**Мухина С.Н.**

**КОМПЬЮТЕРНАЯ МАТЕМАТИКА**

**ДЛЯ ЭКОНОМИСТОВ НА БАЗЕ**

**MATHCAD**

**Учебное пособие**

***Вероятностные схемы***

***Основные законы распределения***

***случайных величин***

***Вариационные ряды***

***Статистические гипотезы***

***Регрессионный и корреляционный анализ***

***Системы массового обслуживания***

**Калининград**

**Издательство БГАРФ**

**2014**

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО РЫБОЛОВСТВУ

Федеральное государственное бюджетное образовательное учреждение

высшего профессионального образования

«Калининградский государственный технический университет»

Балтийская государственная академия рыбопромыслового флота

С.Н. Мухина

**КОМПЬЮТЕРНАЯ МАТЕМАТИКА**

**ДЛЯ ЭКОНОМИСТОВ НА БАЗЕ**

**MATHCAD**

Учебное пособие

Калининград

Издательство БГАРФ

2014

**УДК. 517 (075)**

**М. 13**

**Мухина С.Н.** Компьютерная математика для экономистов на базе MathCAD: Учебное пособие / ***С.Н. Мухина***. − Калининград: Изд-во БГАРФ, 2014. − с.

Учебное пособие охватывает следующие разделы: теорию вероятностей, математическую статистику, корреляционный и регрессионный анализ, теорию массового обслуживания. Содержание теоретического материала соответствует государственному образовательному стандарту по направлению 080100 «Экономика». Структура пособия ориентирована на использование пакетов компьютерной математики. Задачи снабжены подробными решениями и демонстрационными примерами.

Печатается по решению редакционно-издательского совета Балтийской государственной академии рыбопромыслового флота.

|  |  |
| --- | --- |
| **Рецензенты:** | ***Авдеева Н.Н.*,**кандидат педагогических наук, доцент кафедры высшей математики БГАРФ;  ***Лурье И.Н.*,** доктор педагогических наук, профессор Российскогоуниверситета коммерции |

© «БГАРФ» ФГБОУ ВПО «КГТУ» (БГАРФ), 2014

**ОГЛАВЛЕНИЕ**

**Введение**

**Глава 1. Теория вероятностей**

**1.1. Вероятностные схемы**

1.1.1. Непосредственный подсчет вероятностей с применением формул комбинаторики

1.1.2. Статистический и геометрический подходы к определению вероятности

1.1.3. Композиция независимых испытаний при различных вероятностях успеха

1.1.4. Композиция независимых испытаний при одинаковых вероятностях успеха. Схема Бернулли

*Компьютерная лабораторная работа № 1«Вероятностные схемы»*

**1.2 Случайные величины (дискретные и непрерывные). Основные законы распределения случайных величин**

1.2.1. Случайные величины. Способы задания, числовые характеристики

1.2.2. Библиотека стандартных распределений

1.2.3. Примеры построения законов распределения НСВ

1.2.4. Примеры построения законов распределения ДСВ

*Компьютерная лабораторная работа № 2 «Основные законы распределения НСВ»*

*Компьютерная лабораторная работа № 3 «Основные законы распределения ДСВ»*

**Глава 2. Математическая статистика**

**2.1. Математические основы выборочного метода**

2.1.1. Выборка и ее представление

2.1.2. Числовые характеристики вариационного ряда

2.1.3. Точечное и интервальное оценивание параметров распределения

*Компьютерная лабораторная работа № 4 «Вариационные ряды, их числовые характеристики и графическое изображение»*

**2.2. Статистические гипотезы**

2.2.1. Статистическая гипотеза и общая схема ее проверки

2.2.2. Построение теоретического закона распределения по опытным данным. Проверка гипотез о законе распределения

*Компьютерная лабораторная работа № 5 «Проверка гипотезы о виде закона распределения СВ по критерию Пирсона»*

2.2.3. Гипотезы о значениях и сравнении числовых характеристик и однородности выборок

*Компьютерная лабораторная работа № 6 «Статистические гипотезы»*

**2.3. Основы регрессионного и корреляционного анализа**

2.3.1.Функциональная, статистическая и корреляционная зависимость

2.3.2. Линейная регрессия

2.3.3. Нелинейная однофакторная регрессия

*Компьютерная лабораторная работа № 7 «Корреляционно-регрессионный анализ»*

**Глава 3. Системы массового обслуживания (СМО)**

**3.1. Структура и классификация систем массового обслуживания**

**3.2. Марковский случайный процесс в СМО. Уравнения Колмогорова**

**3.3. Расчет показателей эффективности СМО с отказами**

3.3.1. Одноканальные СМО с отказами

3.3.2. Многоканальные СМО с отказами

**3.4. Расчет показателей эффективности СМО с ожиданием (с очередью)**

3.4.1. Одноканальные СМО с неограниченной очередью

3.4.2. Многоканальные СМО с неограниченной очередью

3.4.3. СМО (одноканальные и многоканальные) с ограниченной очередью

*Компьютерная лабораторная работа № 8 «Расчет показателей эффективности систем массового обслуживания»*

**Список использованной литературы**

**Введение**

В деятельности специалиста экономического профиля возникают такие ситуации как социологические опросы, кредиты, страховые полисы, разнообразные банковские начисления. Эти ситуации требуют стратегического мышления и дальнейшего прогнозирования, в основе которых лежат вероятностно-статистические разделы математики.

Очевидно, что проводить аналитические исследования данных ситуаций с учетом дефицита времени и большого объема информации наиболее оптимально, используя средства пакетов компьютерной математики. Поэтому обучение навыкам использования пакетов специализированных компьютерных программ можно отнести к одному из необходимых подходов в современном освоении экономической науки.

Для простых математических моделей (например, линейные модели в математической статистике) можно применять табличный процессор Excel. Но электронные таблицы предоставляют ограниченные возможности для анализа модели. Поэтому в преподавании разделов теории вероятностей, а особенно разделов математической статистики, более оправдано применение пакета компьютерной математики MathCAD.

MathCAD выгодно отличается от других пакетов возможностью свободно компоновать рабочий лист и относительной легкостью изучения. Так же, как с карандашом в руке решается задача на листе бумаги, в принципе, можно оформить и соответствующий MathCAD-документ. Если некоторое время не возникает необходимости работать с MathCAD, то впоследствии навыки пользования пакетом легко восстанавливаются. В других же пакетах компьютерной математики используется очень сложный синтаксис, который быстро забывается, если не работать с этим пакетом постоянно. Кроме того, MathCAD - это универсальная, а не специализированная математическая среда.

В данном пособии представлено решение основных практически значимых задач теории вероятностей и математической статистики с помощью компьютерной математики.Пособие состоит из трех глав: теория вероятностей, математическая статистика и теория систем массового обслуживания. Разработан компьютерный практикум, задания которого необходимо выполнить для приобретения практических навыков решения вероятностно-статистических задач в среде MathCAD**.**

**Глава 1**

**Теория вероятностей**

*Теория вероятностей есть*

*в сущности, ни что иное, как здравый*

*смысл, сведенный к исчислению*.

Лаплас (1814)

*Ни акционерные общества, ни банки, ни биржи*

*не нуждались в теории вероятностей.*

*Спрос на нее появился у перечисленных*

*учреждений лишь в XIX в.,*

*когда методы открытого грабежа сменились*

*методами научного выигрыша.*

Мрочек (1934)

*Сочетая строгость научных доказательств*

*с неопределенностью случая и примиряя,*

*казалось бы противоположные вещи,*

*и, извлекая ее имя из того и другого,*

*можно по праву присвоить ей ошеломляющее*

*название математика случая.*

Паскаль (1654)

**1.1. Вероятностные схемы**

*Вероятность есть степень достоверности*

*и отличается от нее как часть от целого.*

[Якоб Бернулли](javascript:if(confirm('http://www-groups.dcs.st-and.ac.uk/~history/Mathematicians/Bernoulli_Jacob.html%20%20\n\nThis%20file%20was%20not%20retrieved%20by%20Teleport%20Pro,%20because%20it%20is%20addressed%20on%20a%20domain%20or%20path%20outside%20the%20boundaries%20set%20for%20its%20Starting%20Address.%20%20\n\nDo%20you%20want%20to%20open%20it%20from%20the%20server?'))window.location='http://www-groups.dcs.st-and.ac.uk/~history/Mathematicians/Bernoulli_Jacob.html') (1713)

**1.1.1. Непосредственное вычисление вероятностей с применением формул комбинаторики**

Существует несколько подходов к определению вероятности события (классический, геометрический, статистический).

В условиях классического эксперимента с равновозможными исходами *вероятность события А равна отношению числа случаев m, благоприятствующих ему, к общему числу случаев n,* т.е.

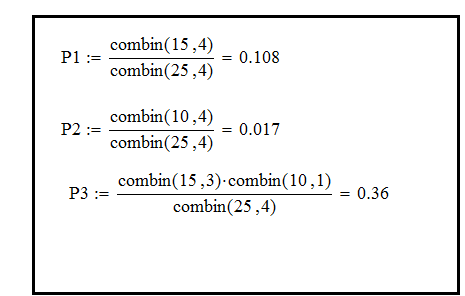
Для решения задач с использованием классического определения вероятности часто используют формулы комбинаторики.

|  |  |  |
| --- | --- | --- |
| **комбинации** | **формула для вычисления** | **встроенная функция**  **MathCAD** |
| сочетания |  | combin(n,m) |
| размещения |  | permut(n,m) |
| перестановки |  | permut(n,n) |

**Пример 1.1.** Среди 25 студентов, из которых 15 девушек, разыгрываются 4 билета, причем каждый может выиграть только один билет. Какова вероятность того, что среди обладателей билета окажутся: четыре девушки; четыре юноши; три девушки и один юноша.

**Решение**

При вычислении числа сочетаний используем встроенную функцию **combin(n,m)**.



*Рис.1.1. Решение примера 1.1 в MathCAD*

Классическая формула для вычисления вероятности события обладает недостатками, ограничивающими ее применение.

1. Какие это недостатки?

2. Как они преодолеваются?

**Задачи для аудиторного решения**

1. В группе 100 студентов, из которых 20 – отличники. Внешней комиссией для последующего тестирования отбирается 10 студентов. Какова вероятность того, что среди отобранных студентов 5 отличников?

2. Кодовый замок имеет 10 кнопок, соответствующих цифрам от 0 до 9. Код замка – двузначный, образованный различными числами. Какова вероятность набора правильного кода при выборе кнопок наугад?

**1.1.2. Статистический и геометрический подходы**

*Статистической вероятностью события А называется число, около которого колеблется относительная частота этого события, приближаясь к нему по мере увеличения числа испытаний.*

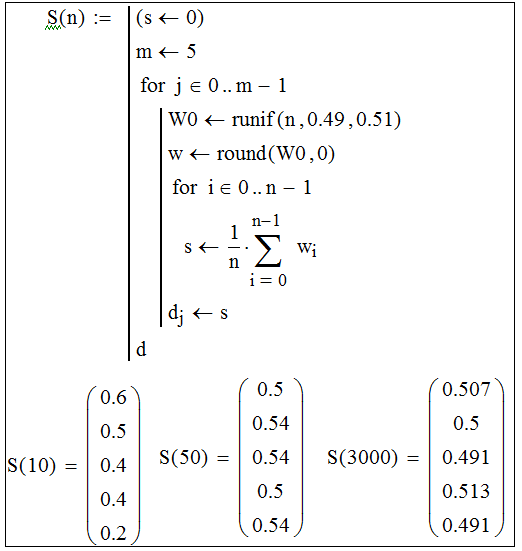
Проиллюстрируем процесс стабилизации относительных частот при увеличении числа испытаний на примере идеализированного эксперимента.

**Пример 1.2.** В эксперименте осуществляется бросание монеты *п* раз; эксперимент повторяется 5 раз.

**Решение**

Процедура имитации результатов бросания и результаты при 10, 50 (500) и 3000 бросаний приведены на рисунке 1.2. Использован генератор независимых случайных равномерно распределенных чисел (встроенная функция **runif**) в интервале (0,49; 0,51). Данные округляются с применением встроенной функции **round.**

Относительные частоты для полученных значений *п* объединим в в одну матрицу *S* (встроенная функция **augment**). Значения элементов матрицы *S* выводим на график (рисунок 1.3).

****

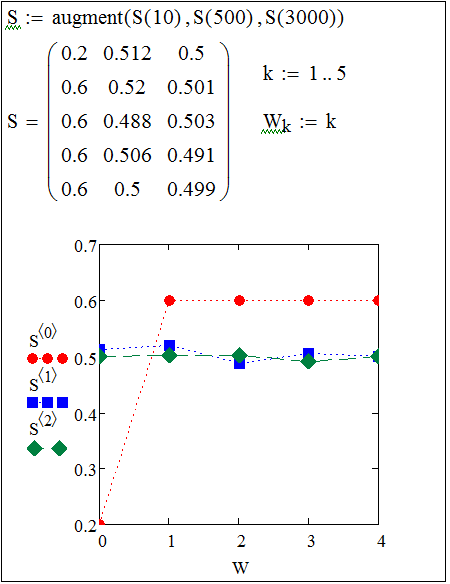
*Рис. 1.2. Имитация эксперимента*

Значения первого столбца матрицы *S* (*п* = 10) сильно отличаются от теоретического значения 0,5. Колебания значений второго столбца матрицы *S* (*п* = 50) вокруг 0,5 происходит в диапазоне 0,04. В третьем случае диапазон отклонений снижается до уровня 0,009 для выбранного числа повторений.

***Замечания:***

*- относительная частота события может служить приближенным значением (оценкой) вероятности данного события;*

*- если эксперимент не может быть отнесен к классическому с равновозможными исходами, то статистический подход служит единственной основой приближенного вычисления вероятности.*

****

*Рис.1.3. Графическая иллюстрация*

При геометрическом подходе рассматривается эксперимент, в котором точка *М* наугад бросается в определенную область *Q* (отрезок, часть плоскости, часть пространства). Характер процесса (наугад, случайно) определяет равновозможность попадания брошенной точки в любую точку конечной области .

*Геометрической вероятностью события А называется отношение меры области q ко всей мере области Q*

**1.1.3. Композиция независимых испытаний при различных вероятностях успеха**

Рассмотрим сложное испытание *В*, образованное повторением, например, трех испытаний *А1*, *А2*, *А3* с различными вероятностями успеха в каждом из них.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **В3,0** | **В3,1** | | | **В3,2** | | | **В3,3** |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |

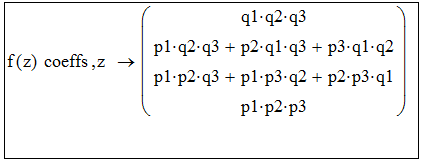
Рассмотрим произведение биномов



Необходимо перемножить биномы и привести подобные. Эти операции можно провести с помощью операторов символьной палитры.

|  |  |
| --- | --- |
| **Операторы** | **Описание** |
| **colleсt** | Разлагает функцию f(z) по степеням z в порядке убывания степени z |
| **expand** | Выводит все члены разложения без приведения подобных в порядке убывания степени z |
| **coeffs** | Выводит вектор коэффициентов разложения в порядке возрастания степени z |

Возможности операторов проиллюстрированы на примере оператора **coeffs** (рисунок 1.4).



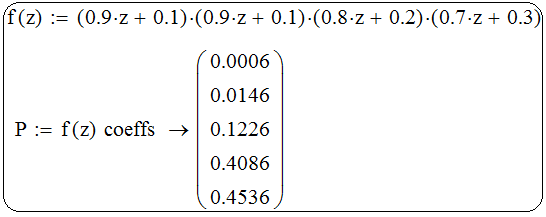
*Рис. 1.4. Символьное преобразование*

*Сравнивая коэффициенты при степенях z с выражениями, приведенными выше для вероятностей , убеждаемся в их полном совпадении. Вывод верен для произвольного числа п испытаний в серии.*

***Замечание.*** *Числовые значения вероятностей можно ввести непосредственно в биномы функции f(z) или задать предварительно.*

**Пример 1.3.** Вероятность того, что студент сдаст первый экзамен, равна 0,9; второй – 0,9; третий – 0,8; четвертый – 0,7. Найти вероятности того, что: студент не сдаст ни один экзамен (событие А1); сдаст только один экзамен (событие А2); сдаст два экзамена (событие А3); сдаст три экзамена (событие А4); сдаст все четыре экзамена (событие А5).

**Решение.** Найдем вероятности перечисленных событий способом сравнения с коэффициентами многочлена 4*-*ой степени (4– число событий **по условию задачи**). Решение отражено на рисунке 1.5.



*Рис. 1.5. Символьные преобразования (п=4)*

P(A1) = 0,0006 - вероятность, что студент не сдаст ни один экзамен;

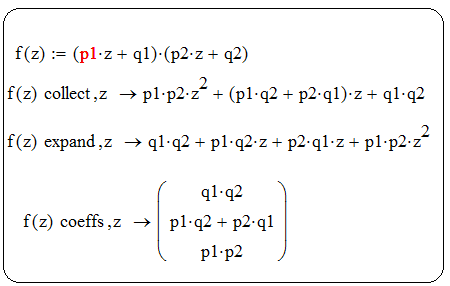
P(A2) = 0,0146 - сдаст только один экзамен;

P(A3) = 0,1226 - сдаст два экзамена;

P(A4) = 0,4086 - сдаст три экзамена;

P(A5) = 0,4536 - сдаст все четыре экзамена.

**Пример 1.4.** Возможности перечисленных операторов проиллюстрируем при .



*Рис.1.6.* *Символьные преобразования (п=2)*

**Задачи для аудиторного решения**

1. В автомобиле обнаружено 4 вида неисправностей, которые проявляются с вероятностями 0,1; 0,2; 0,3; 0,4 соответственно. Осуществлено четыре независимых испытания. Определить вероятности обнаружения одной, двух, трех, четырех неисправностей и вероятность не обнаружить ни одну из них. Найти вероятность обнаружения не менее двух неисправностей.

2. Раскройте скобки с помощью операторов символьной палитры.

**1.1.4. Композиция независимых испытаний при одинаковых вероятностях успеха. Схема Бернулли**

Рассмотрим сложное испытание *В*, образованное повторением, например, трех однотипных испытаний с двумя исходами (успех, неудача).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **В3,0** | **В3,1** | | | **В3,2** | | | **В3,3** |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |

1. Объясните следующую запись:

.

2. Вычислите, используя встроенную функцию **combin**, Сравните результаты с коэффициентами многочлена третьей степени.

3. Формула - бином Ньютона. Какое получится тождество при ?

4. Пусть – вероятность, что событие появится *k* раз в серии из *п* независимых испытаний с вероятностью в каждом испытании, равной *р* и не появится с вероятностью . Очевидно, что

Проанализируйте выражение при

5. Запишите бином Ньютона при . Сравните с пунктом 4. Сделайте вывод.

Вероятности носят *название биномиальных вероятностей*.

6**.**  - формула Бернулли. Позволяет получить вероятность появления события ровно *k* раз.

***Следствия формулы Бернулли***

1. Вероятность того, что событие *А* произойдет не более, чем *k* раз в *п* испытаниях (кумулятивная или накопленная вероятность)

2. Вероятность того, что событие *А* произойдет более, чем *k* раз в *п* испытаниях

3. Вероятность того, что событие *А* произойдет хотя бы один раз

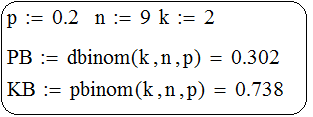
.

В MathCAD имеются встроенные функции для решения данного класса задач.

|  |  |
| --- | --- |
| **формула** | **встроенная функция**  **MathCAD** |
| Бернулли | dbinom(k,n,p) |
| Кумулятивная вероятность по формуле Бернулли | pbinom(k,n,p) |

**Пример 1.5.** В среднем 20% пакетов акций на аукционах продаются по первоначально заявленной цене. Найти вероятность того, что из 9 пакетов акций в результате торгов по первоначально заявленной цене будет продано 2 пакета; не более 2 пакетов.

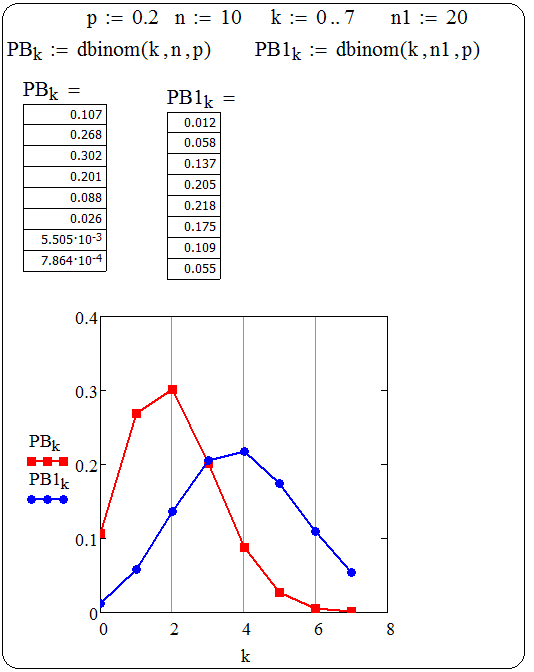
**Решение.** На рис.1.7 представлено решение задачи в среде MathCAD. Вводим исходные данные. Через РВ обозначена вероятность , КВ – кумулятивная вероятность .

****

*Рис.1.7.* *Решение примера 5 в MathCAD*

**Пример 1.6.** Провести исследование влияния числа испытаний на вероятности появления ряда (0, 1, 2, …, k) успехов. Сделать вывод.

**Решение.** Из приведенного на рис. 1.8 графика видно, что вероятности (обозначенные через ) довольно быстро уменьшаются до нуля. С увеличением числа испытаний вероятности медленнее стремятся к своим установившимся значениям, наблюдается стремление к одинаковому максимуму в нескольких точках (полимодальность). Кривая приобретает максимум при k=2, кривая при k=4. Этот факт отражает закономерность, согласно которой максимум биномиальных вероятностей достигается при *k=np*.

****

*Рис.1.8. Изменение вероятностей в схеме Бернулли*

**Асимптотическое приближение при повторении испытаний**

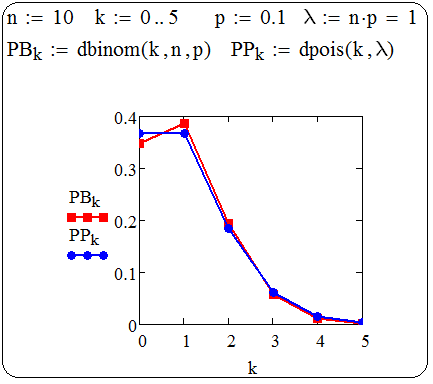
Практическое использование формулы Бернулли сопряжено с трудностями уже при Значительного упрощения алгоритмов вычисления удается достичь, как было показано выше, применением компьютерной математики. Но и этот путь не может в полной мере решить проблему операций над очень большими числами, возникающую при анализе десятков и сотен испытаний.

Ниже рассмотрены асимптотические приближения в схеме Бернулли: формула Пуассона, локальная теорема Муавра-Лапласа, интегральная теорема Лапласа.

|  |  |  |
| --- | --- | --- |
| **формула** | **условие для применения** | **встроенная функция**  **MathCAD** |
| Пуассона |  | dpois(k,λ) |
| Кумулятивная вероятность по формуле Пуассона |  | ppois(k,λ) |
| Локальная формула Муавра-Лапласа  - локальная функция Лапласа (значения могут быть получены из специальных таблиц) |  | dnorm(x,0,1)  значение локальной функции Лапласа в точке *х* |
| Интегральная формула Лапласа  - функция Лапласа (значения берутся из специальных таблиц) |  | pnorm(x,0,1)-0.5 значение интегральной функции Лапласа в точке *х* |

**Пример 1.7.** Сопоставить результаты вычисления по формулам Бернулли и Пуассона для одинаковых значений *k, n, p*.

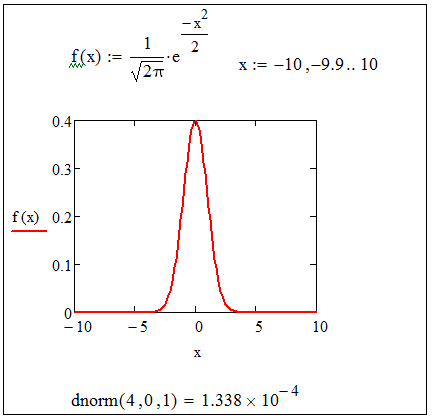
**Решение.** На рисунке 1.9 проведено графическое сопоставление результатов для k=0, 1, 2, 3, 4, 5 при n=10, р=0,1. При увеличении *k* и том же значении *n* значения вероятностей, полученные по формулам Бернулли и Пуассона, достаточно быстро приближаются друг к другу.

****

*Рис. 1.9. Пример сопоставления распределений*

**Пример 1.8.** Изучение свойств локальной функции Лапласа. Построить график локальной функции Лапласа. Используя график, отметить характерные особенности функции. Вычислить значение функции в точке х=4.

