ESTIMATING THE SLOPE OF A STRAIGHT LINE
WHEN BOTH VARIABLES ARE SUBJECT TO ERROR

by

Clifford Spiegelman

FSU Statistics Report M387

October, 1976
The Florida State University
Department of Statistics
Tallahassee, Florida
ESTIMATING THE SLOPE OF A STRAIGHT LINE WHEN BOTH VARIABLES ARE SUBJECT TO ERROR

by

Clifford Spiegelman
The Florida State University

ABSTRACT

This paper provides an estimate for the slope of a straight line when both the independent and dependent variables are subject to error. This is the first such estimate which is shown to have an asymptotically unbiased normal distribution with variance $O(1/n)$, under weak assumptions. In addition, a generalization of a theorem of Bernstein (1941) is given for determining the normality of the underlying variable.

This research was partially supported by the Mathematics Department at Northwestern University, as well as the National Science Foundation under grant number GSOC - 7103704-03.
Introduction

\((\Omega, \mathcal{A}, P)\) is a probability space. \((X, Y), (X_1, Y_1), \ldots, (X_n, Y_n)\) are independent and identically distributed pairs of random variables on this space. \((U, W, V), (U_1, W_1, V_1), \ldots, (U_N, W_N, V_N)\) are independent and identically distributed unobservable random vectors on this space and the median of \(V\) and \(W\) equals 0. \(X_i = U_i + V_i, V_i = \alpha + \beta U_i + V_i, i = 1, \ldots, N\), where \(\alpha\) and \(\beta\) are constants. The problem of estimating \(\beta\) is the classical errors in variables problem, in structural form. If the sequence of random variables \(U_i, i = 1, \ldots, N\) is replaced by a sequence of constants, one obtains the functional form of the errors in variables problem, which is not dealt with here. The structural form appears often in social program evaluation models, examples and references of such models may be found in Campbell and Borch (1975).

Much of the work done on this problem was done in the 1940's and 1950's. However most of the work done on this problem occurred in the 1940's and 1950's, and most of the work was devoted to finding a consistent estimate of \(\beta\), and very few results are available demonstrating any stronger properties.

Some of the more notable contributions follow.

Reiersøl (1950) proved that if \(U\) has a normal distribution, \(\beta\), is identifiable if, and only if \(W\) and \(V\) have a distribution which is not divisible by a normal distribution, and if \(U\) has a non normal distribution \(\beta\), is always identifiable.

Neyman (1951) gave a consistent estimator for \(\beta\) when \(U\) has a non normal distribution. While this is the first consistent estimator for \(\beta\) given under weak assumptions, it is extremely complicated and requires working with eight disjoint subsets of observations simultaneously. Each subset is used in a distinct manner.

This research was partially supported by the Mathematics Department at Northwestern University, as well as the National Science Foundation under grant number GSC - 7103704-03.
Wolfowitz (52, 53, 54a, 54b, 57) gave a method of estimating $\beta$ when it is identifiable. This method is often reasonable if the distributions of $U$ and $V$ belong to a sufficiently small finite dimensional class, for example when $U$ and $V$ have a joint normal distribution.

Kiefer and Wolfowitz (1956) showed that the maximum likelihood estimates for this problem are consistent, if suitable regularity conditions attain. It is not clear whether the usual optimality properties of likelihood estimates apply to this problem.

There are in the literature simple estimators for $\beta$, which are consistent under stringent conditions. Among contributors to these methods are Scott (1950), Wald (1940) who deals with the functional problem, and Neyman and Scott (1951) who deal with both the functional and structural problem. A survey of these and other contributions may be found in Moran (1971).

This paper gives estimates for $\beta$, when $U$ has a non normal distribution. The estimates have asymptotic normal distributions with mean $\beta$, and variance $O(1/n)$. An example is given in which the estimation method of this paper is applicable for estimating treatment effect in the presence of hidden variables. In addition it is shown that the non normality of $U$ is determinable by a generalization of a theorem of Bernstein (1941).

Assumptions:

1) $U$ has a non normal distribution.

2) $|\beta| \neq 0$. (If $\beta = 0$, $X$ and $Y$ are independent).

