Только для студентов, магистрантов и аспирантов НГТУ

Б.Ю. ЛЕМЕШКО
И.В. ВЕРЕТЕЛЬНИКОВА

КРИТЕРИИ ПРОВЕРКИ ГИПОТЕЗ О СЛУЧАЙНОСТИ И ОТСУТСТВИИ ТРЕНДА

РУКОВОДСТВО ПО ПРИМЕНЕНИЮ

2021
Новосибирск

УДК 519.23

Данное руководство предназначено для использования в качестве одного из учебных пособий для магистрантов факультета прикладной математики и информатики Новосибирского государственного технического университета, осваивающих курс “Компьютерные технологии моделирования и анализа данных”, которое должно способствовать критическому восприятию возможностей множества классических критериев при проверке гипотез о случайности и отсутствия тренда, пониманию возможностей методов статистического моделирования при исследовании вероятностных и статистических закономерностей.

Книга рассчитана на специалистов, интересующихся вопросами применеия статистических методов для анализа различных аспектов и тенденций окружающей действительности, соприкасающихся в своей деятельности с обработкой результатов экспериментов, с необходимостью анализа данных.

В руководстве рассматриваются вопросы применения статистических критериев, ориентированных на проверку гипотез об отсутствии тренда в анализируемых выборках. Неотклонение такой гипотезы даёт основание рассматривать анализируемые данные как выборки независимых одинаково распределенных случайных величин.

Рассматривается множество специальных критериев, ориентированных на проверку такого рода гипотез, а также множество критериев однородности законов, однородности средних и однородности дисперсий, которые также могут применяться в указанных целях.

Подчеркиваются недостатки и преимущества различных критериев, рассматривается применение критериев в условиях нарушения стандартных предположений. Приводятся оценки мощности критериев, что позволяет ориентироваться при выборе наиболее предпочтительных критериев.

Следование рекомендациям обеспечит корректность и повысит обоснованность статистических выводов при анализе данных.

Книга будет полезна инженерам, научным сотрудникам, специалистам различного профиля (медикам, биологам, социологам, экономистам, и др.), сталкивающимся в своей деятельности с необходимостью статистического анализа результатов экспериментов. Руководство будет полезно преподавателям вузов, аспирантам и студентам.

Оглавление

ПРЕДИСЛОВИЕ

В различных приложениях часто возникает необходимость убедиться в том, что в рядах измерений некоторой величины или в наблюдаемой последовательности не появилось значимых изменений, свидетельствующих о нарушении ранее установленных закономерностей. Например, хотят убедиться в отсутствии дрейфа в значениях измеряемого параметра, или убедиться в неизменности характеристик точности измерительной системы в течение определённого периода времени.

В прикладной математической статистике накопился достаточно обширный арсенал критериев (параметрических и непараметрических), предназначенных для проверки гипотез об отсутствии каких-либо изменений в наблюдаемой последовательности измерений. Множество таких критериев иногда объёдиняют названием “критерии проверки случайности или отсутствии тренда”, понимая под этим, что если проверяемая гипотеза не отклонена, то анализируемый ряд можно рассматривать как выборку независимых одинаково распределённых случайных величин.

Следует отметить, что для решения этих же и близких задач, может использоваться множество критериев проверки однородности законов, однородности средних, однородности дисперсий [134].

Вместе с тем приходится констатировать, что на практике оба эти множества критериев используется достаточно редко, что связано с рядом причин. Во-первых, заинтересованные специалисты просто не обладают необходимыми знаниями о совокупности таких критериев и, тем более, о реальных свойствах этих критериев. Во-вторых, в нестандартных условиях приложений реальные свойства критериев могут существенно отличаться от асимптотических. В-третьих, многие из критериев затруднительно использовать без соответствующего программного обеспечения. А если говорить о применении критериев в условиях нарушения стандартных предположений, то необходимы специальные программные средства, обеспечивающие корректность статистических выводов в нестандартных условиях.

В основу настоящего руководства легли наши результаты, полученные при исследовании свойств рассмотренных в работе критериев. Это численные исследования, опирающиеся на интенсивное использование методов статистического моделирования. Обобщение результатов в программной системе [148] обеспечивает корректность статистических выводов при использовании системы для проверки гипотез по рассмотренным критериям (в том числе, в нестандартных условиях).

Надеюсь, что данное руководство, как рекомендации [145, 146] и предшествующие книги [130, 132, 131, 133, 134], окажет реальную помощь специалистам, заинтересованным в корректности проводимого статистического анализа.

Я очень признателен Лемешко С.Б., в немалой степени усилиями которого, развивается система [148], благодаря чему и стала возможна подготовка данного руководства.

Б.Ю. Лемешко
февраль 2021 г.

ВВЕДЕНИЕ

Под временным рядом понимается упорядоченная последовательность $\{X_i,\ i=\overline{1,n}\}$ результатов измерений значений некоторой переменной, произведенных, например, через равные промежутки времени. С анализом временных рядов сталкиваются при обработке измерений в технических областях, в экономике, сельском хозяйстве, метеорологии [100]. Временные ряды могут отражать наличие тренда, систематической составляющей, сезонной составляющей, а также случайной составляющей (шума). Под трендом понимается основная тенденция изменения временного ряда, определяющая направление преимущественного движения исследуемой переменной.

Рассматриваемое в данном руководстве множество критериев иногда называют критериями проверки гипотез о случайности и отсутствии тренда [109]. По существу проверяемая гипотеза H_0 в таких критериях, заключается в том, что наблюдаемый ряд измерений представляет собой выборку независимых, одинаково распределённых случайных величин (н.о.р.с.в.). Именно в этом смысле надо понимать проверку гипотезы о случайности в контексте данного руководства.

Отклонение проверяемой гипотезы может являться признаком наличия в наблюдаемой случайной последовательности измерений некоторой неслучайной закономерности, коррелированности наблюдений, присутствия грубых ошибок (выбросов), и всего того, что заставляет нас усомниться в том, что мы имеем дело с выборкой н.о.р.с.в.

Причина отклонения проверяемой гипотезы H_0 может быть связана с наличием тренда, но сам факт отклонения ничего не говорит о характере тренда. Определение характера тренда это уже предмет другой задачи.

Для проверки гипотезы о случайности и отсутствии тренда в разное время предложено множество параметрических и непараметрических критериев. Однако имеющиеся источники не позволяют судить о преимуществах тех или иных критериев, не содержат четких рекомендаций, очерчивающих область применения и предпосылки, выполнение которых обеспечивает корректность статистических выводов при использовании рассматриваемых критериев.

Достаточно полный перечень критериев, ориентированных на проверку гипотезы о случайности и об отсутствии тренда, представлен в работе [109], которую можно рассматривать как справочное пособие, охватывающее самое широкое (в отечественных источниках по прикладной статистике) множество критериев проверки статистических гипотез. Естественно, что книга [109] не дает ответа на сформулированные выше вопросы. Более того, пользоваться ею надо осторожно из-за существенного числа ошибок, допущенных в описаниях критериев и в примерах их применения.

Основной предпосылкой, обеспечивающей корректное применение параметрических критериев, как правило, является предположение о нормальном законе распределения шума, что далеко не всегда выполняется на практике. Возникает вопрос, что произойдет с распределением статистики этого критерия в случае нарушения предположения о нормальности? Насколько будут оставаться корректными выводы, осуществляемые на основании классических результатов?

Использование непараметрических критериев опирается на асимптотические распределения статистик этих критериев. При ограниченных объемах выборок распределения статистик параметрических и непараметрических критериев могут существенно отличаться от соответствующих предельных и асимптотических распределений статистик, используемых в процедуре проверки гипотезы. В случае непараметрических критериев проблема зачастую усугубляется из-за ярко выраженной дискретности статистики. В таких ситуациях использование при проверке гипотезы предельного (асимптотического) распределения статистики вместо “истинного” распределения этой статистики может приводить к неверному выводу.

Исследователя в связи с необходимостью решения задачи статистического анализа может интересовать критерий, обладающий наибольшей мощностью против заданной конкурирующей гипотезы. В зависимости от конкурирующей гипотезы более предпочтительными по мощности могут оказаться (и оказываются) различные критерии. В то же время, как правило, параметрические критерии показывают более высокую мощность по сравнению со своими непараметрическими аналогами, особенно, если конкурирующая гипотеза связана с наличием тренда в характеристиках рассеяния наблюдаемого процесса.

Таким образом, любого исследователя, с одной стороны, интересуют “действительные” свойства критериев, интересует, когда и при каких условиях (при выполнении каких стандартных предположений) он может получить корректные выводы, используя классические результаты. С другой стороны, он заинтересован в возможности осуществления корректных выводов и в ситуациях, когда стандартные предположения нарушаются и использование классических результатов невозможно.

Также исследователя может интересовать наличие или отсутствие тренда в математическом ожидании (в среднем) и в дисперсии (в характеристиках рассеяния).

При проверке гипотезы об **отсутствии тренда в математическом ожидании** задача формулируется следующим образом. Предполагается, что наблюдается временной ряд значений $X_1, X_2, ..., X_n$ взаимно независимых случайных величин с математическими ожиданиями $\mu_1, \mu_2, ..., \mu_n$ и одинаковыми (но неизвестными) дисперсиями. Проверяется гипотеза $H_0: \ \mu_i = \mu$, $i = 1, 2, ..., n$, о том, что все выборочные значения принадлежат к одной генеральной совокупности со средним μ, против конкурирующей гипотезы о наличии тренда $H_1: |\mu_{i+1} - \mu_i| > 0$, $i = 1, 2, ..., n-1$.

Проверяемая **гипотеза об отсутствии тренда в дисперсии** формулируется аналогичным образом (проверяется $H_0: \sigma_i = \sigma$, $i = 1, 2, ..., n$, против $H_1: |\sigma_{i+1} - \sigma_i| > 0$, $i = 1, 2, ..., n-1$).

При **проверке отсутствия сдвига в дисперсии** (в характеристиках рассеяния) предполагается, что наблюдаемая последовательность измерений $X_1, X_2, ..., X_n$ имеет одно и то же среднее μ. Проверяется гипотеза $H_0: \ \sigma_1^2 = ... = \sigma_n^2 = \sigma_0^2$ (σ_0^2 неизвестно) против конкурирующей гипотезы

$$H_1: \ \sigma_1^2 = \sigma_2^2 = ... = \sigma_k^2 = \sigma_0^2; \ \sigma_{k+1}^2 = \sigma_{k+2}^2 = ... = \sigma_n^2 = \sigma_0^2 + \delta, \ (\delta > 0),$$

утверждающей, что значение дисперсии меняется в некоторой неизвестной точке, то есть k неизвестно ($1 \le k \le n-1$).

В настоящем руководстве рассматриваются все перечисленные ситуации. По результатам исследования свойств критериев и оценок их мощности относительно различных конкурирующих гипотез можно судить о способности критериев обнаруживать наличие тренда в наблюдаемых процессах.

В ходе исследований, проведенных при подготовке руководства и направленных на сравнительный анализ свойств критериев, применяемых при проверке гипотез о случайности или об отсутствии тренда, преследовались следующие цели:

- выяснить, как влияет объем выборки n на распределения статистик критериев при справедливости проверяемой гипотезы H_0;
- уточнить, начиная с каких значений n, можно использовать асимптотическое (предельное) распределение статистики соответствующего критерия (при условии существования такого распределения) вместо действительного распределения статистики, имеющего место при данном n;
- при необходимости уточнить таблицы критических значений (процентных точек) статистик критериев;
- для параметрических критериев выяснить, как влияет на распределения статистик (при справедливости проверяемой гипотезы H_0) нарушение стандартных предположений (отклонение наблюдаемого закона от нормального);
- оценить мощность критериев по отношению к различным близким конкурирующим гипотезам и на основании сравнительного анализа указать, применение каких критериев являются предпочтительным в тех или иных ситуациях.

Для проверки гипотезы об отсутствии тренда могут использоваться многочисленные критерии проверки однородности законов, однородности средних, однородности дисперсий [136]. Именно поэтому в данном руководстве рассмотрены и критерии такого вида. Рекомендации по применению множества критериев однородности, учитывающие реальные свойства этих критериев, наиболее полно отражены в руководстве [134].

В целях поиска различий в закономерностях на участках временного ряда в программной системе [148] предусмотрена специальная возможность разбиения анализируемой выборки на произвольные части, к которым могут применяться любые критерии однородности, рассмотренные в [134] и в разделах 3–5 настоящего руководства.

1. ОБЩИЕ ПОЛОЖЕНИЯ

1.1. Общие сведения о проверке статистических гипотез

С каждым из используемых для проверки гипотезы H_0 критериев связана некоторая статистика S, которая в соответствии с некоторой мерой измеряет отклонение в наблюдаемом процессе от ситуации, соответствующей H_0.

В силу случайности извлекаемых выборок случайными оказываются и значения статистики S, вычисляемые в соответствии с этими выборками. При справедливости проверяемой гипотезы H_0 статистика S подчиняется некоторому распределению $G(S|H_0)$.

Схема проверки гипотезы заключается в следующем. Область определения статистики разбивается на два подмножества, одно из которых представляет собой критическую область, и попадание в которую при справедливости H_0 маловероятно. При попадании вычисленного по выборке $X_1, X_2, ..., X_n$ значения S^* статистики S в критическую область проверяемая гипотеза H_0 отклоняется (отвергается). В противном случае – нет оснований для отклонения гипотезы H_0.

С проверкой статистических гипотез связывают ошибки двух видов. Справедливая гипотеза H_0 может быть отклонена и этим самым совершена ошибка 1-го рода. При проверке гипотез вероятность ошибки 1-го рода α (уровень значимости), как правило, задают, допуская тем самым возможность отклонения H_0 и возможность такой ошибки. С другой стороны, может быть справедлива некоторая конкурирующая гипотеза H_1. Если при справедливости H_1 в процессе проверки гипотеза H_0 не была отклонена, то этим самым совершена ошибка 2-го рода. Вероятность ошибки 2-го рода, как правило, обозначают β.

При построении критериев стремятся к использованию одномерных статистик, что упрощает построение критической области. При этом критерии могут быть правосторонними, левосторонними и двусторонними, что определяет построение критической области.

В случае правостороннего критерия граница критической области (критическое значение) $S_{1-\alpha}$, определяется уравнением

$$\alpha = \int\limits_{S_{1-\alpha}}^{\infty} g(s|H_0)ds = 1 - G(S_{1-\alpha}|H_0), \tag{1.1}$$

где $g(s|H_0)$ − условная плотность распределения статистики при справедливости H_0.

Для используемых на практике критериев в благоприятных случаях известны асимптотические (предельные) распределения $G(S|H_0)$ соответствующих статистик при справедливости гипотезы H_0. В тех ситуациях, когда распределения статистик существенно зависят от объёмов выборок n, информация о законе распределения статистики бывает представлена таблицей процентных точек (квантилей распределения $G(S|H_0)$). Критическое значение $S_{1-\alpha}$ вычисляют в соответствии с $G(S|H_0)$ или берут из соответствующей таблицы процентных точек.

В случае правостороннего критерия в принятой практике статистического анализа обычно полученное значение статистики S^* сравнивают с критическим значением $S_{1-\alpha}$ при заданном уровне значимости α. Проверяемую гипотезу H_0 отклоняют, если $S^* > S_{1-\alpha}$ (см. рис. 1.1).

Больше информации о степени соответствия выборки теоретическому закону можно почерпнуть из «достигнутого уровня значимости» (p_{value}): вероятности возможного превышения полученного значения статистики при справедливости H_0

$$P\{S > S^*\} = \int\limits_{S^*}^{\infty} g(s|H_0)ds = 1 - G(S^*|H_0). \tag{1.2}$$

Именно эта вероятность позволяет судить о том, насколько хорошо выборка согласуется с проверяемой гипотезой H_0 (см. рис. 1.2). Проверяемую гипотезу H_0 не отвергают, если $P\{S > S^*\} > \alpha$.

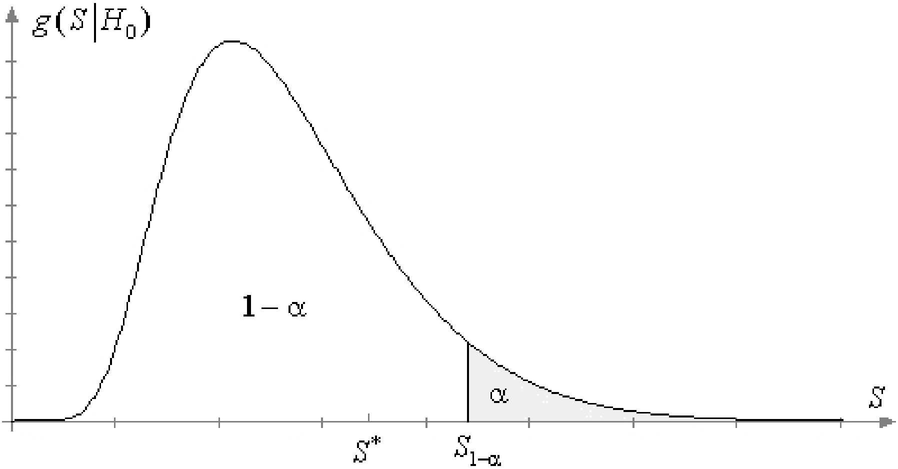


Рис. 1.1. Плотность распределения статистики при справедливости гипотезы H_0 и критическое значение для правостороннего критерия

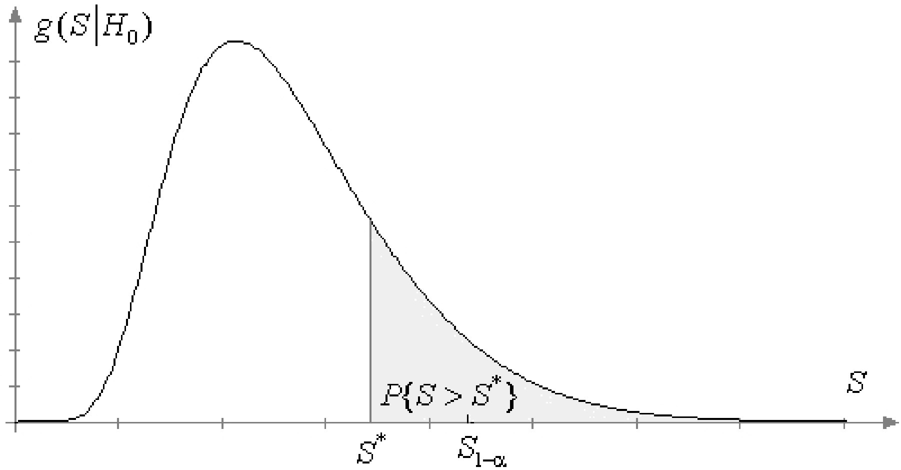


Рис. 1.2. Плотность распределения статистики при справедливости гипотезы H_0 и достигнутый уровень значимости

В случае двустороннего критерия критическая область состоит из двух частей. И проверяемая гипотеза H_0 отклоняется, если $S^* < S_{\alpha/2}$ или $S^* > S_{1-\alpha/2}$ (см. рис. 1.3). А достигнутый уровень значимости p_{value} в этом случае определяется соотношением

$$p_{value} = 2\min\left\{G(S^*|H_0),\, 1 - G(S^*|H_0)\right\}. \quad (1.3)$$

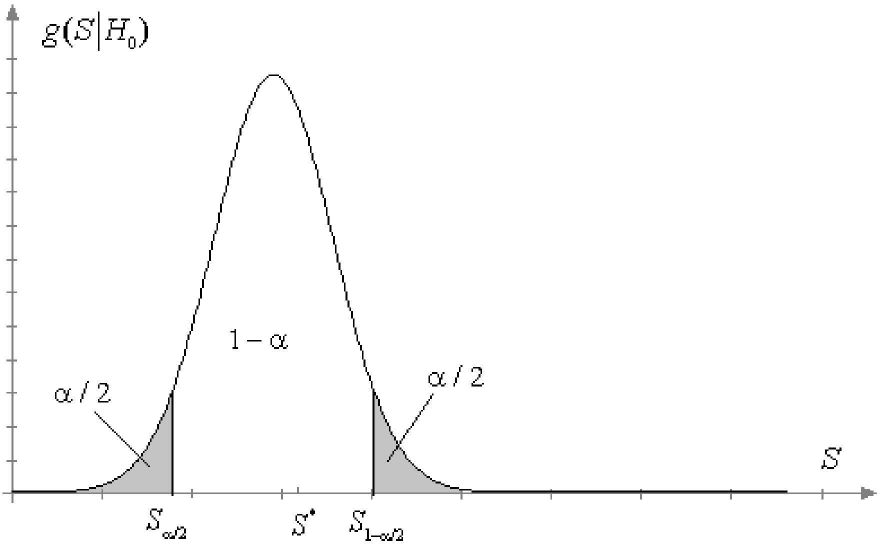


Рис. 1.3. Плотность распределения статистики при справедливости гипотезы H_0 и критические значения для двустороннего критерия

Обычно, используя критерии проверки гипотез, не рассматривают конкретную конкурирующую гипотезу. В таком случае при проверке гипотез о случайности или об отсутствии тренда можно считать, что конкурирующая гипотеза связана с наличием, например, какого-то тренда в наблюдаемом процессе.

Если же гипотеза H_1 задана и связана, например, с наличием тренда некоторого заданного вида, то задание величины α для используемого критерия проверки гипотез определяет и вероятность ошибки

2-го рода β. Вероятность ошибки 2-го рода β для правостороннего критерия определяется выражением

$$\beta = \int_{-\infty}^{S_{1-\alpha}} g(s|H_1)ds \,, \tag{1.4}$$

а для двустороннего – соотношением

$$\beta = \int_{S_{\alpha/2}}^{S_{1-\alpha/2}} g(s|H_1)ds \,. \tag{1.5}$$

Для конкретной альтернативы H_0 и H_1 задание вероятности ошибки 1-го рода определяет и вероятность ошибки 2-го рода. Рис. 1.4 поясняет это для правостороннего критерия. На рис. 1.4 $g(s\,|\,H_0)$ отображает плотность распределения статистики S при справедливости гипотезы H_0, а $g(s\,|\,H_1)$ – плотность распределения при справедливости H_1.

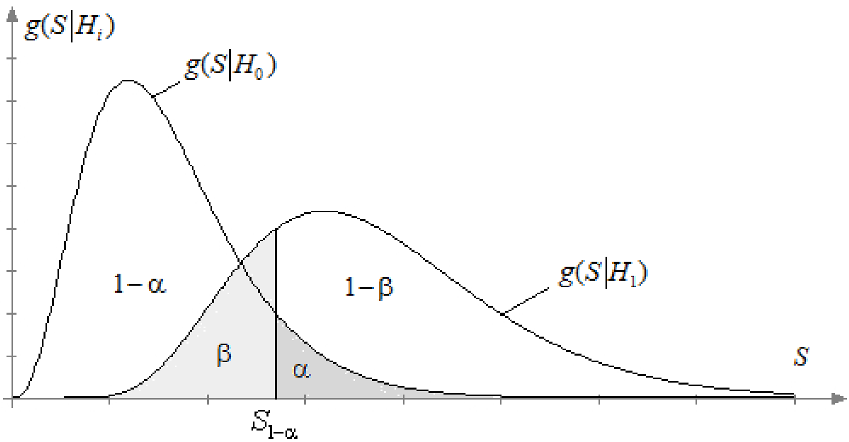


Рис. 1.4. Плотности распределения статистик при справедливости соответственно гипотез H_0 и H_1 в случае правостороннего критерия

Мощность критерия представляет собой величину $1-\beta$. Очевидно, что чем выше мощность используемого критерия при заданном значении α, тем лучше он различает гипотезы H_0 и H_1. Особенно важно, чтобы используемый критерий хорошо различал близкие конкурирующие гипотезы. Графически требование максимальной мощности критерия означает, что на рис. 1.4 плотности распределений статистики $g(s|H_0)$ и $g(s|H_1)$ должны быть максимально "раздвинуты".

Аналогичным образом можно проиллюстрировать вероятности ошибок 2-го рода и мощности для двустороннего критериев (см. рис.1.5).

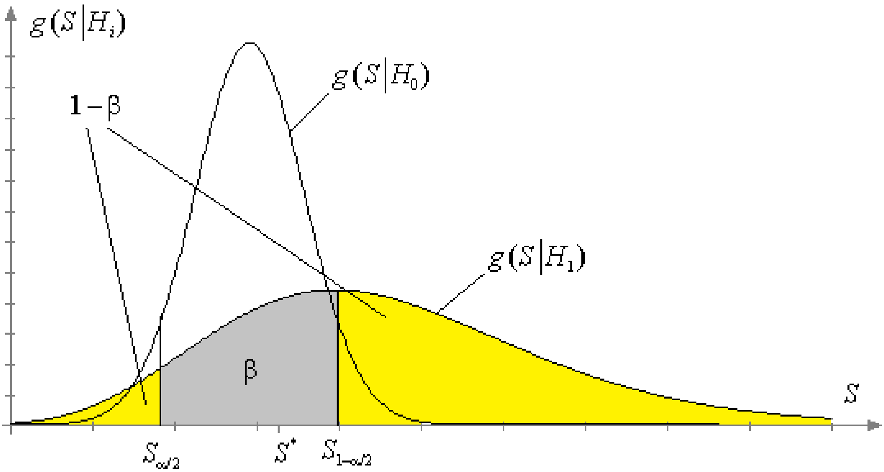


Рис. 1.5. Плотности распределения статистик при справедливости соответственно гипотез H_0 и H_1 в случае двустороннего критерия

Очевидно, что при проверке любой статистической гипотезы желательно использовать наиболее мощный критерий, который для заданной вероятности α ошибки 1-го рода обеспечивает минимальную вероятность β ошибки 2-го рода относительно любой конкурирующей гипотезы H_1.

Лучше всего использовать равномерно наиболее мощный критерий, который для любого заданного α обеспечивает минимальное значение β. Однако, существование такого критерия для проверки конкретной гипотезы H_0 является редчайшим исключением.

1.2. Конкурирующие гипотезы, рассматриваемые при анализе мощности критериев

Анализ мощности рассматриваемых в руководстве критериев проводился для ситуации принадлежности наблюдаемых случайных величин нормальному закону. Проверяемой гипотезе H_0 соответствует выполнение предположения о независимости наблюдаемых случайных величин (отсутствие тренда).

В качестве конкурирующих гипотез рассматривались различные ситуации, соответствующие наличию тренда в математическом ожидании или в дисперсии.

При исследовании мощности в качестве конкурирующих гипотез с наличием тренда в математических ожиданиях рассматривались модели с заданием линейного, периодического и смешанного тренда [121, 91, 103] вида:

$$x_i = a \cdot t_i + b \cdot \sin(2k\pi t_i) + \xi_i,$$

где ξ_i представляют собой независимые случайные величины, распределённые по заданному закону (например, по стандартному нормальному закону), $t \in [0,1]$.

В этом случае справедливой проверяемой гипотезе H_0 соответствует значение параметров $a = 0$, $b = 0$.

Отсутствию периодической составляющей тренда соответствует значение параметра $b = 0$, а отсутствию линейной – $a = 0$. Моменты отсчетов t_i вычислялись в соответствии с выражением $t_i = (i-1)\Delta t$, где шаг $\Delta t = 1/n$ связывался с объемом выборки n.

При этом, если не указывалась конкретная параметрическая модель закона, то псевдослучайные величины ξ_i генерировались в соответствии со стандартным нормальным законом.

Оценки мощности рассматриваемых в руководстве критериев исследовалась относительно конкретных конкурирующих гипотез с линейным трендом при заданных значениях параметра a = 0.5; 1.5; 4. Соответствующие конкурирующие гипотезы обозначены в дальнейшем как:

$$H_1: \; X_i = 0.5t_i + \xi_i;$$

$$H_2: \; X_i = 1.5t_i + \xi_i;$$

$$H_3: \; X_i = 4t_i + \xi_i.$$

Пример реализации временного ряда с наложенным трендом, соответствующим гипотез H_1, H_2, H_3, при объеме выборки $n = 100$ приведен на рис. 1.6.

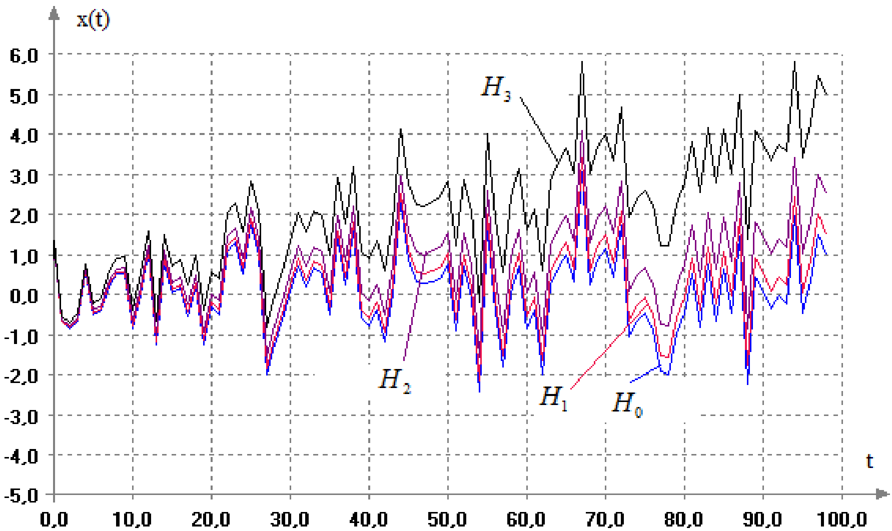


Рис. 1.6. Временной ряд с наложенным линейным трендом

Заметим, что при единичной дисперсии ($\sigma^2 = 1$) наблюдаемой случайной величины, а именно такая ситуация, как правило, для определённости рассматривалась в настоящих исследованиях, гипотеза H_1 очень близка к гипотезе H_0. То есть такой линейный тренд трудно обнаружить. Но при уменьшении дисперсия эти гипотезы становятся более далёкими, и различие между ними обнаруживаются легче.

При анализе мощности относительно периодического тренда в математическом ожидании рассматривались конкурирующие гипотезы:

$$H_4: \; X_i = 0.5\sin(2\pi t_i) + \xi_i;$$

$$H_5: \; X_i = 0.25\sin(4 \cdot 2\pi t_i) + \xi_i;$$

$$H_6: \; X_i = 0.5\sin(4 \cdot 2\pi t_i) + \xi_i;$$

$$H_7: \; X_i = \sin(4 \cdot 2\pi t_i) + \xi_i.$$

Пример реализации временного ряда при справедливости H_0 и H_4 представлен на рис. 1.7, а при справедливости H_0 и H_7 – на рис. 1.8.

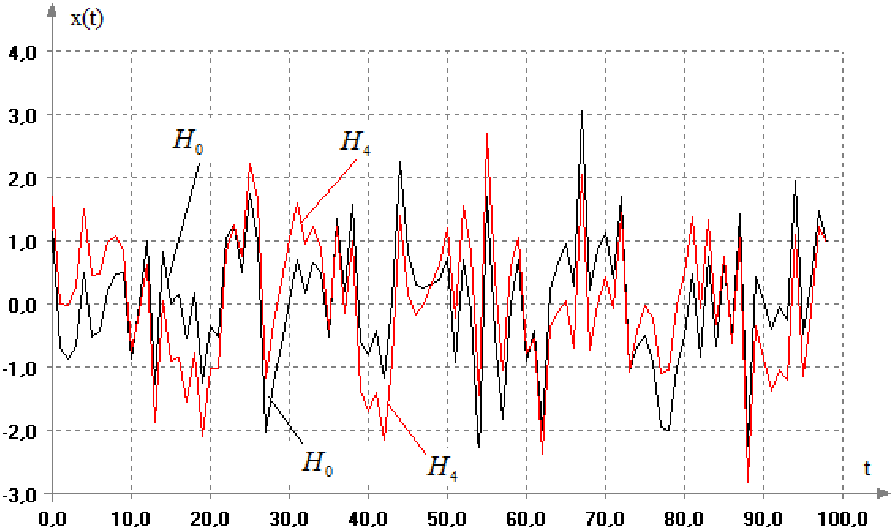


Рис. 1.7. Наложенный периодический тренд при гипотезе H_4

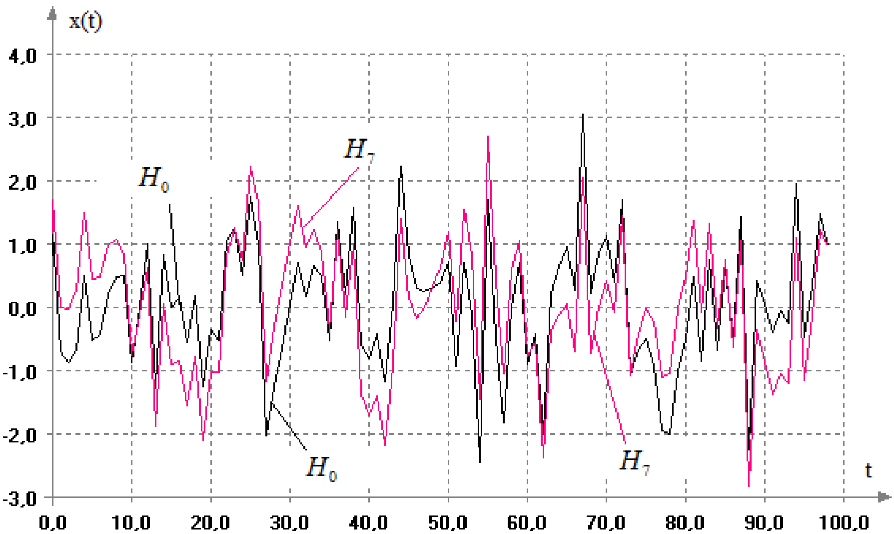


Рис. 1.8. Наложенный периодический тренд при гипотезе H_7

Для критериев обнаружения изменения дисперсии в неизвестной точке при анализе мощности критериев в качестве конкурирующих гипотез (при нормальном распределении случайных величин) рассматривались следующие близкие к H_0 гипотезы, когда в некоторый момент стандартное отклонение увеличивалось на 5, 10, 15%:

$$H_8\ :\ \sigma_1^2 = \ldots = \sigma_k^2 = 1;\ \ \sigma_{k+1}^2 = \ldots = \sigma_n^2 = 1.1025;$$

$$H_9\ :\ \sigma_1^2 = \ldots = \sigma_k^2 = 1;\ \ \sigma_{k+1}^2 = \ldots = \sigma_n^2 = 1.21;$$

$$H_{10}\ :\ \sigma_1^2 = \ldots = \sigma_k^2 = 1;\ \ \sigma_{k+1}^2 = \ldots = \sigma_n^2 = 1.3225,$$

где $k = n/2$. Это очень близкие конкурирующие гипотезы, которые трудно отличить от H_0 при малых объёмах выборок.

В качестве более далеких гипотез рассматривались конкурирующие гипотезы

$$H_{11}\ :\ \sigma_1^2 = \ldots = \sigma_k^2 = 1;\ \sigma_{k+1}^2 = \ldots = \sigma_n^2 = 2;$$

$$H_{12}\ :\ \sigma_1^2 = \ldots = \sigma_k^2 = 1;\ \sigma_{k+1}^2 = \ldots = \sigma_n^2 = 4.$$

На рис. 1.9 показаны реализации временного ряда, связанные с изменением дисперсии, в случае гипотез H_0, H_8 и H_{12}.

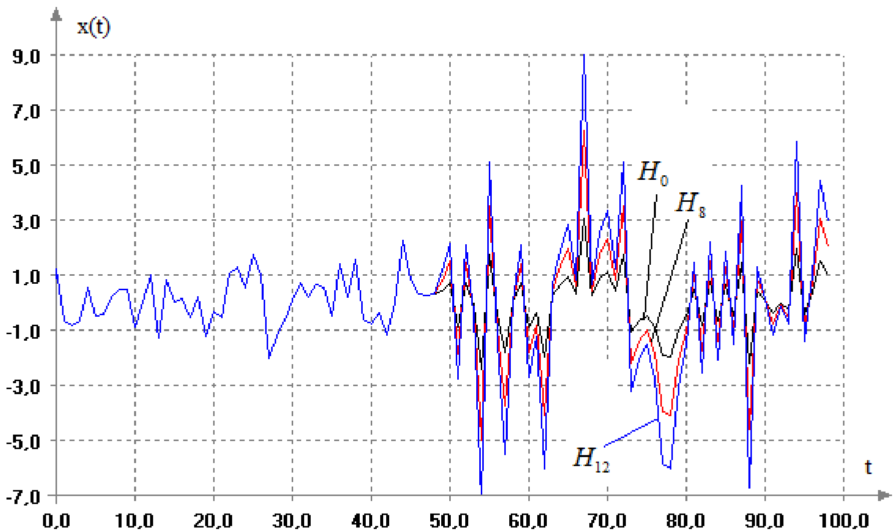


Рис. 1.9. Временной ряд при изменении дисперсии скачком

Наличие линейного тренда в характеристиках рассеяния наблюдаемого ряда случайных величин (изменение масштабного параметра) на интервале $t \in [0,1]$ может моделироваться в соответствии с соотношением

$$X_i = \xi_i (1 + c t_i),$$

где $c \in (-1, \infty)$, $t_i = (i-1)\Delta t$, $\Delta t = 1/n$ Справедливой проверяемой гипотезе H_0 соответствует значение параметра $c = 0$.

В случае наличия периодического тренда в характеристиках рассеяния случайные величины могут моделироваться, например, в соответствии с соотношением

$$X_i = \xi_i (1 + d \cdot \sin(2k\pi t_i))$$

при $|d| < 1$. В случае смешанного тренда – в соответствии с выражением

$$X_i = \xi_i (1 + c t_i + d \sin(2k\pi t_i)),$$

при $|d| < 1$, если $c \ge 0$, и при $|d| < 1 + c$, если $c \in (-1, 0)$. Отсутствию периодической составляющей тренда соответствует значение параметра $d = 0$, а отсутствию линейной – $c = 0$.

При анализе мощности относительно линейного, периодического и смешанного тренда в характеристиках рассеяния (в дисперсии) случайной величины рассматривались конкурирующие гипотезы:

$$H_{13}: \ X_i = \xi_i (1 + c t_i), \ c = 1;$$

$$H_{14}: \ X_i = \xi_i (1 + d \cdot \sin(2k\pi t_i)), \ d = 0.8, \ k = 2;$$

$$H_{15}: \ X_i = \xi_i (1 + c t_i + d \sin(2k\pi t_i)), \ c = 1, \ d = 0.8, \ k = 2.$$

Примеры реализации временных рядов при справедливости H_0 и наложенном тренде, соответствующем гипотезам H_{13}, H_{14}, H_{15}, представлены, соответственно, на рис. 1.10, 1.11 и 1.12.

Так как гипотезы H_{13}, H_{14} и H_{15} относительно близки к H_0, то при малых объёмах выборок n мощность относительно этих гипотез у рассматриваемых критериев может быть очень мала.

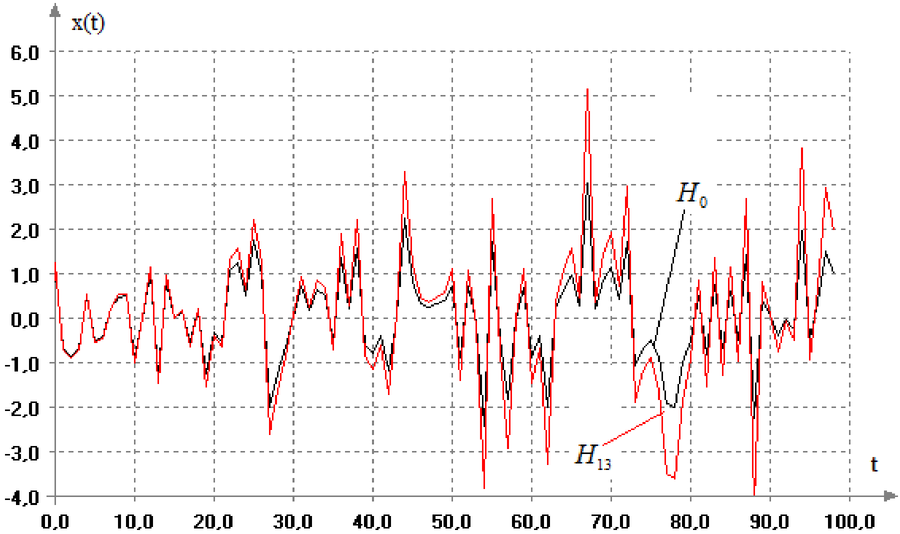


Рис. 1.10. Временной ряд при наложенном линейном тренде в дисперсии H_{13}

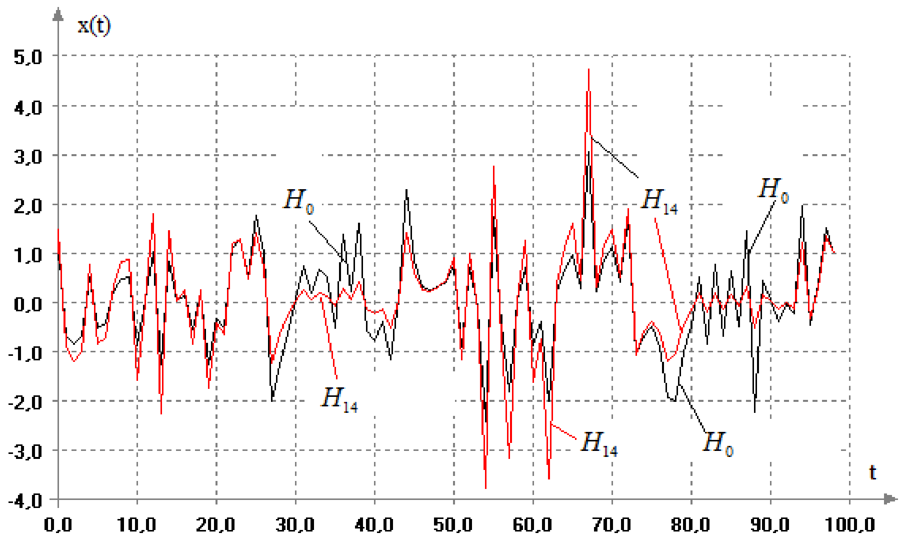


Рис. 1.11. Временной ряд при периодическом тренде в дисперсии H_{14}

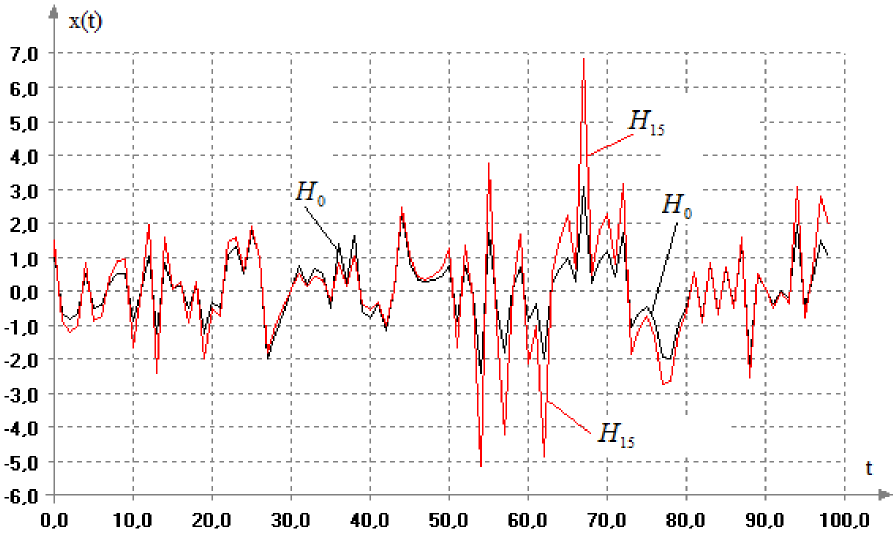


Рис. 1.12. Смешанный тренд в характеристиках рассеяния при H_{15}

В общем случае, при наличии смешанного тренда в математическом ожидании и в характеристиках рассеяния элементы временного ряда могут моделироваться в соответствии с выражением:

$$X_i = at_i + b\sin(2k_1\pi t_i) + \xi_i(1 + ct_i + d\sin(2k_2\pi t_i)) .$$

В качестве конкурирующей гипотезы с наличием смешанного тренда в математическом ожидании и в характеристиках рассеяния может рассматриваться гипотеза, задаваемая выражением:

$$H_{16}:\ X_i = at_i + b\sin(2k_1\pi t_i) + \xi_i(1 + ct_i + d\sin(2k_2\pi t_i)) ,$$

где $a = 1$, $b = 0.5$, $k_1 = 2$, $c = 1$, $d = 0.8$, $k_2 = 2$. Пример реализации показан на рис. 1.13.

В программной системе [148], в рамках которой осуществлялись исследования свойств критериев и оценка мощности относительно различных конкурирующих гипотез, реализована возможность наложения на случайные последовательности различных видов тренда, включая линейный, нелинейный и периодический тренд, накладываемые на математическое ожидание и дисперсию, а также возможность скачкообразного изменения характеристик (см. рис. 1.14).

