SILESIAN UNIVERSITY OF TECHNOLOGY PUBLISHING HOUSE

SCIENTIFIC PAPERS OF SILESIAN UNIVERSITY OF TECHNOLOGY 2025
ORGANIZATION AND MANAGEMENT SERIES NO. 226

THE MODAL REGRESSION MODEL AS A MANAGERIAL INSTRUMENT IN LOGISTICS AND MARKETING

Joanna BRUZDA

Faculty of Economic Sciences and Management, Nicolaus Copernicus University in Toruń;
Joanna.Bruzda@umk.pl, ORCID: 0000-0002-7451-3592

Purpose: The purpose of this paper is to identify possible directions of applications of the modal regression model in enterprise management, with a focus on the areas of logistics and marketing, and to illustrate the use of such model in constructing helpful managerial instruments. The modal regression model serves the purpose of modeling the mode of a given dependent variable conditionally on certain covariates, and, thus, it supplements both the traditional analysis based on modeling the conditional mean and the quantile regression approach. In this paper, we show that this statistical technique can be of help in developing useful decision instruments for logistics and marketing management.
Design/methodology/approach: The paper is based on a review of the relevant statistical literature, conceptual work, simulations, and two empirical studies.
Findings: It is shown that the modal regression model can find uses, among other things, in demand forecasting and logistics planning, optimization of industrial processes, and designing marketing mix strategies. It will enable a direct computation of the demand forecasts minimizing the total logistics cost, modeling the nominal value of an industrial process in a way guaranteeing its high capability to meet the highest quality standards, and building marketing mix models for a target consumer group without the need for market segmentation.
Research limitations/implications: In further research, it will be interesting to investigate possible directions of applications of the multivariate dynamic counterpart of the model studied in this paper, namely, the dynamic vector mode regression, in particular, in discovering the most probable future scenarios for an enterprise.
Practical implications: Our findings suggest that the modal regression model should be popularized among logistics and marketing managers and included in their managerial statistical toolbox.
Originality/value: This paper, for the first time in the literature, draws attention to the empirical modal regression model as a potential instrument in enterprise management and illustrate the use of modal regression in logistics and marketing management. The findings included in the paper should benefit both researchers looking for the most appropriate statistical approach for their research as well as logistics and marketing managers building their decision instruments.
Keywords: modal regression, managerial instruments, marketing mix models, demand forecasting, statistical process control.
Category of the paper: general review, conceptual paper.

1. Introduction

The statistical literature over the past several years has established a formal statistical framework for a new analytical tool – the modal regression model, which serves the purpose of modeling the mode of the distribution of a given response variable or a vector of such variables of interest to the researcher. The concept of modal regression was introduced gradually, starting with linear modal regression for iid (independent and identically distributed) samples (Kemp, Santos Silva, 2012; Yao, Li, 2014) and nonparametric modal regression (Yao et al., 2012; Chen et al., 2016; Yang et al., 2020; Xiang, Yao, 2022), through modal regression for panel data (Ullah et al., 2021) and nonlinear modal regression for dependent samples (Ullah et al., 2022), and ending up with dynamic vector mode regression (Kemp et al., 2020) and the nonparametric estimation of the modal volatility function (Ullah, Wang, 2025).

On the other hand, applications of this concept, especially in economic sciences, are rare and mainly serve illustrating the statistical approach without discussing its merits in economic rational decision-making. Past applications of modal regression in the economic context encompass modeling per capita expenditures on public education in the U.S. (Yao et al., 2012), macroeconomic forecasting and investigating dynamic interrelations between inflation, unemployment, and interest rates in the U.S. (Kemp et al., 2020), modeling house prices in Boston (Yang et al., 2020), modeling the productivity of public capital in the U.S. (Ullah et al., 2021), and forecasting the industrial output in the OECD countries (Bruzda, 2024). As a motivation to introduce modal regression, authors such as Kemp and Santos Silva (2012) indicate econometric analyses serving the purpose of modeling such economic variables as income, prices, and expenditures, which are characterized by positive skewness. These authors also invoke the notion of growth regression, which, according to the researchers dealing with the empirics of growth theory, should fit the majority of the data. As noted by Kemp and Santos Silva (2012), in order to realize this goal, in the absence of appropriate tools for modeling the modal value of growth rates, they rely on robust regression.

Since the author of this paper is not aware of any theoretical or empirical work devoted to applications of modal regression in the fields of management and marketing, this paper aims to fill this research gap and indicate such applications jointly with their advantages and benefits. For this purpose, we present two empirical analyses on data from some publically available databases (as well as a supplementary simulation experiment), which exemplify the use of linear modal regression in logistics and marketing, showing the potential of the modal regression model in developing useful decision instruments for enterprise management. In addition, we also compare the modal and robust regression models, indicating their similarities and dissimilarities.

Further in this paper, the statistical fundamentals of the modal regression model and its relationship with the robust regression model are presented, and the above mentioned example managerial applications are discussed in detail. The last section briefly summarizes the main points of our discussion.

