

ЭТАПЫ ОБРАБОТКИ ДАННЫХ ПРИ МАЛЫХ ВЫБОРКАХ

Введение

Малые выборки широко используются для решения задач, связанных с испытанием статистических гипотез, особенно гипотез о средних величинах. Использование малых выборок в ряде случаев обусловлено характером обследуемой совокупности. Производственный и экономический эксперимент, клинические испытания, связанные с экономическими затратами, также проводятся на небольшом числе испытаний.

В выборках небольшого объема $n \leq 30$ характер распределения генеральной совокупности сказывается на распределении ошибок выборки. Поэтому, существует проблема получения корректных результатов после обработки данных с помощью статистических критериев, рассчитанных на обработку больших объемов исследуемых данных.

Основные положения, а также статистические критерии, необходимые для обработки результатов измерений, описаны в [1]. Область применения ИСО 5725 – точность стандартизованных методов измерений, в том числе предназначенных для целей испытаний продукции, позволяющих количественно оценить характеристики свойств (показателей качества и безопасности) объекта испытаний. Именно поэтому результаты измерений характеристик образцов, взятых в качестве выборки из партии изделий, являются основой для получения результатов испытаний всей партии. Тем не менее, в [1] ничего не говорится о влиянии объемов выборок, на основе которых проводятся все дальнейшие вычисления, и ведется обработка данных с последующими выводами, на достоверность применяемых статистических критериев обработки.

В работе [2] была исследована проблема влияния количества исследуемых групп данных и их объемов на мощность критерия Кохрена, предложена статистическая оценка мощности (СОМ) критерия Кохрена

Стандарт [1] не позволяет его пользователям ориентироваться в том, какой из критериев является предпочтительней, какой из них оказывается наиболее мощным, против каких альтернатив и при каких объемах выборок конкретный критерий обладает преимуществом или наоборот. Поэтому представляет исследовательский интерес задача определения зависимости между объемом обрабатываемой выборки и корректностью полученных выводов при применении конкретного критерия, а также

обратная задача, обработка заданного количества данных, полученных по малым выборкам, и степень достоверности полученных результатов.

Постановка задачи

Целью исследования является разработка алгоритма определяющего последовательность действий исследователя при обработке данных малого объема, сравнительный анализ зависимостей мощностей критериев Граббса и Кохрена от объема исследуемых данных и решение проблемы выбора критерия в заданных условиях.

Определение аномальных значений в выборках малого объема

Первым этапом после получения результатов измерений является проверка на наличие аномальных значений, так называемых «выбросов», которые могут исказить истинную картину распределения данных в выборке. Для этой цели в [1] рекомендуется использовать критерий Кохрена при проверке внутрилабораторных расхождений и критерий Граббса при проверке межлабораторных расхождений. Однако конкретных доводов, почему это так не приводится. Критерий Граббса еще рекомендуется использовать в качестве критерия сравнения средних значений выборок между собой, а также для анализа данных в пределах одной лаборатории, для которой в результате проверки с использованием критерия Кохрена обнаруживается сомнительность значения стандартного отклонения.

В статье рассмотрены ситуации использования критериев Кохрена и Граббса при малых объемах выборок. А именно, смоделирована ситуация получения лабораториями равного количества результатов измерений – от

3 до 10 измерений в каждой, $A = \begin{bmatrix} a_{11} \dots a_{1i} \\ \dots \dots \dots \\ a_j \dots a_{ji} \end{bmatrix}$. Выборки независимы, и не

содержат аномальных значений, что подтверждалось проверкой того, что все значения лежат в пределах 2σ . Определяется максимальное значение по всем данным – $\max(A_{ij})$, причем это значение еще не является выбросом. Затем мы вводим искусственный выброс заменой $new = \max(A_{ij}) + k * \sigma$, величина k определялась минимальным значением, необходимым для того, чтобы новое максимальное значение выходило за рамки 2σ , после чего выброс определялся с помощью критерия Кохрена, по всем лабораториям

$$C = \frac{s_{\max}^2}{\sum_{i=1}^p s_i^2},$$

а так же с помощью критерия Граббса, применительно к каждой из лабораторий, проводилась проверка максимального значения в выборке

$$G_j = \frac{(a_{\max j} - \bar{a}_j)}{s_j}.$$

При этом были получены ситуации, в которых критерий Граббса безошибочно определял введенный выброс, в то время как критерий Кохрена не чувствовал его, поскольку выборка, содержащая выброс, не всегда имеет наибольшую дисперсию.

