

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ
ФЕДЕРАЦИИ**

**Федеральное государственное автономное образовательное учреждение
высшего образования
«Севастопольский государственный университет»**

МЕТОДИЧЕСКИЕ УКАЗАНИЯ

**по выполнению лабораторных занятий по
дисциплине «Исследование и анализ рынков»
студентов специальности
38.03.02 «Менеджмент»
всех форм обучения**

Севастополь

2025

УДК 658

Методические указания по выполнению лабораторных занятий по дисциплине «Исследование и анализ рынков» для всех форм обучения специальности 38.04.02 «Менеджмент» // сост. И.А. Гребешкова – 2025. – 43с.

Целью методических указаний является систематизация теоретического материала и практических рекомендаций для выполнения лабораторных занятий по дисциплине «Исследование и анализ рынков» для студентов всех форм обучения специальности 38.03.02 «Менеджмент»

Утверждена и введена в действие на заседании кафедры Менеджмента и бизнес-аналитики от «____» _____ 2025г., протокол №____.

Содержание

Практическое занятие № 1 Анализ данных: оценка различий	4
Практическое занятие № 2 Регрессионный анализ	18
Практическое занятие № 3 Использование математических методов многомерного анализа для решения проблемы размещения объекта	26
Практическое занятие № 4 Анализ рынка	41
Литература	43

Лабораторное занятие № 1

Анализ данных: оценка различий

ТЕОРЕТИЧЕСКИЕ ВОПРОСЫ АНАЛИЗА ДАННЫХ: ОЦЕНКА РАЗЛИЧИЙ

В маркетинговых исследованиях часто возникают ситуации, когда аналитик должен определить, соответствует ли определенный образец поведения, о котором свидетельствуют данные, тому образцу, который ожидалось обнаружить, когда исследование задумывалось.

1. Проверка согласия по критерию хи-квадрат – статистическая проверка для определения, соответствует какой-то наблюдавшийся образец частот распределению гипотетической генеральной совокупности.

2. Проверка Колмогорова-Смирнова – статистическая проверка, проводимая по выборке из упорядоченных данных для определения, соответствует ли определенный наблюдавшийся образец частот ожидаемому образцу; используется так же для определения того, взяты ли две независимые выборки из одной и той же генеральной совокупности или из совокупностей с одним и тем же распределением.

3. Гипотеза о значении среднего для одной выборки.

4. Гипотеза о значении среднего двух выборок

5. Гипотеза о двух долях.

Проверка согласия по критерию хи-квадрат

Множество значений, принимаемых интересующей нас переменной, разбивается на k взаимоисключающих интервалов. Каждое наблюдение логически попадает в один из этих интервалов. Предполагается, что испытания независимы и объем выборки велик.

Необходимо определить значения вероятностей попадания в рассматриваемые интервалы для значений из гипотетической генеральной совокупности (ожидаемое число) и сравнить их с числом значений из выборки действительно попавших в соответствующие интервалы (наблюдавшимся числом событий), используя формулу:

$$\chi^2 = \sum_{i=1}^k \frac{[O_i - E_i]^2}{E_i},$$

где O_i - наблюдавшееся число событий, попавших в i -ый интервал, E_i - ожидаемое число событий, попавших в i -ый интервал, k - количество интервалов.

При этом осуществляется проверка семейства гипотез:

H_0 : наблюдающийся образец частот соответствует распределению гипотетической генеральной совокупности;

H_1 : наблюдающийся образец частот не соответствует распределению гипотетической генеральной совокупности

Расчетное значение χ_p^2 сопоставляется с табличным значением χ_m^2 с числом степеней свободы $k-1$. Под термином число степеней свободы подразумевается такое количество параметров, характеризующих состояние некоторого объекта, которые могут меняться независимо.

Если расчетное значение хи-квадрат больше табличного, то нулевая гипотеза отвергается с соответствующим числом степеней свободы.

Проверка Колмогорова-Смирнова

Проверка Колмогорова-Смирнова подобна проверке согласия по критерию хи-квадрат в том, что в ней используется сравнение наблюдавшейся и ожидаемой частот для определения того, находятся ли наблюдавшиеся результаты в согласии с заявленной нулевой гипотезой. Но в проверке Колмогорова-Смирнова используется преимущество природы упорядоченных данных.

Гипотеза о значении среднего для одной выборки

Периодически возникающей проблемой маркетинговых исследований является необходимость делать определенные предположения о среднем значении генеральной совокупности. Отметим, что когда берутся выборки из какой-то генеральной совокупности с известной дисперсией, то распределение выборочных средних совпадает с распределением генерального среднего, а разброс выборочных средних равен разбросу генерального среднего, поделенному на объем выборки.

Для проверки гипотезы о значении среднего для одной выборки используется следующие статистические критерии:

Статистика z приемлема, если выборка берется из нормальной генеральной совокупности, или если переменная распределена в генеральной совокупности не по нормальному закону, но выборка достаточно велика для того, чтобы подпадать под действие центральной предельной теоремы.

Когда разброс генеральной совокупности неизвестен, конечно, неизвестна и стандартная ошибка среднего $\sigma_{\bar{x}}$, поскольку она равна $\sigma/\sqrt{n}$, где n – размер выборки. Стандартная ошибка среднего должна в таком случае оцениваться по выборочным данным. В этом случае тестовой статистикой становится t критерий Стьюдента.

Гипотеза о значении среднего двух выборок

Проверка данного рода гипотез рассматривается в рамках зависимых и независимых выборок. При рассмотрении гипотезы равенства средних двух выборок при предположении о независимости этих выборок подлежат рассмотрению три случая (переменная распределена по нормальному закону в каждой генеральной совокупности):

- Дисперсии обеих генеральных совокупностей известны;
- Дисперсии генеральных совокупностей неизвестны, но считаются равными;

- Дисперсии генеральных совокупностей неизвестны и не являются равными.

Проверка различия двух долей

Проверка различия двух долей генеральной совокупности является в основном проблемой объема выборок. Каждая выборка должна быть достаточно большой. Чтобы для явно биномиального распределения выборочных долей могла использоваться нормальная аппроксимация. Практически это означает, что для любой выборки np и nq должны быть больше 10, где n – доля «успехов», p – доля «неудач» в выборке, q – объем выборки. Рассмотрим два вида гипотез:

$$H_0 : \pi_1 = \pi_2,$$

$$H_1 : \pi_1 \neq \pi_2,$$

где 1 относится к первой выборке, а 2 – ко второй выборке. Статистикой проверки является:

$$z = \frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sigma_{p_1 - p_2}},$$

где $\sigma_{p_1 - p_2} = pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$ – стандартная ошибка разности двух выборочных долей, p – суммарное число успехов в двух выборках/суммарный объем наблюдений в двух выборках и $q=1-p$.

ОЦЕНКА РАЗЛИЧИЙ В СИСТЕМЕ MINITAB И MICROSOFT EXCEL

Проверка согласия по критерию хи-квадрат (Microsoft Excel)

ХИ2ТЕСТ

Возвращает тест на независимость. ХИ2ТЕСТ возвращает значение для распределения хи-квадрат (χ^2). Критерий χ^2 используется для определения того, подтверждается ли гипотеза экспериментом.

Синтаксис

ХИ2ТЕСТ(фактический_интервал;ожидаемый_интервал)

Фактический_интервал – это интервал данных, которые содержат наблюдения, подлежащие сравнению с ожидаемыми значениями.

Ожидаемый_интервал – это интервал данных, который содержит отношение произведений итогов по строкам и столбцам к общему итогу.

Замечания

- Если фактический_интервал и ожидаемый_интервал имеют различное количество точек данных, то функция ХИ2ТЕСТ возвращает значение ошибки #Н/Д.

- Критерий χ^2 сначала вычисляет χ^2 статистику, а затем суммирует разности между фактическими значениями и ожидаемыми значениями. Уравнение для этой функции имеет следующий вид: $\text{HI2TEST} = p(X > \chi^2)$. HI2TEST возвращает вероятность для χ^2 статистики и степеней свободы df , где $df = (r - 1)(c - 1)$.

Гипотеза о значении среднего для одной выборки

Для проведения проверки статистической гипотезы в пакете MINITAB необходимо:

- 1) ввести данные.
- 2) из меню **Stat > Basic Statistics** выбрать тест, необходимый для проверки статистической гипотезы:
 1. **1-Sample z** - тест используется при проверке статистической гипотезы о равенстве средней определенному значению в случае, когда известен закон распределения (известно среднееквадратическое отклонение).
 2. **1-Sample t** - тест используется при проверке статистической гипотезы о равенстве средней определенному значению.

Гипотеза о значении среднего двух выборок

Для проведения проверки статистической гипотезы в пакете MINITAB необходимо:

- 1) ввести данные.
- 2) из меню **Stat > Basic Statistics** выбрать тест, необходимый для проверки статистической гипотезы:

2-Sample t - тест используется при проверке статистической гипотезы о равенстве средних двух выборок (выборки независимы). MINITAB позволяет рассмотреть два случая:

 - Дисперсии генеральных совокупностей неизвестны, но считаются равными – в этом случае в диалоговом окне отмечаете **Assume equal variances** – предположить равенство дисперсий;
 - Дисперсии генеральных совокупностей неизвестны и не являются равными.

Microsoft Excel позволяет проверить гипотезу при условии что дисперсии обеих генеральных совокупностей известны:

Сервис/ Анализ данных/ Двухвыборочный z тест для средних

Гипотеза о значении среднего двух зависимых выборок в пакете Statistica выполняется следующим образом:

Открыть пакет STATISTICA. В диалоговом окне открыть Basic Statistics при помощи клавиши Switch too. Ввести данные: выборки должны располагаться в

разных столбцах. После этого выбрать на панели инструментов Analysis/t-test for dependent samples. В диалоговом окне выбрать переменные – Variables.

ОЦЕНКА РАЗЛИЧИЙ (ПРИМЕРЫ)

Проверка согласия по критерию хи-квадрат

Поставщиком готовых завтраков был разработан рецепт новой каши. Каша поставлялась в трех размеров стандартных упаковках: малом, большом и семейном. В прошлом этот поставщик установил, что на каждую одну проданную упаковку малого размера продавались три большого и две семейного размеров. Ему захотелось посмотреть, сохраняется та же тенденция для новой каши, поскольку изменение образцов потребления могло существенно отразиться на производстве. По этой причине был проведен рыночный тест для определения относительных частот приобретения покупателями новой каши в разных упаковках.

В процессе надлежащим образом проведенного рыночного теста недельной продолжительности было продано 1200 коробок новой каши, и что распределение этого объема продаж по размерам коробок выглядело следующим образом (таблица 1).

Таблица 1 – Результаты рыночного теста приобретения каши.

Количество упаковок			
Малых	Больших	Семейных	Итого
240	575	385	1200

Основываясь на результатах исследований, можно ли говорить о том, что фирма должна ожидать изменение образцов спроса на новый вид каши в упаковках различного размера.

