

Министерство образования и науки России

**Федеральное государственное бюджетное**

**образовательное учреждение высшего образования «Казанский национальный исследовательский технологический университет» (ФГБОУ ВО «КНИТУ»)**



**Реферат по дисциплине «Системный анализ»**

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

**Тема:**

**«Применение теории вероятностей в химии»**

Выполнил: студент группы 4381-22 Тимушев Д.А

Проверил: Гильмутдинов Руслан Фаридович

Казань 2021

Содержание

[Введение 2](#_Toc70341203)

[1. Предмет и методы математической статистики 3](#_Toc70341204)

[2. Основные понятия математической статистики 5](#_Toc70341205)

[Основные понятия выборочного метода 5](#_Toc70341206)

[Выборочное распределение 7](#_Toc70341207)

[Теория вероятности в Химии 9](#_Toc70341208)

[Заключение 11](#_Toc70341209)

[Список использованной литературы 12](#_Toc70341210)

## Введение

Математическая статистика — наука о математических методах систематизации и использования статистических данных для научных и практических выводов. Во многих своих разделах математическая статистика опирается на теорию вероятностей, позволяющую оценить надежность и точность выводов, делаемых на основании ограниченного статистического материала (напр., оценить необходимый объем выборки для получения результатов требуемой точности при выборочном обследовании).

В теории вероятностей рассматриваются случайные величины с заданным распределением или случайные эксперименты, свойства которых целиком известны. Предмет теории вероятностей — свойства и взаимосвязи этих величин (распределений).

Но часто эксперимент представляет собой черный ящик, выдающий лишь некие результаты, по которым требуется сделать вывод о свойствах самого эксперимента. Наблюдатель имеет набор числовых (или их можно сделать числовыми) результатов, полученных повторением одного и того же случайного эксперимента в одинаковых условиях.

При этом возникают, например, следующие вопросы: если мы наблюдаем одну случайную величину — как по набору ее значений в нескольких опытах сделать как можно более точный вывод о ее распределении?

Примером такой серии экспериментов может служить социологический опрос, набор экономических показателей или, наконец, последовательность гербов и решек при тысячекратном подбрасывании монеты.

Все вышеприведенные факторы обуславливают актуальность и значимость тематики работы на современном этапе, направленной на глубокое и всестороннее изучение основных понятий математической статистики.

В связи с этим целью данной работы является систематизация, накопление и закрепление знаний о понятиях математической статистики.

## 1. Предмет и методы математической статистики

Математическая статистика — наука о математических методах анализа данных, полученных при проведении массовых наблюдений (измерений, опытов). В зависимости от математической природы конкретных результатов наблюдений статистика математическая делится на статистику чисел, многомерный статистический анализ, анализ функций (процессов) и временных рядов, статистику объектов нечисловой природы. Существенная часть статистики математической основана на вероятностных моделях. Выделяют общие задачи описания данных, оценивания и проверки гипотез. Рассматривают и более частные задачи, связанные с проведением выборочных обследований, восстановлением зависимостей, построением и использованием классификаций (типологий) и др.

Для описания данных строят таблицы, диаграммы, иные наглядные представления, например, корреляционные поля. Вероятностные модели обычно не применяются. Некоторые методы описания данных опираются на продвинутую теорию и возможности современных компьютеров. К ним относятся, в частности, кластер-анализ, нацеленный на выделение групп объектов, похожих друг на друга, и многомерное шкалирование, позволяющее наглядно представить объекты на плоскости, в наименьшей степени исказив расстояния между ними.

Методы оценивания и проверки гипотез опираются на вероятностные модели порождения данных. Эти модели делятся на параметрические и непараметрические. В параметрических моделях предполагается, что изучаемые объекты описываются функциями распределения, зависящими от небольшого числа (1-4) числовых параметров. В непараметрических моделях функции распределения предполагаются произвольными непрерывными. В статистике математической оценивают параметры и характеристики распределения (математическое ожидание, медиану, дисперсию, квантили и др.), плотности и функции распределения, зависимости между переменными (на основе линейных и непараметрических коэффициентов корреляции, а также параметрических или непараметрических оценок функций, выражающих зависимости) и др. Используют точечные и интервальные (дающие границы для истинных значений) оценки.

