TESTING A NONLINEAR REGRESSION SPECIFICATION: A NONREGULAR CASE

by

A. RONALD GALLANT

Institute of Statistics
Mimeograph Series No. 1017
Raleigh - June 1975

TESTING A NONLINEAR REGRESSION SPECIFICATION:

A NONREGULAR CASE

by

A. RONALD GALLANT*

*Associate Professor of Statistics and Economics, Department of Statistics, North Carolina State University, Raleigh, North Carolina 27607. Presently on leave and with the Graduate School of Business, University of Chicago, Chicago, Illinois 60637.

ABSTRACT

The article considers a statistical test of whether or not a nonlinear regression specification should contain an additional additive nonlinear term in the response function. The regularity conditions used to obtain the asymptotic distributions of the usual test statistics for parameters of nonlinear regression models are violated when the null hypothesis - that this additional term is not present - is true. Moreover, standard iterative algorithms are likely to perform poorly when, in fact, the data support the null hypothesis. Methods designed to circumvent these mathematical and computational difficulties are described in the article and are illustrated with examples.

1. INTRODUCTION

In nonlinear regression analysis it is helpful to be able to choose between two model specifications:

$$H: \quad y_t = g(x_t, \Psi) + e_t$$

and

$$A: \quad y_t = g(x_t, \Psi) + \tau h(x_t, \omega) + e_t$$

on the basis of sample data: observed responses y_t to k-dimensional inputs x_t $(t = 1, 2, \ldots, n)$. The unknown parameters are Ψ, u-dimensional, τ, univariate, and ω which is v-dimensional. The functional forms $g(x, \Psi)$ and $h(x, \omega)$ are known and $h(x, \omega)$ does, in fact, vary with ω. The errors e_t are assumed to be normally and independently distributed with mean zero and unknown variance σ^2. Parametrically, the situation described is equivalent to testing:

$$H: \tau = 0 \quad \text{against} \quad A: \tau \neq 0$$

regarding Ψ, ω, and σ^2 as nuisance parameters.

A natural impulse is to employ one of the nonlinear regression analogues of the tests used in linear regression, either the Likelihood Ratio Test or the test based on the asymptotic normality of the least squares estimator [8]. Both of these tests - they are not equivalent in nonlinear regression - depend on the unconstrained least squares estimator $\hat{\theta} = (\hat{\Psi}, \hat{\omega}, \hat{\tau})$. When the null hypotheses is true, this dependence causes two difficulties:

1. Likely, the attempt to fit the full model

$$y_t = g(x_t, \Psi) + \tau h(x_t, \omega) + e_t$$

using one of the standard first derivative iteritive alogirthms [10, 14] will fail or, at best, converge very slowly.

2. The regularity conditions [12,13,4] used to obtain the asymptotic properties of the unconstrained least squares estimator $\hat{\theta}$ are violated; these asymptotic properties are needed to derive the asymptotic null distribution of these two test statistics [8].

The first difficulty can be illustrated using the model

$$y_t = \theta_1 x_{1t} + \theta_2 x_{2t} + \theta_4 e^{\theta_3 x_{3t}} + e_t \,.$$

Consider an attempt to determine from the data whether the exponential term should be included; the correspondences with the notation above are:

$$g(x,\Psi) = \Psi_1 x_1 + \Psi_2 x_2 \,,$$

$$h(x,\omega) = e^{\omega x_3} \,,$$

$$\Psi = (\theta_1, \theta_2)' \,,$$

$$\tau = \theta_4 \,,$$

$$\omega = \theta_3 \,.$$

Table 1 illustrates how the performance of the modified Gauss-Newton method deteriorates as H becomes more nearly true.

Table 1 was constructed as follows. A sample of thirty normally distributed errors with mean zero and variance .001 was generated. To this sample (fixed throughout) was added the response function

$$f(x,\theta) = \theta_1 x_1 + \theta_2 x_2 + \theta_4 e^{\theta_3 x_3}$$

using the parameter choices shown in Table 1 and the design points, x_t, given in the Appendix of [7]. The least squares estimator $\hat{\theta}$ was determined by grid search. From the start value

