

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ**

**ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ**

**«ДОНСКОЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

**(ДГТУ)**

Кафедра «Организация перевозок и дорожного движения»

**МЕТОДИЧЕСКИЕ УКАЗАНИЯ**

**ДЛЯ ПРАКТИЧЕСКИХ ЗАНЯТИЙ ПО ДИСЦИПЛИНЕ**

**«МАТЕМАТИЧЕСКИЕ МЕТОДЫ В ТРАНСПОРТНЫХ ИССЛЕДОВАНИЯХ»**

Направление подготовки 23.04.01 «Технология транспортных процессов»

Программа: «Транспортная логистика», «Организация перевозок на автомобильном транспорте», «Интеллектуальные транспортные системы»

Ростов-на-Дону

ДГТУ

2024

УДК 656

Составитель: Роговенко Т.Н.

Методические указания для практических занятий по дисциплине «Математические методы в транспортных исследованиях» –

Ростов-на-Дону: Донской гос. техн. ун-т, 2024. – 16 с.

УДК 656

Методические указания предназначены для магистрантов очной и заочной формы обучения по направлению подготовки 23.04.01 «Технология транспортных процессов», программы «Транспортная логистика», «Организация перевозок на автомобильном транспорте», «Интеллектуальные транспортные системы». Приведены примеры и задачи, связанные с применением математических методов в исследованиях транспортных процессов.

Печатается по решению редакционно-издательского совета Донского государственного технического университета

Ответственный за выпуск зав. кафедрой «Организации перевозок и дорожного движения» д-р техн. наук, профессор В.В. Зырянов

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

В печать \_\_\_.\_\_\_.2024 г.

Формат 60×84/16. Объем \_\_\_ усл. п. л.

Тираж \_\_\_ экз. Заказ № \_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Издательский центр ДГТУ

Адрес университета и полиграфического предприятия:

344000, г. Ростов-на-Дону, пл. Гагарина, 1

© Донской государственный

технический университет, 2024

**Введение**

Для исследования транспортных процессов традиционно применяется широкий круг математических методов, и в зависимости от предметной области и конкретной задачи могут быть модифицированы и скомбинированы.

Знания, полученные при изучении дисциплины «Математические методы в транспортных исследованиях», в полной мере могут быть применены при интеллектуальном анализе данных о транспортных потоках, в научно-исследовательской работе магистранта и при написании магистерской диссертации.

Последовательное выполнение практических заданий по дисциплине «Математические методы в транспортных исследованиях» позволит получить необходимые магистранту компетенции и навыки научных исследований.

Для выполнения некоторых заданий следует получать исходные данные у преподавателя. Все примеры должны быть выполнены таким образом, чтобы совпали все результаты, включая графики. Все результаты следует сохранять в файле с именем, состоящим из номера группы, фамилии студента и аббревиатуры дисциплины, например «АМЗИТС11\_Иванов\_ММвТИ» или «AMITS\_Li\_MMTR».

**Сбор экспериментальных данных**

**Пример 1**. Для исследования влияния времени суток на время в пути проводим эксперимент с использованием сервиса определения времени в пути на виртуальной карте. В соответствии с требованиями, предъявляемыми к факторам и отклику, введем такие параметры: *x* – разница между часом пик и временем отъезда, мин; *y* – время в пути, мин. Получить 10 значений *y*.

Выберем время , соответствующее часу пик, в качестве начала отсчета и точки начала и конца пути таким образом, чтобы путь проходил по одной улице. Каждые 5 минут повторяем запрос и заносим прогнозируемое время в пути в таблицу. Все данные подтверждаем снимком экрана.