Для ответа на этот вопрос сформулируем нулевую гипотезу: состав продаж каши, отличающейся размерами упаковки, будет повторять нормальное для производителя соотношение (то есть на продажу каждой малой упаковки будет приходится продажа трех упаковок большого размера и двух семейного размера). В терминах долей всех продаж это будет означать

$$\text{малый размер: } \frac{1}{1+3+2} = \frac{1}{6} ;$$

$$\text{большой размер: } \frac{3}{1+3+2} = \frac{3}{6};$$

$$\text{семейный размер: } \frac{2}{1+3+2} = \frac{2}{6}.$$

Следовательно, что если продажа 1200 коробок во время тестирования рынка подчинялась нормальному или ожидаемому образцу потребления, то ожидаемое потребление было бы таким, как представлено в таблице 2.

Таблица 2 – Ожидаемое потребление каши

Количество упаковок			
Малых	Больших	Семейных	Итого
200	600	400	1200

Таким образом, в соответствии с формулой хи-квадрат расчетное составит 9,6. Табличное значение хи-квадрат составляет 5,99 с уровнем значимости 5% и числом степеней свободы 2.

Так как расчетное значение показателя больше табличного. То нулевая гипотеза отвергается. Заключение состоит в том, что выборочный результат вряд ли является всего лишь случайным. Результаты предварительного рыночного тестирования показывают, что продажа каши будет идти иным образом, чем считалось типичным для данной продукции.

Для работы в Excel сформируем таблицу и используем соответствующую функцию (рисунок 1).

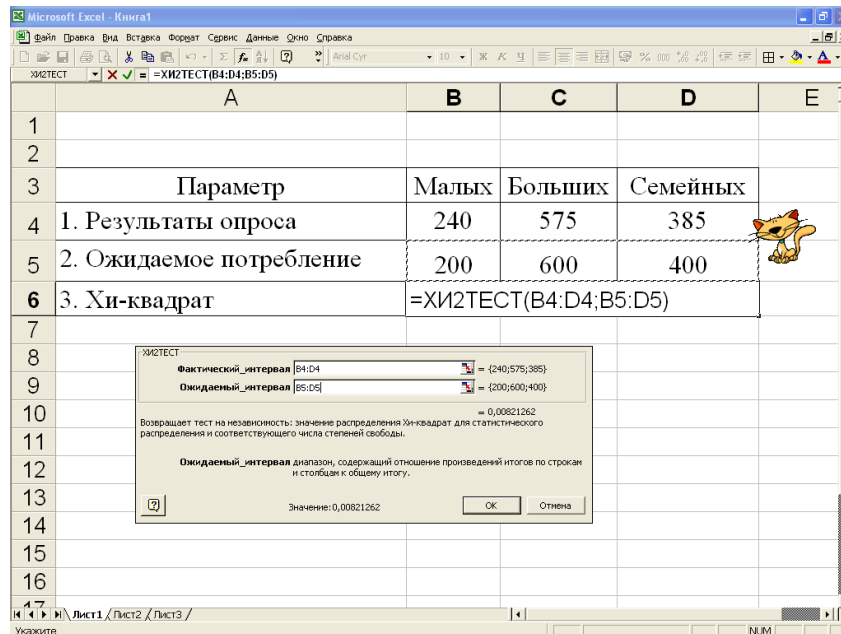


Рисунок 1 – Расчет хи-квадрата в Excel

Расчетное значение составило 0,008. Для проверки выбран уровень значимости 5%. Так как Расчетное значение вероятности меньше 0,05, то нулевая гипотеза отвергается, что подтверждено расчетами выше.

Проверка Колмогорова-Смирнова

Изготовитель косметики, который проверяет четыре различных цвета какого-то основного состава: очень светлый, светлый, средний и темный. Компания наняла фирму, специализирующуюся на маркетинговых исследованиях для определения того, существуют ли какие-то различия в предпочтении одной из крайностей. Если таковое существует, компания будет производить только предпочтительные тона. В противном случае на рынок поступят все четыре тона. Предположим, что в выборке из ста человек очень светлый тон предпочитает 50, светлый – 30, средний – 15, темный – 5 человек. Могут ли эти результаты указать на какого-то рода предпочтения?

Результаты необходимых расчетов сформируем в таблицу 3.

Таблица 3 – Результаты опроса.

Тон	Наблюдавшееся число	Наблюдавшаяся доля	Накопленная доля	Теоретическая доля	Теоретическая накопленная доля
Очень светлый	50	0.5	0.5	0.25	0.25
Светлый	30	0.3	0.8	0.25	0.5
Средний	15	0.15	0.95	0.25	0.75
Темный	5	0.05	1	0.25	1

Поскольку тона представляют собой естественную упорядоченность, для проверки гипотезы о предпочтении может быть использована проверка Колмогорова-Смирнова. Эта проверка включает в себя спецификацию функции накопленного распределения, которое могло бы иметь место в рамках нулевой гипотезы, и ее сравнение с функцией наблюдавшегося накопленного распределения. Определяется точка, в которой две функции показывают максимальное отклонение, и значение этого отклонения является статистикой проверки.

Рассмотрим следующие гипотезы:

H_0 : предпочтения различных тонов не существует;

H_1 : предпочтения различных тонов существует.

Для принятия либо отвержения гипотезы используем величину D в проверке Колмогорова-Смирнова, которая равна абсолютному значению этого максимального отклонения наблюдавшейся накопленной доли от теоретической доли, составляет $0.8-0.5=0.3$. При $\alpha = 0.05$ критическое значение D для больших выборок задается как $1.36/\sqrt{n}$, где n - объем выборки. В нашем случае объем выборки равен 100, следовательно, критическое значение составляет 0.136. Расчетное значение D превышает критическое значение и, таким образом, нулевая

гипотеза об отсутствии предпочтения тонов отвергается. Данные указывают на статистически значимое предпочтение более светлых тонов.

Гипотеза о значении среднего для одной выборки

Рассмотрим сеть супермаркетов, которая исследует желательность добавления какого-то нового товара к тому, что уже есть на полках ее магазинов. Поскольку многие товары должны конкурировать между собой за обладание ограниченным пространством полок, запасы определялись исходя из того, что недельная продажа в каждом магазине 100 единиц данной позиции ассортимента является достаточной гарантией прибыльности. Предположим, что исследовательский отдел решил изучить оборот по рассматриваемой позиции, представив данные в виде случайной выборки по десяти магазинам за какой-то ограниченный период времени. Предположим, что средний недельный объем продаж одного магазина оказался таким, как представлено в таблице 4.

Таблица 4 - Недельный объем продаж опытного товара

Магазин	Продажи	Магазин	Продажи
1	86	6	130
2	97	7	132
3	114	8	116
4	108	9	105
5	123	10	120

Поскольку разброс объема продаж одним магазином неизвестен и должен оцениваться, то используется t статистика? Предположение о нормальности распределения представляется разумным и может быть проверено с использованием одного из критериев согласия. Используем критерий Колмогорова-Смирнова. Выполним команду Stat > Basic Statistics > Normality Test в системе MINITAB. В диалоговом окне следует только указать переменную, проверяемую на нормальность, и тест, который используется для проверки – критерий Колмогорова-Смирнова. Результаты расчетов представлены на рисунке 1. В соответствии с рисунком 1, принимается нулевая гипотеза о том, что распределение объема продаж подчинено нормальному закону распределения ($p > 0.15$), что подтверждает график.

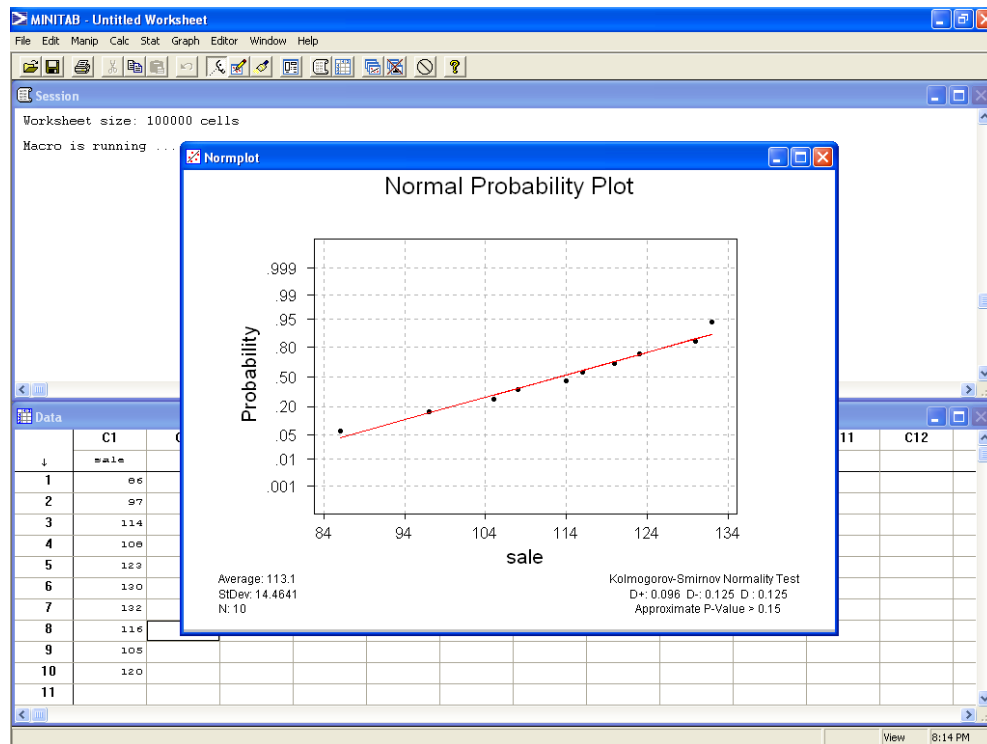


Рисунок 1 - Проверка на нормальность среднего недельного объема продаж

Так как объем продаж одного магазина в неделю составит, по крайней мере, 100 единиц, этот товар будет выведен на рынок в национальном масштабе. Нулевая и альтернативная гипотезы:

$$H_0 : \mu = 100$$

$$H_1 : \mu > 100$$

Уровень значимости примем равным 0.05 или 5%. Результаты расчетов приведены на рисунке 2.

T-Test of the Mean

Test of $\mu = 100.00$ vs $\mu > 100.00$

Variable	N	Mean	StDev	SE Mean	T	P
C1	10	113.10	14.46	4.57	2.86	0.0093

Рисунок 2 - Результаты проверки семейства гипотез

Следовательно, так как $p < \alpha$, то примем альтернативную гипотезу о том, что недельный объем продаж опытного товара будет выше 100 единиц. Прогноз ожидаемых недельных продаж одного магазина при условии вывода товара на рынок в национальном масштабе можно получить, рассчитав доверительный интервал:

выборочное среднее $\pm t$ или $\bar{x} \pm t_{s_{\bar{x}}}$.

Для 95% доверительного интервала и 9 степеней свободы ($n-1$) $t=1.833$. Таким образом, 95% доверительный интервал составляет, либо

$$104.72 \leq \mu \leq 121.48.$$

Проверка различия двух долей

Проверка различия двух долей генеральной совокупности является в основном проблемой объема выборок. Каждая выборка должна быть достаточно большой. Чтобы для явно биномиального распределения выборочных долей могла использоваться нормальная аппроксимация. Практически это означает, что для любой выборки np и nq должны быть больше 10, где n – доля «успехов», p – доля «неудач» в выборке, q – объем выборки.

