Министерство образования и науки Российской Федерации
Ярославский государственный университет им. П. Г. Демидова
Кафедра радиофизики

А. С. Гвоздарёв
Т. К. Артёмова

Планирование и обработка результатов инженерного эксперимента

Методические указания

Рекомендовано
Научно-методическим советом университета
для студентов, обучающихся по направлению Радиотехника

Ярославль
ЯрГУ
2014

УДК 69.162.2.012-52(072)
ББК 3973.2я73
Г25

Рекомендовано
Редакционно-издательским советом университета
в качестве учебного издания. План 2014 года

Рецензент
кафедра радиофизики ЯрГУ им. П. Г. Демидова

Гвоздарёв, Алексей Сергеевич.
Г25 Планирование и обработка результатов инженерного эксперимента : методические указания / А. С. Гвоздарёв, Т. К. Артёмова ; Яросл. гос. ун-т им. П. Г. Демидова. — Ярославль : ЯрГУ, 2014. — 60 с.

Методические указания содержат изложение теоретического материала в объёме, необходимом для решения задач и выполнения лабораторных работ, авторские задачи, задания к лабораторным работам и рекомендации по их выполнению.

Предлагаемые студентам задания охватывают основные методы планирования эксперимента и анализа его результатов и позволяют освоить материал на примере часто встречающихся в практике радиоинженеров задач.

Методические указания предназначены для студентов, обучающихся по направлению 210400.62 Радиотехника (дисциплина «Планирование и обработка результатов инженерного эксперимента», цикл Б3), очной формы обучения.

УДК 69.162.2.012-52(072)
ББК 3973.2я73

Оглавление

Введение

Эксперименты разного рода – необходимая составляющая труда инженера. Грамотное планирование и эффективная обработка результатов позволяют говорить о достоверности, надёжности результатов, гарантировать в дальнейшем качество работы испытанных элементов, устройств, систем.

Методические указания предназначены для освоения курса «Планирование и обработка результатов инженерного эксперимента» и содержат краткое изложение теоретического материала, отсылки к литературе, которую авторы могут рекомендовать для углублённого изучения, авторские задачи, задания к лабораторным работам.

Предлагаемые задания позволяют освоить материал на примере часто встречающихся в практике радиоинженеров задач.

В силу специфики предмета рекомендуется выполнять предложенные задания, а также прорабатывать приводимые примеры в пакетах программ, предназначенных для статистической обработки данных, пользуясь, например, изданиями [1] – [3].

1. Краткие теоретические сведения

1.1. Понятие и классификация инженерных экспериментов

Инженерный эксперимент – это система операций, воздействий и (или) наблюдений, направленных на получение информации об объекте исследования. Основы теории экспериментов приведены в [4].

Инженерные эксперименты имеют общие особенности:

– эксперименты планируются;

– чтобы уменьшить объем эксперимента, стремятся сократить число рассматриваемых переменных;

– ход эксперимента желательно контролировать;

– принимают меры по исключению или уменьшению влияния случайных внешних воздействий;

– оценивают точность измерительных приборов и точность получения данных;

– в процессе эксперимента анализируют и интерпретируют полученные результаты.

1.1.1. Классификация инженерных экспериментов

По цели проведения и форме представления полученных результатов инженерные эксперименты делят следующим образом:

– ***качественный*** – устанавливает только факт существования какого-либо явления; по окончании такого эксперимента получают словесное описание результатов;

– ***количественный*** – устанавливает соотношения количественных характеристик явления и внешних воздействий на объект исследования – ***факторов***.

Факторы бывают:

– ***контролируемые***, уровень которых можно регистрировать (делятся на ***управляемые***, чей уровень можно ещё и задавать, и ***неуправляемые*** – экспериментатор не может задавать их конкретный уровень);

– ***неконтролируемые***, уровни которых экспериментатор не может зарегистрировать; это могут быть и неизвестные ему факторы.

Целью инженерного эксперимента является получение вида и, если возможно, количественного описания ***функции отклика*** – зависимости некоторой статистической характеристики S наблюдаемой случайной величины Y от различных факторов:

$$S_y = f(x_i, h_j) + \varepsilon_\delta,$$

где x_i – контролируемые и управляемые факторы; h_j – контролируемые, но неуправляемые факторы; ε_δ – ошибка эксперимента, учитывающая влияние неконтролируемых факторов.

По тому, какой группой факторов располагает исследователь, количественные эксперименты можно разделить

– на ***пассивные количественные эксперименты*** – с контролируемыми неуправляемыми факторами; при этом $S_y = f(h_j) + \varepsilon_\delta$;

– ***активные количественные эксперименты***, в которых уровни факторов в каждом опыте задаются исследователем и ищется функция отклика $S_y = f(x_i, h_j) + \varepsilon_\delta$ либо сочетание уровней управляемых факторов x_i, при котором достигается оптимальное (экстремальное – минимальное или максимальное) значение функции отклика; в последнем случае эксперимент называется ***поисковым*** (экстремальным).

По условиям проведения инженерного эксперимента различают:

– ***лабораторный эксперимент*** – проводится в лабораторных условиях, отличается более тщательной подготовкой, возможностью обеспечения большей стерильности условий (меньшим влиянием случайных погрешностей), точности проведения опытов, воспроизводимостью результатов, возможностью изменять уровни факторов в большом диапазоне; при прочих равных условиях для установления некоторого факта в таких условиях требуется выполнить меньшее число опытов;

– ***натурный эксперимент*** – промышленный, в условиях полигона или в реальной обстановке; как правило, является завершающим этапом исследования.

1.1.2. Планирование экспериментов

Сначала выделяют факторы и устанавливают цель эксперимента, определяют измеряемые или наблюдаемые характеристики исследуемого явления и строят целевую функцию.

При планировании активного эксперимента также определяют степень точности результатов и степень соответствия уровня факторов требуемым.

Необходимо также продумать:

– надежность результатов – требуемую величину и то, чем она будет подтверждаться;

– число, условия и порядок реализации опытов;

– сроки, имеющиеся в распоряжении исследователя;

– средства, имеющиеся в распоряжении исследователя, в том числе составляют схему экспериментальной установки, устанавливают временную диаграмму воздействий и снятия показаний.

Результаты эксперимента фиксируются, обрабатываюся и анализируются. Если необходимо, проводятся дополнительные или повторные эксперименты.

Эксперименты заканчиваются представлением результатов, формулировкой выводов, выдачей рекомендаций в виде графиков, схем, таблиц, формул, статистических данных или словесных описаний.

1.1.3. Методическое обеспечение процесса обработки экспериментальных данных

Существует много методов обработки и анализа результатов эксперимента. Наиболее востребованы на практике:

- методы оценивания:
  - точечного:
    - метод наименьших квадратов;
    - метод максимального правдоподобия;
    - метод моментов;
    - баейсовский подход к оцениванию;
  - интервального;
- методы проверки статистических гипотез:
  - простых;
  - сложных;
  - параметрических;
- корреляционный анализ;
- кластерный анализ;

■ регрессионный анализ;
■ дисперсионный анализ (ANOVA);
■ дискриминантный анализ (классификация с учителем);
■ методы снижения размерности;
  – классификация;
  – группирование;
  – метод главных компонент.

Выбирают конкретный метод в зависимости
■ от поставленной цели обработки данных;
■ предположений относительно статистических свойств:
  – наблюдаемых явлений;
  – присутствующих факторов;
■ удобного математического аппарата.

Подробнее ознакомиться с методами обработки экспериментальных данных и их статистическо-вероятностными характеристиками можно в [4] – [13].

1.2. Распределения вероятностей случайных величин

1.2.1. Дискретные распределения

Дискретные распределения описывают вероятность счётных событий, характеризуются **функцией вероятности** (ФВ) $w_\xi(x)$ и **функцией распределения** (ФР) $F_\xi(x)$:

$$w_\xi(x) \equiv \mathbb{P}(\xi = x), \quad x = 0,\ldots,n, \quad F_\xi(x) \equiv \mathbb{P}(\xi \leqslant x).$$

Наиболее часто встречающиеся дискретные распределения описаны ниже (см., например, [6], [7], [11]). Их ФР и ФР приведены на рисунках в табл. 1.1, а параметры – в табл. 1.2. Задействованные в описаниях специальные функции можно найти в [14].

1. **Биномиальное распределение (распределение Бернулли).** Если событие осуществляется в единичном испытании с некоторой вероятностью p, то число появлений события в последовательности n независимых испытаний подчинено биномиальному распределению $Bin(n, p)$. Оно характеризуется ФВ и ФР:

$$w_\xi(x)=\binom{n}{x}p^x(1-p)^{n-x},\ F_\xi(x)=\sum_{k=0}^{x}\binom{n}{k}p^k(1-p)^{n-k},\ x=0,\dots n,$$

где $\binom{n}{x},\binom{n}{k}$ – биномиальные коэффициенты.

2. **Распределение Пуассона.** Если вероятность появления независимых событий в малом промежутке времени Δt пропорциональна Δt, то число их появлений подчинено распределению Пуассона $Pois(\lambda)$ с ФВ и ФР:

$$w_\xi(x)=\frac{\lambda^x}{x!}e^{-\lambda},\ \ F_\xi(x)=e^{-\lambda}\sum\nolimits_{i=0}^{x}\frac{\lambda^i}{i!}=\frac{\Gamma(x+1,\lambda)}{x!},$$

где $\Gamma(a,z)=\int\limits_{z}^{\infty}e^{-t}t^{a-1}dt$ – верхняя неполная гамма-функция, а $\lambda=v\Delta t$, где v – средняя интенсивность потока событий.

3. **Отрицательное биномиальное распределение.** Если p – вероятность появления события в единичном испытании, то число неудачных испытаний до появления m-го успеха подчиняется отрицательному биномиальному распределению с ФВ и ФР:

$$w_\xi(x)=\binom{m+x-1}{m}p^m(1-p)^x,\ F_\xi(x)=\begin{cases}I_p(n,\lfloor k\rfloor+1), & k\ge 0,\\ 0, & k\le 0,\end{cases}$$

$$m,\ x=0,\ 1,\ \dots,$$

где $I_p(n,\lfloor k\rfloor+1)$ – неполная бета-функция.

4. **Геометрическое распределение (распределение Фарри).** Если p – вероятность появления события в одном испытании, то число испытаний x до появления события подчинено геометрическому распределению $Geom(p)$ с ФВ и ФР:

$$w_\xi(x)=(1-p)^{x-1}p,\qquad F_\xi(x)=1-(1-p)^x.$$

5. **Гипергеометрическое распределение.** Если в серии из n_{tot} экспериментов реализуется n_s желаемых исходов, то вероятность появления x желаемых исходов в серии объёма n будет подчинена гипергеометрическому распределению с ФВ и ФР:

$$w_\xi(x) = \frac{\binom{n_s}{x}\binom{n_{tot}-n_s}{n-x}}{\binom{n_{tot}}{n}}, \qquad F_\xi(x) = \sum\nolimits_{i=0}^{x} \frac{\binom{n_s}{i}\binom{n_{tot}-n_s}{n-i}}{\binom{n_{tot}}{n}}.$$

В основе биномиального, геометрического и отрицательного биномиального распределений лежит последовательность независимых испытаний Бернулли (***схема Бернулли***); при этом механизм возникновения распределений определяется способом, которым обрывается последовательность.

Биномиальное распределение имеет место в тех случаях, когда последовательность испытаний обрывается после проведения фиксированного числа испытаний n. При этом под биномиальной случайной величиной x понимается число успехов в серии из n испытаний Бернулли.

Геометрическое распределение возникает при обрыве серии испытаний сразу же после первого успеха. При этом рассматриваются две случайные величины: число неудач, предшествовавших первому успеху, и число испытаний до первого успеха включительно.

Отрицательное биномиальное распределение возникает, когда последовательность испытаний обрывается сразу же после m-го успеха.

Пример 1. Вероятность появления дефектной радиостанции в производстве равна 0.05. Вычислить вероятность появления в партии из 50 изделий не более 5 дефектных.

Решение. Имеем $p = 0.1$, $n = 60$, $x = 10$. Искомая величина – $F_\xi(5)$ для распределения Бернулли:

$$F_\xi(5) = \sum_{i=0}^{5} \binom{i}{50} \cdot 0.05^i \cdot 0.95^{50-i} = 0.962\,.$$

Пример 2. На мини-атс автоперевозчика поступают вызовы со средней интенсивность v =40 вызовов в час. Найти вероятность того, что за 5 минут на станцию поступит ровно 10 вызовов; вероятность того, что за 5 минут придёт хотя бы один вызов; вероятность того, что за 5 минут придёт не менее 10 вызовов.

Решение. Считая, что число вызовов распределено по закону Пуассона, найдём среднее число вызовов за 5 минут: $\lambda = 5 \cdot \nu / 60 = 10 / 3$. Тогда вероятность поступления ровно 10 вызовов равна: $w_\xi(10) = \lambda^{10} e^{-\lambda} / 10! = 3.333^{10} e^{-3.333}/10! = 0.0017$.

Вероятность того, что за 5 минут придет хотя бы один вызов:

$$\mathbb{P}(\xi \geq 1) = 1 - \mathbb{P}(\xi \leq 1) = 1 - F_\xi(1) = 1 - e^{-\lambda} = 1 - e^{-10/3} = 0.845.$$

Вероятность того, что за 5 минут придёт не менее 10 вызовов:

$$\mathbb{P}(\xi \geq 10) = 1 - \mathbb{P}(\xi \leq 10) = 1 - F_\xi(10) = 0.0007.$$

Пример 3. Вероятность сборки дефектного телевизора равна 0.01. Найти вероятность того, что будут произведены 200 годных телевизоров до появления 5-го дефектного; вероятность того, что до появления 2-го дефектного телевизора будут произведены не более 10 годных.

Решение. Вероятность того, что потребуется произвести 200 телевизоров до появления 5-го дефектного, подчиняется отрицательному биномиальному закону с $m = 5$, $p = 0.01$ и $x = 200$:

$$w_\xi(200) = \binom{5 + 200 - 1}{5} \cdot 0.01^5 \cdot 0.99^{200} = 0.0009.$$

Вероятность того, что до появления 2-го дефектного телевизора будет произведено не более 10 годных изделий:

$$F_\xi(10) = \sum_{i=1}^{10} w_\xi(i) = \sum_{i=1}^{10} \binom{2 + i - 1}{i} \cdot 0.01^2 \cdot 0.99^i = 0.0062.$$

Пример 4. Вероятность безотказной работы робота равна 0.9. Найти вероятность того, что для получения одного отказа необходимо испытать 20 роботов, и вероятность того, что для получения первого отказа понадобится испытать не более 10 роботов.

Решение. Вероятность того, что для получения одного отказа необходимо испытать 20 роботов (обрыв схемы Бернулли после первого испытания соответствует геометрическому распределению), равна $w_\xi(20) = 0{,}1 \cdot 0{,}9^{19} = 0.0135$.

Вероятность того, что для получения первого отказа понадобится испытать не более 10 роботов, равна $F_\xi(10) = 1 - (1 - 0{,}1)^{10} = 0.686$.

1.2.2. Непрерывные распределения

Непрерывным распределениям подчиняются непрерывные случайные величины. Распределения задаются функцией распределения $F_\xi(x)$ и **функцией плотности вероятности (ФПВ)** $w_\xi(x) = \dfrac{dF_\xi(x)}{dx}$. Графики ФР и ФПВ наиболее часто встречающихся распределений приведены на рисунках в табл. 1.3, параметры – в табл. 1.4 (см., например, [6], [7], [12], [13]).

