded that our conclusions concerning what happens to $b_2$ and $b_3$ for large values of $h$ are based upon the assumption that the errors are sufficiently small. We know that the absolute bias of these two estimators increases very rapidly with increasing $h$ when the errors are small; it seems reasonable to believe that it behaves in the same way for large errors. On the other hand, the absolute bias of $b_1$ and $b_0$ does not depend upon $h$ at all. It appears likely then that the analysis we have just made for small errors is essentially valid for large errors as well, and we feel that it would be hard to justify using $b_2$ or $b_3$ whenever $h$ is larger than $\sqrt{3}$.

We shall make a final assessment and interpretation of these results in section 6, which will be devoted to overall conclusions for the entire dissertation.

5.5. Comparison of Mean Square Error for Estimators of $\beta$

We have already seen in section 2.3 that the leading term of $\text{Var}_A$ is exactly the same for $b_1$, $b_2$, $b_3$, and $b_0$. Consequently, it should not surprise us to find that these estimators all have approximately the same mean square error. After all, the leading term in the expression for $M$ is exactly the same for each of these estimators, and by now we have

\[ \text{Var}_A \]
amassed considerable evidence, all pointing to the fact that
the higher order terms do not change B or M very much.

We have, moreover, computed \( \frac{M_1}{M_0} \), \( \frac{M_3}{M_0} \), and \( \frac{M_Q}{M_0} \) for \( n = 8 \),
\( \omega = 1 \), and symmetric spacing. We obtained

\[
0.99 < \frac{M_1}{M_0} < 1.07
\]

\[
1.02 < \frac{M_3}{M_0} < 1.03 \quad \text{when} \quad h = 1
\]

\[
\frac{M_3}{M_0} = 0.99 \quad \text{when} \quad h = 1/2
\]

\[
1.07 < \frac{M_Q}{M_0} < 1.10
\]

We have not actually computed \( \frac{M_2}{M_0} \) because of the great labor
involved, but we feel certain that it too would be close to
unity for the reasons set forth in section 5.3.2. For larger
values of \( n \), for other values of \( \omega \), and for lower values of
\( h \) the convergence is even more rapid, and we would find these
ratios still closer to unity than in the case we have con­sidered. As for large values of \( h \), we are obliged to restrict
the range of the \( e_{it} \) in order to consider them, and this
alone greatly improves the convergence of the series em­ployed. We would therefore expect our findings to be similar
in these instances also. In short the estimators we have
considered show little variation in their mean square error; if we seek a criterion for choosing among them, we must clearly look elsewhere. One possibility is to consider the bias as was done in section 5.4.
6. CONCLUSIONS

This dissertation is an outgrowth of an actual problem submitted to John Gurland involving a linear functional relation with error in both variables. The conclusions reached here are those of a pragmatist desiring to recommend a practicable method of estimation to experimenters who must deal with this not-infrequent problem. Our conclusions apply for the most part to those problems where the errors in the $x_i$ are independent of the errors in the $y_i$. Although in section 3 we have examined the large-sample properties of the various estimators when the errors are correlated, we have as yet been unable to undertake an investigation of the small-sample properties.

Let us first consider the situation when there is no replication and the errors are independent. It is not an attractive situation, but neither is it as bleak as some statisticians seem to feel. In fact there are at least four reasonable ways of obtaining estimators of $\beta$.

Firstly, one could assume the form of the distribution of the errors to be known and could then apply the method of maximum likelihood. For those cases where $\beta$ is identifiable we are assured of getting a consistent estimator for it. The drawback in this method is the difficulty of obtaining an explicit solution for the estimator. In particular, when the
errors are normally distributed, the explicit solution of the likelihood equations has never appeared in print, as far as we know.

Secondly, we could assume the form of the distribution of the errors to be known and could then apply the minimum distance method of Wolfowitz (59, 60, 61).

Thirdly, we could simply ignore the random nature of the $x_i$ and use ordinary least squares; that is, we could make use of the estimator $b_Q$ of section 4.4. However, we must reckon with the fact that the bias of $b_Q$ increases rapidly with the sample size; where bias is of importance, $b_Q$ is not recommended.

Fourthly, provided that the errors are not too large to prevent us from correctly ordering the observations, we could use one of the family of ratio-of-linear-forms estimators $b_L$ discussed in section 4.3. To be specific one might use Wald's estimator $b_W$, Bartlett's estimator $b_B$, Housner and Brennan's estimator $b_H$, or better yet, the optimal estimator $b_0$ when it is attainable. Except when the $X_i$ are highly skewed or bunched, $b_H$ behaves surprisingly well in that its bias is very much less than that of $b_Q$ (roughly $-\frac{1}{n-3}$ as great) while its mean square error is not much greater. There is also the advantage of consistency on the side of $b_H$.

In most of the cases examined $b_H$ proved superior to
both $b_w$ and $b_B$ - usually quite markedly so. In fact $b_H$
appears to be surprisingly robust with respect to the spac­
ing of the $X_i$ provided that the $X_i$ do not exhibit excessive
skewness or bunchiness. However, when the $X_i$ fall into two
groups, each of which is closely bunched, $b_w$ is superior to
both $b_H$ and $b_B$; when the $X_i$ fall into three groups, each of
which is closely bunched, $b_B$ is the superior estimator.
When we have more than three groups of closely bunched
observations, estimators of the type suggested in section
4.5.3 are better than any of the foregoing.

What to do in the case of excessive skewness is a sub-
ject of continuing research; it appears that other estimators
belonging to the family $b_L$ can be found which behave reason­
ably well.