Изготовитель косметики интересуется сравнением использования распыляемых жидкостей для ухода за волосами мужчинами из числа студентов колледжей и тех, кто в колледжах не учится. Предположим, были взяты выборки из 100 студентов и 100 не студентов города, для которых определялась интенсивность пользования распыляемыми жидкостями для ухода за волосами в течении последних трех месяцев. Из выбранного числа мужчин 30 студентов и 20 не студентов в этот период такими жидкостями пользовались. Доказывает ли это, что доля студентов, пользующихся распыляемыми жидкостями, значительно выше пользующихся ими не студентов?

Рассмотрим два вида гипотез:

$$H_0 : \pi_1 = \pi_2,$$

$$H_1 : \pi_1 \neq \pi_2,$$

где 1 относится к популяции мужчин-студентов колледжей, а 2 – к популяции мужчин, которые студентами не являются. Выборочные доли $p_1=0.3$ и $p_2=0.2$, поэтому $n_1p_1=30$, $n_1q_1=70$, $n_2p_2=20$, $n_2q_2=80$. Статистикой проверки является:

$$z = \frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sigma_{p_1 - p_2}},$$

где $\sigma_{p_1 - p_2} = pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$ – стандартная ошибка разности двух выборочных долей,

p – суммарное число успехов в двух выборках/суммарный объем наблюдений в двух выборках и $q=1-p$.

Следовательно

$$p = \frac{30 + 20}{100 + 100} = \frac{50}{200} = 0.25$$

и

$$\sigma_{p_1 - p_2} = 0.25 \cdot 0.75 \left(\frac{1}{100} + \frac{1}{100} \right) = 0.061.$$

Расчетное z находится следующим образом:

$$z = \frac{(0.3 - 0.2) - (0)}{0.061} = \frac{0.1}{0.061} = 1.61.$$

$z_t = 1.96$ при $\alpha = 0.05$.

Следовательно, выборочные результаты не указывают на существование различий в использовании распыляемых жидкостей для ухода за волосами между мужчинами-студентами колледжей и мужчинами, не являющимися студентами.

ЗАДАНИЕ

1. В крупном издательстве провели опрос с целью выявить читательские пристрастия пожилых людей. Компания издала четыре журнала, специально ориентированные таким образом, чтобы в полной мере соответствовать потребностям пожилой части читательской аудитории. Менеджмент издательства выдвинул гипотезу, согласно которой различий в предпочтении того или иного журнала отмечено не было. Интервью с 1600 пожилыми гражданами из числа жителей города показали следующие данные о предпочтении четырех журналов (таблица 5):

Таблица 5 – Опрос жителей

Издание	Частота предпочтения						
	1	2	3	4	5	6	7
1. Golden years (Золотые годы)	350	300	550	350	380	400	390
2. Maturation (Зрелость)	500	450	450	400	470	480	370
3. High Serenity (Верх безмятежности)	450	400	350	380	370	300	450
4. Time of Living (Время жизни)	300	450	250	470	380	420	390
Итого	1600	1600	1600	1600	1600	1600	1600

Менеджменту необходима ваша экспертная оценка: определения того, имеет ли место предпочтение журналов пожилыми гражданами. (Рекомендации: заявите нулевую и альтернативную гипотезы, должна ли быть отвергнута нулевая гипотеза при уровне значимости 5%, 10%).

2. Silken-Shine Company – это среднего размера изготовитель шампуней. Для увеличения своей доли рынка в последние годы компания расширяла разнообразие продукции, имеющей общее название шампунь Silken-Shine, с трех марок до пяти. Менеджмент провел опрос, имевший целью сравнение объемов продаж шампуня Silken-Shine с шампунями конкурентов Rapunzel и So-Soft. Выборка из 1800 домохозяек показала следующие частоты в отношении приобретаемых ими недавно шампуней (таблица 6):

Таблица 6 – Предпочтения потребителей

Шампунь	Число покупок						
	1	2	3	4	5	6	7
1. Silken-Shine	425	854	580	900	500	456	360
2. Rapunzel	1175	456	956	805	890	950	890
3. So-Soft	200	490	264	95	410	394	550
Итого	1800	1800	1800	1800	1800	1800	1800

Опыт показал, что шампунь Rapunzel предпочитался втрое большим числом домохозяек, чем шампунь Silken-Shine, но последнему отдавало предпочтение вдвое больше домохозяек, чем шампуню So-Soft. Менеджмент желает определить, сохраняется ли исторически сложившаяся тенденция в условиях увеличения компанией ассортимента ее шампуней на рынке.

3. Президент сети универсальных магазинов обещал менеджерам магазинов 8% премию, если они увеличат среднемесячный объем своих продаж на 300 тыс.руб. или более. Средняя выборка из 12 магазинов показала следующее увеличение объемов продаж (таблица 7).

Таблица 7 – Объемы продаж

Магазин	Продажи, тыс.руб.						
	1	2	3	4	5	6	7
1	320	308	353	305	221	505	379
2	230	196	231	435	231	400	362
3	400	371	407	243	313	392	501
4	450	490	477	370	222	571	386
5	280	481	166	173	321	131	232
6	320	542	388	217	148	281	392
7	380	92	445	385	297	155	380
8	280	316	375	473	278	314	504
9	420	469	395	256	390	362	253
10	360	218	434	407	350	323	409
11	440	264	484	351	398	277	423
12	320	149	428	459	324	268	501

Президент задается вопросом, показывает ли эта случайная выборка магазинов, что совокупность магазинов в целом добивается поставленной цели.

4. Come-and-Go Company, крупное туристическое агентство, пожелала изучить предпочтения клиентов в отношении своих комплексных туров в страны Востока.

В течение предыдущих пяти лет Come-and-Go Company предлагала два одинаковых по цене тура на Восток, отличавшихся только включенными в них местами для посещений. Была осуществлена случайная выборка приобретений путевок за пять месяцев из числа этих пяти лет. Ниже представлено количество клиентов, которые приобретали путевки в течении года (таблица 8).

Таблица 8 - Приобретение туров клиентами.

Месяц	x1	x2	x3	x4	x5	x6	x7	x8
1	855	956	546	550	890	640	680	230
2	795	880	500	490	800	640	545	780
3	810	860	482	480	450	700	600	650
4	931	900	490	498	500	450	700	700
5	830	749	530	538	350	470	560	600
6	710	708	600	610	700	540	860	560
7	790	810	650	650	890	900	600	800
8	830	890	620	640	640	800	540	450
9	879	897	593	599	600	740	500	560
10	798	819	562	580	640	700	650	560
11	765	786	589	580	540	760	760	700
12	856	789	562	560	550	500	450	580

Таблица 9 - Варианты заданий

Показатель	Вариант						
	1	2	3	4	5	6	7
Комплексный тур 1	x1	x2	x3	x4	x5	x6	x7
Комплексный тур 2	x2	x3	x4	x5	x6	x7	x8

Менеджмент Come-and-Go Company нуждается в вашей помощи для определения того, существуют ли различия в предпочтении того либо другого туров.

5 Задание выполняется в соответствии с вариантом таблицы 10.

Таблица 10 – Варианты заданий

Задание	Вариант
5.1	1,2,3,4
5.2	5,6,7

5.1 Творческое ателье агентства разработало 2 различных подхода, обозначаемых А и И, к организации искушения рекламой непосредственно по почте для первого своего главного клиента. Чтобы проверить эффективность разных способов искушения, отделу исследований было поручено провести почтовый эксперимент. Были составлены две независимые случайные выборки $n=2000$, а по истечении двухмесячного периода ожидания полученные от каждой выборки заказы были сведены в таблицы. Подходу А соответствовало 257 заказов, подход В дал 230 заказов. Отличаются ли эффективности подходов.

5.2. Изготовитель кассет для записи музыкальных произведений желает проверить четыре типа кассет, отличающихся длиной ленты: 30-, 60-, 90- и 120-минутные. Компания наняла Вас для определения того, отдают ли потребители ее продукции предпочтение любой из крайностей. Если имеет место предпочтение той или другой из них, то компания будет производить кассеты только предпочтительной длины; в противном случае компанией планируется выход на рынок с кассетами всех четырех длин ленты. Выборка из 1000 потребителей указывает следующие предпочтения (таблица 11).

Таблица 11 – Результаты опроса предпочтений потребителей

Длина ленты	Частота предпочтения
30 минут	150
60 минут	250
90 минут	425
120 минут	175
Всего	1000

6 Индивидуальное задание.

Выполняется после защиты основного материала лабораторной работы и защищается отдельно от лабораторной с предоставлением отчета по исследованию. Наличие презентации приветствуется в дополнении к отчету.

Необходимо выполнить проверку на вопрос предпочтения торговых марок товара (услуги). Количество ТМ не менее трёх. В работе должна быть дана характеристика ТМ товара (услуги). Количество респондентов 50. Опрос проводить со слов респондента. При защите индивидуального задания к отчёту прилагается сводная таблица по опросу респондентов. При опросе фиксировать дополнительно следующие характеристики респондентов: пол, возраст (возможна интервальная оценка), социальный статус. В отчете в рамках выводов по исследованию предложить мероприятия для улучшения ситуации по ТМ.

Лабораторное занятие № 2

Регрессионный анализ

1 Многофакторный регрессионный анализ

Регрессионный анализ объединяет широкий круг задач, связанных с построением функциональных зависимостей между двумя группами числовых переменных: $x_1, \dots, x_p$ и $y_1, \dots, y_q$. Для краткости объединим $x_1, \dots, x_p$ в многомерную переменную x . $y_1, \dots, y_q$ — в переменную y , и будем говорить об исследовании зависимости между x и y . При этом будем считать x независимой переменной, влияющей на значения y . В связи с этим будем называть y откликом, а $x = (x_1, \dots, x_p)$ — факторами, влияющими на отклик.

Многомерная линейная регрессия позволяет построить и проверить модель линейной связи между одной зависимой (эндогенной) и несколькими независимыми (экзогенными) переменными.

Уравнение множественной регрессии (то есть регрессии с двумя и более переменными) имеет вид:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \xi,$$

где Y_i — зависимая переменная, X_i — независимая переменная, $\beta_0, \beta_1, \beta_2, \beta_k$ — коэффициенты регрессионного уравнения, ξ — ошибка с нормальным законом распределения, средним равным нулю и стандартным отклонением σ . Матричная запись модели имеет вид:

$$Y = X\beta + \xi,$$

где $Y = (Y_1, Y_2, \dots, Y_n)^T$ — вектор-столбец значений зависимой переменной, X — детерминированная матрица объясняющих переменных (регрессоров), $\beta = (\beta_1, \beta_2, \dots, \beta_k)^T$ — вектор столбец параметров модели, $\xi = (\xi_1, \xi_2, \dots, \xi_n)^T$ — вектор столбец ошибок.

Статистический подход к задаче построения функциональной зависимости y от x основывается на предположении, что нам известны некоторые исходные (экспериментальные) данные (x_i, y_i) , где y_i — значение отклика при заданном значении фактора x_i , i изменяется от 1 до n . Пару значений (x_i, y_i) часто называют результатом одного измерения, а n — числом измерений.