В математической статистике есть общая теория проверки гипотез и большое число методов, посвященных проверке конкретных гипотез. Рассматривают гипотезы о значениях параметров и характеристик, о проверке однородности (то есть о совпадении характеристик или функций распределения в двух выборках), о согласии эмпирической функции распределения с заданной функцией распределения или с параметрическим семейством таких функций, о симметрии распределения и др.

Большое значение имеет раздел математической статистики, связанный с проведением выборочных обследований, со свойствами различных схем организации выборок и построением адекватных методов оценивания и проверки гипотез.

Задачи восстановления зависимостей активно изучаются более 200 лет, с момента разработки К. Гауссом в 1794 г. метода наименьших квадратов. В настоящее время наиболее актуальны методы поиска информативного подмножества переменных и непараметрические методы.

Разработка методов аппроксимации данных и сокращения размерности описания была начата более 100 лет назад, когда К. Пирсон создал метод главных компонент. Позднее были разработаны факторный анализ и многочисленные нелинейные обобщения.

Различные методы построения (кластер-анализ), анализа и использования (дискриминантный анализ) классификаций (типологий) именуют также методами распознавания образов (с учителем и без), автоматической классификации и др.

Математические методы в статистике основаны либо на использовании сумм (на основе Центральной Предельной Теоремы теории вероятностей) или показателей различия (расстояний, метрик), как в статистике объектов нечисловой природы. Строго обоснованы обычно лишь асимптотические результаты. В настоящее время компьютеры играют большую роль в математической статистике. Они используются как для расчетов, так и для имитационного моделирования (в частности, в методах размножения выборок и при изучении пригодности асимптотических результатов).

## 2. Основные понятия математической статистики

## Основные понятия выборочного метода

Пусть 1306472882_1— случайная величина, наблюдаемая в случайном эксперименте. Предполагается, что вероятностное пространство задано (и не будет нас интересовать).

Будем считать, что, проведя 1306472883_2раз этот эксперимент в одинаковых условиях, мы получили числа 1306472883_3, 1306472883_4, 1306472883_5, 1306472884_6— значения этой случайной величины в первом, втором, и т.д. экспериментах. Случайная величина 1306472884_7имеет некоторое распределение 1306472884_8, которое нам частично или полностью неизвестно.

Рассмотрим подробнее набор 1306472884_9, называемый выборкой.

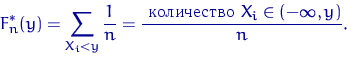
В серии уже произведенных экспериментов выборка — это набор чисел. Но если эту серию экспериментов повторить еще раз, то вместо этого набора мы получим новый набор чисел. Вместо числа 1306472883_3появится другое число — одно из значений случайной величины 1306472884_7. То есть 1306472883_3(и 1306472883_4, и 1306472884_10, и т.д.) — переменная величина, которая может принимать те же значения, что и случайная величина 1306472884_7, и так же часто (с теми же вероятностями). Поэтому до опыта 1306472883_3— случайная величина, одинаково распределенная с 1306472884_7, а после опыта — число, которое мы наблюдаем в данном первом эксперименте, т.е. одно из возможных значений случайной величины 1306472883_3.

Выборка 1306472884_9объема 1306472883_2— это набор из 1306472883_2независимых и одинаково распределенных случайных величин («копий 1306472884_7»), имеющих, как и 1306472884_7, распределение 1306472884_8.

Что значит «по выборке сделать вывод о распределении»? Распределение характеризуется функцией распределения, плотностью или таблицей, набором числовых характеристик — 1306472885_11, 1306472885_12, 1306472885_13и т.д. По выборке нужно уметь строить приближения для всех этих характеристик.

