LIKELIHOOD INFERENCE FOR LINEAR REGRESSION MODELS

TECHNICAL REPORT NO. 21

T. J. DICICCIO

NOVEMBER 1987

U. S. ARMY RESEARCH OFFICE
CONTRACT DAAG29-85-K-0239
THEODORE W. ANDERSON, PROJECT DIRECTOR

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

LIKELIHOOD INFERENCE FOR LINEAR REGRESSION MODELS

T. J. DiCiccio
Stanford University

TECHNICAL REPORT NO. 21
November 1987

U. S. Army Research Office
Contract DAAG29-85-K-0239
Theodore W. Anderson, Project Director

Department of Statistics
Stanford University
Stanford, California

1. Introduction

Consider observed random variables $Y_1,\dots,Y_n$ which follow a linear regression model; that is, let $Y_i = \mu_i + \sigma E_i$, where

$$\mu_i = \sum_{j=1}^{p} \beta_j x_{ij} \quad (i=1,\dots,n)$$

and $E_1,\dots,E_n$ are independent and identically distributed with known probability density function f. The vectors $x_i = (x_{i1},\dots,x_{ip})^T$ $(i=1,\dots,n)$ of covariate values are assumed to be known, and the vector $\beta = (\beta_1,\dots,\beta_p)$ of regression coefficients is to be estimated. Based on the sample $Y = (Y_1,\dots,Y_n)$, the log likelihood function for β and $\theta = \log \sigma$ is

$$l(\beta,\theta;Y) = -n\theta + \sum_{i=1}^{n} \log f\{e^{-\theta}(Y_i - \beta x_i)\} .$$

Several authors, including Fisher (1934), Fraser (1979), and Verhagen (1961), have argued that inferences concerning the parameters β and σ should be made conditionally on the sample configuration $A = (A_1,\dots,A_n)$ consisting of the standardized residuals $A_i = (Y_i-\hat{\beta}x_i)/\hat{\sigma}$ $(i=1,\dots,n)$, where $\hat{\beta}$ and $\hat{\sigma}$ are the maximum likelihood estimators of β and σ. The joint conditional density of the pivotal statistics $P_1 = (\beta-\hat{\beta})/\hat{\sigma}$ and $P_2 = \theta-\hat{\theta}$ given the configuration A is determined by

$$f_{P_1,P_2|A}(p_1,p_2) \propto \exp\{l(\hat{\beta}+\hat{\sigma}p_1, \hat{\theta}+p_2;Y)\} \propto \exp\{l(p_1,p_2;A)\} \quad . \qquad (1)$$

If the scale parameter is known and has the value 1, then a similar expression obtains. In that case, the conditional density of $P = \beta - \hat{\beta}$ given the configuration $A = (A_1, \ldots, A_n)$ of residuals $A_i = Y_i - \hat{\beta} x_i$ $(i=1,\ldots,n)$ satisfies

$$f_{P|A}(p) \propto \exp\{l(\hat{\beta}+p;Y)\} \propto \exp\{l(p;A)\} , \qquad (2)$$

where $l(\beta;Y)$ is the log likelihood function for β based on Y.

The use of (1) or (2) for exact conditional inference usually requires numerical integration over several dimensions, which can be cumbersome. Accurate approximate methods for inference that require less computation are therefore of interest, and log likelihood ratio statistics are a possible basis for such methods.

Suppose that q of the parameters $\beta_1, \ldots, \beta_p, \theta$ are of interest. The number of nuisance parameters is $p - q + 1$ or $p - q$, depending on whether the scale parameter is unknown or not. To simplify the notation, let $\phi = (\omega_1, \ldots, \omega_q)$ consist of the parameters of interest, let ψ consist of the nuisance parameters, and let $l(\omega)$ be the log likelihood function for $\omega = (\phi, \psi)$ based on Y. Fixing the value of ϕ, the log likelihood ratio statistic for that parameter is

