

Selection of Appropriate Statistical Methods for Research Analysis

CAPT. DEEPAK KUMAR ROUSTRAY

Designation-Head, Department of Statistics, Ispat Autonomous College Rourkela-3

ABSTRACT

The aim of the paper is to provide algorithms which allow to choosing methods of statistical data processing and development of a model for acquiring results about statistical methods and mastering skills of competent knowledge application in various research activities in various fields. The method of modelling is a suitable approach to the study of this problem. It focuses on various methodologies for the selection of appropriate statistical model for the efficient processing of the Mathematical research results. The paper focuses an algorithm that allows choosing an appropriate method of statistical data processing: general algorithm of statistical methods application in research, statistical problems systematization based on which there have been outlined conditions for specific research methods and its application. To make a final decision concerning the statistical method at the stage of data received and statistical tasks of the research defined.

Keywords: *statistical processing of the research results, statistical methods, research, statistical criteria, algorithm.*

I.INTRODUCTION

The statistical data analysis is a burning demand for the researcher. The statistical methods are one of the major, generic methods of modern science, which are applied in various subject areas. A large numbers of statistical data processing methods cause problems of adequate comparison, correlation and analysis of different research results. Inappropriate choice of a method for the data analysis can lead to dangerous conclusions, incorrect interpretation of the research results, lead to the loss of the scientific value of such research results and the loss of informativity.

If we take example of choosing the most effective statistical method, implies that defining the objective and characteristics of each method, a list of requirements to information and statistics. In this context, it is important not only to select suitable statistical methods, but to improve the skills of applying this knowledge in various research analyses.

Statistical methods are some of the methods of the applied mathematical statistics used for the processing of the experimental results (Vocational Education, 1999). In this modern day, all kinds of statistical methods are used in various academic fields, depending on the experimental data and the tasks that the researcher has to solve.

In modern demography statistical methods are used in different areas: to obtain information on population and demographic processes, reconstruction using incomplete data set and then to process data and provide statistical

definition of the demographic processes; to analyze the demographic models and socio-demographic relations; to accumulate the characteristics of the demographic processes and calculate some aggregates of reproduction and population movement.

In demography statistical methods are mostly applied in the study of demographic processes versus specific socio-economic factors and for this correlation and regression analyses are used (for example, correlation between fertility or nuptiality and living conditions, etc.). To put it differently, we study the correlation between the characteristics of individuals or families (households), groups of population or subpopulations.

Objective of the Study:

- a) Development and study of general purpose methods without considering specificity of applications.
- b) Development and analysis of suitable statistical methods and models of real phenomena and processes in accordance with the needs of a particular area of research.
- c) The selection and application of statistical methods and models in the statistical analysis of specific data.
- d) Study the general algorithm for statistical methods application in scientific research.

II. REVIEW OF LITERATURE

G. Nikolayev and M. P. Degtyareva (2013) solved the problem of unambiguous identification of literary texts based on the plot study with the help of the systemic analysis of the text object involving the use of statistical methods for identifying texts subjects, methods of systemic analysis, graph theory, functional analysis.

Y.D. Bluvshstejn (1981) widely used Statistical methods in criminology, namely in the criminological statistics and legal statistics, here statistical methods allow a comprehensive qualitative analysis of the legal quantitative phenomena: 1) to give a numerical rating of the condition, level, structure and dynamics of crime and law enforcement combating it, that is to answer the questions about a current situation (descriptive function); 2) to identify statistical relationships, regularities in condition, structure and dynamics of crime, as well as in law enforcement, that is to explore to a certain extent the causes of a particular situation (explanatory function); 3) to identify trends in the development of crime, to make statistical criminological forecast, that is to envisage at least approximately what is expected, what are the prospects (predictive function); 4) to identify the "worrying" signs in the characterization of crime, positive features and shortcomings in the work of law enforcement bodies, "bottlenecks", vulnerabilities (low level of crime detection, lengthy periods and low quality of the investigation and court proceedings etc.) (Organizational, administrative function).

