17. Матанцев, А. Н. Анализ рынка: Настольная книга маркетолога / А. Н. Матанцев. – М. : Изд-во «Альфа-пресс», 2007. – 552 с.

18. Моосьюллер, Г. Маркетинговые исследования с SPSS : учеб. пособие / Г. Моосьюллер, Н. Н. Ребик. – М. : ИНФРА-М, 2007. – 160 с.

19. Просветов, Г. И. Маркетинговые исследования: задачи и решения : учеб.-практ. пособие / Г. И. Просветов. – М. : Изд-во «Альфа-пресс», 2008. – 240 с.

20. Сондерс, М. Методы проведения экономических исследований / М. Сондерс, Ф. Льюис, Э. Торнхилл ; пер. с англ. – 3-е изд. – М. : Эксмо, 2006. – 640 с.

21. Статистика : учеб. / под ред. И. И. Елисеевой. – М. : ТК Велби, Изд-во Проспект, 2005. – 448 с.

22. Токарев, Б. Е. Маркетинговые исследования : учеб. / Б. Е. Токарев. – М. : Экономистъ, 2007. – 624 с.

23. Черчилль, Г. Маркетинговые исследования / Г. Черчилль, Т. Браун. ; пер. с англ. ; под ред. Г. Л. Багиева. – 5-е изд. – СПб. : Питер, 2007. – 704 с.

Министерство образования и науки Российской Федерации

Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
«Комсомольский-на-Амуре государственный технический университет»

Е. В. Чепухалина

МАРКЕТИНГОВЫЕ ИССЛЕДОВАНИЯ
(компьютерный практикум)

Утверждено в качестве учебного пособия
Учёным советом Федерального государственного бюджетного
образовательного учреждения высшего профессионального образования
«Комсомольский-на-Амуре государственный технический университет»

Комсомольск-на-Амуре
2012

Практическое занятие 11
ПРИМЕНЕНИЕ МЕТОДОВ АНАЛИЗА ПАРНЫХ ВЗАИМОСВЯЗЕЙ МЕЖДУ ПРИЗНАКАМИ В МАРКЕТИНГОВЫХ ИССЛЕДОВАНИЯХ. ИЗУЧЕНИЕ ТЕСНОТЫ СВЯЗИ МЕЖДУ КОЛИЧЕСТВЕННЫМИ ПРИЗНАКАМИ.....54

Практическое занятие 12
СТАТИСТИЧЕСКИЙ МЕТОД ОПРЕДЕЛЕНИЯ ОБЪЕМА ВЫБОРКИ С ПОМОЩЬЮ СРЕДНЕГО.....58

Практическое занятие 13
СТАТИСТИЧЕСКИЙ МЕТОД ОПРЕДЕЛЕНИЯ ОБЪЕМА ВЫБОРКИ С ПОМОЩЬЮ ДОЛИ.....63

Практическое занятие 14
ОДНОФАКТОРНЫЙ ДИСПЕРСИОННЫЙ АНАЛИЗ 66

Практическое занятие 15
МНОГОФАКТОРНЫЙ ДИСПЕРСИОННЫЙ АНАЛИЗ.....72

ЗАКЛЮЧЕНИЕ.....78

БИБЛИОГРАФИЧЕСКИЙ СПИСОК.....79

тическою) – 3,4. Следовательно, нулевая гипотеза принимается – между наличием скидок и степенью интенсивности рекламы нет связи.

15.2. Задания

Маркетологи проверяют влияние пола респондента и частоты туристических поездок на предпочтения туристов (табл. 15.3). Половая принадлежность закодирована следующим образом: 1 – мужчины; 2 – женщины. Частота туристических поездок закодирована следующим образом: 1 – редко; 2 – периодически; 3 – часто. Предпочтения туристов измерялись по 9-балльной шкале, от 1 – полное отсутствие предпочтений, до 9 – безоговорочное предпочтение.

Проведите двухфакторный дисперсионный анализ в той же последовательности, что и в представленном примере.

Таблица 15.3

Номер респондента	Пол	Туристические поездки (частота)	Предпочтения	Номер респондента	Пол	Туристические поездки (частота)	Предпочтения
1	1	1	2	16	2	1	6
2	1	1	3	17	2	1	7
3	1	1	4	18	2	1	6
4	1	1	4	19	2	1	5
5	1	1	2	20	2	1	7
6	1	2	4	21	2	2	3
7	1	2	5	22	2	2	4
8	1	2	5	23	2	2	5
9	1	2	3	24	2	2	4
10	1	2	3	25	2	2	5
11	1	3	8	26	2	3	6
12	1	3	8	27	2	3	6
13	1	3	8	28	2	3	6
14	1	3	7	29	2	3	7
15	1	3	7	30	2	3	8

15.3. Контрольные вопросы

1) Что такое многофакторный дисперсионный анализ?
2) Перечислите этапы двухфакторного дисперсионного анализа.
3) Приведите пример для применения двухфакторного дисперсионного анализа.

УДК 339.138(07)
ББК 65.050.9(2)24я7
Ч-446

Рецензенты:
О. С. Малохатько, кандидат экономических наук,
генеральный директор ООО «Инженерно-технический центр»;
кафедра экономики ФГБОУ ВПО «Амурский гуманитарно-педагогический государственный университет», и. о. зав. кафедрой
кандидат педагогических наук, доцент Г. Т. Анурова

Чепухалина, Е. В.

Ч-446 Маркетинговые исследования (компьютерный практикум) : учеб. пособие / Е. В. Чепухалина. – Комсомольск-на-Амуре : ФГБОУ ВПО «КнАГТУ», 2012. – 80 с.

ISBN 978-5-7765-0968-1

В учебном пособии в краткой и доступной форме рассмотрены основные методы анализа маркетинговой информации, даны примеры и практические упражнения. С помощью данного пособия студенты закрепят теоретические знания и приобретут практические навыки анализа рыночной информации с использованием разнообразных методов, начиная с построения рядов распределения и нахождения разнообразных средних и заканчивая многофакторным анализом.

Учебное пособие предназначено для изучения предмета «Маркетинговые исследования» студентами специальности 080111 – «Маркетинг» всех форм обучения. Кроме того, учебное пособие может использоваться магистрантами, обучающимися по программе 8030001 – «Коммерческая деятельность на рынке товаров и услуг», а также студентами специальности 080301 – «Коммерция (торговое дело)» всех форм обучения при изучении дисциплины «Исследование рынка».

УДК 339.138(07)
ББК 65.050.9(2)24я7

ISBN 978-5-7765-0968-1

© ФГБОУ ВПО «Комсомольский-на-Амуре государственный технический университет», 2012

ВВЕДЕНИЕ 5

Практическое занятие 1
АНАЛИЗ ДАННЫХ. ПОСТРОЕНИЕ РЯДОВ РАСПРЕДЕЛЕНИЯ И НАХОЖДЕНИЕ СРЕДНИХ ЗНАЧЕНИЙ 6

Практическое занятие 2
РАСЧЕТ ПОКАЗАТЕЛЕЙ РАЗМЕРА И ИНТЕНСИВНОСТИ ВАРИАЦИЙ 11

Практическое занятие 3
ИСПОЛЬЗОВАНИЕ МЕТОДА ПАРНЫХ СРАВНЕНИЙ ДЛЯ ОЦЕНКИ ПОЛОЖЕНИЯ ФИРМЫ НА РЫНКЕ 15

Практическое занятие 4
ПРОГНОЗИРОВАНИЕ ПУТЕМ АНАЛИЗА ТРЕНДОВ 21

Практическое занятие 5
ИСПОЛЬЗОВАНИЕ МЕТОДА ОТНОШЕНИЯ К ДВИЖУЩЕМУ СРЕДНЕМУ ДЛЯ ЦЕЛЕЙ ПРОГНОЗА ВРЕМЕННОГО РЯДА ПРИ НАЛИЧИИ СЕЗОННОЙ КОМПОНЕНТЫ 25

Практическое занятие 6
ПОИСК ОПТИМАЛЬНОГО СЕГМЕНТА НА ОСНОВЕ СБАЛАНСИРОВАННОЙ СИСТЕМЫ ПОКАЗАТЕЛЕЙ 30

Практическое занятие 7
ИСПОЛЬЗОВАНИЕ МЕТОДА РАНГОВОЙ КОРРЕЛЯЦИИ ПРИ ОБРАБОТКЕ ЭКСПЕРТНЫХ ОЦЕНОК 34

Практическое занятие 8
ПРОВЕРКА ЗНАЧИМОСТИ РЕЗУЛЬТАТОВ МАРКЕТИНГОВЫХ ИССЛЕДОВАНИЙ.
ГИПОТЕЗЫ О ЧАСТОТЕ РАСПРЕДЕЛЕНИЯ 39

Практическое занятие 9
ПРОВЕРКА ЗНАЧИМОСТИ РЕЗУЛЬТАТОВ МАРКЕТИНГОВЫХ ИССЛЕДОВАНИЙ.
ГИПОТЕЗЫ О СРЕДНИХ ВЕЛИЧИНАХ 45

Практическое занятие 10
ПРИМЕНЕНИЕ МЕТОДОВ АНАЛИЗА ПАРНЫХ ВЗАИМОСВЯЗЕЙ МЕЖДУ ПРИЗНАКАМИ
В МАРКЕТИНГОВЫХ ИССЛЕДОВАНИЯХ.
ИЗУЧЕНИЕ ТЕСНОТЫ СВЯЗИ МЕЖДУ КАЧЕСТВЕННЫМИ И КОЛИЧЕСТВЕННЫМИ ПРИЗНАКАМИ 48

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ 5

Практическое занятие 1
АНАЛИЗ ДАННЫХ. ПОСТРОЕНИЕ РЯДОВ РАСПРЕДЕЛЕНИЯ И НАХОЖДЕНИЕ СРЕДНИХ ЗНАЧЕНИЙ 6

Практическое занятие 2
РАСЧЕТ ПОКАЗАТЕЛЕЙ РАЗМЕРА И ИНТЕНСИВНОСТИ ВАРИАЦИЙ 11

Практическое занятие 3
ИСПОЛЬЗОВАНИЕ МЕТОДА ПАРНЫХ СРАВНЕНИЙ ДЛЯ ОЦЕНКИ ПОЛОЖЕНИЯ ФИРМЫ НА РЫНКЕ 15

Практическое занятие 4
ПРОГНОЗИРОВАНИЕ ПУТЕМ АНАЛИЗА ТРЕНДОВ 21

Практическое занятие 5
ИСПОЛЬЗОВАНИЕ МЕТОДА ОТНОШЕНИЯ К ДВИЖУЩЕМУ СРЕДНЕМУ ДЛЯ ЦЕЛЕЙ ПРОГНОЗА ВРЕМЕННОГО РЯДА ПРИ НАЛИЧИИ СЕЗОННОЙ КОМПОНЕНТЫ 25

Практическое занятие 6
ПОИСК ОПТИМАЛЬНОГО СЕГМЕНТА НА ОСНОВЕ СБАЛАНСИРОВАННОЙ СИСТЕМЫ ПОКАЗАТЕЛЕЙ 30

Практическое занятие 7
ИСПОЛЬЗОВАНИЕ МЕТОДА РАНГОВОЙ КОРРЕЛЯЦИИ ПРИ ОБРАБОТКЕ ЭКСПЕРТНЫХ ОЦЕНОК 34

Практическое занятие 8
ПРОВЕРКА ЗНАЧИМОСТИ РЕЗУЛЬТАТОВ МАРКЕТИНГОВЫХ ИССЛЕДОВАНИЙ.
ГИПОТЕЗЫ О ЧАСТОТЕ РАСПРЕДЕЛЕНИЯ 39

Практическое занятие 9
ПРОВЕРКА ЗНАЧИМОСТИ РЕЗУЛЬТАТОВ МАРКЕТИНГОВЫХ ИССЛЕДОВАНИЙ.
ГИПОТЕЗЫ О СРЕДНИХ ВЕЛИЧИНАХ 45

Практическое занятие 10
ПРИМЕНЕНИЕ МЕТОДОВ АНАЛИЗА ПАРНЫХ ВЗАИМОСВЯЗЕЙ МЕЖДУ ПРИЗНАКАМИ
В МАРКЕТИНГОВЫХ ИССЛЕДОВАНИЯХ.
ИЗУЧЕНИЕ ТЕСНОТЫ СВЯЗИ МЕЖДУ КАЧЕСТВЕННЫМИ И КОЛИЧЕСТВЕННЫМИ ПРИЗНАКАМИ 48

ЗАКЛЮЧЕНИЕ

Учебное пособие «Маркетинговые исследования (компьютерный практикум)» является дополнением к учебным изданиям по маркетинговым исследованиям. В пособии дается теоретические основы и практические задания по основным темам дисциплины. Приобретенные знания пригодятся будущим специалистам маркетинга и коммерции при проведении маркетинговых исследований, при обработке и анализе собранной информации.

Первые два занятия посвящены простейшим методам обработки и анализа маркетинговых данных, которые, тем не менее, используются практически в каждом исследовании рынка – построению рядов распределения, нахождению разнообразных средних и расчету показателей размаха и размера вариации признака. На четвертом и пятом занятиях студенты получают навыки прогнозирования путем анализа тренда при наличии и отсутствии сезонной компоненты. Третье, шестое и седьмое занятия знакомят студентов с экспертными методами сбора и анализа информации. Умению выдвигать гипотезы, а также умению проверять значимость маркетинговых исследований посвящены восьмое и девятое занятия. Следующие два занятия обучают студентов анализу парных взаимосвязей между признаками: между качественными и количественными – в шестом и между количественными – в одиннадцатом. Занятия 12 и 13 посвящены важному навыку маркетолога – определению объема выборки по среднему и по доле. И два последних занятия посвящены важному и распространенному виду анализа маркетинговых данных – дисперсионному анализу: однофакторному и многофакторному.

В данном пособии представлен далеко не полный список методов анализа маркетинговой информации. Безусловно, конкретная рыночная ситуация и управленческая проблема предлагают требуют поиска и применения разных методов анализа. И помимо данного пособия, студентам можно порекомендовать изучить труды авторов, перечисленных в библиографическом списке.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Аакер, Д. Маркетинговые исследования / Д. Аакер, В. Кумар, Дж. Дэй. – 7-е изд. – СПб. : Питер, 2004. – 848 с.

2. Анурин, В. Маркетинговые исследования потребительского рынка / В. Анурин, И. Муромкина, Е. Евтушенко. – СПб. : Питер, 2004. – 270 с.

3. Беляевский, И. К. Маркетинговое исследование: информация, анализ, прогноз : учеб. пособие. – М. : Финансы и статистика, 2001. – 320 с.

4. Березин, И. С. Маркетинговые исследования. Как это делают в России / И. С. Березин. – М. : Вершина, 2005. – 432 с.

5. Божук, С. Г. Маркетинговые исследования / С. Г. Божук, Л. Н. Ковалик. – СПб. : Питер, 2004. – 304 с.

6. Власова, М. Л. Социологические методы в маркетинговых исследованиях : учеб. пособие для вузов / М. Л. Власова. – М. : Издательский дом ГУ ВШЭ, 2006. – 710 с.

7. Герасимов, Б. И. Маркетинговое исследование рынка : учеб. пособие / Б. И. Герасимов, Н. Н. Мозгов. – М. : ФОРУМ, 2009. – 336 с.

8. Голубков, Е. П. Маркетинговые исследования: теория, методология и практика : учебник / Е. П. Голубков. – 3-е изд., перераб. и доп. – М. : Изд-во «Финпресс», 2003. – 496 с.

9. Девятко, И. Ф. Методы социологического исследования / И. Ф. Девятко. – 4-е изд. – М. : КДУ, 2006. – 296 с.

10. Дойль, П. Маркетинг-менеджмент и стратегии / П. Дойль. – 3-е изд. – СПб. : Питер, 2003. – 544 с.

11. Дэвис, Джоэл Дж. Исследования в рекламной деятельности : теория и практика / Джоэль Дж. Дэвис. – М. : Издательский дом «Вильямс», 2003. – 864 с.

12. Зозулев, А. В. Маркетинговые исследования: теория, методология, статистика : учеб. пособие / А. В. Зозулев, С. А. Солнцев. – М. : Рыбари; Киев : Знання, 2008. – 643 с.

13. Каменева, Н. Г. Маркетинговые исследования : учеб. пособие / Н. Г. Каменева, В. А. Поляков. – М. : Вузовский учебник, 2005. – 439 с.

14. Котлер, Ф. Основы маркетинга / Ф. Котлер, Г. Армстронг, Д. Сондерс, В. Вонг. – 2-е европ. изд. – М.; СПб. : Издательский дом «Вильямс», 2003. – 944 с.

15. Ламбен, Ж.-Ж. Стратегический маркетинг / Ж.-Ж. Ламбен ; пер с фр. – СПб. : Наука, 1996. – 589 с.

16. Малхотра, Нэреш К. Маркетинговые исследования : практ. руководство / Нэреш К. Малхотра. – 4-е изд. ; пер. с англ. – М. : Издательский дом «Вильямс», 2007. – 1200 с.

В ячейку «Число строк для выборки» необходимо проставить число, равное количеству строк с одним и тем же значением фактора (в данном примере – 5: пять строк со значением фактора «скидка есть» и 5 строк со значением фактора «скидки нет»). После нажатия ОК на новом листе появится вся информация для получения аналогичных выводов (рис. 15.4).

	A	B	C	D	E	F	G	H
1	Двухфакторный дисперсионный анализ с повторениями							
2								
3	ИТОГИ	высокая	средняя	низкая	Итого			
4	*скидка есть*							
5	Счет	5	5	5	15			
6	Сумма	46	38	27	111			
7	Среднее	9,2	7,6	5,4	7,4			
8	Дисперсия	0,7	1,3	1,3	3,54285714			
9								
10	*скидки нет*							
11	Счет	5	5	5	15			
12	Сумма	37	24	10	71			
13	Среднее	7,4	4,8	2	4,73333333			
14	Дисперсия	1,3	0,7	0,5	5,92380952			
15								
16	*Итого*							
17	Счет	10	10	10				
18	Сумма	83	62	37				
19	Среднее	8,3	6,2	3,7				
20	Дисперсия	1,78889	3,06667	4,01111				
21								
22								
23	Дисперсионный анализ							
24	*Источник вариации*	*SS*	*df*	*MS*	*F*	*P-Значение*	*F критическое*	
25	Выборка	53,3333	1	53,3333	55,1724138	1,1439E-07	4,259677214	
26	Столбцы	106,067	2	53,0333	54,862069	1,1169E-09	3,402826105	
27	Взаимодействие	3,26667	2	1,63333	1,68965517	0,20580923	3,402826105	
28	Внутри	23,2	24	0,96667				
29								
30	Итого	185,867	29					
31								
32								

Рис. 15.4. Окно с итоговой информацией

Интерпретация результатов в окне итоговой информации.

1) Строка «Выборка» характеризует связь между скидками и объемом продаж. Расчетное значение F = 55,172 больше табличного (критического) – 4,26. Следовательно, нулевая гипотеза отвергается, принимается альтернативная – между наличием скидок и объемом продаж есть связь.

