العنوان:	Statistical Inference for Some Robust Regression Estimators
المصدر:	مجلة البحوث التجارية
الناشر:	جامعة الزقازيق - كلية التجارة
المؤلف الرئيسي:	Al Assar, Moshera Adel Soliman
المجلد/العدد:	مج38, ع2
محكمة:	نعم
التاريخ الميلادي:	2016
الشهر:	يوليه
الصفحات:	23 - 46
رقم MD:	771802
نوع المحتوى:	بحوث ومقالات
اللغة:	English
قواعد المعلومات:	EcoLink
مواضيع:	الإحصاء، الاستدلال الإحصائي، الانحدار الخطي، القيم المتطرفة
رابط:	http://search.mandumah.com/Record/771802

21.   Pitselis, G., (2013): " A review on Robust Estimators Applied to Regression Credibility ". Journal of Computational and Applied Mathematics., Vol. 239., 231-249.

22.   Rousseeuw, P.J., (1984): " Least Median of Squares Regression ". JASA., Vol. 79., No.388., 871-880.

23.   Rousseeuw, P.J., and V.J. Yohai., (1984): " Robust Regression by Means of S- estimators, in Robust and Nonlinear Time Series ". Lecture Notes in Statistics., Vol. 26., 256-272.

24.   Rousseeuw, P.J., and A.M. Leroy., (1987): " Robust Regression and Outlier Detection ". John Wiley& Sons.

25.   Ryan, T.P., (2009): " Modern Regression Methods: Second Edition ". John Wiley& Sons., New York.

26.   Salibian – Barrera, M., and R.H. Zamar., (2002): " Bootstrapping Robust Estimates of Regression". The Annals of Statistics 30.

27.   Staudte, R.G., and S. J. Sheather., (1990): " Robust Estimation and Testing ". 1st ed., John Wiley& Sons., Inc., New York.

28.   Willems, G., A, S.V., R.H. Zamar., (2013): " Robust and Efficient Estimation of the Residual Scal in Linear Regression ". Journal of Multivariate Analysis., Vol. 239., 231-249.

29.   Yohai, V. J., (1987): " High Breakdown- Point and High Efficiency Robust Estimates for Regression ". The Annals of Statistics., Vol. 15., No. 2., 642-656.

10.   Fox, J., and S. Weisberg., (2012): " Bootstrapping Regression Models in R". An Appendix to An R Companion to Applied Regression, Second Edition.

11.   Hample, F.R., R, E.M., R, P.J., and W.A. Stahel., (1986): " Robust Statistics: The Approach Based on Influence Functions ". John Wiley& Sons., New York.

12.   Hample, F., (2000): " Robust Inference ". Research Report., No. 93. ftp://ftp.stat.math.eth.ch.

13.   Hogg, R. V., (1979): " Statistical Robustness one view of its use in applications Today". The American Statistical., Vol. 33., 108-115.

14.   Huber, P.J., (1964): " Robust Estimation of a Location Parameter ". The Annals of Mathematical Statistics., Vol. 35., 73-101.

15.   Kenndy, P., (1988): "A Guide to Econometrics ". 4th ed., Black Well Publishers., Inc., USA.

16.   Koenker, R., and G. Bassett., (1982): " Tests of Linear Hypotheses and L1 Estimation ". Econometrica., Vol. 50., No. 6., 1577-1583.

17.   Koller, M., and W.A. Shatel., (2011): " Sharpening Wald- Type Inference in Robust for Small Samples". Computational Statistics and Data Analysis., Vol. 55., 2504-2515.

18.   Montgomery, D.C., and Peck, E.A., (1982): " Introduction to Linear Regression Analysis". John Wiley.

19.   Onder, O., and Ozet., (2001): " Least Median Squares: A Robust Regression Technique ". EgeAkademikBakis.

20.   Onder, A. O., and A. Zaman., (2005): " Robust Tests for Normality of Errors in Regression Models". Economics Letters 86., 63-68.

References

1. Abed El-Salam, M.El., (2000): " Methods of Estimation in Linear Regression Analysis with Outliers and Non-normal Errors ". P.H.D., Department of Statistics., Faculty of Commerce., Zagazig University., Egypt.

2. Alma, O.G., (2011): " Comparison of Robust Regression Methods in Linear Regression ". Int.J. Contemp. Math. Sciences., Vol. 6., No. 9., 409-421.