Мультиколлинеарность

Одним из условий регрессионного анализа является независимость объясняющих переменных. При анализе реальных, экономических процессов это условие часто нарушается. Например, в работе оценщика цена продаваемого дома может быть связана со значениями таких независимых переменных, как год

постройки дома, его жилая площадь. Количество жилых комнат и прочих помещений, общее количество продаваемых домов, индекс, учитывающий качество конструкции продаваемого дома. В таких случаях жилую площадь, количество жилых комнат и прочих помещений следует рассматривать вместе. Если одно из этих значений возрастает, другие тоже, как правило, увеличиваются.

Линейная зависимость между двумя или более переменными называется **мультиколлинеарностью**. Степень мультиколлинеарности измеряется **фактором роста дисперсии (VIF)**:

$$VIF_j = \frac{1}{1 - R_j^2}, j=1,2, \dots, k$$

Здесь R_j^2 — это коэффициент детерминации из регрессии j -ой независимой переменной по оставшимся $(k-1)$ независимым переменным. Если j -я независимая переменная x_j не связана с остальными x , $R_j^2=0$ и $VIF_j = 1$. Если зависимость имеет место, то $VIF_j > 1$.

Значение VIF_j , близкое к 1, говорит о том, что для этой переменной проблемы мультиколлинеарности не существует. Оценка ее коэффициента и значение t -статистики не изменяется значительно, если другие независимые переменные будут добавлены в уравнение регрессии или удалены из него. Значение VIF_j , гораздо большее 1, указывает, что оценка коэффициента при этой независимой переменной неустойчива. В сущности, большая величина фактора VIF_j означает, что имеется избыточная информация о независимых переменных. Информация, предоставляемая переменной с большим VIF_j , уже присутствует в других независимых переменных. При наличии мультиколлинеарности интерпретировать влияние отдельной независимой переменной на зависимую гораздо сложнее.

2 Многофакторный регрессионный анализ в системе MINITAB

Для проведения регрессионного анализа необходимо:

- 1) ввести данные;
- 2) **Stat > Regression > Regression** — позволяет выполнить простую или множественную регрессию (построение зависимости между откликом и переменной), и заполнить следующие ячейки:

Response — переменная отклика;

Predictors — независимые переменные;

Options — опции;

Graphs — построение графиков;

Storage — сохранение полученных результатов;

Stat > Regression > Stepwise — позволяет выполнить прямую (включение) и обратную (исключение независимых переменных) пошаговую регрессию;

Stat > Regression > Best Subsets – позволяет осуществить выбор влияющих переменных из потенциального множества переменных на основе R^2 критерия.

Stat > Regression > Fitted Line Plot — позволяет построить линейную, квадратическую или кубическую регрессии от одной переменной. Пункты диалогового окна:

Response [Y] — переменная отклика или независимая переменная,

Predictor [X] — независимая переменная.

Type of Regression Model (типы моделей регрессии):

Linear — модель линейной регрессии,

Quadratic — модель квадратичной регрессии,

Cubic — модель кубической регрессии.

3) При проведении многофакторного регрессионного анализа t-критерий Стьюдента обеспечивает эффективную проверку предельного вклада каждой переменной при допущении, что все другие переменные уже включены в уравнение. Если объясняющая способность независимых переменных перекрывается (присутствует мультиколлинеарность), то предельный вклад в объяснение каждой из переменных может быть небольшим (t – статистика для переменных не значима), в то же время, как F-статистика указывает на адекватность.

4) В MINITAB для проверки явления мультиколлинеарности необходимо в команде **Stat > Regression > Regression** в опции **Options** отметить **Variance inflation factors**: — фактором роста дисперсии (VIF). Если значение VIF очень велико, то оно свидетельствует о том, что данная переменная связана с стальными переменными. О наличии связи между переменными может указывать выборочный коэффициент корреляции. На следующем этапе удаляется та переменная из модели, у которой значение t-статистики будет наименьшим.

3 Задание по выполнению лабораторной работы

Лабораторная работа предполагает выполнение всех заданий. К каждой ситуации необходимо сформировать проблему, требующую решения, и проблему, требующую исследования.

Задание 1. Изготовитель шариковых ручек заинтересован в исследовании эффективности маркетинговых усилий своей фирмы. Компания использует оптовых торговцев для распределения продукции и в дополнение к их усилиям прибегает к персональным продажам и коротким рекламным роликам. Компания планирует использовать в качестве меры оценки эффективности ежегодный объем продаж по территориям. Эти данные и информация о количестве торговых представителей, обслуживающих территорию, всегда готовы для обработки и находятся в регистрационных файлах компании. Другие характеристик, которые изготовитель считает возможным связать с объемом продаж — короткие телеролики и эффективность оптовиков — определить труднее. Чтобы получить

информацию по телевизионной рекламе на территории, исследователи должны проанализировать графики показа роликов и изучить охват территории телевизионным каналом, что позволит определить, каких именно регионов может достичь каждая передача. Эффективность деятельности оптовиков требует установления их рейтингов по ряду критериев и расчета совокупного рейтинга. В нем 4 означает выдающуюся эффективность, 3 – хорошую, 2 – среднего уровня, 1 – слабую. Компания решила анализировать только определенную выборку из общего числа торговых представителей.

Проведите анализ влияния структурных переменных на объем продаж. Сделайте прогноз объемов продаж при следующих значениях структурных переменных: число показов рекламных роликов в месяц 15, число торговых представителей 5 при разной работе оптовиков. В процессе анализа введите фиктивные переменные для оценки влияния оптовиков, согласно таблице 3.1. (Фиктивная переменная – переменная, принимающая одно из двух значений: 1 или 0, и которая используется для обеспечения численного представления атрибутов или характеристик, не являющихся количественными по своей сути).

Таблица 3.1 – Фиктивные переменные для анализа эффективности оптовиков

	x_3	x_4	x_5
Если оптовик плохой	0	0	0
Если оптовик приемлемый	1	0	0
Если оптовик хороший	0	1	0
Если оптовик превосходный	0	0	1

Таблица 3.2 – Данные для анализа

Территория	Продажа (тыс. штук) y	Реклама (число показов в месяц) x_1	Число торговых представителей x_2	Индекс эффективности оптовика x_3
005	260.3	5	3	4
019	286.1	7	5	2
033	279.4	6	3	3
039	410.8	9	4	4
061	438.2	12	6	1
082	315.3	8	3	4
091	565.1	11	7	3
101	570	16	8	2
115	426.1	13	4	3

118	315	7	3	4
133	403.6	10	6	1
149	220.5	4	4	1
162	343.6	9	4	3
164	644.6	17	8	4
178	520.4	19	7	2
187	329.5	9	3	2
189	426	11	6	4
205	343.2	8	3	3
222	450.4	13	5	4
237	421.8	14	5	2
242	245.6	7	4	4
251	503.3	16	6	3
260	375.7	9	5	3
266	265.5	5	3	3
279	620.6	18	6	4
298	450.5	18	5	3
306	270.1	5	3	2
332	368	7	6	2
347	556.1	12	7	1
358	570	13	6	4
362	318.5	8	4	3
370	260.2	6	3	2
291	667	16	8	2
408	618.3	19	8	2
412	525.3	17	7	4
430	332.2	10	4	3
442	393.2	12	5	3
467	283.5	8	3	3
471	376.2	10	5	4
488	481.8	12	5	2

Главный менеджер сети магазинов хотела бы составить прогноз времени, затрачиваемого персоналом на обслуживание одного покупателя. Она решила использовать для анализа следующие данные: количество приобретаемых единиц товара, общая стоимость покупок. Составьте прогноз времени на обслуживание покупателя, приобретающего 14 единиц товара общей стоимостью 70 руб. Определите число работающих касс супермаркета в двух сменах.

Таблица 3.3.1 – Данные для расчета

Показатель	Смена	
	1	2
1. Средняя стоимость покупки (руб)	30	60
2. Количество единиц товара	8	12
3. Число посетителей супермаркета	530	740
4. Продолжительность смены (час)	8	

Таблица 3.3.2 – Данные для анализа

Покупатель	Время обслуживания (мин)	Стоимость покупок (руб)	Количество единиц товара
1	3	36	9
2	1.3	13	5
3	0.5	3	2
4	7.4	81	14
5	5.9	78	13
6	8.4	103	16
7	5	64	12
8	8.1	67	11
9	1.9	25	7
10	6.2	55	11
11	0.7	13	3
12	1.4	21	8
13	9.1	121	21
14	0.9	10	6
15	5.4	60	13
16	3.3	32	11
17	4.5	51	15
18	2.4	28	10

Таксомоторную компанию интересует зависимость между средним пробегом машины в расчете на один галлон топлива и возрастом машины. Были взяты 12 автомашин одной марки, находящейся в достаточно хорошем состоянии, поддерживаемом благодаря регулярному техническому обслуживанию. Поскольку водителями в компании работают и мужчины, и женщины, предполагалось, что какая-то часть изменчивости пробега автомобиля в расчете на один галлон топлива может объясняться разной техникой вождения у мужчин и женщин. При прочих равных условиях у водителя-женщины средний пробег машины на галлон топлива будет больше, чем у мужчины. Данные были получены посредством

случайного выбора 12 машин, на которых водителями работают пять женщин и семь мужчин. Значения среднего пробега на один галлон топлива были рассчитаны на основе сведений о расходе горючего после прохождения машинами расстояния в 300 миль. Данные оформлены в таблице 3.4.

Таблица 3.4 – Результаты опроса

Y (миль на один галлон)	X1 (возраст машины)	X2 (пол: 0 – мужчина, 1 — женщина)
22.3	3	0
22	4	1
23.7	3	1
24.2	2	0
25.5	1	1
21.1	5	0
20.6	4	0
24	1	0
26	1	1
23.1	2	0
24.8	2	1
20.2	5	0

Менеджер по продажам фирмы, занимающейся реализацией запчастей к автомобилям, хотел бы спрогнозировать годовой объем продаж в регионе. Если этот объем продаж можно будет спрогнозировать для региона, то можно будет составить прогноз продаж и для всей компании в целом. Количество пунктов розничной торговли данной компании в регионе и количество автомобилей, зарегистрированных в регионе. На этапе анализа менеджер решил включить в исследование новую независимую переменную — общий доход населения в регионе. Данные об этой переменной приведены в итоговой таблице.

Таблица 3.5 – Данные для анализа продаж фирмы

Регион	Годовой объем продаж (тыс. руб.)	Количество пунктов обслуживания	Количество зарегистрированных автомобилей	Общий доход населения (млн. руб)
1	52.3	2011	24.6	98.5
2	26	2850	22.1	31.1
3	20.2	650	7.9	34.8
4	16	480	12.5	32.7
5	30	1694	9	68.8

6	46.2	2302	11.5	94.7
7	35	2214	20.5	67.6
8	3.5	125	4.1	19.7
9	33.1	1840	8.9	67.9
10	25.2	1233	6.1	61.4
11	38.2	1699	9.5	85.6

Индивидуальное задание

Сформировать задачу, требующую решения и исследования таким образом, чтобы в рамках задачи, требующей исследования, необходимо было бы применение регрессионного анализа как метода анализа информации.

Осуществить сбор данных.

Выполнить анализ данных и сделать выводы.

Подготовить отчет о результатах. Наличие презентации по исследованию приветствуется.