2) Строка «Столбцы» характеризует связь между рекламой и объемом продаж. Расчетное значение F = 54,86 больше табличного (критического) – 3,4. Следовательно, нулевая гипотеза отвергается, принимается альтернативная – между интенсивностью рекламы и объемом продаж есть связь.

3) Строка «Взаимодействие» отражает связь между двумя независимыми признаками. Расчетное значение F = 1,69 меньше табличного (кри-

ВВЕДЕНИЕ

Руководителю любого предприятия, функционирующего на конкурентном рынке, приходится принимать множество разнообразных управленческих решений: стоит ли включить в ассортиментную линейку данный товар, какой способ продвижения товара выбрать, в какой части города открыть магазин и т.д. Чтобы управленческое решение действительно было эффективным, необходим большой массив информации о рынке, получение которого невозможно без проведения маркетинговых исследований.

Важной частью маркетинговых исследований, влияющей на степень эффективности принятого впоследствии управленческого решения, является решение об объеме выборки объектов исследования, решение о применяемом методе обработки и анализа собранной информации. Будет ли это построение рядов распределения или необходим дисперсионный анализ, достаточно ли будет опросить 100 респондентов или выборка составит более 1000 человек, насколько важно учитывать при построении регрессионного уравнения тот или иной фактор – вот те вопросы, с которыми ежедневно сталкивается современный маркетолог. Ответы на них и должно помочь найти данное пособие.

Дисциплина «Маркетинговые исследования» является одной из дисциплин, формирующих профессиональную подготовку студентов коммерческих и маркетинговых специальностей. Предметом дисциплины является изучение теории и практики маркетинговых исследований. Цели изучения дисциплины – рассмотреть теоретические и практические основы проведения маркетинговых исследований, изучить этапы исследования рынка, освоить методы обработки и анализа собранной информации.

Учебное пособие «Маркетинговые исследования (коммерческий практикум)» включает основную часть теоретического и практического материала дисциплины «Маркетинговые исследования». С помощью данного пособия студенты закрепят теоретические знания и приобретут практические навыки обработки и анализа собранной маркетинговой информации.

Пособие включает 15 практических занятий. Каждое занятие содержит основные теоретические сведения в определенной области маркетинговых исследований. Получив эти знания, студент закрепляет их в процессе выполнения практических упражнений и при ответе на контрольные вопросы.

После изучения данного пособия студент должен:

- знать разнообразные методы и подходы к обработке и анализу собранной информации;
- уметь правильно использовать и применять приемы обработки и анализа информации о рынке;

прямых. Две крайние точки прямоугольников замыкаются по оси абсцисс на середины интервалов, в которых частоты равны нулю. Также вариационные ряды представляют в виде различных диаграмм: столбчатых, линейных, круговых.

Средние величины в статистическом понимании – это обобщающие показатели совокупности однотипных явлений по какому-либо количественному признаку. Определяют средние значения, чтобы ослабить влияние факторов на изучаемый показатель и получить сводный показатель, описывающий данную совокупность в целом.

Средние величины разделяются на *математические и структурные***.** К *математическим* средним относят среднеарифметическую (простая и взвешенная), среднегеометрическую и гармоническую, к *структурным* – моду и медиану.

Средняя геометрическая определяется по следующей формуле [21]:

$$\bar{x} = \sqrt[n]{x_1 * x_2 * ... * x_n}, \qquad (1.1)$$

где $x_1 \ldots x_n$ – значение признаков, представляющее собой, как правило, относительные величины динамики; *n* – количество значений признаков.

Средняя геометрическая применяется в тех случаях, когда индивидуальные значения признака представляют собой, как правило, относительные величины динамики, построенные в виде цепных величин, как отношение к предыдущему уровню каждого уровня в ряду динамики, т.е. характеризует средний коэффициент роста. Наиболее широкое применение средняя геометрическая получила для определения средних темпов изменения в рядах динамики, а также в рядах распределения, как показано в примере (табл. 1.4).

Таблица 1.4

Объем продаж, р.	Темп роста
2000	–
2500	1,25
4200	1,68

Применяя формулу (1.1) и пользуясь данными табл. 1.4, рассчитываем значение средней геометрической:

$$\bar{x} = \sqrt[2]{1{,}25 * 1{,}68} = 1{,}45.$$

Средняя гармоническая – величина, обратная средней арифметической из обратных значений признака [21]:

Таблица 15.1

Номер магазина	Степень интенсивности внутримагазинной рекламы	Наличие скидок	Объем продаж, тыс. р.	Номер магазина	Степень интенсивности внутримагазинной рекламы	Наличие скидок	Объем продаж, тыс. р.
1	Высокая	Есть	10	16	Высокая	Нет	8
2	Высокая	Есть	9	17	Высокая	Нет	9
3	Высокая	Есть	10	18	Высокая	Нет	7
4	Высокая	Есть	8	19	Высокая	Нет	7
5	Высокая	Есть	9	20	Высокая	Нет	6
6	Средняя	Есть	8	21	Средняя	Нет	4
7	Средняя	Есть	8	22	Средняя	Нет	5
8	Средняя	Есть	7	23	Средняя	Нет	5
9	Средняя	Есть	9	24	Средняя	Нет	6
10	Средняя	Есть	6	25	Средняя	Нет	4
11	Низкая	Есть	5	26	Низкая	Нет	2
12	Низкая	Есть	7	27	Низкая	Нет	3
13	Низкая	Есть	6	28	Низкая	Нет	2
14	Низкая	Есть	4	29	Низкая	Нет	1
15	Низкая	Есть	5	30	Низкая	Нет	2

Зависимая переменная (*признак, Y*) – объем продаж, тыс. р.

Независимые переменные (*факторы,* X_1 и X_2) – степень интенсивности внутримагазинной рекламы и наличие скидки.

Этап 2. Построение матрицы двухфакторного дисперсионного анализа (рис. 15.1).

	X_{11}	X_{12}	X_{13}	…	X_{1n}
X_{21}	Y_1	Y_1	Y_1	Y_1	Y_1
X_{22}	Y_2	Y_2	Y_2	Y_2	Y_2
…	…	…	…	…	…
X_{2n}	Y_n	Y_n	Y_n	Y_n	Y_n

Рис. 15.1. Матрица двухфакторного дисперсионного анализа

В нашем примере эта матрица будет выглядеть следующим образом (табл. 15.2).

Таблица 15.2

		Интенсивность рекламы		
		высокая	средняя	низкая
Скидки	Скидка есть	10	8	5
	Скидка есть	9	8	7
	Скидка есть	10	7	6
	Скидка есть	8	9	4
	Скидка есть	9	6	5
	Скидки нет	8	4	2
	Скидки нет	9	5	3
	Скидки нет	7	5	2
	Скидки нет	7	6	1
	Скидки нет	6	4	2

Этап 3. Выдвижение гипотез.

Так как в нашем примере изучается влияние двух факторов и, кроме того, возможно взаимодействие факторов между собой, то выдвигается три пары гипотез:

1) *Связь между интенсивностью рекламы и объемами продаж.*

- *нулевая гипотеза* – между интенсивностью рекламы и объемом продаж нет связи, т.е. средние значения объемов продаж в трех группах магазинов, разделенных по степени интенсивности рекламы, равны;
- *альтернативная гипотеза* – между интенсивностью рекламы и объемом продаж есть связь, т.е. средние значения объемов продаж в трех группах магазинов, разделенных по степени интенсивности рекламы, отличаются.

2) *Связь между наличием скидок и объемами продаж.*

- *нулевая гипотеза* – между наличием скидок и объемом продаж нет связи, т.е. средние значения объемов продаж в двух группах магазинов, разделенных по наличию скидок, равны;
- *альтернативная гипотеза* – между наличием скидок и объемом продаж есть связь, т.е. средние значения объемов продаж в двух группах магазинов, разделенных по наличию скидок, отличаются.

3) *Связь между наличием скидок и интенсивностью рекламы.*

- *нулевая гипотеза* – между наличием скидок и степенью интенсивности рекламы нет связи;
- *альтернативная гипотеза* – между наличием скидок и степенью интенсивности рекламы есть связь.

Этап 4. Проверка истинности нулевых гипотез с использованием Excel – сравнение расчетного значения F-статистики с табличным (критическим).

Для этого необходимо в «Анализе данных» выбрать функцию «Двухфакторный дисперсионный анализ с повторением» и в появившемся окне заполнить ячейку «входной интервал» (рис. 15.2).

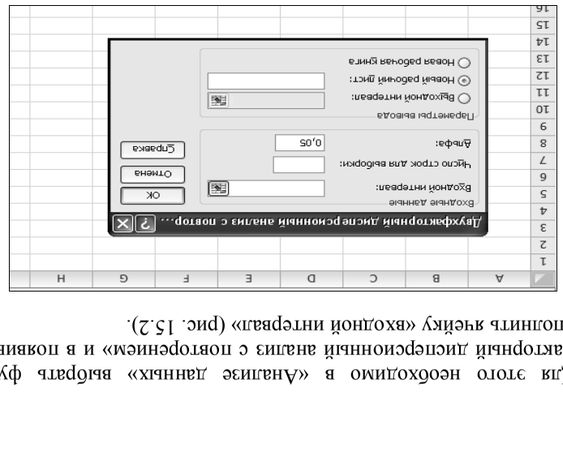


Рис. 15.2. Окно двухфакторного дисперсионного анализа

Входной интервал – это массив всех наблюдений из матрицы двухфакторного дисперсионного анализа, включая названия уровней независимых переменных (рис. 15.3).

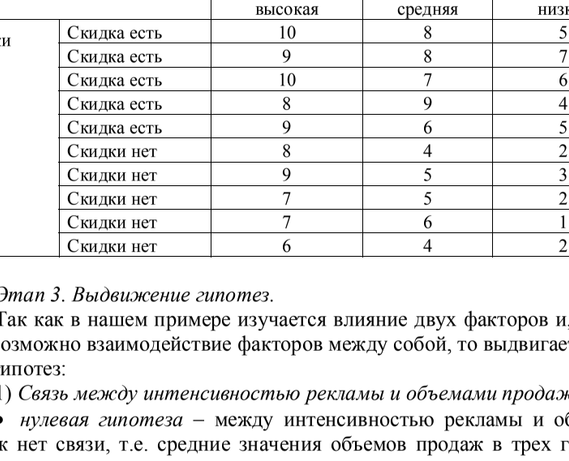


Рис. 15.3. Входной интервал

- иметь представление о технологии проведения маркетинговых исследований;
- владеть навыками расчета объема выборки;
- уметь использовать полученные знания в своей дальнейшей деятельности.

Изучение данного пособия является важным этапом в рамках освоения дисциплины «Маркетинговые исследования» при подготовке студентов специальности 080111 «Маркетинг» всех форм обучения. Кроме того, пособие может использоваться магистрантами, обучающимися по программе 030001 «Коммерческая деятельность на рынке товаров и услуг», а также студентами специальности 08030l «Коммерция (торговое дело)» всех форм обучения при изучении дисциплины «Исследование рынка».

Практическое занятие 1

АНАЛИЗ ДАННЫХ. ПОСТРОЕНИЕ РЯДОВ РАСПРЕДЕЛЕНИЯ И НАХОЖДЕНИЕ СРЕДНИХ ЗНАЧЕНИЙ

Цель работы: получить навыки построения рядов распределения и нахождения средних значений.

1.1. Основные сведения

Самой распространенной процедурой при анализе данных маркетинговых исследований является *группировка* – объединение единиц совокупности в некоторые группы, имеющие свои характерные особенности, общие черты и сходные размеры изучаемого признака [21].

Результатом группировки выступает ряд распределения (табл. 1.1). Для этого рассчитываются частота и частость элементов с определенным признаком. *Частота* – количество элементов совокупности, которые имеют значение данного признака. *Частость* – доля каждой выявленной группы в общей совокупности.

Таблица 1.1

Пол	Частота, чел.	Частость, %
Женский	65	75
Мужской	22	25
Всего	87	100

Различают два вида рядов распределения: атрибутивный и вариационный. В *атрибутивном ряду* распределение происходит по признаку, не имеющему количественной меры, например, пол потребителя (см. табл. 1.1), социальное положение, товар и т.д.

Вариационный ряд распределяет данные по признаку, имеющему количественную меру. Вариационные ряды разделяются на дискретные и интервальные.

В *дискретном* ряду признак отличается друг от друга некоторой величиной (табл. 1.2).

Таблица 1.2

Возраст, лет	Частота, чел.	Частость, %
20	18	21
21	13	15
22	20	23
23	15	17
24	10	11
25	11	13
Всего	87	100

В *интервальном ряду* значение признака представляется в виде интервалов (табл. 1.3). *В качестве основного показателя интервала используется середина интервала.*

Таблица 1.3

Возраст, лет	Частота, чел.	Частость, %
20 – 22	51	59
23 – 25	36	41
Всего	87	100

Для графического изображения рядов распределения применяется гистограмма (рис. 1.1).

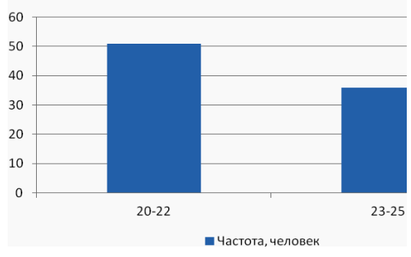


Рис. 1.1. Гистограмма рядов распределения

Гистограмма может быть преобразована в полигон распределения, если середины верхних сторон прямоугольников соединяются отрезками

14.3. Контрольные вопросы

1) Что такое дисперсионный анализ? Чем однофакторный дисперсионный анализ отличается от многофакторного?

2) Перечислите условия применения дисперсионного анализа.

3) Охарактеризуйте зависимую и независимую переменную. Приведите пример.

4) Перечислите этапы однофакторного дисперсионного анализа.

5) С помощью какого статистического коэффициента проверяется гипотеза в однофакторном дисперсионном анализе?

6) Какой коэффициент позволяет определить степень влияния независимой переменной на зависимую? О чем говорит значение коэффициента, равное 50 %?

Практическое занятие 15

МНОГОФАКТОРНЫЙ ДИСПЕРСИОННЫЙ АНАЛИЗ

Цель работы: освоить многофакторный дисперсионный анализ с использованием Excel.

15.1. Основные сведения

Многофакторный дисперсионный анализ – изучается влияние двух и более факторов (независимых переменных). Двухфакторный дисперсионный анализ имеется в стандартном наборе пакета «Анализ данных» программы Excel.

Рассмотрим на примере его применение.

Этапы выполнения двухфакторного дисперсионного анализа

Этап 1. Определение зависимой переменной (*признака*) *и двух независимых переменных* (*фактора*).

Изучается влияние интенсивности внутримагазинной рекламы и наличия скидок на объем продаж в магазине. Маркетологи выбрали три степени внутримагазинной рекламы: высокую, среднюю и низкую, а также два варианта со скидками: есть или нет. Случайным образом были выбраны 30 магазинов одной сети, также случайно были выбраны три группы по 10 магазинов, в которых в течение месяца применялась разная степень интенсивности внутримагазинной рекламы и две группы по 15 магазинов, в которых в течение месяца использовались или нет скидки (табл. 15.1).

$$\bar{x} = \frac{\sum M_i}{\sum(\frac{M_i}{x_i})},$$

где M_i – объемное значение признака; x_i – варианты значения признака.

Средняя гармоническая применяется, когда известны варианты признака, его объемное значение, но не известны частоты.

Рассмотрим это на примере. Необходимо найти среднюю цену товара на трех локальных рынках (табл. 1.5).

Таблица 1.5

Номер рынка	Выручка, тыс. р.	Цена единицы товара, тыс. р.
1	200	2
2	460	2,5
3	110	2,1

Среднеарифметическую (2.2 тыс. р.) находить в данном случае не совсем корректно, так как мы не знаем, сколько было продано данного товара на каждом рынке. Можно разделить выручку на цену и найти это количество, а затем рассчитать средневзвешенное (в качестве веса используя количество товара). Эти расчеты будут аналогичны нахождению средней гармонической:

$$\bar{x} = \frac{200+460+110}{\frac{200}{2}+\frac{460}{2,5}+\frac{110}{2,1}} = \frac{770}{336} = 2,3.$$

Мода — наиболее типичное, чаще всего встречаемое значение признака. В табл. 1.1, где представлен атрибутивный ряд распределения, модой является признак – женский пол, так как он встречается чаще всего. В случаях дискретных рядов модой считается значение признака с наибольшей частотой (в табл. 1.2 это количество респондентов в возрасте 22 года), а в интервальных рядах находится по формуле [6, 21]

$$M_0 = x_{M_0} + i_{M_0} * \frac{f_{M_0} - f_{M_{0-1}}}{(f_{M_0} - f_{M_{0-1}}) + (f_{M_0} - f_{M_{0+1}})}, \qquad (1.2)$$

где x_{M0} – нижняя граница модального интервала; i_{M0} – величина модального интервала; f_{M0}, f_{M0-1}, f_{M0+1} – частоты (частости) соответственно модального, домодального и послемодального интервалов.

Модальный интервал – это интервал, имеющий наибольшую частоту (частость) [6, 21]. Рассмотрим нахождение моды (табл. 1.6).

однородлен, чем второй. Для изучения данного отличия применяются показатели размера и интенсивности вариации.

Таблица 2.1

Признак	Совокупность 1	Совокупность 2
Возраст, лет	20	15
	21	20
	22	23
	23	27
	24	30
Средний возраст	23	23

Вариация признака (колеблемость признака) – различие индивидуальных значений признака внутри изучаемой совокупности в определенный момент времени [3, 13]. Чем больше в совокупности представлено признаков, тем больше вариация признака.

Показатели размера и интенсивности вариации применяются:

- при организации выборочного исследования (чем больше вариация признака внутри генеральной совокупности, тем больший объем должна иметь выборка);
- для сравнения двух совокупностей;
- для характеристики совокупности.

При изучении вариации рассчитывают: размер вариации; интенсивность вариации.

Для характеристики *размера вариации* в статистике применяются абсолютные показатели вариации: *размах вариации, среднее линейное отклонение, среднее квадратическое отклонение, дисперсия.*

Размах вариации (*размах колеблемости*) представляет собой разность между максимальным и минимальным значениями признака в совокупности:

$$R = X_{\max} - X_{\min},$$

где X_{max} – максимальное значение признака; X_{min} – минимальное значение признака.

Например, совокупность представлена десятью рынками, возраст которых колеблется от 20 до 25 лет, значит: $R = 25 - 20 = 5$.

Размах вариации зависит от величины только крайних значений признака. Чем больше признаков представлено в изучаемой совокупности, тем больше значение размаха вариации.

Для вариационных рядов с открытыми первым и последним интервалами, когда неизвестны реальные минимальное и максимальное значения признака в совокупности, расчет размаха вариации некорректен.

Проиллюстрируем расчет F-статистики для нашего примера. Сначала рассчитаем групповые средние, а также общее среднее (табл. 14.3).