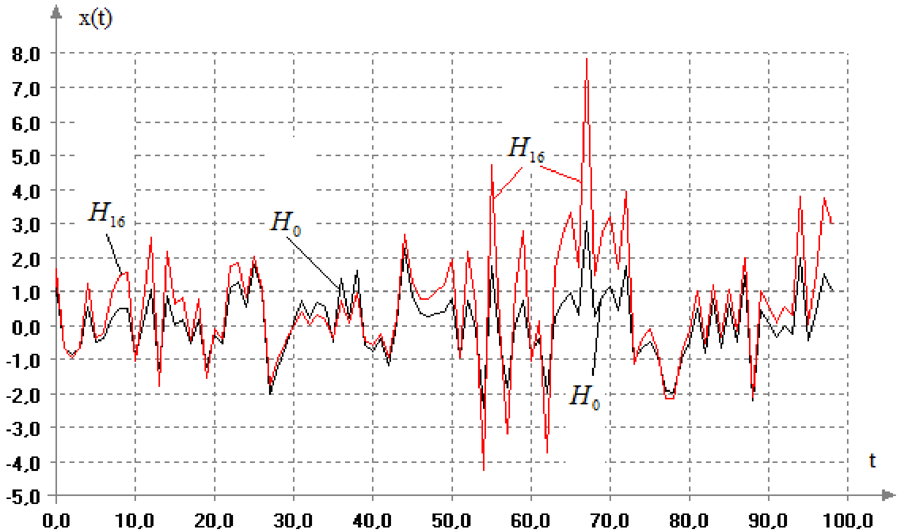


Рис. 1.13. Смешанный тренд в математическом ожидании и в характеристиках рассеяния при H_{16}

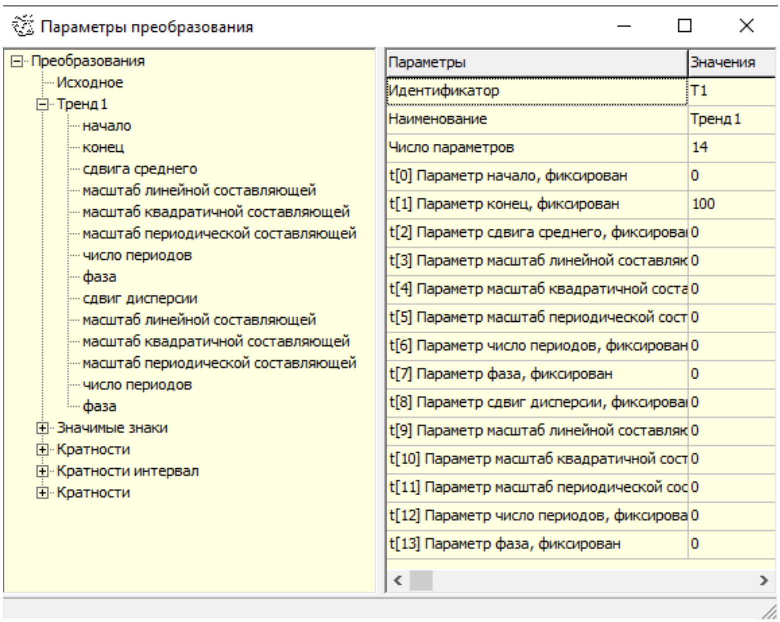


Рис. 1.14. Задание тренда при исследованиях в системе

2. КРИТЕРИИ ПРОВЕРКИ СЛУЧАЙНОСТИ И ОТСУТСТВИЯ ТРЕНДА

2.1. Критерий автокорреляции

Наличие коррелированности может говорить о существовании случайной или неслучайной связи между элементами временного ряда. Если элементы выборки $X_1, X_2, ..., X_n$ некоррелированы, то значение каждого ее элемента не должно зависеть от величины предшествующего и последующего членов. Для проверки этой независимости используется критерий со статистикой [36, 37]

$$r_{1,n} = \frac{n\sum_{i=1}^{n-1} X_i X_{i+1} - \left(\sum_{i=1}^{n} X_i\right)^2 + nX_1 X_n}{n\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2}, \tag{2.1}$$

представляющей собой оценку коэффициента корреляции первого порядка между элементами первичной выборки ($X_1, X_2, ..., X_n$) и элементами выборки, полученной из первичной сдвигом на одну единицу ($X_2, X_3, ..., X_n, X_1$).

При справедливости проверяемой гипотезы статистика $r_{1,n}$ распределена асимптотически нормально с математическим ожиданием и дисперсией

$$E[r_{1,n}] = -\frac{1}{n-1}, \quad D[r_{1,n}] = \frac{n(n-3)}{(n+1)(n-1)^2}.$$

Применяя критерий, обычно используют нормализованную статистику

$$r_{1,n}^* = \frac{r_{1,n} - E[r_{1,n}]}{\sqrt{D[r_{1,n}]}}. \tag{2.2}$$

Гипотеза о некоррелированности (об отсутствии тренда) отклоняется при больших по модулю значениях статистик (2.1) и (2.2) соответственно.

Распределение статистики (2.2) быстро сходится к асимптотическому закону. Отличием распределения статистики (2.2) от стандартного нормального закона можно практически пренебречь при $n > 30$ (см. рис. 2.1), когда реальное распределение статистики практически не отличается от асимптотического [121],

Критерий автокорреляции относится к параметрическим критериям. Это означает, что распределения статистик (2.1) и (2.2) зависят от вида закона распределения вероятностей, которому принадлежат наблюдения (измерения) случайной величины. При нарушении стандартного предположения о принадлежности наблюдений нормальному закону распределения статистик (2.1) и (2.2), соответствующие справедливости проверяемой гипотезы H_0 становятся другими.

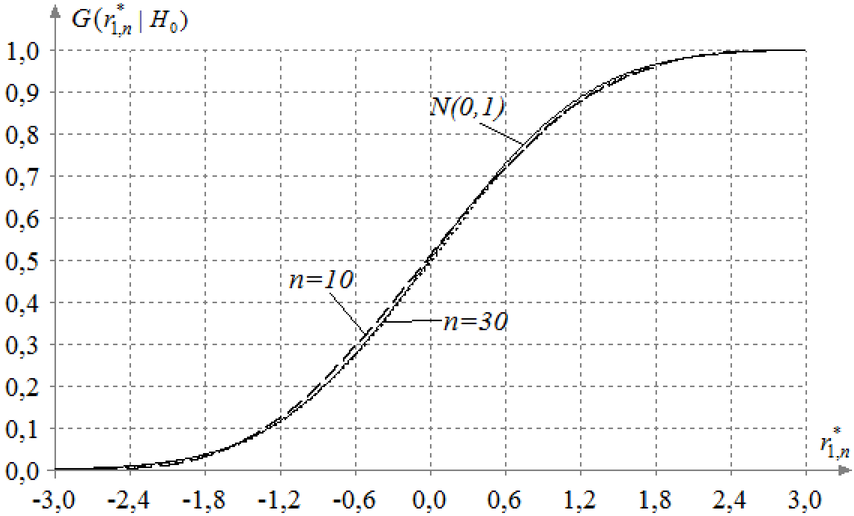


Рис. 2.1. Сходимость к стандартному нормальному закону распределения статистики (2.2) критерия автокорреляции

Распределения статистики (2.2) критерия исследовались для ситуаций принадлежности случайной величины различным законам, в том числе для случая принадлежности наблюдений обобщенному нормальному закону с плотностью

$$f(x) = \frac{\theta_2}{2\theta_1 \Gamma(1/\theta_2)} \exp\left(-\left(\frac{|x-\theta_0|}{\theta_1} \right)^{\theta_2} \right) \tag{2.3}$$

со значениями параметра формы $\theta_2 = 0{,}2$; 0,5; 1; 1,5; 2; 4; 8. При таком изменении этого параметра вид закона меняется от близкого к распределению Коши, до близкого к равномерному закону. При $\theta_2 = 2$ выражение (2.3) дает плотность нормального закона распределения.

Полученные в результате моделирования распределения статистики критерия автокорреляции в случае принадлежности случайной величины законам распределения семейства (2.3) при различных параметрах формы представлены на рис. 2.2.

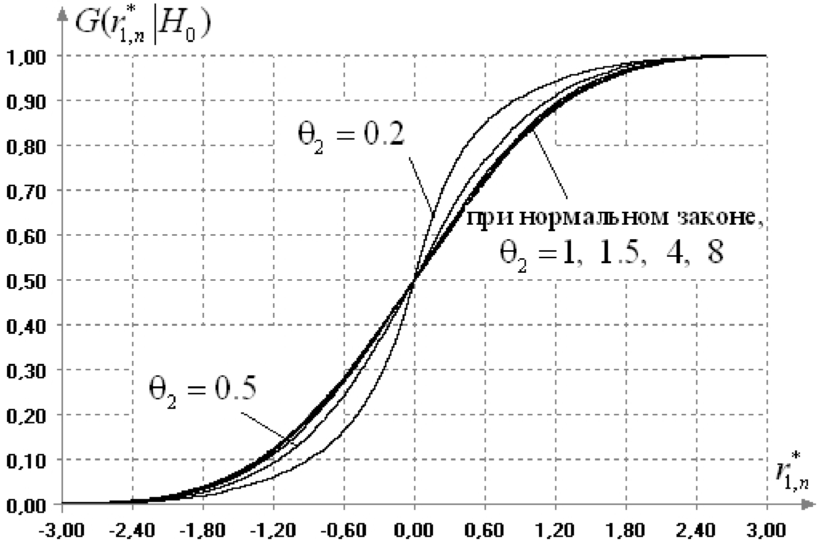


Рис. 2.2. Функции распределения статистики критерия автокорреляции в зависимости от параметра формы семейства (2.3) при $n = 25$

Как следует из представленной на рисунке картины, в случае принадлежности выборок $X_1, X_2, ..., X_n$ достаточно широкому кругу законов распределение статистики критерия автокорреляции существенно не отличается от распределения, имеющего место в случае принадлежности $X_1, X_2, ..., X_n$ нормальному закону. Если закон, которому принадлежат случайные величины, симметричен и с не слишком тяжелыми хвостами, то распределение статистики не отличается значимо от “классического”.

При сильной асимметричности закона распределения случайных величин (например, в случае показательного закона) распределение статистики становится отличным от “классического”. В то же время, асимметричность закона влияет на распределение статистики менее значимо, чем “тяжесть” хвостов. В случае принадлежности выборок $X_1, X_2, ..., X_n$ асимметричным законам экстремальных значений (минимального или максимального) распределения статистики практически не отличаются от “классического”.

Таблица 2.1

Мощность критерия автокорреляции относительно гипотез $H_1 - H_7$

n	α	H_1	H_2	H_3	H_4	H_5	H_6	H_7
10	0.1	0.103	0.126	0.295	0.141	0.102	0.113	0.120
	0.05	0.051	0.066	0.175	0.077	0.051	0.056	0.104
	0.025	0.026	0.034	0.097	0.041	0.025	0.026	0.049
	0.01	0.011	0.015	0.043	0.018	0.010	0.009	0.014
25	0.1	0.105	0.191	0.837	0.180	0.098	0.102	0.216
	0.05	0.053	0.115	0.739	0.108	0.048	0.049	0.107
	0.025	0.027	0.070	0.631	0.065	0.024	0.024	0.048
	0.01	0.011	0.035	0.483	0.033	0.009	0.008	0.014
50	0.1	0.108	0.288	0.993	0.1	0.228	0.112	0.197
	0.05	0.056	0.194	0.985	0.05	0.147	0.058	0.120
	0.025	0.029	0.128	0.969	0.025	0.093	0.030	0.072
	0.01	0.012	0.073	0.935	0.01	0.05	0.013	0.036
100	0.1	0.113	0.455	1.000	0.320	0.123	0.308	0.950
	0.05	0.058	0.342	1.000	0.221	0.066	0.210	0.912
	0.025	0.030	0.251	1.000	0.150	0.036	0.141	0.862
	0.01	0.013	0.163	1.000	0.090	0.016	0.082	0.781
200	0.1	0.119	0.696	1.000	0.479	0.139	0.474	0.999
	0.05	0.064	0.588	1.000	0.364	0.078	0.359	0.997
	0.025	0.034	0.483	1.000	0.270	0.043	0.265	0.993
	0.01	0.015	0.360	1.000	0.176	0.020	0.172	0.983
300	0.1	0.126	0.838	1.000	0.609	0.155	0.606	1.000
	0.05	0.069	0.757	1.000	0.493	0.089	0.490	1.000
	0.025	0.037	0.669	1.000	0.390	0.051	0.387	1.000
	0.01	0.016	0.550	1.000	0.276	0.024	0.274	0.999

Отметим, что влияние закона распределения наблюдаемых случайных величин на распределение статистики критерия автокорреляции такое же, как для критериев, связанных с проверкой гипотез о парной корреляции [123, 124, 125].

Аналогичную картину можно наблюдать для параметрических критериев, используемых при проверке гипотез об однородности средних, также проявляющих устойчивость к отклонениям от нормального закона распределения. Там также такое исключение имеет место для законов с «тяжелыми» хвостами и для асимметричных законов [46, 113, 126, 127].

Оценки мощности критерия автокорреляции, полученные относительно конкурирующих гипотез $H_1 - H_7$ представлены в таблице 2.1.

Конкурирующие гипотезы $H_1 - H_7$ соответствуют наличию линейного ($H_1 - H_3$) и периодического ($H_4 - H_7$) тренда в математическом ожидании. В таблице приведены оценки мощности при объёмах выборок $n = 10, 25, 50, 100, 200, 300$ для заданных вероятностей ошибок 1-го рода $\alpha = 0.1, 0.05, 0.025, 0.01$.

Оценки мощности для всех рассматриваемых в руководстве критериев получены, как правило, при числе имитационных экспериментов $N = 10^6$.

2.2. Критерий Морана

Статистика критерия Морана представляет собой одно из нормализующих преобразований статистики критерия автокорреляции (2.1) [75]:

$$r_{1,n}^M = (n-1)^{1/2} \frac{n r_{1,n} + 1}{n-2}. \tag{2.4}$$

Гипотеза об отсутствии тренда отклоняется при больших и малых значениях этой статистики.

Исследование распределений статистики (2.4) при справедливости проверяемой гипотезы в зависимости от объема анализируемых выборок в случае выполнения предположения о нормальности показало, что эти распределения достаточно хорошо согласуются со стандартным нормальным законом при $n > 50$ (см. рис. 2.3).

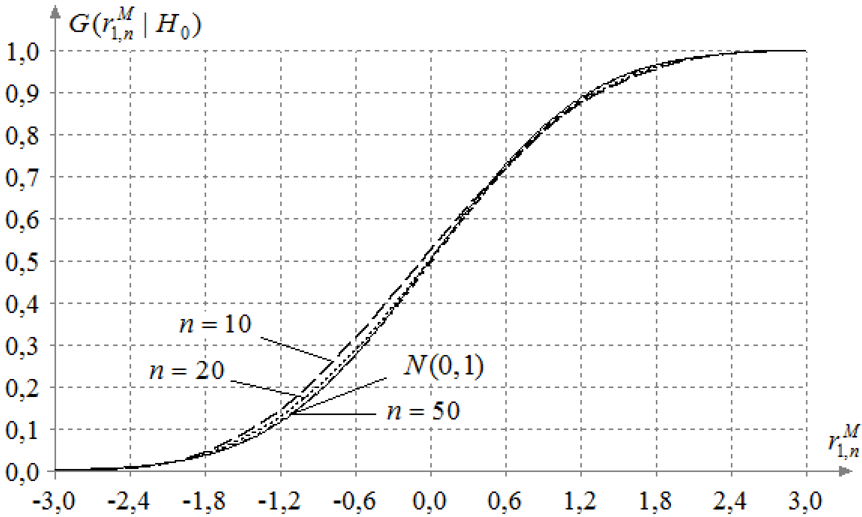


Рис. 2.3. Сходимость распределения статистики критерия Морана к стандартному нормальному закону

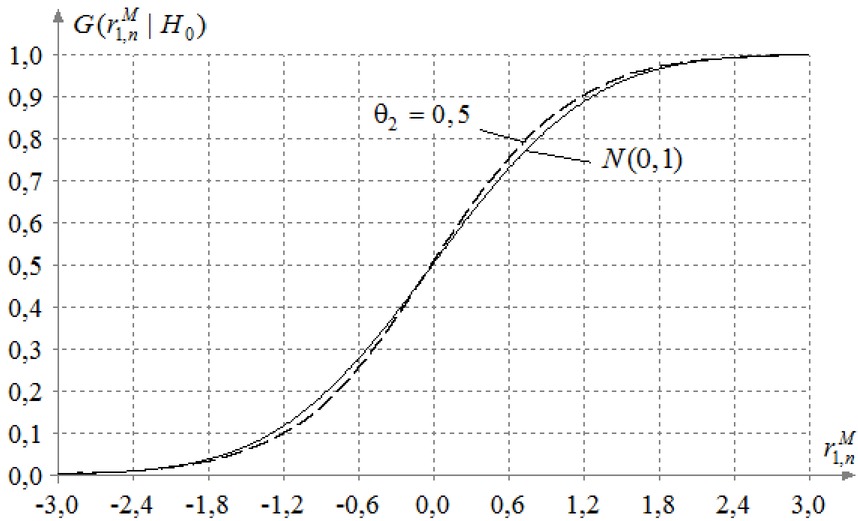


Рис. 2.4. Распределение статистики критерия Морана в случае принадлежности случайных величин семейству (2.3) с параметром формы $\theta_2 = 0.5$ при $n = 100$

Распределения статистики критерия Морана (как и исходного критерия автокорреляции) зависят от вида закона, которому принадлежит анализируемая выборка, но достаточно устойчивы к нарушению предположения о принадлежности наблюдаемой случайной величины нормальному закону. Характер зависимости распределений статистики от вида наблюдаемого закона аналогичен зависимости распределения статистики (2.2): на распределение статистики влияет асимметричность и “тяжелые хвосты” закона, соответствующего анализируемой выборке. В качестве примера на рис. 2.4 показано распределение статистики Морана в случае принадлежности наблюдаемых величин семейству (2.3) с параметром формы $\theta_2 = 0.5$.

Мощность критерия Морана эквивалентна мощности исходного критерия автокорреляции.

2.3. Критерий Люнга–Бокса

Статистика критерия Люнга–Бокса представляет собой другое нормализующее преобразование статистики (2.1) [69]:

$$r_{1,n}^{LB} = \left[\frac{n(n+2)}{n-1} \right]^{1/2} r_{1,n} . \qquad (2.5)$$

Гипотеза об отсутствии тренда отклоняется при больших и малых значениях статистики.

Исследование распределения данной статистики показало, что оно *очень медленно сходится* к стандартному нормальному закону (см. рис. 2.5), в этой связи использование последнего для определения достигаемого уровня значимости при ограниченных объёмах выборок приводит к большим ошибкам, что в свою очередь может являться причиной неверных выводов.

Критерий Люнга–Бокса, как и исходный критерий автокорреляции, относится к параметрическим критериям, поэтому распределение его статистики зависит от закона, которому принадлежат анализируемые выборки. Эта зависимость аналогична той, что демонстрируется на рис. 2.2.

Плохая сходимость распределения статистики (2.5) к стандартному нормальному закону, используемому для вычисления достигаемого уровня значимости, делает нецелесообразной рекомендацию применения критерия Люнга–Бокса в приложениях.

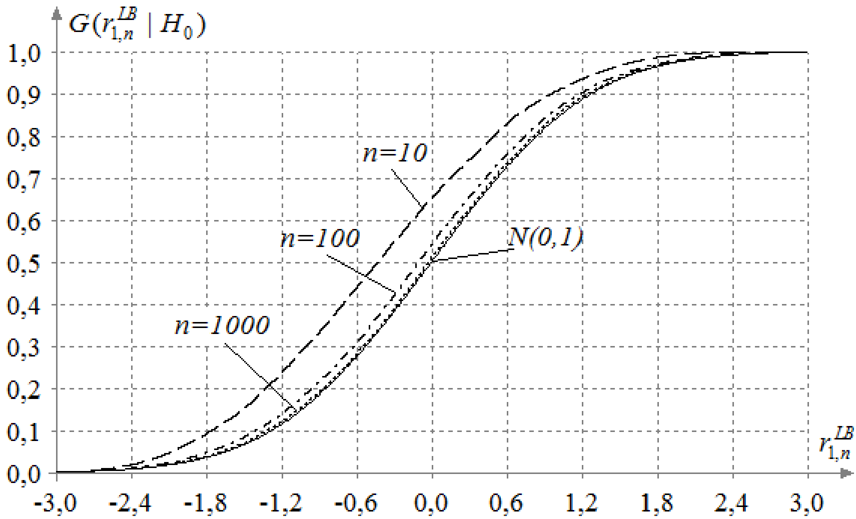


Рис. 2.5. Сходимость распределения статистики критерия Люнга–Бокса к стандартному нормальному закону

Мощность критерия Люнга–Бокса эквивалентна мощности исходного критерия автокорреляции и критерия Морана.

2.4. Критерий Дюффа–Роя

Статистика критерия Дюффа–Роя представляет собой наиболее удачное нормализующее преобразование статистики (2.1) критерия автокорреляции [16]:

$$r_{1,n}^{DR} = \left[\frac{n-1}{n(n-2)} \right]^{1/2} (n r_{1,n} + 1) . \qquad (2.6)$$

Проверяемая гипотеза отклоняется при больших по модулю значениях статистики.

При выполнении предположения о нормальности анализируемых выборок и справедливости проверяемой гипотезы о случайности и отсутствии тренда распределение статистики Дюффа–Роя быстро сходится к стандартному нормальному закону. Отличием распределения статистики от стандартного нормального закона можно пренебречь при $n > 17$. Сходимость распределения статистики (2.6) к стандартному нормальному закону демонстрирует рис. 2.6.

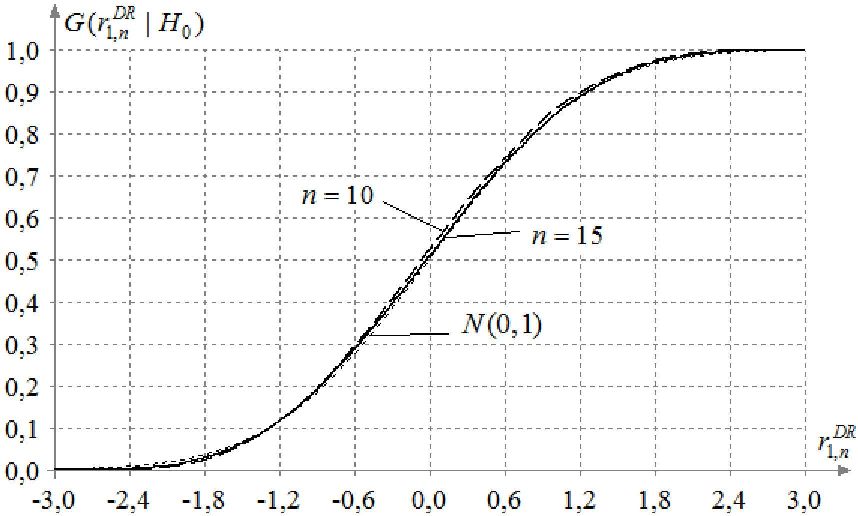


Рис. 2.6. Сходимость распределения статистики критерия Дюффа–Роя к стандартному нормальному закону

Критерии Морана, Дюффа–Роя и Люнга–Бокса являются модификациями критерия автокорреляции и обладают эквивалентными свойствами. При использовании реальных распределений статистик критериев результат проверки гипотезы по всем критериям приведет к одному и тому же достигаемому уровню значимости (p_{value}). Численные исследования подтверждают, что относительно заданных конкурирующих гипотез эти критерии обладают одинаковой мощностью.

А так как распределение статистики Дюффа–Роя быстрее распределений статистик других модификаций сходится к асимптотическому стандартному нормальному закону, то целесообразно рекомендовать применение именно этой модификации. Мощность критерия Дюффа–Роя эквивалентна мощности исходного критерия автокорреляции и мощности критериев Морана и Люнга–Бокса.

2.5. Модификация критерия автокорреляции

В работе [36] рассматривается модификация критерия автокорреляции, статистика которого представляет собой сумму оценок коэффициентов корреляции первого и второго порядков:

$$r_{1,2} = \frac{\sum_{i=1}^{n-1}(X_i - \bar{X})(X_{i+1} - \bar{X}) + \sum_{i=1}^{n-2}(X_i - \bar{X})(X_{i+2} - \bar{X})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}. \tag{2.7}$$

При справедливой проверяемой гипотезе распределение статистики (2.7) асимптотически нормально со средним $E[r_{1,2}]$ и дисперсией $D[r_{1,2}]$, где

$$E[r_{1,2}] = -\frac{2n-3}{n(n-1)},$$

$$D[r_{1,2}] = \frac{2n^4 - 13n^3 + 15n^2 + 28n - 34}{n^2(n+1)(n-1)^2}.$$

Нормализованная статистика имеет вид

$$r_{1,2}^* = \frac{r_{1,2} - E[r_{1,2}]}{\sqrt{D[r_{1,2}]}}. \tag{2.8}$$

Гипотеза об отсутствии тренда отклоняется при больших по модулю значениях статистики.

Сходимость распределения статистики (2.8) к стандартному нормальному закону $N(0,1)$ при выполнении стандартного предположения о нормальности анализируемых данных иллюстрирует рис. 2.7. Для разборчивости на рисунке не приведены распределения статистик при объемах выборок больше 100.

Исследования показали, что распределение статистики (2.8) не очень хорошо сходится к асимптотическому закону. Отличием распределения статистики (2.8) от стандартного нормального закона можно пренебречь лишь при $n > 200$, когда не отвергается гипотеза о согласии распределения статистики со стандартным нормальным законом. При меньших объемах выборок использование стандартного нормального закона в качестве распределения статистики может приводить к ощутимым ошибкам при вычислении достигнутого уровня значимости (p_{value}).

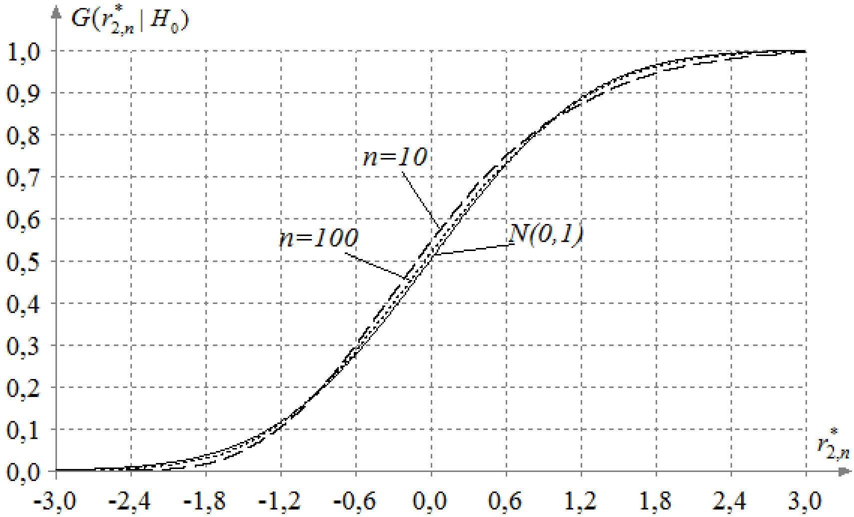


Рис. 2.7. Сходимость к стандартному нормальному закону распределения статистики (2.7)

Это параметрический критерий и распределения его статистики также зависят от вида наблюдаемого закона случайных величин. Характер этой зависимости аналогичен картине, показанной на рис. 2.2 для критерия автокорреляции. Существенные отклонения распределения статистики (2.8) от имеющих место при нормально распределенных случайных величинах проявляются в случае принадлежности наблюдений асимметричным законам или законам с “тяжёлыми хвостами”.

В таблице 2.2 представлены оценки мощности модифицированного критерия автокорреляции относительно конкурирующих гипотез, соответствующих наличию линейного $H_1 - H_3$ и периодического $H_4 - H_7$ тренда в математическом ожидании.

Таблица 2.2

Мощность модификации критерия автокорреляции относительно гипотез $H_1 - H_7$ (наличия линейного и периодического тренда)

n	α	H_1	H_2	H_3	H_4	H_5	H_6	H_7
10	0.1	0.109	0.190	0.708	0.162	0.094	0.078	0.039
	0.05	0.056	0.114	0.570	0.094	0.046	0.037	0.016
	0.025	0.029	0.068	0.435	0.054	0.023	0.018	0.006
	0.01	0.012	0.032	0.273	0.024	0.009	0.007	0.002
25	0.1	0.113	0.289	0.980	0.239	0.091	0.071	0.034
	0.05	0.059	0.194	0.958	0.156	0.044	0.032	0.012
	0.025	0.031	0.128	0.921	0.100	0.022	0.015	0.004
	0.01	0.013	0.073	0.851	0.056	0.009	0.005	0.001
50	0.1	0.117	0.433	1.000	0.321	0.113	0.220	0.792
	0.05	0.062	0.323	1.000	0.225	0.059	0.134	0.682
	0.025	0.033	0.236	0.999	0.155	0.030	0.079	0.563
	0.01	0.014	0.151	0.997	0.092	0.012	0.038	0.408
100	0.1	0.125	0.651	1.000	0.463	0.141	0.425	0.990
	0.05	0.067	0.542	1.000	0.353	0.079	0.315	0.980
	0.025	0.036	0.441	1.000	0.264	0.044	0.228	0.964
	0.01	0.016	0.326	1.000	0.176	0.021	0.146	0.931
200	0.1	0.138	0.879	1.000	0.675	0.173	0.664	1.000
	0.05	0.077	0.813	1.000	0.567	0.103	0.554	1.000
	0.025	0.043	0.739	1.000	0.466	0.061	0.453	1.000
	0.01	0.020	0.632	1.000	0.350	0.030	0.336	0.999
300	0.1	0.152	0.961	1.000	0.809	0.203	0.805	1.000
	0.05	0.087	0.932	1.000	0.724	0.126	0.719	1.000
	0.025	0.050	0.892	1.000	0.634	0.077	0.627	1.000
	0.01	0.024	0.826	1.000	0.516	0.040	0.509	1.000

Мощность модифицированного критерия автокорреляции заметно выше мощности критерия со статистиками (2.1) – (2.2), а, следовательно, выше также мощности критериев Морана, Люнга–Бокса и Дюффа–Роя.

2.6. Критерий Вальда–Вольфовица

Статистика критерия Вальда–Вольфовица, предложенная в [92], основана на коэффициенте сериальной корреляции и имеет вид

$$R_1 = \sum_{i=1}^{n-1} X_i X_{i+1} + X_n X_1 . \qquad (2.9)$$

Величина R_1 распределена асимптотически нормально с математическим ожиданием

$$E[R_1] = (S_1^2 - S_2) / (n-1)$$

и дисперсией

$$D[R_1] = \frac{S_2^2 - S_4}{n-1} + \frac{S_1^4 - 4S_1^2 S_2 + 4S_1 S_3 + S_2^2 - 2S_4}{(n-1)(n-2)} - \frac{(S_1^2 - S_2)^2}{(n-1)^2} ,$$

где $S_r = X_1^r + ... + X_n^r$.

Нормализованная статистика

$$R_1^* = \frac{R_1 - E[R_1]}{\sqrt{D[R_1]}} \qquad (2.10)$$

асимптотически подчинена стандартному нормальному закону $N(0,1)$.

Критерий Вальда–Вольфовица – двусторонний, гипотеза об отсутствии тренда отклоняется при больших по модулю значениях статистики.

Исследование распределений статистики (2.10) в зависимости от объема выборки для случая выполнения предположения о нормальности анализируемых выборок показало быструю сходимость распределения статистики R_1^* к стандартному нормальному закону (см. рис. 2.8). При объёмах выборок $n > 20$ отклонением распределения статистики (2.10) от стандартного нормального закона можно пренебречь.

Критерий Вальда–Вольфовица относится к параметрическим критериям. При нарушении основного предположения о нормальности анализируемых выборок распределение статистики (2.10) отличается от стандартного нормального закона. В качестве примера на рис. 2.9 показаны распределения статистики (2.10) критерия Вальда–Вольфовица в случае принадлежности выборок семейству (2.3) со значениями параметра формы $\theta_2 = 0.2, 0.5, 1, 2, 4, 8$ при объеме выборок $n = 25$ (при $\theta_2 = 2$ распределение семейства (2.3) соответствует нормальному закону).

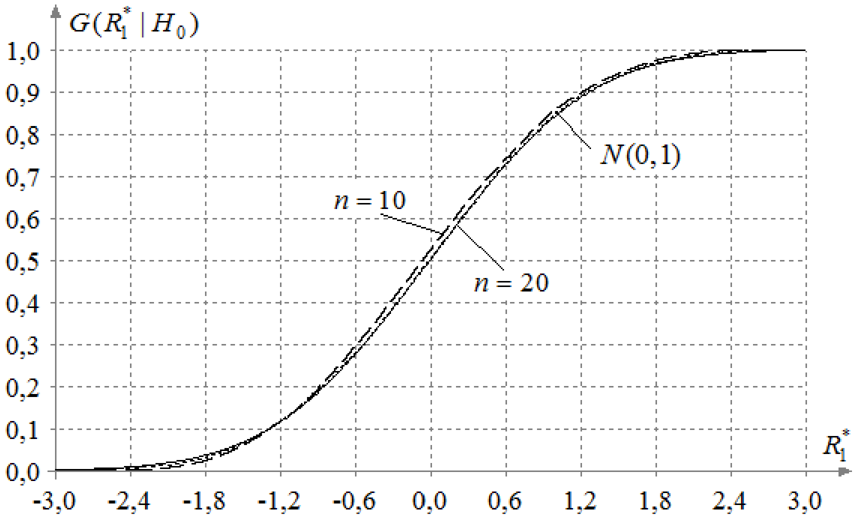


Рис. 2.8. Сходимость распределения статистики (2.10) критерия Вальда–Вольфовица к стандартному нормальному закону

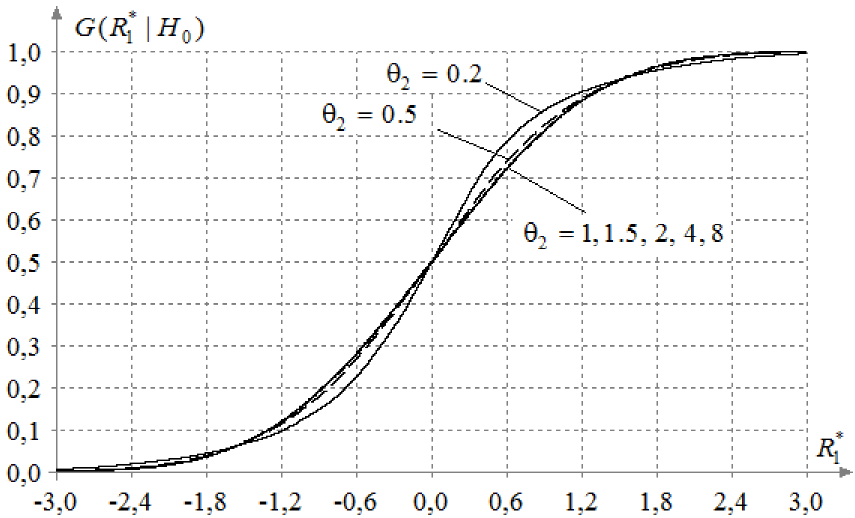


Рис. 2.9. Функции распределения статистики (2.10) критерия Вальда–Вольфовица в зависимости от параметра формы семейства (2.3) при $n = 25$

В то же время по картине, представленной на рис. 2.9 можно сделать вывод об относительной устойчивости критерия по отношению к отклонениям анализируемых данных от нормального закона. Лишь при асимметричности наблюдаемых законов и “тяжелых хвостах” при использовании данного критерия нельзя пренебрегать нарушением этого стандартного предположения.

Таблица 2.3

Мощность критерия Вальда-Вольфовица относительно гипотез $H_1 - H_7$ (наличия линейного и периодического тренда)

n	α	H_1	H_2	H_3	H_4	H_5	H_6	H_7
10	0.1	0.102	0.125	0.276	0.140	0.102	0.112	0.191
	0.05	0.051	0.065	0.156	0.076	0.050	0.054	0.094
	0.025	0.026	0.034	0.083	0.040	0.025	0.026	0.041
	0.01	0.010	0.014	0.035	0.018	0.010	0.009	0.011
25	0.1	0.105	0.190	0.835	0.180	0.098	0.102	0.213
	0.05	0.053	0.115	0.735	0.108	0.048	0.049	0.104
	0.025	0.027	0.069	0.624	0.064	0.023	0.023	0.046
	0.01	0.011	0.034	0.475	0.032	0.009	0.008	0.013
50	0.1	0.108	0.287	0.993	0.228	0.112	0.197	0.714
	0.05	0.056	0.194	0.984	0.147	0.059	0.120	0.600
	0.025	0.029	0.128	0.969	0.093	0.030	0.072	0.486
	0.01	0.012	0.073	0.935	0.051	0.013	0.036	0.353
100	0.1	0.113	0.455	1.000	0.320	0.123	0.308	0.950
	0.05	0.058	0.341	1.000	0.221	0.066	0.210	0.912
	0.025	0.030	0.250	1.000	0.150	0.036	0.141	0.861
	0.01	0.013	0.163	1.000	0.090	0.016	0.082	0.781
200	0.1	0.119	0.696	1.000	0.479	0.139	0.474	0.999
	0.05	0.064	0.587	1.000	0.364	0.078	0.359	0.997
	0.025	0.034	0.482	1.000	0.270	0.043	0.265	0.993
	0.01	0.015	0.359	1.000	0.176	0.020	0.172	0.983
300	0.1	0.126	0.838	1.000	0.609	0.155	0.606	1.000
	0.05	0.069	0.757	1.000	0.493	0.089	0.491	1.000
	0.025	0.037	0.669	1.000	0.390	0.051	0.387	1.000
	0.01	0.016	0.550	1.000	0.276	0.024	0.274	0.999

Заметим, что такая устойчивость характерна для многих параметрических критериев, ориентированных на проверку гипотез об отсутствии корреляционных связей, или на проверку гипотез о математических ожиданиях, или на проверку гипотез о равенстве (об однородности) математических ожиданий случайных величин, соответствующих различным выборкам.

Оценки мощности критерия Вальда–Вольфовица относительно конкурирующих гипотез, соответствующих наличию линейного $H_1 - H_3$ и периодического $H_4 - H_7$ тренда в математическом ожидании наблюдаемых случайных величин, приведены в таблице 2.3.

Как можно заметить, оценки мощности этого критерия не превосходят и практически совпадают с оценками мощности критерия автокорреляции.

2.7. Ранговый критерий Вальда–Вольфовица

В работе [92] была отмечена возможность построения непараметрического аналога сериального коэффициента корреляции. Пусть по выборке $X_1, X_2, ..., X_n$ построен вариационный ряд $x_{(1)} \le x_{(2)} \le ... \le x_{(n)}$ (упорядоченный по возрастанию ряд элементов исходной выборки $X_1, X_2, ..., X_n$). Тогда рангом R_i измерения X_i является его номер в вариационном ряду.

Статистика рангового критерия сериальной корреляции Вальда–Вольфовица имеет вид [92]:

$$R = \sum_{i=1}^{n-1} \left(R_i - \frac{n+1}{2} \right) \left(R_{i+1} - \frac{n+1}{2} \right). \tag{2.11}$$

Распределение статистики (2.11) асимптотически нормально с математическим ожиданием и дисперсией

$$E[R] = 0 ,$$

$$D[R] = \frac{n^2 (n+1)(n-3)(5n+6)}{720} .$$

Критерий двусторонний, гипотеза об отсутствии тренда отклоняется при больших по модулю значениях статистики

$$R^* = \frac{R}{\sqrt{D[R]}} . \tag{2.12}$$

Результаты статистического моделирования показали, что дискретностью распределения статистики практически можно пренебречь. В то же время эти результаты показали, что распределения статистики (2.12) при ограниченных объёмах выборок существенно смещены относительно стандартного нормального закона и *очень медленно сходятся* к последнему (см. рис. 2.10).

Это означает, что при ограниченных объёмах выборок использование стандартного нормального закона для вычисления достигнутого уровня значимости (p_{value}) может приводить к неверному выводу относительно результатов проверки гипотезы.

Результаты исследования распределений статистики (2.12) при различных объемах выборок n, выполненные с использованием методов статистического моделирования, позволили предложить [104] следующую модификацию статистики

$$R^{**} = R^{*} + 1.1216 n^{-0.523} . \tag{2.13}$$

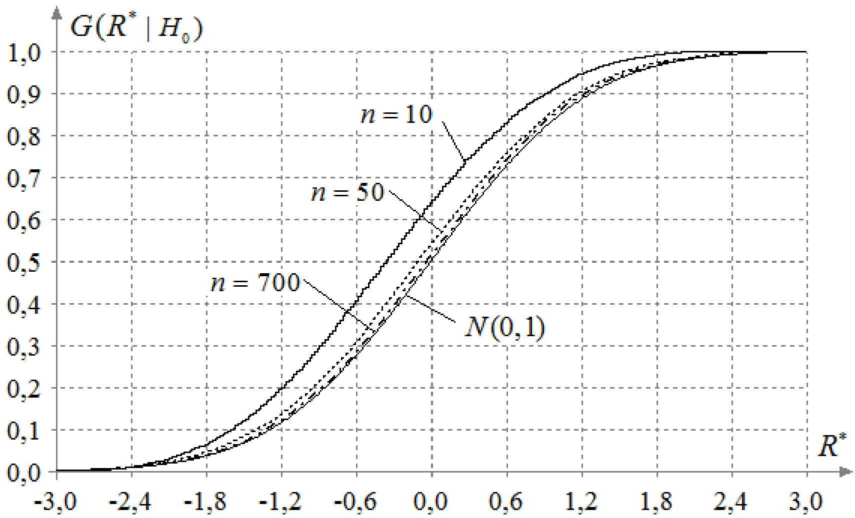


Рис. 2.10. Сходимость распределения статистики (2.12) к стандартному нормальному закону

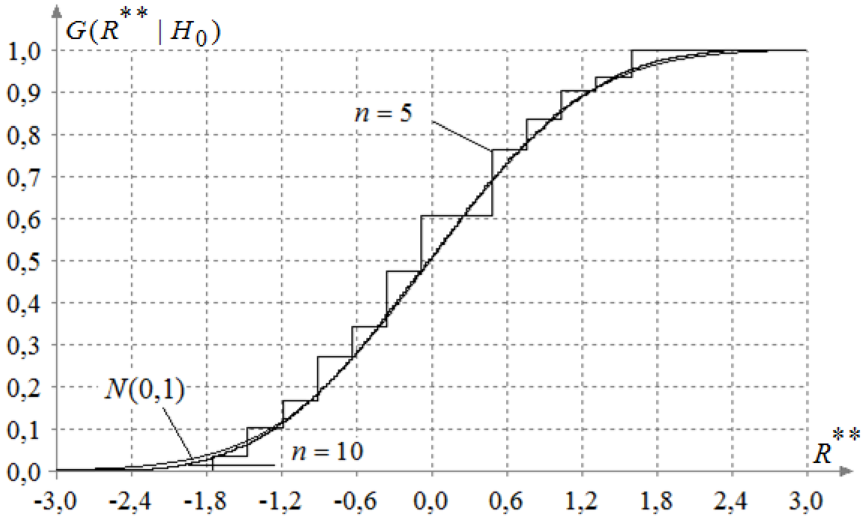


Рис. 2.11. Сходимость распределения статистики (2.13) к стандартному нормальному закону

Распределение модифицированной статистики (2.13) хорошо согласуется со стандартным нормальным законом уже при $n > 10$ (см. рис. 2.11).

Использование в ранговом критерии Вальда–Вольфовица модифицированной статистики (2.13) ликвидирует отмеченный выше (очень существенный) недостаток критерия со статистикой (2.12).

Оценки мощности рангового критерия Вальда–Вольфовица относительно конкурирующих гипотез, соответствующих наличию линейного $H_1 - H_3$ и периодического $H_4 - H_7$ тренда в математическом ожидании наблюдаемых случайных величин, приведены в таблице 2.4.

Ранговый критерий показывает мощность даже несколько выше мощности критерия Вальда–Вольфовица.

Таблица 2.4

Мощность рангового критерия Вальда–Вольфовица относительно гипотез $H_1 - H_7$ (наличия линейного и периодического тренда)

n	α	H_1	H_2	H_3	H_4	H_5	H_6	H_7
10	0.1	0.104	0.155	0.545	0.146	0.104	0.119	0.217
	0.05	0.052	0.088	0.399	0.083	0.051	0.059	0.118
	0.025	0.028	0.051	0.276	0.048	0.026	0.030	0.062
	0.01	0.011	0.022	0.143	0.022	0.010	0.010	0.018
25	0.1	0.107	0.216	0.921	0.182	0.099	0.105	0.231
	0.05	0.054	0.135	0.861	0.109	0.048	0.049	0.118
	0.025	0.028	0.084	0.787	0.066	0.024	0.023	0.053
	0.01	0.012	0.045	0.672	0.034	0.009	0.008	0.016
50	0.1	0.110	0.305	0.997	0.227	0.112	0.197	0.720
	0.05	0.056	0.210	0.993	0.146	0.058	0.120	0.610
	0.025	0.029	0.142	0.986	0.093	0.030	0.072	0.501
	0.01	0.012	0.083	0.968	0.051	0.013	0.037	0.372
100	0.1	0.112	0.461	1.000	0.314	0.122	0.302	0.950
	0.05	0.058	0.349	1.000	0.217	0.066	0.206	0.913
	0.025	0.030	0.258	1.000	0.147	0.035	0.138	0.865
	0.01	0.013	0.170	1.000	0.088	0.016	0.081	0.788
200	0.1	0.119	0.691	1.000	0.468	0.137	0.463	0.999
	0.05	0.063	0.584	1.000	0.354	0.076	0.350	0.997
	0.025	0.034	0.481	1.000	0.262	0.042	0.258	0.993
	0.01	0.014	0.359	1.000	0.170	0.019	0.167	0.984
300	0.1	0.125	0.830	1.000	0.593	0.151	0.590	1.000
	0.05	0.067	0.748	1.000	0.477	0.086	0.475	1.000
	0.025	0.037	0.659	1.000	0.375	0.049	0.373	1.000
	0.01	0.016	0.542	1.000	0.267	0.0235	0.264	0.999

2.8. Ранговый критерий Дюффа–Роя

В [16] предложен непараметрический аналог критерия автокорреляции, статистика которого имеет вид:

$$r_1 = \sum_{i=1}^{n-1}\left(R_i - \frac{n+1}{2}\right)\left(R_{i+1} - \frac{n+1}{2}\right) \Bigg/ \sum_{i=1}^{n}\left(R_i - \frac{n+1}{2}\right)^2 . \qquad (2.14)$$

Распределение статистики (2.14) асимптотически нормально с математическим ожиданием и дисперсией

$$E[r_1] = -\frac{1}{n},$$

$$D[r_1] = \frac{5n^3 - 19n^2 + 10n + 16}{5n^2(n-1)^2}.$$

Гипотеза об отсутствии тренда отклоняется при больших по модулю значениях статистики

$$r_1^* = \frac{r_1 - E[r_1]}{\sqrt{D[r_1]}}. \qquad (2.15)$$

Результаты моделирования показали, что при справедливости проверяемой гипотезы об отсутствии тренда распределение статистики рангового критерия Дюффа–Роя быстро сходится к стандартному нормальному закону. Отличие распределения статистики (2.15) от стандартного нормального закона можно не учитывать при $n > 17$. Дискретностью же распределения статистики можно практически пренебречь при $n > 10$.