1. **Равномерное распределение**. Равномерному распределению $\mathcal{U}(a,b)$ подчиняются случайные величины, имеющие одинаковую вероятность появления (например, погрешность измерений с округлением). Оно имеет следующие ФПВ и ФР:

$$w_\xi(x) = \begin{cases} \dfrac{1}{b-a}, & x \in [a,b], \\ 0, & \text{иначе} \end{cases} \qquad F_\xi(x) = \begin{cases} 0, & x < a, \\ \dfrac{x-a}{b-a}, & x \in [a,b), \\ 1, & x \ge b. \end{cases}$$

2. **Нормальное распределение.** Нормальное распределение (Гаусса) $\mathcal{N}(\mu,\sigma)$ – распределение вероятностей с ФР и ФПВ:

$$w_\xi(x) = \frac{1}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),$$

$$F_\xi(x) = \Phi\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{2}[1 + erf(\frac{x-\mu}{\sqrt{2\sigma^2}})],$$

где $\phi(.)$ – функция Лапласа, а $\Phi(.)$ – интеграл Лапласа [7].

Нормальное распределение является краеугольным камнем математической статистики в силу ряда причин:

– схема его возникновения соответствует многим реальным физическим процессам;

– как следствие центральной предельной теоремы теории вероятностей при возрастании объёма выборки предельное распределение для большинства распределений является нормальным;

– нормальное распределение обладает рядом благоприятных математико-статистических свойств (легко нормируется и аппроксимируется, обладает свойством аддитивности).

3. **Логнормальное распределение.** Если случайная величина η распределена нормально, то случайная величина $\xi = \ln(\eta)$ подчинена логарифмически нормальному («логнормальному») закону $\ln\mathcal{N}(\mu,\sigma^2)$ с ФПВ и ФР:

$$w_\xi(x) = \frac{1}{\sigma}\phi\left(\frac{\ln x - \mu}{\sigma}\right) = \frac{1}{x\sigma\sqrt{2\pi}}e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}},$$

$$F_\xi(x) = \Phi(\frac{\ln x - \mu}{\sigma}) = \frac{1}{2}\left[1 + erf\left(\frac{\ln x - \mu}{\sigma\sqrt{2}}\right)\right], x > 0.$$

Логнормальное распределение используется для описания износовых отказов невосстанавливаемых приборов (некоторые типы электронных ламп, полупроводниковые приборы, печатные платы), или в задачах распространения радиоволн – для описания параметров, выражаемых при измерениях в дБ, например мощность или напряжённость поля. Оно показывает, что значения переменной являются результатом мультипликативного действия многочисленных причин, которые по отдельности большого влияния не оказывают.

4. **Усечённое нормальное распределение.** Если из генеральной совокупности, имеющей нормальное распределение $\mathcal{N}(\mu,\sigma)$, изъять все элементы, меньшие или большие определённых граничных значений a и b, то образуется совокупность, подчиненная усеченному нормальному распределению с ФПВ и ФР:

$$w_\xi(x) = \frac{\frac{1}{\sigma}\phi(\frac{x-\mu}{\sigma})}{\Phi(\frac{b-\mu}{\sigma}) - \Phi(\frac{a-\mu}{\sigma})}, \quad F_\xi(x) = \frac{\Phi(\frac{x-\mu}{\sigma}) - \Phi(\frac{a-\mu}{\sigma})}{\Phi(\frac{b-\mu}{\sigma}) - \Phi(\frac{a-\mu}{\sigma})}.$$

5. **Распределение хи-квадрат (χ^2).** Пусть $\eta_1, \dots, \eta_k$ – совместно независимые *стандартные нормальные* случайные величины: $\eta_i \sim \mathcal{N}(0,1)$. Тогда случайная величина $\xi = \eta_1^2 + \dots + \eta_k^2$ имеет распределение хи-квадрат с k степенями свободы χ_k^2 с ФПВ и ФР:

$$w_\xi(x) = \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)}x^{\frac{k}{2}-1}e^{-\frac{x}{2}}, \qquad F_\xi(x) = \frac{1}{\Gamma\left(\frac{k}{2}\right)}\gamma\left(\frac{k}{2},\frac{x}{2}\right),$$

где $\gamma(\cdot,\cdot)$ – нижняя неполная гамма-функция.

Распределение хи-квадрат используется:

– при оценивании дисперсии с помощью доверительного интервала;

– при проверке гипотез согласия экспериментальных данных с теоретическими законами распределения;

– при проверке гипотез однородности;

– при проверке гипотез независимости, прежде всего для качественных (категоризованных) переменных, принимающих конечное число значений;

– при проверке статистических гипотез.

Распределение хи-квадрат также является предельным для многих выборочных статистик.

6. **Распределение Стьюдента.** Пусть $\eta_0, \eta_1, \ldots, \eta_\nu$ – независимые стандартные нормальные случайные величины $\eta_i \sim \mathcal{N}(0,1)$, $i = 0, \ldots, \nu$. Тогда случайная величина $\xi = \eta_0 \Big/ \sqrt{\frac{1}{\nu}\sum_{i=1}^{\nu}\eta_i^2}$ имеет распределение Стьюдента $\mathrm{t}(\nu)$ с ν степенями свободы с ФПВ и ФР (в таблице параметров $K_\nu(\cdot)$ – функция Макдональда):

$$w_\xi(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \quad F_\xi(x) = \frac{1}{2} + x\Gamma\left(\frac{\nu+1}{2}\right)\frac{{}_2F_1\left(\frac{1}{2},\frac{\nu+1}{2};\frac{3}{2};-\frac{x^2}{\nu}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)}.$$

Распределение Стьюдента используется для точечного оценивания, построения доверительных интервалов и тестирования гипотез о неизвестном среднем статистической выборки из нормального распределения. В частности, если $\bar{X}$ – выборочное среднее выборки независимых случайных величин $X_i \sim \mathcal{N}(\mu,\sigma^2)$, $i = 1, \ldots, \nu$, а S^2 – выборочная оценка её дисперсии, тогда $\frac{\bar{X}-\mu}{S/\sqrt{\nu}} \sim \mathrm{t}(\nu-1)$.

При $\nu \to \infty$ распределение Стьюдента совпадает со стандартным нормальным (хорошая аппроксимация достигается уже при $\nu \to 30$).

7. **Распределение Фишера–Снедекора.** Пусть η_1, η_2 – две независимые случайные величины, имеющие распределение хи-квадрат: $\eta_i \sim \chi_{d_i}^2$, где $d_i \in \mathbb{N}$, $i = 1, 2$. Тогда распределение случайной величины $\xi = \dfrac{\eta_1 / d_1}{\eta_2 / d_2}$ называется распределением Фишера (Фишера–Снедекора) $\xi \sim \mathrm{F}_{d_1,d_2}$ со степенями свободы d_1 и d_2 и описывается ФПВ и ФР (в табл. 1.4 $U(\cdot,\cdot,\cdot)$ – вырожденная гипергеометрическая функция):

$$w_\xi(x) = \frac{1}{\mathrm{B}\left(\frac{d_1}{2}, \frac{d_2}{2}\right)} \left(\frac{d_1}{d_2}\right)^{\frac{d_1}{2}} x^{\frac{d_1}{2}-1} \left(1 + \frac{d_1}{d_2} x\right)^{-\frac{d_1+d_2}{2}}, \quad F_\xi(x) = I_{\frac{d_1 x}{d_1 x + d_2}}\left(\frac{d_1}{2}, \frac{d_2}{2}\right),$$

где $B(\cdot)$ – бета-функция.

Распределение Фишера–Снедекора играет фундаментальную роль в статистике и появляется в первую очередь как распределение отношения двух выборочных дисперсий. Если $X_1, \ldots, X_m$ и $Y_1, \ldots, Y_m$ – выборки из нормальных совокупностей $\mathcal{N}(a_1, \sigma_1^2)$ и $\mathcal{N}(a_2, \sigma_2^2)$ и выражения

$$s_1^2 = \frac{1}{m-1} \sum_i (X_i - \overline{X})^2, \quad s_2^2 = \frac{1}{n-1} \sum_j (Y_j - \overline{Y})^2,$$

где $\overline{X} = \dfrac{1}{m} \sum\limits_i X_i$ и $\overline{Y} = \dfrac{1}{n} \sum\limits_j Y_j$, служат оценками дисперсий σ_1^2 и σ_2^2. Тогда так называемое ***дисперсионное отношение*** $F_{m,n} = \dfrac{s_1^2}{s_2^2}$ имеет *F-распределение* с $(m-1)$ и $(n-1)$ степенями свободы. На F-статистике основан ***F-критерий***, используемый, в частности, для проверки гипотезы равенства дисперсий двух совокупностей в дисперсионном, регрессионном и многомерном статистическом анализе.

Универсальность F-распределения подчёркивается связями с другими распределениями. При $m = 1$ квадрат величины $F_{m,n}$ имеет распределение Стьюдента с n степенями свободы.

8. **Экспоненциальное распределение.** Экспоненциальное распределение $Exp(\lambda)$ – распределение интервала времени между двумя последовательными однотипными событиями, его ФПВ и ФР:

$$w_\xi(x) = \lambda e^{-\lambda x} 1_{x\geq 0}, \qquad F_\xi(x) = \left(1 - e^{-\lambda x}\right) 1_{x \geq 0},$$

где $\mathbf{1}_{x\geq 0}$ – индикаторная функция.

Это одно из наиболее часто встречающихся распределений в теории надёжности и в теории массового обслуживания; оно используется для описания: внезапных отказов, когда износом изделия можно пренебречь; наработки на отказ невосстанавливаемых изделий; наработки между соседними отказами (ремонтами) у восстанавливаемых изделий в случае простейшего потока отказов; наработки на отказ большой многокомпонентной системы при любом распределении наработки на отказ компонентов системы.

Отличительная особенность экспоненциального распределения – постоянство интенсивности отказов λ = const – в теории надёжности интерпретируется как независимость вероятности отказа от наработки, что эквивалентно отсутствию износа.

9. **Распределение Рэлея.** Векторная сумма ξ двух нормальных случайных величин $\eta_1 \sim \mathcal{N}(0,\sigma^2)$ и $\eta_2 \sim \mathcal{N}(0,\sigma^2)$ с равными дисперсиями подчинена распределению Рэлея *Rayleigh*(σ) с ФПВ и ФР:

$$w_\xi(x) = \frac{x}{\sigma^2}\exp\left(-\frac{x^2}{2\sigma^2}\right), \qquad F_\xi(x) = 1 - \exp\left(\frac{-x^2}{2\sigma^2}\right).$$

Распределению Рэлея подчиняются погрешности формы (овальность, конусообразность), ошибки взаимного расположения поверхностей (эксцентриситет, разностенность). Оно находит широкое применение в баллистике, теории надёжности, а также в радиотехнике для описания амплитудных флуктуаций радиосигнала, в том числе в многолучевых каналах, для описания случайной огибающей узкополосного случайного процесса (шума), в радиолокации при оценке погрешности обнаружения объектов.

10. **Распределение Райса.** Если η_1 и η_2 – независимые случайные величины, имеющие нормальное распределение

с одинаковыми дисперсиями и ненулевыми математическими ожиданиями $\mu\eta_1$, $\mu\eta_2$: $\eta_1 \sim \mathcal{N}(\mu_{\eta_1}, \sigma^2)$, $\eta_2 \sim \mathcal{N}(\mu_{\eta_2}, \sigma^2)$, то величина $\xi = \sqrt{\eta_1{}^2 + \eta_2{}^2}$ имеет распределение Райса $Rice(\nu, \sigma)$. Оно является обобщением распределения Рэлея и описывается ФПВ и ФР (в табл. 1.4 $L_\nu(\cdot)$ – полиномы Лагерра):

$$w_\xi(x) = \frac{x}{\sigma^2}\exp\left(\frac{-(x^2+\nu^2)}{2\sigma^2}\right) I_0\left(\frac{x\nu}{\sigma^2}\right), \quad F_\xi(x) = 1 - Q_1\left(\frac{\nu}{\sigma}, \frac{x}{\sigma}\right),$$

где $I_0(\cdot)$ – модифицированная функция Бесселя 1-го рода, а $Q_1(\cdot)$ – Q-функция Маркума. Таково распределение амплитуды сигнала при квадратурном приёме.

11. **Распределение Коши.** Если η_1, η_2 – независимые центрированные нормальные случайные величины $\eta_i \sim \mathcal{N}(0,1)$, $i = 1, 2$, то случайная величина $\xi = \eta_1 / \eta_2$ подчинена распределению Коши C(0,1) с ФПВ и ФР:

$$w_\xi(x) = \frac{1}{\pi\gamma\left[1 + \left(\frac{x - x_0}{\gamma}\right)^2\right]}, \quad F_\xi(x) = \frac{1}{\pi}\operatorname{arctg}\left(\frac{x - x_0}{\gamma}\right) + \frac{1}{2},$$

где x_0 и γ – параметры сдвига и масштаба.

Для более детального ознакомления с существующими вероятностными моделями случайных величин и их оценок можно обратиться к книгам [9] – [11].

Пример 6. Погрешность вольтметра в диапазоне измерений лежит в интервале [–10; 10] мВ. Найти: вероятность того, что прибор занижает показания, и вычислить погрешность, вероятность превышения которой равна 0.95.

Решение. Имеем равномерно распределённую случайную величину с параметрами распределения $a = -10$ мВ и $b = 10$ мВ. Прибор занижает показания, если погрешность отрицательна; вероятность этого: $\mathbb{P}(\mathrm{x}<0)=F(0)=1/2$.

