The Profit Trend of Cosmetics Products Using Bootstrap Edgeworth Approximation

Edlira Donefski, Lorenc Ekonomi, Tina Donefski

***Abstract*—**Edgeworth approximation is one of the most important statistical methods that has a considered contribution in the reduction of the sum of standard deviation of the independent variables' coefficients in a Quantile Regression Model. This model estimates the conditional median or other quantiles. In this paper, we have applied approximating statistical methods in an economical problem. We have created and generated a quantile regression model to see how the profit gained is connected with the realized sales of the cosmetic products in a real data, taken from a local business. The Linear Regression of the generated profit and the realized sales was not free of autocorrelation and heteroscedasticity, so this is the reason that we have used this model instead of Linear Regression. Our aim is to analyze in more details the relation between the variables taken into study: the profit and the finalized sales and how to minimize the standard errors of the independent variable involved in this study, the level of realized sales. The statistical methods that we have applied in our work are Edgeworth Approximation for Independent and Identical distributed (IID) cases, Bootstrap version of the Model and the Edgeworth approximation for Bootstrap Quantile Regression Model. The graphics and the results that we have presented here identify the best approximating model of our study.

***Keywords*—**Bootstrap, Edgeworth approximation, independent and Identical distributed, quantile.

I. Introduction

THE econometrics is related to the application of some statistical methods in the solution of an economical problem. The difference between the econometrist's and statistician's points of view has to do with the stochastic relation that is essentially considered by the econometrist. So, this stochastic relation is treating the errors done during the creation of relations between the variables that are taken for study. Every researcher's aim is to build a model that has a minimum level of the error, and one of the best methods that helps in this objective's arrival is the Edgeworth Approximation. The Edgeworth expansions for the bootstrap version is applied in a Quantile Regression model between the generated profits during and the realized sales for the cosmetic products during the past two years, through the EViews10 software package. The bootstrapping technique for the reduction of standard error is the method that is used to evade the severe estimation. Some kinds of bootstrap methods are given in [6], [15] and [8]. Even though the realized sales have a considered statistical importance in the generated profits, the focus of our work was the identification of the standard errors of the independent variable (the realized sales) and finding the best method of improving it and increasing the rightness of the relation's prediction. In this paper, we have done applications for the Bootstrap version, Edgeworth Approximation for IID cases and the Edgeworth approximation for Bootstrap Quantile Regression Model. We have concretized them with graphics and so, we can easily identify the best approximating model of this study. At the same time, we will give a better evidence of a comparison of the results for the two past years.

E. Donefski and L. Ekonomi are with the Faculty of Natural and Human Sciences, Department of Mathematics and Physics, "Fan S. Noli" University, Korça, Albania (e-mail: edonefski@unkorce.edu.al, lorencekonomi@yahoo.co.uk).

T. Donefski is with is with the Faculty of Economy, Department of Finance-Accounting, "Fan S. Noli" University, Korça, Albania (e-mail: tdonefski@unkorce.edu.al).

II. Definition of the Mean Squared Error Estimate

Let $X = \{X_1, ..., X_n\}$ denote a random sample of size n drawn from a distribution with distribution function F, and write

$$\hat{F}(x) = n^{-1}\sum_{i=1}^{n} I(X_i \le x)$$

for the empirical distribution function of the sample, [10]. The bootstrap estimate of the pth quantile of F, $\xi_p = \hat{F}^{-1}(p)$, is

$$\hat{\xi}_p = \hat{F}^{-1}(p) = \inf\{x : \hat{F}(x) \ge p\} = X_{nr} \tag{1}$$

where $X_{n1} \le ... \le X_{nn}$ denote the order statistics of X and $r = [np]$ is the largest integer not greater than np. The mean squared error of ξ_p is given by

$$\begin{aligned}\tau^2 &= E\{(\hat{\xi}_p - \xi_p)^2\} \\ &= \frac{n!}{(r-1)!(n-r)!}\int_{-\infty}^{\infty}(x-\xi_p)^2 F(x)^{r-1}\{1-F(x)\}^{n-r}\,dF(x) \\ &= r\binom{n}{r}\int_0^1 \{F^{-1}(u) - F^{-1}(p)\}^2 u^{r-1}(1-u)^{n-r}\,du,\end{aligned}$$

of which the bootstrap estimate is

$$\begin{aligned}\hat{\tau}^2 &= r\binom{n}{r}\int_0^1 \{\hat{F}^{-1}(u) - \hat{F}^{-1}(p)\}^2 u^{r-1}(1-u)^{n-r}\,du \\ &= \sum_{j=1}^{n}(X_{nj} - X_{nr})^2 w_j,\end{aligned} \tag{2}$$

where

$$w_j = r\binom{n}{r}\int_{(j-1)/n}^{j/n} u^{r-1}(1-u)^{n-r}\,du.$$

The bootstrap variance estimate differs in details, as follows. Bootstrap estimates of the mean and mean square of $\hat{\xi}_p$ are respectively

$$r\binom{n}{r}\int_0^1 \hat{F}^{-1}(u)u^{r-1}(1-u)^{n-r}\,du = \sum_{j=1}^{n} X_{nj}w_j,$$
$$r\binom{n}{r}\int_0^1 \hat{F}^{-1}(u)^2 u^{r-1}(1-u)^{n-r}\,du = \sum_{j=1}^{n} X_{nj}^2 w_j.$$

