**Правительство Российской Федерации**

**Федеральное государственное автономное образовательное учреждение высшего профессионального образования**

**"Национальный исследовательский университет**

**"Высшая школа экономики"**

Департамент социальных наук

**Программа дисциплины**

Базовые методы анализа данных и работа со статистическими пакетами

для направления

44.06.01 Образование и педагогические науки,

Профиль 13.00.01 Общая педагогика, история педагогики и образования

подготовки научно-педагогических кадров в аспирантуре

Автор программы: Хавенсон Т.Е., [tkhavenson](mailto:tkhavenson@hse)@hse,ru

Москва, 2014

*Настоящая программа* ***не*** *может быть использована другими подразделениями универ- ситета и другими вузами без разрешения кафедры-разработчика программы.*

**1 Область применения и нормативные ссылки**

Настоящая программа учебной дисциплины устанавливает минимальные требования к знаниям и умениям аспиранта и определяет содержание и виды учебных занятий и отчетности.

Программа предназначена для преподавателей, ведущих данную дисциплину, учебных ассистентов и аспирантов направления 44.06.01 Образование и педагогические науки, обучающихся в Аспирантской школе по образованию и изучающих дисциплину «Базовые методы анализа данных и работа со статистическими пакетами».

Программа разработана в соответствии с:

* Образовательным стандартом НИУ ВШЭ;
* Образовательной программой по направлению 44.06.01 Образование и педагогические науки
* Учебным планом подготовки аспирантов по направлению 44.06.01 Образование и педагогические науки, профиль 13.00.01 Общая педагогика, история педагогики и образования, утвержденным в 2014 г.

**2 Цели освоения дисциплины**

Целями освоения дисциплины «Базовые методы анализа данных и работа со статистическими пакетами» являются

* сформировать у аспирантов теоретические представления об основных современных методах анализа данных в социальных науках;
* выработать навыки практического применения методов, как к самостоятельно собираемым данным, так и к базам данных;
* выработать у аспирантов представления о том, какие теоретические модели заложены в различных методах анализа данных;
* сформировать умение сопоставлять эти модели с задачами конкретного исследования и правильно выбирать метод в соответствии с его целями, задачами, гипотезами и имеющимися данными.

**3 Компетенции обучающегося, формируемые в результате освоения дисци плины**

В результате освоения дисциплины аспирант должен:

*Знать*

* основные теоретические представления о современных методах анализа данных в психологии и исследованиях в области образования
* правила применения современных методов анализа данных в социальных науках
* основные методы одномерной статистики и алгоритмы их реализации в па- кете SPSS
* основные методы многомерной статистики и алгоритмы их реализации в пакете SPSS
* основные принципы визуализации получаемой в ходе анализа данных ин- формации
* основные источники информации, необходимые для реализации исследований (базы данных, тематические веб-сайты, главные научные журналы по данной тематике)
* принципы написания аналитических текстов

*Уметь*

* + - * реализовывать основные одномерные и многомерные методы анализа данных
      * анализировать специфику использования методов математики и статистики для изучения психологических и образовательных явлений
      * интерпретировать результаты анализа данных, полученных в ходе исследований в психологии и в области образования
      * ставить задачи для анализа данных в той или иной ситуации в зависимости от типа данных и от исследовательской задачи
      * учитывать ограничения различных методов анализа данных, оценивать качество (валидность и надежность) полученной эмпирической информации
      * представлять результаты исследований для разных аудиторий слушателей критически анализировать информационные источники, научные тексты, результаты других исследований
      * участвовать в проектных формах работы
      * реализовывать самостоятельные аналитические проекты
      * подкреплять свою точку зрения знаниями, полученными на основе статистического анализа существующих или собранных специально для этих целей данных
      * работать в пакете для статистической обработки данных SPSS и специализированная программа для обработки данных международных сравнительных исследований в области образования IDB Analyzer (IEA)

*Иметь навыки (приобрести опыт)*

* + - * работы в программной оболочке для обработки и анализа информации на компьютере – SPSS
      * работы в программной оболочке для обработки данных международных сравнительных исследований в области образования IDB Analyzer (IEA)
      * расчета простых статистических показателей вручную
      * практического применения методов к самостоятельно собираемым данным
      * практического применения методов к существующим базам данных
      * участия в проектных формах работы
      * реализации самостоятельных аналитических проектов
      * написания аналитических текстов с применением результатов анализа социологических данных

В результате освоения дисциплины аспирант осваивает следующие компетенции:

|  |  |  |  |
| --- | --- | --- | --- |
| Компетенция | Код по ФГОС/ НИУ | Дескрипторы – основные признаки освоения (показатели достижения результата) | Формы и методы обучения,  способствующие формированию и развитию компетенции |
| Способность собирать, анализировать, обрабатывать и хранить данные в соответствии с общепринятыми научными и этическими стандартами | УК-4 | Использует программу SPSS, демонстрирует умение работать с базами международных исследований, проводит статистический анализ в соответствие со спецификой исследований TIMSS, PIRLS, PISA | Работа в SPSS, работа с базами данных международных исследований (TIMSS, PIRLS и PISA) |
| Способность к выполнению междисциплинарных исследований в области социально-психологических, социокультурных, социально-экономических аспектов образования, социализации, взросления | ПК-2 | Использует, демонстрирует, владеет,  применяет, представляет связи, распознает, обосновывает, интерпретирует оценивает | Работа в SPSS.  Работа над домашними заданиями, включающими необходимость не только проанализировать данные, но и дать содержательную интерпретацию. |
| Способность самостоятельно осуществлять научно-исследовательскую деятельность в области образования с использованием современных методов исследования и информационно-коммуникационных технологий | ОПК-1 | Распознает, использует, демонстрирует, владеет, применяет, представляет связи, распознает, обосновывает. | Работа над домашними заданиями, включающими необходимость не только проанализировать данные, но и дать содержательную интерпретацию Чтение научных статей по результатам эмпирических исследований.  Написание реферата. |
| Способность представлять результаты своих исследований на международных научных конференциях | ОПК-4 | Дает определение, воспроизводит,  распознает, использует, демонстрирует, владеет, применяет, представляет связи, распознает, обосновывает, интерпретирует, оценивает | Задачи на освоение методов  одномерной и многомерной статистики на базах данных реально проведенных исследований.  Выступление с докладами по выполненным домашним задания и другим формам проектной работы  Анализ данных (готовых баз). |

**4 Место дисциплины в структуре образовательной программы**

Настоящая дисциплина относится к циклу дисциплин по выбору и читается на 1-м году обучения.

Для освоения учебной дисциплины, аспиранты должны владеть следующими знаниями и компетенциями:

 Школьный уровень владения математикой

 Параллельное освоение курса по методам исследования

**5 Тематический план учебной дисциплины**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| № | Название темы | Всего часов по  дисциплине | Аудиторные часы | | Самостоятель-  ная работа |
| Лекции | Семинары и  практич. занятия |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 1. | Общие представления об анализе  данных | 5 | 1 | 0 | 4 |
| 2. | Основы теории вероятностей и  математической статистики | 8 | 1 | 1 | 6 |
| 3. | Основные виды распределений  вероятностей | 10 | 1 | 1 | 8 |
| 4. | Основы работы с пакетом SPSS | 12 | 0 | 4 | 8 |
| 5. | Методы описательной статисти-  ки. Визуализация данных. | 10 | 1 | 1 | 8 |
| 6. | Анализ связи между двумя при-  знаками | 10 | 1 | 1 | 8 |
| 7. | Общие принципы проверки ста-  тистических гипотез (параметри- чекие критерии) | 11 | 1 | 2 | 8 |
| 8. | Линейный регрессионный анализ | 11 | 1 | 2 | 8 |
| 9. | Логистическая регрессия | 11 | 1 | 2 | 8 |
| 10. | Дисперсионный анализ | 11 | 1 | 2 | 8 |
| 11. | Факторный анализ | 11 | 1 | 2 | 8 |
| 12. | Методы классификации призна-  ков. Кластерный анализ и дере- вья классификации | 15 | 2 | 1 | 12 |
| 13. | Основы непараметрической ста-  тистики | 15 | 2 | 1 | 12 |
| 14. | Написание отчета по проведен-  ному анализу данных | 12 | 1 | 2 | 10 |
|  | Итого: | 152 | 14 | 22 | 116 |

**6 Формы контроля знаний аспирантов**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Тип кон-  троля | Форма кон-  троля | 1 год | | | | Параметры \*\* |
| 1 | 2 | 3 | 4 |
| Текущий | Домашние  задания |  | \* | \* | \* | Корректность  решения задач. Корректность выполнения процедур в SPSS. Описание и интер- претация полученных результатов. |
| Итого-  вый | Зачет |  |  |  | \* | Защита проекта |

Домашние задания во 2-м модуле задаются на срок до следующего занятие. В 3-4 сроки каждый раз оговариваются отдельно. Если сдача задания просрочена на одно занятие, его вклад в итоговую оценку снижается на 50%. Если на два и более, то вес в оценке – 0.

**6.1 Критерии оценки знаний, навыков**

Реферат должен отражать способность к самостоятельному освоению новых методов ис- следования, изменению научного и научно-производственного профиля своей деятельности.

Домашнее задание: должно отражать способность использовать современные средства ИКТ для поиска и обработки информации, работы с базами данных профессиональной инфор- мации, оформить и представить результаты своей деятельности в виде отчета по научной и практической работе и презентации.

Оценки по домашним заданиям выставляются в 100-балльной шкале, что обусловлено особенностями работы в ЛМС.

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

Выдача заданий через ЛМС, частичная проверка работ и выставление оценок.

**Итоговая работа: исследовательский проект**

Для выполнения исследовательского проекта по курсу аспирантам предлагается на выбор несколько баз данных реальных исследований. Выбрав одну из них, необходимо сформулировать постановку содержательной цели исследования, определить задачи, сформулировать гипотезы, подобрать соответствующие методы анализа данных, отобрать необходимые данные, про- анализировать и проинтерпретировать их.

В работе обязательно должны быть задействованы все методы анализа данных, входящие в программу курса.