There exist estimators $T_{1n}$, $T_{2n}$ and constants (finite) $\tau_1$ and $\tau_2$ such that $T_{1n} \xrightarrow{a.s.} \tau_1$, $T_{2n} \xrightarrow{a.s.} \tau_2$, as $n \to \infty$, $\tau_1 < \beta < \tau_2$, and both $\tau_1$ and $\tau_2$ have the same sign.
When $E|X^2 + Y^2| < \infty$, \( \sum_{i=1}^{n} \frac{(X_i - \bar{X})(Y_i - \bar{Y})}{(X_i - \bar{X})^2} \), and \( \sum_{i=1}^{n} (Y_i - \bar{Y})^2/ \sum_{i=1}^{n} (Y_i - \bar{Y})(X_i - \bar{X}) \)

are useful choices for $T_{1n}$ and $T_{2n}$. If $E|X^2 + Y^2| = \infty$, other estimates may be used, for example variations of the Neyman and Wolfowitz estimators cited earlier.

Notation:

Let $g(\cdot)$ be a Borel measurable function on the line. ($g(\cdot)$ will be an exponential function (real or complex), or a cosine function.)

Define:

\[
\phi_n(g, t_1, t_2) = \frac{1}{n} \sum_{j=1}^{n} g(t_1 x_j + t_2 y_j), \quad \phi(g, t_1, t_2) = E g(t_1 x + t_2 y)
\]

\[
\psi_{nx}(g, t_1) = \phi_n(g, t_1, 0), \quad \psi_x(g, t_1) = \phi(g, t_1, 0)
\]

\[
\psi_{ny}(g, t_2) = \phi_n(g, 0, t_2), \quad \psi_y(g, t_2) = \phi(g, 0, t_2)
\]

For notational convenience the subscript $g$ will be deleted.

In the work that follows it will be important to find compact intervals $I$, sometimes depending on $\omega$, such that $\omega.p.l.$

\[
|\psi_x(t_1)\psi_y(t_2)| \wedge (\min_{T_1 < b < T_2} |\psi_x(t_2/b)\psi(t_1/b)|)
\]

is bounded away from zero, for $(t_1, t_2) \in I \times I$. If $g(x) = e^x$, then any compact interval satisfies condition 1 (even if expression (1) is infinite).

Generally the following is done: (a less complicated procedure is possible if the characteristic function $\psi_x(t)$ $\psi_y(t)$ does not have level sets).

Let $g(x) = e^{iX}$

\[
D = [0, 1]
\]

\[
\Lambda_n(\omega) = \max(r \in D) ||\psi_{nx}(t_1)\psi_{ny}(t_2)| > 1/2
\]
for \(-r \leq t_1, t_2 \leq r\), (The value 1/2 above could be replaced by any other number in \((0, 1)\)).

\[
A_o(\omega) = 1
\]

\[
C_n = \min_{1 \leq i \leq n} (A_i(\omega))
\]

\[
I_n(\omega) = [-C_n, C_n] \min(|T_{1n}|, |T_{2n}|, |\frac{1}{T_{1n}}|, |\frac{1}{T_{2n}}|)
\]

Due to the monotonicity of \(C_n(\omega)\), there exists intervals \(C(\omega)\) such that

\[
\lim_{n \to \infty} C_n(\omega) = C(\omega)
\]

\(C(\omega)\) is non degenerate almost surely because sample characteristic functions converge uniformly on compact intervals \(\omega, p.1\).

Define \(I(\omega) = \lim_{n \to \infty} I_n(\omega)\).

One may easily verify that expression (1) is bounded away from zero for \((t_1, t_2) \in I(\omega) \times I(\omega)\).

A measure of how far \(b\) is from \(\beta\) is given by \(Z_n(b, \omega)\), which is defined as

\[
Z_n(b, \omega) = \iint_{I_n(\omega) \times I_n(\omega)} \phi_n(t_1, t_2) \nu_X(t_1) \nu_Y(t_2) - \phi_n(bt_2, t_1/b) \nu_X(bt_2) \nu_Y(t_1/b) dt_1 dt_2 \text{ for } T_{1n} \leq b \leq T_{2n}.
\]

For \(T_1 < b < T_2\), \(Z_\infty(b, \omega)\) is defined to be the right hand side of (2) when the \(n\) subscripts are deleted.

