

**М. А. Суворова**

### **Методика организации лабораторных занятий по курсу «Методы прикладной статистики»**

В статье представлен способ организации лабораторных работ по курсу «Методы прикладной статистики». Современный человек должен владеть компьютером, знать его возможности и уметь применять его для решения профессиональных задач – в частности, использовать базы данных для сбора информации, применять прикладные программы для обработки информации.

При изучении этого курса сталкиваемся со следующими основными проблемами: 1) студенты не знают в полной мере всех возможностей прикладных программных продуктов, даже таких широко распространенных, как MS Excel; 2) студенты не умеют ставить гипотезы исследования и интерпретировать полученные результаты; 3) студенты не могут оценить правильность полученных результатов.

В разработанном курсе лабораторных работ исходными данными являются реальные протоколы соревнований по биатлону. На одних и тех же исходных данных предлагается проверить различные статистические гипотезы, то есть провести полноценный статистический анализ. Так как эти соревнования смотрит большое число людей, то полученные статистические выводы легко оцениваются и интерпретируются.

Ключевые слова: интерпретация результатов, лабораторный практикум, прикладная программа MS Excel, прикладная статистика, статистические критерии, эмпирическая информация.

**M. A. Suvorova**

### **Organization of Laboratory Studies in the Course «Methods of Applied Statistics»**

The article presents a way of organizing the laboratory work in the course «Methods of Applied Statistics». Modern man must own a computer, know its capabilities and be able to apply it to solve professional problems. For example: use the database to collect information, to use applications for data processing.

In the study of this course we are faced with the following main problems: 1) students do not know the full extent of all the possibilities of software applications, even those widely used such as a MS Excel; 2) students are not able to put the research hypothesis and interpret the results; 3) students cannot evaluate correctness of the results.

In the developed laboratory work initial data are actual protocols biathlon competitions. The same initial data are proposed to test various statistical hypotheses, i.e., to conduct a full statistical analysis. Since this competition is being watched by many people, received statistical inference is easily evaluated and interpreted.

In the developed course labs input data are real protocols of the competitions in biathlon. The same initial data are proposed to test various statistical hypotheses, i.e., to conduct a full statistical analysis. Since this competition is watched by a large number of people, the resulting statistical inferences are easily estimated and interpreted.

Keywords: interpretation of results, laboratory practical, application of Excel, applied statistics, statistical tests, empirical information.

Прикладная статистика – наука о методах обработки статистических данных. «В СССР термин “прикладная статистика” вошел в широкое употребление в 1981 г. после выхода массовым тиражом сборника “Современные проблемы кибернетики (прикладная статистика)”. В этом сборнике обосновывалась трехкомпонентная структура прикладной статистики. Во-первых, в нее входят ориентированные на прикладную деятельность статистические методы анализа данных. Однако прикладную статистику нельзя целиком относить к математике. Она включает две нематематические области: методологию организации статистического исследования и органи-

зацию компьютерной обработки данных, в том числе разработку и использование баз данных и электронных таблиц» [2].

Современный человек должен владеть компьютером, знать его возможности и уметь применять его для решения профессиональных задач. В частности, использовать базы данных для сбора информации, применять прикладные программы для обработки информации.

В соответствии с образовательным стандартом направления подготовки 39.03.01 «Социология» выпускник должен быть готов решать следующие профессиональные задачи в научно-исследовательской деятельности: «Обработка

социальной, демографической, экономической и другой релевантной эмпирической информации с привлечением широкого круга источников на основе использования современных информационных технологий, средств вычислительной техники, коммуникаций и связи».

Многие сегодняшние студенты используют компьютер лишь как пишущую машинку или как средство общения. Даже сидя за компьютером, простейшие вычисления пытаются делать на калькуляторе или на мобильном телефоне. Это говорит о том, что студенты не знают в полной мере всех возможностей прикладных программных продуктов. В связи с этим одной из целей лабораторного практикума является ознакомление студентов с прикладной программой MS Excel, предназначенной для обработки информации с помощью электронных таблиц.

Лабораторные работы проводятся в среде MS Excel. Это приложение разработано специально для того, чтобы предоставлять разнообразные инструменты для вычислений, анализа и управления данными, а также визуализации результатов. Возможности электронных таблиц Excel очень широки. К их числу относятся решение вычислительных задач, визуализация данных с помощью диаграмм, обработка и анализ статистических данных и т. д. Знание программы позволяет своевременно получать нужные данные и уметь их представлять в наглядном виде, что крайне важно для работы квалифицированного специалиста в любой сфере деятельности.

Следующая проблема изучения методов прикладной статистики в умении студентов ставить гипотезы исследования и интерпретировать полученные результаты.

