DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

EXACT AND APPROXIMATE DISTRIBUTIONS OF THE MAXIMUM LIKELIHOOD ESTIMATOR OF A SLOPE COEFFICIENT: THE LIML ESTIMATOR FOR A KNOWN COVARIANCE MATRIX

T. W. Anderson
California Institute of Technology and
Stanford University

Takamitsu Sawa
Kyoto University and
the University of Illinois

SOCIAL SCIENCE WORKING PAPER 326

June 1980

EXACT AND APPROXIMATE DISTRIBUTIONS OF THE MAXIMUM LIKELIHOOD ESTIMATOR OF A SLOPE COEFFICIENT: THE LIML ESTIMATOR FOR A KNOWN COVARIANCE MATRIX*

by

T. W. Anderson** and Takamitsu Sawa***

1. Introduction

In the model of a linear functional relationship with independent normal errors of observation with equal variances the maximum likelihood estimator of the slope is the slope of the line fitted to minimize the sum of squared distances of the observed points from the line. It is mathematically equivalent (under an appropriate transformation) to the Limited Information Maximum Likelihood (LIML) estimator of a structural coefficient when there are two endogenous variables and the covariance matrix of the reduced form is known or known to within a proportionality factor (Anderson [1976]). The exact density of the estimator has been given by Mariano [1969], but it is not in a form that is easy for computation of probabilities. (The exact density of the LIML estimator when the covariance matrix is estimated has been given by Mariano and Sawa [1972], but its form is of necessity more complicated.) In this paper the density is given in another form. A more important purpose of this paper is to express the cumulative distribution function in a form

*This work was supported in part by National Science Foundation Grant SOC77-14944 and SES79-13976 at the Institute for Mathematical Studies in the Social Sciences, Stanford University, and in part by NSF Grant SOC76-22232. The first author was a Sherman Fairchild Distinguished Scholar at the California Institute of Technilogy when this paper was completed.

**Stanford University.

***Kyoto University and the University of Illinois.

The authors are indebted to Naoto Kunitomo, Kimio Morimune, and Kanemi Ban for assistance in this research.

suitable for computation; it is a convergent infinite series of incomplete beta functions when the number of observations on the linear functional relationship is even.

If the number of observations is two (the number of excluded exogenous variables is one), the estimator reduces to a ratio of normal variables. The exact distribution (given by Marsaglia [1965], for example) simply involves univariate and bivariate normal distributions and a univariate normal distribution is an approximation.

Anderson [1974] has given an asymptotic expansion of the cumulative distribution function of the LIML estimator when the covariance matrix is estimated, up to terms of order $-3/2$ power of the noncentrality parameter. (The degrees of freedom in the estimator of the covariance matrix is assumed to be proportional to the noncentrality parameter.) The expansion to terms of this order holds also for the case of the estimate of the covariance matrix being replaced by the matrix itself. The exact distribution of the LIML estimator, say the LIMLK estimator in this latter case,that is obtained in this paper can be used for computations to compare the exact and approximate distributions of the LIMLK estimator.

In this paper an expansion of the distribution is obtained to a term of one higher order. Another asymptotic expansion is based on the doubly noncentral F-distribution. The latter will be shown accurate enough to be regarded as virtually exact, if the noncentrality parameter is moderately large.

In Section 5 a geometric interpretation in $N - 1$ dimensions is given, where N is the number of observations. This discussion throws light on the effect of the parameters on the distributions.

2. The Model and Estimators

Let the pairs (x_g, y_g) be independently normally distributed, each pair with covariance matrix $\sigma^2 \underset{\sim}{I}$, and suppose $\mathcal{E}x_g = \mu_g$ and $\mathcal{E}y_g = \nu_g$, $g = 1,\ldots,N$, where

$$\nu_g = \gamma + \beta\mu_g \quad , \quad g = 1,\ldots,N \quad . \tag{2.1}$$

The maximum likelihood estimator of β is

$$\hat{\beta} = \frac{s_{yy} - s_{xx} + \sqrt{(s_{yy} - s_{xx})^2 + 4s_{xy}^2}}{2s_{xy}} = \frac{2s_{xy}}{s_{xx} - s_{yy} + \sqrt{(s_{xx} - s_{yy})^2 + 4s_{xy}^2}} \quad , \tag{2.2}$$

where

$$s_{xx} = \sum_{g=1}^{N} (x_g - \bar{x})^2 \quad , \quad s_{yy} = \sum_{g=1}^{N} (y_g - \bar{y})^2 \quad , \quad s_{xy} = \sum_{g=1}^{N} (x_g - \bar{x})(y_g - \bar{y}) \, , \tag{2.3}$$

$$\bar{x} = \frac{1}{N}\sum_{g=1}^{N} x_g \quad , \quad \bar{y} = \frac{1}{N}\sum_{g=1}^{N} y_g \quad . \tag{2.4}$$

If θ is the angle between the line (2.1) and the μ-axis, the maximum likelihood estimator of θ is a solution of $\hat{\beta} = \tan\hat{\theta}$. We shall find densities, distributions, and asymptotic expansions of the distributions of $\hat{\beta}$ and of $\hat{\theta}$ when

$$\delta^2 = \frac{1}{\sigma^2}\sum_{g=1}^{N} (\mu_g - \bar{\mu})^2 \tag{2.5}$$

(where $\bar{\mu} = \sum_{g=1}^{N} \mu_g/N$) tends to infinity with N fixed.

It has been shown (Anderson [1976]) that the model in the above paragraph is equivalent to the structural equation model with two endogenous variables and $K_2 = n$ $(= N - 1)$ excluded exogenous variables. The estimator $\hat{\beta}$ here has been called the standardized coefficient $\hat{\alpha}$ and δ^2 the noncentrality parameter in Anderson [1974], [1976] and in Anderson and Sawa [1975], [1977], [1978], [1979].

3. Exact Densities and Distributions

3.1 Densities

If each pair of variables (x_g, y_g) is subjected to a given rotation, the new pair has the same covariance matrix $\sigma^2 \underset{\sim}{I}$ and the distribution of $\hat{\theta} - \theta$ is unchanged. (We identify angles differing by multiples of π in order to avoid defining the interval within which $\hat{\theta}$ and θ are to lie). We can find the distribution or density of $\hat{\beta}$ or $\hat{\theta}$ for $\beta = 0$ and then transform by a rotation to the case of an arbitrary value of β.

Since the estimator $\hat{\beta}$ is homogeneous of degree 0 in (x_g, y_g), it is invariant with respect to scale transformations; hence, there is no loss in generality to take $\sigma^2 = 1$.

The variables s_{yy}, s_{xx}, s_{xy} have the noncentral Wishart distribution (Anderson and Girshick [1944]) with density (when $\beta = 0$)

$$\frac{e^{-\frac{\delta^2}{2}} e^{-\frac{(s_{yy}+s_{xx})}{2}} (s_{yy}s_{xx} - s_{xy}^2)^{\frac{n-3}{2}}}{2^n \pi^{\frac{1}{2}} \Gamma(\frac{n-1}{2})} \sum_{j=0}^{\infty} \frac{[(\frac{\delta}{2})^2 s_{xx}]^j}{j!\Gamma(\frac{n}{2}+j)} \tag{3.1}$$

for $s_{yy} \geq 0$, $s_{xx} \geq 0$, and $s_{xy}^2 \leq s_{yy}s_{xx}$. Let us make the following variate-transformation: $s_{yy} = (1-s)y$, $s_{xy} = ry$, and $s_{xx} = sy$, where $y \geq 0$, $0 \leq s \leq 1$, $-1/2 \leq r \leq 1/2$, and

$$r^2 \leq s(1-s) \quad . \tag{3.2}$$

The Jacobian is y^2. The density of y, s, and r is

$$\frac{e^{-\frac{\delta^2}{2}} [s(1-s) - r^2]^{\frac{n-3}{2}}}{2^n \pi^{\frac{1}{2}} \Gamma(\frac{n-1}{2})} \sum_{j=0}^{\infty} \frac{[(\frac{\delta}{2})^2 s]^j}{j!\Gamma(\frac{n}{2}+j)} e^{-\frac{y}{2}} y^{n+j-1} \quad . \tag{3.3}$$