**Решение**

****

*Рис. 1.10. График локальной функции Лапласа*

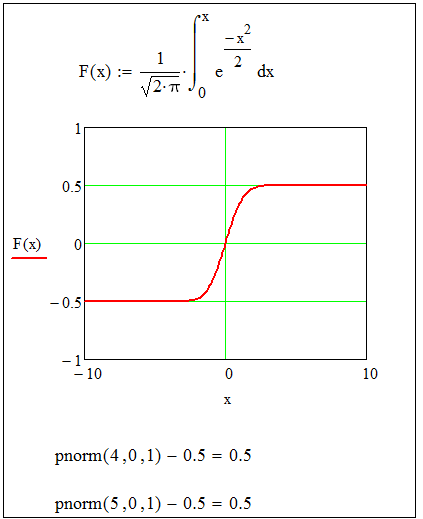
***Свойства локальной функции Лапласа***

1. Функция является четной (график симметричен относительно оси ординат), т.е. .

2. Функция - монотонно убывающая при положительных значениях *х*, причем при  Практически можно считать, что уже при 

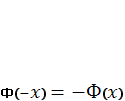
**Пример 1.9.** Изучение свойств функции Лапласа. Построить график функции Лапласа. Используя график, отметить характерные особенности функции. Вычислить значения функции при х=4; х=5.

**Решение.** Проведем анализ графика функции Лапласа (рис. 1.11).

******

*Рис. 1.11. График функции Лапласа*

***Свойства функции Лапласа***

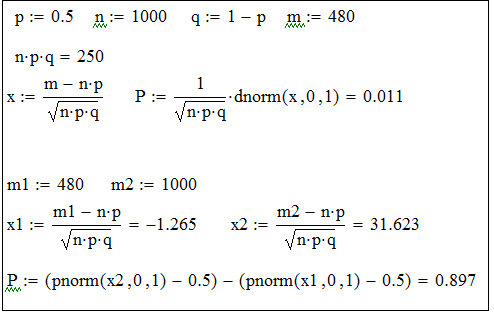
1. Функция является нечетной (график симметричен относительно начала координат), т.е. .

2. Функция – монотонно возрастающая.

Причем при . Практически можно считать, что уже при 

**Пример 1.10.** По результатам проверок налоговыми инспекциями установлено, что в среднем каждое второе малое предприятие региона имеет нарушение финансовой дисциплины. Найти вероятность того, что из 1000 зарегистрированных в регионе малых предприятий имеют нарушения финансовой дисциплины: а) 480 предприятий; б) не менее 480 предприятий.

**Решение.** Решение задачи приведено на рис. 1.12.

****

*Рис. 1.12. Решение примера 5 в MathCAD*

**Задачи для аудиторного решения**

1. В новом микрорайоне поставлено 10000 кодовых замков на входных дверях домов. Вероятность выхода из строя одного замка в течение месяца равна 0,02%. Найти вероятность того, что за месяц откажут два, три, пять замков.

2. Обувной магазин продал за месяц 2000 пар обуви. В среднем 1% обуви возвращается покупателями в магазин в течение месяца. Найти вероятность того, что из проданных пар будет возвращено ровно 4 пары; от 5 до 10 пар.

3. Четыре покупателя приехали на оптовый склад. Вероятность того, что каждому из этих покупателей потребуется холодильник, равна 0,4. Найти вероятность того, что холодильник потребуется: всем четырем покупателям; не более, чем трем; не менее, чем двум.

**Компьютерная лабораторная работа № 1**

***Вероятностные схемы***

**Задания к компьютерной лабораторной работе № 1**

*Умение решать задачи –*

*такое же практическое искусство,*

*как умение плавать или бегать на лыжах.*

*Ему можно научиться только путём*

*подражания или упражнения*.  
Д. Пойа

**Задание 1. Решить задачи, используя встроенные функции MathCAD для формул комбинаторики.**

**1.1**

1. В бригаде 4 женщины и 3 мужчин. Среди членов бригады разыгрываются 4 билета в театр. Какова вероятность того, что среди обладателей билетов окажется 2 женщины и 2 мужчин?

2. В классе 40 учеников, из которых 10 отличников. Класс наудачу разделен на две равные группы. Какова вероятность того, что в каждой группе по 5 отличников?

3. В урне 6 белых и 4 черных шара. Из этой урны наудачу извлечены 5 шаров. Какова вероятность того, что 2 из них белые, а 3 черные?

4. В команде из 12 спортсменов 5 мастеров спорта. По жеребьевке из команды выбирают 3 спортсменов. Какова вероятность того, что все выбранные спортсмены являются мастерами спорта?

5. Студент успел подготовить к экзамену 20 вопросов из 25. Какова вероятность того, что из 3 наудачу выбранных вопросов студент знает не менее двух?

6. В урне 10 шаров, из которых 2 белых, 3 черных, 5 синих. Наудачу извлечены три шара. Какова вероятность того, что все шары разного цвета?

7. В ящике 10 деталей, среди которых 7 окрашенных. Сборщик наудачу достает 4 детали. Найдите вероятность, что все взятые детали окрашены.

8. В команде из 10 спортсменов 4 мастера спорта. По жеребьевке из команды выбирают 3 спортсменов. Какова вероятность того, что все выбранные спортсмены являются мастерами спорта?

9. В бригаде 6 женщин и 5 мужчин. Среди членов бригады разыгрываются 4 билета в театр. Какова вероятность того, что среди обладателей билетов окажется 2 женщины и 2 мужчин?

10. В урне 10 шаров, из которых 2 белых, 3 черных, 5 синих. Наудачу извлечены три шара. Какова вероятность того, что все шары одного цвета?

**1.2.** В партии из *N* изделий *n* изделий имеют скрытый дефект. Какова вероятность того, что из взятых наугад *т* изделий *k* изделий являются дефектными? Данные по вариантам указаны ниже.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| вариант | *N* | *n* | *m* | *k* |
| 1 | 20 | 4 | 5 | 2 |
| 2 | 30 | 5 | 5 | 3 |
| 3 | 20 | 5 | 4 | 2 |
| 4 | 25 | 6 | 5 | 3 |
| 5 | 15 | 4 | 3 | 2 |
| 6 | 20 | 6 | 4 | 1 |
| 7 | 30 | 4 | 3 | 2 |
| 8 | 16 | 4 | 3 | 2 |
| 9 | 18 | 6 | 5 | 3 |
| 10 | 12 | 5 | 4 | 2 |

**Задание 2. Решить задачу, используя операторы символьной палитры** **coeffs, colleсt, expand.**

Мастер обслуживает 4 станка, работающих независимо друг от друга. Вероятность того, что первый станок в течение смены потребует внимания рабочего, равна р1, второй – р2, третий – р3 и четвертый – р4. Найти вероятность того, что в течение смены: ни один станок не потребует внимания мастер; один станок потребует внимания; два станка потребуют внимания; три станка; все станки потребуют внимания. Данные по вариантам указаны ниже.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| вариант | р1 | р2 | р3 | р4 |
| 1 | 0,32 | 0,61 | 0,4 | 0,25 |
| 2 | 0,33 | 0,12 | 0,2 | 0,11 |
| 3 | 0,12 | 0,24 | 0,21 | 0,09 |
| 4 | 0,04 | 0,62 | 0,15 | 0,205 |
| 5 | 0,03 | 0,42 | 0,37 | 0,41 |
| 6 | 0,17 | 0,18 | 0,42 | 0,37 |
| 7 | 0,43 | 0,27 | 0,51 | 0,33 |
| 8 | 0,07 | 0,08 | 0,11 | 0,22 |
| 9 | 0,32 | 0,43 | 0,23 | 0,25 |
| 10 | 0,3 | 0,6 | 0,4 | 0,25 |

**Задание 3. Решить задачу, используя встроенные функции MathCAD в схеме Бернулли.**

1. Магазин получил 100 бутылок минеральной воды. Вероятность того, что при перевозке бутылка окажется разбитой, равна 0,003. Найдите вероятность того, что магазин получил одну разбитую бутылку.

2. В среднем левши составляют 1%. Какова вероятность того, что среди 200 студентов найдется ровно 4 левши?

3. Кинотеатр вмещает 730 зрителей. Найдите вероятность того, что 3 зрителя родились в один день (скажем, 1 марта).

4. 100 станков работают независимо друг от друга, причем вероятность бесперебойной работы каждого из них в течение смены равна 0,8. Найдите вероятность того, что в течение смены бесперебойно проработают от 75 до 85 станков.

5. Вероятность того, что покупателю требуется обувь 41-ого размера, равна 0,2. Найдите вероятность того, что среди 100 покупателей потребуют обувь 41-ого размера 25 человек.

6. Вероятность того, что денежный приемник автомата при опускании монеты сработает неправильно, равна 0,03. Найдите наиболее вероятное число случаев правильной работы автомата, если будет опущено 150 монет и вероятность этого события.

7. По данным технического контроля 2% изготовленных станков нуждаются в дополнительной регулировке. Найдите вероятность того, что из 6 изготовленных станков 4 нуждаются в дополнительной регулировке.

8. Вероятность попадания в мишень при одном выстреле для данного стрелка равна 0,8 и не зависит от номера выстрела. Требуется найти вероятность того, что при 5 выстрелах произойдет ровно 2 попадания в мишень.

9. Вероятность того, что покупателю требуется обувь 41-ого размера, равна 0,2. Найдите вероятность того, что среди 100 покупателей потребуют обувь 41-ого размера 30 человек.

10. 100 станков работают независимо друг от друга, причем вероятность бесперебойной работы каждого из них в течение смены равна 0,9. Найдите вероятность того, что в течение смены бесперебойно проработают от 70 до 80 станков.

**Задание 4.** В опытах Бюффона (18 век) относительная частота появления герба при 4040 подбрасываниях монеты оказалась равной 0,5069; в опытах Пирсона (19 век) при 23000 подбрасываниях – 0,5005. Проиллюстрируйте процесс стабилизации относительных частот, осуществив имитацию экспериментов в MathCAD.

**Задание 5.** Используя данные условия примера 1.6, проведите графический анализ влияния числа испытаний на кумулятивные вероятности в схеме Бернулли. Измените вероятность успеха в каждом испытании: р=0,2; р=0,5; р=0,6.

**Задание 6.** В примере 1.7 установлено, что при увеличении *k* и одном и том же значении *п* значения вероятностей, полученные по формулам Бернулли и Пуассона, достаточно быстро приближаются друг к другу. Проведите сопоставительный анализ влияния числа испытаний *п* на точность асимптотического приближения Пуассона.

**1.2. Случайные величины (дискретные и непрерывные).**

**Основные законы распределения случайных величин**

*Много из математики не остается*

*в памяти, но когда поймешь ее,*

*тогда легко при случае вспомнить забытое.*  
М.В. Остроградский

**1.2.1. Случайные величины. Способы задания, числовые характеристики**

*Случайной величиной Х* (СВ *Х*)называется величина, которая в результате опыта может принять любые заранее неизвестные значения. Различают дискретные и непрерывные случайные величины.

*Дискретной* *случайной величиной* (ДСВ) называется величина, которая может принимать конечное или бесконечное счетное множество значений, т.е. значения принципиально могут быть перечислены.

*Непрерывной случайной величиной* (НСВ) называется случайная величина, если существует такая неотрицательная интегрируемая на всей числовой оси функция , называемая **плотностью распределения** вероятностей, интеграл от которой в пределах от *a* до *b* определяет вероятность попадания значений СВ *Х* в произвольный промежуток от (*a*; *b*)

.

***Замечание****. Для НСВ Х равноправны*

Всякое соотношение, устанавливающее связь между возможными значениями ДСВ *Х* и соответствующими им вероятностями, называют *законом распределения*. Этот закон может быть задан:

- таблично (ряд распределения);

- графически (приводится график, соединяющий точки который носит название *многоугольник распределения*;

- аналитически (указываются аналитические выражения для значений СВ *Х* и их вероятностей).

Основные понятия, связанные с законами распределения ДСВ *Х* и НСВ *Х* представлены ниже.

|  |  |
| --- | --- |
| **Дискретные**  **случайные величины** | **Непрерывные**  **случайные величины** |
| Ряд распределения   |  |  |  |  |  | | --- | --- | --- | --- | --- | |  |  | … |  | … | |  |  | … |  | … | | Плотность распределения |
| Функция распределения и  кумулятивная вероятность  3) | Функция распределения    3) = |

**Числовыми характеристиками**  случайной величины называют **неслучайные** численные параметры, позволяющие в предельно сжатой, компактной форме представить основные особенности закона распределения.

**Типы числовых характеристик**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **СВ *X*** | **Характеристики положения** | | | | |
| **М(х)=mx**  математическое ожидание | **Ме(х)**  медиана | | | **Мо(х)**  мода |
| **ДСВ** |  | **\_** | | | наиболее вероятное значение |
| **НСВ** |  | квантиль  (делит площадь под плотностью вероятности пополам) | | | точка максимума функции  плотности |
|  | **Характеристики рассеяния** | | | | |
| **D(X)**  дисперсия | среднее  квадратическое отклонение | | | коэффициент вариации |
| **ДСВ** |  |  | | | оценка относительного рассеяния СВ *Х* по сравнению со средним значением |
| **НСВ** |  |
|  | **Теоретические моменты** | | | | |
| **Начальные моменты**  **k-го порядка** | | **Центральные моменты**  **k-го порядка** | | |
| **ДСВ** |  | |  | | |
| **НСВ** |  | |  | | |
|  | **Характеристики формы** | | | | |
| асимметрия | | | эксцесс | |
| **ДСВ** | характеризует степень несимметричности распределения | | | характеризует степень крутизны распределения | |
| **НСВ** |

***Замечания.***

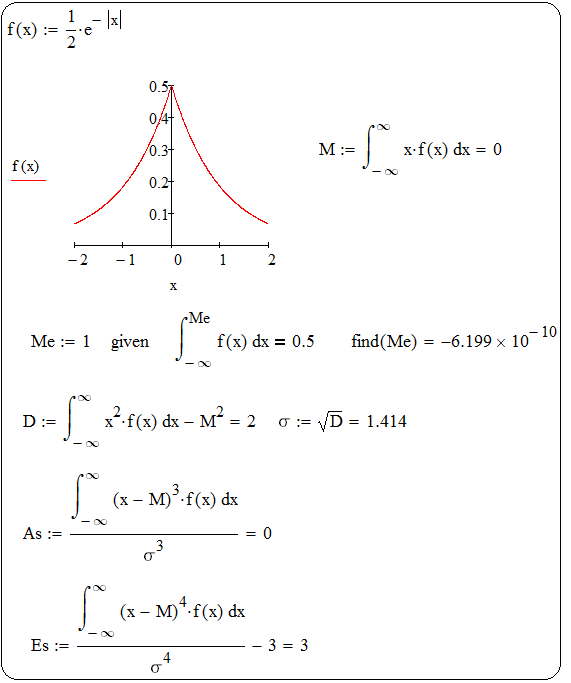
1***. Квантиль*** *порядка р – такое значение СВ Х, для которого*

2*. Среди законов распределения «эталоном формы» признается стандартное нормальное распределение с функцией плотности (локальная функция Лапласа, график изображен на рис. 1.10).*

**Пример 1.11.** Непрерывная СВ *Х* задана функцией плотности (закон Лапласа). Построить график, найти числовые характеристики.

**Решение.** График функции плотности изображен на рис. 1.13. Полученные числовые характеристики сведены в таблицу.

|  |  |  |  |
| --- | --- | --- | --- |
| Характеристики положения | | | |
|  | (при х=0 функция достигает максимума) | |  |
| Характеристики рассеяния | | | |
|  | |  | |
| Характеристики положения | | | |
| (распределение СВ *Х* симметрично относительно оси ординат) | | (эксцесс распределения положителен, что говорит об островершинности кривой распределения) | |

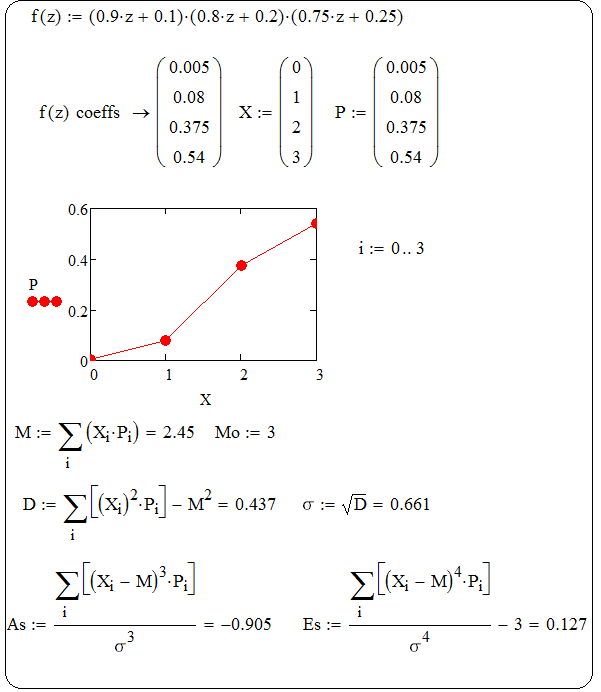


*Рис. 1.13. Решение примера 1.11*

**Обратите внимание.** Значение медианы определено путем решения уравнения с помощью вычислительного блока **«given-find»**. Предварительно задается приближение искомой величины.

**Пример 1.12.** Рабочий обслуживает 3 станка. Вероятность того, что в течение часа станок не потребует внимания рабочего, для первого станка равна 0,9, для второго – 0,8, для третьего – 0,75. Составить закон распределения СВ *Х* – числа станков, которые не потребуют внимания рабочего в течение часа. Построить многоугольник распределения, вычислить числовые характеристики.

**Решение**

****

*Рис. 1.14 Решение примера 1.12*

Полученные числовые характеристики сведем в таблицу.

|  |  |  |  |
| --- | --- | --- | --- |
| Характеристики положения | | | |
|  | | (наиболее вероятное значение) | |
| Характеристики рассеяния | | | |
|  |  | | (если рассеяние составляет десятки процентов, то пренебрегать  им нельзя) |
| Характеристики положения | | | |
| (левосторонняя асимметрия, наблюдается небольшое смещение относительно математического ожидания) | | (распределение островершинное) | |

**Задачи для аудиторного решения**

1. СВ *Х* задана плотностью вероятности

Построить график функции плотности. Найти числовые характеристики.

2. СВ *Х* задана плотностью вероятности

. Определить параметр *а*, числовые характеристики.

3. Два покупателя независимо друг от друга делают по одной покупке. Вероятность того, что покупку сделает первый покупатель, равна 0,8, что второй – 0,6. СВ *Х* – число покупок, сделанных покупателями. Описать закон распределения СВ *Х*, построить многоугольник распределения, найти числовые характеристики.

4. Партия содержит 20 телевизоров, среди которых шесть с дефектом. Купили два телевизора. Составить ряд распределения исправленных телевизоров среди купленных. Построить многоугольник распределения, найти числовые характеристики.

**1.2.2. Библиотека стандартных распределений**

Для работы со случайными величинами (непрерывными и дискретными) в MathCAD есть библиотека встроенных функций наиболее распространенных распределений. Каждое распределение представлено в библиотеке: функцией распределения; функцией, вычисляющей вероятность заданного значения (для дискретных распределений); плотностью вероятностей (для непрерывных распределений); квантили распределения порядка *k*; генератором случайных чисел с заданным законом распределения.

Все функции распределения вероятностей в MathCAD начинаются с буквы «**р**». Функции, имена которых начинаются с буквы «**d**», задают плотность распределения вероятностей (или вычисляют вероятность в заданной точке «k» для дискретных величин); с «**q**» – квантили и с «**r**» – генерируют вектор *m* случайных чисел с соответствующим законом распределения.

Затем идет название закона, в скобках записывают (x, par), где par – параметры закона.

Рассмотрим наиболее часто встречающиеся в вероятностных приложениях законы распределения случайных величин.

**Библиотека основных распределений НСВ**

|  |  |  |  |
| --- | --- | --- | --- |
| **Функция плотности** | **Функция распределения** | **Числовые характеристики** | |
| **Равномерное распределение (параметры *- а, в*)** | | | |
| dunif(*x,a,b*) | рunif(*x,a,b*) |  | |
| **Экспоненциальное распределение (параметр - λ)** | | | |
| dexp(*x,λ*) | pexp(*x,λ*) |  | |
| **Нормальное распределение (параметры *- а*, σ)** | | | |
| dnorm(*x,a,σ*) | где Ф(t) – функция Лапласа  pnorm(*x,a,σ*) |  | |
| **Логнормальное распределение (параметры *- а*, σ)** | | | |
| dlnorm(*x,a,σ*) | где Ф(t) – функция Лапласа  plnorm(*x,a,σ*) | |  |
| **Распределение**  **(k - число степеней свободы)** | | | |
| dchisq(*x,k*) | pchisq(*x,k*) | |  |
| **Распределение Стьюдента (k - число степеней свободы)** | | | |
| dt(*x,k*) | pt(*x,k*) |  | |
| **Распределение Фишера (k1, k2 - две степени свободы)** | | | |
| dF(*x,k1,k2*) | pF(*x,k1,k2*) |  | |

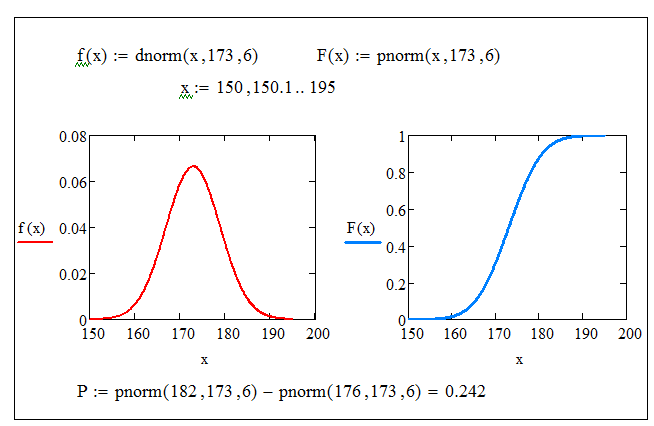
**Библиотека основных распределений ДСВ**

|  |  |  |
| --- | --- | --- |
| **Кумулятивная (накопленная) вероятность** | **Вероятность заданного значения** | **Числовые характеристики** |
| **Биномиальный закон (параметр - *р*)** | | |
| pbinom(*k,n,p*) | dbinom(*k,n,p*) |  |
| **Закон Пуассона (параметр – λ)** | | |
| ppois(*k,**λ*) | dpois(*k,**λ*) |  |
| **Геометрическое распределение (параметр - *р*)** | | |
| pgeom(*k,p*) | dgeom(*k,p*) |  |
| **Гипергеометрическое распределение (параметры – *n1, n2, n*)** | | |
| phypergeom(*k,n1,n2,n*) | dhypergeom(*k,n1,n2,n*) |  |

**1.2.3. Примеры построения графиков функций плотности и распределения непрерывных случайных величин**

**Пример 1.13.** Полагая, что рост мужчин определенной возрастной группы есть нормально распределенная величина *Х* с параметрами =36, построить графики функции плотности и функции распределения СВ *Х*; рассчитать долю костюмов 4-го роста (176-182), которые нужно предусмотреть в общем объеме производства.

**Решение.** Для построения графиков, используем встроенные функции **dnorm(x,*a,*σ) –** функция плотности и **pnorm(x,*a*,σ)** - функция распределения**.** Решение отображено на рис. 1.15.

****

*Рис. 1.15. Нормальный закон распределения*

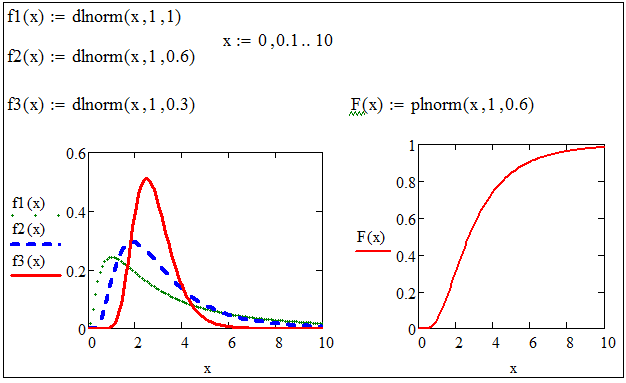
Следовательно, на долю костюмов четвертого роста приходится 24% костюмов от общего объема производства.

**Пример 1.14.** Построить график функции плотности случайной величины Х распределенной по логнормальному закону с параметрами *а =1*, σ = 1; *а =1*, σ = 0,6 на интервале . Сделать вывод об изменении формы кривой, отметить характерные особенности кривой.Построить график функции распределения СВ *Х*.

**Решение.** Для построения графиков, используем встроенные функции **dlnorm(x,*a,*σ)** – функция плотности и **plnorm(x,*a*,σ)** - функция распределения**.** Решение отображено на рис. 1.16.

**Вывод.** Чем меньше σ, тем кривая распределения ближе к симметрии. Если в нормальном законе параметр *а* выступает в качестве *среднего значения* случайной величины, то в логнормальном - в качестве *медианы*.