The integrand of \(Z_n(b, \omega)\) is an estimate of the integrand of \(Z_\infty(b, \omega)\). If \(g\) is an exponential function, or a cosine function

\[
\frac{\phi(t_1, t_2)}{\psi_X(t_1) \psi_Y(t_2)} = \frac{\psi_U(t_1 + t_2)}{\psi_U(t_2) \psi_U(t_2)} \text{ for } (t_1, t_2) \in I \times I.
\]
By direct calculation, using (3) the integrand of \( Z_n(b, \omega) \) equals

\[
\frac{v_U(t_1 + bt_2)}{v_U(bt_2)} - \frac{v_U(bt_2 + t_1 \frac{\beta}{b})}{v_U(bt_2)} \leq 0
\]

If \( g \) is an exponential function, or \( g \) is a cosine function and \( \log v_U(t) \) is not a quadratic function then expression (4) is identically zero for \((t_1, t_2) \in I \times I\) if and only if \( b = \beta \). The if part follows by direct computation, and the only if part follows from either Lemma 1.5.1 Kagan, Linnik, and Rao (1973), or the methodology of Reiers\o l (1950). It should be noticed that \( Z_n(b, \omega) \) requires no explicit estimate for the distribution function of the random variables \( U, V, \) and \( W \) which are nuisance parameters of this problem. This is a major computational convenience, of the methods given here.

Define \( b_n(\omega) \) as the minimizer of \( Z_n(b, \omega) \) i.e. \( Z_n(b_n(\omega), \omega) = \min_{1 \leq b \leq T_2n} Z_n(b, \omega) \).

If \( Z_n(b, \omega) \) does not have a unique minimum, some inconvenience in the definition of \( b_n(\omega) \) is encountered. If such a case occurs define \( b_n(\omega) \) as the average of the smallest and largest minimizers of \( Z(b, \omega) \), or any other convex combination of the largest and smallest minimizers.

Some additional notation is given before stating the main results of this paper.

Define \( m(a_1, a_2) = \begin{cases} a_1 & \text{if } |a_1| \geq |a_2| \\ a_2 & \text{if } |a_2| \geq |a_1| \end{cases} \)

\[(5) \quad f_n(t_2, t_1, b, \omega) = \phi_n(b t_2, t_1) / m(v_{x_1}(bt_2) v_{x_2}(t_2/b), 1/4)\]

\( f(t_2, t_1, b) \) is defined to be the right hand side of (5) when the subscripts, \( n, \) are deleted.
The partial derivative of any differentiable function \( l(t_1, t_2) \) will be denoted by

\[
\frac{\partial l(s, t_2)}{\partial t_1} = \frac{\partial l(t_1, t_2)}{\partial t_1} \bigg|_{t_1=s}
\]

when no confusion is likely from this notation.

If \( I(\omega) \) is to be estimated, the observations \( (X_i, Y_i) \) \( i = 1, \ldots, n \) are divided into two independent groups, the first group containing \( n_1 = o(n) \) observations, and the second group contains the remainder, \( n_2 \) observations.

\( n_1 + n_2 = n \). The first group is used to estimate \( I(\omega) \), and the second group is used to compute the integrand in \( Z_n(b, \omega) \).

It is useful to think of \( I_{n_1}(\omega) \) and the integrand in \( Z_n(b, \omega) \) to be defined on the probability space \( (\Omega_1 \times \Omega_2, A_1 \times A_2, P_1 \times P_2) \), \( I_{n_2}(\omega) \) is a measurable map from \( \Omega_1 \) to the compact intervals of \([-1, 1]\), and the integrand in \( Z_n(b, \omega) \) is a measurable map from \( \Omega_2 \), to the continuous functions on the interval \([-1, 1]\).

For fixed \( \omega \in \Omega_1 \lim_{n_1 \to \infty} I_{n_1}(\omega) = I(\omega) \).

Let \( I_{n_1}(\omega_1) = [A_{n_1}(\omega_1), B_{n_1}(\omega_1)] \) and \( I(\omega_1) = [A(\omega_1), B(\omega_1)] \).

The first theorem, Theorem 1, which follows states that \( b_n(\omega) \overset{P}{\to} \beta \) as \( n \to \infty \).

**Theorem 1:** If

a) \( g(X) = e^{iX} \)

or

b) \( g(X) = \cos X \) and \( \log E \cos tU \) is not a quadratic function of \( t \) (this is stronger of course than non normality)

or
c) \( g(X) = e^X \) and \( Ee^{t_1X_1 + t_2X_2} < \infty \), for all \( t_1 \) and \( t_2 \), and assumptions 1 and 2 hold, then \( b_n(\omega) \overset{p}{\to} \beta \) as \( n \to \infty \).

The next theorem states that the conditional distribution of \( b_n(\omega) \) given \( I_n(\omega_1) \) is asymptotically normal with mean \( \beta \) and finite variance. Again the possibility for level sets of \( |\Psi_X(t)\Psi_Y(t)| \) made it necessary for the author to state a weaker result, then can be stated if \( |\Psi_X(t)\Psi_Y(t)| \) has no level sets.

