NOVATEUR PUBLICATIONS
JournalNX- A Multidisciplinary Peer Reviewed Journal
ISSN No: 2581 - 4230
VOLUME 7, ISSUE 5, May. -2021

INFORMATION TECHNOLOGIES FOR FORECASTING CAR DEPOT INVENTORY USING CORRELATION ANALYSIS

Nurmukhamedov Tolaniddin Ramziddinovich
Doctor of Technical Sciences, Professor of the Department
"Information Systems and Technologies in Transport", Tashkent State Transport University,
e-mail ntolaniddin@mail.ru

Shukurov Farrukh Doniyorovich,
Assistant of the Department "Information Systems and Technologies in Transport",
Tashkent State Transport University,
e-mail farrux.ttymi13@gmail.com

Khakimov Shoyatbek Khalimjonovich,
Assistant of the Department "Information Systems and Technologies in Transport",
Tashkent State Transport University,
e-mail x.shoyatbek.x@mail.ru

ANNOTATION:

The analysis of the statistics of expenses of components and spare parts of mobile units of the car depot is carried out, the trend of changes in expenses is determined and the mathematical apparatus for forecasting the need for spare parts is selected, which provides the maximum accuracy of the forecast.

Forecasting of changes in the parameters of various systems is performed using information about the factors that affect the predicted parameter, using regression models. A multivariate regression model is proposed that takes into account the factors that affect the consumption of spare parts and establishes the density of the relationship between the factors themselves.

Keywords: car depo, analysis, correlation, model, regression, equipment, warehouse.

INTRODUCTION:

Over the past decades, Uzbekistan has paid great attention to the construction of new and reconstruction of existing railway lines. Currently, the movement of trains in the local direction is carried out on its territory, which was facilitated by the construction of railway lines connecting the center (Tashkent) with the Ferghana Valley (east), Khorezm (north), Surkhandarya (south). Cargo and passenger traffic increases from year to year. The increase in passenger traffic on railway transport is the result of positive competition for road and air transport: modern, comfortable passenger cars, the opportunity to observe the flat and mountainous landscape, and transport safety.

In the logistics chain, one of the key places is occupied by car depot warehouses. The warehouse is a part of the logistics chain, which sets the main requirements for warehouse processes, determines the goals and objectives of the warehousing system within the car depot. The main task of car depot warehouses is aimed at equalizing the rhythm of performing repair and equipment work and "releasing" the car into the transportation process, which makes it possible to carry out the continuity of production work and supply.

NOVATEUR PUBLICATIONS
JournalNX- A Multidisciplinary Peer Reviewed Journal
ISSN No: 2581 - 4230
VOLUME 7, ISSUE 5, May. -2021

The car depot of JSC "Uztemiryulyulovchi" carries out repair and equipment work with mobile units that carry out the movement of passengers by rail, both in their own republic and in the territories of the states of the near and far abroad. The acquisition of inventory of the warehouse with inventory items (goods and materials), spare parts, components (hereinafter referred to as goods)and their release to the employees of the car depot ultimately contributes to the equalization of the rhythm of passenger traffic. Determining the availability of the necessary goods in the warehouse, predicting their quantity, both in the near future and in the future with the use of appropriate computer software packages based on information technologies is a very urgent task.

THE MAIN PART:

The timeliness of performing repair and equipment works in a car depot requires solving the following problem: it is necessary to analyze the statistics of expenses of components and spare parts of mobile units of the car depot in order to show the trend of changes in costs and choose a mathematical tool for predicting the need for spare parts, which will ensure maximum accuracy of the forecast.

In order to predict changes in the parameters of various systems, information is used about the factors that affect this parameter, and regression models are often used. Taking into account the influence of various factors on the consumption of spare parts and establishing the density of relationships between the factors themselves requires the construction of a multivariate regression model.

The consumption of goods in this case will be the effective feature, and the other variables will be the factor features $x_1,\ldots, x_m$.

The regression equation for predicting the demand for goods is presented as follows:

$$y = a_0 + a_1x_1 + a_2x_2 + \ldots + a_mx_m, \qquad (1)$$

$x_1,\ldots, x_m$ – factorial signs.

This model will include only those factors, the quantitative accounting and forecasting of which changes can be taken into account in the warehouse of the car depot. The development of the effective feature of the multivariate regression model, let's start with the preliminary selection of the factor features of the model. The pair correlation coefficients (r_{yx1}, r_{yx2}, r_{y1x2}), are determined by the formula:

$$r_{yx1} = \frac{\overline{yx_1} + \overline{yx_2}}{\sigma_y \sigma_{x1}} \qquad (2)$$

σ_y and σ_{x1} – the root-mean-square errors of the corresponding samples.

The root-mean-square errors of the corresponding samples are determined using the following formulas:

$$\sigma_y = \sqrt{\frac{\sum(y-\bar{y})^2}{n}}; \qquad \sigma_x = \sqrt{\frac{\sum(x-\bar{x})^2}{n}} \qquad (3)$$

Factors of the pair correlation coefficient that have values below the specified level relative to the effective feature will not be taken into account in the model. In practical calculations of the need for goods, the given level of significance is assumed to be equal to 0.05.