Таблица 14.3

Номер магазина	Степень интенсивности внутримагазинной рекламы		
	высокая	средняя	низкая
1, 6, 11	10	8	5
2, 7, 12	9	8	7
3, 8, 13	10	7	6
4, 9, 14	8	9	4
5, 10, 15	9	6	5
Суммы по колонкам	46	38	27
Групповые средние	46/5 = 9,2	38/5 = 7,6	27/5 = 5,4
Общее среднее	(46+38+27)/15 = 7,4		

$$SS_x = 5*(9{,}2-7{,}4)^2 + 5*(7{,}6-7{,}4)^2 + 5*(5{,}4-7{,}4)^2 = 36{,}4,$$

$$SS_{ошибки} = (10-9{,}2)^2 + (9-9{,}2)^2 + ... + (8-7{,}6)^2 + (8-7{,}6)^2 + ... + (5-5{,}4)^2 = 13{,}2,$$

$$F = \frac{36{,}4/(3-1)}{13{,}2/(15-3)} = 16{,}5.$$

Критическое (табличное) значение для данных степеней свободы при уровне значимости 0,05 равно 3,89 [1, 16, 23]. А раз расчетное значение F-критерия больше, чем табличное, то нулевая гипотеза отклоняется и принимается альтернативная, которая утверждает, что между интенсивностью рекламы и объемом продаж есть связь, т.е. среднее значение объемов продаж в трех группах магазинов, различенных по степени интенсивности рекламы, отличается.

5) Рассчет степени влияния фактора на зависимую переменную.

Степень зависимости фактора на зависимую переменную определяется с помощью коэффициента *эта-квадрат*:

$$\eta^2 = \frac{SS_x}{SS_y} = \frac{SS_y - SS_{ошибки}}{SS_y},$$

$$SS_y = \sum(Y_i - \bar{Y})^2 = SS_x + SS_{ошибки},$$

где SS_y – вариация переменной Y, связанная с вариацией внутри каждой группы переменной X.

Для нашего примера:

$$SS_y = (10-7{,}4)^2 + ... + (5-7{,}4)^2 = 36{,}4 + 13{,}2 = 49{,}6,$$

$$\eta^2 = \frac{36{,}4}{49{,}6} = 0{,}73.$$

Это значит, что в 73 % случаев из 100 вариация в продажах есть следствие влияния внутримагазинной рекламы, оставшиеся 27 % – следствие влияния других факторов.

Аналогичный расчет можно выполнить с использованием программы Excel. Для этого необходимо в «Анализе данных» выбрать функцию «однофакторный дисперсионный анализ» и в появившемся окне заполнить ячейку «входной интервал» (рис. 14.1).

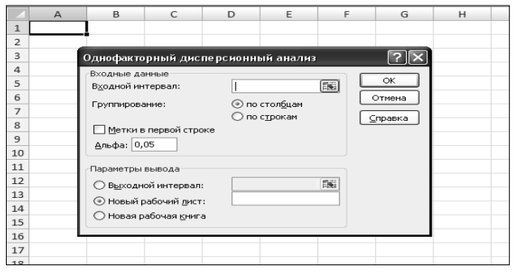


Рис. 14.1. Окно однофакторного дисперсионного анализа

Входной интервал – это массив всех наблюдений из матрицы дисперсионного анализа, которую и необходимо выделить (рис. 14.2).

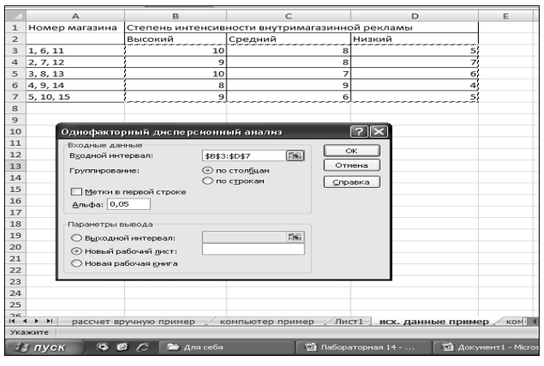


Рис. 14.2. Входной интервал

После нажатия ОК, на новом листе появится вся информация для получения аналогичных выводов (рис. 14.3).

	A	B	C	D	E	F	G
1	Однофакторный дисперсионный анализ						
2							
3	ИТОГИ						
4		*Группы*	*Счет*	*Сумма*	*Среднее*	*Дисперсия*	
5		Столбец 1	5	46	9,2	0,7	
6		Столбец 2	5	38	7,6	1,3	
7		Столбец 3	5	27	5,4	1,3	
8							
9							
10	Дисперсионный анализ						
11	Источник вариации	SS	df	MS	F	P-значение	F критическое
12	Между группами	36,4	2	18,2	16,54545455	0,000355266	3,885293835
13	Внутри групп	13,2	12	1,1			
14							
15	Итого	49,6	14				
16							
17							

Рис. 14.3. Окно с итоговой информацией

14.2. Задания

Маркетологи проверяли влияние скидок на объем продаж. Для этого случайным образом были выбраны 20 магазинов, в 10 из которых предоставлялись скидки, в других 10 магазинах нет. По результатам работы магазинов за месяц были измерены объемы продаж товара в штуках (табл. 14.4).

Таблица 14.4

Номер магазина	Наличие скидки	Объем продаж, шт.	Номер магазина	Наличие скидки	Объем продаж, шт.
1	Скидка имеется	115	11	Скидки нет	79
2	Скидка имеется	98	12	Скидки нет	82
3	Скидка имеется	101	13	Скидки нет	74
4	Скидка имеется	79	14	Скидки нет	93
5	Скидка имеется	92	15	Скидки нет	60
6	Скидка имеется	88	16	Скидки нет	45
7	Скидка имеется	91	17	Скидки нет	51
8	Скидка имеется	76	18	Скидки нет	49
9	Скидка имеется	72	19	Скидки нет	62
10	Скидка имеется	68	20	Скидки нет	40

Оцените, есть ли зависимость объема продаж от скидок и рассчитайте степень этой зависимости, проведя расчеты как вручную, так и с использованием Excel. Расчеты выполнять в той же последовательности, что и рассчитанный пример.

Таблица 1.6

Возраст, лет	Частота, чел.	Частость, %
20 – 21	31	36
22 – 23	**35**	**40**
24 – 25	21	24
Всего	87	100

Используя данные табл. 1.6 и формулу (1.2), находим значение моды:

$$M_0 = 22 + 2 * \frac{35 - 31}{(35 - 31) + (35 - 21)} = 22.$$

Это значит, что наиболее часто в выборке встречаются люди с возрастом 22 года.

Медиана – вариант, расположенный в середине упорядоченного вариационного ряда, делящий его на две равные части таким образом, что половина единиц совокупности имеют значения признака меньше, чем медиана, а половина – больше, чем медиана [6, 21]:

$$M_e = x_{M_e} + i_{M_e} * \frac{0{,}5 * (\sum f + 1) - S_{M_{e-1}}}{f_{M_e}} \qquad (1.3)$$

где x_{M_e} – начало медианного интервала; i_{M_e} – величина медианного интервала; $\sum f$ – сумма частот (частостей) вариационного ряда; f_{M_e} – частота (частость) медианного интервала; $S_{M_{e-1}}$ – сумма накопленных частот (частостей) в домедианном интервале.

Медианный интервал – это интервал, в котором находится порядковый номер медианы. Для его определения необходимо подсчитать сумму накопленных частот (частостей) до числа, превышающего половину объема совокупности (табл. 1.7).

Таблица 1.7

Возраст, лет	Частота, чел.	Накопленная частота, чел.
20 – 21	31	31
22 – 23	**35**	**66**
24 – 25	21	87
Всего	87	

Используя данные табл. 1.7, находим значение медианы по формуле (1.3)

$$M_e = 22 + 2 * \frac{0.5 * (87 + 1) - 31}{35} = 22.$$

Это значит, что примерно половина всех респондентов имеет возраст от 20 до 22 лет, другая половина – от 23 до 25 лет.

Так как медиана находится для упорядоченного ряда, то рассчитывается только для вариационных рядов распределения.

1.2. Задания

Необходимо построить ряды распределения (табличная форма и графическая), а также найти значения структурных средних для следующих признаков: пол, возраст респондента, доход на одного члена семьи респондента, род занятий респондента.

Оформить работу соответствующими выводами после каждого этапа расчета. По полученным данным описать целевой сегмент потребителей - по наиболее часто встречающимся признакам.

Анализируемый массив данных получить у преподавателя.

1.3. Контрольные вопросы

1) Что такое группировка данных? Что значит «построить ряд распределения»?

2) Перечислите виды рядов распределения. Охарактеризуйте каждый.

3) Перечислите типы математических средних. Охарактеризуйте каждый тип математической средней с описанием ситуации, когда применение каждой из них желательно.

4) Охарактеризуйте структурные средние. В чем разница между модой и медианой?

5) Какие признаки, помимо перечисленных в работе, можно использовать для описания целевых потребителей?

Практическое занятие 2

РАСЧЕТ ПОКАЗАТЕЛЕЙ РАЗМЕРА И ИНТЕНСИВНОСТИ ВАРИАЦИЙ

Цель работы: ознакомиться с показателями размера и интенсивности вариаций.

2.1. Основные сведения

Средние значения признака изучаемой совокупности не всегда корректно характеризуют эти совокупности. В табл. 2.1 представлены две совокупности, для которых рассчитаны средние значения признака – возраст респондентов. Среднее значение признака для обеих совокупностей одинаково, но разброс возрастов существенно отличается: первый ряд более

Таблица 14.2

Номер магазина	Степень интенсивности внутримагазинной рекламы		
	высокая	средняя	низкая
1, 6, 11	10	8	5
2, 7, 12	9	8	7
3, 8, 13	10	7	6
4, 9, 14	8	9	4
5, 10, 15	9	6	5

3) *Выдвижение гипотезы.*

В многомерном анализе (а дисперсионный анализ относится к многомерному) проверяются гипотезы о наличии или отсутствии связи.

Нулевая – предположение о том, что между определенными параметрами изучаемой совокупности ***не*** существует связи. *Отказ от нулевой гипотезы ведет к желательному результату.*

Альтернативная – предположение о том, что между параметрами изучаемой совокупности ***есть*** связь.

В нашем примере: так как изучаемым признаком в дисперсионном анализе является среднее значение, то:

- *нулевая гипотеза* – между интенсивностью рекламы и объемом продаж нет связи, т.е. средние значения объемов продаж в трех группах магазинов, разделенных по степени интенсивности рекламы, равны;
- *альтернативная гипотеза* – между интенсивностью рекламы и объемом продаж есть связь, т.е. средние значения объемов продаж в трех группах магазинов, разделенных по степени интенсивности рекламы, отличаются.

4) *Проверка истинности нулевой гипотезы* – сравнение расчетного значения F-статистики с табличным (критическим).

F-статистика находится по формуле

$$F = \frac{SS_x/(c-1)}{SS_{ошибки}/(N-c)},$$

где $$SS_x = \sum n*(\bar{Y}_j - \bar{Y})^2,$$

$$SS_{ошибки} = \sum\sum(Y_{ij} - \bar{Y}_j)^2,$$

где SS_x – вариация переменной Y, связанная с различиями средних между группами переменной X; $SS_{ошибки}$ – дисперсия ошибки; c – число уровней (групп) факторного признака; N – число наблюдений зависимой переменной.

Среднее линейное отклонение находится по формуле

$$d = \frac{\sum\left|x_i - \bar{x}\right| * f_i}{\sum f_i},$$

где x_i – значение признака в i-й группе; $\bar{x}$ – среднее значение признака; f_i – частота i-го признака (интервала); $\sum f_i$ – сумма частот i-го признака.

Если вариационный ряд является интервальным, то в качестве значения признака применяется его середина.

Покажем это на примере. Представлена совокупность респондентов в возрасте от 20 до 22 лет (табл. 2.2).

Таблица 2.2

Возраст, лет	Частота, чел.	$\left\|x_i - \bar{x}\right\|$	$\left\|x_i - \bar{x}\right\| * f_i$
20	30	0,66	19,66
21	35	0,34	12,07
22	21	1,34	28,24
Всего	87	-	59,97

Среднее значение признака:

$$\bar{x} = \frac{20*31 + 21*35 + 22*21}{87} = 21.$$

Среднее линейное отклонение:

$$d = \frac{59,97}{87} = 0,7 \text{ лет}.$$

Это значит, что в среднем возраст всех респондентов отличается от среднего на 0,7 лет.

Использование среднего линейного отклонения связано с определенными неудобствами – приходится иметь дело не только с положительными, но и отрицательными величинами. Поэтому более широкое применение получило использование *среднего квадратического отклонения*:

$$\sigma = \sqrt{\frac{\sum(x_i - \bar{x})^2 * f_i}{\sum f_i}}.$$

В нашем примере: σ = 0,80 лет.

Среднее квадратическое отклонение показывает, насколько в среднем возраст каждого респондента отличается от среднего возраста респондентов. В нашем примере – в среднем на 0,8 года от среднего возраста.

не только тогда, когда число объектов велико, но и в тех случаях, когда различия между объектами настолько малы, что непосредственное ранжирование или оценка не обеспечивают их разумного упорядочивания. Чаще всего при парном сравнении двух объектов отранжировать простой констатацией того, что один из них предпочтительнее другого.

Для оценки своей деятельности фирма должна проводить независимую оценку своего положения на рынке. Путем сравнения фирмы с конкурентами необходимо создать ранжированный ряд впереди ушедших и следующих за ней фирм.

Этапы оценки положения фирмы на рынке методом парных сравнений

Этап 1. Заполнение экспертных формы для сбора данных.

Экспертам предлагается попарно сравнить объекты. Пример формы для сравнения представлен в табл. 3.1. В данном случае идет сравнение четырех фирм. Каждый эксперт выражает одно из трех возможных мнений:

> – фирма, указанная в левом столбце, лучше, чем фирма, указанная в правом столбце;

< – фирма, указанная в левом столбце, хуже, чем фирма, указанная в правом столбце;

= – эксперт не может выбрать между фирмами, они для него равнозначны.

Таблица 3.1

Наименование фирмы	Сравнение (>, <, =)	Наименование фирмы
Фирма 1	>	Фирма 2
Фирма 1	>	Фирма 3
Фирма 1	=	Фирма 4
Фирма 2	<	Фирма 3
Фирма 2	>	Фирма 4
Фирма 3	=	Фирма 4

Этап 2. Построение матрицы сравниваемых фирм и первое «грубое» ранжирование.

Ответы экспертов колируются, т.е. каждому ответу присваиваются следующие значения (коды) сравнения:

0, если фирма i хуже фирмы j;

1, если фирма i равносильна j;

2, если фирма i лучше фирмы j.

Эту кодировку можно применить еще на первом этапе, предлагая экспертам сразу в форму для сбора данных занести коды сравнения, но это может вызвать затруднение и путаницу при заполнении таблиц.

max генеральной совокупности (когда объем выборки составляет 5 % и более от объема генеральной совокупности), он начинает оказывать влияние. Поэтому объем выборки корректируется по формуле, которая включает объем генеральной совокупности, что влечет обычно уменьшение объема выборки.

Очевидно, что при незначительных значениях N корректировка не осуществляется.

Продолжение примера 13.1. Пусть в нашем случае объем генеральной совокупности составил 500 человек, тогда объем выборки составляет более 18 % от объема генеральной совокупности и необходимо скорректировать объем по формуле

$$n = \frac{92*500}{500+92-1} = 78 \text{ чел}.$$

13.2. Задания

Предприятие хочет открыть платный детский сад и для этого собирается провести опрос среди потенциальных покупателей услуги.

Необходимо:

1) Рассчитать объем выборки тремя способами:

а) когда в качестве генеральной доли выбрана доля женщин в совокупности;

б) когда в качестве генеральной доли выбрана доля работающих женщин в совокупности;

в) когда в качестве генеральной доли выбрана доля работающих женщин и мужчин в совокупности, имеющих детей в возрасте до 6 лет включительно.

Коэффициент доверия равен 1,96, а предельная ошибка составляет 10 %.

На основе расчетов выбрать окончательный объем выборки.

2) Сформировать две простые выборки среди всей совокупности с использованием функции «случайное число». Сделать вывод о наиболее репрезентативной выборке. В качестве критерия выбрать рассчитанную ошибку по выборке: чем меньше ошибка, тем репрезентативнее метод. За генеральную долю брать долю тех респондентов, чьи характеристики соответствуют максимальной выборке.

Исходный массив данных по генеральной совокупности получить у преподавателя.

Среднее квадратическое отклонение по величине всегда больше среднего линейного отклонения. Обе эти величины измеряются в тех же значениях, что и сами изучаемые признаки (признак в годах, значит и оба отклонения в годах).

Дисперсия определяется по формуле

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2 * f_i}{\sum f_i}.$$

Дисперсия применяется при расчете объема выборки и для сравнения совокупностей. Единицы измерения дисперсия не имеет.

Для оценки *интенсивности вариации* чаще используют *коэффициент вариации:*

$$v = \frac{\sigma}{\bar{x}} * 100\%.$$

В нашем примере: v = 0,80/20,65 * 100 = 3,9 %.

По величине коэффициента вариации можно судить об однородности состава изучаемой совокупности (табл. 2.3). Чем больше коэффициент, тем больше разброс значений признака вокруг средней, тем больше неоднородность совокупности. В нашем примере совокупность однородная.

Таблица 2.3

Коэффициент вариации, %	Степень однородности совокупности
До 30	Однородная
30-60	Средняя
60 и более	Неоднородная

2.2. Задания

Взять анализируемый объем данных у преподавателя.

Выполнить следующие задания:

1) Рассчитать показатели размера и интенсивности вариации: по возрасту для всех респондентов; по доходу для всех респондентов; по возрасту для мужчин и женщин отдельно; по доходу для мужчин и женщин отдельно.

2) Дать характеристику всей совокупности по величинам показателей размера и интенсивности вариации.

3) Сравнить совокупности из мужчин и женщин по показателям размера и интенсивности вариации и сделать соответствующие выводы.

2.3. Контрольные вопросы

1) Что такое вариация? Для чего исследователю необходимо знать вариацию признака изучаемой совокупности?

2) Какие существуют показатели размера и интенсивности вариации? Что означают значения этих показателей?

3) Что можно сказать о совокупности, коэффициент вариации которой равен 65 %?

4) Что можно сказать о совокупности респондентов с возрастом от 40 до 50, если среднее квадратическое отклонение равно 2?

5) Определите, для какой совокупности объем выборки будет больший, если дисперсия изучаемого признака первой совокупности равна 256, а второй – 126.

Практическое занятие 3

ИСПОЛЬЗОВАНИЕ МЕТОДА ПАРНЫХ СРАВНЕНИЙ ДЛЯ ОЦЕНКИ ПОЛОЖЕНИЯ ФИРМЫ НА РЫНКЕ

Цель работы: освоить метод парных сравнений для расчета рейтинга фирмы на рынке.

3.1. Основные сведения

Ранжирование различных показателей очень часто используется в маркетинговых исследованиях, оно применяется тогда, когда объекты ранжирования нельзя оценить через количественные показатели. Эксперт лишь может указать, какой из объектов ранжирования лучше или хуже других. С помощью ранжирования можно оценивать различные объекты: марку, упаковку, товар, конкурентов, каналы сбыта и т.д.