Therefore, the bootstrap variance estimate is

$$\sum_{j=1}^{n} X_{nj}^2 w_j - (\sum_{j=1}^{n} X_{nj}w_j)^2.$$

III. Edgeworth Expansion for the Studentized Bootstrap Quantile Estimate

The bootstrap versions of ξ_p and τ^2, defined in (1) and (2) are $\hat{\xi}_p$ and $\hat{\tau}^2$. The Standard Normal distribution characterized the limit of $(\hat{\xi}_p - \xi_p)/\hat{\tau}$, because $\hat{\xi}_p$ is Normal distributed (asymptotically) $N(0,\tau^2)$ [4], and $\hat{\tau}^2/\tau^2 \to 1$ converges to 1, in probability. The Edgeworth expansion that we will see now for the degree of convergence to this limit and the properties of terms in this expansion are not similar with those mentioned in Chapter 2 of [10]. The reason is that the order of $\hat{\tau}^2\tau^{-2} - 1$ is $n^{-1/4}$, not $n^{-1/2}$. The order of the first term in the expansion is of size $n^{-1/2}$, but the polynomial of that term is not even or odd; in the cases that we have seen before, the polynomial for the term of order $n^{-1/2}$ term was usually even.

We define

$$\begin{aligned} q(x) &= \tfrac{1}{4}\{\pi p(1-p)\}^{-1/2} x(x^2+1+2^{3/2}) \\ &+ \tfrac{1}{6}\{p(1-p)\}^{-1/2}(1+p)(x^2-1) \\ &- \Big[\{p(1-p)\}^{-1/2} p + \{p(1-p)\}^{1/2} f'(\xi_p) f(\xi_p)^{-2}\Big] x^2 \\ &- \{p(1-p)\}^{-1/2}\{\tfrac{1}{2}(1-p)+r-np\} \end{aligned}$$

to see the result in a clear way. Furthermore

$$P\Big\{(\hat{\xi}_p - \xi_p)/\hat{\tau} \le x\Big\} = \Phi(x) + n^{-1/2}q(x)\phi(x) + O(n^{-3/4})$$

as $n \to \infty$; see [11]. The degree of the polynomial q is 3 and it is not even or odd function.

The Edgeworth expansion of a Studentized quantile is created from a "basic" series of terms diminishing in orders of $n^{-1/2}$, emanating from the numerator in the Studentized ratio, together with a "new" series emanating from the denominator. The order of the "new" series diminishes in $n^{-1/4}$ because the relative error of the variance estimate is $n^{-1/4}$. The first term in the "new" series flees. The jth term is even or odd according to whether j is odd or even, respectively, for the both series. In the joining series, the term of degree $n^{-1/2}$ includes the first, even term of the "basic" series and the second, odd term of the "new" series.

The Edgeworth expansion for the distribution of a Studentized quantile when the estimation of the standard deviation is based clarity on a density estimator, see [12], [9], [2].

IV. The Model

The methods of modeling the conditional distribution in other aspects represent special interest, except the regression models that analyze the conditional mean of a dependent variable. Quantile regression that models the quantiles of the dependent variable for a set of conditioning variables, is one of the most interesting models [16], [17].

Quantile regression estimates the linear relationship between regressors X and Y, a specified quantile of the dependent variable, see [13]. The least absolute deviations (LAD) estimator is a particular case of quantile regression which adjusts the conditional median of the dependent variable.

The description of the fact that how the median or the percentiles of the response variable are affected by regressors variable is offered by the quantile regression [16], [17].

Independent and Identical

Koenker and Bassett [13] derive asymptotic normality results for the quantile regression estimator in the i.i.d. setting, showing that under mild regularity conditions,

$$\sqrt{n}(\hat{\beta}(\tau) - \beta(\tau)) \sim N(0, \tau(1-\tau)s(\tau)^2 J^{-1}) \qquad (3)$$

where

$$J = \lim_{n\to\infty}(\sum_i X_i X_i'/n) = \lim_{n\to\infty}(X'X/n)$$
$$s(\tau) = F^{-1}{}'(\tau) = 1/f(F^{-1}(\tau))$$

and $s(\tau)$ is the derivative of the quantile function or the inverse density function evaluated at the τ-quantile and it is called the sparsity function or the quantile density function, [1], [16], [17].