## Выборочное распределение

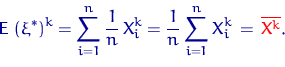
Рассмотрим реализацию выборки на одном элементарном исходе 1306472885_14— набор чисел 1306472885_15, 1306472883_5, 1306472886_16. На подходящем вероятностном пространстве введем случайную величину 1306472886_17, принимающую значения 1306472883_3, 1306472883_5, 1306472884_6с вероятностями по 1306472886_18(если какие-то из значений совпали, сложим вероятности соответствующее число раз). Таблица распределения вероятностей и функция распределения случайной величины 1306472886_17выглядят так:



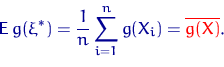
Распределение величины 1306472886_17называют эмпирическим или выборочным распределением. Вычислим математическое ожидание и дисперсию величины 1306472886_17и введем обозначения для этих величин:

1306472887_21

Точно так же вычислим и момент порядка 1306472887_22



В общем случае обозначим через 1306472887_24величину



Если при построении всех введенных нами характеристик считать выборку 1306472883_3, 1306472883_5, 1306472884_6 набором случайных величин, то и сами эти характеристики — 1306472888_26, 1306472888_27, 1306472888_28, 1306472888_29, 1306472889_30 — станут величинами случайными. Эти характеристики выборочного распределения используют для оценки (приближения) соответствующих неизвестных характеристик истинного распределения.

Причина использования характеристик распределения 1306472886_17для оценки характеристик истинного распределения 1306472884_7(или 1306472883_3) — в близости этих распределений при больших 1306472883_2.

Рассмотрим, для примера, 1306472883_2подбрасываний правильного кубика. Пусть 1306472889_31— количество очков, выпавших при 1306472889_32-м броске, 1306472889_33. Предположим, что единица в выборке встретится 1306472889_34раз, двойка — 1306472890_35раз и т.д. Тогда случайная величина 1306472886_17будет принимать значения **1**, 1306472883_5, **6** с вероятностями 1306472890_36, 1306472883_5, 1306472890_37 соответственно. Но эти пропорции с ростом 1306472883_2приближаются к 1306472890_38согласно закону больших чисел. То есть распределение величины 1306472886_17в некотором смысле сближается с истинным распределением числа очков, выпадающих при подбрасывании правильного кубика.

Мы не станем уточнять, что имеется в виду под близостью выборочного и истинного распределений. В следующих параграфах мы подробнее познакомимся с каждой из введенных выше характеристик и исследуем ее свойства, в том числе ее поведение с ростом объема выборки.

## Теория вероятности в Химии

Понимание природы химических реакций, динамического равновесия также невозможно без статистических представлений. Вся физическая химия, ее математический аппарат и предлагаемые ею модели являются статистическими.

О [положении электрона](https://www.chem21.info/info/129236) в атоме можно судить только с [точки зрения теории](https://www.chem21.info/info/1107356) вероятности.

Важность методов математической статистики для химической технологии подчеркивает Веллер: «При опытном испытании поведения какой-либо системы надо принимать во внимание, например, допуски продукции, свойства сырья, а также субъективные различия характеров обслуживающего персонала и изменения в планируемых расходах на обслуживание. Благодаря этому все задачи контроля у инженера сводятся к проблемам вероятности. Тот факт, что инженеры этого еще не уяснили себе, объясняется тем, что литература по вопросам применения теории вероятностей перегружена примерами из биологии и сельского хозяйства, с которыми инженеру нечего делать».