В данной работе мы будем оценивать положение фирмы на рынке относительно конкурентов. Смысл ранжирования состоит в сравнении тех или иных параметров, характеризующих фирму и конкурентов, и придании каждому объекту определенного ранга (места) в ряду всех объектов. Но в случае, когда объектов для сравнения много, простое ранжирование затруднительно для эксперта. В этом случае лучше использовать *метод парных сравнений* [13, 22].

В методе парных сравнений объекты сопоставляются попарно экспертом (экспертами), а затем выбирается один из них. В общем случае эксперт может установить равенство объектов или зафиксировать свои предпочтения на некоторой шкале. Производить парное сравнение удобно

13.3. Контрольные вопросы

1) В чем отличие определения объема выборки для доли от определения объема выборки для среднего? Перечислите этапы определения объема выборки для доли.

2) Приведите пример исследования, когда необходимо рассчитать объем выборки для доли.

3) Если у совокупности имеется несколько признаков, важных для исследователя, то как в этом случае будет определяться объем выборки?

Практическое занятие 14

ОДНОФАКТОРНЫЙ ДИСПЕРСИОННЫЙ АНАЛИЗ

Цель работы: освоить однофакторный дисперсионный анализ (расчет без использования Excel и с его использованием).

14.1. Основные сведения

Дисперсионный анализ (ANOVA) – статистический метод изучения связи между переменными для двух и более совокупностей. Для этого изучаются различия между выборочными средними у двух и более совокупностей [1, 3, 8, 13, 16, 23].

Условия применения дисперсионного анализа:

1) Зависимая переменная (признак) должна быть метрической (измеренной с помощью интервальной или относительной шкалы).

2) Должна быть одна или более независимых переменных (факторов), которые должны быть неметрическими (измеренными с помощью номинальной или порядковой шкалы).

Однофакторный дисперсионный анализ – изучается влияние только одного фактора (независимой переменной) на зависимую.

Например, есть ли связь между полом студента (независимая переменная, номинальная неметрическая шкала) и его успеваемостью (зависимая переменная, относительная метрическая шкала)? Суть метода будет состоять в сравнении среднего значения признака (среднего балла успеваемости) для разных вариантов фактора (два варианта: мужчины и женщины).

Для более полного понимания однофакторного дисперсионного анализа на данном практическом занятии приведен пример его без применения Excel, а также с его использованием.

Этапы выполнения однофакторного дисперсионного анализа

1) *Определение зависимой переменной* (*признака*) *и независимой переменной* (*фактора*).

Изучается влияние интенсивности внутримагазинной рекламы на объем продаж в магазине. Маркетологи выбрали для внутримагазинной рекламы три степени интенсивности: высокую, среднюю и низкую. Случайным образом были выбраны 15 магазинов одной сети, также случайно были выбраны три группы по пять магазинов, в которых в течение месяца применялась разная степень интенсивности внутримагазинной рекламы (табл. 14.1).

Зависимая переменная (*признак, Y*) – объем продаж, тыс. р.

Независимая переменная (*фактор, X*) – степень интенсивности внутримагазинной рекламы: высокая, средняя, низкая (X_1, X_2, X_3).

Таблица 14.1

Номер магазина	Степень интенсивности внутримагазинной рекламы	Объем продаж, тыс. р.	Номер магазина	Степень интенсивности внутримагазинной рекламы	Объем продаж, тыс. р.
1	Высокая	10	9	Средняя	9
2	Высокая	9	10	Средняя	6
3	Высокий	10	11	Низкая	5
4	Высокая	8	12	Низкая	7
5	Высокая	9	13	Низкая	6
6	Средняя	8	14	Низкая	4
7	Средняя	8	15	Низкая	5
8	Средняя	7			

2) *Построение матрицы дисперсионного анализа* (рис. 14.1).

X_1	X_2	X_3	…	X_n
Y_1	Y_1	Y_1	Y_1	Y_1
Y_2	Y_2	Y_2	Y_2	Y_2
…	…	…	…	…
Y_n	Y_n	Y_n	Y_n	Y_n

Рис. 14.1. Матрица дисперсионного анализа

В нашем примере эта матрица будет выглядеть следующим образом (табл. 14.2).

Продолжение примера 13.1. Исследователь остановился на значении коэффициента доверия 1,96 при уровне достоверности 0,95. Это значит, что в 95 случаях из 100 выборка действительно будет отражать характеристики генеральной совокупности.

Этап 3. Определение генеральной доли.

Генеральная доля (π) – доля единиц генеральной совокупности, обладающей каким-либо признаком.

Генеральную долю можно определить несколькими способами:

1) установить на основе мнения исследователя;

2) использовать данные предыдущих исследований;

3) рассчитать, проведя предварительное пилотное исследование.

Продолжение примера 13.1. Исследователь воспользовался данными предыдущего исследования, которое он проводил год назад. Тогда дисперсия составила 60 % (0,6 в долях). Если бы этих данных у исследователя не было, тогда ему пришлось бы провести пилотное исследование (непрезентативное исследование на малой выборке) и на основе собранных данных рассчитать долю лояльных потребителей среди выборки, которую и принял бы за генеральную долю.

Этап 4. Определение объема выборки.

В данном случае (когда изучаемая совокупность представлена долей) формула для определения объема выборки выглядит следующим образом:

$$n = \frac{\pi * (1-\pi) * t^2}{\Delta^2}.$$

Продолжение примера 13.1.

Согласно вышеприведенной формуле

$$n = \frac{0{,}6 * (1-0{,}6) * 1{,}96^2}{0{,}10^2} = 92 \text{ чел.}$$

Этап 5. Окончательная коррекция объема выборки.

Если рассчитанный объем выборки превысит объем генеральной совокупности более чем на 5 %, то необходимо пересчитать объем выборки по формуле

$$n = \frac{n * N}{N + n - 1}.$$

где N – объем генеральной совокупности.

Хотя размер генеральной совокупности – далеко не самый важный фактор, влияющий на размер выборки, но при достаточно больших объе-

Затем для каждого эксперта заполняется матрица сравниваемых фирм Z_{ij} (табл. 3.2), где i – номер строки матрицы, j – номер столбца матрицы.

Таблица 3.2

		1	**2**	**3**	**4**	$\sum Z_{ij}$	P_{ij}
Z_{ij}	**1**	1	2	2	1	6	0,375
	2	0	1	0	2	3	0,1875
	3	0	2	1	1	4	0,25
	4	1	0	1	1	3	0,1875
$\sum\sum Z_{ij}$						**16**	**1,00**

Оценка i-й фирмы относительно других производится с использованием *коэффициента сравнения фирм:*

$$P_{ij} = \frac{\sum Z_{ij}}{\sum\sum Z_{ij}},$$

где $\sum Z_{ij}$ – построчное суммирование кодов сравнения.

Значения коэффициентов сравнения показывают, что фирма, чье значение коэффициента больше – лучше. Например, в нашем случае, фирма 1 лучшая среди всех фирм для данного эксперта, причем, она лучше фирмы 3 (второй по величине коэффициента сравнения) в 1,5 раза или на 50 % (0,375/0,25). Таким образом, данный метод позволяет не просто указать ранги фирм (лучше или хуже), но и определить, насколько лучше или хуже одна фирма относительно другой.

Если расположить фирмы в порядке убывания коэффициента сравнения, то получим первое «грубое» ранжирование фирм.

В нашем случае: 1-е место – фирма 1 (P_1 = 0,375), 2-е место – фирма 3 (P_3 = 0,25), 3-е и 4-е места – фирмы 2 и 4 (P_4 = P_2 = 0,1875).

Если в сравнении участвуют несколько экспертов, то прежде, чем определять ранг фирмы, необходимо рассчитать среднее арифметическое значений коэффициентов сравнения всех экспертов и только после этого проранжировать фирмы.

Для ускорения расчетов, при использовании программы Excel, в ячейки матрицы сравниваемых фирм можно ввести следующие формулы:

- в ячейки правого верхнего поля (выше единиц по диагонали) сделать ссылки на ячейки из преобразованной таблицы с кодами вместо знаков сравнения (рис. 3.1);
- в ячейках левого нижнего поля (ниже единиц по диагонали) использовать функцию «ЕСЛИ». Значения в этих ячейках зеркальны значениям в ячейках верхнего правого поля.

В нашем случае:

$$Z_2 = \begin{vmatrix} 2 & 8 & 5 & 8 \\ 2 & 1 & 2 & 4 \\ 1 & 4 & 2 & 6 \\ 2 & 4 & 4 & 3 \end{vmatrix} * \begin{vmatrix} 1 & 2 & 2 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 2 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 10 & 22 & 17 & 31 \\ 6 & 9 & 10 & 10 \\ 7 & 10 & 10 & 17 \\ 5 & 16 & 11 & 17 \end{vmatrix} = \begin{vmatrix} 80 \\ 35 \\ 44 \\ 49 \end{vmatrix} = \begin{vmatrix} 0.3846 \\ 0.1682 \\ 0.2115 \\ 0.2356 \end{vmatrix}.$$

Матрица второго уровня «тонкого» ранжирования уже меньше отличается от предыдущей, очередность мест практически сохранилась – порядок рейтингов практически стабилизировался. Чем больше производить умножение полученных матриц на исходную, тем точнее будут значения коэффициентов сравнения. Остановиться можно тогда, когда полученный результат (по рангам) совпадет с предыдущим этапом. В нашем случае результат несколько отличается (вывелись 2-е и 3-е места), поэтому лучше еще раз перемножить последнюю матрицу на исходную:

$$Z_3 = \begin{vmatrix} 10 & 22 & 17 & 31 \\ 6 & 9 & 10 & 10 \\ 7 & 10 & 10 & 17 \\ 5 & 16 & 11 & 17 \end{vmatrix} * \begin{vmatrix} 1 & 2 & 2 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 2 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 41 & 76 & 68 & 102 \\ 16 & 41 & 32 & 44 \\ 24 & 44 & 41 & 54 \\ 22 & 48 & 38 & 65 \end{vmatrix} = \begin{vmatrix} 287 \\ 133 \\ 163 \\ 173 \end{vmatrix} = \begin{vmatrix} 0.3796 \\ 0.1759 \\ 0.2156 \\ 0.2288 \end{vmatrix}.$$

Полученные результаты близки к предыдущему (по очередности рангов), поэтому стоит на этом остановиться.

3.2. Задания

Необходимо оценить (сравнить) пять фирм. В опросе участвовали три эксперта, оценки которых по каждой фирме представлены в табл. 3.3.

Таблица 3.3

Наименование фирмы	Сравнение (>, <, =)			Наименование фирмы
	Эксперт 1	Эксперт 2	Эксперт 3	
Фирма 1	>	=	<	Фирма 2
Фирма 1	>	=	=	Фирма 3
Фирма 1	>	<	<	Фирма 4
Фирма 1	>	=	>	Фирма 5
Фирма 2	=	>	=	Фирма 3
Фирма 2	<	>	<	Фирма 4
Фирма 2	<	=	<	Фирма 5
Фирма 3	<	>	>	Фирма 4
Фирма 3	<	>	>	Фирма 5
Фирма 4	<	<	=	Фирма 5

Продолжение табл. 12.1

1	*2*	*1*	*2*	*1*	*2*	*1*	*2*	*1*	*2*
7	995	27	1095	47	1272	67	1398	87	1141
8	997	28	990	48	969	68	920	88	999
9	1000	29	1200	49	1299	69	1089	89	1022
10	1059	30	998	50	1091	70	1101	90	907
11	1037	31	1298	51	1251	71	1070	91	1158
12	892	32	1000	52	900	72	989	92	988
13	809	33	899	53	930	73	880	93	999
14	1099	34	1111	54	1030	74	1020	94	1247
15	1212	35	850	55	905	75	1400	95	1023
16	965	36	1015	56	920	76	1399	96	1090
17	1098	37	1045	57	1190	77	1520	97	1050
18	987	38	1020	58	909	78	950	98	1005
19	1103	39	1211	59	1075	79	1198	99	1070
20	1061	40	1000	60	1089	80	1099	100	920

Необходимо выполнить следующее:

1) Рассчитать объем выборки для предоставленной совокупности. Допустимую погрешность принять по формуле для всей генеральной совокупности. Уровень достоверности 0,95, предельная ошибка составляет 5 % от средней величины изучаемого признака.

2) Сформировать две простые выборки предприятий с использованием функции Excel «случайное число». Сделать вывод о наиболее репрезентативной выборке. В качестве критерия выбрать рассчитанную ошибку по выборке: чем меньше ошибка, тем репрезентативнее метод.

Простая выборка с использованием функции «случайное число» формируется следующим образом:

- пронумеруйте все элементы генеральной совокупности от 1 до N и введите эти порядковые номера в первый столбец таблицы Excel (рис. 12.1);

	A	B	C
1	Номер	Признак	
2	1	100	
3	2	120	
4	3	50	
5	4	23	
6	5	56	
7	6	59	
8	7	12	
9	8	89	
10	9	250	
11	10	212	
12			
13			

Рис. 12.1. Нумерация элементов генеральной совокупности

• в верхнюю ячейку второго столбца введите формулу «=СЛЧИС( )» (рис. 12.2) и скопируйте эту ячейку вниз по столбцу, чтобы получить случайное число напротив каждого номера (рис. 12.3);

	A	B	C	D
1	Номер	Признак		
2	1	100	=СЛЧИС()	
3	2	120		
4	3	50		
5	4	23		
6	5	56		
7	6	59		
8	7	12		
9	8	89		
10	9	250		
11	10	212		
12				
13				

Рис. 12.2. Функция «случайное число»

	A	B	C	D
1	Номер	Признак		
2	1	100	0,571983	
3	2	120	0,625889	
4	3	50	0,463298	
5	4	23	0,101159	
6	5	56	0,781504	
7	6	59	0,810337	
8	7	12	0,715288	
9	8	89	0,884035	
10	9	250	0,205261	
11	10	212	0,279615	

Рис. 12.3. Результат копирования

	A	B	C	D
1	Номер	Признак		
2	6	59	0,010451	
3	1	100	0,087305	
4	9	250	0,253254	
5	8	89	0,419111	
6	2	120	0,52431	
7	4	23	0,541682	
8	5	56	0,576079	
9	3	50	0,58726	
10	7	12	0,76061	
11	10	212	0,874239	
12				

Рис. 12.4. Числа, упорядоченные случайным образом

• скопируйте полученный массив на новый лист (специальная вставка, только значения) и отсортируйте его по столбцу со случайными числами. После этого числа в первом столбце будут упорядочены случайным образом (рис. 12.4), и для получения искомой случайной выборки достаточно будет взять первые n элементов (n – объем выборки).

12.3. Контрольные вопросы

1) Что такое выборка? Что такое генеральная совокупность?
2) Почему не всегда проводят сплошное наблюдение?
3) Какие Вы знаете подходы для определения объема выборки?
4) Перечислите этапы определения объема выборки для среднего.
5) Как увеличение задаваемой ошибки влияет на объем выборки?
6) Что измеряет дисперсия и как её величина влияет на объем выборки?

суммированию строк и построение нового ряда значений коэффициентов сравнения. При «тонком» ранжировании коэффициенты сравнения принимают более четкие значения, и рейтинг фирм выглядит более убедительно.

В нашем случае:

$$Z_1 = \begin{vmatrix} 1 & 2 & 2 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 2 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{vmatrix} * \begin{vmatrix} 1 & 2 & 2 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 2 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 8 & 5 & 8 \\ 2 & 1 & 2 & 4 \\ 1 & 4 & 2 & 6 \\ 2 & 4 & 4 & 3 \end{vmatrix} = \begin{vmatrix} 23 \\ 9 \\ 13 \\ 13 \end{vmatrix} = \begin{vmatrix} 0.3966 \\ 0.1552 \\ 0.2241 \\ 0.2241 \end{vmatrix}.$$

После «тонкого» ранжирования места распределились следующим образом: 1-е место – фирма 1 (0,3966), 2-е и 3-е места – фирмы 3 и 4 (0,2241), 4-е место – фирма 2 (0,1552).

Как видно, изменилось и положение фирм относительно друг друга, и отношение коэффициентов сравнения.

При перемножении матриц можно использовать возможности Excel. Для этого необходимо выполнить следующие действия:

• расположить две матрицы одну под другой (рис. 3.3);

	A	B	C	D	E	F
1	A	1	2	2	1	
2		0	1	0	2	
3		0	2	1	1	
4		1	0	1	1	
5	Б	1	2	2	1	
6		0	1	0	2	
7		0	2	1	1	
8		1	0	1	1	
9		=МУМНОЖ(B1:E4;B5:E8)				

Рис. 3.3. Перемножение матриц

• ввести в ячейку B9 формулу: =МУМНОЖ(B1:E4;B5:E8);

• последовательно выполнить следующие действия: выделить диапазон B9:E12 (соответствует новой матрице); нажать F2; нажать CTRL+SHIFT+ВВОД. В результате этих действий в ячейках B9:E12 появится новая матрица (рис. 3.4).

Полученной формой можно пользоваться для расчета произведения матриц с тем же количеством строк и столбцов, подставляя новые значения матриц в имеющиеся поля.

Существенные изменения после перемножения матриц должны стать сигналом к последующим преобразованиям – умножению новой матрицы на первоначальную.

	A	B	C	D	E
1	A	1	2	2	1
2		0	1	0	2
3		0	2	1	1
4		1	0	1	1
5	Б	1	2	2	1
6		0	1	0	2
7		0	2	1	1
8		1	0	1	1
9		2	8	5	8
10		2	1	2	4
11		1	4	2	6
12		2	4	4	3
13					

Рис. 3.4. Получение новой матрицы

СТАТИСТИЧЕСКИЙ МЕТОД ОПРЕДЕЛЕНИЯ ОБЪЕМА ВЫБОРКИ С ПОМОЩЬЮ ДОЛИ

Цель работы: освоить технику определения объема выборки в рамках статистического подхода с помощью доли.

13.1. Основные сведения

Изучаемая совокупность может быть представлена как средним значением, так и долей. Например, исследователь пытается оценить средний доход потребителей – это средняя, а если исследователя интересует, сколько мужчин и сколько женщин являются потребителями, – это доля.

В данной работе представлены этапы расчета выборки для совокупности, представленной долей [1, 3, 8, 13, 16, 23].

Пример 13.1. Исследователь хочет определить долю лояльных потребителей. Для этого ему нужно определиться с объемом выборки.

Этап 1. Определение степени точности вычисления объема выборки.

Степень точности вычисления объема выборки оценивается посредством *предельной* (*заданной исследователем*) *ошибки выборки* (Δ) – максимально допустимое различие между выборочным средним и генеральным средним. Чем большую ошибку задает исследователь, тем меньший объем будет иметь выборка/

Продолжение примера 13.1. Исследователь хочет, чтобы полученный при выборочном исследовании результат (доля лояльных потребителей) расходился со средним значением генеральной совокупности не более чем на 10 % (0,10 в долях), т.е. в диапазоне от ± 10 %.