The direct estimation of the coefficient covariance matrix is forthright, for the value of the sparsity at a given quantile. The formula for the asymptotic covariance in (3) is similar with the covariance of ordinary least squares in the i.i.d. case, with $\tau(1-\tau)s(\tau)^2$ that express the error variance in the ordinary formula. The Edgeworth approximation is expressed with the ordinary i.i.d. case and it helps to minify the standard error of the independent coefficient.

Sparsity Estimation

The expression for the asymptotic covariance matrix of the quantile regression estimates for i.i.d. data shows the importance of the sparsity function. Since, the sparsity is a function of the unknown distribution F, it is necessary the estimation of a concern quantity [16], [17].

The three methods that EViews ensures, for estimating the scalar sparsity $s(\tau)$ are: two Siddiqui [9] difference quotient methods [14]; Bassett and [3] and one kernel density estimator [5]-[7].

Bootstrapping

Bootstrapping technique for the estimation of the covariance matrix is a method that avoids the estimation of the asymptotic covariance matrix that requires the estimation of the sparsity concern parameter, at a single point or conditionally for each sampling.

Four different bootstrap methods that EViews provides are: the residual bootstrap (Residual), the design, or XY-pair, bootstrap (XY-pair), and two variants of the Markov Chain Marginal Bootstrap (MCMB and MBMB-A).

Details for different bootstrap methods are given in [6], [8], [15]-[17].

Residual Bootstrap

By resampling (with replacement) separately from the residuals $u_i(\tau)$ and from the X_i, we can create the residual bootstrap.

Let X^* be a *mxp* matrix of independently resampled X, and let u^* be an *m*-vector of resampled residual. We create the dependent data and then we create a bootstrap estimate of $\beta(\tau)$ using Y^* and X^*, from $Y^* = X^*\hat{\beta}(\tau) + u^*$, [16], [17]. Repeating this procedure for *M* bootstrap replications, then the estimator of the asymptotic covariance matrix is of the form:

$$V(\hat{\beta}) = n\left(\frac{m}{n}\right)\frac{1}{B}\sum_{j=1}^{B}(\hat{\beta}_j(\tau) - \overline{\beta(\tau)})(\hat{\beta}_j(\tau) - \overline{\beta(\tau)})' \tag{4}$$

where, the mean of the bootstrap elements is $\overline{\beta(\tau)}$. $V(\hat{\beta})$ that is the bootstrap covariance matrix, is a (scaled) estimate of the sample variance of the bootstrap estimates of $\beta(\tau)$.

The independence of the *u* and the X is required for the validity of using separate draws from X_i and $u_i(\tau)$.

XY-Pair (Design) Bootstrap

The usual appearance of bootstrap resampling is the XY-pair bootstrap, and it is useful when *u* and X are not independent. For the usage of the XY-pair bootstrap, we simply create B randomly drawn (with replacement) subsamples of size *m* from the original data, then compute estimates of $\beta(\tau)$ using the (y^*, X^*) for each subsample. Using (4), from sample variance of the bootstrap results, can be estimated the asymptotic covariance matrix, see [16], [17].

Quantile Process Testing

A very interesting issue may be the analysis of the quantile regression model not only for a single quantile, τ. We may require creating joint hypotheses using coefficients for more than one quantile. The general category of quantile process analysis treats the consideration and the importance of more than one quantile regression at the same time [16], [17].

We must first contour the appropriated distributional theory, before operating to the hypothesis tests of interest. The process coefficient vector is defined:

$$\beta = (\beta(\tau_1)', \beta(\tau_2)', ..., \beta(\tau_K)')'$$

Then

$$\sqrt{n}(\hat{\beta} - \beta) \sim N(0, \Omega)$$

where Ω has blocks of the form:

$$\Omega_{ij} = \left[\min(\tau_i, \tau_j) - \tau_i\tau_j\right] H^{-1}(\tau_i) J H^{-1}(\tau_j) \tag{5}$$

In the i.i.d. setting, Ω simplifies to,

$$\Omega = \Omega_0 \otimes J \tag{6}$$

where Ω_0 has representative element:

$$\omega_{ij} = \frac{\min(\tau_i, \tau_j) - \tau_i\tau_j}{f(F^{-1}(\tau_i))(f(F^{-1}(\tau_j)))} \tag{7}$$

By using (5)-(7) or implementing one of the bootstrap methods, we can estimate Ω.

V. Applications of the Methods in the Profit Trend of Cosmetics Products

We have considered 130 observations of the sales of different cosmetics products and the generated profits for 2019 and 2020 from a local business in our country. These observations are used to make the necessary generations through EViews10 package, for identifying the profit trend of cosmetics products.

The reason why we had not applied the statistical methods of the Linear Regression model, is because the relation between the variables taken in study is not linear, and is not free of heteroskedasticity and autocorrelation, because the particular probability of the Observations*Squared is 0.0000 < 0.05. In this way the null hypothesis that verifies the inexistence of the autocorrelation and heteroscedasticity cannot be accepted. So, the fact that the Chi-Square probabilities in the tests of autocorrelations and heteroskedasticity for the both years are 0.0000, lower than 5%, means that the two models are not free of heteroskedasticity and autocorrelations.

