Thailand Statistician
January 2024; 22(1): 192-201
http://statassoc.or.th
Contributed paper

Intra and Outer Bootstrap Methods in Deming Regression Analysis

Necati Alp Erilli
Sivas Cumhuriyet University, Department of Econometrics, Turkey.
Corresponding author; e-mail: aerilli@cumhuriyet.edu.tr

Received: 20 January 2021
Revised: 7 April 2022
Accepted: 13 April 2022

Abstract

Deming regression is a type of regression method that fits a regression line when the measurements of both the explanatory variable and the response variable are assumed to be subject to normally distributed errors. Recall that in ordinary least squares regression, the explanatory variable is assumed to be measured without error. Bootstrapping is a type of resampling where large numbers of smaller samples of the same size are repeatedly drawn, with replacement, from a single original sample. In this study, two different bootstrap methods are introduced as intra and outer bootstrap. It is proposed to use the introduced bootstrap methods together with Deming regression. This study provides an investigation of outer bootstrap Deming methods in cases where outliers are high, and intra bootstrap Deming methods in cases where the central spread is high. In the application part, the proposed methods on 7 different data sets previously used in the literature were used. It was seen that the intra bootstrap method results had less mean square error value than the classical Deming regression results in all datasets except one. According to the results obtained from the study, it was seen that the estimation values made with the intra bootstrap method gave more successful results than the classical bootstrap and outer bootstrap. Intra bootstrapping method will be a guide for researchers who will work with Deming regression and data with few observations.

Keywords: Type-II models, resampling, mean square error, quartiles, measurement error.

1. Introduction

Regression analysis is the most used estimation method among statistical methods. Regression analysis can be summarized as the study of analyzing the value of the response variable with the help of one or more explanatory variables. The two main functions of the method that are used are parameter estimation and prediction. For these functions to be performed successfully, some assumptions must be provided.

One of these assumptions is that the explanatory variable is measured with no error. For the standard regression model with a single explanatory variable measured with error, it is well known that the estimated regression coefficient, on average, is attenuated toward the origin, and measurement error in the independent variable leads to bias in the estimation of a regressor coefficient. The following examples can be given for these situations: The variables may not be measurable (e.g. mental conditions, feelings, color, ability etc.) , the variables are clearly defined, but it is hard to take correct observations (for example, income status is often reported as full value or $1,000 and its multiples), the variable is well understood, but it is qualitative in nature (the likes of a movie are evaluated by IMDB scores) or the variable is conceptually well defined, but it is not possible to take a correct observation on it (job experience is measured by the number of years of working). Regression techniques applied in situations where explanatory variables are assumed to contain no measurement error are called the Type-I regression technique. They are the most powerful estimation methods, as assumptions are provided. The ordinary least squares method can be cited as the best example of this.

In some studies, it is not known whether the measurement error is caused by the response variable or the explanatory variable, or both. Type-II regression techniques are regression techniques that calculate the measurement errors in the response and explanatory variables at the same time. In general, these techniques are based on the logic of taking into account the errors in both variables as a result of taking the distances of the observation values calculated perpendicular to the regression equation obtained or depending on the amount of error. Orthogonal regression, Deming regression, York regression, Bland-Altman regression techniques can be considered as examples for Type-II regression.

The number of studies conducted with Deming regression and bootstrap technique is very few in the literature. Francq (2014) compared six different bootstrap procedures to improve the coverage probabilities of the approximate confidence intervals for the parameters of the Deming and bivariate least square regressions. The bootstrapping in both methods provide very similar results. Benni et al. (2018) tested the single and double bootstrap methods on Deming regression and random coefficient model in their study. As for the results, single bootstrap results were found more efficient than double bootstrap in all methods. Fitrianto et al. (2020) compared jackknife and bootstrap resampling methods in orthogonal regression analysis similar to Deming regression. They showed that jackknife performed better in constructing confidence interval than the bootstrap in their study.

2. Method

2.1. Deming regression analysis

Deming regression is a technique for fitting a straight line to two-dimensional data where both variables, response, and explanatory, are measured with error. This is different from simple linear regression where only the response variable, Y, is measured with error. Deming regression is often used for method comparison studies in clinical chemistry or biostatistics to look for systematic differences between two measurement methods.