3. Angela, S. R., P. C. D., N. N. S and D. M. Mount., (2007): " On The Least Trimmed Squares Estimator ".

4. Bagheri, A.M.H., G, M., and S. Eftekhari., (2010): " A Comparison of Various Influential Points Diagnostic Methods and Robust Regression Approaches: Reanalysis of Interstitial Lung Disease Data ". Applied Mathematical., Vol. 4., No. 28., 1367-1386.

5. Bassett, G. W., (1991):" Equivariant, Monotonic, 50% Breakdown Estimators ". The American Statistical Association., Vol. 45., No. 2., 135-137.

6. Ben-Gel, I., (2005): " Outlier Detection ". Data Mining and Knowledge Discovery Handbook: A Complete Guide for Practitioners and Researchers., Kluwer Academic Publishers., ISBN., WWW.eng.tau.ac.il/bengal/outlier.pdf.

7. Draper, N.R., and H. Smith., (1998): " Applied Regression Analysis ". 3rd ed., John Wiley& Sons., INC., New York.

8. Efron, B.,(1979): " Bootstrap Methods: Another Look at The Jackknife". The Annals of Statistics., Vol. 7., No. 1., 1-26.

9. Enany, M.A.S., (1989): " Alternatives to Least Squares Model Fitting: An Application of Robust Regression ". Journal of Commercial Research., Zagazig University., Vol. 11., No. 3.

(6)Summary and Conclusions

- According to our empirical study results there are some conclusions which can presented as; For all sample sizes, MSE decreases with increasing sample size for all estimation methods under all error distributions as shown in the tables in appendix (B) for sample sizes 50 and 100. In the case of normal errors distribution, S is the most efficiency compared other methods, but MM is the more efficiency compared S method and is much better in outlier generating error distribution. Although robust methods are efficiency compared OLS but The performance of MM method is high efficiency in the case of Normal errors distribution.Efficiency at 10% contamination is higher than 30% contamination for small sample sizes.

- In the case of Chi-square errors distribution, MSE decreases with increasing sample size, MH is the most efficiency compared other methods when the sample size is small, but the greater the sample size to be MM is the most efficiency. RE for all methods increase as sample sizes increases.

- In the case of Cauchy errors distribution, MSE decreases with increasing sample size, RE increase as sample size increases. MM gives better efficiency of all parameter estimates than other methods then S is the most efficiency.

- The performance of MM method is high efficiency in the case of Cauchy errors distribution, and can be used S method Where it also gives high efficiency.

	MSE	478.5	0.081	0.084	0.064	0.091	0.115	0.067
	RE	1.00	5907.4	5696.4	7476.6	5258.2	4160.9	7141.8
	$\hat{\beta}_1$	0.964	0.995	0.994	0.997	1.005	1.000	1.004
n=50	Bias	-0.036	-0.005	-0.006	-0.003	0.005	0.000	0.004
	MSE	65.37	0.018	0.019	0.014	0.023	0.025	0.016
	RE	1.00	3631.7	3660.5	4669.3	2842.2	2614.8	4085.6
	$\hat{\beta}_2$	1.812	1.504	1.502	1.502	1.505	1.502	1.500
	Bias	0.312	0.004	0.002	0.002	0.005	0.002	0.000
	MSE	89.34	0.029	0.031	0.021	0.032	0.036	0.021
	RE	1.00	3080.7	2881.9	4254.3	2791.9	2481.7	4254.3
	$\hat{\beta}_0$	2.213	0.494	0.491	0.490	0.498	0.492	0.495
	Bias	1.713	-0.006	-0.009	-0.010	-0.002	-0.008	-0.005
	MSE	1925.0	0.037	0.039	0.031	0.044	0.064	0.033
	RE	1.00	52027	49359	62096	43750	30078	58333
	$\hat{\beta}_1$	1.750	1.006	1.005	1.004	0.992	0.992	0.997
n=100	Bias	0.750	0.006	0.005	0.004	-0.008	-0.008	-0.003
	MSE	277.8	0.012	0.013	0.009	0.013	0.018	0.009
	RE	1.00	23150	21369	30867	21369	15433	30867
	$\hat{\beta}_2$	2.302	1.499	1.496	1.497	1.503	1.501	1.499
	Bias	0.802	-0.001	-0.004	-0.003	0.003	0.001	-0.001
	MSE	514.6	0.011	0.011	0.009	0.012	0.016	0.009
	RE	1.00	46782	46782	57178	42883	32163	57178