2. Modal regression – the statistical fundamentals

In this paper, we focus is on the linear modal regression model for dependent samples in the following form:

$$\begin{aligned} y_t &= \mathbf{x}_t\boldsymbol{\beta}^T + \varepsilon_t = \beta_0 + \beta_1 x_{1t} + \ldots + \beta_k x_{kt} + \varepsilon_t \\ Mo(y_t \mid \mathbf{x}_t) &= \mathbf{x}_t\boldsymbol{\beta}^T = \beta_0 + \beta_1 x_{1t} + \ldots + \beta_k x_{kt} \end{aligned} \tag{1}$$

where $t = 1,2,...$ is the index denoting the consecutive time units, y_t is the dependent variable, $\mathbf{x}_t = [1, x_{1t}, ..., x_{kt}]$ is a vector of independent variables, $\boldsymbol{\beta} = [\beta_0, \beta_1, ..., \beta_k]$ is a vector of structural parameters, $Mo(y_t \mid \mathbf{x}_t)$ denotes the conditional modal value of y_t, and ε_t denotes an iid sequence fulfilling the condition $Mo(\varepsilon_t \mid \Omega_t) = 0$ a.s., where Ω_t is the σ-algebra generated by $\{x_{is}, \varepsilon_r\}_{i=1,...,k, s\le t, r<t}$. Moreover, it is assumed that the conditional distribution of $\varepsilon_t = y_t - \mathbf{x}_t\boldsymbol{\beta}^T$ is absolutely continuous, supported on the real line, and has a well-defined global mode (Ullah et al., 2022). It is allowed that the errors ε_t are conditionally heteroskedastic, and, furthermore, it is assumed that the process $\{y_t, \mathbf{x}_t\}$ is stationary and fulfills certain conditions related to its memory (the so-called mixing conditions).

Having a sample $(y_t, x_{1t}, ..., x_{kt}, t = 1, ..., n)$, in order to estimate the structural parameters, one can solve the optimization problem formulated as:

$$\frac{1}{n}\sum_t \phi_h\left(y_t - \mathbf{x}_t\boldsymbol{\beta}^T\right) \to \max , \tag{2}$$

where $\phi_h(\varepsilon) = \frac{1}{h}\phi\left(\frac{\varepsilon}{h}\right)$, and $\phi(\varepsilon)$ is a smooth symmetric kernel function with a weak maximum at zero, while *h* is its bandwidth parameter (Yao, Lee, 2014; Ullah et al., 2022). Alternatively, the estimation problem can be formulated as (Kemp, Santos Silva, 2012):

$$\sum_i \left\{1 - \gamma \frac{1}{n} \sum_i \phi\left(\frac{y_i - \mathbf{x}_i\boldsymbol{\beta}^T}{h}\right)\right\} \to \min, \tag{3}$$

where $\gamma = \phi^{-1}(0)$. The usual choice for ϕ is the Gaussian kernel, i.e., the probability density function of the standard normal distribution. This is also the choice later in this paper.

In order to obtain consistency for the parameters of the conditional modal line, the bandwidth parameter needs to slowly converge to zero at a certain prespecified rate. Namely, it is required that, assuming $n \to \infty$, it holds $h \to 0$ and $nh^5 \to \infty$. Then, we get (for example Yao, Li, 2014):

$$\left\| \hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \right\| = O((nh^3)^{-0,5} + h^2). \tag{4}$$

The asymptotic normality result further strengthens the requirements regarding this rate of convergence. Namely, it is assumed that $nh^7 = O(1)$ (Ullah et al., 2022), which leads to the following large sample property:

$$\sqrt{nh^3}\left[\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} - \frac{h^2}{2} J^{-\mathbf{1}} K\{1 + o(1)\} \right] \Rightarrow N\left\{0, \nu_2 J^{-\mathbf{1}} L J^{-\mathbf{1}}\right\}, \tag{5}$$

where $\nu_2 = \int t^2 \phi^2(t)dt$, $J = E\left[f''(0 \mid \mathbf{x}_t)\mathbf{x}_t\mathbf{x}_t^T\right]$, $K = E\left[f'''(0 \mid \mathbf{x}_t)\mathbf{x}_t\right]$, $L = E\left[f(0 \mid \mathbf{x}_t)\mathbf{x}_t\mathbf{x}_t^T\right]$, and $f(\cdot)$ denotes the probability density function of the error term. If $nh^7 \to 0$, the large sample result (5) simplifies as follows:

$$\sqrt{nh^3}\left[\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right] \Rightarrow N\left\{0, \nu_2 J^{-\mathbf{1}} L J^{-\mathbf{1}}\right\}. \tag{6}$$

To solve the optimization problem (2), Yao and Lee (2014) introduced a modification of the Expectation-Maximization (EM) algorithm that iteratively maximizes the expected value of the log-likelihood function. The suggested modal variation of the EM algorithm, called the Modal Expectation-Maximization (MEM) algorithm, in the case of the Gaussian kernel is equivalent to the iteratively reweighted least squares method.

