

Учреждение образования “Белорусский государственный экономический университет”

УТВЕРЖДАЮ

Ректор Учреждения образования
“Белорусский государственный
экономический университет”

_____ В.Н.Шимов

“27” _____ 06. _____ 2014 г.

Регистрационный № УД 1599-14 /баз.

**КОМПЬЮТЕРНЫЕ ТЕХНОЛОГИИ АНАЛИЗА
ДАНЫХ В СОЦИОЛОГИИ**

Учебная программа учреждения высшего образования по учебной
дисциплине для специальности
1-23 01 05 «Социология»

СОСТАВИТЕЛИ:

Романова С.П., заведующая кафедрой экономической социологии Учреждения образования «Белорусский государственный экономический университет», кандидат социологических наук, доцент;

Бородачева Е.М., ассистент кафедры экономической социологии Учреждения образования «Белорусский государственный экономический университет»

РЕЦЕНЗЕНТЫ:

Лашук И.В., старший научный сотрудник сектора социокультурных инноваций государственного научного учреждения «Институт социологии Национальной академии наук Беларуси», кандидат социологических наук;

Кузьменко Т.В., старший научный сотрудник сектора социологии села государственного научного учреждения «Институт социологии Национальной академии наук Беларуси», кандидат социологических наук, доцент.

РЕКОМЕНДОВАНА К УТВЕРЖДЕНИЮ:

Кафедрой экономической социологии Учреждения образования «Белорусский государственный экономический университет»
(протокол № 10 от 07.05. 2014 г.);

Научно-методическим советом Учреждения образования «Белорусский государственный экономический университет»
(протокол № 5 от 18.06. 2014 г.)

Ответственный за редакцию: Бородачева Е.М.

Ответственный за выпуск: Бородачева Е.М.

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Учебная программа предназначена для преподавания спецкурса «Компьютерные технологии анализа данных в социологии» студентам специальности «Социология», специализации – «Экономическая социология» Института социально-гуманитарного образования Белорусского государственного экономического университета в рамках цикла дисциплин специальности.

Базовая учебная программа дисциплины «Компьютерные технологии анализа данных в социологии» разработана в соответствии с действующими образовательным стандартом и типовым учебным планом по специальности 1-23 01 05 «Социология».

Для проведения анализа массива данных социологи используют большое число различных математических методов, позволяющих полно и всесторонне анализировать собранную информацию. В современной социологии для этой цели активно применяются компьютерные программы математико-статистической обработки данных. Курс «Компьютерные технологии анализа данных в социологии» посвящен изучению и отработке на практике технологии компьютерной обработки результатов экспериментальных социологических исследований.

Цели курса: является формирование у студентов теоретических знаний и практических навыков по применению современных пакетов программ для обработки социологических данных в социологии. В процессе изучения курса студенты овладевают практическими навыками в использовании компьютеров на различных этапах проведения социологического исследования.

Задачи курса:

- сформировать навыки работы с базовым программным обеспечением, используемым в научных исследованиях;
- ознакомить студентов с основными программными средствами, используемыми на сегодняшний день в социологии;
- обучить принципам формирования матрицы данных;
- обучить процедурам первичной обработки данных прикладного социологического исследования с помощью электронной таблицы Excel и пакета программ SPSS;
- обучить технике расчета одномерных распределений, получения группировок признаков, представления данных социологического исследования в графическом виде с использованием электронной таблицы Excel и пакета программ SPSS;

- ознакомить студентов с основами методами многомерного статистического анализа данных;
- ознакомить студентов с основными положениями факторного анализа и кластерного анализа;
- рассмотреть способы реализации многомерного статистического анализа с помощью пакета программ SPSS.

Полученные навыки могут использоваться социологами для выполнения следующих работ: планирование исследования, сбор данных, формирование матрицы данных, пригодной для дальнейшего анализа с помощью одной из систем статистического анализа (например, электронной таблицы Excel и пакета программ SPSS), проведение первичной обработки, визуализация полученных результатов, анализ данных, а также проведение многомерного анализа.

В результате изучения учебной дисциплины студенты должны:

знать:

- этапы и процедуры первичной обработки данных прикладного социологического исследования;
- специфику практического применения непараметрических методов при анализе социологических данных;
- принципы построения индексов и снижения размерности пространства признаков;
- основные виды многомерного статистического анализа (регрессионный анализ, дисперсионный анализ, факторный анализ, кластерный анализ, дискриминантный анализ).

уметь:

- создавать матрицу данных и вводить данные прикладного социологического исследования в пакет SPSS;
- выполнять процедуры описательной статистики в пакете SPSS и электронной таблице Excel;
- осуществлять визуальное представление результатов социологического исследования;
- использовать непараметрические методы при анализе социологических данных;
- строить индексы и выполнять процедуру снижения размерности пространства признаков.