The marginal density of s and r is

$$\frac{e^{-\frac{\delta^2}{2}} [s(1-s) - r^2]^{\frac{n-3}{2}}}{\pi^{\frac{1}{2}} \Gamma(\frac{n-1}{2})} \sum_{j=0}^{\infty} \frac{\Gamma(n+j)}{j!\Gamma(\frac{n}{2}+j)} (\frac{\delta^2}{2} s)^j \tag{3.4}$$

We can transform to polar coordinates u $(0 \leq u \leq 1/2)$ and $\hat{\psi}$ $(-\pi \leq \hat{\psi} \leq \pi)$ according to $r = u \sin \hat{\psi}$, $s = (1/2) + u \cos \hat{\psi}$; that is,

$$u^2 = r^2 + (s - \tfrac{1}{2})^2 = r^2 + s(s-1) + \tfrac{1}{4} \quad , \tag{3.5}$$

$$\tan \hat{\psi} = \frac{r}{s - \frac{1}{2}} = \frac{2s_{xy}}{s_{xx} - s_{yy}} = \tan 2\hat{\theta} \quad . \tag{3.6}$$

Now the joint density of u and $\hat{\psi}$ is

$$\frac{e^{-\frac{\delta^2}{2}} [\frac{1}{4} - u^2]^{\frac{n-3}{2}}}{\sqrt{\pi}\, \Gamma[\frac{n-1}{2}]} \sum_{j=0}^{\infty} \frac{\Gamma(n+j)}{j!\Gamma(\frac{n}{2}+j)} (\frac{\delta^2}{2})^j \sum_{i=0}^{j} \frac{j!}{i!(j-i)!} \left(\frac{1}{2}\right)^{j-i} u^{i+1} \cos^i \hat{\psi} \tag{3.7}$$

for $0 \leq u \leq 1/2$ and $-\pi \leq \hat{\psi} \leq \pi$. Let $u = v^{1/2}/2$ to obtain a beta integration. Then the joint density of v and $\hat{\psi}$ is

$$\frac{e^{-\frac{\delta^2}{2}} (1-v)^{\frac{n-3}{2}}}{\sqrt{\pi}\, 2^n \Gamma[\frac{n-1}{2}]} \sum_{j=0}^{\infty} \frac{\Gamma(n+j)}{\Gamma(\frac{n}{2}+j)} (\frac{\delta^2}{4})^j \sum_{i=0}^{j} \frac{v^{\frac{i}{2}} \cos^i \hat{\psi}}{i!(j-i)!} \tag{3.8}$$

for $0 \leq v \leq 1$ and $-\pi \leq \hat{\psi} \leq \pi$. The integral of (3.9) with respect to v from 0 to 1 yields the density of $\hat{\psi}$ $(= 2\hat{\theta})$ as

$$\frac{e^{-\frac{\delta^2}{2}}}{\sqrt{\pi}\, 2^n} \sum_{j=0}^{\infty} \frac{\Gamma(n+j)}{\Gamma(\frac{n}{2}+j)} (\frac{\delta^2}{4})^j \sum_{i=0}^{j} \frac{\Gamma[\frac{i}{2}+1] \cos^i \hat{\psi}}{\Gamma[\frac{n+1+i}{2}] i!(j-i)!} \quad . \tag{3.9}$$

Finally, the density of $\hat{\beta} = \tan(\hat{\psi}/2)$ is obtained by substituting into (3.9) $\cos\hat{\psi} = (1-\hat{\beta}^2)/(1+\hat{\beta}^2)$ and multiplying by the Jacobian $2/(1+\hat{\beta}^2)$ to obtain

$$\frac{e^{-\frac{\delta^2}{2}}}{\sqrt{\pi}\,2^{n-1}(1+\hat{\beta}^2)}\sum_{j=0}^{\infty}\frac{\Gamma(n+j)}{\Gamma(\frac{n}{2}+j)}\left(\frac{\delta^2}{4}\right)^j\sum_{i=0}^{j}\frac{\Gamma(\frac{i}{2}+1)}{\Gamma[\frac{n+1+i}{2}]\,i!\,(j-i)!}\,\frac{(1-\hat{\beta}^2)^i}{(1+\hat{\beta}^2)^i} \tag{3.10}$$

for $-\infty < \hat{\beta} < \infty$. Letting $j - i = k$, rearranging the terms, and using the duplication formula for the Gamma function, we can write (3.10) alternatively as

$$\frac{e^{-\frac{\delta^2}{2}}}{\sqrt{\pi}\,(1+\hat{\beta}^2)}\sum_{i=0}^{\infty}\frac{\Gamma(\frac{n+i}{2})}{\Gamma(\frac{n}{2}+i)\Gamma(\frac{i+1}{2})}\left[\frac{\delta^2}{4}\frac{(1-\hat{\beta}^2)}{(1+\hat{\beta}^2)}\right]^i {}_1F_1(n+i,\tfrac{n}{2}+i;\tfrac{\delta^2}{4}) \tag{3.11}$$

where

$${}_1F_1(a,b;x) = \sum_{j=0}^{\infty}\frac{\Gamma(a+j)\Gamma(b)x^j}{\Gamma(b+j)\Gamma(a)j!}\ . \tag{3.12}$$

Incidentally, it can be seen that $E|\hat{\beta}| = \infty$ since $\int_{-\infty}^{\infty}|x|/(1+x^2)dx = \infty$

As noted at the beginning of this section, the distribution of the estimator of the slope for an arbitrary parameter value can be derived from the distribution for parameter value zero, because the distribution of the difference between the estimate of the angle and its parameter value is independent of the parameter value. Let $\hat{\beta}_M = \tan\hat{\theta}_M$ be the maximum likelihood estimator of the slope when $\beta = \tan\theta$ is the parameter (and $\hat{\beta} = \tan\hat{\theta}$ is the estimator when the parameter value is zero). Then $\hat{\theta}_M = \hat{\theta} + \theta$ and

$$\begin{aligned}\hat{\beta}_M - \beta &= \tan\hat{\theta}_M - \tan\theta\\ &= \tan(\hat{\theta}+\theta) - \tan\theta\\ &= \frac{\tan\hat{\theta}+\tan\theta}{1-\tan\hat{\theta}\tan\theta} - \tan\theta\\ &= \frac{\tan\hat{\theta}(1+\tan^2\theta)}{1-\tan\hat{\theta}\tan\theta}\\ &= \frac{(1+\beta^2)\hat{\beta}}{1-\beta\hat{\beta}}\ ,\end{aligned} \tag{3.13}$$

or

$$\hat{\beta} = \frac{\hat{\beta}_M - \beta}{1+\beta\hat{\beta}_M}\ . \tag{3.14}$$

The density function of $\hat{\beta}_M$ is obtained by replacing $\hat{\beta}$ by $(\hat{\beta}_M - \beta)/(1+\beta\hat{\beta}_M)$ in (3.11), multiplying by $(1+\beta^2)/(1+\hat{\beta}_M\beta)^2$, and replacing δ^2 by $\mu^2 = (1+\beta^2)\delta^2$. We obtain

$$\frac{e^{-\frac{\mu^2}{2}}}{\sqrt{\pi}(1+\hat{\beta}_M^2)}\sum_{j=0}^{\infty}\frac{\Gamma(n+j)}{\Gamma(\frac{n}{2}+j)}\left(\frac{\mu^2}{4}\right)^j\sum_{i=0}^{j}\frac{\Gamma(\frac{i}{2}+1)}{\Gamma(\frac{n+1+i}{2})\,i!\,(j-i)!}$$

$$\times\left\{\frac{1-\beta^2+4\beta\hat{\beta}_M-(1-\beta^2)\hat{\beta}_M^2}{(1+\beta^2)(1+\hat{\beta}_M^2)}\right\}^i$$

$$= \frac{e^{-\frac{\mu^2}{2}}}{\sqrt{\pi}(1+\hat{\beta}_M^2)}\sum_{i=0}^{\infty}\frac{\Gamma(\frac{n+i}{2})}{\Gamma(\frac{n}{2}+i)\Gamma(\frac{i+1}{2})}\left\{\frac{\mu^2}{4}\cdot\frac{(1-\beta^2)(1-\hat{\beta}_M^2)+4\beta\hat{\beta}_M}{(1+\beta^2)(1+\hat{\beta}_M^2)}\right\}^i$$

$$\cdot\,{}_1F_1(n+i,\tfrac{1}{2}n+i;\tfrac{\mu^2}{4})\ . \tag{3.15}$$

3.2 Distribution Functions

The inequality $\hat{\beta} > z$ is

$$\frac{2s_{xy}}{s_{xx} - s_{yy} + \sqrt{(s_{xx} - s_{yy})^2 + 4s_{xy}^2}} > z \quad . \tag{3.16}$$

