Biometrika (1987), **74**, 2, *pp.* 423-5
Printed in Great Britain

Miscellanea

Correction factors for F ratios in nonlinear regression

BY DAVID HAMILTON AND DOUGLAS WIENS
Department of Mathematics, Statistics and Computing Science, Dalhousie University, Halifax, Nova Scotia, Canada B3H 4H8

SUMMARY

Multiplicative correction factors are derived for the limiting F distributions of two test statistics for parameter subsets in nonlinear regression. The factors depend on the first and second derivatives of the model and are related to measures of intrinsic nonlinearity. An example is given for which the correction is substantial. A similar factor is obtained for the lack of fit test in nonlinear regression.

Some key words: Intrinsic nonlinearity; Lack of fit; Limiting distribution; Nonlinear regression; Parameter subsets.

1. INTRODUCTION

The nonlinear regression model is $y=\eta(\theta)+\varepsilon$, where y, $\eta(\theta)$ and ε are n-dimensional vectors respectively containing the observed responses, their expectations and a spherically normal error with zero mean and variance $\sigma^2 I$. The expectation vector depends nonlinearly on the p parameters $\theta^{\mathrm{T}}=(\theta_1^{\mathrm{T}}, \theta_2^{\mathrm{T}})$, and also on some explanatory variables. Two statistics for testing the hypothesis H_0: $\theta_1=\theta_{1,0}$ are based on the likelihood ratio

$$F_1=\frac{(|\tilde{e}|^2-|\hat{e}|^2)/p_1}{|\hat{e}|^2/(n-p)}$$

and on the efficient scores

$$F_2=\frac{|(\tilde{P}-\tilde{P}_2)\tilde{e}|^2/p_1}{|(I-\tilde{P})\tilde{e}|^2/(n-p)}.$$

In these expressions $e=y-\eta(\theta)$ is the residual vector for a given θ. The matrix $P=V(V^{\mathrm{T}}V)^{-1}V^{\mathrm{T}}$ denotes the projection matrix onto the subspace spanned by the columns of $\partial\eta/\partial\theta=V=(V_1, V_2)$, and $P_2=V_2(V_2^{\mathrm{T}}V_2)^{-1}V_2^{\mathrm{T}}$. The tilde above a character indicates evaluation using $\theta_{1,0}$ and θ_2 equal to its restricted maximum likelihood estimate $\tilde{\theta}_2$. Similarly, a circumflex denotes evaluation with $\hat{\theta}_1$ and $\hat{\theta}_2$, the maximum likelihood estimates. Both statistics have approximate $F_{n-p}^{p_1}$ distributions, and the approximation is exact in some situations for the score-based statistic (Hamilton, Watts & Bates, 1982; Hamilton, 1986). Multiplicative correction factors are derived to improve the approximation in general.

2. THE CORRECTION FACTORS

Examination of the joint characteristic function of the numerator and denominator of each $p_1F_i/(n-p)$, to terms of $O_p(\sigma^3)$, shows that they are independently distributed as $(1+\alpha_{i1}\sigma^2/p_1)\chi_{p_1}^2$ and $\{1+\alpha_{i2}\sigma^2/(n-p)\}\chi_{n-p}^2$. In these expressions, $\alpha_{ij}=\sigma^{-4}E(f_{ij})$, where f_{ij} is the fourth-order term in the polynomial expansion in ε for the numerator ($j=1$) or denominator ($j=2$) of $p_1F_i/(n-p)$. To the same order of approximation, F_i is distributed as $(1-\gamma_i\sigma^2)F_{n-p}^{p_1}$, where $\gamma_i=\alpha_{i2}/(n-p)-\alpha_{i1}/p_1$.

The α_{ij} depend on $\partial\eta/\partial\theta$ and $\partial^2\eta/\partial\theta^2$ and are

$$\alpha_{11} = \alpha_{21} + \alpha_{22} - \alpha_{12}, \tag{2·1}$$

$$\alpha_{12} = -\frac{1}{2}\sum_{i=1}^{n-p} \operatorname{tr}(A_i^2) + \frac{1}{4}\sum_{i=1}^{n-p} \{\operatorname{tr}(A_i)\}^2, \tag{2·2}$$

$$\alpha_{21} = -\frac{1}{2}\sum_{i=1}^{p_1} \operatorname{tr}(A_{0,i}^2) + \frac{1}{4}\sum_{i=1}^{p_1} \{\operatorname{tr}(A_{0,i})\}^2, \tag{2·3}$$

$$\alpha_{22} = -\frac{1}{2}\sum_{i=p_1+1}^{n-p_2} \operatorname{tr}(A_{0,i}^2) + \frac{1}{4}\sum_{i=p_1+1}^{n-p_2} \{\operatorname{tr}(A_{0,i})\}^2, \tag{2·4}$$

where the $p \times p$ matrices A_i and $p_2 \times p_2$ matrices $A_{0,i}$ are faces of the intrinsic acceleration arrays (Bates & Watts, 1980) for, respectively, the full model and restricted model obtained by fixing θ_1 at the null values. The basis for the sample space is chosen so that $A_{0,i}$, for $i = 1, \ldots, p_1$, measure the intrinsic curvature of the restricted model which is in the tangent space of the full model. Such a basis can be obtained using a modification of the algorithm of Bates, Hamilton & Watts (1983), which also eliminates the dependence on n of the upper indices of summation in (2·2) and (2·4). Further details of the derivation and calculation of the correction factors are contained in a technical report available from the authors.

