Plagiarism Checker X Originality Report

Similarity Found: 7%

Efektivitas Regresi Kuantil dalam Mengatasi Pontential Pencilan The Effectiveness of Quantile Regression in Dealing with Potential Outliers Netti Herawati 1* 1 Department of Mathematics, Faculty of Mathematics and Natural Sciences, University of Lampung, Bandar Lampung, 35141, Indonesia e-mail: netti.herawati@fmipa.unila.ac.id Corresponding Author*

Abstrak Regresi kuantil sebagai metode regresi robust dapat digunakan untuk mengatasi dampak kasus yang tidak biasa pada estimasi regresi seperti keberadaan pencilan pada data.

Tujuan dari penelitian ini adalah untuk mengevaluasi efektivitas regresi kuantil untuk menangani pencilan potensial dalam regresi linear berganda dibandingkan dengan metode kuadrat terkecil (MKT). Penelitian ini menggunakan data simulasi pada model regresi berganda dengan jumlah variabel independen (p=3) pada ukuran sampel yang berbeda (n = 20, 40, 60, 100, 200) dan $\beta_0 = 0$ and $\beta_1 = \beta_2 = \beta_3 = 1$ diulang 1000 kali. Efektivitas metode regresi kuantil dan MKT dalam pendugaan parameter ß diukur dengan Mean square error (MSE) dan model terbaik dipilih berdasarkan nilai Akaike Information Criterion (AIC) terkecil.

Hasil penelitian menunjukkan bahwa berbeda dengan OLS, regresi kuantil mampu menangani potensial outlier dan memberikan estimator yang lebih baik dengan nilai MSE yang lebih kecil. Dibandingkan dengan MKT dan kuantil lainnya, studi ini juga memberikan hasil yang cukup untuk memastikan bahwa kuantil 0,5 memberikan estimasi parameter terbaik dan model terbaik berdasarkan nilai MSE dan AIC terkecil. Kata Kunci : AIC, MSE, pencilan, regresi kuantil Abstract Quantile regression as a robust regression method can be used to overcome the impact of unusual cases on regression estimates such as the presence of potential outliers in the data.

The purpose of this study was to evaluate the effectiveness of quantile regression in dealing with potential outliers in multiple linear regression compared to ordinary least square (OLS). This study used simulation data in multiple regression model with the number of independent variables (p=3) for different sample sizes (n = 20, 40, 60, 100, 200) and $\beta_0 = 0$ and $\beta_1 = \beta_2 = \beta_3 = 1$ repeated 1000 times. The effectiveness of the quantile regression method and OLS in estimating ß parameters was measured by Mean square error (MSE) and the best model is chosen based on the smallest Akaike Information Criterion (AIC) value.

The results showed that in contrast to OLS, quantile regression was able to deal with potential outliers and provided a better estimator with a smaller mean mean square error. Compared to OLS and other quantiles, this study also provides sufficient results that quantile 0.5 provides the best parameter estimate and the best model based on the smallest MSE and AIC values. Keywords: AIC, MSE, outliers, quantile regression
 INTRODUCTION Classical linear regression estimates the mean response of the dependent variable dependent on the independent variables. The method usually use to estimates in classical linear regression is ordinary least square [1, 2, 3].

However, it is a parametric model and relies on assumptions of certain distribution in residuals that are often not met. There are many cases that the conditional mean behavior fails to entirely capture the patterns in the data when the data is skewed, multimodal, or contains outliers. In this condition, the residual distribution assumptions in the classical linear model will not be met, especially the normality distribution assumption. For this type of data is better estimated using methods that do not require any residual distribution.

One such method is quantile regression. Quantile regression has been used in many studies [4, 5, 6]. Quantile regression method propose a technique for estimating models for the conditional median function and the full range of other conditional quantile function. Just like the least squares regression, quantile regression is interested in studying the linear relationship between a response variable and one or more independent or explanatory variables.

However, the main purpose of the least squares regression is to determine the conditional mean of the response variable Y while the quantile regression models is related to the conditional t ((0,1) with t is quantile level of Y . In addition, quantile regression allows multiple quantiles to be modelled. Quantile regression is offering more comprehensive analysis of the data to be carried out compared to OLS where only

the mean is considered. Quantile regression makes no assumptions about the distribution of the residuals [7, 8, 9]. It also lets you explore different aspects of the relationship between the dependent variable and the independent variables.