When there is replication and the errors are independent
we have at least four estimators of $\beta$ in addition to those
already discussed; namely, $b_1$, $b_2$, $b_3$, and $b_4$, discussed in
section 2 and section 5. The comments made above concerning
the maximum-likelihood estimator, the minimum-distance esti­
mator, and $b_Q$, in the case of unreplicated observations, are
equally applicable here. As for the other estimators, let
us summarize what we have found about them. To begin with,
$b_1$, $b_2$, $b_3$, and $b_0$ have essentially the same mean square
error, so if one is to choose among them, it will have to be
on some other grounds. A reasonable criterion for making
such a choice is the bias, particularly so inasmuch as the estimators differ considerably in this respect. Suppose then that we agree to use the bias as a criterion; we recommend the following procedures for replicated observations, depending upon the information available.

**Situation 1**

It is known that the $X_i$ are not highly skewed

- **Some idea of $h$ is available**
  - When $h > \sqrt{2}$, use $b_h$ if there is no bunching
  - $b_w$, $b_B$, or some other member of the family $b_L$ if there is bunching; see page 172

- When $h < \sqrt{2}$, use $b_2$

**Situation 2**

Either it is known that the $X_i$ are highly skewed or nothing is known about the spacing of the $X_i$

- **Some idea of $h$ is available**
  - When $h > \sqrt{3}$, use $b_1$
  - When $h < \sqrt{3}$, use $b_2$

- **Nothing is known about $h**
  - Use $b_1$
These recommendations are, of course, a consequence of the discussion in section 5.4 and section 6.

One might also try to choose among the various estimators of \( \beta \) using \( \text{Var}_A \) as the criterion. For spacings of the \( X_i \) normally encountered this does not seem to be a judicious criterion because the estimators differ in their \( \text{Var}_A \) only by terms of higher order; see section 2.3.1. Nevertheless it is interesting to note that use of \( \text{Var}_A \) as a criterion leads to essentially the same conclusions as use of the small-sample bias. To see this one should assume that nothing is known about the spacing of the \( X_i \) (since such information is not used in \( \text{Var}_A \)), but that some idea of \( h \) is available. We found, in section 2.6, that \( \text{Var}_A(b_1) - \text{Var}_A(b_2) < 0 \) when \( h > 1 \) and \( \text{Var}_A(b_1) - \text{Var}_A(b_2) > 0 \) when \( h < 1 \); in other words, use \( b_1 \) when \( \beta \) is small relative to \( \lambda \) - use \( b_2 \) when \( \beta \) is large relative to \( \lambda \).

We also found in section 2.6 that in some situations \( b_4 \) is better than \( b_1 \) or \( b_2 \) with regard to \( \text{Var}_A \). This suggests the possibility that \( b_4 \) might have smaller bias than either \( b_1 \), \( b_2 \), or \( b_0 \), at least for certain values of the parameters. Unfortunately we have been unable to find practicable means for studying the small-sample bias of \( b_4 \) and must leave this question open.

We have contented ourselves with the usual estimator of \( \alpha \); that is, \( \alpha = \bar{y} - b\bar{x} \). The main justification for this
estimator seems to be its consistency; however, it is conceivable that other consistent estimators of $\alpha$ might have superior small-sample properties. There is much less interest among experimenters in estimating $\alpha$ than in estimating $\beta$, and consequently the former problem has not received much attention.

As for the situation where the errors in the $x_i$ are correlated with those in the $y_i$, there is unfortunately little that we can say. We have not at this time been able to examine the small-sample properties of the estimators; we have examined $\text{Var}_A$ in detail, but it does not furnish a very satisfactory basis for choosing among the estimators since it is almost the same for all of them, at least for the spacings of the $X_i$ that are usually encountered. Still, $\text{Var}_A$ is the only criterion that we have here, and for lack of a better one, we shall state the conclusion to which it leads:

- when $h > 1$, use $b_1$;
- when $h < 1$, use $b_2$,

where $b_1$ and $b_2$ are defined in section 3.2.

Almost every investigation, experiment, or dissertation poses more numerous and more difficult problems than it ever solves; this dissertation is no exception. We shall conclude it now by asking some of the more troubling questions that have emerged:
1. How does the bias of $b_\delta$ compare with that of $b_1$, $b_2$, and $b_0$?

2. How do $b_1$ and $b_2$ behave when we permit the errors to be really large?

3. What happens to the bias of $b_H$ when 
   \[ \text{Prob} \mid e_{it} \mid > \min_i |X_{i+1} - X_i| \neq 0 \]

4. Is it feasible to investigate non-linear functional relations along the lines employed for linear relations?
7. BIBLIOGRAPHY


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


17. Drion, E. F. Estimation of the parameters of a straight line and of the variances of the variables, if they are both subject to error. Indagationes Mathematicae 13: 256-260.


20. Geary, R. C. Determination of linear relations between systematic parts of variables with errors of observation the variances of which are unknown. Econometrica 17: 30-59. 1949.


35. Kummell, Chas. H. Reduction of observation equations which contain more than one observed quantity. The Analyst (Des Moines) 6: 97-105. 1879.


40. Nair, K. R. and Banerjee, K. S. Note on fitting of straight lines if both variables are subject to error. Sankhya 6: 331. 1942.


49. Roos, C. F. A general invariant criterion of fit for lines and planes where all variates are subject to error. Metron 13: 3-20. 1937.


57. Wald, Abraham. The fitting of straight lines if both variables are subject to error. Annals of Mathematical Statistics 11: 284-300. 1940.


8. ACKNOWLEDGEMENT

The author would like to acknowledge the kind assistance of Dr. John. Gurland, who suggested this investigation, along with many of the ideas incorporated herein. The author is further indebted to the United States Air Force, which supported a part of this research under Air Force Contract AF49(638)-43.