

Министерство образования и науки Российской Федерации  
Федеральное агентство по образованию  
ГОУВПО  
«Амурский государственный университет»

**Е.П. Селькова**

**УЧЕБНО-МЕТОДИЧЕСКИЙ КОМПЛЕКС  
ПО ДИСЦИПЛИНЕ  
«Анализ данных в социологии»  
Учебное пособие**

Благовещенск, 2007

**Печатается по решению  
редакционно-издательского совета  
факультета социальных наук  
Амурского государственного  
университета**

Учебно-методический комплекс по дисциплине «Анализ данных в социологии» для студентов очной формы обучения по специальности 040201 «Социология». /Сост. Е.П. Селькова. Благовещенск: Амурский гос. ун-т, 2007

**Учебное пособие составлено для специальности «социология» 040201 и включает наименование тем, цели и содержание лекционных, семинарских и практических занятий; тестовые задания для контроля изученного материала; задания для самостоятельной работы; вопросы для итоговой оценки знаний; список рекомендуемой литературы.**

**Рецензент:**

© Амурский государственный университет

## СОДЕРЖАНИЕ

Введение	4
1. Цели и задачи дисциплины, ее место в учебном процессе	5
2. Содержание дисциплины	7
2.1 Учебно-методический план дисциплины	7
2.2 Наименование тем лекционных занятий, их содержание	8
2.3 Семинарские занятия	10
2.4 Самостоятельная работа	16
2.5 Вопросы для подготовки к зачету	16
2.6 Критерии оценки знаний студентов	18
2.7 Тест	19
3. Учебно-методические материалы по дисциплины	24
3.1 Рекомендуемая литература	24
3.2 Краткое содержание лекционного курса	25

## **ВВЕДЕНИЕ**

Анализ социологических данных является заключительным и одним из самых сложных этапов любого социологического исследования, так как он основывается на знании законов логики, тенденций развития социума и основ математической статистики.

Данный курс поможет студентам составить представление о стратегиях и логике анализа данных в социологии, а также об отдельных его методах. Помимо этого получить практические навыки и умения по применению математического аппарата для анализа социологической информации и построению адекватных выводов о социальной реальности.

Данное пособие составлено с учетом рекомендаций учебно-методического отдела АмГУ и включает следующие разделы:

- цели и задачи дисциплины, ее место в учебном процессе;
- содержание дисциплины;
- учебно-методические материалы по дисциплине.

## 1. ЦЕЛИ И ЗАДАЧИ ДИСЦИПЛИНЫ, ЕЕ МЕСТО В УЧЕБНОМ ПРОЦЕССЕ

Целью курса является формирование у студентов знаний, умений и навыков работы в русле адекватных представлений:

- 1) об основных этапах конкретного социологического исследования с акцентом на этапах подготовки к обработке (редактирование, кодирование, контроль качества данных, ввод их в ЭВМ), обработке на ЭВМ и анализа полученной в ходе полевого исследования социологической информации;
- 2) о методах и приемах превращения данной информационной базы (через ее обобщение, анализ, научную интерпретацию, а также прогнозирование социальных процессов) в действенную силу, способную эффективно служить различным аспектам управления / менеджмента;
- 3) о принципах формулирования выводов и практических рекомендаций, в т.ч. прогностического характера в отчетах, аналитических справках, научных публикациях представляющих собой форму подведения итогов данного социологического исследования;
- 4) об элементах структуры указанной социологической отчетности:
  - а) описание всех разделов программы КСИ;
  - б) изложение использованных концепций и методик;
  - в) фиксации и анализа полученных данных;
- 5) об особенностях применения проанализированных результатов конкретных социологических исследований в практике управленческой деятельности.

Задачи курса.

- 1) ознакомление студентов с основными элементами структуры эмпирико-прикладной социологии в целом и различных видов социологических исследований;
- 2) предоставление студентам знаний о методах подготовки к обработке, обработки и анализа социологической информации, полученной в виде определенных данных по ходу каждого социологического исследования;

- 3) формирование у студентов умений и навыков практической работы по адекватному применению соответствующих методов, техник и процедур при подготовке к обработке, обработке, анализе, научной интерпретации социологической информации, а также на этой основе – прогнозирования возможных социальных процессов, событий, явлений;
- 4) выявление на занятиях главных аспектов социологического отчета (аналитической справки, научной публикации) в форме обобщающих выводов, практических рекомендаций, прогнозов и т.д., а также подготовка подобной качественной социологической отчетности;
- 5) формирования представления у студентов об использовании результатов социологических исследований в практике менеджмента.

## 2. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

### 2.1 Учебно-методический план курса: «Анализ данных в социологии»

№	Лекции	Часы	№	Практические занятия//семинары	Часы	Самост. работа	Контроль	Всего
1	Подготовка первичной социологической информации к обработке на ЭВМ	2	1	Техника обработки социологической информации на ЭВМ	2	4	О.К	8
2	Понятие измерения в современной эмпирической социологии	2	2	Последовательность действий при качественно-количественном анализе социологических данных	2	4	О.К	8
3	Типологизация социологических данных (в аспекте КСИ)	2	3	Стратегия качественного анализа данных КСИ	2	4	О.К	8
4	Характерные черты и условия осуществления различных видов / методов анализа	4	4	Интерпретация данных КСИ	2	2	О.К	8
			5	Графическая интерпретация социолого-эмпирических зависимостей	2	2	О.К	4
5	Группировка социолого-исследовательских наблюдений как фактор элементарных понятий статистики	2	6	Средние величины и характеристика значений рассеяния признака: в аспекте анализа / интерпретации данных КСИ	2	6	О.К	10
6	Поиск взаимосвязи между переменными: в русле работы с итогами сбора социоинформации	2	7	Показатели колеблемости / вариации значений признаков. Статистические взаимосвязи и их социологический анализ	2	6	О.К	10
7	Социальный эксперимент как аналитический метод верификации научной гипотезы	2	8	Анализ данных повторных и сравнительных КСИ	2	6	О.К	10
8	Статистические методы обработки социологической информации	2	9	Многомерный анализ социологической информации	2	6	О.К	10
Всего	Лекции	18	Всего	Практические занятия// семинары	18	40	зачет	76

## **2.2 Наименование тем лекционных занятий, их содержание**

### **Тема 1 Подготовка первичной социологической информации к обработке на ЭВМ.**

1. Принципы кодирования закрытых / комбинированных вопросов.
2. Контроль качества анкет / бланков интервью, наблюдений и т.д.
3. Полнота заполнения опросников и др. видов инструментария.
4. Точность заполнения анкет респондентами.
5. Адекватность участия в интервью опрашиваемых.
6. Популярные и эффективные социологические программы ЭВМ.

### **Тема 2 Понятие измерения в современной эмпирической социологии.**

1. Определение и сущностные характеристики измерения.
2. Социальные показатели и их измерение.
3. Неоднозначность шкальных значений, допустимые преобразования показателей.
4. Типы шкал в аспекте измерения различных видов социальных показателей.
5. Социологические индикаторы, единицы измерения и социальные показатели.
6. Адекватность математических методов (для анализа полученных через измерение чисел).

### **Тема 3 Типологизация социологических данных (в аспекте КСИ).**

1. Классификация и группировка единиц социологической информации.
2. Простая и перспективная группировка социологических данных.
3. Эмпирическая типологизация как прием социологического анализа по относительному плану.
4. Теоретическая типологизация как обобщение признаков социальных явлений.
5. Объяснение и описание / интерпретация результатов КСИ.

**Тема 4 Характерные черты и условия осуществления различных видов / методов анализа.**

1. Анализ социологических данных как неотъемлемый компонент всех типов КСИ.
2. Социальное прогнозирование, планирование и управление в аспекте анализа результатов КСИ.
3. Социологический анализ и способы, стадии измерений.
4. Надежность инструмента измерения и его составляющие.
5. Степени устойчивости социальных объектов и процедура анализа.
6. Методы социологического анализа: генетический, исторический, сравнительный / сравнительно-исторический.

**Тема 5 Группировка социолого-исследовательских наблюдений как фактор элементарных понятий статистики.**

1. Ряды распределения и статистический ряд.
2. Сгруппированные данные и интервальный ряд.
3. Простые статистические таблицы и обобщение исходных данных.
4. Сложные статистические таблицы и осуществление группировки единиц совокупности по нескольким признакам.
5. Правила построения статистических таблиц.