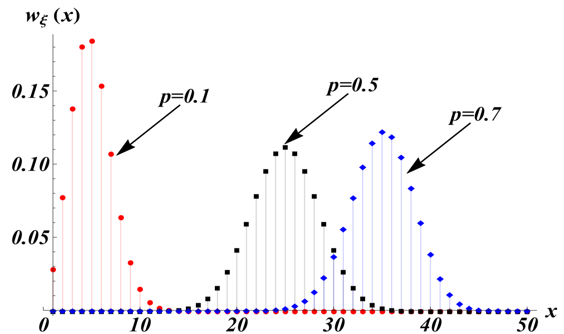
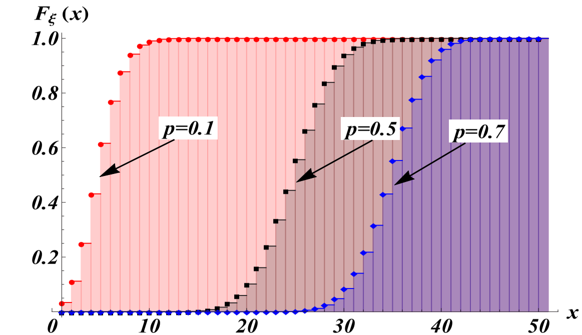
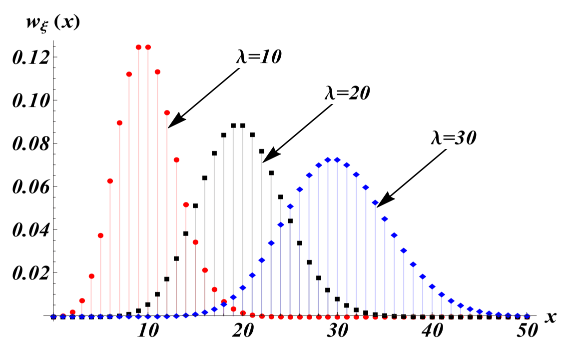
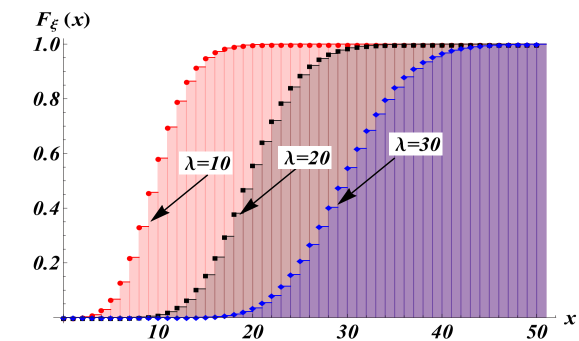
Значение случайной величины p, вероятность превышения которого равна 0.95, находим из условия

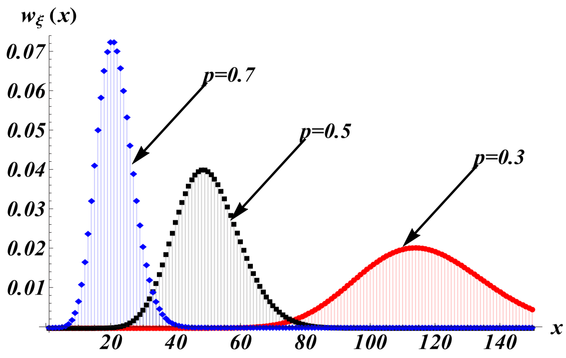
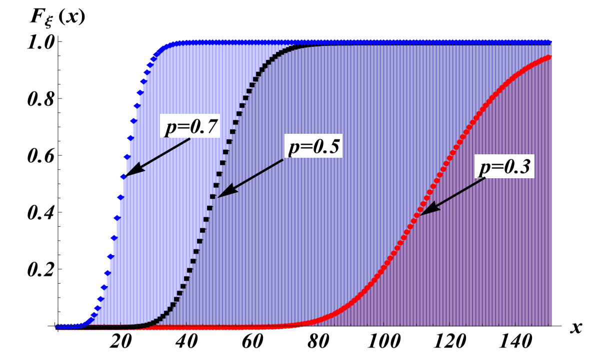
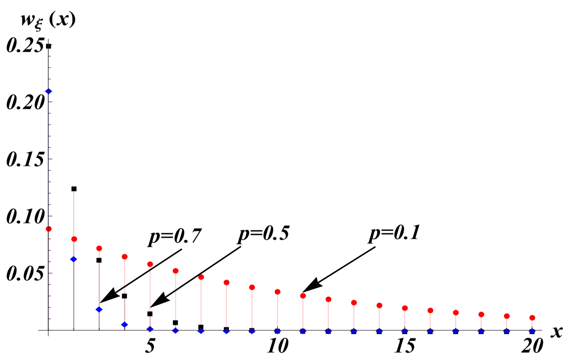
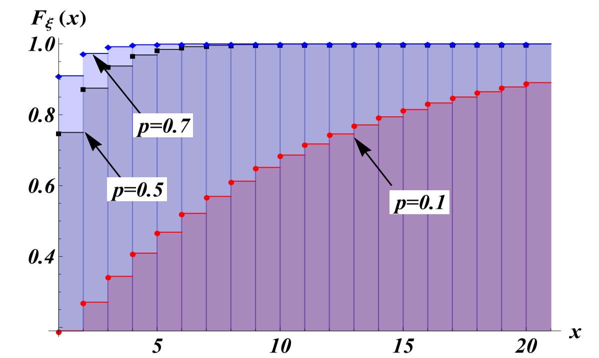
$$\mathbb{P}(\mathrm{x} \le -\mathrm{p} \cup \mathrm{x} \ge \mathrm{p}) = F(-p) + (1 - F(p)) = \frac{-p-a}{b-a} + 1 - \frac{p-a}{b-a} = \frac{-2p-a+b}{b-a} = 0.95,$$

откуда $p = 0.5$ мВ.

Таблица 1.1

Функции вероятности и распределения дискретных распределений

	Функция вероятности	*Функция распределения*
Биномиальное (Бернулли) при $n = 50$		
Пуассона		

Отрицательное биномиальное, $n=50$		
Геометрическое		

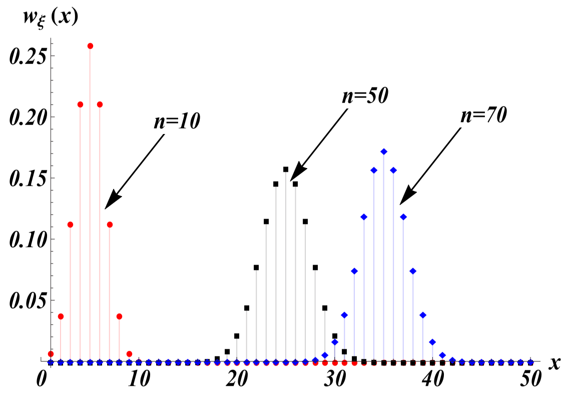
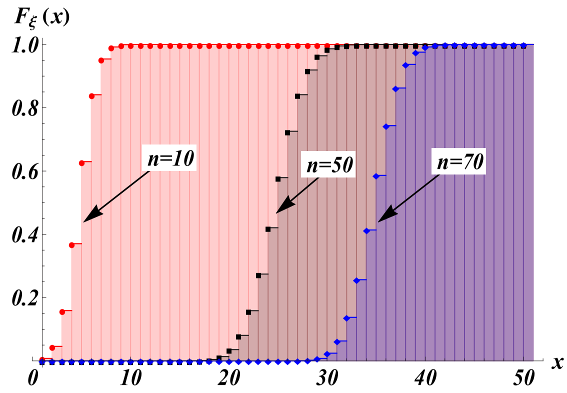
Гипергеометрическое, $n_{tot}=100$, $n_s=50$		

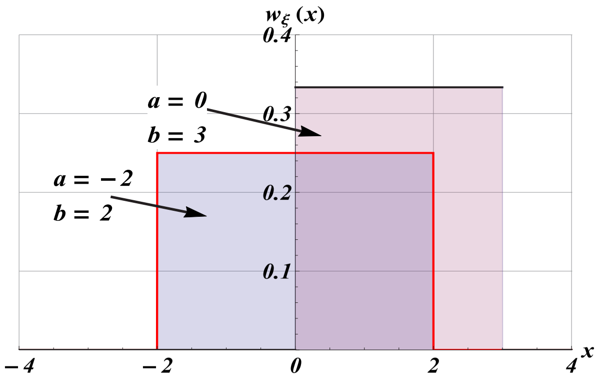
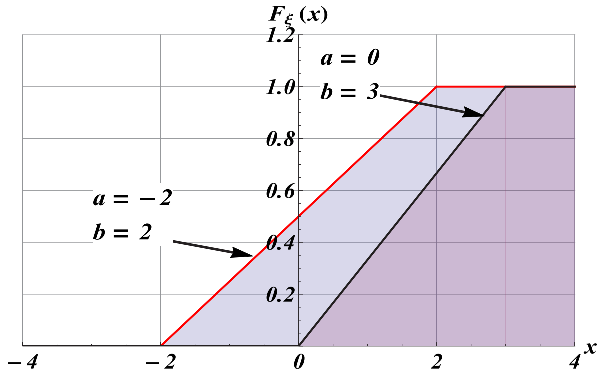
Таблица 1.2

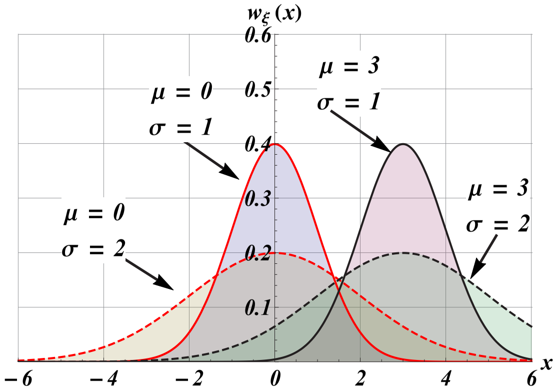
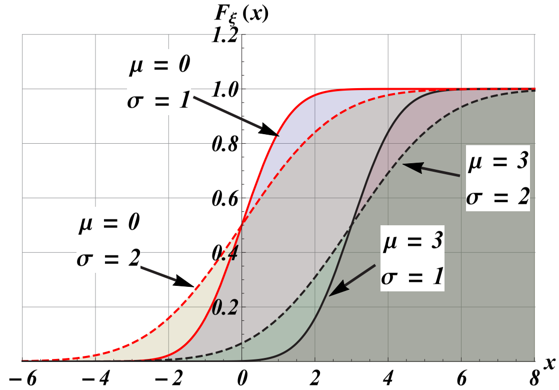
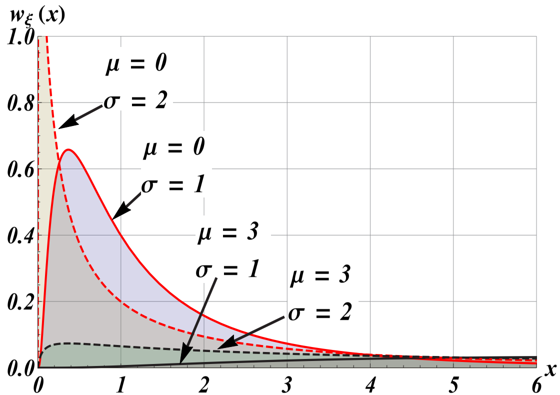
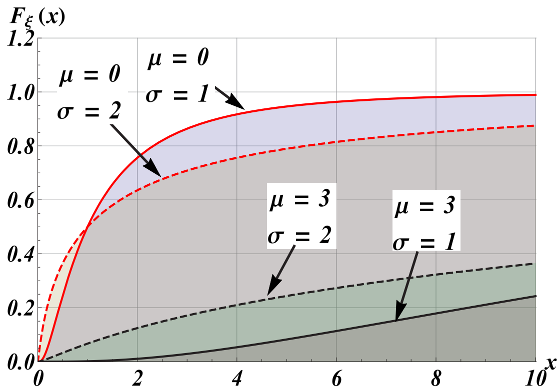
Параметры дискретных распределений

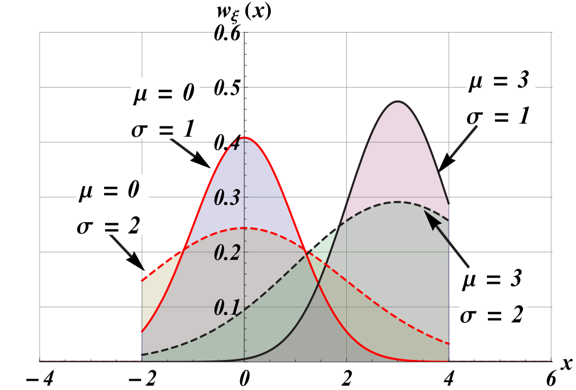
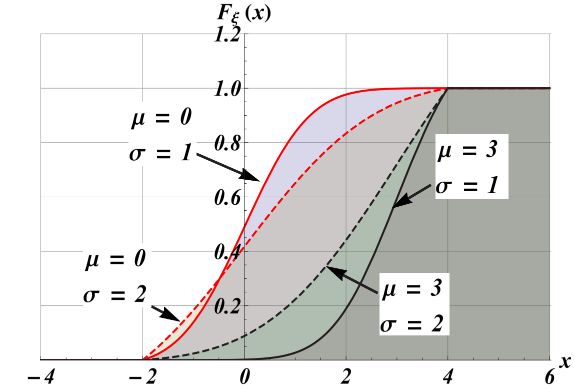
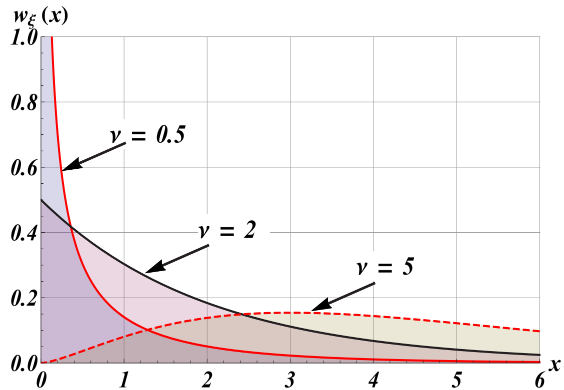
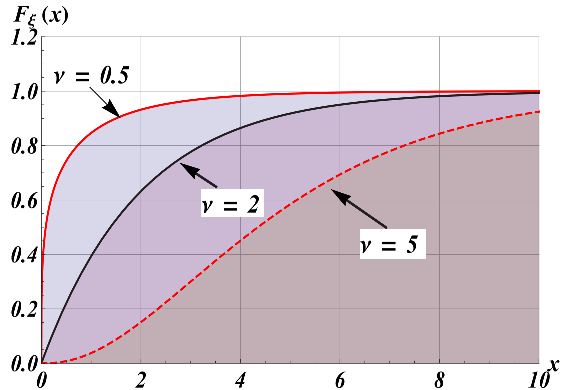
Параметр	*Вид распределения*				
	Бернулли	*Пуассона*	*Отрицательное биномиальное*	*Геометрическое*	*Гипергеометрическое*
Математическое ожидание	np	λ	$m(1-p)/p$	$1/p$	nn_sn_{tot}
Медиана	$(\lfloor np \rfloor - 1) \wedge (\lfloor np \rfloor) \wedge (\lfloor np \rfloor + 1)$	$\approx \lfloor \lambda + 1/3 - 0.02/\lambda \rfloor$	—	—	—
Мода	$\lfloor (n+1)p \rfloor$	$\lfloor \lambda \rfloor$	$\lfloor (m-1)(1-p)/p \rfloor$, $m>0$, 0, если $m \le 1$	1	$\left\lfloor \frac{(n_s+1)(n+1)}{n_{tot}+2} \right\rfloor$
Дисперсия	npq	λ	$\frac{m(1-p)}{p^2}$	$\frac{1-p}{p^2}$	$n(n_s/n_{tot})(1-n_s/n_{tot})(n_{tot}-n)(n_{tot}-1)$
Коэффициент асимметрии	$\frac{1-2p}{\sqrt{np(1-p)}}$	$\lambda^{-1/2}$	$\frac{2-p}{\sqrt{m(1-p)}}$	$\frac{2-p}{\sqrt{1-p}}$	$\frac{(n_{tot}-2n_s)(n_{tot}-1)^{1/2}(n_{tot}-2n)}{[nn_s(n_{tot}-n_s)(n_{tot}-n)]^{1/2}(n_{tot}-2)}$
Коэффициент эксцесса	$\frac{1-6p(1-p)}{np(1-p)}$	λ^{-1}	$\frac{6}{r}+\frac{p^2}{m(1-p)}$	$6+\frac{p^2}{1-p}$	—
Характеристическая функция	$((1-p)+pe^{it})^n$	$\exp(\lambda(e^{it}-1))$	$\left(\frac{p}{1-(1-p)e^{it}}\right)^m$	$\frac{pe^{it}}{1-(1-p)e^{it}}$	$\frac{\genfrac{}{}{0pt}{}{n_{tot}-n_s}{n}}{\genfrac{}{}{0pt}{}{n_{tot}}{n}}\,{}_2F_1(-n,-n_s;n_{tot}-n_s-n+1;e^{it})$