$$W(\phi) = 2\{l(\hat{\omega}) - l(\phi, \tilde{\psi}_\phi)\} ,$$

where $\hat{\omega}$ is the unconstrained maximum likelihood estimator of ω and $(\phi, \tilde{\psi}_\phi)$ is the restricted maximum likelihood estimator of ω corresponding to the specified value of ϕ. For the case $q = 1$ in which a scalar parameter is of interest, the signed square root of the log likelihood ratio statistic for $\phi = \omega_1$ is defined by

$$R(\omega_1) = \mathrm{sgn}(\omega_1 - \hat{\omega}_1)\{W(\omega_1)\}^{1/2} .$$

It is shown in the usual large-sample maximum likelihood theory that the marginal distributions of W and R tend under mild conditions to the chi-squared distribution with q degrees of freedom χ_q^2 and the standard normal distribution $N(0,1)$ respectively, as the sample size n increases. Hinkley (1978) has shown that these limits also hold for the conditional distributions of W and R.

The standard normal approximation to the conditional distribution of R has error of order $O_p(n^{-1/2})$, and the approximate confidence limits thus obtained are correct to that order. The error of the $N(0,1)$ approximation can be reduced to order $O_p(n^{-1})$ by taking the conditional mean of R into account, and this error can be further reduced to $O_p(n^{-3/2})$ by adjusting for both the conditional mean and variance of R. Formulae for these mean and variance adjustments are given in section 2. Such corrections to the signed roots of log likelihood ratio statistics for scalar parameters have been discussed very generally by Barndorff-Nielsen (1986) and in the case of location-scale models by DiCiccio (1987).

The chi-squared approximation to the conditional distribution of W has error of order $O_p(n^{-1})$, and this error can be reduced to order $O_p(n^{-3/2})$ by the use of a scaling factor which accounts for the conditional mean of W. A formula for this adjustment factor is given in section 3. Such corrections for log likelihood ratio statistics, often referred to as Bartlett adjustments, have been discussed very generally by Barndorff-Nielsen and Cox (1984).

2. Approximate Inference for a Scalar Parameter

Adopting the notation of Sprott (1980), let

$$I_{ab} = [-\partial^2 l(\omega)/\partial\omega_a \partial\omega_b]_{\omega=\hat{\omega}} \; , \quad I_{abc} = [\partial^3 l(\omega)/\partial\omega_a \partial\omega_b \partial\omega_c]_{\omega=\hat{\omega}} \; ,$$

$$I_{abcd} = [\partial^4 l(\omega)/\partial\omega_a \partial\omega_b \partial\omega_c \partial\omega_d]_{\omega=\hat{\omega}} \; .$$

Thus $I = ((I_{ab}))$ is the observed information matrix for ω, and $I^{-1} = ((I^{ab}))$ is its inverse. In the expressions that follow, it is convenient to make use of the notation convention for which summation over every index appearing both as a subscript and a superscript is understood. The range of summation is from 1 to $p+1$ if the scale parameter is unknown and from 1 to p otherwise.

Sprott (1980) has shown that the Taylor expansion of $W(\omega_1)$ about $\hat{\omega}_1$ is

$$W(\omega_1) = U^2 - \frac{1}{3} AU^3 - \frac{1}{12} BU^4 + o_p(n^{-3/2}) \; , \qquad (3)$$

where

$$U = (\omega_1 - \hat{\omega}_1)/(I^{11})^{1/2} , \quad A = I_{abc} I^{a1} I^{b1} I^{c1}/(I^{11})^{3/2} ,$$

$$B = (I_{abcd} I^{a1} I^{b1} I^{c1} I^{d1} + 3 S_a S_b J^{ab})/(I^{11})^2 ,$$

$$S_a = I_{abc} I^{b1} I^{c1} , \quad J^{ab} = I^{ab} - I^{a1} I^{b1}/I^{11} .$$

Note that A is $O_p(n^{-1/2})$ and B is $O_p(n^{-1})$. It follows that $R(\omega_1)$ has the expansion

$$R(\omega_1) = U - \frac{1}{6} AU^2 - \frac{1}{72} (3B + A^2) U^3 + O_p(n^{-3/2}) . \tag{4}$$