		OLS	MH	MHP	MM	LTS	LMS	S
	$\hat{\beta}_2$	1.508	1.504	1.505	1.504	1.484	1.497	1.505
	Bias	0.008	0.004	0.005	0.004	-0.016	-0.003	0.005
	MSE	0.036	0.028	0.030	0.029	0.107	0.115	0.063
	RE	1.00	1.286	1.200	1.241	0.336	0.313	0.571
n=100	$\hat{\beta}_0$	4.496	4.194	4.358	4.070	3.036	3.059	3.233
	Bias	3.996	3.694	3.858	3.570	2.536	2.559	2.733
	MSE	16.06	13.73	14.97	12.85	6.755	6.863	7.630
	RE	1.00	1.170	1.073	1.250	2.377	2.340	2.105
	$\hat{\beta}_1$	0.992	0.995	0.992	0.996	1.013	1.020	1.007
	Bias	-0.008	-0.005	-0.008	-0.004	0.013	0.020	0.007
	MSE	0.021	0.016	0.017	0.016	0.068	0.060	0.036
	RE	1.00	1.313	1.235	1.313	0.309	0.350	0.583
	$\hat{\beta}_2$	1.492	1.497	1.495	1.499	1.497	1.503	1.503
	Bias	-0.008	-0.003	-0.005	-0.001	-0.003	0.003	0.003
	MSE	0.023	0.018	0.019	0.018	0.071	0.066	0.040
	RE	1.00	1.278	1.211	1.278	0.324	0.348	0.575

Table (4): The Results of Estimation parameters, Bias, MSE and RE for The Cauchy Distribution : Cauchy (50,0,1) and Cauchy (100,0,1)

Out. Per.	Crite ria	OLS	Some Robust Methods					
			MH	MHP	MM	LTS	LMS	S
	$\hat{\beta}_0$	-0.089	0.499	0.499	0.496	0.495	0.506	0.492
	Bias	-0.589	-0.001	-0.001	-0.004	-0.005	0.006	-0.008

P0=30%		OLS	MH	MHP	MM	LTS	LMS	S
	Bias	14.80	9.467	0.058	0.061	0.105	0.244	0.094
	MSE	224.4	138.2	0.665	0.671	2.465	2.573	0.827
	RE	1.00	1.624	337.4	334.4	91.03	87.21	271.3
	$\hat{\beta}_1$	0.732	0.896	1.007	1.006	0.993	0.990	0.978
	Bias	-0.268	-0.104	0.007	0.006	-0.007	-0.010	-0.022
	MSE	1.416	1.039	0.003	0.003	0.008	0.007	0.115
	RE	1.00	1.363	472.0	472.0	177.0	202.3	12.31
	$\hat{\beta}_2$	0.997	1.272	1.447	1.451	1.454	1.484	1.501
	Bias	-0.503	-0.228	-0.053	-0.049	-0.046	-0.016	0.001
	MSE	1.306	0.965	0.004	0.003	0.009	0.007	0.044
	RE	1.00	1.353	326.5	435.3	145.1	186.6	29.68

Table (3): The Results of Estimation parameters, Bias, MSE and RE for The Chi-Square Distribution : Chisq (50,4) and Chisq (100,4)

Out. Per.	Criteria	OLS	Some Robust Methods					
			MH	MHP	MM	LTS	LMS	S
n=50	$\hat{\beta}_0$	4.528	4.231	4.383	4.113	3.272	3.263	3.364
	Bias	4.028	3.731	3.883	3.613	2.772	2.763	2.864
	MSE	16.40	14.09	15.25	13.24	8.330	8.307	8.571
	RE	1.00	1.164	1.075	1.239	1.969	1.974	1.913
	$\hat{\beta}_1$	0.986	0.990	0.990	0.991	1.005	1.003	0.997
	Bias	-0.014	-0.010	-0.010	-0.009	0.005	0.003	-0.003
	MSE	0.054	0.042	0.047	0.043	0.160	0.156	0.094
	RE	1.00	1.286	1.149	1.236	0.338	0.346	0.574