**Тема 6 Поиск взаимосвязи между переменными: в русле работы с итогами сбора социоинформации.**

1. Перекрестная группировка и правила процентирувания.
2. Анализ взаимосвязи двух переменных с помощью контрольного фактора.
3. Метод обнаружения и интерпретирующей и объясняющей связи.
4. Матрица интеркорреляций переменных и факторная матрица.
5. Графики и диаграммы.

**Тема 7 Социальный эксперимент как аналитический метод верификации научной гипотезы.**

1. Логика экспериментального анализа как формы проверки объяснительной гипотетической модели.

2. Натурный, мыслительный и контролируемый виды социологического эксперимента.
3. Основные социолого-экспериментальные приемы выравнивания индивидуальных характеристик объектов КСИ.
4. Разновидности контролируемых натуральных экспериментов.
5. Модели мысленного, созидательного, проектирующего эксперимента и детерминационные отношения.

## **Тема 8 Статистические методы обработки социологической информации.**

1. Статистика как собрание и классификация количественной информации.
2. Факторный и корреляционный анализ.
3. Математическое моделирование как теоретические и эмпирические представления о свойствах социальной действительности.
4. Методы: монографический, парных сравнений, экспертных оценок.
5. Типологизация и квантификация.

### **2.3 Семинарские занятия**

#### **Семинар 1** Техника обработки социологической информации на ЭВМ.

1. Процедура получения социологической информации и ее обработки / анализа как составляющая КСИ.
2. Искажение и манипуляция социальной информацией.
3. Значение особенностей социологической информации.
4. Основные виды обработки социальной информации: оперативная, поэтапная, итоговая.
5. Способы и условия обработки социальной информации.
6. Социальная информация и многообразие ресурса общественного развития: аспект расширения специализаций в современной социологии.

#### **Вопросы // задания для самоконтроля:**

1. Для чего социологу нужна математика?
2. Что такое математическая статистика?
3. Назовите и охарактеризуйте основные типы эмпирических данных.

4. Какие достоинства и недостатки имеет каждый тип с точки зрения исследователя?
5. Назовите четыре типа показателей, которыми оперирует социолог.
6. Что такое маргинальные частоты, накопленные частоты?
7. Какие трудности создают при обработке данных пропущенные значения?
8. Основные программы статистической обработки социологических данных и их возможности.
9. Моделирование социальных явлений на ЭВМ.

**Семинар 2** Последовательность действий при качественно-количественном анализе социологических данных.

1. Цель КСИ как определитель уровня анализа социологических данных.
2. Стадии последовательных действий в процессе анализа эмпирико-социологических данных.
3. Общая логика анализа данных КСИ.
4. Принципиальный порядок аналитических мероприятий / шагов в условиях операций с данными КСИ.
5. Интерпретационная схема как фактор обеспечения убедительности и значимости результатов КСИ.

**Вопросы // задания для самоконтроля:**

1. С чем связано появление термина «анализ социологических данных»?
2. Что такое восходящая и нисходящая стратегия анализа данных?
3. Охарактеризуйте этап описания результатов качественно-количественного анализа социологических данных.
4. В чем отличие этапа «уплотнения» данных и этапа «углубления интерпретации»?
5. Охарактеризуйте этап прогноза качественно-количественного анализа социологических данных. Назовите методы прогнозирования.

**Семинар 3** Стратегия качественного анализа данных КСИ.

1. Соотношение и особенности применения количественного и качественного анализа данных КСИ в мировой социологической практике.

2. Недостатки количественных и преимущество качественных методов анализа данных КСИ.
3. Виды качественных методов социологического анализа: сущность, характерные черты, научные параметры.
4. Процесс и основные этапы реализации стратегии качественного анализа данных КСИ.
5. Типологизация качественных социолого-аналитических стратегий.

#### **Вопросы // задания для самоконтроля:**

1. Каковы предпосылки возникновения «восходящей» стратегии анализа данных?
2. Методы анализа нечисловых данных в социологическом исследовании.
3. Таблицы сопряженности, методы визуализации, анализ текстовых данных, открытых вопросов.

#### **Семинар 4 Интерпретация данных КСИ.**

1. Общие подходы к интерпретации социологических данных.
2. Процедура интерпретации полученных в ходе КСИ данных.
3. Выявление взаимосвязи признаков события / явления – процесса в аспекте верификации гипотез.
4. Метод исключения (в русле интерпретации данных КСИ).
5. Общие правила оформления результатов интерпретации данных КСИ.
6. Значение интерпретации данных КСИ в структуре / системе социологической практики.

#### **Вопросы // задания для самоконтроля:**

1. Почему гипотеза является основным методологическим инструментом анализа данных?
2. Как взаимосвязаны интерпретация социологических данных и тип исследования?
3. В чем суть внутреннего и внешнего соотнесения?
4. Влияние социолога на процесс объяснения результатов.
5. Логика построения отчета.

6. Для кого социолог пишет аналитические записки?
7. Социальные группы и организации, заинтересованные в результатах исследования. Способы представления результатов исследования общественности.

**Семинар 5** Графическая интерпретация социолого-эмпирических зависимостей.

1. Диаграммы и графики при изображении частотных распределений.
2. Круговые диаграммы.
3. Столбиковые диаграммы и виды интерпретационных рисунков.
4. Полигон и гистограмма.
5. Практическое применение социолого-интерпретационной графики: в отчетах и научных публикациях по итогам КСИ.

**Вопросы // задания для самоконтроля:**

1. Для каких социологических данных строят полигон и гистограмму?
2. Постройте гистограммы для метрических шкал.
3. Сделать презентацию результатов социологического исследования.

**Семинар 6** Средние величины и характеристика значений рассеяния признака: в аспекте анализа / интерпретации данных КСИ.

1. Основные этапы статистического анализа данных КСИ.
2. Получение обобщающих характеристик особенностей.
3. Виды средних величин и показателей колеблемости признака в математической статистике.
4. Среднее значение признака. Медиана, мода, среднее арифметическое.
5. Условия целесообразности использования различных типов средней величины. Форма распределения признака.

**Вопросы // задания для самоконтроля:**

1. В чем отличие средней арифметической от взвешенной средней арифметической?
2. Кумулята. Медиана и квартильный размах.

3. Построить по накопленной частоте гистограмму для порядковой шкалы и изобразить кумуляту и геометрически определить медиану и квартильный размах.
4. Меры средней тенденции.
5. Энтропия как мера качественной вариации.
6. Что такое индекс?

**Семинар 7** Показатели колеблемости / вариации значений признаков.  
Статистические взаимосвязи и их социологический анализ.

1. Виды показателей колеблемости как характеристики разброса значений признака.
2. Дисперсия и ее исчисление.
3. Среднее абсолютное исчисление как мера вариации и ее расчет.
4. Корреляционный и регрессивный анализ в эмпирической социологии.
5. Представление о форме, направлении и тесноте / плотности статистической взаимосвязи величин: в аспекте регрессионного анализа.

**Вопросы // задания для самоконтроля:**

1. Меры разброса признака.
2. Что такое мера связи и коэффициенты связи?
3. В чем отличие корреляционной связи от функциональной?
4. Как понимается статистическая зависимость – статистическая независимость между двумя признаками?
5. Что такое нулевая гипотеза?
6. Какие коэффициенты связи позволяют проверить гипотезу о статистической независимости?
7. Что значит ложное и истинное значение коэффициента связи?

**Семинар 8** Анализ данных повторных и сравнительных КСИ.

1. Виды повторных и сравнительных КСИ.
2. Общие правила осуществления и обработки / анализа данных повторных и сравнительных КСИ.

3. Проблема идентичности / сопоставимости методик сбора данных в повторных и сравнительных КСИ.
4. Оценка существенных различий в сравнительном и повторном КСИ как задача методического / методологического характера.
5. Инструментарии проведения и анализа результатов повторных и сравнительных КСИ.

**Вопросы // задания для самоконтроля:**

1. Охарактеризуйте лонгитюдные (монографические), панельные, когортные и трендовые исследования.
2. Кто из исследователей впервые провел лонгитюдные и панельные исследования?
3. Что такое мониторинг?
4. Через какой период времени целесообразно проводить сравнительные исследования?
5. Назовите критерии оценки существенности различий в сравнительных исследованиях.