The study of handling outliers has been done by many researches [10, 11, 12, 13]. In this study we will investigate the behavior of quantile regression in handling outliers compare to OLS using simulated data based on MSE and AIC. METHOD The quantile regression is an extension of ordinary quantiles ideas. Classical linear regression method is based on minimizing the sum of squared residuals to model the conditional mean of the target variable against the covariates. On the other hand, quantile regression provides estimates of a range of conditional quantiles to model conditional percentiles of the target variable against the covariates.

It is a useful tool for estimating not only upper or lower tail but also the center of the conditional distribution of interest [4, 8, 9, 14, 15]. Consider a general regression function with y is the response variable and $?? = (?? 1 , ?? 2 ,..., ?? ??) ??$ is a set of predictors. To obtain the sample mean, least square regression model solve min ???R ??=1 ?? ?? ?? -?? 2 (1) as an estimate of the unconditional population mean, EY.

By replacing the scalar μ with a parametric function μ (x, ß), we can solve min ???R?? ??=1 ?? ?? ?? -??(?? ?? ,??) 2 (2) to find the estimate of the conditional expectation function ?? ?? ?? . The least square estimate for (2) is given by ?? = ?? ' ?? -1 ?? ' ?? [16]. In term of quantile regression, consider a continuous real valued random variable Y characterized by the following distribution function in the ordinary quantile ?? ?? ?? =??(??=??) (3) The ??-th quantile of Y for any ???(0,1) is defined as ?? ?? =inf?(??: ?? ?? ?? =??) (4) When Q(1/2), it is equal to median [17, 18]. Just like the distribution function F, the quantiles function provides a complete characterization of the random variable Y [8].

When estimating quantiles, the value of y in the sample data corresponding to a given probability ?? has to be determined. The ?? ??h quantile in a sample of data refers to the probability of ?? for a value y, such that ?? ?? ?? ?? =??. It can also write as ?? ?? = ?? ?? -1 (??)where ?? ?? is such that an inverse of the function ?? ?? (??) for a probability ??. The 100 ??% quantile (say ?? =0.5) of the conditional distribution of the response (y) given covariates (x) based on independent observations (?? ?? , ?? ??) ??=1 ?? , the conditional ??-quantile is estimated by minimizing ??=1 ?? ?? ?? ?? ?? ?? - ?? 0 - ?? ?? ?? ?? (5) Where ?? ?? ?? =?? ?? + +(1-??)??_ is the check function with subscript s '+' and '-' stand for the positive and negative parts, respectively [9]. The estimation of selected significant predictors in quantile regression use ?? 1 [8].

The ?? 1 quantile regression model is estimated by ?? (?? 1 ????????)= arg min ?? 0 ??

$$\sum_{i=1}^{n} ??\ ??\ ??\ ??\ ?? - ??\ 0 - ??\ ??\ ??\ ?? + ??\ ??\ 1 \quad (6)$$ where $?? 1 = \sum_{i=1}^{n} ??\ ??\ ??$ is the $?? 1$-norm penalty (or lasso penalty) on $??$. When $??$ is chosen appropriately, some components of $??$ will be shrunk to exact zero. Since the check loss function is piecewise linear, the quantile regression estimator is inherently robust in handling extreme value point and outliers. To get a measure of how close the regression line was to a set of points, the Mean Square Error (MSE) of the regression coefficient $??$ was examined.

The MSE is defined by $?????? ?? = 1 ?? \sum_{l=1}^{??} ??\ (??) - ?? 2$ where $?? (??)$ is the estimated parameter in the l-th simulation. The slope and intercept are correctly estimated when MSE approaches to zero. To determine the performance of the proposed estimate, the Akaike Information Criterion (AIC) was used. AIC can be written as $?????? ?? = 2?? - 2\ln ?(\ ??\ )$ where $?? = ??\ ??\ ??\ , ??\ ,$ $??$ is the value that maximize the likelihood function, n = sample size, and k = the number of parameters [19, 20]. A good estimation model was indicated by the lowest AIC value.

Simulated data was carried out in this study with five different sample sizes (n=20, 40, 60, 100, 200) for three independent variables (p=3) using a package for quantile regression developed by [21]. Dependent variable $(??)$ for each $??$ independent variables was from $?? = ???? + ??$ with $x_i \sim N(0, 1)$ and $e \sim N(0, 1)$ contaminated with various number of outliers (10%, 15%, 20%). ß parameters were chosen with ß 0 =0, and ß 1 , ß 2 , ß 3 =1. After simulating each data, we fitted the OLS and quantile regression and measured the efficiency of both methods in estimating the regression coefficient and determined the best estimation model using AIC.