Using calculations similar to those described by Hinkley (1978), expressions (1) and (2) allow the conditional cumulants of R to be determined with error of order $O_p(n^{-3/2})$. Ignoring terms of that order, the conditional mean and variance of R are

$$\begin{aligned} m &= \frac{1}{2} I_{abc} I^{a1} I^{bc}/(I^{11})^{1/2} - \frac{1}{6} I_{abc} I^{a1} I^{b1} I^{c1}/(I^{11})^{3/2} \\ &= \frac{1}{6} I_{abc} I^{a1} (2I^{bc} + J^{bc})/(I^{11})^{1/2} , \end{aligned} \tag{5}$$

$$\begin{aligned} s^2 = 1 + \frac{1}{4} I_{abcd}(I^{ab} I^{cd} - J^{ab} J^{cd}) + \frac{1}{12} I_{abc} I_{def} \{ 3(I^{ab} I^{cd} I^{ef} - J^{ab} J^{cd} J^{ef}) \\ + 2(I^{ad} I^{be} I^{cf} - J^{ad} J^{be} J^{cf}) \} - m^2 , \end{aligned}$$

and the higher-order conditional cumulants of R are zero. Note that m is of order $O_p(n^{-1/2})$ and involves second- and third-order derivatives, whereas s^2 is $1+O_p(n^{-1})$ and involves derivatives up to the fourth order. Since the $N(0,1)$ approximation to the conditional distribution of $R-m$ has error of order $O_p(n^{-1})$ and can be used to find approximate confidence limits correct to that order, it is thus possible to improve the accuracy of the usual large-sample approximation by taking third-order derivatives into account. The $N(0,1)$ approximation for $(R-m)/s$ has error of order $O_p(n^{-3/2})$, but the improvement in accuracy achieved by using s requires fourth-order derivatives and a substantial increase in computation.

When testing a single hypothesized value of ω_1, the approximate procedures based on R are not difficult to implement; however, when setting approximate confidence limits for ω_1, these procedures can be complicated by the necessity of calculating the restricted maximum likelihood estimate of ω for each of several values of ω_1. By inverting expansion (4) and ignoring terms of order $O_p(n^{-3/2})$, the α quantile of $U=(\omega_1-\hat{\omega}_1)/(I^{11})^{1/2}$ is found to be

$$r_\alpha + \frac{1}{6}Ar_\alpha^2 + \frac{1}{72}(3B+5A^2)r_\alpha^3 ,$$

where $r_\alpha = m + sz_\alpha$ and z_α is the α quantile of the $N(0,1)$ distribution. Thus, correct to order $O_p(n^{-3/2})$, the upper α confidence limit for ω_1 is

$$\hat{\omega}_1 + (I^{11})^{1/2}\{r_\alpha + \frac{1}{6}Ar_\alpha^2 + \frac{1}{72}(3B+5A^2)r_\alpha^3\} . \qquad (6)$$

Expression (6) provides the same order of accuracy as the direct use of the $N(0,1)$ approximation for $(R-m)/s$ in setting approximate confidence limits, and it is usually much simpler to apply. However, in very small samples, expression (6) may suffer from failure of monotonicity and produce inaccurate results. In such cases, the use of $(R-m)/s$ is preferable.

It is evident from (6) that correct to order $O_p(n^{-1})$ the upper α confidence limit for ω_1 is

$$\hat{\omega}_1 + (I^{11})^{1/2}\{(z_\alpha+m) + \frac{1}{6} A(z_\alpha+m)^2\} . \tag{7}$$

For general scalar parameter models, Cox (1980) and McCullagh (1984) have derived expansions of conditional confidence limits correct to order $O_p(n^{-1})$, and Barndorff-Nielsen (1985, 1986) has derived a similar expansion correct to order $O_p(n^{-3/2})$. In the present regression context, if there are no nuisance parameters present, expansion (6) is equivalent to Barndorff-Nielsen's approximate limit, and (7) is equivalent to the limits of Cox and McCullagh.