$$_{(0)}\theta_i = \hat{\theta}_i - .1 \qquad (i = 1, 2, 3, 4) \,,$$

1. PERFORMANCE OF THE MODIFIED GAUSS-NEWTON METHOD

True Parameter Values					Least Squares Estimate					Number of Modified Gauss-Newton
θ_1	θ_2	θ_3	θ_4	σ^2	$\hat{\theta}_k$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_4$	s^2	Iterations From A Start of $\hat{\theta}_1$ - .1
0	1	-1	-.5	.001	-.0259	1.02	- 1.12	-.505	.00117	4
0	1	-1	-.3	.001	-.0260	1.02	- 1.20	-.305	.00117	5
0	1	-1	-.1	.001	-.0265	1.02	- 1.71	-.108	.00118	6
0	1	-1	-.05	.001	-.0272	1.02	- 3.16	-.0641	.00117	7
0	1	-1	-.01	.001	-.0275	1.02	-28.7	-.163	.00117	20
0	1	-1	-.005	.001	-.0268	1.01	- .0971	.0106	.00119	a
0	1	-1	-.001	.001	-.0266	1.01	- .134	.0132	.00119	202
0	1	-1	0	.001	-.0266	1.01	- .142	.0139	.00119	69

[a] Algorithm had failed to converge after 500 iterations

an attempt was made to recompute the least squares estimator using the modified Gauss-Newton method and the stopping rule: Stop when two successive iterations, ${}_{(i)}\theta$ and ${}_{(i+1)}\theta$, do not differ in the fifth significant digit (properly rounded) of any component.

In each case considered in Table 1, all factors are held fixed save τ^* , the true value of τ[1/] , and the computational task is the same. The performance of the algorithm is seen to deteriorate as $|\tau|$ $(= |\theta_4|)$ becomes smaller. The inference to be drawn from Table 1 is: If a non-linear regression specification is not supported by the data one can expect problems with iterative nonlinear least squares algorithms. This inference is also supported by Tables 1 and 2 of [5].

The mathematical difficulties are caused by the violation of two standard [6,4,12] regularity conditions; the first, that the (almost sure) limit

$$\lim_{n\to\infty} (1/n) \Sigma_{t=1}^{n} [f(x_t,\theta) - f(x_t,\theta^*)]^2$$

has a unique minimum at the true value θ^* of θ and, the second, that the (almost sure) limit matrix with typical element

$$\lim_{n\to\infty} (1/n) \Sigma_{t=1}^{n} [(\partial/\partial\theta_i)f(x_t,\theta^*)][(\partial/\partial\theta_j)f(x_t,\theta^*)]$$

is non-singular. The consequence of these violations is that it does not appear possible to derive the asymptotic properties of $\hat{\omega}$ when H is true under assumptions having sufficient generality to be useful in applications. Specifically, $\hat{\omega}$ will not converge in probability to a constant so that as n becomes large the responses are not constrained to lie in the linear space defined by the gradient of the response function evaluated at θ^* . As a result, the methods of proof employed in, e.g., [4,6,8,12] cannot be used to deduce that, under H , test statistics depending on $\hat{\theta}$ will eventually behave as their linear regression analogs.

While it may be possible to discover the asymptotic properties of $\hat{\theta}$ using other methods of proof, it would seem a fruitless task in view of the computational problems noted previously. For this reason, this article does not consider tests for H against A which depend on the computation of $\hat{\theta}$. The approach, rather, shall be to extend Hartley's method [7, Sec. 4] to this problem.

It is instructive to consider when the situation of testing H against A using data which support H is likely to arise. This is the situation which causes computational difficulties and is the situation where the methods suggested here will be the most useful.

Computational problems of the nature discussed above are unlikely to arise, in the author's experience, when plots of the data visually suggest the nonlinear model. For example, in the cases considered in Table 1, plots of the observed response y_t against the input x_{3t} fail to reveal any visual impression of exponential growth for values of $|\tau|$ smaller than .1 . It is improbable, therefore, that one would attempt to fit a nonlinear model which is not supported visually if one is merely attempting to represent data parametrically without reference to the substantive discipline(s) involved. Consequently, substantive considerations will likely have suggested A rather than data analytic considerations; moreover, a statistical test of H against A will likely have substantive relavance. As we shall see, it will be helpful if these same substantive considerations also imply probable values for ω .

2. AN EXTENSION OF HARTLEY'S TEST

Hartley's test [7, Sec. 4] would serve for testing H against A if the model were linear in the remaining parameters once τ was specified.

The linearity requirement may be eliminated, as is seen below, when $g(x,\Psi)$ satisfies regularity conditions sufficient to obtain, asymptotically, the linear regression results on which Hartley's test is based. With these results in hand, Hartley's idea may be applied: Replace the original problem, testing H against A , with a new problem, testing H against A# , where A# is chosen so that the test statistic for the new problem will be analytically and computationally tractable when H is true yet retain good power when A is true.