A. N. Leontiev (1959) in his study give the statement that the need - the purpose - the conditions and correlating with them activities - actions - operations are the principal elements of the activity. Any activity is carried out involving various methods, so the statistical scientific method comprises several techniques. According to him Statistical research techniques include the steps of collecting, processing and presenting research results.

III.METHODOLOGICAL FRAMEWORK

Statistical methods are widely used in various scientific fields like economics psychology, in biology and medicine, in history, in criminology, in Cultural Studies etc. The major types of statistical methods are general-purpose methods, methods applied in accordance with the needs of a particular area of activity, the methods of statistical analysis of specific data. Applicable scope of specific statistical methods is much less than of general-purpose methods, but its importance in analyzing a particular situation is much greater. Scientific results, the significance of which is estimated in accordance with general scientific criteria, correspond to the general-purpose works, as for the works focused on the analysis of specific data it is essential to ensure successful solution of specific problems in a particular area of application (economics, sociology, medicine, history, criminology, etc.). Meanwhile, regardless of the application sphere, it is necessary to correctly apply statistical methods while implementing scientific research, thus guaranteeing scientifically valid and reliable results of data processing.

Statistics distinguish the most commonly applied statistical methods among the existing ones: descriptive statistics; design of experiments; sampling; hypothesis testing; regression, correlation and factor analysis; time series analysis; statistically specified tolerances; analysis of the measurements accuracy; statistical process control; Statistical control of processes; reliability analysis; analysis of the causes of nonconformities; process capability analysis.

The application of appropriate statistical method in economics plays an important role, which deals with the processing and analysis of large amounts of information on socio-economic variable, economic studies analysis the problem of identifying the factors that determine the level and dynamics of the economic process. Generally it is economic statistics which studies the quantitative characteristics of socio economic variable and processes in the economy by means of analysis and statistical data interpretation. Its main methods are descriptive, analytical and comparison methods.

In psychology various applications of statistical methods are as follows 1) descriptive statistics, which includes the grouping of data, tabulation of data, graphical representation of data and a quantitative description of the statistical data 2) the theory of statistical inference used in psychological research to estimate the results of the samples survey (inductive statistics) 3) the experimental design theory serves to find and verify the formal relationships between variables (analytical statistics).

In biological and medical research Statistical methods are significantly and widely used .There are many research areas in biology where the application of statistical methods are more; which comprises biometrics, biostatistics; in case of medical science statistical methods are used for the sample collection, analysis of sample experimental data and clinical observations, biomedical statistics ,health statistics. In ecology they also apply statistical methods and models – methods of variation statistics allowing to explore the whole for example phytocenosis, population, productivity in its particular population using data from survey sites and to assess the degree of the results accuracy.



In history by using the statistical data methods one can trace the dynamics of the society development, changes in its population, social background, political opinion, economic conditions etc. For example, the area of agro-historical research is the widest field of factor analysis application (Litvak, 1985). Cliometrics that appeared in the late 1950s and has been developing ever since is an area in the historical studies, suggesting the systematic use of statistical and mathematical models. In addition, statistical methods and models have been successfully used in archeology to decipher the inscriptions in ancient languages and culture.

Statistical methods are most widely used in criminology, namely in the criminological statistics and legal statistics: criminal and administrative legal statistics, where statistical methods allow a comprehensive qualitative analysis of the legal quantitative phenomena to give a numerical rating of the condition, level, structure and dynamics of crime and law enforcement combating it, that is to answer the questions about a current situation, to identify statistical relationships, regularities in condition, structure and dynamics of crime, as well as in law enforcement, that is to explore to a certain extent the causes of a particular situation, to identify trends in the development of crime, to make statistical criminological forecast, that is to envisage at least approximately what is expected, what are the prospects, to identify the "worrying" signs in the characterization of crime, positive features and shortcomings in the work of law enforcement bodies etc.