Deming proposed minimizing the function that will give the correct equation that best fits the observation values in case both variables have measurements with the error (Deming 1943). The error sum of squares value to be minimized in the Deming regression technique can be calculated as given in (1).

$$SSE = \sum \left\{ \left(x_i - X_i \right)^2 + \lambda \left(y_i - Y_i \right)^2 \right\}. \quad (1)$$

To predict the regression line with the Deming technique, it is necessary to know the value λ, which is the ratio of the square analytical standard deviations of the X and Y methods:

$$\lambda = \frac{\sigma_x^2}{\sigma_y^2}. \quad (2)$$

This calculated λ value allows us to determine the angle by minimizing the squared deviation sums on the line. The formulas and expressions used in the parameter estimates are as given in (3) and (4) (Linnet 1993):

$$\beta_1 = \frac{(\lambda q - u) + \sqrt{(u - \lambda q)^2 + 4\lambda p^2}}{2\lambda p}, \tag{3}$$

$$\begin{aligned} \beta_0 &= \bar{y} - \bar{x}\hat{\beta}_1, \\ u &= \sum (x_i - \bar{x})^2, \\ q &= \sum (y_i - \bar{y})^2, \\ p &= \sum (x_i - \bar{x})(y_i - \bar{y}). \end{aligned} \tag{4}$$

Many studies have been conducted on Deming regression analysis, especially in the field of clinical chemistry and biostatistics. The general purpose of these studies is to produce algorithms to find the coefficient calculations that give the optimal λ value and the minimum error value (Linnet 1998; Martin 2000; Schall et al. 1980; Payne 1984; Smith et al. 1980).

2.2. Bootstrap method

The bootstrap method was first proposed by (Efron 1979). This method aims to obtain smaller standard errors, more reliable parameter estimators, and create narrower confidence intervals by replacing and resampling randomly from the existing data set to generate very large data sets. The Bootstrap method is far from dense mathematical formulas, limited assumptions, and easy to understand and use. This method creates a strong potential in situations where it is difficult or impossible to obtain the sampling distribution of the estimator by asymptotic methods (Simon and Bruce 1991).

The Bootstrap method is a data-based simulation method for statistical inference which can be used to study the variability of estimated characteristics of the probability distribution of a set of observations and provide confidence intervals for parameters in situations where these are difficult or impossible to derive in the usual way (Everitt 1995).

The basic idea of the bootstrap involves repeated random sampling with replacement from the original data $x = x_1, x_2, ..., x_n$ to produce random samples of the same size n of the original sample. Here the aim is to mimic in an appropriate manner the way the sample is collected from the population in the bootstrap samples from the observed data. The "with replacement" means that any observation can be sampled more than once in each bootstrap sample. It is important because sampling without replacement would simply give a random permutation of the original data, with many statistics such as the mean being exactly the same (Campbell 2001). The point about the bootstrap is that it produces a variety of values, whose variability reflects the standard error that would be obtained if samples were repeatedly taken from the whole population (Walters and Campbell 2005). The confidence interval derived from the bootstrap agrees very closely with the one derived from statistical theory. Bootstrap methods are intended to simplify the calculation of inferences, producing them in an automatic way even in situations much more complicated (Efron and Tibshirani 1991).

Using the Monte Carlo approximation, bootstrapping can be applied to many practical problems such as parameter estimation in time series, regression, and analysis of variance problems, and even to problems involving small samples. There is also a strong temptation to apply the bootstrap to several complex statistical problems where we cannot resort to classical theory to resort to. At least for some of these problems, it is recommended that the practitioner try the bootstrap. Only for cases where there is theoretical evidence that the bootstrap leads us astray would advise against its use (Chernick 2008).

The bootstrap method does not require any assumptions regarding the underlying population distribution like other methods. Unlike the normal theory method, the bootstrap method can be applied for any sample size. The only assumption underlying the bootstrap method is that the sample selected is representative of the population that it is sampled from. However, this underlying assumption is inherent in any statistical procedure including the normal theory method and the specific confidence interval methods. Thus, the bootstrap method provides an alternative means for estimating the audit value from skewed populations and small sample sizes (Muralidhar et al. 1991). A main concern in small samples is that with only a few values to select from, the bootstrap sample will underrepresent the true variability since observations are frequently repeated and bootstrap samples, themselves, can repeat (Chernick 2008).