Consider, then, the nonlinear regression model

$$A\#: \quad y_t = g(x_t,\Psi) + z_t'\delta + e_t$$

where the additional regressors z_t $(w \times 1)$ do not depend on any unknown parameters. The function $g(x,\Psi)$ is assumed to be such that, and the regressors z_t are chosen to be such that, the regularity conditions of [6, Sec. 2] are satisfied by the model A#; see [7, Sec. 2] for a less formal statement of the regularity conditions. Notice that satisfaction of the regularity conditions is not affected by whether or not $\delta = 0$ since the additional regressors do not depend on unknown parameters.

The Likelihood Ratio Test statistic T# for testing

$$H: \quad y_t = g(x_t,\Psi) + e_t$$

against A# may be derived as follows. Define $(\Psi\#,\delta\#)$ to be the least squares estimator obtained by minimizing

$$\Sigma_{t=1}^{n} [y_t - g(x_t,\Psi) - z_t'\delta]^2$$

and set

$$(\sigma^2)\# = (1/n)\, \Sigma_{t=1}^{n} [y_t - g(x_t,\Psi\#) - z_t'\delta\#]^2 \; .$$

Define $\tilde{\Psi}$ to be the least squares estimator obtained by minimizing

$$\Sigma_{t=1}^{n} [y_t - g(x_t,\Psi)]^2$$

and set

$$\tilde{\sigma}^2 = (1/n)\ \Sigma_{t=1}^{n}\ [y_t - g(x_t,\tilde{\Psi})]^2\ .$$

The test statistic is

$$T\# = \tilde{\sigma}^2/(\sigma^2)\#$$

and H is rejected when $T\#$ is larger than the appropriate size α critical point.

The asymptotic critical point is

$$c\# = 1 + wF_\alpha/(n - u - w)$$

where F_α denotes the upper $\alpha\cdot 100$ percentage point of an F random variable with w numerator degrees freedom and $n - u - w$ denominator degrees freedom. If the regularity conditions of [6, Sec. 2] are satisfied by the model $A\#$ then the additional assumption of [8, Sec. 2] will automatically be satisfied when H is true with the consequence that

$$\lim_{n\to\infty} P[T\# > c\# | H] = \alpha\ .$$

Thus, the test - reject H when $T\# > c\#$ - is, asymptotically, a size α test of H against A.

It remains to consider how the additional regressors z_t should be chosen. Relevant are two restrictions implied by the regularity conditions:

1. Every component of the w-vector z_t which is a random variable must be distributed independently of the errors e_t.

2. Let $G(\Psi)$ be the n by u matrix with typical element $(\partial/\partial\Psi_j)g(x_t,\Psi)$ where t is the row index and j is the column index; let Z be the n by w matrix with $t^{\underline{th}}$ row z_t'. The (almost sure) limit matrix

$$\lim_{n\to\infty} (1/n)[G(\Psi^*) \vdots Z]'[G(\Psi^*) \vdots Z]$$

must have full rank. Thus, the columns of Z should not be exact linear combinations of the columns of $G(\Psi)$ for admissible values of Ψ [2/].

The objective governing the choice of the additional regressors is to find those which will maximize the power of the test when A is true. One should attempt to find those z_t which best approximate $h(x_t, \omega^*)$ - in the sense of maximizing the ratio

$$h'(\omega^*)Z(Z'Z)^{-1}Z'h(\omega^*)/h'(\omega^*)h(\omega^*)$$

where

$$h(\omega^*) = [h(x_1,\omega^*), \ldots, h(x_n,\omega^*)]' \qquad (n \times 1)$$

- while attempting, simultaneously, to keep the number of columns in Z as small as possible; see Hartley [9]. We consider, next, how this might be done in applications.

In a situation where substantive considerations or previous experimental evidence suggest a single point estimate $\omega^{\#}$ for ω the natural choice is $z_t = h(x_t, \omega^{\#})$.

If, instead of a point estimate, ranges of plausible values for the components of ω are available then a representative selection of values for ω

$$\{\omega_i^{\#}: \quad i = 1, 2, \ldots, K\}$$

whose components fall within these ranges can be chosen - either deterministically or by random sampling from a distribution defined on the plausible values - and the vectors $h(\omega_i^{\#})$ made the columns of Z . If, following this procedure, the number of columns of Z would be unreasonably large, their number may be reduced as follows. Decompose the matrix

$$H = [h(\omega_1^{\#}) \vdots \ldots \vdots h(\omega_K^{\#})]$$

into its principal component vectors and choose the first few of these to make up Z ; equivalently, obtain the singular value decomposition [1]

$H = USV'$ and choose the first few columns of U to make up Z .