Since the denominator is positive (with probability 1), we write (3.16) as

$$zs_{yy} - zs_{xx} + 2s_{xy} > z\sqrt{(s_{xx} - s_{yy})^2 + 4s_{xy}^2} \quad . \tag{3.17}$$

The inequality (3.17) for $z \geq 0$ is satisfied if and only if

$$z(s_{yy} - s_{xx}) + 2s_{xy} > 0 \tag{3.18}$$

and

$$s_{xy}^2(1 - z^2) > zs_{xy}(s_{xx} - s_{yy}) \quad . \tag{3.19}$$

In terms of the new coordinates, the inequalities (3.18) and (3.19) are

$$r > z(s - \tfrac{1}{2}) \quad , \tag{3.20}$$

$$(1 - z^2)r^2 > 2zr(s - \tfrac{1}{2}) \quad . \tag{3.21}$$

The intersection of the three inequalities (3.2), (3.20), and (3.21) is the shaded area in the Figures 1 and 2, respectively, for $0 < z \leq 1$ and $z > 1$.

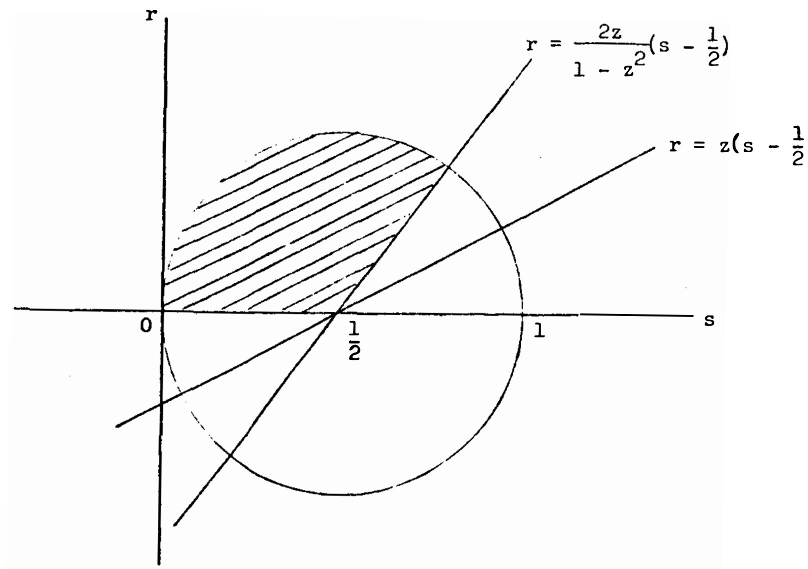


Figure 1: $\{(r,s) \mid \hat{\beta} > z,\ 0 \leq z \leq 1\}$

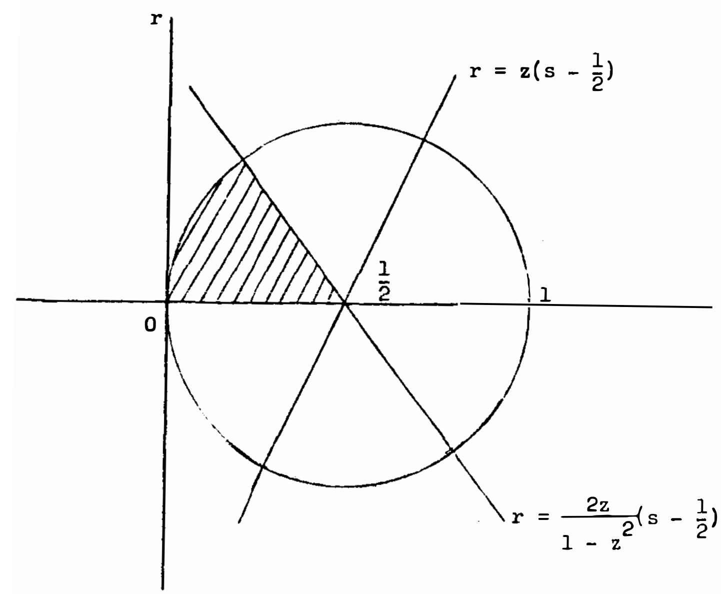


Figure 2: $\{(r,s) \mid \hat{\beta} > z,\ 1 \leq z \leq \infty\}$

Integrating (3.4) with respect to r and s over the shaded region for a fixed value of z, we can evaluate numerically the complement of the distribution function $1 - F(z) = \Pr\{\hat{\beta} > z\}$, $0 \le z < \infty$. By the symmetry of the distribution, the cdf for a negative value of z may be evaluated as $\Pr\{\hat{\beta} < z\} = \Pr\{\hat{\beta} > -z\} = 1 - F(-z)$, $-\infty < z < 0$.

The intersection of (3.2) as an equality and (3.21) as an equality is the point

$$\left(\frac{1}{1+z^2}, \frac{z}{1+z^2}\right) \quad . \tag{3.22}$$

The integral of (3.4) over the region (3.2), $r \ge 0$, and $s \le 1/(1+z^2)$ is

$$\frac{e^{-\delta^2/2}}{\pi^{1/2}\Gamma[(n-1)/2]} \sum_{j=0}^{\infty} \frac{\Gamma(n+j)}{j!\,\Gamma(n/2+j)} \left(\frac{\delta^2}{2}\right)^j \int_0^{1/(1+z^2)} \int_0^{\sqrt{s(1-s)}} [s(1-s) - r^2]^{(n-3)/2} s^j \, dr\, ds$$

$$= \frac{e^{-\delta^2/2}}{2\Gamma(n/2)} \sum_{j=0}^{\infty} \left(\frac{\delta^2}{2}\right)^j \frac{\Gamma(n+j)}{j!\,\Gamma(n/2+j)} \int_0^{1/(1+z^2)} s^{n/2+j-1}(1-s)^{n/2-1} ds$$

$$= \frac{1}{2} e^{-\delta^2/2} \sum_{j=0}^{\infty} \left(\frac{\delta^2}{2}\right)^j \frac{1}{j!} I_{1/(1+z^2)}(n/2+j, n/2) \quad , \tag{3.23}$$

where

$$I_x(p,q) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \int_0^x t^{p-1}(1-t)^{q-1} dt \tag{3.24}$$

is the incomplete beta function. The integral of (3.4) over a triangle $0 \le r \le a(s - 1/2)$ and $1/2 \le s \le b$ for $a > 0$ can be evaluated when n is odd. Let $q = (n-3)/2$. Then the integral is

$$\frac{e^{-\delta^2/2}}{\pi^{1/2}\Gamma(q+1)} \int_{\frac{1}{2}}^{b} \int_0^{a(s-\frac{1}{2})} \sum_{i=0}^{q} \frac{q!}{i!(q-i)!} (-1)^i r^{2i} s^{q-i} (1-s)^{q-i}$$

$$\times \sum_{j=0}^{\infty} \frac{\Gamma(n+j)}{j!\,\Gamma(n/2+j)} \left(\frac{\delta^2}{2} s\right)^j dr\, ds$$