Методы анализа данных:

* Методы описательной статистики
* Коэффициенты парной связи
* Дисперсионный анализ
* Регрессионный анализ
* Факторный анализ
* Методы классификации

Возможные базы данных исследований:

Мониторинг образовательных траекторий аспирантов (ИРО ГУ-ВШЭ. Руководитель По- пов Д.С.)

TIMSS 1999, 2003, 2007, 2011. Доступны данные по анкетам школьников, учителей, ро- дителей, директоров школ для 4 и 8 классов школы. (IEA [http://timss.bc.edu](http://timss.bc.edu/))

PIRLS 2001, 2006. Доступны данные по анкетам школьников, учителей, родителей, ди- ректоров школ. (IEA [http://pirls.bc.edu](http://pirls.bc.edu/))

PISA 2000, 2003, 2006, 2009. Доступны данные по анкетам школьников, директоров школ. (OECD [http://www.pisa.oecd.org](http://www.pisa.oecd.org/))

Дополнительные материалы, примеры задач:

[http://www.hse.ru/jesda/mathbase](http://www.hse.ru/jesda/mathbase/)/

[http://www.hse.ru/jesda/mathbase/tasks/tasks\_psyc](http://www.hse.ru/jesda/mathbase/tasks/tasks_psycho)ho

**7 Содержание дисциплины**

**Тема 1. Общие представления об анализе данных**

Лекция.

Вводная тема, целью которой является общее погружение аспирантов в проблематику ис- следований в социальных науках в общем, и в психологии и образовании в частности. Обсуж- даются основные типы исследований, цели анализа данных. Соответствие определенных мето- дов анализа данных целям и задачам исследования. Разработка стратегии анализа данных. Про- цесс анализа данных в исследовании.

Основная литература

1. Gay L. R., Mills G.E., Airasian, P. Educational research: Competencies for analysis and applications. Upper Saddle River, NJ: Merrill Prentice Hall, 2009. Ch.1, 7-11

2. Tabachnick B.G. & Fidell L.S Using Multivariate Statistics. Pearson Education. Inc Boston, MA. 2007, ch.1,2

3. Учебно-методические материалы для преподавания курсов по применению количе- ственного инструментария к решению социально-экономических задач

[http://www.hse.ru/jesda/mathbase](http://www.hse.ru/jesda/mathbase/)/

**Тема 2. Основы теории вероятностей и математической статистики**

Классическое и статистическое определение понятия вероятности. Правила сложения и умно- жения вероятностей. Формула полной вероятности. Условная вероятность. Зависимость/независимость событий.

Основная литература:

* 1. Bluman A. Elementary statistics. McGraw-Hill, 2008
  2. Bluman A. Probability demystified. McGraw-Hill, 2005. Стр. 1-77

**Тема 3. Основные виды распределений вероятностей**

Случайные величины - дискретные и непрерывные. Общие представления о распределе- нии вероятностей. Нормальное распределение.

Произвольное и стандартное нормальное распределение. Стандартизация. Работа с таб- лицами нормального распределения. Квантили распределения.

Работа с таблицами других статистических распределений (распределение Стьюдента, хи-квадрат распределение, распределение Фишера), нахождение критических значений для проверки статистических гипотез.

1. Bluman A. Elementary statistics. McGraw-Hill, 2008.

2. Tabachnick B.G., Fidell L.S Using Multivariate Statistics. Pearson Education. Inc Boston, MA. 2007. Appendix A

3. Анастази А., Урбина С. Психологическое тестирование. СПб, Питер, 2003. Гл. 3, стр.

75 - 84.

4. Гласс Дж., Стэнли Дж. Статистические методы. Прогресс, 1976. Стр.91-98, 206-215

5. Толстова Ю.Н. Математико-статистические модели в социологии. М.: ИД ГУ-ВШЭ,

2007. Стр. 51-64, 71-94

Дополнительная литература:

1. Bluman A. Probability demystified. McGraw-Hill, 2005. Стр.

2. Kalid Azad An Intuitive Guide To Exponential Functions & e //

[http://betterexplained.com/articles/an-intuitive-guide-to-exponential-functions-](http://betterexplained.com/articles/an-intuitive-guide-to-exponential-functions-e/)e/

3. Kalid Azad Easy Permutations and Combinations //

[http://betterexplained.com/articles/easy-permutations-and-combinati](http://betterexplained.com/articles/easy-permutations-and-combinations/)ons/

4. Айвазян С. А., Мхитарян В. С. Теория вероятностей и прикладная статистика, ч.1.

М.: Юнити, 2001

5. Вадзинский Р.Н. Справочник по вероятностным распределениям. - СПб.: Наука,

2001. - 295с.

6. Тюрин Ю.Н., Макаров А.А., Симонова Г.И. Теория вероятностей: учебник для эко- номических и гуманитарных специальностей. М.: МЦНМО, 2009

**Тема 4. Основы работы с пакетом SPSS**

Семинар.

Начало работы в SPSS. Правила создания макета анкеты (опросного документа). Ввод данных.

Работа с переменными – кодирование, автоматическое и ручное, вычисление новых пе- ременных, свойства переменных. Функции Recode, Count, Compute и д.р..

Работа с данными – сортировка, отбор случаев, извлечение случайной выборки, агреги- рование, чистка данных. Функции Sort cases, Select cases, Aggregate, Split file и д.р..