The methods discussed in this section can be applied for approximate inference concerning the α quantile y_α of an observation Y corresponding to a specified set of covariate values $x=(x_1,\ldots,x_p)^T$. This application is achieved by formulating the model so that y_α appears as one of the regression coefficients. For instance, if β_1 is an intercept term and $x_{11}=\ldots=x_{n1}=x_1=1$, then $y_\alpha=\beta_1+\beta_2x_2+\ldots+\beta_px_p+\sigma e_\alpha$, where e_α is the α quantile of E_1. The regression model can be written as $Y_i = y_\alpha + \beta_2(x_{i2}-x_2) + \ldots + \beta_p(x_{ip}-x_p) + \sigma(E_i-e_\alpha)$ $(i=1,\ldots,n)$, with y_α replacing β_1. Using this formulation of the model, the mean and variance adjustments to the signed root of the log likelihood ratio statistic for y_α can be derived from expression (5), and expressions (6) and (7) provide approximate confidence limits.

Example 1. In location-scale models, for which $p = 1$ and $\mu_i=\mu$ $(i=1,\ldots,n)$, the accuracy of the approximate methods can feasibly be assessed by comparison with exact conditional results. Fraser (1979, p. 26) has presented a location-scale analysis of Darwin's data assuming that the error variables have Student's distribution with λ degrees of freedom. Darwin's data consists of 15 observations. Using likelihood methods, Fraser concluded that values of λ in the range 1 to 9 are well supported by the data. Fraser's analysis includes the one-sided significance level for the hypothesis $\mu = 0$ and 95% confidence intervals for μ and σ in each of the cases $\lambda = 1,3,6,9,$ and ∞. For these values of λ, the exact significance levels are 0.041%, 0.300%, 0.765%, 1.114%, and 2.485% respectively; the levels obtained from the $N(0,1)$ approximation for $(R-m)/s$ are 0.034%, 0.310%, 0.774%, 1.102%, and 2.437%. Table 1 shows the approximate 95% confidence intervals for μ and σ obtained from $(R-m)/s$ and expression (6). Each endpoint of the approximate intervals is accompanied by its true conditional significance level determined by numerical integration. Both methods give fairly accurate approximations, and similar accuracy is obtained for the approximate intervals of other confidence levels. Sprott (1980, 1982) has described a different approach to approximate conditional inference in location-scale models, and he also considered Darwin's data.

3. Bartlett Adjustments

Suppose that the q-dimensional vector $\phi = (\omega_1, \ldots, \omega_q)$ is of interest, and let

$$I = \begin{bmatrix} I_{\phi\phi} & I_{\phi\psi} \\ I_{\psi\phi} & I_{\psi\psi} \end{bmatrix}$$

be the partitioned observed information matrix for $\omega = (\phi, \psi)$. By using an expansion of $W(\phi)$ about $\hat{\phi}$ in conjunction with (1) and (2), it can be shown that to error of order $O_p(n^{-3/2})$ the conditional expectation of the log likelihood ratio statistic for ϕ is

$$b_\phi = q + \frac{1}{4} I_{abcd}(I^{ab}I^{cd} - K^{ab}K^{cd}) + \frac{1}{12} I_{abc}I_{def}\{3(I^{ab}I^{cd}I^{ef} - K^{ab}K^{cd}K^{ef}) + 2(I^{ad}I^{be}I^{cf} - K^{ad}K^{be}K^{cf})\}, \tag{8}$$

where $K = ((K^{ab}))$ is defined by

$$K = \begin{bmatrix} 0 & 0 \\ 0 & I_{\psi\psi}^{-1} \end{bmatrix}.$$

The χ_q^2 approximation to the conditional distribution of $W(\phi)$ has error of order $O_p(n^{-1})$, while the error in the approximation for $(b_\phi/q)^{-1}W(\phi)$ is of order $O_p(n^{-3/2})$. Dividing $W(\phi)$ by (b_ϕ/q) produces a quantity whose conditional distribution is better approximated by the χ_q^2 distribution. In the case $q=1$, $K^{ab} = I^{ab} - I^{a1}I^{b1}/I^{11} = J^{ab}$ and the adjustment (8) equals $s^2 + m^2$, where s^2 and m are as defined in (5).