The computational process is as follows. First, it is assumed that we found a good robust initial estimate of $\boldsymbol{\beta}$, denoted $\hat{\boldsymbol{\beta}}^{(0)}$, as an input to the further computation procedure. Then, two types of computation steps, the so-called E-step and M-step, are repeated until convergence. These steps are as follows (Yao, Li, 2014):

E-step: Calculate the weights:

$$\pi\left(t \mid \hat{\boldsymbol{\beta}}^{(k)}\right) = \frac{\phi_h\left(y_t - \mathbf{x}_t^T \hat{\boldsymbol{\beta}}^{(\mathbf{k})}\right)}{\sum_t \phi_h\left(y_t - \mathbf{x}_t^T \hat{\boldsymbol{\beta}}^{(\mathbf{k})}\right)}. \tag{7}$$

M-step: Update the estimates:

$$\hat{\boldsymbol{\beta}}^{(k+1)} = \arg\max_{\hat{\boldsymbol{\beta}}} \sum_t \left\{\pi\left(t \mid \hat{\boldsymbol{\beta}}^{(k)}\right) \log \phi_h\left(y_t - \mathbf{x}_t^T \hat{\boldsymbol{\beta}}\right)\right\} = \left(\mathbf{X}^T \mathbf{W}_k \mathbf{X}\right)^{-1} \mathbf{X}^T \mathbf{W}_k \mathbf{Y}, \tag{8}$$

where $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2 ..., \mathbf{x}_n)^T$, $Y = (y_1, y_2 ..., y_n)^T$, and

$$\mathbf{W}_k = \begin{pmatrix} \pi\left(1 \mid \hat{\boldsymbol{\beta}}^{(k)}\right) & & & \mathbf{0} \\ & \pi\left(2 \mid \hat{\boldsymbol{\beta}}^{(k)}\right) & & \\ & & \ddots & \\ \mathbf{0} & & & \pi\left(n \mid \hat{\boldsymbol{\beta}}^{(k)}\right) \end{pmatrix}. \tag{9}$$

In the computations later in the simulation and empirical parts of the paper, the initial estimates in the MEM algorithm are set to those obtained with the robust linear regression estimation with the Huber function.

Yao and Li (2014) also derived the value of the bandwidth parameter *h*, which minimizes the asymptotic mean squared error (MSE) of the modal regression estimator (see the formula (9) in Yao, Li, 2014). We use this solution in our computations using the usual plugged-in estimates of the elements *K* and *L* defined in (5). In nonparametric density estimation, we rely on the method introduced in Botev et al. (2010), as suggested in the Matlab codes for modal regression made available by W. Yao.

3. Modal regression vs. robust regression

Popular estimators of the parameters in a regression equation, such as the least squares method, are sensitive to the presence of atypical observations in data samples. This will take place both when the atypical value concerns a certain independent variable, in which case the so-called high-leverage observation is identified, and when there is an atypical value of the response variable, in which case we talk of an outlying observation. In order to enable the proper examination of the dependence in the presence of atypical observations, one is advised to apply robust regression methods. Among the most frequently used estimations techniques from this group, the following estimation approaches can be indicated: M-estimators, GM-estimators, S-estimators, and MM-estimators (for example Hampel et al., 1986; Maronna et al., 2006).

M-estimators of the parameters in a regression equation, introduced by Huber in 1973 (see Huber, 1973), serve the purpose of removing or elevating the problem of atypical values in the response (the problem of large residuals). The M-estimation of the parameters in a regression line generalizes the estimation approach with the least squares method or the quantile regression method, whereas the term ‘M-estimation’ itself is an abbreviation of the phrase ‘Maximum likelihood-like estimation’.

However, when the M-estimation is applied, the so-called influence function, which measures the influence of the ‘contamination’ of the distribution on the obtained estimates, turns out to be unbounded with respect to high-leverage observations. GM-estimators, which is an abbreviation of Generalized M-estimators, were introduced for the purpose of elevating the impact of high-leverage observations through appropriately defined weights reducing their influence on estimates. S-estimators of regression equations are procedures characterized by a high finite sample breakdown point. This means that they admit a large fraction of atypical observations before the estimator returns aberrant results. They are based on a robust estimation of the scale parameter, from which the name ‘S-estimator’ comes. Finally, MM-estimators are procedures joining S-estimation (in the first step) and M-estimation (in the second step), which are able to combine the merits of S-estimators with high efficiency for the normal distribution.

In the M-estimation of a regression line, one is looking for estimates of parameters of the equation in the form:

$$y_t = \beta_0 + \beta_1 x_{t1} + \ldots + \beta_k x_{tk} + \varepsilon_t$$

minimizing the sum:

$$\sum_t \rho_c\left(\frac{y_t - b_0 - b_1 x_{t1} - \ldots - b_k x_{tk}}{\hat{\sigma}}\right). \quad (10)$$

It is usually assumed that the objective function ρ_c is convex, admits the minimum value of zero at zero and has a derivative ψ_c. The quantity $\hat{\sigma}$ is a robust estimate of the standard deviation of the error term. Its presence in the function (10) results from the fact that the estimation result obtained through the minimization of the expression (10) is not scale equivariant, that is, it depends on the measurement units used (for example Maronna et al., 2006, Chapter IV). The solution to the optimization problem defined through (10) leads to the so-called studentized M-estimator.