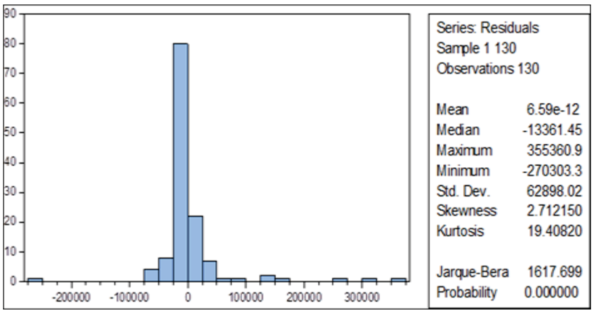

Fig. 1 The Normality Test for the Linear Regression Model of 2019 year's data

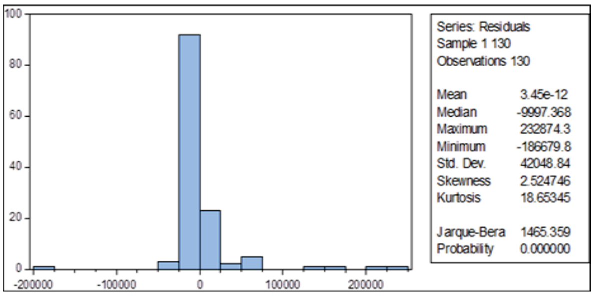

Fig. 2 The Normality Test for the Linear Regression Model of 2020 year's data

Breusch-Godfrey Serial Correlation LM Test (2019)

F-statistic	11.02827	Prob. F(2,126)	0.0000
Obs*R-squared	19.36659	Prob. Chi-Square(2)	0.0001

Test Equation:
Dependent Variable:RESID
Method: Least Squares
Date: 02/10/21 Time: 14:59
Sample: 1 130
Included observations: 130
Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1438.918	5599.924	0.256953	0.7976
SALES_UNIT2020	-6.320355	9.617993	-0.657139	0.5123
RESID(-1)	0.415617	0.089324	4.652913	0.0000
RESID(-2)	-0.099451	0.088556	-1.123037	0.2636

R-squared	0.148974	Mean dependent var	6.59E-12
Adjusted R-squared	0.128711	S.D. dependent var	62898.02
S.E. of regression	58710.8	Akaike info criterion	24.82892
Sum squared resid	4.34E+11	Schwarz criterion	24.91715
Log likelihood	-1609.88	Hannan-Quinn criter.	24.86477
F-statistic	7.35218	Durbin-Watson stat	1.970731
Prob(F-statistic)	0.00014		

Fig. 3 The Serial Correlation test for the Linear Regression Model of 2019 year's data

In the first model of 2019, when the median value of the sold units has an increase of 1%, the profits will increase with 143.14%. So, it is easy to identify that the relation between these two variables is strong. The statistical importance of the sold units in the generated profits is considerable, because the probability is 0.0000 < 0.05. This strong relation between these variables is identified even in the model of 2020, but these we have a different tendency. So, when the median value of the sold units of cosmetics products has an increase of 1%, the generated profits will increase with 141%. This changed tendency is because of the pandemic situations that affected the demand of girl for cosmetics products. So, we have a decrease in the purchases of cosmetics products, because at least some of these products are useless, because of wearing masks during the day. However, the sold units have a statistical importance on the generated profits.

Breusch-Godfrey Serial Correlation LM Test (2020)

F-statistic	8.597132	Prob. F(2,126)	0.0003
Obs*R-squared	15.60994	Prob. Chi-Square(2)	0.0004

Test Equation:
Dependent Variable:RESID
Method: Least Squares
Date: 02/10/21 Time: 15:01
Sample: 1 130
Included observations: 130
Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	729.0175	3757.082	0.194038	0.8465
SALES_UNIT2020	-4.596872	8.537285	-0.538447	0.5912
RESID(-1)	0.364978	0.089529	4.076639	0.0001
RESID(-2)	-0.058543	0.088891	-0.658592	0.5114

R-squared	0.120076	Mean dependent var	3.45E-12
Adjusted R-squared	0.099126	S.D. dependent var	42048.84
S.E. of regression	39910.4	Akaike info criterion	24.05695
Sum squared resid	2.01E+11	Schwarz criterion	24.14518
Log likelihood	-1559.702	Hannan-Quinn criter.	24.09280
F-statistic	5.731422	Durbin-Watson stat	1.975683
Prob(F-statistic)	0.001041		

Fig. 4 The Serial Correlation test for the Linear Regression Model of 2020 year's data

Heteroskedasticity Test: Breusch-Pagan-Godfrey (2019)

F-statistic	25.22631	Prob. F(1,128)	0.0000
Obs*R-squared	21.40246	Prob. Chi-Square(1)	0.0000
Scaled explained SS	190.9758	Prob. Chi-Square(1)	0.0000