Работа с файлами – слияние нескольких файлов, экспорт и импорт данных. Функции

Merge file и др.

1. Tabachnick B.G., Fidell L.S Using Multivariate Statistics. Pearson Education. Inc Boston, MA. 2007. Ch.1,2

2. Беликова А. Очистка персональных данных часть 1 // BaseGroup Labs [http://www.basegroup.ru/library/cleaning/person\_data\_pa](http://www.basegroup.ru/library/cleaning/person_data_part1/)rt1/

3. Бююль А., Цёфель П., SPSS версия 10. СПб.: ООО "ДиаСофтЮП", 2001. Главы 1, 3,

4. Стр. 14-18, 26-81.

4. Leech N.L., Barret K.C., Morgan G.A., SPSS for intermediate statistics. Lawrence Erlbaum

Associates, Publishers. London. 2005. Стр. 27-35

Дополнительная литература

Беликова А. Очистка персональных данных часть 2 // BaseGroup Labs [http://www.basegroup.ru/library/cleaning/person\_data\_pa](http://www.basegroup.ru/library/cleaning/person_data_part2/)rt2/

**Тема 5. Методы описательной статистики. Визуализация данных**

Одномерные частотные таблицы, абсолютные и относительные частоты (процент, доля), накопленная частота. Основные типы шкал и соответствующие им меры средней тенденции и меры разброса. Линейное и нелинейное преобразование шкал. Стены, станайны. Точечное и ин- тервальное оценивание.

Принципы графического представления данных. Наиболее популярные виды графиков:

гистограмма, диаграмма рассеивания, диаграмма «ящик с усами» и др.

Правила оформления результатов описательной статистики.

SPSS: Descriptive Statistics: Descriptives, Explore, Q-Q plots, Frequencies. Graphs

Основная литература

1. Thorkildsen T.A. Fundamentals of Measurement in Applied Research. Allyn & Bacon,

2004

2. Бююль А., Цёфель П., SPSS версия 10. СПб.: ООО "ДиаСофтЮП", 2001. Главы 6,9.

Стр. 91-103, 164-169

3. Малхотра Н. Маркетинговые исследования. М.: Вилямс, 2003. Глава 8. Стр. 317-323;

глава 15, стр. 552-561

4. Наследов А.Д. Математические методы психологического исследования. Анализ и интерпретация данных. Изд-во «Речь». СПб., 2006

5. Толстова Ю.Н. Измерение в социологии. Москва: КДУ, 2009

6. Kalid Azad How To Analyze Data Using the Average //

[http://betterexplained.com/articles/how-to-analyze-data-using-the-average](http://betterexplained.com/articles/how-to-analyze-data-using-the-average/)/

**Тема 6. Анализ связи между двумя признаками**

Таблицы сопряженности. Возможное содержание ячеек таблицы. Условные и безуслов- ные частоты. Повтор правила умножение вероятностей.

Коэффициенты парной связи для различных типов шкал. Критерий Хи-квадрат и осно- ванные на нем коэффициенты. Коэффициенты корреляции. Проверка значимости корреляцион- ной связи. Работа с таблицами множественных ответов.

Правила оформления результатов.

SPSS: Correlate, Crosstabs, Custom tables, Multiple Response

Основная литература

1. Gay L. R., Mills G.E., Airasian, P. Educational research: Competencies for analysis and applications. Upper Saddle River, NJ: Merrill Prentice Hall, 2009. Ch.12

2. Tabachnick B.G., Fidell L.S Using Multivariate Statistics. Pearson Education. Inc Boston, MA. 2007. Ch.3

3. Бююль А., Цёфель П., SPSS версия 10. СПб.: ООО "ДиаСофтЮП", 2001. Главы

11,12,15. Стр. 180-200, 207-219, 256-260

4. Малхотра Н. Маркетинговые исследования. М.: Вилямс, 2003. Главы 15, 17, стр. 562-

599, 640-648.

5. Наследов А.Д. Математические методы психологического исследования. Анализ и интерпретация данных. Изд-во «Речь». СПб., 2006

**Тема 7. Общие принципы проверки статистических гипотез (параметричекие критерии)**

Общие правила проверки статистических гипотез.

Алгоритмы проверки наиболее важных гипотез. Гипотезы о равенстве средних: тесты для одной выборки (z-test, t-test) и двух выборок (зависимые и независимые). Гипотеза о равен- стве долей. Проверка значимости коэффициента корреляции. Доверительный интервал и уро- вень значимости. Ошибки 1 и 2 рода. Правила оформления результатов.

