Robustness of confidence intervals for the maximum point of a quadratic regression against autocorrelation

Dorota Kozioł, Wojciech Zieliński

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

SUMMARY

Consider a problem of interval estimation of a maximum of a quadratic regression function in situation, when random errors are correlated. Our aim is to examine the robustness of confidence intervals for maximum of a quadratic regression function. In the paper lengths and confidence levels of confidence intervals are compared with respect to the correlation. The investigations are made on the basis of computer simulations.

1. Introduction

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 normally distributed random variables such that $E\varepsilon_i = 0$ and $D^2\varepsilon_i = \sigma^2$. The problem is in interval estimation of $\varphi = -\beta_1/2\beta_2$, assuming $\beta_2 < 0$, i.e. the point at which the regression function achieves its maximum.

Under assumption of independence of random errors at least two confidence intervals for φ are known: exact confidence interval and approximate Student confidence interval. But in many practical applications it appears that ε's are not independent, for example in growth's models. The question is what are the properties of the above mentioned confidence intervals in case of correlated errors. In what follows length and confidence level as a functions of correlated errors are investigated. The model with repeated measurements is considered, i.e. the model

$$Y_{ij} = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \varepsilon_{ij}, \quad i = 1, \ldots, m, \quad j = 1, \ldots, k$$

and it is assumed that $E\varepsilon_{ij} = 0$, $D^2\varepsilon_{ij} = \sigma^2$,

$$E\varepsilon_{i_1 j_1}\varepsilon_{i_2 j_2} = \begin{cases} \varrho^{|i_1 - i_2|}, & \text{if } j_1 = j_2 \\ 0, & \text{if } j_1 \neq j_2. \end{cases}$$

Results are obtained on the basis of computer simulations.

2. Confidence intervals

In matrix notation the considered model is of the form

$$(*) \qquad \boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

where $\boldsymbol{Y} = (Y_{11}, \ldots, Y_{1m}, \ldots, Y_{k1}, \ldots, Y_{km})'$ is the vector of observations, $\boldsymbol{X} = \mathbf{1}_k \otimes \boldsymbol{U}$ with $\boldsymbol{U} = [1 \quad x_i \quad x_i^2]_{i=1,\ldots,m}$ is the design matrix, $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)'$ is the vector of regression coefficients and $\boldsymbol{\varepsilon} = (\varepsilon_{11}, \ldots, \varepsilon_{1m}, \ldots, \varepsilon_{k1}, \ldots, \varepsilon_{km})'$ is the vector of random errors. Assume that matrix $\boldsymbol{U}$ is of full rank. If so, there exists the matrix $(\boldsymbol{X}'\boldsymbol{X})^{-1}$. Denote the elements of $(\boldsymbol{X}'\boldsymbol{X})^{-1}$ by ν^{ij}, i.e. $(\boldsymbol{X}'\boldsymbol{X})^{-1} = [\nu^{ij}]_{i,j=0,1,2}$. Let

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

be LSE estimators of $\boldsymbol{\beta}$ and σ^2, respectively ($n = km$). Assuming $\boldsymbol{\varepsilon} \sim N_n(\mathbf{0}, \sigma^2\boldsymbol{I}_n)$ we have

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

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

The first construction of confidence interval for φ is based on Fieller theorem (Fieller, 1940), and is called exact confidence interval. This confidence interval has the following form:

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

where

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

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

and $t(\alpha, n-3)$ is the critical value of the t distribution with $n-3$ degrees of freedom.

The second construction is based on the fact (Serfling 1980) that

$$\hat{\varphi} = \hat{\beta}_1/2\hat{\beta}_2 \sim AN(\varphi; \sigma^2\omega^2),$$

where

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

Application of the classical Student technique gives the following approximate confidence interval:

$$(S) \qquad (\hat{\varphi} \pm t(\alpha, n-3)S\hat{\omega}),$$

where $\hat{\omega^2} = (4\hat{\varphi}^2\nu^{22} + 4\hat{\varphi}\nu^{12} + \nu^{11})/(2\hat{\beta}_2)^2$.

3. Correlated errors

In many practical applications the problem of estimating φ occurs, but with not independent errors. Consider the growth of a plant. It is known that during its growth a plant achieves a maximum of its "possibilities". For example, during the evolution of a fruit the contents of some components (water, microelements, etc.) increases at the beginning of the growth, achieves maximum and then decreases. The example of the yield of tomatoes of specimen *Fireball* (Krzyśko 1990)is given in Figure 1.

