Listy Biometryczne – Biometrical Letters
Vol. 40 (2003), No. 2, 57-64

Comparison of confidence intervals for maximum of a quadratic regression function

Dorota Kozioł, Wojciech Zieliński

Department of Econometrics and Computer Sciences
Agricultural University ul. Nowoursynowska 166, PL-02-787 Warszawa
dorota.koziol@ekonometria.info, wojtek.zielinski@statystyka.info

SUMMARY

Consider a problem of interval estimation of a maximum point of a quadratic regression function. In the paper lengths of three confidence intervals are compared with respect to the location of a point of maximum, function flatness and experimental design. The comparison is made on the basis of a Monte Carlo experiment.

KEY WORDS: quadratic regression function, point of maximum, interval estimation.

1. Introduction

We consider a quadratic regression model

$$Y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \varepsilon_i, \qquad i = 1, ..., n,$$

where ε's are independent, normally distributed random variables such that $E(\varepsilon_i) = 0$ and $D^2(\varepsilon_i) = \sigma^2$. The problem is to estimate $\varphi = -\beta_1/2\beta_2$, assuming $\beta_2 \neq 0$, i.e. the point at which regression function achieves its maximum or minimum.

In matrix notation, the considered model is of the form

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

where $\mathbf{Y} = (Y_1, \ldots, Y_n)'$ is the vector of observations, $\mathbf{X} = [\mathbf{1} \quad \mathbf{x} \quad \mathbf{x}^2]$, where $\mathbf{1}$ is the n–vector of ones, $\mathbf{x} = (x_1, \ldots, x_n)'$ and $\mathbf{x}^2 = (x_1^2, \ldots, x_n^2)'$, is the experimental matrix, $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)'$ is the vector of regression coefficients and $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_n)'$ is the vector of random errors. Under above assumptions $\boldsymbol{\varepsilon} \sim N_n(\mathbf{0}, \sigma^2\mathbf{I})$. Assume that matrix $\mathbf{X}$ is of full rank. Note that the matrix $\mathbf{X}$ is of full rank iff there are at least three different x_i's. If so, there exists the matrix $(\mathbf{X}'\mathbf{X})^{-1}$. Denote elements of

$(\mathbf{X}'\mathbf{X})^{-1}$ by ν^{ij}, i.e. $(\mathbf{X}'\mathbf{X})^{-1} = [\nu^{ij}]$ with $i,j = 0,1,2$. Let

$$\boldsymbol{\beta}' = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \qquad \text{and} \qquad S^2 = \mathbf{Y}'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Y}/(n-3)$$

be LSE estimators of $\boldsymbol{\beta}$ and σ^2 respectively. Note that

$$\hat{\boldsymbol{\beta}} \sim N_3(\boldsymbol{\beta}, \sigma^2(\mathbf{X}'\mathbf{X})^{-1}), \qquad (n-3)S^2 \sim \sigma^2\chi^2(n-3),$$

$\hat{\boldsymbol{\beta}}$ and S^2 are stochastically independent. Let $\hat{\varphi} = -\hat{\beta}_1/2\hat{\beta}_2$ be the point estimator of maximum φ of the regression function.

As far as the methods of obtaining a confidence interval for maximum of regression function are concerned, the first confidence interval for φ was proposed by Filler (1941). Some properties of this confidence interval as well as some large samples confidence intervals were investigated by Bouanaccorsi (1983) and Buonaccorsi and Iyer (1984). We are interested in properties of known confidence intervals for maximum of a quadratic regression and our aim was to make the most exhaustive investigations of them. Our investigations were based on computer simulations. In simulations it is assumed that underlying distribution is normal, because we were not interested in robustness of length of confidence intervals against nonnormality. Some remarks on robustness may be found in Buonaccorsi and Iyer (1984).