SPSS: Compare means (One-Sample T Test, Independent-Samples T Test)

Основная литература

1. Bluman A.G. Elementary statistics. Brown Publishers, 2008, pp.300-337, 344-372, 419-422

2. Tabachnick B.G., Fidell L.S Using Multivariate Statistics. Pearson Education. Inc Boston, MA. 2007. Ch.3

3. Малхотра Н. Маркетинговые исследования. М.: Вилямс, 2003. Гл.15

4. Наследов А.Д. Математические методы психологического исследования. Анализ и интерпретация данных. Изд-во «Речь». СПб., 2006

5. Толстова Ю.Н. Математико-статистические модели в социологии. М.: ИД ГУ-ВШЭ,

2007. С.101-144

6. Stevens J.P. Applied multivariate statistics for the social sciences. 4th ed 2002. Chapter 1.2

Type I error, Type II error, and Power, pp.3-6; chapter 1.4 Statistical Significance versus

Practical Significance, pp.9-12

**Тема 8. Дисперсионный анализ**

Формальная модель заложенная в методе. Одномерный и многомерный дисперсионный анализ. Множественные сравнения. Интерпретация результатов.

Правила оформления результатов.

Основная литература

1. Bluman A.G. Elementary statistics. Wm. C. Brown Publishers, 2008

2. Tabachnick B.G., Fidell L.S Using Multivariate Statistics. Pearson Education. Inc Bos- ton, MA. 2007. Ch.3

3. Бююль А., Цёфель П., SPSS версия 10. СПб.: ООО "ДиаСофтЮП", 2001. Глава 17, стр. 323-346

4. Малхотра Н. Маркетинговые исследования. М.: Вилямс, 2003. Глава 16, стр. 604-

640.

5. Наследов А.Д. Математические методы психологического исследования. Анализ и интерпретация данных. Изд-во «Речь». СПб., 2006

**Тема 9. Линейный регрессионный анализ**

Цели применения регрессионных моделей. Парный и множественный линейный регрес- сионный анализ. Выбор зависимых и независимых признаков и оценка качества построенной модели. Интерпретация коэффициентов регрессии. Регрессия с фиктивными переменными. Ограничения линейной регрессии.

Правила оформления результатов.

Основная литература

1. Bluman A. Elementary statistics. McGraw-Hill, 2008 Ch. 10 Correlation and Regression

2. Tabachnick B.G., Fidell L.S Using Multivariate Statistics. Pearson Education. Inc Boston, MA. 2007. Ch.5

3. Малхотра Н. Маркетинговые исследования. М.: Вилямс, 2003. Гл.17, стр. 648-678.

4. Наследов А.Д. Математические методы психологического исследования. Анализ и интерпретация данных. Изд-во «Речь». СПб., 2006

5. Бююль А., Цёфель П., SPSS версия 10. СПб.: ООО "ДиаСофтЮП", 2001. Глава 16.

Стр. 269-302.

Дополнительная литература

Stanton J.M. Galton, Pearson, and the Peas: A Brief History of Linear Regression for Statistics Instructors // Journal of Statistics Education Volume 9, Number 3 (2001) – история изобретения ре- грессионного анализа.

**Тема 10. Логистическая регрессия**

Логистическая регрессия. Оценка качества полученной модели. Интерпретация получен- ных результатов.

Основная литература

1. Tabachnick B.G., Fidell L.S Using Multivariate Statistics. Pearson Education. Inc Boston, MA. 2007. Ch.12

2. Бююль А., Цёфель П., SPSS версия 10. СПб.: ООО "ДиаСофтЮП", 2001

3. Petersen T, A Comment on Presenting Results from Logit and Probit Models. American

Sociological Review, Vol. 50, No. 1 (Feb., 1985), pp. 130-131

Дополнительная литература

Garson G. D., Logistic Regression, [http://faculty.chass.ncsu.edu/garson/PA765/logist](http://faculty.chass.ncsu.edu/garson/PA765/logistic.htm)ic.htm (Рекомендации по использованию регрессионных моделей. Дэвид Гарсон. Библиотека Универ- ситета Северной Каролины).

**Тема 11. Факторный анализ**

Метод главных компонент. Цели применения метода. Алгоритм проведения анализа, требования к исходным данным, факторные нагрузки, вращение осей. Интерпретация результа- тов.

Основная литература

1. Малхотра Н. Маркетинговые исследования. М.: Вилямс, 2003. Глава 19. Стр. 717-741.

2. Бююль А., Цёфель П., SPSS версия 10. СПб.: ООО "ДиаСофтЮП", 2001. Глава 19.

Стр. 368-384.

3. Наследов А.Д. Математические методы психологического исследования. Анализ и интерпретация данных. Изд-во «Речь». СПб., 2006

4. Митина О.В., Михайловская И.Б. Факторный анализ для психологов. Уч.пос. М., УМК «Психология», 2001 [on-line[: http://psychlib.ru/mgppu/mit/MIT-001-](http://psychlib.ru/mgppu/mit/MIT-001-.HTM).HTM]

5. Tabachnick B.G., Fidell L.S Using Multivariate Statistics. Pearson Education. Inc Boston, MA. 2007. Ch.5

**Тема 12. Методы классификации признаков. Кластерный анализ и деревья классификации.**

Основные цели, решаемые кластерным анализом, сфера применения. Иерархический

кластерный анализ. Способы вычисления расстояний между объектами. Способы кластериза- ции.

Неиерархический кластерный анализ, метод k-средних. Совместное применение иерар- хических и неиерархических методов кластеризации.

Основная литература

Кластерный анализ

1. Rencher A.C. Methods of Multivariate Analysis. John Wiley & Sons. 2002. Ch. 14, p. 471

2. Бююль А., Цёфель П., SPSS версия 10. СПб.: ООО "ДиаСофтЮП", 2001. Глава 20.