Для более точного результата исследователь может снизить предельную ошибку выборки, например, до 5 %, но это приведет к увеличению объема выборки – чем меньше ошибка, тем больше точность, тем больше объем выборки. Часто исследователь рассчитывает несколько значений объема выборки при разных значениях предельной ошибки и выбирает окончательное значение исхода из времени и средств, имеющихся для проведения исследования.

Этап 2. Определение коэффициента доверия (*нормированного отклонения*).

Коэффициент доверия (t) зависит от уровня достоверности. В маркетинговых исследованиях уровень достоверности обычно задается равным 0,95, при таком значении коэффициент доверия равен 1,96 [1, 3, 8, 13, 16, 23].

	A	B	C	D	E	F	G	H	I
1	Наименование фирмы	Сравнение (>,<,=)	Наименование фирмы						
2	Фирма 1	2	Фирма 2						
3	Фирма 1	2	Фирма 3						
4	Фирма 1	1	Фирма 4						
5	Фирма 2	0	Фирма 3						
6	Фирма 2	2	Фирма 4						
7	Фирма 3	1	Фирма 4						
8		Фирмы	1	2	3	4	ΣZ_{ij}	P_{ij}	
9		1	1	2	2	1	6	0.375	
10		2	=ЕСЛИ(D9=2;"0";ЕСЛИ(D9=0;"2";ЕСЛИ(D9=1;"1")))						
11		3	0	2	1	1	4	0.25	
12	Z_{ij}	4	1	0	1	1	3	0.1875	
13	$\Sigma\Sigma Z_{ij}$						16	1.00	
14									

Рис. 3.2. Подстановка функции «ЕСЛИ»

Этап 3. «Тонкое» ранжирование фирм.

Эмпирическим путем определено, что первое «грубое» ранжирование не всегда дает корректные результаты. Поэтому на этом этапе осуществляется перемножение матрицы сравнения саму на себя с получением новой матрицы. С новой матрицей проводится аналогичная работа по

	A	B	C	D	E	F	G	H
1	Наименование фирмы	Сравнение (>,<,=)	Наименование фирмы					
2	Фирма 1	2	Фирма 2					
3	Фирма 1	2	Фирма 3					
4	Фирма 1	1	Фирма 4					
5	Фирма 2	0	Фирма 3					
6	Фирма 2	2	Фирма 4					
7	Фирма 3	1	Фирма 4					
8		Фирмы	1	2	3	4	ΣZ_{ij}	P_{ij}
9		1	1	=B2	2	1	6	0.375
10		2	0	1	0	2	3	0.1875
11		3	0	2	1	1	4	0.25
12	Z_{ij}	4	1	0	1	1	3	0.1875
13	$\Sigma\Sigma Z_{ij}$						16	1.00
14								

Рис. 3.1. Ссылки на ячейки преобразованной таблицы

Например, значение ячейки C10 зеркально (противоположно) значению ячейки D9. Следовательно, в ячейку C10 необходимо ввести функцию (рис. 3.2):

=ЕСЛИ(D9=2;"0";ЕСЛИ(D9=0;"2";ЕСЛИ(D9=1;"1"))).

Теперь, подставляя в преобразованную таблицу с кодами новые значения, мы автоматически получаем значения коэффициентов сравнения.

Этап 5. Окончательная коррекция объема выборки.

Если рассчитанный объем выборки превысит объем генеральной совокупности более чем на 5 %, то необходимо пересчитать объем выборки по формуле

$$n = \frac{n * N}{N + n - 1},$$

где N – объем генеральной совокупности.

Хотя размер генеральной совокупности – далеко не самый важный фактор, влияющий на размер выборки, но при достаточно больших объемах генеральной совокупности (когда объем выборки составляет 5 % и более от объема генеральной совокупности), он начинает оказывать влияние. Поэтому объем выборки корректируется по формуле, которая включает объем генеральной совокупности, что влечет обычно уменьшение объема выборки. *Очевидно, что при неизвестных значениях N корректировка не осуществляется.*

Продолжение примера 12.1. Пусть в нашем случае объем генеральной совокупности составил 500 человек, тогда объем выборки составляет более 15 % от объема генеральной совокупности и необходимо скорректировать объем по формуле

$$n = \frac{77 * 500}{500 + 77 - 1} = 67 \text{ чел.}$$

12.2. Задания

Исследователь должен сформировать выборку из розничных магазинов, если известно, что всего таких предприятий 100, а в качестве признака изучаемой совокупности был выбран товарооборот розничного магазина за месяц (табл. 12.1). Сплошной опрос невозможен по причине ограниченности средств и времени.

Таблица 12.1

Номер магазина	Товарооборот, тыс. р.	Номер магазина	Товарооборот, тыс. р.	Номер магазина	Товарооборот, тыс. р.	Номер магазина	Товарооборот, тыс. р.	Номер магазина	Товарооборот, тыс. р.
1	*2*	*1*	*2*	*1*	*2*	*1*	*2*	*1*	*2*
1	1163	21	1072	41	1257	61	1222	81	1034
2	955	22	850	42	999	62	1099	82	979
3	1102	23	920	43	989	63	1000	83	960
4	999	24	1050	44	970	64	1111	84	1009
5	1060	25	1301	45	972	65	999	85	1020
6	1020	26	1140	46	1301	66	1200	86	888

3.3. Контрольные вопросы

1) Опишите суть метода парных сравнений. Почему и в каких случаях используют парное сравнение?

2) Опишите последовательность метода парных сравнений.

3) Чем «грубое» ранжирование отличается от «тонкого»? С какой целью осуществляется «тонкое» ранжирование?

4) Если бы Вы выявляли лучший развлекательный клуб города, по каким критериям Вы бы оценивали объекты и кто выступал бы в качестве экспертов?

5) Где, на Ваш взгляд, возможно применение метода парных сравнений?

Практическое занятие 4

ПРОГНОЗИРОВАНИЕ ПУТЕМ АНАЛИЗА ТРЕНДОВ

Цель работы: освоить методику анализа трендов для прогнозирования различных показателей.

4.1. Основные сведения

В условиях рыночного окружения, когда на предприятие воздействует изменяемая внешняя среда, одним из важных направлений маркетинговой деятельности является прогнозирование. Главная задача прогноза – получение некоторого предвидения будущего, связанного непосредственно с деятельностью предприятия.

Маркетинговые прогнозы имеют дело с рыночными факторами, оказывающими влияние на результаты деятельности предприятия. Основными составляющими прогнозирования являются: прогноз рыночных тенденций; прогноз развития отрасли; прогноз развития регионов и страны в целом; прогноз изменения рыночной конъюнктуры; прогноз спроса на товары и услуги, производимые предприятием и конкурентами; прогноз изменений у конкурентов; прогноз спроса и др. Задача исследователя при составлении прогноза: выбрать (или разработать) подходящую модель прогнозирования.

Метод анализа трендов относится к количественным методам прогнозирования, т.е. основан на количественных расчетах [2, 3, 19]. *Тренд* – это повышение или понижение каких-либо показателей за достаточно большой промежуток времени. Если показатели имеют тенденцию к повышению, то тренд называется восходящим, если к понижению, то нисходящим.

Среднее абсолютное отклонение находится по формуле

$$d = \frac{\sum_{i=1}^{n} |y_i - y_i|}{n}.$$

Этап 5. Прогноз показателей на будущее период. Путем подстановки в регрессионное уравнение вместо t номеров будущих периодов получают величины прогнозируемых показателей.

Этап 6. Проверка качества модели прогнозирования. Проверку качества осуществляют двумя путями: 1) сравнивают полученные прогнозы с действительными показателями будущих периодов, для чего придется подождать несколько лет; 2) пользуются *методом расщепления данных*.

Этапы метода расщепления данных:

1) построение модели тренда (уравнения регрессии), используя только часть исторических данных (например 7, а не 10 периодов);

2) с помощью полученного уравнения делают прогноз на неиспользованные при моделировании периоды (три в данном случае);

3) вычисляют отклонения прогнозных значений от исторических (y_i - f_i) в эти периоды (в данном случае таких отклонений будет три). Поскольку сохраняется правило минимальности этих отклонений. Кроме того, необходимо обратить внимание на знаки отклонений. Если все отклонения имеют один знак, то это говорит о том, что тренд отклоняется в одну сторону от кривой с истинными данными и может говорить о неадекватах выбранной модели;

4) рассчитывают величину *уклонения модели* (УМ) как среднеарифметическую полученных ранее отклонений (с учетом знаков):

$$УМ = \frac{\sum (y_i - f_i)}{n}.$$

Правила проверки по УМ те же, что и для отклонений.

4.2. Задания

Имеются данные о годовых объемах продаж четырех подразделений торговой компании за 16 лет (табл. 4.1).

Следать прогноз работы каждого подразделения торговой компании на ближайшие три года в точности с описанными этапами прогнозирования путем анализа трендов.

Таблица 11.5

Компания	Ранг качества подготовки персонала	Ранг рыночной доли	Компания	Ранг качества подготовки персонала	Ранг рыночной доли
А	6	3	К	4	4
В	12	6	Л	9	15
С	2	1	М	10	13
Д	5	7	Н	3	12
Е	14	8	О	11	11
Ж	1	9	П	8	10
З	15	2	Р	7	5
И	13	15			

2) Оцените тесноту связи между количественными признаками: ценой товара и объемом его продаж (табл. 11.6).

Таблица 11.6

Магазин	Объем продаж товара, тыс. р.	Цена товара, р.	Магазин	Объем продаж товара, тыс. р.	Цена товара, р.
1	1500	10,2	8	1700	9,8
2	1100	11	9	1200	10,5
3	950	12	10	1300	10,0
4	1000	12	11	1100	10,0
5	980	12	12	1200	11,0
6	975	12	13	1000	12,5
7	1200	10	14	1080	10,8

11.3. Контрольные вопросы

1) Зачем определяют тесноту связи между признаками?

2) Какие существуют подходы к оценке тесноты связи между признаками? От чего зависит выбор того или иного подхода?

3) Назовите показатели оценки тесноты связи между количественными признаками для разных шкал признака.

4) Приведите примеры применения показателей тесноты связи: между количественными признаками для порядковой шкалы; между количественными признаками для относительной шкалы.

Практическое занятие 12

СТАТИСТИЧЕСКИЙ МЕТОД ОПРЕДЕЛЕНИЯ ОБЪЕМА ВЫБОРКИ С ПОМОЩЬЮ СРЕДНЕГО

Цель работы: освоить технику определения объема выборки в рамках статистического подхода с помощью среднего.

12.1. Основные сведения

Изучаемая совокупность может быть представлена как средним значением, так и долей. Например, исследователь пытается оценить средний доход потребителей – это среднее, а если исследователя интересует, сколько мужчин и сколько женщин является потребителями – это доля. В данной работе представлены этапы расчета выборки для совокупности, представленной средним [1, 3, 8, 13, 16, 23].

Пример 12.1. Исследователь хочет определить ежемесячный расход потребителей на покупку продуктов. Для этого ему нужно определиться с объемом выборки.

Этап 1. Определение степени точности вычисления объема выборки.

Степень точности вычисления объема выборки оценивается посредством *предельной* (*заданной исследователем*) *ошибки выборки* (Δ) – максимально допустимое различие между выборочным средним и генеральным средним. Чем большую ошибку задает исследователь, тем меньший объем будет иметь выборка

Продолжение примера 12.1. Исследователь хочет, чтобы полученный при выборочном исследовании результат (ежемесячный расход потребителей на покупку продуктов) расходился со средним значением генеральной совокупности не более чем на 10 р., т.е. в диапазоне от ±10 р. Значит предельная (задаваемая исследователем) ошибка выборки будет равна 10 р. Для более точного результата исследователь может снизить предельную ошибку выборки, например, до 5 р., но это приведет к увеличению объема выборки – чем меньше ошибка, тем больше точность, тем больше объем выборки. Часто исследователь рассчитывает несколько значений объема выборки при разных значениях предельной ошибки и выбирает окончательное значение исходя из времени и средств, имеющихся для проведения исследования.

Этап 2. Определение коэффициента доверия (*нормированного отклонения*).

Коэффициент доверия (t) зависит от уровня достоверности. В маркетинговых исследованиях уровень достоверности обычно задается равным 0,95, при таком значении коэффициент доверия равен 1,96 [12, 17, 18].

Продолжение примера 12.1. Исследователь остановился на значении коэффициента доверия 1,96 при уровне достоверности 0,95. Это значит, что в 95 случаях из 100 выборка действительно будет отражать характеристики генеральной совокупности.

Этап 3. Определение дисперсии изучаемого признака.

Дисперсия изучаемого признака (σ^2) показывает, насколько вариируется (колеблется) изучаемый признак. Чем более вариативен признак, тем больше дисперсия и тем больше объем выборки. Дисперсия не имеет единиц измерения.

Дисперсию можно определить несколькими способами: установить на основе мнения исследователя; использовать данные предыдущих исследований; рассчитать по формуле (при предварительном проведении пилотного исследования):

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2 * f_i}{\sum f_i},$$

где x_i – значение признака у i-го респондента; $\bar{x}$ – среднее значение признака; f_i – частота i-го признака; $\sum f_i$ – сумма частот i-го признака.

Продолжение примера 12.1. Исследователь воспользовался данными предыдущего исследования, которое он проводил год назад. Тогда дисперсия составила 2000. Если бы эти данных у исследователя не было, тогда ему пришлось бы провести пилотное исследование (нерепрезентативное исследование на малой выборке) и на основе собранных данных рассчитать объем выборки. Значением признака каждого респондента был бы ежемесячный расход потребителей на покупку продуктов.

Этап 4. Определение объема выборки.

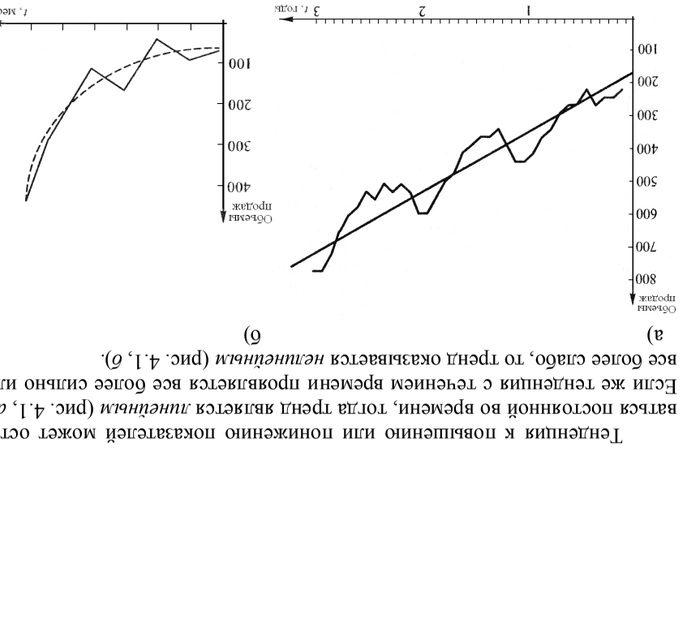
В данном случае (когда изучаемая совокупность представлена средними значениями) формула для определения объема выборки выглядит следующим образом:

$$n = \frac{t^2 * \sigma^2}{\Delta^2}.$$

Продолжение примера 12.1.

В соответствии с вышеприведенной формулой

$$n = \frac{1{,}96^2 * 2000}{10^2} = 77 \text{ чел.}$$



а) б)

Рис. 4.1. Линейный (а) и нелинейный (б) тренды

Технология прогнозирования путем анализа трендов признана:

- установить наличие тренда в показателях (исторических данных, представленных в виде временного ряда);
- смоделировать тренд;
- дать прогноз показателей будущего периода.

Этапы прогнозирования путем анализа трендов

Этап 1. Решение о виде тренда. Используя ретроспективные данные показателей, прогноз которых исследователь собирается проводить, необходимо принять решение о типе тренда. Для этого необходимо расположить точки, характеризующие имеющиеся данные в системе координат, где на оси абсцисс откладывают периоды, а на оси ординат – прогнозируемый показатель. По совокупности полученных точек принимается решение о типе тренда: линейный или нелинейный.

Этап 2. Моделирование тренда. После определения типа тренда используется с помощью регрессионного анализа строит модель тренда (регрессионное уравнение).

Уравнение регрессии для линейного тренда:

$$f_t = a + b * t, \qquad (4.1)$$

где f_t – значение тренда на оси ординат в момент времени t; a – точка, в которой тренд пересекает ось ординат; b – коэффициент наклона линии тренда.

Точка, в которой тренд пересекает ось ординат, вычисляется по формуле

$$a = \frac{\sum_{t=1}^{n} y_t - b * \sum_{t=1}^{n} t}{n}.$$

Коэффициент наклона линии тренда вычисляется по формуле

$$b = \frac{\left(\sum_{t=1}^{n} t * y_t\right) - \frac{\left(\sum_{t=1}^{n} y_t\right) * \left(\sum_{t=1}^{n} t\right)}{n}}{\left(\sum_{t=1}^{n} t^2\right) - \frac{\left(\sum_{t=1}^{n} t\right)^2}{n}},$$

где y_t – фактическое значение показателя в историческом ряду в момент времени t.

Этап 3. Графическое изображение тренда. Расположить на графике кривую, описывающую исторические данные прогнозируемого показателя, и кривую тренда (используя уравнение (4.1)).

Этап 4. Проверка регрессионного уравнения на адекватность. Данная оценка включает расчет суммы квадратов отклонений, среднеквадратического или абсолютного отклонения.

Принято считать наилучшим такой тренд (т.е. уравнение с такими коэффициентами a и b), при котором *сумма квадратов отклонений* исторических данных относительно линии тренда оказывается минимальной:

$$\sum_{t=1}^{n} \left(y_t - f_t\right)^2 \Rightarrow \min.$$

Отклонения возводятся в квадрат, так как могут быть не только положительными, но и отрицательными. *Для оптимизированного тренда сумма всех отклонений всегда равна* 0.

Аналогично выполняются правила минимальности среднеквадратического и абсолютного (линейного) отклонений, говорящие о высоком качестве прогноза и, следовательно, неплохом качестве построенной модели тренда.

Среднеквадратическое отклонение находится по формуле

$$\sigma = \sqrt{\frac{\sum_{t=1}^{n} \left(y_t - f_t\right)^2}{n}}.$$

Таблица 5.3

Сезон	Период t	Объем продаж y_t, тыс. р.	*ДС*	$ДС_ц$ $(f_t * c_t)$
Зима – весна	1	40		
			30	
Лето – осень	2	20		30,5
			31	
Зима – весна	3	42		31,25
			31,5	
Лето – осень	4	21		32,25
			33	
Зима – весна	5	45		33,5
			34	
Лето – осень	6	23		

Этап 3. Расчет вклада кратковременных факторов $(s_t * i_t)$.

Произведение кратковременных факторов (*отношение к центрированному движущему среднему*) вычисляется путем деления величин исходного ряда на соответствующие величины центрированных движущих средних:

$$s_t * i_t = \frac{y_t}{f_t * c_t}.$$

Для нашего примера произведение кратковременных факторов рассчитано и представлено в табл. 5.4.