2. Confidence intervals

A. *Exact Confidence Interval for* φ. Consider a random variable $(\hat{\beta}_1 + 2\varphi\hat{\beta}_2)$. It is easily seen that its distribution is $N(0, \sigma_u^2)$ with $\sigma_u^2 = \sigma^2(\nu^{11} + 4\varphi\nu^{12} + 4\varphi^2\nu^{22})$. This random variable is stochastically independent on S^2, hence

$$\frac{(\hat{\beta}_1 + 2\varphi\hat{\beta}_2)/\sigma_u}{\sqrt{S^2/\sigma^2}}$$

is distributed as t with $n-3$ degrees of freedom. Let $t(\alpha; n-3)$ be the critical value of t distribution. Then, the exact confidence interval is obtained as a solution of the inequality:

$$\left(\frac{\hat{\beta}_1 + 2\varphi\hat{\beta}_2}{\sqrt{S^2(\nu^{11} + 4\varphi\nu^{12} + 4\varphi^2\nu^{22})}}\right)^2 < (t(\alpha; n-3))^2.$$

Confidence interval is of the form

$$\begin{cases} (r_1; r_2) & \text{for } c_{22} > 0 \text{ and } D \geq 0 \\ (-\infty; \infty) & \text{for } c_{22} \leq 0 \text{ and } D < 0 \\ (-\infty; r_1) \cup (r_2; \infty) & \text{for } c_{22} \leq 0 \text{ and } D \geq 0 \end{cases}$$

where

$$r_1 = \frac{-c_{12} - \sqrt{D}}{2c_{22}}, \qquad r_2 = \frac{-c_{12} + \sqrt{D}}{2c_{22}}, \qquad D = c_{12}^2 - c_{11}c_{22},$$

$$c_{ij} = \hat{\beta}_i\hat{\beta}_j - (t(\alpha; n-3))^2 S^2 \nu^{ij} \qquad (i,j = 1,2).$$

Note that confidence interval is finite iff hypothesis $H_0 : \beta_2 = 0$ is rejected. The hypothesis is rejected when following inequality holds

$$\frac{|\hat{\beta}_2|}{S\sqrt{\nu^{22}}} > t(\alpha; n-3)$$

i.e. $c_{22} = \hat{\beta}_2^2 - (t(\alpha; n-3))^2 S^2 \nu^{22} > 0$. The above construction is similar to the one based on Filler theorem from 1941 which states that in considered model the random variable

$$\frac{(\hat{\beta}_1 - \varphi\hat{\beta}_2)}{\sqrt{S^2(\nu^{11} - 2\varphi\nu^{12} + \varphi^2\nu^{22})}}$$

is distributed as t with $n-3$ degrees of freedom. The confidence limits obtained here are as above with

$$r_1 = \frac{c_{12} - \sqrt{D}}{c_{22}} \qquad \text{and} \qquad r_2 = \frac{c_{12} + \sqrt{D}}{c_{22}}.$$

B. *Approximate Student Confidence Interval for* φ. Serfling (1980) proved (theorem A in §3.3) that if $\boldsymbol{\xi}_n$ are asymptotically normal $N(\boldsymbol{\mu}, b_n\boldsymbol{\Sigma})$ ($b_n \to 0$) random m–vectors and $\mathbf{g}(\mathbf{u}) : \mathbf{R}^m \to \mathbf{R}^k$ is a function such that $\mathbf{D}[\partial \mathbf{g}/\partial \mathbf{u}]|_{\mathbf{u}=\boldsymbol{\mu}}$ is not zero, then $\mathbf{g}(\boldsymbol{\xi}_n)$ is asymptotically normal $N(\mathbf{g}(\mathbf{u}), b_n\mathbf{D}\boldsymbol{\Sigma}\mathbf{D}')$.