Таблица 1. Результаты виртуального эксперимента.

|  |  |  |
| --- | --- | --- |
| Номер опыта | *x*, мин | *y*, мин |
| 1 | 5 | 7 |
| 2 | 10 | 7 |
| 3 | 15 | 7 |
| 4 | 20 | 5 |
| 5 | 25 | 7 |
| 6 | 30 | 6 |
| 7 | 35 | 5 |
| 8 | 40 | 5 |
| 9 | 45 | 4 |
| 10 | 50 | 3 |

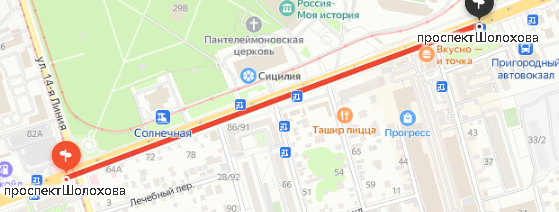


Рис. 1. Снимок экрана для опыта №1

**Задача 1.** Производится доставка грузов от склада до двух пунктов выдачи. Для исследования влияния времени суток и расположения пунктов выдачи на время в пути проводим эксперимент с использованием сервиса определения времени в пути на виртуальной карте. Место расположения склада и пунктов выдачи выбрать самостоятельно так, чтобы от склада до пунктов было больше одного маршрута.

В соответствии с требованиями, предъявляемыми к факторам и отклику, введем такие параметры: *x1* – разница между часом пик и временем отправки груза со склада, мин; *x2* – расстояние от склада до пункта доставки, км; *y* – время в пути, мин.

Интервал варьирования фактора *x1* выбрать самостоятельно. Учитывать все предлагаемые картой маршруты, для каждого *x1* должно быть не менее двух значений *x2* для каждого склада. Результаты наблюдений занести в таблицу 2 и подтвердить снимками экрана. Количество опытов в таблице должно быть не меньше 20.

Таблица 2. Результаты наблюдений к задаче 1.

|  |  |  |  |
| --- | --- | --- | --- |
| Номер опыта | *x1*, мин | *x2*, км | *y*, мин |
| … | … | … | … |

**Задача 2.** Для условий задачи 1 с учетом полученных при ее решении результатов составить план ПФЭ и провести его [1,2]. Результаты занести в таблицу и подтвердить снимками экрана.

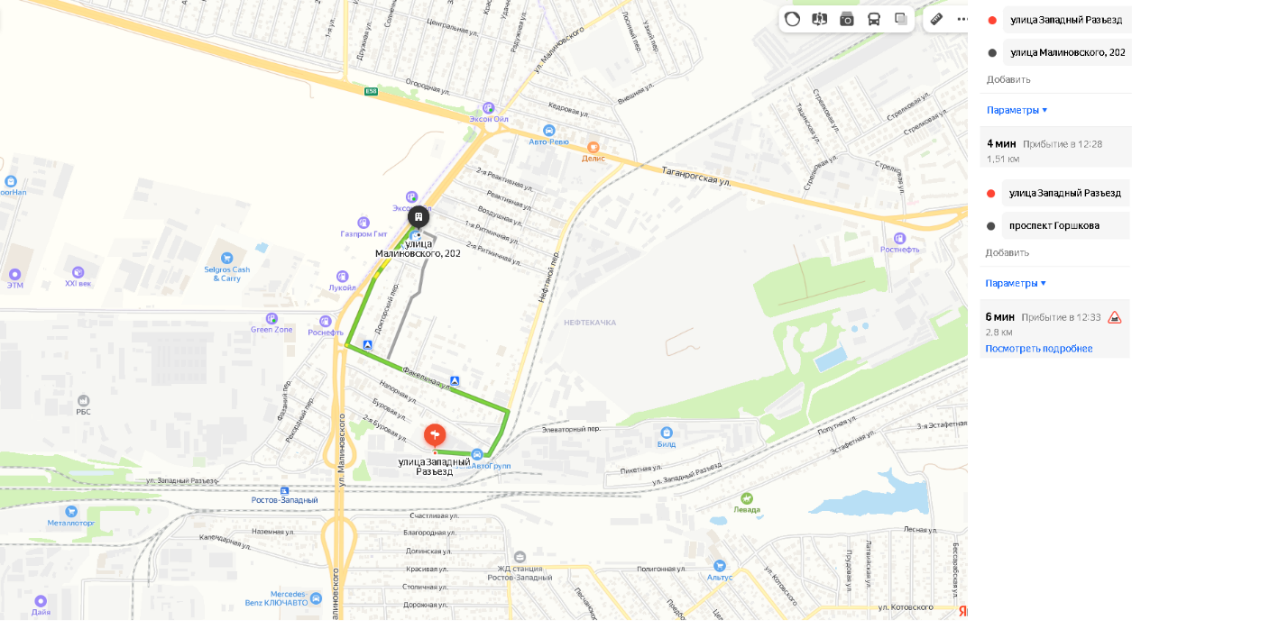


Рис. 2. Пример снимка экрана для задачи 2.

