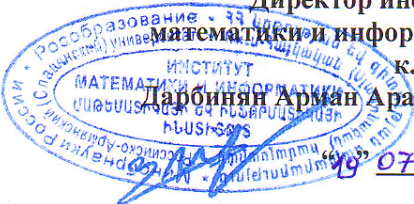


ГОУ ВПО РОССИЙСКО-АРМЯНСКИЙ УНИВЕРСИТЕТ

Составлен в соответствии с государственными требованиями к минимуму содержания и уровню подготовки выпускников по направлению 01.03.02 Прикладная математика и информатика и Положением «Об УМКД РАУ».

УТВЕРЖДАЮ:
Директор института
математики и информатики
к.ф.-м.н.,
Дарбинян Арман Араикович
07 2023г.



Институт Математики и информатики

Кафедра: Математической кибернетики

Автор(ы): к.ф.-м.н., доцент Петросян Артак Норикович

УЧЕБНО-МЕТОДИЧЕСКИЙ КОМПЛЕКС

**Дисциплина: Б1.В.ДВ.03.02 Специальный курс МК
«Регрессионные модели»**

Направление: «Прикладная математика и информатика» 01.03.02

ЕРЕВАН

1. Аннотация

Регрессионный анализ - один из основных концепций и методов исследования и количественной интерпретации закономерностей в современной математической статистике. Действенность данного метода базируется на построении регрессионных моделей, популярность которых объясняется следующими причинами:

- относительная простота регрессионных моделей и соответствующего математического аппарата;
- богатство интерпретации регрессионных моделей;
- применимость регрессионного анализа практически к любым экспериментальным данным (типа прямоугольных таблиц, содержащих зарегистрированные значения независимых и зависимых переменных);
- большая потребность в статистической обработке массивов данных (как с целью свертки, так и для извлечения из них дополнительной информации).

Многие вычислительные процедуры в регрессионном анализе автоматизированы, но основные практические аспекты применения регрессионного анализа еще не формализованы. Дело в том, что многие ключевые моменты построения и применения регрессионных моделей (в первую очередь, собственно, выбор вида модели; интерпретация модели; проблемы выбросов и влиятельных наблюдений и т.д.) требуют не только чисто статистического анализа экспериментальных данных, но и учета трудно формализуемой содержательной информации, относящейся к изучаемой задаче.

2. Цели и задачи дисциплины

Целью данного курса является ознакомление студентов с практическими приемами использования регрессионного анализа. В курсе предусматриваются методические рекомендации для практического применения регрессионного анализа.

Задачи курса – формирование у студентов умения:

- анализа качества регрессионного уравнения;
 - выбора «наилучшего» уравнения регрессии;
- владения:
- навыками проверки значимости (надежности) параметров уравнения регрессии;
 - навыками проверки согласия модели с экспериментальными данными (адекватности модели);
 - другими навыками проверки качества регрессионной модели;
 - навыками интерпретации регрессионной модели.

Специальный курс "Регрессионные модели" разработан для выполнения курсовых работ по дисциплине "Статистическое моделирование с использованием регрессионного анализа" для студентов направления подготовки "Математика и прикладная математика", а также студентов иных технических специальностей, изучающих основы машинного обучения.

3. Трудоемкость дисциплины и виды учебной работы по учебному плану

Виды учебной работы	Всего часов	Количество часов по семестрам							
		1 сем. м.	2 сем.	3 сем.	4 сем.	5 сем.	6 сем.	7 сем. м.	8 сем.
1	2	3	4	5	6	7	8	9	10
1. Общая трудоемкость изучения дисциплины по семестрам, в т. ч.:	36						36		
1.1. Аудиторные занятия, в т. ч.:	36						36		
1.1.1. Лекции	36						36		
1.1.2. Практические занятия, в т. ч.	0						0		
3. Самостоятельная работа, в т. ч.:									
4. Кредиты	2						2		
4. Форма итогового контроля: Экзамен/Зачет	зачет						зачёт		

4. Распределение объема дисциплины по темам и видам учебной работы

Разделы и темы дисциплины	Всего (ак. часов)	Лекции (ак. часов)	Практ. занятия (ак. часов)	Семинары (ак. часов)	Лабор. (ак. часов)	Другие виды занятий (ак. часов)
1	2	3	4	5	6	7
Модуль 1.						
Введение						
Раздел 1. Основные понятия. Предварительный анализ экспериментальной информации.	18		18			
Тема 1. Последовательность действий при построении регрессионной модели.	2		2			
Тема 2. Набор эффектов. Вид модели.	2		2			