The next stage of the study is to check for a strong correlation between factorial (multicollinear) features. In the model, we take into account only one attribute with the highest coefficient r_{yxn}. Of each pair of such features in the model, we take into account only the feature with the highest coefficient r_{yxn}. The selection of factors in the product demand forecasting model using the multicollinearity attribute of factors must be performed when the pair correlation coefficient is more than 0.8.

The parameters of the model $a_0, a_1,..,a_n$ of the regression equation are estimated using the least squares method, which is convenient to represent in matrix form. Let assume the following notation:

$$\alpha=(\alpha_j)\ j = 0,1, ..., m$$

m – Vector of unknown parameters;
a = (a_j) – vector of parameter estimates;

$$y = (y_i),\ i = l,..., n$$

n – Vector of values of the dependent variable;
X = (x_{ij}) – matrix of values of independent variables with dimension
n*(m+1);
ε=ε(i)– error vector in the model;
e = (e_i) – the error vector in the equation with the estimated parameters.
In the usual notation, a vector is understood as a column vector, that is, a matrix with dimension n*1.
The regression equation with the estimated parameters has the form:

$$\hat{y} = X * a + e. \quad (4)$$

Note that calculating the parameters of the regression equation is quite a time-consuming process. The development of hardware and software of modern computers using application software packages allows you to calculate the parameters of the regression equation quickly and automatically.
After calculating the partial correlation coefficients, we determine the multiple correlation coefficient ry, which characterizes the density of the relationship between the effective and factor features and, in general, is determined by the formula:

$$r_y = \sqrt{\frac{\sigma_{y1,2,...m}}{\sigma_y^2}} = \sqrt{1 - \frac{\sigma_{y(2,...m)}}{\sigma_y^2}}, \quad (5)$$

$\sigma_{y1,2,...m}$ – factor variance;
$\sigma_{y(2,...m)}$ – final variance;
σ_y^2– Variance of the resulting attribute.

$$\sigma^2_{y1,2...m} = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})}{n-1}; \qquad \sigma^2_{y(1,2...m)} = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-1};$$
$$\sigma_y^2 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1}, \quad (6)$$

$\hat{y}_i$ – calculated value of the effective attribute;
$\bar{y}$ – the average value of the effective attribute.
The accepted form of writing indexes means:
$\sigma^2_{y1,2,...m}$– the variance $\hat{y}_i$, obtained by taking into account the factors $x_1, x_2,... x_m$;
$\sigma^2_{y(1,2,...m)}$ – variance $\hat{y}_i$, obtained by excluding unknowns $x_1, x_2,... x_m$ from the system of equations (ellimination).

To check the quality of the regression model, it is necessary to evaluate the significance of the multiple correlation coefficients. This estimation is carried out using the Student's t-statistics [1] by checking for the equality of the k-th regression coefficient to zero (k=1, 2, ..., m). The calculated value of the t-criterion with the number of degrees of freedom (n – m - 1) is found by distributing the k-th regression coefficient by the standard deviation of this coefficient δ_{ry}. In this case:

$$\sigma_{ry} = \frac{\sqrt{1-r^2}}{\sqrt{n-m-1}};$$
$$t_{payment} = \frac{r_y}{\sigma_{ry}} = r_y * \sqrt{\frac{n-m-1}{1-r^2}}. \quad (7)$$

The calculated value of $t_{payment}$ is compared with the critical t_{kr}, which is taken from the table of values of t of the Student's criterion, taking into account the given level of significance and the number of degrees of freedom k. If the calculated value of t is greater than the critical value, then the correlation coefficient is considered significant and the relationship between the effective feature and the set of factor features is close.

We analyze a multiple regression model by evaluating its significance (the model) using the Fischer F-Test [2, 3]. At the same time, the hypothesis is put forward that the coefficients of the regression equation are equal to zero ($a_1 = a_2 = ... a_m = 0$) – the model is insignificant.

The actual value of the Fischer F-Test is determined by the formula:

$$F_{payment} = \frac{r_y^2}{1-r} * \frac{n-m-1}{m}, \quad (8)$$

payment

m- is the number of parameters of the regression equation.

The value of $F_{payment}$ is compared with F_{kr}, the value of which is determined by the table of the F -Criterion, taking into account the accepted level of significance and the number of degrees of freedom $k_1 = n - 1$ and $k_2 = m - 1$. If the calculated value of the criterion is greater than the critical one, then the competing hypothesis will be valid, that is, the multivariate model is significant.

In practice, not all programs provide a detailed analysis of the constructed regression model. When using standard application software packages for calculating the demand for goods and materials, components, and spare parts in a car depot warehouse, the packages that will evaluate the quality and significance of the regression model will prevail.

The determination coefficient $D = r*y^2$ is used to estimate the part of the change (variation) of the effective feature under the influence of factor features [2, 3].