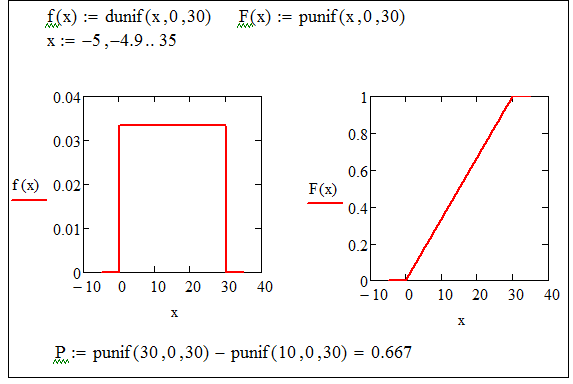
***Замечание****.* *Логнормальное распределение используется для описания распределения доходов, банковских вкладов, месячной заработной платы, долговечности изделий в режиме износа и др.*

****

*Рис. 1.16. Логнормальный закон распределения*

**Пример 1.15.** Автобусы подходят к остановке регулярно с интервалом 30 минут. СВ *Х* – время ожидания автобуса на временном отрезке [0;30] имеет равномерный закон распределения. Построить график функции плотности равномерного закона и график функции распределения. Найти вероятность прождать более 10 минут.

**Решение**



*Рис. 1.17. Равномерный закон распределения*

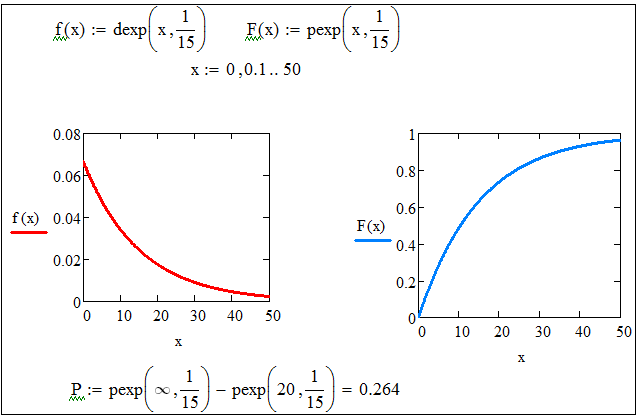
**Пример 1.16.** Установлено, что время ремонта телевизоров есть случайная величина *Х*, распределенная по показательному закону. Определить вероятность того, что на ремонт телевизора потребуется не менее 20 дней, если среднее время ремонта составляет 15 дней. Построить график плотности вероятности и график функции распределения.

**Решение**

По условию математическое ожидание 1/λ, следовательно, параметр показательного закона λ=1/15.

Для построения графиков используем встроенные функции **dexp(x,λ)** и **pexp(x,λ)**.

Решение отображено на рис. 1.18.



*Рис. 1.18. Показательный закон распределения*

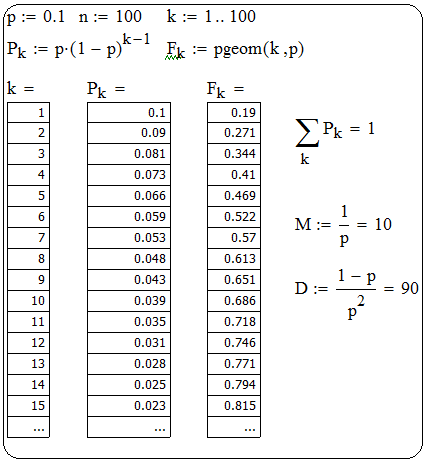
**1.**

**2.4. Примеры построения законов распределения ДСВ**

**Пример 1.17.** Проводится проверка партии деталей (100 изделий) до обнаружения бракованной. Составить закон распределения числа проверенных деталей. Найти его математическое ожидание и дисперсию, если известно, что вероятность брака для каждой детали равна 0.1. Составить функцию распределения.

**Решение**

СВ *Х* – число проверенных деталей до обнаружения бракованной – имеет геометрическое распределение с параметром р = 0,1. Решение приведено на рис. 1.19.



*Рис.1.19. Геометрический закон распределения*

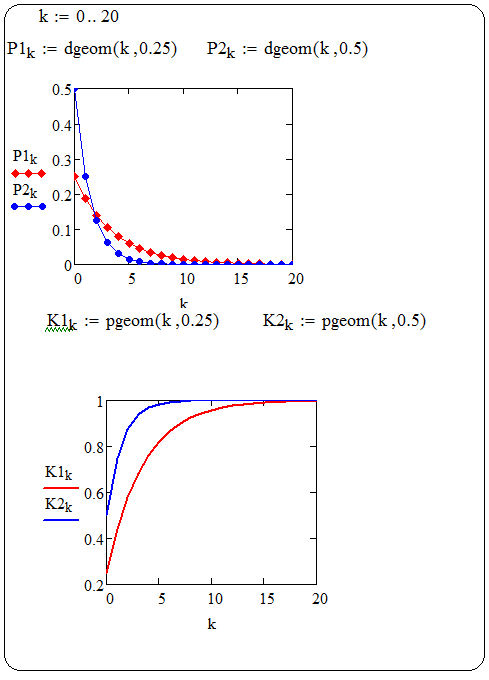
**Пример 1.18.** Проведем анализ геометрического распределения.

**Решение**

Построим многоугольники распределения для двух значений вероятностей появления события в одиночном испытании р1=0,25 и р2=0,5.

Из графиков на рис. 1.20 видно, что вероятности при росте *k* имеют убывающий характер. Скорость убывания вероятностей успеха растет с увеличением значения *р*.

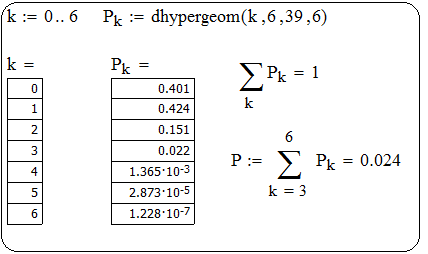
Характер нарастания кумулятивных вероятностей геометрического распределения при росте *k* также зависит от *р*.

****

*Рис. 1.20. Характеристики геометрического распределения*

**Пример 1.19.** В лотерее «Спортлото 6 из 45» денежные призы получают участники, угадавшие 3, 4, 5, 6 видов спорта из отобранных случайно 6 видов из 45. Найти закон распределения СВ *Х* – число угаданных видов спорта среди случайно отобранных шести. Какова вероятность получения денежного приза?

**Решение.** СВ *Х* – число угаданных видов спорта среди случайно отобранных шести – имеет гипергеометрический закон распределения. Решение приведено на рис. 1.21.



*Рис. 1.21. Гипергеометрический закон распределения*

***Замечание.*** *Проанализируйте результаты решения примера 1.19. Затем внимательно прочитайте следующие высказывания знаменитых математиков.*

1. *Лотерея, по сути, налог на несчастливых самонадеянных дураков (*УильямПетти, 1662).

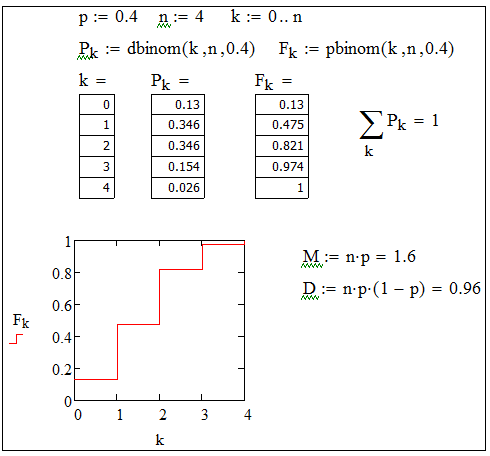
2. *Такой-то, говорите вы, получил первую ставку и вдруг приобрел состояние – почему такая же ставка не достанется и мне? Вы невольно сравниваете себя с выигравшим и не хотите подумать, что гораздо естественнее поместить себя в число проигравших, потому что их несравненно более* (Михаил Остроградский, 1847).

3. *Рулетка не имеет ни воли, ни памяти* (Жан Бертранд, 1888).

**Пример 1.20.** В магазин поступила обувь с двух фабрик в соотношении 2:3. Куплено 4 пары обуви. Найти закон распределения числа купленных пар обуви, изготовленной первой фабрикой. Найти математическое ожидание, дисперсию этой случайной величины. Составить функцию распределения.

**Решение.** Вероятность того, что случайно выбранная пара обуви изготовлена первой фабрикой, равна

СВ *Х* – число пар обуви среди четырех, имеет биномиальный закон распределения с параметрами *п = 4, р = 0,4*. Решение приведено на рисунке 1.22.



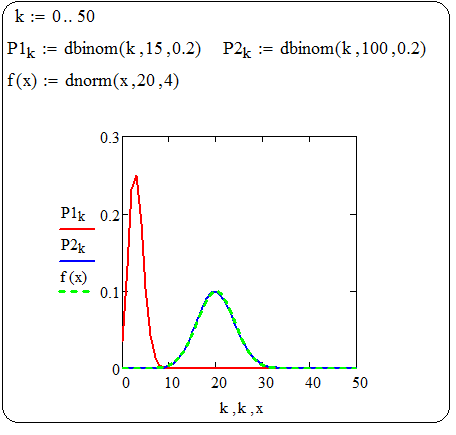
*Рис. 1.22. Биномиальный закон распределения*

**Пример 1.21.** Проведем анализ биномиального распределения.

**Решение.** Рассмотрим две совокупности параметров биномиального распределения ( Графики представлены на рисунке 1.23.

Анализируя результаты, заметим, что с ростом числа испытаний *п* биномиальное распределение по форме приближается к нормальному распределению. Для иллюстрации этого факта построен график функции плотности нормального распределения с параметрами второго варианта биномиального распределения.

Отмеченная близость биномиального распределения при больших *п* к нормальному закону служит косвенным основанием применения асимптотических формул Муавра-Лапласа и Лапласа.

****

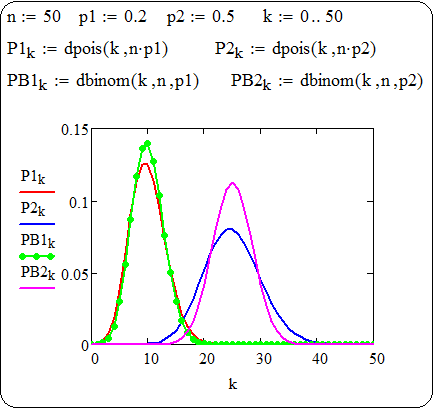
*Рис. 1.23. Характеристики биномиального закона*

**Пример 1.22.** Проведем анализ закона Пуассона.

**Решение.** На графике приведены значения распределения Пуассона для λ=10 (Р1=0,2) и λ=20 (Р2=0,5). При увеличении значения вероятности успеха *р* математическое ожидание λ=*пр* увеличивается, центр распределения Пуассона смещается вправо. Это сопровождается «растеканием» зависимости вероятности успеха от числа успехов *k*, т.е. увеличением дисперсии.

На рисунке представлены также графики вероятности, полученные по формуле Бернулли (РВ1, РВ2) для каждого случая. Из полученных результатов следует, что отличие закона Пуассона от биномиального закона растет с увеличением значений вероятности успеха *р*.

***Замечание.*** *Математическое ожидание и дисперсия СВ Х, распределенной по закону Пуассона, равны между собой. Этот факт можно использовать на практике в качестве критерия при проверке гипотезы о том, что исследуемая СВ распределена по закону Пуассона. Если эти значения близки между собой, то имеются основания полагать, что выдвинутая гипотеза истинна.*

****

*Рис. 1.24. Характеристики закона Пуассона*

**Компьютерная лабораторная работа № 2**

***Основные законы распределения НСВ***

*Иногда шансы не подчиняются*

*никакому ощутимому закону и*

*кривая возможностей [плотность]*

*может принимать*

*наиболее причудливые формы.*

Адольф Кетле (1846)

**Задания к компьютерной лабораторной работе № 2**

**Задание 1.** Выяснить, как будет меняться нормальная кривая



при изменении параметров *а* и *σ*.

1. σ = const, изменяется параметр *а* .

На одной координатной плоскости построить с помощью встроенной функции **dnorm(x,*a,*σ)** графики функции плотности нормального закона, если σ = 2, 

2. *а* = const, изменяется параметр σ . .

На одной координатной плоскости построить с помощью встроенной функции **dnorm(x,*a,*σ)** графики функции плотности нормального закона, если **.**

3. Сделать вывод о том, какой параметр характеризует форму, какой положение нормальной кривой. Как меняется форма, как меняется положение нормальной кривой при изменении параметров нормального закона?

4. Построить график функции распределения нормального закона с параметрами  и σ = 2, используя встроенную функцию **pnorm(*x*,*a,*σ).**

**Задание 2.** Построить функцию плотности закона распределения (хи–квадрат) при k=2; k=5; k=10 на интервале . Использовать встроенную функцию **dchisq(*x,k***). Сделать вывод об изменении формы кривой с увеличением параметра k, отметить характерные особенности кривой.

***Замечание.*** *Распределением*  *с k степенями свободы называется распределение суммы квадратов k**независимых случайных величин, каждая из которых подчинена нормальному закону распределения с математическим ожиданием, равным нулю и дисперсией, равной единице. Аналитически это распределение выражается через гамму-функцию. Закон распределения* *используется во многих задачах математической статистики.*

**Задание 3.** Построить функцию плотности закона распределения Стьюдента (t-распределение) с помощью встроенной функции **dt(x,k)**

при k=2; k=30 на интервале 0. Сравнить графически с нормальной кривой *N*(0;1). Сделать вывод.

***Замечание.*** *Распределением Стьюдента или «t»-распределением называется распределение отношения*

*где Z – случайная величина, распределенная по стандартному нормальному закону, т.е. N(0;1). Распределение Стьюдента широко используется в математической статистике, дисперсионном анализе.*

**Задание 4.**Построить функцию плотности закона распределения Фишера-Снедекора (F-распределение) с помощью встроенной функции **dF(x,k1,k2**) при:

1) k1=1, k2=4;

2) k1=10, k2=50 на интервале . Сделать вывод.

**Задание 5.** Проведенное исследование показало, что вклады населения в данном банке могут быть описаны СВ Х, распределенной по логнормальному закону с параметрами *а* = 530, σ = 0,64. Пояснить смысл параметра *а*; найти средний размер вклада; долю вкладчиков, размер вклада которых составляет не менее 1000 ден. ед. Построить графики функции плотности и распределения.

**Задание 6**. Составьте полную библиотеку основных распределений случайных величин, которые представлены в системе MathCAD, заполнив таблицу. Графики постройте с помощью встроенных функций, задавая параметры закона самостоятельно.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Распределение | Встроенные функции | Функция плотности и ее график | Функция распределения и ее график | *M(X),*  *D(X)* |

**Компьютерная лабораторная работа № 3**

***Основные законы распределения ДСВ***

**Задания к компьютерной лабораторной работе № 3**

*Если вы хотите научиться плавать,*

*то смело входите в воду,*

*а если хотите научиться решать задачи,*

*то решайте их.*

Джордж Пойа

**Задание 1.**  В городе имеется *N* оптовых баз. Вероятность того, что требуемого сорта товар отсутствует на этих базах одинакова и равна *р.* Составить закон распределения числа баз, на которых искомый товар отсутствует в данный момент времени. Вычислить числовые характеристики случайной величины. Составить функцию распределения.

*Указание: использовать встроенные функции для биномиального распределения.*

|  |  |  |
| --- | --- | --- |
| **вариант** | ***N*** | ***p*** |
| 1 | 5 | 0,12 |
| 2 | 6 | 0,13 |
| 3 | 7 | 0,21 |
| 4 | 4 | 0,07 |
| 5 | 5 | 0,11 |
| 6 | 6 | 0,05 |
| 7 | 7 | 0,06 |
| 8 | 5 | 0,08 |
| 9 | 6 | 0.14 |
| 10 | 4 | 0,32 |

**Задание 2.** Найти закон распределения, функцию распределения, вычислить числовые характеристики СВ *Х*.Построить многоугольник распределения. Решить задачу, используя встроенные функции MathCAD.

2.1. Из пяти гвоздик две белые. Составить закон распределения случайной величины, выражающей число белых гвоздик среди двух одновременно взятых.

2.2. В коробке 20 одинаковых деталей, из них – 4 детали с дефектом. Наудачу вынимают две детали. Найти закон распределения числа деталей с дефектом.

2.3. Из 6 телевизоров на выставке 3 оказались фирмы «Сони». Наудачу для осмотра выбрано 2. Составить закон распределения числа телевизоров фирмы «Сони» среди 2 отобранных.

2.4. Среди 15 лотерейных билетов имеется 4 билета с выигрышем. Наудачу покупают 3 билета. Написать закон распределения вероятностей числа выигрышных билетов среди купленных.

2.5. В коробке 20 одинаковых катушек с ниток, из них – 4 катушки с белыми нитками. Наудачу вынимают две катушки. Найти закон распределения числа катушек с белыми нитками.

2.6. Из партии в 20 изделий, среди которых имеется 6 бракованных, выбраны случайным образом 3 изделия для проверки их качества. Построить ряд распределения случайной величины *Х* – числа бракованных изделий среди отобранных.

2.7. В партии из 25 кожаных курток 5 имеют скрытый дефект. Покупают три куртки. Найти закон распределения числа дефектных курток среди купленных.

2.8. Из 8 телевизоров на выставке 5 оказались фирмы «Сони». Наудачу для осмотра выбрано 3. Составить закон распределения числа телевизоров фирмы «Сони» среди 3 отобранных.

2.9. Среди 10 лотерейных билетов имеется 4 билета с выигрышем. Наудачу покупают 2 билета. Написать закон распределения вероятностей числа выигрышных билетов среди купленных.

2.10. Группа студентов состоит из 15 юношей и 10 девушек. Разыгрывается 4 билета в театр. Составить закон распределения случайной величины *Х* – число юношей, пошедших в театр.

**Задание 3.** Решить задачу, используя встроенные функции MathCAD.

3.1. Вероятность того, что в библиотеке необходимая студенту книга свободна, равна 0,3. Составить закон распределения числа библиотек, которые посетит студент, если в городе 4 библиотеки. Найти функцию распределения, математическое ожидание, дисперсию.

3.2. На пути движения автомобиля 4 светофора, каждый из них или разрешает, или запрещает дальнейшее движение с вероятностью 0,5. Найдите закон распределения случайной величины *Х*, равной числу светофоров, пройденных автомобилем до первой остановки. Составить функцию распределения, вычислить математическое ожидание и дисперсию случайной величины.

3.3. На пути движения судна 4 шлюза, каждый из них или разрешает, или запрещает дальнейшее движение с вероятностью 0,5. Найдите закон распределения случайной величины *Х*, равной числу шлюзов, пройденных судном до первой остановки. Составить функцию распределения, вычислить математическое ожидание и дисперсию случайной величины.

3.4. Вероятность того, что необходимая студенту формула находится в данном справочнике 0,3. Составить закон распределения числа справочников, которые возьмет студент, если в библиотеке 8 справочников по данной дисциплине. Найти функцию распределения, математическое ожидание, дисперсию.

3.5 Ученый проводит опыты до первого успеха, но не более 10. Вероятность успеха в каждом опыте равна 0,4. Составить закон распределения случайной величины *Х* – число опытов, проведенных ученым до первого успеха, функцию распределения, вычислить математическое ожидание, дисперсию.

3.6. Бросается монета до первого появления «герба», но не более 10 раз. Составить закон распределения случайной величины *Х* – число подбрасываний монет до выпадения герба, функцию распределения, вычислить математическое ожидание, дисперсию.

3.7. Вероятность того, что необходимая студенту формула находится в данном справочнике 0,4. Составить закон распределения числа справочников, которые возьмет студент, если в библиотеке 15 справочников по данной дисциплине. Найти функцию распределения, математическое ожидание, дисперсию.

3.8. Стрелок имеет 7 патронов, стреляет по мишени до первого попадания с вероятностью попадания при каждом выстреле 0,6. Составить закон распределения случайной величины *Х* – число выстрелов до первого попадания, функцию распределения, вычислить математическое ожидание, дисперсию.

3.9. Ученый проводит опыты до первого успеха, но не более 10. Вероятность успеха в каждом опыте равна 0,2. Составить закон распределения случайной величины *Х* – число опытов, проведенных ученым до первого успеха, функцию распределения, вычислить математическое ожидание, дисперсию.

3.10. Бросается монета до первого появления «герба», но не более 15 раз. Составить закон распределения случайной величины *Х* – число подбрасываний монет до выпадения герба, функцию распределения, вычислить математическое ожидание, дисперсию.

*Как бы машина хорошо ни работала,*

*она может решать все требуемые от нее задачи,*

*но она никогда не придумает ни од­ной.*

Альберт Эйнштейн

**Задание 4.** Придумайте и решите задачу, в которой СВ *Х* распределена по закону Пуассона (закон редких событий).

**Глава 2**

**Математическая статистика**

*История — это движущаяся статистика,*

*а статистика — это застывшая история.*

Август Щлецер (1804)

*Знание статистики схоже со знанием*

*иностранных языков или алгебры;*

*оно может оказаться полезным в любое время*

*и при любых обстоятельствах.*

Bowley (1901)

*В каждой науке, в каждом искусстве и объекте изучения*

*существует статистическая или количественная часть.*

*Сознаюсь, я люблю числа настолько же, насколько другие их*

*ненавидят. Но что мне нравится, это не накоплять числа, а выяснять, до*

*какой степени надлежит отбирать значения, обсуждать их и, иначе*

*говоря, подчинять их законам логики и здравого смысла.*

Alph.DeCandolle (1855)

*Грубую прозу статистики они [поэты]*

*когда-нибудь облекут в стихи,*

*потому что цифрами открывается*

*сила, власть, людские слабости,*

*пути истории и много других ... сторон мира ...*

Дмитрий Менделеев (1888)

**2.1. Математические основы выборочного метода**

*Разум не способен полностью представить себе*

*значение хоть сколько-нибудь значительного*

*количества числовых данных*.

Роналд Фишер (1925)

**2.1.1. Выборка и ее представление**

Совокупность возможных значений СВ *Х*, которые принципиально могут быть получены в результате экспериментов, называют **генеральной совокупностью**.

В математической статистике свойства СВ *Х* изучаются на основе некоторого ограниченного множества данных, полученных в результате экспериментов. Это ограниченное множество называют **выборка (выборочная совокупность)**.

Объем выборки – число элементов выборки. Различают малые и большие выборки. Примерной границей служит

Важным принципом формирования выборки является обеспечение случайного отбора в одинаковыхусловиях**.**

**Вариационный ряд** – совокупность *всех* элементов выборки, записанных в неубывающем порядке. В вариационном ряду могут присутствовать одинаковые элементы.

**Статистический ряд** – последовательность различных элементов, записанных в возрастающем порядке и соответствующие им частоты или относительные частоты. Удобно представлять в виде таблицы.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | … |  |  |
|  |  | … |  |  |
|  |  | … |  |  |

Решение задач в условиях большого объема выборки может потребовать «сжатия» полученных данных и перехода к интервальному, а затем к группированному статистическому ряду.

**Группированный статистический ряд –** совокупность середин промежутков и соответствующих им частот.

Группированный статистический ряд также удобно представлять в виде таблицы. Группированный статистический ряд преимущественно используется при анализе вида распределения СВ *Х* по данным наблюдений. Искажая информацию, ведет к появлению дополнительных погрешностей, однако уровень таких погрешностей не является значительным.

Важным вопросом является определение ширины промежутков (интервалов) *h* и их числа *m*.

***Рекомендации по выбору взаимосвязанных значений h и m***

1. ; .

. – формула Стерджесса,

или , тогда

***Замечание.*** *В формулах квадратные скобки**- целая часть числа. Формула*

Графическим изображением статистического ряда служит **полигон частот** или относительных частот – аналог многоугольника распределения дискретной случайной величины.

Интервальный вариационный ряд графически изображают с помощью **гистограммы** частот или относительных частот. Гистограмма – это ступенчатая фигура из прямоугольников, основание которых равно *h*, а высота каждого прямоугольника - или .

Гистограммы обычно используются для предварительного анализа законов распределения исследуемой СВ *Х*, выдвижения гипотезы о виде распределения. С этой целью строят **гистограмму приведенных относительных частот** (гистограмма оценки функции плотности). Площадь всей гистограммы, построенной с использованием равна единице, что позволяет непосредственно сопоставлять такую гистограмму с графиками функций плотности.

*Диаграммы могут облегчить предварительное исследование*

*большинства исходных данных.*

*Они ничего не доказывают, но наглядно показывают*

*отличительные черты.*

*Они поэтому не заменяют таких критических исследований,*

*которым могут быть подвергнуты исходные данные,*

*но они ценны тем, что подсказывают такие исследования и*

*поясняют выводы, основанные на них.*

Роналд Фишер (1925)

В среде MathCAD для обработки выборочных данных и построения гистограммы используются функции, перечисленные ниже.