RESULTS AND DISCUSSION The simulation results for identifying the effectiveness of quantile regression presented in Table 1 in terms of standard error of parameter estimates. It shows that for n=20, 40, 60, 100, 200 with 10% outliers, standard error of estimates using quantile regression was lower than OLS. Similar results were obtained for n= 20, 40, 60, 100, 200 with 20% and 30% outliers. It indicates that quantile regression provides better parameter estimates than OLS for all sample sizes and various number of outliers being studied. We can also see that quantile 0.5

gives the lowest standard error of parameter estimates compare to quantile 0.25 and quantile 0.75 for n=20, 40, 60, 100, 200 with various number of outliers. It proves that the quantile 0.5 is the most accurate parameter estimates than quantile 0.25 and quantile 0.75. Table 1. $??$ and SE ($??\ ??$) for different sample sizes and various number of outliers Sample size _Method _10% outliers _20% outliers _30% outliers _ _ _ _SE ($?? 1$) _SE ($?? 2$) _SE ($?? 3$) _SE ($?? 1$) _SE ($?? 2$) _SE ($?? 3$) _SE ($?? 1$) _SE ($?? 2$) _SE ($?? 3$) _ _n=20 _OLS _2.8046 _2.3233 _2.4495 _2.3998 _3.3882 _4.0557 _4.8388 _2.8077 _4.2070 _ _ _QR 0.25 _2.1241 _2.1766 _2.0195 _3.1991 _3.1465 _3.1412 _3.6997 _3.7711 _3.8961 _

_ _QR 0.50 _2.0933 _2.1944 _2.0252 _3.1989 _3.1462 _3.1402 _3.6990 _3.7710 _3.8961 _ _ _QR 0.75 _2.1241 _2.1766 _2.0195 _3.3116 _3.3660 _3.3260 _3.9607 _3.7982 _4.0457 _ _ _ _ _ _ _ _ _ _ _ _ _ _ _n=40 _OLS _1.9844 _1.2343 _1.8596 _2.3226 _2.4937 _2.9881 _3.0259 _3.2395 _2.0931 _ _ _QR 0.25 _1.6162 _1.6174 _1.6894 _2.2147 _2.2566 _2.2404 _2.7026 _2.7457 _2.7248 _ _ _QR 0.50 _1.6160 _1.6170 _1.6894 _2.2140 _2.2489 _2.2400 _2.7020 _2.7455 _2.7240 _ _ _QR 0.75 _1.7222 _1.7072 _1.7262 _2.2405 _2.2566 _2.2610 _2.7287 _2.7476 _2.7589 _ _ _ _ _ _ _ _ _ _ _ _ _ _n=60 _OLS _1.3450 _1.2208 _1.1639 _1.7725 _1.8507 _1.77421 _2.3037 _2.3443 _1.9384 _ _ _QR 0.25 _1.2146 _1.2215 _1.1880 _1.7267 _1.7341 _1.7091 _2.0941 _2.1605 _2.0852 _ _ _QR 0.50 _1.2144 _1.2210 _1.1880 _1.7266 _1.7339 _1.7081 _2.0931 _2.1600 _2.0850 _ _ _QR 0.75 _1.2387 _1.2343 _1.2070 _1.7623 _1.7585 _1.7591 _2.1920 _2.2037 _2.1573 _ _ _ _ _ _ _ _ _ _ _ _ _ _n=100 _OLS _1.0329 _1.0482 _1.1334 _1.2240 _1.4546 _1.3552 _1.6024 _1.7452 _1.7047 _ _ _QR 0.25 _1.0125 _1.0382 _1.0609 _1.3018 _1.3199 _1.3123 _1.6564 _1.6474 _1.6322 _ _ _QR 0.50 _1.0125 _1.0272 _1.0125 _1.3011 _1.3189 _1.3120 _1.6562 _1.6470 _1.6321 _ _ _QR 0.75 _1.0382 _1.0480 _1.0468 _1.3314 _1.3636 _1.3405 _1.6763 _1.6894 _1.6728 _ _ _ _ _ _ _ _ _ _ _ _ _ _n=200 _OLS _0.6429 _0.7979 _0.6867 _0.9394 _0.9394 _0.9394 _1.1853 _1.3932 _1.1794 _ _ _QR 0.25 _0.6835 _0.6978 _0.6921 _0.9826 _0.9826 _0.9826 _1.1404 _1.1552 _1.1425 _ _ _QR 0.50 _0.6825 _0.6970 _0.6911 _0.9394 _0.9394 _0.9394 _1.1400 _1.1542 _1.1422 _ _ _QR 0.75 _0.7075 _0.7111 _0.7156 _0.9826 _0.9826 _0.9826 _1.1799 _1.1748 _1.1776 _ _