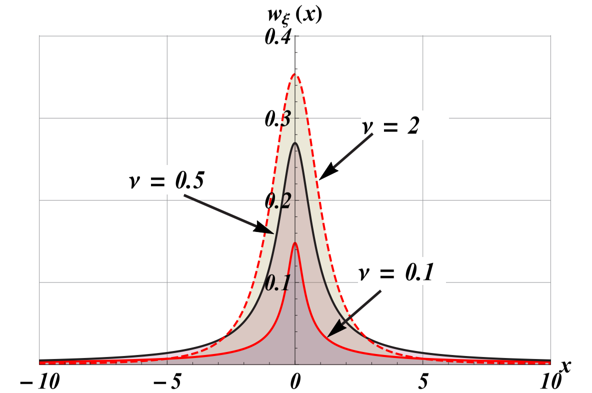
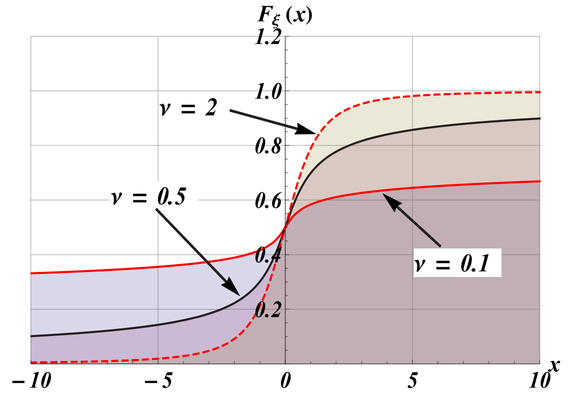
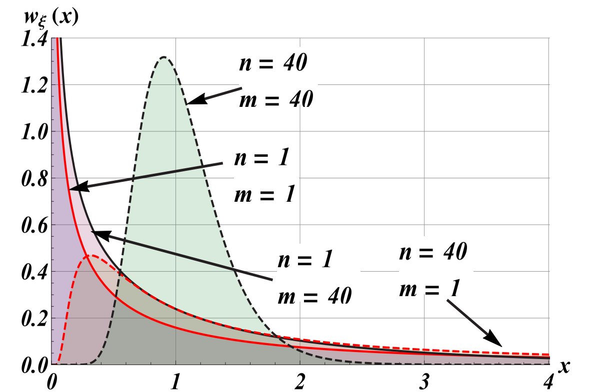
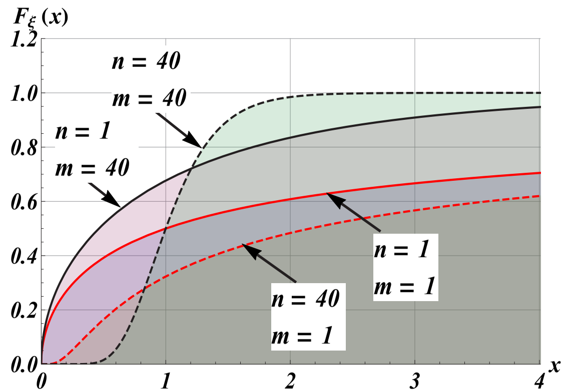
Таблица 1.3

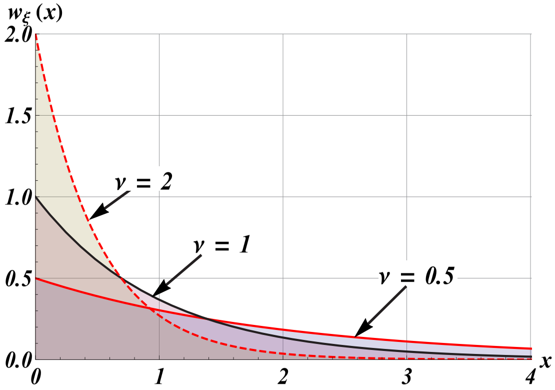
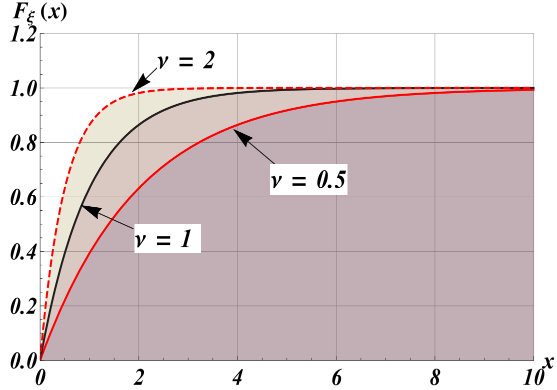
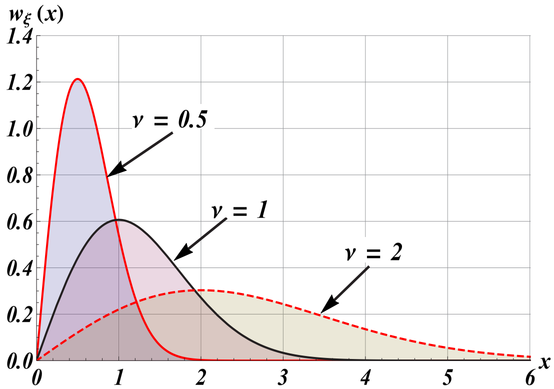
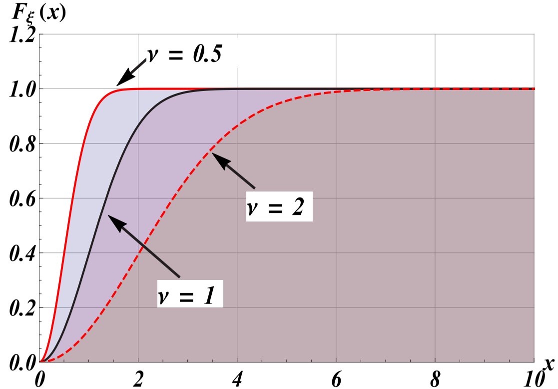
ФПВ и ФР непрерывных распределений

	Функция вероятности	*Функция распределения*
Равномерное		

Нормальное		
Логнормальное		

Усечённое нормальное		
Хи-квадрат		

Стьюдента		
Фишера–Снедекора		

Экспоненциальное		
Рэлея		

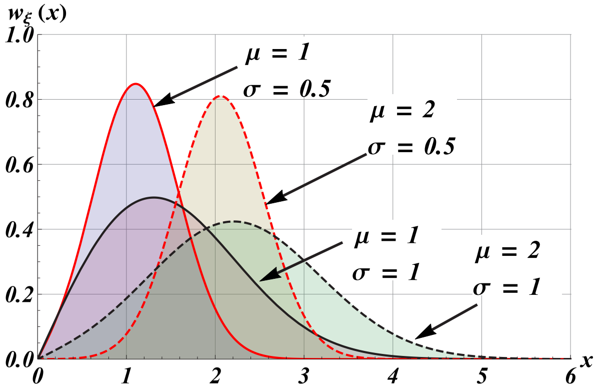
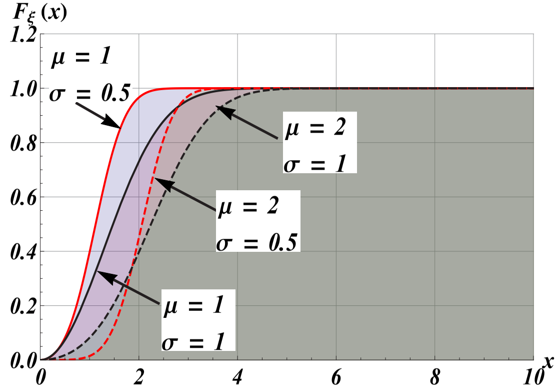
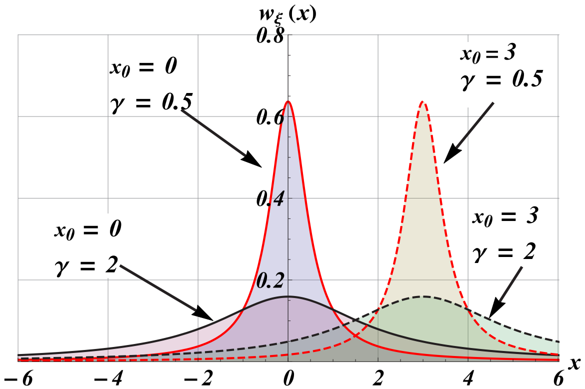
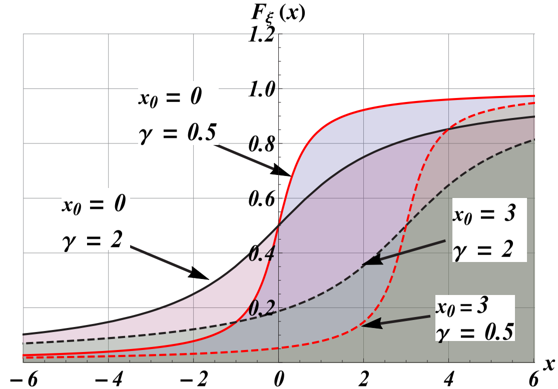
<table>
<tr>
<td>Райса</td>
<td>
</td>
<td>
</td>
</tr>
<tr>
<td>Коши</td>
<td>
</td>
<td>
</td>
</tr>
</table>

Таблица 1.4

Параметры непрерывных распределений

Параметр	*Вид распределения*			
	Равномерное	*Нормальное*	*Логнормальное*	*Геометрическое*
Математическое ожидание	$\frac{1}{2}(a+b)$	μ	$e^{\mu+\sigma^2/2}$	$\mu+\frac{\phi(\frac{a-\mu}{\sigma})-\phi(\frac{b-\mu}{\sigma})}{\Phi(\frac{b-\mu}{\sigma})-\Phi(\frac{a-\mu}{\sigma})}\sigma$
Медиана	$\frac{1}{2}(a+b)$	μ	e^{μ}	
Мода	любой $x \in [a,b]$	μ	$e^{\mu-\sigma^2}$	—
Дисперсия	$\frac{1}{12}(b-a)^2$	σ^2	$(e^{\sigma^2}-1)e^{2\mu+\sigma^2}$	$\sigma^2\left[1+\frac{(\alpha-1)\phi(\alpha)-(\beta-1)\phi(\beta)}{\Phi(\beta)-\Phi(\alpha)}\right]$ $\alpha=(a-\mu)/\sigma, \beta=(b-\mu)/\sigma$
Коэффициент асимметрии	0	0	$(e^{\sigma^2}+2)\sqrt{e^{\sigma^2}-1}$	—
Коэффициент эксцесса	$-6/5$	0	$e^{4\sigma^2}+2e^{3\sigma^2}+3e^{2\sigma^2}-6$	—
Характеристическая функция	$\frac{e^{itb}-e^{ita}}{it(b-a)}$	$\mathrm{e}^{\mu i t-\frac{\sigma^2 t^2}{2}}$	$\sum_{n=0}^{\infty}\frac{(it)^n}{n!}e^{n\mu+n^2\sigma^2/2}$	—

Продолжение таблицы 1.4

Параметр	*Вид распределения*		
	Хи-квадрат	*Стьюдента*	*Фишера-Снедекора*
Математическое ожидание	k	0, если $\nu > 0$, иначе – не определено	$\frac{d_2}{d_2-2}, \quad d_2 > 2$
Медиана	$\approx k(1-\frac{2}{9k})^3$	0	—
Мода	$\max\{k-2, 0\}$	0	$\frac{d_1-2}{d_1}\frac{d_2}{d_2+2}, \quad d_1 > 2$
Дисперсия	$2k$	$\begin{cases} \nu/(\nu-2), & \nu > 2, \\ \infty, & 1 < \nu \le 4, \\ \text{не определена} & \text{иначе} \end{cases}$	$\frac{2d_2^2(d_1+d_2-2)}{d_1(d_2-2)^2(d_2-4)}, \quad d_2 > 4$
Коэффициент асимметрии	$\sqrt{8/k}$	—	—
Коэффициент эксцесса	$12/k$	—	—
Характери-стическая функция	$(1-2it)^{-k/2}$	$\frac{K_{\nu/2}\left(\sqrt{\nu}\lvert t\rvert\right)\cdot\left(\sqrt{\nu}\lvert t\rvert\right)^{\nu/2}}{\Gamma(\nu/2)2^{\nu/2-1}}, \quad \nu > 0$	$\frac{\Gamma(\frac{d_1+d_2}{2})}{\Gamma(\frac{d_2}{2})}U\left(\frac{d_1}{2}, 1-\frac{d_2}{2}, -\frac{d_2}{d_1}it\right)$

Продолжение таблицы 1.4

Параметр	*Экспоненциальное*	*Рэлея*	*Райса*	*Коши*
Математическое ожидание	λ^{-1}	$\sqrt{\frac{\pi}{2}}\sigma$	$\sigma\sqrt{\pi/2}\,L_{1/2}(-\nu^2/2\sigma^2)$	не существует
Медиана	$\lambda^{-1}\ln(2)$	$\sigma\sqrt{\ln(4)}$	—	x_0
Мода	0	σ	—	x_0
Дисперсия	λ^{-2}	$(2-\pi/2)\sigma^2$	$2\sigma^2+\nu^2-\frac{\pi\sigma^2}{2}L_{1/2}^2\left(\frac{-\nu^2}{2\sigma^2}\right)$	не существует
Коэффициент асимметрии	2	$\frac{2\sqrt{\pi}(\pi-3)}{(4-\pi)^{3/2}}$	—	не существует
Коэффициент эксцесса	6	$-\frac{6\pi^2-24\pi+16}{(4-\pi)^2}$	—	не существует
Характеристическая функция	$\left(1-\frac{it}{\lambda}\right)^{-1}$	$1-\sigma t e^{-\sigma^2 t^2/2}\sqrt{\frac{\pi}{2}}\left(erfi\left(\frac{\sigma t}{\sqrt{2}}\right)-\right.$	—	$\exp(x_0 it-\gamma\|t\|)$

1.3. Критерии согласия

Процесс подбора теоретического распределения для экспериментальных данных состоит из трёх основных этапов [5]:

1) подбор вида распределения (гипотетического закона распределения);

2) подбор параметров распределения;

3) проверка гипотезы о согласованности эмпирического и гипотетического законов распределения по ***критериям***:

– общим, применимым к самой общей формулировке гипотезы – о согласии наблюдаемых результатов с любым априорно предполагаемым распределением вероятностей, или

– специальным, проверяющим гипотезы о согласии с определённой формой распределения вероятностей.

Критерии согласия можно разделить на две большие группы (подробнее см. [7], [8]).

1 группа – критерии, основанные на изучении *разницы* между теоретической плотностью распределения и эмпирической гистограммой:

– критерий Пирсона (хи-квадрат);

– критерий числа пустых интервалов;

– квартильный критерий Барнетта–Эйсена.

2 группа – критерии, основанные на *расстоянии* между теоретической и эмпирической ФР. Это критерии:

– Джини;

– Крамера–фон Мизеса;

– Колмогорова–Смирнова;

– Реньи (R-критерий);

– Смирнова–Крамера–фон Мизеса (критерий омега-квадрат);

– Андерсона–Дарлинга;

– Купера;

– Ватсона;

– Фроцини.

