

**Применение пакетов прикладных статистических программ
в образовательном процессе в техническом вузе**

Прихач Н.К., Прусова И.В.

Белорусский национальный технический университет

Аннотация:

Описывается применение прикладных статистических программ (на базе пакетов MS Excel и STATISTICA) в образовательном процессе студентов специальности «Метрология, стандартизация и сертификация».

Использование статистических методов обработки данных стали привычным и широко распространенным аппаратом для инженеров. Работа инженера связана с научными исследованиями производственных процессов и технологий, поэтому важно, чтобы студенты в техническом вузе умели пользоваться различными прикладными пакетами и решать исследовательские задачи.

Математическая статистика – раздел математики, посвященный математическим методам сбора, систематизации, обработки и интерпретации статистических данных и использование их для научных или практических выводов. Правила и процедуры математической статистики опираются на теорию вероятностей, которая позволяет оценить точность и надежность выводов, получаемых в каждой задаче на основании статистического материала.

Задачи математической статистики не ограничиваются расчетом частот и оценкой вероятности наступления интересующего студента события. Основным объектом изучения математической статистики является случайная величина X , над которой проводятся наблюдения (эксперименты, испытания) с целью получения данных для анализа статистического вывода о случайной величине X : о ее законе распределения, числовых характеристиках и т. д.

Статистические методы в настоящее время широко используются на различных этапах контроля и управления качеством продукции, способствуя созданию условий для принятия рациональных управленческих решений с использованием накопленной информации, их эффективного взаимодействия с внешней средой [1].

Создание универсальных и специализированных пакетов прикладных программ для проведения статистического анализа данных (таких как MS Excel и STATISTICA) стало следующим этапом внедрения методов математической статистики в практику проведения исследований.

Умение делать правильные выводы на основе полученных с помощью программ данных определяется, в первую очередь, квалификацией и знаниями специалиста, применяющего пакет. Основные требования к специалисту, который использует статистические пакеты следующие: способность правильно выбрать метод решения задачи, учесть его возможности и ограничения, грамотно применить метод, а также верно интерпретировать полученные результаты. Поэтому дисциплина «Прикладная математика», разработанная на кафедре «Инженерная математика» ПСФ БНТУ для специальности 1-54 01 01 «Метрология, стандартизация и сертификация» имеет следующие цели: познакомить студентов с теоретическими основами статистической обработки данных, научить их оптимизировать и автоматизировать данный процесс с помощью компьютера.

В учебно-методическом пособии для лабораторных работ излагаются основные понятия, приемы и методы статистического анализа данных. В качестве основного инструмента статистического анализа используются возможности пакетов MS Excel и STATISTICA.

Весь изучаемый материал разбит на восемь лабораторных работ. На каждом занятии студент получает индивидуальное задание, которое выполняет самостоятельно под руководством преподавателя. Варианты заданий приведены в конце каждой лабораторной работы. Кроме того, лабораторный практикум содержит также разработанные в Excel и STATISTICA документы с многочисленными примерами.

Статистические расчеты без использования универсальных или специализированных программ являются не столько сложными, сколько затратными по времени, а также требуют повышенного внимания в связи с применением многочисленных таблиц функций, квантилей стандартного распределения и прочего. Поэтому такая практика «ручного» вычисления не позволяет за своей громоздкостью увидеть что-то новое в изучаемом материале. Однако использование компьютера при выполнении статистических расчетов может полностью скрыть от студента алгоритм вычислений, показать лишь конечный результат. Поэтому в рамках обучения методам математической статистики необходимо находить баланс между автоматизированными и «ручными» расчетами.

Остановимся подробнее на содержании лабораторных работ. Первая работа посвящена первичной обработке данных и точечному оцениванию параметров распределения.

Первичная обработка позволяет представить первичный числовой массив в сжатой форме, выделить основные закономерности в изучаемой совокупности случайных величин и, соответственно, также в генеральной совокупности.