	RE	1.00	0.433	0.765	0.684	0.149	0.351	0.112
	$\hat{\beta}_1$	1.009	1.003	1.009	1.006	1.058	1.071	0.979
	Bias	0.009	0.003	0.009	0.006	0.058	0.071	-0.021
	MSE	0.00009	0.00001	0.00008	0.00003	0.0035	0.005	0.0014
P0=0%	RE	1.00	9.00	1.125	3.00	0.026	0.018	0.064
	$\hat{\beta}_2$	1.447	1.464	1.455	1.459	1.514	1.479	1.472
	Bias	-0.053	-0.036	-0.045	-0.041	0.014	-0.021	-0.028
	MSE	0.003	0.001	0.002	0.002	0.0003	0.001	0.002
	RE	1.00	3.00	1.50	1.50	10.00	3.00	1.50
	$\hat{\beta}_0$	5.458	0.739	0.538	0.543	0.459	0.457	0.560
	Bias	4.958	0.239	0.038	0.043	-0.041	-0.043	0.060
	MSE	26.93	0.064	0.003	0.0029	0.0154	0.024	0.017
	RE	1.00	420.8	8977	8977	1749	1122	1584
	$\hat{\beta}_1$	0.901	1.008	1.009	1.006	1.040	1.039	0.979
	Bias	-0.099	0.008	0.009	0.006	0.040	0.039	-0.021
P0=10%	MSE	0.562	0.002	0.0004	0.0003	0.003	0.004	0.002
	RE	1.00	281.0	1405	1873	187.3	140.5	281.0
	$\hat{\beta}_2$	1.279	1.458	1.451	1.457	1.453	1.437	1.483
	Bias	-0.221	-0.042	-0.049	-0.043	-0.047	-0.063	-0.017
	MSE	0.524	0.003	0.0026	0.002	0.0066	0.006	0.002
	RE	1.00	174.7	201.5	262.0	79.39	87.33	262.0
	$\hat{\beta}_0$	15.30	9.967	0.558	0.561	0.605	0.744	0.594
	Bias	14.80	9.467	0.058	0.061	0.105	0.244	0.094

		OLS	MH	MHP	MM	LTS	LMS	S
	$\hat{\beta}_2$	1.459	1.716	1.717	1.722	1.706	1.731	1.738
	Bias	-0.041	0.216	0.217	0.222	0.206	0.231	0.238
	MSE	1.594	0.050	0.048	0.0499	0.045	0.056	0.059
	RE	1.00	31.88	33.21	31.94	35.42	28.46	27.02
P0=30 %	$\hat{\beta}_0$	15.19	10.309	0.556	0.566	0.359	0.221	0.454
	Bias	14.69	9.809	0.056	0.066	-0.141	-0.279	-0.046
	MSE	226.4	154.2	3.201	3.82	6.798	2.483	5.605
	RE	1.00	1.468	70.73	59.27	33.30	91.18	40.39
	$\hat{\beta}_1$	0.549	0.680	0.833	0.822	0.716	0.724	0.759
	Bias	-0.451	-0.320	-0.167	-0.178	-0.284	-0.276	-0.241
	MSE	3.038	2.259	0.0542	0.070	0.129	0.087	0.292
	RE	1.00	1.345	56.05	43.40	23.55	34.92	10.40
	$\hat{\beta}_2$	1.205	1.465	1.712	1.718	1.741	1.756	1.749
	Bias	-0.295	-0.035	0.212	0.218	0.241	0.256	0.249
	MSE	3.689	2.860	0.090	0.113	0.106	0.080	0.282
	RE	1.00	1.290	40.99	32.65	34.80	46.113	31.08

Table (2): The Results of Estimation parameters, Bias, MSE and RE for Outliers Generating for The Normal Distribution :Norm(100,0,1)

Out. Per.	Criteria	OLS	Some Robust Methods					
			MH	MHP	MM	LTS	LMS	S
	$\hat{\beta}_0$	0.536	0.550	0.542	0.544	0.407	0.465	0.539
	Bias	0.036	0.050	0.042	0.044	-0.093	-0.035	0.039
	MSE	0.0013	0.003	0.0017	0.0019	0.0087	0.0037	0.0116

Table (1): The Results of Estimation parameters, Bias, MSE and RE for Outliers Generating for The Normal Distribution :Norm(50,0,1)