$$= \frac{e^{-\delta^2/2}}{\pi^{1/2}} \int_{\frac{1}{2}}^{b} \sum_{i=0}^{q} \frac{(-1)^i}{i!(q-i)!} \frac{a^{2i+1}}{2i+1} \left(s - \frac{1}{2}\right)^{2i+1} s^{q-i} (1-s)^{q-i} \sum_{j=0}^{\infty} \frac{\Gamma(n+j)}{j!\,\Gamma(n/2+j)} \left(\frac{\delta^2}{2} s\right)^j ds$$

$$= \frac{e^{-\delta^{1/2}}}{\pi^{1/2}} \sum_{i=0}^{q} \frac{1}{i!(q-i)!} \frac{(-1)^i a^{2i+1}}{2i+1} \sum_{k=0}^{2i+1} \frac{(2i+1)!}{k!(2i+1-k)!} \left(-\frac{1}{2}\right)^k \sum_{j=0}^{\infty} \frac{\Gamma(n+j)}{j!\,\Gamma(n/2+j)}$$

$$\times \left(\frac{\delta^2}{2}\right)^j \frac{\Gamma(q+i+j-k+2)\Gamma(q-i+1)}{\Gamma(n+j-k)} [I_b(q+i+j-k+2, q-i+1)$$

$$- I_{1/2}(q+i+j-k+2, q-i+1)] \quad . \tag{3.25}$$

Then $\Pr\{\hat{\beta} > z\}$ for $0 < z \le 1$ is (3.23) less (3.25) for $a = 2z/(1-z^2)$ and $b = 1/(1+z^2)$. Since the integral of (3.4) over a triangle $0 \le r \le a(s - 1/2)$ and $b \le s \le 1/2$ for $a < 0$ is the negative of (3.25), $\Pr\{\hat{\beta} > z\}$ for $1 \le z < \infty$ is (3.23) minus (3.25) for $a = 2z/(1-z^2)$ and $b = 1/(1+z^2)$. Thus we obtain the cdf of $\hat{\beta}$ when $\beta = 0$ and n is odd.

The distribution of $\hat{\beta}_M - \beta$ is the distribution of $(1+\beta^2)\hat{\beta}/(1 - \beta\hat{\beta})$. Hence, for $\mu^2 = \delta^2(1+\beta^2)$ the cdf of the normalized estimator is ($\sqrt{1+\beta^2}/\delta$ is the asymptotic standard deviation)

$$\Pr\left\{\frac{\delta}{\sqrt{1+\beta^2}}(\hat{\beta}_M - \beta) \le t\right\}$$

$$= \Pr\left\{\frac{\mu\hat{\beta}}{1-\beta\hat{\beta}} \le t\right\}$$

$$= \Pr\left\{\hat{\beta} \le \frac{t}{\beta t + \mu}\right\} + \Pr\left\{\frac{1}{\beta} < \hat{\beta}\right\}, \quad \beta(\beta t + \mu) > 0 ,$$

$$= \Pr\left\{\frac{1}{\beta} < \hat{\beta} < \frac{t}{\beta t + \mu}\right\}, \quad \beta(\beta t + \mu) < 0 . \tag{3.26}$$

Alternatively, the density (3. 9) of $\hat{\psi}$ can be used since $\Pr\{\hat{\beta} > z\}$ is the probability of the shaded area in Figure 1 or 2 and hence is $\Pr\{\hat{\psi} > \psi\}$ where $\psi = \arctan 2z/(1 - z^2)$, $0 \le \psi \le \pi$, that is, $\cos\psi = (1 - z^2)/(1 + z^2)$. For $0 \le a \le b \le \pi/2$

$$\int_a^b \cos^i x dx = \frac{1}{2}\frac{\Gamma[\frac{1}{2}(i+1)]\Gamma(\frac{1}{2})}{\Gamma(\frac{1}{2}i+1)}\left\{I_{\cos^2 a}[\tfrac{1}{2}(i+1),\tfrac{1}{2}] - I_{\cos^2 b}[\tfrac{1}{2}(i+1),\tfrac{1}{2}]\right\} , \tag{3.27}$$

and for $\pi/2 \le a \le b \le \pi$ the integral on the left-hand side of (3.27) is $(-1)^{i+1}$ times the right-hand side. The integral of (3.9) on $\hat{\psi}$ from ψ to $\pi/2$ for $0 \le \psi \le \pi/2$ $(0 \le z \le 1)$ is

$$\frac{e^{-\delta^2/2}}{2^{n+1}}\sum_{j=0}^{\infty}\frac{\Gamma(n+j)}{\Gamma(\frac{n}{2}+j)}\left(\frac{\delta^2}{4}\right)^j\sum_{i=0}^{j}\frac{\Gamma[(i+1)/2]}{\Gamma[(n+1+i)/2]i!(j-i)!} \times I_{\cos^2\psi}[(i+1)/2, 1/2] , \tag{3.28}$$

where $\cos\psi = (1 - z^2)/(1 + z^2)$. The integral of (3.9) on $\hat{\psi}$ from ψ to π for $\pi/2 \le \psi \le \pi$ $(1 \le z < \infty)$ is

$$\frac{e^{-\delta^2/2}}{2^{n+1}}\sum_{j=0}^{\infty}\frac{\Gamma(n+j)}{\Gamma(\frac{n}{2}+j)}\left(\frac{\delta^2}{4}\right)^j\sum_{i=0}^{j}\frac{\Gamma[(i+1)/2](-1)^i}{\Gamma[(n+1+i)/2]i!(j-i)!} \times \{1 - I_{\cos^2\psi}[(i+1)/2, 1/2]\} . \tag{3.29}$$

Then $\Pr\{\hat{\beta} > z\}$ for $0 \le z \le 1$ is (3.28) plus (3.29) with $\psi = \pi/2$, that is, $\cos^2\psi = 0$, and { } in (3.29) being 1, and $\Pr\{\hat{\beta} > z\}$ for $1 \le z < \infty$ is (3.29).

4. Asymptotic Expansions of the Distributions

4.1 An Asymptotic Expansion with Five Terms

Anderson [1974] gave an asymptotic expansion of the Limited Information Maximum Likelihood (LIML) estimator. As shown by Anderson [1976], the distribution of the LIML estimator is the distribution of the maximum likelihood estimator of the slope of the linear functional relationship if the covariance matrix of error is estimated by use of a sample covariances matrix distributed independently of (x_g, y_g), $g = 1,\ldots,N$. The asymptotic expansion to terms of order μ^{-3}, where $\mu^2 = \delta^2(1 + \beta^2)$, is valid when the estimator of the covariance matrix is replaced by the covariance matrix itself. (In the proof of Anderson [1974] $\underset{\sim}{X}$ can be replaced by $\underset{\sim}{0}$ since the expectations with respect to $\underset{\sim}{X}$ in that proof are o.) Hence it holds for the LIMLK estimator.

The method of Anderson [1974] will be modified to give the expansion to order μ^{-4}. It will be convenient to treat first the case of $\beta = 0$. Let

$$W = \tan 2\hat{\theta} = \frac{2s_{xy}}{s_{xx} - s_{yy}} \quad . \tag{4.1}$$

The distribution of s_{xx}, s_{yy}, and s_{xy} is that of

$$s_{xx} = \sum_{g=2}^{n} u_g^2 + (u_1 + \delta)^2 \ , \quad s_{yy} = \sum_{g=1}^{n} v_g^2 \ , \quad s_{xy} = \sum_{g=2}^{n} u_g v_g + (u_1 + \delta)v_1 \ , \tag{4.2}$$

where the random variables $u_1, v_1, u_2, v_2, \ldots, u_n, v_n$ are independently normally distributed with means 0 and variances 1 (Anderson [1974], for example). The cumulative distribution function of δW is

$$\Pr\{\delta W \leq x\} = \Pr\{2\delta s_{xy} \leq x(s_{xx} - s_{yy})\} + O(e^{-\delta}) \tag{4.3}$$

because $\Pr\{s_{xx} \leq s_{yy}\} = O(e^{-\delta})$ as $\delta \to \infty$. The first term on the right-hand side of (4.3) can be written