Существует множество пособий [1, 3] по проведению лабораторных работ по курсу статистики в MS Excel. В этих пособиях на конкретных примерах описываются возможности решения конкретных задач средствами MS Excel. При получении исходных данных для статистического анализа многие студенты спрашивают, какой критерий нужно использовать. На лабораторных занятиях в качестве исследуемого материала выбраны биатлонные гонки. Протоколы соревнований содержат разные типы данных (числовые – количество промахов, временные – время бега, стрельбы и т. д.). Соревнования проводятся на разных трассах, в разных погодных условиях, то есть студентам предоставляется широкое поле для творчества при постановке гипотез. На одних и тех же исходных данных можно проверить

большое количество статистических гипотез. Так как эти соревнования смотрит большое количество людей, учимся интерпретировать полученные результаты.

Содержание лабораторного практикума охватывает все темы курса «Методы прикладной статистики», а именно:

1. Первичный статистический анализ эксперимента.
2. Сравнение распределений по критерию Пирсона.
3. Статистические критерии.
4. Корреляционный и регрессионный анализ.

Многие задания лабораторных работ снабжены подробными методическими указаниями по выполнению этих заданий. Для похожих заданий пояснений нет. До занятия необходимо ознакомиться с соответствующим теоретическим материалом по теме лабораторной работы. Во время занятия в компьютерном классе студенты самостоятельно выполняют задания, результаты работы сохраняют в виде файла. После выполнения всех заданий лабораторной работы и собеседования по ней с преподавателем студенту выставляется зачет по данной работе.

#### **Лабораторная работа № 1**

Целью первой лабораторной работы является знакомство с основными принципами работы в MS Excel и выполнение первичного статистического анализа исходных данных.

#### **Лабораторная работа № 2**

Целью работы является изучение статистического критерия Пирсона для сравнения эмпирических распределений и сравнения эмпирического и теоретического распределений. Студенты учатся строить ось значимости и интерпретировать полученные результаты.

#### **Лабораторная работа № 3**

Изучение параметрических и непараметрических статистических критериев для зависимых и независимых выборок.

#### **Лабораторная работа № 4**

Целью четвертой лабораторной работы является изучение корреляционного и регрессионного видов анализа средствами MS Excel.

В качестве примера рассмотрим организацию лабораторной работы по первичному статистическому анализу. Работа посвящена основам работы в MS Excel, а именно приемам оформления таблиц, способам форматирования исходных данных, изменению формата представления дан-

ных, переводу данных из одного формата в другой. Особое внимание уделяется работе с формулами в электронных таблицах, также осваивается важное понятие электронных таблиц – *адресация ячеек*, изучается абсолютная, относительная и смешанная адресация. В этой же работе изучаются приемы визуализации данных: построение диаграмм и графиков.

Также в первой лабораторной работе изучаются способы представления статистической информации, а также основные возможности MS Excel по обработке статистических данных.

**Задание 1.** Выполните первичный статистический анализ биатлонной гонки по следующим критериям: чистое время бега, чистое время стрельбы, количество промахов.

Первичный статистический анализ предлагается выполнять двумя способами: 1) используя стандартные функции, 2) используя пакет анализа.

При выполнении этого задания студенты должны научиться интерпретировать полученные результаты, правильно указывать формат результатов: при анализе времени основные характеристики положения, среднее квадратическое отклонение также должны быть временного формата, при анализе числовых данных (количество промахов) результаты должны быть числового формата.

Вторая группа заданий нацелена на разные способы представления исходных данных: в виде дискретного и интервального рядов. Построить дискретный вариационный ряд по числу промахов и интервальный вариационный ряд по времени бега.

При построении дискретных и интервальных вариационных рядов студенты должны отработать понятия варианты, частоты и объема выборки, научиться правильно выбирать способ разбиения на интервалы.

**Далее студенты должны построить дискретный и интервальный ряды по времени стрельбы.**

**Следующая группа заданий нацелена на вычисление основных характеристик вариационных рядов.** При выполнении этих заданий отрабатываются алгоритмы вычисления основных характеристик табличным способом. Студентам важно понять, что для данных, представленных в виде дискретного и интервального ряда, нельзя применять стандартные статистические функции.