A model of acquiring techniques of statistical methods and skills of competent knowledge application in the above variety of scientific research areas is proposed for consideration. That model, in general is the system. The system represents an integrity composed of individual elements and correlation between them, which includes the components like motivational, content-related, procedural and evaluative. The model also incorporates suitable procedures for the appropriate selection of statistical methods for the accurate analysis of the research results (Ganieva et al., 2014; Zaripova et al., 2014; Masalimova & Nigmatov, 2015). Here, the aspiration to choose and prepare for the scientific and professional activities can serve as the main objective of conscious learning associated with awareness of its objectives. It is advisable to use the following approaches in order to teach statistical methods and develop their ability to make an appropriate choice of the model 1. Methodological, having an effect on goals and learning process. 2. Systemic, which affects both the content and the process of learning. 3. Activity-algorithmic approach influencing the processual aspect of learning. 4. Process-oriented approach affects the learning process, primarily carrying out experiments and statistical analysis.

The methodological approach basically represents a scientific process of knowing the method, peculiarities of which are exemplified by the historical-scientific material. This approach defines the purpose of learning: introduction to the scientific process of knowing the method, acquirement of certain research techniques. Experiment and scientific research are used in training statistical methods in accordance with this approach. It basically says that the methodological approach also affects the learning process.

Activity-algorithmic approach has much contribution to the development of statistical methods teaching process. From the perspective of the activity approach the objectives of training statistical methods are formulated with the help of objective, tasks, activities and methods, by taking the objective, when the task of a



situation is taken in which you need to reach a certain goal, the activities are the process of achieving the goal, and the appropriate method is the way to implement activities.

According to the theory of A. N. Leontiev (1959), the need - the purpose - the conditions and correlating with them activities - actions - operations are the principal elements of the activity. Any activity is carried out involving various methods or ways, so the statistical scientific method comprises several techniques. Statistical research techniques include the steps of collecting, processing and presenting research results.

Techniques for statistical data processing are mostly tied to the use of algorithms. The application of the algorithms in the learning and analysis process was studied by B. V. Biryukov (1974), L. Lund (1966), N. Rosenberg (1979), and others. An algorithm is an incremental description of mechanically step by step performed uniform and relying on a finite set of rules procedure for solving the problem. In training statistical research methods algorithms are used in the form of regulations to address the educational tasks with a provision of operational procedure. Each algorithm is nothing but a model by following which the student obtained his knowledge of a particular studied portion of educational material and thereby labels it as learned.

An algorithmic approach is implemented through examining the order of evaluation of statistical indicators using different statistical formula. Algorithms description is possible through both inductive and deductive process. In the first case, students study a formula, divide it into constituent parts that are analysis, and then combine the actions that are synthesis. In the second case, the formulae are derived from the objective set, they define the steps to solve it and analysis of the problem by choosing appropriate statistical model, and then derive the formula for analysis.

The training of choosing statistical methods is carried as follows

1. Setting objectives of certain skills formation and the adoption of these objectives the teacher explains the purpose of the method, its capabilities, process of conversation.
2. Instruction on the subject and methods of research aimed at mastering the training skills in which teacher explains the appropriate procedure to choose statistical methods, to choose formulas - the order of calculations following the selected formula – interpretation and analysis of the results.
3. Practical problem to simulate the activities – students solve common tasks.
4. In connection with Monitoring of skills formation progress, teacher checks the process of the studied method using Q&A technique, tests, etc. The application of the acquired skills in a variety of practical situations students solve statistical problems that have different questions wording, conduct statistical experiments and interpretation and design the experiment by choosing appropriate statistical techniques.
4. The process of consolidating the skills and independent application of the skills - students perform statistical research working individually from the initial stage of project development to presenting the results.

The content of statistical methods training is developed on the basis of a systematic approach, systematic approach is an area in the methodology of scientific cognition and social practice, which is based on approaching the objects as the systems, focusing research on disclosing the integrity of the object, on identifying the diverse types of bonds within it and consolidating them into a common theoretical picture.