A portion of the sample data may be used to refine this procedure, if desired. Select, by inspecting the inputs x_t , a representative subset

$$(y_{t_j}, x_{t_j}) \qquad (j = 1, 2, \ldots, n^*)$$

of the sample $- u + v/^{+1}$ observations will suffice. If the data consist of replicates of a designed experiment one of the replicates may be chosen randomly and used as the subsample. Attempt to fit the model

$$y_{t_j} = g(x_{t_j}, \Psi) + \tau h(x_{t_j}, \omega) + e_{t_j}$$

by least squares to obtain a point estimate $\omega^{\#}$ for use as described above. The average $\bar{\omega}$ of the $\omega_i^{\#}$ $(i = 1, 2, \ldots, K)$ may be used as the start value for the iterations; convergence may be checked by trying a few of the $\omega_i^{\#}$ as start values to see if convergence to the same solution obtains. As seen previously, the least squares estimator probably can be found without difficulty when the data support A ; if the computations prove troublesome, find the minima

$$M_i = \Sigma_{j=1}^{n^*} [y_{t_j} - g(x_{t_j}, \Psi) - \tau h(x_{t_j}, \omega_i^{\#})]^2$$

with respect to (Ψ, τ) and retain as the point estimate that $\omega_i^{\#}$ associated with the smallest M_i . One may, if desired, retain those $\omega_i^{\#}$ whose associated M_i are very near the smallest of the M_i and proceed as above with this smaller set $\{\omega_i^{\#}:\ i = 1, \ldots, K^*\}$ replacing $\{\omega_i^{\#}:\ i = 1, 2, \ldots, K\}$. To preserve independence, the data points

$$(y_{t_j}, x_{t_j}) \qquad (j = 1, 2, \ldots, n^*)$$

used to acquire information concerning ω should be deleted prior to the computation of $T^{\#}$.

The reason that it is easier to compute $(\Psi\#,\delta\#)$ for

$$A\#:\quad y_t = g(x_t,\Psi) + z_t'\delta + e_t$$

than it is to compute $(\hat{\Psi},\hat{\omega},\hat{\tau})$ for

$$A:\quad y_t = g(x_t,\Psi) + \tau h(x_t,\omega) + e_t$$

from data which support

$$H:\quad y_t = g(x_t,\Psi) + e_t$$

is that δ enters the model $A\#$ linearly and z_t does not depend on any unknown parameters. The length $\|\delta\|$ of δ can move toward zero in the iterations without causing other parameters to become indeterminate or the matrix $[G(\Psi)\vdots Z]'[G(\Psi)\vdots Z]$ to become nearly singular.

The Likelihood Ratio Test statistic $T\#$ was selected to test H against $A\#$ rather than a test statistic $S\#$ based on the asymptotic normality of $\delta\#$ because the Monte-Carlo evidence in Tables 1 and 2 of [8] support this choice; we are acting on the presumption that, for well chosen $A\#$, the more powerful test of H against $A\#$ will be the more powerful test of H against A. If desired, a test using $S\#$ may be substituted and the considerations above remain valid.

One proceeds as follows. Let $(\Psi\#,\delta\#)$ be the least squares estimate for

$$A\#:\quad y_t = g(x_t,\Psi) + z_t'\delta + e_t$$

and let

$$(s^2)\# = \Sigma_{t=1}^{n}\,[y_t - g(x_t,\Psi\#) - z_t'\delta\#]^2/(n - u - w)\ .$$

Evaluate the matrix $G(\Psi)$ at $\Psi = \Psi\#$ and put

$$C\# = \{[G(\Psi\#)\vdots Z]'[G(\Psi\#)\vdots Z]\}^{-1}\ .$$

Let $C\#_{22}$ be the matrix formed by deleting the first u rows and columns of C ; then

$$S\# = \frac{(\delta\#)'(C\#_{22})^{-1}\delta\#/w}{(s^2)\#} .$$

H is rejected when $S\#$ exceeds the upper $\alpha \cdot 100$ percentage point F_α of an F random variable with w numerator degrees freedom and $n - u - w$ denominator degrees freedom.

3. EXAMPLES

The first example is a reconsideration of the testing problem described in [5, Sec. 5]. The data shown in Figure A are preschool boy's weight/height ratios plotted against age and were obtained from [2]; the tabular values are given in the Appendix. The question to be decided is whether the data support the choice of a three segment quadratic-quadratic-linear polynomial response function as opposed to a two segment quadratic-linear response function. In both cases, the response function is required to be continuous in x (= age) and to have a continuous first derivative in x. Formally,

$$H: \quad y_t = \theta_1 + \theta_2 x_t + \theta_3 T_2(\theta_4 - x_t) + e_t$$

and

$$A: \quad y_t = \theta_1 + \theta_2 x_t + \theta_3 T_2(\theta_4 - x_t) + \theta_5 T_2(\theta_6 - x_t) + e_t$$

where

$$T_k(z) = \begin{cases} z^k & \text{when} \quad z \geq 0 \\ 0 & \text{when} \quad z < 0 \end{cases} ;$$

see [5, Sec. 2].