Данный критерий по мощности эквивалентен ранговому критерию Вальда–Вольфовица: они обладают одинаковой мощностью и результаты проверки гипотезы приводят к одинаковым значениям p_{value}.

2.9. Критерий Бартелса

Пусть R_i – ранг наблюдения X_i в последовательности n измерений $X_1, X_2, ..., X_n$.

Статистика рангового критерия, предложенного Бартелсом [3], имеет вид

$$B = \sum_{i=1}^{n}\left(R_i - R_{i+1}\right)^2 \Bigg/ \sum_{i=1}^{n}\left(R_i - \frac{n+1}{2}\right)^2. \qquad (2.16)$$

Гипотеза об отсутствии тренда отклоняется при больших по модулю значениях нормализованной статистики

$$B^* = \frac{B - E[B]}{\sqrt{D[B]}} = \frac{B-2}{2\sqrt{5/(5n+7)}}, \qquad (2.17)$$

которая в условиях отсутствия тренда приближенно подчиняется стандартному нормальному закону.

Исследования показали, что распределение статистики (2.17) достаточно быстро сходится к стандартному нормальному закону (см. рис. 2.12) и при $n > 10$ практически не отличается от стандартного нормального закона.

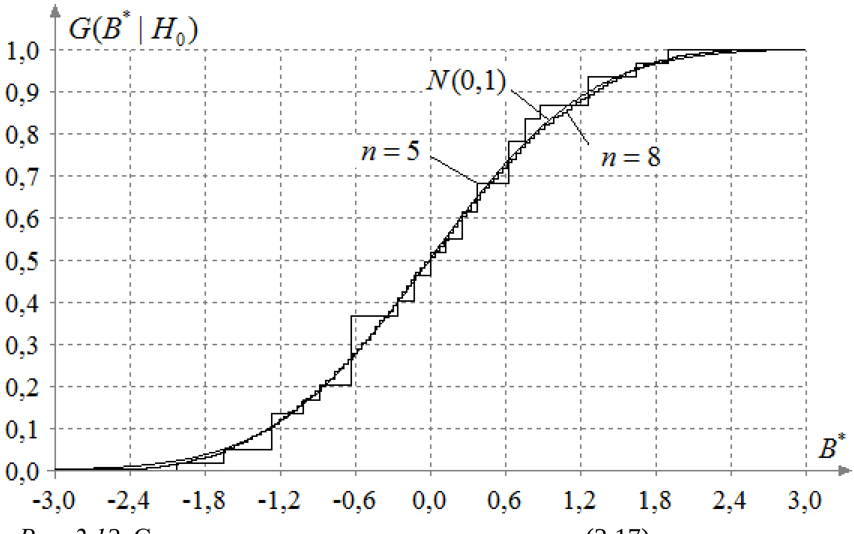


Рис. 2.12. Сходимость распределения статистики (2.17) к стандартному нормальному закону

Оценки мощности критерия Бартелса относительно конкурирующих гипотез, соответствующих наличию линейного $H_1 - H_3$ и периодического $H_4 - H_7$ тренда в математическом ожидании наблюдаемых случайных величин, представлены в таблице 2.5.

Т а б л и ц а 2 . 5

Мощность критерия Бартелса относительно гипотез $H_1 - H_7$ (наличия линейного и периодического тренда)

n	α	H_1	H_2	H_3	H_4	H_5	H_6	H_7
10	0.1	0.105	0.163	0.624	0.138	0.105	0.122	0.231
	0.05	0.055	0.095	0.492	0.075	0.052	0.062	0.131
	0.025	0.028	0.056	0.381	0.042	0.027	0.031	0.070
	0.01	0.013	0.027	0.250	0.019	0.011	0.012	0.026
25	0.1	0.107	0.223	0.937	0.176	0.098	0.103	0.220
	0.05	0.055	0.143	0.890	0.106	0.048	0.049	0.112
	0.025	0.028	0.090	0.826	0.063	0.024	0.023	0.051
	0.01	0.012	0.049	0.729	0.032	0.009	0.008	0.015
50	0.1	0.110	0.313	0.998	0.224	0.112	0.195	0.711
	0.05	0.056	0.216	0.995	0.143	0.058	0.118	0.599
	0.025	0.029	0.147	0.989	0.090	0.030	0.070	0.488
	0.01	0.012	0.088	0.975	0.049	0.013	0.036	0.360
100	0.1	0.113	0.468	1.000	0.311	0.122	0.300	0.948
	0.05	0.059	0.355	1.000	0.214	0.065	0.203	0.910
	0.025	0.031	0.264	1.000	0.145	0.035	0.136	0.861
	0.01	0.013	0.175	1.000	0.086	0.015	0.079	0.782
200	0.1	0.119	0.696	1.000	0.465	0.137	0.461	0.999
	0.05	0.063	0.589	1.000	0.352	0.076	0.347	0.996
	0.025	0.034	0.487	1.000	0.261	0.042	0.256	0.993
	0.01	0.014	0.364	1.000	0.169	0.019	0.166	0.983
300	0.1	0.125	0.832	1.000	0.591	0.151	0.589	1.000
	0.05	0.067	0.750	1.000	0.475	0.086	0.472	1.000
	0.025	0.037	0.662	1.000	0.374	0.049	0.371	1.000
	0.01	0.016	0.545	1.000	0.265	0.023	0.262	0.999

Так как распределения статистики критерия Бартелса обладают существенной дискретностью, представленные в таблицах оценки мощности получены при аппроксимации дискретных распределений непрерывными. Как правило, эмпирические распределения статистик, полученные в результате моделировании в условиях справедливости проверяемой или конкурирующих гипотез, наилучшим образом приближались нормальным законом. Таким образом, в таблице представлены оценки “асимптотической мощности”. Это позволяет сравнивать оценки мощности критерия Бартелса с мощностью других критериев.

Подобный приём использован при вычислении оценок мощности ряда других критериев, распределения статистик которых обладают ярко выраженной дискретностью (Фостера–Стюарта, Кокса–Стюарта, сериальных критериев).

2.10. Критерий кумулятивной суммы

Данный непараметрический критерий проверки отсутствия тренда основан на сумме [71, 96]

$$V = \sum_{i=0}^{n} \delta(X_i - \tilde{X}), \tag{2.18}$$

где $\tilde{X}$ - выборочная медиана и $\delta(z) = 1$, если $z \geq 0$, иначе $\delta(z) = -1$. Оценка $\tilde{X}$ находится по вариационному ряду $x_{(1)} \leq x_{(2)} \leq ... \leq x_{(n)}$, соответствующему исходной выборке $X_1, X_2, ..., X_n$: при n четном $\tilde{X} = (x_{(n/2)} + x_{(n/2+1)})/2$; при n нечетном $\tilde{X} = x_{((n-1)/2+1)}$.

Статистикой критерия CS является число переходов через ноль суммы V (2.18). Критерий двусторонний. При заданном уровне значимости α гипотеза о случайности ряда не отвергается, если $CS_{\alpha/2} \leq CS \leq CS_{1-\alpha/2}$, где $CS_{\alpha/2}$ и $CS_{1-\alpha/2}$ – верхние и нижние критические значения. Критические значения статистики для заданных уровней значимости α=0.1, 0.05, 0.01, полученные в результате статистического моделирования и расширяющие таблицу, приведенную в [71], представлены в таблице 2.6. Решение отклонить гипотезу принимается при большом (большие значения статистики могут являться результатом наличия какой-то детерминированной зависимости) или малом (это может быть связано с наличием тренда во входных данных) числе переходов суммы V через ноль.

Как показано на рис. 2.13, распределения статистики CS числа переходов кумулятивной суммы (2.28) через ноль отличаются большой дискретностью и сильно зависят от объема выборки.

Таблица 2.6

Критические значения статистики критерия кумулятивной суммы

n	$1-\alpha$					
	0.9		0.95		0.99	
	$CS_{\alpha/2}$	$CS_{1-\alpha/2}$	$CS_{\alpha/2}$	$CS_{1-\alpha/2}$	$CS_{\alpha/2}$	$CS_{1-\alpha/2}$
3	0	1	0	1	0	1
4	1	2	1	2	1	2
5	0	2	0	2	0	2
6	1	3	1	3	1	3
7	0	3	0	3	0	3
8	1	4	1	4	1	4
9	0	4	0	4	0	4
10	1	5	1	5	1	5
11	0	5	0	5	0	5
12	1	6	1	6	1	6
13	0	5	0	6	0	6
14	1	6	1	7	1	7
15	0	6	0	6	0	7
16	1	7	1	7	1	8
17	0	7	0	7	0	8
18	1	8	1	8	1	9
19	0	7	0	8	0	9
20	1	8	1	9	1	10
25	1	9	0	9	0	11
30	2	11	1	11	1	13
40	2	13	1	14	1	16
50	2	14	2	16	1	18
100	3	22	2	24	1	27
150	4	27	3	30	1	35
200	5	32	3	35	2	41
300	6	39	4	43	2	51

С этим связаны определённые проблемы применения критерия, так как для формирования корректного вывода о результатах проверки необходимо знание распределения и возможность вычисления достигнутого уровня значимости по этому распределению. При наличии только таблицы процентных точек такая возможность отсутствует.

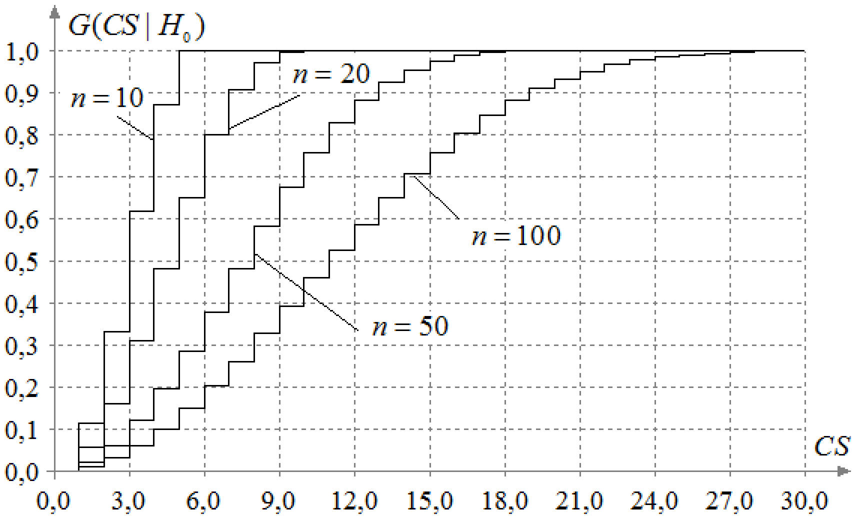


Рис. 2.13. Функции распределения статистики переходов кумулятивной суммы (2.18) через ноль при различных объемах выборок

Оценки мощности критерия кумулятивной суммы относительно конкурирующих гипотез, соответствующих наличию линейного $H_1 - H_3$ и периодического $H_4 - H_7$ тренда в математическом ожидании наблюдаемых случайных величин, представлены в таблице 2.7.

Как можно видеть, относительно периодического тренда, соответствующего конкурирующим гипотезам $H_4 - H_7$, возможности критерия, мягко говоря, сильно ограничены. Относительно линейного тренда (среди рассматриваемого множества критериев) критерий кумулятивной суммы занимает среднюю позицию.

Таблица 2.7

Мощность критерия кумулятивной суммы относительно гипотез $H_1 - H_7$ (наличия линейного и периодического тренда)

n	α	H_1	H_2	H_3	H_4	H_5	H_6	H_7
10	0.1	0.120	0.075	0.004	0.093	0.135	0.156	0.239
	0.05	0.120	0.075	0.004	0.093	0.135	0.156	0.239
	0.025	0.120	0.075	0.004	0.093	0.135	0.156	0.239
	0.01	0.120	0.075	0.004	0.093	0.135	0.156	0.239
25	0.1	0.069	0.025	0.000	0.128	0.107	0.115	0.119
	0.05	0.057	0.024	0.000	0.021	0.052	0.046	0.030
	0.025	0.021	0.009	0.000	0.008	0.019	0.015	0.007
	0.01	0.006	0.002	0.000	0.002	0.005	0.004	0.001
50	0.1	0.103	0.270	0.891	0.055	0.095	0.087	0.065
	0.05	0.066	0.267	0.891	0.044	0.047	0.045	0.048
	0.025	0.011	0.001	0.000	0.002	0.012	0.008	0.001
	0.01	0.005	0.000	0.000	0.001	0.006	0.003	0.000
100	0.1	0.124	0.535	1.000	0.081	0.078	0.072	0.081
	0.05	0.051	0.281	0.992	0.024	0.032	0.026	0.029
	0.025	0.047	0.281	0.902	0.024	0.025	0.022	0.029
	0.01	0.004	0.000	0.000	0.000	0.005	0.002	0.000
200	0.1	0.241	0.846	1.000	0.178	0.099	0.102	0.148
	0.05	0.098	0.554	1.000	0.048	0.036	0.034	0.053
	0.025	0.039	0.291	1.000	0.014	0.013	0.011	0.017
	0.01	0.037	0.291	1.000	0.014	0.009	0.010	0.017
300	0.1	0.308	0.918	1.000	0.205	0.102	0.111	0.169
	0.05	0.163	0.742	1.000	0.078	0.044	0.048	0.078
	0.025	0.094	0.559	1.000	0.036	0.020	0.024	0.040
	0.01	0.037	0.293	1.000	0.011	0.007	0.007	0.012

2.11. Знаково-ранговый критерий Холлина

В работах [27, 28, 29, 30, 31] рассмотрены проблемы использования рангов в критериях проверки случайности. В [31] предложено некоторое обобщение рангового критерия Вальда–Вольфовица и знакового критерия кумулятивной суммы. Предложенный Холлиным знаково-ранговый критерий основан на статистике

$$r = \frac{1}{k(n-1)} \sum_{i=2}^{n} \delta\left[(X_i - \tilde{X})(X_{i-1} - \tilde{X})\right] R_i R_{i-1}, \qquad (2.19)$$

где k – коэффициент, зависящий от объема выборки (рекомендуемые автором значения k приведены в таблице 2.8); $\tilde{X}$ – выборочная медиана, определяемая по вариационному ряду $x_{(1)} \le x_{(2)} \le \ldots \le x_{(n)}$, построенному по выборке $X_1, X_2, ..., X_n$ (при n четном $\tilde{X} = (x_{(n/2)} + x_{(n/2+1)})/2$, при n нечетном $\tilde{X} = x_{((n-1)/2+1)}$); R_i – ранг величины $z_i = |X_i - \tilde{X}|$ в упорядоченном по возрастанию ряду значений $z_1 \le z_2 \le \ldots \le z_n$;

$$\delta(y) = \begin{cases} 1, & y > 0; \\ -1, & y < 0; \\ 0, & y = 0. \end{cases}$$

Таблица 2.8

Значения k для знаково–рангового критерия Холлина

n	5	10	20	50	100	200	400
k	10.11	36.95	140.62	851.62	3370	13407	53480

Ряд значений x_i признается случайным, если $|r| < r_\alpha$. Критические значения r_α можно найти, например, в [109].

Распределения статистики критерия (2.19) в зависимости от объема выборки показаны на рис. 2.14. Критерий является двусторонним. В то же время, как можно видеть, распределения статистики критерия не являются симметричными относительно 0.

Поэтому при использовании критерия необходимо иметь в виду, что вследствие ***асимметричности*** распределения статистики использование представленных в [109] процентных точек *может приводить к некорректным статистическим выводам* (к ошибочным отклонениям справедливой гипотезы о случайности).

Критерий Холлина не относится к параметрическим критериям. Однако проведенное исследование распределений его статистики (при различных параметрических моделях законов распределения наблюдаемых случайных величин и справедливости проверяемой гипотезы H_0 об отсутствии тренда) позволяет сделать следующие замечания, которые целесообразно учитывать при использовании критерия.

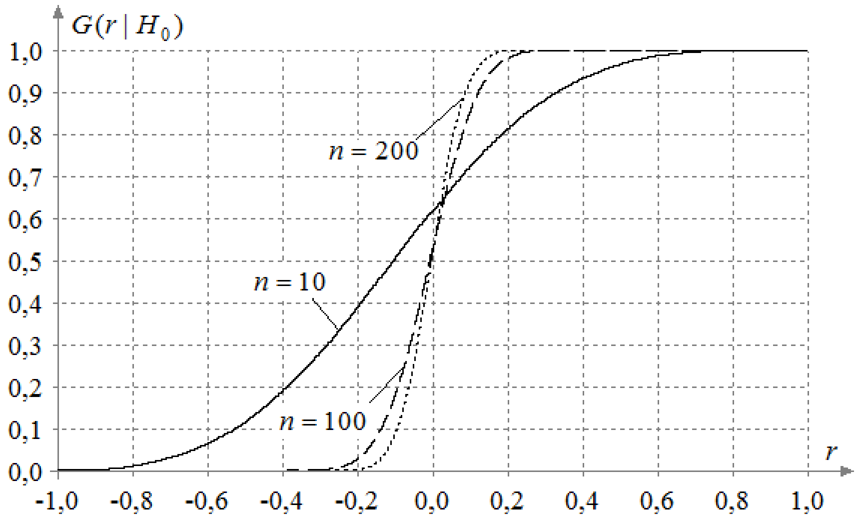


Рис. 2.14. Функции распределения статистики (2.19) критерия Холлина при различных объемах выборок

В случае принадлежности выборок $X_1, X_2, ..., X_n$ симметричным законам распределения вероятностей существенные отклонения наблюдаемого закона, например, от нормального практически не влияют на распределения статистики критерия.

В то же время асимметричность закона, которому соответствует наблюдаемая выборка $X_1, X_2, ..., X_n$, сильно сказывается на распределении статистики критерия (свойство непараметричности теряется!).

Сказанное иллюстрирует рис. 2.15, на котором приведены распределения статистики (2.19) в случае принадлежности выборок объёмом $n = 100$ семейству (2.3) с параметром формы $\theta_2 = 2$ и $\theta_2 = 0.5$, равномерному закону и экспоненциальному закону.

Исследование распределений статистики (2.19) показало, что оно хорошо описывается нормальным законом с параметром масштаба $\sigma = \sqrt{n}$ и отрицательным значением параметра сдвига, стремящимся к 0 с ростом n.

Это позволило преобразовать статистику (2.19) к виду

$$r^* = \frac{2.9992}{\sqrt{n}(n+1)(n-1)} \sum_{i=2}^{n} \delta\left[(X_i - \tilde{X})(X_{i-1} - \tilde{X}) \right] R_i R_{i-1} + \vartheta(n), \quad (2.20)$$

где $\vartheta(n) = n^{-0.5853}$ – поправка, компенсирующая смещение распределения статистики от стандартного нормального закона.

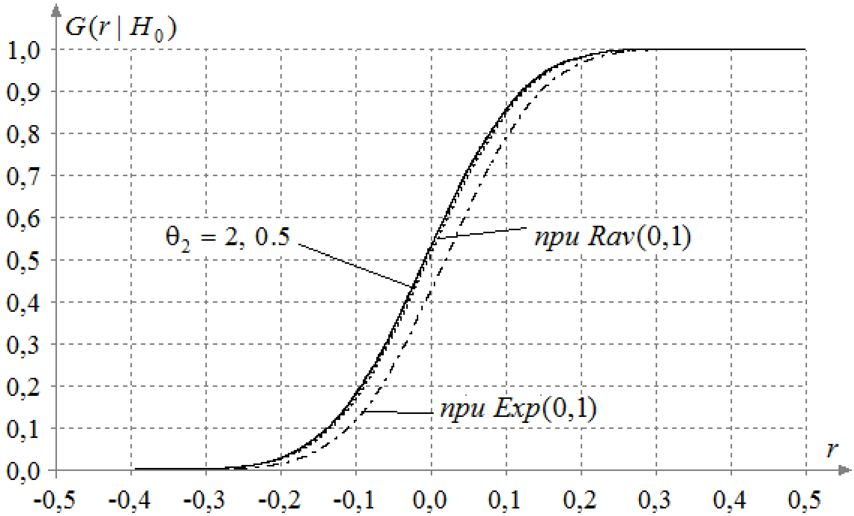


Рис. 2.15. Функции распределения статистики (2.19) критерия Холлина при $n = 100$ в зависимости от вида наблюдаемого закона

Мощность критерия Холлина со статистикой (2.20) совпадает с мощностью критерия с исходной статистикой (2.19). А преимущество в том, что для вычисления оценок достигнутого уровня значимости **p**-value можно использовать функцию распределения стандартного нормального закона.

Оценки мощности критерия Холлина относительно конкурирующих гипотез, соответствующих наличию линейного $H_1 - H_3$ и периодического $H_4 - H_7$ тренда в математическом ожидании наблюдаемых случайных величин, представлены в таблице 2.9.

Таблица 2.9

Мощность критерия Холлина относительно гипотез $H_1 - H_7$ (наличия линейного и периодического тренда)

n	α	H_1	H_2	H_3	H_4	H_5	H_6	H_7
10	0.1	0.105	0.157	0.540	0.151	0.105	0.120	0.221
	0.05	0.053	0.089	0.414	0.087	0.052	0.059	0.120
	0.025	0.027	0.050	0.289	0.049	0.025	0.028	0.058
	0.01	0.011	0.023	0.148	0.023	0.010	0.010	0.019
25	0.1	0.111	0.216	0.887	0.923	0.099	0.105	0.234
	0.05	0.059	0.136	0.806	0.866	0.048	0.050	0.120
	0.025	0.029	0.085	0.714	0.794	0.024	0.023	0.055
	0.01	0.011	0.045	0.593	0.681	0.009	0.008	0.016
50	0.1	0.113	0.306	0.997	0.997	0.112	0.198	0.722
	0.05	0.057	0.210	0.993	0.993	0.058	0.120	0.613
	0.025	0.028	0.142	0.984	0.986	0.030	0.073	0.505
	0.01	0.011	0.084	0.972	0.969	0.013	0.037	0.376
100	0.1	0.116	0.462	1.000	0.313	0.122	0.302	0.950
	0.05	0.058	0.350	1.000	0.216	0.066	0.206	0.914
	0.025	0.030	0.258	1.000	0.147	0.035	0.138	0.866
	0.01	0.012	0.170	1.000	0.088	0.016	0.081	0.790
200	0.1	0.119	0.692	1.000	1.000	0.137	0.463	0.999
	0.05	0.063	0.585	1.000	1.000	0.076	0.350	0.997
	0.025	0.034	0.483	1.000	1.000	0.042	0.259	0.993
	0.01	0.014	0.360	1.000	1.000	0.019	0.167	0.984
300	0.1	0.125	0.830	1.000	0.593	0.151	0.591	1.000
	0.05	0.067	0.747	1.000	0.477	0.086	0.474	1.000
	0.025	0.037	0.659	1.000	0.375	0.049	0.373	1.000
	0.01	0.016	0.542	1.000	0.267	0.024	0.264	0.999

2.12. Критерии Фостера–Стюарта

В зависимости от вида используемой статистики этот непараметрический критерий может применяться для проверки гипотез об отсутствии тренда в средних значениях или в дисперсиях (в характеристиках рассеяния). Соответствующие статистики критерия имеют вид [21]:

$$d = \sum_{i=2}^{n} d_i \ , \tag{2.21}$$

$$S = \sum_{i=2}^{n} S_i \ , \tag{2.22}$$

где

$d_i = u_i - l_i$, $S_i = u_i + l_i$;

$u_i = 1$, если $X_i > X_{i-1}, X_{i-2}, ..., X_1$, иначе $u_i = 0$;

$l_i = 1$, если $X_i < X_{i-1}, X_{i-2}, ..., X_1$, иначе $l_i = 0$.

Очевидно, что $0 \le S \le n-1$ и $-(n-1) \le d \le n-1$.

Критерий со статистикой d используется для обнаружения тренда в средних значениях (в математическом ожидании), а критерий со статистикой S – для проверки наличия тренда в дисперсиях.

В том и другом случае проверяемая гипотеза H_0 заключается в отсутствии тренда. При отсутствии тренда нормализованные статистики:

$$t = \frac{d}{\hat{\sigma}_d} \ , \tag{2.23}$$

$$\tilde{t} = \frac{S - \mu}{\hat{\sigma}_S} \ , \tag{2.24}$$

где

$$\mu = 2\sum_{i=2}^{n} \frac{1}{i}, \quad \hat{\sigma}_d = \sqrt{\mu} \approx \sqrt{2\ln n - 0{,}8456} \ ,$$

$$\hat{\sigma}_S = \sqrt{\mu - 4\sum_{i=2}^{n} \frac{1}{i^2}} \approx \sqrt{2\ln n - 3{,}4253} \ ,$$

приближенно описываются распределением Стьюдента с $\nu = n$ степенями свободы. Проверяемая гипотеза H_0 об отсутствии соответствующего тренда отклоняется при больших по модулю значениях статистик (2.23), (2.24).

На самом деле областью определения статистик t и $\tilde{t}$ является область дискретных значений. Исследование распределений статистик показало [121], что даже при достаточно больших объемах выборок порядка $n = 100, 200$ дискретные распределения статистик критерия существенно отличаются от распределения Стьюдента с n степенями свободы. На рис. 2.16 показаны функции распределения статистики (2.23) t, а на рис. 2.17 – распределения статистики (2.24) $\tilde{t}$, которые отличается еще большей дискретностью.

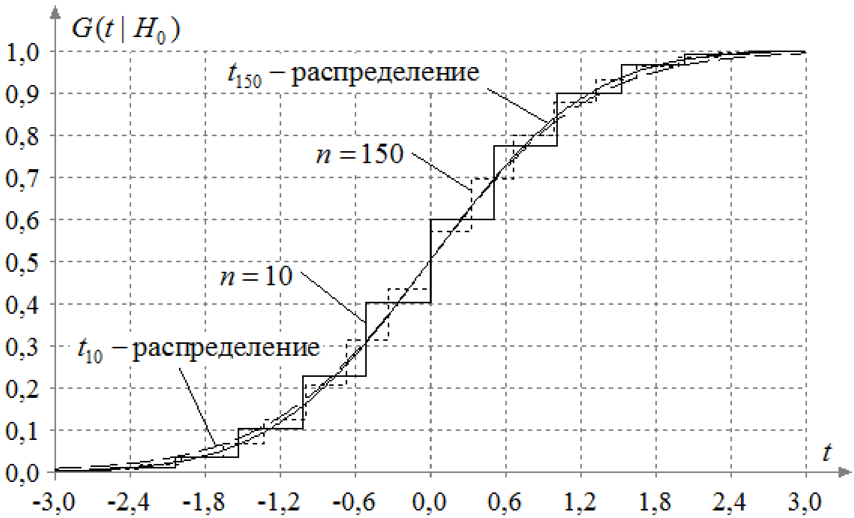


Рис. 2.16. Функции распределения статистики (2.23) критерия Фостера-Стюарта в зависимости от объёмов выборок

Следовательно, использование распределений Стьюдента для вычисления достигнутого уровня значимости вместо действительных (дискретных) распределений этих статистик может приводить к ошибкам.

Оценки мощности критерия Фостера–Стюарта со статистикой (2.23) относительно конкурирующих гипотез, соответствующих наличию линейного $H_1 - H_3$ и периодического $H_4 - H_7$ тренда в математическом ожидании наблюдаемых случайных величин, представлены в таблице 2.10.

Оценки мощности критерия Фостера–Стюарта со статистикой

(2.24) относительно близких конкурирующих гипотез $H_8 - H_{15}$, соответствующих наличию тренда в дисперсии наблюдаемых случайных величин, представлены в таблицах 2.11–2.12.

Таблица 2.10

Мощность критерия Фостера–Стюарта относительно гипотез $H_1 - H_7$ наличия линейного и периодического тренда в математическом ожидании

n	α	H_1	H_2	H_3	H_4	H_5	H_6	H_7
10	0.1	0.100	0.317	0.491	0.118	0.038	0.030	0.013
	0.05	0.050	0.152	0.301	0.033	0.038	0.030	0.013
	0.025	0.025	0.152	0.176	0.033	0.008	0.006	0.002
	0.01	0.010	0.054	0.063	0.007	0.086	0.082	0.089
25	0.1	0.104	0.290	0.679	0.089	0.031	0.029	0.031
	0.05	0.050	0.162	0.530	0.032	0.010	0.008	0.009
	0.025	0.025	0.162	0.418	0.032	0.126	0.124	0.136
	0.01	0.010	0.078	0.270	0.010	0.056	0.054	0.062
50	0.1	0.112	0.374	0.756	0.131	0.007	0.007	0.009
	0.05	0.055	0.242	0.636	0.059	0.124	0.081	0.088
	0.025	0.029	0.140	0.537	0.023	0.038	0.036	0.041
	0.01	0.011	0.073	0.401	0.008	0.038	0.036	0.041
100	0.1	0.117	0.310	0.798	0.086	0.082	0.107	0.113
	0.05	0.059	0.201	0.699	0.039	0.037	0.053	0.058
	0.025	0.033	0.201	0.610	0.039	0.037	0.024	0.027
	0.01	0.012	0.120	0.480	0.016	0.015	0.010	0.012
200	0.1	0.159	0.370	0.884	0.114	0.109	0.064	0.070
	0.05	0.090	0.259	0.814	0.058	0.054	0.031	0.035
	0.025	0.047	0.169	0.722	0.027	0.025	0.014	0.016
	0.01	0.023	0.102	0.614	0.012	0.011	0.103	0.057
300	0.1	0.177	0.401	0.903	0.130	0.064	0.030	0.013
	0.05	0.104	0.289	0.844	0.069	0.031	0.006	0.002
	0.025	0.057	0.196	0.765	0.034	0.014	0.082	0.089
	0.01	0.029	0.125	0.670	0.015	0.012	0.029	0.031

Таблица 2.11

Мощность критерия Фостера–Стюарта со статистикой (2.24) относительно гипотез $H_8 - H_{12}$, соответствующих изменению дисперсии скачком

n	α	H_8	H_9	H_{10}	H_{11}	H_{12}
10	0.1	0.138	0.154	0.171	0.268	0.418
	0.05	0.138	0.154	0.171	0.268	0.418
	0.025	0.027	0.033	0.039	0.078	0.153
	0.01	0.027	0.033	0.039	0.078	0.153
20	0.1	0.146	0.163	0.185	0.321	0.513
	0.05	0.044	0.055	0.068	0.151	0.295
	0.025	0.044	0.055	0.068	0.151	0.295
	0.01	0.011	0.016	0.021	0.057	0.137
30	0.1	0.139	0.168	0.200	0.383	0.596
	0.05	0.067	0.080	0.097	0.215	0.400
	0.025	0.020	0.027	0.036	0.103	0.232
	0.01	0.020	0.027	0.036	0.103	0.232
50	0.1	0.106	0.126	0.154	0.336	0.556
	0.05	0.085	0.127	0.144	0.335	0.556
	0.025	0.040	0.053	0.071	0.198	0.390
	0.01	0.014	0.021	0.030	0.104	0.246
100	0.1	0.106	0.130	0.165	0.394	0.625
	0.05	0.085	0.117	0.157	0.394	0.625
	0.025	0.045	0.062	0.087	0.263	0.484
	0.01	0.018	0.028	0.043	0.161	0.347
150	0.1	0.099	0.134	0.180	0.448	0.556
	0.05	0.058	0.078	0.107	0.319	0.556
	0.025	0.047	0.071	0.104	0.319	0.390
	0.01	0.021	0.035	0.055	0.211	0.246
200	0.1	0.161	0.208	0.269	0.578	0.785
	0.05	0.083	0.120	0.168	0.446	0.679
	0.025	0.041	0.065	0.099	0.322	0.559
	0.01	0.021	0.034	0.054	0.218	0.436
300	0.1	0.152	0.211	0.285	0.619	0.816
	0.05	0.084	0.129	0.187	0.495	0.724
	0.025	0.048	0.076	0.116	0.374	0.616
	0.01	0.012	0.021	0.036	0.179	0.387

Таблица 2.12

Мощность критерия Фостера–Стюарта со статистикой (2.24) относительно гипотез $H_{13} - H_{15}$, соответствующих наличию линейного, периодического тренда и смешанного тренда в дисперсии

n	α	H_{13}	H_{14}	H_{15}
10	0.1	0.204	0.142	0.157
	0.05	0.204	0.142	0.157
	0.025	0.052	0.008	0.028
	0.01	0.052	0.008	0.028
25	0.1	0.207	0.112	0.081
	0.05	0.205	0.026	0.054
	0.025	0.090	0.016	0.020
	0.01	0.033	0.013	0.008
50	0.1	0.272	0.087	0.065
	0.05	0.270	0.022	0.049
	0.025	0.149	0.016	0.020
	0.01	0.073	0.001	0.006
100	0.1	0.342	0.059	0.072
	0.05	0.341	0.020	0.065
	0.025	0.218	0.013	0.029
	0.01	0.128	0.002	0.011
200	0.1	0.551	0.072	0.198
	0.05	0.416	0.025	0.110
	0.025	0.294	0.008	0.055
	0.01	0.193	0.003	0.026
300	0.1	0.609	0.076	0.266
	0.05	0.482	0.034	0.164
	0.025	0.359	0.016	0.093
	0.01	0.166	0.002	0.023
500	0.1	0.588	0.075	0.278
	0.05	0.470	0.036	0.180
	0.025	0.357	0.016	0.108
	0.01	0.257	0.007	0.061

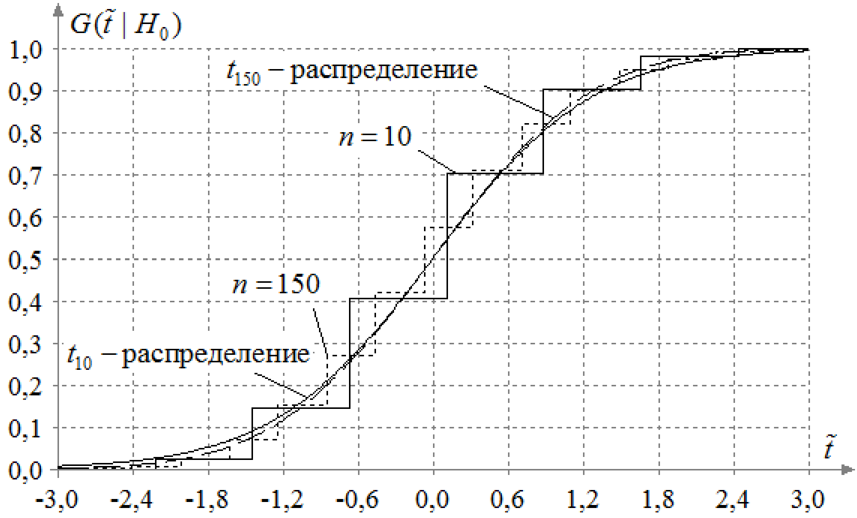


Рис. 2.17. Функции распределения статистики (2.24) критерия Фостера–Стюарта в зависимости от объёмов выборок

2.13. Критерий Кокса–Стюарта

Непараметрический критерий Кокса–Стюарта [13] может использоваться для проверки последовательности измерений $X_1, X_2, ..., X_n$ на предмет наличия тренда и в среднем, и в дисперсии.

Для проверки гипотезы об отсутствии тренда в средних значениях по выборке объема n используется критерий со статистикой

$$S_1 = \sum_{i=1}^{[n/2]} (n - 2i + 1) h_{i,n-i+1} , \tag{2.25}$$

где $h_{i,j} = 1$, если $X_i > X_j$, и $h_{i,j} = 0$, если $X_i \le X_j$ ($i < j$).

Нормализованная статистика

$$S_1^* = \frac{S_1 - E[S_1]}{\sqrt{D[S_1]}} , \tag{2.26}$$

где

$$E[S_1]=\frac{n^2}{8},\ D[S_1]=\frac{n(n^2-1)}{24},$$

при справедливости проверяемой гипотезы об отсутствии тренда приближенно описывается стандартным нормальным законом.

Распределение статистики (2.26) является дискретным и при малых n следует учитывать его отличие от стандартного нормального закона (см. рис. 2.18 при $n=10$). Однако при $n>40$ отличием распределения статистики (2.26) от стандартного нормального закона можно практически пренебречь.

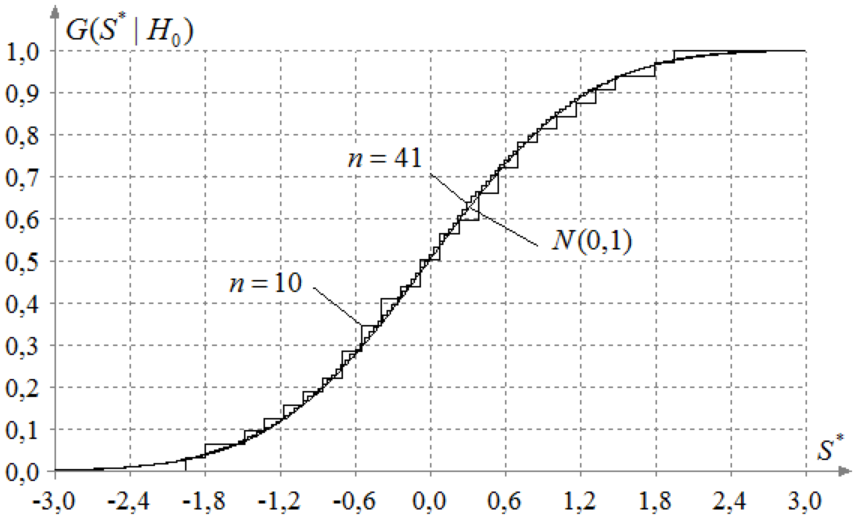


Рис. 2.18. Сходимость распределения статистики (2.26) критерия Кокса-Стюарта к стандартному нормальному закону

При меньших объемах выборок для вычисления достигнутого уровня значимости, соответствующего полученному значению статистики S_1^*, целесообразно использовать реальное распределение статистики, имеющее место при данном объеме выборки n.

"Асимптотические" оценки мощности критерия Кокса–Стюарта со статистикой (2.26) относительно конкурирующих гипотез, соответствующих наличию линейного H_1-H_3 и периодического H_4-H_7 тренда в математическом ожидании наблюдаемых случайных величин, представлены в таблице 2.13.

Таблица 2.13

Мощность критерия Кокса–Стюарта относительно гипотез $H_1 - H_7$ наличия линейного и периодического тренда в математическом ожидании

n	α	H_1	H_2	H_3	H_4	H_5	H_6	H_7
10	0.1	0.114	0.157	0.515	0.222	0.098	0.102	0.106
	0.05	0.057	0.002	0.190	0.127	0.031	0.030	0.027
	0.025	0.028	0.002	0.041	0.127	0.031	0.030	0.027
	0.01	0.012	0.002	0.003	0.127	0.031	0.030	0.027
25	0.1	0.154	0.471	0.984	0.288	0.100	0.094	0.075
	0.05	0.084	0.321	0.953	0.187	0.046	0.042	0.029
	0.025	0.049	0.216	0.906	0.117	0.023	0.021	0.013
	0.01	0.020	0.130	0.812	0.069	0.010	0.008	0.003
50	0.1	0.206	0.757	1.000	0.442	0.102	0.101	0.100
	0.05	0.123	0.636	1.000	0.319	0.050	0.050	0.042
	0.025	0.075	0.515	0.999	0.224	0.025	0.024	0.017
	0.01	0.039	0.367	0.997	0.136	0.010	0.009	0.005
100	0.1	0.308	0.957	1.000	0.686	0.107	0.125	0.172
	0.05	0.210	0.915	1.000	0.562	0.054	0.064	0.085
	0.025	0.135	0.858	1.000	0.447	0.027	0.032	0.040
	0.01	0.077	0.759	1.000	0.317	0.011	0.013	0.014
200	0.1	0.491	0.999	1.000	0.915	0.118	0.169	0.319
	0.05	0.366	0.998	1.000	0.853	0.062	0.095	0.190
	0.025	0.264	0.994	1.000	0.778	0.033	0.053	0.107
	0.01	0.166	0.984	1.000	0.666	0.014	0.024	0.047
300	0.1	0.634	1.000	1.000	0.981	0.131	0.215	0.462
	0.05	0.509	1.000	1.000	0.959	0.071	0.129	0.308
	0.025	0.395	1.000	1.000	0.927	0.038	0.075	0.193
	0.01	0.274	0.999	1.000	0.866	0.017	0.036	0.095

Критерий Кокса–Стюарта при проверке гипотезы об отсутствии тренда в дисперсии (в характеристиках рассеяния) строится следующим образом.

Исходная выборка $X_1, X_2, ..., X_n$ разбивается на $[n/k]$ подвыборок объемом k элементов $X_1,...,X_k$; $X_{k+1},...,X_{2k}$; $X_{2k+1},...,X_{3k}$; ...; $X_{n-k+1},...,X_n$ (если n не делится на k, отбрасывается необходимое число измерений в центре). Для каждой i-й подвыборки находится размах w_i ($1 \le i \le r$, $r = [n/k]$). Далее полученная последовательность размахов w_i проверяется на наличие тренда в средних значениях тем же критерием со статистикой (2.26).

Величину k предлагается выбирать, исходя из следующих рекомендаций авторов [13]:

$$n \geq 90 \rightarrow k = 5; \quad 64 \leq n < 90 \rightarrow k = 4;$$
$$48 \leq n < 64 \rightarrow k = 3; \quad n < 48 \rightarrow k = 2.$$

Дискретность распределения статистики S_1^* при обнаружении тренда в дисперсии заметно выше (см. рис. 2.19). Это естественно, так как анализируемая выборка размахов содержит лишь $[n/k]$ элементов.

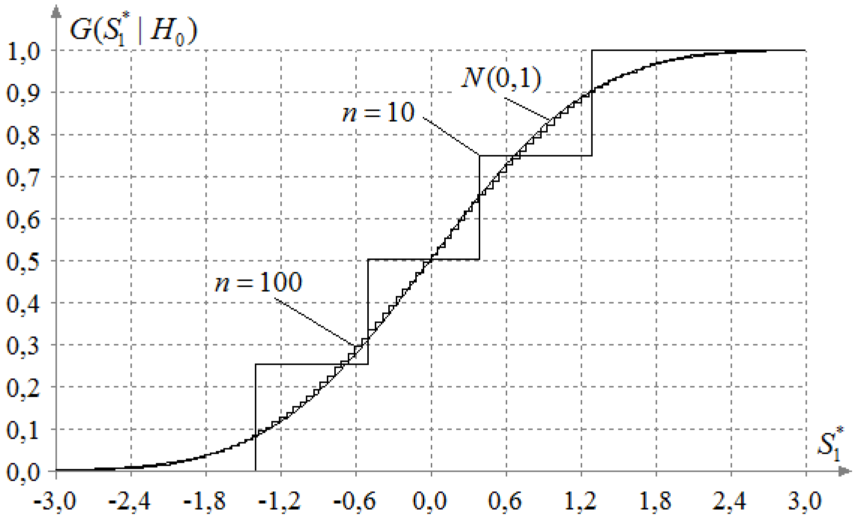


Рис. 2.19. Сходимость к стандартному нормальному закону распределения статистики (2.26) критерия Кокса-Стюарта (при обнаружении тренда в дисперсиях)

При использовании критерия Кокса–Стюарта для обнаружения тренда в дисперсиях отличием дискретного распределения статистики от стандартного нормального закона можно пренебрегать практически лишь при $n > 170$.

Оценки мощности критерия Кокса–Стюарта со статистикой (2.26) относительно близких конкурирующих гипотез $H_8 - H_{15}$, соответствующих наличию тренда в дисперсии наблюдаемых случайных величин, представлены в таблицах 2.14–2.15.