Далее студентам необходимо изучить точечные оценки числовых характеристик и методы оценок параметров распределения. Оценки параметров вычисляются с помощью стандартных функций пакета Excel, а также с помощью статистической процедуры *Описательная статистика* из *Пакета анализа*. Также для нахождения точечных оценок используются возможности пакета STATISTICA – вычисления происходят с помощью модуля *Basic Statistics/Tables*.

В заключении необходимо сделать вывод о предполагаемом законе распределения случайной величины.

Во второй лабораторной работе решается вопрос согласования результатов оценивания с опытными данными: выдвигается гипотеза, что наблюдаемая случайная величина подчиняется нормальному закону. Далее рассчитываются доверительные интервалы для параметров нормального распределения – математического ожидания и среднего квадратического отклонения.

В третьей лабораторной работе проводится проверка гипотез о параметрах распределения. Даются основные понятия и принципы проверки параметрических гипотез. Также представлены основные направления сравнения двух и более выборок. На примерах показаны особенности реализации этих методов с помощью инструментов из *Пакета анализа MS Excel* и из *Basic Statistics/Tables* пакета STATISTICA.

В следующей лабораторной работе с помощью *дисперсионного анализа* (ANOVA) изучается влияние одного или нескольких качественных факторов на наблюдаемую случайную величину.

Проверка значимости различия между средними является целью дисперсионного анализа. Данная проверка проводится при помощи сравнения (анализа) дисперсий. Разделение общей дисперсии на несколько источников позволяет сравнить дисперсию, вызванную различием между группами, с дисперсией, вызванной внутригрупповой изменчивостью. При истинности нулевой гипотезы (например, о равенстве средних в нескольких группах наблюдений, выбранных из генеральной совокупности), оценка дисперсии, связанной с внутригрупповой изменчивостью, должна быть близкой к оценке межгрупповой дисперсии [2].

Пятая и шестая лабораторная работы знакомят студентов с элементами регрессионного и корреляционного анализа, с помощью которых можно изучать статистические связи между переменными. Выявление наличия связи между случайными величинами и оценка тесноты данной связи – основная задача корреляционного анализа. В случае регрессионного анализа основная задача – это установление формы зависимости и изучение этой формы между переменными.

Следующая лабораторная работа посвящена методам непараметрической статистики. На практике при проведении измерений иногда появляется необходимость в обработке данных «низкого качества», из выборок малого объема и/или без каких-либо (либо очень малым количеством) данных о распределении случайной величины. Для обработки таких данных могут применяться непараметрические методы. Данные методы разработаны для случаев, когда о параметрах исследуемой величины известно или мало, или вообще ничего. При описании выборочного распределения интересующей величины оценка параметров (например, среднее или стандартное отклонение) не используется в непараметрических методах.

В последней лабораторной работе курса «Прикладная математика» приведены основные подходы к прогнозированию на основе анализа временных рядов: методы сглаживания (скользящее среднее, метод экспоненциального сглаживания), метод экстраполяции и способы его реализации в MS Excel и STATISTICA.

Таким образом, предложенные лабораторные работы по курсу «Прикладная математика» существенно расширяют объем изучаемого материала по математической статистике. С другой стороны, студенты получают наглядное представление решенной задачи.

На основании вышеизложенного, данные лабораторные работы позволяют студентам использовать современные информационные технологии в изучении математической статистики.

Литература

1. Статистические методы контроля качества: учебно-методический комплекс для студентов специальности 1-54 01 01-01 «Метрология, стандартизация и сертификация» (машиностроение и приборостроение) / Белорусский национальный технический университет, кафедра «Стандартизация, метрология и информационные системы»; сост. В.Л. Соломахо. – Минск: БНТУ, 2017. – 202 с.

2. Дубровина, О.В. Прикладная математика: методическое пособие по выполнению практических и лабораторных работ для студентов заочного отделения специальности 1-54 01 01 «Метрология, стандартизация и сертификация» / О.В. Дубровина, Н.К. Прихач, В.М. Романчак. – 2009. – 70 с.