To predict the need for spare parts using a multivariate regression econometric model for L steps ahead, you need to know the predicted values of all the factors included in the model. These values can be obtained based on extrapolation methods, for example, using the average absolute increments of factor features. The forecast values of the factors are substituted into the model and point-based forecast estimates of the need for goods of the required nomenclature are obtained. If, in the course of further research, it is revealed that the use of a linear multi-factor model is impractical, then the dependence of the need for spare parts on factor characteristics will be described by nonlinear equations. In this case, to estimate the regression parameters, it is necessary to reduce the regression equation to a linear form, which can be provided by logarithm [4].

In general, the estimation of nonlinear regression parameters is carried out using the nonlinear least squares method. To do this, the sum of the squared deviations of the calculated f(a1, a2) and the actual values of yi of the resulting attribute is minimized:

$$Q = \sum e_i^2 = \sum \left[y_i - f(a_1, a_2, \dots)\right]^2. \quad (9)$$

by differentiating Q by parameters a_j.

As a result, we get a system of normal equations that is liberalized by decomposing into a Taylor series, and then we use the linear least squares method.

In the conditions of workshops, sections of the car depot, when it is necessary to predict the need for goods for a short period, the most important are the latest indicators of the issue, release of spare parts. In this case, it is advisable to use adaptive forecasting models that take into account the uneven levels of the time series.

To predict the demand for goods, we use an adaptive forecasting model, which is based on the moving average scheme. According to the moving average scheme, the assessment of the current level is the weighted average of all previous levels, and the weights for observations decrease in the measure of distance from the next level, that is, the information value of the observation is considered the greater, the closer to the end of the observation interval.

The response to the forecast error in the model, which is based on this scheme, is determined using the smoothing parameters (adaptation), the value of which can vary from zero to one. A high value of these parameters (above 0.5) means giving more weight to the last levels of the series, and a lower value (less than 0.5) means giving more weight to the previous observation. As shown by the analysis of the series of values of the demand for goods and materials, components and spare parts of the car depot enterprises, the value of the smoothing parameters should be in the range from zero to 0.5.

For direct calculations, it is assumed to use the basic model according to the moving average scheme-the Brown model, which represents the development process as a linear trend with constantly non-constant parameters. To construct a linear adaptive Brown model, a time series of length N is needed. For example, when forecasting the demand for goods and materials, components and spare parts for the months of the year, the specified series will be made up of the values of the sale of spare parts in the previous months of the enterprise. The forecast of the demand value for k steps is carried out according to the formula:

$Y(t + k) = A_0 + A_1*k,$ (10)

A_0– the coefficient is a value that is close to the last value of the value of the need for spare parts, and is a natural component of this value;

A_1 - the coefficient determines the growth that was formed mainly before the end of the observation period (it shows at least the growth rate at earlier stages).

The initial values A_0 and A_1 of the model parameters are estimated using the least squares method for linear approximation using the first five points of the series: $Y_p(t,k)=A_0(t)+A_1*k$ (t=1,2,...,5). Using the obtained values of the parameters A_0 and A_i, we find a forecast for one step (k=1):

$Y_p(t, k)=A_0 (t)+A_1(t) *k.$ (11)

The estimated value of the need for spare parts is compared with its actual indicator and the amount of their discrepancy (error) is calculated.

For k = 1, we have: e (t + 1) = Y(t +1l) - Pym(t, 1).Accordingly, the model parameters are adjusted for this value. The modification is carried out as follows:

$A_0 (t) =A_0(t-1)+A_1(t-1)+\alpha^2e(t);$

$A_1(t)=A_1(t -1)+\alpha^2e(t),$ (12)

α – the smoothing parameter, the optimal value of which is found iteratively, that is, by repeatedly constructing the model for different values of α and selecting the best one;

e(t) – the error of predicting the level Y(t), calculated at a time one step ahead.

This method provides good results when there is a constant trend in the flow rate of the spare part, close to a linear relationship.

To implement these calculations using a personal computer, you need information about the costs of goods and materials, components and spare parts in the previous months of the company's operation (at least 12 months).

CONCLUSIONS:

1.A mathematical model for forecasting inventory values, spare parts, and components of the car depot of JSC "Uztemiryulyulovchi" is proposed.

2. Based on the mathematical model, you can calculate the forecast data of the availability of goods in the warehouse of the car depot. Determining the inventory of goods based on the method of correlation analysis will eventually allow you to evenly perform repair and equipment work with mobile units.

REFERENCES:

1) Kramer G. Mathematical methods of statistics, 2nd ed., trans. from English; ed. by A. N. Kolmogorov. - M.: Mir, 1975. -648 p.
2) Bondarenko P. S. Probability theory and mathematical statistics: a textbook / P. S. Bondarenko, G. V. Gorelova, I. A. Katsko; edited by I. A. Katsko, A. I. Trubilin. - M.: KNORUS, 2019. - 390 p.
3) Shikhalev A.M. Correlation analysis. Nonparametric methods / A.M. Shikhalev. - Kazan: Kazan. un-t, 2015. - 58 p.
4) Gmurman, V. E. Probability theory and mathematical statistics: textbook and practice for academic undergraduate studies / V. E. Gmurman. - 12th ed. - Moscow: Yurayt Publishing House, 2014. - 479 p.