МРНТИ 28.23.33, 06.39.21

G.N. Pachshenko

The International Information Technology University,
Almaty c., Kazakhstan

STATISTICAL METHODS FOR BUSINESS AND ENTERPRISE CONTROL

Abstract. Many statistics used in business for enterprise control and quality improvement applications are included in a common set of basic statistical data, including chi-square test, t-tests and methods based on a common linear model. However, there are also a number of methods developed for specific business needs for enterprise control and quality improvement applications. This article discusses such statistical methods as time series analysis and decision analysis. A comparative analysis and advantages of each method are given.
Keywords: enterprise control, business, statistical methods.

• • •

Аннотация. Многие статистические данные, используемые в бизнесе, для приложений управления предприятием и повышения качества, входят в общий набор основных статистических данных, включая критерий хи-квадрат, t-тесты и методы, основанные на общей линейной модели. Тем не менее, существует также ряд методов, разработанных для конкретных потребностей бизнеса, для управления предприятием и приложений для повышения качества. В данной статье рассматриваются такие статистические методы, как анализ временных рядов и анализ решений. Приводится сравнительных анализ и преимущества каждого метода.
Ключевые слова: управление предприятием, бизнес, статистические методы.

• • •

Түйіндеме. Бизнесте қолданылатын кептеген статистикалық мәліметтер кәсіпорынды басқару және сапаны арттыру қосымшалары үшін жалпы сызықтық модельге негізделген хи-квадрат, t-тесттер мен әдістерді қоса алғанда, непзп статистикалық деректердің жалпы жиынтығына кіреді. Дегенмен, бизнестің нақты қажеттіліктері мен кәсіпорынды басқару және сапаны арттыру үшін әзірленген бірқатар әдістер бар. Бұл мақалада уақытша қатарларды, шешімдерді талдау сияқты статистикалық әдістер қарастырылады. Салыстырмалы талдау және әр әдістің артықшылықтары келтіріледі.
Түйінді сөздер: кәсіпорынды басқару, бизнес, статистикалық әдістер.

Introduction. Nowadays many statistics used in business for enterprise control and quality improvement applications are included in a common set of basic statistical data, including chi-square test, t-tests and methods based on a common linear model. However, there are also a number of techniques developed for the specific needs of business and quality improvement applications. They are time series analysis, decision analysis and decision tree [1, 2]. The aim of investigation in this article is to make a comparative analysis of methods and show advantages of each method for enterprise control and business.

Research methods. Statistics is a branch of science that deals with data collection, organization, analysis, interpretation and presentation. This requires a proper design of the study, an appropriate selection of the study sample and choice of suitable statistical tests and methods. In this article the comparative analysis for some statistical methods is carried out.

Research results. Time series analysis. Time series analysis are used frequently in business statistics to chart the changes in some quantity over time. Strictly speaking, a time series is just a sequence of measurements of some quantity taken at different times, often at equally spaced intervals. The previous example of the number of automobiles manufactured in the years 1986–2005 would qualify, as would the measurements discussed later in this chapter in the section on control charts. Time series may be used for either descriptive or inferential purposes; the latter includes forecasting, i.e., predicting values for time periods that have not yet occurred. Time series analysis is a complex topic with many specialized techniques, and that this section can only introduce some of the terminology and a few simple examples.

One characteristic of time series data is that data points in sequence are assumed to not be independent, as would be required for standard General Linear Model and many other analytical techniques, but to be autocorrelated. This means that the value for a given time point is expected to be related to the points before and after it, and perhaps to points more distant in the series as well.

Additive models are often used to describe the components of a time series, i.e.:

$$Yt = Tt + Ct + St + Rt \quad (1)$$

The components of the trend *Yt* in this model are:

Tt – Secular or long-term trend, i.e., the overall trend over the time studied.

Ct – The cyclical effect, i.e., fluctuations about the secular trend due to business or economic conditions, such as periods of general economic recession or expansion.

St – The seasonal effect, i.e., fluctuations due to time of year, for instance the summer versus the winter months.

Rt – The residual or error effect, i.e., what remains after the secular, cyclical and seasonal effect have been accounted for; it may include both random effects and effects due to rare events such as hurricanes or epidemics.

Much of time series analysis is devoted to resolving the variance observed over time into these components. The concept is similar to partitioning the variance in ANOVA models, although the mathematics involved is different.

Exact measurements plotted over time, also known as raw time series, will almost always show a great deal of minor variation that may obscure major trends that could help explain the pattern and make accurate future forecasts. Various types of smoothing have been devised to deal with this problem. They can be divided into two types: moving average or rolling average techniques, which involve taking some kind of average over a series of consecutive points and substituting this average for the raw values, and exponential techniques, in which an exponential series is used to weight the data points.

Example. To calculate a simple moving average (SMA), take the unweighted mean of a specified number of data points (n) prior to the time point in question. The size of n is sometimes described as a window because the idea is that a window including n data points (a window of width n) is used to calculate the moving average. As you progress forward in time through the data, the window moves so you can see different data points each time, and the average is calculated using the points included in the window for each time point. For instance, a five-point SMA would be the average of a given value and the previous four data points.

The SMA for each new data point drops only one value and adds only one new value, reducing the fluctuation from point to point. This attribute gave rise to the term rolling average, because the last value “rolls off’ the series as the new value “rolls on.” This is similar to the methodology used to compute player standings on professional tennis tours, although in that case a total rather than an average is computed. Each player’s total points in a given week is the sum of their points from the previous 52 weeks, and each week the total is recalculated as the oldest week’s points are dropped and the newest week’s points added in.