2.3. Intra bootstrap-Outer bootstrap

While resampling in the bootstrap method, sampling is done from all the data within the framework of certain rules. In some data structures, observations are spread around the mean, while in some data there may be more spread in the extreme regions. Particularly in data with small samples, spreading to certain parts may cause deviations in the data. The effect of these spreads on the regression can sometimes cause the calculated estimates to be biased. In this study, instead of classical bootstrap results for data with such observations, new bootstrap approaches are proposed for observations that spread to certain regions. Thus, it is aimed to minimize the deviations arising from the general bootstrap calculations by making bootstrap according to the regions separated by quartiles for both the observations spreading around the mean and the data concentrated in the extreme regions. For this, the regions separated by 4 quartiles were determined as the calculation area. The reason for making bootstrap to regions separated by quartiles is that the algorithm proposed in this study will be studied especially for data with small observation volumes. For data with larger observation volumes, deciles can be tried instead of quartile. The region from the first observation to the first quartile and the region from the third quartile to the last observation were determined as the external region, and between the first quartile and the third quartile as the inner region, and the calculations were made separately for these regions.

Let's divide a data set into 4 equal parts with 3 quartiles. Let the data between the Q_1 and Q_3 quartiles be defined as intra and the data outside as outer. When we do the classic bootstrap process from intra data, let's name it intra bootstrap. Similarly, when we do from outer data, let's call it outer bootstrap. In this way, instead of choosing a random sample from all of the data, it will be possible to derive bootstrap data suitable for the distribution of the data by selecting data from regions separated by quartiles. Outer bootstrap is likely to be more suitable for data with high outliers and intra bootstrap sampling for data close to normal distribution. Bootstrap selections defined in Figure 1 are shown.

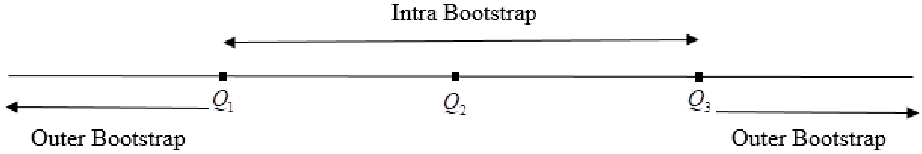

Figure 1 Structure of intra-outer bootstrap

In this way, the effect of the general distribution of observations in the data on the data is also revealed. For the data concentrated around the arithmetic mean, intra, and for the data concentrated around the initial or trailing outliers, the effect of outer bootstrap methods will be greater.

3. Application

In the application part, seven different data sets were used, and bootstrap Deming regression methods were compared according to the classical method. For each data set, the average square error values of the models obtained by the classical Deming regression method, classical bootstrap, intra bootstrap, and outer bootstrap methods obtained from all data were compared. Classical Deming regression analysis was performed with the "Deming" package of the R package program. Deming regression with Intra and outer bootstrap methods performed with Microsoft Excel codes which authors prepared. Each data set is also given in the Appendix. The mean square error (MSE) model selection criterion was used to compare the predicted values obtained from all data. MSE uses squared units rather than the natural data units, the interpretation is less intuitive (Frost 2019). MSE is an estimator measures the average of the squares of the errors and its formula is given in (5).

$$MSE = \frac{\sum (y_i - \hat{y}_i)^2}{n}. \quad (5)$$

Bootstrap data generation (classical, intra and outer bootstraps) is done by writing code in the R package program. For each data set, samples were drawn up to 50% of the data using the Bootstrap technique. The predictions of the expanded data with Bootstrap were realized by simulating 500 trials. The arithmetic means of the Deming regression parameter and lambda results obtained with 500 trials were taken as the final bootstrapped Deming regression results. Parameter estimation values obtained from each model and MSE values calculated for each model are given in Table 1.