Test Equation:
Dependent Variable:RESID^2
Method: Least Squares
Date: 02/10/21 Time: 15:00
Sample: 1 130
Included observations: 130

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.03E+09	1.48E+09	0.699362	0.4856
SALES_UNIT2020	12636064	2515851	5.02258	0.0000

R-squared	0.164634	Mean dependent var	3.93E+09
Adjusted R-squared	0.158108	S.D. dependent var	1.69E+10
S.E. of regression	1.55E+10	Akaike info criterion	49.78319
Sum squared resid	3.08E+22	Schwarz criterion	49.82731
Log likelihood	-3233.908	Hannan-Quinn criter.	49.80112
F-statistic	25.22631	Durbin-Watson stat	1.50046
Prob(F-statistic)	0.000002		

Fig. 5 The heteroscedasticity test for the Linear Regression Model for 2019 year's data

Heteroskedasticity Test: Breusch-Pagan-Godfrey (2020)

F-statistic	24.91218	Prob. F(1,128)	0.0000
Obs*R-squared	21.17937	Prob. Chi-Square(1)	0.0000
Scaled explained SS	181.2366	Prob. Chi-Square(1)	0.0000

Test Equation:
Dependent Variable:RESID^2
Method: Least Squares
Date: 02/10/21 Time: 15:01
Sample: 1 130
Included observations: 130

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.04E+08	6.39E+08	0.944908	0.3465
SALES_UNIT2020	7194458	1441426	4.99121	0.0000

R-squared	0.162918	Mean dependent var	1.75E+09
Adjusted R-squared	0.156379	S.D. dependent var	7.40E+09
S.E. of regression	6.80E+09	Akaike info criterion	48.13265
Sum squared resid	5.91E+21	Schwarz criterion	48.17676
Log likelihood	-3126.62	Hannan-Quinn criter.	48.15057
F-statistic	24.91218	Durbin-Watson stat	1.657908
Prob(F-statistic)	0.000002		

Fig. 6 The heteroscedasticity test for the Linear Regression Model for 2020 year's data

Dependent Variable: PROFIT2019
Method: Quantile Regression (Median)
Date: 02/09/21 Time: 21:01
Sample: 1 130
Included observations: 130
Huber Sandwich Standard Errors & Covariance
Sparsity method: Kernel (Epanechnikov) using residuals
Bandwidth method: Hall-Sheather, bw=0.19179
Estimation successfully identifies unique optimal solution

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	12909.71	1769.685	7.294921	0.0000
SALES_UNIT_2019	143.1429	6.387266	22.41067	0.0000

Pseudo R-squared	0.390078	Mean dependent var	55550.89
Adjusted R-squared	0.385313	S.D. dependent var	91839.30
S.E. of regression	64835.38	Objective	1768740.
Quantile dependent var	22070.00	Restr. Objective	2899943.
Sparsity	40509.30	Quasi-LR statistic	223.3963
Prob(Quasi-LR stat)	0.000000		

Fig. 7 The Quantile Regression Model of the 2019 year's data

Dependent Variable: PROFIT_2020
Method: Quantile Regression (Median)
Date: 02/09/21 Time: 21:14
Sample: 1 130
Included observations: 130
Huber Sandwich Standard Errors & Covariance
Sparsity method: Kernel (Epanechnikov) using residuals
Bandwidth method: Hall-Sheather, bw=0.19179
Estimation successfully identifies unique optimal solution

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	7075.527	1074.773	6.583273	0.0000
SALES_UNIT2020	141.0085	9.263846	15.22138	0.0000

Pseudo R-squared	0.406264	Mean dependent var	36395.49
Adjusted R-squared	0.401625	S.D. dependent var	63516.50
S.E. of regression	44147.96	Objective	1179896.
Quantile dependent var	12044.10	Restr. objective	1987239.
Sparsity	24475.46	Quasi-LR statistic	263.8865
Prob(Quasi-LR stat)	0.000000		

Fig. 8 The Quantile Regression Model of the 2020 year's data

The adjusted R square expresses the percentage of the generated profits variance from the sold units of cosmetics products, So, the both models have approximately the same percentage, 38.5% for 2019 and 40.2% for 2020. The 61.5% of the variance of the generated profits for 2019 is explainable by other variables that are not considered in this model. The same logic is followed for 2020 too, related to the adjusted R-square. The particular median variances for 2019 and 2020 are high, because of the R-square results and followed logic. So, the relation is strong, but not optimal, because there are other influences or variables that together can describe in a most appropriate way the model.

In Fig. 9 is presented the graphics of the distribution of the standard error of the coefficient of the independent variable that is the units of cosmetics products that are sold, for every one of the 10 selected quantiles, for the quantiles' test process.

From Figs. 7 and Fig. 8, it is evident that the standard error of the 2020 model is higher than the 2019 one. That is why the distribution of standard errors of the sold units in 2020 have higher values for the first to the 10-th quantile, compared with 2019.

