

# 統計学

第 104 号

---

## 論文

教育用擬似マイクロデータの作成

—平成16年全国消費実態調査を例として—

..... 山口 幸三・伊藤 伸介・秋山 裕美 (1)

Analysis of IO-based Annual Supply and Use Tables for the Development of QNA

..... Takeshi SAKURAMOTO (16)

## 研究ノート

Introduction of the Theory of Correlation into Russia and E. Slutsky

..... IRINA ELISEEVA (41)

## 本会記事

支部だより..... (52)

経済統計学会内規・編集委員会規程・投稿規程・執筆要綱・投稿原稿査読要領..... (57)

---

2013年3月

経済統計学会

## 創刊のこ と ば

社会科学の研究と社会的実践における統計の役割が大きくなるにしたがって、統計にかんする問題は一段と複雑になってきた。ところが統計学の現状は、その解決にかならずしも十分であるとはいえない。われわれは統計理論を社会科学の基礎のうえにおくことによって、この課題にこたえることができると考える。このためには、われわれの研究に社会諸科学の成果をとりいれ、さらに統計の実際と密接に結びつけることが必要であろう。

このような考えから、われわれは、一昨年来経済統計研究会をつくり、共同研究を進めてきた。そしてこれを一層発展させるために本誌を発刊する。

本誌は、会員の研究成果とともに、研究に必要な内外統計関係の資料を収めるが同時に会員の討論と研究の場である。われわれは、統計関係者および広く社会科学研究者の理解と協力をえて、本誌をさらによりよいものとするを望むものである。

1955年4月

## 経 済 統 計 研 究 会

## 経 済 統 計 学 会 会 則

第1条 本会は経済統計学会（JSES : Japan Society of Economic Statistics）という。

第2条 本会の目的は次のとおりである。

1. 社会科学に基礎をおいた統計理論の研究
2. 統計の批判的研究
3. すべての国々の統計学界との交流
4. 共同研究体制の確立

第3条 本会は第2条に掲げる目的を達成するために次の事業を行う。

1. 研究会の開催
2. 機関誌『統計学』の発刊
3. 講習会の開催、講師の派遣、パンフレットの発行等、統計知識の普及に関する事業
4. 学会賞の授与
5. その他本会の目的を達成するために必要な事業

第4条 本会は第2条に掲げる目的に賛成した以下の会員をもって構成する。

- (1) 正会員
- (2) 院生会員
- (3) 団体会員
- 2 入会に際しては正会員2名の紹介を必要とし、理事会の承認を得なければならない。
- 3 会員は別に定める会費を納入しなければならない。

第5条 本会の会員は機関誌『統計学』等の配布を受け、本会が開催する研究大会等の学術会合に参加することができる。

- 2 前項にかかわらず、別に定める会員資格停止者については、それを適用しない。

第6条 本会に、理事若干名をおく。

- 2 理事から組織される理事会は、本会の運営にかかわる事項を審議・決定する。
- 3 全国会計を担当する全国会計担当理事1名をおく。
- 4 渉外を担当する渉外担当理事1名をおく。

第7条 本会に、本会を代表する会長1名をおく。

- 2 本会に、常任理事若干名をおく。
- 3 本会に、常任理事を代表する常任理事長を1名おく。
- 4 本会に、全国会計監査1名をおく。

第8条 本会に次の委員会をおく。各委員会に関する規程は別に定める。

1. 編集委員会
2. 全国プログラム委員会
3. 学会賞選考委員会
4. ホームページ管理運営委員会
5. 選挙管理委員会

第9条 本会は毎年研究大会および会員総会を開く。

第10条 本会の運営にかかわる重要事項の決定は、会員総会の承認を得なければならない。

第11条 本会の会計年度の起算日は、毎年4月1日とする。

- 2 機関誌の発行等に関する全国会計については、理事会が、全国会計監査の監査を受けて会員総会に報告し、その承認を受ける。

第12条 本会会則の改正、変更および財産の処分は、理事会の審議を経て会員総会の承認を受けなければならない。

付 則 1. 本会は、北海道、東北、関東、関西、九州に支部をおく。

2. 本会に研究部会を設置することができる。
3. 本会の事務所を東京都町田市相原4342法政大学日本統計研究所におく。

1953年10月9日（2010年9月16日一部改正[最新]）

# Introduction of the Theory of Correlation into Russia and E. Slutsky

IRINA ELISEEVA\*

## Summary

The end of 19th – beginning of 20th century was marked by a burst of development in statistical methodology. As a result of the appearance of the Russian-language exposition of Pearsonian ideas by E. Slutsky in 1912, Russian statisticians had separated into two groups – those who supported the application of the correlation theory in social researches and those who rejected it. This split of the Russian statistical community had great consequences for the evaluation of statistics in education and research in Russia. This paper mainly considers Slutsky's contribution to the distribution curves and theory of correlation and also his position in these development.