Out. Per.	Criteria	OLS	Some Robust Methods					
			MH	MHP	MM	LTS	LMS	S
P0=0 %	$\hat{\beta}_0$	0.435	0.336	0.398	0.354	0.391	0.278	0.225
	Bias	-0.065	-0.164	-0.102	-0.146	-0.109	-0.222	-0.275
	MSE	0.0043	0.027	0.0104	0.0212	0.0122	0.054	0.079
	RE	1.00	0.159	0.413	0.203	0.352	0.080	0.054
	$\hat{\beta}_1$	0.836	0.783	0.812	0.781	0.734	0.783	0.744
	Bias	-0.164	-0.217	-0.188	-0.219	-0.266	-0.217	-0.256
	MSE	0.027	0.047	0.035	0.048	0.071	0.048	0.066
	RE	1.00	0.574	0.771	0.563	0.380	0.563	0.409
	$\hat{\beta}_2$	1.711	1.728	1.717	1.727	1.698	1.676	1.729
	Bias	0.211	0.228	0.217	0.227	0.198	0.176	0.229
	MSE	0.044	0.052	0.047	0.0517	0.039	0.034	0.053
	RE	1.00	0.846	0.936	0.851	1.128	1.294	0.830
P0=10 %	$\hat{\beta}_0$	5.256	0.578	0.410	0.379	0.313	0.264	0.208
	Bias	4.756	0.078	-0.899	-0.121	-0.187	-0.236	-0.292
	MSE	26.57	0.025	0.011	0.017	0.043	0.059	0.093
	RE	1.00	1063	2415.5	1562.9	617.9	450.3	285.7
	$\hat{\beta}_1$	0.759	0.810	0.817	0.801	0.741	0.801	0.733
	Bias	-0.241	-0.190	-0.183	-0.199	-0.259	-0.199	-0.267
	MSE	1.246	0.039	0.0344	0.409	0.069	0.043	0.074
	RE	1.00	31.95	36.22	3.046	18.06	28.98	44.50

The empirical methods are compared using the criteria of estimation parameters, bias, mean square errors (MSE) and relative efficiency. When comparing to the MSE of the ordinary least squares estimate (OLS) for each robust methods. Relative efficiency is calculated by:

$$R.E = \frac{MSE\ (OLS)}{MSE\ (Robust Method)}$$

Where $MSE = \frac{1}{R} \sum_{i=1}^{R} \left(\hat{\beta} - \beta \right)$, Where $\hat{\beta}$ is the estimated value of β, β is the true value and R denotes the number of replications. When R.E is greater than 1, the robust method is considered more efficient than OLS.(Koller and Stahel, 2011)

The empirical results are presented in tables (1) to (5) displaying the properties of different robust estimation methods for different distributions of errors with different percentages of outliers and different sample sizes. Tables (1 and 2) presents the empirical results for normal distribution. While table (3and 4) present the empirical results for chi-square distribution, and Cauchy distribution. In addition, the tables in this section present the parameter estimates, bias, MSE and RE.

where $\sqrt{\hat{\tau}}$ is equal to the function (8), analogous to the one that is derived from the Z-test. For finite sample sizes, this asymptotic approximation can be insufficient. To better align the actual coverage probability with its nominal value (=95%), substitute the value 1.96 with the quintile value of the corresponding t distribution (i.e., we take into account that the standard deviation σ must be estimated).

(5) Empirical Results

I generated random samples using the following multiple linear regression model of order $p=3$: $\hat{Y}_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + e_i$, $i=1,2,\ldots\ldots,n$, I set the coefficients values are $\beta_0 = 0.5,\ \beta_1 = 1,\ \beta_2 = 1.5$, Obtain the error term using five possibilities, (i) normal distribution N (0,1), (ii) Chi-square distribution with (DF=4), and (iii) Cauchy distribution(0,1) where the values 0, 1 are the location and scale parameters for Cauchy distribution, The values X_{1i} , X_{2i} are drawn from uniform distribution on interval (0, 1), I select a sample size of $n = 50$ and100 and consider that the sample may contain outliers , finallyto investigate the robustness of the methods against outliers, we randomly generated different percentages of outliers (P0= 0%, 10%, and 30%).

All empirical results are based on 1000 replications. We compare six estimators in this case: (1) Least Squares Estimators (OLS), (2) M- Huber (MH), (3) M-Hampel (MHP), (4) MM- Estimator (MM), (5) Least Median of Squares (LMS), (6) Least Trimmed Squares (LTS)and S- Estimators (S). All computations are obtained based on the R language.

$$Z_i = \frac{e_i}{s}$$

$$e_i = Y_i - X_i' \hat{\beta}_i \qquad (16)$$

Staudte and Sheather (1990), illustrated a reasonable approximation of the covariance matrix of $\hat{\beta}_M$ as:

$$n s^2 \frac{\sum \psi^2(Z_i)}{[\sum \psi'(Z_i)]^2} (X'X)^{-1} \qquad i = 1,2,\dots\dots n \qquad (17)$$