Таблица 2.14

Мощность критерия Кокса–Стюарта со статистикой (2.26) относительно гипотез $H_8 - H_{12}$, соответствующих изменению дисперсии скачком

n	α	H_8	H_9	H_{10}	H_{11}	H_{12}
10	0.1	0.220	0.193	0.170	0.087	0.024
	0.05	0.220	0.193	0.170	0.087	0.024
	0.025	0.220	0.193	0.170	0.087	0.024
	0.01	0.220	0.193	0.170	0.087	0.508
20	0.1	0.091	0.093	0.100	0.182	0.0001
	0.05	0.023	0.016	0.012	0.002	0.0001
	0.025	0.023	0.016	0.012	0.002	0.0001
	0.01	0.023	0.016	0.012	0.002	0.0001
30	0.1	0.093	0.098	0.109	0.214	0.485
	0.05	0.061	0.065	0.073	0.161	0.418
	0.025	0.022	0.024	0.029	0.087	0.306
	0.01	0.005	0.003	0.002	0.000	0.000
50	0.1	0.099	0.120	0.153	0.418	0.816
	0.05	0.055	0.070	0.095	0.313	0.735
	0.025	0.030	0.040	0.059	0.247	0.690
	0.01	0.012	0.017	0.026	0.154	0.593
100	0.1	0.123	0.188	0.284	0.808	0.997
	0.05	0.066	0.109	0.178	0.647	0.967
	0.025	0.031	0.060	0.107	0.530	0.947
	0.01	0.016	0.033	0.066	0.439	0.927
150	0.1	0.134	0.233	0.371	0.922	1.000
	0.05	0.072	0.143	0.253	0.849	0.999
	0.025	0.040	0.089	0.173	0.769	0.997
	0.01	0.017	0.044	0.095	0.633	0.985
200	0.1	0.149	0.281	0.456	0.973	1.000
	0.05	0.082	0.179	0.328	0.939	1.000
	0.025	0.046	0.115	0.233	0.893	1.000
	0.01	0.021	0.062	0.142	0.807	0.999
300	0.1	0.173	0.363	0.595	0.997	1.000
	0.05	0.103	0.254	0.470	0.992	1.000
	0.025	0.059	0.169	0.354	0.982	1.000
	0.01	0.028	0.098	0.237	0.957	1.000

Таблица 2.15

Мощность критерия Кокса–Стюарта со статистикой (2.26) относительно гипотез $H_{13}-H_{15}$, соответствующих наличию линейного, периодического тренда и смешанного тренда в дисперсии

n	α	H_{13}	H_{14}	H_{15}
10	0.1	0.142	0.166	0.131
	0.05	0.142	0.166	0.131
	0.025	0.142	0.166	0.131
	0.01	0.142	0.166	0.131
25	0.1	0.143	0.120	0.099
	0.05	0.053	0.012	0.040
	0.025	0.004	0.003	0.002
	0.01	0.004	0.003	0.002
50	0.1	0.229	0.112	0.053
	0.05	0.147	0.006	0.021
	0.025	0.084	0.003	0.017
	0.01	0.031	0.0001	0.011
100	0.1	0.485	0.120	0.055
	0.05	0.339	0.025	0.021
	0.025	0.214	0.001	0.004
	0.01	0.125	0.000	0.003
200	0.1	0.771	0.458	0.052
	0.05	0.654	0.245	0.023
	0.025	0.523	0.147	0.010
	0.01	0.373	0.012	0.003
300	0.1	0.907	0.757	0.058
	0.05	0.835	0.553	0.026
	0.025	0.741	0.341	0.011
	0.01	0.607	0.135	0.003
500	0.1	0.987	0.938	0.067
	0.05	0.971	0.849	0.029
	0.025	0.943	0.725	0.013
	0.01	0.888	0.523	0.005

2.14. Критерий Хсу обнаружения “сдвига дисперсии”

При использовании данного критерия отклонение проверяемой гипотезы об отсутствии тренда может свидетельствовать об обнаружении “сдвига дисперсии”. Статистика критерия Хсу имеет вид [33]

$$H = \frac{\sum_{i=1}^{n}(i-1)(X_i - \tilde{X})^2}{(n-1)\sum_{i=1}^{n}(X_i - \tilde{X})^2}, \quad 0 \le H \le 1, \tag{2.27}$$

где $\tilde{X}$ – выборочная медиана, которая может быть найдена по вариационному ряду $x_{(1)} \le x_{(2)} \le ... \le x_{(n)}$, построенному по выборке $X_1, X_2, ..., X_n$. При четном n оценка медианы принимает вид $\tilde{X} = (x_{(n/2)} + x_{(n/2+1)})/2$, при нечетном – $\tilde{X} = x_{((n-1)/2+1)}$.

В предположении, что математические ожидания последовательности случайных величин имеют одно и то же значение, проверяется гипотеза о неизменности дисперсий. В качестве конкурирующей гипотезы может рассматриваться изменение дисперсии наблюдаемых величин в некоторый (неизвестный) момент (начиная с некоторого элемента выборки). Критерий двусторонний: проверяемая гипотеза об отсутствии сдвига в дисперсии отклоняется при малых и больших значениях статистики.

Обычно критерий используется в нормализованной форме

$$H^* = \frac{H - 1/2}{\sqrt{D[H]}}, \tag{2.28}$$

где $D[H] = \dfrac{n+1}{6(n-1)(n+2)}$.

Статистика (2.28) при справедливости гипотезы об отсутствии изменения дисперсии в асимптотике подчиняется стандартному нормальному закону. Результаты моделирования показали [121], что при $n > 30$ распределение статистики (2.28) достаточно хорошо согласуется со стандартным нормальным законом (см. рис. 2.20).

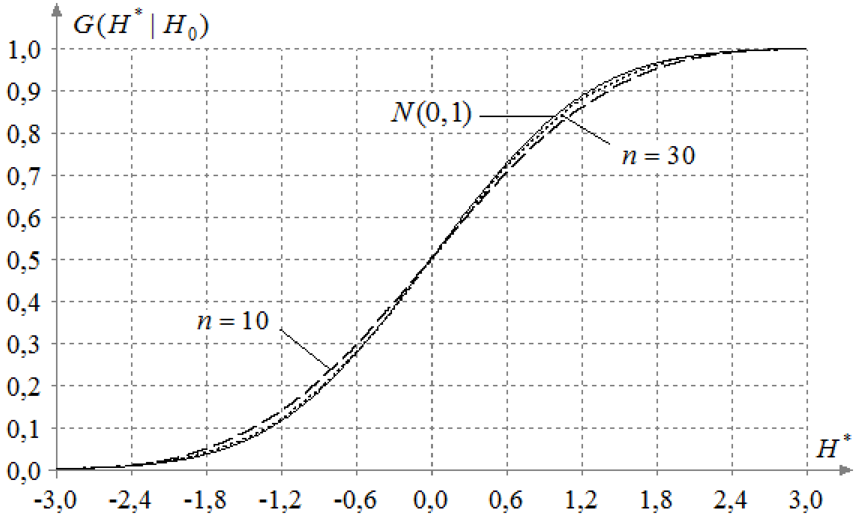


Рис. 2.20. Сходимость распределения статистики (2.28) критерия Хсу к стандартному нормальному закону

Критерий Хсу относится к параметрическим критериям. Поэтому, как и в случае любого параметрического критерия, связанного с проверкой гипотез о дисперсиях, распределения его статистики существенно зависят от закона, которому принадлежат анализируемые случайные величины (от закона распределения шума).

Полученные в результате моделирования распределения статистики (2.28) в случае принадлежности случайных величин законам распределения семейства (2.3) при различных значениях параметра формы представлены на рис. 2.21.

Как можно видеть, распределения статистики (2.28) сильно зависят от закона распределения, которому принадлежат случайные величины. При этом наибольшее отклонение от стандартного нормального закона наблюдается в случае принадлежности случайных величин законам с тяжелыми хвостами. Существенно влияет на распределение статистики и асимметричность закона [105].

Оценки мощности критерия Хсу со статистикой (2.27) относительно близких конкурирующих гипотез $H_8 - H_{15}$, соответствующих наличию тренда в дисперсии наблюдаемых случайных величин, представлены в таблицах 2.16–2.17.

Таблица 2.16

Мощность критерия Хсу со статистикой H (2.27) относительно гипотез $H_8 - H_{12}$, соответствующих изменению дисперсии скачком

n	α	H_8	H_9	H_{10}	H_{11}	H_{12}
10	0.1	0.103	0.111	0.123	0.231	0.473
	0.05	0.052	0.056	0.064	0.134	0.316
	0.025	0.026	0.029	0.033	0.076	0.204
	0.01	0.011	0.012	0.014	0.035	0.111
20	0.1	0.108	0.131	0.166	0.451	0.845
	0.05	0.055	0.070	0.093	0.309	0.717
	0.025	0.028	0.037	0.051	0.202	0.574
	0.01	0.011	0.016	0.023	0.110	0.400
30	0.1	0.114	0.153	0.213	0.633	0.967
	0.05	0.059	0.085	0.126	0.487	0.917
	0.025	0.031	0.047	0.074	0.360	0.839
	0.01	0.013	0.021	0.036	0.228	0.704
50	0.1	0.182	0.265	0.383	0.903	1.000
	0.05	0.126	0.198	0.303	0.857	0.999
	0.025	0.067	0.118	0.197	0.757	0.997
	0.01	0.036	0.069	0.125	0.644	0.989
100	0.1	0.156	0.304	0.500	0.991	1.000
	0.05	0.088	0.200	0.371	0.977	1.000
	0.025	0.049	0.128	0.266	0.953	1.000
	0.01	0.023	0.069	0.164	0.901	1.000
150	0.1	0.184	0.400	0.652	1.000	1.000
	0.05	0.110	0.282	0.526	0.998	1.000
	0.025	0.064	0.194	0.410	0.996	1.000
	0.01	0.031	0.114	0.282	0.988	1.000
200	0.1	0.214	0.490	0.766	1.000	1.000
	0.05	0.132	0.366	0.657	1.000	1.000
	0.025	0.080	0.263	0.544	1.000	1.000
	0.01	0.040	0.164	0.406	0.999	1.000
300	0.1	0.271	0.637	0.900	1.000	1.000
	0.05	0.175	0.513	0.835	1.000	1.000
	0.025	0.111	0.399	0.756	1.000	1.000
	0.01	0.059	0.274	0.641	1.000	1.000

Т а б л и ц а 2.17

Мощность критерия Хсу со статистикой H (2.27) относительно гипотез $H_{13} - H_{15}$, соответствующих наличию линейного, периодического тренда и смешанного тренда в дисперсии

n	α	H_{13}	H_{14}	H_{15}
10	0.1	0.164	0.220	0.132
	0.05	0.092	0.152	0.077
	0.025	0.051	0.104	0.045
	0.01	0.024	0.062	0.023
25	0.1	0.323	0.344	0.126
	0.05	0.217	0.257	0.069
	0.025	0.141	0.191	0.038
	0.01	0.077	0.127	0.018
50	0.1	0.554	0.504	0.137
	0.05	0.426	0.403	0.074
	0.025	0.316	0.319	0.039
	0.01	0.205	0.232	0.017
100	0.1	0.831	0.725	0.164
	0.05	0.738	0.631	0.093
	0.025	0.635	0.541	0.051
	0.01	0.498	0.435	0.023
200	0.1	0.982	0.924	0.218
	0.05	0.962	0.877	0.132
	0.025	0.932	0.822	0.080
	0.01	0.874	0.741	0.039
300	0.1	0.999	0.981	0.269
	0.05	0.996	0.964	0.174
	0.025	0.991	0.940	0.111
	0.01	0.978	0.898	0.058
500	0.1	1.000	0.999	0.364
	0.05	1.000	0.998	0.254
	0.025	1.000	0.995	0.174
	0.01	1.000	0.989	0.102

Способность критериев при малых объемах выборок отличать от H_0 такие близкие конкурирующие гипотезы достаточно невысока.

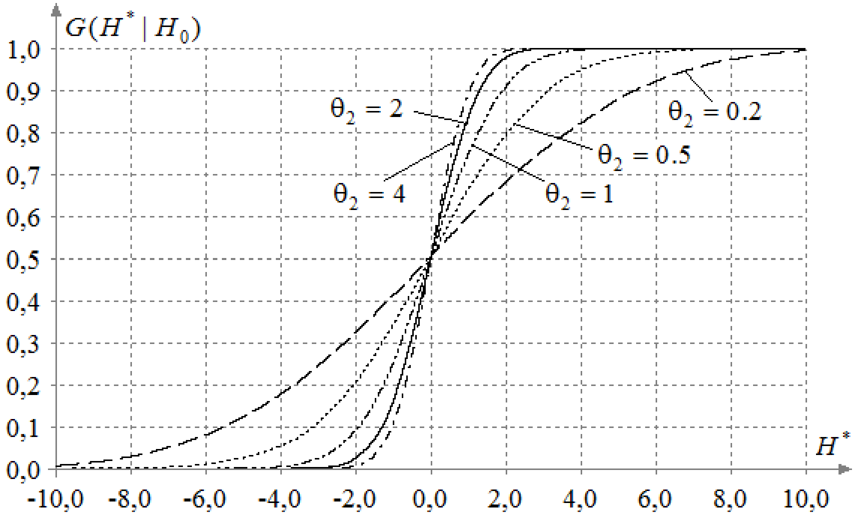


Рис. 2.21. Распределения статистики (2.28) в случае принадлежности случайных величин законам семейства (2.3) при различных значениях параметра формы при $n = 100$

Например, для того, чтобы различать гипотезы H_0 и H_{10} с заданной вероятностью ошибки 1-го рода $\alpha = 0.1$ и с вероятностью ошибки 2-го рода $\beta \leq 0.1$ в случае критерия Хсу требуются объемы выборок $n \geq 900$, а для различения в таких же условиях H_0 и H_8 объемы выборок должны быть ещё больше ($n \geq 6000$).

2.15. G-критерий Хсу обнаружения точки “сдвига дисперсии”

Критерий, позволяющий определить точку изменения дисперсии (в случае принадлежности наблюдений нормальному закону), предложен в [33]. Статистика этого критерия строится следующим образом. Пусть для $k = 1, 2, ..., n-1$

$$w_k = \sum_{i=1}^{k} (X_i - \tilde{X})^2 ,$$

$$W_k = \frac{w_n - w_k}{w_k} \frac{k}{n-k},$$

где k соответствует искомой точке изменения дисперсии, $\tilde{X}$ – выборочная медиана. В случае принадлежности X_i нормальному закону величины W_k, $k = 1, 2, ..., n-1$, принадлежат соответствующим $F_{n-k,k}(W)$ - распределениям Фишера с $n-k$ и k степенями свободы.

Далее по соответствующим функциям распределения находим $\gamma_k = F_{n-k,k}(W_k)$, где при отсутствии "сдвига в дисперсии" γ_k должны подчиняться равномерному закону.

Статистика *G-критерия* имеет вид

$$G = \frac{1}{n-1} \sum_{k=1}^{n-1} \gamma_k , \; 0 \le G \le 1. \qquad (2.29)$$

Гипотеза об отсутствии изменения дисперсии отклоняется с уровнем значимости α, если $G < G_{\alpha/2}$ или $G > G_{1-\alpha/2}$. В этом случае значение k, которому соответствует максимальная величина $|\gamma_k - 1/2|$, дает оценку искомой точки изменения значения дисперсии в наблюдаемом ряду. При $X_1 = \tilde{X}$ значение $w_1 = 0$, значит $W_1 = \infty$ и $\gamma_1 = 1$.

Изменение распределения статистики (2.29) в зависимости от объема выборки (для случая выполнения предположения о нормальности анализируемых выборок) иллюстрирует рис. 2.22.

В первоисточниках вид предельного распределения статистики (2.29) не приводится, даны лишь процентные точки.

На основе результатов статистического моделирования нами было показано, что хорошей моделью предельного распределения статистики (2.28) является бета-распределение 1-го рода с плотностью

$$f(x) = \frac{1}{\theta_2 \mathrm{B}(\theta_0, \theta_1)} \left(\frac{x-\theta_3}{\theta_2} \right)^{\theta_0 - 1} \left(1 - \frac{x-\theta_3}{\theta_2} \right)^{\theta_1 - 1}$$

и значениями параметров $\theta_0 = 2.7663$, $\theta_1 = 2.7663$, $\theta_2 = 1$, $\theta_3 = 0$.

Опираясь на этот закон можно находить процентные точки $G_{\alpha/2}$ и $G_{1-\alpha/2}$ или значения p_{value}.

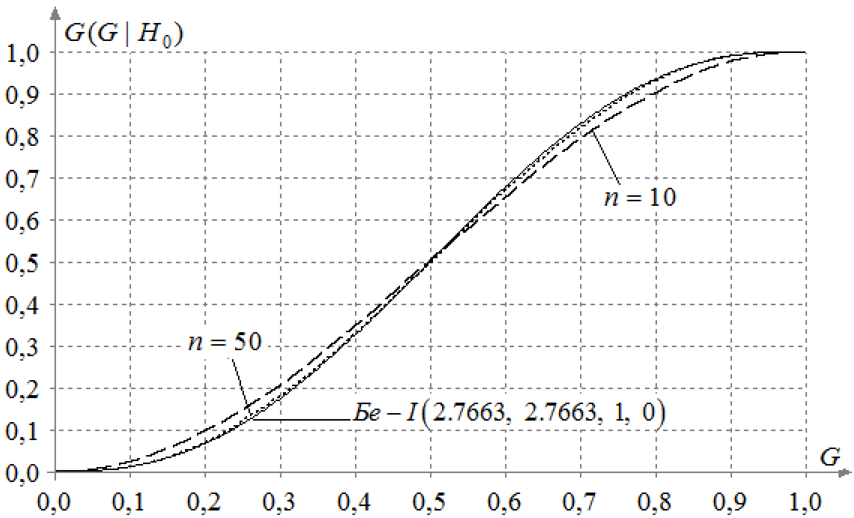


Рис. 2.22. Сходимость распределения статистики (2.29) G-критерия к бета-распределению 1-го рода

G-критерий также относится к параметрическим критериям. Поэтому распределения его статистики существенно зависят от вида наблюдаемого закона. Распределения статистики критерия для случая принадлежности случайной величины обобщенному нормальному закону (2.3), частными случаями которого при различных значениях параметра формы являются распределения нормальное и Лапласа, представлены на рисунке 2.23.

В случае нарушения стандартного предположения о нормальности X_i законами величин W_k, $k = 1, 2, ..., n-1$, не будут распределения Фишера. Для корректного применения критерия в нестандартных условиях должны быть найдены и использоваться при вычислении γ_k действительные распределения величин W_k, $k = 1, 2, ..., n-1$ (что предпочтительней, так как распределение G-статистики будет то же), либо могут использоваться $F_{n-k,k}(W)$-распределения Фишера, но тогда в этих условиях должно находиться неизвестное распределение G-статистики (что менее предпочтительно, но проще при реализации, так как надо найти только одно распределение).

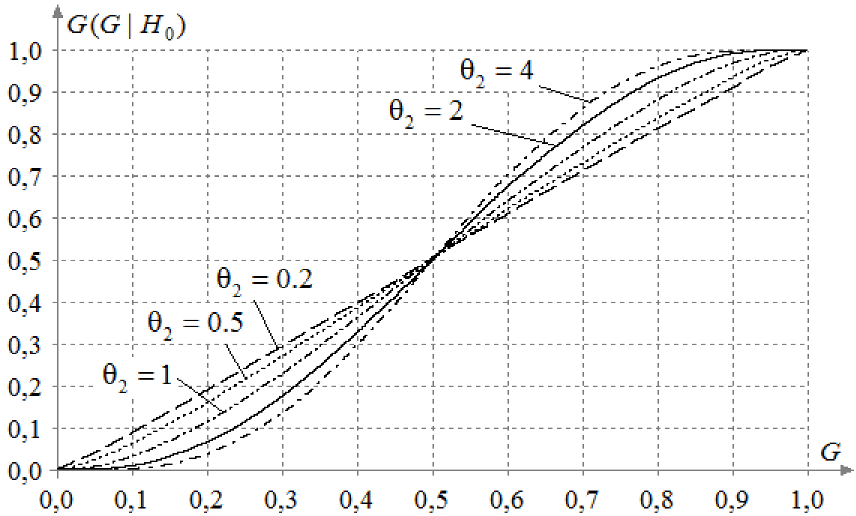


Рис. 2.23. Распределения статистики (2.29) G-критерия Хсу в случае принадлежности случайных величин законам семейства (2.3) при различных значениях параметра формы при $n = 100$

О возможности интерактивного исследования распределений статистик при использовании критериев в условиях нарушений стандартных предположений будет сказано ниже.

Оценки мощности G-критерия Хсу со статистикой (2.29) относительно близких конкурирующих гипотез $H_8 - H_{15}$, соответствующих наличию тренда в дисперсии наблюдаемых случайных величин, представлены в таблицах 2.18–2.19.

Таблица 2.18

Мощность G-критерия Хсу со статистикой (2.29) относительно гипотез $H_8 - H_{12}$, соответствующих изменению дисперсии скачком

n	α	H_8	H_9	H_{10}	H_{11}	H_{12}
10	0.1	0.103	0.111	0.123	0.227	0.467
	0.05	0.052	0.056	0.064	0.133	0.321
	0.025	0.026	0.029	0.033	0.076	0.216
	0.01	0.010	0.012	0.014	0.036	0.122
20	0.1	0.107	0.127	0.158	0.400	0.735
	0.05	0.055	0.068	0.088	0.268	0.594
	0.025	0.028	0.036	0.048	0.172	0.461
	0.01	0.011	0.015	0.020	0.090	0.307
30	0.1	0.112	0.146	0.196	0.545	0.854
	0.05	0.058	0.080	0.114	0.396	0.743
	0.025	0.030	0.043	0.065	0.274	0.620
	0.01	0.012	0.019	0.030	0.159	0.465
50	0.1	0.122	0.183	0.270	0.740	0.946
	0.05	0.064	0.105	0.168	0.597	0.881
	0.025	0.034	0.059	0.100	0.458	0.794
	0.01	0.014	0.027	0.050	0.306	0.663
100	0.1	0.147	0.269	0.430	0.933	0.993
	0.05	0.080	0.167	0.296	0.854	0.975
	0.025	0.044	0.107	0.195	0.750	0.941
	0.01	0.020	0.050	0.108	0.596	0.871
150	0.1	0.171	0.347	0.560	0.980	0.999
	0.05	0.097	0.229	0.412	0.942	0.993
	0.025	0.055	0.145	0.291	0.880	0.980
	0.01	0.025	0.077	0.174	0.766	0.943
200	0.1	0.195	0.422	0.664	0.993	1.000
	0.05	0.114	0.290	0.517	0.976	0.998
	0.025	0.065	0.192	0.382	0.940	0.992
	0.01	0.031	0.108	0.245	0.862	0.974
300	0.1	0.240	0.547	0.805	0.999	0.999
	0.05	0.146	0.401	0.676	0.995	0.998
	0.025	0.088	0.282	0.541	0.982	0.997
	0.01	0.043	0.167	0.376	0.946	0.994

Таблица 2.19

Мощность G-критерия Хсу со статистикой (2.29) относительно гипотез $H_{13}-H_{15}$, соответствующих наличию линейного, периодического тренда и смешанного тренда в дисперсии

n	α	H_{13}	H_{14}	H_{15}
10	0.1	0.163	0.204	0.125
	0.05	0.091	0.130	0.069
	0.025	0.051	0.081	0.038
	0.01	0.022	0.043	0.017
25	0.1	0.319	0.321	0.104
	0.05	0.210	0.224	0.055
	0.025	0.133	0.150	0.029
	0.01	0.070	0.084	0.013
50	0.1	0.542	0.436	0.084
	0.05	0.408	0.323	0.043
	0.025	0.296	0.229	0.022
	0.01	0.183	0.136	0.009
100	0.1	0.813	0.558	0.063
	0.05	0.705	0.434	0.030
	0.025	0.588	0.325	0.015
	0.01	0.435	0.208	0.006
200	0.1	0.973	0.678	0.040
	0.05	0.938	0.542	0.017
	0.025	0.885	0.423	0.008
	0.01	0.787	0.290	0.003
300	0.1	0.996	0.743	0.028
	0.05	0.988	0.605	0.011
	0.025	0.970	0.481	0.005
	0.01	0.925	0.337	0.002
500	0.1	1.000	0.821	0.015
	0.05	0.999	0.682	0.005
	0.025	0.997	0.550	0.002
	0.01	0.989	0.398	0.001

2.16. Ранговые критерии обнаружения “сдвига дисперсий” Клотца и Сэвиджа

Ранговые критерии обнаружения изменения параметра масштаба (характеристики рассеяния) в неизвестной точке опираются на использование семейства ранговых статистик вида [32]

$$S_R = \sum_{i=1}^{n} i a_n(R_i), \tag{2.30}$$

где R_i - ранги выборочных значений в упорядоченном ряду измерений.

Критерии различаются используемыми метками $a_n(\cdot)$. Их вид и определяет название критерия. Часто используются:

– метки Клотца $a_{1n}(i) = U^2_{i/(n+1)}$, где U_γ – γ-квантиль стандартного нормального закона;

– метки Сэвиджа $a_{2n}(i) = \sum_{j=1}^{i} \frac{1}{n-j+1}$.

При справедливости проверяемой гипотезы H_0 критерии со статистиками $S_{R,j} = \sum_{i=1}^{n} i a_{jn}(R_i)$, $j = 1,2$ свободны от распределения и симметричны относительно $E[S_{R,j}] = \frac{n+1}{2}\sum_{i=1}^{n} a_{jn}(i)$.

Обычно используются нормализованные критерии со статистиками вида

$$S^*_{R,j} = \frac{S_{R,j} - E[S_{R,j}]}{\sqrt{D[S_{R,j}]}}, \tag{2.31}$$

где

$$E[S_{R,1}] = \frac{n+1}{2}\sum_{i=1}^{n} U^2_{i/(n+1)}, \quad E[S_{R,2}] = \frac{n(n+1)}{2};$$

$$D[S_{R,1}] = \frac{n(n+1)}{12}\sum_{i=1}^{n} U^4_{i/(n+1)} - \frac{1}{3n+3}\left[E[S_{R,1}]\right]^2;$$

$$D[S_{R,2}] = \frac{n(n+1)}{12}(n - \sum_{j=1}^{n} \frac{1}{j}) .$$

Статистики (2.31) приближенно подчиняются стандартному нормальному закону. Сходимость распределений статистик к стандартному закону исследовалась в [122].

Исследование методами статистического моделирования распределений статистики критерия с метками Клотца показало (см. рис. 2.24), что при $n > 20$ распределения статистики критерия достаточно хорошо приближаются стандартным нормальным законом. Распределения статистики критерия с метками Сэвиджа также хорошо согласуются со стандартным нормальным законом, но при $n > 30$ (см. рис. 2.25).

Оценки мощности рангового критерия с метками Клотца и статистикой (2.30) относительно близких конкурирующих гипотез $H_8 - H_{15}$, соответствующих наличию тренда в дисперсии наблюдаемых случайных величин, представлены в таблицах 2.20–2.21.

По мощности данный критерий несколько уступает критерию Хсу (см. п. 2.15).

Оценки мощности рангового критерия с метками Сэвиджа и статистикой (2.31) относительно близких конкурирующих гипотез $H_8 - H_{15}$, соответствующих наличию тренда в дисперсии наблюдаемых случайных величин, представлены в таблицах 2.22–2.23.

Как можно видеть, ранговый критерий с метками Сэвиджа относительно рассмотренных конкурирующих гипотез $H_8 - H_{15}$ заметно уступает критерию с метками Клотца.

В случае принадлежности случайных величин нормальному закону критерии обнаружения изменения дисперсии (параметра масштаба) в неизвестной точке можно по мощности упорядочить следующим образом: критерий Хсу (H^*) $\succ$ критерий Клотца ($S^*_{R,1}$) $\succ$ критерий Сэвиджа ($S^*_{R,2}$).

Критерий с метками Клотца является локально наиболее мощным ранговым критерием в случае принадлежности выборок нормальному закону [107], что объясняется выбором меток в критерии.

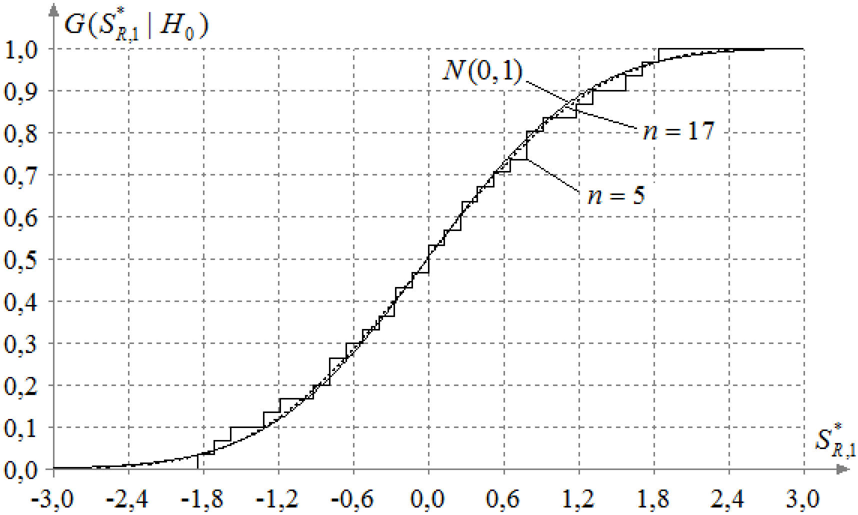


Рис. 2.24. Сходимость к стандартному нормальному закону распределения статистики $S^*_{R,1}$ рангового критерия с метками Клотца

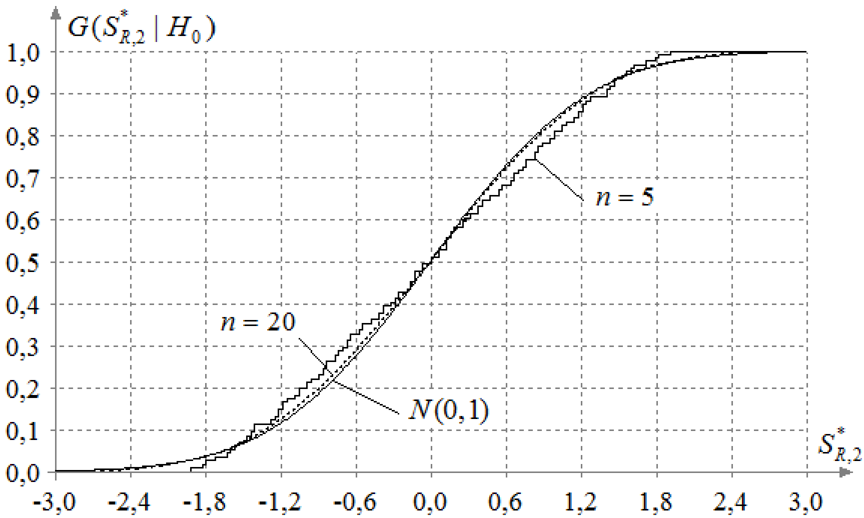


Рис. 2.25. Сходимость к стандартному нормальному закону распределения статистики $S^*_{R,2}$ рангового критерия с метками Сэвиджа

Таблица 2.20

Мощность критерия с метками Клотца со статистикой (2.31) относительно гипотез $H_8 - H_{12}$, соответствующих изменению дисперсии скачком

n	α	H_8	H_9	H_{10}	H_{11}	H_{12}
10	0.1	0.102	0.109	0.119	0.200	0.369
	0.05	0.052	0.056	0.062	0.113	0.231
	0.025	0.026	0.028	0.032	0.062	0.135
	0.01	0.011	0.012	0.013	0.028	0.067
20	0.1	0.107	0.127	0.156	0.387	0.753
	0.05	0.054	0.067	0.086	0.257	0.606
	0.025	0.028	0.035	0.047	0.163	0.458
	0.01	0.011	0.015	0.021	0.086	0.293
30	0.1	0.112	0.146	0.198	0.559	0.927
	0.05	0.058	0.081	0.116	0.415	0.852
	0.025	0.030	0.044	0.066	0.292	0.747
	0.01	0.012	0.019	0.031	0.173	0.584
50	0.1	0.123	0.186	0.279	0.798	0.996
	0.05	0.065	0.109	0.179	0.681	0.988
	0.025	0.034	0.063	0.110	0.558	0.971
	0.01	0.015	0.029	0.056	0.400	0.924
100	0.1	0.151	0.287	0.469	0.982	1.000
	0.05	0.085	0.186	0.342	0.958	1.000
	0.025	0.047	0.118	0.241	0.920	1.000
	0.01	0.022	0.063	0.146	0.844	1.000
150	0.1	0.179	0.382	0.623	0.999	1.000
	0.05	0.105	0.264	0.493	0.996	1.000
	0.025	0.061	0.178	0.376	0.991	1.000
	0.01	0.029	0.102	0.250	0.975	1.000
200	0.1	0.208	0.471	0.741	1.000	1.000
	0.05	0.127	0.346	0.626	1.000	1.000
	0.025	0.076	0.245	0.508	0.999	1.000
	0.01	0.037	0.150	0.369	0.997	1.000
300	0.1	0.263	0.618	0.887	1.000	1.000
	0.05	0.169	0.491	0.817	1.000	1.000
	0.025	0.107	0.378	0.737	1.000	1.000
	0.01	0.057	0.258	0.620	1.000	1.000

Таблица 2.21

Мощность критерия с метками Клотца со статистикой (2.31) относительно гипотез $H_{13}-H_{15}$, соответствующих наличию линейного, периодического тренда и смешанного тренда в дисперсии

n	α	H_{13}	H_{14}	H_{15}
10	0.1	0.143	0.096	0.094
	0.05	0.078	0.047	0.045
	0.025	0.041	0.025	0.021
	0.01	0.018	0.009	0.008
25	0.1	0.283	0.214	0.094
	0.05	0.179	0.125	0.045
	0.025	0.111	0.073	0.022
	0.01	0.056	0.034	0.008
50	0.1	0.505	0.397	0.093
	0.05	0.374	0.265	0.045
	0.025	0.266	0.171	0.021
	0.01	0.162	0.093	0.008
100	0.1	0.802	0.682	0.107
	0.05	0.697	0.540	0.053
	0.025	0.585	0.410	0.026
	0.01	0.443	0.268	0.010
200	0.1	0.977	0.932	0.141
	0.05	0.953	0.866	0.074
	0.025	0.915	0.779	0.038
	0.01	0.847	0.646	0.016
300	0.1	0.998	0.989	0.178
	0.05	0.995	0.971	0.099
	0.025	0.988	0.938	0.054
	0.01	0.971	0.871	0.023
500	0.1	1.000	1.000	0.256
	0.05	1.000	0.999	0.154
	0.025	1.000	0.997	0.091
	0.01	1.000	0.991	0.042

Таблица 2.22

Мощность критерия с метками Сэвиджа со статистикой (2.31) относительно гипотез $H_8 - H_{12}$, соответствующих изменению дисперсии скачком

n	α	H_8	H_9	H_{10}	H_{11}	H_{12}
10	0.1	0.100	0.101	0.102	0.109	0.126
	0.05	0.050	0.051	0.051	0.056	0.067
	0.025	0.025	0.025	0.025	0.028	0.036
	0.01	0.010	0.010	0.010	0.012	0.016
20	0.1	0.100	0.103	0.105	0.127	0.170
	0.05	0.050	0.051	0.054	0.066	0.094
	0.025	0.025	0.026	0.026	0.034	0.050
	0.01	0.010	0.010	0.011	0.014	0.012
30	0.1	0.101	0.105	0.110	0.154	0.229
	0.05	0.051	0.053	0.056	0.084	0.135
	0.025	0.026	0.026	0.029	0.044	0.077
	0.01	0.010	0.011	0.012	0.019	0.036
50	0.1	0.103	0.111	0.124	0.213	0.351
	0.05	0.052	0.057	0.064	0.124	0.229
	0.025	0.026	0.029	0.033	0.070	0.143
	0.01	0.010	0.013	0.014	0.032	0.074
100	0.1	0.110	0.129	0.159	0.364	0.610
	0.05	0.055	0.068	0.089	0.242	0.468
	0.025	0.028	0.036	0.049	0.156	0.345
	0.01	0.012	0.016	0.023	0.084	0.220
150	0.1	0.112	0.146	0.195	0.499	0.782
	0.05	0.058	0.081	0.115	0.362	0.663
	0.025	0.030	0.044	0.067	0.253	0.540
	0.01	0.012	0.019	0.031	0.149	0.388
200	0.1	0.118	0.165	0.234	0.618	0.886
	0.05	0.061	0.094	0.143	0.479	0.801
	0.025	0.032	0.053	0.085	0.356	0.700
	0.01	0.014	0.024	0.042	0.230	0.558
300	0.1	0.127	0.201	0. 324	0.781	0.969
	0.05	0.069	0.121	0.219	0.667	0.943
	0.025	0.037	0.071	0.114	0.549	0.909
	0.01	0.016	0.035	0.080	0.402	0.849

Таблица 2.23

Мощность критерия с метками Сэвиджа со статистикой (2.31) относительно гипотез $H_{13} - H_{15}$, соответствующих наличию линейного, периодического тренда и смешанного тренда в дисперсии

n	α	H_{13}	H_{14}	H_{15}
10	0.1	0.103	0.113	0.109
	0.05	0.051	0.058	0.057
	0.025	0.025	0.029	0.030
	0.01	0.010	0.011	0.012
25	0.1	0.119	0.119	0.102
	0.05	0.062	0.063	0.052
	0.025	0.032	0.033	0.026
	0.01	0.013	0.014	0.011
50	0.1	0.158	0.142	0.097
	0.05	0.088	0.079	0.048
	0.025	0.048	0.044	0.024
	0.01	0.020	0.020	0.009
100	0.1	0.244	0.199	0.097
	0.05	0.150	0.119	0.048
	0.025	0.090	0.070	0.024
	0.01	0.044	0.034	0.009
200	0.1	0.406	0.309	0.101
	0.05	0.284	0.205	0.0502
	0.025	0.191	0.132	0.025
	0.01	0.109	0.072	0.010
300	0.1	0.547	0.413	0.107
	0.05	0.415	0.291	0.054
	0.025	0.305	0.199	0.027
	0.01	0.193	0.116	0.011
500	0.1	0.750	0.587	0.120
	0.05	0.635	0.457	0.063
	0.025	0.521	0.345	0.033
	0.01	0.381	0.228	0.014

По оценкам мощности мы видим, что относительно очень близких конкурирующих гипотез и при малых объемах выборок критерий практически не уступает в мощности параметрическому критерию Хсу. Однако с удалением конкурирующих гипотез и возрастанием объемов выборок преимущества критерия Хсу становятся очевидными.

Мощность всех рассматриваемых критериев обнаружения изменения в дисперсии относительно тех же конкурирующих гипотез повышается с “облегчением” по сравнению с нормальным законом хвостов распределений наблюдаемых случайных величин и понижается в противном случае. В качестве примера в таблице 2.24 приведены оценки мощности критериев относительно конкурирующей гипотезы H_{10} в случае принадлежности случайных величин распределениям семейства (2.3) при различных значениях параметра формы. При $\theta_2 = 1$ мы имеем распределение Лапласа с более «тяжелыми» чем у нормального закона хвостами, а при $\theta_2 = 4$ – закон, более плосковершинный и с более “легкими” хвостами.

Мощность критерия с метками Сэвиджа во всех рассмотренных выше случаях оказывается существенно ниже. В [107] говорится о пригодности критерия Сэвиджа только для законов с плотностями положительными на полупрямой $0 < x < \infty$. В частности утверждается, что в случае принадлежности случайных величин экспоненциальному закону $f(x) = \frac{1}{\theta}\exp\{-x/\theta\}$ критерий является локально наиболее мощным ранговым критерием против конкурирующей гипотезы, связанной с увеличением параметра масштаба.

В таблице 2.25 приведены оценки мощности критериев относительно конкурирующих гипотез H_8, H_9, H_{10}, H_{12} в случае принадлежности случайных величин экспоненциальному закону распределении при объемах выборок n=50, которые подтверждают определенные преимущества критерия Сэвиджа в данной ситуации и, в то же время, показывают в этом случае относительно низкую мощность критерия с метками Клотца.

В таблице 2.26 представлены оценки мощности относительно конкурирующей гипотезы H_6 в случае принадлежности случайных величин распределению Вейбулла с плотностью

$$f(x) = \frac{\theta_0 (x-\theta_2)^{\theta_0 - 1}}{\theta_1^{\theta_0}} \exp\left\{-\left(\frac{x-\theta_2}{\theta_1}\right)^{\theta_0}\right\}$$

с параметром формы $\theta_0 = 4$, сдвига $\theta_2 = 0$ при объемах выборок $n = 50$. Экспоненциальный закон является частным случаем этого распределения при $\theta_0 = 1$.

В данном случае обращаем внимание на возрастание мощности критерия Хсу и критерия с метками Сэвиджа, а также на смещённость критерия с метками Клотца.

Таблица 2.24

Мощность критериев обнаружения “сдвига дисперсии” относительно конкурирующей гипотезы H_{10} в случае принадлежности наблюдений семейству (2.3)

θ_2	α	Критерий					
		Хсу		Клотца		Сэвиджа	
		n=10	n=30	n=10	n=30	n=10	n=30
1	0.1	0.124	0.154	0.124	0.159	0.107	0.116
	0.05	0.062	0.084	0.064	0.086	0.052	0.061
	0.01	0.014	0.019	0.013	0.019	0.010	0.012
4	0.1	0.135	0.211	0.141	0.215	0.109	0.129
	0.05	0.073	0.127	0.073	0.124	0.053	0.070
	0.01	0.014	0.027	0.015	0.026	0.010	0.014

Таблица 2.25

Мощность критериев обнаружения “сдвига дисперсии” относительно гипотез H_8, H_9, H_{10} и H_{12} в случае принадлежности наблюдений экспоненциальному закону

Конкурирующая гипотеза	α	Критерий Хсу	Критерий Клотца	Критерий Сэвиджа
H_8	0.1	0.123	0.112	0.128
	0.05	0.062	0.058	0.069
	0.01	0.012	0.012	0.013
H_9	0.1	0.147	0.122	0.158
	0.05	0.074	0.063	0.090
	0.01	0.016	0.013	0.018
H_{10}	0.1	0.170	0.131	0.188
	0.05	0.091	0.069	0.111
	0.01	0.020	0.015	0.024
H_{12}	0.1	0.583	0.229	0.740
	0.05	0.394	0.121	0.607
	0.01	0.140	0.033	0.295

Таблица 2.26

Мощность критериев обнаружения “сдвига дисперсии” относительно гипотезы H_{12} в случае принадлежности наблюдений распределению Вейбулла

Конкурирующая гипотеза	α	Критерий Хсу	Критерий Клотца	Критерий Сэвиджа
H_{12}	0.1	0.889	0.0448	1
	0.05	0.815	0.0135	1
	0.01	0.560	0.0014	1

Как правило, распределения статистик $G(S^*_{R,j}|H_i)$ при справедливости соответствующей конкурирующей гипотезы H_i сдвигаются с ростом объемов выборок относительно $G(S^*_{R,j}|H_0)$ влево или вправо в зависимости от вида конкурирующей гипотезы H_i. При таком поведении $G(S^*_{R,j}|H_i)$ за счет увеличения n критерий всегда способен с заданными вероятностями ошибок различать H_0 и H_i.

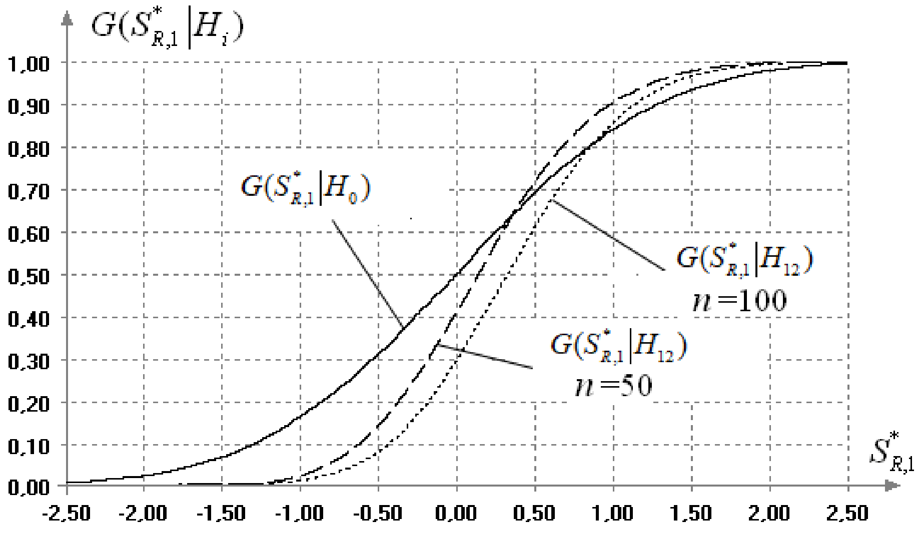


Рис. 2.26. Распределения статистики критерия Клотца в случае принадлежности наблюдений закону Вейбулла

В данном же случае критерий с метками Клотца не способен различать H_0 и H_{12}, что иллюстрирует рис. 2.26, на котором приведены распределения $G(S^*_{R,1}|H_0)$ и $G(S^*_{R,1}|H_{12})$ при $n=50$ и $n=100$.

2.17. Критерий инверсий

Инверсия имеет место, если в выборке значений $X_1, X_2, ..., X_n$, записанных в порядке их появления, за некоторым значением X_i следует меньшее по величине, т.е. $X_i > X_j$, где $i < j \le n$.

Статистикой критерия является общее число инверсий I в выборке $X_1, X_2, ..., X_n$ [149].