**Задача 3.** Для условий задачи 1 с учетом полученных при ее решении результатов составить план ДФЭ и провести его [1,2]. Результаты занести в таблицу и подтвердить снимками экрана.

**Задача 4.** Повторить эксперимент из задачи 3 еще два раза. Результат занести в таблицу 3.

Таблица 3. Результаты наблюдений к задаче 3.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Номер опыта | *x1*, мин | *x2*, км | *y1*, мин | *y*2, мин | *y3*, мин |
| … | … | … | … |  |  |

**Статистическая обработка экспериментальных данных**

**Пример 2.** Получить характеристики описательной статистики для выборки данных, полученной в примере 1.

Используем надстройку «Анализ данных»-«Описательная статистика». Результаты представлены в таблице 4.

Таблица 4. Описательная статистика

|  |  |
| --- | --- |
| Статистика | Значение |
| Среднее | 5,6 |
| Стандартная ошибка | 0,452155 |
| Медиана | 5,5 |
| Мода | 7 |
| Стандартное отклонение | 1,429841 |
| Дисперсия выборки | 2,044444 |
| Эксцесс | -0,78686 |
| Коэффициент асимметрии | -0,53594 |
| Объем выборки | 4 |

**Задача 5.** Получить характеристики описательной статистики для откликов, полученных в задачах 1 - 4.

**Пример 3.** Применить метод наименьших квадратов для оценки коэффициентов регрессии по экспериментальным данным из примера 1. Построить иллюстрирующий график.

В соответствии с [2] найдем коэффициенты линейного однофакторного регрессионного уравнения: *b1*=-0,0824, *b0*=7,8667. График уравнения и экспериментальные значения представлены на рисунке 3.

**Задача 5.** Получить коэффициенты линейных однофакторных регрессионных моделей *y*(*x1*), *y*(*x2*) для экспериментальных данных задачи 1. Построить иллюстрирующие графики аналогично примеру 3.

**Задача 6.** Получить коэффициенты линейных однофакторных регрессионных уравнений *yi*(*x1*) и *yi* (*x2*), *i*=2,3 для экспериментальных данных из задачи 3. Построить иллюстрирующие графики.

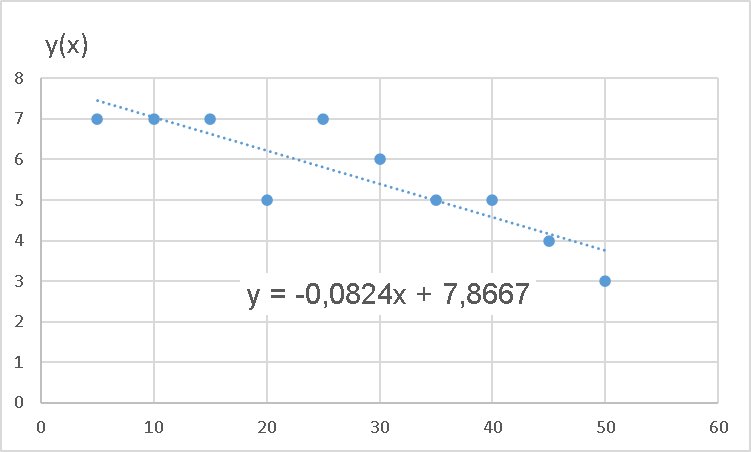
****

Рис. 3. График уравнения линейной регрессии для примера 3

**Пример 4.** Вычислить коэффициент детерминации и коэффициент корреляции для данных из примера 1.

Коэффициент корреляции вычисляем по формуле из [2]: *r*=0,872656. Коэффициент детерминации *R*=*r2* = 0,7615. Используем надстройку «Анализ данных»-«Корреляция». Получаем корреляционную матрицу для двух наборов: *yi* и *y(xi*) (в матрице «Столбец 1» и «Столбец 2», соответственно).

