

Программа учебной дисциплины «Регрессионный анализ данных в SPSS»

Утверждена

Академическим советом ООП

Протокол № от «__» ____ 20__ г.

Автор	Меликян Алиса Валерьевна
Число кредитов	2
Контактная работа (час.)	32
Самостоятельная работа (час.)	44
Курс	любой
Формат изучения дисциплины	Очный (full time)

I. ЦЕЛЬ, РЕЗУЛЬТАТЫ ОСВОЕНИЯ ДИСЦИПЛИНЫ И ПРЕРЕКВИЗИТЫ

Основные цели освоения дисциплины «Регрессионный анализ данных в SPSS»:

- формирование у слушателей целостного представления о возможностях проведения регрессионного анализа данных в SPSS;
- развитие практических навыков построения регрессионных моделей в SPSS.

В результате освоения дисциплины студент должен:

знать:

- основные задачи и методы регрессионного анализа;
- существующие в программе SPSS возможности для проведения регрессионного анализа за данных.

уметь:

- применять изученные методы регрессионного анализа к решению различных социально-экономических задач;
- выбрать подходящий метод регрессионного анализа данных в зависимости от особенностей данных и исследовательской задачи;
- интерпретировать результаты анализа в SPSS и представлять их в доступном для широкой аудитории виде;
- понимать результаты регрессионного анализа данных, приведённые в ведущих рецензируемых журналах социально-экономической направленности.

Владеть навыками:

- проведения регрессионного анализа данных в SPSS;
- составления аналитического отчёта по результатам проведённого регрессионного анализа данных.

Изучение дисциплины «Регрессионный анализ данных в SPSS» базируется на следующих дисциплинах:

- математическая статистика и теория вероятностей;
- информатика в объёме средней школы.

Для освоения учебной дисциплины студенты должны владеть следующими знаниями и компетенциями:

- знать основы математической статистики;

- знать основы теории вероятностей;
- обладать навыками работы с компьютерными программами.

Основные положения дисциплины должны быть использованы в дальнейшем при изучении следующих дисциплин:

- Научно-исследовательский семинар.

II. СОДЕРЖАНИЕ УЧЕБНОЙ ДИСЦИПЛИНЫ

Тема 1. Линейная регрессия.

Цели применения регрессионного анализа. Основные понятия и задачи регрессионного анализа. Простая линейная регрессия. Множественная линейная регрессия. Выбор зависимых и независимых признаков. Оценка параметров регрессионной модели. Свойства коэффициента детерминации. Интерпретация регрессионных коэффициентов. Мультиколлинеарность. Гетероскедастичность. Ограничения линейной регрессии. Регрессионная модель с использованием фиктивных переменных. Диагностика регрессионной модели.

Тема 2. Логистическая регрессия.

Бинарная логистическая регрессия. Мультиномиальная логистическая регрессия. ROC-кривая. Probit и Logit модели. Порядковая регрессия. Оценка параметров регрессионной модели.

Тема 3. Приближение с помощью кривых. Нелинейная и взвешенная регрессия.

Аппроксимация зависимостей с помощью кривых. Нелинейная регрессия.

Тема 4. Анализ временных рядов.

Стационарные и нестационарные временные ряды. АРПСС (Бокс и Дженкинс). Прогнозирование значений на будущие периоды. Авторегрессионные модели, интегральные модели и модели скользящего среднего.

Тема 5. Анализ панельных данных.

Преимущества и проблемы использования панельных данных. Классификация моделей панельных данных. Методы оценки регрессий по панельным данным. Модели с детерминированными и случайными эффектами.

III. ОЦЕНИВАНИЕ

Предусмотрены следующие формы контроля знаний студентов:

Тип контроля	Форма контроля	Параметры
Текущий	Домашнее задание	Провести анализ данных и подготовить отчет по результатам.
Текущий	Контрольная работа	Работа на компьютере.
Итоговый	Экзамен	Работа на компьютере. Оценка результатов работы в течение 3-х дней.

На текущем контроле студент должен продемонстрировать навыки анализа данных с использованием регрессионного метода анализа данных в программе SPSS.

На итоговом контроле студент должен продемонстрировать навыки самостоятельного поиска подходящего метода анализа данных, который наиболее подходит для решения поставленного исследовательского вопроса, интерпретации и представления результатов анализа, формулировки выводов на основе проведённого анализа данных.

Оценки по всем формам текущего контроля выставляются по 10-ти балльной шкале.

Преподаватель оценивает работу студентов на семинарских занятиях: активность при ответе на вопросы преподавателя, правильность выполнения заданий на семинарах. Оценки за работу на семинарских занятиях преподаватель выставляет в рабочую ведомость. Накопленная оценка по 10-ти балльной шкале за работу на семинарских занятиях определяется перед итоговым контролем - $O_{аудиторная}$.

Преподаватель оценивает самостоятельную работу студентов: правильность выполнения домашнего задания. Накопленная оценка по 10-ти балльной шкале за домашнее задание определяется перед итоговым контролем – $O_{дом. задание}$.

Преподаватель осуществляет текущий контроль знаний студентов в форме контрольной работы в конце 3-го модуля – $O_{к/р}$.

Способ округления накопленной оценки текущего контроля: в пользу студента.