***Формирование массива данных***

|  |  |
| --- | --- |
| **Параметры матрицы** | |
| число строк массива М | rows(M) |
| число столбцов массива М | cols(M) |
| максимальный элемент массива М | max(M) |
| минимальный элемент массива М | min(M) |
| выноска столбца матрицы М (находится в панели «матрицы») |  |
| **Образование новых матриц** | |
| объединение матриц А и В «бок в бок» (матрицы должны иметь одинаковое число строк) | augment(A,B) |
| объединение матриц А и В «сверху вниз» (матрицы должны иметь одинаковое число столбцов) | stack(A,B) |
| **Сортировка массива** | |
| сортировка элементов массива по возрастанию | sort(M) |
| перестановка элементов в обратном порядке | reverse(M) |

**Пример 2.1.** Имеются следующие данные о размерах основных фондов (в млн руб.) 30 предприятий:

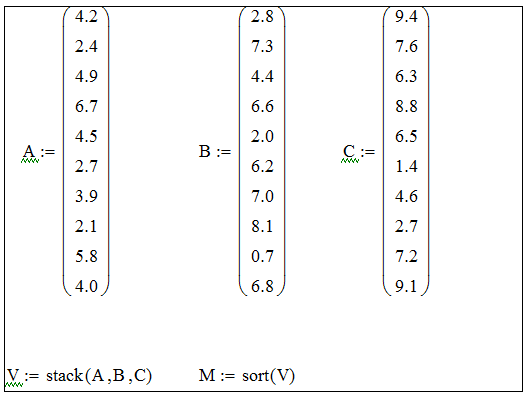
4,2; 2,4; 4,9; 6,7; 4,5; 2,7; 3,9; 2,1; 5,8; 4,0;

2,8; 7,3; 4,4; 6,6; 2,0; 6,2; 7,0; 8,1; 0,7; 6,8;

9,4; 7,6; 6,3; 8,8; 6,5; 1,4; 4,6; 2,0; 7,2; 9,1.

**Решение**

1. Формируем массив данных. Для этого вводим данные с помощью, например, трех матриц A, B, C. Используя встроенную функцию **stack(A,B,С)**, создаем матрицу *V*, содержащую все статистические данные. Выполняем сортировку элементов матрицы *V* по возрастанию с помощью функции **sort(*V*).**

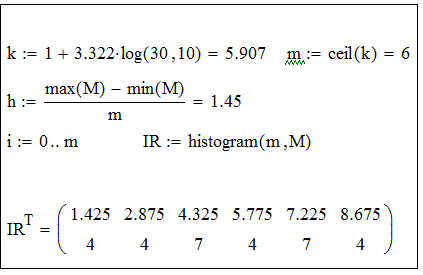
****

*Рис. 2.1. Формирование массива данных*

***Группированный статистический ряд***

|  |  |
| --- | --- |
| **встроенная функция**  **MathCAD** | **описание** |
| hist(x,M) | Вектор частот попадания данных в заданные интервалы.  x – матрица значений случайной величины;  М – массив данных после ранжирования |
| histogram(m,M) | Матрица гистограммы, где  m – число интервалов случайной величины;  М – массив данных после ранжирования |

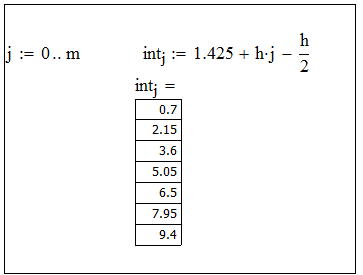
2. Составляем группированный статистический ряд *IR*, рассчитав число интервалов *т* по формуле Стерджеса, округлив до ближайшего целого числа. Находим шаг *h.* Нумеруем концы интервалов, используем встроенную функцию **histogram(m,M).**

****

*Рис. 2.2. Построение группированного статистического ряда*

Получен ряд, разбитый на 6 интервалов. Элементы первой строки – середины интервалов; элементы второй строки – частоты (количество вариант, попавших в данный интервал).

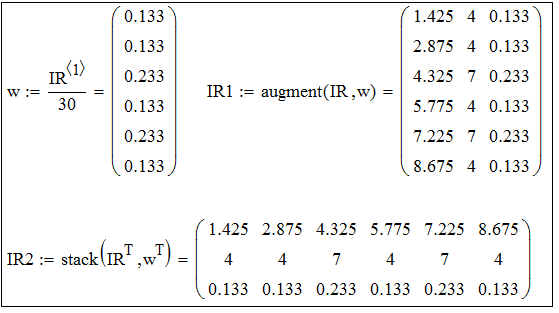
3. Зададим границы интервалов. Значения вариант – концов интервалов имеют вид **intj** (рисунок 2.3).



*Рис. 2.3. Значения вариант – концов интервалов*

4. Вычислим относительные частоты *wi*. Построим группированный статистический ряд частот и относительных частот.

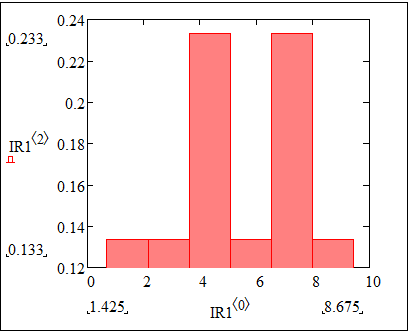
***Замечание.*** *Обратите внимание! Нумерация строк и столбцов матрицы начинается с нуля*.



*Рис. 2.4. Группированный статистический ряд частот и*

*относительных частот*

5. Строим гистограмму частот или относительных частот, распределив по горизонтали интервалы, по вертикали – частоты или относительные частоты. Затем форматируем график (тип линии – solidbar – панель заливок).



*Рис. 2.5. Построение гистограммы относительных частот*

**2.1.2. Числовые характеристики вариационного ряда**

*Знание среднего значения*

*это скудная частица информации. ...*

*Трудно понять, почему обычно*

*ограничивают свои исследования средними*

*и не наслаждаются более исчерпывающим взглядом.*

*Они, видимо плохо воспринимают очарования разнообразия.*

Фрэнсис Гальтон (1889)

В теории вероятностей рассматривались числовые характеристики случайных величин. Среди этих характеристики выделялись характеристики положения (математическое ожидание, мода, медиана), характеристики рассеивания (дисперсия, среднее квадратическое отклонение, коэффициент вариации), характеристики формы (асимметрия, эксцесс).

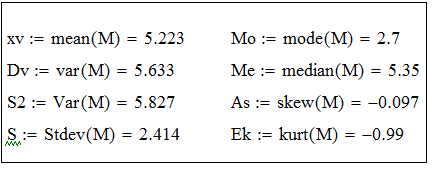
Применительно к выборочным распределениям в математической статистике введены аналогичные характеристики (статистические), значения которых вычисляется по результатам преобразований выборочных данных.

В MathCAD существуют встроенные функции, которые реализуют соответствующие вычисления выборочных числовых характеристик.

|  |  |  |
| --- | --- | --- |
| **характеристика** | **формула** | **встроенная функция**  **MathCAD** |
| средняя арифметическая (выборочная средняя) |  | mean(M) |
| выборочная мода | значение признака, которому соответствует наибольшая частота (существует не всегда) | mode(M) |
| выборочная медиана | значение признака, приходящееся на середину ранжированного ряда наблюдений и зависит от четности объема выборки | median(M) |
| размах вариации |  |  |
| выборочная  дисперсия |  | var(M) |
| исправленная  выборочная  дисперсия  (для |  | Var(M) |
| исправленное среднее квадратическое отклонение |  | Stdev(M) |
| выборочная  асимметрия |  | skew(M) |
| выборочный  эксцесс |  | kurt(M) |

**Пример 2.2.** Найтичисловые характеристики массива(среднее значение, моду, медиану, исправленную дисперсию, исправленное среднее квадратическое отклонение, коэффициенты асимметрии и эксцесса) по данным примера 2.1.

**Решение.** Используем соответствующие встроенные функции.



*Рис. 2.6. Вычисление статистических характеристик*

**2.1.3. Точечное и интервальное оценивание параметров распределения**

При заданном виде закона распределения СВ *Х* неизвестные параметры этого распределения можно оценить. Оценки параметров генеральной совокупности, полученные на основе выборки, называются статистическими.

***Виды статистических оценок***

1. Точечные – оценка характеризуется одним числом.

2. Интервальные – оценка характеризуется двумя числами - концами интервала.

Одним из методов получения точечных оценок является **метод моментов**, предложенный Пирсоном в 1894 г.

**Суть метода**: сопоставление выборочных начальных и центральных моментов с соответствующими теоретическими моментами и . Число уравнений должно соответствовать числу неизвестных параметров:

;

,

|  |  |  |
| --- | --- | --- |
| **Закон** | **Параметры** | **Уравнения** |
| Нормальный | 2 параметра – |  |
| Показательный | 1 параметр – λ |  |
| Равномерный | 2 параметра – |  |
| Биномиальный | 1 параметр – *р* |  |
| Пуассона | 1 параметр – λ |  |

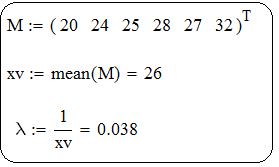
**Пример 2.3.** На предприятии изготавливается определенный вид продукции. Ежемесячный объем выпуска этой продукции является случайной величиной, для характеристики которой принят показательный закон распределения

В течение шести месяцев проводился замер объемов выпуска, продукции, получены следующие данные:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| месяц | 1 | 2 | 3 | 4 | 5 | 6 |
| объем выпуска | 20 | 24 | 25 | 28 | 27 | 32 |

Найти оценку параметра

**Решение.** Так как закон распределения содержит лишь один параметр то для его оценки требуется составить одно уравнение. На рисунке 2.7 приведены результаты оценки параметра в среде MathCAD.



*Рис. 2.7. Оценка параметра*

*показательного распределения*

Пусть точечная оценка, использующая данные выборки, получена. **Точность** такой оценки можно охарактеризовать неравенством

,

причем чем меньше значение , тем выше точность полученной оценки.

**Надежность** (степень доверия) оценки характеризуется величиной , с которой выполняется неравенство

Перейдя к двустороннему неравенству:

.

Вероятность - доверительная вероятность (надежность). В практике решения задач интервального оценивания наибольшее распространение получили значения , равные 0,9; 0,95; 0,99.

Часто вместо доверительной вероятности задается ее дополнение до единицы или **уровень значимости** Уровень значимости определяет вероятность нахождения неизвестного параметра за границами доверительного интервала.

***Алгоритм нахождения границ доверительного интервала***

1. Получить точечную оценку искомого параметра.

2. Задать статистику, содержащую параметр и его точечную оценку таким образом, чтобы функция распределения статистики была известна (точно или приближенно).

3. Задать уровень значимости . Принять вероятности за пределами доверительного интервала слева и справа равные .

4. Определить квантили и распределения случайной величины .

5. Записать неравенство , выразить из него искомый параметр .

Многолетней практикой решения задач математической статистики выработаны рекомендации по заданию статистик в каждой конкретной задаче (например, стандартное нормальное распределение, хи-квадрат, Фишера, Стьюдента).

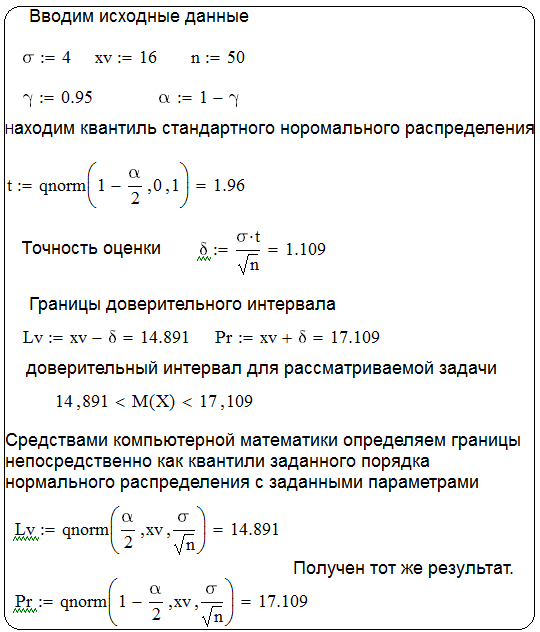
Использование средств компьютерной математики позволяет в некоторых случаях упростить алгоритм.Средствами компьютерной математики можно определять квантили для распределений с произвольными параметрами, поэтому применение статистик в виде стандартного нормального распределения становится необязательным*.*

|  |  |
| --- | --- |
| **Интервальная оценка математического ожидания нормального распределения с известной дисперсией** | |
| С помощью нормированной статистики | Применение средств компьютерной математики |
| Статистика t имеет нормальное распределение  Определяются квантили стандартного нормального распределения и , связанные в данном случае соотношением  .  Квантили можно определить по специальной таблице или с помощью встроенной функции MathCAD **qnorm(,0,1)**.  Доверительный интервал | Статистика z имеет нормальное распределение  Определяются квантили и нормального распределения.  Квантили можно определить с помощью встроенной функции MathCAD  **qnorm(,,)**,  **qnorm(,,)**, которые и являются границами интервала. |
| **Интервальная оценка математического ожидания для случайной величины с произвольным законом распределения с известной дисперсией** | |
| Построение интервальной оценки может осуществляться двумя способами, как и в предыдущем случае по тем же алгоритмам при условии большой выборки | |
| **Интервальная оценка дисперсии (среднего квадратического отклонения) нормального распределения** | |
| Статистика распределена по закону хи-квадрат с (*п*-1) степенями свободы  Определяются квантили  , .  Квантили можно определить по специальной таблице или с помощью встроенной функции MathCAD  **qchisq, qchisq**  Доверительный интервал | |
| **Интервальная оценка дисперсии (среднего квадратического отклонения) произвольно распределенной случайной величины** | |
| Статистика *m* имеет нормальное распределение Определяются квантили стандартного нормального распределения и , связанные в данном случае соотношением  .  Квантили можно определить по специальной таблице или с помощью встроенной функции MathCAD  m **= qnorm(,0,1)**.  Вычисляем параметр , где - выборочные центральные моменты.  Доверительный интервал  Если , то | |
| **Интервальная оценка математического ожидания нормального распределения при неизвестной дисперсии и произвольного распределения (при большом объеме выборки)** | |
| Статистика распределена по закону Стьюдента с (*п*-1) степенями свободы. Определяются квантили t-распределения  и ,  связанные соотношением .  Квантили можно определить по специальной таблице или с помощью встроенной функции MathCAD **qt(,n-1)**.  Доверительный интервал | |

**Пример 2.3.** Найти доверительный интервал с надежностью 0,95 для оценки математического ожидания нормально распределенной СВ *Х*, если известны ее среднее квадратическое отклонение, выборочная средняя и объем выборки *n=*16.

**Решение**

На рисунке 2.8 приведены результаты решения (традиционный подход) и средствами компьютерной математики.



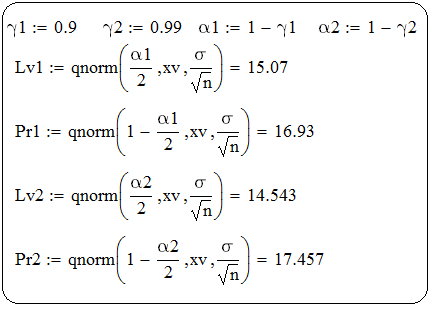
*Рис. 2.8. Решение задачи интервального оценивания*

**Пример 2.4.** Рассмотрим пример 2.3 при

**Решение.** Результаты сведем в таблицу.

|  |  |  |  |
| --- | --- | --- | --- |
| Границы | Значение доверительной вероятности | | |
|  |  |  |
| левая | 15,07 | 14,891 | 14,543 |
| правая | 16,93 | 17,109 | 17,457 |

Анализируя данные, отмечаем тенденцию к расширению границ доверительных интервалов при увеличении значения .



*Рис. 2.9. Вычисление границ интервалов*

**Задание для аудиторного решения**

Обследование оплаты труда 50 рабочих данного завода дало следующие результаты (в условных единицах):

214, 204, 212, 201, 190, 222, 226, 216, 228, 240,

224, 220, 260, 204, 240, 190, 218, 232, 254, 224,

204, 221, 256, 260, 228, 232, 204, 182, 230, 214,

242, 222, 260, 198, 216, 198, 232, 242, 216, 226,

208, 221, 202, 204, 222, 196, 222, 238, 224, 223.

1. Провести статистическую обработку выборочных данных: сформировать массив, составить вариационный ряд, группированный статистический ряд, построить гистограмму оценки функции плотности. Вычислить с помощью встроенных функций числовые характеристики выборки.

2. Провести визуальный сравнительный анализ неизвестного распределения данной СВ *Х* с нормальным распределением. Для этого на одном графике построить гистограмму оценки функции плотности изучаемой СВ *Х* и график функции плотности нормального закона. В качестве параметров нормального закона взять их наилучшие оценки по выборке.

***Замечание.*** *Близость средней арифметической величины, медианы и моды указывает на вероятное соответствие изучаемого распределения нормальному закону.*

3. Найти доверительные интервалы для оценки математического ожидания и дисперсии с надежностью 0,99.

**Компьютерная лабораторная работа № 4**

***Вариационные ряды, их числовые характеристики и***

***графическое изображение***

**Задания к компьютерной лабораторной работе № 4**

**Задание 1.** При заданном виде закона распределения СВ *Х* оценить неизвестные параметры этого распределения.

1. При условии показательного распределения СВ *Х* произведена выборка

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *хi* | 3 | 4 | 10 | 12 | 15 |
| *ni* | 3 | 3 | 6 | 4 | 4 |

Найти оценку параметра λ.

2. При условии равномерного распределения СВ *Х* произведена выборка

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *хi* | 2 | 3 | 4 | 5 | 6 |
| *ni* | 4 | 6 | 5 | 12 | 8 |

Найти оценку параметров *a* и *b*.

3. При условии показательного распределения СВ *Х* произведена выборка

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *хi* | 3 | 5 | 6 | 8 | 10 |
| *ni* | 2 | 3 | 5 | 10 | 10 |

Найти оценку параметра λ.

4. При условии равномерного распределения СВ *Х* произведена выборка

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *хi* | 3 | 5 | 7 | 9 | 11 | 13 | 15 | 17 | 19 | 21 |
| *ni* | 4 | 6 | 5 | 5 | 8 | 6 | 7 | 4 | 8 | 6 |

Найти оценку параметров *a* и *b*.

5. При условии равномерного распределения СВ *Х* произведена выборка

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| *хi* | 3 | 5 | 7 | 9 | 11 | 13 | 15 |
| *ni* | 6 | 9 | 6 | 5 | 7 | 6 | 8 |

Найти оценку параметра

6. СВ *Х* распределена по биномиальному закону. Статистическое распределение выборки представлено в таблице:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *хi* | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| *ni* | 2 | 3 | 10 | 22 | 26 | 20 | 12 | 5 |

Найти оценку параметра *р*.

7. СВ *Х* распределена по закону Пуассона. Статистическое распределение выборки представлено в таблице:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *хi* | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| *ni* | 112 | 87 | 42 | 21 | 9 | 3 | 1 | 1 |

Найти оценку параметра λ.

8. СВ *Х* распределена по биномиальному закону. Статистическое распределение выборки представлено в таблице:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *хi* | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| *ni* | 1 | 3 | 9 | 17 | 20 | 18 | 12 | 5 |

Найти оценку параметра *р*.

9. При условии показательного распределения СВ *Х* произведена выборка

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| *хi* | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| *ni* | 26 | 19 | 16 | 15 | 8 | 6 | 4 |

Найти оценку параметра λ

10. При условии равномерного распределения СВ *Х* произведена выборка

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *хi* | 1 | 3 | 5 | 7 | 9 | 11 | 13 | 15 | 17 | 19 |
| *ni* | 4 | 6 | 5 | 5 | 8 | 6 | 7 | 4 | 8 | 6 |

Найти оценку параметров *a* и *b*.

**Задание 2.** Провести статистическую обработку выборочных данных: сформировать массив, составить вариационный ряд, группированный статистический ряд, построить гистограмму оценки функции плотности. Вычислить с помощью встроенных функций числовые характеристики выборки.

**Задание 3.** Провести визуальный сравнительный анализ неизвестного распределения данной СВ *Х* с нормальным распределением. Для этого на одном графике построить гистограмму оценки функции плотности изучаемой СВ *Х* и график функции плотности нормального закона. В качестве параметров нормального закона взять их наилучшие оценки по выборке.

**Задание 4.** Найти доверительные интервалы для оценки математического ожидания и дисперсии. Надежность выбрать самостоятельно.

**Задание 5.** Используя данные примера 2.3 и образец решения примера 2.4 проведите анализ влияния объема выборки на поведение границ доверительного интервала.

Статистические данные по вариантам представлены ниже.

**1.** 176; 184; 169; 176; 182; 162; 181; 167; 176; 167; 166; 176; 162; 172; 175; 182; 166; 171; 176; 174; 174; 176; 178; 179; 178; 175; 169; 176; 172; 168; 186; 166; 163; 175; 164; 171; 178; 170; 168; 168; 180; 167; 165; 177; 169; 176; 177; 176; 174; 173.

**2.** 333,2; 332,5; 328,3; 325,5; 332,9; 331,5; 320,7; 337,9; 337,2; 340,4, 338,1; 336,2; 327,5, 329,1; 336,8; 337,9; 333,6; 327,9; 328,4; 331,4; 338,6; 349,6; 330,7; 335,6; 335,6; 348,9; 344,8; 307,1; 336,6; 325,6; 325,3; 330,9; 323; 333,5; 321,3; 349; 302,7; 338,7; 351,8; 340,8; 341,8; 320,3; 314,2; 334,6; 344,7; 341,7; 322,4; 335,1; 335; 322.

**3.** 117; 118; 121; 119; 128; 120; 121; 142; 128; 152; 105; 126; 130; 146; 133; 103; 107; 156; 135; 125; 114; 112; 104; 111; 140; 137; 135; 111; 120; 146; 125; 121; 119; 130; 131; 124; 126; 130; 115; 123; 106; 128; 147; 113; 135; 123; 110; 134; 129; 114.

**4.** 61,1; 50,5, 61,1; 44,4, 59,1; 55,2; 60,4; 93,9; 58,6; 51,1; 79,7; 86,4; 32,1; 84,5; 37,8; 33,4; 50,2; 56,6; 72,2; 56,1; 30,1; 59,6; 67; 74,8; 90,7; 48,4; 62,6; 37,4; 10,7; 88,7; 51,4; 25,7; 28,9; 64,3; 69,8; 42,6; 91,2; 71,3; 84,6; 69,2; 35,9; 78,9; 37,8; 59,5; 53,8; 41,7; 27,9; 66,6; 72,2; 55,9.

**5.** 27,39; 19,59; 17,04; 4,06; 12,38; 28,65; 18,62; 17,75; 28,97; 21,84; 20,26; 15,53; 23,3; 14,63; 12,24; 25,9; 7,13; 16,38; 15,34; 17,72; 25,14; 25,33; 13,51; 10,31; 15,03; 11,27; 25,63; 20,63; 24,13; 28,55; 19,7; 24,43; 32,19; 6,82; 16,19; 21,75; 18,06; 21,65; 25,39; 21,85; 16,54; 16,72; 30,89; 12,13; 17,02; 10,22; 16,81; 3,64; 29,77; 11,72.

**6.** 418; 455; 521; 517; 476; 473; 506; 398; 509; 464; 431; 529; 430; 436; 588; 507; 549; 511; 534; 477; 561; 385; 511; 486; 524; 543; 490; 545; 425; 590; 561; 503; 504; 438; 450; 494; 485; 543; 423; 527; 489; 500; 524; 450; 479; 502; 486; 517; 434; 545.

**7.** 100,86; 99,05; 98,61; 100,09; 99,43; 97,44; 101,52; 99,22; 97,59; 97,93; 99,96; 100,06; 100,53; 100,98; 100,48; 98,63; 96,77; 102,82; 101,94; 97,67; 100,29; 103,39; 99,36; 95,94; 97,18; 97,35; 98,97; 100,12; 98,81; 99,34; 101,08; 100,44; 100,81; 97,69; 97,79; 97,79; 99,18; 97,37; 94,32; 102,55; 101,58; 100,29; 97,99; 98,15; 98,24; 100,56; 99,20; 102,60; 97,24; 100.76.

**8.** 50,88;53,52; 48,71; 51,10; 53,09; 46,36; 52,63; 47,91; 50,92; 47,89; 47,52; 50,85; 46,17; 49,75; 50,83; 52,93; 47,54; 49,23; 50,95; 50,37; 50,39; 51,04; 51,65; 51,86; 51,79; 50,52; 48,61; 50,93; 49,70; 48,27; 54,23; 47,53; 46,61; 50,77; 47,14; 49,18; 51,75; 48,86; 48,18; 48,27; 52,19; 47,95; 47,50; 51,25; 48,53; 51,08; 51,33; 51,04; 50,29; 49,90.

**9.** 27,14; 22,62; 21,52; 40,24; 23,57; 18,60; 28,80; 23,04; 18,98; 19,82; 24,90; 25,16; 26,33; 27,46; 26,19; 21,58; 16,93; 32,06; 29,86; 19,22; 25,72; 33,49; 23,40; 14,84; 17,94; 18,37; 22,42; 25,31; 22,02; 23,35; 27,70; 26,09; 27,03; 19,22; 19,47; 19,48; 22,96; 18,43; 10,80; 31,38; 28,96; 25,72; 19,97; 20,37; 20,61; 26,39; 22,99; 31,51; 18,11; 26,90.

**10.** 10,63; 10,50; 9,66; 9,09; 10,59; 10,30; 8,13; 11,59; 11,43; 12,09; 11,62; 11,23; 9,51; 9,83; 11,37; 11,58; 10,72; 9,58; 9,69; 10,27; 11,73; 13,92; 10,14; 11,13; 11,11; 13,78; 12,96; 5,41; 11.33; 9,12; 9,05; 10,18; 8,60; 10,69; 8,25; 13,80; 4,55; 11,74; 14,36; 12,17; 12,35; 8,06; 6,85; 10,93; 12,95; 12,33; 8,47; 11,02.