**Семинар 9** Многомерный анализ социологической информации.

1. Принцип системности в анализе социологической информации.
2. Многомерное признаковое пространство: в аспекте логических и графических методов.
3. Многомерный анализ эмпирической информации и строение семантического пространства.
4. Симметричная корреляционная матрица и построение корреляционного графа.
5. Коэффициенты корреляции: частный, множественный, канонический и их познавательные возможности.
6. Многомерный регрессионный и многомерный факторный анализ.
7. Метод главных компонент как актуальная аналитико-математическая техника.

### **Вопросы // задания для самоконтроля:**

1. Назовите коэффициенты связи, основанные на величине «хи-квадрат».
2. Какие задачи можно решить в ходе анализа номинальных признаков?
3. Назовите методы многомерного анализа.
4. На основе имеющегося массива социологической информации сделать корреляционный анализ и проинтерпретировать полученные данные.
5. На основе имеющегося массива социологической информации сделать факторный анализ и проинтерпретировать полученные данные.
6. На основе имеющегося массива социологической информации сделать регрессионный анализ и проинтерпретировать полученные данные.

#### **2.4 Самостоятельная работа**

Нормативный объем самостоятельной работы студентов по данной дисциплине установлен в 40 час. Самостоятельная работа предполагает перманентную подготовку по основным теоретическим проблемам курса, выработку навыков и приемов ввода, обработки и анализа массива первичной социологической информации.

#### **2.5 Вопросы для подготовки к зачету**

1. Конкретные социологические исследования / КСИ как инструмент познания.
2. Метод, методика, техника, процедура КСИ.
3. Программа КСИ.
4. Рабочий и стратегический план КСИ.
5. Выборка в КСИ.
6. Социальные показатели и их измерение.
7. Виды КСИ.
8. Логический анализ основных понятий.
9. Измерение социальных явлений и показателей.
10. Научно-организационный менеджмент в осуществлении различных этапов КСИ.
11. Логическая последовательность действий при разработке и реализации программы.

12. Общая характеристика процедурно-методической части программы КСИ.
13. Выборочная и генеральная совокупность. Репрезентативность.
14. Расчет случайной ошибки репрезентативности.
15. Расчет объема выборочной совокупности.
16. Первичное районирование для общероссийской территориальной выборки.
17. Таблица случайных чисел.
18. Основные принципы расчета стоимости КСИ.
19. Отчет / аналитическая справка / научная публикация по итогам КСИ.
20. Общая характеристика организационной части программы КСИ.
21. Подготовка первичной социологической информации к обработке на ЭВМ.
22. Обобщение и изображение результатов КСИ.
23. Общая характеристика процесса обработки социологической информации.
24. Корреляционный анализ: сущность процедура, значение.
25. Факторный анализ: содержание, этапы.
26. Регрессионный анализ: характерные черты, процесс реализации, интенции.
27. Многомерный анализ социологической информации.
28. Количественный анализ данных КСИ.
29. Качественный анализ данных КСИ.
30. Качественно-количественный анализ данных КСИ.
31. Общая характеристика методов анализа социологической информации.
32. Общая характеристика методов обработки социологической информации.
33. Статистические методы обработки данных КСИ.
34. Техника ручной / на ЭВМ обработки социологической информации.
35. Группировка и типологизация: аспект анализа данных КСИ.
36. Поиск взаимосвязей между переменными: аспект анализа данных КСИ.
37. Социальный эксперимент как метод проверки научной гипотезы: в русле анализа данных КСИ.
38. Анализ данных повторных и сравнительных КСИ.
39. Последовательность действий при качественно- количественном анализе данных КСИ.

40. Стратегия качественного анализа данных КСИ.
41. Программное обеспечение КСИ.
42. Разработка социологической задачи.
43. Сбор социологической информации и ввод ее в ЭВМ.
44. Методология оперативного КСИ.
45. Общая характеристика статистической обработки.
46. Понятие измерения.
47. Элементарные понятия статистики.
48. Графическая интерпретация социолого-эмпирических зависимостей.
49. Средние величины и характеристики рассеяния значений признака.
50. Показатели колеблемости / вариации значений признаков: аспект социолого-аналитической работы.
51. Статистические взаимосвязи и их социологический анализ.
52. Интерпретация данных КСИ: общая характеристика.
53. Прогнозирование социальных процессов в современной социологии.
54. Использование результатов КСИ в сфере управления/менеджмента.
55. Эмпирическая социология в современной России: состояние и перспективы.
56. Обработка и анализ данных сравнительных КСИ в современной России.
57. Современные технологии инновации в КСИ.
58. Перспективы развития аналитического раздела социологической науки.

## **2.6 Критерии оценки знаний студентов**

Оценка знаний, умений, навыков студентов осуществляется в соответствии со следующими критериями:

- 1) уровень глубины изучения и усвоения содержания рекомендуемой литературы по данному курсу;
- 2) степень усвоения основной проблематики, закономерностей, категориально - понятийного аппарата изучаемого курса;
- 3) достаточность знания главных направлений, структуры, разделов, места, роли и значения курса «Анализ данных в социологии» в системе

современного научного знания, в аспекте качественной подготовки социолога-профессионала;

- 4) системность и качество подготовки к семинарским занятиям, а также - выполнения курсовых работ по данному курсу.

### 2.7 Тест

Применяется в качестве формы контроля остаточных знаний, а также – рубежного контроля. В тесте – 30 вопросов и по 4 варианта ответов на каждый из них. Возможны один или несколько правильных вариантов ответа. В соответствии с федеральным нормативом оценка «отлично» выставляется за 90 % правильных ответов; «хорошо» - за 75 % правильных ответов; «удовлетворительно» - за 50 % правильных ответов.

#### Тест

1. Процедура, с помощью которой измеряемый объект сравнивается с некоторым эталоном и получает числовое выражение в определенном масштабе или шкале, называется:

1. Верификация
2. Измерение
3. Классификация
4. Типологизация

2. «Кафетерием Лайкерта» называется следующая процедура:

1. Суммарные оценки по ряду номинальных шкал
2. Суммарные оценки по ряду интервальных шкал
3. Суммарные оценки по ряду ранговых шкал
4. Суммарные оценки по ряду шкал отношений

3. Характеристики, изменяемые в ходе эксперимента исследователем называется:

1. Контрольные
2. Факторные
3. Нейтральные (переменные)
4. Нейтральные (постоянные)

4. Степень разброса значений признака показывает:
  1. Ряд распределения
  2. Среднее арифметическое
  3. Дисперсия
  4. Коэффициенты корреляции
5. Графическое отображение непрерывных рядов распределения – это:
  1. Полигон
  2. Гистограмма
  3. Диаграмма
  4. Корреляционный граф
6. Связывание данных предварительно упорядоченных по двум признакам (свойствам, переменным):
  1. Простая классификация
  2. Перекрестная группировка
  3. Эмпирическая типологизация
  4. Теоретическая типологизация
7. Всестороннее объяснение социологических данных именуется:
  1. Группировка
  2. Типологизация
  3. Интерпретация
  4. Анализ
8. Степень положительной или отрицательной напряженности социальной установки респондента к социальному объекту определяется с помощью:
  1. Шкалы Лайкерта
  2. Шкалы Гутмана
  3. Шкалы Терстоуна
  4. Семантического дифференциала
9. Поиск устойчивых сочетаний свойств социальных объектов (или явлений), рассматриваемых в соответствии с описательными гипотезами в нескольких измерениях одновременно:

1. Простая группировка
  2. Перекрестная группировка
  3. Эмпирическая типологизация
  4. Теоретическая типологизация
10. К мерам средней тенденции относится:
1. Энтропия
  2. Медиана
  3. Индекс
  4. Регрессия
11. Первичное измерение качественных признаков:
1. Ряд распределения
  2. Группировка
  3. Верификация
  4. Квантификация
12. Графическая форма отображения дискретных рядов распределения:
1. Полигон
  2. Гистограмма
  3. Таблица
  4. Корреляционный граф
13. Упорядочение данных по одному признаку – это:
1. Простая группировка
  2. Перекрестная группировка
  3. Эмпирическая типологизация
  4. Теоретическая типологизация
14. Мода – это:
1. Наиболее часто встречающееся значение признака
  2. Значение признака, который делит вариационный ряд пополам
  3. Мера разброса признака
  4. Среднее значение признака
15. Нулевая гипотеза ( $H_0$ ) говорит о том, что:

1. Связь между рассматриваемыми переменными отсутствует
  2. Связь между рассматриваемыми переменными наблюдается
  3. Связь между рассматриваемыми переменными функциональная
  4. Связь между рассматриваемыми переменными ложная
16. Анализ социологических данных определяется:
1. Проблемой исследования
  2. Гипотезами исследования
  3. Количеством вопросов в анкете
  4. Уровнем измерения социальных характеристик
17. Шкалами низкого уровня измерения являются:
1. Номинальные
  2. Порядковое
  3. Интервальные
  4. Шкалы отношений
18. Степень неопределенности распределения признака – это:
5. Энтропия
  6. Медиана
  7. Индекс
  8. Регрессия
19. Регрессионный анализ номинальных признаков решает задачи типа:
1. "Признак - признак"
  2. "Признак - (группа признаков)"
  3. "(Группа признаков) - (группа признаков)"
  4. Анализ системы признаков
20. Вычисление коэффициентов парной связи номинальных признаков решает задачи типа:
1. "Признак - признак"
  2. "Признак - (группа признаков)"
  3. "(Группа признаков) - (группа признаков)"
  4. Анализ системы признаков

№ вопроса	Вариант ответа	№ вопроса	Вариант ответа
1	2	11	4
2	1	12	2
3	2	13	1
4	3	14	1
5	1	15	1
6	2	16	2,4
7	3	17	1,2
8	3	18	1
9	3	19	2
10	2	20	1

### 3. УЧЕБНО-МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЫ «АНАЛИЗ ДАННЫХ В СОЦИОЛОГИИ»

#### 3.1 Рекомендуемая литература

##### Основная литература

1. Оперативные социологические исследования: учеб. пособие / Д. Г. Ротман и др. - Минск: Веды, 1997. – 205 с.
2. Российская социологическая энциклопедия / под общ. ред. Г.В.Осипова. - М.: НОРМА-ИНФРА-М, 1998. – 666 с.
3. Социология: учеб. пособие / под ред. А. Н. Елсукова. - Минск: ТетраСистемс, 2004. - 541 с.
4. Татарова Г.Г. Методология анализа данных в социологии (введение): учебник для вузов. – М.: NOTA BENE, 1999, - 224 с.
5. Толстова Ю.Н. Анализ социологических данных: Методология, дескриптивная статистика, изучение связей между номинальными признаками: учеб. пособие. Рек. Мин. обр. РФ - М.: Научный мир, 2000. – 350 с.
6. Толстова Ю.Н. Измерение в социологии: курс лекций. - М.: ИНФРА-М, 1998. – 223 с.

##### Дополнительная литература

1. Градосельская Г.В. Сетевые измерения в социологии: учеб. пособие. - М.: Новый учебник, 2004. - 248 с.
2. Девятко И.Ф. Методы социологического исследования: учеб. пособие. – 2-е изд., испр. – М.: Университет, 2002. – 296 с.
3. Добреньков В.И. Фундаментальная социология: в 15 т: учеб. пособие. Т. 2. - Эмпирическая и прикладная социология. - 2004. - 986 с.
4. Карлинский И.З. Анализ социального и правового положения бездомных в современной России. - СПб.: Дельта, 2004. - 62 с.
5. Магницкий Ю.Г. Измерения и анализ данных в социологии: учеб. – метод. пособие. – Благовещенск: Изд-во Амур. гос. ун-та, 2003. – 26 с.
6. Наследов А.Д. SPSS: компьютерный анализ данных в психологии и

- социальных науках: самоучитель. - СПб.: Питер, 2005. - 416 с.
7. Рабочая книга социолога / Ред., предисл. Г.В. Осипова. - М.: Едиториал УРСС, 2003. - 477 с.
  8. Торопчина Г.Н. Элементы кластерного анализа: учеб. пособие / Г. Н. Торопчина, Н. Н. Двоерядкина, Г. П. Вохминцева. - Благовещенск: Изд-во Амур. гос. ун-та, 2006. - 39 с.
  9. Шалак В.И. Современный контент-анализ: приложения в области политологии, психологии, социологии, культурологии, экономики, рекламы. - М.: Омега-Л, 2004. - 272 с.
  10. Ядов В.А. Стратегия социологического исследования: описание, объяснение, понимание социальной реальности: учебник для вузов. Рек. Мин. обр. РФ. - М.: Добросвет, 1999. - 596с.

### **3.2 Краткое содержание лекционного курса**

#### **Тема 1 Подготовка первичной социологической информации к обработке на ЭВМ.**

Подготовка первичной социологической информации к обработке осуществляется в несколько этапов:

- 1) присвоение кодов вариантам ответов на закрытые и полужакрытые вопросы на стадии разработки методического инструментария. Кодирование закрытых вопросов имеет две ситуации: кодирование вопросов линейной и табличной форм. Кодирование вопросов линейной формы происходит путем присвоения вариантам ответов на закрытые вопросы кодов в виде натурального числа. Табличные варианты обычно аккумулируют несколько линейных вопросов. Их кодирование происходит в соответствии с расположением шкалы измерения. В перечневых таблицах шкала измерения располагается горизонтально, в соответствии с чем коды и выстраиваются в горизонтальный ряд натуральных чисел. В функциональных таблицах, измеряющих разные типы поведения респондента, коды располагаются вертикально. Респондент, выбирая вариант ответа, одновременно кодирует его;
- 2) контроль качества (полнота и точность) анкет. Если отсутствие ответов

характерно менее чем на 10 % содержащихся в нем вопросов, то, опираясь на ответы респондента на остальные вопросы анкеты, можно логически восстановить ответы на «пропущенные» вопросы. Если отсутствие ответов характерно для 20 % ответов или 2-3 вопросов в социально-демографической части, то такие анкеты исключаются из основного массива как искажающие социологическую информацию. При контроле анкет целесообразно проверить их на предмет противоречивых ответов на вопросы. В случае исключения из обработки 20 % и более вопросов анкета подлежит выбраковке;

3) кодирование открытых вопросов. Различают два вида открытых вопросов: перечневые и текстовые. Перечневые открытые вопросы содержат многократно повторяющуюся и легко поддающуюся группировке информацию. Ответы текстового характера предварительно следует выписать, отмечая частоту повторяемости суждений. После подсчета частоты идентичных по смыслу суждений, близкие по содержанию суждения объединяют в одну группу, сведя, таким образом, многозначность суждений к 5-10 укрупненным. Основой подобной группировки служат задачи исследования. Каждой группе присваивается код;

4) ввод первичной информации в ЭВМ. Для этого целесообразно знать пакет программ SPSS для компьютерной статистической обработки социологической информации и Excel для форматирования таблиц и построения графиков и диаграмм.

## **Тема 2 Понятие измерения в современной эмпирической социологии.**

Поскольку социолог имеет дело с массовыми процессами, он по необходимости оперирует различными числовыми показателями, выражающими частоты, протяженности и напряженность связи между различными социальными характеристиками. Предпосылка всех операций с количественными выражениями свойств социальных объектов и процессов — первичное измерение качественных признаков или их квантификация. Любое измерение начинается с поиска простейших качественных признаков, отношения между которыми могли бы быть выражены в некотором числовом

масштабе.

Измерение — это процедура, с помощью которой измеряемый объект сравнивается с некоторым эталоном и получает числовое выражение в определенном масштабе или шкале.

Конструирование эталона измерения - шкалы. Поиск эталона измерения.

Первая стадия — качественная классификация объектов.

Вторая стадия — поиск протяженности выделенных в качественном анализе свойств. Следует установить, обладают ли эти свойства прерывной или непрерывной протяженностью, можно ли их представить в виде различных последовательных состояний измеряемого качества.

Третья стадия — установление эмпирических индикаторов или внешних признаков тех свойств объекта, которые поддаются расположению в континуум.

Индикатор — внешне хорошо различимый показатель измеряемого признака. С его помощью устанавливается наличие или отсутствие признака, его состояние.

Одно и то же свойство можно фиксировать с помощью нескольких индикаторов, которые сводятся в общий показатель — индекс. Это случай сложного первичного измерения, по технике напоминающего приемы вторичных измерений.

Четвертая стадия заключается в том, чтобы уяснить, все ли единицы, составляющие измеряемый объект, укладываются в ранжируемый ряд, все ли они обладают свойством занимать определенное место в континууме отношений по принятым индикаторам.

В итоге устанавливается эталон, или шкала измерения.