Критерий двусторонний. Гипотеза о случайности (об отсутствии тренда) не отклоняется, если $I_{\alpha/2} < I < I_{1-\alpha/2}$. Число инверсий принимает целые значения, множество которых зависит от объема выборки.

Математическое ожидание и дисперсия статистики I имеют соответственно вид [149].

$$E[I] = n(n-1)/4, \quad D[I] = (2n^3 + 3n^2 - 5n)/72.$$

Нормализованная статистика

$$I^* = \frac{I - E[I]}{\sqrt{D[I]}} \tag{2.32}$$

приближенно описывается стандартным нормальным законом. Гипотеза о случайности отклоняется при больших по модулю значениях статистики (2.31). При объемах выборок $n \ge 30$ дискретностью распределения статистики можно практически пренебречь (см. рис. 2.27) и опираться на стандартный нормальный закон как на распределение статистики.

Иногда рассматривают критерий со статистикой T, которая определяет число *обратных инверсий* ($X_i < X_j$, $i < j$), или критерий со статистикой $K = T - I$.

Оценки мощности критерия инверсий со статистиками I, I^* и K, T относительно конкурирующих гипотез, соответствующих наличию линейного $H_1 - H_3$ и периодического $H_4 - H_7$ тренда в математическом ожидании наблюдаемых случайных величин, представлены в таблицах 2.27 –2.29.

Таблица 2.27

Мощность критериев инверсий со статистикой I, I^*, K, T относительно гипотез $H_1 - H_3$ (наличия линейного тренда)

n	α	Критерии инверсий I, I^*			Критерии инверсий K, T		
		H_1	H_2	H_3	H_1	H_2	H_3
10	0.100	0.102	0.265	0.883	0.128	0.337	0.926
	0.050	0.068	0.198	0.822	0.089	0.263	0.883
	0.025	0.026	0.095	0.637	0.037	0.140	0.740
	0.010	0.015	0.060	0.517	0.022	0.095	0.637
25	0.100	0.173	0.640	1.000	0.185	0.659	1.000
	0.050	0.097	0.499	0.999	0.104	0.520	1.000
	0.025	0.054	0.374	0.998	0.059	0.395	0.998
	0.010	0.027	0.259	0.994	0.031	0.277	0.995
50	0.100	0.253	0.897	1.000	0.258	0.900	1.000
	0.050	0.163	0.827	1.000	0.167	0.831	1.000
	0.025	0.100	0.737	1.000	0.103	0.743	1.000
	0.010	0.053	0.614	1.000	0.055	0.621	1.000
100	0.100	0.401	0.994	1.000	0.403	0.994	1.000
	0.050	0.283	0.986	1.000	0.285	0.986	1.000
	0.025	0.195	0.972	1.000	0.196	0.972	1.000
	0.010	0.116	0.941	1.000	0.117	0.942	1.000
200	0.100	0.633	1.000	1.000	0.634	1.000	1.000
	0.050	0.509	1.000	1.000	0.509	1.000	1.000
	0.025	0.396	1.000	1.000	0.399	1.000	1.000
	0.010	0.273	1.000	1.000	0.274	1.000	1.000
300	0.100	0.785	1.000	1.000	0.785	1.000	1.000
	0.050	0.680	1.000	1.000	0.682	1.000	1.000
	0.025	0.574	1.000	1.000	0.576	1.000	1.000
	0.010	0.441	1.000	1.000	0.443	1.000	1.000

Таблица 2.28

Мощность критериев инверсий со статистикой I, I^* относительно гипотез $H_4 - H_7$ (наличия периодического тренда)

n	α	H_4	H_5	H_6	H_7
10	0.150	0.248	0.289	0.115	0.075
	0.100	0.187	0.230	0.077	0.047
	0.050	0.136	0.178	0.050	0.027
	0.025	0.062	0.098	0.018	0.008
	0.010	0.038	0.069	0.010	0.004
25	0.150	0.437	0.602	0.147	0.138
	0.100	0.358	0.529	0.100	0.091
	0.050	0.235	0.400	0.047	0.039
	0.025	0.150	0.297	0.023	0.017
	0.010	0.088	0.208	0.010	0.007
50	0.150	0.659	0.857	0.172	0.219
	0.100	0.572	0.807	0.116	0.147
	0.050	0.444	0.714	0.061	0.076
	0.025	0.327	0.611	0.031	0.037
	0.010	0.213	0.486	0.013	0.014
100	0.150	0.884	0.356	0.964	0.354
	0.100	0.837	0.269	0.944	0.266
	0.050	0.741	0.148	0.896	0.152
	0.025	0.642	0.085	0.834	0.087
	0.010	0.503	0.044	0.735	0.037
200	0.150	0.990	0.585	0.999	0.585
	0.100	0.982	0.490	0.999	0.485
	0.050	0.962	0.330	0.996	0.335
	0.025	0.931	0.198	0.992	0.220
	0.010	0.873	0.105	0.979	0.118
300	0.150	0.999	0.739	1.000	0.746
	0.100	0.999	0.653	1.000	0.659
	0.050	0.996	0.497	1.000	0.509
	0.025	0.990	0.361	1.000	0.372
	0.010	0.977	0.227	1.000	0.229

Таблица 2.29

Мощность критериев инверсий со статистикой K, T относительно гипотез $H_4 - H_7$ (наличия периодического тренда)

n	α	H_4	H_5	H_6	H_7
10	0.15	0.195	0.095	0.109	0.066
	0.1	0.140	0.060	0.072	0.039
	0.05	0.097	0.036	0.046	0.023
	0.025	0.040	0.010	0.016	0.007
	0.01	0.023	0.005	0.009	0.003
25	0.15	0.417	0.123	0.141	0.127
	0.1	0.340	0.081	0.095	0.083
	0.05	0.220	0.035	0.044	0.035
	0.025	0.138	0.015	0.021	0.015
	0.01	0.079	0.006	0.009	0.006
50	0.15	0.653	0.135	0.169	0.213
	0.1	0.565	0.086	0.113	0.142
	0.05	0.437	0.042	0.060	0.074
	0.025	0.321	0.019	0.030	0.036
	0.01	0.208	0.007	0.013	0.014
100	0.15	0.883	0.164	0.217	0.354
	0.1	0.835	0.116	0.157	0.267
	0.05	0.739	0.057	0.082	0.147
	0.025	0.639	0.030	0.045	0.084
	0.01	0.500	0.015	0.024	0.044
200	0.15	0.990	0.186	0.261	0.586
	0.100	0.982	0.132	0.191	0.485
	0.050	0.962	0.070	0.110	0.336
	0.025	0.931	0.035	0.062	0.222
	0.010	0.872	0.015	0.028	0.119
300	0.150	0.999	0.206	0.366	0.738
	0.100	0.998	0.148	0.288	0.653
	0.050	0.996	0.081	0.181	0.496
	0.025	0.990	0.047	0.110	0.361
	0.010	0.977	0.023	0.056	0.227

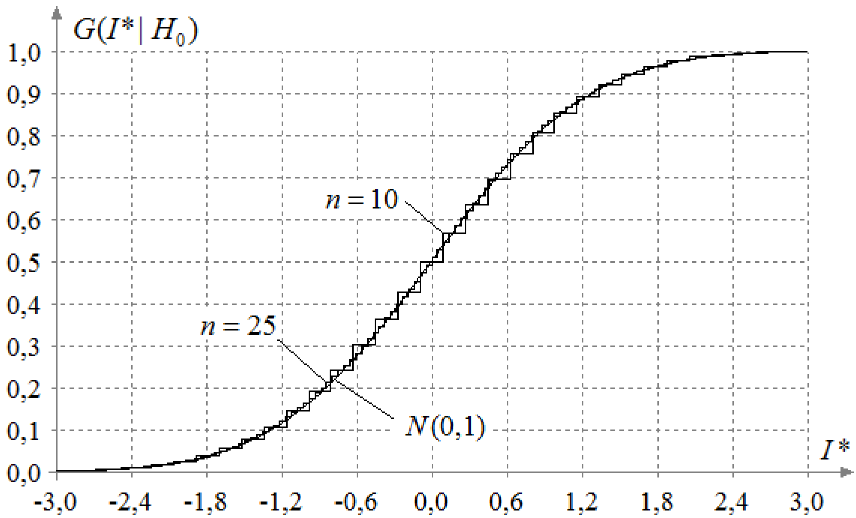


Рис. 2.27. Сходимость распределения статистики (2.32) к стандартному нормальному закону

2.18. Сериальный критерий Вальда–Вольфовица

Пусть имеются выборки нескольких случайных величин. Упорядочим их по возрастанию в общей последовательности. Тогда под серией понимается последовательность элементов одной из этих выборок в упорядоченной по возрастанию объединенной выборке, ограниченная с обеих сторон элементами других выборок (если это на границах последовательности, то с одной стороны).

Статистика сериального критерия Вальда–Вольфовица формируется следующим образом. Пусть имеется выборка $X_1, X_2, ..., X_n$ значений случайной величины X в порядке их появления, и пусть $\tilde{X}$ – выборочная медиана. Поставим в соответствие значениям $X_i \geq \tilde{X}$ символ a, а значениям $X_i < \tilde{X}$ – символ b. В результате получим последовательность символов a и b, состоящую из серий элементов a и b различной длины. В качестве статистики N_S сериального критерия Вальда–Вольфовица берётся общее число получившихся серий элементов a и b [92].

Проверяемая гипотеза H_0 о случайности ряда (об отсутствии тренда) не отвергается с вероятностью α, если $N_{S,\alpha/2} < N_S < N_{S,1-\alpha/2}$.

Критические значения $N_{S,\alpha/2}$ и $N_{S,1-\alpha/2}$ можно найти в [101, 109]. Расширенная таблица критических значений статистики, полученная в результате статистического моделирования, представлена в таблице 2.30.

Нормализованная статистика имеет вид

$$N_S^* = \frac{N_s - (2n_a n_b / (n_a + n_b) + 1)}{\sqrt{2n_a n_b (2n_a n_b - n_a - n_b) / ((n_a + n_b)^2 (n_a + n_b - 1))}}, \qquad (2.33)$$

где n_a и n_b – количества элементов a и b, соответствующих исходной последовательности $X_1, X_2, ..., X_n$. Очевидно, что n_a и n_b не превосходят величины $n/2$. В качестве аппроксимации распределения статистики (2.33) при $n_a, n_b > 20$ рекомендуется использовать стандартный нормальный закон. Проверяемая гипотеза H_0 отклоняется при больших по модулю значениях статистики N_S^*.

Однако исследования распределений статистики (2.33) критерия показали, что даже при относительно больших объемах выборок ($n = 700$) дискретное распределение статистики существенно отличается от (непрерывного) стандартного нормального закона (см. рис. 2.28).

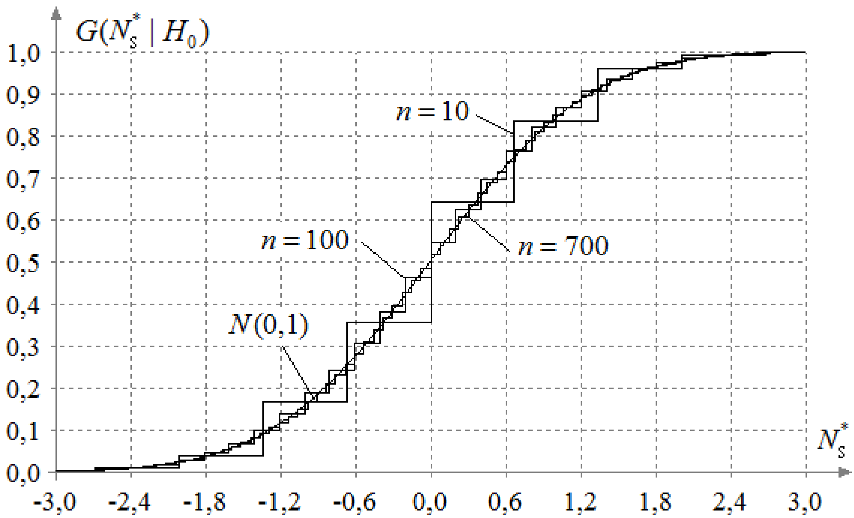


Рис. 2.28. Сходимость распределения статистики (2.33) сериального критерия Вальда-Вольфовица к нормальному закону

Таблица 2.30

Критические значения статистики сериального критерия Вальда–Вольфовица

n	$1-\alpha$					
	0.90		0.95		0.99	
	$N_{\alpha/2}$	$N_{1-\alpha/2}$	$N_{\alpha/2}$	$N_{1-\alpha/2}$	$N_{\alpha/2}$	$N_{1-\alpha/2}$
3	0.707	2.828	0.707	2.828	0.707	2.828
4	-1.225	1.225	-1.225	1.225	-1.225	1.225
5	-0.655	2.619	-0.655	2.619	-0.655	2.619
6	-1.826	1.826	-1.826	1.826	-1.826	1.826
7	-1.334	2.062	-1.334	2.910	-1.334	2.910
8	-1.334	2.062	-1.334	2.910	-1.334	2.910
9	-1.125	2.490	-1.125	2.490	-1.847	3.213
10	-1.342	1.342	-2.012	2.012	-2.683	2.683
11	-0.991	2.216	-0.991	2.216	-1.633	2.858
12	-1.817	1.817	-1.817	1.817	-2.422	2.422
13	-0.897	2.019	-1.480	2.602	-2.063	3.185
14	-1.669	1.669	-2.225	2.225	-2.225	2.225
15	-1.364	2.405	-1.364	2.405	-1.903	2.944
16	-1.553	1.553	-2.070	2.070	-2.588	2.588
17	-1.272	2.248	-1.272	2.248	-2.278	2.751
18	-1.458	1.458	-1.944	1.944	-2.430	2.430
19	-1.196	2.118	-1.670	2.592	-2.143	3.065
20	-1.838	1.838	-1.838	1.838	-2.297	2.297
25	-1.031	1.834	-1.441	2.243	-2.259	3.061
30	-1.486	1.486	-1.858	1.858	-2.601	2.601
40	-2.602	2.602	-1.922	1.922	-2.563	2.563
50	-1.715	1.715	-2.000	2.000	-2.572	2.572
100	-1.608	1.608	-2.010	2.010	-2.613	2.613
150	-1.639	1.639	-1.966	1.966	-2.622	2.622
200	-1.701	1.701	-1.985	1.985	-2.552	2.552
300	-1.619	1.619	-1.966	1.966	-2.545	2.545

Поэтому для вычисления достигнутого уровня значимости p_{value} целесообразно использовать реальное распределение статистики. Это распределение может находиться, например, в результате интерактивного моделирования при заданном объеме выборки.

Оценки мощности сериального критерия Вальда–Вольфовица со статистикой (2.33) относительно конкурирующих гипотез, соответствующих наличию линейного $H_1 - H_3$ и периодического $H_4 - H_7$ тренда в математическом ожидании наблюдаемых случайных величин, представлены в таблицах 2.31–2.32.

Таблица 2.31

Мощность сериального критерия Вальда–Вольфовица со статистикой (2.33) относительно гипотез $H_1 - H_3$ (наличия линейного тренда)

n	α	H_1	H_2	H_3
10	0.100	0.103	0.198	0.348
	0.050	0.052	0.061	0.237
	0.025	0.026	0.061	0.152
	0.010	0.010	0.005	0.090
25	0.100	0.110	0.179	0.687
	0.050	0.056	0.112	0.565
	0.025	0.027	0.035	0.446
	0.010	0.012	0.019	0.320
50	0.100	0.106	0.154	0.927
	0.050	0.053	0.111	0.873
	0.025	0.026	0.055	0.796
	0.010	0.010	0.038	0.687
100	0.100	0.109	0.291	0.997
	0.050	0.055	0.174	0.994
	0.025	0.027	0.117	0.986
	0.010	0.010	0.058	0.965
200	0.100	0.094	0.412	1.000
	0.050	0.051	0.313	1.000
	0.025	0.026	0.226	1.000
	0.010	0.012	0.155	1.000
300	0.100	0.115	0.571	1.000
	0.050	0.055	0.430	1.000
	0.025	0.032	0.347	1.000
	0.010	0.013	0.247	1.000

Таблица 2.32

Мощность сериального критерия Вальда–Вольфовица со статистикой (2.33) относительно гипотез $H_4 - H_7$ (наличия периодического тренда)

n	α	H_4	H_5	H_6	H_7
10	0.100	0.208	0.222	0.269	0.449
	0.050	0.055	0.053	0.070	0.146
	0.025	0.055	0.053	0.070	0.146
	0.010	0.005	0.007	0.005	0.002
25	0.100	0.183	0.150	0.138	0.178
	0.050	0.090	0.065	0.057	0.071
	0.025	0.035	0.022	0.017	0.007
	0.010	0.014	0.007	0.005	0.002
50	0.100	0.131	0.089	0.115	0.420
	0.050	0.078	0.047	0.066	0.318
	0.025	0.043	0.023	0.034	0.223
	0.010	0.023	0.010	0.017	0.150
100	0.100	0.203	0.115	0.196	0.762
	0.050	0.108	0.049	0.103	0.636
	0.025	0.076	0.030	0.072	0.565
	0.010	0.035	0.010	0.031	0.420
200	0.100	0.259	0.101	0.256	0.947
	0.050	0.179	0.055	0.176	0.913
	0.025	0.117	0.029	0.115	0.866
	0.010	0.072	0.014	0.071	0.803
300	0.100	0.365	0.125	0.364	0.992
	0.050	0.250	0.062	0.248	0.982
	0.025	0.185	0.037	0.184	0.971
	0.010	0.111	0.016	0.110	0.944

2.19. Критерий Рамачандрана–Ранганатана

В отличие от сериального критерия Вальда–Вольфовица, данный непараметрический критерий учитывает не только количество, но и длины серий (число элементов в сериях). Статистика критерия имеет вид [109]

$$RR = \sum_{j} j^2 n_j \ , \tag{2.34}$$

где j – длина серии, n – объем выборки, n_j – количество серий длины j.

Критерий правосторонний, гипотеза о случайности отвергается при больших значениях статистики (2.34). Критические значения статистики приводятся в [149, 109] для объемов выборок $n = 6 \div 30$ с шагом 2, что затрудняет его использование. Расширенная таблица приведена в таблице 2.33.

Дискретные распределения статистики существенно зависят от объема выборки (см. рис. 2.29). Поэтому применение критерия для произвольных объёмов выборок сопряжено с понятными трудностями.

Для вычисления достигнутого уровня значимости можно опираться на реальное (истинное) распределению статистики (2.34), получаемое в результате интерактивного моделирования (при данном объеме выборки).

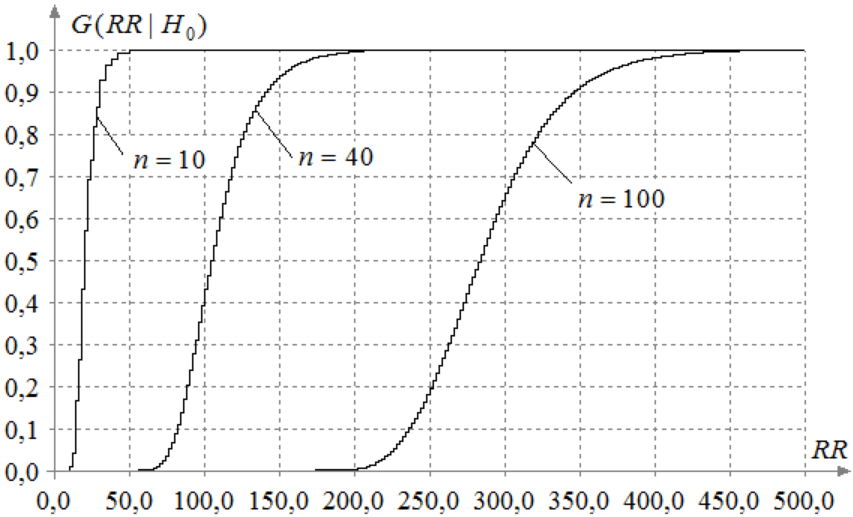


Рис. 2.29. Функции распределения статистики (2.34) критерия Рамачандрана–Ранганатана при различных объемах выборок

Оценки мощности критерия Рамачандрана–Ранганатана со статистикой (2.34) относительно конкурирующих гипотез, соответствующих наличию линейного $H_1 - H_3$ и периодического $H_4 - H_7$ тренда в математическом ожидании наблюдаемых случайных величин, представлены в таблицах 2.34–2.35.

Таблица 2.33

Критические значения статистики RR критерия Рамачандрана–Ранганатана

n	1–α				
	0.80	0.85	0.90	0.95	0.99
3	5	5	5	5	5
4	8	8	8	8	8
5	11	13	13	13	13
6	14	14	14	18	18
7	17	19	21	25	25
8	20	20	24	26	32
9	23	25	27	33	41
10	26	28	30	34	42
11	31	31	35	39	51
12	32	34	38	44	56
13	37	39	43	47	63
14	40	42	46	52	66
15	43	45	49	57	71
16	46	50	52	60	76
17	49	53	57	65	81
18	52	56	62	68	86
19	57	61	65	73	93
20	60	64	68	76	96
25	77	81	87	97	121
30	94	98	106	116	144
40	126	132	140	154	188
50	160	166	176	192	230
100	322	332	346	370	422
150	482	494	512	540	602
200	640	654	674	706	774
300	954	972	996	1034	1114

Таблица 2.34

Мощность критерия Рамачандрана–Ранганатана со статистикой (2.34) относительно гипотез $H_1 - H_3$ (наличия линейного тренда)

n	α	H_1	H_2	H_3
10	0.100	0.109	0.296	0.510
	0.050	0.056	0.249	0.422
	0.025	0.029	0.139	0.359
	0.010	0.011	0.073	0.282
25	0.100	0.130	0.062	0.846
	0.050	0.070	0.368	0.808
	0.025	0.039	0.289	0.779
	0.010	0.017	0.190	0.737
50	0.100	0.133	0.125	0.948
	0.050	0.072	0.068	0.936
	0.025	0.042	0.496	0.928
	0.010	0.019	0.408	0.914
100	0.100	0.139	0.291	0.984
	0.050	0.076	0.204	0.980
	0.025	0.047	0.123	0.977
	0.010	0.020	0.658	0.974
200	0.100	0.240	0.578	1.000
	0.050	0.173	0.448	1.000
	0.025	0.098	0.345	1.000
	0.010	0.055	0.235	1.000
300	0.100	0.187	0.839	1.000
	0.050	0.109	0.782	1.000
	0.025	0.062	0.679	1.000
	0.010	0.030	0.577	1.000

Таблица 2.35

Мощность критерия Рамачандрана–Ранганатана со статистикой (2.34) относительно гипотез $H_4 - H_7$ (наличия периодического тренда)

n	α	H_4	H_5	H_6	H_7
10	0.100	0.225	0.128	0.108	0.058
	0.050	0.145	0.067	0.055	0.026
	0.025	0.083	0.036	0.027	0.009
	0.010	0.065	0.022	0.016	0.005
25	0.100	0.273	0.104	0.097	0.077
	0.050	0.181	0.052	0.044	0.022
	0.025	0.120	0.026	0.022	0.010
	0.010	0.067	0.010	0.009	0.005
50	0.100	0.347	0.133	0.231	0.652
	0.050	0.241	0.067	0.126	0.477
	0.025	0.166	0.033	0.064	0.315
	0.010	0.099	0.013	0.023	0.149
100	0.100	0.459	0.173	0.407	0.947
	0.050	0.335	0.094	0.277	0.897
	0.025	0.245	0.052	0.187	0.833
	0.010	0.157	0.023	0.104	0.723
200	0.100	0.615	0.213	0.598	0.998
	0.050	0.490	0.127	0.469	0.995
	0.025	0.383	0.075	0.360	0.989
	0.010	0.270	0.037	0.246	0.976
300	0.100	0.723	0.240	0.714	1.000
	0.050	0.609	0.148	0.597	1.000
	0.025	0.499	0.089	0.486	0.999
	0.010	0.375	0.046	0.361	0.998

2.20. Критерий числа серий знаков первых разностей

Для выборки $X_1, X_2, ..., X_n$ вычисляются $n-1$ значений вида [150]

$$z_i = \begin{cases} -1, & X_{i+1} < X_i; \\ 1, & X_{i+1} > X_i; \\ 0, & X_{i+1} = X_i. \end{cases} \tag{2.35}$$

В ряду значений z_i фиксируется количество серий S, которое и является статистикой рассматриваемого критерия.

Гипотеза случайности ряда не отклоняется при $S_{\alpha/2} < S < S_{1-\alpha/2}$. Таблица критических значений для статистики S доступна в [109, 142].

Утверждается, что при $n > 30$ распределение статистики S удовлетворительно аппроксимируется нормальным распределением с математическим ожиданием и дисперсией:

$$E[S] = \frac{2n-1}{3}, \quad D[S] = \frac{16n-29}{90}.$$

В критерии, как правило, используют нормализованную статистику

$$S^* = \frac{S - E[S]}{\sqrt{D[S]}}, \qquad (2.36)$$

приближенным распределением которой при справедливости проверяемой гипотезы о случайности (об отсутствии тренда) является стандартный нормальный закон. Проверяемая гипотеза отклоняется при больших по модулю значениях S^*.

Однако, применяя данный критерий, следует иметь в виду, что распределение его статистики, как и статистики критерия Вальда–Вольфовица, остается дискретным даже при больших объемах выборок n (см. рис. 2.30).

Отсюда следует, что при проверке гипотезы оценку достигнутого уровня значимости p_{value} предпочтительнее вычислять по реальному (дискретному) распределению данной статистики, а не по стандартному нормальному закону.

Оценки мощности критерия числа серий знаков первых разностей со статистикой (2.36) относительно конкурирующих гипотез, соответствующих наличию линейного $H_1 - H_3$ и периодического $H_4 - H_7$ тренда в математическом ожидании наблюдаемых случайных величин, представлены в таблицах 2.36 –2.37.

Исследования мощности критерия относительно конкурирующих гипотез, связанных с наличием различных видов тренда, результаты которых представлены, в том числе в таблицах 2.36 –2.37, показали, что применение его в целях обнаружения тренда малоперспективно. Критерий оказался нечувствителен к присутствию тренда, соответствующего конкурирующим гипотезам вида $H_1 - H_7$.

В то же время использование критерия в качестве критерия проверки случайности вполне себя оправдывает, так как он хорошо реагирует на наличие в выборках серий убывающих и возрастающих последовательностей, отклоняя гипотезу H_0.

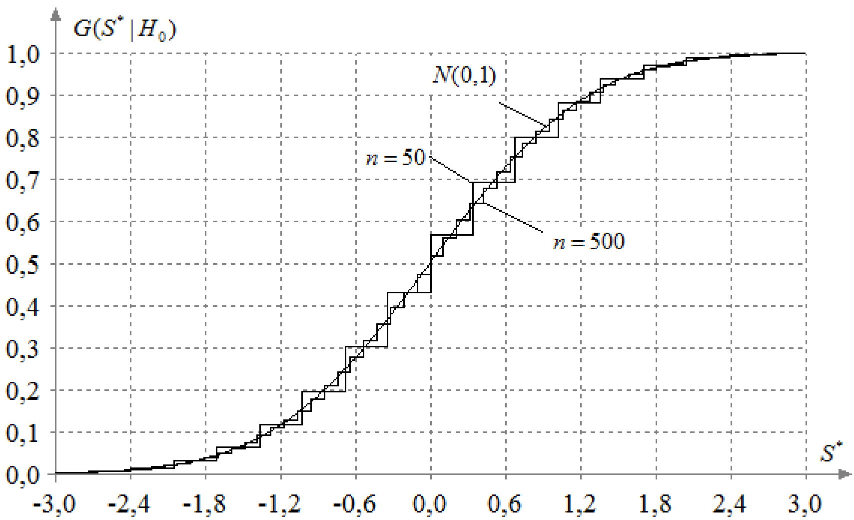


Рис. 2.30. Сходимость распределения статистики (2.36) критерия числа серий знаков первых разностей к нормальному закону

Таблица 2.36

Мощность критерия числа серий знаков первых разностей со статистикой (2.36) относительно гипотез $H_1 - H_3$ (наличия линейного тренда)

n	α	H_1	H_2	H_3
10	0.100	0.175	0.229	0.151
	0.050	0.036	0.169	0.051
	0.025	0.036	0.037	0.051
	0.010	0.028	0.037	0.022
25	0.100	0.085	0.027	0.086
	0.050	0.064	0.170	0.060
	0.025	0.025	0.085	0.026
	0.010	0.018	0.064	0.017
50	0.100	0.091	0.025	0.091
	0.050	0.042	0.018	0.041
	0.025	0.034	0.178	0.033
	0.010	0.014	0.091	0.013
100	0.100	0.100	0.042	0.100
	0.050	0.060	0.033	0.060
	0.025	0.023	0.013	0.023
	0.010	0.012	0.159	0.012
200	0.100	0.131	0.100	0.092
	0.050	0.093	0.060	0.043
	0.025	0.044	0.023	0.024
	0.010	0.029	0.012	0.011
300	0.100	0.097	0.131	0.100
	0.050	0.046	0.093	0.048
	0.025	0.027	0.044	0.029
	0.010	0.011	0.029	0.012

Таблица 2.37

Мощность критерия числа серий знаков первых разностей со статистикой (2.36) относительно гипотез $H_4 - H_7$ (наличия периодического тренда

n	α	H_4	H_5	H_6	H_7
10	0.100	0.167	0.194	0.247	0.432
	0.050	0.038	0.032	0.023	0.007
	0.025	0.038	0.032	0.023	0.007
	0.010	0.027	0.025	0.018	0.006
25	0.100	0.085	0.087	0.096	0.187
	0.050	0.063	0.062	0.059	0.096
	0.025	0.025	0.026	0.031	0.086
	0.010	0.018	0.018	0.017	0.037
50	0.100	0.091	0.091	0.092	0.122
	0.050	0.042	0.042	0.042	0.061
	0.025	0.033	0.033	0.031	0.038
	0.010	0.014	0.013	0.012	0.015
100	0.100	0.100	0.099	0.098	0.098
	0.050	0.060	0.059	0.058	0.059
	0.025	0.023	0.023	0.023	0.026
	0.010	0.012	0.012	0.012	0.014
200	0.100	0.093	0.093	0.092	0.093
	0.050	0.044	0.043	0.043	0.043
	0.025	0.029	0.029	0.028	0.029
	0.010	0.011	0.011	0.011	0.011
300	0.100	0.099	0.099	0.099	0.100
	0.050	0.047	0.047	0.047	0.047
	0.025	0.024	0.024	0.024	0.024
	0.010	0.011	0.011	0.011	0.011

2.21. Сравнительный анализ критериев, используемых для проверки гипотезы об отсутствии тренда в математическом ожидании

Для проверки гипотезы о случайности и отсутствии тренда в математическом ожидании наблюдаемого процесса может использоваться достаточно представительный ряд критериев.

В таблице 2.38 критерии, которые могут применяться для обнаружения тренда в математическом ожидании, упорядочены по мощности $1-\beta$, которую они проявляют относительно гипотезы H_1 (см. раздел 1.2), соответствующей наличию линейного тренда. Ранжирование критериев проведено по оценкам мощности, полученным при объёме выборок $n=100$ и заданной вероятности ошибки 1-го рода $\alpha=0.1$. В этой же таблице приведены мощности критериев относительно более далёких альтернатив H_2, H_3. Пример реализации рядов с трендом, соответствующим гипотезам H_0-H_3, представлен на рис. 2.31.

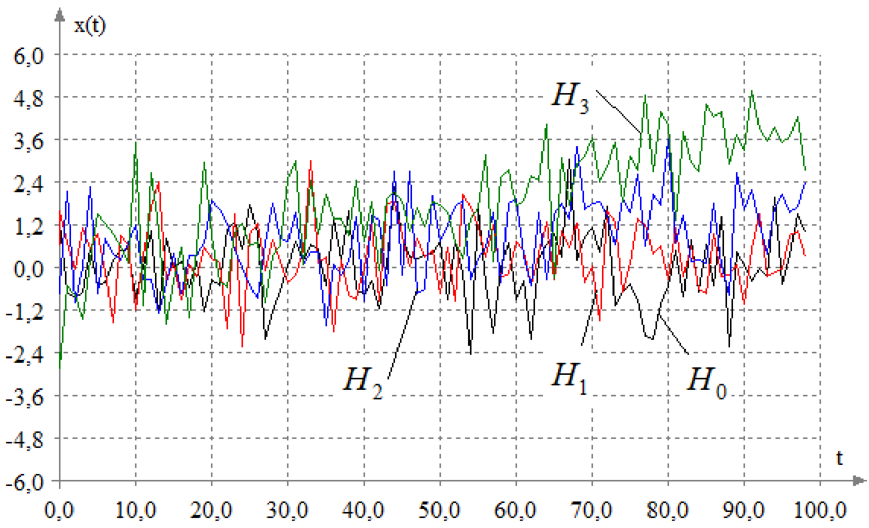


Рис. 2.31. Наложение линейного тренда на временной ряд

Таблица 2.38

Упорядоченность по мощности критериев проверки гипотез об отсутствии линейного тренда в математическом ожидании

№ п/п	Критерий	Оценки мощности $1-\beta$ относительно гипотез		
		H_1	H_2	H_3
1	Критерий обратных инверсий	0.403	0.994	1.000
2	Критерий К-инверсий	0.403	0.994	1.000
3	Критерий инверсий	0.401	0.994	1.000
4	Кокса–Стюарта	0.308	0.957	1.000
5	Рамачандрана–Ранганатана	0.139	0.291	0.984
6	Модификация критерия автокорреляции	0.125	0.651	1.000
7	Кумулятивной суммы	0.124	0.535	1.000
8	Фостера–Стюарта	0.117	0.310	0.798
9	Холлина	0.116	0.462	1.000
10	Критерий автокорреляции	0.113	0.455	1.000
11	Морана	0.113	0.455	1.000
12	Люнга–Бокса	0.113	0.455	1.000
13	Дюффа–Роя	0.113	0.455	1.000
14	Вальда–Вольфовица	0.113	0.455	1.000
15	Ранговый критерий Вальда–Вольфовица	0.112	0.461	1.000
16	Ранговый критерий Дюффа–Роя	0.112	0.455	1.000
17	Бартелса	0.112	0.461	1.000
18	Сериальный критерий Вальда–Вольфовица	0.109	0.291	0.997
19	Критерий числа серий знаков первых разностей	0.100	0.042	0.100

В таблице 2.39 в таких же условиях это же множество критериев упорядочено по мощности относительно гипотезы H_4, и приведены оценки мощности относительно гипотез $H_5 - H_7$, соответствующих наличию нелинейного (периодического) тренда в математическом ожидании. Отметим, что относительно конкурирующих гипотез, соответствующей наличию смешанного тренда вида $X_i = 0.5t_i + 0.5\sin(4\cdot 2\pi t_i) + \xi_i$, включающего линейную и периодическую составляющие, рассматриваемые критерии упорядочиваются по мощности точно так же, как и относительно гипотезы H_7. Пример реализации рядов с трендом, соответствующим гипотезам $H_0,\ H_4,\ H_6$, иллюстрирует рис. 2.32.

Таблица 2.39

Упорядоченность по мощности критериев проверки гипотез об отсутствии периодического тренда в математическом ожидании

№ п/п	Критерий	Оценки мощности $1-\beta$ относительно гипотез наличия нелинейного тренда			
		H_4	H_5	H_6	H_7
1	Критерий инверсий	0.837	0.356	0.964	0.354
2	Критерий обратных инверсий	0.835	0.116	0.157	0.267
3	Критерий К-инверсий	0.835	0.116	0.157	0.267
4	Кокса–Стюарта	0.686	0.107	0.125	0.172
5	Модификация критерия автокорреляции	0.463	0.141	0.425	0.990
6	Рамачандрана–Ранганатана	0.344	0.172	0.407	0.947
7	Критерий автокорреляции	0.320	0.123	0.308	0.950
8	Морана	0.320	0.123	0.308	0.950
9	Люнга–Бокса	0.320	0.123	0.308	0.950
10	Дюффа–Роя	0.320	0.123	0.308	0.950
11	Вальда–Вольфовица	0.320	0.123	0.308	0.950
12	Ранговый критерий Вальда–Вольфовица	0.314	0.122	0.302	0.950
13	Ранговый критерий Дюффа–Роя	0.314	0.122	0.302	0.950
14	Холлина	0.313	0.122	0.302	0.950
15	Бартелса	0.311	0.122	0.299	0.948
16	Сериальный критерий Вальда–Вольфовица	0.203	0.115	0.196	0.762
17	Критерий числа серий знаков первых разностей	0.100	0.099	0.098	0.098
18	Фостера–Стюарта	0.086	0.082	0.107	0.113
19	Кумулятивной суммы	0.081	0.078	0.072	0.081

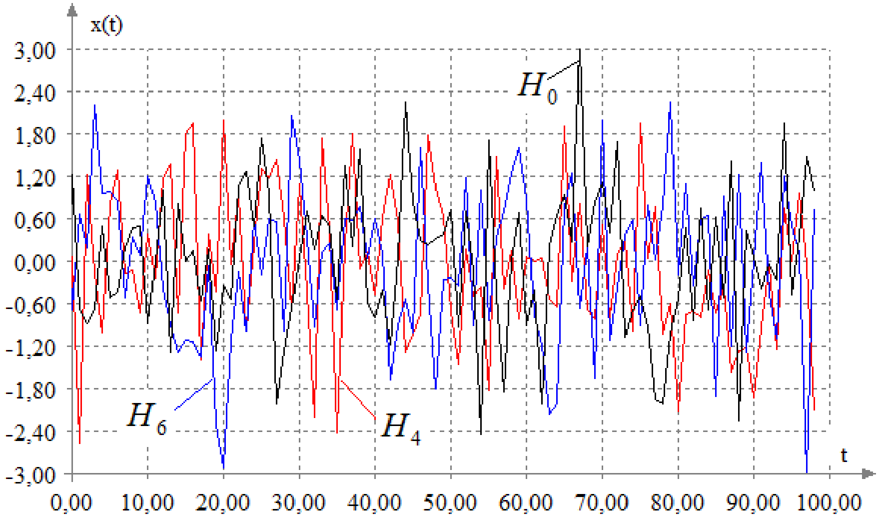


Рис. 2.32. Наложение периодического тренда на временной ряд

Как видим, по оценкам мощности, проявленной относительно близкой конкурирующей гипотезы с линейным трендом, можно отметить явное преимущество критериев, опирающихся на подсчёт инверсий (см. табл. 2.38).

Далее следует критерий Кокса–Стюарта, затем, заметно уступая, – модификация критерия автокорреляции и критерий Рамачандрана–Ранганатана.

Ниже расположена группа эквивалентных по мощности критериев, в которую входят критерий автокорреляции и критерии, построенные на нормализующих преобразованиях его статистики (Морана, Люнга–Бокса, Дюффа–Роя). Из этих критериев наиболее предпочтительно применение критерия Дюффа–Роя, так как распределение его статистики при справедливости H_0 наилучшим образом аппроксимируется стандартным нормальным законом. К этой же группе можно отнести критерий Вальда–Вольфовица и знаково-ранговый критерий Холлина.

Практически не уступают упомянутой группе ранговые критерии Вальда–Вольфовица и Дюффа–Роя, являющиеся эквивалентными по мощности, и критерий Бартелса.

Очень низкую мощность демонстрируют сериальный критерий Вальда–Вольфовица и критерий числа серий знаков первых разностей.

Из сериальных в лучшую сторону по мощности выделяется лишь критерий Рамачандрана–Ранганатана, учитывающий длины серий, но эффективное применение этого критерия затруднительно вследствие сильной зависимости распределения статистики от объема выборки.

Относительно гипотез с линейным трендом, более существенным по сравнению с H_1 (но более близких, чем H_2), упорядоченность критериев изменяется незначительно. Только критерии Рамачандрана–Ранганатана и Фостера-Стюарта смещаются вниз таблицы.

В таблице 2.39 приведены оценки мощности относительно конкурирующих гипотез, связанных с наличием нелинейного периодического тренда. Это достаточно близкие к H_0 гипотезы (см. рис. 2.32). Гипотезы $H_5 - H_7$ предполагают наличие периодического тренда с четырьмя периодами на участке наблюдения, а конкурирующей гипотезе H_4 соответствует только один период. Как показали исследования, с повышением частоты периодического тренда критерии инверсий хуже обнаруживают наложенный тренд, а, следовательно, мощность критериев падает.

Основные достоинства и недостатки критериев, применяемых при проверке гипотезы об отсутствии тренда в математическом ожидании, сформулированы в таблице 2.40, где критерии также упорядочены по уменьшению мощности.

Таблица 2.40

Основные достоинства и недостатки критериев, применяемых при проверке гипотез об отсутствии тренда в математическом ожидании

<table>
<tr><th>№ п/п</th><th>Критерий</th><th>Достоинства</th><th>Недостатки</th></tr>
<tr><td>1</td><td>Инверсий</td><td rowspan="3">Высокая мощность по отношению к линейному тренду. При $n \ge 30$ дискретностью нормализованных статистик можно пренебречь.</td><td rowspan="3">При $n < 30$ необходимо учитывать дискретность нормализованных статистик.</td></tr>
<tr><td>2</td><td>Обратных инверсий</td></tr>
<tr><td>3</td><td>К-инверсий</td></tr>
<tr><td>4</td><td>Кокса–Стюарта</td><td>Мощность выше среднего. При $n \ge 40$ дискретностью нормализованной статистики можно пренебречь.</td><td>При $n < 40$ необходимо учитывать дискретность нормализованной статистики.</td></tr>
</table>

Продолжение таблицы 2.40

№ п/п	Критерий	Достоинства	Недостатки
5	Модификация критерия автокорреляции	Относительно неплохая мощность.	Отличием распределения нормализованной статистики от стандартного нормального закона можно пренебречь лишь при $n > 200$.
6	Рамачандрана–Ранганатана	Относительно неплохая мощность.	Сильная зависимость распределения статистики от n. Необходимость использования таблицы критических значений.
7	Дюффа–Роя	При $n > 17$ отличием дискретного распределения нормализованной статистики от стандартного нормального закона можно пренебречь.	Невысокая мощность критерия.
8	Автокорреляции	Отклонением распределения нормализованной статистики от стандартного нормального закона можно пренебречь при $n > 30$.	Невысокая мощность критерия.
9	Морана		Невысокая мощность критерия. Отклонением распределения статистики от стандартного нормального закона можно пренебречь лишь при $n > 50$.
10	Люнга–Бокса		Невысокая мощность критерия. Распределение статистики очень медленно сходится к стандартному нормальному закону.

Продолжение таблицы 2.40

№ п/п	Критерий	Достоинства	Недостатки
11	Вальда–Вольфовица	При объёмах выборок $n > 20$ отклонением распределения нормализованной статистики от стандартного нормального закона можно пренебречь.	Невысокая мощность критерия.
12	Холлина	Средняя мощность.	Распределения статистики зависят от n. Критерий непараметрический, но распределение статистики реагирует на асимметричность наблюдаемого закона.
13	Ранговый Вальда–Вольфовица	При $n > 10$ в качестве распределения предложенной модификации нормализованной статистики можно использовать стандартный нормальный закон.	По мощности несколько уступает критериям Дюффа–Роя и Вальда–Вольфовица. Эквивалентен ранговому критерию Дюффа–Роя.
14	Ранговый Дюффа–Роя	Распределение статистики при $n > 17$ хорошо аппроксимируется стандартным нормальным законом. Дискретностью распределения статистики можно пренебречь при $n > 10$.	По мощности несколько уступает критериям Дюффа–Роя и Вальда–Вольфовица. Эквивалентен ранговому критерию Вальда–Вольфовица.
15	Бартелса	При $n > 10$ отличием дискретного распределения нормализованной статистики от стандартного нормального закона можно пренебречь	Невысокая мощность критерия.

Окончание таблицы 2.40

№ п/п	Критерий	Достоинства	Недостатки
16	Фостера–Стюарта		Высокая дискретность распределения статистики, которая сохраняется при больших n. Использование асимптотического t_n -распределения Стьюдента для оценки p-value приводит к большим погрешностям. Мощность ниже среднего относительно линейного тренда и низкая относительно нелинейного.
17	Кумулятивной суммы	Хорошая мощность относительно линейного тренда.	Дискретность распределения статистики и зависимость его от n. Очень низкая мощность относительно нелинейного тренда.
18	Сериальный Вальда–Вольфовица		Долго сохраняется дискретность распределения нормализованной статистики. Низкая мощность.
19	Числа серий знаков первых разностей		Дискретность распределения нормализованной статистики даже при больших объёмах выборок. Очень низкая мощность.

Использование стандартного нормального закона в качестве распределений статистик при объемах выборок n меньше указанных в таблице 2.40 может приводить к существенной ошибке при определении достигаемого уровня значимости p_{value}.

Применение параметрических критериев в условиях нарушения предположений о принадлежности анализируемых выборок нормальному закону также может приводить к ошибке при принятии решения [103] о результатах проверки.

Предотвратить возможную некорректность выводов при использовании параметрических и непараметрических критериев в условиях нарушения стандартных предположений можно за счет использования при проверке гипотезы реального распределения статистики, имеющего место при справедливости проверяемой гипотезы (в условиях реального приложения рассматриваемого критерия: при заданном объеме выборки, при конкретном законе распределения вероятностей наблюдаемой случайной величины). Такое распределение может быть найдено методами статистического моделирования в процессе проверки соответствующей гипотезы (в интерактивном режиме).

Такой режим для рассмотренных в руководстве критериев реализован в развиваемой программной системе, используемой для исследования и применения множеств различного рода критериев [148].