**2.2. Статистические гипотезы**

*Во время экспериментирования*

*начальная гипотеза никогда не доказывается*

*и не устанавливается,*

*но возможно не одобряется.*

*Можно сказать, что каждый опыт*

*существует лишь для того,*

*чтобы дать фактам возможность*

*проверить начальную гипотезу.*

Рональд Фишер (1935)

**2.2.1. Статистическая гипотеза и общая схема ее проверки**

***Статистической гипотезой***  называется любое предположение о виде или параметрах неизвестного закона распределения.

Проверяемую гипотезу называют **нулевой** . Противоположную ей гипотезу называют **альтернативной, конкурирующей**.

*Нулевая и альтернативная гипотезы – это две возможности выбора, осуществляемого в задачах проверки статистических гипотез.*

***Схема проверки нулевой гипотезы***

1. Рассматривая выборочные данные и учитывая конкретные условия задачи, выдвигают нулевую гипотезу и альтернативную.

2. Так как решение о справедливости нулевой гипотезы принимается на основе выборочных данных, могут возникнуть ошибки двух родов

|  |  |  |
| --- | --- | --- |
| *Гипотеза* | *Принимается* | *Отвергается* |
| верна | правильное решение | ошибка 1-го рода;  вероятность ошибки 1-го рода равна уровню значимости α |
| неверна | ошибка 2-го рода;  вероятность ошибки 2-го рода равна β | правильное решение |

***Замечание.*** *Возможно* ***одновременное*** *уменьшение α и β лишь при увеличении объема выборки.*

Пользуясь терминологией статистического контроля качества продукции:

вероятность α – риск поставщика (по результатам выборочного контроля вся партия, удовлетворяющая стандарту, забракована);

вероятность *β* – риск потребителя (принятие по анализу выборки партии, не удовлетворяющей стандарту).

Применяя юридическую терминологию:

α – вероятность вынесение судом обвинительного приговора невиновному;

*β* – вероятность вынесение судом оправдательного приговора виновному.

Зарождения понятия об ошибках 1-го и 2-го родов можно проследить в следующих записях.

*«Любой скорее оправдает преступника, чем признает виновным невинного».* Аристотель (951b 0).

*«Много лучше освободить десять виновных, чем приговорить к смерти одного невинного».* (Воинский устав России, 1716).

3. Задается уровень значимости α, значение которого обычно находится в интервале (0,001; 0,1). Задание таких значений связано с необходимостью использовать таблицы функций распределений. Применение средств компьютерной математики позволяет находить характеристики распределений при произвольных значениях α.

4. Используя выборочные данные, вводится статистический критерий – некоторая функция *К*. Эти функции подчинены известному закону распределения (хи-квадрат распределение, *t*-распределение, нормальное и др.).

5. Определяется тип критической области и критические точки.

|  |  |
| --- | --- |
|  | **Критическая область** |
|  | **-** левосторонняя |
|  | **-** правосторонняя |
|  | **-** двусторонняя |

6. По выборочным данным вычисляется числовое значение критерия . Применяется правило:

- ***,*** то гипотеза отвергается и принимается альтернативная;

- ***,*** то гипотеза принимается.

*Принятие гипотезы следует расценивать не как раз и навсегда установленный, абсолютно верный содержащийся в ней факт, а лишь как достаточно правдоподобное, не противоречащее опыту утверждение.*

По своему прикладному содержанию статистические гипотезы можно подразделить на несколько основных типов:

- о законе распределения;

- о числовых значениях параметров;

- о равенстве числовых характеристик генеральных совокупностей;

- об однородности выборок (т.е. принадлежности их одной и той же генеральной совокупности).

**2.2.2. Построение теоретического закона распределения по опытным данным. Проверка гипотез о законе распределения**

*Ответственные решения должны приниматься не интуитивно,*

*а на основе предварительных прикидок, математических расчётов.*

*И не случайно именно в наше время*

*отмечается бурный рост математических методов*

*во всех областях практики.*

*Вместо того чтобы «пробовать и ошибаться» на реальных объектах,*

*люди предпочитают делать это на математических моделях.*

*Построение таких моделей, их анализ и вывод рекомендаций –*

*одна из важнейших задач прикладной математики.*Е.С. Вентцель

При решении практических задач (особенно экономических) модель закона распределения в общем случае заранее неизвестна, возникает необходимость выбора модели закона распределения, согласующейся с результатами выборочных наблюдений. Существует много подходов к решению этой задачи. Наибольшее распространение получила схема, разработанная Пирсоном (1900 г.), дополненная Фишером (1924 г.).

***Схема проверки гипотезы по критерию Пирсона***

1**.** По выборке объема *п* строят статистический ряд (дискретный или интервальный) и его графическое изображение – полигон или гистограмму.

2. По данным предыдущих исследований или по предварительным данным выдвигают гипотезу о модели закона распределения СВ *Х* с функцией распределения *F(x)*.

3. Параметры выбранного закона неизвестны, поэтому их заменяют наилучшими оценками по выборке.

4. Используя гипотетическую функцию распределения *F(x)*, определяют теоретические значения вероятностей или

5. Рассчитывают теоретические частоты ,

6**.** Рассчитывают наблюдаемое значение критерия Пирсона

где - эмпирические частоты,

7. Задаваясь уровнем значимости α, находят критическую область (она всегда правосторонняя) Значениеопределяют по специальным таблицам или с помощью встроенной функции MathCAD – **qchisq**(, где *m* - количество промежутков, *r* - число параметров выбранного закона, оцениваемых по выборке.

8. Если *w*, то гипотеза отклоняется; если  *w* **-** проверяемая гипотеза согласуется с выборочными данными и она принимается.

***Замечания.***

*1.Если в каком-нибудь интервале число теоретических частот < 5, необходимо объединить соседние интервалы.*

*2. На практике если решение об отклонении нулевой гипотезы принято при близких значениях и****,*** *перед переходом к проверке другой гипотезы целесообразно повторить проверку с увеличением объема выборки.*

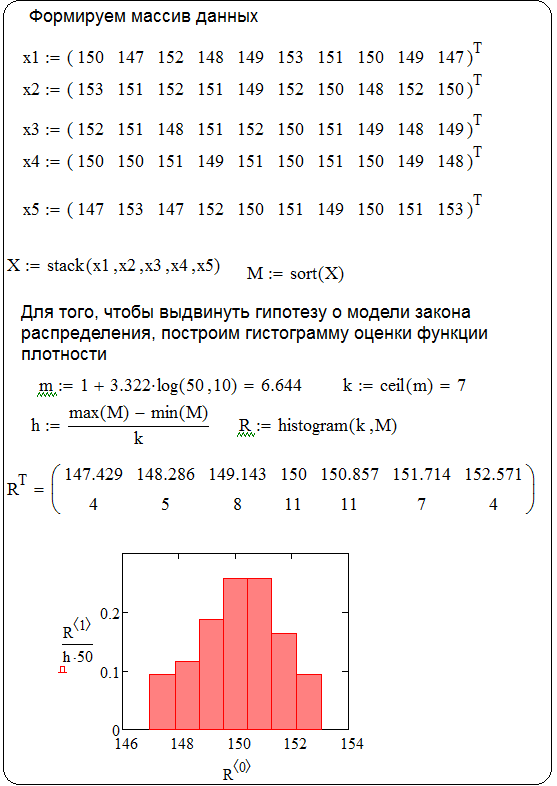
**Пример 2.5.** Результаты взвешивания 50 случайно отобранных пачек чая приведены ниже (в граммах). Оценить закон распределения СВ *Х* – массы пачки чая – для уровня значимости 0,05.

**Решение**

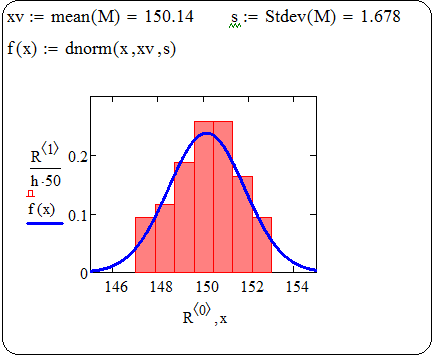
Формирование массива данных и построение гистограммы отображено на рисунке 2.10.

По виду гистограммы и по условию задачи в качестве модели закона распределения выберем нормальный закон, число параметров которого равно 2: *а* – математическое ожидание, σ – среднее квадратическое отклонение. По выборочным данным получим оценки параметров нормального закона распределения:

На одном графике строим гистограмму оценки функции плотности и график плотности распределения нормального закона с найденными параметрами (рисунок 2.11).



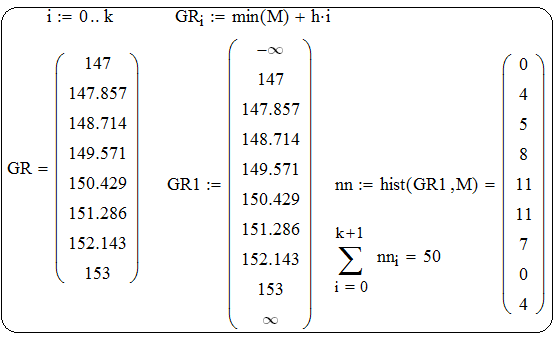
*Рис. 2.10. Построение гистограммы оценки функции плотности*



*Рис. 2.11. Визуальный анализ теоретического закона*

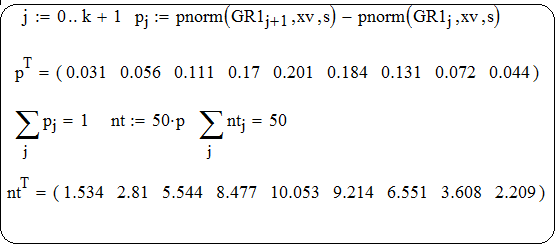
*с эмпирическим*

Определим границы интервалов GR и наблюдаемые частоты (nn).

**

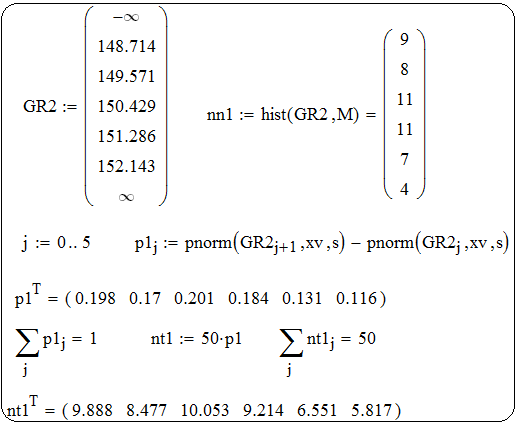
*Рис. 2.12. Предварительный анализ наблюдаемых частот*

Определим вероятности *p* попаданийв отдельные промежутки и теоретические частоты (nt).



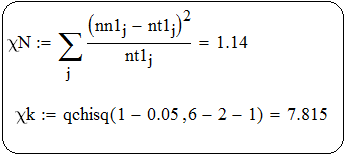
*Рис. 2.13. Предварительный анализ теоретических частот*

Анализ полученных теоретических частот показывает, что условие  *5* будет выполнено, если первые три промежутка объединить в один общий и последние два также в один общий. На рисунке 2.14 обозначены: nn1- наблюдаемые частоты, nt1 – теоретические частоты после объединения.

****

*Рис. 2.14. Объединение частот*

Найдем наблюдаемое значение критерия χN. Результирующий вариант имеет 6 интервалов. Квантиль порядка 1-α =1- 0,05 распределения хи-квадрат определен с помощью встроенной функции **qchisq**.

****

*2.15. Результат решения*

Наблюдаемое значение не входит в критическую область , поэтому гипотеза о том, что СВ *Х* – масса пачки чая – подчинена нормальному закону распределения, согласуется с выборочными данными.

**Пример 2.6.** Дано следующее распределение успеваемости 125 студентов, сдавших три экзамена:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Число сданных экзаменов | 0 | 1 | 2 | 3 |
| Число студентов | 3 | 5 | 47 | 70 |

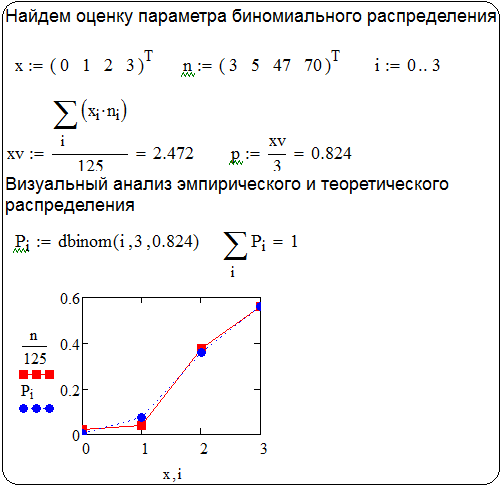
Проверить гипотезу о биномиальном распределении числа сданных экзаменов при уровне значимости 5%.

**Решение**

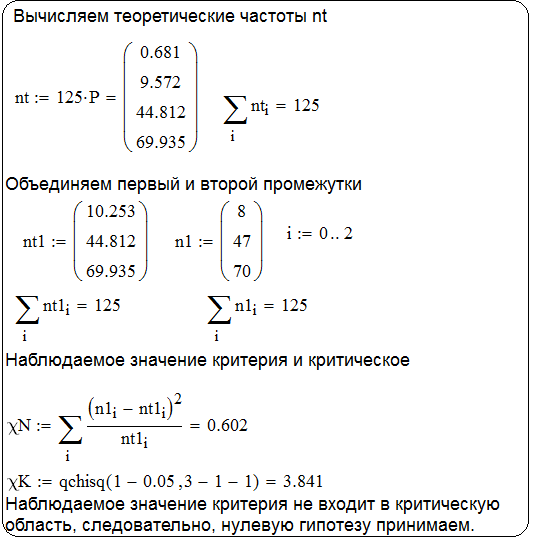
Для проверки гипотезы о биномиальном законе необходимо найти оценку параметра выбранного закона распределения.

Биномиальный закон имеет один параметр – *p*, который найдем из условия

где *п* – число испытаний в каждом опыте (не путать с объемом выборки!).



*Рис. 2.16. Оценка параметра биномиального закона*

**

*Рис. 2.17. Проверка гипотезы о биномиальном*

*законе распределения*

**Выполните задания самостоятельно**

1. Коммерсант предполагает, что объем продаж нового вида продукции в каждой из пяти торговых точек, расположенных в различных районах города, будет одинаков. Фактический объем продаж оказался разным:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Район | 1 | 2 | 3 | 4 | 5 |
| Объем продаж | 105 | 117 | 84 | 111 | 83 |

Оценить, значимы или нет различия между наблюдаемыми и ожидаемыми объемами продаж при уровне значимости 0,01.

2. При принятии на работу фирма предлагает 4 теста. СВ *Х* - число решенных тестов десятью претендентами: 0, 1, 1, 2, 2, 3, 3, 3, 4, 4. Проверить гипотезу о биномиальном распределении СВ *Х* при уровне значимости 0,05.

3. Время безотказной работы 60 элементов: 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 4, 4, 4, 5, 5, 5, 5, 6, 6, 6, 6, 6, 6, 7, 7, 7, 7, 7, 7, 7, 8, 8, 9, 9, 10, 11, 12, 13, 14, 15, 16, 16, 16, 17, 18, 20, 21, 21, 22,23. Проверить гипотезу о показательном законе распределения СВ *Х* – время безотказной работы при уровне значимости 0,05.

**Компьютерная лабораторная работа № 5**

***Проверка гипотезы о виде закона распределения СВ***

***по критерию Пирсона***

**Задания к компьютерной лабораторной работе № 5**

**Задание 1.** Используя результаты заданий 2 и 3 компьютерной лабораторной работы № 4, проверить гипотезу о нормальном законе распределения СВ *Х*.

**Задание 2.** К имеющемуся эмпирическому распределению подобрать теоретический закон распределения. Осуществить проверку по критерию Пирсона.

**1.** 5,009; 10,76; 6,219; 9,973; 7,128; 5,64; 6,031; 11,92; 5,834; 5,062; 8,722; 9,212; 6,164; 8,156; 5,399; 10,483; 8,639; 11,132; 11,691; 8,775; 8,235; 11,036; 10,458; 11,978; 9,28; 6,863; 10,881; 7,631; 9,74; 5,062; 6,931; 9,115; 10,863; 10,206; 8,206; 10,211; 9,193; 10,145; 9,007; 6,061; 7,976; 8,62; 10,261; 6,183; 8,443; 9,898; 6,353; 9,095; 7,452; 8,395.

**2.** 4,493; 5,623; 6,029; 5,033; 6,566; 2,52; 2,427; 4,66; 1,78; 1,755; 8,452; 8,157; 2,819; 4,286; 6,025; 4,613; 5,783; 7,838; 5,998; 5,526; 2,474; 5,441; 2,943; 5.838; 5,677; 4.956; 6,926; 5,963; 7,436; 5,608; 8,292; 6,821; 6,342; 3,52; 3,447; 1,869; 7,81; 2,239; 1,635; 6,128; 5.361; 4,272; 4,724; 2,221; 6,904; 7,614; 7,987; 3,401; 2,018; 7,28.

**3**. 11,974; 2,301; 12,769; 2,887; 5,088; 17,559; 12,93; 9,866; 9,943; 10,167; 13,387; 11,913; 1,106; 2,811; 16,538; 14,452; 7,842; 10,951; 18,203; 4,179; 3,37; 18,131; 1,501; 2,007; 3,37; 16,558; 13,99; 1,264; 13,742; 4,916; 4,041; 7,138; 7,616; 15,443; 10,478; 15,367; 3,607; 8,243; 3,458; 2,191; 11,32; 10,865; 6,604; 15,544; 6,117; 17,114; 14,388; 7,501; 5,113; 5,13.

**4.** 0; 4; 3; 2; 4; 12; 1; 9; 4; 4; 6; 6; 8; 3; 6; 4; 5; 7; 3; 9; 5; 9; 3; 7; 3; 5; 4; 6; 2; 2; 3; 6; 2; 7; 3; 3; 3; 8; 1; 2; 5; 3; 5; 4; 2; 2; 3; 7; 4; 7.

**5.** 13,809; 9,14; 7,736; 11,415; 6,046; 27,683; 28,731; 13,034; 38,794; 39,336; 1,183; 1,856; 24,69; 14,831; 7,752; 13,249; 8,574; 2,615; 7,839; 9,439; 28,188; 9,809; 23,588; 8,393; 8,947; 11,739; 5,002; 7,957; 3,625; 9,193; 1,544; 5,299; 6,73; 19,252; 19,746; 37,005; 2,685; 31,084; 42,232; 7,412; 10,113; 14,902; 12,742; 31,326; 5,064; 3,172; 2,257; 20,062; 34,363; 4,035.

**6.** 4,35; 3,16; 3,28; 3,56; 4,04; 1,63; 4,15; 2,75; 3,03; 1,27; 4,93; 3,46; 2,47; 4,44; 1,97; 1,20; 1,69; 1,88; 4,15; 1,34; 1,55; 3,25; 3,53; 2,82; 2,34; 2,13; 4,49; 4,47; 4,18; 3,54; 3,38; 1,78; 1,93; 1,96; 2,63; 2,09; 2,52; 1,83; 3,42; 2,45; 4,90; 4,58; 4,81; 1,81; 2,42; 2,16; 1,39; 3,84; 3,69; 1,02.

**7.** 0,22; 1,51; 1,93; 0,41; 0,85; 11,25; 0,51; 0,57; 0,22; 5,52; 0,60; 4,18; 6,60; 3,20; 1,15; 3,60; 2,91; 5,29; 0,47; 4,99; 7,74; 14,96; 6,96; 9,66; 2,49; 8,13; 1,36; 2,26; 8,73; 0,19; 8,11; 15,25; 16,62; 12,48; 2,82; 7,60; 5,94; 4,44; 6,90; 3,38; 0,06; 10,53; 1,73; 6,11; 3,22; 3,30; 0,94; 11,33; 0,27; 3,73.

**8.** 1; 2; 1; 1; 2; 3; 3; 2; 4; 0; 1; 2; 3; 3; 1; 2; 6; 3; 2; 0; 3; 2; 3; 2; 4; 1; 0; 1; 4; 2; 1; 6; 2; 2; 3; 3; 1; 2; 1; 2; 4; 0; 7; 2; 2; 4; 2; 5; 2; 2.

**9.** 9,28; 6,863; 10,881; 7,631; 9,74; 5,062; 6,931; 9,115; 10,863; 10,206; 8,206; 10,211; 9,193; 10,145; 9,007; 6,061; 7,976; 8,62; 10,261; 6,183; 8,443; 9,898; 6,353; 9,095; 7,452; 8,395; **.** 5,009; 10,76; 6,219; 9,973; 7,128; 5,64; 6,031; 11,92; 5,834; 5,062; 8,722; 9,212; 6,164; 8,156; 5,399; 10,483; 8,639; 11,132; 11,691; 8,775; 8,235; 11,036; 10,458; 11,978.

**10.** 1,501; 2,007; 3,37; 16,558; 13,99; 1,264; 13,742; 4,916; 4,041; 7,138; 7,616; 15,443; 10,478; 15,367; 3,607; 8,243; 3,458; 2,191; 11,32; 10,865; 6,604; 15,544; 6,117; 17,114; 14,388; 7,501; 5,113; 5,13; 11,974; 2,301; 12,769; 2,887; 5,088; 17,559; 12,93; 9,866; 9,943; 10,167; 13,387; 11,913; 1,106; 2,811; 16,538; 14,452; 7,842; 10,951; 18,203; 4,179; 3,37; 18,131.

**2.2.3. Гипотезы о значениях и сравнении числовых характеристик и однородности выборок**

***Проверка гипотез о числовых значениях параметров***

На практике часто требуется оценить, соответствуют ли действительности рекламные данные о параметрах того или иного товара. В этом случае возникает задача сравнения выборочной средней с анонсируемым значением этого параметра. Аналогичные задачи могут возникнуть, например, в финансовом анализе, когда по данным выборки надо установить, можно ли считать доходность актива равным заданному числу; или по результатам выборочной аудиторской проверки однотипных документов нужно убедиться, можно ли считать процент допущенных ошибок равным номиналу.

При проверке качества функционирования измерительных устройств необходимо проверить гипотезы о равенстве среднего значения *a* и дисперсии определенным числам . Если – номинальное значение измеряемого параметра и , то это означает, что прибор дает систематическую ошибку. Точность прибора определяется значением и, если , то это означает, что качество прибора не отвечает стандартным требованиям. Или, например, что доля бракованных изделий, производимых станком, равна заданной величине .

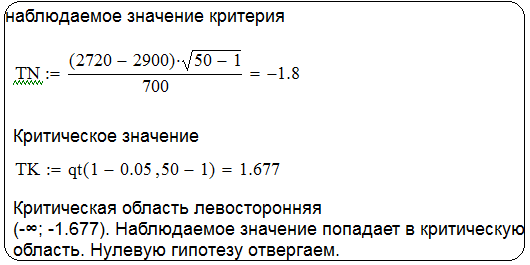
Схемы проверки гипотез о числовых значениях сведем в таблицу.

|  |  |  |  |
| --- | --- | --- | --- |
| Гипотеза | Условие | Статистический критерий | |
|  |  |
|  | неизвестна |  | *qt(1-)* |
|
|
|  | известна |  | (1-) |
|  | –  произвольное |  |  |
|  |  |  | (1-) |

**Пример 2.6.** Фирма-поставщик в рекламном буклете утверждает, что средний срок безотказной работы предлагаемого изделия – 2900 часов. Для выборки из 50 изделий средний срок безотказной работы оказался равным 2720 ч при выборочном среднем квадратическом отклонении 700 ч. При 5%-м уровне значимости проверить гипотезу о том, что значение 2900 ч является математическим ожиданием.

**Решение.** Предположим, что случайная величина срока безотказной работы подчинена нормальному закону распределения. Требуется проверить гипотезу о числовом значении математического ожидания (генеральной средней) при неизвестной генеральной дисперсии. Имеем: .

Другие альтернативные гипотезы нецелесообразны (потребитель обычно обеспокоен лишь тем, что срок службы изделия может оказаться меньше рекламируемого). Исходя из альтернативной гипотезы, критическая область – левосторонняя.



*Рис. 2.16. Проверка гипотезы*

Можно сделать вывод о том, что фирма в рекламе завышает срок безотказной работы изделия.

***Проверка гипотез о равенстве числовых характеристик двух совокупностей***

На практике часто встречаются ситуации, когда среднее значение

данных одного эксперимента отличается от среднего значения данных другого, хотя условия эксперимента являются схожими. Тогда возникает вопрос, можно ли считать это расхождение незначимым, т.е. чисто случайным, или оно вызвано существенным различием двух генеральных совокупностей. Например, такие вопросы возникают при исследовании надежности технических систем, где результаты сравниваются с предыдущими измерениями; при контроле качества изделий, изготовленных на разных предприятиях; в финансах – при сравнении уровня доходности различных активов.

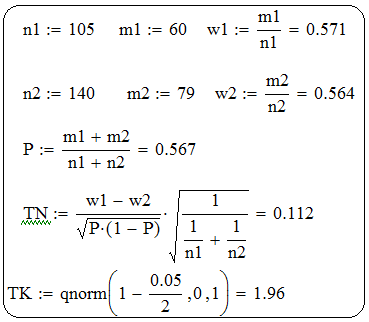
Гипотезы о равенстве дисперсий возникают часто, поскольку дисперсия характеризует такие важные показатели, как точность приборов, технологических процессов, риск, связанный с отклонением доходности от заданного уровня, и т.д.