$$\begin{aligned} &\Pr\{xv_1^2 + 2\delta(u_1+\delta)v_1 + 2\delta \sum_{g=2}^{n} u_g v_g - x[\delta^2 + 2\delta u_1 + \sum_{g=1}^{n} u_g^2] + x\sum_{g=2}^{n} v_g^2 \leq 0\} \\ &= \Pr\Bigg\{v_1 \leq \frac{1}{x}\Bigg[-(\delta^2 + \delta u_1) + \delta^2\Bigg(1 + \frac{1}{\delta}2u_1 + \frac{1}{\delta^2}\{x^2 + u_1^2\} \\ &\quad + \frac{1}{\delta^3}\{2x^2 u_1 - 2x\sum_{g=2}^{n} u_g v_g\} + \frac{1}{\delta^4}\{x^2 \sum_{g=1}^{n} u_g^2 - x^2 \sum_{g=2}^{n} v_g^2\}\Bigg)^{\frac{1}{2}}\Bigg]\Bigg\} + O(e^{-\delta}) \\ &= \mathcal{E}\Phi\Big\{\frac{1}{x}[-(\delta^2 + \delta u_1) + \delta^2(1 + \frac{1}{\delta}2u_1 + \frac{1}{\delta^2}\{x^2 + u_1^2\} \\ &\quad + \frac{1}{\delta^3}\{2x^2 u_1 - 2x\sum_{g=2}^{n} u_g v_g\} + \frac{1}{\delta^4}\{x^2 \sum_{g=1}^{n} u_g^2 - x^2 \sum_{g=2}^{n} v_g^2\})^{\frac{1}{2}}]\Big\} + O(e^{-\delta}) \ , \end{aligned} \tag{4.4}$$

where the expectation is taken with respect to the random variables $u_1, u_2, v_2, \ldots, u_n, v_n$. The argument of Φ in (4.4) can be expanded (details given in Anderson [1975]) to yield

$$\Pr\{\delta W \leq x\} = \mathcal{E}\Phi(\tfrac{1}{2}x + \Delta) + O(\delta^{-5}) \ , \tag{4.5}$$

where

$$\begin{aligned} \Delta &= \frac{1}{\delta}[\tfrac{1}{2}xu_1 - \sum_{g=2}^{n} u_g v_g] + \frac{1}{\delta^2}[-\frac{x^3}{8} + \tfrac{1}{2}x(\sum_{g=2}^{n} u_g^2 - \sum_{g=2}^{n} v_g^2) + u_1 \sum_{g=2}^{n} u_g v_g] \\ &\quad + \frac{1}{\delta^3}[-\frac{x^3 u_1}{8} - \tfrac{1}{2}xu_1(\sum_{g=2}^{n} u_g^2 - \sum_{g=2}^{n} v_g^2) + (\tfrac{1}{2}x^2 - u_1^2)\sum_{g=2}^{n} u_g v_g] + \frac{1}{\delta^4}[\frac{x^5}{16} \\ &\quad - (\frac{x^3}{4} - \frac{xu_1^2}{2})(\sum_{g=2}^{n} u_g^2 - \sum_{g=2}^{n} v_g^2) - (\tfrac{1}{2}x^2 u_1 - u_1^3)\sum_{g=2}^{n} u_g v_g - \tfrac{1}{2}x(\sum_{g=2}^{n} u_g v_g)^2] \ . \end{aligned} \tag{4.6}$$

The next step is to expand $\Phi[(1/2)x + \Delta]$ in a Taylor's series around the value $(1/2)x$ to obtain

$$\Phi(\tfrac{1}{2}x + \Delta) = \Phi(\tfrac{1}{2}x) + \phi(\tfrac{1}{2}x)\{\Delta - \frac{x}{4}\Delta^2 + (\frac{x^2}{24} - \frac{1}{6})\Delta^3 - (\frac{x^3}{192} - \frac{x}{16})\Delta^4\} + O(\Delta^5) \tag{4.7}$$

The expected values of the powers of Δ are

$$\mathcal{E}\Delta = -\frac{1}{\delta^2}\frac{x^3}{8} + \frac{1}{\delta^4}[\frac{x^5}{16} - \tfrac{1}{2}(n-1)x] \ , \tag{4.8}$$

$$\mathcal{E}\Delta^2 = \frac{1}{\delta^2}[\frac{x^2}{4} + n - 1] + \frac{1}{\delta^4}[\frac{x^6}{64} - \frac{x^4}{8} + 3(n-1)] + O(\delta^{-5}) \ , \tag{4.9}$$

$$\mathcal{E}\Delta^3 = \frac{1}{\delta^4}[-\frac{3}{32}x^5 - \frac{3}{8}(n-1)x^3 - 3(n-1)x] + O(\delta^{-5}) \quad , \qquad (4.10)$$

$$\mathcal{E}\Delta^4 = \frac{1}{\delta^4}[\frac{3}{16}x^4 + \frac{3}{2}(n-1)x^2 + 3(n^2-1)] + O(\delta^{-5}) \quad . \qquad (4.11)$$

From this we obtain

$$\begin{aligned}\Pr\{\delta W \le x\} &= \Pr\{\delta \tan 2\hat{\theta} \le x\} \\ &= \Phi(\tfrac{1}{2}x) + \phi(\tfrac{1}{2}x)\{-\frac{1}{\delta^2}[\frac{3}{16}x^3 + \frac{n-1}{4}x] \\ &\quad + \frac{1}{\delta^4}[-\frac{9}{1024}x^7 - \frac{6n-37}{256}x^5 - \frac{(n-1)^2}{64}x^3 \\ &\quad + \frac{3}{16}(n-1)(n-3)x]\} + O(\delta^{-5}) \quad . \qquad (4.12)\end{aligned}$$

We find the asymptotic expansion of the distribution of $\delta\hat{\theta}$,

$$\Pr\{\delta\hat{\theta} \le z\} = \Pr\{\delta \tan 2\hat{\theta} \le \delta \tan \frac{2z}{\delta}\} \quad , \qquad (4.13)$$

by substituting $\delta \tan (2z/\delta)$ into (4.12) for x and expanding in a Taylor's series to obtain (see Anderson [1975])

$$\begin{aligned}\Pr\{\delta\hat{\theta} \le z\} &= \Phi(z) - \phi(z)\Big\{\frac{1}{\delta^2}[\frac{z^3}{6} + \frac{n-1}{2}z] + \frac{1}{\delta^4}[\frac{z^7}{72} + \frac{10n-11}{120}z^5 \\ &\quad + \frac{(n-1)(3n+13)}{24}z^3 - \frac{3(n-1)(n-3)}{8}z]\Big\} + O(\delta^{-6}) \quad . \qquad (4.14)\end{aligned}$$

The error is $O(\delta^{-6})$ because the term of order δ^{-5} is a polynomial with only even powers of z and the distribution of $\hat{\theta}$ is symmetric (Anderson [1976]).

We find the asymptotic expansion of the distribution of $\delta\hat{\beta} = \delta \tan \hat{\theta}$,

$$\Pr\{\delta\hat{\beta} \le w\} = \Pr\{\delta \tan \hat{\theta} \le w\} = \Pr\{\delta\hat{\theta} \le \delta \arctan \frac{w}{\delta}\} \quad , \qquad (4.15)$$

by substituting $\delta \arctan (w/\delta)$ into (4.14) and expanding in a Taylor's series to obtain

$$\begin{aligned}\Pr\{\delta\hat{\beta} \le w\} &= \Phi(w) - \phi(w)\{\frac{1}{2\delta^2}[w^3 + (n-1)w] + \frac{1}{8\delta^4}[w^7 + (2n-5)w^5 \\ &\quad + (n-1)(n+3)w^3 - 3(n-1)(n-3)w]\} + O(\delta^{-6}) \quad . \qquad (4.16)\end{aligned}$$

Now consider the model with an arbitrary slope coefficient β. Denote the maximum likelihood estimator in this general case by $\hat{\beta}_M$. Then the distribution of $\hat{\beta}_M - \beta$ is the distribution of $(1+\beta^2)\hat{\beta}/(1-\beta\hat{\beta})$. Hence

$$\begin{aligned}\Pr\left\{\frac{\delta}{\sqrt{1+\beta^2}}(\hat{\beta}_M - \beta) \le t\right\} &= \Pr\left\{\frac{\mu\hat{\beta}}{1-\beta\hat{\beta}} \le t\right\} \\ &= \Pr\left\{\mu\hat{\beta} \le \frac{t}{1+\frac{\beta}{\mu}t}\right\} + O(e^{-\delta}) \quad . \qquad (4.17)\end{aligned}$$