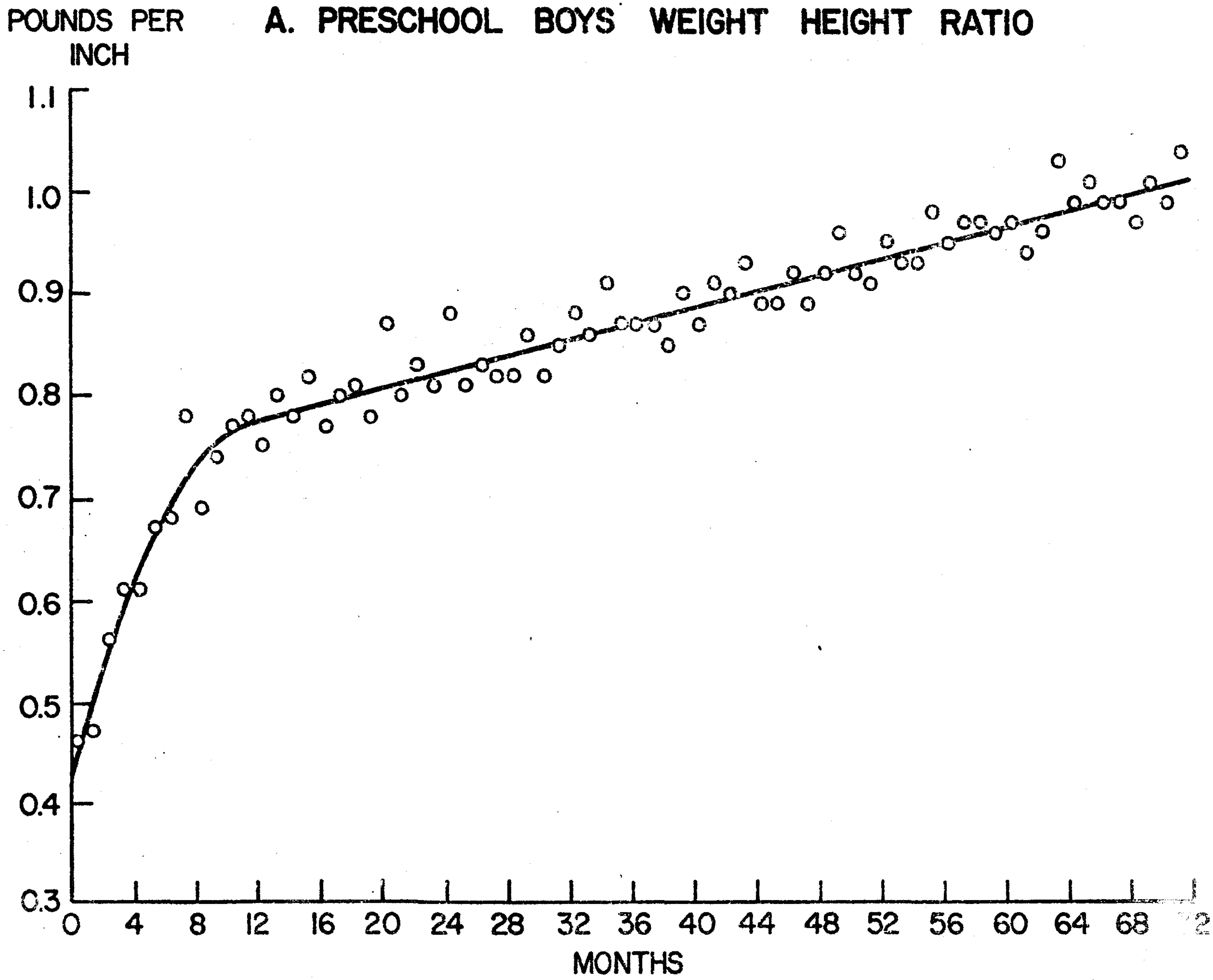
The correspondences with the notation of Section 1 are:

$$g(x,\Psi) = \Psi_1 + \Psi_2 x + \Psi_3 T_2(\Psi_4 - x) ,$$

$$h(x,\omega) = T_2(\omega - x) ,$$

$$\Psi = (\theta_1, \theta_2, \theta_3, \theta_4) ,$$

A. PRESCHOOL BOYS WEIGHT HEIGHT RATIO



$$\tau = \theta_5 ,$$

$$\omega = \theta_6 .$$

Choosing as plausible values for ω the points $\omega_1^\# = 4$, $\omega_2^\# = 8$, $\omega_3^\# = 12$ the matrix H of Section 2 has typical row

