Министерство образования и науки Российской Федерации

НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

Компьютерные технологии моделирования и анализа данных:

Исследование относительной эффективности критериев проверки статистических гипотез

Методические указания
к выполнению курсовых проектов
для студентов II-го курса ФПМИ
по направлению 010400.68
дневного отделения

Новосибирск, 2020

Методические указания предназначены для студентов, выполняющих курсовые проекты по курсу "Компьютерные технологии моделирования и анализа данных " в третьем семестре (направление 010400.68 – Прикладная математика и информатика, магистерская программа – *Математическое и программное обеспечение информационных технологий моделирования и анализа данных*). Указания содержат необходимые сведения для выполнения курсового проекта, порядок выполнения, структуру оформления пояснительной записки и примерное содержание её разделов, варианты заданий.

Составители: доктор техн. наук, проф. *Б.Ю. Лемешко*,
доктор техн. наук, доц. *С.Н. Постовалов*,
доктор техн. наук, доц. *Е.В. Чимитова*

Работа подготовлена на кафедре
теоретической и прикладной информатики

Исследование относительной эффективности критериев проверки статистических гипотез

1. Цель выполнения работы

Исследование статистических критериев. Определение необходимого объема выборки для проверки статистической гипотезы. Вычисление относительной эффективности статистических критериев. Определение оптимального критерия по правилам Вальда и Сэвиджа.

2. Методические указания

Курсовой проект (КП) носит исследовательский характер. Варианты заданий могут быть связаны с известными методами или критериями, изложение которых в литературных источниках не позволяет однозначно охарактеризовать их свойства по сравнению с аналогами, либо остается неясной картина корректности статистических выводов в случае нарушения стандартных предположений. Задание может быть связано с исследованием свойств оценок или критериев в нестандартных условиях, связанных с конкретной областью приложения, либо в заданных нестандартных условиях.

Предпочтение отдается вариантам заданий, ориентированным на развитие аппарата прикладной математической статистики и создание программного обеспечения задач статистического анализа.

В процессе выполнения КП необходимо ознакомиться с историческими аспектами, состоянием и тенденциями развития в соответствующем разделе прикладной математической статистики, предпосылками, определившими интерес и потребности к исследованиям в данном направлении.

Выполнение экспериментальной части КП может предусматривать:

- самостоятельную разработку некоторого программного обеспечения;
- развитие существующего программного обеспечения;
- использование различных математических пакетов и программных систем, например, для исследования реальных свойств классических методов и критериев статистического анализа в условиях нарушения стандартных предположений.

При анализе результатов численных экспериментов целесообразно использование доступных математических пакетов, программных систем статистического анализа и средств графической визуализации.

Тематика КП может быть связана с анализом и развитием статистических методов, статистических критериев, характером использования методов и критериев в статистических пакетах, с корректностью применения статистических методов в приложениях и в программном обеспечении.

2.1 Статистические гипотезы и критерии их проверки

Статистической гипотезой называется любое утверждение о виде или свойствах распределения наблюдаемых в эксперименте случайных величин. Обычно статистические гипотезы делят на следующие виды: *однородности*, если имеется две или более выборок случайных величин; *независимости*, если имеется выборка многомерной случайной величины; *случайности*, если есть предположения о независимости и одинаковом распределении наблюдений в выборке; *о виде распределения*, если есть предположения о законе распределения случайной величины (рисунок 1).

Проверка статистической гипотезы состоит в том, чтобы сформулировать такое правило, которое позволило бы по результатам проведенных наблюдений принять или отклонить гипотезу. Правило, согласно которому гипотеза принимается или отвергается, называется *критерием* проверки статистической гипотезы.