Методом Монте–Карло были получены и усреднены данные для 1000 смоделированных ситуаций, по этим данным построена круговая диаграмма, отображающая количество определенных выбросов каждым из критериев по отношению к общему количеству смоделированных ситуаций.

Величины P_{grabbs} и $P_{kochren}$ – это частоты определения каждым из сравниваемых критериев выброса по отношению к общему количеству смоделированных случаев (1000 случаев): $P_{grabbs} = \frac{N_{grabbs}}{1000}$ и

$$P_{kochren} = \frac{N_{kochren}}{1000}.$$

Как видно из рис. 1, критерий Кохрена определяет искусственно введенный выброс в 84% случаев, тогда как критерий Граббса не чувствует выброс. Нечувствительность критерия Граббса к выбросу при объеме выборки в 3 элемента обусловлено влиянием, вносимым в выборку выбросом [3].

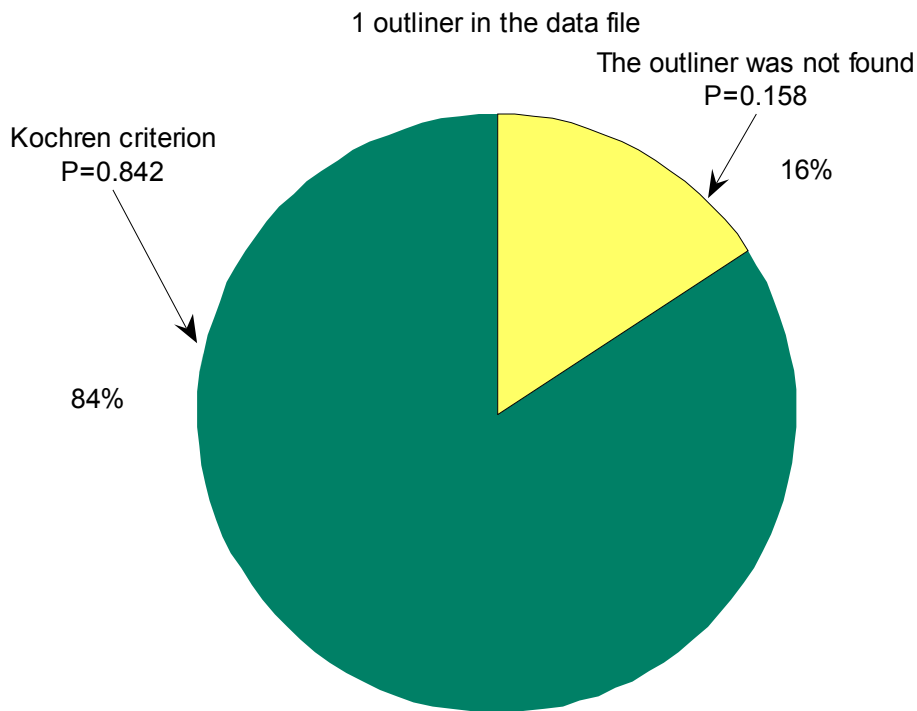


Рис. 1. Сравнительный анализ критериев Кохрена и Граббса:
количество лабораторий 3, количество измерений в каждой 3

На рис. 2 представлена аналогичная диаграмма, для 5 лабораторий, реализующих по 5 измерений каждая, и здесь критерий Граббса имеет большую чувствительность, определяя выброс в 21% случаев, тогда как критерий Кохрена только в 13%. Это означает, что в данном случае критерий Граббса может быть использован наравне с критерием Кохрена.