The left-handside can be evaluated to within an error of order $O(\delta^{-6})$ by replacing w in (4.16) by $t/(1+\beta t/\mu)$. An asymptotic expansion in this general case may be obtained by replacing w in (4.16) by the Taylor's series approximation for $t/(1+\beta t/\mu)$. and expanding the function as a Taylor' series in $1/\delta$. We obtain

$$\Pr\left\{\frac{\delta}{\sqrt{1+\beta^2}}(\hat{\beta}_M - \beta) \leq t\right\} = \Phi(t) - \phi(t)\Big\{\frac{\beta}{\mu}t^2 + \frac{1}{2\mu^2}[\beta^2 t^5 + (1-2\beta^2)t^3 + (n-1)t] + \frac{\beta}{6\mu^3}[\beta^2 t^8 + (3-7\beta^2)t^6 + (6\beta^2 + 3n - 12)t^4 - 3(n-1)t^2] + \frac{1}{24\mu^4}[\beta^4 t^{11} + 3(2\beta^2 - 5\beta^4)t^9 + 3(16\beta^4 + (2n-20)\beta^2 + 1)t^7 - 3(8\beta^4 + (10n-34)\beta^2 + (2n-5))t^5 + 3(n-1)(4\beta^2 + n + 3)t^3 - 9(n-1)(n-3)t]\Big\} + O(\mu^{-5}) \tag{4.18}$$

The term of order μ^{-5} could be obtained by using the Taylor's series for $t/(1 + \beta t/\mu)$ to one more term, but the result is too complicated.

4.2 The Doubly-Noncentral F-distribution

Another approach is the use of the doubly-noncentral F-distribution. There is no loss in generality in assuming that $\beta \geq 0$. In Section 3.2 we showed that $\hat{\beta} > z$ (> 0) if and only if the inequalities (3.18) and (3.19) simultaneously hold. Now we want to argue for $\beta \geq 0$ that as $\delta \to \infty$ the difference between the probability of (3.18) and (3.19) and that of

$$(1 - z^2)s_{xy} > z(s_{xx} - s_{yy}) \tag{4.19}$$

is $O(e^{-\delta})$ for $0 < z$ and hence can be neglected asymptotically. Let

$$s_{xx} = (u_1 + \delta)^2 + \sum_{g=2}^{n} u_g^2 \;, \quad s_{yy} = (v_1 + \beta\delta)^2 + \sum_{g=2}^{n} v_g^2$$

$$s_{xy} = (u_1 + \delta)(v_1 + \beta\delta) + \sum_{g=2}^{n} u_g v_g \;. \tag{4.20}$$

(See Section 5.) Then $\operatorname{plim}_{\delta\to\infty} \hat{\beta} = \beta$. Thus we need only to consider z arbitrarily close to β and hence can confine z to positive values.

We first consider the case when $\beta = 0$. By (4.20) we can see that $\Pr\{s_{xx} - s_{yy} > 0\} = 1 - O(e^{-\delta})$ if $\beta = 0$. Then (4.19) for $0 < z < 1$ asymptotically implies $s_{xy} > 0$, and hence (4.19) is asymptotically equivalent to (3.19) and also implies (3.18) with probability $1 - O(e^{-\delta})$. Conversely, (3.18) implies $s_{xy} > 0$ and hence (3.19) is asymptotically equivalent to (4.19). Therefore, (3.18) and (3.19) are asymptotically equivalent to (4.19).

Next we consider the case where $\beta > 0$. In this case $\Pr\{s_{xy} > 0\} = 1 - O(e^{-\delta})$ by use of (4.20). Along with $s_{xy} > 0$, (4.19) implies and is implied by (3.18) and (3.19). Therefore, (4.19) is asymptotically equivalent to (3.18) and (3.19). The consequence of the above reasoning is that $\hat{\beta} > z$ is asymptotically equivalent to (4.19) for $0 < z$ and $\beta \geq 0$.

We now consider the distribution of

$$(1 - z^2)s_{xy} - zs_{xx} + zs_{yy} \tag{4.21}$$

when s_{xy}, s_{xx}, and s_{yy} have the distribution implied by (4.20). The quantity (4.21) is a quadratic form in $2n$ variables and can be diagonalized by an orthogonal transformation. The matrix of the quadratic form is

$$\begin{bmatrix} -z\underset{\sim}{I} & \frac{1}{2}(1 - z^2)\underset{\sim}{I} \\ \frac{1}{2}(1 - z^2)\underset{\sim}{I} & z\underset{\sim}{I} \end{bmatrix} \tag{4.22}$$

with characteristic roots $\lambda_1 = (1/2)(1 + z^2)$ and $\lambda_2 = -(1/2)(1 + z^2)$ each with multiplicity n.

There exists an orthogonal matrix $\underset{\sim}{L}$ such that

$$\underset{\sim}{L}' \begin{bmatrix} -z\underset{\sim}{I} & \frac{1}{2}(1 - z^2)\underset{\sim}{I} \\ \frac{1}{2}(1 - z^2)\underset{\sim}{I} & z\underset{\sim}{I} \end{bmatrix} \underset{\sim}{L} = \begin{bmatrix} \lambda_1\underset{\sim}{I} & \underset{\sim}{0} \\ \underset{\sim}{0} & \lambda_2\underset{\sim}{I} \end{bmatrix} . \tag{4.23}$$

Then (4.21) has the distribution of $\lambda_1 V_1 + \lambda_2 V_2$, where V_1 and V_2 have independent noncentral χ^2-distribution with n degrees of freedom and noncentrality parameters δ_1 and δ_2, respectively. We have

$$\delta_1 + \delta_2 = \delta^2(1 + \beta^2) \quad , \tag{4.24}$$

$$\begin{aligned} \lambda_1\delta_1 + \lambda_2\delta_2 &= (\delta \quad \beta\delta) \begin{bmatrix} -z & \frac{1}{2}(1 - z^2) \\ \frac{1}{2}(1 - z^2) & z \end{bmatrix} \begin{bmatrix} \delta \\ \beta\delta \end{bmatrix} \\ &= \delta^2 \{-z + z\beta^2 + (1 - z^2)\beta\} \\ &= \delta^2 \{z(\beta^2 - 1) + \beta(1 - z^2)\} \quad . \end{aligned} \tag{4.25}$$

Then

$$\begin{aligned} \delta_1 &= \delta^2\left[\frac{1}{2}(1 + \beta^2) + \frac{(\beta z + 1)(\beta - z)}{1 + z^2}\right] \quad , \\ \delta_2 &= \delta^2\left[\frac{1}{2}(1 + \beta^2) - \frac{(\beta z + 1)(\beta - z)}{1 + z^2}\right] \quad . \end{aligned} \tag{4.26}$$

The inequality $\hat{\beta} \leq z$ is equivalent to $\lambda_1 V_1 + \lambda_2 V_2 \leq 0$ which is equivalent to $V_1 - V_2 \leq 0$, or $V_1 \leq V_2$ or $V_1/V_2 \leq 1$. The ratio V_1/V_2 has the doubly noncentral F-distribution with n and n degrees of freedom and δ_1 and δ_2 as noncentrality parameters.

We want $\Pr\{\delta(\hat{\beta}_M - \beta)/\sqrt{1 + \beta^2} \leq t\}$. Hence we let $z = \beta + t\sqrt{1 + \beta^2}/\delta$. Then

$$\begin{aligned} \delta_1 &= \mu^2\left[\frac{1}{2} - \frac{t(\mu + \beta t)}{(t^2 + (\mu + \beta t)^2)}\right] \quad , \\ \delta_2 &= \mu^2\left[\frac{1}{2} + \frac{t(\mu + \beta t)}{(t^2 + (\mu + \beta t)^2)}\right] \quad . \end{aligned} \tag{4.27}$$

It should be noted that the present approximation is legitimate over the region $t > -\delta\beta/\sqrt{1 + \beta^2}$, since $\hat{\beta} > z$ was shown to be equivalent to (4.19) only for the range $0 < z$.