С проверкой статистических гипотез связывают ошибки двух типов. *Ошибкой первого рода* называют событие, когда верная проверяемая гипотеза отвергается критерием. *Ошибкой второго рода* называют событие, когда неверная проверяемая гипотеза принимается критерием. Вероятности ошибок первого и второго рода обозначают α и β, соответственно. Вероятность ошибки второго рода зависит от выдвигаемой конкурирующей гипотезы. Вероятность отклонения ложной проверяемой гипотезы, т.е. принятия правильного решения в пользу конкурирующей, называется *мощностью*, и она равна $1-\beta$. Вероятность ошибки первого рода также называют *уровнем значимости* критерия.

Есть небольшая терминологическая тонкость в использовании выражения «гипотеза принимается критерием». Часто, вместо этого говорят «гипотеза не отвергается критерием», понимая, что по выборке большего объема гипотеза может быть отвергнута этим же критерием, или же говорят «нет оснований для отвержения гипотезы по данной выборке». В дальнейшем, для упрощения текста будет использоваться самый краткий вариант.

Гипотезу, которую мы проверяем, будем называть *основной* или *нулевой* гипотезой, и будем всегда обозначать H_0. Альтернативные или *конкурирующие* гипотезы будем обозначать $H_1,\ H_2,\ \ldots,\ H_m$.