Сравнение долей признака в двух совокупностях является важной для практики задачей, так как значимое расхождение между долями признаков характеризует различие между генеральными совокупностями. Схемы проверки гипотез о равенстве числовых характеристик сведем в таблицу, в которой альтернативная гипотеза выбирается двусторонней.

|  |  |  |  |
| --- | --- | --- | --- |
| Гипотеза | Условие | Статистический критерий | |
|  |  |
|  | Дисперсии  и известны |  | *qnorm* |
|
|
|  | Дисперсии неизвестны, но равны |  |  |
|  |  | (в числителе  должно быть большее значение) |  |
|  |  |  | *qnorm* |

**Пример 2.7.** Контрольную работу по математике по индивидуальным вариантам выполняли студенты двух групп первого курса. В первой группе было предложено 105 задач, из которых верно решено 60, во второй группе из 140 предложенных задач верно решено 80. Проверить гипотезу об отсутствии существенных различий в усвоении математики студентами обеих групп.

**Решение.** Имеем гипотезу , т.е. доли решенных задач студентами первой и второй групп равны. В качестве альтернативной возьмем гипотезу . При конкурирующей гипотезе выбираем двустороннюю критическую область.



*Рис. 2.17. Проверка гипотезу о равенстве долей*

Критическая область . Так как TN (наблюдаемое значение критерия) не входит в критическую область, то нулевую гипотезу принимаем.

***Проверка гипотез об однородности выборок***

В практике статистической обработки данных существуют задачи, связанные с необходимостью объединения нескольких массивов выборочных данных, полученных в различное время, из разнородных источников, от различных технических средств. Необходимость объединения диктуется желанием получить более полную информацию, повысить достоверность оценок.

*Необходимым условием однородности выборок служит равенство или достаточная близость математических ожиданий и дисперсий*.

***Критерий***  ***для проверки гипотезы об однородности выборок***

1. Объединить выборки в один массив.

2. Для полученного массива рассчитать число интервалов, шаг, получить интервальный вариационный ряд.

3. При анализе однородности двух выборок, статистика имеет вид

- число элементов первой и второй выборок, попавших в j-й интервал;

- общее число элементов двух выборок, попавших в j-й интервал;

*-* объемы первой и второй выборок, общий объем выборки.

4. Найти наблюдаемое значение критерия по специальной таблице или использовать встроенную функцию qchisq(1-α,k-1), k – число интервалов.

5. Задаваясь правосторонней критической областью, сделать вывод.

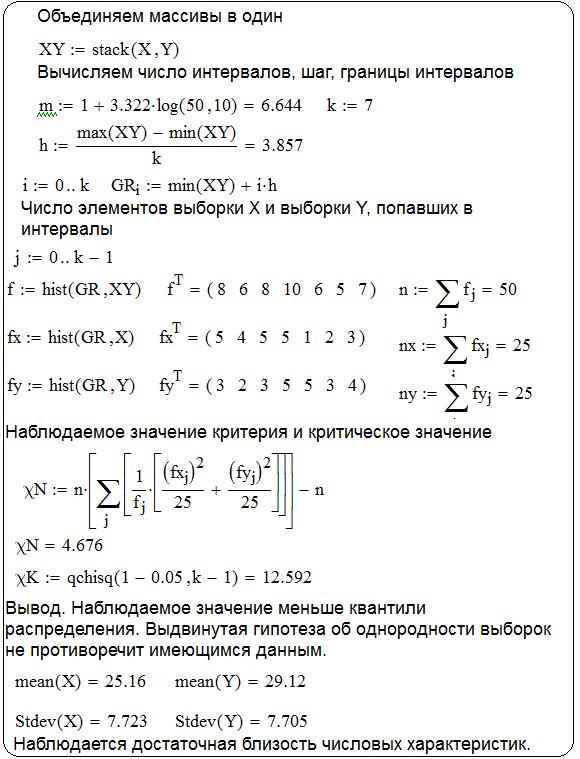
**Пример 2.8**. Из выборочных данных образованы две выборки равного объема. Проверить однородность выборок, используя критерий .

**Решение**

В среде MathCAD перечисленные пункты выполнены с применением функций stack (объединение массивов) и hist (определение частот, попавших в данный интервал). Получено 7 интервалов с границами, составляющими вектор GR.

Выборочное значение получилось меньше значения квантили распределения уровня 0,05 с k-1 степенями свободы.

Ввод выборок *X* и *Y* на рис. 2.18 не приводится.



*Рис. 2.18. Проверка однородности выборок по критерию*

**Задачи для аудиторного решения**

1. Акционерное общество (АО) выпускает печенье «Русские узоры» в пачках, на которых написано: масса нетто 200 г. Осуществлена выборка для оценки средней массы печенья в пачках, выпущенных московской и санкт-петербургской фабриками АО. Результаты выборок таковы (указана масса пачек печенья «Русские узоры»):

*Московская фабрика*

201, 195, 197, 199, 202, 198, 199, 203, 195, 196, 198, 199, 194, 203, 195, 202, 197

*Санкт-петербургская фабрика*

203, 207, 191, 193, 197, 201, 196, 192, 194, 195, 198, 196.

Определить:

а) средние выборочные и исправленные средние квадратические каждой фабрики;

б) для 0,05 значимо или нет различие между средними выборочными (сначала проверить гипотезу );

в) является ли величина 200 г математическим ожиданием массы при 5%-м уровне значимости?

Можно ли считать выборки однородными?

**Компьютерная лабораторная работа № 6**

***Статистические гипотезы***

**Задание 1.**

1.1.Из большой партии ананасов одного размера случайным образом отобрано 36 штук. Выборочная средняя масса одной штуки при этом оказалась равной 930 г. Проверить гипотезу, что средняя масса одного ананаса (по утверждению поставщика) составляет 1 кг, если среднее квадратическое отклонение неизвестно, а выборочное составляет 250 г.

1.2.По утверждению руководства фирмы, средний размер дебиторского учета равен 187,5 тыс. руб. Ревизор составляет случайную выборку из 10 счетов и обнаруживает, что средняя арифметическая выборка равна 175 тыс. руб. при среднем квадратичном отклонении 35 тыс. руб. Может ли оказаться в действительности правильным объявленный размер дебиторского счета?

1.3.Автомат, работающий со стандартным отклонением г, фасует чай в пачки со средним весом г. В случайной выборке объема *n* = 25 пачек средний вес г. Надо ли отрегулировать автомат?

1.4. Автомат, работающий со стандартным отклонениемг, фасует чай в пачки со средним весом *a =* 80 г. В случайной выборке объема *n =* 16 пачек средний вес г. Надо ли отрегулировать автомат?

1.5.Станок, работающий со стандартным отклонением 0,5 мм, производит детали средней длины *a =* 20 мм. В случайной выборке объема *n =* 16 деталей средняя длина19,8 мм. Правильно ли настроен станок?

1.6. Станок, работающий со стандартным отклонением 0,4 мм, производит детали средней длины *a =* 30 мм. В случайной выборке объема *n =* 25 деталей средняя длина 30,1 мм. Правильно ли настроен станок?

1.7.Производитель утверждает, что средний вес пачки чая не меньше *a* = 100 г. Инспектор отобрал 10 пачек чая и взвесил. Их вес оказался 97, 102, 103, 98, 96, 105, 98, 100, 101, 99 г соответственно. Не противоречит ли это утверждению производителя? Предполагается, что вес пачек чая распределен нормально.

1.8. Производитель утверждает, что средний вес плитки шоколада не меньше *a* = 50 г. Инспектор отобрал 10 плиток шоколада и взвесил. Их вес оказался 49, 50, 51, 52, 48, 47, 49, 52, 48, 51 г соответственно. Не противоречит ли это утверждению производителя? Предполагается, что вес плитки шоколада распределен нормально.

1.9. Поставщик двигателей утверждает, что средний срок их службы равен 800 ч. Для выборки из 17 двигателей средний срок службы оказался равным 865 ч при выборочном среднем квадратическом отклонении 120 ч. Проверить нулевую гипотезу при уровне значимости 5%.

1.10.Фирма-изготовитель женских украшений, выпустив новый товар, утверждает, что 40% покупателей купят эти украшения. В ходе 10-дневной рекламной распродажи в среднем приобрели украшения 29,5% покупателей, выборочное среднее квадратическое отклонение составило 16,5%. При 5%-м уровне значимости оценить утверждение изготовителя товара.

**Задание 2.**

2.1.Точность работы станка-автомата проверяется по дисперсии контролируемого размера изделий, которая не должна превышать 0,15. По данным из 25 отобранных изделий вычислена несмещенная дисперсия . Выяснить, обеспечивает ли станок требуемую точность.

2.2.Партия изделий принимается в том случае, если вероятность того, что изделие окажется соответствующим стандарту, будет не менее 0,95. Среди случайно отобранных 100 изделий оказалось 98 соответствующих стандарту. Можно ли принять партию?

2.3.Производитель утверждает, что доля бракованных изделий не превосходит 3 %. В случайной выборке объема *n* = 100 изделий оказалось 5 бракованных изделий. Не противоречит ли это утверждению производителя?

2.4. Производитель утверждает, что доля бракованных изделий не превосходит 7 %. В случайной выборке объема *n =* 150 изделий оказалось 16 бракованных изделий. Не противоречит ли это утверждению производителя?

2.5.Точность работы станка-автомата проверяется по дисперсии контролируемого размера изделий, которая не должна превышать 0,12. По данным из 35 отобранных изделий вычислена несмещенная дисперсия . Выяснить, обеспечивает ли станок требуемую точность.

2.6.Партия изделий принимается в том случае, если вероятность того, что изделие окажется соответствующим стандарту, будет не менее 0,92. Среди случайно отобранных 150 изделий оказалось 108 соответствующих стандарту. Можно ли принять партию?

2.7.Точность работы станка-автомата проверяется по дисперсии контролируемого размера изделий, которая не должна превышать 0,14. По данным из 25 отобранных изделий вычислена несмещенная дисперсия . Выяснить, обеспечивает ли станок требуемую точность.

2.8.Партия изделий принимается в том случае, если вероятность того, что изделие окажется соответствующим стандарту, будет не менее 0,97. Среди случайно отобранных 120 изделий оказалось 112 соответствующих стандарту. Можно ли принять партию?

2.9.Производитель утверждает, что доля бракованных изделий не превосходит 2 %. В случайной выборке объема *n* = 100 изделий оказалось 4 бракованных изделий. Не противоречит ли это утверждению производителя?

2.10. Производитель утверждает, что доля бракованных изделий не превосходит 5 %. В случайной выборке объема *n =* 150 изделий оказалось 16 бракованных изделий. Не противоречит ли это утверждению производителя?

**Задание 3.**

3.1.Было произведено измерений диаметра вала (в мм). При этом оказалось, что среднее , а стандартное среднее квадратичное отклонение . Затем вал поместили в условия с высокой температурой и провели еще измерений диаметра его оси. Среднее на этот раз оказалось равным, а стандартное отклонение . Можно ли сделать вывод, что диаметр вала существенно увеличивается при увеличении температуры?

3.2.Расходы сырья и на единицу продукции по старой и новой технологиям приведены в таблице.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | По старой технологии | | | | По новой технологии | | | | |
| Расход  сырья |  | 304 | 307 | 308 |  | 303 | 304 | 306 | 308 |
| Число  изделий |  | 1 | 4 | 4 |  | 2 | 6 | 4 | 1 |

Предполагается, что генеральные совокупности *X* и *Y* имеют нормальные распределения с одинаковыми дисперсиями и средними и . Требуется проверить гипотезу против .

3.3.Контрольную работу по высшей математике выполняли студенты двух групп. В первой группе было предложено 100 задач, из которых были правильно решены 58, во второй группе из 120 задач верно решены 65. Проверить гипотезу о том, что материал одинаково усвоен студентами обеих групп.

3.4.Инвестиция 1 рассчитана на лет, выборочная дисперсия ежегодных прибылей составляет 20 %. Инвестиция 2 рассчитана на лет, выборочная дисперсия ежегодных прибылей составляет 30 %. Предполагается, что распределение ежегодных прибылей на инвестиции подчиняется нормальному закону распределения. Равны ли риски 1 и 2?

3.5.Инвестиция 1 рассчитана на 14 лет, выборочная дисперсия ежегодных прибылей составляет 15 %. Инвестиция 2 рассчитана на =12 лет, выборочная дисперсия ежегодных прибылей сосатвляет 20 %. Предполагается, что распределение ежегодных прибылей на инвестиции подчиняется нормальному закону распределения. Раны ли риски 1 и 2?

3.6.На двух токарных станках обрабатываются втулки. Отобраны две пробы: из втулок, сделанных на первом станке, 15 шт., на втором станке -шт. По данным этих выборок рассчитаны выборочные дисперсии - 8,5 (для первого станка) и 6,3 (для второго станка). Полагая, что размеры втулок подчиняются нормальному закону распределения, выяснить, можно ли считать, что станки обладают различной точностью.

3.7.Автомат 1 и автомат 2 фасуют чай в пачки. Стандартные

отклонения г и г соответственно. В случайной выборке объема пачек для автомата 1 средний вес =101 г. В случайной выборке пачек для автомата 2 средний вес =98 г. Верно ли, что оба автомата фасуют чай в пачки одинакового среднего веса?

3.8. Проводились испытания нового лекарства. В эксперименте участвовали 3000 мужчин и 3500 женщин. У 50 мужчин и 110 женщин наблюдались побочные эффекты. Можно ли утверждать, что побочные эффекты от нового лекарства у женщин возникают чаще, чем у мужчин?

3.9.Контрольную работу по математике по индивидуальным вариантам выполняли студенты двух групп первого курса. В первой группе было предложено 105 задач, из которых верно решено 60, во второй группе из 140 предложенных задач верно решено 69. Проверить гипотезу об отсутствии существенных различий в усвоении математики студентами обеих групп.

3.10**.** Проводились испытания нового лекарства. В эксперименте участвовали 3500 мужчин и 3000 женщин. У 150 мужчин и 120 женщин наблюдались побочные эффекты. Можно ли утверждать, что побочные эффекты от нового лекарства у мужчин возникают чаще, чем у женщин?

**2.3. Основы регрессионного и корреляционного анализа**

*Числа не управляют миром,*

*но показывают, как управляется мир.*И. Гёте

**2.3.1. Функциональная, статистическая и корреляционная зависимости**

***Функциональная зависимость*** – это зависимость, при которой каждому значению одной (или нескольких) переменной соответствует определенное значение другой переменной (зависимой).

Пусть наблюдению подлежит случайная величина *Y*, зависящая от одной или нескольких других случайных величин *X1*, *X2*, …, *Xk*, которые называются **факторами**. В общем случае число факторов может быть неизвестно. Исследователь выбирает *k* наиболее существенных факторов. В этих условиях функциональная зависимость между *Y* и *X* недостижима, так как неучтено влияние неопределенных факторов, т.е. или , где - стохастическая переменная, включающая влияние неучтенных факторов в модели. Говорят, что между *Y* и *X* существует ***стохастическая (или статистическая, вероятностная) связь.*** Пример статистической связи – зависимость урожайности от количества внесенных удобрений и механизации предприятия. В силу неоднозначности статистической зависимости для исследователя представляет интерес усредненная по *Х* схема зависимости.

***Корреляционной зависимостью*** между двумя переменными величинами называется функциональная зависимость между значениями одной из них и условным математическим ожиданием другой.

Корреляционная зависимость может быть представлена в виде уравнения, которое называется модельным уравнением регрессии (или просто уравнением регрессии).

***Задача корреляционного анализа*** – выявление связи между случайными переменными и оценка ее тесноты.

***Задача регрессии*** - выбор модели зависимости междупеременными и определение оценок неизвестных параметров этой модели.

Выбор модели регрессионных зависимостей осуществляется исходя из теоретических представлений о возможной взаимосвязи между переменнымиили из визуального анализа графиков наблюдений.

В зависимости от количества включенных в модель факторов *Х* модели делятся на *однофакторные* (парная модель регрессии) и *многофакторные* (модель множественной регрессии).

В зависимости от вида функции модели делятся на *линейные* и *нелинейные*.

**2.3.2. Линейная регрессия**

|  |  |
| --- | --- |
| ***Однофакторная регрессия***  (на наблюдаемую переменную *Y* влияет один фактор *X*)  ***Замечание***. *Уравнение называется простой линейной регрессией или парной линейной регрессией.* | ***Множественная регрессия***  (на наблюдаемую переменную *Y* влияют несколько факторов  ) |
| ***Для оценки неизвестных параметров β применяют***  ***метод наименьших квадратов (МНК)***,  суть которого состоит в минимизации суммы квадратов отклонений фактических значений результатного признака от его расчетных значений , т.е.: | |
| ***Алгоритм МНК в форме обобщенного обращения матрицы*** | |
| 1. Ввести исходные данные – массивы *Y* и *X*.  2. Составить матрицу , *n* -число наблюдений, 2- число неизвестных параметров  . | 1. Ввести исходные данные – массивы *Y* и .  2. Составить матрицу , *n* -число наблюдений, *р+1*- число неизвестных параметров  . |
| 3. Матрица неизвестных параметров | |
| В среде MathCAD для определения параметров простой линейной регрессии можно использовать встроенные функции  1.  ;  *2. line(X,Y)* |  |
| ***Качество уравнения регрессии*** ***определяется по величине средней***  ***ошибки аппроксимации***  (уравнение можно использовать как прогностическую модель, если ). | |
| ***Влияние совокупности факторов на результат Y*** | |
| ***Выборочный линейный коэффициент корреляции*** *(характеризует степень взаимосвязи пары случайных величин, если зависимость между ними соответствует прямой линии)*  – линейная функциональная связь, ;  - *Y* и *X* некоррелированы.  В среде MathCAD используют встроенную функцию *corr(X,Y).* | ***Выборочный сводный коэффициент корреляции*** *(характеризует связь Y со всеми факторами, входящими в уравнение)*  - *Y* имеет функциональную связь с совокупностью факторов;  - *Y* некоррелирован ни с одним из факторов. |
| ***Проверка значимости выборочного коэффициента корреляции***  ***(t-критерий Стьюдента)***  Н0 – изучаемый фактор (факторы) не оказывает существенного влияния на результат, т.е. коэффициент корреляции генеральной совокупности равен 0.  Н1 - коэффициент корреляции генеральной совокупности отличен от 0.  *p* – число факторов, влияющих на результат,  *п* – число измерений,  *w* – критическая область двусторонняя | |
|  |  |
| ***Проверка значимости уравнения регрессии (F-критерий Фишера)*** – *установить, соответствует ли* *математическая модель экспериментальным данным и достаточно ли включенных в уравнение факторов (одного или нескольких) для описания зависимой переменной.*  - уравнение регрессии не надежное;  - уравнение регрессии надежное  *p* – число факторов, влияющих на результат,  *п* – число измерений,  *w* – критическая область левосторонняя. | |
|  |  |
| ***Значимость отдельных (кроме свободного члена) коэффициентов регрессии (t-критерий Стьюдента)***  – коэффициент статистически не значим;  – коэффициент статистически значим  *p* – число факторов, влияющих на результат,  *п* – число измерений,  *w* – критическая область двусторонняя.  *Если коэффициент статистически не значим, то фактор, соответствующий этому коэффициенту следует исключить из модели (при этом ее качество не ухудшится).* | |
| ***Коэффициенты эластичности и детерминации***  1***.***Коэффициент эластичности показывает, на сколько процентов в среднем изменяется результативный признак *Y* при изменении факторного признака *Xi* на 1%. Высокий уровень эластичности означает сильное влияние независимой переменной на объясняемую переменную.  2. Коэффициент детерминации ( показывает долю вариации результативного признака, объясненную вариацией факторного признака. Чаще всего, давая интерпретацию коэффициента детерминации, его выражают в процентах. | |

***Особенности практического применения***

***линейных множественных регрессионных моделей***

Одним из условий регрессионной модели является предположение о линейной независимости объясняющих переменных, т. е., решение задачи возможно лишь тогда, когда столбцы и строки матрицы ис­ходных данных линейно независимы. Для экономических показате­лей это условие выполняется не всегда.

Под ***мультиколлинеарностью*** понимается высокая взаимнаякоррелированность объясняющих переменных (факторов), которая приводит к линейной зависимости нормальных уравнений. Существует несколько способов для определения наличия или отсутствия мультиколлинеарности. Один из подходов заключается в анализе коэффициентов парной корреляции.

1. Факторные признаки, у которых исключают из модели.

2. Считают явление мультиколлинеарности в исходных данных установленным, если коэффициент парной корреляции между двумя переменными (факторами) больше 0,8. В этом случае одну переменную исключают из рассмотрения. При этом какую пе­ременную оставить, а какую удалить из анализа, решают в первую очередь на основании экономических соображений. Если с эконо­мической точки зрения ни одной из переменных нельзя отдать предпочтение, то оставляют ту из двух переменных, которая имеет больший коэффициент корреляции с зависимой переменной.

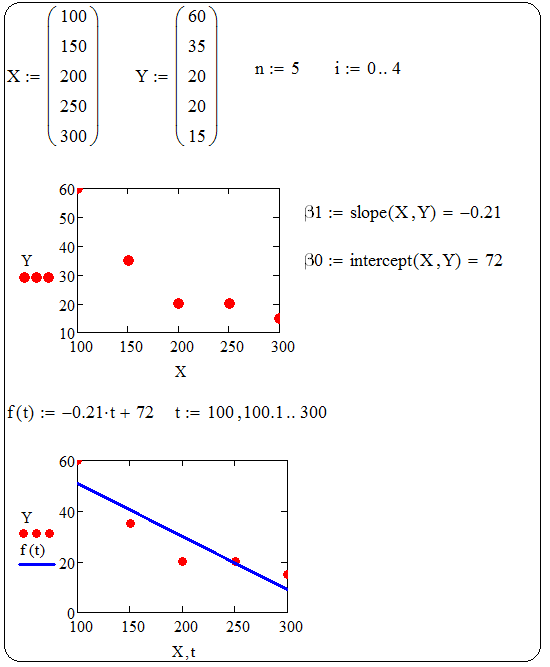
**Пример 2.9.** С целью анализа взаимного влияния зарплаты и текучести рабочей силы на пяти однотипных фирмах проведены измерения уровня месячной зарплаты *X* и числа уволившихся за год рабочих *Y:*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *X* | 100 | 150 | 200 | 250 | 300 |
| *Y* | 60 | 35 | 20 | 20 | 15 |

Найти линейную регрессию *Y* на *X.* Провести статистический анализ.

**Решение**

На рис. 2.19 решена задача простой линейной регрессии. Для нахождения коэффициентов линейной регрессии использованы встроенные функции ***slope, intercept***. Проведен сравнительный визуальный анализ расположения экспериментальных точек и полученной функциональной зависимости.

****

*Рис. 2.19 Однофакторная линейная регрессия*

Проведем статистический анализ (рис. 2.20).

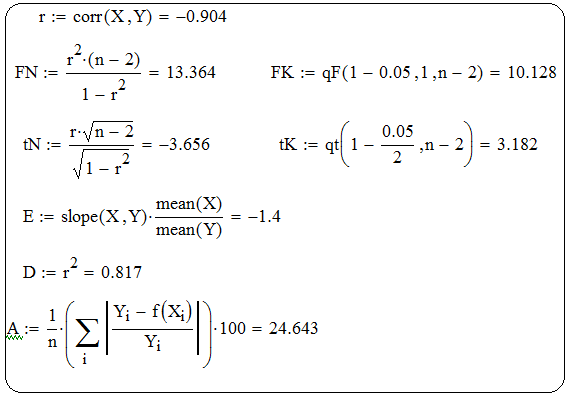
1. Коэффициент корреляции равен (-0,904), следовательно, связь тесная и обратная. Коэффициент корреляции значим, т.е. зарплата оказывает существенное влияние на текучесть рабочей силы.

2. Полученное уравнение регрессии является значимым, т.е. для описания зависимой переменной достаточно одного фактора.

3. Коэффициент эластичности равен (-1,4), т.е. при увеличении зарплаты на 1% текучесть рабочей силы уменьшается в среднем на 1,4%.

4. Коэффициент детерминации равен 0,817, т.е. на 81,7% текучесть рабочей силы зависит от заработной платы и на 18,3% от других неучтенных факторов.

5. Ошибка аппроксимации составляет 24,6%. Следовательно, полученную функциональную зависимость не следует использовать для прогнозирования. Рекомендуется подобрать другую однофакторную модель регрессии (нелинейную).

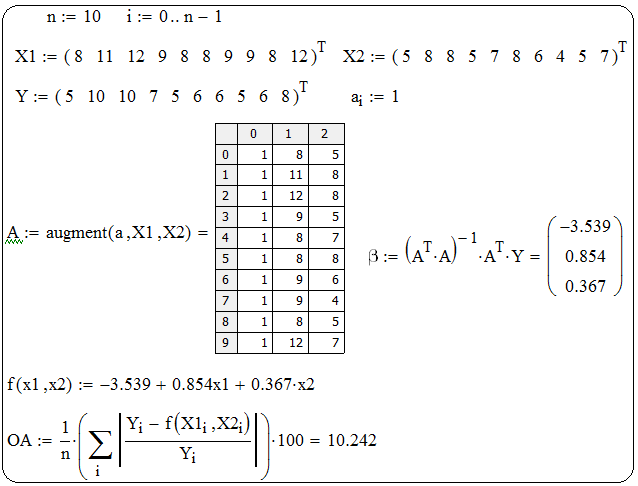


*Рис. 2.20. Статистический анализ*

**Пример 2.10.** Имеются следующие данные о сменной добыче угля на одного рабочего *Y* (т), мощности пласта *X*1(м) и уровне механизации работ *X2* (%), характеризующие процесс добычи угля в 10 шахтах. Предполагая, что между переменными *Y, X*1 и *X*2существует линейная корреляционная зависимость, найти ее аналитическое выражение (уравнение регрессии *Y* по *X*1 и *X*2*)*. Провести статистический анализ.