Пусть $x_1, \ldots, x_n$ – выборка из распределения случайной величины ξ, задаваемой ФР $F_\xi(x)$. Будем считать, что x_i, $i \in \mathbb{N}$, – независимые случайные величины. ***Выборочной (эмпирической) функцией распределения*** (ЭФР) называется величина [8, 9]

$$F_n^*(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{X_i \le x\}},$$

где $\mathbf{1}_A$ – индикатор события.

Разобьём числовую прямую на $k + 2$ непересекающихся интервала $-\infty < a_0 < a_1 < \ldots < a_i < \ldots < a_{k-1} < a_k < \infty$. ***Выборочной (эмпирической) плотностью распределения*** (***нормализованной гистограммой***) называется величина

$$w_n^*(x) = \frac{n_i}{n\Delta a_i},$$

где $n_i = \sum_{j=1}^{n} \mathbf{1}_{\{X_j \in (a_{i-1}, a_i]\}}$ – число элементов выборки, попавших в i-й интервал $\Delta a_i = a_i - a_{i-1}$.

1.3.1. Критерий Колмогорова

В качестве меры расхождения ЭФР $F_n{}^*(x)$ и предполагаемой ФР $F(x)$ случайной величины ξ возьмём расстояние D_n как точную верхнюю грань множества $|F_n{}^*(x) - F_\xi(x)|$:

$$D_n = \sup_x |F_n^*(x) - F_\xi(x)|.$$

Обозначим H_0 гипотезу о том, что выборка подчиняется распределению $F_\xi(X)$. Тогда по теореме Колмогорова величина $\sqrt{n}D_n$ подчиняется распределению Колмогорова:

$$\forall x > 0: \quad \lim_{n\to\infty} P(\sqrt{n}D_n \leqslant x) = K(x) = \sum_{k=-\infty}^{+\infty} (-1)^k e^{-2k^2x^2}.$$

Принятие решения по критерию Колмогорова заключается в следующем. Если $\sqrt{n}D_n$ превышает процентную точку K_α распределения Колмогорова заданного уровня значимости α, то гипотеза H_0 о соответствии закону $F_\xi(x)$ отвергается. Иначе гипотеза принимается на уровне α:

$$\sqrt{n}D_n \underset{H_1}{\overset{H_0}{\lesseqgtr}} K_\alpha .$$

При $\alpha \approx 1$ K_α можно приблизительно рассчитать по формуле:

$$K_\alpha \approx \sqrt{-\frac{1}{2}\ln\frac{1-\alpha}{2}}.$$

1.3.2. Критерий Пирсона (критерий χ^2)

Пусть H_0 – гипотеза о том, что выборка подчиняется распределению $F_\xi(X)$. В качестве меры расхождения ЭФР $F_n^*(x)$ и предполагаемой ФР $F_\xi(x)$ случайной величины ξ примем статистику:

$$\chi^2 = \sum_{i=1}^{k} \frac{\left(n_i - np_i^{\mathrm{H}_0}\right)^2}{np_i^{\mathrm{H}_0}},$$

где $p_i^{H_0} = F_\xi(x_i) - F_\xi(x_{i-1})$ – предполагаемая вероятность попадания в i-й интервал $[x_{i-1}, x_i)$, n_i – количество реализаций, попавших в него, а k – количество интервалов, используемых для построения ЭФР.

Принятие решения по критерию Пирсона заключается в следующем. Если статистика χ^2 превышает квантиль уровня значимости $1-\alpha$ распределения Пирсона с $k-1$ степенями свободы, то нулевая гипотеза H_0 о соответствии закону $F_\xi(x)$ отвергается. Иначе гипотеза принимается на уровне $1-\alpha$, т. е.

$$\chi^2 \underset{H_1}{\overset{H_0}{\lesseqgtr}} \chi^2_{k-1,1-\alpha} .$$

1.4. Методы оценки параметров

Одной из ключевых задач обработки результатов экспериментов является задача оценки неизвестных параметров. В общем случае её можно сформулировать следующим образом.

Пусть для обработки доступны реализации наблюдений $\vec{\xi} = \left\{\xi_j\right\}_{j=1}^{J}$ измеряемой величины $\vec{s}$, зависящей от совокупности известных параметров $\vec{\lambda}^{(1)} = \left\{\lambda_{i_1}^{(1)}\right\}_{i_1=1}^{I_1}$, неизвестных не интересующих инженера («сопутствующих» или «мешающих») параметров $\vec{\lambda}^{(2)} = \left\{\lambda_{i_2}^{(2)}\right\}_{i_2=1}^{I_2}$ и неизвестных интересующих исследователя параметров $\vec{\lambda}^{(3)} = \left\{\lambda_{i_3}^{(3)}\right\}_{i_3=1}^{I_3}$:

$$\vec{\xi} = f\left(\vec{s}\left(\vec{\lambda}^{(1)}, \vec{\lambda}^{(2)}, \vec{\lambda}^{(3)}\right), \vec{n}\right),$$

где $\vec{n}$ – вектор реализаций шумов, помех или ошибок, присутствующих в эксперименте, а $f(\cdot)$ – оператор, определяющий вид связи $\vec{s}$ и $\vec{n}$ в $\vec{\xi}$.

Под ***процедурой оценки*** неизвестной величины подразумевается выбранный на основе какого-либо критерия (называемого ***критерием качества/оптимальности оценки*** [1], [7] – [9]) способ определения по доступной для наблюдения реализации $\vec{\xi}$ оценочных значений (***оценок***) $\hat{\vec{\lambda}}^{(3)}$ интересующей исследователя совокупности параметров $\vec{\lambda}^{(3)}$ (***точечные оценки*** [9]) или указания интервала, в котором с заданной вероятностью лежат оцениваемые параметры (***интервальная оценка*** [9]). На практике точечные оценки чаще удаётся получить в виде замкнутых аналитических выражений, чем интервальные, поэтому в дальнейшем под оценками будем подразумевать именно точечные оценки.

К основным характеристикам точечных статистических оценок относят [9]:

– ***состоятельность*** оценки, которая проявляется в том, что при неограниченном увеличении объёма выборки наблюдений ($J \to \infty$) оценка должна по вероятности сходиться к истинному значению;

– ***смещение*** оценки $b = \mathrm{E}\left[\hat{\vec{\lambda}}\right] - \vec{\lambda}$, где $\mathrm{E}[\cdot]$ – операция математического усреднения. Если $b = 0$ (т. е. среднее значение оценки совпадает с истинным значением оцениваемого параметра), то говорят, что получаемая оценка ***несмещённая***;

– ***эффективность*** оценки, которая определяется по величине её дисперсии $\mathrm{D}\left[\hat{\vec{\lambda}}\right]$: эффективной называется состоятельная несмещённая оценка с минимальной среди всех других возможных оценок дисперсией.

Наличие или отсутствие информации относительно вероятностно-статистических характеристик $\vec{n}$ и/или вида оператора $f(\cdot)$ позволяет построить метод оценивания $\vec{\lambda}^{(3)}$:

– ***параметрический***, если эта информация доступна**;**

– ***непараметрический***, если нет.

В современной радиофизической и радиотехнической литературе развито много методов оценивания, однако с практической точки зрения наиболее важны семейства **максимально правдоподобных оценок** (МП) и оценок, получаемых на основе **метода моментов** (ММ) [7] – [9], [1].

Пусть исследуемая выборка $\vec{\xi}$ представляет собой совокупность независимых, одинаково распределённых реализаций.

Для получения МП оценки сформируем ***функцию правдоподобия*** $L\left(\vec{\xi} \mid \vec{\lambda}^{(3)}\right)$ (совместную ФПВ вектора выборки) как функцию вектора оцениваемых параметров. Она позволяет качественно сравнить, насколько вероятны распределения $\vec{\xi}$ с разными параметрами $\vec{\lambda}^{(3)}$. ***МП оценкой*** $\hat{\vec{\lambda}}_{МП}$ вектора параметров $\vec{\lambda}^{(3)}$ называется оценка, максимизирующая $L\left(\vec{\xi} \mid \vec{\lambda}^{(3)}\right)$ [1], т. е.

$$\hat{\vec{\lambda}}_{МП} = \arg\max_{\vec{\lambda}^{(3)}} L\left(\vec{\xi} \mid \vec{\lambda}^{(3)}\right).$$

Так как элементы выборки независимы и одинаково распределены, т. е. $L\left(\vec{\xi} \mid \vec{\lambda}^{(3)}\right) = \prod_{j=1}^{J} w\left(\xi_j \mid \vec{\lambda}^{(3)}\right)$, для удобства проведения процедуры максимизации обычно ищется аргумент максимума не самой функции правдоподобия, а её логарифма (т. к. процедура логарифмирования в силу её монотонности не смещает положения экстремума), т. е.

$$\hat{\vec{\lambda}}_{МП} = \arg\max_{\vec{\lambda}^{(3)}} \ln L\left(\vec{\xi} \mid \vec{\lambda}^{(3)}\right) = \arg\max_{\vec{\lambda}^{(3)}} \left(\ln \prod_{j=1}^{J} w\left(\xi_j \mid \vec{\lambda}^{(3)}\right)\right) = \arg\max_{\vec{\lambda}^{(3)}} \left(\sum_{j=1}^{J} \ln w\left(\xi_j \mid \vec{\lambda}^{(3)}\right)\right).$$

Для получения ММ оценки необходимо сформировать систему I_3 уравнений (I_3 – число оцениваемых параметров, т. е. размерность вектора $\vec{\lambda}^{(3)}$), приравняв выражения для теоретических (функционально зависящих от оцениваемых параметров) и выборочных моментов произвольных порядков наблюдаемой величины $\vec{\xi}$. Решение системы относительно $\vec{\lambda}^{(3)}$ даёт искомую ММ оценку $\vec{\lambda}^{(3)}$ $\vec{\lambda}^{(3)}_{ММ}$ [1].

МП оценки проявляют «хорошее» поведение, в особенности при большом объеме выборки (большом времени наблюдения при анализе непрерывных реализаций) или малом уровне шумов

/помех/ошибок (большом отношении сигнал/шум): они являются гарантированно состоятельными, асимптотически эффективными и асимптотически нормальными [1]. ММ оценки легко получать: многие измерительные приборы выдают результат в виде, удобном для подстановки в систему уравнений (например, в виде усреднённого (готовое выборочное среднее) или эффективного (среднеквадратическое) значения, а математические расчёты в ММ отличаются простотой и ясностью [7]. Данные свойства обусловили широкое применение этих двух классов методов.

Однако они имеют и недостатки: неявность и порой чрезмерную вычислительную сложность получения МП оценок и негарантированность достижения желаемого качества (несмещённости, эффективности, а иногда даже и состоятельности) ММ оценок [1].

Пример 7. Наблюдаются флуктуации напряжения $\xi_1,\ldots,\xi_J \sim \mathcal{N}(\mu,\sigma^2)$, представляющие собой J независимых реализаций нормально распределённой случайной величины с неизвестными средним и дисперсией, т. е. $\vec{\lambda}^{(3)} = (\mu,\sigma^2)^{\mathrm{T}}$. Построить МП оценку $\hat{\vec{\lambda}}^{(3)}_{МП} = (\hat{\mu}_{МП},\hat{\sigma}^2_{МП})^{\mathrm{T}}$ и ММ оценку $\hat{\vec{\lambda}}^{(3)}_{ММ} = (\hat{\mu}_{ММ},\hat{\sigma}^2_{ММ})^{\mathrm{T}}$ для неизвестного вектора параметров.

Решение. Логарифм функции правдоподобия принимает вид:

$$L(\vec{\xi} \mid \mu,\sigma^2) = -\frac{J}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{j=1}^{J}(\xi_j-\mu)^2 .$$

Тогда МП оценка двух параметров $\hat{\vec{\lambda}}_{МП} = \arg\max_{\mu,\sigma^2} L\left(\vec{\xi} \mid \mu,\sigma^2\right)$ задаётся решением системы двух уравнений:

$$\begin{cases} \left.\dfrac{\partial}{\partial\mu}L(\vec{\xi} \mid \mu,\sigma^2)\right|_{\substack{\mu=\hat{\mu}_{МП}\\ \sigma^2=\hat{\sigma}^2_{МП}}} = 0 \\ \left.\dfrac{\partial}{\partial\sigma^2}L(\vec{\xi} \mid \mu,\sigma^2)\right|_{\substack{\mu=\hat{\mu}_{МП}\\ \sigma^2=\hat{\sigma}^2_{МП}}} = 0 \end{cases} \Rightarrow \begin{cases} \left.\left(\dfrac{1}{\sigma^2}\left(\sum_{j=1}^{J}\xi_j - J\mu\right)\right)\right|_{\substack{\mu=\hat{\mu}_{МП}\\ \sigma^2=\hat{\sigma}^2_{МП}}} = 0 \\ \left.\left(-\dfrac{J}{2\sigma^2} + \dfrac{1}{2\left(\sigma^2\right)^2}\sum_{j=1}^{J}(\xi_j-\mu)^2\right)\right|_{\substack{\mu=\hat{\mu}_{МП}\\ \sigma^2=\hat{\sigma}^2_{МП}}} = 0 \end{cases} .$$

Откуда выражем $\hat{\mu}_{МП} = \frac{1}{J}\sum_{j=1}^{J}\xi_j$ и $\hat{\sigma}^2_{МП} = \frac{1}{J}\sum_{j=1}^{J}(\xi_j - \mu_{МП})^2$, т. е. МП оценки среднего и дисперсии равны выборочному среднему и выборочной дисперсии соответственно.

Система уравнений метода моментов – равенства выражений для математического ожидания μ и дисперсии σ^2 нормально распределённой величины (см. таблицу 1.4) и соответствующих выборочных моментов μ^*, $(\sigma^2)^*$:

$$\begin{cases} \mu\Big|_{\substack{\mu=\hat{\mu}_{ММ} \\ \sigma^2=\hat{\sigma}^2_{ММ}}} = \mu^*, \\ \sigma^2\Big|_{\substack{\mu=\hat{\mu}_{ММ} \\ \sigma^2=\hat{\sigma}^2_{ММ}}} = \left(\sigma^2\right)^*. \end{cases}$$

Из неё непосредственно следует тот же результат, что и для МП метода.

1.5. Методы анализа экспериментальных данных

Среди широкого спектра методов анализа экспериментальных данных можно особо отметить связку дисперсионного, корреляционного и регрессионного, являющихся последовательными ступенями при изучении связей между случайными величинами [5], [7] – [9].

Дисперсионный анализ устанавливает наличие влияния заданного фактора на изучаемый процесс, отображаемый наблюдаемой статистической совокупностью выборочных данных.

Корреляционный анализ позволяет оценить силу такой связи.

Регрессионный анализ позволяет выбрать конкретную математическую модель и оценить адекватность отражения ею установленной взаимосвязи случайных величин.

1.5.1. Дисперсионный анализ

Дисперсионный анализ применяется для анализа результатов наблюдений, зависящих от различных одновременно действующих факторов, с целью выбора наиболее значимых факторов и оценки их влияния на исследуемый процесс [7].

Влияние различных факторов на изучаемые величины (например, влияние технологического способа изготовления или режима нагрузки на долговечность изделия) приводит к изменению значений параметров распределения вероятностей этих величин – среднего, дисперсии, моментов более высокого порядка.