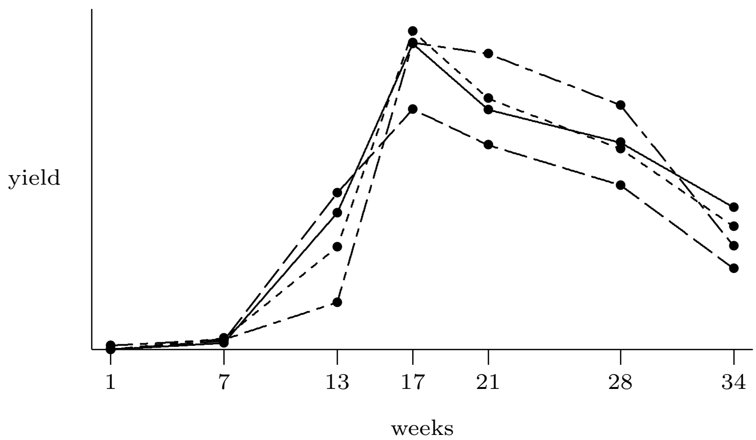

Figure 1. Yield of four plants of tomatoes *Fireball*

The practical problem is in estimation of the moment in which plant gains the maximum of its powers.

The main difference between the previous model and the growth model lies in properties the of random errors. In growth models those errors are not independent. We assume now that the correlation matrix of random errors ε is of the form $\boldsymbol{\sigma}^2(\boldsymbol{I}_k \otimes \boldsymbol{\Sigma})$ with $\boldsymbol{\Sigma} = [\varrho^{|i-j|}]_{i,j=1,\dots,m}$. Such a correlation structure is typical for $AR(1)$ process.

We are interested in robustness of confidence level and length of confidence intervals (E) and (S) against correlation. Note that if $\varrho \neq 0$, then

$$D^2\hat{\boldsymbol{\beta}} = \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'(D^2\boldsymbol{Y})\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1} = \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'(\boldsymbol{I}_k \otimes \boldsymbol{\Sigma})\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}.$$

Hence elements ν^{ij} in formulae (E) and (S) now depend on ϱ. Analytical form of this dependence is rather complicated and intractable (these are polynomials of m-th degree). So to estimate the confidence level as well as the length the Monte Carlo method was applied.

4. Simulation studies

In simulation studies we confine ourselves to $x \in [-1;1]$ interval. Note that every finite interval for x may be reduced to $[-1;1]$. We chose

$$x_i = -1 + \frac{2}{9}i, \qquad i = 0, 1, \dots, 9,$$

i.e. ten equally distributed points over the considered interval ($m = 10$). Such a choice should model time points which are equidistant (for example ten weeks). We observe $k = 5$ courses of regression function. Hence we have 5×10 observations. Also the $\sigma^2 = 0.1$ was taken. On such observations we build confidence interval for maximum and note its length and the fact if it hits a true maximum point. This procedure was repeated 1000 times and as the result we note the mean length as well as the empirical confidence level.

In our simulations we consider 100 quadratic regression functions given bellow:

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	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9

In columns there are functions with the same maximum point and of different flatness. In rows there are functions of the same flatness and of different maximum points. Because of symmetry we consider only positive maximum points. The constant β_0 is not important and we put it equal to 10.

Also we have to put some values for a correlation. To make our simulations as enhanced as possible we consider correlations from -0.9 to 0.9 by 0.1, i.e. we consider 19 values of correlation.

For our investigations we chose the approximate Student confidence interval, for two reasons. The first one is in similar properties of (E) and (S) in basic model ($\varrho = 0$). Properties of those confidence intervals were studied widely by many authors (Buonaccorsi (1985), Buonaccorsi and Iyer (1984), Kozioł and Zieliński (2003). The second reason is the fact that approximate Student confidence interval always exists.

5. Results

The results of simulations are shown in figures. For presentation we chose only functions with different maximum points and the same flatness ($\beta_2 = -5$). For other functions results are similar.

The comparison of length (Fig.2) of confidence intervals for different correlations shows that length of interval does not react significantly to correlation. The widest intervals are for correlation equal to zero (i.e., in the initial model). Then the length reduces. The minimal length of confidence interval is for maximal correlations. Such a behavior may be considered as a positive one: non zero correlation results in shortening of confidence intervals. Hence the length may be considered as a robust one against correlation.