Лабораторное занятие № 3

Использование математических методов многомерного анализа для решения проблемы размещения объекта

Теоретические вопросы методики определения оптимального расположения автостоянок

Основой успешной деятельности организации является ее эффективное размещение. Для определения местоположения применяются различные математические методы: метод гравитации и метод минимума затрат, метод зоны тяготения, розничная гравитационная модель, модель Рейсли.

Методика определения оптимального расположения автостоянок позволяет получить непрерывное распределение потребностей. На первом этапе методики определения оптимального расположения стоянки необходимо получить данные о потребностях потребителей в стоянках с помощью опросов, анкетирования, наблюдений и других методов проведения маркетинговых исследований. Сбор данных о потребностях потребителей предполагает также определение координат потребностей.

На втором этапе данные о потребностях потребителей в k стоянках описываются суммой двумерных нормальных законов распределения [Ошибка! Источник ссылки не найден.], умноженных на коэффициенты N'_j :

$$n'(x,y) = \sum_{j=1}^k N'_j \frac{1}{2\pi\sigma_{xj}\sigma_{yj}\sqrt{1-r_{xyj}^2}} \exp \left\{ -\frac{1}{2(1-r_{xyj}^2)} \left[\frac{(x-m_{xj})^2}{\sigma_{xj}^2} - \frac{2r_{xyj}(x-m_{xj})(y-m_{yj})}{\sigma_{xj}\sigma_{yj}} + \frac{(y-m_{yj})^2}{\sigma_{yj}^2} \right] \right\},$$

где n' — теоретические потребности потребителей в стоянках; x и y — координаты потребностей; N'_j — коэффициенты перехода от относительных единиц к абсолютным; m_{xj} , m_{yj} — математические ожидания; σ_{xj} , σ_{yj} — средние квадратические отклонения; r_{xyj} — коэффициент корреляции.

Введем следующие допущения: $\sigma_{xj} = \sigma_{yj} = \sigma$, $r_{xyj} = 0$. Смысл коэффициента σ заключается в том, что он отражает готовность людей идти пешком от своей автомашины на определенное расстояние. Коэффициент σ — это психологический параметр. Для упрощения решения задачи коэффициент σ принимается равным для всех потребностей в стоянках. Определяется коэффициент σ в процессе аппроксимации экспериментальных данных по методу наименьших квадратов. Коэффициент N'_j определяется как сумма всех фактических потребностей в каждой стоянке. Числа m_{xj} и m_{yj} задают координаты j мест, в которые стремятся попасть потребители, например, крупные торговые точки, культурные центры, различные государственные учреждения и т. п.

Таким образом, зависимость потребностей потребителей в стоянках со всеми введенными упрощениями будет иметь вид:

$$n'(x,y) = \sum_{j=1}^k N'_j \frac{1}{2\pi\sigma^2} \exp \left\{ -\frac{(x - m_{xj})^2 + (y - m_{yj})^2}{2\sigma^2} \right\} \quad (1)$$

На третьем этапе методики определяется оптимальное месторасположение стоянки по максимуму функции (1). Если максимум оказался одинаковым в нескольких координатах, то для выбора одного местоположения необходимо теперь руководствоваться дополнительными факторами (экономическими, экологическими, политическими и др.). Масштаб стоянки можно также первоначально определить из максимального значения функции, а в дальнейшем скорректировать с учетом экономических и других соображений.

На четвертом этапе осуществляется имитация размещения стоянки. Для этого необходимо величины N'_j в формуле (1) заменить на величины N''_j , которые показывают пропорциональное снижение потребностей потребителей в j -ой стоянке из-за появления новой стоянки:

$$N''_j = N'_j - N_n \frac{1/r_j}{\sum_{j=1}^k 1/r_j}, \quad (2)$$

где N_n — масштаб новой стоянки; r_j — расстояние от новой стоянки до максимального значения потребностей в j -ой стоянке:

$r_j = \sqrt{(x_n - m_{xj})^2 + (y_n - m_{yj})^2}$, где x_n и y_n — координаты новой стоянки.

На последнем этапе методики полученная зависимость анализируется и при необходимости изменяется. Можно изменить выбранное месторасположение стоянки, или проимитировать размещение нескольких стоянок.

Зависимость (1) является простейшим случаем рассматриваемой проблемы, потому что она не учитывает различные факторы: специфику рассматриваемого района, карту дорог, рельеф, административные ограничения и т. п. Всякий учет вышеперечисленных факторов повлечет за собой усложнение полученной зависимости.

Определение оптимального расположения автостоянок в системе STATISTICA

Для наглядности рассмотрим сначала одномерный случай решения проблемы размещения автостоянки. В таблице 1 показаны гипотетические данные

потребностей потребителей в двух стоянках. В координатах 3 и 8 наблюдаются максимумы потребностей.

Таблица 1 – Гипотетические данные потребностей в стоянках

х – координата	n – гипотетические потребности в стоянках, чел.
1	100
2	150
3	200
4	150
5	100
6	100
7	150
8	200
9	150
10	100

Предварительно построим график в пакете STATISTICA (возможно использование и другого ПО), чтобы визуально убедиться, что предположение суммы нормальных законов распределения не лишено смысла: **2D Graphs** ⇒ **Bar/Column plots**. В появившемся окне выбрать Graph type: regular, в следующем подокне: select variables: var1, var2. Данные на графике имеют следующий вид (рисунок 1).

Данные аппроксимируем суммой одномерных нормальных законов распределения, умноженных на коэффициенты N'_j :

$$n' = \sum_{j=1}^k N'_j \frac{1}{\sigma\sqrt{2\pi}} \exp \left\{ -\frac{(x - m_j)^2}{2\sigma^2} \right\}$$

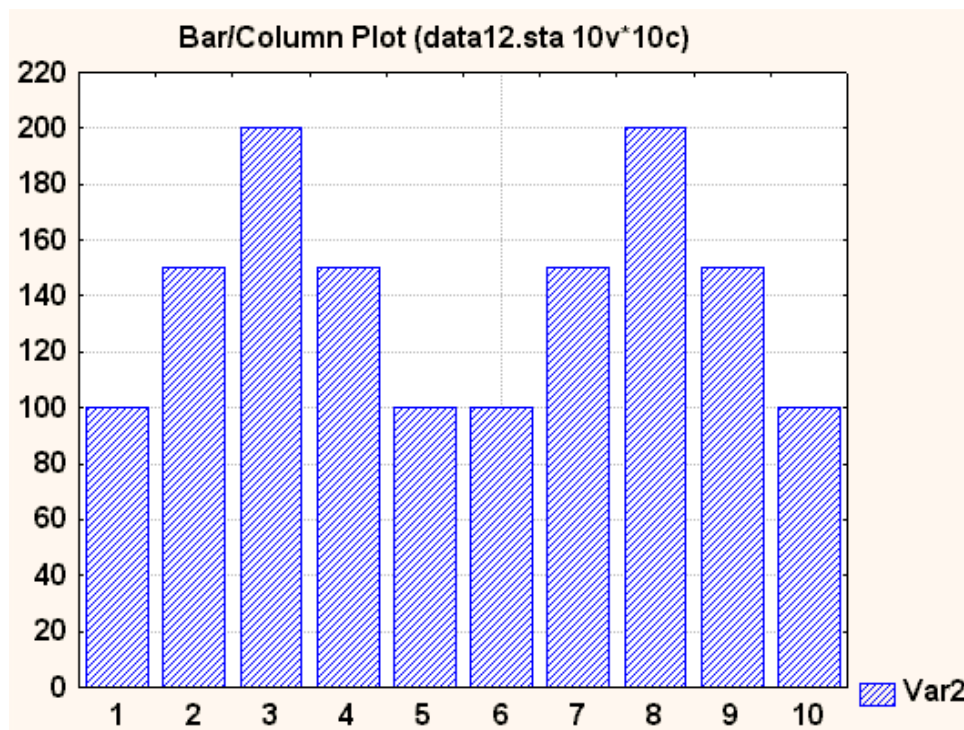


Рисунок 1 – Гипотетические потребности потребителей в стоянках

С помощью аппроксимации мы получим значение параметра σ , который нам неизвестен. Параметры N'_j определяем как сумму всех фактических потребностей в 1 и 2 стоянках: $N'_1 = N'_2 = 700$. В пакете STATISTICA для получения сумм необходимо выделить суммируемые данные, нажать правую кнопку мыши и выбрать **Statistics** $\Rightarrow$ **Columns** $\Rightarrow$ **Sums**. Теперь наша оцениваемая функция выглядит так:

$$n' = 700 \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{\frac{-(x-3)^2}{2\sigma^2}\right\} + 700 \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{\frac{-(x-8)^2}{2\sigma^2}\right\}$$

Для аппроксимации данных используем модуль **Nonlinear Estimation** — Нелинейное оценивание. Выбираем метод обработки **User-specified regression, least squares** — Определенная пользователем регрессия, метод наименьших квадратов. В подокне **Estimated function** задаем с помощью формул функцию, параметры которой мы хотим оценить (рисунок 2).

В подокне **Model Estimation** задается **Estimation method** — метод оценивания, **Start values** — начальные значения для неизвестных параметров, **Maximum number of iterations** — максимальное число итераций, **Convergence criterion** — критерий сходимости (рисунок 3).

В начальных значениях оцениваемого нами параметра (**Start values**) установим 1. Необходимо отметить, что используемые в модуле вычислительные процедуры носят итеративный характер. Они строят оценки последовательно, шаг за шагом, начиная с некоторого начального приближения. И если начальные

значения выбраны плохо, очень плохим может оказаться и окончательный результат.

После того как оценивание завершено, появляется окно **Results** (рисунок 4).

Нажимая последовательно на кнопки в этом окне, мы получим результаты оценивания. Можно пронаблюдать за вычислениями компьютера, нажав на кнопку **Iteration history** во вкладке **Quick**. В таблице будут показаны **Loss Function** — Функция потери на каждой итерации и **b1** — оценка параметра. Процесс оценивания нашей функции сошелся на 6-ой итерации (таблица 2).