In Figs. 10 and 11 are presented the results of the Quantile Regression Model after applying the Ordinary IID method that is related to Edgeworth Approximation. The results are sensitively improved, because of the standard errors' minimization for the both years.

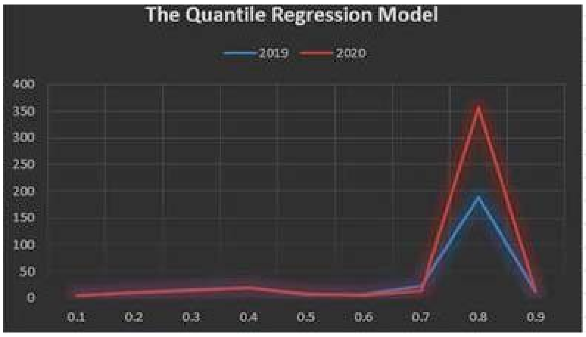

Fig. 9 The graphic of the standard errors' distribution for the Quantile Regression Model of the 2019 and 2020 years

Dependent Variable: PROFIT2019
Method: Quantile Regression (Median)
Date: 02/09/21 Time: 21:03
Sample: 1 130
Included observations: 130
Ordinary (IID) Standard Errors & Covariance
Sparsity method: Kernel (Epanechnikov) using residuals
Bandwidth method: Hall-Sheather, bw=0.19179
Estimation successfully identifies unique optimal solution

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	12909.71	1950.852	6.617473	0.0000
SALES_UNIT_2019	143.1429	3.321782	43.09220	0.0000

Pseudo R-squared	0.390078	Mean dependent var	55550.89
Adjusted R-squared	0.385313	S.D. dependent var	91839.30
S.E. of regression	64835.38	Objective	1768740.
Quantile dependent var	22070.00	Restr. Objective	2899943.
Sparsity	40968.23	Quasi-LR statistic	220.8938
Prob(Quasi-LR stat)	0.000000		

Fig. 10 The application of Edgeworth approximation for IID cases in 2019 data

Dependent Variable: PROFIT_2020
Method: Quantile Regression (Median)
Date: 02/09/21 Time: 21:18
Sample: 1 130
Included observations: 130
Ordinary (IID) Standard Errors & Covariance
Sparsity method: Siddiqui using residuals
Bandwidth method: Hall-Sheather, bw=0.19179
Estimation successfully identifies unique optimal solution

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	7075.527	1042.195	6.789060	0.0000
SALES_UNIT2020	141.0085	2.350360	59.99445	0.0000

Pseudo R-squared	0.406264	Mean dependent var	36395.49
Adjusted R-squared	0.401625	S.D. dependent var	63516.50
S.E. of regression	44147.96	Objective	1179896.
Quantile dependent var	12044.10	Restr. objective	1987239.
Sparsity	22166.23	Quasi-LR statistic	291.3776
Prob(Quasi-LR stat)	0.000000		

Fig. 11 The application of Edgeworth approximation for IID cases in 2020 data

In Fig. 12 are graphically presented the results of the model after applying the Edgeworth approximation. In this case it is easy to identify the positive results of reducing the standard errors of the sold units' coefficient in the two years.

In Figs. 13 and 14 is presented the Bootstrap Version for the Quantile Regression Model for 2019 and 2020. The method used is Bootstrap Residual, for 10,000 replications. Even in this case the standard error for the independent variable taken is minimized and the curve of standard errors is decreased compared with the curve of the 2019 and 2020 quantile regression model.

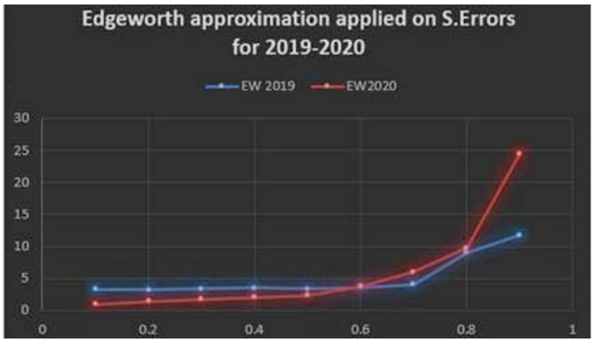

Fig. 12 The graphic of the standard errors' distribution after the application of the Edgeworth for IID Cases for 2019 and 2020 years

Dependent Variable: PROFIT2019
Method: Quantile Regression (Median)
Date: 02/09/21 Time: 21:05
Sample: 1 130
Included observations: 130
Bootstrap Standard Errors & Covariance
Bootstrap method: Residual, reps=10000, rng=kn, seed=633957423
Sparsity method: Siddiqui using residuals
Bandwidth method: Hall-Sheather, bw=0.19179
Initial Values: C(1)=0.00000, C(2)=0.00000
Estimation successfully identifies unique optimal solution

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	12909.71	1619.185	7.972970	0.0000
SALES_UNIT_2019	143.1429	3.329511	42.99217	0.0000