Таблица 5. Матрица корреляции для примера 4.

|  |  |  |
| --- | --- | --- |
|  | Столбец 1 | Столбец 2 |
| Столбец 1 | 1 |  |
| Столбец 2 | 0,872656 | 1 |

**Задача 7.** Вычислить коэффициенты корреляции для экспериментальных данных задач 1-4. Результаты представить в табличном виде аналогично примеру 4.

**Пример 5.** Провести однофакторный дисперсионный анализ времени в пути по данным серии экспериментов из примера 1.

Исходными положениями дисперсионного анализа являются: случайный и независимый характер выборки; нормальное распределение значений изучаемого признака в генеральной совокупности; равенство дисперсий в сравниваемых генеральных совокупностях. В качестве исходных данных рассматриваем три выборки по 10 наблюдений (таблица 7).

Таблица 7. Исходные данные для примера 5.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Номер опыта | *x*, мин | *y1*, мин | *y2*, мин | *y3*, мин |
| 1 | 5 | 7 | 7 | 7 |
| 2 | 10 | 7 | 7 | 7 |
| 3 | 15 | 7 | 7 | 7 |
| 4 | 20 | 5 | 7 | 5 |
| 5 | 25 | 7 | 5 | 5 |
| 6 | 30 | 6 | 6 | 5 |
| 7 | 35 | 5 | 5 | 5 |
| 8 | 40 | 5 | 5 | 5 |
| 9 | 45 | 4 | 4 | 5 |
| 10 | 50 | 3 | 4 | 4 |

Используем надстройку «Анализ данных»-«Однофакторный дисперсионный анализ». Результаты анализа трех выборок при уровне значимости α=0,05представлены на рисунке 4.

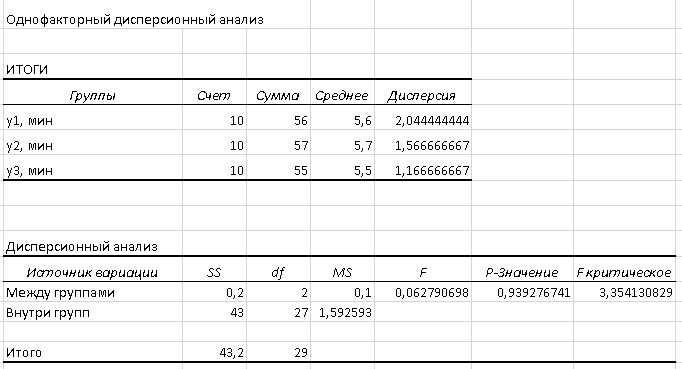


Рис. 4. Результат работы надстройки «однофакторный дисперсионный анализ»

Нулевая гипотеза *H0*состоит в равенстве средних значений выборки. Если *F*>*Fкритическое*, то нулевая гипотеза *H0*не принимается.

Так как *F* =0.06279  меньше  *Fкритическое* =3.3541 , то гипотеза о равенстве средних значений трех выборок не отвергается. Между выборками откликов из таблицы 7 нет статистически значимых различий, следовательно их можно использовать для дальнейших исследований.

**Задача 8**. Провести однофакторный дисперсионный анализ данных о времени в пути, полученных в задачах 1- 4. Результаты внести в таблицу 8.

Таблица 8. Результаты попарного дисперсионного анализа.

|  |  |  |  |
| --- | --- | --- | --- |
| Группы | *F* | *F критическое* | Вывод о гипотезе *H0*  принимается /не принимается |
| *…* | … | … | … |
| *…* |

**Методы статистического оценивания**

**Пример 6.** Провести статистическое оценивание экспериментальных данных из примера 5.

Объединяем в одну выборку значения откликов из примера 5, получим выборку *yi* (*i*=1, 2,…, *n*) объема *n*=30. Описательная статистика для полученной выборки представлена в таблице 9. Для построения гистограммы требуется задать значения границ интервалов, которые в электронных таблицах называются «карманом». Количество интервалов можно определить по формуле Стёрджеса: *k=*1+[3.322*⋅lg(n)*]*=*1+[3.322*⋅lg*(30)]=1+4=5.