**Решение**

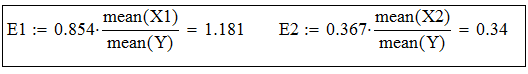
1.Матрицу *А* формируем объединением столбцов с элементами 1, *Х*1, *Х*2 с применением встроенной функции ***augment***. По алгоритму получаем оценки параметров β, что позволяет составить уравнение двухфакторной линейной регрессии. Ошибка аппроксимации составляет 10,24%, что позволяет использовать полученное уравнение в прогностических целях.



*Рис. 2.21. Вариант двухфакторной линейной регрессии*

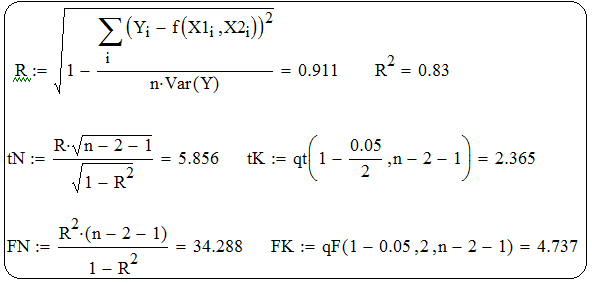
Уравнение множественной регрессии показывает, что при увеличении только мощности пласта *Х*1 (при неизменном *Х*2) на 1 м, добыча угля на одного рабочего увеличивается в среднем на 0,854 т, а при увеличении только уровня механизации работ *Х*2 (при неизменном *Х*1) – в среднем на 0,367 т.

2. Проведем сравнение раздельного влияния на добычу угля двух факторов – мощности пласта *Х*1 и уровня механизации работ *Х*2. Для этого вычислим коэффициенты эластичности.

  
*Рис. 2.22. Коэффициенты эластичности*

Увеличение факторов на 1% (от своих средних значений) приводит в среднем к росту добычи угля соответственно на 1,18% и 0,34%. Таким образом, на добычу угля большее влияние оказывает фактор *Х*1 (мощность пласта) по сравнению с фактором *Х*2 (уровень механизации).

3. Определим множественный коэффициент корреляции и проверим его значимость и значимость полученного уравнения регрессии.



*Рис. 2.23. Вычисление множественного коэффициента*

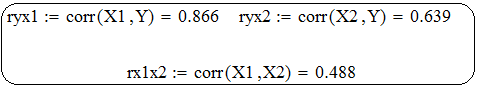
*корреляции*

Значение , близкое к 1, указывает на тесную взаимосвязь переменной *Y* и факторов *Х*1 и *Х*2. Коэффициент детерминации свидетельствует о том, что вариация исследуемой зависимой переменной на 83% объясняется изменчивостью включенных в модель факторов.

Проверим значимость коэффициента корреляции. Наблюдаемое значение tN принадлежит критической области, следовательно, гипотезу о равенстве нулю коэффициента корреляции генеральной совокупности отвергаем.

Проверим значимость уравнения регрессии. Наблюдаемое значение FN принадлежит критической области, следовательно, гипотезу «уравнение регрессии ненадежно» отвергаем.

4. Мультиколлинеарность.



*Рис. 2.24. Парные коэффициенты корреляции*

Так как , то явление мультиколлинеарности не установлено;, следовательно, факторы не исключаются из модели.

**Задачи для аудиторного решения**

1. Дана корреляционная таблица зависимости между суточной выработкой продукции *Y* (т) и величиной основных производственных фондов *Х* (млн руб.) для 50 однотипных предприятий.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Величина ОПФ, млн руб. (*Х*)** | **Суточная выработка продукции, т (*Y*)** | | | | |
| **7-11** | **11-15** | **15-19** | **19-23** | **23-27** |
| **20-25** | 2 | 1 | - | - | - |
| **25-30** | 3 | 6 | 4 | - | - |
| **30-35** | - | 3 | 11 | 7 | - |
| **35-40** | - | 1 | 2 | 6 | 2 |
| **40-45** | - | - | - | 1 | 1 |

Найти параметры простой линейной регрессии. Провести статистический анализ.

2. В процессе тренировок многоборцев анализировались причины промахов спортсменов на этапе стрельбы. Для десяти спортсменов перед стрельбой фиксировались: сила бокового ветра (фактор *Х*1), отклонение частоты пульса спортсменов от 80 в течение двух минут (фактор *Х*2) и максимальное отклонение попаданий (см) в серии выстрелов (*Y*).

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *Х*1 | -1,49 | -0,45 | -2,77 | -2,2 | -2,78 | 2,18 | 2,03 | -0,6 | -2,23 | -0,6 |
| *Х*2 | -0,65 | -1,12 | 0,77 | 6,98 | -1,19 | 8,61 | 3,86 | -8,27 | -0,47 | 9,08 |
| *Y* | -0,1 | -0,33 | -0,64 | 0,42 | -0,75 | 1,57 | 0,75 | -0,7 | 0,82 | 0,84 |

Найти параметры множественной линейной регрессии. Провести статистический анализ.

**2.3.3. Нелинейная однофакторная регрессия**

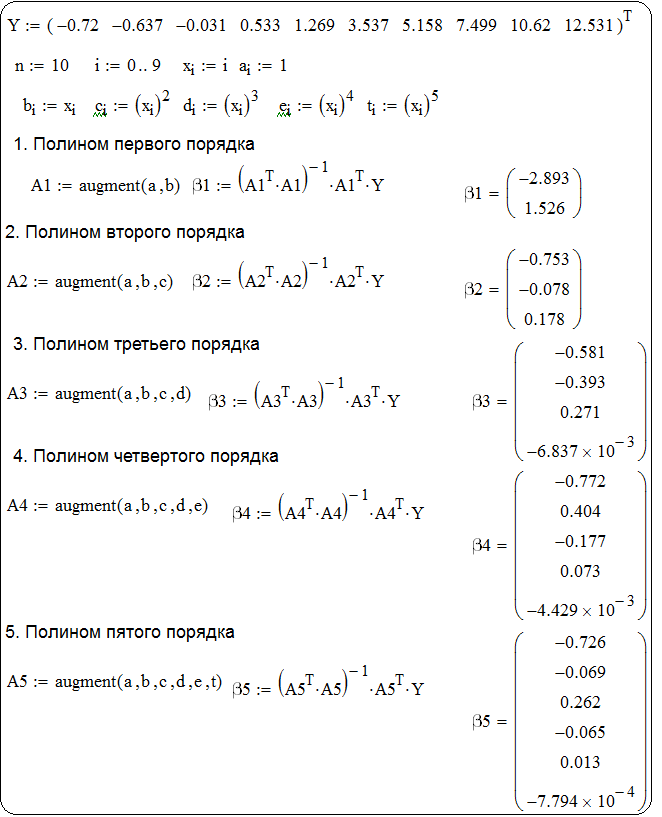
|  |  |
| --- | --- |
| ***Регрессии, нелинейные относительно включенных факторов, но линейные по оцениваемым параметрам*** | ***Регрессии, нелинейные по оцениваемы параметрам*** |
| Полиномы разных степеней, например,  .  ***Заменяя переменные, например, , получим уравнение множественной линейной регрессии.*** | 1.Внутренне линейные, например, функции:  степенная  показательная  экспоненциальная .  ***С помощью соответствующих преобразований приводятся к простой линейной регрессии.***  2. Внутренне нелинейные, например, функции:  ***Используют итеративные процедуры.*** |

**Пример 2.11.** В контрольный период проведено 10 измерений некоторой величины. Измерения проводились через фиксированный интервал времени. Результаты наблюдений *Y* сведены в таблицу. Решить задачу регрессии, используя в качестве зависимостей полиномы с 1 по 5 степеней. Сопоставить точности вариантов.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| № | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| *Y* | -0,72 | -0,637 | -0,031 | 0,533 | 1,269 | 3,537 | 5,158 | 7,499 | 10,62 | 12,531 |

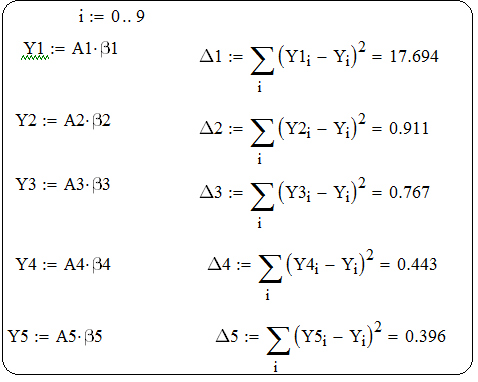
**Решение**

На рис. 2.25 приведены результаты решения в среде MathCAD. Решения для полиномов разных степеней проведены по одному алгоритму.



*Рис. 2.25. Варианты регрессии*

Для сопоставления полиномиальных регрессионных выражения (от полинома первого порядка до пятого порядка включительно) вычислены остаточные суммы квадратов, которые обозначены Δ (рис. 2.26). Проведенные вычисления подтверждают качественную зависимость остаточной суммы от порядка полинома.

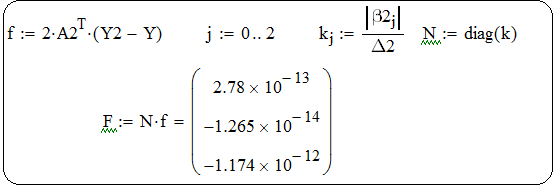
****

*Рис. 2.26. Сопоставление полиномиальных регрессий*

При выборе регрессионной модели путем сопоставления остаточных сумм квадратов (при условии их близких значений) выбирают более простой вариант модели. В качестве дополнительного упрощения выбранной регрессионной модели применяется ***функция чувствительности*** *,*с помощью которой определяют вклад оценки каждого параметра в результирующее значение.

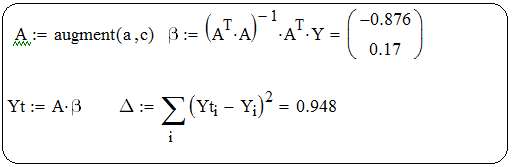
Функция чувствительности *f* представляет собой *m*-мерный вектор-столбец. Умножая вектор *f* слева на диагональную матрицу *N*, элементы диагонали которой составляют вектор , получаем вектор относительных функций чувствительности

Пусть выбранная регрессионная модель – полином второго порядка . На рис. 2.27 получен вектор относительной чувствительности для полинома второго порядка.

****

*Рис. 2.27. Получение вектора относительной чувствительности*

Анализ вектора*F* показывает, что наименьшая функция чувствительности у коэффициента , следовательно, возможно применить упрощенную регрессионную модель .

****

*Рис. 2.28. Решение упрощенной задачи регрессии*

Значение остаточной суммы квадратов**,** равное 0,948, незначительно отличается от значения 0,911. Таким образом, выбранная регрессионная модель имеет вид .

**Задачи для аудиторного решения**

1. Исследовать зависимость урожайности зерновых культур *Y* (ц/га) от количества осадков *Х* (см), выпавших в вегетационный период.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *X* | 25 | 27 | 30 | 35 | 36 | 38 | 39 | 41 | 42 | 45 | 46 | 47 | 50 | 52 | 53 |
| *Y* | 23 | 24 | 27 | 27 | 32 | 31 | 33 | 35 | 34 | 32 | 29 | 28 | 25 | 24 | 25 |

***Указание.*** *При выборе регрессионной модели можно учесть, что увеличение количества осадков приводит к увеличению урожайности до некоторого предела, после чего урожайность будет снижаться.*

2. Имеются следующие данные о выработке литья на одного работающего *Х*1 (т), браке литья *Х*2 (%) и себестоимости одной тонны литья *Y* (руб.) по 25 литейным цехам завода:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **№** | *Х*1 | *Х*2 | *Y* | **№** | *Х*1 | *Х*2 | *Y* |
| **1** | 14,6 | 4,2 | 239 | **14** | 75,8 | 3,4 | 172 |
| **2** | 13,5 | 6,7 | 254 | **15** | 27,6 | 1,1 | 201 |
| **3** | 21,5 | 5,5 | 262 | **16** | 88,4 | 0,1 | 130 |
| **4** | 17,4 | 7,7 | 251 | **17** | 16,6 | 4,1 | 251 |
| **5** | 44,8 | 1,2 | 158 | **18** | 33,4 | 2,3 | 195 |
| **6** | 111,9 | 2,2 | 101 | **19** | 17,0 | 9,3 | 282 |
| **7** | 20,1 | 8,4 | 259 | **20** | 33,1 | 3,3 | 196 |
| **8** | 28,1 | 1,4 | 186 | **21** | 30,1 | 3,5 | 186 |
| **9** | 22,3 | 4,2 | 204 | **22** | 65,2 | 1,0 | 176 |
| **10** | 25,3 | 0,9 | 1998 | **23** | 22,6 | 5,2 | 238 |
| **11** | 56,0 | 1,3 | 170 | **24** | 33,4 | 2,3 | 204 |
| **12** | 40,2 | 1,8 | 173 | **25** | 19,7 | 2.7 | 205 |
| **13** | 40,6 | 3,3 | 197 |

Найти уравнение множественной линейной регрессии *Y* по *Х*1 и *Х*2. Провести статистический анализ.

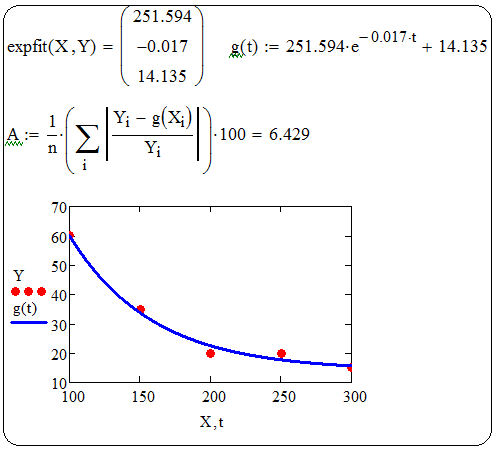
***Библиотека типовых функций регрессии Mathcad***

|  |  |  |
| --- | --- | --- |
| **Регрессионная модель** | **Встроенная функция** | **Замечание** |
|  | **expfit(X,Y,S)** | Выдаёт вектор, содержащий три коэффициента, что лучше всего аппроксимирует данные в векторах x и y. Дополнительный вектор S содержит оценки для этих трех коэффициентов. |
|  | **lgsfit(X,Y,S)** | Выдаёт вектор, содержащий эти 3 коэффициента для логистической кривой, лучше всего аппроксимируя данные в векторах x и y, используя оценки в S. |
|  | **pwrfit(X,Y,S)** | Выдаёт вектор, содержащий коэффициенты для кривой степени, что лучше всего аппроксимирует данные в векторах x и y. Вектор S содержит оценки для этих трех коэффициентов. |
|  | **sinfit(X,Y,S)** | Выдаёт вектор, содержащий коэффициенты для синусоиды, что лучше всего аппроксимирует данные в векторах x и y. Вектор S содержит оценки для этих трех коэффициентов**.** |
|  | **logfit(X,Y,S)** | Выдаёт вектор, содержащий эти три коэффициента для логарифмической кривой, что лучше всего аппроксимирует данные в векторах x и y. Вектор S содержит оценки для этих трех коэффициентов. |
|  | **lnfit(X,Y)** | Выдаёт вектор, содержащий эти 2 коэффициента для логарифмической кривой, что лучше всего аппроксимирует данные в x и y. |

**Пример 2.12.** По условию примера 2.9. подобрать нелинейную

модель регрессии, используя встроенные функции Mathcad.

**Решение**



*Рис. 2.28. Решение задачи нелинейной регрессии*

**Компьютерная лабораторная работа № 7**

***Корреляционно-регрессионный анализ***

**Задания к компьютерной лабораторной работе № 7**

**Задание 1.** Найти выборочное уравнение линейной регрессии Y на *X* на основании корреляционной таблицы. Провести статистический анализ.

1.1.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 10 | 15 | 20 | 25 | 30 | 35 | 40 |
| 30 |  | 4 | 7 |  |  | 1 | 5 |
| 50 | 2 |  | 4 | 6 | 5 |  |  |
| 70 |  | 3 |  | 4 | 5 | 6 |  |
| 90 | 10 |  | 2 |  |  | 5 | 3 |
| 110 | 2 | 4 |  | 8 | 4 |  | 10 |

1.2.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 42 | 46 | 50 | 54 | 58 | 62 |
| 15 |  |  | 4 | 2 | 1 |  |
| 25 | 2 | 1 |  | 3 | 8 | 5 |
| 35 |  | 4 | 2 | 1 |  | 3 |
| 45 | 3 | 2 | 10 |  | 3 | 2 |
| 55 | 1 | 3 |  | 9 |  | 1 |

1.3.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 5 | 10 | 15 | 20 | 25 | 30 | 35 |
| 30 |  | 6 |  | 4 |  | 2 | 5 |
| 40 | 4 |  | 5 |  | 7 | 1 |  |
| 50 |  | 4 | 3 | 5 |  |  | 6 |
| 60 | 5 | 3 |  |  | 10 | 2 |  |
| 70 |  |  | 4 | 10 | 4 | 2 | 8 |

1.4.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 20 | 25 | 30 | 35 | 40 | 45 |
| 30 |  | 6 |  | 4 |  | 2 |
| 40 | 4 | 1 | 5 |  | 7 |  |
| 50 | 3 |  | 4 | 5 |  | 6 |
| 60 | 5 | 3 |  | 10 | 2 |  |
| 70 |  | 2 | 3 |  | 3 | 5 |

1.5.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 10 | 15 | 20 | 25 | 30 | 35 | 40 |
| 10 | 2 |  | 4 | 6 | 5 |  |  |
| 20 |  | 4 | 7 |  |  | 1 | 5 |
| 30 | 3 |  |  | 4 | 5 | 6 |  |
| 40 | 3 | 5 |  | 2 |  |  | 10 |
| 50 |  | 4 | 2 |  | 4 | 10 | 8 |

1.6.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 5 | 10 | 15 | 20 | 25 | 30 |
| 100 |  | 6 | 4 | 2 |  | 2 |
| 110 | 4 | 2 | 8 | 1 | 5 |  |
| 120 |  |  |  | 10 | 7 | 1 |
| 130 | 5 | 3 | 8 |  | 6 | 7 |
| 140 | 9 | 5 |  | 4 |  | 1 |

1.7.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 5 | 10 | 15 | 20 | 25 | 30 | 35 |
| 5 | 10 |  | 3 | 5 |  | 1 | 4 |
| 15 |  | 4 | 10 |  | 2 | 8 |  |
| 25 | 3 | 4 |  | 6 |  |  | 6 |
| 35 |  |  |  | 4 | 7 | 1 | 5 |
| 45 | 2 | 5 |  |  | 10 |  |  |

1.8.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 5 | 10 | 15 | 20 | 25 | 30 | 35 |
| 30 |  | 6 |  | 4 |  | 2 | 5 |
| 40 | 4 |  | 5 |  | 7 | 1 |  |
| 50 |  | 4 | 3 | 5 |  |  | 6 |
| 60 | 5 | 3 |  |  | 10 | 2 |  |
| 70 |  |  | 4 | 10 | 4 | 2 | 8 |

1.9.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 10 | 15 | 20 | 25 | 30 | 35 |
| 105 | 6 | 4 |  |  |  |  |
| 115 |  | 6 | 8 |  |  |  |
| 125 |  |  |  | 20 | 2 | 5 |
| 135 |  |  |  | 5 | 12 | 6 |
| 145 |  |  |  |  | 1 | 5 |

1.10.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 10 | 15 | 20 | 25 | 30 | 35 |
| 15 | 6 | 4 |  |  |  |  |
| 25 |  | 6 | 8 |  |  |  |
| 35 |  |  |  | 21 | 2 | 5 |
| 45 |  |  |  | 4 | 12 | 6 |
| 55 |  |  |  |  | 1 | 5 |

**Задание 2.** По данным изучаемых регионов изучить зависимость общего коэффициента рождаемости (*Y*) от уровня бедности *X*1 (%) и среднедушевого дохода *X*2 (тыс. руб.). Исходные данные для корреляционно-регрессионного анализа приведены в таблице.

|  |  |  |  |
| --- | --- | --- | --- |
| Регион | *Х*1 | *Х*2 | *Y* |
| 1. Орловская область | 7,2 | 19,9 | 9,6 |
| 2. Рязанская область | 8,1 | 17,1 | 9,4 |
| 3. Смоленская область | 8,4 | 17,4 | 9,6 |
| 4. Тамбовская область | 8,6 | 13,5 | 8,9 |
| 5. Тверская область | 8,6 | 14,8 | 10,2 |
| 6. Тульская область | 8,4 | 14,2 | 8,4 |
| 7. Ярославская область | 9,9 | 15,1 | 9,9 |
| 8. Республика Карелия | 10,1 | 17,0 | 10,6 |
| 9. Республика Коми | 16,2 | 14,5 | 11,9 |
| 10. Архангельская область | 11,6 | 16,1 | 11,9 |
| 11. Вологодская область | 10,5 | 14,8 | 11,6 |
| 12. Калининградская область | 11,4 | 12,4 | 10,9 |
| 13. Ленинградская область | 10,6 | 12,6 | 8,3 |
| 14. Мурманская область | 15,2 | 15,5 | 10,3 |
| 15. Новгородская область | 8,6 | 20,3 | 10,7 |
| 16. Псковская область | 7,9 | 17,1 | 9,7 |
| 17. Республика Адыгея | 5,8 | 30,4 | 11,8 |
| 18. Республика Дагестан | 8 | 13,8 | 17 |
| 19. Республика Ингушетия | 4 | 44,8 | 16,7 |
| 20. Кабардино-Балкарская Республика | 6,6 | 18,3 | 12,8 |
| 21. Республика Калмыкия | 4,5 | 44,2 | 14,5 |
| 22. Карачаево-Черкесская Республика | 6,9 | 18,3 | 14,2 |
| 23. Республика Северная Осетия - Алания | 7,9 | 12,9 | 13,6 |
| 24. Чеченская Республика | … | ... | 27,1 |
| 25. Краснодарский край | 9,8 | 19,2 | 11,3 |

**Задание 3.** На 10 опытных участках одинакового размера получены следующие данные об урожайности *Х* (т) и содержании белка *Y* (%) для некоторой культуры:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *Х* | 9,9 | 10,2 | 11,0 | 11,6 | 11,8 | 12,5 | 12,8 | 13,5 | 14,3 | 14,4 |
| *Y* | 10,7 | 10,8 | 12,1 | 12,5 | 12,8 | 12,8 | 12,4 | 11,8 | 10,8 | 10,6 |

Построить эмпирическую линию регрессии *Y* по *Х*. Выровнять полученную зависимость по прямой. Подобрать нелинейную модель регрессии. Определить, какое из полученных уравнений регрессии целесообразно использовать для исследования данной зависимости.

**Глава 3**

**Теория систем массового обслуживания**

Теория систем массового обслуживания (СМО) начала развиваться в начале XX столетия. В 1909 г. шведский математик Эрланг применил теорию вероятностей к исследованию зависимости обслуживания телефонных вызовов от числа поступающих на телефонную станцию вызовов. В нашей стране известный математик А.Я. Хинчин систематизировал основные положения СМО в монографии «Теория очередей». Именно такое название теории СМО используется за рубежом.

В последние годы применение теории СМО в экономике приобрело особую актуальность в связи с использованием ряда ее аспектов в финансово-экономической сфере (банки различных типов, страховые организации, налоговые инспекции, аудиторские службы). Теория СМО широко применяется также в сфере обслуживания (различные системы связи, АЗС, магазины, ремонтные предприятия и т.д.) и в современных высоких технологиях (компьютерные сети, базы данных, военные системы ПВО).

**3.1. Структура и классификация систем массового обслуживания**

СМО представляют собой системы специфического вида. *Системой* называется целостное множество взаимосвязанных элементов, которые нельзя разделить на независимые подмножества.

Основными элементами СМО являются: входной поток заявок; очередь; каналы обслуживания (приборы, операторы, продавцы и пр.); выходной поток заявок (обслуженные заявки).