And multiplying the previous equation in the correction factor $n/(n-k)$, it becomes as follows:

$$\frac{n^2 s^2}{n-k} \frac{\sum \psi^2 \left[((Y_i - X_i'\hat{\beta}_i)/s \right]}{\left[\sum \psi'[(Y_i - X_i'\hat{\beta}_i)/s] \right]^2} (X'X)^{-1} \qquad (18)$$

Hogg (1979), has shown that, the distribution of the robust estimators is approximately normal if the sample size is large as:

$$Z_i = \frac{\hat{\beta}_j - \beta_j}{\sqrt{var(\hat{\beta}_j)}} \sim N(0,1) \quad (19)$$

so, the usual statistical inferences about the unknown parameters can be made.

The asymptotic 95% confidence interval for the parameter β is now given by

$$\hat{\beta} \pm 1.96 Se\left(\hat{\beta}\right) \qquad (20)$$

mit $Se\left(\hat{\beta}\right) = \sqrt{\hat{\tau}} \frac{\hat{\sigma}}{\sqrt{n}}$

$$\hat{\sigma} = S(e_1(\hat{\theta}), \ldots \ldots e_n(\hat{\theta})) \quad (11)$$

The scale estimator can be obtained through the following dispersion minimization problem

$$\frac{1}{n} \sum_{i=1}^{n} \rho\left(\frac{e_i(\beta)}{s}\right) = K \quad (12)$$

K is often put equal to $E_\phi[\rho]$, where $\emptyset$ is the standard normal. The value of β defined as

$$\hat{\beta}_j = \arg\min S\{e_1(\beta), \ldots \ldots e_n(\beta)\} \quad (13)$$

Subject to

$$\hat{\beta}_j = \frac{1}{n} \sum_{i=1}^{n} \rho\left(\frac{e_i(\beta)}{s(\beta)}\right) = k \quad (14)$$

S- estimators are more robustly than the LTS estimator, because S- estimators have smaller asymptotic bias and smaller asymptotic variance in the case contaminated data. (Rousseeuw and Leroy, 1987), and Pitselis, 2013))

(4)Inference for M-Estimators

Birkes and Dodge (1993) pointed out that, the distribution of the M- estimate $\hat{\beta}_M$ cannot be specified exactly, but for large sample size, they suggested that, asymptotic normality for M- estimators with normal distribution.

The offer variance- covariance matrix for robust regression is:

$$\frac{\sigma^2 E\left[\psi^2(Z)\right]}{\left[E(\psi'(Z))\right]^2}(X'X)^{-1} \quad , (15)$$

where

The constant 0.6745 makes s an approximately unbiased estimate of σ if n is large and the distribution is normal, ρ is a symmetric, positive definite function gives the contribution if each residual to the objective function (7).

(3.4) MM-estimator: MM estimation is a special type of M- estimation developed by Yohai (1987). Some properties of MM- estimator are follows as; They are highly efficient when the errors have normal distribution and Their BP is 0.5.

The three stages of computing MM-estimator can be illustrated in detail as follows:

1) Computing an initial consistent estimate $\hat{\beta}$ with high BP but not necessarily efficient. The S- estimator can be used as the initial part of an overall MM-estimate computational strategy.

2) Computing a robust scale $\hat{\sigma}$ of the residuals $s(\beta)$.

3) The final step is using M- estimate the regression parameters.

(3.5) S- Estimators: The S – estimators introduced by Rousseeuw and Yohai (1984) from another class of high BP, are a generalization of LMS and LTS and have the same asymptotic properties corresponding to M- estimators and also handle 50% of the outliers appearing in the data. They are the first high BP regression to achieve the usual $n^{-1/2}$ consistency under appropriate regularity conditions. They are defined by a minimization of the dispersion of the residuals:

$$\min_{\theta} S\left(e_1(\theta), \ldots e_n(\theta)\right) (10)$$

With the final scale estimate

outlying data points are trimmed , this method is computationally equivalent to LS . However, if there are more outlying data points that are trimmed, LTS is not as efficient.