According to I. Y. Lerner's concept (Lerner, 1981), content of education is pedagogically oriented system of knowledge, work methods, experience of creativity and emotional and sensory education, assimilation of which facilitates personality formation. By systemic mastering of statistical methods we understand gradual familiarization of students with the function, content, objective and application of methods, systematic exercises of statistical methods application, conscious use of methods as the training techniques in the educational activities and analysis.

Current statistical method is implemented on the basis of the previous one. For example, the study of the correlation is not possible without mastering the methods of studying the variations of the characteristic, and the study of variations is impossible without a study of averages, and so on. Here we can outline the following principles of designing the content of training skills to make appropriate choice of statistical methods of the research: 1. the principle of the content compliance with the level of modern statistical methods and one's own research.

2. The principle of consideration the integrity between the content and processual aspects of education.

3. Structural consistency of the education content at different levels of its formation from general to special.

Experience of the statistical methods application is acquired gradually by familiarization with the individual data of statistical methods, the formation of a systemic knowledge about appropriate statistical methods, and conscious application of the methods in different situations. Repeated application of statistical methods in different situations that is standard, modified and new allows us to approach a statistical method from different perspectives, apply it to different objects. This gives a scholar an opportunity to choose essential features of the statistical method, and thus, apply it in different situations. Experience of emotionally-valuable relation to statistical work includes attitudes, beliefs, and values. This experience includes the motives of activity, moral problems, which are reflected in the scholar's behavior. In the process of statistical generalization and analysis the students develop their own attitude to the phenomenon, understand the operations made. Otherwise, they disengage from the educational process; students make operations mechanically, which significantly reduces the effectiveness of the training. Further, we will dwell on the method of choosing appropriate statistical methods necessary for the successful development and implementation of the scientific research and to achieve the goal.

IV.FINDINGS

The above paper focused and concludes that the statistical methods of data analysis are used in virtually all areas of human activity. They are used whenever necessary to obtain and justify any judgments about the group of objects or subjects with some internal heterogeneity.

It is expedient to distinguish three kinds of scientific and applied activities in the field of statistical methods of data analysis those are:

- 1) Development and study of general purpose methods without considering specificity of applications;
- 2) Development and analysis of appropriate statistical methods and models of real phenomena and processes in accordance with the objective of a particular area of activity.

3) The application of statistical methods and models in the statistical analysis of specific data.

Based upon the analysis of the above we established the following relationship between the statistical problems tackled in the research, conditions and statistical methods applied to research result which are focused in the below table.

Relationship between statistical tasks addressed in research, conditions and statistical methods applied to them

Tasks	Conditions		Methods
a) identifying differences in the level of the studied statistical feature	Two sampling populations	feature is distributed normally	Tests: t, F
		feature distribution is different from normal	Tests: McNemar, m, Q, U, χ^2 , ϕ^*
	Three and more sampling populations		Tests: χ^2 , S, H
b) identification of the significance and direction of the shift in the level of the studied feature	two measurements on one and the same sample population	feature is distributed normally	Tests: t, F
		feature distribution is different from normal	Tests: T, G, ϕ^*
	and more measurement (treatment methods etc.	feature is distributed normally	repeated measures analysis of variance
	one and the same sample population	feature distribution is different from normal	Tests: χ^2_r , L
c) identification of differences in the distribution of the statistical feature	when comparing the empirical distribution with the theoretical one		Tests: χ^2 , λ
	when comparing two empirical		Tests: χ^2 , λ , ϕ^*