Тема 3. Формализация априорной информации.	2		2			
Тема 4. Графический анализ экспериментальных данных.	2		2			
Тема 5. Выбор вида модели. Линейная (однофакторная, многофакторная) модель.	2		2			
Тема 6. Метод наименьших квадратов. Предпосылки Гаусса-Маркова.	2		2			
Тема 7. Проверка гипотезы о значимости параметров модели.	2		2			
Тема 8. Проверка согласия модели с экспериментальными данными (адекватности модели).	2		2			
Модуль 2.						
Раздел 2. Анализ качества модели.	18		18			
Тема 9. Анализ остатков. Коэффициент детерминации.	4		4			
Тема 10. Влиятельные наблюдения.	2		2			
Тема 11. Мультиколлинеарность.	2		2			
Тема 12. Прогноз значения откликов в области эксперимента.	4		4			
Тема 13. Другие способы проверки качества модели.	4		4			
Тема 14. Выводы.	2		2			
ИТОГО	36		36			

5. Содержание

5.1. Основные вопросы дисциплины.

В курсе приводятся методические указания по получению качественных хорошо интерпретируемых регрессионных моделей. Указанные методические приемы используются на практике при применении регрессионного анализа. Главные рассматриваемые вопросы: выбор модели; анализ качества модели; прогнозная (предсказывающая) способность модели; возможные недостатки регрессионных моделей и способы их выявления.

5.2. Содержание разделов и тем дисциплины

Введение.

Раздел 1. Основные понятия. Предварительный анализ экспериментальной информации: определены основные понятия регрессионного анализа, приведено описание этапов предварительного анализа экспериментальных данных, где заложены основы для успешного построения подходящей регрессионной модели.

Тема 1. Последовательность действий при построении регрессионной модели: описана поэтапная процедура построения адекватной экспериментальным данным регрессионной модели.

Тема 2. Набор эффектов. Вид модели: рассматривается круг (набор) исследуемых признаков (величин, переменных) - какие факторы и какие выходы исследуются; область существования факторов; число уровней и интервал варьирования каждого фактора в эксперименте; описаны различные классы моделей.

Тема 3. Формализация априорной информации: перечислены, систематизированы и формализованы известные из теоретических предпосылок и эмпирических исследований; формы взаимосвязи (взаимодействия) между различными факторами и/или различными откликами.

Тема 4. Графический анализ экспериментальных данных: наглядная (графическая) иллюстрация данных; график точечного рассеяния; график однофакторной зависимости.

Тема 5. Выбор вида модели: предварительный выбор вида модели осуществляется на основе априорной информации. В частности, выбирают и строят линейную модель, как правило, в следующих случаях: при малом количестве опытов (например, если количество опытов примерно равно количеству факторов, то это - отсеивающий эксперимент, когда, требуется по ограниченному количеству опытов выбрать факторы, оказывающие наибольшее влияние на отклик); при узких интервалах варьирования факторов; при априорном предположении о линейном характере взаимосвязи отклика, и факторов; в случаях, когда однофакторные зависимости отклика, полученные при графическом анализе экспериментальных данных имеют вид, близкий к прямой линии.

Тема 6. Метод наименьших квадратов. Предпосылки Гаусса-Маркова: приводятся ряд подходов (в том числе – метод наименьших квадратов) для оценки параметров (коэффициентов) линейной регрессионной модели; приводятся критерия оптимальности оценок; в условиях Гаусса-Маркова устанавливается оптимальность МНК-оценок.

Тема 7. Проверка гипотезы о значимости параметров модели: в этой теме изложена методика анализа степени воздействия (значимости) отдельно взятого фактора модели только в статистическом смысле (не касаясь интерпретации модели); оцениваются значимости отдельных параметров (коэффициентов) выбранных факторов применением критерия Стьюдента.

Тема 8. Проверка согласия модели с экспериментальными данными (адекватности модели):

применение критерия Фишера для проверки гипотезы об адекватности модели; оценка дисперсии воспроизводимости; случай неадекватности модели.