**Theorem 2:**

If \( E|X^4 + Y^4| < \infty \), and the assumptions of Theorem 1 hold, then the asymptotic distribution of \( n^{1/2}(b_n(\omega) - \beta) \) given \( A_n(\omega_1), B_n(\omega_1) \) for fixed \( \omega_1 \in \Omega_1 \) is normal mean zero and variance \( \sigma^2(A(\omega_1), B(\omega_1)) < \infty \).

**COMMENT 1.** If \( U \) is not normal, then \( U_1 - U_2 \) is not normal and the characteristic function of \( U_1 - U_2 \) is real. Therefore by pairing observations and working with independent and identically distributed random variables having the same distribution as \( X_1 - X_2 \) and \( Y_1 - Y_2 \) condition b) is satisfied.

**COMMENT 2.** Assumption 2 is not restrictive, because if

\[
\bar{b}_n(\omega) = \begin{cases} 
  b_n(\omega) & \text{if } \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})/\sum_{i=1}^n (X_i - \bar{X})^2 > n^{-1/4} \\
  0 & \text{otherwise.}
\end{cases}
\]

and the conditions of Theorem 2 hold, except that \( \beta \) may equal zero, then \( n^{1/2}(\bar{b}_n(\omega) - \beta) \) is conditionally asymptotically normal mean zero, and variance =

\[
\begin{cases} 
  \sigma^2(A(\omega_1), B(\omega_1)) & \text{if } \beta \neq 0 \\
  0 & \text{if } \beta = 0
\end{cases}
\]
distributed.
Proofs: Theorem 1:

From the construction of $I_n(\omega_1)$,

$$\min_{(t_1, t_2) \in (I_n \times I_n) \cup (I \times I)} |\psi_{nX}(t_1)\psi_{nY}(t_2)| < 1/4 \to 0 \text{ as } n \to \infty.$$ 

Define $Z'_n(b, \omega) = \int \int_{I_n \times I_n} f_n(t_1, t_2, 1, \omega_2) - f_n(t_2, t_1, b, \omega_2)^2 dt_1 dt_2$

and $Z'_\omega(b, \omega) = Z'_\omega(b, \omega)$.

$$\sup_{T_{1n} < b < T_{2n}} |Z'_n - Z'_\omega| = 0$$

as $n \to \infty$. Therefore in order to show

$$\sup_{T_{1n} < b < T_{2n}} |Z_n(b, \omega) - Z_\omega(b, \omega)| \to 0 \text{ as } n \to \infty,$$

it is sufficient to show $\sup_{T_{1n} < b < T_{2n}} |Z'_n(b, \omega) - Z'_\omega(b, \omega)| \to 0 \text{ as } n \to \infty$.

$$\sup_{t_1, t_2 \in I_n \times I_n} E|\phi_n(t_1, t_2) - \phi(t_1, t_2)|^2 = 0(n^{-1}),$$

therefore, for $T_{1n} < b < T_{2n}$, and $t_1, t_2 \in I_n(\omega_1) \cup I(\omega_1)$

$$\sup_{t_1, t_2, b} E|f_n(t_2, t_1b, \omega_2) - f(t_2, t_1, b)|^2 = 0(n^{-1}),$$

and

$$\sup_{t_1, t_2} E|f_n(t_1, t_2, 1, \omega_2) - f(t_1, t_2, 1)|^2 = 0(n^{-1})$$
\begin{align}
&|Z_n'(b, \omega) - Z_n'(b, \omega)| \\
&\leq \int \int_{I_n \times I_n} \left| f_n(t_1, t_2, 1, \omega_2) - f_n(t_2, t_1, b, \omega_2) \right|^2 dt_1 dt_2 \\
&- \left| f(t_1, t_2, 1) - f(t_2, t_1, b) \right|^2 dt_1 dt_2 \\
&+ \int \int_{(I_n \times I_n) \Delta (I_n \times I_n)} \left| f_n(t_1, t_2, 1, \omega_2) - f_n(t_2, t_1, b, \omega_2) \right|^2 + \\
&\left| f(t_1, t_2, 1) - f(t_2, t_1, b) \right|^2 dt_1 dt_2.
\end{align}

By taking the expectation of both sides of (6), using the independence of $I_n(\omega_1)$ from $f_n(\cdot, \cdot, \omega_2)$, Fubini's theorem, and Lebesgue's dominated convergence theorem, both sides of (6) are seen to converge to zero independently of the value of $b$. It is assumed that $b_n(\omega)$ is the unique minimizer of $Z_n(b, \omega)$, if it is not then it is sufficient to note that either the largest or the smallest minimizers of $Z_n(b, \omega)$ may be substituted for $b_n(\omega)$.