Table 1 Predicted parameter and MSE values for all data sets

	β_0	β_1	MSE
		Data 1	
Deming	0,1052	0,9687	0,1544
Deming Bootstrap	0,1692	0,9659	0,1564
Deming Intra Bootstrap	0,6167	0,8853	0,1491
Deming Outer Bootstrap	0,0708	0,9706	0,1553
		Data 2	
Deming	0,0179	0,9183	0,00668
Deming Bootstrap	0,014	0,919	0,00663
Deming Intra Bootstrap	0,0123	0,9179	0,00661
Deming Outer Bootstrap	0,0142	0,9395	0,00688
		Data 3	
Deming	–0,0589	1,0545	0,0249
Deming Bootstrap	–0,056	1,0495	0,0248
Deming Intra Bootstrap	–0,0248	1,0301	0,0245
Deming Outer Bootstrap	–0,0632	1,0538	0,0249
		Data 4	
Deming	0,412	0,8806	1,0345
Deming Bootstrap	0,4099	0,8689	1,0250
Deming Intra Bootstrap	0,4645	0,8693	1,0234
Deming Outer Bootstrap	0,3711	0,8921	1,0501

Table 1 (Continued)

	β_0	β_1	MSE
		Data 5	
Deming	0,3908	0,9422	0,5458
Deming Bootstrap	0,0847	0,9743	0,5370
Deming Intra Bootstrap	1,1487	0,8584	0,5001
Deming Outer Bootstrap	0,4722	0,9291	0,5097
		Data 6	
Deming	−14,722	1,0279	1443,9098
Deming Bootstrap	−29,6802	1,0538	1282,0427
Deming Intra Bootstrap	1,1289	0,9989	1197,2528
Deming Outer Bootstrap	−29,5376	1,0541	1278,2799
		Data 7	
Deming	−10,5641	0,3544	8,4793
Deming Bootstrap	−11,9587	0,3702	8,7194
Deming Intra Bootstrap	−10,7558	0,3585	8,5320
Deming Outer Bootstrap	−12,1572	0,3694	8,7369

As for the results in Table 1, we can see that the best MSE value in the first 6 datasets was obtained by the Deming intra bootstrap method. In the last data set, the best MSE value is obtained from the classical Deming regression model, while the Deming intra bootstrap MSE value has the second-best result. According to these results, it can be said that the effect of the Deming Intra Bootstrap method, namely the bootstrap samples taken between Q_1 and Q_3, on the model prediction results seems quite successful.

4. Discussion

The OLS method is still the most used in prediction studies. One of the basic assumptions of this method is that the explanatory variable has no measurement error. In many studies, it is not known whether the measurement error is caused by the response or explanatory variable. Type II regression techniques that do not require assumptions should be used when the measurements of response or explanatory variables contain errors. Deming regression technique is one of the most used Type II regression techniques.

In this study, new bootstrap techniques-intra and outer bootstrap methods are proposed to be used in Deming regression estimates. The results obtained from these bootstrap methods, which can also be named as inside and outside quartiles, were compared with the classical Deming regression results. It can be said that the intra bootstrap method was quite successful in studies conducted with 7 different data used in previous studies. In this study, it was seen that the prediction values obtained by the Intra method gave lower error values, especially in data structures with a small number of observations. This method is thought to be a guide for researchers who will work with Deming regression. The investigation of similar bootstrap structures in other Type-II regression analyzes such as York, Bland-Altman, orthogonal regression structures is left for future studies.

Acknowledgements

The author would like to thank the referees for their great scientific comments and suggestions that yielded great contributions to this study.

References

Benni PB, MacLeod D, Ikeda K, Lin HM. A validation method for near-infrared spectroscopy-based tissue oximeters for cerebral and somatic tissue oxygen saturation measurements. J Clin Mon Comp. 2018; 32(2): 269-284.

Campbell MJ. Statistics at square two: understanding modem statistical applications in medicine, London: Blackwell Publishing; 2006.

Chernick MR. Bootstrap methods: a guide for practitioners and researchers, New Jersey: John Wiley & Sons; 2008.

Cleophas TJ, Zwinderman AH. Regression analysis in medical research, for starters and 2nd levelers. Switzerland: Springer; 2018.

Deming EW. Statistical adjustment of data. New York: Dover Publishing; 1943.

Efron B, Tibshirani RJ. An introduction to the bootstrap. Washington DC: Chapman & Hall; 1993.

Efron B. Bootstrap methods: another look at the jackknife. Ann Stat. 1979; 7(1): 1-26.

Everitt BS. The Cambridge dictionary of statistics in the medical sciences. Cambridge: Cambridge University Press; 1995.

Fitrianto A, Yun TS, Wan Ahmad WZ. Application of resampling techniques in orthogonal regression. Int J Eng Res Tech. 2020; 13(12): 4118-4124.

Francq BG. Bootstrap in errors in variables regressions applied to methods comparison studies. Inf Med Slov. 2014; 19(1-2): 1-11.