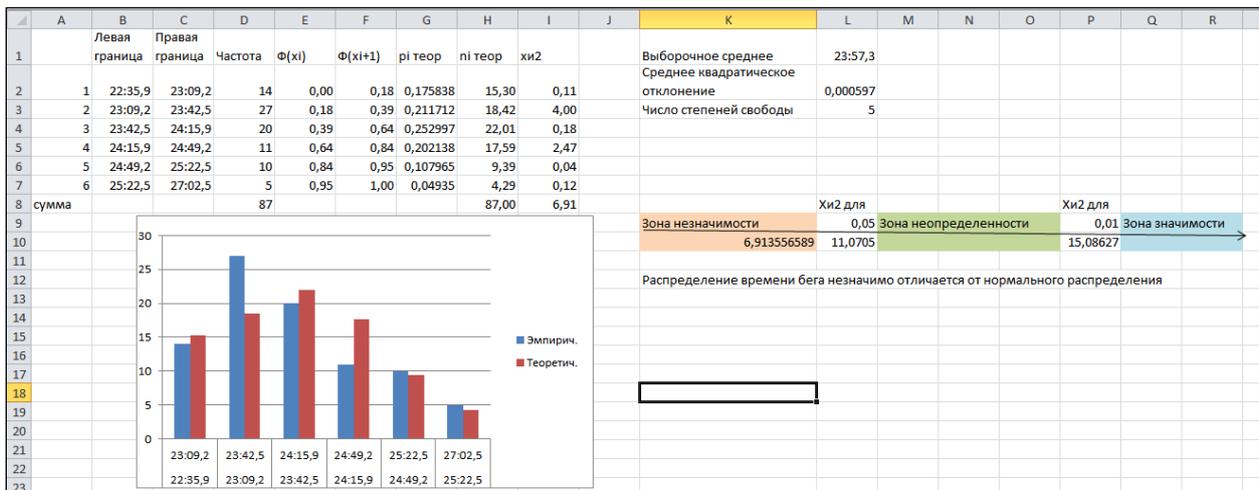
После выполнения этих заданий важно научить проверять полученные результаты. Так как вычисления производились дважды разными методами, то и результаты, вычисленные с помощью стандартных функций для исходных выборок и результаты, вычисленные табличным способом для дискретного ряда должны совпадать, а для интервального ряда – должны быть близки между собой.

**Последняя группа заданий нацелена на изучение графических возможностей MS Excel, а именно на построение полигонов частот для дискретных рядов и гистограмм для интервальных рядов.**

Итак, при выполнении первой лабораторной работы студенты знакомятся с основными возможностями работы в MS Excel. При выполнении следующих работ они используют основные знания, умения и навыки для полноценного статистического анализа результатов.

**Пример:** проанализировать, соответствует ли распределение чистого времени бега нормальному закону распределения.

Построение интервального ряда и вычисление основных характеристик по нему студенты выполняли на первой лабораторной работе, результаты которой являются исходными данными для текущей работы. Значит, при выполнении этой работы исходные данные требуется лишь скопировать. При проверке на соответствии нормальному закону распределения на практических занятиях студенты вычисляли нормализованные варианты, находили по таблицам значения функции Лапласа, что требовало постоянного внимания. При использовании электронных таблиц MS Excel студенты используют лишь функцию НОРМРАСП, и трудоемкие вычисления выполнены. После выполнения всех расчетов строится ось значимости. Причем критические значения также находятся без таблиц, используется встроенная функция ХИ2ОБР, зона расположения эмпирического значения также определяется автоматически с использованием встроенной функции ЕСЛИ. Для подтверждения полученных математических результатов строится гистограмма, на которой хорошо видно, отличается распределение от нормального закона или нет.



Таким образом, при выполнении лабораторных работ студенты осваивают основные принципы работы в MS Excel, учатся выполнять статистическую обработку результатов эксперимента, интерпретировать полученные результаты и все это выполнять на реальных данных большого объема.

По завершении курса «Методы прикладной статистики» студенты должны разработать проект по одной из выбранных тем социологического исследования, где основными требованиями являются правильный выбор исходных данных, аргументированный выбор статистических критериев и корректная интерпретация результатов. При выполнении этого проекта они могут применять изученные возможности MS Excel.

### Библиографический список

1. Макарова, Н. В. Статистика в Excel [Текст] : учеб. пособие / Н. В. Макарова, В. Я. Трофимец. – М. : Финансы и статистика, 2002. – 368 с.
2. Орлов, А. И. Прикладная статистика [Текст] / А. И. Орлов. – М. : Экзамен, 2004.
3. Сапегин, А. Г. Психологический анализ в среде Excel. Математические методы и инструментальные средства [Текст] / А. Г. Сапегин. – М. : Ось-89, 2005. – 144 с.

### Bibliograficheskij spisok

1. Makarova, N. V. Statistika v Excel [Tekst] : ucheb. posobie / N. V. Makarova, V. Ja. Trofimec. – M. : Finansy i statistika, 2002. – 368 s.
2. Orlov, A. I. Prikladnaja statistika [Tekst] / A. I. Orlov. – M. : Jekzamen, 2004.
3. Sapegin, A. G. Psihologicheskij analiz v srede Excel. Matematicheskie metody i instrumental'nye sredstva [Tekst] / A. G. Sapegin. – M. : Os'-89, 2005. – 144 s.