Изменения дисперсии результатов эксперимента при изменении уровней изучаемого фактора устанавливаются с помощью дисперсионного анализа. Если дисперсии будут отличаться значимо, то влияние фактора на среднее значение наблюдаемой случайной величины считается *значимым*.

Классические методы дисперсионного анализа основываются на следующих предпосылках:

• распределение исходных случайных величин нормально;

• дисперсии экспериментальных данных одинаковы (неразличимы) для всех условий эксперимента (т. е. для экспериментов, выполненных на различных уровнях изучаемого фактора), выполнение которых следует предварительно проверить.

1.5.1.1. Однофакторный дисперсионный анализ

Анализ начинается с построения таблицы анализа факторов (табл. 1.5). Предположим, что анализируется влияние фактора А, изучаемого на k уровнях ($A_1, A_2, \ldots, A_i, \ldots, A_k$). На каждом уровне Ai проведены n наблюдений ($x_{i1}, x_{i2}, \ldots, x_{in}$). Следовательно, всего произведено $N = kn$ наблюдений. Сразу впишем суммы $X_i = \sum_{j=1}^{n} x_{ij}$.

Таблица 1.5

Номер наблюдения	*Уровни фактора А*				
	A_1	$\cdots$	A_i	$\cdots$	A_k
1	x_{11}	$\cdots$	x_{i1}	$\cdots$	x_{k1}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
j	x_{1j}	$\cdots$	x_{ij}	$\cdots$	x_{kj}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	x_{1n}	$\cdots$	x_{in}	$\cdots$	x_{kn}
$\sum$	X_1	$\cdots$	X_i	$\cdots$	X_k

По таблице анализа факторов можно определить основные виды выборочных дисперсий [8]:

– выборочную дисперсию по фактору:

$$S_i^2 = \frac{1}{n-1}\sum_{j=1}^{n}(x_{ij} - \overline{x}_i)^2 = \frac{1}{n-1}\left[\sum_{j=1}^{n}x_{ij}^2 - \frac{1}{n}\left(\sum_{j=1}^{n}x_{ij}\right)^2\right],$$

где $\overline{x} = \frac{1}{n}\sum_{j=1}^{n}x_{ij}$;

– остаточную выборочную дисперсию:

$$S_0^2 = \frac{1}{k}\sum_{i=1}^{k}S_i^2 = \frac{1}{k(n-1)}\left[\sum_{i=1}^{k}\sum_{j=1}^{n}x_{ij}^2 - \frac{1}{n}\sum_{i=1}^{k}\left(\sum_{j=1}^{n}x_{ij}\right)^2\right];$$

– общую выборочную дисперсию:

$$S^2 = \frac{1}{kn-1}\sum_{i=1}^{k}\sum_{j=1}^{n}(x_{ij} - \overline{\overline{x}})^2 = \frac{1}{kn-1}\left[\sum_{i=1}^{k}\sum_{j=1}^{n}x_{ij}^2 - \frac{1}{kn}\left(\sum_{i=1}^{k}\sum_{j=1}^{n}x_{ij}\right)^2\right],$$

где $\overline{\overline{x}} = \frac{1}{kn}\sum_{i=1}^{k}\sum_{j=1}^{n}x_{ij}$,

и проверить выполнение равенства $S_1^2 = S_2^2 = \ldots = S_k^2$ с помощью одного из критериев сравнения, например критерия (см. [7]):

– Бартлетта;

– Кохрана;

– Неймана–Пирсона (критерия отношения правдоподобия);

– Блисса–Кохрана–Тьюки;

– Хартли;

– Кэдуэлла–Лесли–Брауна;

– Самиуддина.

Введём оценку дисперсии S_A^2, характеризующей изменение средних $\overline{x}_i$, связанное с влиянием фактора А:

$$S_A^2 = \frac{n}{k-1}\sum_{i=1}^{k}(\overline{x}_i - \overline{\overline{x}})^2.$$

Дисперсии S_A^2 и S_0^2 как суммы квадратов совместно независимых стандартных нормальных случайных величин подчиняются распределению хи-квадрат с $k-1$ и $N-k$ степенями свободы соответственно, а дисперсионное отношение S_A^2 / S_0^2 – распределению Фишера–Снедекора с этими же степенями свободы $\mathrm{F}(k-1,\ N-k)$. Поэтому проверка влияния фактора A на изменение средних может быть сведена к сравнению дисперсий S_A^2 / S_0^2: влияние фактора A признается значимым с вероятностью α (значимостью $1-\alpha$), если

$$\frac{S_A^2}{S_0^2} > F_\alpha(k-1, k(n-1)),$$

где $F_\alpha(\cdot,\cdot)$ – α-квантиль F-распределения (можно найти, используя таблицы, например, в [7]).

Алгоритм определения влияния фактора ***A*** **при использовании дисперсионного анализа** можно свести к следующим этапам.

1 этап: вычисляем вспомогательные суммы

$$Q_1 = \sum_{i=1}^{k}\sum_{j=1}^{n} x_{ij}^2; \quad Q_2 = \frac{1}{n}\sum_{i=1}^{k} X_i^2; \quad Q_3 = \frac{1}{N}\left(\sum_{i=1}^{k} X_i\right)^2.$$

Если на различных уровнях A_i фактора A проводятся различные количества наблюдений n_i, то вычисление суммы Q_2 ведётся по формуле:

$$Q_2 = \sum_{i=1}^{k} \frac{X_i^2}{n_i}.$$

2 этап: находим выборочные дисперсии

$$S_0^2 = \frac{Q_1 - Q_2}{N-k}; \quad S_A^2 = \frac{Q_2 - Q_3}{k-1}.$$

3 этап: устанавливаем наличие влияния фактора A:

– если $\dfrac{S_A^2}{S_0^2} > F_\alpha(k-1, k(n-1))$, то влияние фактора A признается значимым;

– в противном случае всю выборку наблюдений можно считать однородной с общей дисперсией $S^2 = \dfrac{Q_1 - Q_3}{kn-1}$.

Пример. Провести дисперсионный анализ данных, представленных табл. 1.6, при доверительной вероятности $\alpha = 0{,}95$.

Таблица 1.6

Номер наблюдения	*Уровни фактора A*				
	A_1	A_2	A_3	A_4	A_5
1	3.2	2.6	2.9	3.6	3.0
2	3.1	3.1	2.6	3.4	3.4
3	3.1	2.7	3.0	3.2	3.2
4	2.8	2.9	3.1	3.3	3.5
5	3.3	2.7	3.0	3.5	2.9
6	3.0	2.8	2.8	3.3	3.1
$\sum$	18.5	16.8	17.4	20.3	19.1

Вычисляем ($k = 5$, $n = 6$, $N = 30$):

$$Q_1 = \sum_{i=1}^{5}\sum_{j=1}^{6} x_{ij}^2 = 284.87, \quad Q_2 = \frac{1}{6}\sum_{i=1}^{5} X_i^2 = 284.025; \quad Q_3 = \frac{1}{5\cdot 6}\left(\sum_{i=1}^{5} X_i\right)^2 = 282.747;$$

$$S_0^2 = \frac{284.87 - 284.025}{5\cdot(6-1)} = 0.0338; \quad S_A^2 = \frac{284.025 - 282.747}{5-1} = 0.319; \quad \frac{S_A^2}{S_0^2} = 9.45.$$

0,95-квантиль распределения Фишера со степенями свободы $k - 1 = 4$ и $N - k = 25$ равен $F_{0,95}(4;25) = 2.8$. Так как $S_A^2 / S_0^2 = 9.45 > F_{0,95}(4;25) = 2.8$, то влияние фактора A на поведение наблюдаемых случайных величин следует признать значимым.

1.5.1.2. Двухфакторный дисперсионный анализ

При учёте влияния двух и более факторов проводится двухфакторный (многофакторный) дисперсионный анализ, что позволяет установить не только влияние, но и взаимовлияние факторов.

Предположим, что анализируется влияние фактора A, изучаемого на k уровнях ($A_1, A_2, \ldots, A_k$), и фактора B, изучаемого на m уровнях ($B_1, B_2, \ldots, B_m$). На каждом уровне A_i, $i = 1, \ldots, k$ проведены n наблюдений ($x_{i1}, x_{i2}, \ldots, x_{iv}, \ldots, x_{in}$).

Алгоритм дисперсионного анализа двухфакторных таблиц можно свести к следующим этапам [7], [8].

1 этап: вычисляем суммы:

$$Q_1 = \sum_{i=1}^{k}\sum_{j=1}^{m} x_{ij}^2 ;$$

$$Q_2 = \frac{1}{m}\sum_{i=1}^{k} X_i^2 \text{, где } X_i = \sum_{j=1}^{m} x_{ij} ;$$

$$Q_3 = \frac{1}{k}\left(\sum_{j=1}^{m} X_{j'}\right)^2 \text{, где } X_{j'} = \sum_{i=1}^{k} x_{ij} ;$$

$$Q_4 = \frac{1}{km}\left(\sum_{i=1}^{k} X_i\right)^2 = \frac{1}{km}\left(\sum_{j=1}^{m} X_{j'}\right)^2 ;$$

$$Q_5 = \sum_{i=1}^{k}\sum_{j=1}^{m}\sum_{\nu=1}^{n} x_{ij\nu}^2 .$$

2 этап: находим

$$S_0^2 = \frac{Q_1 + Q_4 - Q_2 - Q_3}{(k-1)(m-1)};$$

$$S_A^2 = \frac{Q_2 - Q_4}{k-1};$$

$$S_B^2 = \frac{Q_3 - Q_4}{m-1};$$

$$S_{AB}^2 = \frac{Q_5 - nQ_1}{mk(n-1)}.$$

3 этап: устанавливаем наличие влияния факторов и их взаимодействия:

– при $\dfrac{S_A^2}{S_0^2} > F_\alpha(k-1,(k-1)(m-1))$ влияние фактора *A* признается значимым;

– при $\dfrac{S_B^2}{S_0^2} > F_\alpha(m-1,(k-1)(m-1))$ влияние фактора *B* признается значимым;

– при $\dfrac{nS_0^2}{S_{AB}^2} > F_\alpha((m-1)(k-1), mk(n-1))$ взаимодействие факторов *A* и *B* признается значимым.

1.5.2. Корреляционный анализ

Корреляционный анализ позволяет установить наличие зависимости между случайными величинами и количественно оценить степень неслучайности их совместного изменения.

Пусть результат детерминированного преобразования $\varphi(\cdot)$ случайной величины X измеряется в присутствии набора случайных факторов ε: $Y = \varphi(X) + \varepsilon$. Тогда изменение случайной величины Y, соответствующее изменению величины X, разбивается на две составляющие:

– ***стохастическую***, связанную с неслучайной зависимостью Y от X;

– ***статистическую*** (случайную), связанную со случайным характером поведения X и Y.

Связь между X и Y может быть ***линейной*** или ***нелинейной***.

Для независимых случайных величин выполняется равенство [1]:

$$\mathrm{D}[X+Y] = \mathrm{D}[X] + \mathrm{D}[Y],$$

где $\mathrm{D}[\cdot]$ – оператор вычисления дисперсии, однако в общем случае:

$$\mathrm{D}[X+Y] = \mathrm{E}[(X+Y-(\mathrm{E}[X]+\mathrm{E}[Y]))^2] = \mathrm{D}[X] + \mathrm{D}[Y] + 2\underbrace{\mathrm{E}[(X-\mathrm{E}[X])(Y-\mathrm{E}[Y])]}_{\mathrm{cov}[X,Y]},$$

т. е. существует ковариация между X и Y:

$$\mathbf{cov}[X,Y] = \mathbf{E}\left[(X-\mathbf{E}[X])(Y-\mathbf{E}[Y])\right] = \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y],$$

по величине которой можно судить о наличии зависимости между X и Y. На практике удобнее работать с *относительной долей* этой связи.

Доля стохастической составляющей *в линейной связи* между X и Y характеризуется ***коэффициентом корреляции*** [7]:

$$\rho = \frac{\mathrm{E}[(X-\mathrm{E}[X])(Y-\mathrm{E}[Y])]}{\sqrt{\mathrm{D}[X]\mathrm{D}[Y]}}.$$

Так как $|\,\mathbf{cov}[X,Y]\,| \leqslant \sqrt{\mathbf{D}[X]\mathbf{D}[Y]}$, то

$$-1 < \rho = \frac{\mathbf{cov}[X,Y]}{\sqrt{\mathbf{D}[X]\mathbf{D}[Y]}} < 1 \cdot$$

Коэффициент корреляции показывает, насколько связь между случайными величинами близка к строго линейной. Для нормально распределённых величин X и Y при $\rho = 0$ линейная связь между ними отсутствует, при $\rho = \pm 1$ связь между ними строго линейна (знак указывает на направление связи).

Меру *нелинейной связи* между случайными величинами характеризуют ***корреляционным отношением К. Пирсона*** [7] – [9].

При одновременном наблюдении X и Y, когда на каждом уровне одной переменной x_i наблюдаются n_i значений другой переменной y_{ij} ($j = 1, ..., n_i$), корреляционное отношение определяется следующим образом:

$$\eta_{YX}^2 = \frac{S_0^2}{S^2},$$

где S_0^2 – дисперсия рассеяния y_{ij}, связанная с влиянием группировки значений y_{ij} по i уровням x, S^2 – дисперсия рассеяния значений y_{ij} без учета их группировки по уровням переменной x.

Перестановкой переменных можно определить корреляционное отношение X по Y η^2_{XY} (когда на каждом уровне переменной y_i наблюдается группа значений другой переменной x_{ij}). В общем случае $\eta^2_{XY} \neq \eta^2_{YX}$, $0 \leq \rho^2 \leq \eta^2 \leq 1$.

На практике возможны следующие ситуации:

– если Y и X связаны строго линейно, то $\rho^2 = \eta^2 = 1$;

– если между ними существует линейная стохастическая связь, то $\rho^2 = \eta^2 < 1$;

– при нелинейной стохастической связи $\rho^2 < \eta^2 < 1$.

Предположим, что мы имеем n пар нормально распределенных случайных величин x_i и y_i – (x_1, y_1), …, (x_i, y_i), …, (x_n, y_n). Определим $\overline{x} = \sum_{i=1}^{n} x_i$ и $\overline{y} = \sum_{i=1}^{n} y_i$. ***Выборочной оценкой коэффициента корреляции*** ρ является случайная величина

$$r = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \sum_{i=1}^{n}(y_i - \overline{y})^2}} = \frac{\sum_{i=1}^{n} x_i y_i - n\overline{xy}}{\sqrt{\left(\sum_{i=1}^{n} x_i^2 - n\overline{x}\right)\left(\sum_{i=1}^{n} y_i^2 - n\overline{y}\right)}}.$$

При малых n ($n < 15$) лучшей оценкой коэффициента корреляции является $r^* = r\left[1 + \frac{1 - r^2}{2(n-3)}\right]$.