Таблица 9. Описательная статистика

|  |  |
| --- | --- |
| *y, мин* | |
|  |  |
| Среднее | 5,6 |
| Стандартная ошибка | 0,222834406 |
| Медиана | 5 |
| Мода | 5 |
| Стандартное отклонение | 1,220514307 |
| Дисперсия выборки | 1,489655172 |
| Эксцесс | -1,124154909 |
| Асимметричность | -0,126800566 |
| Интервал | 4 |
| Минимум | 3 |
| Максимум | 7 |
| Сумма | 168 |
| Счет | 30 |

На рисунке 5 показан результат работы надстройки «Анализ данных»-«Гистограмма» для 5 интервалов.

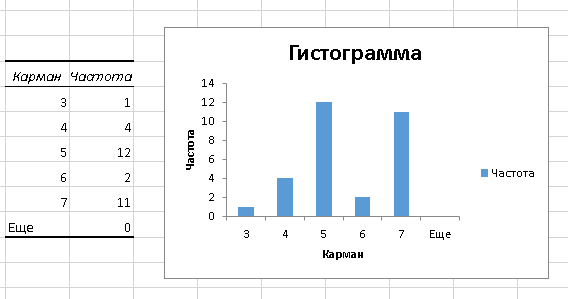


Рис. 5. Результат построения гистограммы для интервалов по формуле Стерджеса

Для построения эмпирической функции распределения (рисунок 6) сначала нужно получить вариационный ряд

*y1≤ y2≤ … ≤ yn*,

затем определить эмпирические вероятности *Fi = i/n* (*i*=1, 2,…, *n*).

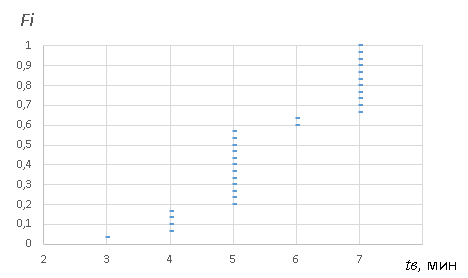
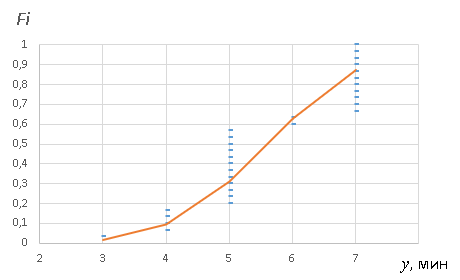


Рис. 6. Эмпирическая функция распределения времени в пути *tв*

Проверим гипотезу о нормальном распределении случайной величины *y* с параметрами *a*=5,6 и *σ*=1,220514307 (рисунок 7).



*F(y)*

Рис. 7. Эмпирическая *Fi* и теоретическая *F(y)* функции распределения времени в пути *y*

Для проверки гипотезы о нормальном распределении применим два критерия: Пирсона и Колмогорова. Следует задать вероятность α=0,05 и определить степень свободы *m*=*k*-1=5-1. Значение критерия *χ2* определяется по формуле:

где *ni* – количество значений выборки, попавших в *i*-ый интервал гистограммы («частота» на рисунке 5), – значение теоретической функции распределения в точке .

Расчетное значение критерия *χ2* сравниваем с критическим значением, полученным с помощью стандартной функции в электронных таблицах для заданных *α* и *m* . Если расчетное значение больше критического, то нулевая гипотеза отклоняется. Для применения критерия Колмогорова следует найти и вычислить критическое значение *.* Если произведение>*Kα*, то нулевая гипотеза отклоняется.

Результаты проверки гипотезы о распределении времени *y* приведены в таблице 10.