[Теория вероятностей](https://www.chem21.info/info/17205) развивалась, чтобы предсказывать до [проведения эксперимента](https://www.chem21.info/info/684290) вероятность того, что [случайная величина](https://www.chem21.info/info/24280) X лежит между двумя значениями xi и хг По [мере развития](https://www.chem21.info/info/1055645) теории [неизбежно](https://www.chem21.info/info/776116) стали появляться также и [некоторые виды](https://www.chem21.info/info/937219) [статистических выводов Статистические выводы](https://www.chem21.info/info/1529800) имеют дело с задачек, являющейся обратной по отношению к [задаче теории](https://www.chem21.info/info/1403162) вероятностей, а именно как использовать данные х, хг, х после эксперимента для того, чтобы [сделать выводы](https://www.chem21.info/info/1670910) о [свойствах случайной](https://www.chem21.info/info/1867205) величины X Предположим, например, что в результате 15 бросаний монеты мы получили 12 гербов и требуется узнать, совместим ли этот результат с предположением о симметричности монеты [Классическое решение](https://www.chem21.info/info/1566336) этой задачи представляет [собой](https://www.chem21.info/info/1795776) пример одного из [ранних способов](https://www.chem21.info/info/1734970) получения выводов, известного теперь под названием [критерия значимости](https://www.chem21.info/info/136414) Решение использует исключительно вероятностные понятия и состоит в [вычислении вероятности](https://www.chem21.info/info/748479) получения 12 или более гербов при допущении гипотезы, что монета симметрична Если эта [вероятность мала](https://www.chem21.info/info/1361404), то она может рассматриваться как веский признак того, что предположение о симметричности монеты ложно, если вероятность велика, то этот результат не противоречит гипотезе о том, что монета симметрична. В упомянутом выше примере вероятность получить 12 или более гербов в 15 бросаниях в предположении, что монета симметрична, равна 0,018, из чего можно заключить, что монета несимметрична

 Из [теоретических методов исследования](https://www.chem21.info/info/1645330), находящих применение в различных [разделах физической химии](https://www.chem21.info/info/1458376), можно [выделить метод](https://www.chem21.info/info/1670821) статистический. Данный метод основан на применении к рассматриваемой системе, состоящей из [очень большого](https://www.chem21.info/info/1034817) [числа частиц](https://www.chem21.info/info/4574), [законов теории](https://www.chem21.info/info/679846) вероятности. Примером такого подхода может служить [кинетическая теория газов](https://www.chem21.info/info/6387). Исходя из допущения о полной [беспорядочности движения](https://www.chem21.info/info/1169475) [отдельных молекул](https://www.chem21.info/info/362978) газа и определяя [наиболее вероятные](https://www.chem21.info/info/6336) сочетания их, для системы, состоящей из [очень большого](https://www.chem21.info/info/1034817) [числа молекул](https://www.chem21.info/info/82239), эта теория дает возможность установить важные [соотношения между различными](https://www.chem21.info/info/1456811) [свойствами газа](https://www.chem21.info/info/120246) и определить их зависимость от [условий существования](https://www.chem21.info/info/26956) газа.

В настоящее время на основе применения статистических методов к решению термодинамических задач сформировался новый важный раздел термодинамики - статистическая термодинамика

## Заключение

Математическая (или теоретическая) статистика в химии опирается на методы и понятия теории вероятностей, но решает в каком-то смысле обратные задачи.

Если мы наблюдаем одновременно проявление двух (или более) признаков, т.е. имеем набор значений нескольких случайных величин -- что можно сказать об их зависимости? Есть она или нет? А если есть, то какова эта зависимость?

Часто бывает возможно высказать некие предположения о распределении, спрятанном в «черном ящике», или о его свойствах. В этом случае по опытным данным требуется подтвердить или опровергнуть эти предположения («гипотезы»). При этом надо помнить, что ответ «да» или «нет» может быть дан лишь с определенной степенью достоверности, и чем дольше мы можем продолжать эксперимент, тем точнее могут быть выводы. Наиболее благоприятной для исследования оказывается ситуация, когда можно уверенно утверждать о некоторых свойствах наблюдаемого эксперимента - например, о наличии функциональной зависимости между наблюдаемыми величинами, о нормальности распределения, о его симметричности, о наличии у распределения плотности или о его дискретном характере, и т.д.

## Список использованной литературы

1. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. М.: Наука, 1965.
2. Боровков А.А. Математическая статистика. М.: Наука, 1984.
3. Коршунов Д.А., Чернова Н.И. Сборник задач и упражнений по математической статистике. Новосибирск: Изд-во Института математики им. С.Л.Соболева СО РАН, 2001.
4. Феллер В. Введение в теорию вероятностей и ее приложения. М.: Мир, Т.2, 1984.
5. Пехелецкий И.Д. Математика: учебник для студентов.