Связь с другими дисциплинами. Данная базовая учебная программа связана с блоками знаний, которые приобретает студент, изучая следующие дисциплины: «Общая социология», «Прикладная статистика в социологии»,

«Информационные технологии», «Методология и методы социологического исследования», «Прикладная статистика в социологии», «Основы информатики».

Структура курса. Курс «Экономическая социология» состоит из трех разделов: «Первичная обработка прикладного социологического исследования», «Практическое применение непараметрических методов при анализе социологических данных», «Многомерный статистический анализ».

Всего часов по дисциплине 138. Из них аудиторных 68, в том числе 30 часов лекционных занятий, 38 часов семинарских занятий.

Формы контроля знаний. Основной рекомендуемой формой текущего контроля знаний являются выступления на семинарских занятиях и выполнение самостоятельных работ. Рекомендуемая итоговая форма контроля знаний - зачет.

ПРИМЕРНЫЙ ТЕМАТИЧЕСКИЙ ПЛАН

№ тем	Наименование разделов и тем	Количество часов		
		всего	лекции	семинары
<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
Раздел I	Первичная обработка прикладного социологического исследования			
1.1	Этапы и процедуры первичной обработки данных прикладного социологического исследования	4	2	2
1.2	Отбор и модификация данных в пакете SPSS	4	2	2
1.3	Описательная статистика в пакете SPSS и электронной таблице Excel	4	2	2
1.4	Визуальное представление результатов социологического исследования	6	2	4
	Итого по разделу:	18	8	10
Раздел II	Практическое применение непараметрических методов при анализе социологических данных			
2.1	Практическое применение пакета SPSS при выявлении различий в уровне исследуемого признака	6	2	4
2.2	Практическое применение пакета SPSS при оценке достоверности сдвига в значениях исследуемого признака	6	2	4
2.3	Практическое применение пакета SPSS при выявлении различий в распределении признака	6	2	4
	Итого по разделу:	18	6	12
Раздел III	Многомерный статистический анализ			
3.1	Введение в многомерный статистический анализ данных	2	2	
3.2	Построение индексов и снижение размерности пространства признаков	4	2	2

3.3	Регрессионный анализ	4	2	2
3.4	Дисперсионный анализ	4	2	2
3.5	Факторный анализ	4	2	2
3.6	Кластерный анализ	4	2	2
3.7	Дискриминантный анализ	4	2	2
3.8	Оформление аналитической записки по результатам статистической обработки данных	6	2	4
	Итого по разделу:	34	16	18
	ВСЕГО:	64	30	38

СОДЕРЖАНИЕ УЧЕБНОГО МАТЕРИАЛА

РАЗДЕЛ 1. ПЕРВИЧНАЯ ОБРАБОТКА ДАННЫХ ПРИКЛАДНОГО СОЦИОЛОГИЧЕСКОГО ИССЛЕДОВАНИЯ

Тема 1.1 Этапы и процедуры первичной обработки данных прикладного социологического исследования

Подготовка данных прикладного социологического исследования к обработке как этап социологического исследования. Основные процедуры этапа.

Кодирование информации. Определение состава и типов переменных в процессе кодирования. Способы кодирования закрытых, многовариантных и открытых вопросов. Кодирование пропущенных значений.

Матрица данных типа «объект-признак». Принципы формирования матрицы данных. Создание макета ввода данных. Ошибки ввода данных и их предотвращение.

Виды обработки информации. Преимущества и основные направления использования компьютеров в социологии. Программное обеспечение статистического анализа социологических данных. Понятие о пакетах прикладных программ, их состав. Достоинства и недостатки пакетов. Обзор отечественных и зарубежных пакетов прикладных программ.

Тема 1.2 Отбор и модификация данных в пакете SPSS

Выбор объектов. Извлечение случайной выборки. Сортировка объектов. Разделение объектов на группы.

Вычисление новых переменных. Перекодировка значений. Вычисление новых переменных в соответствии с определенными условиями. Агрегированные данные. Ранговые преобразования.

Коррекция выборки при отсутствии репрезентативности.

Тема 1.3 Описательная статистика в пакете SPSS и электронной таблице Excel

Частотный анализ. Форматы частотных таблиц. Вывод статистических характеристик.

Таблицы сопряженности. Коэффициенты корреляции. Меры связанности для переменных с номинальной шкалой. Меры связанности для переменных с порядковой шкалой. Ранговые коэффициенты корреляции по Спирману и Кендалу. Частичная корреляция. Внутриклассовый коэффициент корреляции.

Проверка закона распределения.

Тема 1.4 Визуальное представление результатов социологического исследования

Использование мастера диаграмм электронной таблицы Excel для построения стандартных графиков. Создание рисунка на основе ячеек, диаграммы или объекта.