Классификация измерительных эталонов. Шкалы упорядочены по мере повышения их способности удовлетворять требованиям более многообразных операций с числами.

1. Номинальные шкалы:

- номинальная шкала (неупорядоченная шкала наименований);

- частично упорядоченная номинальная шкала;
- порядковая шкала или полностью упорядоченная ординарная шкала (например, шкала рангов, она же — ранговая шкала).

## 2. Метрические шкалы:

- интервальная шкала (шкала равных интервалов, например типа шкалы Терстоуна);
- идеальная или абсолютная шкала (шкала пропорциональных оценок, шкала отношений).

Все эти шкалы предназначены для квантификации одномерных распределений, т. е. измерения некоторой протяженности в одном и только в одном континууме свойств.

Простая номинальная шкала (низкий уровень измерения) служит предпосылкой всех шкальных процедур. Она устанавливает отношения равенства между явлениями, которые включены в один класс. Пункты-шкалы — эталоны качественной классификации свойств.

Частично упорядоченная шкала служит для установления отношений равенства между явлениями в каждом классе и отношений последовательности в терминах  $>$  или  $<$  между несколькими, но не всеми классами (минимум двумя из  $n$  классов, где  $n > 2$ ).

Она обычно используется как промежуточный этап при разработке полностью упорядоченных шкал. Иногда, однако, ранжировать весь ряд не удается.

Порядковая шкала. Полностью упорядоченная шкала наименований устанавливает отношения равенства между явлениями в каждом классе и отношения последовательности в понятиях  $>$  и  $<$  между всеми без исключения классами.

Упорядоченные номинальные шкалы имеют и более сложные конструкции (например, шкала Гутмана), а в простейшем варианте являются составными элементами многих мерительных операций, в особенности методов суммирования оценок по ряду шкал.

Часто употребляемая разновидность шкал этого типа — ранговые. Они предполагают полное упорядочение каких-то объектов от наиболее к наименее важному, значимому, предпочитаемому.

Метрическая шкала равных интервалов. Класс метрических шкал в отличие от номинальных устанавливает отношение между пунктами не просто в понятиях больше-меньше, но позволяет фиксировать величину интервала.

Шкала интервалов представляет собой полностью упорядоченный ряд с измеренными интервалами между пунктами, причем отсчет начинается с произвольно избранной величины.

Главная трудность в построении таких шкал — обоснование равенства или разности дистанций между пунктами.

Шкала пропорциональных оценок.

Идеальная или абсолютная метрическая шкала, напоминающая шкалу равных интервалов, но с одним преимуществом: отсчет в этой шкале начинается не с произвольной точки, а с экспериментально установленного нулевого пункта. Для таких шкал применимы решительно все операции с числами, так как можно определить, насколько или во сколько данный пункт на шкале превышает другой. Подобные шкалы приняты в точных науках, где нулевой пункт (точка отсчета—откуда и происходит название «точные науки») экспериментально зафиксирован.

### **Тема 3 Типологизация социологических данных: (в аспекте КСИ).**

Группировка и классификация — элементарные процедуры упорядочения данных, предваряющих их анализ.

Простая группировка — это классификация или упорядочение, данных, по одному признаку. Связывание фактов в систему осуществляется здесь в соответствии с описательной гипотезой относительно ведущего признака группировки (или признака классификации). Так, в зависимости от гипотез можно сгруппировать выборочную совокупность по возрасту, полу, роду занятий, образованию, по высказанным суждениям и т. д.

Квантифицированные данные или количественные показатели

группируются в ранжированные ряды по возрастанию (убыванию) признака, качественные или атрибутивные группируются по принципу построения неупорядоченных номинальных шкал.

Все операции последующего анализа покоятся на изучении сгруппированных данных.

Статистические приемы поиска средней тенденции (мода, медиана, среднеарифметическая), подсчет дисперсии, отклонения позволяют оцепить сгруппированный ряд в емком показателе и отобразить результаты графически. Простейший анализ группировки — исчисление частот по процентам.

Перекрестная группировка (или перекрестная классификация) — это связывание данных предварительно упорядоченных по двум признакам (свойствам, показателям) с целью: а) обнаружить какие-то взаимозависимости, б) осуществить взаимоконтроль показателей, сформировать новый составной показатель (индекс) на основе совмещения двух свойств или состояний объекта, определить направление связей влияния одного явления (характеристики, свойства) на другое.

Одна из задач перекрестной классификации — поиск устойчивых, связей, выявляющих структурные свойства изучаемого явления.

Наконец, типичный случай использования перекрестной группировки — поиск тенденции, динамики процесса.

Эмпирическая типологизация — наиболее сильный прием анализа по описательному плану. Этот метод можно характеризовать как поиск устойчивых, сочетаний свойств социальных объектов (или явлений), рассматриваемых в соответствии с описательными гипотезами в нескольких измерениях одновременно,

Основную идею подобной типологизации сформулировал применительно к социологии П. Лазарасфельд. Он ввел понятие «пространство свойств». Так, нетрудно вообразить свойства социальной группы в трехмерном физическом пространстве, т. е. в декартовой системе координат. В этом трехмерном пространстве следует теперь определить, какова же упорядоченность свойств.

Теоретическая типологизация — обобщение признаков социальных явлений на основе идеальной теоретической модели и по теоретически обоснованным критериям. Такая типология отличается от рассмотренной выше, где устойчивость свойств типа находится путем многократного перебора, тогда как в теоретической типологии критерии свойств выявляются путем логического анализа.

В современной логике существует понятие «идеализированный» (идеальный) объект, которым обозначают реальный объект или целый класс объектов, отраженных в сознании в виде некоторой абстракции, идеальной системы, воспроизводящей его в упрощенном, схематизированном виде.

Идеальная социальная модель строится на основе абстракций двоякого рода: тех, что логически вытекают из более общих социологических понятий или принципов, а также абстракций на основе наблюдения эмпирических данных. Разумеется, и те, и другие имеют своей посылкой реальную действительность. Именно потому, что конструированная таким путем идеальная модель соотносится с системой теоретического знания, она выполняет важные функции включения теории в непосредственный анализ эмпирических данных.

Модель такого рода обладает рядом особенностей: она определяет идеальные (в смысле абстракции) границы социального объекта; включает критерии (или параметры), на основе которых определяется жесткая, устойчивая связь его свойств и характеристик; если параметры, составляющие модель, представляют континуумы, фиксируются также количественные границы идеализированного объекта.

Анализ эмпирических данных, согласно теоретической типологии, предполагает, во-первых, определение частот распределения по каждому типу; во-вторых, изучение отклонений от идеализированных моделей по отдельным параметрам и, если возможно, измерение интенсивности и вероятности этих отклонений.

Метод теоретической типологизации ведет к объяснению, тогда как

эмпирическая типологизация допускает лишь описание полученных данных и их интерпретацию.

#### **Тема 4 Характерные черты и условия осуществления различных видов / методов анализа.**

Выбор конкретного метода зависит, в первую очередь, от характера исследовательских гипотез. Помимо характера исследовательских гипотез на выбор методов статистического анализа влияет и природа полученных социологом данных. Разные уровни измерения социологических переменных определяют возможности и ограничения анализа. Методы, используемые для анализа связи между двумя номинальными переменными, будут отличаться от методов анализа связи между номинальной переменной и переменной, измеренной на интервальном уровне. Таким образом, выбор той или иной статистики будет зависеть и от целей анализа, и от уровня измерения исследуемых переменных.

Способы проверки процедуры первичного измерения на надежность.

Следующие операции повышения надежности первичного измерения используются лишь на стадии отработки инструмента измерения в процессе пилотажа.

Понятие надежности инструмента измерения (и соответственно надежности данных, фиксируемых этим инструментом), включает три составляющие: 1) обоснованность, 2) устойчивость и 3) правильность измерения.

Обоснованность (валидность) шкалы заключается в том, что с ее помощью целенаправленно измеряют вполне определенное свойство или признак, не смешивая его с другими.

Чтобы повысить обоснованность измерения, используют ряд технических приемов.

1. Наиболее простой способ — логические рассуждения на основе опыта и здравого смысла.

2. Тест по «эталонной группе» — более сильный прием проверки

инструмента на обоснованность. Смысл проверки — в сопоставлении данных, полученных путем измерения по шкале, с достоверными сведениями об объекте измерения.