Стр. 385-409

3. Малхотра Н. Маркетинговые исследования. М.: Вилямс, 2003. Глава 20. Стр. 747-771

4. Наследов А.Д. Математические методы психологического исследования. Анализ и интерпретация данных. Изд-во «Речь». СПб., 2006

Деревья классификации

5. Интерпретация и анализ данных в социологических исследованиях. ред. Андреенков

В.Г., Толстова Ю.Н. М.: Наука, 1987. Глава 6, стр. 136

6. Толстова Ю.Н. Анализ социологических данных. М.: Научный мир, 2000. Глава 2.5, стр. 256-269

**Тема 13. Основы непараметрической статистики**

Цели применения непараметрических методов. Работа с малыми выборками. Непараметрические критерии: критерии t-тест Манна-Уитни, W-тест Уилкоксона.

Основная литература

1. Наследов А.Д. Математические методы психологического исследования. Анализ и интерпретация данных. Изд-во «Речь». СПб., 2006

2. Bluman A.G. Elementary statistics. Brown Publishers, 2008

3. Бююль А., Цёфель П., SPSS версия 10. СПб.: ООО "ДиаСофтЮП", 2001

**Тема 14. Написание отчета по проведенному анализу данных**

Работа с окном выдачи результатов – SPSS Viewer. Редактирование таблиц, графиков. Экспорт объектов в приложения MS Office.

Общие принципы представления результатов применения статистических методов ана- лиза данных. Отбор релевантной информации. Визуализация информации.

Основная литература

1. Бююль А., Цёфель П., SPSS версия 10. СПб.: ООО "ДиаСофтЮП", 2001. Главы

7,8,10, стр. 104-144, 170-179.

2. Желязны Дж. Говори на языке диаграмм: Пособие по визуальным коммуникациям.

М., Манн, Иванов и Фербер, 2007

3. Малхотра Н. Маркетинговые исследования. М.: Вилямс, 2003. Глава 14. Стр.519-551

4. Bluman A.G. Elementary statistics. Brown Publishers, 2008

5. Gay L. R., Mills G.E., Airasian, P. Educational research: Competencies for analysis and applications. Upper Saddle River, NJ: Merrill Prentice Hall, 2009. Ch.21

**8 Образовательные технологии**

Разбор практических задач и кейсов на занятиях на данных международных образова- тельных исследованиях и других. Мастер-классы экспертов и специалистов

**9 Оценочные средства для текущего контроля и аттестации аспиранта**

**9.1 Тематика заданий текущего контроля**

Примерные вопросы/ задания для домашнего задания:

**1. В исследовании** "Мониторинг образовательных и трудовых траекторий выпускников школ и вузов" задавался вопрос "Сколько книг у Вас дома?". Распределение ответов представлено в таблице:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | | Frequency | Percent | Valid Percent | Cumulative Per-  cent |
| Valid | 0-10 книг | 144 | 4,1 | 4,1 | 4,1 |
| 11-25 книг | 311 | 8,9 | 8,9 | 13,1 |
| 26-100 книг | 949 | 27,3 | 27,3 | 40,4 |
| 101-200 книг | 809 | 23,3 | 23,3 | 63,7 |
| 201-500 книг | 716 | 20,6 | 20,6 | 84,3 |
| более 500 книг | 529 | 15,2 | 15,2 | 99,5 |
| нет ответа | 18 | ,5 | ,5 | 100,0 |
| Total | 3476 | 100,0 | 100,0 |  |

Найти:

А) Медиану

Б) Квартильный размах

В) Моду

**2. Тема:** описательная статистика

Найдите достаточно простую русскоязычную статью (отчет по описательному исследованию, например, [http://cim.hse.ru/unimonitorin](http://cim.hse.ru/unimonitoring)g), в которой изложены результаты исследования, включая описание процесса сбора и анализа данных и интерпретацию результатов.

Напишите комментарий к статье (1-2 страницы, до 4000 знаков). По примерному плану:

1. Название рецензируемого материала. Полное библиографическое описание.

2. Цель исследования, основной исследовательский вопрос или вопросы.

3. Методы сбора данных. Выборка.

4. Какие методы анализа данных применялись для получения результатов.

5. Основные полученные выводы и результаты.

6. Какой информации Вам не хватает в публикации.

7. Какие вопросы Вы задали бы авторам исследования.

8. Как бы Вы построили исследование с тем же исследовательским вопросом.

В выполненной работе необходимо указать либо полное библиографическое описание статьи, либо ссылку на сайт с материалами. Если это статья не из интернет-источника, желательно приложить текст статьи.

стране.

**3. Тема:** описательная статистика. **Работа выполняется** на базе TIMSS 2007 (по аспирантам) по вашей

Необходимо выбрать несколько переменных разного уровня измерения и провести анализ методами опи-

сательной статистики.

Построить частотное распределение, найти меры центральной тенденции и разброса, описать полученные результаты для нескольких переменных (как минимум по 1 переменной для каждой шкалы). Если для переменной возможно найти несколько показателей, запишите все.