In LTS, the objective is to minimize the sum of the smallest 50% squared residuals. LTS is a robust estimator with 50% breakdown point, which means that the estimator is insensitive to corruption due to outliers, provided that the outliers represent less than 50% of the data set. LTS has the advantage of being statistically more efficient than LMS. (Angela Y.Wu et al, 2007)

(3.3) M- Estimators: The most common general method of robust regression is M-estimation, introduced by Huber (1964), the method of M-estimation as a generalization to maximum likelihood estimation in context of location models. In simple terms, the M-estimator minimizes some functions of the residuals. As in the case of M-estimation of location, the robustness of the estimator is determine by the choice of weight function.

Consider the following linear model:

$Y = x\beta + u_i$, where Y is the vector of response variable, x is an nxp matrix of independent variables, β is a $p \times 1$ vector of unknown parameters, and u_i is random error with expectation zero and variance σ^2 .

Fox (2002) described The general M-estimator as follows:

$$\min \sum_{i=1}^{n} \rho(e_i) = \min \sum_{i=1}^{n} \rho(y_i - X_i \hat{\beta}) \qquad (7)$$

$$\text{Or } \min \sum_{i=1}^{n} \rho(\frac{e_i}{s}) = \min \sum_{i=1}^{n} \rho(\frac{(y_i - X_i \hat{\beta})}{s}) \qquad (8)$$

Where s is the estimated scale of residuals and a popular choice for s is,

$$S = median \left| e_i - median(e_i) \right| / 0.6745 \, (9)$$

$$w_i = \begin{cases} 1 & if \quad \left|\dfrac{e_i}{s^0}\right| \leq 2.5 \\ 0 & oterwise \end{cases}$$ (4)The outliers points are considered when

their $\dfrac{e_i}{s^0} > 2.5$, where e_i is computed based on the final fit .

Then the final scale estimate S for LMS regression is give by :

$$S = \sqrt{\dfrac{\sum\limits_{i=1}^{n} w_i e_i^2}{\sum\limits_{i=1}^{n} w_i - p}} \quad , \quad (5)$$

The a above estimator is very robust with respect to outliers in Y as well as outlier in x . Unfortunately, the LMS performs poorly from the point of view of asymptotic efficiency. (pitselis, 2013).

(3.2) Least Trimmed Squares: The LTS estimate is defined as follows:

$$\hat{\beta}_{LTS} = \underset{\beta}{min\; imize} \sum_{i=1}^{n} (e_i^2), \; i=1,2,......,h .$$ (6)

Where $(e_i^2)_1 \leq (e_i^2)_2 \leq .... \leq (e_i^2)_n$ are the ordered squared residuals sorted from smallest to largest and h is the number of these terms which are included in the summation called the coverage of the estimator. (Rousseeuw,1984)

LTS is calculated by minimizing the h ordered squares residuals, where $h = [n/2] + [(p+1)/2]$, with n and p being sample size and number of parameters, respectively .The largest squared residual are excluded from the summation in this method, which allows those outlier data points to be excluded completely. Depending on the value of h and the outlier data configuration, LTS can be very efficient. In fact, if the exact numbers of

(3.1) Least Median of Squares: Rousseeuuw (1984), proposed the Least Median of Squares (LMS) has the highest possible BP of 50% , by minimizing the median of squared error as follows:

$$\hat{\beta}_{LMS} = \min_{} \underset{1\le i \le n}{median} \ e_i^2(\beta), \qquad (2)$$

where $\hat{\beta}$ is an estimate of β. The LMS estimator has 50% breakdown point. A robust analysis with LMS may be conducted as follows:

- First step is to obtain the standardized residuals from LMS analysis,
- Then remove the observation corresponding to large outliers, (a above 2.5 standard deviations),
- Then run the OLS on the remaining data. (Onder and Ozet, 2001)

Instead of taking the median of the ordered squared residuals, consider h^{th} order statistic, where $h = [n/2]+[(p+1)/2]$, n denotes the sample size, and p is the number of regression parameters. The symbol [ ] mean " integer portion of ". LMS estimator has BP equal $([n-p]/2]+1)/n$.

Apart from the regression coefficients, the scale parameter S (the dispersion of errors e_i) has also to be estimated in a robust way.

As an initial estimate we use the following

$$s^0 = 1.4826(1+\frac{5}{n-p})\sqrt{med_i e_i^2(\beta)}, \ n \ne p \ (3)$$

Where n is the sample size, p is the regression parameters, and the factor $1.426 = \dfrac{1}{\phi^{-1}(0.75)}$, and $\emptyset$ is the standard normal cumulative function .