$$H_t = [T_2(4 - x_t), T_2(8 - x_t), T_2(12 - x_t)] .$$

The first principal component vector of H , with elements

$$z_t = [(2.08)T_2(4 - x_t) + (14.07)T_2(8 - x_t) + (39.9)T_2(12 - x_t)] \times 10^{-4} ,$$

was chosen as the additional regressor to obtain:

$$\tilde{\sigma}^2 = .0005264 ,$$

$$(\sigma^2)^\# = .0005235 ,$$

$$T^\# = 1.006 ,$$

$$P[T^\# > 1.006 | H] \doteq .485 .$$

These data give little support to A .

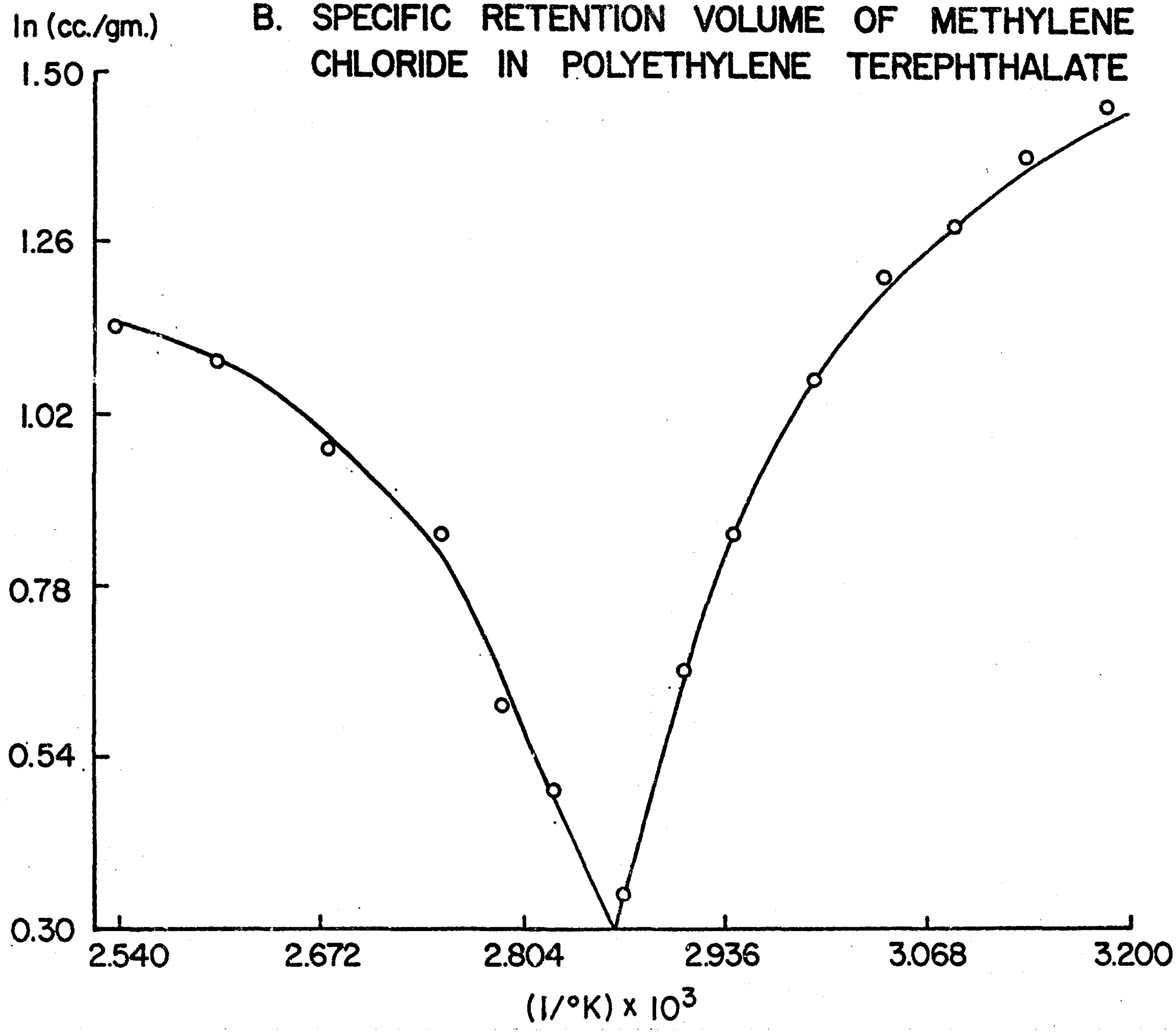
The model

$$A^\#: \quad y_t = \Psi_1 + \Psi_2 x_t + \Psi_3 T_2(\Psi_4 - x_t) + \delta' z_t + e_t$$

does not satisfy the regularity conditions of [6, Sec. 2] because the second derivative $(\partial^2/\partial\Psi_4^2)T_2(\Psi_4 - x) = 2T_0(\Psi_4 - x)$ is not continuous in Ψ_4 ; we are relying on the asymptotic theory in [4] to justify the use of $T^\#$ in this instance.