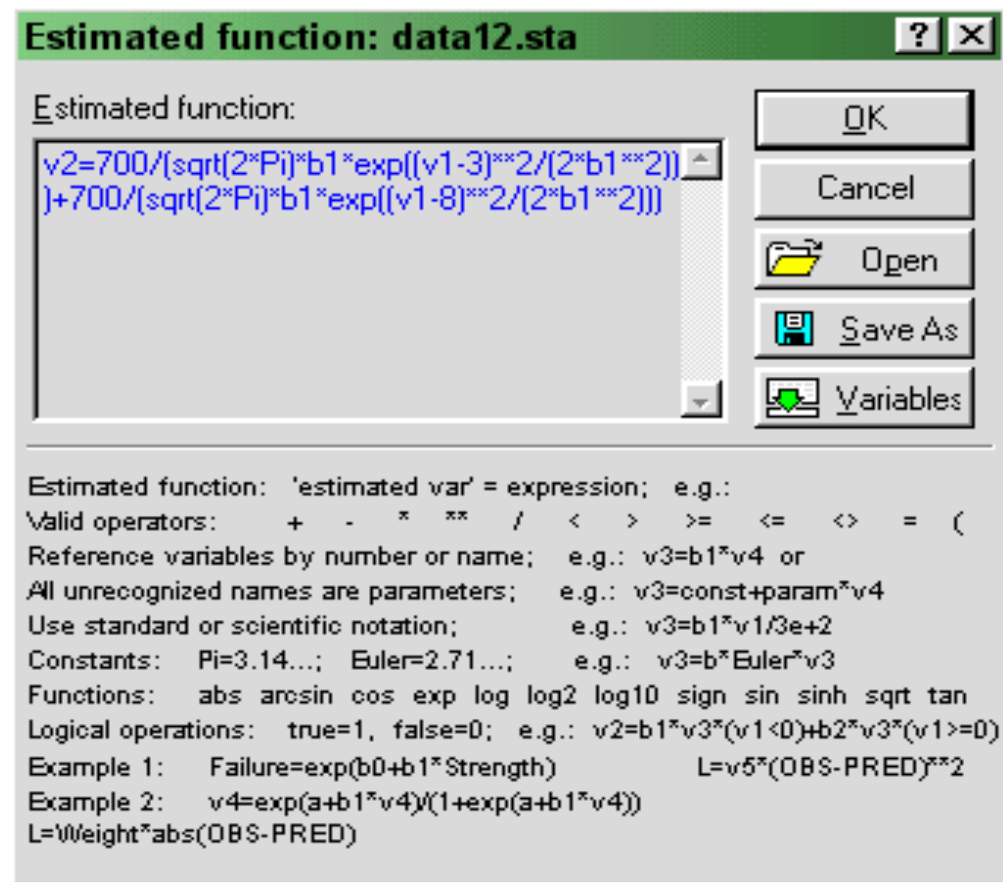


Рисунок 2 – Задание оцениваемой функции

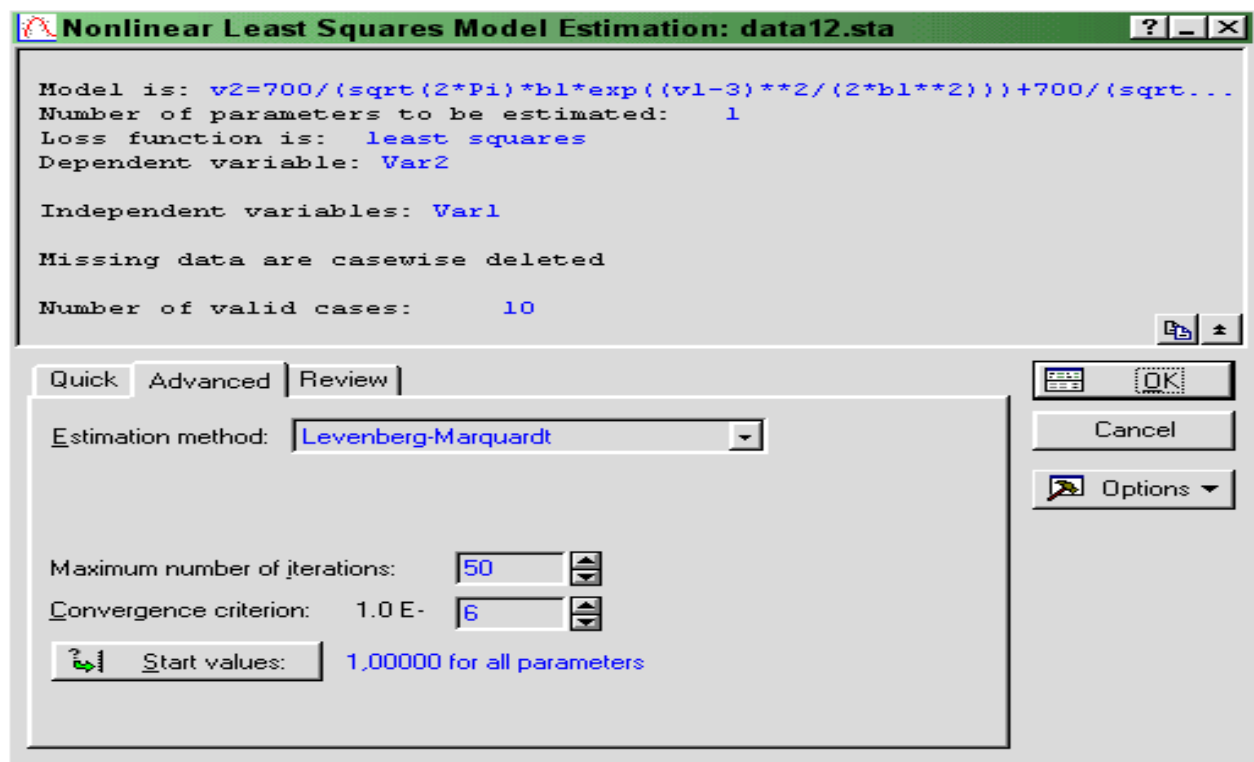


Рисунок 3 – Окно выбора метода и начальных установок

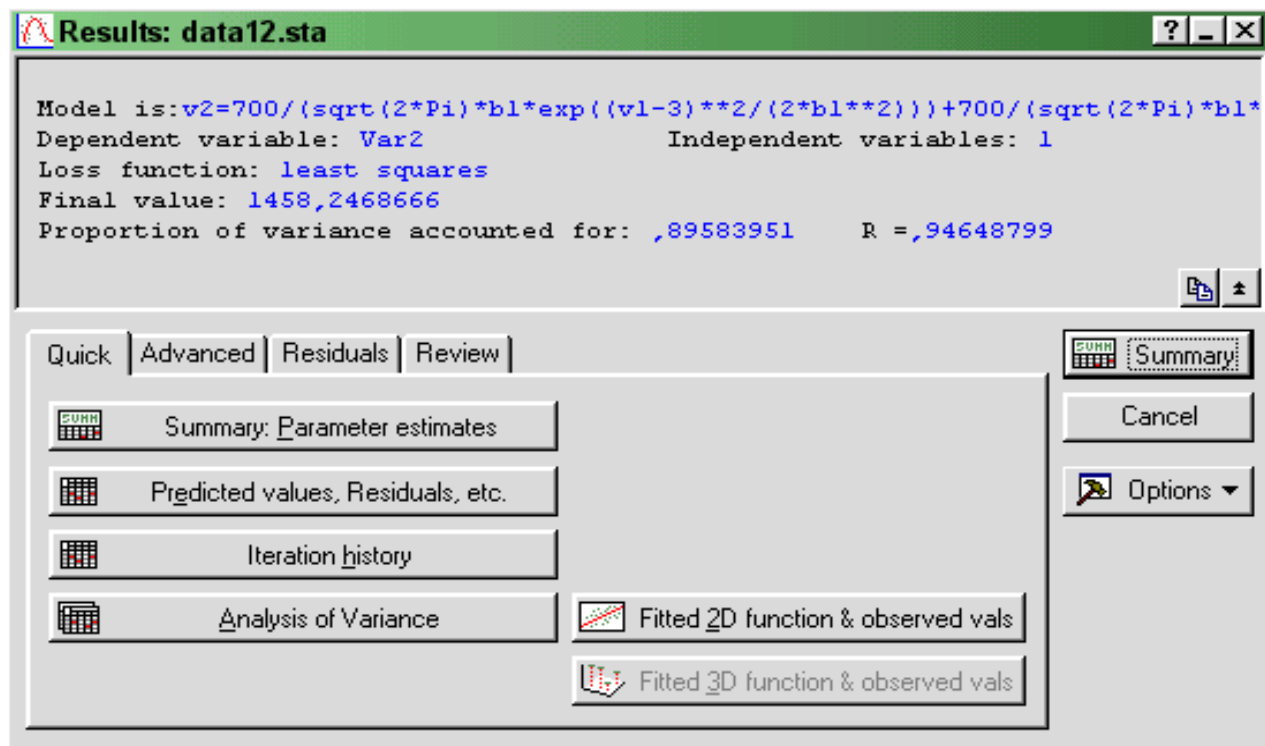


Рисунок 4 – Окно результатов оценивания параметра

Таблица 2 – История оценивания параметра

Iteration	Loss Function	b1
1	169,7052	1,000000
2	53,9032	1,339550
3	38,2432	1,468840
4	38,1871	1,477387
5	38,1870	1,477034
6	38,1870	1,477051

Теперь посмотрим результаты графически, нажав на кнопку **Fitted 2D**. На рисунке будет изображен график подогнанной функции вместе и исходными данными (рисунок 5).

Рассматривая график, визуально можно сказать, что функция достаточно хорошо подошла к данным. Нажав на кнопку **Summary**, мы увидим значение и статистику оцениваемого параметра. Итак, аппроксимация по нашим данным дала значение коэффициента $\sigma = 1,477$. В таблице также показано значение t-критерия Стьюдента для коэффициента (**t-value**). При уровне значимости $\alpha = 5 \%$ принимается гипотеза H_1 о значимости коэффициента ($P < 0.05$).

В таблице **Analysis of Variance** показан дисперсионный анализ, на основе которого при уровне значимости в 5% можно принять гипотезу H_1 об адекватности уравнения регрессии ($P < 0.05$).

Далее посмотрим расчетные значения и остатки, нажав на кнопку **Observed, predicted, residual vals**. Мы увидим таблицу с прогнозируемыми значениями потребностей в стоянках (таблица 3).

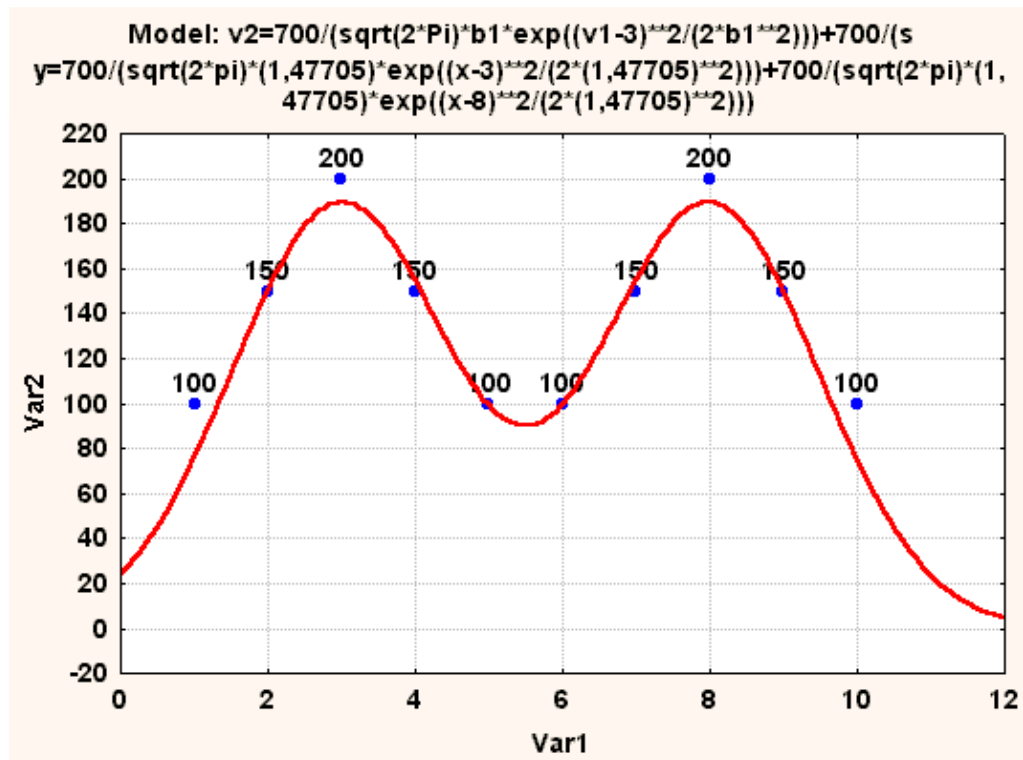


Рисунок 5 – График кривой, подобранной к данным

Таблица 3 – Расчетные данные и остатки по модели

x	Observed	Predicted	Residuals
1	100,0000	75,5960	24,40404
2	150,0000	150,3914	-0,39140
3	200,0000	189,6798	10,32021
4	150,0000	155,1737	-5,17372
5	100,0000	99,6273	0,37268
6	100,0000	99,6273	0,37268
7	150,0000	155,1737	-5,17372
8	200,0000	189,6798	10,32021
9	150,0000	150,3914	-0,39140
10	100,0000	75,5960	24,40404