2.22. Сравнительный анализ критериев, используемых для проверки гипотезы об отсутствии тренда в характеристиках рассеяния

Для проверки гипотезы о случайности и отсутствии тренда в характеристиках рассеяния (в дисперсии) наблюдаемого процесса может использоваться существенно меньшее число критериев. К ним относятся критерии Кокса–Стюарта, Фостера–Стюарта, Клотца, Сэвиджа, H- и G-критерии Хсу.

В ходе исследований методами статистического моделирования (для вероятностей ошибок первого рода $\alpha = 0.15,\ 0.1,\ 0.05,\ 0.01$) выше были получены оценки мощности рассматриваемых критериев относительно конкурирующих гипотез $H_8 - H_{12}$ (см. п. 1.2), соответствующих сдвигу величины дисперсии. Была исследована мощность критериев относительно конкурирующих гипотез H_{13}, H_{14}, H_{15}, соответствующих наличию линейного или нелинейного тренда в характеристиках рассеяния анализируемых процессов.

Для сравнительного анализа мощности в таблицу 2.41 вынесены оценки мощности только при уровне значимости $\alpha = 0.1$ и объеме выборок $n = 100$. Критерии в колонках таблицы расположены в порядке убывания мощности $1 - \beta$.

Таблица 2.41

Упорядоченность по мощности критериев, применяемых для проверки гипотез об отсутствии тренда в характеристиках рассеяния

№ п/п	Относительно H_8		Относительно H_9	
	Критерий	$1-\beta$	Критерий	$1-\beta$
1	H-критерий Хсу	0.156	H-критерий Хсу	0.304
2	Клотца	0.151	Клотца	0.287
3	G-критерий Хсу	0.147	G-критерий Хсу	0.269
4	Кокса–Стюарта	0.123	Кокса–Стюарта	0.188
5	Сэвиджа	0.110	Фостера–Стюарта	0.130
6	Фостера–Стюарта	0.106	Сэвиджа	0.129
№ п/п	Относительно H_{10}		Относительно H_{11}	
	Критерий	$1-\beta$	Критерий	$1-\beta$
1	H-критерий Хсу	0.500	H-критерий Хсу	0.991
2	Клотца	0.469	Клотца	0.982
3	G-критерий Хсу	0.430	G-критерий Хсу	0.933
4	Кокса–Стюарта	0.284	Кокса–Стюарта	0.808
5	Фостера–Стюарта	0.165	Фостера–Стюарта	0.394
6	Сэвиджа	0.159	Сэвиджа	0.364
№ п/п	Относительно H_{12}		Относительно H_{13}	
	Критерий	$1-\beta$	Критерий	$1-\beta$
1	H-критерий Хсу	1.000	H-критерий Хсу	0.836
2	Клотца	1.000	G-критерий Хсу	0.818
3	Кокса–Стюарта	0.997	Клотца	0.807
4	G-критерий Хсу	0.993	Кокса–Стюарта	0.489
5	Фостера–Стюарта	0.625	Фостера–Стюарта	0.346
6	Сэвиджа	0.610	Сэвиджа	0.246

Окончание таблицы 2.41

№ п/п	Относительно H_{14}		Относительно H_{15}	
	Критерий	$1-\beta$	Критерий	$1-\beta$
1	H-критерий Хсу	0.711	H-критерий Хсу	0.162
2	Клотца	0.678	Сэвиджа	0.095
3	G-критерий Хсу	0.545	Фостера–Стюарта	0.082
4	Сэвиджа	0.196	G-критерий Хсу	0.057
5	Кокса–Стюарта	0.143	Кокса–Стюарта	0.052
6	Фостера–Стюарта	0.048	Клотца	0.104

Полученные оценки мощности указывают на преимущество критериев Хсу и Клотца [105, 106]. Им заметно уступают критерии Кокса–Стюарта, Фостера–Стюарта и, особенно, критерий Сэвиджа.

2.23. Технологии проверки гипотез о случайности и отсутствии тренда

Принятие решения о результатах проверки гипотезы H_0 на основании достигнутого уровня значимости (**p**-value) всегда более обосновано (см. раздел 1.2), чем в результате сравнения полученного значения статистики с заданным критическим значением, извлекаемым из соответствующей таблицы процентных точек. В последнем случае наша уверенность в правильности принимаемого решения отклонить (либо не отклонить) проверяемую гипотезу H_0 базируется на более слабом основании.

Вычисление достигнутых уровней значимости в соответствии с соотношениями (1.2) для правостороннего критерия или (1.3) для двустороннего не вызывает труда при известном распределении статистики критерия. Если информация о распределении статистики соответствующего критерия отсутствует и представлена лишь таблицей процентных точек, либо объёмы выборок относительно невелики и таковы, что распределение статистики существенно отличается от предельного (асимптотического), то корректное вычисление достигнутого уровня значимости (***p***-value) представляет собой некоторую проблему.

К сожалению (см. табл. 2.40), асимптотическими распределениями статистик большинства критериев проверки случайности и отсутствия тренда также приходится пользоваться осторожно. Чаще всего из-за дискретности, при ограниченных n реальные распределения статистик $G(S_n|H_0)$ многих из критериев существенно отличаются от асимптотических $G(S|H_0)$.

Обеспечить корректность статистических выводов и повысить их качество можно за счет использования компьютерных технологий, роль которых в программных системах статистического анализа становится всё более весомой. Именно с использованием компьютерных технологий анализа данных удаётся достаточно просто решать те задачи, которые в рамках только аналитических методов требуют годы.

В частности, когда распределение статистики критерия, используемого для проверки некоторой гипотезы, к моменту начала проверки (в силу разных причин) оказывается неизвестным (при данном объёме выборки n), появляется возможность исследования распределения статистики в реальном времени проверки гипотезы (в интерактивном режиме) [53, 54, 55, 56, 57, 129, 130, 131, 132, 133, 134]. Например, в интерактивном режиме можно исследовать неизвестное распределение статистики любого критерия проверки случайности и отсутствия тренда, зависящее от объема выборки, именно при том значении n, которое соответствует анализируемой выборке, и оценить по найденному в результате моделирования эмпирическому распределению статистики достигнутый уровень значимости ***p***-value.

При таком подходе необходимое для проверки гипотезы эмпирическое распределение $G_N(S_n|H_0)$ статистики соответствующего критерия строится в результате статистического моделирования с точностью, зависящей от числа экспериментов N в методе Монте-

Карло [128]. Затем по эмпирическому распределению $G_N(S_n|H_0)$ и вычисленному по анализируемой выборке значению статистики S^* критерия в соответствии с соотношением (1.2) для правостороннего критерия или по соотношению (1.3) для двустороннего критерия определяется оценка достигнутого уровня значимости p_{value}.

При проведении статистического моделирования в интерактивном режиме (в ходе осуществляемого статистического анализа) его результаты могут использоваться при формировании вывода по итогам проверки гипотезы.

Реализация такого интерактивного режима [53, 54] требует наличия развитого программного обеспечения, позволяющего (как в [148]) в целях ускорения распараллеливать процессы моделирования и привлекать доступные вычислительные ресурсы. В условиях распараллеливания время построения распределения $G_N(S_n|H_0)$ статистики критерия оказывается не очень заметным на фоне полного решения задачи статистического анализа.

В случае применения параметрических критериев в условиях нарушения предположения о нормальности, когда распределение статистики может существенно отличаться от “классического”, требуемое для вычисления p_{value} распределение статистики $G_N(S_n|H_0)$ можно моделировать при предполагаемом законе $F(x,\theta)$.

В качестве примера рассмотрим применение критериев проверки случайности и отсутствия тренда в программной системе [148], где требуемая последовательность действий выглядит следующим образом. В меню “Действия”, показанном на рис.2.33, выбирается пункт “Проверка на отсутствие тренда”.

На открывшейся вкладке (см. рис.2.34) появляется возможность загрузить выборку с анализируемой последовательностью (с временным рядом). Далее можно указать применяемые критерии.

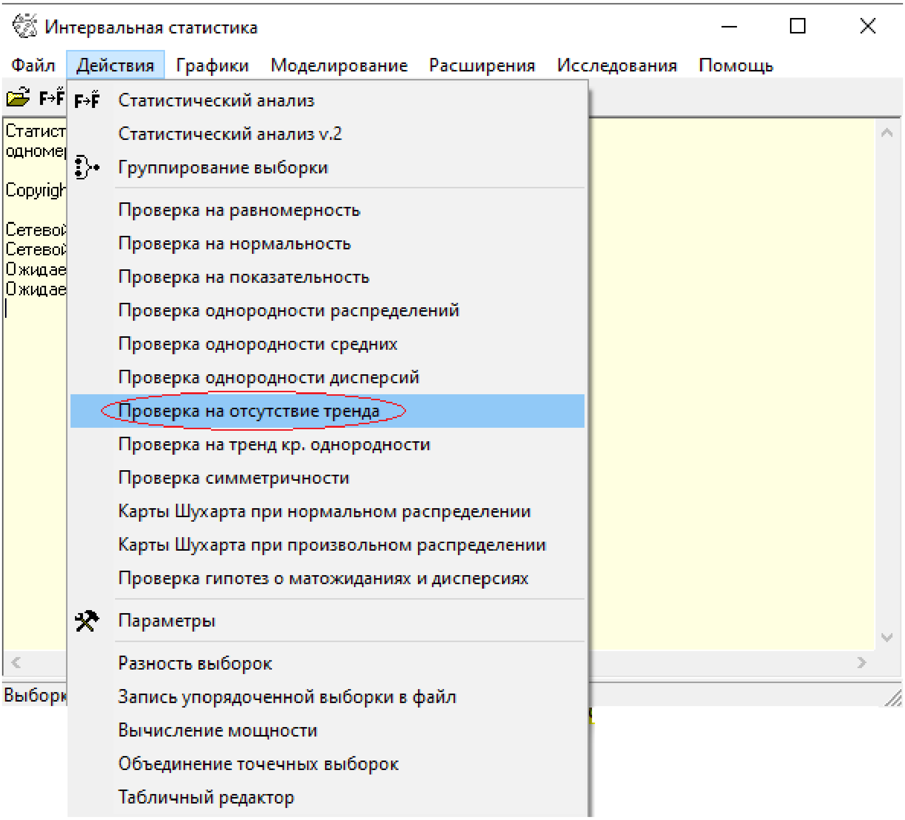


Рис. 2.33. Переход к использованию критериев проверки отсутствия тренда

Если предполагается использование параметрических критериев в условиях нарушения стандартного предположения о принадлежности выборки нормальному закону, то используя вкладку “Параметры” можно указать предполагаемый закон распределения и его параметры.

По команде “Проверить” для выбранных критериев вычисляются значения статистик. Для критериев с известными асимптотическими (предельными) распределениями статистик по этим распределениям вычисляются значения p_{value}. Результаты приводятся в открывшейся форме, представленной на рис. 2.35.

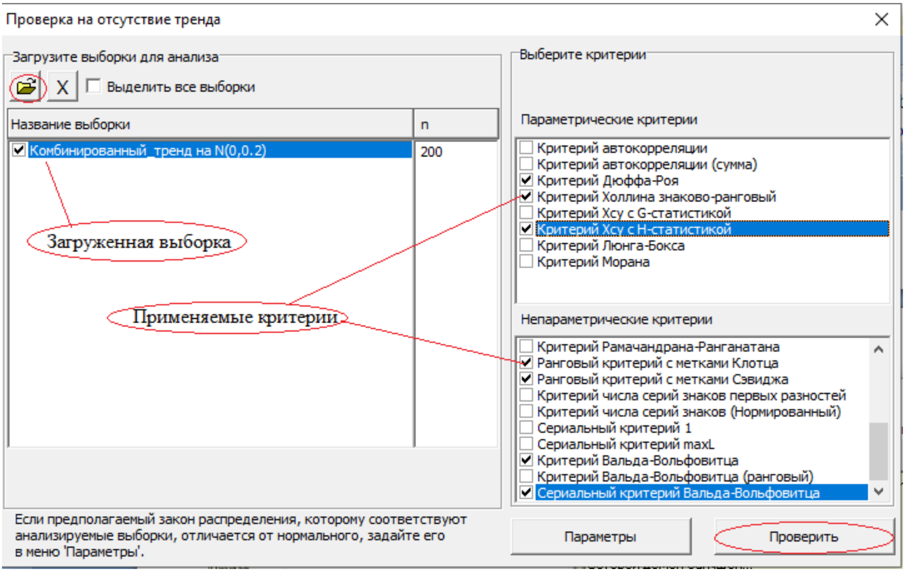


Рис. 2.34. Загрузка анализируемой выборки и выбор используемых критериев

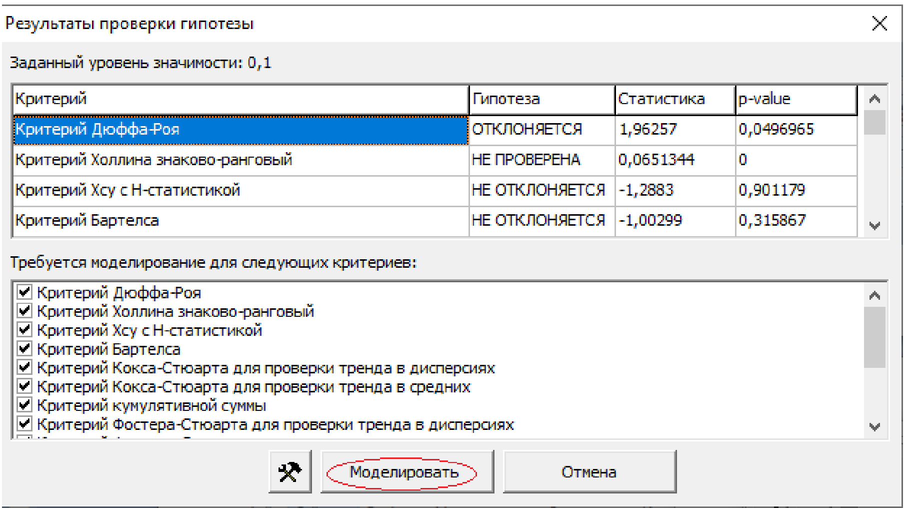
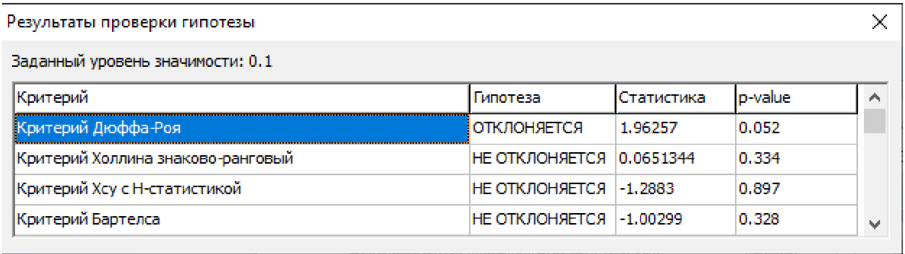


Рис. 2.35. Вычисление статистик, моделирование распределений статистик и оценка p_{value}

В этой форме приводятся значения статистик и значения p_{value}, в случае известных распределений статистик. Здесь же можно указать критерии, для которых оценки p_{value} будут вычисляться в результате статистического моделирования. При этом можно указать параметры для моделирования и запустить моделирование. Окончательные результаты проверки гипотезы по тем критериям, для которых оценки p_{value} уточнялись в процессе моделирования, приводятся в форме с возможностью прокрутки по всем критериям, приведенной на рис. 2.36.


Результаты проверки гипотезы

Заданный уровень значимости: 0.1

Критерий	Гипотеза	Статистика	p-value
Критерий Дюффа-Роя	ОТКЛОНЯЕТСЯ	1.96257	0.052
Критерий Холлина знаково-ранговый	НЕ ОТКЛОНЯЕТСЯ	0.0651344	0.334
Критерий Хсу с Н-статистикой	НЕ ОТКЛОНЯЕТСЯ	-1.2883	0.897
Критерий Бартелса	НЕ ОТКЛОНЯЕТСЯ	-1.00299	0.328

*Рис. 2.36.*Результаты проверки гипотезы

Пример временного ряда в ситуации отсутствии тренда и принадлежности величин нормальному закону приведен на рис. 4.37.

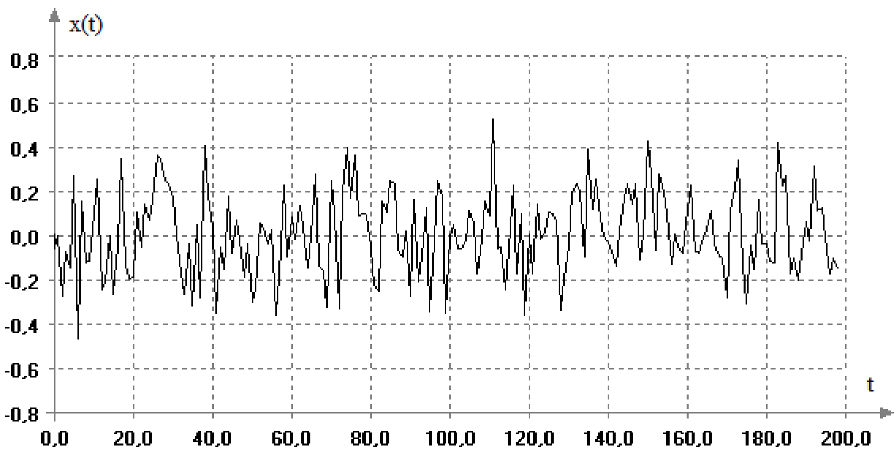


Рис. 2.37. Временной ряд измерений без тренда

Эта последовательность объёмом $n = 200$ сгенерирована в случае принадлежности выборки нормальному закону $N(0, 0.2)$ (с нулевым математическим ожиданием и $\sigma = 0.2$).

Результаты проверки гипотезы об отсутствии тренда в этой последовательности по некоторой совокупности критериев представлены в таблице 2.42. В первой колонке для p_{value} приведены значения, вычисленные по (известным) асимптотическим распределениям статистик критериев (в случае неизвестного распределения в ячейке таблицы стоит прочерк). В правой колонке представлены оценки p_{value}, полученные по результатам статистического моделирования при числе экспериментов $N = 10^5$, в том числе, и для тех критериев, для которых асимптотические распределения известны.

Таблица 2.42

Результаты проверки гипотезы об отсутствии тренда в последовательности, представленной на рис. 2.37

№ п/п	Критерий	Статистика	p_{value}	p_{value}
1	Критерий автокорреляции	1.97554	0.0482	0.04636
2	Критерий автокорреляции (модиф.)	1.98629	–	0.05316
3	Критерий Дюффа–Роя	1.97021	0.0488	0.04636
4	Критерий Холлина	0.15383	–	0.02574
5	Критерий Люнга–Бокса	1.90877	–	0.04636
6	Критерий Морана	1.98013	0.04775	0.04636
7	Критерий Бартелса	-2.10714	0.03511	0.03366
8	Критерий Кокса–Стюарта (в средних)	-1.09987	0.27140	0.27424
9	Критерий кумулятивной суммы	22	–	0.54908
10	Критерий Фостера–Стюарта (в средних)	0.64048	–	0.62556
11	Критерий инверсий	9532	–	0.37484
12	Критерий Вальда–Вольфовица	1.97373	0.04841	0.04656

По этим оценкам очевидна эквивалентность критерия автокорреляции, его вариантов (критериев Дюффа–Роя и Люнга–Бокса), критерия Морана и критерия Вальда–Вольфовица. Эти же критерии, а также критерии Холлина и Бартелса, при $\alpha = 0.05$ отклоняют гипотезу H_0 об отсутствии тренда.

В то же время при анализе этой же последовательности на отсутствие тренда с использованием критериев проверки однородности законов, критериев однородности средних и критериев однородности дисперсий проверяемая гипотеза об отсутствии тренда не отклоняется.

В качестве второго примера рассмотрим аналогичную последовательность, принадлежащую нормальному закону $N(0, 0.2)$, такого же объёма $n = 200$, на которую наложен линейный тренд вида $X_i = 0.5t_i + \xi_i$, где $\xi_i \in N(0, 0.2)$, $t_i = (i-1)\Delta t$, $\Delta t = 1/n$ (см. рис. 2.38).

Результаты проверки гипотезы об отсутствии тренда в этой последовательности по тому же ряду критериев представлены в таблице 2.43. В данном случае при $\alpha < 0.0147$ гипотеза об отсутствии тренда отклоняется по всем применяемым критериям за исключением критериев Кокса–Стюарта, Фостера–Стюарта, кумулятивной суммы и критерия инверсий.

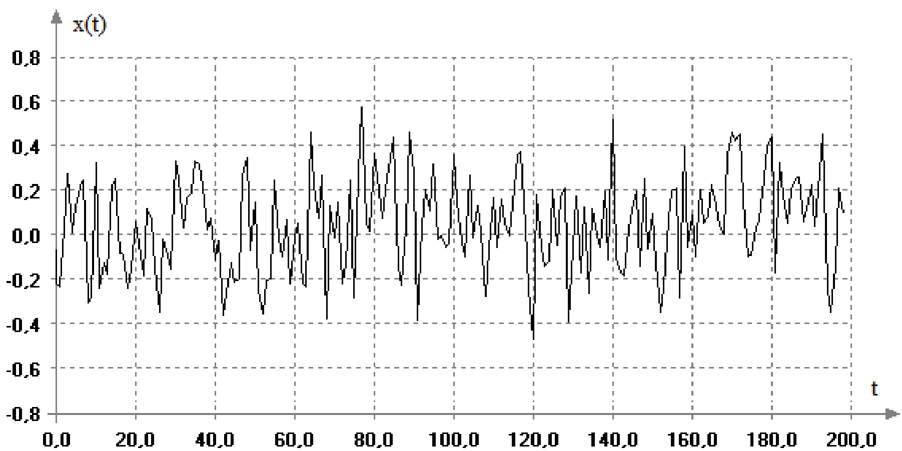


Рис. 2.38. Временной ряд с линейным трендом в математическом ожидании

При использовании для анализа этой же последовательности критериев однородности законов, критериев однородности средних и критериев однородности дисперсий проверяемая гипотеза об отсутствии тренда, как правило, отклоняется.

Таблица 2.43

Результаты проверки гипотезы об отсутствии тренда в последовательности, представленной на рис. 2.38

№ п/п	Критерий	Статистика	p_{value}	p_{value}
1	Критерий автокорреляции	2,69906	0,00695	0.007
2	Критерий автокорреляции (модиф.)	1,69903	–	0.0972
3	Критерий Дюффа–Роя	2,69191	0,00710	0.007
4	Критерий Холлина	0,17469	–	0.0113
5	Критерий Люнга–Бокса	2,63406	–	0.007
6	Критерий Морана	2,70547	0,00682	0.007
7	Критерий Бартелса	-2,44457	0,01450	0.0140
8	Критерий Кокса–Стюарта (в средних)	-1,35448	0,17558	0.1772
9	Критерий кумулятивной суммы	8	–	0.269
10	Критерий Фостера–Стюарта (в средних)	-0,320239	–	0.6303
11	Критерий инверсий	8797	–	0.0147
12	Критерий Вальда–Вольфовица	2,69469	0,00705	0.0071

В завершающем примере на последовательность объёмом $n = 200$, принадлежащую закону $N(0, 0.2)$, наложен комбинированный тренд вида:

$$X_i = at_i + b\sin(2k_1\pi t_i + \varphi_1) + \xi_i(1 + ct_i + d\sin(2k_2\pi t_i + \varphi_2)),$$

где $a = 0.2$, $b = 0.2$, $k_1 = 4$, $\varphi_1 = \pi/2$, $c = 0.2$, $d = 0$, $k_2 = 0$, $\varphi_2 = 0$. Пример реализации последовательности показан на рис. 2.39. В данном случае наличие тренда хорошо прослеживается визуально.

Результаты проверки некоторыми критериями гипотезы об отсутствии тренда в последовательности, показанной на рис.2.39, представлены в таблице 2.44. Не смотря на явное присутствие тренда, по шести критериям гипотеза об отсутствии тренда не будет отклонена. Эти результаты подтверждают показанную ранее невысокую мощность критерия кумулятивной суммы, критериев Фостера–Стюарта и критерия Кокса–Стюарта (для проверки отсутствия тренда в средних).

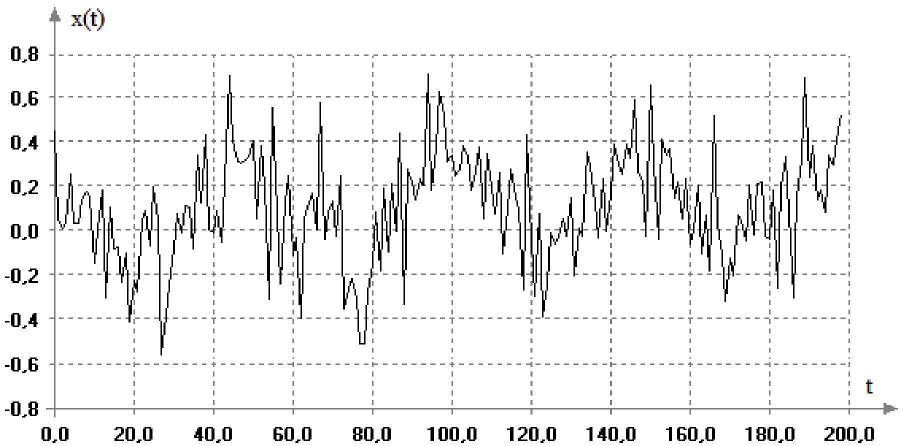


Рис. 2.39. Временной ряд с комбинированным трендом

Критерий Хсу с H-статистикой и критерий Клотца ориентированы на проверку гипотез об отсутствии тренда в дисперсии. А в данном случае тренд наложен, в основном, на математическое ожидание, чем и можно объяснить результаты, показанные данными критериями.

2.24. Выводы по разделу 2

Проведенные исследования свойств множества критериев, предназначенных для проверки гипотез об отсутствии тренда (и случайности) вызывают стойкое чувство неудовлетворенности.

Таблица 2.44

Результаты проверки гипотезы об отсутствии тренда в последовательности, представленной на рис. 2.39

№ п/п	Критерий	Статистика	p_{value}	p_{value}
1	Критерий Дюффа–Роя	5.31837	1.04704E-7	0
2	Критерий Холлина	0.35107	–	0
3	Критерий Хсу с H-статистикой	-0.637567	0.738122	0.73922
4	Критерий Бартелса	-4.95692	7.16198E-7	0
5	Критерий Кокса–Стюарта (в средних)	-1.02666	0.30458	0.30312
6	Критерий Кокса–Стюарта (в дисперсии)	-2.1928	0.0283215	0.02618
7	Критерий кумулятивной суммы	10	–	0.42116
8	Критерий Фостера–Стюарта (в средних)	-0.280453	–	0.66734
9	Критерий Фостера–Стюарта (в дисперсии)	-0.960718	–	0.25884
10	Критерий инверсий	8808	–	0.01566
11	Критерий Клотца	-0.577723	0.563451	0.56454
12	Критерий Сэвиджа	1.97082	0.0487446	0.04802
13	Критерий Вальда–Вольфовица	5.33141	0	0
14	Сериальный критерий Вальда–Вольфовица	-3.68623	0.000227602	0.00016

Во-первых, нельзя дать однозначного ответа на вопрос, какой критерий лучше всего использовать в целях проверки данного вида гипотез. Относительно разного вида конкурирующих гипотез критерии по мощности могут упорядочиваться различным образом.

Во-вторых, применение совокупности критериев для анализа одной и той же последовательности (для проверки одной и той же гипотезы) может приводить к диаметрально противоположным выводам о результатах проверки. В принципе, это естественно, так как статистики критериев опираются на различные меры отклонения анализируемой последовательности от проверяемой гипотезы H_0. А причиной является низкая и различающаяся мощность критериев относительно различных альтернатив.

В такой ситуации для большей объективности статистических выводов можно рекомендовать воспользоваться рядом критериев, обладающих определёнными достоинствами (большей мощностью относительно некоторых конкурирующих гипотез). Но тогда возникает вопрос, что делать, когда выводы по отобранным критериям приводят к той же спорной ситуации? По-видимому, потребуется дополнительный анализ с использованием каких-то других критериев.

В качестве таких критериев для сравнительного анализа частей анализируемой последовательности можно применять критерии однородности законов, однородности математических ожиданий или однородности дисперсий.

В частности, при анализе 3-х рассмотренных в п. 2.24 последовательностей критерии однородности законов дали однозначные ответы: об однородности рассматриваемых частей последовательности в первом случае (об отсутствии какого-либо тренда) и об отклонении однородности во втором и третьем случаях, что говорит о возможном наличии тренда.

Вопросы применения критериев однородности для проверки гипотез об отсутствии тренда рассмотрены в последующих разделах руководства.

3. КРИТЕРИИ ОДНОРОДНОСТИ ПРИ ПРОВЕРКЕ ГИПОТЕЗ ОБ ОТСУТСТВИИ ТРЕНДА

Очевидно, что проверка гипотез о том, что части выборки принадлежат одному и тому же закону, или имеют одно и то же математическое ожидание, или одну и ту же дисперсию, – это эквивалентно проверке гипотез об отсутствии какого-либо тренда в наблюдаемой последовательности, связанного с изменением закона, или математического ожидания, или характеристик рассеяния. Отсюда следует, что для проверки гипотезы об отсутствии тренда могут использоваться многочисленные критерии проверки однородности законов, однородности средних, однородности дисперсий [136]. Рекомендации по применению множества критериев однородности, учитывающие реальные свойства этих критериев, наиболее полно отражены в руководстве [134].

С необходимостью решения задач проверки гипотез о принадлежности двух (или более) выборок случайных величин одной и той же генеральной совокупности (проверки однородности) постоянно сталкиваются при анализе случайных ошибок средств измерений, при статистическом управлении качеством процессов. Такая задача естественно возникает при проверке средств измерений, когда пытаются убедиться в том, что закон распределения случайных ошибок измерений не претерпел существенных изменений по истечении некоторого интервала времени. При обработке экспериментальных данных очень часто эту задачу приходится решать технологам, медикам, биологам.

В программной системе [148] предусмотрена возможность разбиения анализируемой выборки на произвольные части, к которым могут применяться любые критерии однородности, упомянутые ниже.

Задача проверки однородности двух выборок формулируется следующим образом. Пусть имеются две упорядоченные по возрастанию выборки размером m и n:

$$x_1 < x_2 < \ldots < x_m \quad \text{и} \quad y_1 < y_2 < \ldots < y_n .$$

Для определенности обычно полагают, что $m \le n$, но это совсем необязательно. Проверяется гипотеза о том, что две выборки извлечены из одной и той же генеральной совокупности, т. е. H_0:

$F(x) = G(x)$ при любом x.

Как правило, на практике используется либо критерий Смирнова [101], либо критерий Лемана–Розенблатта [101, 43, 79]. Предпочтительность использования данных критериев для проверки однородности обсуждалась в [144]. В русскоязычной литературе практически не упоминается применение двухвыборочного критерия Андерсона–Дарлинга [78] (Андерсона–Дарлинга–Петита) или, тем более, об использовании многовыборочного варианта критерия Андерсона–Дарлинга [83].

Реальные свойства критериев проверки однородности и сравнительный анализ подробно рассмотрены в [134], поэтому в данном руководстве мы ограничимся кратким описанием соответствующих критериев.

3.1. Критерий Смирнова

Критерий однородности Смирнова предложен в работе [147]. Предполагается, что функции распределения $F(x)$ и $G(x)$ являются непрерывными. Статистика критерия Смирнова измеряет расстояние между эмпирическими функциями распределения, построенными по выборкам

$$D_{m,n} = \sup_x \left| G_m(x) - F_n(x) \right|.$$

При практическом использовании критерия значение статистики $D_{m,n}$ рекомендуется вычислять в соответствии с соотношениями [101]:

$$D_{m,n}^{+} = \max_{1 \le r \le m} \left[\frac{r}{m} - F_n(x_r) \right] = \max_{1 \le s \le n} \left[G_m(y_s) - \frac{s-1}{n} \right],$$

$$D_{m,n}^{-} = \max_{1 \le r \le m} \left[F_n(x_r) - \frac{r-1}{m} \right] = \max_{1 \le s \le n} \left[\frac{s}{n} - G_m(y_s) \right],$$

$$D_{m,n} = \max\left(D_{m,n}^{+}, D_{m,n}^{-} \right).$$

Если гипотеза H_0 справедлива, то при неограниченном увеличении объемов выборок [101] $\lim_{m\to\infty} P\left\{\sqrt{\frac{mn}{m+n}}D_{m,n} < S\right\} = K(S)$, т. е. статистика

$$S_C = \sqrt{\frac{mn}{m+n}}D_{m,n} \tag{3.1}$$

в пределе подчиняется распределению Колмогорова $K(S)$ [101] с функцией распределения

$$K(s) = \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2 s^2} . \tag{3.2}$$

Дискретное распределение статистики (3.1) при небольших и умеренных значениях m и n существенно отличается от предельного $K(S)$ и заметно сдвинуто от него влево.

В этой связи предложена [111, 47] следующая простая модификация статистики (3.1):

$$S_{CM} = \sqrt{\frac{mn}{m+n}}\left(D_{m,n} + \frac{m+n}{4.6mn}\right), \tag{3.3}$$

у которой практически отсутствует последний недостаток.

3.2. Критерий Лемана–Розенблатта

Критерий однородности Лемана–Розенблатта представляет собой критерий типа ω^2. Критерий предложен в работе [43] и исследован в [79]. Статистика критерия имеет вид [101]

$$T = \frac{mn}{m+n}\int_{-\infty}^{\infty}\left[G_m(x) - F_n(x)\right]^2 dH_{m+n}(x),$$

где $H_{m+n}(x) = \frac{m}{m+n}G_m(x) + \frac{n}{m+n}F_n(x)$ – эмпирическая функция распределения, построенная по вариационному ряду объединения двух выборок. Статистика T используется в форме [101]

$$T = \frac{1}{mn(m+n)}\left[n\sum_{i=1}^{n}\left(r_i - i\right)^2 + m\sum_{j=1}^{m}\left(s_j - j\right)^2 \right] - \frac{4mn-1}{6(m+n)}, \qquad (3.4)$$

где r_i – порядковый номер (ранг) y_i; s_j – порядковый номер (ранг) x_j в объединенном вариационном ряде.

В [79] было показано, что статистика (3.4) в пределе распределена как $a1(t)$:

$$\lim_{\substack{m\to\infty \\ n\to\infty}} P\left\{T < t\right\} = a1(t) .$$

Функция распределения $a1(t)$ имеет вид [101]:

$$a1(t) = \frac{1}{\sqrt{2s}} \sum_{j=0}^{\infty} \frac{\Gamma(j+1/2)\sqrt{4j+1}}{\Gamma(1/2)\Gamma(j+1)} \exp\left\{ -\frac{(4j+1)^2}{16t} \right\} \times$$

$$\times \left\{ I_{-\frac{1}{4}}\left[\frac{(4j+1)^2}{16t}\right] - I_{\frac{1}{4}}\left[\frac{(4j+1)^2}{16t}\right] \right\}, \qquad (3.5)$$

где $I_{-\frac{1}{4}}(\cdot), I_{\frac{1}{4}}(\cdot)$ – модифицированные функции Бесселя вида

$$I_\nu(z) = \sum_{k=0}^{\infty} \frac{\left(\frac{z}{2}\right)^{\nu+2k}}{\Gamma(k+1)\Gamma(k+\nu+1)}, \ |z| < \infty, \ |\arg z| < \pi .$$

В отличие от критерия Смирнова распределение статистики T быстро сходится к предельному $a1(T)$ [101].

3.3. Критерий Андерсона–Дарлинга

Двухвыборочный критерий Андерсона–Дарлинга (критерий однородности) рассмотрен в работе [78]. Статистика критерия определяется выражением

$$A^2 = \frac{mn}{m+n} \int\limits_{-\infty}^{\infty} \frac{[G_m(x) - F_n(x)]^2}{(1 - H_{m+n}(x)) H_{m+n}(x)} dH_{m+n}(x) . \quad (3.6)$$

Для выборок непрерывных случайных величин выражение для этой статистики принимает простой вид [78]

$$A^2 = \frac{1}{mn} \sum_{i=1}^{m+n-1} \frac{\left(M_i(m+n) - mi\right)^2}{i(m+n-i)} , \quad (3.7)$$

где M_i – число элементов первой выборки, меньших или равных i-му элементу вариационного ряда объединенной выборки.

Предельным распределением статистики (3.7) при справедливости проверяемой гипотезы H_0 является то же самое распределение $a2(t)$ [78], которое является предельным для статистики критерия согласия Андерсона–Дарлинга [101]. Функция распределения $a2(t)$, имеет вид [101]

$$a2(t) = \frac{\sqrt{2\pi}}{t} \sum_{j=0}^{\infty} (-1)^j \frac{\Gamma(j+1/2)(4j+1)}{\Gamma(1/2)\Gamma(j+1)} \exp\left\{ -\frac{(4j+1)^2 \pi^2}{8t} \right\} \times$$

$$\times \int\limits_0^{\infty} \exp\left\{ \frac{t}{8(y^2+1)} - \frac{(4j+1)^2 \pi^2 y^2}{8t} \right\} dy . \quad (3.8)$$

Критерий Андерсона–Дарлинга, как правило, обладает несколько большей мощностью по сравнению с критерием Лемана–Розенблатта.

3.4. Многовыборочный критерий Андерсона–Дарлинга

Многовыборочный вариант критерия однородности Андерсона–Дарлинга предложен в [83].

Задача проверки однородности k выборок формулируется следующим образом. Пусть x_{ij} j-е наблюдение i-й выборки $j = \overline{1, n_i}$, $i = \overline{1, k}$. Предположим, что i-й выборке соответствует непрерывная функция распределения $F_i(x)$. Необходимо проверить гипотезу вида

$H_0: F_1(x) = F_2(x) = \ldots = F_k(x)$ без указания общего для них закона распределения. Обозначим эмпирическую функцию распределения, соответствующую i-й выборке, как $F_{in_i}(x)$, а эмпирическую функцию распределения, соответствующую объединённой выборке объёмом $n = \sum_{i=1}^{k} n_i$, как $H_n(x)$. Статистика k-выборочного критерия Андерсона–Дарлинга определяется выражением

$$A_{kn}^2 = \sum_{i=1}^{k} n_i \int_{B_n} \frac{[F_{in_i}(x) - H_n(x)]^2}{\left(1 - H_n(x)\right) H_n(x)} dH_n(x), \tag{3.9}$$

где $B_n = \{x \in R: H_n(x) < 1\}$. Для $k = 2$ соотношение (3.9) сводится к (3.6). В предположении о непрерывности $F_i(x)$, упорядочив объединённую выборку $Z_1 \le Z_2 \le \ldots \le Z_n$, непосредственно из (3.9) можно получить простое выражение для вычисления статистики [83]:

$$A_{kn}^2 = \frac{1}{n} \sum_{i=1}^{k} \frac{1}{n_i} \sum_{j=1}^{n-1} \frac{\left(nM_{ij} - jn_i\right)^2}{j(n-j)}, \tag{3.10}$$

где M_{ij} – число элементов в i-й выборке, которые не больше чем Z_j. Проверяемая гипотеза H_0 отклоняется при больших значениях статистики (3.10).

В работе [83] таблица верхних процентных точек представлена не для статистики (3.10), а для статистики вида:

$$T_{kn} = \frac{A_{kn}^2 - (k-1)}{\sqrt{D[A_{kn}^2]}}. \tag{3.11}$$

Дисперсия статистики A_{kn}^2 определяется выражением [83]

$$D\left[A_{kn}^2\right] = \frac{an^3 + bn^2 + cn + d}{(n-1)(n-2)(n-3)} \tag{3.12}$$

при

$$a = (4g-6)(k-1)+(10-6g)H\ ,$$

$$b = (2g-4)k^2+8hk+(2g-14h-4)H-8h+4g-6\,,$$

$$c = (6h+2g-2)k^2+(4h-4g+6)k+(2h-6)H+4h\ ,$$

$$d = (2h+6)k^2-4hk\,,$$

где

$$H = \sum_{i=1}^{k}\frac{1}{n_i}\,,\quad h = \sum_{i=1}^{n-1}\frac{1}{i}\,,\quad g = \sum_{i=1}^{n-2}\sum_{j=i+1}^{n-1}\frac{1}{(n-i)\,j}\,.$$

В результате исследований распределений статистики (3.11) методами статистического моделирования (при $n_i = 1000$ и числе экспериментов имитационного моделирования $N = 10^6$) таблица критических значений, приведенная в [83], была уточнена и расширена (см. таблицу 3.1).

Одновременно для предельных распределений статистик (для $k = 2 \div 11$) были построены приближенные модели законов для распределений статистики (3.11) [44, 112]. Хорошими моделями оказались законы семейства бета-распределений III рода с плотностью

$$f(x) = \frac{\theta_2^{\theta_0}}{\theta_3 \mathrm{B}(\theta_0,\theta_1)} \frac{\left(\dfrac{x-\theta_4}{\theta_3}\right)^{\theta_0-1}\left(1-\dfrac{x-\theta_4}{\theta_3}\right)^{\theta_1-1}}{\left[1+(\theta_2-1)\dfrac{x-\theta_4}{\theta_3}\right]^{\theta_0+\theta_1}} \qquad (3.13)$$

при конкретных значениях параметров этого закона $\mathrm{B}_{\mathrm{III}}(\theta_0,\theta_1,\theta_2,\theta_3,\theta_4)$, найденными по выборкам статистик объёмом $N = 10^6$, полученным в результате моделирования.

Представленные в таблице 3.2 модели $\mathrm{B}_{\mathrm{III}}(\theta_0,\theta_1,\theta_2,\theta_3,\theta_4)$, с приведенными значениями параметров, позволяют по значениям статистики, вычисленным в соответствии с соотношением (3.11), находить оценки p_{value} при соответствующем числе k сравниваемых выборок.

Таблица 3.1

Уточненные верхние критические значения $T_{kn}^2(\alpha)$ статистики (3.11)

k	$1-\alpha$				
	0.75	0.90	0.95	0.975	0.99
2	0.325	1.228	1.966	2.731	3.784
3	0.439	1.300	1.944	2.592	3.429
4	0.491	1.321	1.925	2.511	3.277
5	0.523	1.331	1.900	2.453	3.153
6	0.543	1.333	1.885	2.410	3.078
7	0.557	1.337	1.870	2.372	3.017
8	0.567	1.335	1.853	2.344	2.970
9	0.577	1.334	1.847	2.323	2.927
10	0.582	1.3345	1.838	2.306	2.899
11	0.589	1.332	1.827	2.290	2.867
∞	0.674	1.282	1.645	1.960	2.326

Таблица 3.2

Модели предельных распределений статистики (2.11)

k	Модель
2	B_{III} (3.1575, 2.8730, 18.1238, 15.0000, –1.1600)
3	B_{III} (3.5907, 4.5984, 7.8040, 14.1310, –1.5000)
4	B_{III} (4.2657, 5.7035, 5.3533, 12.8243, –1.7500)
5	B_{III} (6.2992, 6.5558, 5.6833, 13.010, –2.0640)
6	B_{III} (6.7446, 7.1047, 5.0450, 12.8562, –2.2000)
7	B_{III} (6.7615, 7.4823, 4.0083, 11.800, –2.3150)
8	B_{III} (5.8057, 7.8755, 2.9244, 10.900, –2.3100)
9	B_{III} (9.0736, 7.4112, 4.1072, 10.800, –2.6310)
10	B_{III} (10.2571, 7.9758, 4.1383, 11.186, –2.7988)
11	B_{III} (10.6848, 7.5950, 4.2041, 10.734, –2.8400)
∞	N(0.0, 1.0)

3.5. Критерии однородности Жанга

Кроме рассмотренных выше, можно указать непараметрические критерии, предложенные в работах Жанга [97, 98, 99], которые дают возможность сравнивать $k \geq 2$ выборок [97]. Предложенные Жангом критерии являются развитием критериев однородности Смирнова, Лемана–Розенблатта и Андерсона–Дарлинга.

Критерии согласия Жанга [97] показывают некоторое преимущество в мощности по сравнению с критериями согласия Крамера–Мизеса–Смирнова и Андерсона–Дарлинга [130], а определённым недостатком, ограничивающим их применение, является зависимость распределений статистик от объёмов выборок.

Аналогичным недостатком (зависимостью от объёмов сравниваемых выборок) обладают и варианты критериев Жанга, предназначенные для проверки однородности законов. В случае применения указанных критериев автор рекомендует для оценивания p_{value} использовать метод Монте–Карло [97].

Задача моделирования распределений статистик критериев однородности Жанга, по сравнению с аналогичной задачей для критериев согласия, оказывается много проще, так как приходится моделировать распределения статистик $G(S|H_0)$ критериев в случае принадлежности анализируемых выборок равномерному закону. Это делает применение критериев очень перспективным.

Пусть $x_{i1}, x_{i2}, \ldots, x_{in_i}$ упорядоченные выборки непрерывных случайных величин с функциями распределения $F_i(x)$, $(i = \overline{1,k})$ и пусть $X_1 < X_2 < \ldots < X_n$, где $n = \sum_{i=1}^{k} n_i$, объединённая упорядоченная выборка.

Обозначим R_{ij} ранг j-го упорядоченного наблюдения x_{ij} i-й выборки в объединённой выборке. Пусть $X_0 = -\infty$, $X_{n+1} = +\infty$, а ранги $R_{i,0} = 1$, $R_{i,n_i+1} = n+1$.