The results of analyzing the effectiveness of quantile regression compares to OLS in estimating parameter ß using simulated data for n=20, 40, 60, 100, 200 and contaminated by 10%, 20%, 30% outliers repeated 1000 times in terms of MSE are given in Table 2. Table 2. MSE of OLS and QR for different sample sizes and various number of outliers Sample size _Method _MSE _ _ _ _10% outliers _20% outliers _30% outliers _ _n-20 n=40 n=60 n=100 n=200 _OLS QR 0.25 QR 0.50 QR 0.75 OLS QR 0.25 QR 0.50 QR 0.75 OLS QR 0.25 QR 0.50 QR 0.75 OLS QR 0.25 QR 0.50 QR 0.75 OLS QR 0.25 QR 0.50 QR 0.75 _0.1531 0.0738 0.0067 0.0103 0.5280 0.0042 0.0005 0.0309 0.0288 0.0088 0.0032 0.0123 0.0850 0.0019 0.0006 0.0145 0.1095 0.0153 0.0005 0.0012 _0.5307 0.0153 0.0061 0.0133 0.1199 0.0052 0.0008 0.0034 0.0191 0.0085 0.0008 0.0108 0.0603 0.0245 0.0022 0.0034 0.1366 0.0157 0.0001 0.0013 _0.5175 0.0086 0.0067 0.0142 0.4259 0.0052 0.0015 0.0048 0.0731 0.0422 0.0004 0.0150 0.0167 0.0063 0.0001 0.0046 0.2466 0.0336 0.0008 0.0047 _ _ It can be seen from Table 2 that for 10 % outliers, OLS gives MSE =0.1531 for n=20, MSE= 0.5280, for n=40, MSE= 0.0288, for n=60, MSE= 0.0850, for n=100 and MSE= 0.1095, for n=200, respectively.

These values are much higher than the MSE of quantile 0.25, 0.50, and 0.75. Similarly, when there is 20% and 30% outliers in the data, the MSE of OLS all sample sizes being studied is much higher than MSE of quantile regression. The MSE of OLS appears to be increasing due to the increasing of the outliers in smaller sample sizes (n=20, 40, 60). However, it reverse for larger sample sizes (n=100, 200). Whereas MSE of quantile 0.25,

0.5 and 0.75 decreased as the number of samples increased. When there are potential outliers, quantile regression does not appear to be affected. This phenomenon occurs in all number of outliers studied. Among all quantiles, the MSE value for quantile 0.5

is the smallest compared to MSE for quantile 0.25 and quantile 0.75 for all sample sizes and various number of outliers. As it shows in Table 2, for 10% outliers, quantile 0.5 has MSE =0.0067 for n=20, MSE= 0.0005, for n=40, MSE= 0.0032, for n=60, MSE= 0.0006, for n=100 and MSE= 0.0005, for n=200, respectively. The MSE values at quantile 0.5 are also the smallest in all sample sizes containing 20% ??and 30% outliers compared to quantile 0.25 and quantile 0.75 as shown in Table 2. To present a more comprehensive result, the MSE values ??of both methods and each quantile are displayed in Figure 1-5. / Figure 1. MSE for n=20 contain various number of outliers / Figure 2. MSE for n=40 contain various number of outliers / Figure 3.

MSE for n=60 contain various number of outliers / Figure 4. MSE for n=100 contain various number of outliers / Figure 5. MSE for n=200 contain various number of outliers In the clear view of Figure 1-5, we can see that the MSE of quantile regression 0.25, 0.5, and 0.75 are significantly lower than OLS for all sample sizes (n=20, 40, 60, 100, 200) and various number of outliers (10%, 20%, 30%). It is clear that MSE of quantile regression does not to be affected by outliers. This indicates that quantile regression is able to handle potential outliers very well and robust to potential outliers up to 30% of the data. Moreover, if we compare the quantiles it becomes clear that the quantile 0.5

gives the lowest MSE value compared to the quantile 0.25 and quantile 0.75. This provides evidence that quantile 0.5 gives better parameter estimates than quantile 0.25 and quantile 0.75. To find the best estimation model, the AIC value for OLS and quantile regression in each number of sample sizes contaminated by outliers was measured. Table 3 shows the AIC values ??for n = 20, 40, 60, 100, 200 and were contaminated by 10%, 20%, 30% outliers which were repeated 1000 times. Table 3.