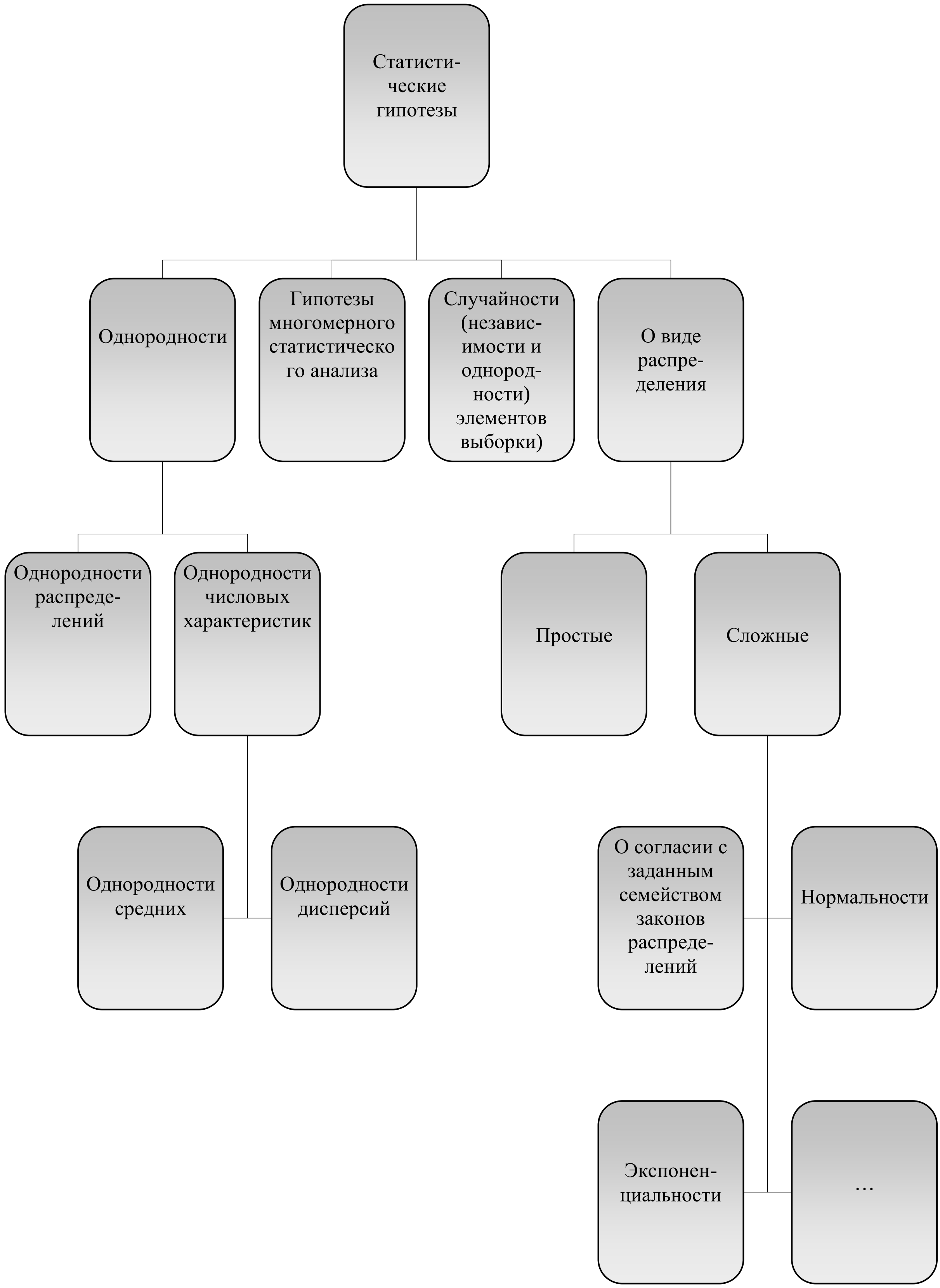

Рисунок 1 – Классификация статистических гипотез

Для проверки одной и той же гипотезы, как правило, существует несколько различных статистических критериев $T_1, T_2, ..., T_k$. Выбор подходящего статистического критерия, вообще говоря, не является тривиальной задачей. Можно сформулировать следующие принципы подбора статистических критериев.

1. Должны выполняться “стандартные” предположения, обуславливающие возможность применения рассматриваемого критерия (например, о виде распределения случайной величины и о наблюдаемых данных). Так, например, нельзя применять критерий Колмогорова по группированным данным, или по наблюдениям дискретной случайной величины.
2. Критерий должен быть *состоятельным*, т.е. его мощность должна стремиться к единице с ростом объема выборки.
3. Критерий должен быть *несмещенным*, т.е. мощность должна быть больше, чем вероятность ошибки первого рода.
4. Критерий должен обладать наибольшей мощностью при заданном объеме выборки и заданном уровне значимости критерия.

Добиться выполнения последнего принципа на практике не представляется возможным, потому что построить наиболее мощный критерий удается только в очень редких случаях, например, когда основная и конкурирующая гипотезы являются *простыми*. Чаще всего, для разных конкурирующих гипотез, для разных уровней значимости, для разных объемов выборки, более мощными оказываются разные критерии.

В этой ситуации для выбора оптимального критерия можно применить классическую теорию *принятия решений в условиях неопределенности*. Критерии являются стратегиями, конкурирующие гипотезы – состояниями среды, функция полезности $u\left(T_i, H_j\right)$ – это мощность критерия (таблица 1). Другим способом определения функции полезности при выборе критерия может быть стоимость проведения эксперимента по различению основной и конкурирующей гипотез с заданными вероятностями ошибок первого и второго рода.

Таблица 1 – Матрица полезности выбора критерия T_i при конкурирующих гипотезах H_j

$T_i \setminus H_j$	H_1	H_2	...	H_m
T_1	$u(T_1, H_1)$	$u(T_1, H_2)$	...	$u(T_1, H_m)$
T_2	$u(T_2, H_1)$	$u(T_2, H_2)$	...	$u(T_2, H_m)$
...	...	...	...	...
T_k	$u(T_k, H_1)$	$u(T_k, H_2)$	...	$u(T_k, H_m)$

Существуют разные подходы к выбору оптимальной стратегии при принятии решения в условиях неопределенности. В случае, когда нет никакой информации о том, какая конкурирующая гипотеза может быть верна, рациональным выглядит выбор критерия по правилу Вальда (известны также такие названия как «критерий крайнего пессимиста» или «критерий осторожного наблюдателя»):

$$T^* = \arg\max_{T_i} \min_{H_j} u\left(T_i, H_j\right). \quad (1)$$

Критерий, выбранный по правилу Вальда, максимизирует полезность против самой «неудобной» конкурирующей гипотезы.