The initial scale estimate s^0 is then used to determine a weight w_i for the ith observation as follows :

leverage points are used to define robust techniques performance in a theoretical sense. Efficiency shows that, how the robust technique performs well relative to LS in case of clean data (without outliers). High efficiency is mostly desired on estimation .BP is a measure for stability of the estimator when the sample contains a large portion of outliers. It gives the minimum portion of outliers which may produce an infinite bias. (Alma, 2011)

Generally, robust regression methods aim to justify the following goals:

i. Consistency, asymptotic normality and high efficiency of the estimators if there are no model violations.

ii. Methods for forming confidence intervals for the unknown parameters and for testing hypotheses about them.

iii. Relative insensitivity of the properties in (i) and results in (ii) to slight violations of the model.

iv. Simplicity of theory.

v. Ease of computation, given a standard computer package (staudte, et al, 1990).

To conclude, if the errors having a non –normal distribution, we might consider a robust regression method, particularly in cases where the error distribution is heavier-tailed distribution and generate more errors than the normal. ALS analysis weights each observation equally in getting parameter estimates. Robust methods enable the observation to be weighted unequally .Essentially, observations that produce large residuals are down-weighted by a robust estimation method.

In this section, we introduce some methods of robust regression estimators as follows:

values and $\varepsilon_n^*(T_n, x)$ is the smallest fraction of contamination that can cause

the estimator T to take values arbitrarily far from T_n.

The breakdown point usually does not depend on the sample value x, but depends only slightly on the sample size n. (pitselis. G, 2013)

Definition 2: Outliers are defined to be observations which deviate from the pattern set defined by the majority of the data and are very dangerous of the data and many classical; statistical procedures. (Hampel et al, 1986)

Outliers in linear regression model are classified into:

Y- Outlier (Vertical Outlier); This is a point that is outlying only because its Y-coordinate is extreme and is called (Vertical Outliers). In this case, the presence of such values affects the LS – estimates, in particular the estimated intercept also, **X – Outlier (Leverage Point);** This type is outlying only in regard to the x- coordinate. Such apoint can cause some robust regression estimators to perform poorly, but the more modern robust estimators are not undermined by x- outliers. Also, they are called (Leverage Point) and **X and Y- Outlier;** This type is outlying in both coordinates may be a regression outlier, or residual outlier (or both) or it may have a very small effect or even on effect on the regression equation. (Ryan, 2009)

(3)The Robust Regression Approach

Robust regression is an important method for analyzing data that are contaminated with outliers. Robust regression analysis provides an alternative to least squares regression when fundamental assumptions are unfulfilled by the nature of the data. The properties of efficiency , breakdown and high

(1) Introduction

It is well known that, The linear regression model is one of the most widely used tools in statistical analysis and the least squares method is a very popular estimation technique for this model. In this respect, the classical estimators for the regression coefficients and error scale are the well-known least squares estimators. These estimators are optimal under normal errors but extremely sensitive to outliers and unusual observations in the data set. This particularly the case for the residual scale estimators, and as a result many more robust estimators have been proposed as alternatives. Much attention has been paid in statistical literature to robust and efficient estimation of the regression parameters. In this paper is organized as follows: Section (2) we present some definition of robust inference. Section (3) we present some robust regression approach. Section (4) we present inference for M-estimators. Section (5) is mainly devoted to an empirical results. Finally, in section (6) we give summary and conclusions.

(2) Robust Inference

We introduce some basic concepts of robust statistics that will be used in this study:

Definition 1: The Finite – Sample Breakdown Pointof the estimator of the observed sample X=($X_1, X_2 ..., X_n$) is defined as

$$\varepsilon_n^*(T, x) = \min(\frac{m}{n}; bais(m; T_n, x) < \infty)(1)$$

where $bais(m; T_n, x)$ is the maximum bias that can be contaminated by the (presence of outliers), m is the number of original points replaced by arbitrary

Abstract

The method of least squares of the most commonly used methods for estimating the parameters of linear regression models,this method requires the availability of several assumptions for the capabilities more efficiently. So robust methods can be used to give better results than OLS when there are outliers. This research discussessome robust regression approach and inference for M-estimators. Empirical study illustrates that robust methods aremore efficiency compare the OLS, when the data contain outliers.

Keywords:Bootstrap, Linear Regression, M- estimators, Outliers, Robust Inference, Statistical Inference.

Statistical Inference for some Robust Regression Estimators

Submitted by

Moshera Adel Soliman El Assar

Demonstrator at Department of Statistics and Insurance

Faculty of Commerce - Zagazig University