Frost J. Regression analysis: an intuitive guide. Philadelphia: Statistics by Jim Publishing; 2019.

Hathaway NR. Some comments on linear regression analysis by the method of W.E. Deming. Clin Chem. 1980; 26(10): 1511.

Konings H. Use of Deming regression in method-comparison studies. Surv Immu Res. 1982; 1(4): 371-374.

Linnet K. Evaluation of regression procedures for methods comparison studies. Clin Chem. 1993; 39(3): 424-432.

Linnet K. Performance of Deming regression analysis in case of misspecified analytical error ratio in method comparison studies. Clin Chem. 1998; 44(5): 1024-1031.

Manuilova E, Schuetzenmeister A, Model F. Mcr: method comparison regression. R package, version 1.2.1 [Internet]. 2015 [cited 2021 August 15]. Available from: https://cran.r-project.org/web/packages/mcr/mcr.pdf.

Martin RF. General Deming regression for estimating systematic bias and its confidence interval in method-comparison studies. Clin Chem. 2000; 46(1):100-104.

Muralidhar K, Ames GA, Sarathy R. Bootstrap confidence intervals for estimating audit value from skewed populations and small samples. Sim. 1991; 56(2):119-127.

Payne RB. Calculation of slope by Deming's method from least-squares regression coefficients and imprecision. Clin Chem. 1984; 30(5): 807.

Schall Jr RF, Kern CW, Tenoso HJ. Test data matrix and results for linear regression analysis by method of W. E. Deming. Clin Chem. 1980; 26(2): 352.

Simon JL, Bruce P. Resampling: a tool for everyday statistical work. Chnc N Dir Stat Comp. 1991; 4(1): 22-32.

Smith DS, Pourfarzaneh M, Kamel RS. Linear regression analysis by Deming's method. Clin Chem. 1980; 26(7): 1105-1106.

Therneau T. Deming: Deming, Theil-Sen, Passing-Bablock and total least squares regression. R package, version 1.4 [Internet]. 2018 [cited 2021 August 20]. Available from: https://cran.r-project.org/web/packages/deming/vignettes/deming.pdf.

Walters SJ, Campbell MJ. The use of bootstrap methods for estimating sample size and analysing health-related quality of life outcomes. Stat Med. 2005; 24(7): 1075-1102.

Wicklin R. Deming regression for comparing different measurement methods [Internet]. 2019 [cited 2021 August 20]. Available from: https://blogs.sas.com/content/iml/2019/01/07/deming-regression-sas.html.

Zaiontz C. Deming Regression, real statistics using excel [Internet]. 2020 [cited 2021 August 15]. Available from: http://www.real-statistics.com/regression/deming-regression.

Appendices

Data 1 ($n = 10$) (Zaiontz 2020)

x	4,5	5,2	4	5,6	5,1	5,6	5,9	6,8	6,6	6,7
y	4,1	4,6	4,7	5,1	5,4	5,6	6,1	6,3	6,6	6,8

Data 2 ($n = 12$) (Konings 1982)

x	0,15	0,5	0,4	1,39	0,75	0,79	0,35	0,45	0,12	0,62	0,6	0,81
y	0,17	0,45	0,3	1,24	0,78	0,88	0,42	0,35	0,16	0,69	0,5	0,64

Data 3 ($n = 108$) (Manuilova et al. 2015)

Serum (x)	Plasma (y)	Serum (x)	Plasma (y)	Serum (x)	Plasma (y)	Serum (x)	Plasma (y)
0,82	0,79	1,2	1,29	0,9	0,87	1,09	1,24
1,83	1,62	1,01	0,87	1,34	1,3	1,27	1,29
1,39	1,36	1,33	1,13	1,32	1,04	2,06	2,08
0,81	1,3	0,87	0,82	1,02	0,92	1,21	1,31
1,72	1,88	1,66	1,33	1,04	1,01	0,77	0,82
3,23	3,35	1,41	1,14	1,27	1,21	1,06	1,28
1,23	1,06	1,08	0,93	1,08	1,11	1,02	1,15
1,37	1,34	1,17	1,09	1,1	1,21	1	1,04
1,72	1,56	1,77	1,89	1,39	1,54	1,47	1,51
0,96	0,94	1,33	1,57	1,07	0,99	0,97	1,13
0,76	0,69	1,23	1,29	0,79	0,79	1,7	1,84
1,15	1,16	1,21	1,13	1,02	1,06	1,78	1,94
0,95	0,71	1,2	1,11	0,94	1,02	0,93	1,32
1	0,83	1,49	1,61	1,99	2,15	0,92	1,36
1,52	1,44	0,89	0,79	0,7	0,56	1,59	1,88
1,56	1,26	1,09	1,19	0,84	1,01	0,68	0,73
2,45	2,36	1,03	0,96	0,66	0,59	0,7	0,78
1,85	1,9	1,07	1,25	1,1	1,17	1,08	1,01