Но подобный выигрыш может быть получен при более значительных временных затратах, нежели при использовании критерия Кохрена, позволяющего сразу охватить весь объем исследуемых данных. Поэтому, в любом случае при выборе статистического критерия для обработки данных следует исходить из конкретно поставленной задачи и объема обрабатываемых данных.

Результаты, полученные для 7 лабораторий по 7 измерений в каждой, показывают, что вносимый выброс не оказывает существенного влияния на разброс данных при большом количестве результатов, что затрудняет его определение с помощью критерия Кохрена.

Рассмотренные результаты моделирующих экспериментов сведены в таблице.

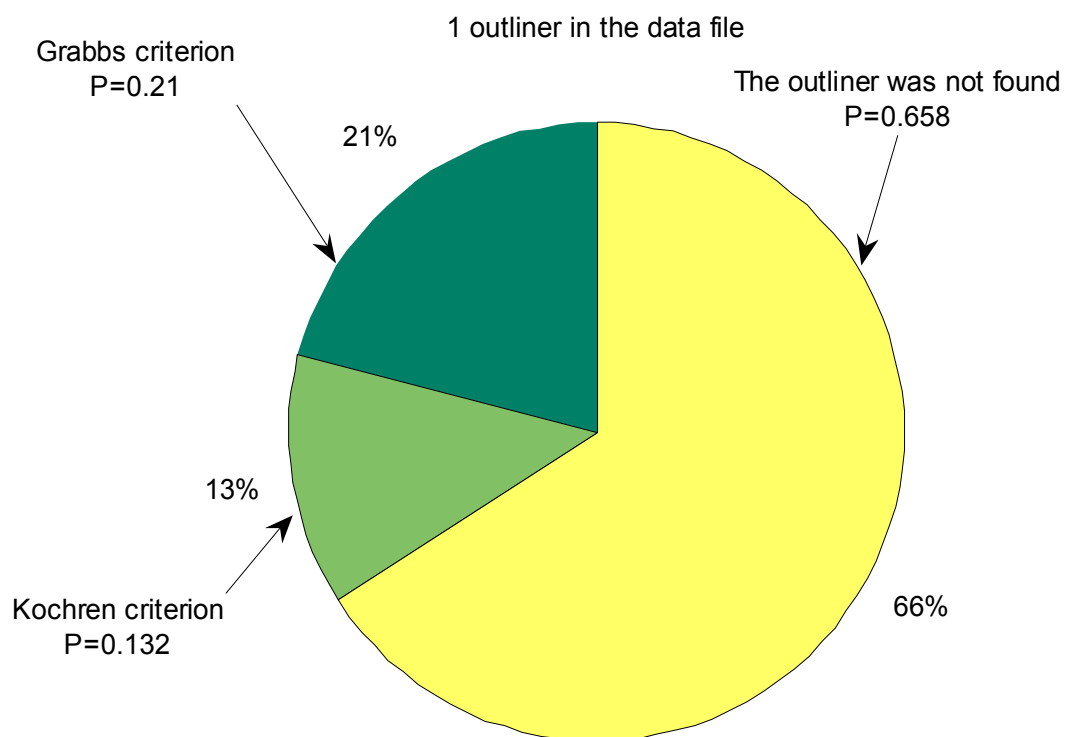


Рис. 2. Сравнительный анализ критериев Кохрена и Граббса: количество лабораторий 5, количество измерений в каждой 5