The second example illustrates how the ideas in Section 2 may be extended in a natural way to a situation studied by Feder [3]. The data shown in Figure B are from [11] and were obtained to study the specific retention volume of the organic liquid methylene chloride in the polymer polyethylene terephthalate at selected temperatures; the tabular values are given in the Appendix. The question is whether the data support a single quadratic polynomial model

B. SPECIFIC RETENTION VOLUME OF METHYLENE CHLORIDE IN POLYETHYLENE TEREPHTHALATE



$$H: \quad y_t = \theta_1 + \theta_2 x_t + \theta_3 x_t^2 + e_t$$

or a segmented quadratic-quadratic polynomial model

$$A: \quad y_t = \theta_1 + \theta_2 x_t + \theta_3 x_t^2 + \theta_4 T_1(\theta_6 - x_t)$$
$$+ \theta_5 T_2(\theta_6 - x_t) + e_t$$

where the latter response function is restricted to be continuous in x but is <u>not</u> restricted to have continuous first derivitive in x - there will be a jump discontinuity in the first derivative of gap $|\theta_4|$ at $x = \theta_6$. The answer to the question is visually obvious from Figure B but, nonetheless, we shall confirm it by means of a statistical test.

Choosing as plausible values for $\omega = \theta_6$ the points $\omega_1^{\#} = 2.8$, $\omega_2^{\#} = 2.85$, and $\omega_3^{\#} = 2.9$ we construct H $(n \times 6)$ with typical row

$$H_t = [T_1(2.8 - x_t),\ T_1(2.85 - x_t),\ T_1(2.9 - x_t)\ ,$$
$$T_2(2.8 - x_t),\ T_2(2.85 - x_t),\ T_2(2.9 - x_t)]\ .$$

The first two principal component vectors of H, with elements

$$z_{1t} = H_t \cdot (.52,\ .68,\ .84,\ .11,\ .17,\ .25)'\ ,$$
$$z_{2t} = H_t \cdot (-6.4,\ -.94,\ 6.6,\ -2.7,\ -3.1,\ -2.8)'\ ,$$

were chosen as additional regressors to obtain:

$$\tilde{\sigma}^2 = .02381\ ,$$
$$(\sigma^2)^{\#} = .003001\ ,$$
$$T^{\#} = 7.935\ ,$$
$$P[T^{\#} > 7.935 | H] = .0000895\ .$$

These data strongly support A which, as was noted, is visually obvious. Observe that the model

$$A^{\#}: \quad y_t = \Psi_1 + \Psi_2 x_t + \Psi_3 x_t^2 + \delta_1 z_{1t} + \delta_2 z_{2t} + e_t$$

is linear in its parameters so that

$$P[T^{\#} > 7.935 | H] = .0000895$$

exactly, not asymptotically - granted normally distributed errors.

For this problem Feder [3] has shown that the Likelihood Ratio Test statistic

$$T = \tilde{\sigma}^2 / \hat{\sigma}^2$$

(normalized as nT) is not asymptotically distributed as a Chi-square random variable when H holds. Moreover, he argues that the distribution of T will depend on the particular arrangement of design points x_t. Thus, it would appear impossible to obtain a single set of tables for the use of T in applications; it would be necessary to find critical points by Monte-Carlo simulations at each instance. The methodology suggested here substitutes the statistic $T\#$ allowing the use of tables of F. In fact, as was noted, for this particular example the model chosen for $A\#$ is linear in the parameters; hence, exact size, not asymptotic size, is achieved by the test - granted the normality assumption.

FOOTNOTES

1/ The asterisk is used in connection with parameters - Ψ^*, τ^*, ω^*, and θ^* - to emphasize that it is the true, but unknown, value which is meant. The omission of the asterisk does not necessarily denote the contrary.

2/ This is technically correct if one wishes to apply the results of [6,8] straightforwardly. It appears that the nonlinear theory could be strengthened to allow Z to have rank less than w. Such a generalization would have no practical utility in the present context.

REFERENCES

[1] Bussinger, P. A. and Golub, G. H., "Singular Value Decomposition of a Complex Matrix", Communications of the ACM, 12 (October 1969), 564-565.