## Key Words

E. Slutsky, K. Pearson, English biometric school, History of Russian statistics, Theory of correlation

## Introduction

The aim of this paper is to consider the preliminary attempts of overcoming the traditions of the German statistical school and E. Slutsky's role in introducing the ideas and methods of the biometrical school of K. Pearson at the beginning of the 20th century in Russia. In this paper, Slutsky's methodology will be looked at through an analysis of his book, *Theory of correlation and elements of distribution curves study (handbook for studying some of the most important elements in modern statistics (1912))*<sup>1</sup>. Under the influence of this book, Russian statisticians had separated into two groups – those who supported the application of the correlation theory and those who rejected it. This split of the Russian statistics community had great consequences

for the evaluation of statistics in education and research in Russia.

## 1. Appearance of correlation theory in Russia

The end of 19th – beginning of the 20th century was marked by a burst of development in statistical methodology. The biometric school of F. Galton and K. Pearson brought into the statistical community of the definition of regression methods and correlation measurement, created the study of distribution curve, suggested the  $\chi^2$ -test for goodness of fit to check the hypothesis of statistical law, and discovered nonparametric techniques. Above all due to such journals as *Biometrika* and at some point to *Proceeding of the Royal Statistical Society*, all of these achievements became available for specialists and could not be left out of consideration by Russian statisticians.

---

\* Dr., Professor, member-in-correspondence of Russian Academy of Sciences

Let us note that at that time despite of Russian professors' loyalty to the German canons of constitutional law and recognition of political economy as only statistical methodological basis, the interest in statistical methods and theory of probability as the basis for statistics started to emerge in universities all around Russia. This interest appeared first among mathematicians who were prone to empirical research such as V. Bunyakovsky (1804-1889), professor at Petersburg University, and also A. Vasiliev (1853-1929), the professor at Kazan University. Then also interest began to emerge among "traditional" professors of statistics such as professor at Moscow University A.I. Tchuprov (1842-1908), who appreciated the value of mathematical education for the research of economic and social phenomenon, and belonged undoubtedly to this second group of "traditional" professors of statistics. His son A.A. Tchuprov (1874-1926) studied first at Moscow University in the mathematics department under P. Nekrasov (1853-1924) and then continued his education at Strasbourg University where he was taught economics by Professor G. Knapp (1842-1926).

In Russia, Pearsonian biometrical ideas appeared to have started from the article "About Pearson's methods of application of the theory of probability to the problems of statistics and biology"<sup>2</sup> written by L. Lakhtin. Recently the same opinion was also expressed by E. Seneta<sup>3</sup>. Strictly speaking, however Lakhtin's article was devoted to the problem of the approximation of the curve of empirical distribution that appears in processing statistical data and the problem of analysis of the relationship between variables has not covered in the article. The publication of the handbook by A. Leontovich in three volumes (1909, 1911, and 1912). *Elementary handbook for application of Gauss and Pearson meth-*

*ods for evaluation of the errors in statistics and biology*<sup>4</sup> can be considered as the next step. This handbook was quite formal, it did not have logical-philosophic basis that became a prominent characteristic of the Russian school of correlation analysis afterwards. These characteristics appeared to the full extent in *Novels on the theory of statistics*<sup>5</sup> by A.A. Tchuprov. Leontovich noted that he was basically forced to work in this field in order to satisfy the necessity to be able to process results of scientific researches. He coincidentally discovered that there were not so many people in Russia who were familiar with the "methods of errors"<sup>6</sup>. There were no writings about the connection between the Gauss' method of errors and Pearson's method in heredity studies in Russia. Leontovich wrote that there were no books in Russia devoted to this method, and this encouraged him to publish his own notes on the connection between Gauss' method of errors and Pearson's method on heredity studies. He emphasized the actual calculation of problems which in turn led his entire third book to be a collection of calculative statistical-mathematical tables. However, as noted by Leontovich, most of its content can be used for the general theory of statistics.

Evaluating Leontovich's work, N.S. Chetverikov (1885-1973) wrote. "It can not be said that this book, whose author later became a famous physiologist and histologist, the member of the Academy of Sciences in Ukrainian Soviet Socialist Republic, distinguished with clarity and a correct presentation of compilable material. Though through this book a Russian reader could learn about the ideas of K. Pearson and his colleagues, and he could also learn about the cited literature of these authors here. This was enough to gain Evgeniy Evgenievich's (Chetverikov means E.E. Slutsky – the author) inter-

est. The speed to which E. Slutsky was able to learn from the originals about quite complicated models of English statisticians, how deep he went to the fundamentals of correlation theory, how he could critically appraise these works and outline the most essential points and also how he pointed out various flaws in Pearson's concepts were all simply remarkable (translated by the author)<sup>7</sup>.

