



Supervisor's review of the bachelor's thesis

by

Daniel Pacák

# LEAST ABSOLUTE DEVIATIONS

Supervisor Jan Ámos Víšek

Thesis is devoted to the study of the *Least Absolute Deviations* (LAD) estimator of the regression coefficients in the linear model. Historically this method was used for the identification of the model earlier than the *Ordinary Least Squares* (OLS) - see Galileo Galilei (1632), Roger Boscovich (1757) or Pierre Simon Laplace (1793), of course no one of them spoke about the regression model because the name was firstly used by Francis Galton (1886) nearly century later. Although LAD has better behavior in the sense of a resistance against contamination of data than OLS, it was (and is) employed less frequently. The reasons are easy to see. OLS has simple geometric interpretation as well as it allows for a formula which is easy to implement. Moreover, the proofs of statistical properties (unbiasedness, consistency or asymptotic normality) is nearly trivial while - as the present thesis very clearly demonstrates - the proofs for LAD are rather complicated. Moreover, the computation is also a bit more intricate than the computation for OLS.

The thesis in its first part proves strong consistency and asymptotic normality of LAD, and discusses unbiasedness. Nice result is the demonstration that LAD is maximum likelihood estimator in the case when the distribution of disturbances is the Laplace one. The rest of thesis is devoted to the numerical study.

First of all, let me stress that Daniel Pacák wrote the thesis completely “independently of me”. He arrived three or four times on some consultations and he discusses with me more or less question from the meta-level, i. e. what is present point of view on point estimation, how far the numerical study should illustrate the theoretical results or which philosophy of contamination is to be used, etc. Sometimes he came to clarify some problems from my lecture on *Robust statistics and econometrics*.

Form of thesis is (nearly) excellent. Thesis is esasy to read because the notations and text is really nice mathematics, moreover written in a compact, dense form. It was really pleasure to read it. The word “nearly” indicates that there is some tax paid to inexperience of writing such text, so that there are some misprints or omissions but the percentage is very low - much lower than in usual bachelor thesis.

Numerical study is well designed and the patterns of results are clearly presented so that it

is easy to orient in them and to learn relevant information. Comparison with the results for OLS confirm the me sensetheoretical ideas. Finally, the results for contaminated data (the philosophy of *breakdown point* was utilized, i. e. thee data which were on the outskirts of data cloud were corrupted) show that LAD is able to cope with the situation much better than OLS.

I suppose that there is no doubt that the thesis should be appreciated by the grade 1 but if possible Daniel surely deserves a praise by dean.

## References

- [1] Boscovich, R. J. (1757): De litteraria expeditione per pontificiam ditionem, et synopsis amplioris operis, ac habentur plura eius ex exemplaria etiam sensorum impressa. *Bononiensi Scientiarum et Artium Instituto Atque Academia Commentarii* 4, 353-396.
- [2] Galilei, G. (1632): *Dialogo dei masimi sistemi*.
- [3] Galton, F. (1886): Regression towards mediocrity in hereditary stature. *Journal of the Anthropological Institute* vol. 15, pp. 246-263.
- [4] Laplace, P. S. (1793): Sur quelques points du systeme du mode. *Memoires de l'Academic Royale des Sciences de Paris*, 1-87.

SUMMARY OF POINTS AWARDED		
CATEGORY		POINTS
Literature	(max. 20 points)	20
Methods	(max. 30 points)	30
Contribution	(max. 30 points)	30
Manuscript Form	(max. 20 points)	18
TOTAL POINTS	(max. 100 points)	98
GRADE	(1 – 2 – 3 – 4)	1

Prague, August 30, 2017

.....  
(signature)