

Уважаемые члены государственной аттестационной комиссии!

**Вашему вниманию представляется выпускная квалификационная
работа студента _____
на тему: «Выявление закономерностей во множествах данных на основе
моделей дисперсионного анализа»**

Актуальность темы исследования заключается в том, что многомерный статистический анализ (МСА) представляет собой самостоятельное направление статистической теории. Интерес к нему связан с широкими возможностями в отображении и моделировании реальных объектов и явлений, имеющих многопризнаковую природу. Многопризнаковая природа различных классов объектов ведет к тому, что их анализ и отображение в моделях возможны при учете всех наиболее существенных характеристик.

Необходимо отметить, что многомерный статистический анализ является логическим развитием традиционной статистики. Однако при анализе многомерных данных используют и методы, иногда не имеющие четкой теоретической трактовки в терминах проверки гипотез, построения доверительных интервалов и т.д.

В основе многомерного анализа лежит предварительная обработка данных, включающая описательную статистику, визуальные средства анализа и формулировку гипотез о законах распределения и статистических закономерностях. Главной же задачей является анализ структуры выборочных совокупностей и выявление закономерностей.

Одним из мощных и широко распространенных методов является дисперсионный анализ. Более того, он занимает особое место в статистических исследованиях, поскольку практически все модели анализа данных опираются на изучение вариации признаков и численные оценки этой вариации (дисперсии). В задачах поиска закономерностей в массивах данных дисперсионный анализ применяется для исследования влияния одной или

нескольких качественных переменных (факторов) на одну (и более) количественную переменную (отклик).

Несмотря на глубокую проработанность теоретических обоснований и широкое использование в прикладных задачах, методы дисперсионного анализа остаются актуальными в современных исследованиях, использующих процедуры статистического анализа больших массивов данных.

Таким образом, вышеперечисленные обстоятельства свидетельствуют об актуальности темы исследования.

Целью выпускной квалификационной работы является изучение методов и моделей дисперсионного анализа и их применение в анализе многомерных данных.

Цель предопределила задачи исследования:

- систематическое изложение основных положений, задач и методов дисперсионного анализа и их использования на различных этапах статистической обработки данных;
- изучение процедур одномерного и многомерного дисперсионного анализа и их иллюстрация с помощью специально разработанных примеров;
- исследование прикладных задач выявления закономерностей в наборах данных с использованием моделей дисперсионного анализа и интерпретация полученных результатов.

Структура выпускной квалификационной работы включает введение, три раздела, которые подразделяются на пункты, заключение и список использованной литературы.

В первом разделе дано систематическое изложение моделей и методов дисперсионного анализа: приведены общая задача дисперсионного анализа и общая линейная модель. Среди задач одномерного дисперсионного анализа рассмотрены подробно однофакторная и двухфакторная модели, модель

анализа с повторными измерениями и основы однофакторного рангового анализа.

Второй раздел посвящен задачам многомерного дисперсионного анализа: приводится многомерная обобщенная линейная модель и критерии проверки многомерной гипотезы.

В третьем разделе рассмотрен ряд задач исследования с использованием различных моделей дисперсионного анализа на множестве данных об успеваемости обучающихся, а также по результатам опроса обучающихся для выявления факторов, влияющих на удовлетворенность содержанием программы обучения.

В выпускной квалификационной работе приведено систематическое изложение основ дисперсионного анализа как метода выявления закономерностей в массивах данных различной природы. Выявлена одна из проблем многофакторного дисперсионного анализа – реализация программы наблюдения.

В теоретической части работы приведена общая линейная модель для одномерного анализа ANOVA и обобщенная линейная модель многомерного анализа MANOVA. Приведены программы статистических процедур проверки одномерных и многомерных гипотез, перечислены основные критерии, представлены структуры итоговых таблиц анализа и правила получения статистических выводов, что представляет интерес для практических задач анализа данных.

В прикладной части работы представлены этапы и результаты исследований фактического материала, касающегося образовательной области. Первое исследование имело целью анализ итогов промежуточной аттестации (показатели успеваемости и качества) по различным направлениям подготовки и четырем курсам программы бакалавриата. Примененные модели одномерного однофакторного и двухфакторного анализа в целом не выявили существенности различия средних показателей успеваемости и качества по группам факторов направления и курса.

Второе исследование проведено по итогам опроса обучающихся по программам магистратуры с целью выявления средних оценок удовлетворенности содержанием обучения. Подготовленные анкеты опроса и его результаты послужили статистическим материалом для анализа. Результаты применения двух одномерных и многомерной модели анализа в целом не выявили существенного различия оценок удовлетворенности программой обучения по категориям обучающихся, совмещающих учебу с работой, как и по показателям их успеваемости.

Сценарии анализа данных и полученные результаты могут использоваться в учебном процессе как примеры проведения исследований с использованием моделей дисперсионного анализа.

Таким образом, цель выпускной квалификационной работы - изучение методов и моделей дисперсионного анализа и их применение в анализе многомерных данных, достигнута.

Благодарю за внимание, готов(а) ответить на Ваши вопросы.