In practice, the M-estimation of the parameters of the regression equation is usually performed with the help of the function ψ_c, called score function. This takes place if the score function ψ_c is smooth. Then, finding the minimum of (10) requires solving the equation:

$$\sum_t x_{ij}\psi_c\left(\frac{y_t - \mathbf{x}_t\mathbf{b}}{\hat{\sigma}}\right) = \sum_t x_{tj}\psi_c\left(\frac{e_t}{\hat{\sigma}}\right) = 0, \qquad j = 1,\ldots,k. \quad (11)$$

The condition (11), after defining an appropriate weight function, reduces the computation procedure to the iterated weighted least squares method.

Table 1.
Chosen objective and score functions in robust estimation and the least squares method

	Objective function	**Score function**
Least squares method	$\varphi(u) = \frac{1}{2}u^2$	$\psi(u) = u$
Huber function	$\varphi_c(u) = \begin{cases} \frac{1}{2}u^2 & \text{for } \lvert u\rvert \le c \\ c\lvert u\rvert - \frac{c^2}{2} & \text{otherwise} \end{cases}$	$\psi_c(u) = \begin{cases} u & \text{for } \lvert u\rvert \le c \\ c\,\mathrm{sign}(u) & \text{otherwise} \end{cases}$
Tukey function	$\varphi_c(u) = \begin{cases} \frac{c^2}{6}\left[1-\left(1-\frac{u^2}{c^2}\right)^3\right] & \text{for } \lvert u\rvert \le c \\ \frac{c^2}{6} & \text{otherwise} \end{cases}$	$\psi_c(u) = \begin{cases} u\left(1-\frac{u^2}{c^2}\right)^2 & \text{for } \lvert u\rvert \le c \\ 0 & \text{otherwise} \end{cases}$
Welsch function	$\varphi_c(u) = \frac{c^2}{2}\left(1-e^{-\frac{u^2}{c^2}}\right)$	$\psi_c(u) = ue^{-\frac{u^2}{c^2}}$

Note. The constant *c* is usually defined as that guaranteeing 95% asymptotic efficiency for the normal distribution.

Source: own elaboration.

Table 1 contains example objective and score functions used in robust estimation and compares them with the corresponding functions from the least squares estimation. Among the compared functions, the Tukey and Welsch objective functions, similar to the objective functions in modal regression estimation, are bounded. In fact, for the Gaussian kernel, the maximization problem (2) is equivalent to the non-studentized M-estimation based on the Welsch function. It should be noted, however, that robust regression serves the purpose of modeling the conditional mean value and assumes a symmetric error distribution. As indicated in the previous section, it turns out that, with a judicious definition of the bandwidth parameter h in the appropriate objective function, a consistent estimation of the parameters in the equation for the conditional mode (1) can be guaranteed even if the distribution of errors is asymmetric.

4. Benefits of using modal regression

Among the potential benefits of using modal regression, one can indicate the following (compare Bruzda, 2024, and references therein):

- Modal regression enables measuring marginal effects for typical units in the population in a cross-section study (the response of a typical enterprise, client, worker etc.) and typical responses of a single examined unit over time in a time series scenario.
- It is robust to outlying observations and tick tails of the conditional distribution.
- It enables a direct calculation of modal forecasts and the most likely trajectories of the examined phenomena over time (Dimitriadis et al., 2024).
- Under certain conditions, it should result in the narrowest interval forecasts and the smallest values of symmetric and bounded forecast loss functions (Yao, Li, 2014; Chen et al., 2016; Bruzda, 2024).
- It can provide more accurate parameter estimates in the conditions of asymmetry of the error term (Yao, Li, 2014).

Thus, it appears that the modal regression model will constitute a significant addendum to the existing analytical toolbox. The potential area of its applications in management, marketing, economics, and finance is broad, encompassing, among other things,

- the evaluation of the impact of different components of marketing mix strategies on sales in order to maximize the effectiveness of promotional and advertising activities,
- the computation of logistics demand forecasts evaluated with bounded and symmetric loss functions, that is, forecasts associated with limited (and similar) overstocking and understocking costs for material or distribution inventories (composed of, for example, the cost of the next transport to replenish or reduce the inventories in the production cell or the point of sale),
- forecasting prices of the underlying financial instruments when using derivatives with payoff functions that depend on the price remaining within a certain band,

- macroeconomic forecasting maximizing credibility measures.