Результирующая оценка выставляется по следующей формуле:

$$O_{результ.} = 0,3 \cdot O_{экзамен} + 0,3 \cdot O_{к/р} + 0,2 \cdot O_{дом. задание} + 0,2 \cdot O_{аудиторная},$$

где $O_{экзамен}$ – оценка за работу, выполненную на письменном экзамене:

Способ округления результирующей оценки: в пользу студента.

На передаче студенту не предоставляется возможность получить дополнительный балл для компенсации оценки за текущий контроль.

На экзамене студент может получить дополнительную практическую задачу, которая оценивается в 1 балл. Таким образом, результирующая оценка, получаемая на передаче, выставляется по формуле

$$O_{результ.} = (0,3 \cdot O_{экзамен} + 0,3 \cdot O_{к/р} + 0,2 \cdot O_{дом. задание} + 0,2 \cdot O_{аудиторная}) + O_{доп. вопрос}$$

IV. ПРИМЕРЫ ОЦЕНОЧНЫХ СРЕДСТВ

Вопросы для оценки качества освоения дисциплины в рамках текущего контроля студентов:

- 1) Свойства коэффициента детерминации, скорректированный коэффициент детерминации.
- 2) Выбор зависимых и независимых признаков для различных регрессионных моделей.
- 3) Оценка значимости уравнения регрессии и отдельных коэффициентов.
- 4) Проблема мультиколлинеарности регрессоров.
- 5) Использование фиктивных переменных.
- 6) Оценивание параметров логистических регрессионных моделей.
- 7) Проверка нормальности распределения ошибок модели.
- 8) Анализ предсказанных моделью значений.

- 9) Принятие решения о включении и исключении переменных из модели.
- 10) Получение уравнения нелинейной регрессии.

Пример домашнего задания

Необходимо сформировать файл с данными в SPSS, содержащий не менее 10 переменных с разными типами шкал и не менее 100 наблюдений. Данные могут быть взяты из любых открытых источников, например, <http://www.hse.ru/rlms>, <http://sophist.hse.ru/>.

Нужно провести мини-исследование, в котором необходимо применить все знания и навыки, полученные в ходе курса:

- Провести описательный анализ данных (частотные таблицы, статистические показатели, графики),
- Рассчитать новые переменные на основе существующих,
- Исследовать взаимосвязи (таблицы сопряженности, корреляции),
- Построить регрессионные модели и оценить их.

По результатам анализа нужно подготовить доклад. В докладе должны быть приведены исследовательские вопросы и гипотезы, а также основные результаты исследования.

V. РЕСУРСЫ

V.1 Основная литература

1. SPSS: искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей: Пер. с нем./ Ахим Бююль, Петер Цефель – СПб.: ООО «ДиасофтЮП», 2002.
2. Введение в эконометрику: учебник для вузов / К. Доугерти; Пер. с англ. О. О. Замкова, Е. Н. Лукаша, О. Ю. Шибалкина; Науч. ред. пер. О. О. Замков. – Изд. 3-е. – М.: ИНФРА-М, 2009. – 464 с. – (Сер. "Университетский учебник") . ISBN 978-5-16-003640

V.2 Дополнительная литература

1. Крыштановский, А. О. Анализ социологических данных с помощью пакета SPSS: учеб. пособие для вузов / А. О. Крыштановский. – М.: ГУ-ВШЭ, 2006. – 281 с. ISBN 5-7598-0373-5.
2. Таганов, Д. SPSS: Статистический анализ в маркетинговых исследованиях / Д. Таганов. – СПб.: Питер, 2005. – 190 с. ISBN 5-469-00640-9.
3. Наследов, А. SPSS 19: профессиональный статистический анализ данных / А. Наследов. – СПб.: Питер, 2011. – 399 с. - НП. - ISBN 978-5-459-00344-4.
4. Connolly Paul. Quantitative Data Analysis in Education : A Critical Introduction Using SPSS [Электронный ресурс] Routledge, 2007. ProQuest Ebook Central. Режим доступа: <https://ebookcentral.proquest.com/lib/hselibrary-ebooks/detail.action?docID=324865>.

V.3 Программное обеспечение

	Наименование	Условия доступа
п/п		
	IBM SPSS Statistics Premium	<i>Из внутренней сети университе-</i>

1.	Campus Edition	<i>та (договор)</i>
----	----------------	---------------------

V.4 Профессиональные базы данных, информационные справочные системы, интернет-ресурсы (электронные образовательные ресурсы)

п/п	Наименование	Условия доступа
<i>Профессиональные базы данных, информационно-справочные системы</i>		
1.	OECD iLibrary	<i>Из внутренней сети университета (договор)</i>
2.	World Bank e-Library	<i>Из внутренней сети университета (договор)</i>
3.	Единый архив экономических и социологических данных	URL: sophist.hse.ru
<i>Интернет-ресурсы (электронные образовательные ресурсы)</i>		
1.	Открытое образование	URL: https://openedu.ru/

V.5 Материально-техническое обеспечение дисциплины

Учебные аудитории для лекционных занятий по дисциплине обеспечивают использование и демонстрацию тематических иллюстраций, соответствующих программе дисциплины в составе:

- ПЭВМ с доступом в Интернет (операционная система, офисные программы, антивирусные программы);
- мультимедийный проектор с дистанционным управлением.

Учебные аудитории для практических занятий по дисциплине оснащены компьютерами, с возможностью подключения к сети Интернет и доступом к электронной информационно-образовательной среде НИУ ВШЭ.