In the considered model $\hat{\boldsymbol{\beta}}$ is normal $N(\boldsymbol{\beta}, \sigma^2(\mathbf{X}'\mathbf{X})^{-1})$ for every n. Note that

$$(\mathbf{X}'\mathbf{X})^{-1} = \frac{1}{n}\begin{bmatrix} 1 & \bar{x}^1 & \bar{x}^2 \\ \bar{x}^1 & \bar{x}^2 & \bar{x}^3 \\ \bar{x}^2 & \bar{x}^3 & \bar{x}^4 \end{bmatrix}^{-1}, \qquad \text{where} \qquad \bar{x}^k = \frac{1}{n}\sum_{i=1}^{n} x_i^k, \ k = 1,2,3,4,$$

hence $b_n = \frac{1}{n}$. For $\mathbf{u} = (u_1, u_2, u_3)'$ let $\mathbf{g}(\mathbf{u}) = -u_1/(2u_2)$. Then $\mathbf{D} = \left[0, -\frac{1}{(2\beta_2)}, \frac{\beta_1}{(2\beta_2^2)}\right]$. Hence $\hat{\varphi} = \mathbf{g}(\hat{\boldsymbol{\beta}}) = -\frac{\hat{\beta}_1}{2\hat{\beta}_2}$ is asymptotically normal with mean φ and variance

$$\omega^2 = \frac{\sigma^2(4\varphi^2\nu^{22} + 4\varphi\nu^{12} + \nu^{11})}{(2\beta_2)^2}.$$

The approximate Student confidence interval is of the form $(\hat{\varphi} \pm t(\alpha, n-3)S\hat{\omega})$,where $\hat{\omega} = (4\hat{\varphi}^2\nu^{22} + 4\hat{\varphi}\nu^{12} + \nu^{11})/(2\hat{\beta}_2)^2$.

Note that, in contradiction to the Exact confidence interval, the Student confidence interval always exists.

C. *Approximate Normal Confidence Interval for* φ. For large n $t(\alpha, n-3)$ in (B) may be replaced by the appropriate $z(\alpha)$ critical value of $N(0,1)$ distribution. Then we obtain approximate Normal confidence interval

$$(\hat{\varphi} \pm z(\alpha) S\hat{\omega}).$$

3. Simulation studies

Our aim was to compare lengths of Exact, Student and Normal confidence intervals with respect to the point of maximum, function flatness and experimental design. The comparison was based on computer simulations. In simulations a one hundred different quadratic functions were investigated. Those functions are given in Table 1.

Table 1. Investigated functions

β_2					β_1					
-0.1	0.00	0.02	0.04	0.06	0.08	0.10	0.12	0.14	0.16	0.18
-0.5	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
-1.0	0.00	0.20	0.40	0.60	0.80	1.00	1.20	1.40	1.60	1.80
-1.5	0.00	0.30	0.60	0.90	1.20	1.50	1.80	2.10	2.40	2.70
-2.0	0.00	0.40	0.80	1.20	1.60	2.00	2.40	2.80	3.20	3.60
-2.5	0.00	0.50	1.00	1.50	2.00	2.50	3.00	3.50	4.00	4.50
-3.0	0.00	0.60	1.20	1.80	2.40	3.00	3.60	4.20	4.80	5.40
-3.5	0.00	0.70	1.40	2.10	2.80	3.50	4.20	4.90	5.60	6.30
-4.0	0.00	0.80	1.60	2.40	3.20	4.00	4.80	5.60	6.40	7.20
-4.5	0.00	0.90	1.80	2.70	3.60	4.50	5.40	6.30	7.20	8.10
-5.0	0.00	1.00	2.00	3.00	4.00	5.00	6.00	7.00	8.00	9.00
$x_{\max}$	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90

All function were observed on the $[-1;1]$ interval. In each column there are functions with the same maximum point. In rows there are functions with the same flatness.

Every function was observed 12 times, because we want to check if the asymptotic on a such small sample works. Standard deviation of the random error was put 0.1. Such small standard deviation provides, almost every time, the finite exact confidence interval (A), even for very flat functions. Every function was investigated till the exact confidence interval was obtained 1000 times.

In simulations different experimental designs were applied. Those designs are shown in Table 2. Notation $(x;n)$ means that at the point x n observations were taken.