The comparison of confidence levels (Fig.3) shows, that:

1. if the correlation is negative, the confidence level is at last the nominal one, i.e., the confidence level in the basic model which was taken to be 95% (such a behavior may be considered as a good one);
2. if the correlation is positive, the confidence level decreases gradually; the smallest confidence level is in situation, when correlation is about 0.6, after that we observe a rapid growth of the confidence level; the minimal value of the confidence level is between 80% of the nominal level (for maximum at 0) and 90% of that level (for maximum at 0.9).

Hence the confidence level may be considered as highly unrobust against correlation.

Our investigations were made for ten x points. It may be expected that increasing the number of x points would not change above conclusions.

Now, the question arises how to construct interval estimator of a maximum of a quadratic regression function in presence of correlation. In general, there are three approaches to the problem. The first one is to consider a model $(*)$ with $\varepsilon \sim N_{km}(0, \sigma^2 \boldsymbol{I}_k \otimes \boldsymbol{\Sigma})$ instead of $\varepsilon \sim N_{km}(0, \sigma^2 \boldsymbol{I}_{km})$. This approach is analytically difficult because unknown correlation ϱ is involved in $\hat{\boldsymbol{\beta}}$ and S^2 in rather complicated way:

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'(\boldsymbol{I}_k \otimes \boldsymbol{\Sigma})^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'(\boldsymbol{I}_k \otimes \boldsymbol{\Sigma})^{-1}\boldsymbol{Y},$$

$$S^2 = \boldsymbol{Y}'((\boldsymbol{I}_k \otimes \boldsymbol{\Sigma})^{-1} - (\boldsymbol{I}_k \otimes \boldsymbol{\Sigma})^{-1}\boldsymbol{X}(\boldsymbol{X}'(\boldsymbol{I}_k \otimes \boldsymbol{\Sigma})^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'(\boldsymbol{I}_k \otimes \boldsymbol{\Sigma})^{-1})\boldsymbol{Y}/(n-3).$$

The second approach relies on "robustification" of known confidence intervals, i.e. in doing such modifications of confidence intervals and/or data after which the minimum over ϱ of confidence level is as near nominal one as possible.

In the third approach confidence intervals are constructed in the model with $\varrho = 0$ which are robust against correlation, i.e. such that $\sup_{\varrho} |\gamma(\varrho) - \gamma(0)|$ is minimal ($\gamma(\varrho)$ is the confidence level in case of correlation ϱ).

Above mentioned approaches are under consideration and results will be presented separately.

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 Standardization of Insulin, Supplement to the Journal of the Royal Statistical Society, 7, 1–64

Krzyśko M. (1990) Analiza dyskryminacyjna, WNT, Warszawa

Serfling R. J. (1980) Approximation Theorems of Mathematical Statistics, John Wiley & Sons.

Kozioł D., Zieliński W. (2003) Comparison of confidence intervals for maximum of a quadratic regression function, Biometrical Letters (submitted).