Другим возможным правилом при выборе оптимального критерия является правило Сэвиджа: это правило минимизации «сожалений». «Сожаление» – это величина, равная отклонению мощности критерия относительно наилучшего возможного значения мощности на альтернативной гипотезе H. Чтобы определить "сожаление", поступаем следующим образом. Для каждого значения $u\left(T_i, H_j\right)$ в таблице 1 находим отклонение от наибольшего значения в j-м столбце $MAX_j,\ j = \overline{1,m}$, т.е.

$$s_{ij} = \left|u\left(T_i, H_j\right) - MAX_j\right|, i = \overline{1,k},\ j = \overline{1,m}. \quad (2)$$

В результате получается матрица "сожалений" (таблица 2).

Таблица 2 – Матрица сожалений

$T_i \setminus H_j$	H_1	H_2	...	H_m
T_1	$s(T_1, H_1)$	$s(T_1, H_2)$	...	$s(T_1, H_m)$
T_2	$s(T_2, H_1)$	$s(T_2, H_2)$	...	$s(T_2, H_m)$
...	...	...	...	...
T_k	$s(T_k, H_1)$	$s(T_k, H_2)$	...	$s(T_k, H_m)$

Выбор оптимального критерия по правилу Сэвиджа заключается в определении критерия, который имеет наименьшее сожаление против "наихудшей" альтернативы:

$$T^* = \arg\min_{T_i} \max_{H_j} s\left(T_i, H_j\right). \quad (3)$$

Таким образом, рекомендация по выбору статистического критерия для формируется на основе правил Вальда и Сэвиджа.

2.2 Относительная эффективность критерия

Предположим, что имеется два критерия T_1 и T_2 для проверки гипотезы H_0 против гипотезы H_1. Тогда *относительная эффективность* первого критерия по отношению ко второму равна

$$RE(T_1, T_2) = \frac{n_2(\alpha, \beta)}{n_1(\alpha, \beta)}, \quad (4)$$

где $n_i(\alpha, \beta)$ - это необходимый объем выборки для различения гипотез H_0 и H_1 с вероятностями ошибок первого и второго рода, равными α и β, соответственно.

Так как стоимость проведения эксперимента, как правило, пропорциональна числу испытаний, то более эффективный критерий будет экономически более выгоден, чем менее эффективный. Поэтому величину

$$u(T_i, H_j) = -\ln n_i(\alpha, \beta, H_j) \quad (5)$$

можно рассматривать как полезность применения критерия T_i при конкурирующей гипотезе H_j.

2.2 Определение необходимого объема выборки

Вычисление необходимого объема выборки $n(\alpha,\beta)$ можно выполнить с помощью компьютерного моделирования. Пусть заданы некоторые значения α и β. Задав некоторое начальное значение n_0 можно вычислить мощность критерия $1-\beta_{n_0}$.

- Если вычисленная мощность критерия меньше $1-\beta$, объем n нужно увеличивать, до тех пор, пока мощность $1-\beta_n$ не станет больше или равна $1-\beta$.
- Если вычисленная мощность критерия больше $1-\beta$, объем n нужно уменьшать, до тех пор, пока мощность $1-\beta_n$ не станет меньше или равна $1-\beta$.

3. Структура пояснительной записки к КП

Рекомендуемый объем записки не должен превышать 30-35 страниц текста.

Примерная структура пояснительной записки имеет следующий вид.

1. Введение

   Во введении указывается цель работы, кратко характеризуется место и значение соответствующего метода в аппарате прикладной математической статистики.

2. Постановка задачи

   Постановка задачи включает алгоритм или последовательность действий, связанную с применением соответствующего метода или критерия, основные соотношения, а также формулировку предпосылок, обуславливающих область корректного использования метода.