Table 2. Experimental designs

Design A	(-1;4), (0;4), (1;4)
Design B	(-1;3), (0;6), (1;3)
Design D	(-1+i/3;3), i=0,1,2,3
Design E	(-1+i/5;2), i=0,1,...,5
Design F	(-1+i/11;1), i=0,1,...,11

Design A is the D–optimal one. There were three allocation points $(-1;0;1)$, at each point we have four observations. Design B is D_1-optimal one: at points $(-1;0;1)$ we put $(3,6,3)$ observations respectively. The other three designs have different numbers of allocation points: the same number of observations were put at uniformly distributed points on $[-1;1]$ interval. The last design is the most intuitive one and seems it should show the function curvature.

4. Results

Results of simulations are presented in tables and figures. Table 3 shows average lengths of Exact, Student and Normal confidence intervals with respect to the point of maximum.

Table 3. Average lengths

Maximum point	Confidence interval		
	Exact	Student	Normal
0.0	0.00452407	0.00452329	0.00391910
0.1	0.00478779	0.00478690	0.00414750
0.2	0.00550373	0.00550249	0.00476749
0.3	0.00652466	0.00652296	0.00565166
0.4	0.00773066	0.00772847	0.00669617
0.5	0.00904807	0.00904536	0.00783716
0.6	0.01043475	0.01043153	0.00903816
0.7	0.01186646	0.01186267	0.01027819
0.8	0.01332868	0.01332436	0.01154462
0.9	0.01481238	0.01480755	0.01282968

It is easily seen that if maximum is more on the right side, the lengths of confidence interval grow. The lengths of Exact and Student confidence intervals are comparable which suggests that for the sample $n = 12$ asymptotic works well. Normal confidence intervals are the shortest, but they do not keep a given confidence level (0.95). Estimated confidence levels are given in the Table 4.

Table 4. Confidence levels

Maximum point	Confidence interval		
	Exact	Student	Normal
0.0	0.9500	0.9500	0.9170
0.1	0.9490	0.9489	0.9218
0.2	0.9520	0.9505	0.9168
0.3	0.9520	0.9505	0.9188
0.4	0.9500	0.9499	0.9151
0.5	0.9550	0.9546	0.9134
0.6	0.9560	0.9556	0.9143
0.7	0.9570	0.9570	0.9146
0.8	0.9560	0.9559	0.9129
0.9	0.9550	0.9564	0.9140

Above results are obtained for the experimental design A, but for other experimental designs the average lengths and confidence levels behave similarly.

A comparison of confidence intervals with respect to the function flatness is presented in Figure 1. As it may be expected, for flat functions (small $|\beta_2|$) confidence intervals are wider than for functions with big $|\beta_2|$.

A comparison of confidence intervals with respect to the experimental design is presented in Figure 2. It may be seen that there is no uniformly the best experimental design in the sense of the length of confidence interval. In fact, there are two designs which may be considered as "good" ones. These are A and B designs. Note that if maximum is near zero, i.e. the middle of x interval, than shortest confidence intervals are obtained under A design. Elsewhere the design B is better.

5. Conclusions

In the simulation study it appeared that instead of exact confidence interval, the Student confidence interval may be applied even for such small samples as $n = 12$. Also the strong dependence of the length on the point of maximum, function flatness and experimental design is showed. The most interesting question is, is it possible to find the uniformly the best experimental design.

It appeared that in some situations the obtained confidence interval goes out of range of observed x, i.e. its right end is greater than 1 (or left is smaller than -1). There are at least two reasons of such effects. The first one is connected with properties of the point estimator of maximum. Note that under model assumptions the distribution of the estimator $\hat{\varphi}$ is the Cauchy one.