Pseudo R-squared	0.390078	Mean dependent var	55550.89
Adjusted R-squared	0.385313	S.D. dependent var	91839.30
S.E. of regression	64835.38	Objective	1768740.
Quantile dependent var	22070.00	Restr. Objective	2899943.
Sparsity	34588.92	Quasi-LR statistic	261.6337
Prob(Quasi-LR stat)	0.000000		

Fig. 13 The application of Edgeworth approximation of Bootstrap Residual in 2019 data

Dependent Variable: PROFIT_2020
Method: Quantile Regression (Median)
Date: 02/09/21 Time: 21:36
Sample: 1 130
Included observations: 130
Bootstrap Standard Errors & Covariance
Bootstrap method: Residual, reps=10000, rng=kn, seed=911414654
Sparsity method: Kernel (Epanechnikov) using residuals
Bandwidth method: Hall-Sheather, bw=0.19179
Initial Values: C(1)=0.00000, C(2)=0.00000
Estimation successfully identifies unique optimal solution

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	7075.527	1054.578	6.709346	0.0000
SALES_UNIT2020	141.0085	2.914924	48.37469	0.0000

Pseudo R-squared	0.406264	Mean dependent var	36395.49
Adjusted R-squared	0.401625	S.D. dependent var	63516.50
S.E. of regression	44147.96	Objective	1179896.
Quantile dependent var	12044.10	Restr. objective	1987239.
Sparsity	24498.41	Quasi-LR statistic	263.6393
Prob(Quasi-LR stat)	0.000000		

Fig. 14 The application of Edgeworth approximation of Bootstrap Residual in 2020 data

Figs. 16 and 17 present the generated results for the 2019 and 2020 after the application of Edgeworth Approximation in the Bootstrap version of the Quantile Regression Model. The results are more improved, because of the sensitively reduction of the standard errors of the sold units' coefficient. So, we can conclude that the Edgeworth Approximation in the Bootstrap Quantile Regression Model is the optimal solution of the model's estimation.

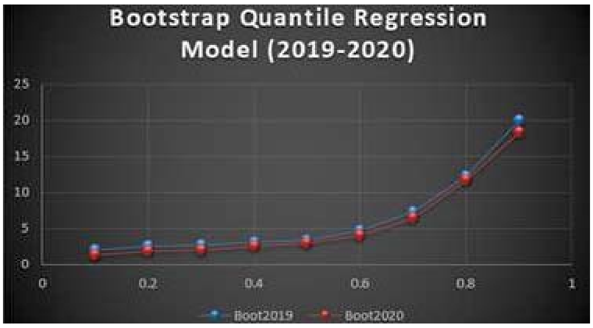

Fig.15 The graphic of the standard errors' distribution after the application of the Bootstrap Residual for 2019 and 2020

Dependent Variable: PROFIT2019
Method: Quantile Regression (Median)
Date: 02/09/21 Time: 21:06
Sample: 1 130
Included observations: 130
Ordinary (IID) Standard Errors & Covariance
Sparsity method: Siddiqui using residuals
Bandwidth method: Hall-Sheather, bw=0.19179
Initial Values: C(1)=0.00000, C(2)=0.00000
Estimation successfully identifies unique optimal solution

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	12909.71	1647.078	7.837947	0.0000
SALES_UNIT_2019	143.1429	2.804536	51.03979	0.0000

Pseudo R-squared	0.390078	Mean dependent var	55550.89
Adjusted R-squared	0.385313	S.D. dependent var	91839.30
S.E. of regression	64835.38	Objective	1768740.
Quantile dependent var	22070.00	Restr. Objective	2899943.
Sparsity	34588.92	Quasi-LR statistic	261.6337
Prob(Quasi-LR stat)	0.000000		

Fig. 16 The application of Edgeworth approximation for Bootstrap Quantile Regression Model in 2019 year's data

Dependent Variable: PROFIT_2020
Method: Quantile Regression (Median)
Date: 02/09/21 Time: 21:34
Sample: 1 130
Included observations: 130
Ordinary (IID) Standard Errors & Covariance
Sparsity method: Siddiqui using residuals
Bandwidth method: Hall-Sheather, bw=0.19179
Initial Values: C(1)=0.00000, C(2)=0.00000
Estimation successfully identifies unique optimal solution

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	7075.527	1042.195	6.789060	0.0000
SALES_UNIT2020	141.0085	2.350360	59.99445	0.0000

Pseudo R-squared	0.406264	Mean dependent var	36395.49
Adjusted R-squared	0.401625	S.D. dependent var	63516.50
S.E. of regression	44147.96	Objective	1179896.
Quantile dependent var	12044.10	Restr. objective	1987239.
Sparsity	22166.23	Quasi-LR statistic	291.3776
Prob(Quasi-LR stat)	0.000000		

Fig. 17 The application of Edgeworth approximation for Bootstrap Quantile Regression Model in 2020 year's data

The optimal results of the Edgeworth Approximation for Bootstrap Quantile Regression Model for 2019 and 2020 are shown in Fig. 18, and the tendency is the approximation of the standard errors' distribution with zero.