В ходе корреляционного анализа также проверяют гипотезы о существенности корреляционной связи и о возможных значениях коэффициента корреляции, для чего используют аппроксимации распределения выборочного коэффициента корреляции. Некоторые часто применяющиеся аппроксимации r и связанных с ним величин [7]:

– при $n > 200$ распределение выборочного коэффициента корреляции аппроксимируется нормальным законом со средним $M(r) = \rho$ и дисперсией $\mathrm{D}(r) = \dfrac{1-\rho^2}{n-1}$;

– при $n > 5$ распределение величины $z = \dfrac{1}{2}\ln\left(\dfrac{1-r}{1+r}\right) = \operatorname{arcth}(r)$ аппроксимируется нормальным распределением с параметрами $\mathrm{M}(z) = \dfrac{1}{2}\ln\left(\dfrac{1-\rho}{1+\rho}\right) = \operatorname{arcth}(r)$ и $\mathrm{D}(z) = \dfrac{1}{n-3}$;

– при $n > 10$ распределение случайной величины $\tau = \dfrac{(r-\rho)\sqrt{n-2}}{\sqrt{(1-r^2)(1-\rho)}}$ аппроксимируется распределением Стьюдента с $f = n - 2$ степенями свободы.

Часто встречается ***задача проверки гипотезы о значимости корреляционной связи*** между случайными величинами, т. е. значимости отклонения коэффициента корреляции ρ от нуля. Выборочное значение r коэффициента корреляции сравнивается с его критическим значением r_α – α-квантилью распределения r при $\rho = 0$ (см. табл. 1.7). Корреляция между случайными величинами признается *значимой* с *доверительной вероятностью* α или *уровнем значимости* $1 - \alpha$, если $|r| > r_\alpha$, т. е. нулевая гипотеза $H_0 : |\rho| = 0$ принимается, а её альтернатива $H_0 : |\rho| \neq 0$ отвергается:

$$|r| \underset{H_0}{\overset{H_1}{\gtrless}} r_\alpha .$$

На основе вышеописанных аппроксимаций получаются оценки r_α [7]:

при $n > 5$

$$r_\alpha = \frac{\exp\left(\dfrac{2}{\sqrt{n-3}} u_{\frac{1+\alpha}{2}}\right) - 1}{\exp\left(\dfrac{2}{\sqrt{n-3}} u_{\frac{1+\alpha}{2}}\right) + 1},$$

где u_α – α-квантиль стандартного нормального распределения;

при $n > 10$

$$r_\alpha = \sqrt{\frac{t^2_{\frac{1+\alpha}{2}}}{n-2+t^2_{\frac{1+\alpha}{2}}}},$$

где t_α – α-квантиль распределения Стьюдента с $f = n - 2$ степенями свободы;

при $n > 200$

$$r_\alpha = \frac{1}{\sqrt{n-1}} u_{\frac{1+\alpha}{2}}.$$

Таблица 1.7

Критические значения значения выборочного коэффициента корреляции для ρ = 0 [7]

n	*Доверительная вероятность α*			n	*Доверительная вероятность α*		
	0.90	0.95	0.99		0.90	0.95	0.99
3	0.988	0.997	1.000	13	0.476	0.553	0.684
4	0.900	0.950	0.990	14	0.457	0.532	0.661
5	0.805	0.878	0.959	15	0.441	0.514	0.641
6	0.729	0.811	0.917	16	0.426	0.497	0.623
7	0.669	0.754	0.874	17	0.412	0.482	0.606
8	0.621	0.707	0.834	18	0.400	0.468	0.590
9	0.582	0.666	0.798	19	0.389	0.456	0.575
10	0.549	0.632	0.765	20	0.378	0.444	0.561
11	0.521	0.602	0.735	21	0.369	0.433	0.549
12	0.497	0.576	0.708	22	0.360	0.423	0.537

Пусть на каждом i-ом из k уровней случайной величины X $(x_1, x_2, \ldots, x_k)$ наблюдаются $n > 5_{\text{i}}$ значений случайной величины Y y_{ij}, $j = 1, \ldots, n_i$ и общее количество наблюдений равно $N = \sum_{i=1}^{k} n_i$. Определим: $\bar{y}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} y_{ij}$, $\bar{y} = \frac{1}{k}\sum_{i=1}^{k}\bar{y}_i$.

***Выборочной оценкой корреляционного отношения** Y **по** X* называется величина:

$$\eta^2_{YX} = \frac{\sum_{i=1}^{k} n_i(\bar{y}_i - \bar{y})^2}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y})^2} = \frac{\sum_{i=1}^{k} n_i\bar{y}_i^2 - N\bar{y}^2}{\sum_{i=1}^{k}\sum_{j=1}^{n_i} y_{ij}^2 - N\bar{y}^2}.$$

При оценке наличия существенной нелинейной связи между переменными Y и X решающая статистика

$$l = \frac{\eta^2 (N-k)}{(k-1)(1-\eta^2)}$$

сравнивается с α-квантилью распределения Фишера с $f_1 = k - 1$ и $f_2 = N - k$ степенями свободы:

$$l \underset{H_0}{\overset{H_1}{\gtrless}} F_\alpha(f_1, f_2)$$

и, если $l > F_\alpha$, то гипотеза $H_0 : \eta^2 = 0$ отвергается с достоверностью α, а альтернатива $H_1 : \eta^2 \neq 0$ принимается.

При проверке существенности линейной связи между переменными решающая статистика

$$l^* = \frac{(\eta^2 - \rho^2)(N-k)}{(k-2)(1-\eta^2)}$$

сравнивается с α-квантилью распределения Фишера с $f_1^* = k - 1$ и $f_2^* = N - k$ степенями свободы:

$$l^* \underset{H_0}{\overset{H_1}{\gtrless}} F_\alpha(f_1^*, f_2^*)$$

и при $l^* > F_\alpha$ гипотеза H_0 о линейности связи отвергается с вероятностью α.

Пример. Результаты эксперимента по снятию амплитудной характеристики (зависимости $U_{вых}$ от $U_{вх}$) некоторого устройства таковы ($n = 10$):

$$U_{вх}, \text{ мВ} : 2 \;\; 4 \;\; 1 \;\; 7 \;\; 3 \;\; 11 \;\; 14 \;\; 15 \;\; 21 \;\; 5$$
$$U_{вых}, \text{ мВ} : 7 \;\; 6 \;\; 4 \;\; 11 \;\; 2 \;\; 21 \;\; 29 \;\; 23 \;\; 40 \;\; 13$$

Необходимо проверить гипотезу о наличии линейной связи между случайными $U_{вых}$ и $U_{вх}$ с достоверностью α =0,95.

Находим оценки коэффициента корреляции:

$$r = \frac{717.2}{\sqrt{398.1 \cdot 1372}} = 0.97; \qquad r^* = 0.97 \cdot \left(1 + \frac{1 - 0.97^2}{2 \cdot (10-3)}\right) = 0.974 .$$

Из таблицы 1.7 для $n = 10$ и α =0,95 находим критическое значение $r_{0,95}$ = 0.632. Так как $r > r_{0,95}$ (и $r^* > r_{0,95}$), то с достоверностью α =0,95 линейная связь между $U_{вых}$ и $U_{вх}$ имеется и значима.

2. Задачи и методические указания к их решению

№ 1. Приведите пример качественного радиотехнического эксперимента. Поясните, какие явления устанавливаются, каким образом. Дайте словесное описание ожидаемых результатов.

№ 2. Мобильный телефон с некоторым программным обеспечением может использоваться как измеритель уровня принимаемого сигнала в выбранном частотном канале. Приведите пример эксперимента, который можно поставить с использованием такого оборудования. Выделите и охарактеризуйте возможные участвующие факторы. Запишите вид функции отклика. К каким классам относится этот эксперимент? Какова, на ваш взгляд, достижимая точность результатов?

№ 3. Одна из программ анализа компьютерных сетей позволяет формировать передаваемые пакеты данных, анализировать принятые пакеты и считать количество неверно принятых или потерянных пакетов. Приведите пример эксперимента, который можно поставить с использованием такого оборудования. Выделите и охарактеризуйте возможные участвующие факторы. Запишите вид функции отклика. К каким классам относится этот эксперимент?

№ 4. Требуется имитировать работу некоторого приёника в условиях помех (гипотетически – гауссовы помехи) и собственных шумов (гипотетически – стандартный нормальный процесс). Спланируйте такой статистический эксперимент:

– определите наблюдаемую величину и её измеряемые характеристики;

– оцените необходимые объёмы выборок;

– выберите средства генерации выборок помех и шумов;

– определите средство проверки статистических характеристик выборок на соответствие гипотетическим распределениям;

– опишите ход эксперимента;

– опишите средства и методику обработки результатов: факторы, целевую функцию, требуемую точность, методы обработки, форму представления результатов.

№ 5. Дано эмпирическое распределение вероятности правильного обнаружения цели (в процентах по отношению к системе предыдущего поколения) системой нового поколения (табл. 2.1). Подберите соответствующее теоретическое распределение и на уровне значимости $1 - \alpha = 0.05$ проверьте гипотезу о согласованности этих двух распределений.

Указание 1: построив гистограмму, убедиться, что в качестве модельного можно использовать нормальное распределение.

Указание 2: для определения вероятности попадания реализации случайной нормальной величины $p_i^{H_0}$ в интервал $[x_i, x_{i+1}]$ воспользоваться функцией Лапласа:

$$P(x_i \le \xi \le x_{i+1}) = \frac{1}{2}\left(\Phi\left(\frac{x_{i+1} - \mu}{\sigma} \right) - \Phi\left(\frac{x_i - \mu}{\sigma} \right) \right).$$

Таблица 2.1

i	*Интервал* $[x_i, x_{i+1}]$	*Эмпирические частоты*	*Вероятности нормального распределения* $p_i^{H_0}$
1	94 – 100	3	0.017
2	100 – 106	7	0.059
3	106 – 112	11	0.141
4	112 – 118	20	0.228
5	118 – 124	28	0.247
6	124 – 130	19	0.182
7	130 – 136	10	0.087
8	136 – 142	2	0.029
	Σ	100	0.99

№ 6. Решить задачу № 5, пользуясь критерием Колмогорова с достоверностью 95 %. В качестве основной принять гипотезу H_0 о нормальном распределении выборки с математическим ожиданием $\mu = 119.2$ и дисперсией $\sigma^2 = 87.48$.

Указание 1: воспользоваться приведёнными в табл. 2.2 критическими значениями λ_α уровня значимости $1 - \alpha$, т. е. уровней, определяемых из условия $P(\lambda_\alpha) = 1 - \alpha$.

Таблица 2.2

$1 - \alpha$	0.40	0.30	0.20	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
λ_α	0.89	0.97	1.07	1.22	1.36	1.48	1.63	1.73	1.95	2.03

Указание 2: для построения теоретической функции распределения нормальной случайной величины воспользоваться значениями функции Лапласа из табл. 2.3: $F(x) = \frac{1}{2} + \frac{1}{2}\Phi\left(\frac{x\_\mu}{\sigma}\right)$.

Таблица 2.3

x	94	100	106	112	128	124	130	136	142
$F_n^*(x)$	0.010	0.030	0.100	0.210	0.410	0.690	0.880	0.980	1.000
$F_n(x)$	0.004	0.021	0.800	0.221	0.449	0.695	0.878	0.964	0.993

№ 7. Результаты наблюдения вероятности обнаружения цели в зависимости от времени на РЛС дальнего обзора приведены в табл. 2.4. Здесь $i = 0, 1, \ldots, 11$ – номер временного двухчасового интервала анализа, k_i – количество целей, обнаруженных в i-й интервал. Общий объём обнаруженных за сутки целей составил 500 единиц. Проверить по уровню значимости $1 - \alpha = 0.05$, согласуются ли эти данные с гипотезой о равномерном распределении числа обнаруживаемых целей в диапазоне (0; 12).

Таблица 2.4

i	0	1	2	3	4	5	6	7	8	9	10	11
k_i	41	34	54	39	49	45	41	33	37	41	48	39

№ 8. Результаты измерений уровня сигнала на выходе приёмника для 7 значений уровня входного сигнала приведены в табл. 2.5. Провести однофакторный дисперсионный анализ данных при доверительной вероятности $\alpha = 0.98$.

Таблица 2.5

Номер наблюдения	*Фактор*						
	A_1	A_2	A_3	A_4	A_5	A_6	A_7
1	4.04	4.11	4.36	4.35	4.21	4.34	4.58
2	4.18	4.02	4.19	4.42	4.28	4.31	4.28
3	4.00	3.96	3.95	4.26	4.37	4.34	4.48
4	4.07	3.99	4.28	4.29	4.36	4.35	4.54
5	4.15	4.12	4.13	4.25	4.23	4.45	4.52
6	3.98	4.15	4.23	4.36	4.35	4.44	4.60
7	3.87	4.04	4.38	4.35	4.42	4.30	4.51
8	3.92	4.18	4.00	4.13	4.29	4.43	4.42
9	4.04	4.08	4.15	4.21	4.32	4.40	4.59
10	4.04	4.06	4.17	4.30	4.40	4.49	4.55

Указание: предполагается, что исходные данные прошли тест на нормальность по критерию Шапиро–Уилка и тест на на равенство дисперсий для различных уровней фактора *A* по критерию Кохрана.

№ 9. В эксперименте измерялась ёмкость полупроводникового элемента, нФ, в зависимости от приложенного напряжения *U* (табл. 2.6). Определить наличие влияния *U* на результаты измерений на уровне значимости $1 - \alpha = 0.05$.

Указание: предполагается, что исходные данные прошли тест на нормальность по критерию Колмогорова–Смирнова и тест на равенство дисперсий для различных уровней фактора *U* по критерию Бартлетта.

Таблица 2.6

Номер наблюдения	*Фактор*				
	U_1	U_2	U_3	U_4	U_5
1	1.00	1.10	1.06	1.17	1.25
2	0.963	0.936	0.964	0.981	1.18
3	0.899	1.11	0.999	0.965	1.13
4	1.08	0.910	1.12	1.05	1.18
5	1.02	0.825	0.955	1.18	1.17
6	1.15	1.13	1.07	1.19	1.27
7	0.966	0.980	1.04	1.15	1.22

№ 10. Для экспериментальных данных из табл. 2.7 определить наличие влияния фактора Ñ на результаты измерений с достоверностью $\alpha = 0.9$.

Указание: считать, что исходные данные прошли тест на нормальность по критерию χ^2 и тест на равенство дисперсий для различных уровней фактора Ñ по критерию Кохрана.

Таблица 2.7

Номер наблюдения	*Фактор*			
	$\tilde{N}_1$	$\tilde{N}_2$	$\tilde{N}_3$	$\tilde{N}_4$
1	8.93	8.48	8.56	8.53
2	8.99	8.54	8.35	8.49
3	8.91	8.45	8.60	8.53
4	9.11	8.56	8.46	8.49
5	8.91	8.36	8.62	8.65
6	8.97	8.39	8.70	8.47
7	9.24	8.65	8.65	8.54
8	8.96	8.49	8.30	8.52

№ 11. Провести двухфакторный дисперсионный анализ данных, представленных в табл. 2.8, при доверительной вероятности $\alpha = 0.95$.

Указание: предполагается, что исходные данные прошли тест на нормальность по критерию Шапиро–Уилка и тест на равенство дисперсий для различных уровней фактора *A* и *B* по критерию Кохрана.

Таблица 2.8

Фактор B	*Фактор A*								
	A_1			A_2			A_3		
B_1	3.6	3.8	4.1	2.9	3.1	3.0	2.6	2.5	2.9
B_2	4.2	4.0	4.1	3.3	2.9	3.0	3.7	3.5	3.6
B_3	3.8	3.5	3.6	3.6	3.7	3.5	3.2	3.0	3.4
B_4	3.4	3.2	3.2	3.4	3.6	3.5	3.6	3.8	3.7

№ 12. Для экспериментальных данных из табл. 2.9 (выходные ВАХ некоторого биполярного транзистора) определить наличие влияния факторов $U_{кэ}$ и $I_Б$ на результаты измерений, а также их взаимное влияние на уровне значимости $1 - \alpha = 0.05$.

