

Министерство образования и науки Российской Федерации
Федеральное агентство по образованию
Ярославский государственный университет им. П.Г. Демидова

О.В. Епархина

Математические методы обработки и анализа социологических данных

Учебное пособие

*Рекомендовано
Научно-методическим советом университета
для студентов специальности Социология*

Ярославль 2007

УДК 316:303.7

ББК С5в6я73

Е 65

Рекомендовано

*Редакционно-издательским советом университета
в качестве учебного издания. План 2007 года*

Рецензенты:

доктор технических наук, профессор А.А. Мурашов;
кафедра политологии ЯГТУ

Епархина, О.В. Математические методы обработки и анализа
Е 65 социологических данных : учеб. пособие / О.В. Епархина; Яросл.
гос. ун-т. – Ярославль : ЯрГУ, 2007. – 132 с.
ISBN 978-5-8397-0527-2

В пособии даны теоретические и практические аспекты использования математики в социологии, описаны конкретные методы анализа социологических данных. Представлены алгоритмы поиска связей между номинальными признаками (коэффициенты связи, многомерные отношения преобладания, сочетания независимых предикторов и т.п.).

Учебное пособие предназначено для студентов, обучающихся по специальности 020300 Социология (блок ЕН) очной формы обучения.

УДК 316:303.7

ББК С5в6я73

© Ярославский
государственный
университет
им. П.Г. Демидова, 2007
© О.В. Епархина, 2007

ISBN 978-5-8397-0527-2

Введение

Настоящая работа является учебным пособием по курсу «Математические методы обработки и анализа социологических данных», читаемому автором для студентов-социологов. Практически в каждом учебном заведении, готовящем социологов, преподается эта дисциплина, что обусловлено возрастанием значения математического аппарата в аналитической работе с данными как в рамках используемых ПК статпакетов, так и в традиционных формах.

В пособии разъясняется специфика анализа социологических данных и показана необходимость использования математики в социологии. В нем содержится описание методов анализа данных, характерных именно для социологии: методов т.н. описательной (дескриптивной) статистики – выборочного представления одномерного вероятностного распределения и расчета его основных параметров (мер средней тенденции и показателей разброса), простейших методов изучения связей между номинальными признаками, а также рассматриваются коэффициенты связи для ранговых признаков, элементы дисперсионного и факторного анализа и т.п.

В отечественной литературе еще в 1970 – 1980-х гг. было представлено много работ, предназначенных для изучения социологами математических методов, использующихся в решении социологических задач. Однако большинство этих работ изданы давно, недоступны студентам, некоторые методы описаны недостаточно подробно.

Приоритет в разработке рассматриваемых методов принадлежит западным ученым. Прежде всего речь идет о

работах А. Агрести, ставших классикой на Западе (логлинейные, логит-, пробит-модели, ряд моделей логистической регрессии, алгоритмы анализа отношений преобладания и т.д.), о монографиях Г. Аптона. В основу данного пособия положены исследования этих крупнейших специалистов, а также Ю.Н. Толстовой.

Предполагается, что студент, приступивший к изучению данного пособия, имеет элементарные знания из курсов по общей социологии, методике социологических исследований, математической статистике.

Содержание

Введение	3
Тема 1 Общие аспекты применения математических методов в социологическом анализе.....	5
1.1. Статистические закономерности в анализе социологической информации	5
1.2. Специфика математико-статистических методов применительно к социологической информации.....	11
1.3. Задачи математики применительно к социологической информации	13
1.4. Сложности использования математических методов в социологии	16
Тема 2 Общая характеристика процедуры анализа данных.....	21
2.1. Социологические данные	21
2.2. Общие принципы анализа данных.....	25
Тема 3 Анализ одномерных распределений	27
3.1. Необходимость анализа одномерных распределений в социологии	27
3.2. Меры средней тенденции	29
3.3. Дисперсия.....	36
3.4. Мера качественной вариации	37
3.5. Энтропийный коэффициент разброса.....	40
Тема 4 Типы шкал и методы анализа информации	43
4.1. Номинальная шкала	43
4.2. Ранговая шкала.....	46
4.3. Интервальная шкала.....	48

Тема 5 Анализ двумерных распределений	60
5.1. <i>Общая характеристика двумерных распределений.....</i>	60
5.2. <i>Показатели связи в двумерных распределениях</i>	63
Тема 6 Анализ связей между номинальными признаками	65
6.1. <i>Общая характеристика подходов к анализу номинальных данных.....</i>	65
6.2. <i>Анализ связей типа «признак – признак».....</i>	68
6.2.1. Коэффициенты связи, основанные на критерии хи-квадрат	68
6.2.2. Коэффициенты связи, основанные на моделях прогноза.....	75
6.2.3. Коэффициенты связи, основанные на понятии энтропии	80
6.2.4. Коэффициенты связи для четырехклеточных таблиц сопряженности	82
6.2.5. Многомерные отношения преобладаний	87
6.3. <i>Анализ связей типа «альтернатива – альтернатива»: ДА.....</i>	91
6.4. <i>Анализ связей типа «группа альтернатив – группа альтернатив».....</i>	94
6.4.1. Анализ фрагментов таблиц сопряженности.....	94
6.4.2. Методы поиска сочетаний значений независимых признаков (предикторов).....	101
6.5. <i>Анализ связей типа «признак – группа признаков».....</i>	106
6.5.1. Номинальный регрессионный анализ (НРА).....	106
6.5.2. Логит- и пробит-модели.....	118
Глоссарий.....	120
Вопросы к зачету	125
Библиография	127
Рекомендуемая зарубежная литература.....	129