To illustrate the potential of modal regression to find forecasts that minimize symmetric and bounded loss functions, the following simplified simulation experiment was conducted. The data were generated according to the model adopted in simulation experiments in Yao and Li (2014), which is as follows:

$$Y = 1 + 3X + \sigma(X)\varepsilon,\ \ X \sim U(0,1),$$

$$\varepsilon \sim 0{,}5N(-1, 2{,}5^2) + 0{,}5N(1, 0{,}5^2),\ \ \sigma(X) = 1 + 2X.$$

Then, we get:

$$E(Y \mid X) = 1 + 3X,\ \ Mod(Y \mid X) \approx 2 + 5X,\ \ Med(Y \mid X) \approx 1{,}67 + 4{,}34X,$$

$$E(\varepsilon) = 0,\quad Mod(\varepsilon) \approx 1,\quad Med(\varepsilon) \approx 0{,}67.$$

For a sample of size 120 and with 1,000 repetitions, information was collected on chosen accuracy measures for one-step ahead forecasts. The obtained values were the following:

- the MSE: 20,3467 for modal regression and 17,8375 for the least squares method,
- the loss function in modal estimation with the bandwidth parameter h set to an estimate of the asymptotically optimal value of h: 0,8531 for modal regression and 0,8977 for the least squares method,
- the rectangular loss function of the form:

$$L(x, y) = \ell(y - x) = \ell(e) = \begin{cases} 0 & \text{for} \quad |e| \le c \\ 1 & \text{for} \quad |e| > c \end{cases} \tag{12}$$

  with c defined as 10% of the mean value of Y: 0,8470 for modal regression and 0,9560 for the least squares method.

The above experiment confirmed that modal regression is a natural solution for a single-step computation of forecasts defined as values minimizing the zero-one loss or, more broadly, a symmetric and bounded loss function. Figure 1, on an example of the gamma distribution with the mean value of 6 and the modal value of 5, shows that this may occur, in particular, if the distribution of forecast errors is strongly unimodal, and the constant c in the loss function (12) does not take too large a value.

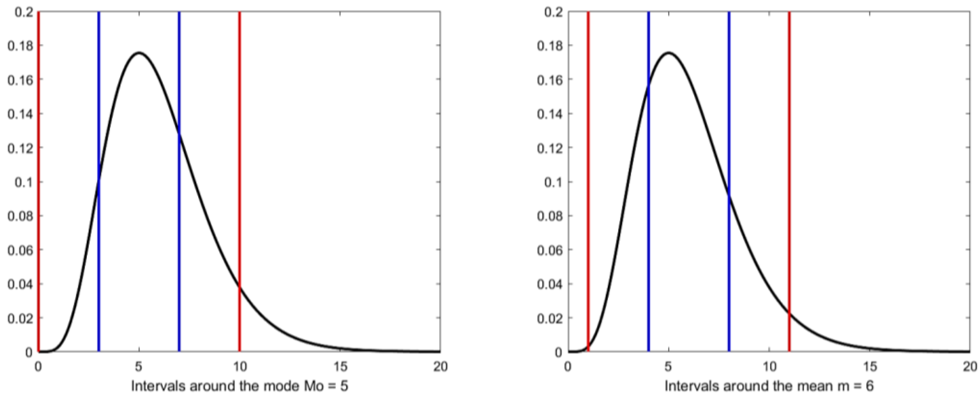

Figure 1. Symmetric intervals of the same width around the mode and mean.

Source: own elaboration.

5. Empirical modal regression models and their applications

Now, we will present two illustrative applications of the modal regression model based on certain real datasets. The first example application concerns modeling the nominal value of an industrial process, assuming that this value depends on the conditions characterizing this process, and more precisely, settings of the device used in the process. Also, we assume that the conditional distribution of the outcome of the process can be asymmetric and even multimodal, but it possesses a well-defined global mode. Under such circumstances, it is expected that modal regression can be employed to model the nominal value of the process in a way guaranteeing that the distribution of the outcome is highly concentrated around this nominal value.

In the example, we use a dataset available on the Kaggle webpade (Guetout, n.d.). It consists of 125 observations on four variables. The first is called ‘roughness’ (denoted further as R) and is considered an important outcome of a turning process since it impacts on the life cycle of the produced steel elements. Thus, this particular outcome of the production process remains of interest to the statistical process control and general quality management. In addition, the dataset includes variables describing the production process such as the cutting depth (X_1), feed (X_2) and cutting and turning speed (X_3).

On this dataset, we estimated parameters of three linear regression models, namely, the ordinary conditional mean equation (estimated by least squares), the quantile regression equation for quantile of order 0.5, that is, median regression (estimated by least absolute deviations), and the modal regression line (estimated as described in the previous sections). After removing the variable X_1, which turned out to be insignificant, the following estimation results were obtained:

$$\hat{R}_1 = \hat{E}(R\,|\,\mathbf{X}) = \underset{(\pm 0{,}362)}{4{,}135} + \underset{(\pm 1{,}618)}{11{,}570}\,X_2 - \underset{(\pm 0{,}002)}{0{,}020}\,X_3, \tag{13}$$

$$\hat{R}_2 = \hat{Me}(R\,|\,\mathbf{X}) = \underset{(\pm 0{,}595)}{2{,}910} + \underset{(\pm 1{,}774)}{7{,}667}\,X_2 - \underset{(\pm 0.004)}{0.011}\,X_3, \tag{14}$$

$$\hat{R}_3 = \hat{Mo}(R\,|\,\mathbf{X}) = \underset{(\pm 0{,}453)}{2{,}447} + \underset{(\pm 2{,}026)}{6{,}993}\,X_2 - \underset{(\pm 0{,}002)}{0{,}008}\,X_3, \tag{15}$$

where $\hat{R}_i$, i = 1, 2, 3, denote the theoretical values of R from the three models, and $\mathbf{X} = [X_1, X_2, X_3]$ is the vector of exogenous variables.