## 2. E. Slutsky and beginning of the new era in Russian statistics

E. Slutsky (1880-1948) became interested in Leontovich's book. At the time of its publication, E. Slutsky had just graduated with excellence from law school at Kiev University for his graduation thesis *Theory of marginal utility*. The involvement in economics did not exclude his interest in empirical researches. His friendship with N.A. Svavitsky (1879-1936), an expert on regional statistics (*zemskaya statistika*), could possibly be the largest contributor to his interest in empirical researches. Aside from his wide sphere of interests, his versatile talents – mathematics, painting and poetry – played a large role in leading him to take a different exposition of statistical methods than Leontovich. A. Kolmogorov even said that E. Slutsky was “exquisite, a smart companion, a literature expert, a poet and a painter”<sup>8</sup>.

The acquaintance with Leontovich's book and with the English biometrical school raised in Slutsky an urge to share his revelations and his understanding with a Russian audience. That was how his book *Theory of correlation and elements of distribution curves study (handbook for studying some of the most important methods of modern statistics)* (Kiev, 1912) came about<sup>9</sup>. He had written: “Common revival of interest towards theoretical statistics allows hoping that

the dissemination of the ideas of the new school to all the countries and all of the spheres where it might be applied is a problem of not the distant future. A humble goal of the author is to contribute to this natural and inevitable process”<sup>10</sup>.

Considering this statement, it is clear that Slutsky saw the beginning of a new era in statistics in his ideas, moreover he declared this beginning from 1900-1920. He saw signs of this new era in improving the old methods and in discovering and developing of new ones that would show the applicability of statistical methods in biology and social sciences. Also he saw this new era in emergence of researchers who would control and manage further development of statistical methods. In this revival of the theory of statistics he gave “the palm of victory” to Karl Pearson (1857-1936), emphasizing that that this ranks him in mathematics together with Laplace, Gauss, and Poisson.

In the introduction to his own book, E. Slutsky already noted that the application of the new methods was relatively simple, aside from work simplified with availability of the special tables that were constructed upon an initiative by K. Pearson. However, it was necessary to understand the limits of the application of this method and the meaning of received results. He stated that this “requires not only knowledge and recipes for calculation, but understanding the spirit of the theory and its mathematical substantiation”<sup>11</sup>. Further he came to the conclusion that “a statistician has to be a mathematician because his science is a mathematical science”<sup>12</sup>. This conclusion is still a debatable and up-to-date issue at least for Russian statisticians. As he explained his exposition style was abundant with formulas and mathematical evidence. Apparently, this conclusion is fair for those who

engaged in development of statistical methods, but these “true statisticians” are inevitably surrounded by a group of scholars who feel that they do not have to be mathematicians in order to be valuable statisticians. On the contrary these “true statisticians” have a role that is quite important for the development of statistical methods and their applications. A changing society causes the emergence of new problems in statistical measurements of economic and social phenomenon, construction of adequate indicators, development of new methods of data mining, construction of large data bases with complex structures and so forth. To find a solution for these problems it is important to have not only statistical mathematicians but also experts in particular economics, sociology, and also it is necessary to have experts who have a statistical way of thinking. Having many different viewpoints on one problem between statistician-economists, statistician-sociologists, and statistician-mathematicians, the speed of development of statistical methods and their applications had increased.

The beginning of the 20th century was the starting point of understanding the importance of statistics and its successful application in different spheres of academics and society. In Russia, these ideas that became the core of quantitative methods of data processing appeared with some delay compared to other countries. It was necessary for Russia to perceive English terms and to build Russian terminology which would better reflect the core of new methods. On that account E. Slutsky noted that because of the scarcity of works and a large confusion over English statistics terminology developing at that time in Russian literature, even common phrases or terms such as “frequency curves” had more than five interpretations. The closest

to the modern translation was used by P. Orzhencky (1872-1923) – “curves of distribution of frequencies”<sup>13</sup>. Slutsky himself used two terms; “distribution curves” and “frequency curves”.

### 3. Slutsky's start from the theory of distribution curves

Even though the main subject of E. Slutsky's work was the theory of correlation, he started his exposition from the theory of distribution curves. This conforms to the logic of the theory of correlation which is based on discovery of variation coherence between two or more variables. That is how E. Slutsky argued for the necessity of distribution curve study: “Considering any group of individuals who possess common measurable characteristics, we notice that these characteristics are not the same numerical values for all individuals. There was a time when statisticians neglected this difference focusing on arithmetic mean a characteristic. At present there is no need to struggle with this out-of-date self-containment. It can be considered as common merit, the idea that arithmetic mean cannot fully describe characteristics of the whole statistical group and that the task of statistics is just to describe the structure of the given group as fully and as accurately as possible”<sup>14</sup>. The statements that were inarguable in the time of A. Quetelet (1796-1874) – when average was the center of attention (remember “average person”) –, but now it is just one of the components of the research where the first step is to describe the distribution of characteristics among the group under study.