Suppose there exists a sequence $\{k_n\}$, $k_n \to \infty$ as $n \to \infty$,

\[b_{k_n}(\omega_1) \to b_0(\omega_1), \quad \text{and} \quad b_0(\omega_1) \neq \beta, \text{ then} \]

\[Z_n'(b_{k_n}(\omega_1), \omega_0) \overset{p}{\to} Z(b_0(\omega_1), \omega_0) \text{ as } n \to \infty, \text{ but} \]

\[Z(b_0(\omega_1), \omega_0) > Z'(\beta, \omega_0) = 0, \text{ as was previously noted.} \]

Proof of Theorem 2:

It is assumed that $b_n(\omega)$ is the unique minimizer of $Z_n(b, \omega)$, if it is not then it is sufficient to note that either the largest or smallest minimizers of $Z_n(b, \omega)$, may be substituted for $b_n(\omega)$.

$Z_n'$ and $Z'$ are defined as in the proof of Theorem 1.

For each $\epsilon > 0$ and each $\omega_1$ in a sure set, there exists a set $\Omega_n \subset \Omega_2$ such that for $n > n_0(\epsilon)$;
\[ \omega \times \Omega_n = \left\{ \omega \left| \frac{\partial Z'_n(b_\omega)}{\partial b} = 0 \right. \right\} \cap \{ \omega \mid Z'_n(b_\omega) = Z_n(b_\omega) \}, \]

and \( P_2(\Omega_2) \geq 1 - \varepsilon. \)

\[ \frac{\partial Z'_n(b_\omega)}{\partial b} \] will expand about the point \( \beta \), for fixed \( \omega \in \omega_1 \times \Omega_n \).

\[ \frac{\partial Z'_n(b_\omega)}{\partial b} - \frac{\partial Z'_n(\beta_\omega)}{\partial b} = \frac{\partial^2 Z_n(b_\omega^*)}{\partial b^2} (b_n - \beta) \]

where \( b_\omega^* \) is between \( \beta \) and \( b_\omega \).

For \( \omega \in \omega_1 \times \Omega_n \)

\[ n^{1/2}(b_\omega - \beta) = -n^{1/2} \frac{\partial Z'_n(\beta_\omega)}{\partial b} \frac{\partial^2 Z_n(b_\omega^*)}{\partial b^2} \]

It is useful to note that for any complex valued function \( \alpha(b) \),

\[ \frac{\partial |\alpha(b)|^2}{\partial b} = 2 \text{Re} \frac{\partial \alpha(b)}{\partial b} \alpha(b). \]

Let \( \alpha(b) \) be the integrand in \( Z_n(b_\omega) \).

\[ \frac{\partial Z_n(\beta_\omega)}{\partial b} = -2 \text{Re} \int_{I_n} \int_{I_n} \frac{\partial f_n(t_2, t_1, \beta_\omega)}{\partial b} \left\{ f_n(t_2, t_1, 1, \omega_2) \right\} \]

\[ -f_n(t_2, t_1, \beta_\omega) \right\} dt_1 dt_2. \]

By arguments similar to those used in the proof of Theorem 1, for \( t_1 \) and \( t_2 \) in \( I_n(\omega_1) \),
\( f_{n_2}(t_2, t_1, \beta, \omega_2) \)
\[ \sup E \left| \frac{\partial f(t_2, t_1, \beta)}{\partial b} - \frac{\partial f(t_2, t_1, \beta)}{\partial b} \right| = 0(n^{-1/2}). \]

Suppose it can be shown that \( f_{n_2}(t_1, t_2, \beta, \omega_2) - f_{n_2}(t_2, t_1, \beta, \omega_2) \) can be written as
\[ \sum_{i=1}^{n_2} \frac{Q_i(t_1, t_2)}{n_2} + T_n \text{ for } \omega_2 \in \Omega_n, \]
where \( Q_i \) are iid random variables, such that \( EQ_i(t_1, t_2) = 0 \) for \( (t_1, t_2) \in \mathbb{R}^2 \) and possessing moments of all order \( \geq 1 \), continuous in \( t_1 + t_2 \), and \( \sup_{(t_1, t_2) \in I_n \times I_n} n^{1/2} T_n \xrightarrow{p} 0 \) as \( n \to \infty \). Then by using 7, Fubini's theorem, and the Cauchy-Schwarz inequality it may be concluded that
\[ n^{1/2} \left| \sum_{i=1}^{n_2} \frac{Q_i(t_1, t_2)}{n_2} \right| dt_1 dt_2 \to 0 \text{ as } n \to \infty. \]