Example 2. Consider the location-scale model of Darwin's data described in Example 1, and let $\lambda = 3$. The approximate 90%, 95%, and 99% confidence intervals for μ determined using the Bartlett adjustment are (12.93, 40.34), (9.72, 43.24), and (2.49, 49.47), which have true conditional coverage probabilities 89.9%, 94.9%, and 99.0%. The approximate intervals for σ are (15.38, 38.83), (14.15, 42.78), and (12.03, 52.01), having true conditional confidence levels 89.9%, 94.9%, and 99.0%. The use of the Bartlett adjustments produces very accurate approximations in this situation.

4. A Summary of Derivatives

This section presents formulae for the calculation of derivatives of the log likelihood function for β and θ evaluated at the maximum likelihood estimators. The second-, third-, and fourth-order derivatives are

$$\hat{l}_{\beta_a\beta_b} = \hat{\sigma}^{-2} \; \Sigma x_{ia}x_{ib}g^{(2)}(A_i) \; ,$$

$$\hat{l}_{\beta_a\theta} = \hat{\sigma}^{-1} \; \Sigma x_{ia}A_i g^{(2)}(A_i) \; ,$$

$$\hat{l}_{\theta\theta} = -n + \Sigma A_i^2 g^{(2)}(A_i) \; ,$$

$$\hat{l}_{\beta_a\beta_b\beta_c} = -\hat{\sigma}^{-3} \; \Sigma x_{ia}x_{ib}x_{ic}g^{(3)}(A_i) \; ,$$

$$\hat{l}_{\beta_a\beta_b\theta} = -\hat{\sigma}^{2} \; \Sigma x_{ia}x_{ib}\{2g^{(2)}(A_i) + A_i g^{(3)}(A_i)\} \; ,$$

$$\hat{l}_{\beta_a\theta\theta} = -\hat{\sigma}^{-1}\Sigma x_{ia}\{3A_i g^{(2)}(A_i) + A_i^2 g^{(3)}(A_i)\} \; ,$$

$$\hat{l}_{\theta\theta\theta} = n - \Sigma\{3A_i^2 g^{(2)}(A_i) + A_i^3 g^{(3)}(A_i)\} \; ,$$

$$\hat{l}_{\beta_a\beta_b\beta_c\beta_d} = \hat{\sigma}^{-4} \; \Sigma x_{ia}x_{ib}x_{ic}x_{id}g^{(4)}(A_i) \; ,$$

$$\hat{l}_{\beta_a\beta_b\beta_c\theta} = \hat{\sigma}^{-3} \; \Sigma x_{ia}x_{ib}x_{ic}\{3g^{(3)}(A_i) + A_i g^{(4)}(A_i)\} \; ,$$

$$\hat{l}_{\beta_a\beta_b\theta\theta} = \hat{\sigma}^{-2} \; \Sigma x_{ia}x_{ib}\{4g^{(2)}(A_i) + 5A_i g^{(3)}(A_i) + A_i^2 g^{(4)}(A_i)\} \; ,$$

$$\hat{l}_{\beta_a\theta\theta\theta} = \hat{\sigma}^{-1}\ \Sigma x_{ia}\{7A_i g^{(2)}(A_i) + 6A_i^2 g^{(3)}(A_i) + A_i^3 g^{(4)}(A_i)\}\ ,$$

$$\hat{l}_{\theta\theta\theta\theta} = -n + \Sigma\{7A_i^2 g^{(2)}(A_i) + 6A_i^3 g^{(3)}(A_i) + A_i^4 g^{(4)}(A_i)\}\ ,$$