Further E. Slutsky moved to the frequency moments that allow getting the characteristic of the data structure. Slutsky noted that the mo-

ments can be calculated using any reference point, but the moments calculated “relative to the arithmetic mean have the maximum value. They characterize the distribution”<sup>15</sup>. Along with that the point on the axis of abscissas which corresponds with the arithmetical mean of the characteristics represents the center of distribution, respectively, moments relative to the vertical axis that go through that point are called the central moments. Since there were no calculative equipment, the methods of calculation were extremely important at that time. Slutsky showed how to simplify calculations by using conditional and central moments when estimating the grouping data where some value of characteristics corresponded to the class mark and possibly close to the centre of distribution could be taken as conditional onset (conditional zero), and the range of interval could be taken as one. The book is accompanied with the formulas that link conditional moments with central ones.

E. Slutsky used probable errors of characteristic distribution (E). The formulas with agreed notation<sup>16</sup> are as follows.

$$E_h = 0,67449 \frac{\sigma}{\sqrt{N}}$$

$$E_\sigma = 0,67449 \frac{\sigma}{\sqrt{2\pi}}$$

$$E_v = 0,67449 \frac{V}{\sqrt{2N}} \sqrt{1 + 2\left(\frac{V}{100}\right)^2},$$

Where  $h$  – arithmetic mean;  $\sigma$  – standard deviation;  $V$  – coefficient of variation (or how Slutsky refers to it, coefficient of variability, “coefficient *izmenchivosti*” in Russian).

All of these formulas were derived from K. Pearson. Slutsky explained the necessity of calculating the probable errors from the position of the application of sampling method to the finite population; “If we studied the vast amount of

observations and calculated, for instance, the mean then calculated the calculated average value for this part that was significantly smaller than the population, and the probable error of the last value should be crucial when answering the question if there is a difference between this part and the whole population. The difference whether it is less or more than the probable error could be explained by random causes”<sup>17</sup>

We devoted so much time to this issue because it represents the supposition of the logic of the statistical conclusion which lies in the basis for mathematical statistics.

E. Slutsky paid to K. Pearson’s contribution to the distribution theory by noting that “Pearson curves” almost always end up with great results “delivering material characteristics in cases when normal (Gauss) curve fails to work for statisticians”<sup>18</sup>. Certainly, he could not leave out the limitations of veracity of the parameters of Pearson distribution curves: “In order to find curve:

$$\frac{1}{y} \frac{dy}{dx} = \frac{x - a}{b_0 + b_1x + b_2x^2},$$

It is necessary to know the four moments of observed distribution and in order to find a curve of bigger communality we find moments of 5th, 6th and bigger degrees”<sup>19</sup>. Here Slutsky talks about K. Pearson’s warning on the fact that errors of moments higher than the fourth range are significant and increase rapidly with the greater the value of the moment’s range. “That’s why the curve’s coefficients calculated with these means also should not be reliable”<sup>20</sup>.

Moving on to find coefficients in the equation  $y = f(x, a_1, a_2, \dots, a_n)$  that provides the best approximation of observable data. E. Slutsky offers a rule in which fairness was justified by K. Pearson in the article “On the Systematic Fitting of Curves”<sup>21</sup>.

In order to select a good “theoretical curve  $y = \phi(x, c_1, c_2, \dots, c_n)$  for the empirical curve, it is necessary to express the square and moments of the first curve in terms of its parameters ( $c_1, c_2, \dots, c_n$ ) and to equate them to the square and moments of empirical curve”<sup>22</sup>.

With this method a statistician was faced with many calculating difficulties, which were simplified with the help of Sheppard corrections that allowed a statistician to find the true values of central moments. In addition E. Slutsky analyzed the technique for finding the coefficients of not just linear functions but also parabolic curves when approximating the distribution curve.

#### 4. Slutsky’s discussions on the theory of correlation

Slutsky started the analysis of the theory of correlation with consideration of the term of correlation dependence: “...we say that several values are correlated if each set of the values of all variables except one variable corresponds with the whole complex values of the last variable, Moreover the arithmetic mean of each variable varies according to the values of others and frequency with every combination of variable values that come across are the function of these values”<sup>23</sup>. Addressing the correlation table he noted that it was nothing but “usual combination tables well-known to every statistician”<sup>24</sup>. The parallel between the correlation and combination tables was also shown later by A. Tchuprov, who emphasized that the correlation table allows not only the ability to bring whole totality of present numbers to “convenient-looking forms”<sup>25</sup>, but also gives an opportunity to overcome calculating difficulties which, as mentioned before, were very problematic at that time. The regression line, adequate to empirical

material within probable errors, can be used as criteria for the determination of the types of regression, linear or nonlinear. Slutsky included the nonlinear regression with undeveloped ones and focused firstly on analyzing linear regressions. Later in the 1930-1960s the linearity of regressions was specifically described by many scientists, but first by A. Goldberger. Goldberger named at least three reasons for preference of linear regression over nonlinear regression; 1) the fact that variation of variables is limited with particular (essential) limits: 2) the fact that primary measurements of variables are made quiet rough and a mathematical “delicacy” such as choosing the type of regression equation is unlikely to significantly improve it; and finally 3) if there is multi-collinearity, i.e. the multiple linear relationships between independent variables, also increases preference of linear regression that contain only original variables and not their functions”<sup>26</sup>. Slutsky started his exposition by identifying relationships between two items on district budget expenses (budgetov uezdbych zemstv) – the share of expenses on education and the share of expenses on supporting the district management which was based on data from 359 districts in 1901. Later he showed that the same methods could be applied when studying connections between the average prices of rye in Yelets and Samara from 1893 to 1903, that is, he showed the applicability of correlation method to cross-section data as well as to time series.