Таблица 5.4

Сезон	Период t	Объем продаж y_t, тыс. р.	*ДС*	$ДС_ц$ $(f_t * c_t)$	$(s_t * i_t)$
Зима – весна	1	40			
			30		
Лето – осень	2	20		30,5	0,656
			31		
Зима – весна	3	42		31,25	1,344
			31,5		
Лето – осень	4	21		32,25	0,651
			33		
Зима – весна	5	45		33,5	1,343
			34		
Лето – осень	6	23			

Этап 4. Вычисление значений сезонных индексов. Для этого полученные значения $(s_t * i_t)$ суммируются по одинаковым сезонам. Затем вычис-

«Мир покупок» и «Все для дома». При покупке товара покупателя просили назвать источник получения информации о товаре. Ответы распределились следующим образом: количество женщин, назвавших издание «Мир покупок», составило 48 человек, «Все для дома» – 45. Мужчины в качестве источника информации назвали «Мир покупок» 70 раз, а «Все для дома» – 35. Сделайте выводы из проведенного исследования.

2) Изучить связь между видом упаковки и популярностью (регулярностью покупок) товара. Данные для анализа представлены в табл. 10.5. Сделать соответствующие выводы.

Таблица 10.5

Марка товара	Активные потребители	Умеренные потребители	Редкие потребители	Случайные потребители
Мягкая упаковка	120	115	150	15
Картонная коробка	228	219	180	20
Стеклянная упаковка	95	120	37	50
На разлив	12	18	20	5

3) Рассчитать тесноту связи между лояльностью покупателей и расстоянием от дома покупателей до магазина. Данные для анализа представлены в табл. 10.6. Сделать соответствующие выводы.

Таблица 10.6

Тип покупатель	Расстояние до магазина, км			
	0,5	1,0	1,5	2,0
Истинная лояльность	120	118	110	102
Ложная лояльность	150	100	51	22
Латентная лояльность	108	92	96	56
Отсутствие лояльности	2	18	5	12

10.3. Контрольные вопросы

1) Зачем определяют тесноту связи между признаками?

2) Какие существуют подходы к оценке тесноты связи между признаками? От чего зависит выбор того или иного подхода?

3) Назовите показатели оценки тесноты связи между качественными признаками.

4) Назовите показатели оценки тесноты связи между качественными и количественными признаками.

5) Приведите примеры применения показателей тесноты связи: между качественными признаками; между качественными и количественными признаками.

Таблица 11.3

Период	Динамика объема продаж Y	Затраты на рекламу X	$Y*X$
1	112	16,0	1792
2	98	9,5	931
3	106	10,3	1091,8
4	96	8,5	816
5	92	9,5	874
Сумма	504	53,8	5504,8
Среднее	100,8	10,76	1100,96

$$\sigma_x = \sqrt{\frac{(16-10{,}76)^2 + ... + (9{,}5-10{,}76)^2}{5}} = 2{,}68,$$

$$\sigma_y = \sqrt{\frac{(112-100{,}8)^2 + ... + (92-100{,}8)^2}{5}} = 7{,}22,$$

$$r = \frac{1100{,}96 - 10{,}76 * 100{,}8}{2{,}68 * 7{,}22} = 0{,}85.$$

Выводы о силе связи можно судить на основе данных табл. 11.4.

Таблица 11.4

Коэффициент корреляции (по модулю)	Сила связи
0,81 – 1,00	Сильная
0,61 – 0,80	Умеренная
0,41 – 0,60	Слабая
0,21 – 0,40	Очень слабая
0,00 – 0,20	Отсутствует

В нашем примере эта связь сильная.

Кроме того, с помощью *коэффициента детерминации* (r^2) можно определить вклад независимого признака в результативный (зависимый). В нашем примере: $r^2 = 0{,}72$.

Это значит, что на 72 % объем продаж зависит от рекламных затрат, остальные 28 % относятся на влияние других факторов (признаков).

11.2. Задания

1) Существует ли связь между признаками: «качество подготовки персонала» и «доля рынка фирмы», если имеются ранговые шкалы этих двух показателей (табл. 11.5).

Таблица 4.1

Номер подразделения	Объем продаж, млн р.															
	Год															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	100	95	94	90	88	86	85	85	84	82	79	75	77	78	75	72
2	18	20	25	35	32	35	38	45	55	58	60	68	70	72	70	80
3	60	65	62	70	75	68	52	60	65	60	62	58	65	63	60	57
4	120	105	110	102	98	95	80	83	75	72	58	35	40	45	35	30

4.3. Контрольные вопросы

1) В чем состоит основная задача прогнозирования, и какие Вы знаете составляющие маркетингового прогнозирования?

2) В чем суть количественных методов прогнозирования?

3) Что такое тренд? Какие типы трендов Вы знаете?

4) Перечислите и охарактеризуйте этапы прогнозирования путем анализа трендов.

5) В чем, на Ваш взгляд, состоят достоинства и недостатки данного метода прогнозирования.

Практическое занятие 5

ИСПОЛЬЗОВАНИЕ МЕТОДА ОТНОШЕНИЯ К ДВИЖУЩЕМУ СРЕДНЕМУ ДЛЯ ЦЕЛЕЙ ПРОГНОЗА ВРЕМЕННОГО РЯДА ПРИ НАЛИЧИИ СЕЗОННОЙ КОМПОНЕНТЫ

Цель работы: освоить методику указанного метода для прогнозирования временного ряда при наличии сезонной компоненты.

5.1. Основные сведения

Сезонность – это такие изменения в показателях, которые повторяются из года в год и которые сопоставимы по величине в одинаковые периоды времени.

Использовать метод анализа тренда в классическом варианте при таком положении вещей будет некорректно, так как этот метод не учитывает сезонные колебания, кривая продаж сглаживается. Поэтому его модифицируют путем дополнения специальными коэффициентами, отражающими сезонную активность покупателей, – сезонными индексами. Такой метод получил название *метод отношения к движущему среднему* [7, 13, 19].

Этапы метода отношения к движущему среднему

Этап 1. Отделение долговременных факторов, отражающих общую тенденцию, от кратковременных (*проявляющих себя в периоды, меньшие, чем один год*). Обычно предполагается, что любые данные из временного ряда могут быть представлены в виде произведений четырех компонентов:

$$y_t = f_t * s_t * c_t * i_t,$$

где y_t – величина во временном ряду, соответствующая моменту времени t; f_t – значение тренда в момент времени t; s_t – значение фактора сезонности в момент времени t; c_t – значение фактора цикличности в момент времени t; i_t – значение фактора случайности в момент времени t.

То есть считается, что временным рядам свойственны четыре характеристики: *тенденция (тренд)*; *сезонность*; *цикличность*; *случайность.*

Сезонность – это такие изменения в показателях, которые повторяются из года в год и которые сопоставимы по величине в одинаковые периоды времени. Сезонные колебания являются предсказуемыми. Чтобы выявить сезонный фактор, исторические данные необходимо представить по кварталам или по месяцам. Если эти данные будут представлены по годам, то сезонность в них не проявится.

Цикличность отличается от сезонности следующими особенностями: период циклических колебаний составляет более одного года; циклические колебания определяются не одним, а несколькими факторами; величины максимумов (и минимумов) циклов могут немного отличаться; длительность циклов может быть разной.

Циклическим колебаниям подвержены макроэкономические показатели, например колебания объема валового национального продукта (ВНП). В циклах ВНП выделяют четыре основные фазы: депрессия, оживление, подъем и спад.

Случайность – это изменения в показателях, которые невозможно предугадать. Они не подчиняются какой-либо закономерности, подобно тренду, сезонности и цикличности. Поэтому иногда их называют «шумом». *Тренд и цикличность являются долговременными факторами, а сезонность и случайность – кратковременными (проявляющимися в периоды, меньшие, чем один год).*

Чтобы отделить долговременные факторы от кратковременных, вычисляют среднее значение временного ряда за n периодов подряд, равных по суммарной продолжительности одному году. И продолжают вычислять его, сдвигая на один период.

Предположим, что имеются данные продаж (табл. 5.1).

Таблица 5.1

Год	Сезон	Период t	Объем продаж y_t, тыс. р.
1	Зима – весна	1	40
	Лето – осень	2	20
2	Зима – весна	3	42
	Лето – осень	4	21
3	Зима – весна	5	45
	Лето – осень	6	23

Имеем два сезона в году: зима – весна и лето – осень, в которые наблюдается повышение и снижение объема продаж, следовательно, $n = 2$.

Среднее значение временного ряда (*движущее среднее* (*ДС*)) *за первый годичный период:*

$$ДС_1 = \frac{40+20}{2} = 30.$$

ДС за второй годичный период и т.д. (табл. 5.2):

$$ДС_2 = \frac{20+42}{2} = 31.$$

Таблица 5.2

Сезон	Период t	Объем продаж y_t, тыс. р.	*ДС*
Зима – весна	1	40	
			30
Лето – осень	2	20	
			31
Зима – весна	3	42	
			31,5
Лето – осень	4	21	
			33
Зима – весна	5	45	
			34
Лето – осень	6	23	

Этап 2. Расчет вклада долговременных факторов ($f_t * c_t$). Для этого рассчитывается центрированное движущее среднее $ДС_ц$.

Центрированное движущее среднее так же, как и движущее среднее, отражает влияние долговременных факторов – тренда и цикличности ($f_t * c_t$). В отличие от движущего среднего, оно приводится к тем же моментам времени, что и исторические данные исходного ряда. Рассчитывается как усредненное значение двух соседних величин движущихся средних. Для нашего примера значения $ДС_ц$ рассчитаны и представлены в табл. 5.3.

Практическое занятие 11

ПРИМЕНЕНИЕ МЕТОДОВ АНАЛИЗА ПАРНЫХ ВЗАИМОСВЯЗЕЙ МЕЖДУ ПРИЗНАКАМИ В МАРКЕТИНГОВЫХ ИССЛЕДОВАНИЯХ. ИЗУЧЕНИЕ ТЕСНОТЫ СВЯЗИ МЕЖДУ КОЛИЧЕСТВЕННЫМИ ПРИЗНАКАМИ

Цель работы: ознакомиться с подходами и методами анализа парных взаимосвязей между признаками и научится применять их в зависимости от вида признака: количественного или качественного.

11.1. Основные сведения

Данная работа является продолжением практического занятия 10.

В предыдущей работе мы учились рассчитывать тесноту связи между качественными признаками, а также между качественными и количественными. В данной работе мы научимся оценивать тесноту связи между количественными признаками [12, 17, 18]. Используемые для этих целей показатели тесноты связи представлены в табл. 11.1.

Таблица 11.1

Результативные показатели	Шкалы измерения признака	Показатели тесноты связи результативных показателей с признаками		
		альтернативными	качественными	количественными
Альтернативные	Наименований	A, k	C	Z
Качественные	Наименований	C	C	ω
Количественные	Порядковые	ω	R	R
	Интервальные и относительные	ω	ω	r

Изучение тесноты связи между количественными признаками.

В этом случае возможны варианты:

- когда количественные признаки представлены в виде рангов (проранжированные переменные, ранговые шкалы);
- когда количественные переменные представлены в виде относительных или интервальных шкал.

При ранговых шкалах применяется *коэффициент корреляции рангов Спирмэна:*

$$R = 1 - \left[\frac{6 * \sum_i (d_i)^2}{n^3 - n}\right], \qquad (11.1)$$

где d_i – разница рангов по двум переменным; n – число ранжированных позиций.

Такие ситуации часто бывают, когда оценка показателей дается экспертно. Например, необходимо установить существует ли связь между рангами компании, присвоенными по качеству производимой ею продукции, и положением компании на рынке (табл. 11.2).

Таблица 11.2

Компания	Ранг качества продукции X	Ранг рыночной доли Y	Разность рангов d_i	Квадрат разности рангов
1	4	3	1	1
2	5	4	-1	1
3	3	5	-2	4
4	2	1	1	1
5	1	2	-1	1
Сумма			0	8

Используя данные табл. 11.2, находим коэффициент корреляции рангов Спирмэна по формуле (11.1):

$$R = 1 - \left[\frac{6*8}{5^3 - 5}\right] = 0{,}6.$$

Коэффициент корреляции рангов Спирмэна находится в диапазоне между (-1) и 1. Чем ближе по модулю коэффициент к единице, тем теснее связь между признаками. В данном случае можно сказать, что связь между качеством продукции и занимаемой долей рынка достаточно сильна.

При относительных или интервальных шкалах применяется *коэффициент корреляции:*

$$r = \frac{\overline{xy} - \bar{x} * \bar{y}}{\sigma_x * \sigma_y},$$

где $\overline{xy}$ – среднее значение произведений двух признаков; $\bar{x}$ – среднее значение признака x; $\bar{y}$ – среднее значение признака y; σ_x – среднеквадратическое отклонение признака x; σ_y – среднеквадратическое отклонение признака y.

Например, исследуется связь между рекламными затратами и объемом продаж (табл. 11.3).

Рассчитаем средние уровни дохода в каждой группе покупателей и в целом. Так как один из признаков (доход) представлен в виде интервальной шкалы, необходимо рассчитать средние значения этих интервалов и взвесить их по частоте покупок:

$$Y_1 = \frac{210*6000+180*8000+160*10000}{550} = 7818 \text{ р.},$$

$$Y_2 = \frac{230*6000+310*8000+390*10000}{930} = 8344 \text{ р.},$$

$$Y_3 = \frac{440*6000+490*8000+550*10000}{1480} = 8149 \text{ р.}$$

Среднее квадратическое отклонение для сгруппированного ряда:

$$\sigma_y = \sqrt{\frac{\sum_i (Y_i - Y)^2 * f}{\sum_i f}}.$$

$$\sigma_y = \sqrt{\frac{(6000-8149)^2*440+(8000-8149)^2*490+(10000-8149)^2*550}{1480}} = 1629.$$

Поскольку в рассматриваемом примере качественный признак может принимать два значения, то можно рассчитать два дисерийных коэффициента:

$$\varpi_1 = \frac{7818-8149}{1629*\sqrt{\frac{1480}{550}-1}} = -0{,}156, \qquad \varpi_2 = \frac{8344-8149}{1629*\sqrt{\frac{1480}{930}-1}} = 0{,}156.$$

Сумма дисерийных коэффициентов должна быть равна нулю. Отрицательное значение говорит (в случае, если его значение больше 0,3) об обратной зависимости признаков. В нашем случае дисерийный коэффициент меньше 0,3, значит, связи между доходом и частотой покупок не существует.

10.2. Задания

Определите взаимосвязь признаков в следующих случаях.

1) Маркетологу необходимо принять решение, в каком издании размещать рекламу в следующем году, если целевой аудиторией товара фирмы являются мужчины. Реклама товара ранее была дана в двух изданиях:

ляются их средние значения, которые и являются значениями сезонных индексов (табл. 5.5).

Таблица 5.5

Год	Зима – весна	Лето – осень
1	-	0,656
2	1,344	0,651
3	1,343	-
Сезонный индекс	1,3435	0,6535

Значение сезонного индекса 1,3435 означает, что объем продаж в этот сезон превышает среднегодовой уровень на 34,35 %.

Этап 5. Устранение фактора сезонности.

Такая процедура обеспечивается делением исторических величин объемов продаж на полученные сезонные индексы (табл. 5.6).

Таблица 5.6

Год	Сезон	Период t	Объем продаж y_t, тыс. р.	Объем продаж с устраненным фактором сезонности, тыс. р.
1	Зима – весна	1	40	29,77
	Лето – осень	2	20	30,60
2	Зима – весна	3	42	31,26
	Лето – осень	4	21	32,13
3	Зима – весна	5	45	33,49
	Лето – осень	6	23	35,20

Этап 6. Построение модели тренда с помощью данных с устраненным фактором сезонности. Используя методику, описанную в предыдущей лабораторной работе, необходимо построить уравнение регрессии (уравнение тренда), используя в качестве временного ряда показатели объема продаж с устраненным фактором сезонности.

Этап 7. Прогноз показателей будущих периодов. Пользуясь полученным уравнением тренда (на 6-м этапе), составляется прогноз объемов продаж будущих периодов.

Этап 8. Корректировка прогнозных показателей с учетом фактора сезонности. Это обеспечивается умножением объемов продаж, полученных на 7-м этапе выполнения практического задания, на сезонные индексы, полученные на 4-м этапе выполнения упражнения.

Этап 9. Графическое представление результатов лабораторной работы – построение на одной плоскости графика изменения объема продаж во времени с учетом фактора сезонности, включая прогнозные периоды, и графика тренда за те же периоды.

(количественные и качественные), то необходимо привести их все к одной базе. Для этого проводится расчет так называемых приведенных коэффициентов. При этом, если следует стремиться к максимизации показателя (т.е. чем выше показатель, тем выгоднее сегмент), то приведенный коэффициент нужно рассчитывать по формуле

$$a_i = \frac{n_i}{n_{\max}},$$

где n_i – значение конкретного показателя для каждого сегмента из представленных; $n_{\max}$ – максимальное значение этого показателя из представленных.

А если следует стремиться к минимизации показателя, то следует использовать формулу

$$a_i = \frac{n_{\min}}{n_i},$$

где $n_{\min}$ – минимальное значение показателя из представленных.

Этап 4. Определение коэффициента оптимальности сегмента. Коэффициент оптимальности сегмента рассчитывается следующим образом:

$$K_0 = \frac{\sum_{i=1}^{k} (a_{ij} * w_i)}{\sum w_i},$$

где w_i – весa показателей; k – количество показателей; j – номер оцениваемого сегмента.

Если сумма весов равна 1 или 100 %, то значениями формулы предпочтения. По величине K_0 и выбирают оптимальный для фирмы сегмент. Чем ближе K_0 к единице, тем более выгоден сегмент для деятельности предприятия.

6.2. Задания

Фирма, работающая на пяти сегментах, решает вопрос, на каком из них сконцентрировать свои усилия (табл. 6.1).

Таблица 6.1

Показатели	Веса	Сегменты				
		1	2	3	4	5
Количественные показатели						
1. Емкость рынка, тыс.р.	0,20	200	250	100	150	220
2. Доля рынка, занимаемая предприятием, %	0,11	18	13	2	15	10

количественную зависимую переменную (объемы продаж). Или, исследуется влияние пола респондента на уровень потребления того или иного вара, следовательно, изучают влияние качественной переменной на количественную. В табл. 10.1 представлены показатели тесноты связи между количественными и качественными признаками, которые могут быть использованы в процессе маркетингового исследования.

Таблица 10.1

Результативные показатели	Шкалы измерения признака	Показатели тесноты связи результатив-ных показателей с признаками		
		альтерна-тивным	качествен-ным	количест-венным
Альтернативные	Наименований	A, k	C	Z
Качественные	Наименований	C	C	ω
Количественные	Порядковые	ω	R	R
	Интервальные и относительные	ω	ω	r

Изучение тесноты связи между качественными и альтернативными признаками

Альтернативные признаки – это признаки, принимающие два возможных значения, например, связь между полом (или женский, или мужской) и увиденным рекламным роликом (или видели, или не видели).