Раздел 2. Анализ качества модели.

Тема 9. Анализ остатков:

проверка нарушения основных предположений; выбросы; коэффициент детерминации; скорректированный коэффициент детерминации

Тема 10. Влиятельные наблюдения:

сила, теснота влияния факторов на результат (отклик); корреляция; множественная корреляция; частная корреляция.

Тема 11. Мультиколлинеарность:

случай сильной взаимосвязи между эффектами, вошедшими в модель, не позволяющей точно оценить их раздельное влияние на отклик (коэффициенты множественной корреляции эффектов в модели по абсолютной единице превышают 0.95).

Тема 12. Прогноз значения откликов в области эксперимента:

предсказывающая способность модели; проверка расчетных значений отклика вместе с 95 %-ными доверительными интервалами.

Тема 13. Другие способы проверки качества модели: метод “складного ножа”; метод "перепроверки"; метод бутстреп.

Тема 14. Выводы: перечислены основные возможные недостатки регрессионных моделей и способы их выявления.

6. Распределение весов по модуля и формам контроля

Формы контролей	Веса форм текущих контролей в результирующих оценках текущих контролей			Веса форм промежуточных контролей в оценках промежуточных контролей			Веса оценок промежуточных контролей и результирующих оценок текущих контролей в итоговых оценках промежуточных контролей			Веса итоговых оценок промежуточных контролей в результирующей оценке промежуточных контролей	Веса результирующей оценки промежуточных контролей и оценки итогового контроля в результирующей оценке итогового контроля
	M1 ¹	M2	M3	M1	M2	M3	M1	M2	M3		
Вид учебной работы/контроля											
Контрольная работа						1					
Тест											
Курсовая работа											
Лабораторные работы											
Письменные домашние задания			1								
Реферат											
Эссе											
<i>Другие формы (Указать)</i>											
<i>Другие формы (Указать)</i>											

¹ Учебный Модуль

Весы результирующих оценок текущих контролей в итоговых оценках промежуточных контролей										0.4		
Весы оценок промежуточных контролей в итоговых оценках промежуточных контролей										0.6		
Вес итоговой оценки 1-го промежуточного контроля в результирующей оценке промежуточных контролей												
Вес итоговой оценки 2-го промежуточного контроля в результирующей оценке промежуточных контролей												
Вес итоговой оценки 3-го промежуточного контроля в результирующей оценке промежуточных контролей											1	
Вес результирующей оценки промежуточных контролей в результирующей оценке итогового контроля												0.4
Экзамен/зачет (оценка итогового контроля)												0.6 (Экзамен/Зачет)
	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$	$\Sigma = 1$

7. Вопросы для самоконтроля

- 1 В чем заключается анализ качества модели?
- 2 Что делать, если модель незначима?
- 3 Что такое «адекватность модели»?
- 4 Может ли регрессионная модель быть значима, но не адекватна?
Адекватна, но не значима?
- 5 Что делать, если модель не адекватна?
- 6 В чем заключается анализ остатков?
- 7 Что такое «выбросы»? Почему нежелательно наличие выбросов в данных?
- 8 Что такое «влиятельные наблюдения»?
- 9 Что такое «мультиколлинеарность»?

8. Учебно-методическое обеспечение дисциплины.

8.1. Основная литература

1. Дрейпер Н., Смит Г. Прикладной регрессионный анализ. В 2-х книгах. М.: Финансы и статистика, т.1 - 1986, т.2 - 1987
2. Себер Дж. Линейный регрессионный анализ. М.: Мир, 1980
3. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика. Основы моделирования и первичная обработка данных. М.: Финансы и статистика, 1983

4. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика. Исследование зависимостей. М.: Финансы и статистика, 1985

8.2. Дополнительная литература

5. Вучков И, Бояджиева Л., Солаков Е. Прикладной линейный регрессионный анализ. М. :Финансы и статистика, 1987

6. Демиденко Е.З. Линейная и нелинейная регрессии. М.: Финансы и статистика, 1981

7. Петрович М.Л. Регрессионный анализ и его математическое обеспечение. М.: Финансы и статистика, 1982

Учебная программа одобрена кафедрой Математической кибернетики.

Зав. кафедрой: Арамян Р.Г



(подпись)