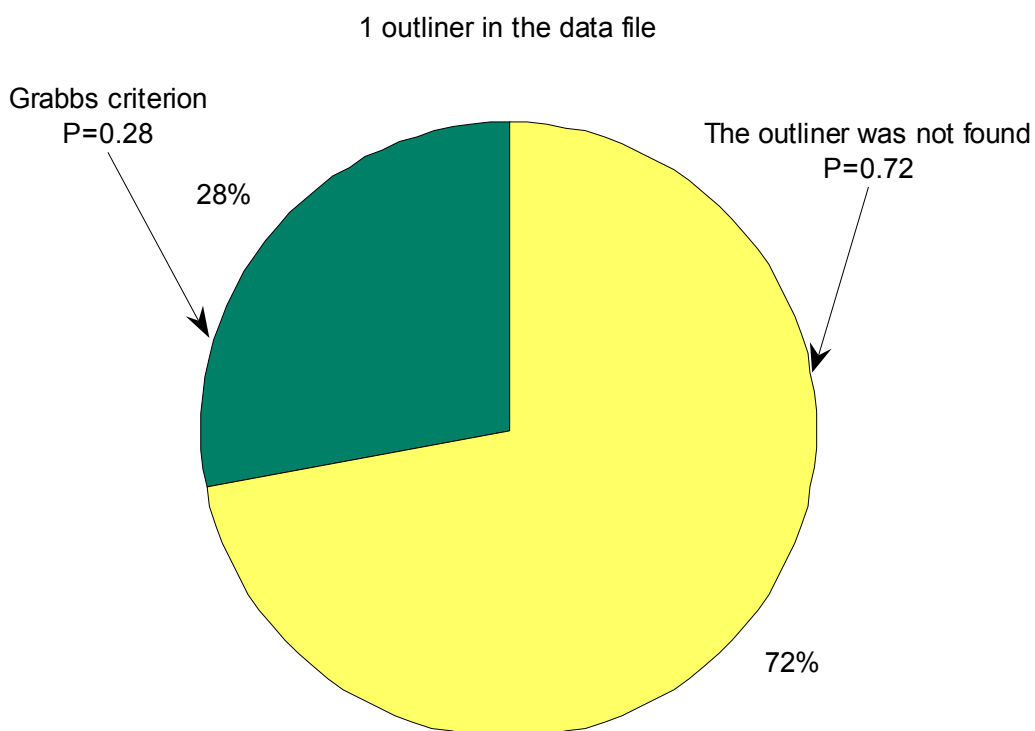


Рис. 3. Сравнительный анализ критериев Кохрена и Граббса: количество лабораторий 7, количество измерений в каждой 7

Таблица.

Мощность критериев Кохрена и Граббса в зависимости от объема сравниваемых выборок

n	3	4	5	6	7
критерий Граббса	$P \rightarrow 0$	$P = 0.158$	$P = 0.21$	$P = 0.214$	$P = 0.28$
критерий Кохрена	$P = 0.84$	$P = 0.504$	$P = 0.13$	$P = 0.012$	$P \rightarrow 0$

Из таблицы четко просматривается изменение вероятности определения выброса с изменением объема выборки, что можно использовать при выборе одного из сравниваемых критериев для дальнейшего использования.

На основе полученных результатов, в алгоритме, представленном на рис. 4, критерий Граббса был поставлен на один уровень с критерием Кохрена, для случая определения аномальных значений в выборке и ориентироваться при выборе одного из критериев необходимо на размер исследуемых выборок.

Возвращаясь к рекомендации в [1], об использовании критерия Граббса для нахождения межлабораторных расхождений, можно сделать предположение о том, что она основана на использовании опорного среднего значения вместо идеального, а опорное среднее, в свою очередь, приближается к идеальному значению при увеличении числа обрабатываемых данных. Поэтому при обработке малых выборок более целесообразным является сравнение средних значений исследуемых лабораторий по отношению к общему среднему по всем данным.

Алгоритм обработки данных для выборок малого объема

Как уже рассматривалось выше, результаты, полученные с помощью критериев Кохрена и Фишера, являются зависимыми от количества измерений в лабораториях. Проблемой критерия Граббса для выборок малого объема является его неспособность определить выброс при количестве элементов в выборке равное трем [3], однако именно критерий Граббса является более чувствительным к выбросу с ростом числа элементов в выборке. На рис. 4 приведен общий алгоритм этапов обработки результатов измерений.