Так, например, проводился опрос среди лиц, купивших определенный товар, видели они или нет определенный рекламный ролик.

Сначала составляется двумерная таблица сопряженности признака, в которой на пересечении строк и столбцов указывается частота значений признаков (табл. 10.2).

Таблица 10.2

Признак 1 (причина) «пол»	Признак 2 (следствие) «рекламный ролик»		Итого
	Первое значение признака 2 «не видели»	Второе значение признака 2 «видели»	
Первое значение признака 1 «мужской»	a (36)	b (17)	53
Второе значение признака 1 «женский»	c (42)	d (87)	129
Итого	78	104	182

Изучается связь между полом и влиянием рекламы.

Для измерения тесноты связи между альтернативными признаками применяются:

5.2. Задания

Осуществить прогноз работы торговой компании на ближайшие два года в точности с описанными этапами прогнозирования с учетом фактора сезонности (табл. 5.7).

Таблица 5.7

Сезон	Объем продаж в соответствующем году, тыс. р.						
	2001	2002	2003	2004	2005	2006	2007
Зима	600	590	661	690	700	742	781
Весна	202	180	255	270	272	280	285
Лето	401	420	460	479	480	485	490
Осень	252	202	255	260	270	274	280

5.3. Контрольные вопросы

1) Опишите четыре характеристики, свойственные временным рядам. Какие из них оказывают кратковременное воздействие, а какие долговременное?

2) В чем проявляется влияние сезонности?

3) Почему при прогнозировании временного ряда с сезонностью нельзя применять метод анализа трендов?

4) Охарактеризуйте этапы метода отношения к движущему среднему. В чем основное отличие этого метода от метода анализа трендов?

5) В чем, на Ваш взгляд, состоят достоинства и недостатки данного метода прогнозирования?

Практическое занятие 6

ПОИСК ОПТИМАЛЬНОГО СЕГМЕНТА НА ОСНОВЕ СБАЛАНСИРОВАННОЙ СИСТЕМЫ ПОКАЗАТЕЛЕЙ

Цель работы: освоить технику оценки потребительских сегментов на основе сбалансированной системы показателей.

6.1. Основные сведения

В связи с ограниченностью ресурсов предприятие предпочитает чаще работать не на всем рынке, а стремится выделить для себя целевые сегменты. *Целевой сегмент* – определенная группа потребителей со схожими потребностями и мотивационными характеристиками, создающими для фирмы благоприятные маркетинговые возможности [14]. Выбранный сегмент (или сегменты) должен иметь ряд характеристик, позволяющих на-

звать этот сегмент для фирмы оптимальным. *Оптимальный сегмент* – сегмент, соответствующий тактическим и стратегическим целям фирмы. Тактической характеристикой можно назвать *выгодность* сегмента в данный момент времени, стратегической – *перспективность* сегмента, т.е. возможность сохранения характеристики *выгодность* в будущем.

Выгодным сегмент рынка будет в том случае, если он приносит сравнительно большой доход, обладает высокой нормой прибыли, умеренной конкуренцией, эффективными каналами распределения и т.д. Перспективным сегмент рынка будет в том случае, если описанные характеристики сохранятся в будущем. Но нельзя забывать, что на практике крайне редко бывает сочетание всех перечисленных характеристик в одном сегменте, поэтому задачей маркетолога чаще является поиск *оптимального* сегмента, т.е. наиболее благоприятного.

Важной задачей в оценке сегментов является выбор показателей оценки. Чем более сбалансирована будет эта система показателей, тем точнее будет выбор оптимального сегмента.

Сбалансированность показателей оценки сегмента заключается в балансе между финансовыми и нефинансовыми показателями, между результатами и факторами достижения этих результатов, в балансе между внешним окружением компании (акционеры, клиенты) и внутренними ее составляющими (персонал, бизнес-процессы).

Этапы оценки сегментов на основе сбалансированной системы показателей [12]

Этап 1. Выбор системы показателей, отвечающей критерию «сбалансированность». Каждое предприятие в зависимости от конкретных условий выбирает свой набор показателей оценки сегментов. Многие авторы [17, 22] рекомендуют четыре направления поиска показателей: финансовый аспект; клиенты и внешнее окружение; внутренние бизнес-процессы; уровень подготовки персонала и степень развития предприятия. Выбор показателей зависит от стратегических целей фирмы, к которым могут быть отнесены: работа на потребительском или промышленном рынках; работа на имеющемся или имеющихся сегментах или поиск сегментов для работы.

Этап 2. Определение весов и значений выбранных показателей. На этом этапе все показатели разбивают на две группы: количественные и качественные. Количественные оценивают с помощью подходящих единиц измерения (денежные, процентные и т.д.), а для оценки качественных прибегают к помощи экспертов, которые по определенной шкале оценивают сравниваемые сегменты. Значение весов показателей также оценивается экспертно.

Этап 3. Расчет приведенных коэффициентов сравнения. Так как полученные значения показателей для сравнения имеют разную размерность

- *коэффициент ассоциации:*

$$A = \frac{a*d - b*c}{a*d + b*c}.$$

- *коэффициент контингенции:*

$$k = \frac{a*d + b*c}{\sqrt{(a+b)*(c+d)*(a+c)*(b+d)}}.$$

Связь считается подтвержденной, если $|A| \geq 0{,}5$; $|k| \geq 0{,}3$.

В нашем случае: $A = 0{,}629$, а $k = 0{,}325$. Можно сделать следующие выводы из приведенного примера:

1) Так как $|A| \geq 0{,}5$ и $|k| \geq 0{,}3$, то существует влияние на покупку мужской или женской аудитории факта просмотра рекламного ролика.

2) В целом женщины покупают данного товара больше, чем мужчины.

3) На женщин ролик оказывает большее значение (87 женщин видели ролик и купили товар), чем на мужчин (только 17).

4) Среди женщин и мужчин, не видевших ролик, ответы распределились примерно поровну.

Изучение тесноты связи между качественными признаками

Для оценки тесноты связи в этом случае применяется *коэффициент сопряженности Пирсона:*

$$C = \sqrt{\frac{\chi^2}{n + \chi^2}},$$

$$\chi^2 = n * \left[\sum_{ij} \frac{(n_{ij})^2}{\sum_i n_i * \sum_j n_j} - 1 \right],$$

где n – общее число единиц совокупности; n_i – совокупность единиц с i-м признаком; n_j – совокупность единиц с j-м признаком; n_{ij} – совокупность единиц с i-м и j-м признаками.

Например, изучается регулярность потребления определенного товара в зависимости от его марки. Ответы распределились следующим образом (табл. 10.3).

Таблица 10.3

Марка товара	Регулярные покупки	Нерегулярные покупки	Случайные покупки (единичные)	Всего
Марка 1	43	53	26	122
Марка 2	107	94	34	235
Марка 3	93	102	37	232
Всего	243	249	97	589

$$\chi^2 = 589 * \left[\frac{42^2}{243*122} + \frac{53^2}{249*122} + \ldots + \frac{37^2}{97*232} - 1 \right] = 5{,}02,$$

$$C = \sqrt{\frac{5{,}02}{589 + 5{,}02}} = 0{,}09.$$

Коэффициент взаимной сопряженности Пирсона должен быть не меньше 0,5. В нашем примере связь между регулярностью покупок и маркой товара не подтверждается, т.е. марка не влияет на регулярность покупок. Распределение частот случайное.

Изучение тесноты связи между количественными и качественными признаками

Если один из взаимосвязанных признаков количественный, а второй качественный, то тесноту связи можно определить на основе *дисерийного коэффициента:*

$$\varpi_i = \frac{Y_i - Y}{\sigma_y * \sqrt{\frac{\sum f}{\sum f_i} - 1}},$$

где Y – среднее значение признака; Y_i – среднее значение признака в i-й группе; σ_y – среднее квадратическое отклонение; f – частота признака.

Для подтверждения связи достаточно, чтобы дисерийный коэффициент был больше 0,3 (*по модулю*).

Пример 10.1. Оценим тесноту связи между уровнем дохода и типом покупателя (табл. 10.4).

Таблица 10.4

Тип покупателя	Уровень доходов, р.			Всего
	5000 – 7000	7001 – 9000	9001 – 11000	
Случайные покупатели f_1	210	180	160	550
Лояльные покупатели f_2	230	310	390	930
Всего f	440	490	550	1480

Исследователь хочет получить ответ на следующие вопросы:

1) Превышает ли осведомленность всех респондентов о товаре уровень 4,0?

2) Превышает ли осведомленность женщин о товаре 4,0?

3) Превышает ли осведомленность мужчин о товаре 4,0?

9.3. Контрольные вопросы

1) Охарактеризуйте этапы проверки гипотез о средних величинах.

2) Какие Вам известны статистические критерии для проверки статистических гипотез о средних величинах? От влияния каких факторов зависит их применение?

3) Приведите пример ситуации, когда необходимо применение таких статистических критериев для проверки гипотез о средних величинах.

Практическое занятие 10

ПРИМЕНЕНИЕ МЕТОДОВ АНАЛИЗА ПАРНЫХ ВЗАИМОСВЯЗЕЙ МЕЖДУ ПРИЗНАКАМИ В МАРКЕТИНГОВЫХ ИССЛЕДОВАНИЯХ. ИЗУЧЕНИЕ ТЕСНОТЫ СВЯЗИ МЕЖДУ КАЧЕСТВЕННЫМИ И КОЛИЧЕСТВЕННЫМИ ПРИЗНАКАМИ

Цель работы: ознакомиться с подходами и методами анализа парных взаимосвязей между признаками и научится применять их в зависимости от вида признака: количественного или качественного.

10.1. Основные сведения

Изучение взаимосвязей между признаками составляют важную часть методологии маркетинговых исследований, является обязательным этапом проведения эксперимента. Примером выявления такой взаимосвязи является влияние рекламы на увеличение спроса, влияние национальности или пола на уровень потребления определенного товара.

Для выявления наличия или отсутствия связи между двумя признаками (парные взаимосвязи) применяется набор статистических показателей: показатели тесноты связи [12, 17, 18]. Применение того или иного показателя зависит от видов переменных, связь между которыми определяется количественными или качественными признаками.

Например, исследуется связь между влиянием рекламных затрат на увеличение объема продаж. В этом случае можно говорить о возможном влиянии количественной независимой переменной (рекламные затраты) на

Продолжение табл. 6.1

Показатели	Веса	Сегменты				
		1	2	3	4	5
3. Производственный потенциал сегмента, шт.	0,08	1900	1500	2100	2000	1000
4. Затраты на систему сбыта и сервисное обслуживание, р.	0,09	28000	32000	21000	13000	15000
5. Затраты, связанные с реагированием предприятия на запросы потребителей, р.	0,06	18500	15200	13000	10000	21000
6. Уровень риска	0,07	2,3	3,1	4,7	2,8	5,0
7. Прибыльность сегмента, р.	0,10	42500	38000	29500	30000	35000
8. Рентабельность продаж, %	0,03	25	16	12	10	25
Качественные показатели						
9. Возможность реализовывать продукцию, доступность сегмента (чем выше показатель, тем доступнее сегмент)	0,04	7,3	6,9	7,5	10,0	5,0
10. Устойчивость сегмента (чем выше показатель, тем устойчивее сегмент)	0,07	6,3	5,75	6,1	5,5	8,2
11. Уровень конкуренции в сегменте (чем выше показатель, тем конкурентнее среда)	0,11	8,23	7,72	8,03	7,0	5,0
12. Оптимальность товарного портфеля на данном сегменте (чем выше показатель, тем портфель оптимальнее)	0,04	8,89	7,57	6,73	5,0	8,2

6.3. Контрольные вопросы

1) Что такое целевой сегмент? В чем суть сегментации рынка?

2) Охарактеризуйте оптимальный сегмент. В чем состоит выгодность сегмента и его перспективность?

3) В чем состоит принцип сбалансированности при поиске показателей оценки сегмента?

4) Укажите четыре направления для поиска показателей оценки оптимального сегмента. Как стратегические цели фирмы могут повлиять на выбор системы показателей?

5) Опишите методику поиска оптимального сегмента с использованием сбалансированной системы показателей.

ПРОВЕРКА ЗНАЧИМОСТИ РЕЗУЛЬТАТОВ МАРКЕТИНГОВЫХ ИССЛЕДОВАНИЙ. ГИПОТЕЗЫ О СРЕДНИХ ВЕЛИЧИНАХ

Цель работы: ознакомиться с методами проверки значимости результатов маркетинговых исследований; научиться применять данные методы в зависимости от области применения.

9.1. Основные сведения

Данное занятие является логическим продолжением 8-го. Но в отличие от предыдущего проверяются гипотезы о средних величинах.

В табл. 9.1 представлены наиболее часто используемые критерии для проверки статистических гипотез о средних величинах.

Таблица 9.1

Объект применения	Число групп или выборочных совокупностей	Вид шкал	Критерий	Специальные требования
Гипотезы о средних величинах	Одна (большая, > 30)	Интервальная или относительная (имеющая начало отчета)	Z-критерий для одной средней	Случайная выборка, $n > 30$
	Одна (малая)	Интервальная или относительная	t-критерий для одной средней	Случайная выборка, $n < 30$
	Две (большие)	Интервальная или относительная	Z-критерий для двух средних	Случайная выборка, $n < 30$
	Две (малые)	Интервальная или относительная	Анализ вариаций	Случайная выборка

Этапы проверки гипотез о средних величинах идентичны этапам проверки гипотез о равах распределения [1, 16, 23].

Пример. Производитель безалкогольных напитков планирует изучить рынок нового напитка. Случайным образом в городе были отобраны 10 магазинов. Им был предложен для продажи в течение определенного промежутка времени новый безалкогольный напиток (тестирование продукта). Компания рассчитывала на продажу 100 бутылок нового напитка в день в каждом магазине. Только в этом случае окупаемая прибыль оправ-

Этапы оценки согласованности экспертных оценок

Этап 1. Расчет среднего значения сумы рангов по всем направлениям (параметрам):

$$\bar{S} = \frac{\sum_{j=1}^{m} S_j}{m},$$

где m – количество оцениваемых направлений (параметров).

Таблица 7.3

Параметр (направление)	Экспертная оценка значимости параметров (ранги)				S_j	$1/S_j$	$\bar{K}_{весj}$
	1	2	3	4	5			
Время работы на рынке	1,5	3	2	1	1	8,5	0,1176	0,3291
Рентгация	1,5	1	1	3	2	8,5	0,1176	0,3291
Форма торговли	3,5	2	3	3	3,5	15	0,0667	0,1865
Местоположение	3,5	4	4	3	3,5	18	0,0556	0,1554
ИТОГО						50	0,3575	1

В нашем примере: $\bar{s} = 50/4 = 12{,}5$.

Этап 2. Расчет отклонения сумы рангов по j-му направлению от среднего значения сумы рангов и возведение его в квадрат:

$$d_j = S_j - \bar{S}.$$

Данные для нашего примера представлены в табл. 7.4.

Таблица 7.4

Параметр	Экспертная оценка значимости параметров (ранги)				S_j	d_j	d_j^2
	1	2	3	4	5			
Время работы на рынке	1,5	3	2	1	1	8,5	-4,00	16,00
Рентгация	1,5	1	1	3	2	8,5	-4,00	16,00
Форма торговли	3,5	2	3	3	3,5	15	2,50	6,25
Местоположение	3,5	4	4	3	3,5	18	5,50	30,25
Итого						50		68,50

Этап 3. Расчет коэффициента конкордации:

$$W = \frac{12 * \sum_{j=1}^{m} d_j^2}{n^2 * (m^3 - m) - n * \sum_{j=1}^{n} T_i},$$

дает расходы на продвижение нового товара. Фактические данные об объемах продаж представлены в табл. 9.2.

Таблица 9.2

Номер магазина	Средний объем продаж в день, бут.	Отклонение от средней величины	Квадрат отклонения от средней величины
1	87	-12	144
2	96	-3	9
3	105	6	36
4	120	21	441
5	86	-13	169
6	104	5	25
7	115	16	256
8	80	19	361
9	100	1	1
10	97	2	4
Сумма	990		1446

Средний объем продаж за неделю (M) составил 99 бут.

Последовательность процедуры расчета:

1) Выдвигаются нулевая и альтернативная гипотезы:

H_0: M ≥ 100 бутылок, следовательно, ожидаемая прибыль оправдает расходы на продвижение нового товара;

H_a: M < 100 бутылок, следовательно, ожидаемая прибыль не оправдает расходы на продвижение нового товара.

2) Устанавливается допустимый уровень ошибки выборки. При α = = 0,05 и числе степеней свободы 9 = 10 – 1 t = 1,833 [1, 16, 21, 23].

3) Находится теоретическое значение критерия. В данном случае это *t-критерий для одной средней (одна малая группа)* (см. табл. 9.1):

$$t = \frac{\left|\tilde{X} - \bar{X}\right|}{S_x},$$

где $\tilde{X}$ – выборочная средняя; $\bar{X}$ – генеральная средняя по H_0; S_x – стандартная ошибка оценки средней величины.

Стандартная ошибка оценки средней величины:

$$S_x = \frac{\sigma_x}{\sqrt{n}},$$

где σ_x – среднее квадратическое отклонение в выборочной совокупности; n – объем выборки.

Среднее квадратическое отклонение в выборочной совокупности определяется по формуле

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \tilde{X})^2}{n-1}}.$$

В нашем случае:

$$\sigma_x = \sqrt{\frac{1446}{10-1}} = 12{,}675, \quad S_x = \frac{12{,}675}{\sqrt{10}} = 4{,}0082, \quad t = \frac{|99-100|}{4{,}0082} = 0{,}2495.$$

4) Формулируются выводы. Так как расчетное значение t меньше табличного, то нулевая гипотеза принимается. На основе этого можно заключить, что при тестировании были допущены ошибки (среднее 99 бутылок), и, следовательно, продажи в каждом магазине будут более 100 бутылок, ожидаемая прибыль оправдает расходы на продвижение нового товара.

9.2. Задания

В ходе опроса исследовалась уровень осведомленности о товаре. Уровень осведомленности определялся по 7-балльной шкале Лайкерта. Были получены следующие данные (табл. 9.3). Пол респондента закодирован: 1 – женщины, 2 – мужчины.

Таблица 9.3

Номер анкеты	Пол респондента	Осведом-ленность	Номер анкеты	Пол респондента	Осведом-ленность
1	2	7	16	2	3
2	1	2	17	1	7
3	1	3	18	1	6
4	2	6	19	1	1
5	2	5	20	1	5
6	2	4	21	2	6
7	1	5	22	2	2
8	1	2	23	1	1
9	2	4	24	1	1
10	1	3	25	2	3
11	2	6	26	2	5
12	2	6	27	2	7
13	1	4	28	1	6
14	2	6	29	1	7
15	2	4	30	2	7

ИСПОЛЬЗОВАНИЕ МЕТОДА РАНГОВОЙ КОРРЕЛЯЦИИ ПРИ ОБРАБОТКЕ ЭКСПЕРТНЫХ ОЦЕНОК

Цель работы: освоить методику указанного метода для оценки экспертных оценок.