Also notice that
\[ \sup_{(t_1, t_2) \in I_n \times I_n} n^{1/2} |T_n| E \left| \frac{f_{n_2}(t_2, t_1, \beta, \omega_2)}{\partial b} \right| dt_1 dt_2 \to 0 \]
by application of Equation 7 and Fubini's theorem. Therefore
\[ |n^{1/2} \theta b + \sum_{i=1}^{n_2} \frac{\text{Re}}{n_2} \left( \frac{\partial f(t_2, t_1, \beta)}{\partial b} Q_i(t_1, t_2) dt_1 dt_2 \right) \xrightarrow{p} 0 \]
as \( n \to \infty \). Now it is shown that expression (8) has the proper form. Expression (8) equals

\[
(f_{n_2}(t_1, t_2, 1, \omega_2) - \frac{\phi(t_1, t_2)}{\psi_X(t_1)\psi_Y(t_2)} + \frac{\phi(t_1, t_2)}{\psi_X(t_1)\psi_Y(t_2)}) - f_{n_2}(t_2, t_1, \beta, \omega_2) \]

it is sufficient to show that

\[
f_n(t_1, t_2, 1, \omega_2) - \frac{\phi(t_1, t_2)}{\psi_X(t_1)\psi_Y(t_2)} \]

has the desired form, because

\[
\frac{\phi(t_1, t_2)}{\psi_X(t_1)\psi_Y(t_2)} = \frac{\phi(\beta t_2, t_1/\beta)}{\psi_X(\beta t_2)\psi_Y(t_1/\beta)}
\]

(straight forward computation will verify this).

For random variables, \( \phi_{n_2}(t_1, t_2) \), \( \psi_{n_2}X(t_1) \), \( \psi_{n_2}Y(t_2) \), possessing moments of all orders, and \( |\psi_X(t_1)\psi_Y(t_2)| \) bounded above zero one has

\[
\phi_{n_2}(t_1, t_2) - \phi(t_1, t_2) = \frac{\phi_{n_2}(t_1, t_2)}{\psi_X(t_1)\psi_Y(t_2)} + \phi(t_1, t_2) \frac{(\psi_{n_2}X(t_1) - \psi_{n_2}X(t_2))}{\psi_X(t_1)\psi_Y(t_2)} + \phi(t_1, t_2) \frac{(\psi_{n_2}Y(t_1) - \psi_{n_2}Y(t_2))}{\psi_Y(t_2)\psi_X(t_2)} + \phi(t_1, t_2) \frac{(\psi_{n_2}X(t_1) + \psi_{n_2}X(t_2))}{\psi_X(t_1)\psi_Y(t_2)} + o_p \left( \frac{1}{n} \right) \sup_{(t_1, t_2) \in I_n \times I_n} (|g(t_1X_1 + t_2Y_2)|^2)
\]
and statement (8) has been verified.

It remains to show that \( \frac{\partial^2 Z_n(b,\omega)}{\partial b^2} \to c_0 \) as \( n \to \infty \), where \( c_0 \) is a positive constant.

In the fashion of the proof of Theorem 1, it may be shown that

\[
\sup_{T_n < b < T_2n} \left| \frac{\partial^2 Z_n(b,\omega)}{\partial b^2} - \frac{\partial^2 Z'(b,\omega)}{\partial b^2} \right| \to 0 \quad \text{as} \quad n \to \infty.
\]

It is useful to note that for a complex valued function \( \alpha(b) \)

\[
\frac{\partial^2 |\alpha(b)|^2}{\partial b^2} = 2 \left( \frac{\partial \alpha(b)}{\partial b} \right)^2 + 2Re \frac{\partial^2 \alpha(b)}{\partial b^2} \alpha(b).
\]

By letting \( |\alpha(b)|^2 \) be the integrand in \( Z'(b,\omega) \), and noting that

\[
Z'(b,\omega) = 0, \quad \frac{\partial^2 Z'(b,\omega)}{\partial b^2} = \int \int I(\omega_1) \left| \frac{\partial f(t_2, t_1, b)}{\partial b} \right| dt_1 dt_2.
\]

If \( \frac{\partial^2 Z'(b,\omega)}{\partial b^2} = 0 \), then \( \frac{\partial}{\partial b} \left| \log \frac{\psi(b t_2 + \frac{b}{\beta} t_1)}{\psi(b t_2) \psi(t_1)} \right| = 0 \), for all \( (t_1, t_2) \in I_n \times I_n \).