For tests of composite hypotheses the correction factors are evaluated using the hypothesized value for θ_1 and the restricted estimate $\tilde{\theta}_2$. For the associated confidence regions the factors are calculated using the unrestricted maximum likelihood estimates $\hat{\theta}_1$ and $\hat{\theta}_2$.

Johansen (1983) derived an equivalent result with $\gamma_1 = n\alpha_{12}/\{p(n-p)\}$ for the likelihood ratio statistic F_1 and $p_1 = p$. This result refines that of Beale (1960), whose more conservative factor replaces α_{12} by $-(p+2)N_{\min}/\sigma^2$, which is obtained from (2·2) by changing the sign of the second term. Beale (1960) showed that $N_{\min}$ is a measure of intrinsic nonlinearity of the model, and Bates & Watts (1980) showed that $N_{\min}$ is one quarter of their mean square intrinsic curvature. Inspection of (2·3) and (2·4) shows that $\alpha_{21} + \alpha_{22}$ gives an expression analogous to that for α_{12} and measures the intrinsic curvature of the restricted model. Therefore α_{11} measures the difference in intrinsic curvature between the full and restricted models. Hamilton (1986) discusses two additional test statistics, based on the scores with $\theta_2 = \hat{\theta}_2$ and on the large-sample normality of $\hat{\theta}_1$. Correction factors for these statistics were found to be prohibitively complicated.

3. APPLICATION TO TESTING FOR LACK OF FIT

The usual test statistic for evaluating lack of fit in regression is the likelihood ratio statistic for the test H_0: $y = \eta(\theta) + \varepsilon$ versus H_a: $y = Z\beta + \varepsilon$, where $\beta_i = E(y|x_i)$, for $i = 1, \ldots, k$, and k is the number of distinct settings of the explanatory variables x, and Z is the matrix of indicators. The model under H_0 is a restriction of that under H_a, and a parameterization $\phi = \phi(\theta)$ can be found for which the null hypothesis is H_0: $\phi_1 = 0$, where ϕ_1 is $(k-p) \times 1$. Thus, the previous result can be applied, with $i = 1$, $p = k$, and $p_1 = p$. Fortunately it is not necessary to find ϕ or the model derivatives with respect to ϕ because of the invariance of the intrinsic arrays and α_{ij} under reparameterizations. The test statistic is approximately distributed as $(1 + \sigma^2\alpha_{12}/p)F_{n-k}^{k-p}$, where α_{12} is calculated for the restricted nonlinear model using (2·2). The multiplicative correction factor for testing lack of fit is always closer to unity than is Johansen's (1983) factor for the likelihood ratio test.

4. EXAMPLES AND DISCUSSION

Data for the model $\eta_t = \theta_1(1 - \theta_2 e^{-\theta_3 x_t})$ is given in problem N of Draper & Smith (1981, p. 524). The multiplicative factors for the likelihood ratio test and the score test were calculated for their data set 1, using parameter estimates $\hat{\theta} = (263{\cdot}71, 0{\cdot}94801, 4{\cdot}86911 \times 10^{-4})$ and $s^2 = 50{\cdot}98$, for all

possible parameter subsets. The factors are similar for the two tests for every subset, and are very close to unity except for the case where only θ_2 is fixed by the hypothesis. In this case the factors are 0·6500 and 0·6499 for the likelihood ratio and score methods respectively. A confidence interval for θ_2 without regard for the correction would be much too large.

Correction factors were also obtained for 89 data sets corresponding to 37 different models with up to 5 parameters. The correction factors for subsets were usually further from 1 than the correction factors for all the parameters, and the smallest subsets usually had the most extreme factors for a given data set. The likelihood ratio factor was usually positive and further from 1 than the score factor, which was usually less than 1. The subset likelihood ratio factors were all between 0·6500 and 1·0612, while the subset score factors were all between 0·6499 and 1·0021.

Acknowledgement

This work was supported by the Natural Sciences and Engineering Research Council of Canada.

References

Bates, D. M., Hamilton, D. C. & Watts, D. G. (1983). Calculation of intrinsic and parameter-effects curvatures for nonlinear regression models. *Comm. Statist. Simul. Computa.* **12**, 469–77.

Bates, D. M. & Watts, D. G. (1980). Relative curvature measures of nonlinearity (with discussion). *J. R. Statist. Soc.* B **42**, 1–25.

Beale, E. M. L. (1960). Confidence regions in nonlinear estimation (with discussion). *J. R. Statist. Soc.* B **22**, 41–88.

Draper, N. R. & Smith, H. (1981). *Applied Regression Analysis*, 2nd ed. New York: Wiley.

Hamilton, D. C. (1986). Confidence regions for parameter subsets in nonlinear regression. *Biometrika* **73**, 57–64.

Hamilton, D. C., Watts, D. G. & Bates, D. M. (1982). Accounting for intrinsic nonlinearity in nonlinear regression parameter inference regions. *Ann. Statist.* **10**, 386–93.

Johansen, S. (1983). Some topics in regression. *Scand. J. Statist.* **10**, 161–94.

[*Received February* 1986. *Revised July* 1986]