Таблица 10. Результаты проверки гипотезы о нормальном распределении

|  |  |
| --- | --- |
| Параметр | Значение |
| Вероятность *α* | 0,05 |
| Критерий Пирсона | |
| Степень свободы *m* | 4 |
| *χ2* расчетный | 14,035 |
| *χ2* критический | 9,488 |
| Вывод: | Нулевая гипотеза о нормальном распределении **отклоняется** |
| Критерий Колмогорова | |
|  | 1,398 |
|  | 1,36 |
| Вывод: | Нулевая гипотеза о нормальном распределении **отклоняется** |

Статистическая обработка данных показала, что объединенная выборка откликов с вероятностью 95% не может быть описана нормальным законом распределения .

**Задача 9.** Провести статистическое оценивание экспериментальных данных из примера 5, использовав следующие параметры: для гистограммы 4 интервала с границами 3, 4, 6, 7; вероятность α=0,1. Сравнить полученные результаты с примером 6.

**Задача 10.** Провести статистическое оценивание времени в пути, в качестве выборки принять набор данных , полученный в задаче 1.

**Задача 11.** Для данных примера 6 проверить гипотезу о логарифмически нормальном распределении.

**Методы корреляционного анализа**

**Пример 8.** Проверить гипотезу о значимости коэффициента корреляции из примера 4.

Проверка значимости коэффициентов корреляции проводится по критерию Стьюдента для уровня значимости α. Нулевая гипотеза *H0*состоит в том, что парный коэффициент корреляции *RXY* не значим.

Критическое значение коэффициента Стьюдента *tкр* определяем с помощью стандартной функции для числа степеней свободы *n*-2=8и α=0,05, *tкр* = 2,306 . Расчетное значение коэффициента Стьюдента:

= 5,0544.

Так как | *t |* > *tкр* , то *H0*не принимается, парный коэффициент корреляции значим.

**Задача 12**. Проверить гипотезу о значимости коэффициентов корреляции из задачи 7. Результаты проверки внести в таблицу 11.

Таблица 11. Результаты проверки гипотезы о значимости *RXY*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *X* | *Y* | *t* | Вывод о гипотезе *H0*  принимается /не принимается | Вывод о значимости *RXY*  значим /не значим |
| … | … | … | *…* | *…* |

**Методы регрессионного анализа**

**Задача 13**. Получить линейные многофакторные регрессионные модели по данным задач 1-4. Использовать надстройку «Анализ данных»-«Регрессия». Все полученные модели проиллюстрировать аналогично примеру 3.

**Задача 14**. Проверить гипотезу о значимости коэффициентов корреляции, полученных в задаче 13.

**Пример 9.** Проверить гипотезу о значимости коэффициентов регрессии из примера 3.

Для проверки значимости коэффициентов регрессии используем распределение Стьюдента. Нулевая гипотеза *H0*состоит в том, что коэффициенты *bj* =0 (*j*=0, 1, 2, …). Критическое значение коэффициента Стьюдента для числа степеней свободы *n-(m-1)=*10-(2-1)=9 (*n* – число наблюдений, *m –* число коэффициентов регрессии)и вероятности α=0,05  *tкр*(0,05,9) = 2,262. Расчетные значения коэффициентов Стьюдента полученные с помощью надстройки «Анализ данных»-«Регрессия»:

|  |
| --- |
| *t(b0* *)* = 15,549 |
| *t(b1 )* = -5,0544 |

Если | *t |* > *tкр*, то *H0*не принимается.

Так как все расчетные значения критерия Стьюдента по модулю больше критического, то нулевая гипотеза не принимается. Коэффициенты регрессии значимы.

**Задача 15.** Проверить гипотезу о значимости коэффициентов регрессии, полученных в задаче 13.

**Пример 10.** Проверить гипотезу об адекватности линейной однофакторной модели по данным примера 5.

Используем надстройку «Анализ данных»-«Регрессия». Проверка адекватность модели проводится по критерию Фишера. Нулевая гипотеза *H0*состоит в том, регрессионная модель не адекватна экспериментальным данным. Если расчетное значение коэффициента Фишера больше критического то, гипотеза отвергается. Критическое значение для вероятности α=0,05 и двух степеней свободы *n-m* и (*m*-1): *Fкр*( 0,05;8;1)=19,37. Так как *F*>*Fкр,* то гипотеза *H0* отвергается, модель адекватна.