Следующим этапом после выявления аномальных значений в выборках является сравнение средних значений, и для этого чаще всего

используется критерий Стьюдента. Однако способность критерия Стьюдента различать расхождение между средними значениями также зависит от объема исследуемых выборок [4]. Например, для минимального объема выборки в 2 элемента, для того, чтобы получить результат о различии между средними с вероятностью $P = 0.95$, разница между средними должна быть $\mu_1 - \mu_2 > 2.5\sigma$, т.е. при одинаковых дисперсиях центры распределений находятся в зоне между 2σ и 3σ , т.е. в зоне наибольших допустимых отклонений друг от друга. Поэтому при использовании критерия Стьюдента, исходя из имеющегося количества данных, можно определить, каким должно быть различие между средними значениями, для того чтобы критерий определил его с необходимым уровнем вероятности ($P = 0.8; 0.9; 0.95; 0.99$).

Критерий Колмогорова – Смирнова, как и многие другие критерии, основан на предельном распределении некоторой статистики при бесконечном увеличении объемов выборок. Поэтому уровень значимости, с которым используется тест при малых объемах выборок, весьма приблизительно отражает вероятность реальной ошибки при отклонении основной гипотезы.

Выводы

Проведенное исследование показало, что результаты обработки данных, полученные при использовании критериев Кохрена, Граббса и Стьюдента являются зависимыми от объемов исследуемых выборок. А именно, определена не способность критерия Граббса, идентифицировать выброс при выборке в 3 элемента, однако при увеличении числа элементов в выборке уже именно критерий Граббса является более чувствительным к выбросу, что позволяет использовать его не как дополнительный инструмент проверки после применения критерия Кохрена, а наравне с ним.

Полученный результат дает возможность, определить какой из критериев, Кохрена или Граббса, может быть применен для выявления аномального значения в имеющихся объемах данных и с какой мощностью, или же, исходя из требуемой мощности, определить необходимое количество элементов выборок и критерий для их обработки.

Разработанный алгоритм позволяет исследователю выбирать на каждом этапе обработки данных, какой из критериев оказывается предпочтительней и при каких условиях.

Дальнейшее изучение методов обработки малых выборок (до 30 элементов) позволит избежать получения недостоверных результатов, как следствие некорректного использования статистических критериев, рассчитанных на большие объемы данных.

Список использованной литература

1. ГОСТ ИСО 5725–2–2002–1 Точность (прецизионность и правильность) результатов измерений. Часть 2. Основной метод определения повторяемости и воспроизводимости стандартного метода измерений.
2. Еременко В.С., Мокийчук В.М., Самойличекно О.В. Исследование мощности критерия Кохрена при ограниченном числе наблюдений // Системы обработки информации, – 2007. –№6(64) – С.35–38.