[2] Eppright, E. S., Fox, H. M., Fryer, B. A., Lamkin, G. H., Vivian, V. M., and Fuller, E. S., "Nutrition of Infants and Preschool Children in the North Central Region of the United States of America", World Review of Nutrition and Dietetics, 14 (1972), 269-332.

[3] Feder, P. I., "The Log Likelihood Ratio in Segmented Regression", The Annals of Statistics, 3 (January 1975) 84-97.

[4] Gallant, A. R., "Inference for Nonlinear Models", Institute of Statistics Mimeograph Series, No. 875. Raleigh: North Carolina State University, 1973.

[5] Gallant, A. R. and Fuller, W. A., "Fitting Segmented Polynomial Regression Models Whose Join Points Have to Be Estimated", Journal of the American Statistical Association, 68 (March 1973), 144-147.

[6] Gallant, A. R., "The Power of the Likelihood Ratio Test of Location in Nonlinear Regression Models", Journal of the American Statistical Association, 70 (March 1975), 198-203.

[7] Gallant, A. R., "Nonlinear Regression", The American Statistician, 29 (May 1975), 73-81.

[8] Gallant, A. R., "Testing a Subset of the Parameters of a Nonlinear Regression Model", Journal of the American Statistical Association, 70 (In press, December 1975).

[9] Hartley, H. O., "Exact Confidence Regions for the Parameters in Non-Linear Regression Laws", Biometrika, 51 (December 1964), 347-353.

[10] Hartley, H. O., "The Modified Gauss-Newton Method for the Fitting of Non-Linear Regression Functions by Least Squares", Technometrics, 3 (May 1961), 269-280.

[11] Hsiung, P. O., Study of the Interaction of Organic Liquids With Polymers by Means of Gas Chromatography. Ph.D. Dissertation, Raleigh: North Carolina State University.

[12] Jennrich, R. I., "Asymptotic Properties of Non-Linear Least Squares Estimators", The Annals of Mathematical Statistics, 40 (April 1969), 633-643.

[13] Malinvaud, E., "The Consistency of Nonlinear Regressions", The Annals of Mathematical Statistics, 41 (June 1970), 956-969.

[14] Marquardt, D. W., "An Algorithm for Least-Squares Estimation of Non-Linear Parameters", Journal of the Society for Industrial and Applied Mathematics, 11 (June 1963), 431-441.

2. BOYS' WEIGHT/HEIGHT VS. AGE

W/H	AGE	W/H	AGE	W/H	AGE
0.46	0.5	0.88	24.5	0.92	48.5
0.47	1.5	0.81	25.5	0.96	49.5
0.56	2.5	0.83	26.5	0.92	50.5
0.61	3.5	0.82	27.5	0.91	51.5
0.61	4.5	0.82	28.5	0.95	52.5
0.67	5.5	0.86	29.5	0.93	53.5
0.68	6.5	0.82	30.5	0.93	54.5
0.78	7.5	0.85	31.5	0.98	55.5
0.69	8.5	0.88	32.5	0.95	56.5
0.74	9.5	0.86	33.5	0.97	57.5
0.77	10.5	0.91	34.5	0.97	58.5
0.78	11.5	0.87	35.5	0.96	59.5
0.75	12.5	0.87	36.5	0.97	60.5
0.80	13.5	0.87	37.5	0.94	61.5
0.78	14.5	0.85	38.5	0.96	62.5
0.82	15.5	0.90	39.5	1.03	63.5
0.77	16.5	0.87	40.5	0.99	64.5
0.80	17.5	0.91	41.5	1.01	65.5
0.81	18.5	0.90	42.5	0.99	66.5
0.78	19.5	0.93	43.5	0.99	67.5
0.87	20.5	0.89	44.5	0.97	68.5
0.80	21.5	0.89	45.5	1.01	69.5
0.83	22.5	0.92	46.5	0.99	70.5
0.81	23.5	0.89	47.5	1.04	71.5

Source: Eppright *et al* [2]

3. SPECIFIC RETENTION VOLUME OF METHYLENE CHLORIDE IN POLY-ETHYLENE TEREPHTHALATE

RECIPROCAL $\times 10^3$ OF TEMPERATURE IN DEGREES KELVIN	NATURAL LOGARITHM OF SPECIFIC VOLUME IN CC. PER GM.
2.54323	1.16323
2.60960	1.10458
2.67952	0.98832
2.75330	0.87471
2.79173	0.62060
2.82965	0.51175
2.87026	0.35371
2.91120	0.66954
2.94637	0.85555
3.00030	1.07086
3.04228	1.22272
3.09214	1.29113
3.13971	1.38480
3.19081	1.46728

Source: Hsiung [11]