3. Поиск независимого критерия как разновидность внешнего контроля надежности для измерения того же самого объекта или свойства.

4. Использование метода судей для отбора пунктов шкалы.

5. Один из широко используемых приемов внутреннего контроля обоснованности — совмещение нескольких показателей для регистрации определенного одного свойства или построение индекса.

Выделяют два существенно разных аспекта обоснованности: теоретический и эмпирический. Первый непосредственно связан с содержательными посылками исследования и предполагает установление значимых связей с широким классом ситуаций, предсказываемых теорией, второй требует доказательства надежной регистрации данных в сравнительно узком секторе, в частном проявлении изучаемых объектов.

Устойчивость измерения выражается в однозначности информации, которую мы извлекаем с помощью данной процедуры.

1. Наиболее распространенный прием контроля по устойчивости — повторное измерение.

2. Использование нескольких лиц для измерения данного свойства.

3. Наконец, третий прием контроля эталона измерения на устойчивость — «расщепление шкалы».

Точность и правильность измерения зависят от а) степени устойчивости измеряемого объекта или свойства, б) чувствительности эталона измерения (дробности пунктов шкалы), в) отсутствия систематических ошибок измерения и, г) от устойчивости измерения.

Социальные объекты, подлежащие измерению, обладают различной степенью устойчивости.

Дробность метрики — чувствительность шкалы — прямо связана с точностью измерения. Шкала в 10 делений измеряет точнее, чем в 5 или 3

деления. Но дробность пунктов шкалы нельзя увеличивать беспредельно. Надо установить оптимум, удовлетворяющий двум требованиям: максимум градаций шкалы при условии высокой устойчивости результатов измерения. Постепенно повышая дробность эталона измерения и параллельно проверяя шкалу на устойчивость, мы найдем границу, за пределами которой дальнейшее повышение дробности влечет понижение устойчивости. Это и есть оптимум чувствительности шкалы с учетом меры устойчивости измеряемого свойства. Таким образом, достижение устойчивых данных при максимальной дробности метрики повышает точность измерения. Оно будет удовлетворительно точным, если абсолютная ошибка измерения не превышает 0,5 деления шкалы. Вместе с тем, если ошибка вообще отсутствует  $|\Delta|=0$ , то не исключено, что шкала обладает заниженной чувствительностью (особенно в случаях, когда мы предполагаем достаточную вариабельность измеряемого свойства).

При квалификации социальных характеристик проблема правильности, т. е. отсутствия отклонений от истинного значения измеряемого свойства, намного сложнее, ибо часто мы в принципе не способны установить, каковы же эти истинные значения измеряемых свойств (скажем, мнений людей по каким-то вопросам). Мы можем лишь, сопоставляя разные способы фиксирования данного свойства, добиваться устранения замеченных систематических ошибок. Каковы же эти систематические ошибки?

Одна из возможных — отсутствие «разброса» информации по шкале вследствие того, что какие-то ее пункты «не работают», т. е. не реагируют на определенное состояние измеряемого свойства.

Другой причиной неправильности может быть плохая различительная способность соседних пунктов шкалы высокой дробности.

Правильность и точность измерения можно повысить путем расчета относительной ошибки измерения.

Относительная ошибка позволяет сопоставлять правильность замеров по двум и более шкалам разной чувствительности и таким путем отработать оптимальный инструмент. В отличие от абсолютной, относительная ошибка

исчисляется не в долях погрешности деления шкалы, а в соизмеримом, определенным образом нормированном показателе.

### **Тема 5 Группировка социолого-исследовательских наблюдений как фактор элементарных понятий статистики.**

Статистическая группировка. Наиболее простая форма обобщения первичной социологической информации — группировка. Она позволяет зачислить респондента в ту или иную группу в соответствии с выбранным признаком (или признаками) группировки.

Выбор признака группировки диктуется задачей социологического исследования, сформулированными ранее гипотезами.

В зависимости от шкалы измерения, в соответствии с которой получены ответы на вопрос, группировка социологической информации может представлять собой:

- зачисление респондентов в номинальные группы (группировка опрошенных по полу, национальности и т. д.);
- упорядочение информации в ранжированном ряду;
- группировку по количественному признаку, в результате которой группы респондентов характеризуются числовой величиной и потому количественно сравнимы между собой.

Операции с номинальными и ранжированными группами осуществляются при помощи математических приемов, соответствующих номинальной и ранговой шкалам. Что касается групп, распределенных по количественному признаку, то они могут быть подвергнуты дальнейшему изучению при помощи любых приемов математической статистики, так как такая группировка осуществляется по интервальной шкале. Группировка, произведенная по одному признаку, называется простой, по двум и более признакам и служащая выявлению их взаимосвязи, называется аналитической.

Когда респондентов распределяют в группы по двум или более признакам, то говорят о комбинационной группировке. В зависимости от решаемых задач она может быть структурной, типологической и

аналитической.

При структурной группировке респондентов классифицируют по некоторому свойственному всей совокупности объективному признаку.

Типологическая группировка – выделение типов на основании субъективного признака, измеренного по оценочной шкале, либо на базе нескольких признаков (многомерная группировка).

Ряд чисел, получаемый в результате группировки, называется рядом распределения. Ряды распределений, отражающие результат группировки респондентов по качественным признакам, называются атрибутивными, а по количественным — вариационными. В соответствии с характером количественных признаков вариационные ряды делятся на дискретные и непрерывные. Последние, как правило, носят интервальный характер. Это значит, что та или иная группа респондентов характеризуется по количественному признаку не одним числом, а числовым интервалом. Это важно учитывать при вычислении среднестатистического показателя. В соответствии с задачами группировки интервалы могут выбираться равные и неравные, с возрастающими и убывающими значениями. Числа, обозначающие интервалы, называются их границами.

Составление таблиц - это форма отображения рядов распределения.

Числовые данные в таблице объясняются заголовками, подлежащим и сказуемым. Заголовки в таблице бывают общие, выступающие в качестве названия таблицы и раскрывающие структуру группировки рассматриваемой совокупности респондентов либо связь между рядами распределения. Наряду с этим в названии таблицы целесообразно указывать место и время, к которым относятся приводимые сведения, а также общие для всех показателей единицы измерения (проценты, число человек, оценочный индекс, номинал денег и т. д.). Содержание строк и столбцов раскрывается внутренними заголовками: боковыми для строк и верхними для столбцов.

Подлежащим в таблице называется объект (совокупность респондентов), характеристики которого — сказуемое — выражены в таблице в числовом

виде. Перечневая таблица – таблица, составленная на основании ряда распределения по одному признаку. Таблицы, отображающие ряды распределений по двум и более признакам называются комбинационными.

### **Тема 6 Поиск взаимосвязи между переменными: в русле работы с итогами сбора социологической информации.**

Статистическая группировка. Наиболее простая форма обобщения первичной социологической информации — группировка. Она позволяет зачислить респондента в ту или иную группу в соответствии с выбранным признаком (или признаками) группировки.

Выбор признака группировки диктуется задачей социологического исследования, сформулированными ранее гипотезами.

В зависимости от шкалы измерения, в соответствии с которой получены ответы на вопрос, группировка социологической информации может представлять собой:

- зачисление респондентов в номинальные группы (группировка опрошенных по полу, национальности и т. д.);
- упорядочение информации в ранжированном ряду;
- группировку по количественному признаку, в результате которой группы респондентов характеризуются числовой величиной и потому количественно сравнимы между собой.

Операции с номинальными и ранжированными группами осуществляются при помощи математических приемов, соответствующих номинальной и ранговой шкалам. Что касается групп, распределенных по количественному признаку, то они могут быть подвергнуты дальнейшему изучению при помощи любых приемов математической статистики, так как такая группировка осуществляется по интервальной шкале. Группировка, произведенная по одному признаку, называется простой, по двум и более признакам и служащая выявлению их взаимосвязи, называется аналитической.

Когда респондентов распределяют в группы по двум или более признакам, то говорят о комбинационной группировке. В зависимости от

решаемых задач она может быть структурной, типологической и аналитической.

При структурной группировке респондентов классифицируют по некоторому свойственному всей совокупности объективному признаку.

Типологическая группировка – выделение типов на основании субъективного признака, измеренного по оценочной шкале, либо на базе нескольких признаков (многомерная группировка).

Графики и диаграммы - способ отображения социологических данных. Чаще всего он имеет вид полигона или гистограммы. Полигон используется для графического отображения непрерывных рядов, а гистограмма — дискретных.