The cdf of the doubly-noncentral F distribution is represented as a doubly-infinite series. The straightforward computation, based on this representation, not only requires prohibitively long computational time but also causes difficulty to control the accuracy when either the noncentrality parameters or the degrees of freedom or both are large. Therefore, we need one more step of approximation to develop a workable computational program for evaluating the cdf of the LIMLK estimator. Our approximation is based on the fact that the cube root of a chi-square variate is well approximated by the Gram-Charlier expansion about the normal distribution. That is, we replace the probability $\Pr\{V_1 - V_2 \leq 0\}$, which is asymptotically equal to $\Pr\{\hat{\beta} \leq t\}$, by $\Pr\{V_1^{1/3} - V_2^{1/3} \leq 0\}$, where V_1 and V_2 are independently distributed as noncentral χ^2. Expanding the cdf of $V_1^{1/3} - V_2^{1/3}$ as a Gram-Charlier series up to the fourth order, we can evaluate the latter probability with little computational time, but enough accuracy. For the details of the Gram-Charlier expansion of the cube root of the noncentral chi-square variate the reader is referred to Anderson and Sawa [1979] or Mudholkar, Chaubey, and Lin [1976].

5. A Geometric Interpretation

The N-dimensional vectors $\underset{\sim}{x} = (x_1,\ldots,x_N)'$, $\underset{\sim}{y} = (y_1,\ldots,y_N)'$, $\underset{\sim}{\mu} = (\mu_1,\ldots,\mu_N)'$, and $\underset{\sim}{\nu} = (\nu_1,\ldots,\nu_N)'$ can be transformed (see Anderson [1976]) by an orthogonal matrix to $(\underset{\sim}{x}^{*\prime}, \sqrt{N}\,\bar{x})'$, $(\underset{\sim}{y}^{*\prime}, \sqrt{N}\,\bar{y})'$, $(\underset{\sim}{\mu}^{*\prime}, \sqrt{N}\,\bar{\mu})'$, and $(\underset{\sim}{\nu}^{*}, \sqrt{N}\,\bar{\nu})'$, respectively. The vectors of $n = N-1$ dimensions satisfy

$$s_{xx} = \underset{\sim}{x}^{*\prime}\underset{\sim}{x}^{*} \quad , \quad s_{xy} = \underset{\sim}{x}^{*\prime}\underset{\sim}{y}^{*} \quad , \quad s_{yy} = \underset{\sim}{y}^{*\prime}\underset{\sim}{y}^{*} \quad , \tag{5.1}$$

$$\delta^2 = \frac{\underset{\sim}{\mu}^{*\prime}\underset{\sim}{\mu}^{*}}{\sigma^2} \quad . \tag{5.2}$$

Then $\underset{\sim}{\nu}^* = \beta\underset{\sim}{\mu}^*$ lies on the line through the origin and $\underset{\sim}{\mu}^*$. If β is positive, the length of $\underset{\sim}{\nu}^*$ is β times the length of $\underset{\sim}{\mu}^*$. Under normality $\underset{\sim}{x}^*$ has a spherical normal distribution in the n-dimensional space with center at $\underset{\sim}{\mu}^*$, and $\underset{\sim}{y}^*$ has such a distribution centered at $\underset{\sim}{\nu}^*$.

To obtain the least squares estimator of β the vector $\underset{\sim}{y}^*$ is projected on the vector $\underset{\sim}{x}^*$; the estimator is the signed length of this projection.

Consider a vector $\underset{\sim}{z} = b_x\underset{\sim}{x}^* + b_y\underset{\sim}{y}^*$, which is in the plane of $\underset{\sim}{x}^*$ and $\underset{\sim}{y}^*$. The sum of the squared distances of $\underset{\sim}{x}^*$ and $\underset{\sim}{y}^*$ to $\underset{\sim}{z}$ is

$$\left[\underset{\sim}{x}^* - \frac{\underset{\sim}{x}^{*\prime}\underset{\sim}{z}}{\underset{\sim}{z}'\underset{\sim}{z}}\underset{\sim}{z}\right]'\left[\underset{\sim}{x}^* - \frac{\underset{\sim}{x}^{*\prime}\underset{\sim}{z}}{\underset{\sim}{z}'\underset{\sim}{z}}\underset{\sim}{z}\right] + \left[\underset{\sim}{y}^* - \frac{\underset{\sim}{y}^{*\prime}\underset{\sim}{z}}{\underset{\sim}{z}'\underset{\sim}{z}}\underset{\sim}{z}\right]'\left[\underset{\sim}{y}^* - \frac{\underset{\sim}{y}^{*\prime}\underset{\sim}{z}}{\underset{\sim}{z}'\underset{\sim}{z}}\underset{\sim}{z}\right]$$
$$= \underset{\sim}{x}^{*\prime}\underset{\sim}{x}^* + \underset{\sim}{y}^{*\prime}\underset{\sim}{y}^* - \frac{(\underset{\sim}{x}^{*\prime}\underset{\sim}{z})^2 + (\underset{\sim}{y}^{*\prime}\underset{\sim}{z})^2}{\underset{\sim}{z}'\underset{\sim}{z}} \quad . \tag{5.3}$$

Minimizing (5.3) is equivalent to maximizing

$$\frac{(\underset{\sim}{x}^{*\prime}\underset{\sim}{z})^2 + (\underset{\sim}{y}^{*\prime}\underset{\sim}{z})^2}{\underset{\sim}{z}'\underset{\sim}{z}} = \frac{(b_x\underset{\sim}{x}^{*\prime}\underset{\sim}{x}^* + b_y\underset{\sim}{y}^{*\prime}\underset{\sim}{x}^*)^2 + (b_x\underset{\sim}{x}^{*\prime}\underset{\sim}{y}^* + b_y\underset{\sim}{y}^{*\prime}\underset{\sim}{y}^*)^2}{(b_x\underset{\sim}{x}^* + b_y\underset{\sim}{y}^*)'(b_x\underset{\sim}{x}^* + b_y\underset{\sim}{y}^*)}$$
$$= \frac{\underset{\sim}{b}'\underset{\sim}{S}^2\underset{\sim}{b}}{\underset{\sim}{b}'\underset{\sim}{S}\,\underset{\sim}{b}} \quad , \tag{5.4}$$

where $\underset{\sim}{b} = (b_x\ b_y)'$ and the symmetric matrix $\underset{\sim}{S}$ has elements s_{xx}, $s_{xy} = s_{yx}$ and s_{yy}. A vector $\underset{\sim}{b}$ maximizing (5.4) maximizes $\underset{\sim}{b}'\underset{\sim}{S}\,\underset{\sim}{b}/\underset{\sim}{b}'\underset{\sim}{b}$. Thus $\underset{\sim}{b}$ satisfies

$$\underset{\sim}{S}\,\underset{\sim}{b} = \lambda_M\underset{\sim}{b} \quad , \tag{5.5}$$

where λ_M is the maximum characteristic root of $\underset{\sim}{S}$. If a vector $\underset{\sim}{b}$ satisfying (5.5) is $\underset{\sim}{b}^M = (b_x^M\ b_y^M)'$, then the ratio of the signed length of the projection of $\underset{\sim}{y}^*$ on $\underset{\sim}{z}$ to that of $\underset{\sim}{x}^*$ is

$$\frac{\underset{\sim}{y}^{*\prime}\underset{\sim}{z}}{\underset{\sim}{x}^{*\prime}\underset{\sim}{z}} = \frac{\underset{\sim}{y}^{*\prime}\underset{\sim}{x}^*b_x^M + \underset{\sim}{y}^{*\prime}\underset{\sim}{y}^*b_y^M}{\underset{\sim}{x}^{*\prime}\underset{\sim}{x}^*b_x^M + \underset{\sim}{x}^{*\prime}\underset{\sim}{y}^*b_y^M} = \frac{\lambda_M b_y^M}{\lambda_M b_x^M} = \frac{b_y^M}{b_x^M} \quad . \tag{5.6}$$

Since $\underset{\sim}{b}^M$ is orthogonal to $\underset{\sim}{b}$ minimizing $\underset{\sim}{b}'S\underset{\sim}{b}/\underset{\sim}{b}'\underset{\sim}{b}$, the ratio (5.6) is $\hat{\beta}$ given by (2.2).