From a modern viewpoint, E. Slutsky’s explanation of correlation coefficient is quite remarkable. He represented it as a square root of the product of regression coefficients  $y$  to  $x$  and  $x$  to  $y$ :  $r_{yx} = \sqrt{tg\alpha \times tg\beta}$ . If there is no correlation then regressions coincide with axis of reference, i.e.

$\operatorname{tg}\alpha = 0$ ,  $\operatorname{tg}\beta = 0$ , therefore, the correlation coefficient equals zero. If correlation turns into linear functional dependence, then lines of regression are coincided, and the sum of angles  $\alpha$  and  $\beta$  will be equal to  $90^\circ$ . In this case  $\sqrt{\operatorname{tg}\alpha \times \operatorname{tg}(90 - \alpha)} = \sqrt{\operatorname{tg}\alpha \times \operatorname{ctg}\alpha} = 1$ . According to the ideas of E. Slutsky, these features make geometrical means of regression coefficients a convenient measure of the correlation. Nowadays this determination of correlation coefficient is rare, it was replaced by its interpretation as a measure of degree of deviation compatibility of variables  $y$  and  $x$  from their means  $\bar{y}$  and  $\bar{x}$ , i.e. in the foreground comes

cross product moment: 
$$\frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n}$$
 and,

therefore, the formula of correlation coefficient that includes this value becomes common:

$$r_{yx} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n\sigma_x\sigma_y}$$

This formula is also mentioned by E. Slutsky and he called it “the basic formula of correlation method”<sup>27</sup>.

E. Slutsky emphasized the conventionality of least-squares method which is usually used to estimate the parameters of regression equations. He noted that the regression could be produced by the absolute sum of distances between points on a line and empirical regression points would be the least, it is possible to find the line of regression for which the sum of squares of distances will be the least. It all depends on what the distance between empirical regression line and theoretical is taken as the common. E. Slutsky placed a lot of emphasis on the next expression:

$$\dot{a} = \frac{1}{N} \left( \sum_{(i)} \dot{o}_i - b \sum_{(i)} x_i \right) = 0$$

Where  $x$  and  $y$  are presented as deviations from their arithmetical means.

Therefore

$$\bar{y} = a + b\bar{x} \text{ or (in case of multiple regression) } \bar{y} = a + b_1\bar{x}_1 + \dots + b_k\bar{x}_k$$

Based on this expression E. Slutsky made the interesting conclusion that “in case of linear regression, presentation of the typical combination of arithmetical means is possible. Quetelet’s “average person” who has the average height for his age, the average sizes of different organs, average abilities and etc., is not something unreal because as range of statistical researches have shown (especially Pearson school) in anthropology it is possible to apply linear formulas to all kinds of characteristics with a very small margin of error”<sup>28</sup>.

In his book E. Slutsky analyzed not only principle and methods of construction of linear regression equation and correlation measurement, but also determination of average error of pair regression equations. Without exaggeration E. Slutsky can be called the true guide of correlation theory ideas in Russia. As P. Klyukin writes, Slutsky “got engaged with mathematical statistics and alone he efficiently developed biometric direction of F. Galton – K. Pearson in Kiev”<sup>29</sup>. Slutsky’s work was simply not just a retelling of the biometric school’s ideas, but contributed to the development of them which can be shown by the fact that E. Slutsky’s paper on “the Criterion of Goodness of Fit of the Regression Lines and the Best Method of Fitting them to the Data”<sup>30</sup> was published in the *Journal of the Royal Statistical Society*, in 1914. The article contained criticism towards the biometric school about solving the problem of regression estimation and that is why his work did not go unnoticed. In a letter on June 19th 1919, O. Anderson (1887-1960), a

fellow of A. Tchuprov, wrote, “as far as I can see E. Slutsky’s work is right by its idea and can be considered to be the first approximation to solving the raised problem and it was the first attempt to approach regression lines in the terms of Pearson’s criteria  $\chi^2$ ”. Further O. Anderson also noted that Slutsky’s paper got a response from not only Pearson’s followers, but Pearson himself. Pearson’s followers’ publications are really just further developments and improvements on the ways determined by Slutsky. As Anderson had written: “if Pearson is right that E. Slutsky’s research can not be considered as the final solution of the problem, then his appraisal of the author’s work is still not fair and is full of personal irritancy with the author and are loosely based on scientific reason. Pearson’s adjustments to E. Slutsky’s formulas made the comparatively minimal changes to the values of coefficients calculated by Slutsky to the both given numeric illustrations and to application of new formulas. I suppose that Slutsky’s work identifies him as a great expert of Pearson’s criteria and the methods of statistical research connected with it and showed that a person equal to Slutsky’s ability can be rarely found in Russia today”<sup>31</sup>.