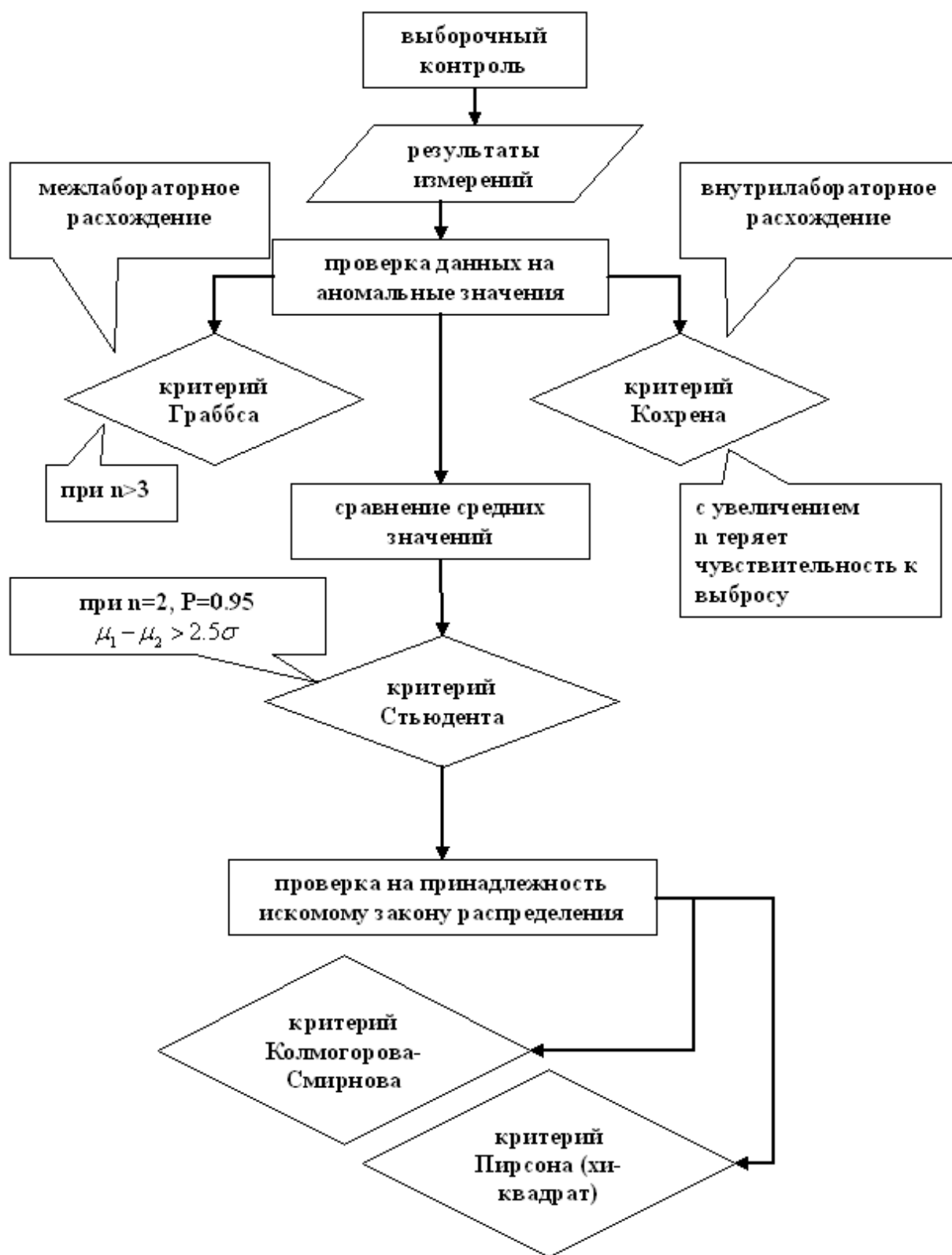


Рис. 4. Алгоритм обработки данных

3. Володарский Е.Т., Козыр Е.В. Границы применимости вероятностно–статистических критериев в условиях малых выборок //6 международная конференция Гиротехнологии, навигация, управление движением и конструирование авиационно–космической техники. Сборник докладов. Часть 2/К.НТУУ «КПИ», 26–27 апреля 2007. – С.70.
4. Володарский Е.Т., Козыр Е.В. Влияние объема выборки при проведении контрольно–сличительных испытаний // Вестник Хмельницкого национального университета, Том 1 – 2007. – №3 – С.147–149.
5. Смирнов Н.В., Дунин–Борковский И.В. Курс теории вероятностей и математической статистики для технических приложений. М.: Наука, 1965.–511с.
6. Гришин В.К. Статистические методы анализа и планирования экспериментов. Издательство Московского университета, 1975. – 127с.
7. Джонсон Н., Лион Ф. Статистика и планирование эксперимента в технике и науке. Методы обработки данных, М.: Мир, 1980. – 610с.