В критериях используется модификация эмпирической функции распределения $\hat{F}(t)$, принимающая в точках разрыва X_m, $m = \overline{1,n}$, значения $\hat{F}(X_m) = (m - 0.5) / n$ [97].

Статистика Z_{K} критерия однородности Жанга имеет вид [97]:

$$Z_{\text{K}} = \max_{1 \le m \le n} \left\{ \sum_{i=1}^{k} n_i \left[F_{i,m} \ln \frac{F_{i,m}}{F_m} + \left(1 - F_{i,m}\right) \ln \frac{1 - F_{i,m}}{1 - F_m} \right] \right\}, \qquad (3.14)$$

где $F_m = \hat{F}(X_m)$, так что $F_m = (m - 0.5)/n$, а вычисление $F_{i,m} = \hat{F}_i(X_m)$ осуществляется следующим образом. В начальный момент значения $j_i = 0$, $i = \overline{1,k}$. Если $R_{i,j_i+1} = m$, то $j_i := j_i + 1$ и $F_{i,m} = (j_i - 0.5)/n_i$, в противном случае если $R_{i,j_i} < m < R_{i,j_i+1}$, то $F_{i,m} = j_i / n_i$.

Критерий *правосторонний*: проверяемая гипотеза H_0 отклоняется при *больших* значениях статистики (3.14).

Статистика Z_{A} критерия однородности k выборок определяется выражением [97]:

$$Z_{\text{A}} = -\sum_{m=1}^{n} \sum_{i=1}^{k} n_i \frac{F_{i,m} \ln F_{i,m} + \left(1 - F_{i,m}\right) \ln\left(1 - F_{i,m}\right)}{(m - 0.5)(n - m + 0.5)}, \qquad (3.15)$$

где F_m и $F_{i,m}$ вычисляются, как определено выше.

Критерий *левосторонний*: проверяемая гипотеза H_0 отклоняется при *малых* значениях статистики (3.15).

Статистика Z_C критерия однородности k выборок определяется выражением [97]:

$$Z_C = \frac{1}{n} \sum_{i=1}^{k} \sum_{j=1}^{n_i} \ln\left(\frac{n_i}{j - 0.5} - 1 \right) \ln\left(\frac{n}{R_{i,j} - 0.5} - 1 \right). \qquad (3.16)$$

Этот критерий также *левосторонний*: проверяемая гипотеза H_0 отклоняется при *малых* значениях статистики (3.16).

Сравнивая мощности критериев Жанга с мощностью критериев Андерсона–Дарлинга, Лемана–Розенблатта и Смирнова, можно обратить внимание на то, что критерии Жанга имеют преимущество в мощности относительно альтернатив, связанных с изменением характеристик масштаба, и уступают в мощности при альтернативах сдвига.

Отсутствие информации о законах распределения статистик и таблиц критических значений в современных условиях не является серьёзным недостатком критериев Жанга, так как в программном обеспечении, осуществляющем поддержку применения критериев, несложно организовать вычисление достигнутых уровней значимостей p_{value}, используя методы статистического моделирования. Поэтому интерес к данным критериям, к сравнению их свойств с традиционными в последнее время возрастает [8].

3.6. Использование двухвыборочных критерии при анализе k выборок

Различные подходы к построению k-выборочных аналогов критериев Смирнова, Лемана–Розенблатта и Андерсона–Дарлинга рассматривались в работе [26]. k-выборочный вариант критерия Колмогорова–Смирнова, основанный на таком подходе, был построен в [10] и рассматривается в последовательных изданиях книги [11]. На использовании такого же подхода построен k-выборочный критерий Андерсона–Дарлинга [83], рассмотренный выше, для которого нами были построены модели предельных распределений, представленные, в том числе, в [112, 44]. В этих критериях, так же как и в критериях однородности Жанга, строится объединённая выборка, а статистики измеряют отклонение эмпирических распределений отдельных выборок от эмпирического распределения, построенного по совокупности анализируемых выборок.

Другой возможный путь заключается в том [62, 63], что для анализа k выборок к каждой паре из них применяется двухвыборочный критерий со статистикой S (всего $(k-1)k/2$ вариантов), а решение принимают по совокупности результатов. В качестве статистики такого k-выборочного критерия (в случае использования правостороннего двухвыборочного критерия) можно рассмотреть, например, статистику вида

$$S_{\max} = \max_{\substack{1 \le i \le k \\ i < j \le k}} \{S_{i,j}\}, \qquad (3.17)$$

где $S_{i,j}$ – значения статистик используемого двухвыборочного критерия, полученные при анализе i-й и j-й выборок.

Проверяемая гипотеза H_0 будет отклоняться при **больших** значениях статистики $S_{\max}$.

Преимуществом такого рода критерия является и то, что в результате будет определена пара выборок, различие между которыми оказывается наиболее значимым с позиций используемого двухвыборочного критерия.

В качестве $S_{i,j}$ можно использовать статистики двухвыборочных критериев Смирнова (лучше в модифицированном виде), Лемана–Розенблата, Андерсона–Дарлинга. В этом случае распределения соответствующих статистик $S_{\max}$ сходятся к некоторым предельным, модели которых могут быть найдены по результатам статистического моделирования.

Конечно, в критериях такого рода можно использовать и двухвыборочный вариант статистики Z_{K} критерия Жанга, однако тогда распределение статистики $S_{\max}$ будет зависеть от объёмов сравниваемых выборок. Следовательно, применение критерия будет возможно лишь при использовании компьютерных технологий, обеспечивающих исследование распределения статистики критерия в интерактивном режиме.

При использовании левосторонних двухвыборочных критериев Жанга со статистиками Z_{A} и Z_C статистика k-выборочного критерия примет вид:

$$S_{\min} = \min_{\substack{1 \le i \le k \\ i < j \le k}} \{S_{i,j}\}. \tag{3.18}$$

В этом случае проверяемая гипотеза H_0 будет отклоняться при **малых** значениях статистики $S_{\min}$. Однако строить варианты критериев, предусматривающие использование двухвыборочных критериев Жанга, большого смысла нет из-за зависимости в таком случае распределений статистик (3.17) и (3.18) от объемов выборок.

3.6.1. k–выборочный критерий Смирнова (max)

В качестве $S_{i,j}$ в (3.17) в этом случае будет рассматриваться модификация статистики Смирнова (3.3), распределение которой всегда ближе к предельному распределению Колмогорова $K(S)$. Статистику $S_{\max}$ в этом случае будем обозначать как $S_{\max}^{Sm}$.

При равных объёмах сравниваемых выборок распределения статистики $S_{\max}^{Sm}$ (как и в двухвыборочном варианте) обладают существенной дискретностью и отличаются от асимптотических (предельных) распределений. Если есть такая возможность, то предпочтительней в качестве n_i выбирать взаимно простые числа, тогда распределения $G(S|H_0)$ статистики $S_{\max}^{Sm}$ практически не будут отличаться от асимптотических.

Верхние критические значения для статистики $S_{\max}^{Sm}$, построенные по эмпирическим распределениям статистик, полученным методом Монте–Карло при количестве имитационных экспериментов $N = 10^6$, представлены в таблице 3.3, а в таблице 3.4 приведены построенные модели асимптотических (предельных) распределений статистики $S_{\max}^{Sm}$ при числе сравниваемых выборок $k = 3 \div 11$ [134, 135, 62, 63].

Представленные в таблице 3.4 модели $B_{III}(\theta_0, \theta_1, \theta_2, \theta_3, \theta_4)$ бета-распределения 3-го рода (2.13) с приведенными значениями параметров, позволяют по значениям статистики, вычисленным в соответствии с соотношением (3.17) с использованием в качестве $S_{i,j}$ статистики Смирнова (3.1) или её модификации (3.3), находить оценки p_{value} при соответствующем числе k сравниваемых выборок.

Таблица 3.3

Верхние критические значения статистики $S_{\max}^{Sm}$ Смирнова

k	$1-\alpha$				
	0.75	0.90	0.95	0.975	0.99
2	1.0192	1.2238	1.3581	1.4802	1.6276
3	1.2059	1.4006	1.5266	1.6396	1.7782
4	1.3194	1.5070	1.6278	1.7372	1.8701
5	1.4002	1.5822	1.6997	1.8057	1.9329
6	1.4620	1.6396	1.7545	1.8573	1.9833
7	1.5127	1.6867	1.7989	1.9009	2.0245
8	1.5546	1.7264	1.8362	1.9362	2.0596
9	1.5911	1.7598	1.8682	1.9672	2.0879
10	1.6224	1.7893	1.8969	1.9940	2.1148
11	1.6506	1.8156	1.9217	2.0182	2.1371

Таблица 3.4

Модели предельных распределений статистики $S_{\max}^{Sm}$

k	Модель
2	$K(S)$
3	B_{III} (6.3274, 6.6162, 2.8238, 2.4073, 0.4100)
4	B_{III} (7.2729, 7.2061, 2.6170, 2.3775, 0.4740)
5	B_{III} (7.1318, 7.3365, 2.4813, 2.3353, 0.5630)
6	B_{III} (7.0755, 8.0449, 2.3163, 2.3818, 0.6320)
7	B_{III} (7.7347, 8.6845, 2.3492, 2.4479, 0.6675)
8	B_{III} (7.8162, 8.9073, 2.2688, 2.4161, 0.7120)
9	B_{III} (7.8436, 8.8805, 2.1696, 2.3309, 0.7500)
10	B_{III} (7.8756, 8.9051, 2.1977, 2.3280, 0.7900)
11	B_{III} (7.9122, 9.0411, 2.1173, 2.2860, 0.8200)

3.6.2. k–выборочный критерий Лемана–Розенблатта (max)

В качестве $S_{i,j}$ в статистике $S_{\max}^{LR}$ вида (3.17) в этом случае используется статистика (3.4) Лемана–Розенблатта.

Верхние критические значения для статистики $S_{\max}^{LR}$, построенные по результатам статистического моделирования при количестве имитационных экспериментов $N = 10^6$, представлены в таблице 3.5.

Построенные модели асимптотических (предельных) распределений статистики $S_{\max}^{LR}$ при числе сравниваемых выборок $k = 3 \div 11$ представлены в таблице 3.6 [134, 135, 62, 63].

Таблица 3.5

Верхние критические значения статистики $S_{\max}^{LR}$ Лемана–Розенблатта

k	$1-\alpha$				
	0.75	0.90	0.95	0.975	0.99
2	0.2094	0.3473	0.4614	0.5806	0.7435
3	0.3306	0.4995	0.6283	0.7581	0.9308
4	0.4206	0.6050	0.7413	0.8770	1.0550
5	0.4924	0.6856	0.8267	0.9653	1.1429
6	0.5521	0.7512	0.8959	1.0365	1.2175
7	0.6037	0.8072	0.9524	1.0947	1.2774
8	0.6481	0.8550	1.0015	1.1444	1.3298
9	0.6876	0.8976	1.0457	1.1902	1.3781
10	0.7237	0.9355	1.0858	1.2303	1.4171
11	0.7563	0.9703	1.1214	1.2655	1.4535

В данном случае наилучшими моделями оказались распределения Sb–Джонсона с плотностью

$$f(x) = \frac{\theta_1 \theta_2}{\sqrt{2\pi}(x-\theta_3)(\theta_2+\theta_3-x)} \exp\left\{-\frac{1}{2}\left[\theta_0 - \theta_1 \ln \frac{x-\theta_3}{\theta_2+\theta_3-x}\right]^2\right\}$$

при конкретных значениях параметров этого закона, обозначенного в таблице 3.6 как $Sb(\theta_0, \theta_1, \theta_2, \theta_3, \theta_4)$. Представленные модели позволяют по значениям статистики $S_{\max}^{LR}$ при соответствующем числе k сравниваемых выборок находить оценки p_{value}.

Таблица 3.6

Модели предельных распределений статистики $S_{\max}^{LR}$

k	Модель
2	$a1(t)$
3	Sb(3.2854, 1.2036, 3.0000, 0.0215)
4	Sb(2.5801, 1.2167, 2.2367, 0.0356)
5	Sb(3.1719, 1.4134, 3.1500, 0.0320)
6	Sb(2.9979, 1.4768, 2.9850, 0.0380)
7	Sb(3.2030, 1.5526, 3.4050, 0.0450)
8	Sb(3.2671, 1.6302, 3.5522, 0.0470)
9	Sb(3.4548, 1.7127, 3.8800, 0.0490)
10	Sb(3.4887, 1.7729, 3.9680, 0.0510)
11	Sb(3.4627, 1.8168, 3.9680, 0.0544)

3.6.3. k–выборочный критерий Андерсона–Дарлинга (max)

В статистике $S_{\max}^{AD}$ вида (3.17) качестве $S_{i,j}$ используется статистика (3.7) Андерсона–Дарлинга.

Верхние критические значения для статистики $S_{\max}^{AD}$, построенные по результатам статистического моделирования при количестве имитационных экспериментов $N = 10^6$, представлены в таблице 3.7.

Для распределений $G(S_{\max}^{AD} | H_0)$ также построены [134, 135, 62, 63] модели асимптотических (предельных) распределений статистики $S_{\max}^{AD}$ для числа сравниваемых выборок $k = 3 \div 11$, которые представлены в таблице 3.8. В этом случае лучшими моделями оказались бета-распределения 3-го рода (3.13), которые в виде $\mathrm{B_{III}}(\theta_0, \theta_1, \theta_2, \theta_3, \theta_4)$ с конкретными значениями параметров приведены в таблице 3.8 и могут использоваться для оценки p_{value} при k сравниваемых выборках.

Таблица 3.7

Верхние критические значения статистики $S_{\max}^{AD}$ Андерсона–Дарлинга

k	$1-\alpha$				
	0.75	0.90	0.95	0.975	0.99
2	1.2479	1.9330	2.4924	3.0775	3.8781
3	1.8535	2.6796	3.31215	3.95176	4.7924
4	2.2990	3.1966	3.8682	4.5368	5.4076
5	2.6514	3.5948	4.2877	4.9686	5.8472
6	2.9431	3.9187	4.6292	5.3175	6.2089
7	3.1943	4.1950	4.9097	5.6063	6.5118
8	3.4135	4.4292	5.1501	5.8531	6.7710
9	3.6094	4.6395	5.3660	6.0733	7.0076
10	3.7867	4.8270	5.5616	6.2720	7.2042
11	3.9466	4.9974	5.7384	6.4512	7.3837

Таблица 3.8

Модели предельных распределений статистики $S_{\max}^{AD}$

k	Модель
2	$a2(t)$
3	B_{III} (4.4325, 2.7425, 12.1134, 8.500, 0.1850)
4	B_{III} (5.2036, 3.2160, 10.7792, 10.000, 0.2320)
5	B_{III} (5.7527, 3.3017, 9.7365, 10.000, 0.3000)
6	B_{III} (5.5739, 3.4939, 7.7710, 10.000, 0.3750)
7	B_{III} (6.4892, 3.6656, 8.0529, 10.500, 0.3920)
8	B_{III} (6.3877, 3.8143, 7.3602, 10.800, 0.4800)
9	B_{III} (6.7910, 3.9858, 7.1280, 11.100, 0.5150)
10	B_{III} (6.7533, 4.2779, 6.6457, 11.700, 0.5800)
11	B_{III} (7.1745, 4.3469, 6.6161, 11.800, 0.6100)

Критерий со статистикой $S_{\max}^{AD}$ демонстрирует более высокую мощность относительно рассматриваемых конкурирующих гипотез по сравнению с критериями со статистиками $S_{\max}^{LR}$ и $S_{\max}^{Sm}$.

2.7. Критерий однородности χ^2

Пусть имеется k выборок $x_{i1}, x_{i2}, \ldots, x_{in_i}$, $i = \overline{1,k}$, непрерывных случайных величин и $n = \sum_{i=1}^{k} n_i$. Общая область, которой принадлежат выборки, разбивается на r интервалов (групп). Пусть η_{ij} – количество элементов i-й выборки, попавших в j-й интервал, тогда $n_i = \sum_{j=1}^{r} \eta_{ij}$.

Статистика критерия однородности χ^2 имеет вид

$$\chi^2 = n\sum_{i=1}^{k}\sum_{j=1}^{r}\frac{(\eta_{ij} - \nu_j n_i / n)^2}{\nu_j n_i} = n\left(\sum_{i=1}^{k}\sum_{j=1}^{r}\frac{\eta_{ij}^{\ 2}}{\nu_j n_i} - 1\right), \qquad (3.19)$$

где $\nu_j = \sum_{l=1}^{k}\eta_{lj}$ – общее число элементов всех выборок, попавших j-й интервал.

Асимптотическим распределением статистики (3.19) является χ^2-распределение с числом степеней свободы $(k-1)(r-1)$ [110].

Сравнивая оценки мощности критерия однородности χ^2 с мощностью других крнтериев [134], можно отметить, что он уступает большинству критериев относительно альтернативы сдвига, но вполне может конкурировать в других ситуациях.

3.7. Замечания о мощности критериев однородности законов

Опираясь на сравнительный анализ мощности критериев однородности, проведенный в [134], можно высказать некоторые соображения о предпочтительности применения тех или иных критериев.

Относительно конкурирующих гипотез, соответствующих изменению параметра сдвига, двухвыборочные критерии Смирнова Лемана–Розенблатта, Андерсона–Дарлинга–Петита, критерии Жанга со статистиками Z_{K}, Z_{A}, Z_C и критерий χ^2 по убыванию мощности располагаются в следующем порядке:

Андерсона–Дарлинга $\succ$ Лемана–Розенблатта $\succ$ Жанга Z_C $\succ$
Жанга Z_{A} $\succ$ Смирнова $\succ$ Жанга Z_{K} $\succ$ χ^2.

Относительно конкурирующих гипотез, соответствующих изменению параметра масштаба, критерии располагаются уже в другом порядке:

Жанга Z_{A} $\succ$ Жанга Z_C $\succ$ Жанга Z_{K} $\succ$ Андерсона–Дарлинга $\succ$ χ^2
$\succ$ Лемана–Розенблатта $\succ$ Смирнова.

При этом разница в мощности критериев Жанга со статистиками Z_{A} и Z_C невелика. Относительно более близких конкурирующих гипотез с меньшим сдвигом критерий Смирнова может оказаться впереди критерия Лемана–Розенблатта.

В ситуации, когда сравниваемые выборки принадлежат очень близким законам, критерии упорядочиваются по мощности следующим образом:

$$\text{Жанга } Z_{\text{K}} \succ \text{Жанга } Z_{\text{A}} \succ \text{Жанга } Z_C \succ \chi^2 \succ \text{Андерсона–Дарлинга} \succ \text{Смирнова} \succ \text{Лемана–Розенблатта}.$$

В случае k выборок в аналогичных ситуациях тот же порядок предпочтения сохраняется для k-выборочных вариантов критериев Андерсона–Дарлинга и Жанга. Однако в этот порядок вмешиваются критерии, использующие статистики $S_{\max}^{Sm}$, $S_{\max}^{LR}$, $S_{\max}^{AD}$. В частности, относительно изменения параметра сдвига порядок предпочтения имеет вид:

$$S_{\max}^{AD} \succ \text{Андерсона–Дарлинга} \succ S_{\max}^{LR} \succ S_{\max}^{Sm} \succ \text{Жанга } Z_C \succ \text{Жанга } Z_{\text{A}} \succ \text{Жанга } Z_{\text{K}} \succ \chi^2.$$

Относительно изменения параметра масштаба –

$$\text{Жанга } Z_C \succ \text{Жанга } Z_{\text{A}} \succ \text{Жанга } Z_{\text{K}} \succ \text{Андерсона–Дарлинга} \succ \chi^2 \succ S_{\max}^{AD} \succ S_{\max}^{Sm} \succ S_{\max}^{LR}.$$

При этом критерии Жанга со статистиками Z_{A} и Z_C практически эквивалентны по мощности, а критерий Андерсона–Дарлинга заметно уступает критериям Жанга.

Относительно ситуации, когда одна из выборок принадлежит другому, но близкому закону, критерии располагаются по мощности в следующем порядке:

$$\text{Жанга } Z_{\text{A}} \succ \text{Жанга } Z_C \succ \text{Жанга } Z_{\text{K}} \succ \chi^2 \succ \text{Андерсона–Дарлинга} \succ S_{\max}^{Sm} \succ S_{\max}^{AD} \succ S_{\max}^{LR}.$$

3.8. Пример применения критериев однородности законов для проверки гипотез об отсутствии тренда

В качестве примера, в котором демонстрируется возможность использования критериев однородности законов для проверки гипотезы об отсутствии тренда, покажем, как это реализовано в программной системе [148].

На рис. 3.1 показана начальная вкладка, по которой осуществляется переход к совокупности критериев однородности, применяемых для проверки гипотез об отсутствии того или иного вида тренда.

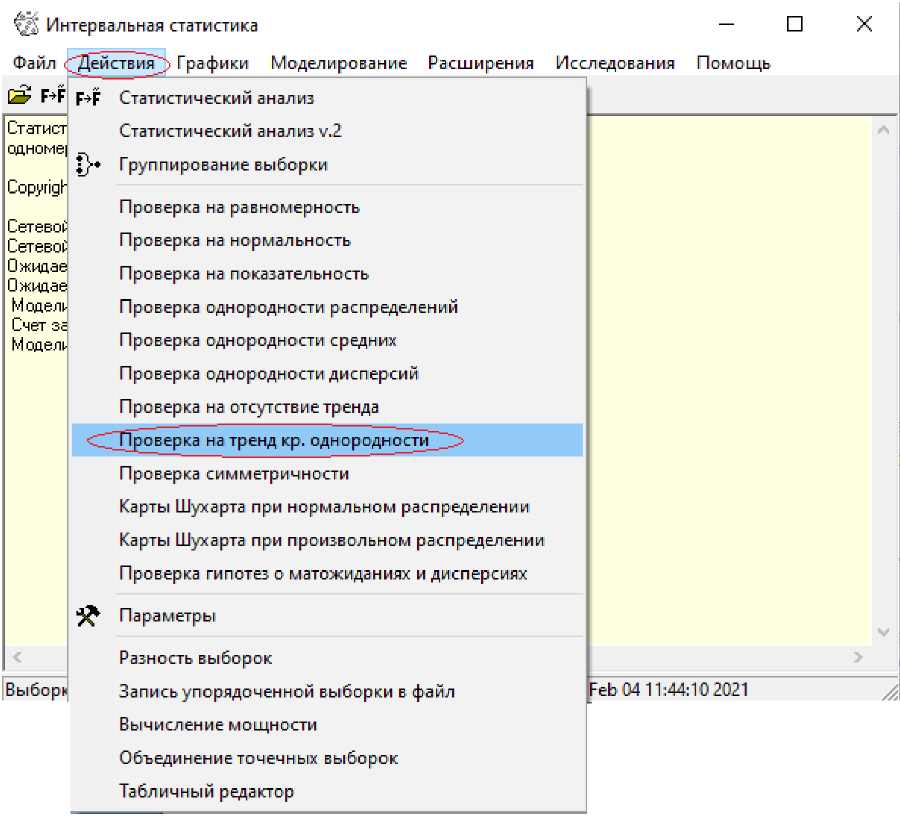


Рис. 3.1. Начальная вкладка допуска к критериям однородности, используемым для проверки отсутствия треда

На следующей вкладке (рис. 3.2) можно загрузить выборки, которые предполагается анализировать. После пометки анализируемой выборки открывается возможность разбить её на части, относительно которых возникают подозрения о различии в закономерностях, связанных с этими частями временного ряда.

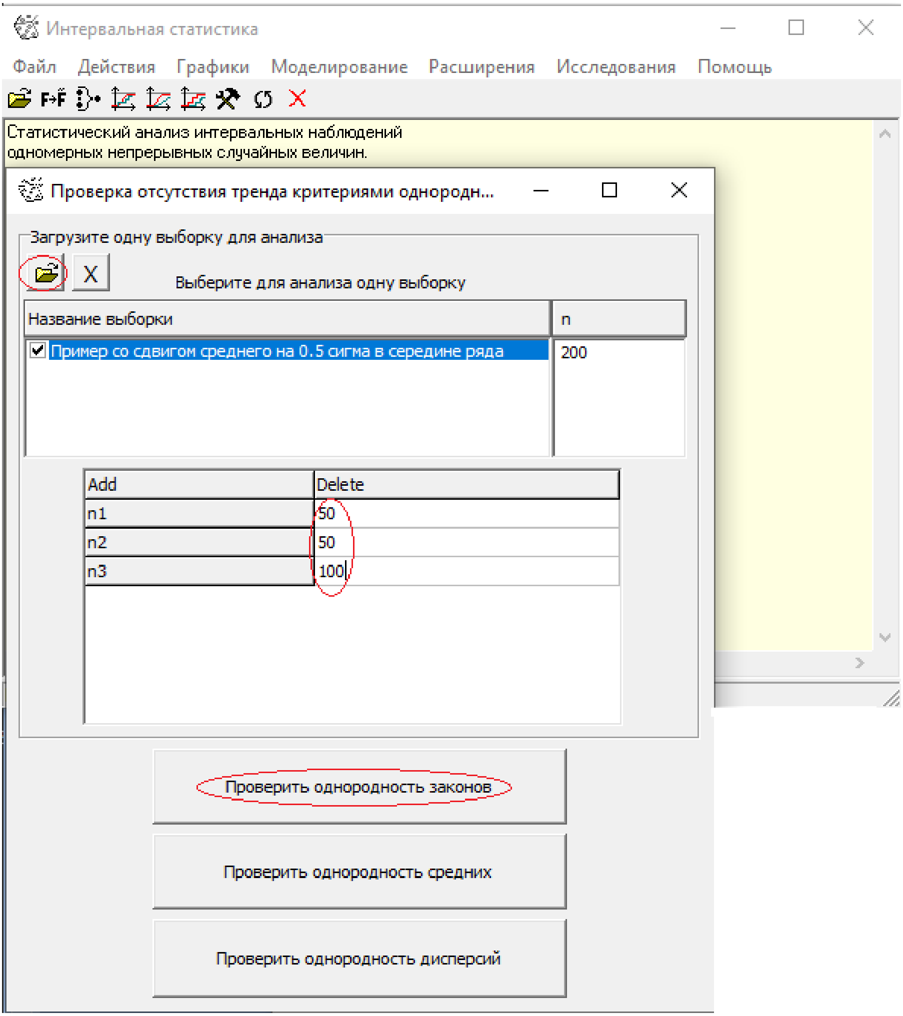


Рис. 3.2. Загрузка выборки с временным рядом измерений и разбиение её на части

Далее осуществляется переход к одной из групп критериев: к критериям однородности законов, к критериям проверки гипотез об однородности средних (о равенстве математических ожиданий), к критериям проверки гипотез об однородности дисперсий. При выборе проверки однородности законов, соответствующих выделенным частям анализируемой выборки, оказываемся на вкладке, показанной на рис. 3.3.

Выделив сравниваемые части (анализируемой выборки) и отметив используемые критерии, запускаем проверку. Результаты проверки отражаются на вкладке, показанной на рис.3.4.

Распределения статистик критериев Жанга зависят от объёмов выборок, и могут быть найдены только методами статистического моделирования. В параметрах этой вкладки можно задать количество экспериментов метода Монте-Карло, от которого зависит точность оценки p_{value}. Результаты проверки гипотезы по критериям Жанга, полученные с использованием смоделированных в интерактивном режиме распределений статистик, отражаются на открывающейся вкладке, показанной на рис. 3.5.

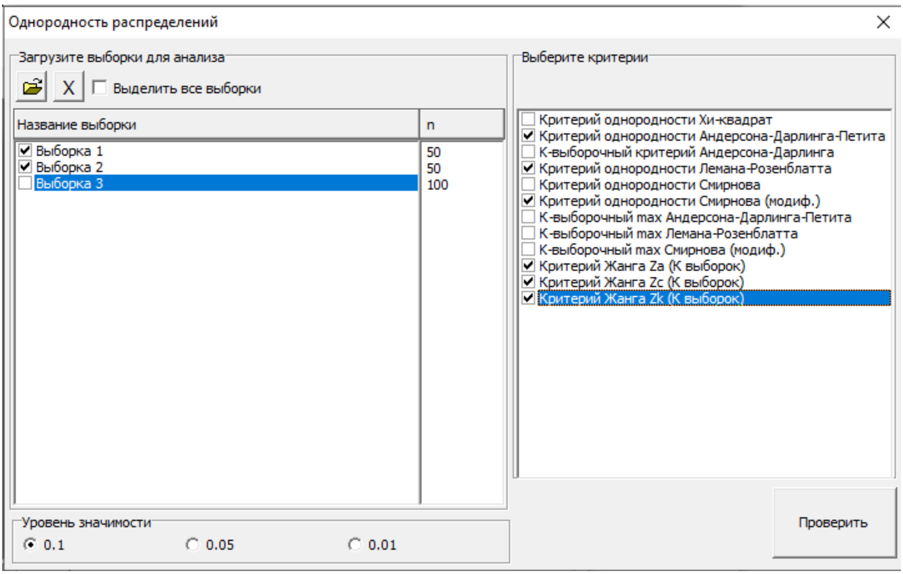


Рис. 3.3. Выбор анализируемых частей выборки используемых критериев

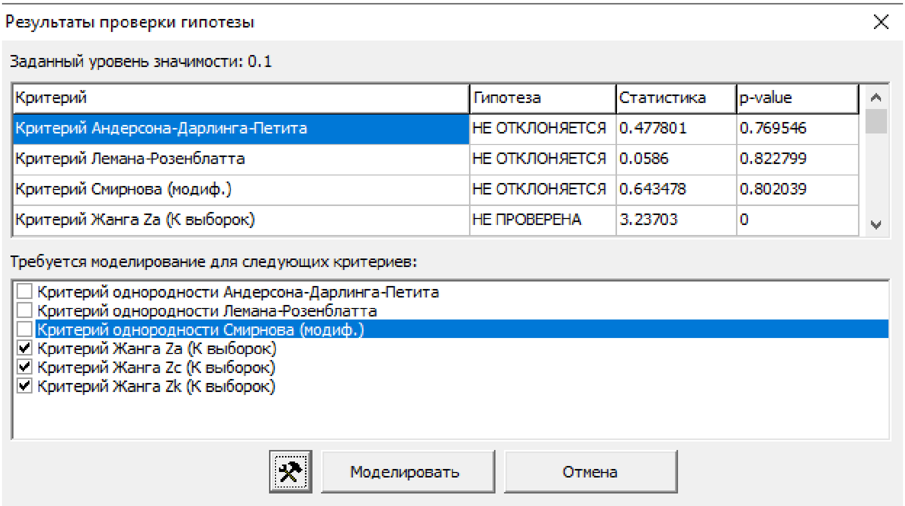


Рис. 3.4. Результаты проверки по критериям с известными асимптотическими распределениями статистик

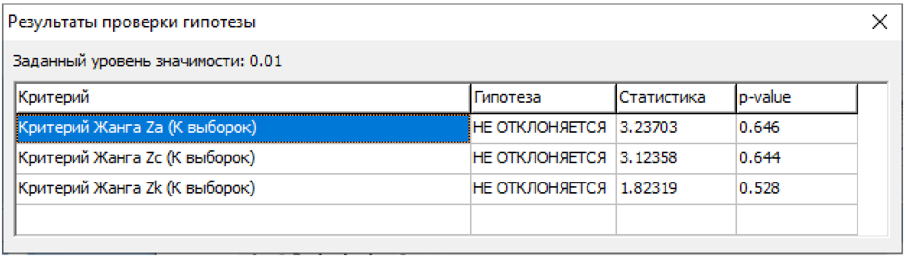


Рис. 3.5. Результаты проверки на основании смоделированных распределений статистик критериев

В данном случае применение критериев однородности для проверки гипотез об отсутствии тренда продемонстрируем на простом примере временного ряда, показанного на рис.3.6. Временной ряд представляет собой последовательность из двухсот псевдослучайных величин, подчиняющихся нормальному закону, первые 100 с нулевым математическим ожиданием, а последние с математическим ожиданием величиной 0.5σ.

В таблице 3.9 приведены результаты проверки однородности первых двух частей ряда по 50 измерений. Результаты свидетельствуют об отсутствии причин для отклонения проверяемой гипотезы об однородности (об отсутствии тренда). Подчеркнём, что в случае критериев Жанга оценки p_{value} находились на основании результатов статистического моделирования.

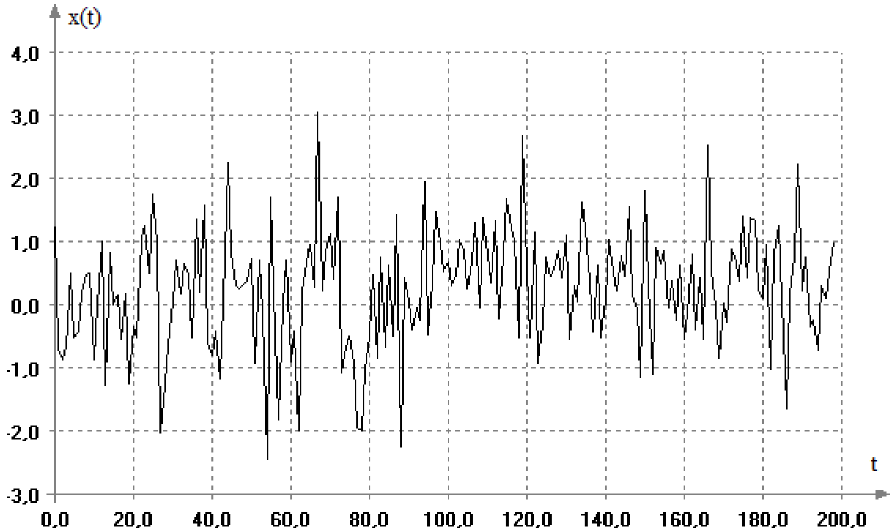


Рис. 3.6. Временной ряд измерений со сдвигом среднего на 0.5σ

Таблица 3.9

Результаты проверки гипотезы об однородности первых 2-х выборок

№ п/п	Критерий	Статистика	p_{value}
1	Андерсона-Дарлинга-Петита	0.4778	0.770
2	Лемана-Розенблатта	0.0586	0.823
3	Смирнова (модиф.)	0.6435	0.802
4	Жанга Za	3.2370	0.646
5	Жанга Zc	3.1236	0.644
6	Жанга Zk	1.8232	0.528

В таблице 3.10 приведены результаты проверки однородности трёх частей, на которые разбит ряд (2 по 50 и 100 измерений). Набор критериев несколько изменился, так как используются только критерии, ориентированные на анализ однородности более 2-х выборок. Полученные оценки p_{value} указывают на то, что при задании вероятности ошибки 1-го рода $\alpha = 0.05$, гипотеза об отсутствии тренда уверенно отклоняется по всем применяемым критериям.

Таблица 3.10

Результаты проверки гипотезы об однородности 3-х выборок

№ п/п	Критерий	Статистика	p_{value}
1	k-выборочный Андерсона-Дарлинга	3.8841	0.0068
2	k-выборочный max Андерсона-Дарлинга	4.5741	0.0131
3	k-выборочный max Лемана-Розенблатта	0.7860	0.0231
4	k-выборочный max Смирнова (модиф.)	1.5965	0.0336
5	Жанга Za	3.1420	0.0046
6	Жанга Zc	3.0673	0.0040
7	Жанга Zk	7.4805	0.0175

Показанное в данном разделе подчеркивает возможность эффективного применения критериев однородности законов при проверке гипотез об отсутствии тренда.

4. КРИТЕРИИ ОДНОРОДНОСТИ СРЕДНИХ ПРИ ПРОВЕРКЕ ГИПОТЕЗ ОБ ОТСУТСТВИИ ТРЕНДА

Очевидно, что для проверки гипотезы об отсутствии тренда в математическом ожидании могут эффективно использоваться критерии проверки однородности средних, реальные свойства которых достаточно подробно отражены в руководстве [134].

В критериях проверки гипотез об однородности средних проверяемая гипотеза о равенстве математических ожиданий, соответствующих k выборкам, имеет вид

$$H_0\colon\ \mu_1 = \mu_2 = \dots = \mu_k$$

при конкурирующей гипотезе

$$H_1\colon\ \mu_{i_1} \neq \mu_{i_2}\ ,$$

где неравенство выполняется хотя бы для одной пары индексов i_1, i_2.

Для проверки гипотезы H_0 может использоваться ряд параметрических критериев: сравнения двух выборочных средних при известных дисперсиях; сравнения двух выборочных средних при неизвестных, но равных дисперсиях (критерий Стьюдента); сравнения двух выборочных средних при неизвестных и неравных дисперсиях; F-критерий. В этих же целях применяется целая совокупность непараметрических критериев: критерий Уилкоксона, критерий Манна–Уитни, критерий Краскела–Уаллиса, критерий Ван дер Вардена, k - выборочный критерий Ван дер Вардена.

Основным предположением, обусловливающим возможность применения параметрических критериев, является принадлежность анализируемых выборок нормальному закону. Непараметрические критерии свободны от этого требования.

Реальные свойства критериев проверки однородности средних и результаты сравнительного анализа мощности подробно изложены в [134], поэтому в данном руководстве мы ограничимся лишь кратким описанием соответствующих критериев.

4.1. Параметрические критерии однородности средних

4.1.1. Критерий сравнения двух выборочных средних при известных дисперсиях

Применение критерия сравнения двух выборочных средних (по двум выборкам) при известных и равных дисперсиях предусматривает вычисление статистики

$$z = \left(\bar{x}_1 - \bar{x}_2 - \mu_1 + \mu_2\right) \Big/ \sqrt{\sigma_1^2 / n_1 + \sigma_2^2 / n_2}\ , \tag{4.1}$$

где $\bar{x}_i = \dfrac{1}{n_i} \sum\limits_{j=1}^{n_i} x_{ij}$, n_i – объем i-й выборки, $i = 1, 2$. В случае принадлежности наблюдений (ошибок измерений) нормальным законам статистика z подчиняется стандартному нормальному закону. Если в статистике (4.1) опустить μ_1 и μ_2, то будет проверяться гипотеза о равенстве математических ожиданий.

Критерий двусторонний. Проверяемая гипотеза отклоняется при больших по модулю значениях статистики (4.1).

4.1.2. Критерий Стьюдента

Это критерий проверки гипотезы о равенстве математических ожиданий при неизвестных, но равных дисперсиях. Критерий предусматривает вычисление статистики t [108]:

$$t = \left(\bar{x}_1 - \bar{x}_2 - \mu_1 + \mu_2\right) \Bigg/ \sqrt{\left[\frac{n_1 + n_2}{n_1 n_2}\right]\left[\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}\right]}\ , \tag{4.2}$$

где

$$s_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} \left(x_{ij} - \bar{x}_i\right)^2 .$$

Аналогично, если в статистике (4.2) опустить μ_1 и μ_2, то проверяется гипотеза о равенстве математических ожиданий.

В случае принадлежности выборок нормальному закону при справедливости гипотезы H_0 эта статистика подчиняется t_ν-распределению Стьюдента с числом степеней свободы $\nu = n_1 + n_2 - 2$. Критерий двусторонний.

4.1.3. Критерий сравнения двух выборочных средних при неизвестных и неравных дисперсиях

При неравных объемах выборок $n_1 \neq n_2$ статистика критерия, предложенная в [93], имеет вид [108]

$$t = \left(\bar{x}_1 - \bar{x}_2 - \mu_1 + \mu_2\right) \Big/ \sqrt{s_1^2/n_1 + s_2^2/n_2}\,. \tag{4.3}$$

Критерий также двусторонний. В случае нормального закона и справедливости гипотезы H_0 статистика (4.3) подчиняется распределению t_ν-Стьюдента с числом степеней свободы

$$\nu = \left(s_1^2/n_1 + s_2^2/n_2\right)^2 \Bigg/ \left[\frac{\left(s_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(s_2^2/n_2\right)^2}{n_2 - 1}\right]. \tag{4.4}$$

Задача, связанная с поиском распределения статистики (4.3), долгое время носила название проблемы Беренса–Фишера [4, 17, 18]. Её настоящее решение получено в работах [94, 81, 82]. Поэтому иногда критерий называют критерием Крамера–Уэлча.

В случае равенства неизвестных дисперсий статистика (4.3) эквивалентна статистике (4.2). При неравенстве дисперсий всегда число степеней свободы $\nu < n_1 + n_2 - 2$. Чем больше разница в дисперсиях, соответствующих анализируемым выборкам, тем сильнее распределение статистики (4.3) отличается от распределения статистики (4.2).

Распределения статистик (4.1) и (4.2) при справедливости H_0 известны точно, а решение проблемы Беренса–Фишера носит приближенный характер. В то же время это решение обладает хорошей точностью.

4.1.4. F-критерий однородности средних

В случае справедливости гипотезы о постоянстве (о равенстве) дисперсий с помощью F-критерия можно проверять гипотезу об однородности математических ожиданий по k выборкам [143].

Пусть у нас имеется k выборок объемом n. Сумма квадратов отклонений по всем выборкам определяется соотношением

$$Q_{kn} = \sum_{i=1}^{k}\sum_{j=1}^{n}\left(x_{ij} - \bar{x}_{kn}\right)^2,$$

где

$$\bar{x}_{kn} = \frac{1}{kn}\sum_{i=1}^{k}\sum_{j=1}^{n} x_{ij} = \bar{\bar{x}}_k$$

– среднее по всем выборкам. Общая сумма Q_{kn} раскладывается на два компонента

$$Q_{kn} = Q_1 + Q_2,$$

$$Q_1 = n\sum_{i=1}^{k}\left(\bar{x}_{in} - \bar{\bar{x}}_k\right)^2 = n\sum_{i=1}^{k}\left(\bar{x}_{in}^2 - k\bar{\bar{x}}_k\right)^2,$$

$$Q_2 = \sum_{i=1}^{k}\sum_{j=1}^{n}\left(x_{ij} - \bar{x}_{in}\right)^2 = (n-1)\sum_{i=1}^{k} s_{in}^2 .$$

Компонент Q_1, например, в задачах контроля качества является мерой различия в уровнях настройки между k выборками, в то время как Q_2 определяет различие в уровнях настройки внутри этих k выборок.

Для проверки гипотезы используется критерий со статистикой

$$F = \frac{Q_1 / (k-1)}{Q_2 / [k(n-1)]} . \qquad (4.5)$$

Если все выборки извлекаются из нормальной генеральной совокупности, то при справедливости гипотезы H_0 статистка (4.5) подчиняется F_{ν_1,ν_2}-распределению Фишера со степенями свободы

$\nu_1 = k-1$ и $\nu_2 = k(n-1)$ [143]. Критерий правосторонний. Проверяемая гипотеза отклоняется при больших значениях статистики (4.5).

4.1.5. k-выборочный вариант критерия Стьюдента

В работах [14, 15] для использования в задачах метрологии предлагается развитие критерия, представленного в п. 4.1.3, на случай k выборок.

В случае анализа k выборок $x_{ij},\ i=\overline{1,k},\ j=\overline{1,n_i}$, для проверки гипотезы об однородности средних статистика критерия имеет вид

$$\chi^2 = \sum_{i=1}^{k} n_i \frac{(\bar{x}_i - y)^2}{s_i^2}, \tag{4.6}$$

где $\bar{x}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}$, $s_i^2 = \frac{1}{n_i-1}\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_i)^2$, $y = \left(\sum_{i=1}^{k} n_i \bar{x}_i / s_i^2\right) \Big/ \sum_{i=1}^{k} \frac{n_i}{s_i^2}$ – взвешенное среднее по всем выборкам.

Критерий правосторонний. Асимптотическим распределением статистики (4.6) является χ^2_{k-1}-распределение.

4.1.6. Об устойчивости параметрических критериев

Принадлежность выборок нормальному закону является основным предположением, обуславливающим возможность применения перечисленных выше параметрических критериев. Именно в предположении о нормальности имеют место указанные распределения статистик соответствующих критериев при справедливости проверяемой гипотезы H_0.

В то же время давно известно, что параметрические критерии, предназначенные для проверки гипотез об однородности средних анализируемых выборок, достаточно устойчивы к нарушению стандартного предположения о нормальности.

В ряде наших работ, например [46, 113, 118, 120], было показано, что распределения статистик (4.1)–(4.3), (4.5) и (4.6) при справедливой проверяемой гипотезе H_0 в случае нарушения предположений о нормальности закона практически не изменяются. На отклонение распределения статистик от асимптотических влияет лишь асимметрия закона, которому прнадлежат выборки, или тяжелые “хвосты” этого закона.

Можно отметить общую закономерность: параметрические критерии, связанные с проверкой гипотез о математических ожиданиях, весьма устойчивы к отклонениям наблюдаемых законов от нормального. Это справедливо даже в случае многомерных случайных величин [120].

4.2. Непараметрические критерии однородности средних

Строго говоря, применение непараметрических критериев проверки гипотез о равенстве математических ожиданий предполагает, что анализируемые выборки принадлежат законам, которые могут отличаться разве что параметрами сдвига. И проверка гипотезы об однородности направлена именно на обнаружение возможного сдвига.

Нарушение данного предположения отражается на распределениях статистик критериев, имеющих место при справедливости проверяемой гипотезы H_0 о равенстве математических ожиданий. Однако на этом факте, как правило, не акцентируют внимания, так как заметные изменения распределений статистик (при справедливости H_0) проявляются при существенно (на порядок) отличающихся параметрах масштаба закона, соответствующего анализируемым выборкам.

4.2.1. Критерии Уилкоксона и Манна–Уитни.

Ранговый критерий Манна и Уитни [70, 73, 26, 11] является непараметрическим аналогом *t*-критерия для сравнения двух средних значений непрерывных распределений. Критерий Манна и Уитни представляет собой развитие критерия Уилкоксона [95].