Тема рефератов для каждого аспиранта утверждается преподавателем в индивидуальном порядке.

1. Linn R.L. Assessments and accountability. Educational researcher. 2000, vol.29, #2. Pp.4-16 (Boxplots)

2. Kloosterman P., Cougan M.C. Students' Beliefs about Learning School Mathematics // The Elementary School Journal, Vol. 94, No. 4, Mar., 1994. pp. 375-388 (Z-scores)

3. Dewey J.D., Montrosse B.E., Schröter D.C., Sullins C.D., Mattox J.R. Evaluator Competencies What’s Taught

Versus What’s Sought // American Journal of Evaluation. Volume 29, Number 3 September 2008. Рр. 268-287

4. Apodaca-Tucker M.T., Slate J.R. School-Based Management: Views from Public and Private Elementary School

Principals // Education Policy Analysis Archives. Volume 10, Number 23, May 2002

5. Bruno J.E. The Geographical Distribution of Teacher Absenteeism in Large Urban School District Settings: Im- plications for School Reform Efforts Aimed at Promoting Equity and Excellence in Education // Education Policy Analysis Archives. Volume 10 Number 32, July 2002

6. O'Leary M., Kellaghan T., Madaus G.F., Beaton A.E. Consistency of Findings Across International Surveys of Mathematics and Science Achievement: A Comparison of IAEP2 and TIMSS // Education Policy Analysis Ar- chives. 2001, Volume 9 Number 42 [http://epaa.asu.edu/ojs/article/view/43](http://epaa.asu.edu/ojs/article/view/434)4

7. Wiest D.J., Wong E.H., Cervantes J.M., Craik L., Kreil D.A. Intrinsic motivation among regular, special, and alternative education high school students // Adolescence. 2001. Vol. 36, Iss. 141; pp. 111-126

8. Basow S.A., Montgomery S. Student Ratings and Professor Self-Ratings of College Teaching: Effects of Gender and Divisional Affiliation // Journal of Personnel Evaluation in Education. Volume 18, Number 2 / May 2005, pp.

91–106

9. Ludlow L.H. A longitudinal approach to understanding course evaluations // Practical Assessment Research & Evaluation, 2005, Vol 10 No 1

[http://pareonline.net/pdf/v10n1.pd](http://pareonline.net/pdf/v10n1.pdf)f

10. Baumert J., Lüdtke O., Trautwein U., Brunner M. Large-scale student assessment studies measure the results of processes of knowledge acquisition: Evidence in support of the distinction between intelligence and student achievement // Educational Research Review. Volume 4, Issue 3, 2009, Pages 165-176

11. Kessler-Sklar S.L., Baker A.J.L. School district parent involvement policies and programs // Elementary School

Journal. 2000, vol.101, #1, pp.101-118.

12. Bolon C. Significance of Test-based Ratings for Metropolitan Boston Schools // Education Policy Analysis Ar- chives. 2001, Volume 9 Number 42

[http://epaa.asu.edu/ojs/article/view/37](http://epaa.asu.edu/ojs/article/view/371)1

Статья длинная. Можно выбирать сразу двоим.

13. Krueger A.B. Experimental Estimates of Education Production Functions // The Quarterly Journal of Economics, Vol. 114, No. 2 (May, 1999), pp. 497-532.

14. Folmer-Annevelink E., Doolaard S., Mascareño M., Bosker R.J. Class Size Effects on The Number and Types of Student-Teacher Interactions in Primary Classrooms // Journal of Classroom Interaction 2010. Vol. 45, Iss. 2; pp. 30-39.

**9.2 Вопросы для оценки качества освоения дисциплины**

*Вопросы приведены исключительно для ознакомления с темами курса, экзамен в форме ответов на теоретические вопросы не проводится*

1. Модели независимости, заложенных в коэффициентах, основанных на критерии хи-квадрат, коэффициентах ранговой корреляции, коэффициенте корреляции Пирсона

2. Т-Тесты :для независимых выборок, для парных выборок, одно выборочный (проверяемая стат.гипотеза. алгоритм проверки, содержательные выводы).

3. Дисперсионный анализ (проверяемая стат.гипотеза. алгоритм проверки, содержательные выводы).

4. Линейная регрессионная модель: оценка качества модели и значения коэффициентов регрессионного уравнения, ограничения

5. Линейная регрессионная модель: стандартизованные коэффициенты регрессионного уравнения

6. Бинарная и множественная логистическая регрессия: оценка качества модели, значение коэффициентов

7. Алгоритм иерархического агломеративного кластерного анализа

8. Кластерный анализ: алгоритм К-средних

9. Разведывательный факторный анализ. Метод главных компонент

**10 Порядок формирования оценок по дисциплине**

Оценка за текущий контроль в 1 и 2 модулях учитывает результаты аспиранта по следу- ющим компонентам:

Домашние задания – 5 баллов Проектная работа – 1 балл Контрольная работа – 4 балла.

*О12мод =* од.з.\*0,5 + опроект\*0,1+ ок.р.\*0,4

**3-4 модули :**

Оценка за текущий контроль в 1 и 2 модулях учитывает результаты аспиранта по следу- ющим компонентам:

Домашние и самостоятельные задания – 5,5 баллов

Контрольная работа – 2 балла

Подготовка доклада/рецензии по статье – 1 балл. Активность (доклады по д.з., ответы у доски) – 1,5 балл.