where

$$\hat{l}_{\beta_a\beta_b} = [\partial^2 l(\beta,\theta;Y)/\partial\beta_a\partial\beta_b]_{(\beta,\theta)=(\hat{\beta},\hat{\theta})}\ ,$$

$$\hat{l}_{\beta_a\theta} = [\partial^2 l(\beta,\theta;Y)/\partial\beta_a\partial\theta]_{(\beta,\theta)=(\hat{\beta},\hat{\theta})}\ ,$$

etc. $(a,b,c,d = 1,\ldots,p)$, $g = \log f$, and each sum is taken over $i = 1,\ldots,n$.

Acknowledgements

The helpful comments of a referee are gratefully acknowledged. This work was supported by the Natural Sciences and Engineering Research Council of Canada.

References

Barndorff-Nielsen, O. E. (1985). Confidence limits from $c|\hat{j}|^{1/2}\bar{L}$ in the one-parameter case. Scand. J. Statist. 12, 83-87.

Barndorff-Nielsen, O. E. (1986). Inference on full or partial parameters based on the standardized signed log likelihood ratio. Biometrika 73, 307-22.

Barndorff-Nielsen, O. E. and Cox, D. R. (1984). Bartlett adjustments to the likelihood ratio statistic and the distribution of the maximum likelihood estimator. J. R. Statist. Soc. B 46, 483-95.

Cox, D. R. (1980). Local ancillarity. Biometrika 67, 279-86.

DiCiccio, T. J. (1987). Approximate inference for the generalized gamma distribution. Technometrics 29, 33-40.

Fisher, R. A. (1934). Two new properties of mathematical likelihood. Proc. R. Soc. A 144, 285-307.

Fraser, D.A.S. (1979). Inference and Linear Models. New York: McGraw-Hill.

Hinkley, D. V. (1978). Likelihood inference about location and scale parameters. Biometrika 65, 253-61.

McCullagh, P. (1984). Local sufficiency. Biometrika 71, 233-41.

Sprott, D. A. (1980). Maximum likelihood in small samples: Estimation in the presence of nuisance parameters. Biometrika 67, 515-23.

Sprott, D. A. (1982). Robustness and maximum likelihood estimation. Comm. Statist. A 11, 2513-29.

Verhagen, A.M.W. (1961). The estimation of regression and error-scale parameters when the joint distribution of the errors is of any continuous form and known apart from a scale parameter. Biometrika 48, 125-32.

Table 1. Approximate conditional 95% confidence intervals for μ and σ from Darwin's data

	μ		σ	
λ	(R-m)/s	(6)	(R-m)/s	(6)
1	13.16,41.08	12.62,41.70	8.11,34.06	8.12,33.88
	(2.72,2.24)	(2.28,1.92)	(2.65,2.54)	(2.66,2.63)
3	9.43,42.95	9.83,43.14	15.16,45.07	15.16,45.10
	(2.48,2.54)	(2.71,2.43)	(2.44,2.56)	(2.45,2.54)
6	5.79,43.01	5.83,42.99	19.50,50.99	19.49,50.90
	(2.48,2.61)	(2.50,2.62)	(2.45,2.60)	(2.44,2.64)
9	4.12,42.85	4.12,42.75	21.69,53.51	21.68,53.43
	(2.52,2.59)	(2.52,2.65)	(2.43,2.62)	(2.42,2.65)
∞	0.13,41.73	0.22,41.65	27.59,59.29	27.58,59.24
	(2.55,2.55)	(2.59,2.59)	(2.44,2.59)	(2.43,2.62)

The true conditional one-sided significance levels of the interval endpoints are shown as percentages in parentheses.