Построение графиков с использованием статистического пакета SPSS. Диаграммы (столбчатые, линейчатые, круговые, биржевые, коробчатые, рассеяния), гистограммы.

Картограммы, схемограммы.

РАЗДЕЛ 2. ПРАКТИЧЕСКОЕ ПРИМЕНЕНИЕ НЕПАРАМЕТРИЧЕСКИХ МЕТОДОВ ПРИ АНАЛИЗЕ СОЦИОЛОГИЧЕСКИХ ДАННЫХ

Тема 2.1 Практическое применение пакета SPSS при выявлении различий в уровне исследуемого признака

Сравнение двух независимых выборок (U-тест по методу Манна и Уитни, тест Мозеса, тест Колмогорова-Смирнова, тест Уалда-Вольфовица). Сравнение более чем двух независимых выборок (H-тест по методу Крускала и Уоллиса, медианный тест).

Тема 2.2 Практическое применение пакета SPSS при оценке достоверности сдвига в значениях исследуемого признака

Сравнение двух зависимых выборок (тест Уилкоксона, знаковый тест, тест хи-квадрат по методу Мак-Немара). Сравнение более чем двух зависимых выборок (тест Фридмана, W Кендала, Q Кохрана).

Тема 2.3 Практическое применение пакета SPSS при выявлении различий в распределении признака

Тест Колмогорова-Смирнова для проверки формы распределения. Биномиальный тест. Анализ последовательностей.

РАЗДЕЛ 3. МНОГОМЕРНЫЙ СТАТИСТИЧЕСКИЙ АНАЛИЗ

Тема 3.1 Введение в многомерный статистический анализ данных

Понятие, цели и задачи многомерного анализа социологических данных. Классификация методов многомерного статистического анализа.

Понятие многомерного пространства признаков. Интерпретация коэффициентов корреляции между признаками как косинуса угла между соответствующими осями пространства. Независимость признаков как ортогональность соответствующих осей. Представление объекта из выборки в виде точ-

ки в пространстве признаков. Координаты объекта в многомерном пространстве.

Матрица корреляций. Использование матрицы корреляций для описания структуры пространства переменных.

Понятие меры сходства между объектами в многомерном пространстве признаков. Виды мер сходства.

Понятие расстояния между объектами. Матрица расстояний. Меры расстояния (многомерное евклидово расстояние, расстояние Хемминга, расстояние Махаланобиса, расстояние, вычисляемое на основе коэффициента корреляции). Достоинства и недостатки мер расстояния.

Тема 3.2 Построение индексов и снижение размерности пространства признаков

Индекс как обобщенный показатель, сформированный из исходных переменных посредством логических и математических операций. Соотношение понятия индекса и индикатора. Классы индексов. Построение индексов как метод решения проблемы измерения. Этапы конструирования индексов.

Семантический дифференциал как метод количественного и качественного индексирования значения объектов. Сущность, цели и техника метода семантического дифференциала. Достоинства и недостатки метода.

Тема 3.3 Регрессионный анализ

Регрессионный анализ как статистический метод, позволяющий прогнозировать значения одной переменной на основании значения другой переменной.

Парная линейная регрессия. Диаграмма рассеяния, режим «подсолнухи». Требования к исходным данным регрессионного анализа. «Нормализация» исходных данных. Построение регрессионного уравнения. Интерпретация коэффициентов. Проверка статистических гипотез о значимости коэффициентов регрессионного уравнения.

Определение силы связи между признаками в линейной регрессионной модели. Анализ, полученной регрессионной модели

Тема 3.4 Дисперсионный анализ

Дисперсионный анализ как метод статистического анализа, изучающий влияние одной или нескольких независимых переменных на изменчивость одной зависимой переменной. Требования к исходным данным дисперсионного анализа.

Одномерный дисперсионный анализ. Команды для проведения одномерного дисперсионного анализа. Чтение и интерпретация итоговых таблиц одномерного дисперсионного анализа. Парные межгрупповые сравнения с помощью Post Hoc гипотез, получение однородных подмножеств.

Тема 3.5 Факторный анализ

Метод главных компонент: основные положения, геометрические и линейные модели, основные условия. Требования, предъявляемые к главным компонентам. Этапы реализации метода главных компонент. Матрица нагрузок главных компонент. Интерпретация нагрузок как коэффициентов корреляции между исходными переменными и главными компонентами. Интерпретация главных компонент в качестве латентных переменных.

Факторный анализ: основные положения, требования к исходным данным, проблемы использования в социологии. Этапы реализации факторного анализа. Представление факторной структуры. Факторные нагрузки и их интерпретация. Матрица отсортированных факторных нагрузок и ее использование при интерпретации факторов.