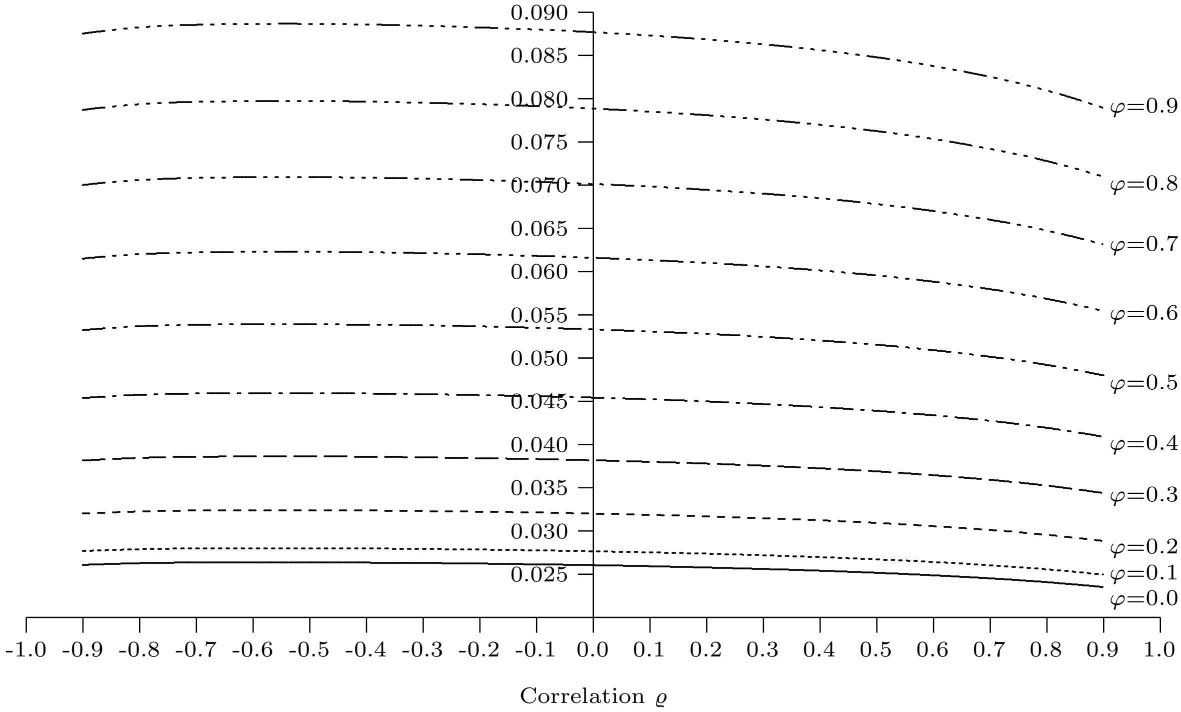

Figure 2. Empirical length of Student confidence interval

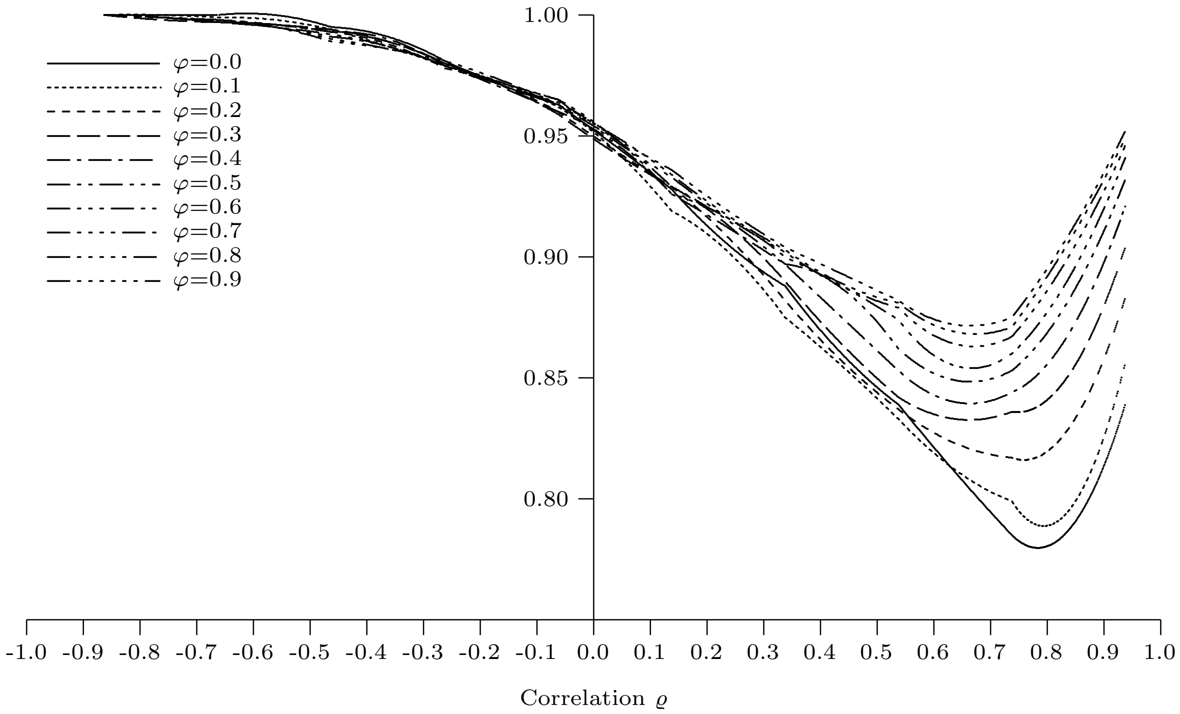

Figure 3. Empirical confidence level of Student confidence interval