**Задача 14.** Проверить гипотезу об адекватности линейных однофакторных регрессионных моделей из задач 6 и 13 для двух уровней α=0,05 и α=0,1.

**Пример 11.** Построить графики логистической функции вида *p*(*x*)=1/(1*+exp*(-(*y*(*x*)-*m*)/*s*), где *y(x)=b0+b1x*, *m* – среднее значение, *s* - среднеквадратическое отклонение для данных примеров 1 и 3.

Результаты вычисления *y(x)* и *p(x)* представлены в таблице 12, графики  *p(y)* и *p(x)* - на рисунке 8.

Таблица 12. Результаты вычисления логистической функции

|  |  |  |  |
| --- | --- | --- | --- |
| *x*, мин | *y*, мин | *y(x)* | *p(x)* |
| 5 | 7 | 7,454545 | 0,785335 |
| 10 | 7 | 7,042424 | 0,732785 |
| 15 | 7 | 6,630303 | 0,672733 |
| 20 | 5 | 6,218182 | 0,606433 |
| 25 | 7 | 5,806061 | 0,535966 |
| 30 | 6 | 5,393939 | 0,464034 |
| 35 | 5 | 4,981818 | 0,393567 |
| 40 | 5 | 4,569697 | 0,327267 |
| 45 | 4 | 4,157576 | 0,267215 |
| 50 | 3 | 3,745455 | 0,214665 |

**Задача 15.** Построить графики логистической функции вида *p*(*x*)*=*1/(1+*exp*(-*y*(*x*))), по данным задачи 6.

**Задача 16.** Построить логит-модель вида *p*(*x*)*=*1/(1+*exp*(-*y*(*x*))), где *y(x)=b0+b1x1+b2x2*, по данным задачи 1.

**Статистическое моделирование**

**Задача 17.** Провести статистическое моделирование для логит модели из задачи 16. Использовать надстройку «Анализ данных»-«Генерация случайных чисел».

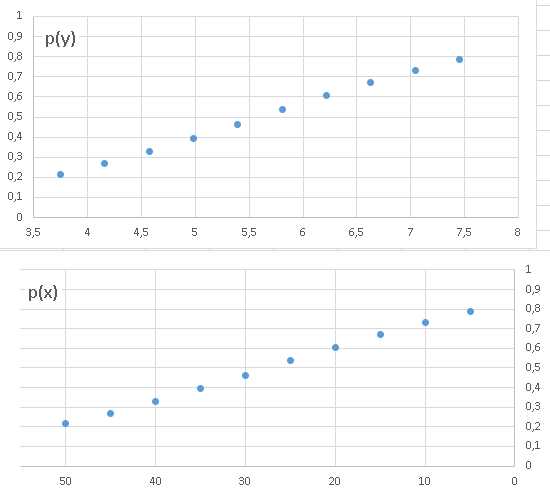


Рис. 8. Графики *p(y)* и *p(x)*

**Список использованных источников**

1. ГОСТ 24026-80. Исследовательские испытания. Планирование эксперимента. Термины и определения.
2. Адлер Ю.П., Маркова Е.В., Грановский Ю.В. Планирование эксперимента при поиске оптимальных условий. - М.: Наука, 1976 – 280 с.
3. Основы теории надежности и планирования эксперимента: учебное пособие / Т. Н. Роговенко, И. А. Серебряная, И. В. Топилин; М-во образования и науки Российской Федерации, Федеральное агентство по образованию, Гос. образовательное учреждение высш. проф. образования Ростовский гос. строит. ун-т. - Ростов-на-Дону : РГСУ, 2006. - 176 с.
4. Организация и математическое планирование эксперимента: учебное пособие / Ю. В. Юдин, М. В. Майсурадзе, Ф. В. Водолазский. - Екатеринбург : Изд-во Урал. ун-та, 2018 - 124 с.
5. Дрейпер Н., Смит Г. Прикладной регрессионный анализ. Книга 1. В 2-х кн. М.: Финансы и статистика, 1986. — 366 с.
6. Численное статистическое моделирование. Методы Монте-Карло: учебное пособие для студ. вузов/ Г. А. Михайлов, А. В. Войтишек. -Москва: «Академия», 2006 - 368 с.