Указание: предполагается, что исходные данные прошли тест на нормальность по критерию Колмогорова–Смирнова и тест на равенство дисперсий для различных уровней фактора $I_Б$ по критерию Бартлетта.

Таблица 2.9

Фактор $I_Б$	*Фактор* $U_{кэ}$											
	$U_{кэ1}$			$U_{кэ2}$			$U_{кэ3}$			$U_{кэ4}$		
$I_{Б1}$	4.06	3.91	4.16	4.45	4.56	4.42	4.81	4.91	4.78	4.94	5.05	4.93
$I_{Б2}$	4.54	4.50	4.29	4.81	5.16	4.99	5.35	5.24	5.31	5.42	5.46	5.50
$I_{Б3}$	4.90	4.87	4.75	5.27	5.11	5.24	5.55	5.64	5.66	5.80	5.80	5.94
$I_{Б4}$	5.22	5.22	5.03	5.57	5.57	5.58	5.98	5.85	5.91	6.30	6.28	6.11

№ 13. Для экспериментальных данных из табл. 2.10 определить наличие влияния факторов *A* и *B* на результаты измерений, а также их взаимное влияние на уровне значимости $1 - \alpha = 0.1$.

Указание: предполагается, что исходные данные прошли тест на нормальность по критерию Колмогорова–Смирнова и тест на равенство дисперсий для различных уровней фактора *B* по критерию Бартлетта.

Таблица 2.10

Фактор B	*Фактор A*				
	A_1	A_2	A_3	A_4	A_5
B_1	3.97	4.32	4.10	4.46	4.65
B_2	5.16	5.35	5.48	5.68	5.74
B_3	6.12	6.14	6.49	6.57	6.58
B_4	7.09	6.99	7.31	7.44	7.63
B_5	7.81	8.04	8.31	8.34	8.43
B_6	9.20	9.58	9.62	9.79	9.80
B_7	8.61	8.71	8.95	9.08	9.07

№ 14. В результате наблюдений реакции *I* нелинейного элемента с кусочно-линейной аппроксимацией ВАХ и большим значением уровня начала характеристики на слабый, подверженный флуктуациям сигнал *U* получена совокупность данных ($n = 10$):

$$U\text{, мВ}: \; 2 \;\; 4 \;\; 1 \;\; 7 \;\; 3 \;\; 11 \;\; 14 \;\; 15 \;\; 21 \;\; 4$$
$$I\text{, мкА}: \; 7 \;\; 6 \;\; 4 \;\; 11 \;\; 2 \;\; 21 \;\; 31 \;\; 23 \;\; 40 \;\; 15$$

Проверить гипотезу о том, что система не находится в состоянии полной отчёски, т. е. о наличии корреляции между случайными величинами *X* и *Y* на уровне значимости $1 - \alpha = 0.05$.

№ 15. В результате снятия участка зависимости выходного тока модулятора от входного напряжения получена следующая совокупность данных:

U, мВ :	4	1	7	3	11	14	15	21	14	18	4	7	4	10
I_1, мкА :	6	4	11	2	21	31	23	40	15	19	20	11	14	16

Проверить с достоверностью $\alpha = 0.95$ гипотезу о наличии линейной связи между *U* и I_1 (о том, что это – линейный участок модуляционной характеристики).

№ 16. В результате снятия участка амплитудно-частотной характеристики неизвестного четырёхполюсника получена следующая совокупность данных:

f_i, кГц :	2	4	9	13	15
A_{ij} :	1, 3, 4	7, 8, 12	14, 19, 21	11, 9, 6	8, 7, 3

Проверить с достоверностью $\alpha = 0.95$ гипотезы о наличии частотной зависимости коэффициента передачи и её линейности.

№ 17. В результате снятия участка амплитудно-частотной характеристики неизвестного четырёхполюсника получена следующая совокупность данных:

f_i, кГц :	1	3	7	10	12	17
A_{ij} :	5,7,6,4,	4,8,11,5	11,21,20,18	18,11,5,17	2,5,10,14	7,6,11,17

Проверить на уровне значимости $1 - \alpha = 0.05$ гипотезы о наличии частотной зависимости коэффициента передачи и её линейности.

№ 18. Для выборки x={8.53, 11.4, 8.42, 9.72, 11.5, 8.86, 9.64, 10.4, 9.73, 9.58} мкВ напряжения белого гауссова шума ξ методами максимального правдоподобия и моментов найти оценки математического ожидания $\hat{\mu}_\xi$ и дисперсии $\hat{\sigma}_\xi^2$.

№ 19. Отдел техобслуживания поставщика радиотелефонной связи 15 дней вёл статистику поломок абонентских терминалов: $x = \{1, 2, 1, 1, 1, 3, 2, 2, 2, 4, 3, 1, 2, 2, 0\}$. Найти оценку параметра распределения методами моментов на основе первого и второго моментов и максимального правдоподобия.

Указание: число поломок подчиняется распределению Пуассона.

3. Задания и методические указания к выполнению лабораторных работ

3.1. Указания к выполнению

Для выполнения работы необходимо получить у преподавателя в соответствии со своим вариантом (варианты 1–10) два файла, содержащих выборки, сгенерированные на основе распределения дискретной и непрерывной случайных величин (табл. 3.1). Каждый файл содержит 10^6 отсчётов.

Обработку данных можно производить в любом удобном пакете, в том числе GNU Octave (или MatLAB) [1].

При сдаче работы необходимо указывать, каким правилом группирования вы пользовались при построении гистограммы.

При построении гистограмм по выборочным данным (объёмом N) количество интервалов группирования k можно выбирать, пользуясь следующими рекомендациями [10]:

– с использованием эвристической формулы Старджесса:

$$k = \log_2 N + 1 = 3.3\lg N + 1;$$

– с использованием формулы Брукса и Каррузера: $k = 5\lg N$;

– с использованием стандартного соотношения: $k = \sqrt{N}$;

– с использованием формулы Манна и Вальда:

$$k = 4\sqrt[5]{2}\left(N / t\right)^{0.4},$$

где t – квантиль стандартного нормального распределения для заданного уровня значимости;

– с использованием модификации формулы Манна и Вальда, предложенной Таушановым, Тоневой, Пеновым: $k = 4\lg N$;

– с использованием модификации формулы Манна и Вальда, предложенной Тоневой: $k = \lg N - 5$;

– с использованием модификации формулы Манна и Вальда, предложенной Алексеевой: $k = \dfrac{4}{\xi}\lg\dfrac{N}{10}$, где ξ – значение контрэксцесса $\xi = \sqrt{\sigma^4 / \mu_4}$.

Желательно приводить на одном графике пары: выборочной и теоретической функций распределения, выборочной и теоретической функций вероятности, гистограммы и функции плотности вероятности.

Заметьте, что нормировка гистограммы производится делением на площадь под кривой.

3.2. Содержание

На основе выборки для каждого из распределений необходимо выполнить следующие пункты задания.

Задание к лабораторной работе № 1

1.1. Построить выборочную функию вероятности/плотности вероятности для дискретного/непрерывного распределений в виде гистограммы.

1.2. Построить выборочную функцию распределения в виде гистограммы.

1.3. Записать выражения и указать значения следующих выборочных моментов для обрабатываемой выборки:

– мода;
– выборочное среднее;
– выборочная дисперсия;
– выборочный коэффициент асимметрии;
– выборочный коэффициент эксцесса.

1.4. Построить вариационный ряд и по нему определить:

– медиану;
– квартили уровня 0.25, 0.5 и 0.75;
– интерквартильное расстояние;
– децили уровня 0.1 и 0.9.

1.5. Найти оценку неизвестного параметра (в соответствии с методом, указанным в таблице 3.2) по всей выборке.

1.6. Построить график распределения веротяности/плотности вероятности со значением неизвестного параметра, равного его полученной оценке (на этом же графике отобразить исходную гистограмму).

1.7. На основе критериев согласия Колмогорова–Смирнова и Пирсона определить на уровне значимости 0.05 состоятельность гипотезы о том, что выборка порождена распределениями из табл. 3.1.

Таблица 3.1

Варианты заданий и модельные распределения для них

№ вар.	*Файлы данных*		*Модельные распределения*	
	Дискретные	*Непрерывные*	*Дискретные*	*Непрерывные*
1	dataD1.dat	dataN1.dat	$\mathrm{Bin}(50, 0.3)$ и $\mathrm{Bin}(20, 0.7)$	$\mathcal{U}(0,5)$ и $\mathcal{U}(1,4)$
2	dataD2.dat	dataN2.dat	$\mathrm{Bin}(70, 0.5)$ и $\mathrm{Bin}(80, 0.4)$	$\mathcal{U}(-2,7)$ и $\mathcal{U}(-1,5)$
3	dataD3.dat	dataN3.dat	$\mathrm{Bin}(75, 0.7)$ и $\mathrm{Bin}(75, 0.1)$	$\mathcal{N}(10,3)$ и $\mathcal{N}(5,4)$
4	dataD4.dat	dataN4.dat	$\mathrm{Geom}(0.4)$ и $\mathrm{Geom}(0.1)$	$\mathcal{N}(-4,1)$ и $\mathcal{N}(0,2)$
5	dataD5.dat	dataN5.dat	$\mathrm{Geom}(0.4)$ и $\mathrm{Geom}(0.6)$	$\chi^2(2.5)$ и $\chi^2(7.5)$
6	dataD6.dat	dataN6.dat	$\mathrm{Geom}(0.8)$ и $\mathrm{Geom}(0.3)$	$\chi^2(5)$ и $\chi^2(4)$
7	dataD7.dat	dataN7.dat	$\mathrm{Pois}(10)$ и $\mathrm{Pois}(20)$	$\mathrm{Exp}(2)$ и $\mathrm{Exp}(5)$
8	dataD8.dat	dataN8.dat	$\mathrm{Pois}(20)$ и $\mathrm{Pois}(15)$	$\mathrm{Exp}(7)$ и $\mathrm{Exp}(4)$
9	dataD9.dat	dataN9.dat	$\mathrm{Pois}(50)$ и $\mathrm{Pois}(30)$	$\mathrm{Rayleigh}(2)$ и $\mathrm{Rayleigh}(10)$
10	dataD10.dat	dataN10.dat	$\mathrm{Pois}(75)$ и $\mathrm{Pois}(100)$	$\mathrm{Rayleigh}(5)$ и $\mathrm{Rayleigh}(7)$

– найти значение статистики критерия для эмпирической функции распределения;

– определить значение процентной точки тестового распределения (Колмогорова–Смирнова или Пирсона);

– принять решение о согласии выборочного и модельного распределений.

Задание к лабораторной работе № 2

2.1. Разбив выборку на группы по 10^3 отсчётов, провести оценку неизвестного параметра в каждой группе (в соответствии с методом, указанным в таблице 3.2).

2.2. Построить гистограмму оценок.

2.3. Определить выборочное и выборочную дисперсию среднее оценки.

2.4. Аппроксимировать (и отобразить на одном графике с гистограммой) распределение оценок нормальным распределением с математическим ожиданием, равным выборочному среднему оценки, и дисперсией, равной выборочной дисперсии оценки.

Таблица 3.2

Методы оценки и оцениваемые параметры

Распределение	*Неизвестный параметр*	*Метод оценки*
$\mathrm{Geom}(p)$	p	Метод моментов
$\mathrm{Pois}(\lambda)$	λ	Метод максимального правдоподобия
$\mathrm{Bin}(n, p)$	n, p	Метод моментов
$\mathcal{U}(a,b)$	a, b	Метод максимального правдоподобия
$\mathcal{N}(\mu,\sigma^2)$	μ, σ^2	Метод максимального правдоподобия
$\chi^2(\nu)$	ν	Метод моментов
$\mathrm{t}^2(n)$	n	Метод моментов
$\mathrm{Rayleigh}(\sigma^2)$	σ^2	Метод максимального правдоподобия
$\mathrm{Exp}(\lambda)$	λ	Метод максимального правдоподобия

Рекомендуемая литература

1. Иглин, С. П. Математические расчёты на базе MATLAB / С. П. Иглин. – СПб. : БХВ-Петербург, 2005.

2. Введение в Octave для инженеров и математиков / Е. Р. Алексеев, О. В. Чеснокова. – М. : ALT Linux, 2012.

3. Статистический анализ данных в системе R : учеб. пособие / А. Г. Буховец, П. В. Москалев, В. П. Богатова, Т. Я. Бирючинская ; под ред. проф. А. Г. Буховца. – Воронеж : ВГАУ, 2010.

4. Спирин, Н. А. Методы планирования и обработки результатов инженерного эксперимента : конспект лекций (отдельные главы из учебника для вузов) / Н. А. Спирин, В. В. Лавров ; под общ. ред. Н. А. Спирина. – Екатеринбург : ГОУ ВПО УГТУ-УПИ, 2004.

5. Дронов, С. В. Многомерный статистический анализ : учебное пособие / С. В. Дронов. – Барнаул : Изд-во Алт. гос. ун-та, 2003.

6. Вадзинский, Р. Н. Справочник по вероятностным распределениям / Р. Н. Вадзинский. – СПб. : Наука, 2001.

7. Кобзарь, А. И. Прикладная математическая статистика : Для инженеров и научных работников / А. И. Кобзарь. – М. : Физматлит, 2006.

8. Соколов, Г. А. Математическая статистика : учебник для вузов / Г. А. Соколов, И. М. Гладких. – М. : Экзамен. 2004.

9. Ивченко, Г. И. Математическая статистика : учеб. пособие для втузов / Г. И. Ивченко, Ю. И. Медведев. – М. : Высш. шк. 1984.

10. Лемешко, Б. Ю. О выборе числа интервалов в критериях согласия типа χ^2 / Б. Ю. Лемешко, Е. В. Чимитова // Заводская лаборатория. Диагностика материалов. 2003. – Т. 69, № 1. – С. 61–67.

11. Johnson, N. L. Univariate discrete distributions / N. L. Johnson, S. Kotz, A. Kemp // John Wiley & Sons, 1992.

12. Johnson, N. L. Continuous univariate distributions / N. L. Johnson, S. Kotz, N. Balakrishnan // John Wiley & Sons, 1994. – Vol. 1.

13. Johnson, N. L. Continuous univariate distributions / N. L. Johnson, S. Kotz, N. Balakrishnan // John Wiley & Sons, 1994. – Vol. 2.

14. Справочник по специальным функциям с формулами, графиками и математическими таблицами / под ред. М. Абрамовица и М. Стиган. – М. : Наука, 1979.

Учебное издание

Гвоздарёв Алексей Сергеевич
Артёмова Татьяна Константиновна

Планирование и обработка результатов инженерного эксперимента

Методические указания

Редактор, корректор М. Э. Левакова
Верстка Е. Б. Половковой

Подписано в печать 09.10.14. Формат 60×84 1/16.
Усл. печ. л. 3,49. Уч.-изд. л. 2,4.
Тираж 30 экз. Заказ

Оригинал-макет подготовлен
в редакционно-издательском отделе ЯрГУ.

Ярославский государственный университет
им. П. Г. Демидова.
150000, Ярославль, ул. Советская, 14.