Serum (x)	Plasma (y)	Serum (x)	Plasma (y)	Serum (x)	Plasma (y)	Serum (x)	Plasma (y)
0,89	0,95	0,89	0,87	0,9	0,82	0,83	0,9
0,82	0,77	1,39	1,36	2,06	2,09	1,17	1,3
1,01	0,82	0,96	0,86	1,31	1,44	0,98	0,96
0,96	0,88	1,06	1,03	1,03	1,22	0,77	0,61
3,38	3,42	1,17	0,86	1,61	1,77	0,91	1,27
1,31	1,15	0,9	0,8	1,52	1,6	1,1	1,05
1,1	0,96	1,39	1,22	1,11	1,13	0,8	1,12
1,26	1,12	0,91	0,86	0,85	1,06	0,85	0,72
1,15	1,02	2,28	2,25	1,2	1,38	0,82	0,87

Data 4 ($n = 30$) (Therneau 2018)

x (aas)	0	0	0,82	0	0,73	1,38	0,9	0,4	1,88	1,94
y (aes)	0	0,37	0,44	0,49	0,66	1,17	1,25	1,29	1,37	1,5
x (aas)	1,27	1,55	1,98	1,75	1,81	0,34	3,69	4,39	3,28	3,66
y (aes)	1,88	2,07	2,16	2,29	2,31	2,32	2,72	3,31	3,4	3,43
x (aas)	2,07	4,64	5,66	5,6	9,39	5,66	8,71	7,01	10,2	19,3
y (aes)	3,5	3,9	4,66	5,44	6,58	7,04	7,35	7,92	12,5	15,9

Data 5 ($n = 10$) (Hathaway 1980)

x	7	8,3	10,5	9	5,1	8,2	10,2	10,3	7,1	5,9
y	7,9	8,2	9,6	9	6,5	7,3	10,2	10,6	6,3	5,2

Data 6 ($n = 17$) (Cleophas and Zwinderman 2018)

x	512	430	520	428	500	600	364	380	658
y	494	395	516	434	476	557	413	442	650
x	445	432	626	260	477	259	350	451	
y	433	417	656	267	478	178	423	427	

Data 7 ($n = 65$) (Wicklin 2019)

micrograms (x)	kiloOhms (y)	micrograms (x)	kiloOhms (y)	micrograms (x)	kiloOhms (y)
169	45,5	76	18,2	69,6	18,8
130,8	33,4	77,8	18,3	66,7	7,4
109	23,8	74,2	15,7	64,4	8,2
94,1	19,8	73,1	13,9	63	15,5
86,3	20,4	182,5	55,5	61,7	13,7
78,4	18,7	144	38,7	61,2	9,2
76,1	16,1	123,8	35,1	62,4	12
72,2	16,7	107,6	30,6	58,4	15,2
70	11,9	96,9	25,7	171,3	48,7

micrograms (x)	kiloOhms (y)	micrograms (x)	kiloOhms (y)	micrograms (x)	kiloOhms (y)
69,8	14,6	92,8	19,2	136,3	36,1
69,5	10,6	87,2	22,4	111,9	28,6
68,7	12,7	86,3	18,4	96,5	21,8
67,3	16,9	84,4	20,7	90,3	25,6
174,7	57,8	83,7	20,6	82,9	16,8
137,9	39	83,3	20	78,1	14,1
114,6	30,4	83,9	18,8	76,5	14,2
99,8	21,1	82,7	21,8	73,5	11,9
90,1	21,7	160,8	49,9	74,4	17,7
85,1	25,2	122,7	32,2	73,9	17,6
80,7	20,6	102,6	19,2	71,9	10,2
78,1	19,3	86,6	14,7	72	15,6
77,8	20,9	76,1	16,6		