Based on all generated results, in Fig. 19 is created a summary form of all the graphics for the both years, and the conclusion is the same, that the Bootstrap Edgeworth Approximation is the best solution.

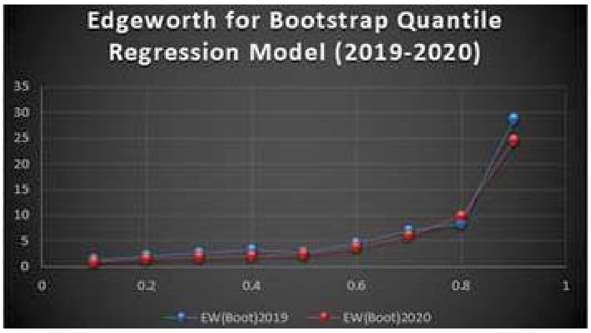

Fig. 18 The graphic of the standard errors' distribution after the application of Edgeworth Approximation for Bootstrap Quantile Regression Model for 2019 and 2020

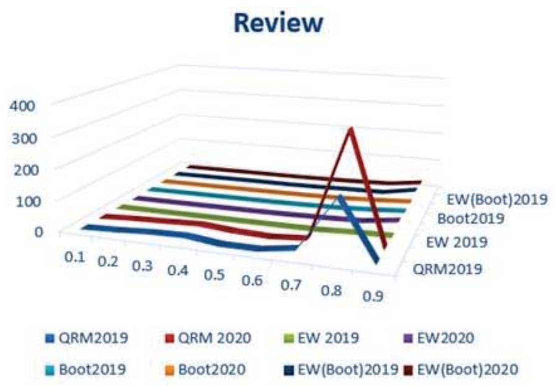

Fig. 19 The summary graph

VI. Conclusion

In this paper, we have treated an economical problem using mathematical tools, like Edgeworth Approximation for independent and identical distributed cases, Bootstrap, Edgeworth Approximation for Bootstrap version. We have applied these methods in our model and with the help of EViews10 software, we have done the necessary simulations. At the end, we conclude that the best approximation method in our study is the Bootstrap Edgeworth Approximation.

References

[1] A. H. Welsh, Asymptotically Efficient Estimation of the Sparsity Function at a Point, Statistics & Probability Letters, 6 (1988), 427-432.
[2] D. A. Bloch and J. L. Gastwirth, On a simple estimate of the reciprocal of the density function, Ann. Math. Statist. 39 (1968), 1083-1085.
[3] G. J. Bassett and R. Koenker, An Empirical Quantile Function for Linear Models with i.i.d. Errors," Journal of the American Statistical Association, 77(378) (1982), 407-415.
[4] H. A. David, Order Statistics, 2nd ed. Wiley, New York, 1981.
[5] J. Powell, Censored Regression Quantiles, Journal of Econometrics, 32 (1986), 143-155.
[6] M. Buchinsky, Estimating the Asymptotic Covariance Matrix for Quantile Regression Models: A Monte Carlo Study, Journal of Econometrics, 68 (1995), 303-338.
[7] M. C. Jones, Estimating Densities, Quantiles, Quantile Densities and Density Quantiles, Annals of the Institute of Statistical Mathematics, 44(4) (1992), 721-727.
[8] M. Kocherginsky, X. He, and Y. Mu, Practical Confidence Intervals for Regression Quantiles, Journal of Computational and Graphical Statistics, 14(1) (2005), 41-55.
[9] M. M. Siddiqui, Distribution of Quantiles in Samples from a Bivariate Population, Journal of Research of the National Bureau of Standards–B, 64(3) (1960), 145-150.
[10] P. Hall, The Bootstrap and Edgeworth Expansion, Springer-Verlag, New York, USA, 1992.
[11] P. Hall and M. A. Martin, On the error incurred using the bootstrap variance estimate when constructing confidence intervals for quantiles, J. Mult. Anal. 38 (1991). 70-81.
[12] P. Hall and S. J. Sheather, On the Distribution of the Studentized Quantile, Journal of the Royal Statistical Society, Series B, 50(3) (1988), 381-391.
[13] R. Koenker and G.J. Bassett, Regression Quantiles, Econometrica, 46(1) (1978), 33-50.
[14] R. Koenker, P. Mandl and M. Huskova, Confidence Intervals for Regression Quantiles, in Asymptotic Statistics, eds., New York: Springer-Verlag, (1994), 349-359.
[15] X. He and F. Hu, Markov Chain Marginal Bootstrap, Journal of the American Statistical Association, 97(459) (2002), 783-795.
[16] http://www.eviews.com/help/helpintro.html#page/content%2Fquantreg-Background.html%23
[17] https://idoc.pub/documents/eviews-8-users-guide-ii-on2397q8dyl0