This implies that

\[
(t_2 - \frac{t_1}{\beta}) \bar{\psi}(\beta t_2 + t_1) = \psi(\beta t_2) t_2 + \bar{\psi}(t_1) \left( -\frac{t_1}{\beta} \right)
\]

for all \( (t_1, t_2) \in I_n \times I_n \), where \( \bar{\psi}(\cdot) = \frac{\partial \log \psi(\cdot)}{\partial t} \).

By taking partial derivatives with respect to \( t_1 \), and multiplying both sides of (10) by \( \beta \), one obtains:
(11) \[ -\tilde{\psi}(\beta t_2 + t_1) + (\beta t_2 - t_1) \frac{\partial \tilde{\psi}}{\partial t_1} (\beta t_2 + t_1) \]
\[ = -\tilde{\psi}(t_1) + \frac{\partial \tilde{\psi}(t_1)}{\partial t_1} (-t_1) \]

for all \((t_1, t_2) \in I_n \times I_n\). On the line segment \(\beta t_2 - t_1 = 0\) equation (11) becomes

\[ \tilde{\psi}(2t_1) = \tilde{\psi}(t_1) + t_1 \frac{\partial \tilde{\psi}(t_1)}{\partial t_1}, \]

which implies that a one term Taylor expansion for \(\tilde{\psi}\)
in a neighborhood of the origin is exact. Therefore

\[ \frac{\partial^2 \tilde{\psi}}{\partial t_1^2} = \frac{\partial^3 \log \tilde{\psi}_U(t_1)}{\partial t_1^3} = 0. \]

This implies that \(\log \tilde{\psi}_U(t)\) is a quadratic contrary to hypothesis.

Recall that

(12) \[ n^{1/2}(b_n(\omega) - \beta) + \]

\[ = \sum_{i=1}^{n^2} \frac{\partial f(t_2, t_1, \beta)}{\partial b} \int_{I_n \times I_n} Q_1(t_1, t_2) dt_1 dt_2 \left| \frac{p}{\partial b} \right| \]

as \(n \to \infty\),

\[ \frac{\partial^2 Z'(\beta, \omega)}{\partial b^2} \]

where \(EQ(t_1, t_2) = 0\), and \(Q_1(t_1, t_2)\) possesses all order moments continuous in \(t_1\) and \(t_2\).

It is only to be noted that the second term in expression (12) is asymptotically normal mean zero, and variance

\[ 4E_2 \left[ \frac{\partial f(t_2, t_1, \beta)}{\partial b} \int_{I(\omega_1) \times I(\omega_1)} Q_1(t_1, t_2) dt_1 dt_2 \right]^2 \]

\[ \frac{\partial^2 Z(\beta, \omega)}{\partial b^2} \]
where $E_2$ is the expectation operator corresponding to $P_2$, because

$$E_2 \left[ \sum_{i=1}^{n_2} \frac{\text{Re}}{n_2^{1/2}} (I_n \times I_n) \Delta (I \times I) \int \int \frac{\partial f(t_2, t_1, \beta)}{\partial b} Q_1(t_1, t_2) dt_1 dt_2 \right]^2 + 0$$

as $n \to \infty$, where $\Delta$ denotes the symmetric difference operator, and the central limit theorem applies to the second term of expression (12) if $I_n$ is replaced by $I$.

**COMMENT:** It is to be noted that $\frac{\partial^2 Z(b_n(\omega), \omega)}{\partial b^2}$ is a statistic for determining the normality of $U$.

An example.

In social experiments there is often considerable error in both the independent and dependent variables; this makes the evaluation of social experiments difficult. Many times social evaluators overlook the measurement error in the independent variable and make incorrect recommendations about the continuation of social programs. A discussion of this phenomenon may be found in Campbell and Boruch (1975).

An example follows in which treatment effect is estimable following a modification of Theorem 2.

$X = U + V$ is a pretest, an errorful measurement of true ability, $U$.