The book *Textbook of mathematical statistics* (1914) by R. Orzhentsky is perhaps less intellectual in its ideas and style, but is still very important for the introduction of the ideas of correlation in Russia.

Orzhentky, similarly to Slutsky, reviewed the theory of distribution curves and later moved on to the analysis of correlation theory. He was not as mathematically talented as Slutsky and maybe that is the reason why his work did not leave as large of an impact as E. Slutsky’s book, however his work still had many positive aspects to it. Aside from its ultramodern title, its

value came mostly from the fact that his work named all of the Russian scientists who had contributed to the introduction of the ideas of variational statistics and theory of correlation. He mentioned V. Kosinsky, V. Vasiliev, A. Leontovich, E. Slutsky. This shows his knowledge of national statistical literature and his aspiration to give credit to the preceding researchers. However, it should be noted that the term “mathematical statistics” defined by R. Orzhentsky doesn’t correspond to the modern understanding of mathematical statistics as a science about the estimation of parameters of population and testing of statistic hypothesis of parameters and characteristics of population. In that sense mathematical statistics formed as a separate area of knowledge a little bit later in the 1920s to 1930s by efforts of R. Fisher (1890-1962).

## 5. Concluding remarks—correlation theory in Russian statistics after Slutsky

In this paper we have clarified Slutsky’s role in introducing correlation methods in Russia. Finishing our review, let us emphasize that in the beginning of the 20th century in Russia the term “correlation” was used to determine the degree of contingencies in changes of two or more variables. And the coefficient of pair correlation, regression equation, and standard error of regression equation were introduced to Russian statistical science. An issue about quality of regression equation, linear and nonlinear correlation was raised, and the research on multiple correlation relationships were started.

The theory of correlation in Russia was faced with a lot of challenges. The theory of correlation became the “apple of discord” between statistician-mathematicians and statistician-non-mathematicians. “Statistician-mathemati-

cians look down on the methods used by non-mathematician statisticians as rough and inconsiderate. Non-mathematician statisticians leave questions unanswered and deny mathematical ways as scientifically senseless amusement for fun to play with numbers and mathematical symbols<sup>32</sup>. In this context we can see A. Tchuprov's vexation to the position of one of the leading St. Petersburg's statisticians A.A. Kaufman (1864-1919) who was quite skeptical about correlation methods<sup>33</sup>.

Skeptical attitude toward correlation methods in social research is not uncommon for Russian scholars. The critical attitude towards the correlation methods by J.M. Keynes is well-known. Keynes thought that since there are no independent events in economic reality, the application of the theory of probability is under question. Moreover, Keynes saw harm in using mathematical methods because they create clear numeric results which can look very persuasive, but these estimations in reality can be unjustifiable and it is necessary to avoid them. Based on this he considered the correlation estimation to be especially dangerous because it can create visual relationships that are conceptually sound but absent in reality<sup>34</sup>. Keynes' skeptical attitude towards the correlation methods would remain as an object of research.

In the end of the 1940s because of the development of the Soviet economic planning systems where there is no place for stochasticity, reacting against the idea of stochasticity of social relationships, theory of correlation was almost completely excluded from social sciences in Russia. Later in the 1950s it ended with the official acceptance at the national Soviet Union statistical conference in 1954<sup>35</sup>. Soviet statisticians accepted the determination of statistics as a separate "social science" that studied the quantitative side of social production in its union of productive forces, productive relations, and phenomenon of cultural and political life. Through the methodological union of statistics that has many areas of application, understanding of statistics as a science about method was destroyed, and the barrier between social knowledge and natural science was raised. The persecution of the correlation theory was forced at the time of persecution of genetics (1948). Elucidation of the social background for acceptance of correlation methods since 1954 in Russia still remained as an object of further research.

It has been many decades since the theory of correlation in Russia took an appropriate place among statistical methods of research including social and economic applications.

### Notes

- 1) Slutsky E.E. *Teoriya correlyatsii i elementy ucheniya o krivyyh raspredeleniya (posobie k izucheniyu nekotorykh vazhneishih metodov sovremennoi statistiki)*. Kiev, 1912. (Slutsky E.E., *Theory of correlation and elements of distribution curves study (handbook for studying some of the most important methods of modern statistics)*, Kiev, 1912.)
- 2) Lakhtin L.K. "O metode Pirsona v prilozheniyah teorii veroyatnosti k zadacham statistiki i biologii", *Matematicheskiy sbornik, izdavaemyi Moskovskim matematicheskim obshchestvom*. T.24. Vyp.3. M., 1994. pp.481-500. (Lakhtin L.K. "About Pearson methods of application of the theory of probability to the problems of statistics and biology", *Mathematical collection, published by Moscow mathematical society*. Vol. 24. No. 3, M.: University Publishing House, 1904. pp.481-500.)
- 3) Seneta E. "Karl Pearson in Russian Contexts", *International Statistical Review*. Vol. 77. No. 1. 2009.