The estimation results show that the regression coefficients in the equation for the mean value (13) are about two times larger than those of the modal regression equation (15), whereas the corresponding coefficients in the median regression line (14) deviate from those in the equation (15) to a lesser extent. But even more important are the shapes of the error distributions in the estimated regressions. Figure 2 presents histograms of the residuals from the three equations, jointly with a histogram for the dependent variable R. The histograms indicate that the residuals from the modal regression model are highly concentrated around zero, much more than in the case of least squares estimation and also more than in the case of median regression, confirming the expectation formulated at the beginning of this empirical study.

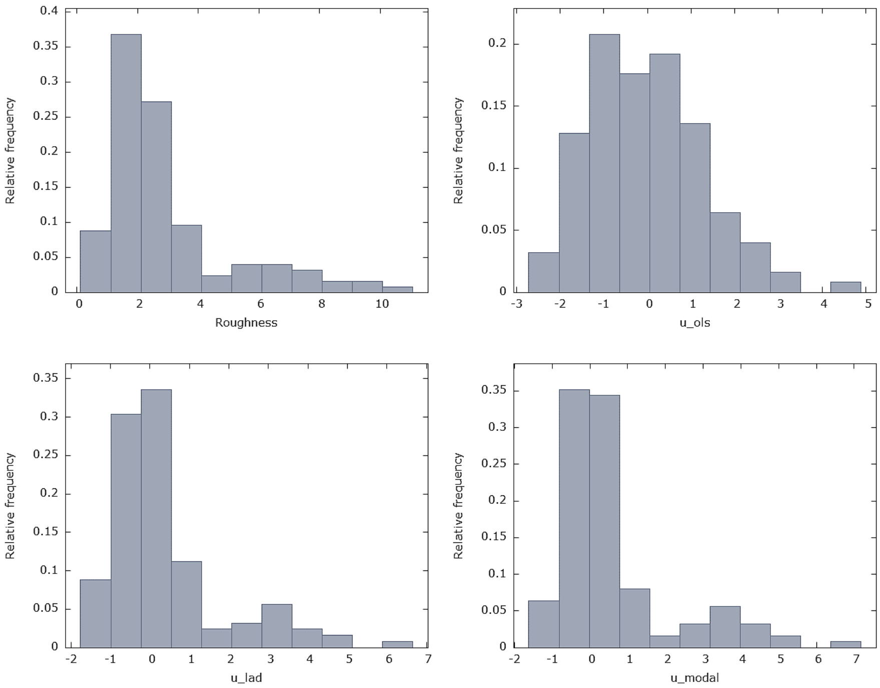

Figure 2. Histograms of the dependent variable and the residuals. Upper left panel – the dependent variable, upper right panel – residuals from least squares estimation, bottom left panel – residuals from median regression, bottom right panel – residuals from modal regression.

Source: own elaboration.

To confirm and quantify this observation, for each of the models, we also found the percentages of residuals with absolute values no greater than a certain constant c, that is, $\frac{1}{n}\sum_{t=1}^{n}\mathbf{1}\left\{\left|R_t - \hat{R}_{it}\right| \le c\right\}$. Such defined quantities can also be computed as $1-\frac{1}{n}\sum_{t=1}^{n}\ell(R_t - \hat{R}_{it})$, where $\ell(\cdot)$ is the loss function defined in (12). The results of these computations for selected values of c are presented in Table 2. The juxtaposition in Table 2 shows that, if the deviations from the nominal value of R which are less in magnitude than c enable classifying the output of the turning process as the highest quality product, the modal regression line, when used in the prediction mode, could be used to model the nominal value of the production process in a way guaranteeing the maximum share of products meeting the highest quality requirements in the total production.

Table 2.
Shares of residuals with absolute values less than c (in %)

c	Mean value regression	Median regression	Modal regression
0,25	17,6	25,6	33,6
0,5	26,4	45,6	52,0
0,75	41,6	62,4	67,2
1	56,8	72,0	78,4
1,25	65,6	80,8	80,8

Source: own computations.

The second discussed application of modal regression is inspired by the marketing mix concept. The dataset used in this empirical illustration comes from the Dominick's Finer Foods database maintained by the Kilts Center for Marketing at the University of Chicago's Booth School of Business, which part is available in the `bayesm` R package of Rossi (2019). From this package, we took the canned food dataset covering sales volumes of seven brands, jointly with the corresponding potential explanatory variables such as display activity (being a measure of the use of so-called display marketing), logarithms of retail prices, and logarithms of wholesale prices. The data have the form of (non-consecutive) weekly observations aggregated to the chain level. The number of available observation is 338. The estimation results presented here concern the brand no. seven (HH Chunk Lite 6.5 oz.). Figure 3 depicts the histogram of the sales volume of this brand, which shows that the examined distribution is strongly positively skewed. This is due to occasional bulk (wholesale) purchases.