Текущий контроль:

*О34мод. =* остат.\*0,1 + од.з.\*0,55+ оакт.\*0,15 + ок.р.\*0,2

Экзаменационная работа – 4 балла.

*О34мод =* од.з.\*0,55 + оакт\*0,15+ ок.р.\*0,2 + ост\*0,1

Итоговая оценка: Оитоговая = О12мод.\*0,4 + О34мод.\*0,6

Домашние задания в 1-2 модулях задаются на срок до следующего занятие. В 3-4 моду- лях сроки каждый раз оговариваются отдельно. За опоздание сдачи работы применяются штра- фы. 1 и 2 модуль: 20% за каждый день просрочки

3 и 4 модуль: если сдача задания просрочена на одно занятие, его вклад в итоговую оценку снижается на 50%. Если на два и более, то вес в оценке – 0.

**11 Учебно-методическое и информационное обеспечение дисциплины**

**11.1 Базовый учебник**

1. Бююль А., Цёфель П., SPSS версия 10. СПб.: ООО "ДиаСофтЮП", 2001

2. Vogt W. P. Quantitative Research Methods for Professionals in Education and Other Fields.

Allyn & Bacon, 2006

3. Gay L. R., Mills G.E., Airasian, P. Educational research: Competencies for analysis and appli- cations. Upper Saddle River, NJ: Merrill Prentice Hall, 2009

Доступна электронная версия базового учебника.

**11.2 Основная литература**

1. Малхотра Н. Маркетинговые исследования. М.: Вилямс, 2003

2. Толстова Ю.Н. Математико-статистические модели в социологии. М.: ИД ГУ-ВШЭ,

2007

3. Bluman A. Elementary statistics. McGraw-Hill, 2008

4. Bluman A. Probability demystified. McGraw-Hill, 2005

5. Girden E.R. Evaluating research articles from start to finish. Sage Publications, 1996. 289 p.

6. Ho R. Handbook of Univariate and Multivariate Data Analysis and Interpretation with SPSS.

T&F, 2006

7. Tabachnick B.G., Fidell L.S Using Multivariate Statistics. Pearson Education, 2007

**11.3 Дополнительная литература**

1. Агабекян Р.Л., Кириченко М.М., Усатиков С.В. Математические методы в социологии.

Анализ данных и логика вывода в эмпирическом исследовании. Ростов н/Д.: Феникс,

2005

2. Айвазян С. А., Мхитарян В. С. Теория вероятностей и прикладная статистика, ч.1. М.: Юнити, 2001

3. Гуц А.К., Фролова Ю.В. Математические методы в социологии. М.: Изд-во ЛКИ, 2007

4. Желязны Дж. Говори на языке диаграмм: Пособие по визуальным коммуникациям. М., Манн, Иванов и Фербер, 2007

5. Интерпретация и анализ данных в социологических исследованиях. ред. Андреенков

В.Г., Толстова Ю.Н. М.: Наука, 1987

6. Крыштановский А.О. Анализ социологических данных с помощью пакета SPSS. М.: ИД ГУ ВШЭ, 2006

7. Толстова Ю.Н. Анализ социологических данных. М.: Научный мир, 2000

8. Толстова Ю.Н. Измерение в социологии. Москва: КДУ, 2009

9. Тюрин Ю.Н., Макаров А.А. Анализ данных на компьютере. – М.: ИД Форум, 2008

10. Тюрин Ю.Н., Макаров А.А., Симонова Г.И. Теория вероятностей: учебник для экономи- ческих и гуманитарных специальностей. М.: МЦНМО, 2009

11. Сидоренко Е.В. Методы математической обработки в психологии. Речь, 2007

12. Хили Дж. Статистика. Социологические и маркетинговые исследования. СПб.: Питер,

2005

13. [http://betterexplained.com/articles/how-to-analyze-data-using-the-average](http://betterexplained.com/articles/how-to-analyze-data-using-the-average/)/

14. Hair J.F. Jr, Black W.C., Babin B.J., Anderson R.E. Multivariate Data Analysis. Pearson. 2009

15. McMillan J.H., Wergin J.F. Understanding and Evaluating Educational Research. Pearson Ed- ucation. 2009

16. PISA 2003 Data Analysis Manual. SPSS Users

17. Thorkildsen T.A. Fundamentals of Measurement in Applied Research. Allyn & Bacon, 2004

18. Thorndike, R. M., Christ, T.M. Measurement and evaluation in psychology and education.

Pearson, 2009

19. Vogt W. P. Dictionary of Statistics & Methodology: A Nontechnical Guide for the Social Sci- ences. Sage Publications. 2005

**11.4 Справочники, словари, энциклопедии**

Wikipedia

**11.5 Программные средства**

Для успешного освоения дисциплины, аспирант использует следующие программные средства:

 SPSS

 IDB-Analyzer

**11.6 Дистанционная поддержка дисциплины**

ЛМС

**12 Материально-техническое обеспечение дисциплины**

Проектор, компьютерный класс