- pp.118-146.
- 4) Leontovich A.V. *Elementarnoe posobie k primeneniyu metodov Gaussa i Pearsona pri otsenke oshibok v statistike i biologii*. Kiev, 1909, 1911, 1912. (Leontovich A.V. *Elementary handbook for application of Gauss and Pearson methods for evaluation of errors in statistics and biology*. Kiev, 1909-1912.)
  - 5) Tchuprov A.A. *Ocherki po teorii statistiki*. St. Petersburg, 1910. (Tchuprov A.A. *Novels of the theory of statistics*. Saint-Petersburg, 1910.)
  - 6) Leontovich A.V. *Op. cit.*, part 1. p.258
  - 7) Chetverikov N.S. “Zhizn i nauchnaya deyatelnost E.E. Slutskogo (1880-1948)”, *Uchyonye zapiski po statistike AN SSSR*. T.5. M., 1959. p.258. (Chetverikov N.S. “Life and scientific work of E.E. Slutsky (1880-1948)”, *Scientific notes on statistics in Academy of Sciences in USSR*. T.5. M., 1959. p.258.)
  - 8) Kolmogorov A.N. “Evgeniy Evgenievich Slutskiy [Nekrolog]”, *Uspehi matematicheskikh nauk*. 1948. T.3. Vyp.4. (Kolmogorov A.N. “Evgeniy Evgenievich Slutsky [Obituary]”, *Success of mathematical sciences*. 1948. Vol. 3, Issue 4.)
  - 9) Slutsky, *Op. cit.*
  - 10) *Ibid*, p.1.
  - 11) *Ibid*, p.2.
  - 12) *Ibid*, p.2.
  - 13) Orzhenskiy R.M. *Uchebnik matematicheskoi statistiki*. St. Petersburg, 1914, p.215. (Orzhenskiy R.M., *Textbook of mathematical statistics*. Saint-Petersburg, 1914. p.215.)
  - 14) Slutsky, *Op. cit.*, p.5.
  - 15) *Ibid*.
  - 16) *Ibid*, p.12-14.
  - 17) *Ibid.*, p.12-14.
  - 18) *Ibid.*, p.13.
  - 19) *Ibid*.
  - 20) *Ibid*.
  - 21) Pearson. K. “On the Systematic Fitting of Curves”, *Biometrika*. Vol. 1, pp.267-271.
  - 22) Slutsky, *Op. cit.*
  - 23) *Ibid.*, p.59.
  - 24) *Ibid.*, p.61.
  - 25) Tchuprov A.A. *Osnovnye problemy teorii korrelatsii*. M., 1926, p.8. (Tchuprov A.A. *Main problems of the theory of correlation*. M., 1926. p.8.)
  - 26) Goldberger A.S. *Econometric Theory*. Wiley, 1964.
  - 27) Slutsky, *Op. cit.*, p.76.
  - 28) Slutsky, *Op. cit.*, p.75.
  - 29) Klyukin P.N. “Tvorcheskaya biografiya E.E. Slutskogo v svete arhivnykh fondov”, *Sbornik izbrannykh statei molodykh uchenykh*. M, 2010, p.278. (Klyukin P.N., “Creative biography of E.E. Slutsky in the light of archive materials”, *Collection of selected papers of young scientists*. Moscow, 2010. pp.278.)
  - 30) Slutsky E.E. “On the Criterion of Goodness of Fit of the Regression Lines and the Best Method of Fitting them to the Data”, *Journal of the Royal Statistical Society*. Vol. LXXVII. Pt. I (Dec., 1913), 1914. pp.78-84
  - 31) Elisseeva I.I., Volkov A.G.E., E. “Slutsky: zhizn i deyatelnost”, *Izvestiya Sankt-Peterburgskogo Universiteta Ekonomiki i Finansov*. 1999, No. 1, p.115. (Elisseeva I.I., Volkov A.G., “E.E. Slutsky: life and work”, *Proceeding of St. Petersburg University of Economics and Finance*. 1991. No. 1. p.115.)
  - 32) Tchuprov A.A. *Osnovnye problemy teorii korrelatsii*. M., 1926, p.7. (Tchuprov A.A., *Main problems of the theory of correlation*. M., 1926. p.7.)
  - 33) Kaufman A.A., “Korrelatsionnye formuly kak orudie statisticheskogo analiza”, *Statisticheskii vestnik*, 1915, kn. 3. (Kaufman A.A., “Correlation formulas as tool of statistic analysis”, *Statistical Newsletter*. 1915. No. 3). Kaufman A.A. “Esche k voprosu o znachenii korrelatsionnykh formul kak orudiya statisticheskogo

analiza (po povodu statei prof. A.V. Leontovicha, prof. *R.M. Orzhentskogo i I.A. Saboneeva*)", *Statisticheskii vestnik*, 1916, kn. 4. (Kaufman A.A., "More to the question of value of correlation formulas as tool of statistical analysis (concerning papers of professor A.V. Leontovich, professor R.M. Orzhentsky and I.A. Saboneev)", *Statistical Newsletter*. 1916. No. 4.)

