COMPREHENSIVE WRITTEN EXAMINATION, PAPER III
FRIDAY AUGUST 20, 2010, 9:00 A.M. – 1:00 P.M.
STOR 664 Question

Suppose we have an experiment to fit the Michaelis-Menten relationship

$$y_i = \frac{\beta_1 x_i}{\beta_2 + x_i} + \epsilon_i, \quad i = 1, 2, ..., n,$$

where $\epsilon_i \sim N[0, \sigma^2]$ (independent for each i). Suppose the ordinary least squares estimates are $\hat{\beta}_1, \hat{\beta}_2$, that s^2 is the estimate of σ^2, and assume $\text{Cov}\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} \approx \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \sigma^2$.

(a) (30 points) Show that, as an approximation, we may take

$$a_{11} = \sum \frac{x_i^2}{(\hat{\beta}_2 + x_i)^4} \Big/ \left[\sum \frac{x_i^2}{(\hat{\beta}_2 + x_i)^2} \sum \frac{x_i^2}{(\hat{\beta}_2 + x_i)^4} - \left\{ \sum \frac{x_i^2}{(\hat{\beta}_2 + x_i)^3} \right\} \right]$$

and give corresponding expressions for $a_{12},\ a_{22}$.

For the remainder of the question, assume that the traditional linear model distribution theory for $(\hat{\beta}_1, \hat{\beta}_2, s^2)$ is also valid in this case — in particular, if k_1 and k_2 are constants and $\text{Var}(k_1\hat{\beta}_1 + k_2\hat{\beta}_2) = k_3^2\sigma^2$ for some constant k_3, then $(k_1\hat{\beta}_1 + k_2\hat{\beta}_2)/k_3 s$ has a t_{n-2} distribution. Suppose we are interested in determining the value $x = x^*$ for which the reaction rate $\beta_1 x^*/(\beta_2 + x^*)$ is some predetermined value c.

(b) (30 points) Suggest an estimator $\hat{x}^*$, and calculate its approximate standard error by the delta method.

(c) (40 points) Show how it is possible to construct a $100(1-\alpha)\%$ confidence set for x^* by means of a sequence of tests of the hypothesis $H_0: \ x^* = x_0$ (for given x_0) at significance level α. Assume t^* is the $(1-\alpha/2)$-quantile of the t_{n-2} distribution. Derive a condition (in terms of $\hat{\beta}_1,\ \hat{\beta}_2,\ c$, etc.) for this confidence set to be a confidence interval, and show that in that case, the length of the interval is

$$\frac{2c\sqrt{\left\{\hat{\beta}_2(\hat{\beta}_1 - c) - s^2 t^{*2} a_{12}\right\}^2 - \left\{(\hat{\beta}_1 - c)^2 - s^2 t^{*2} a_{11}\right\}(\hat{\beta}_2^2 - s^2 t^{*2} a_{22})}}{(\hat{\beta}_1 - c)^2 - s^2 t^{*2} a_{11}}$$

What happens in the cases where the confidence set is not an interval?

SOLUTION

(a) If we write the response $E(y)$ as $g(x;\beta_1,...,\beta_p)$ then the covariance matrix of the OLS $\hat{\beta}$ is approximately $(X^TX)^{-1}\sigma^2$ where the (i,j) entry of X is $\partial g(x_i;\beta_1,...,\beta_p)/\partial\beta_j$ (bookwork, derived in course text). In this case $\partial g/\partial\beta_1 = x/(\beta_2+x)$ and $\partial g/\partial\beta_2 = -\beta_1 x/(\beta_2+x)^2$ so

$$\begin{pmatrix} \sum \frac{x_i^2}{(\beta_2+x_i)^2} & -\beta_1 \sum \frac{x_i^2}{(\beta_2+x_i)^3} \\ -\beta_1 \sum \frac{x_i^2}{(\beta_2+x_i)^3} & \beta_1^2 \sum \frac{x_i^2}{(\beta_2+x_i)^4} \end{pmatrix}.$$

In practice we substitute $\hat{\beta}_j$ for $\beta_j,\ j=1,2$. With this substitution, the expression for a_{11} is the top left entry of $(X^TX)^{-1}$; a_{12} and a_{22} follow similarly.

(b) We solve $\beta_1 x^*/(\beta_2+x^*) = c$ to deduce $x^* = \beta_2 c/(\beta_1 - c)$. In practice we substitute the OLS estimates to get $\hat{x}^* = \hat{\beta}_2 c/(\hat{\beta}_1 - c)$. By the delta method, the approximate variance is

$$\sigma^2 \left\{ \frac{\beta_2^2 c^2}{(\beta_1-c)^4} a_{11} - 2\frac{\beta_2 c^2}{(\beta_1-c)^3} a_{12} + \frac{c^2}{(\beta_1-c)^2} a_{22} \right\}. \quad (1)$$

The standard error follows from (1) by substituting sample estimates for $\sigma^2,\ \beta_1,\ \beta_2$ and taking the square root.

(c) The desired confidence set will consist of all x_0 that are accepted by the hypothesis test. For given x_0, H_0 is the hypothesis that $\theta = x_0\beta_1 - c\beta_2 - cx_0 = 0$. We estimate θ by $\hat{\theta} = x_0\hat{\beta}_1 - c\hat{\beta}_2 - cx_0$ which has variance $\sigma^2(x_0^2 a_{11} + c^2 a_{22} - 2x_0 c a_{12})$ and so the t-test will accept H_0 whenever

$$\frac{|x_0\hat{\beta}_1 - c\hat{\beta}_2 - cx_0|}{s\sqrt{x_0^2 a_{11} + c^2 a_{22} - 2x_0 c a_{12}}} \le t^*,$$

which is rearranged into the quadratic inequality

$$x_0^2 \left\{ (\hat{\beta}_1 - c)^2 - s^2 t^{*2} a_{11} \right\} - 2x_0 c \left\{ \hat{\beta}_2(\hat{\beta}_1 - c) - s^2 t^{*2} a_{12} \right\} + c^2(\hat{\beta}_2^2 - s^2 t^{*2} a_{22}) \ \le \ 0. \quad (2)$$

The inequality (2) becomes an equality at

$$x_0 = \frac{c\left\{\hat{\beta}_2(\hat{\beta}_1 - c) - s^2 t^{*2} a_{12}\right\} \pm c\sqrt{\left\{\hat{\beta}_2(\hat{\beta}_1 - c) - s^2 t^{*2} a_{12}\right\}^2 - \left\{(\hat{\beta}_1 - c)^2 - s^2 t^{*2} a_{11}\right\}(\hat{\beta}_2^2 - s^2 t^{*2} a_{22})}}{(\hat{\beta}_1 - c)^2 - s^2 t^{*2} a_{11}} \quad (3)$$

The required conditions for the resulting solution to be an interval are then (i) $(\hat{\beta}_1 - c)^2 - s^2 t^{*2} a_{11} > 0$, and (ii) the expression inside the square root sign in (3) is positive. In that case, the length of the confidence interval is the distance between the two roots, which is the expression given.

If (ii) is satisfied but not (i), then the confidence set consists of everything *outside* the interval between the two roots (3) — a union of two semi-infinite intervals. If (ii) is violated, then there are no real roots and the confidence set is $(-\infty, \infty)$.