The greater the size of the window used to calculate an SMA, the greater the smoothing since each new data point has less influence relative to the total. At some point the data may become so “smoothed” that important information about the pattern is lost. In addition, the larger the window, the more data points that have to be discarded (because you need more points to calculate each average).

Decision analysis. Decision analysis is a body of professional practices, methodologies, and theories used to systematize the decision-making process in the service of improving the process of decision-making. There are many schools of thought within decision theory, and each may be useful in a particular context: this section concentrates on several of the most common decision analysis methods, which will help to introduce the student to the types of processes involved, as well as providing concrete assistance in particular decision-making contexts.

The decision-making process will be described in terms of financial costs and payoffs, but can be used with other metrics as well, for instance, personal satisfaction or improved quality of life, as long as they can be quantified.

In decision analysis, the process of making a decision is usually conceived of as a series of steps that is not unlike the process involved in hypothesis testing. They are also not that different, except for the selection and application of a mathematical model in steps 5–6, from the ordinary type of decision-making process we engage in every day. Besides the potential to lead to better decisions, going through these steps and justifying and documenting them, should make the

reasons for a particular decision easier to explain and justify to someone who wasn’t involved in the process. The basic steps are:

1. Define the situation or context, including states of nature (any situation in the real world that may influence the outcomes).States of nature must be stated as mutually exclusive and exhaustive alternatives, for instance, strong/medium/weak market, or low rainfall/adequate rainfall.

2. Identify the choices at hand, i.e., the alternative decisions that could be made; these are known as actions.

3. Identify the possible outcomes or consequences.

4. Assign costs and profits associated with all possible combinations of choices and outcomes.

5. Select an appropriate mathematical model.

6. Apply the model using the information from steps 2–4.

7. Make a decision based on the best expected outcome as predicted by the model.

Choice of a decision theory methodology depends in part on how much is known about a situation. There are three types of contexts in which one may apply decision theory:

- Decision-making under certainty
- Decision-making under uncertainty
- Decision-making under risk

Decision-making under certainty means that the future state of nature is known, so the decision-making process requires only stating the alternatives and payoffs in order to be able to pick the choices that will invariably lead to the best outcome. This situation will not be further discussed because no mathematical modeling is required and because there is no uncertainty about what is the best choice.

Decision-making under uncertainty is a more common situation: we do not know the probabilities of each state of nature and must make our decision based only on the gains or losses from different actions under each state. For instance, if we are choosing from several cities in which to open a restaurant, the success of the restaurant depends in part on the economic climate in each city when the restaurant opens, but we may not have good estimates of the future economic climates in the future in these cities. Similarly, when choosing what crop or variety to plant, our success at harvest time depends partly on the amount of rainfall during the growing season, but we may feel we do not have sufficient information to estimate this in advance.

In decision-making under risk, we know the probabilities of each outcome (or have reasonable estimates of them) and can combine this information with that about expected payoffs to determine which decision is optimal.

If the probability of various outcomes, given particular actions, is known, then a decision tree can be constructed that displays the actions and payoffs under different states of nature and can be used to clarify the outcomes of different combinations. The purpose of a decision tree is to display decision-making information, including available actions, states of nature, and expected payoffs in a clear and graphical manner. It does not include any rules for making decisions but can aid decision-making by presenting the relevant information in one graphical summary.

Discussion of results. Time-series data allow to make operational decisions in near real-time, rather than having to wait hours to make decisions based on traditional batch systems. And time-series data allow to look at processes historically, as well as build predictive models for the future. Researcher can prevent problems before they occur, and seize opportunities as they emerge. In the industrial world, in quality improvement in business and for enterprise control a time-series database tends to become a unifying platform, tying together many disparate systems: sensors, control systems, business systems, legacy systems - into a democratizing platform for collaboration, analysis, and decision making [3].

One of the important problem in many industries is the predicting the occurrence of emergency situations in a multidimensional environment of time series. For example, in refineries, hundreds of variable process sensors are installed at different sites of the processing unit. This can lead to a decrease in the quality of the product, or an accident. However, in the problems of forecasting multidimensional time series there are many variables, and the analysis of their interaction is an extremely complex process. Studies in this area indicate that there are no clearly defined regular components in the data, and therefore there is a need to create a predictive model directly from time series. For this purpose, decision trees in time series forecasting problems are used. For enterprise control it is necessary to make decisions about investments, products, resources, suppliers, financing, and many other items. Decision making may be the most important function of management.

Conclusion. It is very important for the researcher to know various statistical methods for statistical data processing and to distinguish the advantages and disadvantages of all these methods. If you do not take into account this fact, the results of studies can lead to inaccurate results and even serious errors in quality improvement in business and for enterprise control. This article provides examples of using time series analysis and decision analysis.

References

1 Boslaugh S.. Statistics in a Nutshell. Publisher: O'Reilly Media, 2012, - 594 p.

2 James T. McClave, P. George Benson. Statistics for Business and Economics. Pearson, 2017, - 714 p.

3 From a Time-Series Database to a Key Operational Technology for the Enterprise. [Electronic resource] / Access mode: - https://blog.colinbreck.com, free, - Screen title.

Pachshenko G.N., Candidate of Engineering Science, Docent, associate Professor