34) Skidelsky G. *John Maynard Keynes*. Vol. 1. M., 2005. p.261-262. (in Russian).

35) "Materialy Vsesoyuznogo Soveschaniya po statistike 1954 goda", *Vestnik statistiki*, 1954, No. 5, p.87. ("Materials of Whole-Soviet Union statistical conference in 1954", *Statistical Newsletter*, 1954. No. 5, p.87.)

## 編集委員会からのお知らせ

山口秋義（編集委員長）

機関誌『統計学』の編集・発行について

1. 常時、投稿を受け付けます。
2. 各号ごとに投稿の締め切りを設けます。その期日までに受け付けた原稿でも、査読の進捗如何によつては、その号に掲載されないことがあります。
3. 投稿に際しては、2012年9月の総会において改正された「投稿規程」、「執筆要綱」、「査読要領」をご熟読願います。
4. 原稿は編集委員長に宛ててお送り願います。
5. 原稿はPDF形式のファイルとして提出してください。また紙媒体での提出も旧規程に準拠して受け付けます。紙媒体の送付先も編集委員長としてください。
6. 原則としてすべての投稿原稿が査読の対象となります。
7. 今後の締め切りは次のとおりです。  
A：「論文」・「研究ノート」；B：その他
  - (1) 第105号（2013年9月30日発行予定）  
A：2013年7月31日；B：2013年8月31日
  - (2) 第106号（2014年3月31日発行予定）  
検討中（学会HPなどでお知らせします）
8. 次年度（2013年4月－2014年3月）編集委員会メンバー（敬称略）は次のとおりです。  
金子治平（関西，委員長，原稿送付先），西村善博（九州，副委員長），山田 満（関東），橋本貴彦（関西），栗原由紀子（関東）

以上

---

### 編集後記

研究成果をご投稿いただいた会員諸氏に御礼申し上げます。また製版と発送の作業を昭和情報プロセス株式会社様と音羽リスマチック株式会社様にお世話になりました。この場をお借りして御礼申し上げます。本号では山口秋義（編集委員長）が責任編集を務め、前田修也（東北支部編集委員）が発行業務を担当しました。

（山口秋義 記）

---

## 執筆者紹介 (掲載順)

山口 幸三 ((独)統計センター)

伊藤 伸介 (明海大学経済学部)

秋山 裕美 ((独)統計センター)

櫻本 健 (松山大学経済学部)

IRINA ELISEEVA (Dr. Professor, Member-in-correspondence of Russian Academy of Sciences)

## 支部名

## 事務局

北海道	062-8605	札幌市豊平区旭町 4-1-40 北海学園大学経済学部 (011-841-1161)	水野谷武志
東北	986-8580	石巻市南境新水戸 1 石巻専修大学経営学部 (0225-22-7711)	深川通寛
関東	192-0393	八王子市東中野 742-1 中央大学経済学部 (042-674-3424)	芳賀寛
関西	525-8577	草津市野路東 1-1-1 立命館大学経営学部 (077-561-4631)	田中力
九州	870-1192	大分市大字旦野原 700 大分大学経済学部 (097-554-7706)	西村善博

## 編集委員

水野谷武志 (北海道)

前田修也 (東北)

岡部純一 (関東)

良永康平 (関西) [副]

山口秋義 (九州) [長]

## 統計学 No.104

---

2013年3月31日 発行	発行所	経済統計学会 〒194-0298 東京都町田市相原町4342 法政大学日本統計研究所内 TEL 042(783)2325 FAX 042(783)2332 <a href="http://www.jsest.jp/">http://www.jsest.jp/</a>
	発行人	代表者 森 博美
	発売所	株式会社 産業統計研究社 〒162-0801 東京都新宿区山吹町15番地 TEL 03(5206)7605 FAX 03(5206)7601 E-mail : <a href="mailto:sangyoutoukei@sight.ne.jp">sangyoutoukei@sight.ne.jp</a> 代表者 品川 宗典

---

# STATISTICS

---

No. 104

2013 March

---

## Articles

Generating Pseudo Microdata for Educational Use in Japan  
..... Kozo YAMAGUCHI, Shinsuke ITO, Hiromi AKIYAMA (1)

Analysis of IO-based Annual Supply and Use Tables for the Development of QNA  
..... Takeshi SAKURAMOTO (16)

## Note

Introduction of the Theory of Correlation into Russia and E. Slutsky  
..... IRINA ELISEEVA (41)

## Activities of the Society

Activities in the Branches of the *Society* ..... (52)

Bylaws of the Society, Regulation of the Editorial Committee, Prospects for the Contribution  
to the Statistics ..... (57)

---

JAPAN SOCIETY OF ECONOMIC STATISTICS

---