Below we present the empirical marketing mix models for the brand analyzed, reduced to significant parameters. As in the previous example, three estimation methods were applied.

$$\hat{S}_1 = \hat{E}(S \mid \mathbf{X}) = \underset{(\pm 3{,}91)}{-28{,}667} - \underset{(\pm 5{,}562)}{46{,}090}\, P - \underset{(\pm 5{,}973)}{37{,}848}\, Pw, \tag{16}$$

$$\hat{S}_2 = \hat{Me}(S \mid \mathbf{X}) = -\underset{(\pm 0{,}859)}{23{,}921}\, P, \tag{17}$$

$$\hat{S}_3 = \hat{Mo}(S \mid \mathbf{X}) = \underset{(\pm 0{,}6034)}{2{,}9934} + \underset{(\pm 0{,}6316)}{3{,}5372}\, D - \underset{(\pm 2{,}2198)}{9{,}6421}\, P, \tag{18}$$

where S denotes the sales volume in thousands, D is a measure of the display activity, P is the logarithm of the retail price, Pw is the logarithm of the wholesale price, $\mathbf{X} = [D, P, Pw]$, and $\hat{S}_i$, i = 1, 2, 3, are the theoretical values from the regression models.

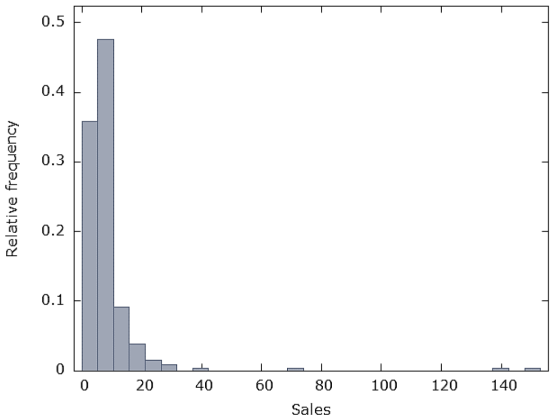

Figure 3. Histogram of the sales volume.

Source: own elaboration.

The estimation results show that the display activity variable is not significant in the equations for the conditional mean and median (16) – (17), whereas it is highly significant in the equation for the conditional mode (18), that is, it impacts on sales in a typical low demand week and can be used in explaining the largest part of the distribution of sales. By contrast, the wholesale price turns out to be significant exclusively in the equation for the mean, which says that this variable may be employed to describe the upper part of the distribution of sales.

The above analysis shows that modal regression makes it possible to build empirical marketing mix models without the need for market segmentation (in a sense, also saving degrees of freedom in estimation) since it naturally concentrates on the most typical sales volumes, reflecting the demand of the target consumer group.

In addition, the empirical modal regression model, such as that given in the equation (18), will make it possible to forecast the most likely volume of sales defined as that associated with the maximum of the conditional probability density function (for example Dimitriadis et al., 2024). Let us express such a forecast as $\hat{S} = a_0 + a_1 D - a_2 P$ for certain positive constants a_0, a_1, and a_2. Then, the corresponding modal forecast of the company's profit can be written as:

$$\begin{aligned} Profit(D,P) &= (P - c_1)\hat{S} - c_2 D - c_0 = \\ &= -a_2 P^2 + a_1 DP + (a_0 + a_2 c_1)P - (c_2 + a_2 c_1)D - a_0 c_1 - c_0, \end{aligned} \tag{19}$$

where c_0, c_1, and c_2 denote, respectively, the fixed cost of the marketing department, the unit production (or purchasing) and distribution cost of the goods, and the unit cost of the display activity. The function (19) has the following partial derivatives:

$$\left[\frac{\partial Profit}{\partial D}, \frac{\partial Profit}{\partial P}\right] = \left[a_1 P - c_2 - a_2 c_1, -2a_2 P + a_1 D + a_0 + a_2 c_1\right], \tag{20}$$

whereas its hessian is given as:

$$H(D,P) = \begin{bmatrix} 0 & a_1 \\ a_1 & -2a_2 \end{bmatrix}. \tag{21}$$

Since the hessian (21) is negative semidefinite, the function (19) is concave, and, thus, it attains its (unique) global maximum at the stationary point of the form:

$$P = \frac{c_2 + a_2 c_1}{a_1},$$
$$D = \frac{2a_2(c_2 + a_2 c_1) - a_1(a_0 + a_2 c_1)}{a_1^2}. \tag{22}$$

The solution (22) should be of help to the company in optimizing its marketing mix strategy in relation to the target consumer group.