3. Аналитический обзор

   Аналитический обзор характеризует состояние соответствующей области на момент появления (создания) метода (критерия). Он включает указания на ключевые работы, связанных с возникновением, развитием и актуальным применением соответствующих методов в приложениях.

4. Результаты исследований

   Приводятся результаты исследований (численных экспериментов) автора КП. Из теста записки должно быть однозначно понятно, с какой целью проводились эксперименты, что для этого было сделано, какие средства использовались, какова точность экспериментов, что было выявлено в результате исследований (что подтвердилось, что не подтвердилось, почему).

5. Выводы

   На основании п.4. приводится краткая формулировка ключевых результатов численных экспериментов. Формулируются общие выводы, в которых приводится сравнение статистических критериев, даются рекомендации по использованию.

6. Список использованных источников

   Приводится список источников, отмеченных в библиографическом обзоре ключевых работ по теме КП, включая Интернет-издания, а также работ, использованных в связи с проведенными экспериментами.

4. Порядок выполнения

1. Сформулировать несколько альтернативных гипотез (не менее трех). Построить графики функций распределения, функций плотности, вычислить расстояния между функциями распределений основной и альтернативной

гипотезы в L_2 . Расстояния для разных альтернатив не должны быть существенно различными.

2. Выбрать доступные и наиболее оптимальные средства для проведения вычислительных экспериментов, при необходимости разработав требуемое программное обеспечение.

3. Вычислить необходимый объем выборки для каждого статистического критерия из варианта и каждой альтернативной гипотезы для заданных значений $\alpha = 0,01$ и $\beta = 0,01$ и заполнить таблицу 1 значениями функции полезности по (5) .

4. Сравнить статистические критерии по правилам Вальда и Сэвиджа.

5. Вычислить относительную эффективность критериев по (4).

6. Сформулировать выводы по работе. Оформить пояснительную записку по работе.

Варианты заданий

№ варианта	Критерии
1.	1,2,3
2.	2,3,4
3.	3,4,5
4.	4,5,6
5.	7,8,9
6.	8,9,10
7.	9,10,11
8.	10,11,12
9.	13,14,15
10.	14,15,16
11.	15,16,17
12.	16,17,18
13.	17,18,19
14.	18,19,20
15.	19,20,21
16.	20,21,22
17.	21,22,23
18.	22,23,24
19.	23,24,25
20.	24,25,26

21.	25,26,27
22.	28,29,30
23.	29,30,31
24.	30,31,32
25.	31,32,33
26.	32,33,34
27.	33,34,35
28.	34,35,36
29.	35,36,37
30.	36,37,38
31.	37,38,39
32.	38,39,40

Статистические критерии

1. Проверка гипотезы о виде распределения по критерию Колмогорова
2. Проверка гипотезы о виде распределения по критерию Крамера-Мизеса-Смирнова
3. Проверка гипотезы о виде распределения по критерию Андерсона-Дарлинга
4. Проверка гипотезы о виде распределения по критерию Жанга Zk
5. Проверка гипотезы о виде распределения по критерию Жанга Zs
6. Проверка гипотезы о виде распределения по критерию Жанга Za

7. Проверка гипотезы равномерности по критерию Шермана
8. Проверка гипотезы равномерности по критерию Морана
9. Проверка гипотезы равномерности по критерию Ченга-Спиринга.
10. Проверка гипотезы равномерности по критерию Саркади-Косика.
11. Проверка гипотезы равномерности по критерию Хегази-Грина
12. Проверка гипотезы равномерности по критерию Гринвуда-Кэсенберри-Миллера