	distributions		
d) identification of the coherence between changes of statistical features	two features		correlation analysis (r _{xy} , τ, r _s , R, φ), paired regression analysis
	three and more features		correlation analysis (r _{xy} , r _s , multiple and partial correlation), multiple regression analysis, factor analysis and cluster analysis
e) identification of a statistical characteristic changes under the influence of controlled conditions (factors)	under the influence of one factor		Критерии (S, L, H), однофакторный дисперсионный анализ, множественное сравнение независимых выборок
	under the influence of two factors simultaneously		two-factor variance analysis
	under the influence of three and more factors (search for hidden causes)		factor analysis
f) distribution of the studies objects into a relatively homogeneous groups		Groups are predefined	discriminant analysis
	Groups are not predefined		cluster analysis

g) analysis of survival data (comparative analysis of efficiency	Two sampling populations, one feature		Tests: Gehan's, Logrank test
--	---------------------------------------	--	------------------------------

V.CONCLUDING REMARKS

The above mentioned algorithm in the table allows to incrementally choosing the appropriate type of statistical analysis, which is to be solved along with conducting the research. Then, on the basis of the choice made there should be defined the conditions of applying certain research methodology. With data obtained and statistical challenges of the study identified a final decision on the selection of appropriate statistical methods, which is made in accordance with the algorithm of applying statistical methods to specific scientific fields, and enabling to quickly and correctly choose appropriate statistical method of statistical data processing. The algorithm suggested in the paper combines the advantages of a separate, previously known algorithms and techniques for selecting appropriate statistical methods to address the specific statistical tasks.

Thus, the methods of the research results processing and guidelines for their implementation have their own specific features, which, above all, are defined by a specific kind of research, its methodology and ways of research results analysis and presentation. Meanwhile, statistical methods can adequately use the available data for decision-making and thus contribute to positive change of some processes in natural, physical, psychological, sociological, economic, etc.

VI.RECOMMENDATIONS

The above paper is focused on choosing and studying different types of statistical methods of experimental data processing, understanding and its application opportunities. Methods of research results processing and guidelines for their implementation for every scientific field related to the experiment have their own specific features, which are predetermined by a certain type of research analysis, its methodology and ways of its results presentation. Although the experiment description for different research areas is rather specific, formal differences hide general methodology and common principles of building results processing systems.

It should be emphasized that the submitted work will help to purposefully and effectively apply the knowledge accumulated by the scholar, teachers, and all those whose professional and scientific activities are related to the analysis of experimental data. The paper is also intended for undergraduates, graduate, research scholar and scientist.

REFERENCES

1. Cochran, W. (1976). *Sampling techniques*, Moscow: Statistika
2. Ermolaev, O. J. (2006). *Mathematical statistics for psychologists: the textbook*. Moscow: Flint.

3. Glass, J. & Stanly, J. (1976). “*Statistical methods in pedagogy and psychology*” Moscow: Progress.
4. Gmurman, V. E. (2003). The theory of probability and mathematical statistics: *a manual for schools. Moscow: Higher School.*
5. Progress Report of the Arbitration Court of the Kirov region. (2014). Reference Form № 1. from <http://kirov.arbitr.ru>
6. Masalimova, A. R. & Nigmatov, Z. G. (2015). “Structural-Functional Model for Corporate Training of Specialists in Carrying Out Mentoring” *Review of European Studies*, 7(4), 39-48.
7. Orlov, A. I. (2001). Development of Methodology of Statistical Methods. Statistical methods of assessment and hypothesis testing. Interuniversity collection of scientific papers. Perm, *Perm: Publishing house of the PSU.*
8. Novikov, D. A. (2004). Statistical methods in educational research (typically). *Moscow: MZ-Press*
9. www.google.com
10. Hollender, M. & Wolfe, D. (1983). Non parametric Statistical Methods. *Moscow: Finance and Statistics.*
11. Vocational Education, (1999)
12. iSER,, (2016) *Mathematics Education*, 11(1), 303-315
13. Nikolaev, A. G. & Degtyarev, M. P. (2013). *Identification of text files by statistical methods* (conventional cases). *Radioelektronni kopm'yuterni i sistemi*, 4 (63) 55-59.