Technical Reports

U.S. Army Research Office - Contracts DAAG29-82-K-0156 and DAAG29-85-K-0239

1. "Maximum Likelihood Estimators and Likelihood Ratio Criteria for Multivariate Elliptically Contoured Distributions," T. W. Anderson and Kai–Tai Fang, September 1982.
2. "A Review and Some Extensions of Takemura's Generalizations of Cochran's Theorem," George P.H. Styan, September 1982.
3. "Some Further Applications of Finite Difference Operators," Kai–Tai Fang, September 1982.
4. "Rank Additivity and Matrix Polynomials," George P.H. Styan and Akimichi Takemura, September 1982.
5. "The Problem of Selecting a Given Number of Representative Points in a Normal Population and a Generalized Mills' Ratio," Kai–Tai Fang and Shu–Dong He, October 1982.
6. "Tensor Analysis of ANOVA Decomposition," Akimichi Takemura, November 1982.
7. "A Statistical Approach to Zonal Polynomials," Akimichi Takemura, January 1983.
8. "Orthogonal Expansion of Quantile Function and Components of the Shapiro–Francia Statistic," Akimichi Takemura, January 1983.
9. "An Orthogonally Invariant Minimax Estimator of the Covariance Matrix of a Multivariate Normal Population," Akimichi Takemura, April 1983.
10. "Relationships Among Classes of Spherical Matrix Distributions," Kai–Tai Fang and Han–Feng Chen, April 1984.
11. "A Generalization of Autocorrelation and Partial Autocorrelation Functions Useful for Identification of ARMA(p,q) Processes," Akimichi Takemura, May 1984.
12. "Methods and Applications of Time Series Analysis Part II: Linear Stochastic Models," T. W. Anderson and N. D. Singpurwalla, October 1984.
13. "Why Do Noninvertible Estimated Moving Averages Occur?" T. W. Anderson and Akimichi Takemura, November 1984.
14. "Invariant Tests and Likelihood Ratio Tests for Multivariate Elliptically Contoured Distributions," Huang Hsu, May 1985.
15. "Statistical Inferences in Cross–lagged Panel Studies," Lawrence S. Mayer, November 1985.
16. "Notes on the Extended Class of Stein Estimators," Suk–ki Hahn, July 1986.
17. "The Stationary Autoregressive Model," T. W. Anderson, July 1986.
18. "Bayesian Analyses of Nonhomogeneous Autoregressive Processes," T. W. Anderson, Nozer D. Singpurwalla, and Refik Soyer, September 1986.
19. "Estimation of a Multivariate Continuous Variable Panel Model," Lawrence S. Mayer and Kun Mao Chen, June 1987.
20. "Consistency of Invariant Tests for the Multivariate Analysis of Variance," T. W. Anderson and Michael D. Perlman, October 1987.
21. "Likelihood Inference for Linear Regression Models," T. J. DiCiccio, November 1987.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 21	2. GOVT ACCESSION NO. N/A	3. RECIPIENT'S CATALOG NUMBER N/A
4. TITLE (and Subtitle) LIKELIHOOD INFERENCE FOR LINEAR REGRESSION MODELS		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) T. J. DiCiccio		8. CONTRACT OR GRANT NUMBER(s) DAAG29-85-K-0239
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics - Sequoia Hall Stanford University Stanford, CA 94305-4065		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office Post Office Box 12211 Research Triangle Park, NC 27709		12. REPORT DATE November 1987
		13. NUMBER OF PAGES 15pp
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

NA

18. SUPPLEMENTARY NOTES

The view, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Asymptotic expansion; Bartlett adjustment; Conditional inference; Confidence limit; Likelihood ratio; Maximum likelihood.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

(See reverse side for abstract.)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

20. ABSTRACT.

Approximate conditional inference based on large-sample likelihood ratio methods is considered for the parameters of linear regression models. Mean and variance adjustments that improve the standard normal approximation to the conditional distribution of the signed square root of the log likelihood ratio statistic for a scalar parameter of interest are given. A Bartlett adjustment factor that improves the chi-squared approximation to the conditional distribution of the log likelihood ratio statistic for a vector parameter of interest is also presented. The accuracy of approximate confidence limits obtained by using the adjustments is demonstrated for a location-scale analysis of Darwin's data.