$Y = \alpha + \beta U + (\theta_0 + \theta_1 U)Z + W$, where $\theta_0$ and $\theta_1$ are constants, and $Z$ is an observable non degenerate Bernoulli random variable, which is independent of $V$ and $W$, $E(|X|^4 + |Y|^4) < \infty$, and everything else is the same as in Section 2. $Y$ has the trade name of posttest.
This example includes first come first serve, and financial need assignment to treatment. It does not include assignment based on pretest.

Theorem 3: If either the conditional distribution of $U$ given $Z = 1$, or $Z = 0$ is not normal, or the distribution of $V$ and $W$ are not divisible by a normal distribution then $\theta_0$ and $\theta_1$ are identifiable.

Proof: The proof follows from the fact that when $\beta$, or $(\theta_1 + \beta)$ is identifiable then the characteristic functions of $U$, $W$, and $V$ are identifiable in the neighborhood of the origin. See Reiersøl (1950).

If $\alpha + \beta U$ is replaced by a measurable function of $U$, consistent and sometimes optimal estimates of $\theta_0$ are given in Ray and Sacks (1974) under moderate assumptions. Spiegelman (1976) generalizes one of the techniques shown there.

The dependence between $X$ and $Z$ is more general in the last two references than is given here, but it would be hard to verify the assumptions necessary to apply these techniques to the example of this section, if full generality is to be maintained.

A characterization of the normality of the distribution of $U$.

Theorem 4:

If $\beta \neq 0$, $X + X_1$ is independent of $Y - Y_1$ if and only if $U$ has a normal distribution.

COMMENT: This theorem is a quick generalization of a result of Bernstein (1941), which states that $U_1 + U_2$ is independent of $U_1 - U_2$ if and only if $U$ has a normal distribution.
Proof of Theorem 4:

Only if part: \( V + V_1 \) is independent of \( W - W_1 \) and therefore the joint characteristic function of \( X + X_1 \) and \( Y - Y_1 \) equals

\[
E \exp i \beta t_1 (U + U_1) + \beta t_2 (U_1 - U_2) E \exp i(t_1(V + V_1))
\]

\[
E \exp i(t_2(W - W_1)) \text{ which equals}
\]

\[
E \exp i \beta t_1 (U + U_1) E \exp i(t_1(V + V_1)) E \exp i(\beta t_2 (U - U_1))
\]

\[
(U - U_1)) E \exp i(t_2(W - W_1)) \text{ only if}
\]

\[
\log E \exp i(\beta t_1 (U + U_1) + \beta t_2 (U - U_1)) = \log(E \exp i(\beta t_2(U + U_1)) E \exp i(\beta t_2(U - U_1))) \text{ in an open neighborhood of the origin. Following a proof of Reiersøl (or applying Lemma 1.5.1 Kagan, Linnik, and Rao (1973)), } U \text{ has a normal distribution. If Part: this follows by an application of Bernstein's theorem.}
\]

Q.E.D.

COMMENT 1: An appropriate test of independence would be in order, before \( \beta \) is estimated, if there is serious doubt about the distribution of \( U \) being non-normal.

COMMENT 2: If \( V \) and \( W \) are normal, and \( U \) is not, it is a straight forward proof to show that \( X_1 - X_2 \) and \( (Y_1 - bX_1) + (Y_2 - bX_2) \) are independent if and only if \( b = \beta \).

Discussion:

Let \( X = U + V, Y = \beta U + \xi, \xi = aV + W \) where \( U, V, \) and \( W \) are as before. If \( a = 0 \) the model is an errors in variables model; if \( a = \beta \) the correlation model is present. The least squares estimate for \( \beta \) converges to \( \beta \sigma_U^2/(\sigma_U^2 + \sigma_V^2) \) \( \sigma_U^2 \) a.s. if \( a = 0 \), and it converges to \( \beta \) if \( a = \beta \). The estimates given in Section 2 converges in probability to \( \beta \) when either model holds.
It is also noted that Monte Carlo experiments were done for the procedures given here (Spiegelman (1976)), with \( g(X) = e^X \). The procedure given here seems to work reasonably well for small to moderate sized samples (25-50 observations), provided that \( U \) is not too close to normality.

There is a need for good estimators of the effect of social programs. With the exception of the few references, and the methods given here, there is very little statistical work which give general procedures for analyzing data with errors in both the independent and dependent variables.

Acknowledgment: The author wishes to thank Professor Jerome Sacks, his advisor, for suggesting many improvements of earlier versions. He would also like to thank Rose Ray for pointing out several errors in previous versions, Professor Donald T. Campbell for introducing him to this problem, and Professor Jacob Wolfowitz for some helpful comments.

REFERENCES


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