In the method of maximum likelihood both $\underset{\sim}{x}^*$ and $\underset{\sim}{y}^*$ are used to estimate the direction of $\underset{\sim}{\mu}^*$. Then the estimator of β is the ratio of the lengths of the projections on a line with this direction.

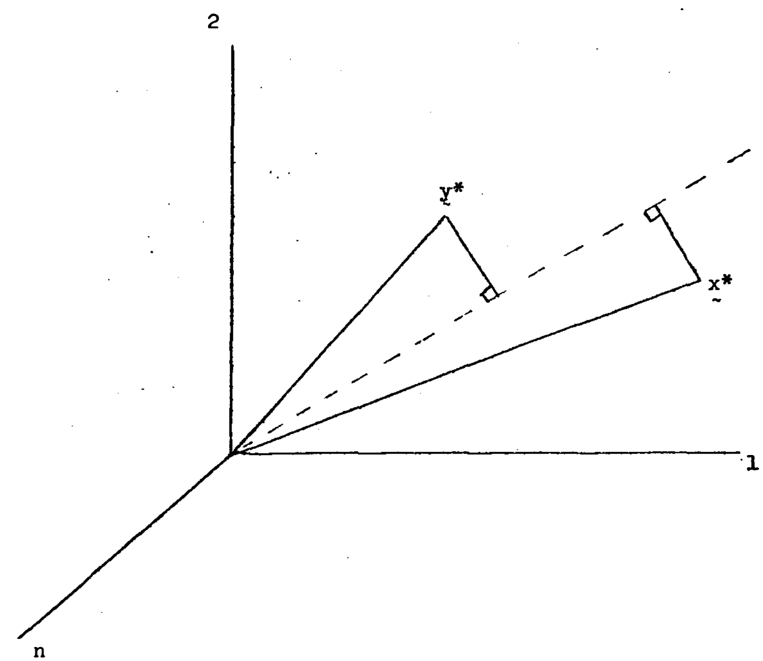
There is an $n \times n$ orthogonal matrix that carries $\underset{\sim}{x}^*$ to $\underset{\sim}{X}$, $\underset{\sim}{y}^*$ to $\underset{\sim}{Y}$, $\underset{\sim}{\mu}^*$ to $(\sigma\delta,0,\ldots,0)'$ and $\underset{\sim}{\nu}^*$ to $(\beta\sigma\delta,0,\ldots,0)$. The components of $\underset{\sim}{X}$ and $\underset{\sim}{Y}$ are independently normally distributed with variances σ^2 and means 0 except that the first component of each vector $\mathcal{E}X_1 = \sigma\delta$ and $\mathcal{E}Y_1 = \beta\sigma\delta$. Only the distributions of X_1 and Y_1 depend on β and δ. This fact suggests that if $\underset{\sim}{X}$ and $\underset{\sim}{Y}$ were known, the best estimate of β would be based on X_1 and Y_1; the other components of $\underset{\sim}{X}$ and $\underset{\sim}{Y}$ give information solely about σ^2. In turn the implication is that for given δ the estimator of β for $n = 1$ would be more accurate than for $n > 1$ and that the accuracy decreases as n increases.

As δ increases, the leading term in s_{xx} is $X_1^2 = \sigma^2(u_1 + \delta)^2$, and the leading term in s_{xy} is $X_1Y_1 = \sigma^2(u_1 + \delta)v_1$ when $\beta = 0$. Thus $\hat{\beta}$ (2.2) is approximately v_1/δ; as $\delta \to \infty$, $\delta\hat{\beta}$ has a limiting normal distribution.

6. Tables

Anderson and Sawa [1978] have given tables of the cdf of $\delta(\hat{\beta}_M - \beta)/\sqrt{1 + \beta^2}$ for β = 0.0, 0.2, 0.5, 1.0, 2.0, 5.0, and 10.0, n = 3, 7, 10, 20, 30, δ^2 = 40, 100, 300, 1000, 3000 and 8000, at values of the argument -4.0, -3.0, -2.5, -2.0(.2)-1.0(.1)1.0(.2)2.0, 2.5, 3.0, 4.0.

Figure 3



Anderson and Sawa [1975] have given tables of the cdf for $\beta = 0$, $n = 1, 3, 7, 11$, $\delta^2 = 2, 4, 10, 20, 40$, at values 0.1(.1)1.0(.2)2.0(.5)7.0, for $\beta = 0.2, 0.5, 1.0, 2.0, 5.0$, $n = 1, 3, 7$, $\delta^2 = 2, 4, 10, 20, 40$, at values -7.0(1.0)-2.0(.5)2.0(1.0)7.0, and for $\beta = 0.0, 0.2, 0.5, 1.0, 2.0, 5.0$, $n = 1$, $\delta^2 = 2, 4, 6, 8, 10, 15$ at values -7.0(.5)-2.0(.2) -1.0(.1)1.0(.2)2.0(.5)7.0. They also give some tables of the errors of approximation using various numbers of terms of the asymtotic expansion in Section 4.1.

References

Anderson, T.W. [1974], "An Asymptotic Expansion of the Distribution of the Limited Information Maximum Likelihood Estimate of a Coefficient in a Simultaneous Equation System," Journal of the American Statistical Association, 69, 565-573.

Anderson, T.W. [1975], "An Asymptotic Expansion of the Distribution of the Maximum Likelihood Estimate of the Slope Coefficient in a Linear Functional Relationship," Technical Report No. 176, Economics Series, Institute for Mathematical Studies in the Social Sciences (IMSSS), Stanford University

Anderson, T.W. [1976], "Estimation of Linear Functional Relationships: Approximate Distributions and Connections with Simultaneous Equations in Econometrics," Journal of the Royal Statistical Society B, 38, 1-36.

Anderson, T.W. and M.A. Girshick [1944], "Some Extensions of the Wishart Distribution," Annals of Mathematical Statistics, 15, 345-375.

Anderson, T.W. and Takamitsu Sawa [1975], "Distribution of a Maximum Likelihood Estimate of a Slope Coefficient: The LIML Estimate for Known Covariance Matrix," Technical Report No. 174, Economics Series, IMSSS, Stanford University.

Anderson, T.W. and Takamitsu Sawa [1977], "Tables of the Distribution of the Maximum Likelihood Estimate of the Slope Coefficient and Approximations," Technical Report No. 234, Economics Series, IMSSS, Stanford University.

Anderson, T.W. and Takamitsu Sawa [1978], "Tables of the Exact Distribution Function of the Limited Information Maximum Likelihood Estimator when the Covariance Matrix is Known," Technical Report No. 275, Economics Series, IMSSS, Stanford University.

Anderson, T.W. and Takamitsu Sawa [1979], "Evaluation of the Distribution Function of the Two-Stage Least Squares Estimate." Econometrica, 47, 163-182.

Mariano, Roberto Sanchez [1969], "On Distributions and Moments of Single-Equation Estimators in a Set of Simultaneous Linear Stochastic Equations," Technical Report No. 27, IMSSS, Stanford University.

Mariano, Roberto Sanchez and Takamitsu Sawa [1972], "The Exact Finite-Sample Distribution of the Limited Information Maximum Likelihood Estimator in the Case of Two Included Endogenous Variables," Journal of the American Statistical Association, 67, 159-163.

Marsaglia, George [1965], "Ratios of Normal Variables and Ratios of Sums of Uniform Variables," Journal of the American Statistical Association, 60, 193-204.

Mudholkar, Govind S., Yogendra P. Chaubey, and Ching-Choung Lin [1976], "Some Approximations for the Noncentral F-distribution," Technometrics, 18, 351-358.