13. Проверка гипотезы экспоненциальности по критерию Большева
14. Проверка гипотезы экспоненциальности по критерию Гнеденко
15. Проверка гипотезы экспоненциальности по критерию Харриса
16. Проверка гипотезы экспоненциальности по критерию Холландера-Прошана
17. Проверка гипотезы экспоненциальности по критерию Гини
18. Проверка гипотезы экспоненциальности по критерию Эпштейна
19. Проверка гипотезы экспоненциальности по критерию Кокса-Оукса
20. Проверка гипотезы экспоненциальности по критерию Эппса-Палли
21. Проверка гипотезы экспоненциальности по критерию Ватсона
22. Проверка гипотезы экспоненциальности по критерию Купера
23. Проверка гипотезы экспоненциальности по критерию Дешпанде

24. Проверка гипотезы экспоненциальности по критерию Клара
25. Проверка гипотезы экспоненциальности по критерию Барингхауса–Хенце
26. Проверка гипотезы экспоненциальности по критерию Хенце
27. Проверка гипотезы экспоненциальности по критерию Хенце–Мейнтаниса

28. Проверка гипотезы нормальности по критерию Фросини
29. Проверка гипотезы нормальности по критерию Хегази–Грина
30. Проверка гипотезы нормальности по критерию Гири
31. Проверка гипотезы нормальности по критерию Дэвида–Хартли–Пирсона
32. Проверка гипотезы нормальности по критерию Шпигельхальтера
33. Проверка гипотезы нормальности по критерию Шапиро–Уилка
34. Проверка гипотезы нормальности по критерию Ройстона
35. Проверка гипотезы нормальности по критерию Эппса–Палли,
36. Проверка гипотезы нормальности по критерию Д'Агостино

37. Проверка гипотезы о распределении Вейбулла-Гнеденко по критерию Майкла
38. Проверка гипотезы о распределении Вейбулла-Гнеденко по критерию Шапиро-Уилка
39. Проверка гипотезы о распределении Вейбулла-Гнеденко по критерию Тайку-Синга
40. Проверка гипотезы о распределении Вейбулла-Гнеденко по критерию Майка-Фертига-Шуйера

Литература

1. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход : [монография] / Б. Ю. Лемешко [и др.]. - Новосибирск, 2011. - 887 с. : ил., табл.
2. Кобзарь А.И. Прикладная математическая статистика для инженеров и научных работников. – М.: Физматлит, 2006. – 816 с.
3. Большев Л.Н. К вопросу о проверке «показательности». Вероятность и ее применения С. 542-544. (есть в электронном виде)
4. Ascher S. A survey of tests for exponentiality. Communications in Statistics - Theory and Methods, 1811-1825 (есть в электронном виде)
5. Henze N. and Meintanis S.G. Recent and classical tests for exponentiality: a partial review with comparisons. Metrika (2005) 61: 29–45 (есть в электронном виде)
6. Лемешко Б.Ю., Постовалов С.Н. Компьютерные технологии анализа данных и исследования статистических закономерностей: Учеб. пособие. – Новосибирск: изд-во НГТУ, 2004. – 120 с.
7. Лемешко Б.Ю. Критерии проверки гипотез об однородности. Руководство по применению : монография / Б.Ю. Лемешко. – М. : ИНФРА-М, 2017. – 208 с. – (Научная мысль). – ISBN: 978-5-16-012557-2 (print), 978-5-16-105463-5(online) DOI: 10.12737/22368
8. Лемешко Б.Ю., Блинов П.Ю. Критерии проверки отклонения распределения от равномерного закона. Руководство по применению: Монография / Б.Ю. Лемешко, П.Ю. Блинов. – М.: ИНФРА-М, 2015. – 183 с. – (Научная мысль). DOI: 10.12737/11304
9. Лемешко Б.Ю. Критерии проверки отклонения распределения от нормального закона. Руководство по применению: Монография / Б.Ю. Лемешко. – М.: ИНФРА-М, 2015. – 160 с. – (Научная мысль). DOI: 10.12737/6086
10. Лемешко Б.Ю. Непараметрические критерии согласия: Руководство по применению: Монография / Б.Ю. Лемешко.– М.: ИНФРА-М, 2014. – 163 с. DOI: 10.12737/11873