На основании этой таблицы мы можем принять решение о месторасположении новой стоянки по максимальному значению расчетных данных. Так как у нас два максимума, равных 189,68, то необходимо выбрать один из них на основании дополнительных данных. Например, у нас есть информация, что в координате $x = 3$ создание новой автостоянки проще из-за наличия подходящей ровной площадки. Сымитируем размещение в этой координате стоянки, например, на 100 машиномест. Для этого необходимо изменить оцениваемую функцию на величины N''_j , рассчитываемые по формулам (2):

$$N_1'' = N_1' - N_n \frac{1/r_1}{1/r_1 + 1/r_2} = 700 - 100 \frac{r_2}{r_1 + r_2} = 600;$$

$$N_2'' = N_2' - N_n \frac{1/r_1}{1/r_1 + 1/r_2} = 700 - 100 \frac{r_1}{r_1 + r_2} = 700,$$

где N_n — масштаб стоянки (100 м/м); r_1 — расстояние от новой стоянки до максимального значения потребностей в первой стоянке:

$r_1 = \sqrt{(x_n - m_1)^2} = \sqrt{(3 - 3)^2} = 0$; r_2 — расстояние от новой стоянки до максимального значения потребностей во второй стоянке: $r_2 = \sqrt{(x_n - m_2)^2} = \sqrt{(3 - 8)^2} = 5$.

После учета расчетов функция распределения примет вид:

$$n'' = 600 \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-3)^2}{2\sigma^2}\right\} + 700 \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-8)^2}{2\sigma^2}\right\}$$

Вводим измененную функцию в окне **Estimated function** пакета STATISTICA и вновь просматриваем результаты нелинейной регрессии. По графику видно, как пропорционально снизились потребности в стоянках из-за появления новой стоянки (рисунок 6).

Значение оцениваемого коэффициента σ немного изменился и равен 1,43. Таблица расчетных значений и оценок дает возможность увидеть численное значение нового непрерывного распределения (таблица 4).

Таблица 4 – Расчетные данные и остатки по модели, учитывающей новую стоянку

x	Observed	Predicted	Residuals
1	100,0000	62,8935	37,10647
2	150,0000	131,1681	18,83194
3	200,0000	167,9651	32,03494
4	150,0000	135,0207	14,97932
5	100,0000	84,4526	15,54744
6	100,0000	91,8546	8,14542
7	150,0000	156,3227	-6,32269
8	200,0000	195,8267	4,17329
9	150,0000	153,0204	-3,02045
10	100,0000	73,3754	26,62459

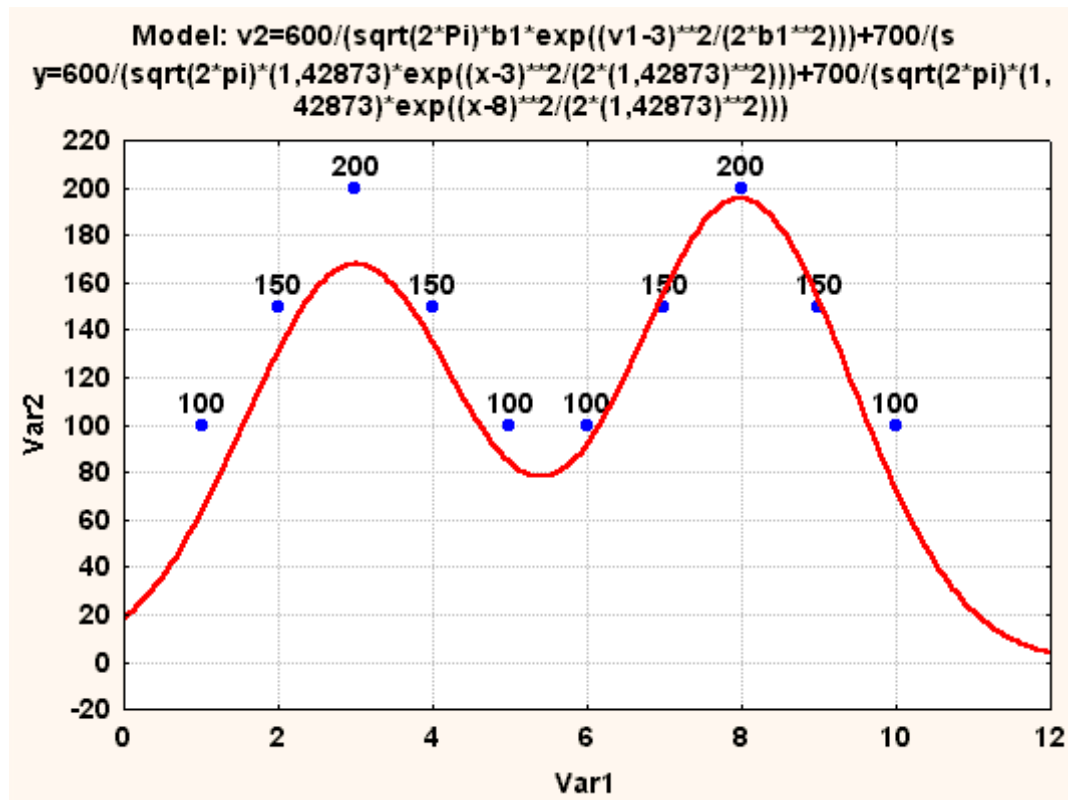


Рисунок 6 – График потребностей с учетом новой стоянки, расположенной в координате 3

В двумерном случае решение задачи нахождения оптимального местоположения при размещении стоянки происходит аналогичным образом.

Задание по выполнению лабораторной работы

1. Ознакомиться с пунктом 1 методических указаний. Выполнить пример определения месторасположения стоянки в одномерном случае из пункта 2 и проанализировать результаты.
2. В соответствии с вариантом задания создать таблицу исходных данных.
3. Выполнить нелинейную регрессию. Выписать результаты (значение оцениваемого параметра, максимумы функции) и построить трехмерный график.
4. Проимитировать размещение новой стоянки. Выписать результаты.
5. Сделать выводы.
6. Представить отчет.
7. Индивидуальное задание. Выполняется отдельно от отчета по работе и защищается с предоставлением отчета, включая исследовательскую и практическую части.

Исследовательская часть. Изучить методы и подходы, используемые для определения месторасположения торговой точки.

Практическая часть. Предоставить решение задачи определения место дополнительной торговой точки двумя методами, одним из которых должен быть метод, рассмотренный в методических материалах, второй – изученный самостоятельно.

Варианты заданий

№	х	1		2		3		4		5		6	
		у	п	у	п	у	п	у	п	у	п	у	п
1.	1	1	100	1	100	1	100	1	100	1	100	1	100
2.	1	2	100	2	100	2	100	2	100	2	100	2	100
3.	1	3	100	3	100	3	100	3	100	3	100	3	100
4.	1	4	100	4	100	4	100	4	100	4	100	4	100
5.	1	5	100	5	100	5	100	5	100	5	100	5	100
6.	2	1	100	1	100	1	100	1	100	1	100	1	100
7.	3	1	100	1	100	1	100	1	100	1	100	1	100
8.	4	1	100	1	100	1	100	1	100	1	100	1	100
9.	5	1	100	1	100	1	100	1	100	1	100	1	100
10.	5	2	100	2	100	2	100	2	100	2	100	2	100
11.	5	3	100	3	100	3	100	3	100	3	100	3	100
12.	5	4	100	4	100	4	100	4	100	4	100	4	100
13.	5	5	100	5	100	5	100	5	100	5	100	5	100
14.	4	5	100	5	100	5	100	5	100	5	100	5	100
15.	3	5	100	5	100	5	100	5	100	5	100	5	100
16.	2	5	100	5	100	5	100	5	100	5	100	5	100
17.	2	2	150	2	150	2	200	2	150	2	150	2	200
18.	2	3	150	3	150	3	200	3	150	3	150	3	200
19.	2	4	150	4	150	4	200	4	150	4	150	4	200
20.	3	2	150	2	150	2	200	2	150	2	150	2	200
21.	3	3	200	3	200	3	250	3	200	3	200	3	250
22.	3	4	150	4	150	4	200	4	150	4	150	4	200
23.	4	2	150	2	150	2	200	2	150	2	150	2	200
24.	4	3	150	3	150	3	200	3	150	3	150	3	200
25.	4	4	150	4	150	4	200	4	150	4	150	4	200
26.	6	1	100	1	100	1	100	6	100	6	100	6	100
27.	6	2	100	2	100	2	100	7	100	7	100	7	100
28.	6	3	100	3	100	3	100	8	100	8	100	8	100
29.	6	4	100	4	100	4	100	9	100	9	100	9	100
30.	6	5	100	5	100	5	100	10	100	10	100	10	100
31.	7	1	100	1	100	1	100	6	100	6	100	6	100
32.	8	1	100	1	100	1	100	6	100	6	100	6	100

33.	9	1	100	1	100	1	100	6	100	6	100	6	100
34.	10	1	100	1	100	1	100	6	100	6	100	6	100
35.	10	2	100	2	100	2	100	7	100	7	100	7	100
36.	10	3	100	3	100	3	100	8	100	8	100	8	100
37.	10	4	100	4	100	4	100	9	100	9	100	9	100
38.	10	5	100	5	100	5	100	10	100	10	100	10	100
39.	9	5	100	5	100	5	100	10	100	10	100	10	100
40.	8	5	100	5	100	5	100	10	100	10	100	10	100
41.	7	5	100	5	100	5	100	10	100	10	100	10	100
42.	7	2	150	2	200	2	150	7	150	7	200	7	150
43.	7	3	150	3	200	3	150	8	150	8	200	8	150
44.	7	4	150	4	200	4	150	9	150	9	200	9	150
45.	8	2	150	2	200	2	150	7	150	7	200	7	150
46.	8	3	200	3	250	3	200	8	200	8	250	8	200
47.	8	4	150	4	200	4	150	9	150	9	200	9	150
48.	9	2	150	2	200	2	150	7	150	7	200	7	150
49.	9	3	150	3	200	3	150	8	150	8	200	8	150
50.	9	4	150	4	200	4	150	9	150	9	200	9	150

Дополнительная информация по объектам

№	Информация
1	Спальный район. Коммуникации отсутствуют, для их создания необходимы инвестиции в 5 млн. руб.
2	Вблизи налаженная инфраструктура, наличие дороги в удовлетворительном состоянии.
3	Место для строительства недалеко от дороги общественного транспорта в строящемся спальном районе
4	Место для парковки находится на территории крупного торгового комплекса. Возможно строительство многоуровневой парковки
5	Стоимость квадратного метра подземной автостоянки составит 50 тыс.рублей. Планируемая площадь до 1000 метров квадратных. Рентабельность вложений составит 20 %. Риск неполучения планируемой рентабельности 30 %.
6	Планируемая стоимость услуг автостоянки 100 руб с 9 до 18, за сутки 200 рублей, цена в месяц 3000 рублей. Вблизи шаговой доступности присутствует наземная парковка. Стоимость услуг конкурентов 120 рублей с 9 до 18 соответственно, загрузка в данный период 80 %. Стоимость услуг в сутки 210 рублей, цена в месяц 3200 рублей.
7	Место для строительства недалеко от дороги общественного транспорта в строящемся спальном районе
8	Планируется строительство в районе новостроек.