***Классификация СМО***

|  |  |
| --- | --- |
| **По числу каналов** | |
| ***одноканальные*** | ***многоканальные*** |
| **По дисциплине обслуживания** | |
| ***с отказами*** (заявка получает отказ при условии занятости каналов, например, вызовов абонента через АТС) | ***с ожиданием*** (***очередью***) (в случае занятости системы заявка поступает в очередь, например, обслуживание покупателей в магазине) |

***Показатели эффективности СМО*** описывают ее возможность справляться с потоком заявок.

|  |  |
| --- | --- |
| **Показатели эффективности** | |
| ***СМО с отказами*** | ***СМО с очередью*** |
| *A -* абсолютная пропускная способность СМО (среднее число заявок, обслуживаемых в единицу времени);  *Q –* относительная пропускная способность (средняя доля пришедших заявок, обслуживаемых СМО);  *Pserv* -вероятность обслуживания (вероятность того, что заявка будет принята на обслуживание);  *Potk* -вероятность отказа (вероятность того, что заявка покинет СМО необслуженной);  - среднее число занятых каналов (для многоканальной системы). | К показателям эффективности СМО с отказами добавляются:  *LСМО* – среднее число заявок в системе;  *Lоч* - среднее число заявок в очереди;  - среднее число заявок, находящихся по обслуживанием;  - среднее время пребывания заявки в очереди;  - среднее время пребывания заявки в СМО ;  *Рзан* -вероятность того, что канал занят(степень загрузки канала). |

**3.2. Марковский случайный процесс в СМО. Уравнения Колмогорова**

Процессы поступления и обслуживания заявок в СМО являются случайными, что обусловлено случайным характером потока заявок и длительности их обслуживания.

***Случайным процессом*** *X(t)* называется процесс, значение которого при любом значении аргумента *t* является случайной величиной. При фиксированном *t=t0 X(t0)* представляет собой обычную величину. Случайные процессы упрощают исследование СМО.

В основе СМО находится ***марковский случайный процесс*** (процесс без последствия), когда вероятность состояния СМО в будущем зависит только от ее состояния в настоящем и не зависит от прошлого (название по имени известного российского математика А.А. Маркова). Условие марковского случайного процесса: необходимо, чтобы все потоки событий, при которых система переходит из одного состояния в другое (потоки заявок, потоки обслуживания и т.д.) были пуассоновскими.

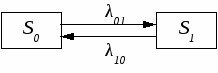
Пуассоновский поток (простейший поток) событий обладает следующими свойствами:

- отсутствие последствия (число событий, попавших на заданный временной интервал, не зависит от числа событий, попавших на другие интервалы);

- ординарности (вероятность попадания на элементарный временной интервал двух и более событий пренебрежимо мала по сравнению с вероятностью попадания одного события);

- стационарности (число событий, попавших на заданный временной интервал, зависит лишь от длины интервала и не зависит от числа событий, попавших на другие интервалы).

Для простейшего потока справедлив закон Пуассона. Плотность вероятности случайной величины при этом , где *λ* – интенсивность потока.

При анализе случайных процессов с дискретными состояниями пользуются *графом состояний*, где прямоугольниками изображают состояния системы, а переходы из состояния в состояние – стрелками. Если у стрелок проставлены интенсивности, то граф состояния называется *размеченным*. Переходы системы из состояния *Si* в состояние *Sj* происходят под воздействием простейших потоков событий с интенсивностями *λij.* Простейший граф состояний представлен на рисунке 3.1.

*Рис. 3.1. Граф состояний*

На рис. 3.1 изображена СМО, состоящая из одного канала обслуживания, который находится в двух возможных состояниях: либо свободен *(S0*), либо занят *(S1)*; *λ01* – интенсивность поступления заявок; *λ10* – интенсивность обслуживания заявок в единицу времени. Стрелка из *S0* в *S1* означает переход системы из состояния «канал свободен» в состояние «канал занят». Стрелка из *S1* в *S0* означает обратный переход.

Анализ состояния СМО сводится к определению вероятности, с которой система пребывает в данном состоянии. В общем случае ***вероятностью i-го состояния pi(t)*** называется вероятность того, что в момент *t* система будет находиться в состоянии *Si*. Для любого момента *t* справедливо соотношение

Определить вероятности состояний СМО можно, решив систему уравнений Колмогорова.

***Правило составления системы уравнений Колмогорова***

1. Слева в уравнениях стоит предельная вероятность данного состояния умноженная на суммарную интенсивность всех потоков, ведущих из данного состояния.

2. Справа – сумма произведений интенсивностей всех потоков, входящих в *i*-е состояние, на вероятности тех состояний, из которых эти потоки исходят.

3. Для решения системы вводится нормировочное уравнение

Для достаточно большого значения времени *t* распределение вероятностей стабилизируется и практически не зависит от времени.

**3.3. Расчет показателей эффективности СМО с отказами**

**3.3.1. Одноканальные СМО с отказами**

Система *S* может находиться в одном из двух состояний: *S0* – канал свободен или *S1* – канал занят. Из состояния *S0* в состояние *S1* систему переводит поток входящих заявок, а из состояния *S1*в состояние *S0* – поток обслуживаний. ***Плотности вероятностей перехода*** из состояния *S0* в состояние *S1* и обратно равны соответственно *λ* и *μ*. Граф состояний СМО показан на рисунке

http://ru.convdocs.org/pars_docs/refs/105/104994/104994_html_m1ee68433.gif

*Рис. 3.2. Граф состояний одноканальной СМО*

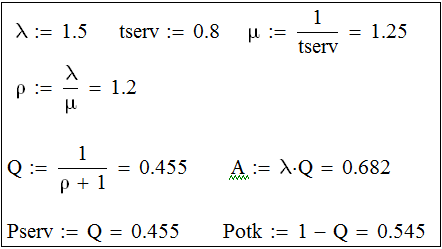
*с отказами*

Показатели эффективности одноканальной СМО с отказами вычисляются по формулам, которые приведены ниже.

|  |  |  |
| --- | --- | --- |
| **Показатели** | **Обозначение** | **Формула для вычисления** |
| Интенсивность входящего потока заявок | *λ* | известно из условия задачи |
| Интенсивность выходящего потока обслуженных заявок | *μ* | известно из условия задачи; |
| Среднее время обслуживания заявки |  | известно из условия задачи; |
| Приведенная интенсивность потока заявок | ρ |  |
| Относительная пропускная способность СМО | Q |  |
| Абсолютная пропускная способность СМО | А |  |
| Вероятность того, что заявка будет обслужена |  |  |
| Вероятность отказа |  |  |
| Среднее время простоя канала |  |  |
| Среднее время пребывания заявки в системе |  |  |

**Пример 3.1.** Телефонная АТС имеет одну линию, на которую в среднем приходит 0,8 вызовов в минуту. Среднее время разговора 1,5 мин. Вызов, пришедший во время разговора, не обслуживается. Считая потоки вызовов пуассоновскими, найти абсолютную и относительную пропускную способность станции и вероятность отказа абоненту.

**Решение.** Телефонную станцию рассматриваем как одноканальную СМО с отказами.



*Рис. 3.3. Показатели эффективности*

*одноканальной СМО с отказами*

Абсолютная пропускная способность СМО *А* оказалась почти вдвое меньше интенсивности μ потока обслуживания, что обусловлено случайным характером потока заявок.

**3.3.2. Многоканальная СМО с отказами**

Система *S* имеет следующие состояния (нумеруем по числу заявок, находящихся в системе): *S0*, *S1*,…, *Sn*, где *Sк* – состояние системы, когда в ней находится *к* заявок, т.е. занято *к* каналов.

Граф состояний СМО показан на рисунке 3.4.

S1

S2

Sk

Sn

S0

…

…

…

…

kµ

3µ

2µ

µ

λ

λ

λ

λ

λ

λ

…

(k+1)µ

nµ

*Рис. 3.4. Граф состояний многоканальной СМО*

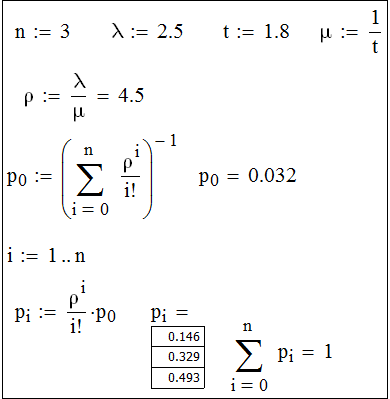
*с отказами*

Показатели эффективности многоканальной СМО с отказами вычисляются по формулам, которые приведены ниже.

|  |  |  |
| --- | --- | --- |
| **Показатели** | **Обозначение** | **Формула для вычисления** |
| Число каналов обслуживания | *n (n>1)* | известно из условия задачи |
| Интенсивность входящего потока заявок | *λ* | известно из условия задачи |
| Интенсивность потока обслуженных заявок, выходящих *из одного канала* | *μ* | известно из условия задачи; |
| Приведенная интенсивность потока заявок | ρ |  |
| Вероятность того, что занято  0, 1, …, *п* каналов, соответственно (формулы Эрланга) |  |  |
| Вероятность отказа |  |  |
| Относительная пропускная способность СМО | Q |  |
| Абсолютная пропускная способность СМО | А |  |
| Вероятность того, что заявка будет обслужена |  |  |
| Среднее число занятых каналов |  |  |

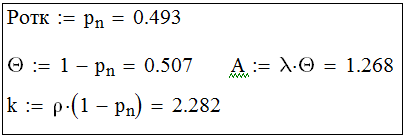
**Пример 3.2.** Имеетсятри поста для мойки автомобилей. Автомобиль, прибывший в момент, когда все посты заняты, – получает отказ в обслуживании. Интенсивность потока автомобилей *λ* = 2,5 (автомобиля в час). Средняя продолжительность обслуживания 1,8 часа. Поток автомобилей и поток обслуживания являются простейшими. Требуется определить в установившемся режиме предельные значения вероятностей, показатели эффективности.

**Решение.** Вводим исходные данные.Вычисляем приведенную интенсивность *ρ*. По формулам Эрланга вычисляем предельные вероятности состояний.

****

*Рис. 3.5.Вычисление предельных вероятностей состояний*

Вычислим вероятность отказа, относительную Q и абсолютную A пропускные способности, а также среднее число занятых каналов .

****

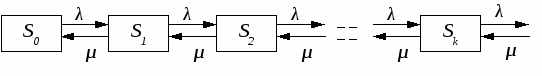
*Рис. 3.6*. *Вычисление показателей эффективности*

**3.4. Расчет показателей эффективности СМО с ожиданием (очередью)**

**3.4.1. Одноканальная СМО с неограниченной очередью**

Имеется одноканальная СМО с очередью, на которую не наложены никакие ограничения (ни по длине очереди, ни по времени ожидания). Система *S* может находится в одном из состояний (по числу заявок, находящихся в СМО): *S0* – канал свободен; *S1* – канал занят (обслуживает одну заявку), очереди нет; *S2* – канал занят, одна заявка в очереди; …; *Sк* – канал занят, (*к*-1) заявок стоят в очереди и т.д.

Граф состояний СМО представлен на рисунке 3.7.

****

*Рис. 3.7. Граф состояний одноканальной СМО*

*с неограниченной очередью*

Показатели эффективности одноканальной СМО с очередью вычисляются по формулам, которые приведены ниже.

|  |  |  |
| --- | --- | --- |
| **Показатели** | **Обозначение** | **Формула для вычисления** |
| Длина очереди | ∞ |  |
| Интенсивность входящего потока заявок | *λ* | известно из условия задачи |
| Интенсивность потока обслуженных заявок, выходящих *из одного канала* | *μ* | известно из условия задачи; |
| Приведенная интенсивность потока заявок | ρ<1 |  |
| Вероятность того, что СМО свободна и может обслужить заявку |  |  |
| Вероятность того, что СМО занята, в очереди нет заявок |  |  |
| Вероятность того, что СМО занята, в очереди находятся 1,…, *п,…* заявок, соответственно |  | 2,... |
| Вероятность того, что заявка получит отказ |  |  |
| Относительная пропускная способность СМО | Q |  |
| Абсолютная пропускная способность СМО | А |  |
| Вероятность того, что заявка будет обслужена |  |  |
| Среднее число заявок, стоящих в очереди |  |  |
| Среднее число заявок, находящихся по обслуживанием |  |  |
| Среднее число заявок в СМО (обслуживаемых и стоящих в очереди) |  |  |
| Среднее время ожидания заявки в очереди |  |  |
| Среднее время пребывания заявки в СМО |  |  |

***Замечание****. Прежде чем записать формулы предельных вероятностей, необходимо быть уверенным в их существовании, ведь в случае, когда время очередь может неограниченно возрастать. Если , т.е. среднее число приходящих заявок меньше среднего числа обслуженных заявок (в единицу времени), то предельные вероятности существуют. Если очередь растет до бесконечности.*

**3.4.2. Многоканальная СМО с неограниченной очередью**

Имеется *п-*канальная СМО с неограниченной очередью. Система *S* может находиться в одном из состояний (по числу заявок, находящихся в СМО): *S0* – в системе нет заявок (все каналы свободны); *S1* – занят один канал, остальные свободны; *S2* –заняты два канала, остальные свободны; …; *Sп* – заняты все *п* каналов (очереди нет); *Sп+1* - заняты все *п* каналов, в очереди одна заявка;…; *Sп+r* - заняты все *п* каналов, в очереди *r* заявок;…

Граф состояний СМО представлен на рисунке 3.8.

λ

λ

Sn+1

λ

Sn

kµ

…

Sk

…

Sk

λ

λ

…

S0

S1

S0

S1

2µ

µ

λ

λ

λ

…

…

…

…

(k+1)µ

…

λ

λ

Sn+r

…

…

…

…

nµ

nµ

nµ

nµ

nµ

*Рис. 3.8. Граф состояний многоканальной СМО*

*с неограниченной очередью*

Показатели эффективности многоканальной СМО с неограниченной очередью вычисляются по формулам, которые приведены ниже.

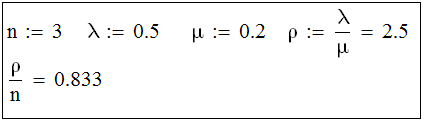
|  |  |  |
| --- | --- | --- |
| **Показатели** | **Обозначение** | **Формула для вычисления** |
| Число каналов обслуживания | *n (n>1)* | известно из условия задачи |
| Длина очереди | ∞ |  |
| Интенсивность входящего потока заявок | *λ* | известно из условия задачи |
| Интенсивность потока обслуженных заявок, выходящих *из одного канала* | *μ* | известно из условия задачи; |
| Приведенная интенсивность потока заявок | ρ |  |
| Предельные вероятности состояний системы (очереди нет) |  |  |
| Предельные вероятности состояний системы (наличие в очереди 1, 2,… заявок) |  | ,…,  ,…, |
| Вероятность того, что заявка окажется в очереди |  |  |
| Вероятность того, что заявка получит отказ |  |  |
| Относительная пропускная способность СМО | Q |  |
| Абсолютная пропускная способность СМО | А |  |
| Вероятность того, что заявка будет обслужена |  |  |
| Среднее число занятых каналов |  |  |
| Среднее число заявок, стоящих в очереди |  |  |
| Среднее число заявок в СМО (обслуживаемых и стоящих в очереди) |  |  |
| Среднее время ожидания заявки в очереди |  |  |
| Среднее время пребывания заявки в СМО |  |  |

***Замечание.*** *При* *предельные вероятности существуют. Если , очередь растет до бесконечности.*

**Пример 3.3.** Система массового обслуживания с ожиданием – билетная касса с тремя окошками (с тремя кассирами). Пассажиров, желающих купить билет, приходит в среднем 5 человек за 10 мин. Поток пассажиров можно считать простейшим. Кассир в среднем обслуживает двух пассажиров за 10 минут. Определите вероятностные

характеристики СМО в стационарном режиме.

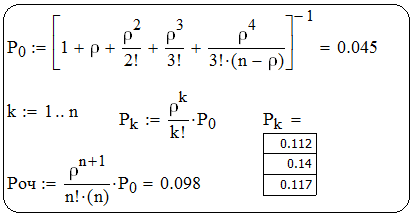
**Решение.** Введем начальные значения: *n* – число каналов обслуживания, λ– интенсивность потока заявок, µ– интенсивность обслуживания. Вычисляем приведенную интенсивность ρ. Проверяем условие существования предельных вероятностей.



*Рис. 3.9. Условие существования*

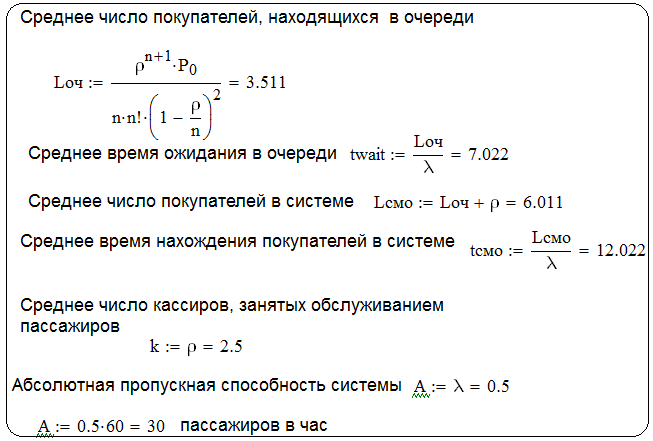
*предельных вероятностей*

Так как , то предельные вероятности существуют. Как видно из вычислений (рисунок 3.10), в среднем 4,5% времени кассиры будут простаивать.



*Рис. 3.10. Вычисление предельных вероятностей системы*

Найдем характеристики обслуживания СМО (рисунок 3.11).



*Рис. 3.11. Характеристики обслуживания СМО*

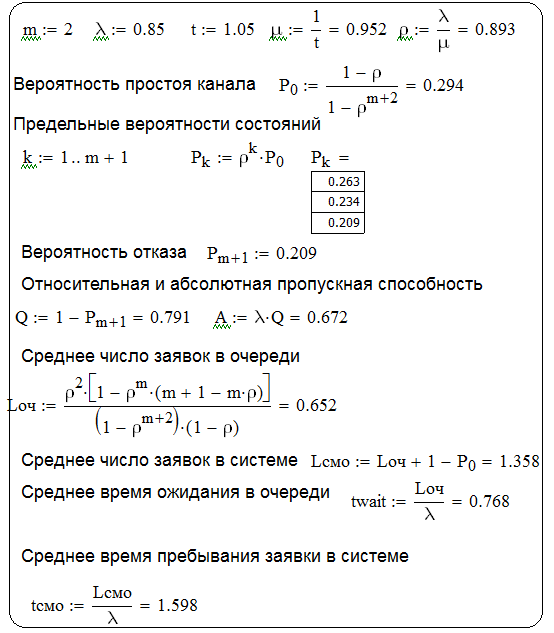
**3.4.3. СМО с ограниченной очередью**

СМО с ограниченной очередью отличаются от рассмотренных выше задач лишь тем, что число заявок в очереди ограничено (не может превосходить некоторого заданного числа *т*). Если новая заявка поступает в момент, когда все места в очереди заняты, она покидает СМО необслуженной. Формулы сведены в таблицу 3.1.

**Пример 3.4.** На вход одноканальной СМО с длиной очереди *m*= 2

поступает поток заявок с интенсивностью 0,85 заявок в час. Среднее время обслуживания одной заявки 1,05 часа. Найти основные характеристики данной СМО.

**Решение.**

****

*Рис. 3.12. Расчет показателей эффективности*

Таблица 3.1.

|  |  |  |
| --- | --- | --- |
| **Показатели** | **Одноканальная СМО**  **с ограниченной очередью** | **Многоканальная СМО**  **с ограниченной очередью** |
| Число каналов  Длина очереди  Приведенная интенсивность | 1  *m* | *п*  *m* |
| Предельные  вероятности |  |  |
| Вероятность  отказа |  |  |
| Относительная  пропускная  способность |  |  |
| Абсолютная  пропускная  способность |  |  |
| Среднее число  заявок, стоящих  в очереди |  |  |
| Среднее число  занятых каналов |  |  |
| Среднее число  заявок в СМО  (обслуживаемых  и стоящих в очереди) |  |  |
| Среднее время  ожидания заявки  в очереди |  |  |
| Среднее время  пребывания  заявки в СМО |  |  |

**Задачи для аудиторного решения**

1. В отделении сберегательного банка кассир обслуживает клиентов с интенсивностью 0,5 чел./мин. Среднее число клиентов, находящихся на обслуживании, равно 0,7. Предполагается, что нет ограничений на длину очереди. Определить показатели эффективности СМО и вероятность того, что ожидают своей очереди не более одного человека.

2. Рассматривается круглосуточная работа пункта проведения профилактического осмотра автомашин с одним каналом (одна группа проведения осмотра). На осмотр и выявление дефектов каждой машины затрачивается в среднем 0,5 часа. На осмотр поступает в среднем 36 машин в сутки. Если машина, прибывшая в пункт осмотра, не застает ни одного канала свободным, она покидает пункт осмотра необслуженной. Определить вероятности состояний и характеристики обслуживания пункта.

3. Решить задачу 2 для случая *п*=4.

4. Решить задачу 2 при условии, что машина, прибывшая на пункт осмотра, покидает этот пункт лишь в случае, если в очереди на осмотр стоят более 5 машин.

5. Решить задачу 3 при условии, что машина, прибывшая на пункт осмотра, покидает этот пункт лишь в случае, если в очереди на осмотр стоят более 5 машин.

**Компьютерная лабораторная работа № 8**

***Расчет показателей эффективности***

***систем массового обслуживания***

**Задание 1.** СМО с отказами представляет собой *n* диспетчеров телефонной станции. Заявка, пришедшая в момент, когда все диспетчеры заняты, получает отказ в обслуживании. Интенсивность потока заявок *λ*. Средняя продолжительность обслуживания . Поток заявок и поток обслуживания являются простейшими. Требуется определить в установившемся режиме предельные значения вероятностей, показатели эффективности.

|  |  |  |  |
| --- | --- | --- | --- |
| Вариант | *n* | *λ* |  |
| 1 | 4 | 5 | 0,9 |
| 2 | 3 | 5 | 0,7 |
| 3 | 4 | 6 | 0,8 |
| 4 | 3 | 4 | 0,9 |
| 5 | 5 | 6 | 0,6 |
| 6 | 4 | 7 | 0,7 |
| 7 | 3 | 4 | 0,8 |
| 8 | 5 | 5 | 0,9 |
| 9 | 4 | 6 | 0,8 |
| 10 | 3 | 5 | 0,7 |

**Задание 2.** СМО с ожиданием – ремонтная мастерская на *n* рабочих мест. Интенсивность потока заявок *λ*. Интенсивность обслуживания . Определите вероятностные характеристики СМО в стационарном режиме и показатели эффективности**.**

|  |  |  |  |
| --- | --- | --- | --- |
| Вариант | *n* | *λ* |  |
| 1 | 4 | 6,4 | 2 |
| 2 | 4 | 7,2 | 2 |
| 3 | 3 | 5,4 | 2 |
| 4 | 3 | 3,8 | 1,6 |
| 5 | 3 | 4,2 | 1,8 |
| 6 | 4 | 7 | 2 |
| 7 | 3 | 2,8 | 2 |
| 8 | 4 | 7,6 | 2 |
| 9 | 3 | 3,8 | 2 |
| 10 | 3 | 4,2 | 2 |

**Задание 3. Решите задачи**

**3.1.** Интенсивность потока телефонных звонков в агентство по заказу железнодорожных билетов, имеющему один телефон, составляет 16 вызовов в час. Продолжительность оформления заказа на билет равна 2,4 минуты. Определить относительную и абсолютную пропускную способность этой СМО и вероятность отказа (занятости телефона). Сколько телефонов должно быть в агентстве, чтобы относительная пропускная способность была не менее 0,75.

**3.2.** Система массового обслуживания — билетная касса с одним окошком и неограниченной очередью. Касса продает билеты в пункты А и В. Пассажиров, желающих купить билет в пункт А, приходит в среднем трое за 20 мин, в пункт В — двое за 20 мин. Поток пассажиров простейший. Кассир в среднем обслуживает трех пассажиров за 10 мин. Время обслуживания — показательное. Вычислить финальные вероятности, среднее число заявок в системе и в очереди, среднее время пребывания заявки в системе, среднее время пребывания заявки в очереди.

**3.3.** Междугородный переговорный пункт имеет четыре телефонных аппарата. В среднем за сутки поступает 320 заявок на переговоры. Средняя длительность переговоров составляет 5 мин. Длина очереди не должна превышать 6 абонентов. Потоки заявок и обслуживаний простейшие. Определить характеристики обслуживания переговорного пункта в стационарном режиме (вероятность простоя каналов, вероятность отказа, вероятность обслуживания, среднее число занятых каналов, среднее число заявок в очереди, среднее число заявок в системе, абсолютную пропускную способность, относительную пропускную способность, среднее время заявки в очереди, среднее время заявки в системе, среднее время заявки под обслуживанием).

**Список использованной литературы**

1. Ивановский Р.И. Теория вероятностей и математическая статистика. Основы, прикладные аспекты с примерами и задачами. – СПб.: БХВ-Петербург, 2008. – 528 с.

2. Красс М.С., Чупрынов Б.П. Математика для экономического бакалавриата: Учебник. – М.: ИНФРА-М, 2011. – 472 с.

3. Кремер Н.Ш. Теория вероятностей и математическая статистика: Учебник для вузов. – М.: ЮНИТИ-ДАНА, 2002. – 543 с.

4. Кремер Н.Ш. Исследование операций в экономике: Учеб. пособие для вузов. – М.: ЮНИТИ, 2006. – 407 с.

4. Сборник задач по высшей математике для экономистов: Учебное пособие / Под ред. В.И. Ермакова. – 2-е изд., испр. – М.: ИНФРА-М, 2008. – 575 с.

5. Черняк А.А., Новиков В.А., Мельников О.И., Кузнецов А.В. Математика для экономистов на базе MathCAD: Учебное пособие. – СПб.: БХВ-Петербург, 2003. – 496 с.