7.1. Основные сведения

Часто при прогнозировании экономических явлений используются экспертные оценки. Например, экспертов часто просят оценить приоритетность развития тех или иных направлений деятельности, секторов рынка, степень значимости для покупателей тех или иных параметров товара.

Для обработки полученных экспертных данных часто используется метод рангової корреляции [19], в рамках которых решаются две задачи:

1) анализируется значимость исследуемых параметров, т.е. рассчитываются коэффициенты весомости каждого параметра и по их величине судят о предпочтительности того или иного направления деятельности или значимости того или иного параметра для потребителей;

2) оценивается согласованность мнений экспертов, т.е. рассчитывается *коэффициент конкордации*, по величине которого судят о степени согласованности мнений экспертов. Только при высоких значениях этого коэффициента (стремящегося к единице) и можно принимать прогнозные значения, данные экспертами.

Этапы анализа значимости исследуемых параметров.

Этап 1. Получение индивидуальных экспертных оценок относительно важности, значимости, предпочтительности оцениваемых параметров или направлений. Оценки экспертов данного в виде весовых коэффициентов, которые могут принимать значение от 0 до 1. Сумма коэффициентов, данных одним экспертом, должна равняться единице. В табл. 7.1 показаны весовые коэффициенты, данные пятью экспертами в отношении параметров, по которым потребитель выбирает магазин, торгующий бытовой техникой.

Таблица 7.1

Параметр (направление)	Экспертная оценка значимости параметров (весомость)				
	1	2	3	4	5
Время работы на рынке	0,3	0,2	0,3	0,4	0,5
Репутация	0,3	0,4	0,4	0,2	0,3
Форма торговли	0,2	0,3	0,2	0,2	0,1
Местоположение	0,2	0,1	0,1	0,2	0,1
ИТОГО	1	1	1	1	1

Этап 2. Ранжирование оценок важности, данных экспертами. Каждая оценка, данная i-м экспертом, выражается рангом R_{ij} – числом натурального ряда – таким образом, что значение, равное единице, дается максимальной оценке (важности), а n – минимальной. Если среди оценок, данных i-м экспертом, есть одинаковые, то им присваивается одинаковый ранг, равный среднему арифметическому соответствующих чисел натурального ряда. Например, в нашем примере (табл. 7.1) первый эксперт присвоил первым двум параметрам (время работы на рынке и репутация) одинаковую максимальную весомость 0,3, т.е. эти параметры занимают первые два места. Ранг каждого из этих параметров можно рассчитать как (1 + 2)/2 = 1,5. Следующие два параметра занимают третье и четвертое места, ранг каждого их них будет рассчитан следующим образом: (3 + 4)/2 = 3,5 (табл. 7.2). Если имеется три одинаковые оценки важности для трех параметров, как у четвертого эксперта, то ранги их будут рассчитаны следующим образом: (2 + 3 + 4)/3 = 3.

Таблица 7.2

Параметр (направление)	Экспертная оценка значимости параметров (ранги) R_{ij}				
	1	2	3	4	5
Время работы на рынке	1,5	3	2	1	1
Репутация	1,5	1	1	3	2
Форма торговли	3,5	2	3	3	3,5
Местоположение	3,5	4	4	3	3,5

Этап 3. Оценка значимости исследуемых параметров с помощью среднего коэффициента весомости:

$$K_{весj} = \frac{1/s_j}{\sum(1/s_j)},$$

где S_j – сумма рангов по каждому направлению (параметру).

$$s_j = \sum_{i=1}^{n} R_{ij}.$$

Полученные значения среднего коэффициента весомости (табл. 7.3) говорят о том, что наиболее важными параметрами для потребителя при выборе магазина являются «время работы на рынке» и «репутация». На втором месте по важности «форма торговли» (18,65 %) и последним является «местоположение» (15,54 %).

Но эти данные можно принимать за данность только после этапа оценки согласованности экспертных оценок, т.е. оценки того, насколько единодушны в своем мнении эксперты.

Табличное значение при α = 0,05 и числе степеней свободы (r - 1)*(k - 1) = (2 - 1)*(3 - 1) = 2 (r – число колонок, количество групп посетителей) равно 5,991 [1, 16, 21, 23]. Так как расчетное значение 13,35 больше табличного, то нулевая гипотеза отвергается, и можно сделать вывод о том, что существуют различия в частоте посещения магазина мужчинами и женщинами.

8.2. Задания

Исследователь хочет выяснить, влияет ли уровень образования на доход респондентов (табл. 8.5).

Выявите взаимосвязь между указанными признаками.

Таблица 8.5

Доход, р.	Среднее образование	Высшее образование	Неоконченное среднее образование	Неоконченное высшее образование
До 5000	195	10	12	122
5001 – 10000	111	20	9	110
10001 – 15000	56	55	15	96
15001 – 20000	62	61	16	98
20001 – 25000	46	125	20	78
25001 – 30000	38	212	21	75
Свыше 30000	36	250	19	76

8.3. Контрольные вопросы

1) Дайте определение гипотезы. Что такое нулевая гипотеза и чем она отличается от альтернативной? Опишите основное правило выдвижения нулевой гипотезы.

2) Охарактеризуйте этапы проверки значимости результатов МИ (если выдвигается гипотеза о частоте распределения).

3) Какие Вам известны статистические критерии для проверки статистических гипотез о частоте распределения? От влияния каких факторов зависит их применение?

4) Приведите пример МИ, когда необходимо применение статистических критериев для проверки гипотез о частоте распределения.

где n – количество экспертов; T_i – коэффициент отклонения рангов, который рассчитывается по формуле

$$T_i = \sum(t_i^3 - t_i).$$

Здесь t_i – количество равных рангов в i-й группе (у одного эксперта).

В нашем примере у первого эксперта *два раза* встречается *по два* равных ранга, следовательно:

$$T_1 = (2^3 - 2) + (2^3 - 2) = 12.$$

У второго и третьего экспертов ни разу не встречаются равные ранги, следовательно:

$$T_2 = T_3 = 0.$$

У четвертого эксперта *три раза* встречается одинаковый ранг, следовательно:

$$T_4 = (3^3 - 3) = 24.$$

У пятого эксперта два раза встречается один ранг, следовательно:

$$T_5 = (2^3 - 2) = 6.$$

Тогда коэффициент конкордации:

$$W = \frac{12 * 68{,}5}{5^2 * (4^3 - 4) - 5 * 42} = 0{,}64.$$

При $W = 1$ наблюдается полная согласованность мнений экспертов, при $W = 0$ – полная несогласованность.

В нашем примере значение коэффициента конкордации достаточно высокое, мнение экспертов можно считать достаточно согласованным, и данные расчета весомости параметров можно принимать за достаточно объективные и верные. Если бы данные были несогласованными, т.е. коэффициент конкордации стремился к нулю, то в этом случае лучше еще раз опросить экспертов для получения более объективной информации.

7.2. Задания

1) Для анализа наиболее значимых атрибутов для посетителей кафе при получении услуги питания была собрана оценка 10 экспертов (табл. 7.5).

На основе имеющихся данных выберите наиболее значимые атрибуты и оцените степень согласованности экспертов.

вследствие к желаемому (полученному при опросе выборки) результату – женщины покупают больше, чем мужчины).

Альтернативная гипотеза (H_a) – утверждение, опровергающее нулевую гипотезу. В нашем примере, в качестве альтернативной гипотезы может быть выдвинуто утверждение, что мужчины покупают меньше, чем женщины.

Если данные проверки гипотезы приводят к отказу от нулевой гипотезы, то принимается альтернативная гипотеза.

При проведении маркетинговых исследований (МИ) чаще всего в процессе анализа данных выборочного наблюдения используются:

- либо средние значения того или иного признака (средняя величина расходов семьи на покупку составляет 5000 р.);
- либо доля единиц обследованной совокупности, обладающих тем или иным признаком (70 % владельцев мобильных телефонов занимаются новым тарифом).

Поэтому чаще в маркетинговых исследованиях выдвигаются два типа гипотез: гипотезы о частоте распределения и гипотезы о средних величинах.

2) *Выбор соответствующих статистических критериев.*

Статистические критерии (*критерии значимости*) – решающее правило, по которому на основе результатов наблюдений (выборки) принимается решение о задаче статистической проверки гипотез. В табл. 8.1 представлены наиболее часто используемые критерии для проверки статистических гипотез о частоте распределения.

Таблица 8.1

Область применения	Число групп или выборочных совокупностей	Виды шкал	Критерий	Специальные требования
Гипотезы о частоте распределения	Одна	Номинальная (атрибутивная)	χ^2	Случайная выборка
	Две и более	Номинальная (атрибутивная)	χ^2	Случайная выборка, независимые выборки
	Одна	Порядковая (ранги)	Критерий Колмогорова-Смирнова	Случайная выборка

3) *Принятие уровня значимости.*

Уровень значимости (α) – показатель, характеризующий вероятность ошибочного заключения (показатель качества статистической проверки гипотез). Поскольку всякое решение, принимаемое на основе ограниченного ряда наблюдений (выборка), неизбежно сопровождается вероятностью ошибочного решения, важно определить, насколько велика эта вероятность. На практике часто используют следующие стандартные значения уровня значимости: 0,1; 0,05; 0,01; 0,005; 0,001.

Т. е. с вероятностью 95 % (при $\alpha = 0{,}05$) нулевая гипотеза может оказаться отвергнутой, в то время как на самом деле она является справедливой (ошибка первого рода), или, наоборот, гипотеза может быть принята, в то время, как она является ошибочной (ошибка второго рода).

4) *Проверка значимости МИ включает:*

- выбор соответствующей формулы для расчета статистических критериев и расчет критерия по этой формуле;
- сравнение расчетной величины с критическими значениями статистических критериев, которые находятся по специальным статистическим таблицам [1, 16, 21, 23] и принятие решения о нулевой гипотезе: либо она отвергается, либо принимается. Нулевая гипотеза принимается, если расчетное значение критерия меньше, чем табличное.

Пример 8.1. Менеджер магазина проанализировал вторичную информацию (табл. 8.2), где собраны данные о количестве посетителей в магазине по месяцам. Менеджеру необходимо проверить эффективность трех мероприятий, проводимых в магазине с целью привлечения покупателей. В качестве критерия эффективности выбрано число посетителей магазина в первой акции.

Таблица 8.2

Мероприятие	Месяц	Число посетителей, тыс. чел.
1	Апрель	12 800
2	Май	13 100
3	Июнь	12 600
Всего		38 500

Менеджер должен выяснить, существенны ли различия между числом посетителей магазина в различные периоды времени, т.е. это различие случайно или вызвано мероприятиями.

Последовательность проведения проверки значимости МИ (проверки гипотез).

1) *Выдвигается нулевая и альтернативная гипотезы:*

H_0: число посетителей магазина во время проведения трех мероприятий существенно не изменилось (практически одинаково), и это колебание не связано с проводимыми мероприятиями;

H_a: существует значительная разница в числе посетителей магазина, и это колебание связано с проводимыми мероприятиями.

2) *Выбирается статистический критерий* для проверки гипотез (из табл. 8.1) χ^2 для одной выборки.

3) *Принимается уровень значимости* $\alpha = 0{,}05$.

4) *Проверяется значимость МИ.*

Выберем соответствующую формулу для расчета статистических критериев и рассчитаем критерий по этой формуле

$$\chi^2 = \sum_{i=1} \frac{(O_i - E_i)^2}{E_i},$$

где χ^2 – критерий для одной группы (хи-квадрат Пирсона); O_i – наблюдаемая частота в каждой категории; E_i – ожидаемая частота (в соответствии с нулевой гипотезой).

Ожидаемое (теоретическое) число посетителей в случае, если нулевая гипотеза верна: $E = 38500/3 = 12833$.

Тогда, расчетное значение χ^2:

$$\chi^2 = \frac{(12800-12833)^2}{12833} + \frac{(13100-12833)^2}{12833} + \frac{(12600-12833)^2}{12833} = 9{,}870.$$

Число степеней свободы равно (k - 1) = (3 - 1) = 2 (k – число строк, число мероприятий). Следовательно, табличное значение χ^2 при α = 0,05 и числе степеней свободы 2 равно 5,991 [1, 16, 21, 23].

Сравним табличное значение χ^2 с расчетным. Так как расчетное значение (9,870) больше, чем табличное (5,991), нулевая гипотеза отвергается. Следовательно, можно сделать вывод с вероятностью 95 %, что посетители магазина реагируют на проводимые мероприятия.

Пример 8.2. Имеется вторичная информация о частоте посещения магазина (табл. 8.3).

Таблица 8.3

Частота посещения магазина за период	Середина интервала x	Мужчины f_m	Женщины f_f	Итого
1 – 6	3	13	41	54
7 – 13	10	72	63	135
14 – 20	17	12	14	26
Итого		97	118	215

Менеджер рассчитал среднее число (средневзвешенное) посещений мужчин и женщин и увидел, что мужчины посещают магазин чаще.

$$x_m = \frac{3*13 + 10*72 + 17*12}{97} = 9{,}93,$$

Практическое занятие 8

ПРОВЕРКА ЗНАЧИМОСТИ РЕЗУЛЬТАТОВ МАРКЕТИНГОВЫХ ИССЛЕДОВАНИЙ. ГИПОТЕЗЫ О ЧАСТОТЕ РАСПРЕДЕЛЕНИЯ

Цель работы: ознакомиться с методами проверки значимости результатов маркетинговых исследований; научиться применять данные методы в зависимости от области применения.

8.1. Основные сведения

В процессе анализа данных у исследователя регулярно возникает вопрос: достаточно ли значимы результаты исследования? Например, результаты исследования говорят о том, что женщин, покупающих данный товар, больше, чем мужчин (60 % женщин против 20 % мужчин). Но можно предположить, что в выборку попали респонденты, которые не представляют генеральную совокупность, т.е. выборка не отражает в полной мере генеральную совокупность. И тогда распределение частот в выборке нельзя использовать для всей генеральной совокупности. Следовательно, необходимо проверить, действительны ли результаты обследования выборки для всей генеральной совокупности; возможно ли по данным, полученным при обследовании выборки, судить обо всей генеральной совокупности, *т.е. значимы ли* (*достоверны ли*) *результаты маркетинговых исследований?*

Для ответа на этот вопрос используются статистические гипотезы.

Гипотеза – предположения или теории, которые исследователь выдвигает относительно некоторых характеристик генеральной совокупности, подлежащей обследованию [16].

Исследователь выдвигает гипотезу и проверяет ее. В процессе проверки гипотеза либо подтверждается, либо опровергается и в зависимости от полученного результата принимается управленческое решение.

Последовательность проверки значимости результатов маркетинговых исследований [1, 16, 23].

1) ***Выдвижение гипотезы (нулевой и альтернативной).***

Нулевая гипотеза (H_0) – утверждение, в котором исследователь констатирует *факт отсутствия* каких-либо отличий либо влияния в исходных данных. В нашем примере в качестве нулевой гипотезы должно быть выдвинуто утверждение, что мужчины больше покупают данный товар, т.е. формулируется нулевая гипотеза так, чтобы отказ от нее приводил к желательному результату (в нашем случае отказ от нулевой гипотезы при-

Таблица 7.5

Атрибуты	Экспертная оценка значимости атрибутов (коэффициенты весомости)									
	1	2	3	4	5	6	7	8	9	10
Работа пресонала	0,4	0,2	0,2	0,3	0,3	0,4	0,3	0,4	0,1	0,4
Качество блюд	0,4	0,3	0,2	0,3	0,4	0,3	0,3	0,2	0,4	0,3
Цены	0,1	0,4	0,3	0,3	0,1	0,2	0,2	0,2	0,3	0,1
Атмосфера получения услуги	0,1	0,1	0,3	0,1	0,2	0,1	0,2	0,2	0,2	0,2

2) Экспертов попросили оценить степень влияния каждого из четырех факторов на перспективу развития малого бизнеса в регионе (табл. 7.6).

Укажите наиболее значимые факторы, способствующие развитию малого бизнеса в регионе, и оцените степень согласованности экспертных оценок.

Таблица 7.6

Атрибуты	Экспертная оценка значимости факторов (коэффициенты весомости)									
	1	2	3	4	5	6	7	8	9	10
Льготное кредитование	0,1	0,2	0,1	0,1	0,1	0,2	0,1	0,1	0,2	0,1
Льготное налогообложение	0,2	0,1	0,2	0,3	0,1	0,2	0,1	0,2	0,2	0,2
Уменьшение количества проверок	0,3	0,3	0,3	0,2	0,4	0,2	0,2	0,3	0,3	0,3
Снижение закупочных цен	0,4	0,4	0,4	0,4	0,4	0,4	0,6	0,4	0,3	0,4

7.3. Контрольные вопросы

1) Для каких целей используется метод рангової корреляции?

2) Какие две задачи решаются в рамках применения метода рангової корреляции?

3) Перечислите этапы расчета среднего коэффициента весомости.

4) Если коэффициент конкордации равен 0,20, то какие можно сделать выводы об обработанных экспертных оценках?

5) Если степень согласованности мнений экспертов невелика (стремится к нулю), то какие действия необходимо далее предпринять для принятия окончательной экспертной оценки. За пример возьмите одну из рассматриваемых выше ситуаций.

Среднее число посещений магазина мужчинами 9,93, женщинами – 8,40. На первый взгляд создается впечатление, что мужчины посещают магазин чаще. Менеджеру необходимо определить, действительно ли есть связь между полом покупателей и частотой посещения магазинов или такое распределение частот случайно.

Последовательность проведения проверки значимости МИ (предполагаемых данных).

1) *Выдвигаются нулевая и альтернативная гипотезы:*

H_0: между полом и частотой посещения магазина связи нет;

H_a: есть связь между двумя переменными: полом и частотой посещения.

2) *Выбирается статистический критерий* для проверки гипотез (из табл. 8.1) χ^2 для двух выборок.

3) *Принимается уровень значимости* $\alpha = 0{,}05$.

4) *Проверяется значимость МИ.*

Рассчитывается χ^2 – *критерий для двух* (*и более*) *независимых групп:*

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{k} \frac{(O_{ij} - E_{ij})^2}{E_{ij}},$$

где O_{ij} – наблюдаемое число в каждой i-й строке j-й колонки; E_{ij} – ожидаемое число в i-й строке j-й колонки (в соответствии с нулевой гипотезой).

Определим ожидаемые частоты для каждой группы, попавшей в исследование (табл. 8.4).

Таблица 8.4

Частота посещений	Мужчины	Женщины
1 – 6	(97*54)/215 = 24,36	(118*54)/215 = 29,64
7 – 13	(97*135)/215 = 60,91	74,07
14 – 20	(97*26)/215 = 11,73	14,27

Рассчитаем теоретическую величину χ^2:

$$\chi^2 = \frac{(13-24{,}36)^2}{24{,}36} + \frac{(41-29{,}64)^2}{29.64} + \frac{(72-60{,}91)^2}{60{,}91} + \frac{(63-74{,}09)^2}{74{,}09} + \frac{(12-11{,}73)^2}{11{,}73} + \frac{(14-14{,}27)^2}{14{,}27} = 13{,}35$$