9	Стоимость квадратного метра наземного паркинга составит 15 тыс.рублей. Планируемая площадь до 700 метров квадратных. Рентабельность вложений составит 15 %. Риск неполучения планируемой рентабельности 20 %.
10	Планируется строительство в районе новостроек.
11	Спальный район. Коммуникации отсутствуют, для их создания необходимы инвестиции в 6 млн. руб.
12	Вблизи налаженная инфраструктура, наличие дороги в неудовлетворительном состоянии. Требуемы вложения 1 млн. руб.
13	Стоимость квадратного метра подземной автостоянки составит 50 тыс.рублей. Планируемая площадь до 1000 метров квадратных. Рентабельность вложений составит 10 %. Риск неполучения планируемой рентабельности 10 %.
14	Место для строительства недалеко от дороги общественного транспорта в строящемся спальном районе
15	Для создания инфраструктуры дополнительные вложения составят 5 млн рублей. Срок окупаемости дополнительно вложенных средств 10 лет.
16	Планируется строительство в районе новостроек.
17	Спальный район. Коммуникации частично присутствуют, для их реконструкции необходимы инвестиции в 0,9 млн. руб.
18	Для создания инфраструктуры дополнительные вложения составят 5 млн рублей. Срок окупаемости дополнительно вложенных средств 3 года. Рентабельность вложений 15 % годовых. Инфляционные ожидания 13 % годовых.
19	Стоимость квадратного метра подземной автостоянки составит 50 тыс.рублей. Планируемая площадь до 1000 метров квадратных. Рентабельность вложений составит 20 %.
20	Вблизи места строительства отсутствуют организации, предоставляющие аналогичные услуги.
21	Место для строительства недалеко от дороги общественного транспорта в строящемся спальном районе
22	Место для парковки находится на территории крупного торгового комплекса. Возможно строительство многоуровневой парковки
23	Для создания инфраструктуры дополнительные вложения составят 5 млн рублей. Срок окупаемости дополнительно вложенных средств 3 года. Рентабельность вложений 15 % годовых. Инфляционные ожидания 10 % годовых.
24	Вблизи планируемого места строительства паркинга присутствует три наземные парковки со стоимостью услуг средней по району. Парковки конкурентов обеспечивают 80 % потребности района.
25	Спальный район. Коммуникации отсутствуют, для их создания

	необходимы инвестиции в 7 млн. руб.
26	Район, в котором планируется строительство, молодёжный. В районе расположена одна автостоянка наземного типа, покрывающая 50 % потребности района на текущий момент.
27	Стоимость квадратного метра наземного паркинга составит 20 тыс.рублей. Планируемая площадь до 800 метров квадратных. Рентабельность вложений составит 40 %. Риск неполучения планируемой рентабельности 30 %.
28	Вблизи места строительства отсутствуют организации, предоставляющие аналогичные услуги.
29	Вблизи налаженная инфраструктура, наличие дороги в неудовлетворительном состоянии. Требуемы вложения 0,5 млн. руб.
30	Проект строительства автостоянки подземного типа предполагает 200 мест. В виду наличия конкурирующих организаций срок окупаемости 5 лет. Рентабельность вложений 10 % в год.
31	Спальный район. Коммуникации частично присутствуют, для их реконструкции необходимы инвестиции в 1 млн. руб.
32	Проект строительства автостоянки подземного типа предполагает 180 мест. В виду наличия конкурирующих организаций срок окупаемости 5 лет. Рентабельность вложений 8 % в год.
33	Вблизи налаженная инфраструктура, наличие дороги в неудовлетворительном состоянии. Требуемы вложения 0,8 млн. руб.
34	Микрорайон, в котором планируется строительство, молодой. Площадь застройки жилым фондом 70 %. В районе идёт строительство торгового центра, расположена одна автостоянка наземного типа, покрывающая 50 % потребности района на текущий момент.
35	Для создания инфраструктуры дополнительные вложения составят 3 млн рублей. Срок окупаемости дополнительно вложенных средств 2 года. Рентабельность вложений 10 % годовых. Инфляционные ожидания 15 % годовых.
36	Спальный район. Коммуникации частично присутствуют, для их реконструкции необходимы инвестиции в 0,5 млн. руб.
37	Вблизи планируемого места строительства паркинга присутствует три наземные парковки со стоимостью услуг выше средней по району. Парковки конкурентов обеспечивают 90 % потребности района.
38	Планируемая стоимость услуг автостоянки 100 руб с 9 до 18, за сутки 200 рублей, цена в месяц 3000 рублей. Вблизи шаговой доступности присутствует наземная парковка. Стоимость услуг конкурентов 80 рублей с 9 до 18 соответственно, загрузка в данный период 100 %. Стоимость услуг в сутки 150 рублей, цена в месяц 2500 рублей.
39	Планируется строительство в районе новостроек.

40	Место для строительства недалеко от дороги общественного транспорта в строящемся спальном районе
41	Вблизи места строительства отсутствуют организации, предоставляющие аналогичные услуги.
42	Место для парковки находится на территории крупного торгового комплекса. Возможно строительство многоуровневой парковки
43	Для создания инфраструктуры дополнительные вложения составят 5 млн рублей. Срок окупаемости дополнительно вложенных средств 3 года. Рентабельность вложений 15 % годовых. Инфляционные ожидания 15 % годовых.
44	Спальный район. Коммуникации отсутствуют, для их создания необходимы инвестиции в 4 млн. руб.
45	Стоимость квадратного метра наземного паркинга составит 17 тыс.рублей. Планируемая площадь до 500 метров квадратных. Рентабельность вложений составит 30 %. Риск неполучения планируемой рентабельности 10 %.
46	Для создания инфраструктуры дополнительные вложения составят 3 млн рублей. Срок окупаемости дополнительно вложенных средств 5 лет.
47	Вблизи налаженная инфраструктура, наличие дороги в неудовлетворительном состоянии. Требуемы вложения 0,7 млн. руб.
48	Планируется строительство в районе новостроек.
49	Место для парковки находится на территории крупного торгового комплекса. Возможно строительство многоуровневой парковки
50	Спальный район. Коммуникации частично присутствуют, для их реконструкции необходимы инвестиции в 0,8 млн. руб.

Контрольные вопросы

1. Что такое нелинейная регрессия? Какие задачи она позволяет решать?
2. В чем суть метода наименьших квадратов?
3. Какие математические методы используются для решения проблемы размещения фирмы?
4. Какой подход используется при применении методики, описанной в работе? Какие его преимущества и недостатки?

Практическое занятие № 4

Разработка и проведение исследований

Данное лабораторное занятие является заключительным. Выполнение данного задания предполагает проведение мини-исследования.

Необходимо выполнить исследование отрасли. Рекомендуется выполнить поисковое исследование, но исследователь сам в праве выбрать проект исследования.

Результаты работы должны быть представлены в виде презентации (устный отчет) и письменный в виде отчета по исследованию. На рисунке 1 представлены этапы проведения исследований.

В презентации необходимо в обязательном порядке указать гипотезы, которые проверяются. Или, если их изначально не было, представить гипотезы уже принятые по результатам исследования рынка конкретной отрасли.

Масштаб анализа отрасли исследователь определяет сам. Возможно исследование отрасли в рамках курсовой работы.

- ▲ Анализ имеющейся информации:
 - о конкурентах;
 - потребителях;
 - рыночных тенденциях;
 - размере и емкости рынка;
 - критериях сегментации и позиционирования продукции предприятия

- ▲ Ответ на вопрос о целесообразности маркетингового исследования:
 - какую информацию возможно получить из открытых источников?
 - насколько эта информация полна, достоверна и точна?

- ▲ Ответ на вопросы:
 - в чем состоит проблема?
 - каковы требования к качеству и объему информации?
 - насколько возможно проведение исследования?
 - каковы временные рамки исследования?
 - каковы целевые ожидания и бюджет исследования?
- ▲ Выбор специализированной организации для проведения маркетинговых исследований.
- ▲ Написание технического задания на проведение

- ▲ Разработка программы исследования.
- ▲ Определение необходимого уровня репрезентативности и достоверности.
- ▲ Формирование выборки.
- ▲ Выбор инструментария проведения исследования.
- ▲ Разработка анкеты или опросника.
- ▲ Подбор и контроль интервьюеров (контрольные бригады).
- ▲ Обработка данных и построение прогнозов.
- ▲ Подготовка итогового отчета о проведенном исследовании

- ▲ Анализ соответствия результатов исследования техническому заданию.
- ▲ Оценка адекватности и логичности сделанных выводов.
- ▲ Проверка должной тщательности проведенного маркетингового исследования



Рисунок 1 – Этапы проведения исследований

ЛИТЕРАТУРА

1. Анализ и прогнозирование рынка : учебник для вузов / А. Н. Асаул, М. А. Асаул, В. Н. Старинский, Г. Ф. Щербина ; под редакцией А. Н. Асау-ла. — 2-е изд., доп. — Москва : Издательство Юрайт, 2023. — 296 с. — (Высшее образование). — ISBN 978-5-534-15179-4. — Текст : электрон-ный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/520260> (дата обращения: 01.04.2023).
2. Божук, С. Г. Маркетинговые исследования : учебник для вузов / С. Г. Божук. — 2-е изд., испр. и доп. — Москва : Издательство Юрайт, 2023. — 304 с. — (Высшее образование). — ISBN 978-5-534-08764-2. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/510093> (дата обращения: 01.04.2023).
3. Галицкий, Е. Б. Маркетинговые исследования. Теория и практика : учебник для вузов / Е. Б. Галицкий, Е. Г. Галицкая. — 2-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2022. — 570 с. — (Бакалавр. Академический курс). — ISBN 978-5-9916-3225-6. — Текст : электрон-ный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/488325> (дата обращения: 01.04.2023).
4. Карасев, А. П. Маркетинговые исследования и ситуационный анализ : учебник и практикум для вузов / А. П. Карасев. — 2-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2023. — 315 с. — (Высшее обра-зование). — ISBN 978-5-534-05189-6. — Текст : электронный // образо-вательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/511401> (дата обращения: 10.04.2023).