6. Conclusions

The modal regression model, introduced by Kemp and Santos Silva (201) and Yao and Li (2014) and developed further, among others, by Kemp et al. (2020) and Ullah et al. (2022), is a statistical concept with a broad range of potential applications. However, to the best of our knowledge, empirical applications of this concept remain somewhat underdeveloped and are limited so far exclusively to certain macro- and microeconomic phenomena. In this paper, for the first time in the literature, we have drawn attention to some of possible uses of modal regression in enterprise management, and more precisely, in logistics and marketing management. As has been shown in the paper, the empirical modal regression model should help to develop decision instruments which can be utilized in demand forecasting (and logistics planning), managing the quality of industrial processes, and planning marketing mix strategies. This, however, certainly does not exhaust the topic. We believe that this direct statistical approach can be beneficial to managers in many areas of an enterprise and in many businesses. For this reason, it is recommended to include this tool in the managerial statistical toolbox.

Acknowledgements

The paper is an extended and revised version of an unpublished manuscript created with financial support from the National Science Center under grant no. 2017/27/B/HS4/01025.

References

1. Botev, Z.I., Grotowski, J.F., Kroese, D.P. (2010). Kernel density estimation via diffusion. *Annals of Statistics, Vol. 38, No. 5*, pp. 2916-2957, doi: 10.1214/10-AOS799
2. Bruzda, J. (2024). Does modal (auto)regression produce credible forecasts of macroeconomic indicators? *Wiadomości Statystyczne* [*The Polish Statistician*]*, Vol. 69, No. 10*, pp. 1-27, doi: 10.59139/ws.2024.10.1
3. Chen, Y.-C., Genovese, C.R., Tibshirani, R.J., Wasserman, L. (2016). Nonparametric modal regression. *Annals of Statistics, Vol. 44, No. 2*, pp. 489-514, doi: 10.1214/15-AOS1373
4. Dimitriadis, T., Patton, A.J., Schmidt, P.W. (2024). *Testing forecast rationality for measures of central tendency*. Retrieved from: https://arxiv.org/pdf/1910.12545, 23.01.2025.
5. Guetout, M.I. Effect of cutting parameters on roughness. Retrieved from: https://www.kaggle.com/datasets/mohamedikbalguetout/effect-of-cutting-parameters-on-roughness, 23.01.2025.
6. Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J., Stahel, W.A. (1986). *Robust statistics. The approach based on influence functions*. New York: Wiley.
7. Huber, P.J. (1973). Robust regression: Asymptotics, conjectures, and Monte Carlo. *Annals of Statistics, Vol. 1, No. 5*, pp. 799-821, doi: 10.1214/aos/1176342503.
8. Kemp, G.C.R., Parente, P.M.D.C., Santos Silva, J.M.C. (2020). Dynamic vector mode regression. *Journal of Business and Economic Statistics, Vol. 38, No. 3*, pp. 647-661, doi: 10.1080/07350015.2018.1562935
9. Kemp, G.C.R., Santos Silva, J.M.C. (2012). Regression towards the mode. *Journal of Econometrics, Vol. 170, No. 1*, pp. 92-101, doi: 10.1016/j.jeconom.2012.03.002
10. Koenker, R., Bassett, G. Jr. (1978). Regression quantiles. *Econometrica, Vol. 46, No. 1*, pp. 33-50, doi: 10.2307/1913643.
11. Maronna, R.A., Martin, R.D., Yohai, V.J. (2006). *Robust statistics. Theory and methods*. Chichester: Wiley.
12. Rossi, P. (2023). *bayesm: Bayesian inference for marketing/micro-econometrics. Version 3.1-6*. Retrieved from: https://cran.r-project.org/web/packages/bayesm/index.html, 23.01.2025.
13. Ullah, A., Wang, T. (2025). Modal volatility function. *Journal of Time Series Analysis (Early View)*. Retrieved from: https://onlinelibrary.wiley.com/doi/10.1111/jtsa.12790, 23.01.2025.
14. Ullah, A., Wang, T., Yao, W. (2021). Modal regression for fixed effects panel data. *Empirical Economics, Vol. 60, No. 1*, pp. 261-308, doi: 10.1007/s00181-020-01999-w

15. Ullah, A., Wang, T., Yao, W. (2022). Nonlinear modal regression for dependent data with application for predicting COVID-19. *Journal of the Royal Statistical Society. Series A: Statistics in Society, Vol. 185, No. 3*, pp. 1424-1453, doi: 10.1111/rssa.12849
16. Xiang, S., Yao, W. (2022). Nonparametric statistical learning based on modal regression. *Journal of Computational and Applied Mathematics, Vol. 409, Iss. August 2022, No. C, Article No. 114130*, doi: 10.1016/j.cam.2022.114130
17. Yang, J., Tian, G., Lu, F., Lu, X. (2020). Single-index modal regression via outer product gradients. *Computational Statistics and Data Analysis, Vol. 144, Iss. April 2020, Article No. 106867*, doi: https://doi.org/10.1016/j.csda.2019.106867
18. Yao, W., Lindsay, B.G., Li, R. (2012). Local modal regression. *Journal of Nonparametric Statistics, Vol. 24, No. 3*, pp. 647-663, doi: 10.1080/10485252.2012.678848
19. Yao, W., Li, L. (2014). A new regression model: modal linear regression. *Scandinavian Journal of Statistics, Vol. 41, No. 3*, pp. 656-671, doi: 10.1111/sjos.12054