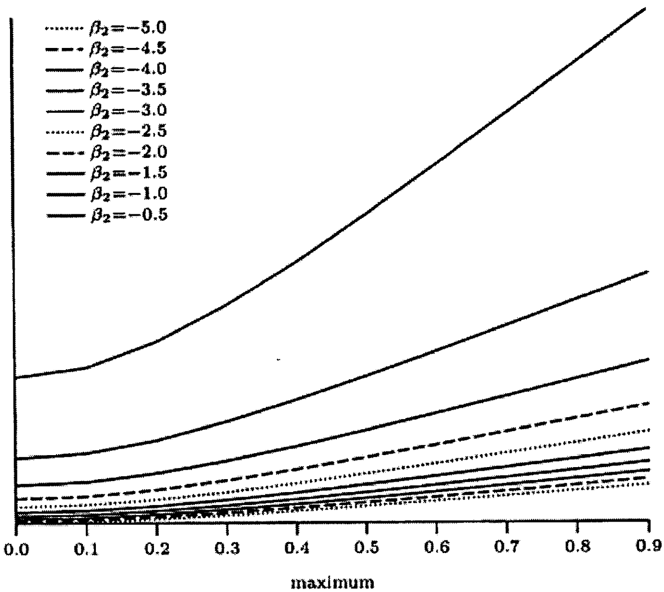

Figure 1. Dependence of length on function flatness

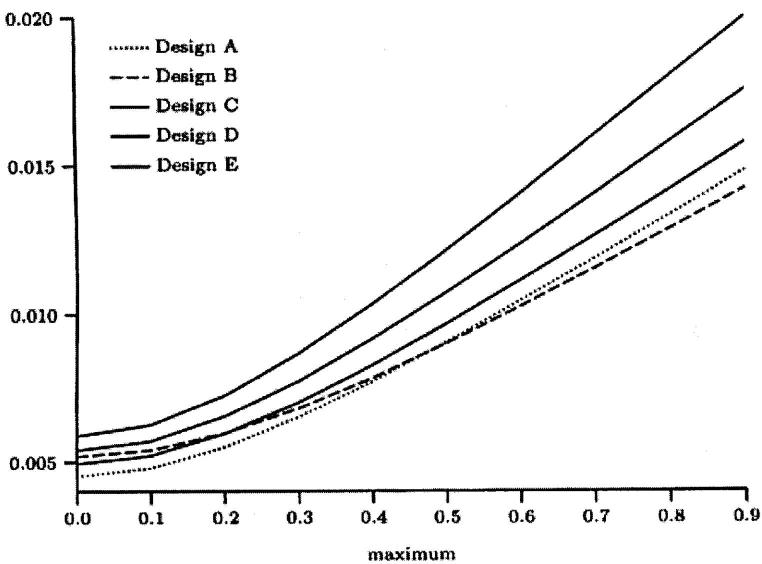

Figure 2. Dependence of length on experimental design

The second reason is as follows. The regression function was observed on the finite $[-1; 1]$ interval, but presented confidence intervals do not take it into account. Formally the model of quadratic regression $Y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \varepsilon_i$ $(i = 1, ..., n)$, should be considered with additional constrains: $-1 \leq x_i \leq 1$ for $i = 1, \ldots, n$.

References

Buonaccorsi J. P. (1985). Ratios of linear combinations in the general linear model. *Comm. in Statist.* **14**, 635–650.

Buonaccorsi J. P., Iyer H. K. (1984). A comparison of confidence regions and designs in estimation of a ratio. *Comm. in Statist.* **13**, 723–741.

Filler E. C. (1941). The biological standarization of insulin. Supplement to the *Journal of the Royal Statistical Society* **7**, 1–64.

Serfling R. J. (1980). *Approximation theorems of mathematical statistics.* John Wiley & Sons.

Received 20 February 2003; revised 12 September 2003

Porównanie przedziałów ufności dla maksimum kwadratowej funkcji regresji

Streszczenie

W pracy rozważany jest problem przedziałowej estymacji punktu maksimum kwadratowej funkcji regresji. Porównane zostały długości oraz poziomy ufności trzech przedziałów ufności w zależności od położenia punktu maksimum, spłaszczenia funkcji oraz planu doświadczenia. Porównania zostały przeprowadzone w oparciu o symulacje komputerowe.

Słowa kluczowe: kwadratowa funkcja regresji, punkt maksimum, estymacja przedziałowa.