AIC of OLS and QR contain various number of potential outliers Sample size _Method _AIC _ _ _ _10% outliers _20% outliers _30% outliers _ _n=20 n=40 n=60 n=100 n=200 _OLS QR 0.25 QR 0.50 QR 0.75 OLS QR 0.25 QR 0.50 QR 0.75 OLS QR 0.25 QR 0.50 QR 0.75 OLS QR 0.25 QR 0.50 QR 0.75 OLS QR 0.25 QR 0.50 QR 0.75 _-1.5767 -2.3064 -4.7075 -4.2701 -0.4884 -5.3203 -7.3308 -3.3266 -3.4450 -4.6247 -5.6176 -4.2980 -2.4040 -6.1631 -7.3252 -4.1727 -2.1814 -4.1446 -9.8145 -4.3161 _-0.3334 -3.8767 -4.7910 -4.0184 -1.9706 -5.0946 -6.9583 -5.5247 -3.8574 -4.6674 -7.0122 -4.4191 -2.7482 -3.6482 -6.0498 -5.6019 -1.9603 -4.1241 -10.8246 -4.2661 _-0.3586 -4.4456 -4.7027 -3.9476 -0.7035 -5.0946 -6.2911 -5.1728 -2.5154 -3.0630 -7.5020 -4.0996 -4.0269 -4.9946 -8.7135 -5.3185 -1.3696 -3.3624 -7.0512 -5.3121 _ _ From Table 3 it is

visible that there is a significant difference between OLS and quantile regression in the AIC value. Quantile regression produces AIC value much lower than OLS in all sample sizes and number of outliers.

That is, compared to OLS, quantile regression provides a better regression model if there are potential outliers. It proves that quantile regression model is robust to potential outliers than OLS. The study also shows that quantile regression 0.5 gives the best estimates than other quantiles. That is one of the advantages of using quantile regression where one can estimate parameters using different quantiles which are not provided in OLS. In this study, quantile 0.5 provided the best parameter estimate and the best model compared to OLS and the other quantiles for all sample sizes and various number of outliers.

CONCLUSION The study showed that quantile regression performanced far better than OLS based on MSE and AIC for n=20, 40, 60, 100, 200 and number of potential outliers 10%, 20%, 30%. We concluded that in contrast to OLS, quantile regression model was more effective in dealing with potential outliers for different sample sizes and various number of potential outliers. In addition, quantile 0.5 gives the best parameter estimate and the best model based on the smallest MSE and AIC values ??compared to quantile 0.25 and quantile 0.75 for all sample sizes and number of outliers studied. REFERENCES [1] M.H. Kutner, C.J. Nachtsheim and J.

Neter, Applied Linear Regression Models, 4th Ed., New York: McGraw-Hill Co. Inc., 2004. [2] D.C. Montgomery, E.A. Peck and G.G. Vining, Introduction to Linear Regression Analysis, 3rd Ed., USA: John Wiley & Sons, Inc., 2001. [3] N.R. Draper and H. Smith, Applied Regression Analysis, 3rd Ed., USA: John Wiley & Sons, Inc., 1998. [4] R. Koenker and K.F. Hallock, "Quantile regression," Journal of Economic Perspectives, vol. 15, no.4, pp 143-156, 2001. [5] G.A.Haile and A.N. Nguyen, "Determinant of academic attainment in the United States: A quantile regression analysis of test scores, " Education Economics, vol. 16, pp. 29-57, 2008. [6] J. Staruss, K. Beege, B. Sikoki, A. Dwiyanto, Y. Herawati and F.

Witoelar, The third wave of the Indonesia family life survey: Overview and field report, Santa Monica, CA: RAND Corp, 2004. [7] L. Hao and D.Q. Naiman, Quantile Regression, Thousand Oaks: Sage Publications Inc., 2007. [8] R. Koenker, Quantile regression, New York: Cambridge University Press, 2005. [9] R. Koenker and G. Bassett, "Regression quantile," Econometrica, vol. 46, pp.33-50, 1978. [10] K. Nisa and N. Herawati, . 'Robust Estimation of Generalized Estimating Equation when data contain outliers,' Insist, vol. 2, No. 1, pp., 1-5, 2017. [11] P.J. Rousseeuw and A.M. Leroy, Robust Regression and Outlier Detection, New York: Wiley, 1987. [12] E. Setiawan, N.