Процесс многомерного анализа начинается с определения взаимосвязи исходных признаков исследуемого объекта. Многие из методов многомерного анализа базируются на использовании совокупности разных коэффициент корреляции, носящих имя К.Пирсона.

Вся совокупность корреляционных коэффициентов образует симметричную корреляционную матрицу, в которой содержится  $m(m-1)/2$  различных элементов.

Применение специальных статистических критериев позволяет выбрать из всей совокупности коэффициентов корреляции те, которые отличны от нуля, то есть указывают на существование зависимости между конкретными переменными. Осуществляя эту статистическую процедуру по отношению ко всему множеству из  $m(m-1)/2$  парных коэффициентов корреляции можно сделать первый вывод многомерного характера о типе "взаимоотношений" между переменными.

Совокупность логических и инструментальных операций, позволяющих выявить природу латентных переменных и измерить силу их влияния на наблюдаемые признаки, составляет содержание многомерного факторного анализа.

Ч. Спирменом была сформулирована идея существования единственной латентной (скрытой) переменной (генеральный фактор) и предложен эмпирический способ ее определения (однофакторная модель Спирмена). Стремление к обобщению одномерной модели на случай многомерности привели

к созданию многомерных схемы факторного анализа. Л.Терстоун и ряд американских математиков обосновали алгоритмы извлечения латентных переменных.

Сведенные в таблицы или отображенные при помощи графиков и рисунков результаты социологического исследования являют собой числовые величины, пригодные для выводов о характере и признаках изучаемого явления. Первое условие, позволяющее должным образом использовать социологические данные, — их всестороннее и правильное объяснение, именуемая социологами интерпретацией.

В процессе логического превращения усредненных числовых величин в показатели большое значение имеют ранее выдвинутые гипотезы. Характер проверки гипотезы в прикладной социологии предопределен видом исследования. В разведывательном исследовании гипотеза проверяется непосредственно путем соотнесения предполагаемого утверждения с выявленной числовой величиной.

В описательном исследовании она предполагает интерпретацию усредненных величин, полученных на основе обобщения характеристик разнородного по составу объекта анализа.

Один из наиболее распространенных способов интерпретации данных в описательном исследовании – сравнение рядов распределения по относительно однородным группам, выделенным в обследованной совокупности. Она может осуществляться двумя путями: внутренним и внешним соотнесением. Под внутренним соотнесением понимают сравнение между собой элементов числового ряда, а под внешним соотнесением – сравнение двух или нескольких рядов распределения, построенных по двум или нескольким признакам, из которых один общий для соотносимых рядов.

Логика доказательства гипотез в аналитическом исследовании преимущественно основан на поиске связи между характеристиками объекта, их взаимовлияния, тенденций и причин изменения. Она состоит из двух последовательных, взаимосвязанных этапов интерпретации полученных

данных. Первый из них основан на методе сравнения числовых рядов распределения, второй на поиске факторного признака (признаков). Один из наиболее используемых методов для второго случая – метод последовательного исключения.

### **Тема 7 Социальный эксперимент как аналитический метод верификации научной гипотезы.**

Общая логика эксперимента заключается в том, чтобы, выбрав экспериментальную группу (или группы) и поместив ее в необычную экспериментальную ситуацию (под воздействие определенного фактора), проследить направление, величину и устойчивость изменения характеристик, называемых контрольными.

При формировании эксперимента следует соблюдать три требования:

1) в качестве контрольных выбираются характеристики, самые важные с точки зрения изучаемой проблемы;

2) изменение контрольных характеристик должно зависеть от тех характеристик экспериментальной группы (либо среды проведения эксперимента), которые вводятся или изменяются самим исследователем. Такие характеристики называются факторными. Характеристики, «не участвующие» в эксперименте, носят название нейтральных. Они в процессе эксперимента может сложиться по-разному. В одних случаях они могут без «постороннего» воздействия измениться (тогда они получают название переменных), в других - остаться неизменными (тогда их называют постоянными). Для конечных результатов эксперимента и то, и другое не имеет большого значения. Главное, чтобы изменения в нейтральных характеристиках не отражались на контрольных;

3) на протекание эксперимента не должны оказывать воздействие те явления, которые не относятся к экспериментальной ситуации, но потенциально способны изменить ее.

Виды эксперимента.

По характеру экспериментальной ситуации эксперименты делятся на

полевые и лабораторные.

В полевом эксперименте объект (группа) находится в естественных условиях своего функционирования. При этом члены группы могут быть осведомлены, либо не осведомлены о том, что они участвуют в эксперименте.

В лабораторном эксперименте ситуация, а часто и сами группы формируются искусственно. Поэтому члены группы, как правило, осведомлены об эксперименте.

По логической структуре доказательства гипотез различают линейный и параллельный эксперимент.

В линейном эксперименте анализу подвергается одна и та же группа, являющаяся и контрольной (ее первоначальное состояние), и экспериментальной (ее состояние после изменения одной или нескольких характеристик). То есть еще до начала эксперимента четко фиксируются все контрольные, факторные и нейтральные характеристики объекта. После этого изменяются факторные характеристики группы (или условия ее функционирования) и по истечении определенного, заранее заданного срока вновь измеряется состояние объекта по его контрольным характеристикам.

В параллельном эксперименте одновременно участвуют две группы: контрольная и экспериментальная. Их состав должен быть идентичен по всем контрольным, а также по нейтральным характеристикам, которые могут повлиять на исход эксперимента (в первую очередь это социально-демографические признаки). Характеристики контрольной группы остаются постоянными в течение всего периода эксперимента, а экспериментальной - изменяются. По итогам эксперимента контрольные характеристики двух групп сравниваются, и делается вывод о причинах и величине происшедших изменений.

Для успешного проведения эксперимента большую роль играет правильный подбор его участников. Используют три метода: попарного отбора, структурной идентификации и случайного отбора (рандомизация).

Ко всем трем методам предъявляется одно общее требование: они

должны гарантировать правомерность распространения результатов эксперимента на тот объект, изменение характеристик (или условий функционирования) которого предусматривается в последующем. Этот объект выступает как генеральная совокупность, из состава которой выбирается экспериментальная группа. Кроме того, использование соответствующего метода отбора группы (групп) в некоторой степени предопределено моделью планируемого эксперимента.

Метод попарного отбора используется преимущественно в параллельном эксперименте. Суть его заключается в следующем. Из генеральной совокупности отбираются две группы таким образом, чтобы они были идентичны по нейтральным и контрольным, но отличались по факторным характеристикам. Для обеих групп создаются одинаковые условия, а через некоторое время измеряется эффект эксперимента путем фиксации и сравнения параметров контрольных признаков в обеих группах.

Как в линейном, так и в параллельном эксперименте можно использовать метод структурной идентификации. В этом случае в линейном эксперименте группа отбирается так, чтобы она представляла собой микромодель генеральной совокупности по нейтральным и контрольным характеристикам. Такой отбор может быть осуществлен по принципу квотной выборки. В свою очередь, в параллельном эксперименте по тем же характеристикам выравниваются структуры экспериментальной и контрольной групп.

Метод случайного отбора идентичен методам вероятностной выборки с заранее заданным объемом. Как правило, его используют в полевых экспериментах при большой (до нескольких сот) численности экспериментальной группы.

Планирование и логика осуществления эксперимента.

Подготовка и проведение эксперимента предполагает последовательное решение ряда вопросов:

- 1) определение цели эксперимента;
- 2) выбор объекта (объектов), используемого в качестве

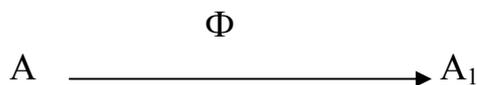
экспериментальной, а также контрольной группы (групп);

- 3) выделение предмета эксперимента;
- 4) выбор контрольных, факторных и нейтральных признаков;
- 5) определение условий эксперимента и создание экспериментальной ситуации;
- 6) формулировку гипотез и определение задач;
- 7) выбор индикаторов и способа контроля протекания эксперимента;
- 8) определение метода фиксации результатов;
- 9) выбор критерия оценки эффективности эксперимента.