Тема 3.6 Кластерный анализ

Использование кластерного анализа в социологических и маркетинговых исследованиях. Предпосылки к применению кластерного анализа, теоретические аспекты объединения отдельных наблюдений в кластеры. Требования к исходным данным для получения статистически устойчивой многомерной модели. Подготовка данных к кластерному анализу. Реализация кластерного анализа с помощью команд пакета SPSS PC. Построение дендрограммы, выбор оптимального числа кластеров, интерпретация полученных кластеров.

Тема 3.7 Дискриминантный анализ

Использование дискриминантного анализа в социологических и маркетинговых исследованиях. Предпосылки к применению дискриминантного анализа, теоретические аспекты построения дискриминантной функции. Требования к исходным данным для получения статистически устойчивой многомерной модели. Подготовка данных к дискриминантному анализу. Реализация дискриминантного анализа с помощью команд пакета SPSS. Получение и описание коэффициентов

Тема 3.8 Оформление аналитической записки по результатам статистической обработки данных

Общие требования к написанию аналитической записки по итогам проведенного статистического анализа данных. Структура аналитической записки. Основные разделы статистической записки. Использование таблиц и графиков. Преобразование рабочих таблиц, полученных в ходе статистической обработки данных в презентационные таблицы. Система ссылок и приложений аналитической записки.

ИНФОРМАЦИОННО-МЕТОДИЧЕСКАЯ ЧАСТЬ ЛИТЕРАТУРА

Основная

1. Бююль, А. SPSS: Искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей: Пер. с нем. / А. Бююль, А. Цёфель П. – СПб: ООО «ДиаСофтЮП», 2001. – 608 с.
2. Ковалёва, Г.Д. Анализ социологических данных с применением статистического пакета SPSS: Учебно-методическое пособие / Г.Д. Ковалёва, П.С. Ростовцев. – Новосибирск: НГУ, 2002. – 160 с.
3. Добренков, В.И. Методы социологического исследования: Учебник / В.И. Добренков, А.И. Кравченко. – М.: Инфра-М, 2009. – 768 с.
4. Сидоренко, Е.В. Методы математической обработки в психологии / Е.В. Сидоренко. – СПб: ООО «Речь», 2003. – 350 с.
5. Симчера, В.М. Методы многомерного анализа статистических данных: Учебное пособие / В.М. Симчера. – М: Финансы и статистика, 2008. – 400 с.

Дополнительная:

1. Агабекян, Р.Л., Математические методы в социологии: Анализ данных и логика вывода в эмпирическом исследовании: Учебное пособие для вузов / Р.Л. Агабекян, М.М. Кириченко, С.В. Усатилов. – Ростов-на-Дону: Феникс, 2005. – 192 с.
2. Алексахин, С.В. Прикладной статистический анализ данных. Теория. Компьютерная обработка. Области применения. Учебно-практическое пособие для вузов. В 2-х книгах / С.В. Алексахин, А.В. Балдин, В.В. Криницин и др. – М.: Приор, 1998. – 352 с.
3. Гайдышев, И. Анализ и обработка данных: специальный справочник / И. Гайдышев. – СПб.: Питер, 2001. – 752 с.
4. Девятко, И.Ф. Методы социологического исследования / И.Ф. Девятко. – Екатеринбург: Издательство Уральского университета, 1998. – 208 с.
5. Дюран, Б. Кластерный анализ: Пер с англ. / Б. Дюран, П. Одел. – М.: Статистика, 1977. – 128 с.
6. Жуковская, В.М. Факторный анализ в социально-экономических исследованиях / В.М. Жуковская, И.Б. Мучник. – М.: Статистика, 1976. – 154 с.
7. Коцюбинский, А. Excel для менеджера и экономиста в примерах / А. Коцюбинский, С. Грошев – ГроссМедиа, 2007. – 304 с.
8. Леонов, Н. Математические методы анализа нечисловых баз данных и их применение в социологии / Н. Леонов, В. Студент, Д. Давыденко. – Минск: ФУАинформ, 2003. – 112 с.
9. Окунь, Я. Факторный анализ / Я. Окунь. – М.: Статистика, 1974. – 107 с.

10. Резник, А.Д. Книга для тех, кто не любит статистику, но вынужден ею пользоваться / А.Д. Резник. – СПб.: Речь, 2008. – 264 с.
11. Суходольский, Г.В. Основы математической статистики для психологов: Учебное пособие / Г.В. Суходольский. – СПб.: изд.-во Санкт-Петербургского университета, 1998. – 439 с.
12. Ядов, В.А. Стратегия социологического исследования. Описание, объяснение, понимание социальной реальности / В.А. Ядов. – 7-е изд. – М.: Добросвет, 2003. – 596 с.