Herawati, K. Nisa, Nyusirwan, and S. Saidi, "Handling full multicollinearity and various numbers of outliers using robust ridge regression," Sci.Int.(Lahore), vol 31, no. 2, pp. 201-204, 2019. [13] V. Barnett and T Lewis, Outlier in Statistical Data, New York :John Wiley & Sons, 1984. [14] W. Gilchrist, Statistical modelling with quantile functions, FL: Chapman and Hall/CRC, Boca Raton, 2000. [15] C.M. Kuan, An Introduction to Quantile Regression. Institute of Economics, Canada: Academia Sinica, 2007. [16] W.H. Greene, Econometric Analysis, 5th ed., New Jersey: Prentice Hall, 2002. [17] B. Fitzenverger, R. Koenker, J.A.F. Machado, (Eds.), Economic Applications of Quantile Regression, Berlin Heidelberg GmbH: Springer-Verlag, 2002. [18] M. Furno, Parameter Instability in Quantile Regressions, NewYork: Elsevier, 2007. [19] H.

Akaike, Information theory and an extension of the maximum likelihood principle, In B.N. Petrow and F. Csaki (eds), Second International symposium on information theory (pp.267-281), Budapest: Academiai Kiado, 1973. [20] H. Akaike, "A new look at the statistical model identification," IEEE Transactions on Automatic Control, vol. 19, pp.716-723, 1974. [21] R. Koenker, Quantreg: Quantile Regression, Version R package version 4.22, 2008.

INTERNET SOURCES:

<1% - https://avstatistik.blogspot.com/feeds/posts/default
<1% - https://www.researchgate.net/publication/45529647_Quantile_Regression_Applications_and_Current_Research_Areas
1% - https://support.sas.com/resources/papers/proceedings15/3412-2015.pdf
<1% - https://id.123dok.com/document/yrk9xnpz-python-for-data-science-for-dummies-pdf-pdf.html
<1% - https://www.annualreviews.org/doi/full/10.1146/annurev.publhealth.23.100901.140546
<1% - https://www.mit.edu/~vchern/papers/ch_IVQR_2001_rev_Oct24_2004.pdf
1% - https://arxiv.org/pdf/1402.2492.pdf
<1% - https://wikimili.com/en/Quantile_regression
<1% - https://www.sciencedirect.com/science/article/pii/S0378112718317572
1% - https://abigaildai.wordpress.com/author/abigaildai/page/2/
<1% - https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4466087/
<1% - https://www.sciencedirect.com/science/article/pii/S0378779620304399
<1% - https://www.cliffsnotes.com/study-guides/statistics/principles-of-testing/point-estimates-and-confidence-intervals

<1% - https://www.tutorialspoint.com/r/r_multiple_regression.htm

<1% - https://core.ac.uk/download/pdf/6408188.pdf

<1% - https://link.springer.com/referenceworkentry/10.1057/978-1-349-95121-5_2795-1

1% - http://www.columbia.edu/~my2550/papers/smqr.final.pdf

<1% - https://www.sciencedirect.com/science/article/pii/S0304407606000546

<1% - http://statweb.stanford.edu/~susan/courses/s60/split/node60.html

<1% - https://www.researchgate.net/topic/Akaike-Information-Criterion-AIC

<1% - https://www.sciencedirect.com/science/article/pii/S0026271419300125

<1% - https://www.researchgate.net/publication/247312065_Quantile_Regression_An_Introduction

<1% - https://www.studocu.com/en-us/document/brigham-young-university-idaho/engineering-statistics/other/math-330-engineering-statistics/6981910/view

<1% - https://idoc.pub/documents/regression-modeling-strategies-with-applications-to-linear-models-logistic-and-ordinal-regression-and-survival-analysispdf-wl1pjw8ymvlj

<1% - https://www.twirpx.com/files/science/mathematics/tvms/math_stats/regression_analysis/

<1% - https://www.amazon.co.uk/Applied-Regression-Analysis-Probability-Statistics/dp/0471170828

<1% - http://statisticaloutsourcingservices.com/Outlier2.pdf

1% - http://www.aee.odu.edu/manuscripts/AES00509N.doc