Логика эксперимента всегда подчинена поиску причин, характера изменения интересующего исследователя социального явления или процесса. Непременное условие их нахождения — изменение параметров экспериментальной группы под воздействием некоторого фактора. Так, выбранная нами группа имеет состояние  $A$ . В процессе эксперимента вводится фактор  $\Phi$  (изменение одного из обычных условий функционирования группы, либо каких-то собственных ее характеристик). В результате группа приобретает состояние  $A_1$ . Тогда задача исследователя заключается в определении величины этого изменения, а также в выяснении степени воздействия фактора  $\Phi$  на изменение характеристик группы.

Рассмотрим логическую модель проведения линейного и параллельного видов эксперимента. Линейный эксперимент.

Схематично процесс изменения состояния объекта в линейном эксперименте удобно изобразить так:



Под новым состоянием объекта подразумевается изменение одной или нескольких его характеристик. В соответствии с этим процесс определения воздействующего фактора (основной причины изменения характеристик объекта), а также эффекта его воздействия (степени изменения контрольных характеристик) может быть отображен при помощи следующей

схематической модели:

$$A(a, \bar{b}, v) \longrightarrow A_1(\gamma, d)$$

$$A(a, \bar{b}, v_1) \longrightarrow A_1(\gamma, d_1)$$

$$\text{Следствие: } \Phi = v \longrightarrow d_1$$

Здесь **а, б, в, г, д** - различные характеристики экспериментальной группы (пол, возраст, стаж работы, профессия, производительность труда, этнические установки и т.д.).

Приведенная модель расшифровывается следующим образом: под влиянием характеристик **а, б, в** исходного состояния (**А**) экспериментального объекта получаем его результирующее состояние (**А<sub>1</sub>**) с характеристиками **г** и **д**, после изменения характеристики **в** результирующее состояние объекта изменилось по характеристике **д**. Следовательно, вывод - причиной изменения характеристики **д** на **д<sub>1</sub>** является изменение характеристики **в** на **в<sub>1</sub>**

Параллельный эксперимент. Логическая модель параллельного эксперимента имеет две разновидности: модель по методу единственного сходства и модель единственного различия.

Модель по методу единственного сходства может быть проиллюстрирована при помощи следующей логической схемы сравнения некоторой экспериментальной группы **А** и контрольной группы **Б**:

$$A(a, \bar{b}, v) \longrightarrow A_1(\gamma, d)$$

$$B(e, \bar{z}, v) \longrightarrow B_1(z, d)$$

$$\text{Следствие: } \Phi = v \longrightarrow d$$

Приведенная модель имеет следующую расшифровку: пусть некоторая экспериментальная группа (**А**) с характеристиками **а, б** и **в** в некотором ином состоянии имеет свойства **г** и **д**, а некоторая контрольная группа (**Б**) с характеристиками **е, ж** и **в** в том же ином состоянии имеет свойства **з** и **д**. В соответствии с этим уместно предположить, что единственной причиной качества **д** является характеристика **в**, так как только она неизменно присутствует среди характеристик обоих объектов.

Модель по методу единственного различия имеет следующую

логическую схему:

$A(a, б, в) \longrightarrow A_1(г, д)$

$B(a, б, е) \longrightarrow B_1(г, ж)$

Следствие:  $\Phi = е \longrightarrow ж$

Приведенная модель имеет следующую расшифровку: пусть некоторая экспериментальная группа (А) с характеристиками **а**, **б** и **в** в некотором ином состоянии имеет свойства **г** и **д**, а некоторая контрольная группа (Б) с характеристиками **а**, **б** и **е** в том же ином состоянии имеет свойства **г** и **ж**. В соответствии с этим уместно предположить, что единственной причиной качества **ж** является характеристика **е**, так как именно она отсутствует среди характеристик первого объекта.

## **Тема 8 Статистические методы обработки социологической информации.**

Существует два основных класса задач, решаемых с помощью статистических методов анализа. Задачей дескриптивной (описательной) статистики является описание распределения переменной-признака в конкретной выборке. Методы дескриптивной статистики позволяют также анализировать взаимосвязь между различными переменными. Другой класс задач, связанный с необходимостью вывести свойства большой совокупности, основываясь на имеющейся информации о свойствах выборки из этой совокупности, решается с помощью методов индуктивной статистики, или теории статистического вывода, основанной на вероятностном подходе к принятию решений. Воспользовавшись какой-то моделью для анализа полученных выборочных данных, социолог обычно также применяет некоторые методы статистического вывода, позволяющие определить, выполняются ли обнаруженные им при анализе данных отношения на уровне большой совокупности, из которой была извлечена выборка.

На фазе программирования эмпирического исследования, и в момент анализа собранной информации социолог имеет дело с совокупностью отдельных характеристик, или, другими словами, сначала, конструирует, а

затем, исследует многомерное признаковое пространство.

Многомерный анализ в социологии – это использование специальных многопараметрических методов математической статистики для изучения строения и свойств социального многомерного пространства.

Процесс многомерного анализа начинается с определения взаимосвязи исходных признаков исследуемого объекта. Многие из методов многомерного анализа базируются на использовании совокупности разных коэффициентов корреляции, носящих имя К.Пирсона.

Вся совокупность корреляционных коэффициентов образует симметричную корреляционную матрицу, в которой содержится  $m(m-1)/2$  различных элементов.

Применение специальных статистических критериев позволяет выбрать из всей совокупности коэффициентов корреляции те, которые отличны от нуля, то есть указывают на существование зависимости между конкретными переменными. Осуществляя эту статистическую процедуру по отношению ко всему множеству из  $m(m-1)/2$  парных коэффициентов корреляции можно сделать первый вывод многомерного характера о типе "взаимоотношений" между переменными.

Трехмерным дополнением этому первичному анализу станет графическое представление корреляционной матрицы (метод построения корреляционного графа).

Построение корреляционного графа — это всегда синтез научной логики и исследовательской интуиции. Только они помогают найти то пороговое значение коэффициента корреляции, при использовании которого граф оказывается простым в интерпретации структуры зависимостей между исследуемыми социальными переменными. Если "порог" слишком низкий, то в графе будет много ребер, они будут взаимопересекаться и создавать сложную для анализа ситуацию. С другой стороны, слишком высокое пороговое значение станет причиной излишнего упрощения корреляционного графа: в нем будет меньшее количество вершин и ребер. Как следствие, семантика графа окажется крайне бедной.

Коэффициенты корреляции измеряют некое абсолютное значение связи между  $x_i$  и  $x_j$ , но не учитывающие их связи с другими переменными.

Частный коэффициент корреляции даст меру связи между  $x_i$  и  $x_j$  при условии, что значения какого-то третьего признака (или набора признаков) фиксированы.

Формула для вычисления частного коэффициента корреляции между  $x_i$  и  $x_j$  при фиксированном значении одной переменной ( $x_q$ ):

$$r_{ijq} = \frac{r_{ij} - r_{iq} \cdot r_{jq}}{\sqrt{(1 - r_{ij}^2) \cdot (1 - r_{jq}^2)}}$$

Измерения связи между одним из признаков объекта исследования и набором переменных можно осуществить, вычисляя и анализируя множественный коэффициент корреляции. Его значение измеряет совокупную детерминацию любого из признаков, входящих в ансамбль переменных, другими элементами этого же набора.

Мерой коррелированности одной группы изучаемых переменных с другой является канонический коэффициент корреляции. Его значение указывает, в какой степени изменчивость одной совокупности признаков может быть предсказана, обусловлена совместной изменчивостью другой совокупности социологических показателей.

Однако познавательные возможности коэффициентов корреляции, устанавливающих зависимость между одним признаком и группой признаков, а также между одной совокупностью переменных и другой ограничены в том отношении, что они дают слишком обобщенную, интегральную характеристику связи между переменными.

Более полную картину взаимосвязи позволяет построить многомерный регрессионный анализ, он объединяет в себе систему весьма глубоких логических построений и математико-вычислительных алгоритмов, синтезирующих опыт исследования в различных областях науки о зависимости между переменными.

Совокупность логических и инструментальных операций, позволяющих выявить природу латентных переменных и измерить силу их влияния на наблюдаемые признаки, составляет содержание многомерного факторного анализа.

Ч. Спирменом была сформулирована идея существования единственной

латентной (скрытой) переменной (генеральный фактор) и предложен эмпирический способ ее определения (однофакторная модель Спирмена). Стремление к обобщению одномерной модели на случай многомерности привели к созданию многомерных схемы факторного анализа. Л.Терстоун и ряд американских математиков обосновали алгоритмы извлечения латентных переменных.