Для вычисления статистики Уилкоксона две независимые выборки объединяют в одну объёмом в $n_1 + n_2$ значений и упорядочивают. По объединенной выборке определяют сумму рангов R_1, соответствующую элементам первой выборки, и сумму рангов второй R_2. Статистика критерия Уилкоксона [95] имеет вид

$$U = \min\{U_1, U_2\}, \qquad (4.7)$$

где

$$U_1 = n_1 n_2 + n_1(n_1+1)/2 - R_1,$$
$$U_2 = n_1 n_2 + n_2(n_2+1)/2 - R_2.$$

Дискретные распределения U-статистики сильно зависят от размера выборок, что затрудняет применение критерия.

В критерии Манна–Уитни (Манна–Уитни–Уилкоксона) вместо U-статистики используется нормализованная статистика

$$\tilde{z} = \frac{U - n_1 n_2/2}{\sqrt{n_1 n_2 (n_1 + n_2 + 1)/12}}. \qquad (4.8)$$

Дискретное распределение статистики (4.8) в случае справедливости проверяемой гипотезы H_0 хорошо приближается стандартным нормальным законом при $n_1 + n_2 > 60$, если объем каждой из выборок не слишком мал: $n_1 \ge 8$, $n_2 \ge 8$.

4.2.2. Критерий Краскела–Уаллиса

H-критерий Краскела–Уаллиса [38, 39] является развитием U-критерия для проверки гипотезы о равенстве средних по k выборкам. Объединенную выборку объёмом $n = \sum_{i=1}^{k} n_i$ упорядочивают и вычисляют суммы рангов R_i для i-й выборки, $i = \overline{1,k}$. Статистика для проверки гипотезы H_0 имеет вид

$$H = \left[\frac{12}{n(n+1)}\right]\left[\sum_{i=1}^{k} \frac{R_i^2}{n_i}\right] - 3(n+1). \qquad (4.9)$$

Статистика H представляет собой дисперсию ранговых сумм. При больших n_i и k в случае справедливости проверяемой гипотезы H_0 распределение статистики хорошо аппроксимируется χ^2_{k-1}-распределением [39]. В описаниях критерия говорится, что χ^2_{k-1}-распределением практически можно пользоваться при $n_i \ge 5$, $k \ge 4$.

Исследования показали, что на практике в случае $k = 2$ можно пренебречь дискретностью при $n_i \ge 30$.

С ростом числа выборок влияние дискретности быстро убывает.

4.2.3. Критерий Ван дер Вардена

Критерий предназначен для анализа 2-х выборок. Статистика непараметрического критерия Ван дер Вардена вычисляется в соответствии с выражением [102]:

$$V = \sum_{i=1}^{n_2} u_{R_i/(n_1+n_2+1)}\,, \tag{4.10}$$

где u_γ – γ-квантиль стандартного нормального закона, R_i, $i=\overline{1,n_2}$ – ранг i-го наблюдения, например, как в (4.10), второй выборки в общем вариационном ряду объединенной выборки из n_1+n_2 наблюдений.

Считается [109], что при $n_1, n_2 \ge 20$ распределение статистики (4.10) удовлетворительно описывается нормальным законом с нулевым математическим ожиданием и дисперсией

$$D[V] = \frac{n_1 n_2}{(n_1+n_2)(n_1+n_2-1)} \sum_{i=1}^{n_1+n_2} u^2_{i/(n_1+n_2+1)}\,.$$

Нормализованная статистика

$$V^* = \frac{V}{\sqrt{D[V]}} \tag{4.11}$$

должна подчиняться стандартному нормальному закону.

Критерий двусторонний, проверяемая гипотеза отклоняется при больших по модулю значениях статистики (4.11).

Исследование распределений статистики (4.11) методами статистического моделирования показало, что при $n_1+n_2 \ge 40$ отличием распределения $G(V^* | H_0)$ статистики (4.11) от стандартного нормального закона можно пренебречь. Величина отклонения не имеет практического значения уже при $n_1+n_2 \ge 20$.

Из непараметрических критериев однородности средних критерий Ван дер Вардена, по-видимому, является наиболее предпочтительным .

4.2.4. Критерий Фишера–Йэйтса–Терри–Гёфдинга

Этот критерий, рассмотренный в работах [19, 86, 25], очень близок к критерию Ван дер Вардена. В качестве меток в критерии выбраны математические ожидания соответствующих порядковых статистик в выборке объёмом $n = n_1 + n_2$ из стандартного нормального закона. Статистика критерия имеет вид:

$$S = \sum_{i=1}^{n_2} u_{(R_i - 3/8)/(n+1/4)} , \tag{4.12}$$

где u_γ – γ-квантиль стандартного нормального закона, R_i, $i = \overline{1, n_2}$ – ранг i-го наблюдения, например, второй выборки в общем вариационном ряду объединенной выборки из $n_1 + n_2$ наблюдений.

Так же как и в случае критерия Ван дер Вардена, статистика (4.12) достаточно хорошо описывается нормальным законом с нулевым математическим ожиданием и дисперсией

$$D[S] = \frac{n_1 n_2}{(n_1 + n_2)(n_1 + n_2 - 1)} \sum_{i=1}^{n_1+n_2} u^2_{(i-3/8)/(n_1+n_2+1/4)} ,$$

а нормализованная статистика

$$S^* = \frac{S}{\sqrt{D[S]}} \tag{4.13}$$

– стандартным нормальным законом.

По своим свойствам и мощности критерий со статистикой (4.13) эквивалентен критерию Ван дер Вардена [137].

3.2.5. Многовыборочный критерий Ван дер Вардена

Статистика критерия Ван дер Вардена для проверки гипотезы о равенстве средних по k выборкам имеет вид

$$T = (n-1) \sum_{i=1}^{k} \frac{1}{n_i} \left(\sum_{j=1}^{n_i} u_{R_{ij}/(n+1)} \right)^2 \Bigg/ \sum_{i=1}^{n} u^2_{i/(n+1)} , \tag{4.14}$$

где $n=\sum_{i=1}^{k} n_i$, u_γ – γ-квантиль стандартного нормального закона, R_{ij}, $j=\overline{1,n_i}$, – ранг j-го элемента i-й выборки в вариационном ряду объединённой выборки объёма n.

При справедливости проверяемой гипотезы H_0 статистика (4.14) хорошо описывается χ^2_{k-1}-распределением. Отклонением распределения статистики от χ^2_{k-1}-распределения можно пренебречь при $n_i > 30$.

k-выборочный критерий Ван дер Вардена демонстрирует более высокую мощность по сравнению с критерием Краскела–Уаллиса [137].

4.3. Сравнительный анализ мощности критериев

Подробный анализ мощности критериев однородности средних представлен в [134], из которого можно сделать следующие выводы.

Во-первых, очевидно, что параметрические критерии обладают несколько большей мощностью по сравнению с непараметрическими, но это преимущество весьма незначительно.

t-критерий Стьюдента со статистикой (4.2) слегка уступает в мощности z-критерию со статистикой (4.1) при известных дисперсиях. В рассматриваемой ситуации равенства дисперсий критерий со статистикой (4.3) эквивалентен критерию со статистикой (4.2) и имеет ту же мощность.

F-критерий эквивалентен k-выборочному критерию Стьюдента. Можно полагать, что в случае анализа 2-х выборок оба они эквивалентны критерию со статистикой (4.3), применяемому при неравных и неизвестных дисперсиях.

Следует отметить, что непараметрические критерии Ван дер Вардена практически не уступают в мощности параметрическому F-критерию [137].

В случае анализа 2-х выборок ранговые $\tilde{z}$-критерий Манна–Уитни и H-критерий Краскела–Уаллиса асимптотически эквивалентны и слегка уступают в мощности критериям Ван дер Вардена и чуть больше – параметрическим критериям.

Некоторый разнобой в оценках мощности $\tilde{z}$-критерия Манна–Уитни и H-критерия Краскела–Уаллиса, отраженный в результатах анализа, является следствием дискретности распределений этих статистик. Из-за дискретности действительные уровни значимости для этих критериев отличаются от заданных значений вероятности ошибок 1-го рода α и несколько превышают их. Этим же объясняется “преимущество” в некоторых случаях H-критерия Краскела–Уаллиса перед параметрическими критериями.

Можно заметить, что с ростом числа сравниваемых выборок k-выборочный критерий Стьюдента начинает уступать не только F-критерию, но и критерию Ван дер Вардена.

Аналогично, возрастает преимущество в мощности критерия Ван дер Вардена по сравнению с критерием Краскела–Уаллиса.

4.4. Пример применения критериев однородности средних для проверки гипотез об отсутствии тренда

Продемонстрируем применение критериев однородности средних для проверки гипотезы об отсутствии тренда в математическом ожидании. Сделаем это на примере реализации такой возможности в программной системе [148].

Начало процедуры применения, связанной с загрузкой анализируемой выборки и разбиением её на части, перспективные для целей обнаружения тренда, не отличается от процедуры, описанной в разделе 3.8 и показанной на рис.3.1 и 3.2. В данном случае анализируемая выборка (см. рис. 4.1) разбивается на 4 равные части.

После выбора раздела меню “Проверить однородность средних” открывается вкладка (см. рис. 4.2), позволяющая выбирать выделенные для анализа части временного ряда и применяемые для этого критерии однородности средних.

Выделив сравниваемые части (анализируемой выборки) и отметив используемые критерии, запускаем проверку. Результаты проверки отражаются на вкладке, показанной на рис.4.3.

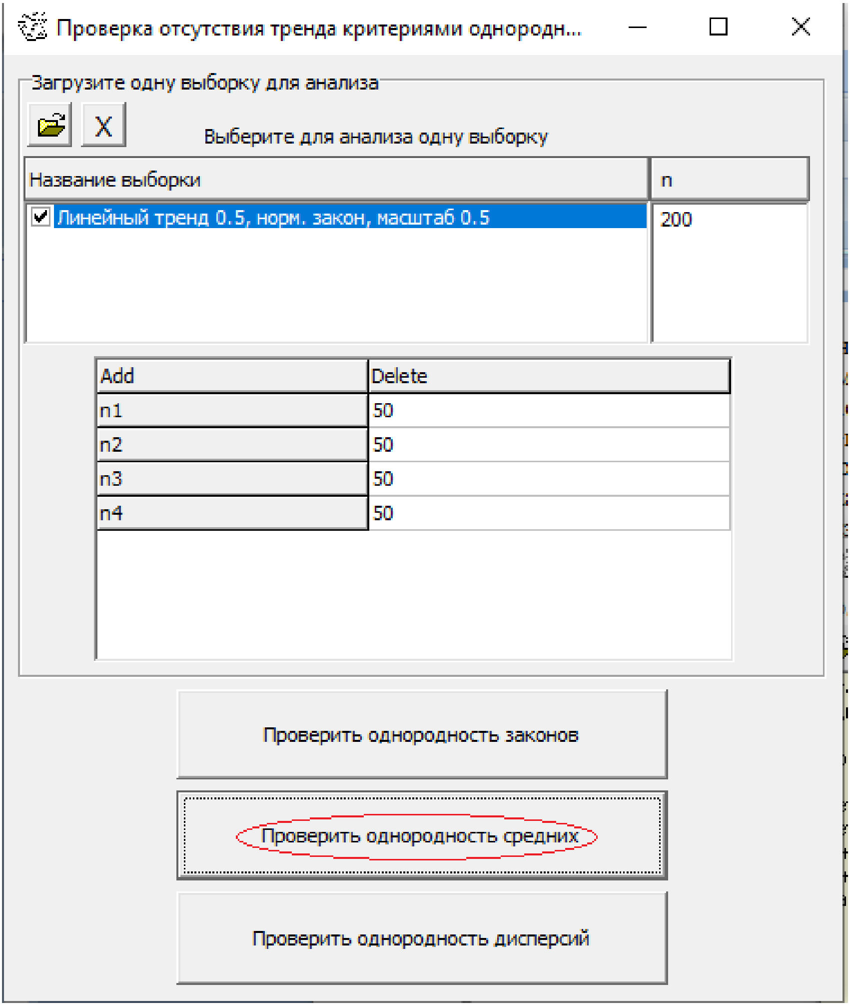


Рис. 4.1. Загрузка выборки с временным рядом измерений и разбиение её на части

Применение критериев однородности в средних для проверки гипотезы об отсутствии тренда в математическом ожидании продемонстрируем на примере временного ряда, показанного на рис. 4.4.

Временной ряд представляет собой последовательность из двухсот псевдослучайных величин ξ_i , подчиняющихся нормальному закону с нулевым математическим ожиданием и дисперсией $\sigma^2 = 0.25$, на который наложен тренд вида $X_i = 0.5t_i + \xi_i$. Здесь $t_i = (i-1)\Delta t$, где шаг $\Delta t = 1/n$ связывался с объемом выборки n (см. раздел 1.2).

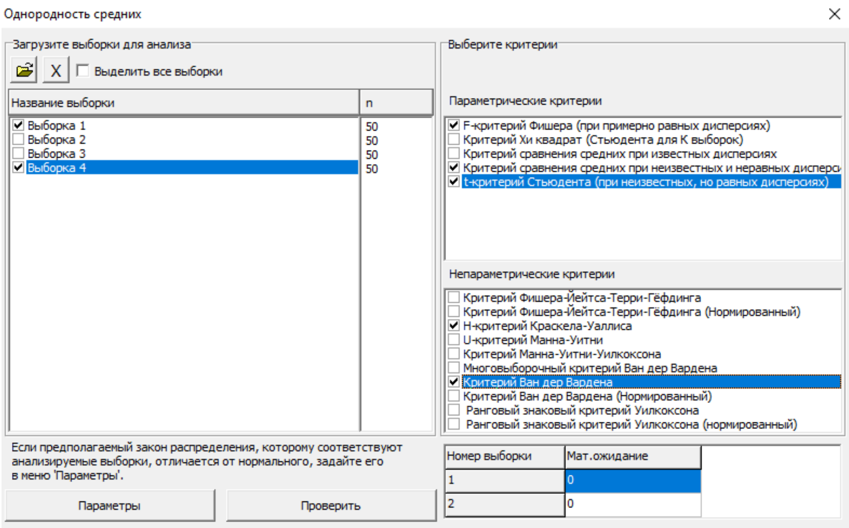


Рис. 4.2. Выбор анализируемых частей выборки используемых критериев

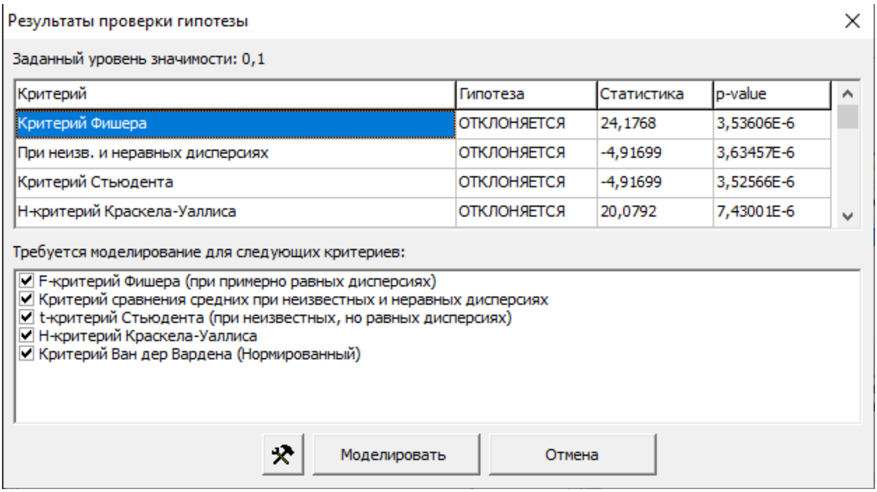


Рис. 4.3. Результаты проверки по критериям с известными асимптотическими распределениями статистик

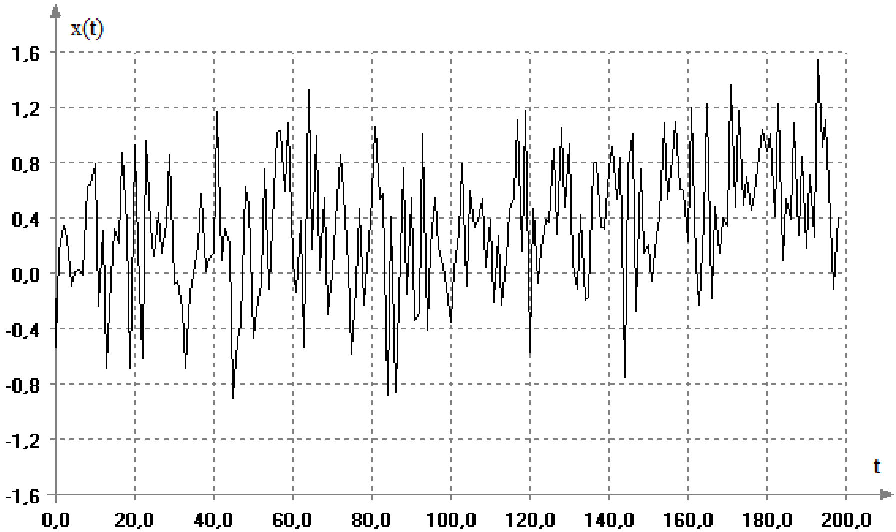


Рис. 4.4. Временной ряд измерений с линейным трендом в математическом ожидании

В данном случае временной ряд разбит на 4 равные выборки.

В таблице 4.1 приведены результаты проверки однородности (по некоторым из критериев) первых двух частей ряда по 50 измерений. Оценки p_{value} вычислены по известным асимптотическим распределениям. Но могут находиться на основании статистического моделирования. Результаты свидетельствуют об отсутствии причин для отклонения проверяемой гипотезы об однородности (об отсутствии тренда).

Таблица 4.1

Результаты проверки однородности средних первых 2-х выборок

№ п/п	Критерий	Статистика	p_{value}
1	F-критерий Фишера	0.871329	0.352882
2	При неизвестных и неравных дисперсиях	-0.93345	0.353054
3	t-критерий Стьюдента	-0.93345	0.352882
4	H-критерий Краскела–Уаллиса	0.639493	0.423894
5	Манна-Уитни-Уилкоксона	0.799683	0.423894
6	Ван дер Вардена (нормализованный)	0.93018	0.352278

В таблице 4.2 при использовании того же перечня критериев приведены полученные результаты проверки однородности первой и последней 4-й части. В данном случае, судя по оценкам p_{value}, гипотеза об отсутствии тренда **отклоняется** по всем применяемым критериям.

В заключение в таблице 4.3 приведены результаты проверки однородности всех 4-х частей по критериям, позволяющим сравнивать более 2-х выборок. И здесь по всем критериям гипотеза об отсутствии тренда должна быть отклонена.

Таблица 4.2

Результаты проверки однородности средних 1-й и 4-й выборок

№ п/п	Критерий	Статистика	p_{value}
1	F-критерий Фишера	24.1768	0.0000035
2	При неизвестных и неравных дисперсиях	-4.91699	0.0000036
3	t-критерий Стьюдента	-4.91699	0.0000035
4	H-критерий Краскела–Уаллиса	20.0792	0.0000074
5	Манна-Уитни-Уилкоксона	4.48098	0.0000074
6	Ван дер Вардена (нормализованный)	4.52951	0.0000059

Таблица 4.3

Результаты проверки однородности средних всех 4-х выборок

№ п/п	Критерий	Статистика	p_{value}
1	F-критерий Фишера	8.08983	0.00004
2	Хи-квадрат (Стьюдента для k выборок)	26.8661	0.00004
3	H-критерий Краскела–Уаллиса	21.0275	0.00010
4	k-выборочный Ван дер Вардена	23.0607	0.00004

Для k-выборочного критерия Стьюдента (χ^2-критерия) оценка p_{value} находилась на основании результатов статистического моделирования, так как распределение статистики этого критерия при справедливости H_0 далеко от χ_k^2-распределения.

Таким образом, критерии однородности средних можно эффективно применять в качестве критериев об отсутствии тренда в математическом ожидании.

5. КРИТЕРИИ ОДНОРОДНОСТИ ДИСПЕРСИЙ ПРИ ПРОВЕРКЕ ГИПОТЕЗ ОБ ОТСУТСТВИИ ТРЕНДА

В критериях проверки однородности дисперсий проверяемая гипотеза о постоянстве дисперсий k выборок имеет вид

$$H_0: \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2 ,$$

а конкурирующая с ней гипотеза

$$H_1: \sigma_{i_1}^2 \neq \sigma_{i_2}^2 ,$$

где неравенство выполняется, по крайней мере, для одной пары индексов i_1, i_2.

Реальные свойства множества критериев, используемых при проверке гипотез об однородности характеристик рассеяния, наиболее полно рассмотрены в руководстве [134]. Очевидно, что множество этих критериев может эффективно использоваться для проверки гипотезы об отсутствии тренда в дисперсии. С этой целью критерии проверки однородности дисперсий могут применяться для сравнения характеристик выбираемых частей временного ряда.

Совокупность критериев, применяемых для проверки гипотез однородности дисперсий, делится на два подмножества. Возможность использования классических результатов, связанных с применением параметрических критериев, обусловлено принадлежностью анализируемых выборок нормальным законам. Для непараметрических критериев такое требование отсутствует, но по умолчанию предполагается, что выборки принадлежат одному и тому же виду закона. Только в последнем случае обеспечивается корректность статистических выводов.

Распределения статистик параметрических критериев очень чувствительны к нарушению стандартного предположения о нормальности, в связи с этим и исключается возможность использования классических результатов. Но не исключается применение критериев в нестандартных условиях [53, 54]. Ясно, что для такого применения должна быть обеспечена возможность нахождения распределения статистики $G(S_n | H_0)$ критерия и вычисления p_{value} в этих нестандартных условиях. Именно таким образом с опорой на численные методы статистического моделирования в программной системе [148] обеспечивается корректность выводов при использовании параметрических критериев однородности дисперсий в условиях нарушения стандартного предположения о нормальности.

5.1. Параметрические критерии однородности дисперсий

5.1.1. Критерий Бартлетта

Статистика критерия Бартлетта [2, 101] вычисляется в соответствии с соотношением

$$\chi^2 = M\left[1+\frac{1}{3(k-1)}\left(\sum_{i=1}^{k}\frac{1}{\nu_i}-\frac{1}{N}\right)\right]^{-1}, \tag{5.1}$$

где

$$M = N\ln\left(\frac{1}{N}\sum_{i=1}^{k}\nu_i S_i^2\right)-\sum_{i=1}^{k}\nu_i \ln S_i^2 ;$$

k – количество выборок; n_i – объемы выборок; $\nu_i = n_i$, если математическое ожидание известно, и $\nu_i = n_i - 1$, если не известно; $N=\sum_{i=1}^{k}\nu_i$; S_i^2 – оценки выборочных дисперсий. При неизвестном математическом ожидании оценки $S_i^2 = \frac{1}{n_i-1}\sum_{j=1}^{n_i}\left(X_{ji}-\overline{X}_i\right)^2$, $\overline{X}_i = \frac{1}{n_i}\sum_{j=1}^{n_i}X_{ji}$, где X_{ij} – j-е наблюдение в i-й выборке.

Если гипотеза H_0 верна, все $\nu_i > 3$ и выборки извлекаются из нормальной генеральной совокупности, то статистика (5.1) приближенно подчиняется χ^2_{k-1}-распределению. Если вычисленное значение статистики $\chi^{2*} > \chi^2_{1-\alpha,k-1}$, то проверяемая гипотеза отклоняется на заданном уровне значимости α.

При нормально распределенных результатах измерений распределение статистики (4.4) практически не зависит от изменения объемов выборок [45, 119]. При малых величинах ν_i можно воспользоваться таблицей процентных точек для критерия Бартлетта, приводимой в [101]. Использование этой таблицы несколько затруднительно.

5.1.2. Критерий Кокрена

Статистика критерия Кокрена [9] выражается формулой

$$Q = \frac{S_{\max}^2}{S_1^2 + S_2^2 + \cdots + S_k^2}, \tag{5.2}$$

где $S_{\max}^2 = \max\left(S_1^2,\ S_2^2, \ldots,\ S_k^2\right)$; k – число независимых оценок дисперсий (число выборок); S_i^2, $i = \overline{1,k}$, – оценки выборочных дисперсий. Проверяемая гипотеза отклоняется при больших значениях статистики.

В [101] критерий Кокрена называется более простым и менее мощным по сравнению с критерием Бартлетта. Второе не соответствует действительности: при выполнении стандартного предположения о нормальности при $k = 2$ эти критерии по мощности эквивалентны, а при $k > 2$ преимущество в мощности за критерием Кокрена.

Распределения статистики Кокрена сильно зависят от объемов наблюдаемых выборок. Аналитический вид распределений статистики неизвестен. В справочной литературе можно найти только таблицы процентных точек для ситуации равных объёмов сравниваемых выборок $n_i = n,\ i = \overline{1,k}$ [101], которые и используются при проверке гипотез.

Критические значения статистики критерия Кокрена, полученные нами методами статистического моделирования при числе экспериментов $2{\times}10^6$ для числа анализируемых выборок $k = 2 \div 6$ при равных объемах выборок $n_i = n$, представлены в [134].

5.1.3. Критерий Хартли

Статистика критерия Хартли [23], применяемого для проверки гипотезы об однородности дисперсий, имеет вид

$$F = \frac{S_{\max}^2}{S_{\min}^2}, \tag{5.3}$$

где $S_{\max}^2 = \max\left(S_1^2,\ S_2^2, \ldots,\ S_k^2\right)$; $S_{\min}^2 = \min\left(S_1^2,\ S_2^2, \ldots,\ S_k^2\right)$; k – число независимых оценок дисперсий (число выборок); $S_i^2,\ i = \overline{1,k}$ – оценки выборочных дисперсий. Проверяемая гипотеза отклоняется при больших значениях статистики.

Распределения статистики существенно зависят от числа сравниваемых выборок k и их объемов $n_i,\ i = \overline{1,k}$.

Полученные с использованием методов статистического моделирования при числе экспериментов $2{\times}10^6$ критические значения статистики критерия Хартли для числа анализируемых выборок $k = 2 \div 6$ при равных объемах выборок $n_i = n$ представлены в [134].

5.1.4. Критерий Левене

Статистика критерия Левене [65] имеет вид

$$W = \frac{N-k}{k-1} \frac{\sum_{i=1}^{k} n_i \left(\overline{Z}_{i\bullet} - \overline{Z}_{\bullet\bullet}\right)^2}{\sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(Z_{ij} - \overline{Z}_{i\bullet}\right)^2}, \tag{5.4}$$

где k – количество выборок; n_i − объем i-й выборки; $N = \sum_{i=1}^{k} n_i$; X_{ij} – j-е наблюдение в i-й выборке; $Z_{ij} = \left|X_{ij} - \overline{X}_{i\bullet}\right|$, в котором $\overline{X}_{i\bullet}$ − среднее в i-й выборке; $\overline{Z}_{i\bullet}$ − среднее Z_{ij} по i-й выборке; $\overline{Z}_{\bullet\bullet}$ − среднее Z_{ij} по всем выборкам. Проверяемая гипотеза отклоняется при больших значениях статистики. Существуют модификации этой статистики [6, 24, 66, 76, 90].

Критические значения статистики критерия Левене для числа анализируемых выборок $k = 2 \div 6$ при равных объемах выборок $n_i = n$, построенные с использованием методов статистического моделирования при числе экспериментов 2×10^6, представлены в [134].

4.1.5. Критерий Фишера

Критерий Фишера используется для проверки гипотезы о равенстве дисперсий *двух* выборок случайных величин. Статистика критерия имеет простой вид

$$F = \frac{s_1^2}{s_2^2}, \tag{5.5}$$

где s_1^2 и s_2^2 – несмещенные оценки дисперсий, вычисленные по выборкам.

В отличие от критериев, рассмотренных выше, распределение статистики критерия Фишера при выполнении стандартного предположения известно точно. В случае принадлежности выборок нормальному закону и справедливости $H_0: \sigma_1^2 = \sigma_2^2$ эта статистика подчиняется F_{ν_1,ν_2}-распределению Фишера с числом степеней свободы $\nu_1 = n_1 - 1$ и $\nu_2 = n_2 - 1$, где n_1 и n_2 – объемы сравниваемых выборок. Проверяемая гипотеза отклоняется при малых $F^* < F_{\alpha/2,\nu_1,\nu_2}$ или больших $F^* > F_{1-\alpha/2,\nu_1,\nu_2}$ значениях статистики.

Как и остальные, критерий Фишера очень чувствителен к отклонениям от нормальности. Критерий применяется только при анализе двух выбок и в этих условиях по мощности эквивалентен критериям Бартлетта, Кокрена, Хартли, Неймана–Пирсона и Z-критерию Оверолла–Вудворда.

5.1.6. Критерий Неймана-Пирсона

Статистика критерия Неймана–Пирсона (критерия отношения правдоподобия) [109] определяется отношением арифметического среднего всех оценок дисперсий s_i^2 к их геометрическому среднему:

$$h = \frac{1}{k}\sum_{i=1}^{k} s_i^2 \Bigg/ \left(\prod_{i=1}^{k} s_i^2\right)^{\frac{1}{k}}, \tag{5.6}$$

где k – количество выборок, $s_i^2 = \frac{1}{n_i - 1}\sum_{j=1}^{n_i}\left(x_{ij} - \bar{x}_i\right)^2$ – оценки выборочных дисперсий, $\bar{x}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}$ – выборочное среднее значение, x_{ij} – j-й элемент i-й выборки. Обычно предполагается, что $n_1 = n_2 = \ldots = n_k = n$. Критерий правосторонний. Проверяемая гипотеза H_0 отклоняется при больших значениях статистики (5.5), когда $h > h_{1-\alpha}$.

Распределения статистики (5.6) зависят от объёма n и от числа анализируемых выборок k.

Уточненные значения процентных точек (при выполнении стандартного предположения о нормальности) приведены в [134].

Критерий Неймана–Пирсона крайне чувствителен к отклонениям от предположения о нормальности.

5.1.7. Критерий О`Брайена.

При формировании статистики критерия [87] каждый j-й элемент i-й выборки x_{ij} преобразуется в соответствии с формулой

$$V_{ij} = \frac{(n_i - 1.5)n_i\left(x_{ij} - \bar{x}_i\right)^2 - 0.5 s_i^2 (n_i - 1)}{(n_i - 1)(n_i - 2)},$$

где n_i – объём, $\bar{x}_i$– среднее значение, s_i^2 – оценка дисперсии i-й выборки.

Статистика критерия имеет вид

$$V = \frac{\frac{1}{k-1}\sum_{i=1}^{k} n_i\left(\overline{V_i} - \overline{\overline{V_i}}\right)^2}{\frac{1}{N-k}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(V_{ij} - \overline{V_i}\right)^2}, \tag{5.7}$$

где $\overline{V_i} = \dfrac{1}{n_i}\sum_{j=1}^{n_i} V_{ij}$, $\overline{\overline{V_i}} = \dfrac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{n_i} V_{ij}$, $N = \sum_{i=1}^{k} n_i$.

Критерий правосторонний, и проверяемая гипотеза H_0 отклоняется при больших значениях статистики (5.7).

Предельным распределением статистики критерия О'Брайена при справедливости H_0 является $F_{k-1,N-k}$-распределение Фишера с $k-1$ и $N-k$ степенями свободы [87]. Однако проведенные исследования показывают, что распределение статистики (5.7) критерия О'Брайена достаточно медленно сходится к соответствующему распределению Фишера. Например, в случае $k=2$ отличием реального распределения $G(V|H_0)$ статистики (5.7) от соответствующего $F_{1,N-k}$-распределения Фишера можно пренебречь лишь при $n_1 = n_2 = n \geq 80$. При малых объёмах выборок существенное отличие распределения статистики $G\left(V|H_0\right)$ от $F_{k-1,N-k}$-распределения Фишера наблюдается при больших значениях V, поэтому использование процентных точек $F_{k-1,N-k}$-распределения приводит к увеличению вероятности β ошибок второго рода (вследствие уменьшения уровня значимости по сравнению с заданным α).

Чтобы обеспечить возможность корректного применения критерия и при малых объемах выборок, в [134] приведены полученные верхние процентные точки распределений статистики критерия О'Брайена при различном количестве m сравниваемых выборок для $n_1 = n_2 = n \leq 80$.

Исследования показали, что распределения статистик критериев О'Брайена, Левене и модифицированного Z-критерия Оверолла–Вудворда достаточно устойчивы к нарушению стандартного предположения о принадлежности выборок нормальному закону.

5.1.8. Критерий Линка

Критерий Линка (критерий отношения размахов) является аналогом критерия Фишера и используется только при анализе 2-х выборок ($m=2$). Статистика критерия имеет вид [68]:

$$F^* = \frac{\omega_{n_1}}{\omega_{n_2}}, \qquad (5.8)$$

где $\omega_{n_1} = x_{1,\max} - x_{1,\min}$, $\omega_{n_1} = x_{2,\max} - x_{2,\min}$ – размахи, а $x_{1,\max}$, $x_{2,\max}$, $x_{1,\min}$, $x_{2,\min}$ – максимальные и минимальные элементы сравниваемых выборок.

Критерий двусторонний. Проверяемая гипотеза отклоняется с уровнем значимости α, если $F^* > F^*_{1-\alpha/2}$ или $F^* < F^*_{\alpha/2}$, где $F^*_{1-\alpha/2}$ и $F^*_{\alpha/2}$ – верхнее и нижнее критические значения статистики.

Распределение статистики критерия существенно зависит от объемов сравниваемых выборок.

Уточненные в ходе исследований нижние и верхние процентные точки для статистики (5.8) критерия отношения размахов в случае принадлежности выборок нормальному закону при $n_1, n_2 \le 20$ приведены в [134].

5.1.9. Критерий Ньюмана

Статистика критерия Ньюмана (стьюдентизированного размаха) имеет вид [77]:

$$q = \frac{\omega_{n_1}}{s_{n_2}}, \tag{5.9}$$

где $\omega_{n_1} = x_{1,\max} - x_{1,\min}$, $s_{n_2} = \sqrt{\frac{1}{n_2 - 1}\sum_{i=1}^{n_2}\left(x_{2i} - \bar{x}_2\right)^2}$.

Как и предыдущий, этот критерий также является двусторонним. Проверяемая гипотеза H_0 о равенстве дисперсий отклоняется, если $q < q_{\alpha/2}$ или $q > q_{1-\alpha/2}$, где $q_{\alpha/2}$ и $q_{1-\alpha/2}$ – нижнее и верхнее критические значения статистики при заданном уровне значимости α.

Распределения статистики (5.9) критерия Ньюмана при справедливости H_0 зависят от объёмов анализируемых выборок.

Уточненные нижние и верхние критические значения статистики (5.9), выход за которые приводит к отклонению нулевой гипотезы, приведены в [134].

Критерий также очень чувствителен к любым отклонениям от нормальности.

5.1.10. Критерий Блиса–Кокрена–Тьюки

Статистика критерия [5], предложенного в качестве аналога критерия Кокрена, имеет вид

$$c = \frac{\max\limits_{1 \le i \le k} \omega_i}{\sum\limits_{i=1}^{k} \omega_i}, \tag{5.10}$$

где k – количество сравниваемых выборок, $\omega_i = \max\limits_{1 \le j \le n_i} x_{ij} - \min\limits_{1 \le j \le n_i} x_{ij}$ – размах i-й выборки.

Критерий правосторонний. Если статистика $c > c_{1-\alpha}$, где $c_{1-\alpha}$ – верхнее критическое значение при заданном уровне значимости α, то проверяемая гипотеза H_0 о равенстве дисперсий отклоняется.

Распределение статистики критерия сильно зависит от объема выборок и числа сравниваемых выборок.

Значения верхних процентных точек для статистики критерия при некоторых объёмах $n_1 = n_2 = n \le 20$ и числе $k \le 10$ сравниваемых выборок и выполнении стандартного предположения о нормальности приведены в [134]. Формально критерий Блисса–Кокрена–Тьюки можно применять и при неравных объёмах n_i анализируемых выборок, однако следует иметь ввиду, что в этом случае критические значения статистики будут отличаться от приведенных в [134].

Как и в случае критерия Кокрена, распределения статистики данного критерия сильно зависят от закона, которому принадлежат анализируемые выборки

5.1.11. Критерий Кадуэлла–Лесли–Брауна

Этот критерий [64] был предложен в качестве аналога критерия Хартли с заменой в статистике отношений оценок дисперсий на отношения размахов

$$K = \frac{\max\limits_{1 \le i \le k} \omega_i}{\min\limits_{1 \le i \le k} \omega_i}, \tag{5.11}$$

где k – количество выборок, ω_i – размах i-й выборки.

Критерий правосторонний. При $K > K_{1-\alpha}$, где $K_{1-\alpha}$ – верхнее критическое значение статистики при заданном уровне значимости α, проверяемая гипотеза H_0 отклоняется.

Распределения статистики критерия Кадуэлла–Лесли–Брауна так же, как и статистики критерия Блисса–Кокрена–Тьюки, существенно зависят и от объема выборок, и от закона распределения, которому подчиняются выборки.

В [134] приведены критические значения $K_{1-\alpha}$ для количества выборок $k \le 10$ при равных объемах выборок $n_1 = n_2 = n \le 20$, $i = \overline{1,m}$, полученные методами статистического моделирования. Формально критерий Кадуэлла–Лесли–Брауна можно применять и при неравных объёмах n_i анализируемых выборок.

5.1.12. Z–критерий Оверолла–Вудворда

Статистика критерия имеет вид [88]:

$$Z = \frac{1}{k-1}\sum_{i=1}^{k} Z_i^2 , \tag{5.12}$$

где $Z_i = \sqrt{\dfrac{c_i\left(n_i-1\right)s_i^2}{MSE}} - \sqrt{c_i\left(n_i-1\right)-\dfrac{c_i}{2}}$, $MSE = \dfrac{1}{N-k}\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(x_{ij}-\overline{x}_i\right)^2$,

m – количество выборок, $c_i = 2 + 1/n_i$, n_i – размер i-й выборки, s_i^2 – несмещенная оценка дисперсии i-й выборки, $N = \sum_{i=1}^{k} n_i$, x_{ij} – j-й элемент i-й выборки, $\overline{x}_i$ –среднее значение i-й выборки.

При справедливости проверяемой гипотезы H_0 об однородности дисперсий и принадлежности анализируемых выборок нормальному закону предельное распределение статистики (5.12) не зависит от размера выборки и подчиняется $F_{k-1,\infty}$-распределению Фишера. Однако при малых объемах n_i распределение статистики Z–критерия Оверолла–Вудворда заметно отличается от предельного $F_{k-1,\infty}$-распределения Фишера.

Различием между реальным распределением статистики критерия и предельным $F_{k-1,\infty}$-распределением можно пренебречь при объемах выборок $n_i \ge 50$. Поэтому (в предположении о принадлежности анализируемых выборок нормальному закону) для объемов выборок $n_i \le 50$ методами статистического моделирования (для разного количества сравниваемых выборок m и для различных $n_i = n$ объемов выборок) были вычислены верхние критические значения $Z_{1-\alpha}$, представленные в [134].

Распределение статистики (5.12) Z–критерия очень чувствительно к нарушению предположения о нормальности.

5.1.13. Модифицированный Z–критерий

В целях построения критерия, более устойчивого к нарушению стандартного предположения о нормальности, в [89] предложена модификация Z–критерия, статистика которого отличается вычислением величин c_i:

$$c_i = 2.0\left[\frac{1}{K_i}\left(2.9+\frac{0.2}{n_i}\right)\right]^{\frac{1.6\left(n_i-1.8K_i+14.7\right)}{n_i}}, \qquad (5.13)$$

где $K_i = \dfrac{1}{n_i-2}\sum\limits_{j=1}^{n_i} G_{ij}^4$ – оценка коэффициента эксцесса i-й выборки, $G_{ij} = \left(x_{ij}-\bar{x}_i\right)\Big/\sqrt{\dfrac{n_i-1}{n_i}s_i^2}$.

Распределение статистики модифицированного Z–критерия с ростом объема выборок очень медленно сходится к $F_{k-1,\infty}$-распределению Фишера. Даже при больших объемах выборок распределение статистики критерия не согласуется с $F_{k-1,\infty}$-распределением, хотя в области больших значений статистики отличие её распределения от $F_{k-1,\infty}$-распределения не играет существенного значения.

Для корректного применения критерия при малых объёмах выборок [134] представлены критические значения, полученные с использованием методов статистического моделирования.

Распределение статистики модифицированного Z–критерия действительно обладает большей стабильностью к нарушению стандартного предположения о нормальности.

Однако, борьба за устойчивость привела не только к ухудшению сходимости распределения статистики модифицированного Z–критерия к $F_{k-1,\infty}$-распределению, но и к некоторому снижению мощности критерия.

5.1.14. Критерий Миллера

Миллер [72] предложил критерий однородности дисперсий, базирующийся на F-преобразовании Фишера для выборочных дисперсий. Лайард [41] обобщил двухвыборочный критерий Миллера на случай k выборок.

Статистика k-выборочного критерия Миллера имеет вид

$$M = \frac{\sum\limits_{i=1}^{k} n_i \left(\overline{U}_{i\bullet} - \overline{U}_{\bullet\bullet}\right)^2 \Big/ (k-1)}{\sum\limits_{i=1}^{k} \sum\limits_{j=1}^{n_i} \left(U_{ij} - \overline{U}_{i\bullet}\right)^2 \Big/ (n-k)}, \qquad (5.14)$$

где

$$U_{ij} = n_i \ln S_i^2 - (n_i - 1) \ln S_{i(j)}^2;$$

$$S_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} \left(x_{ij} - \bar{x}_i\right)^2; \quad \bar{x}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij};$$

$$S_{i(j)}^2 = \frac{1}{n_i - 2} \sum_{l \neq j} \left(x_{il} - \bar{x}_{i(j)}\right)^2; \quad \bar{x}_{i(j)} = \frac{1}{n_i - 1} \sum_{l \neq j} x_{il};$$

$$\overline{U}_{i\bullet} = \frac{1}{n_i} \sum_{j=1}^{n_i} U_{ij}; \quad \overline{U}_{\bullet\bullet} = \frac{1}{n} \sum_{i=1}^{k} \sum_{j=1}^{n_i} U_{ij}; \quad n = \sum_{i=1}^{k} n_i.$$

При справедливости проверяемой гипотезы H_0 об однородности дисперсий и выполнении предположений о принадлежности выборок нормальным законам статистика (5.14) должна подчиняться F-

распределению Фишера с числами степеней свободы ($k-1$) и ($n-k$). Проверяемая гипотеза отклоняется при больших значениях статистик.

Отметим, что при малых объёмах выборок распределение $G(M_n | H_0)$ статистики (5.14) Миллера заметно отличается от соответствующего F-распределения. Реально отклонением распределения $G(M_n | H_0)$ статистики от распределения Фишера с ($k-1$) и ($n-k$) степенями свободы можно пренебречь при $n_i > 40 \div 50$.

Распределения статистики (5.14) чувствительны к нарушению предположения о нормальности анализируемых выборок.

5.1.15. Критерий Лайарда

Лайард [41] представил критерий со статистикой, в которой используется функция эксцесса нескольких выборок для проверки однородности дисперсий.

Статистика критерия Лайарда определяется выражением

$$L = \sum_{i=1}^{k} (n_i - 1) \frac{\left(\ln S_i^2 - T\right)^2}{\delta^2}, \tag{5.15}$$

где

$$T = \left[\sum_{i=1}^{k} (n_i - 1) \ln S_i^2 \right] \Big/ (n-k); \quad n = \sum_{i=1}^{k} n_i; \quad \delta^2 = 2 + \gamma [1 - k/n];$$

$$\gamma = \frac{1}{n} \sum_{i=1}^{k} n_i^2 \left[\sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^4 \Big/ \left(\sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 \right)^2 \right] - 3.$$

Здесь γ – взвешенное среднее коэффициентов эксцесса k выборок.

Автор критерия далее предпочёл использовать в статистике (5.15) несколько другую оценку коэффициента эксцесса:

$$\hat{\gamma} = n \sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^4 \Big/ \left[\sum_{i=1}^{k} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 \right]^2 - 3.$$

Проверяемая гипотеза отклоняется при больших значениях статистики (5.15), которая при справедливости проверяемой гипотезы об однородности дисперсий и выполнении предположений о принадлежности выборок нормальным законам <u>асимптотически</u>

подчиняется χ^2_{k-1}-распределению.

Однако сходимость распределения статистики $G(L_n|H_0)$ к χ^2_{k-1}-распределению достаточно медленная. Реально отклонением $G(L_n|H_0)$ от χ^2_{k-1}-распределения, например, при $k=2$ можно пренебречь лишь при $n_i > 300$.

Следует отметить так же, что для вычисления статистики (5.15) всё-таки предпочтительней использовать оценку γ, а не $\hat{\gamma}$, так как в последнем случае сходимость $G(L_n|H_0)$ к χ^2_{k-1}-распределению несколько хуже.

По своим асимптотическим свойствам (и мощности) критерий Лайарда очень близок критерию Миллера.

5.2. Непараметрические критерии однородности дисперсий

5.2.1. Критерий Ансари–Бредли

Непараметрические аналоги критериев проверки однородности дисперсий предназначены для проверки гипотез о принадлежности двух выборок с объемами n_1 и n_2 общей генеральной совокупности с одинаковыми характеристиками рассеяния. При этом, как правило, предполагается равенство средних.

Статистика критерия Ансари–Бредли [1] может быть вычислена следующим образом:

$$S = \sum_{i=1}^{n_1} \left\{ \frac{n_1 + n_2 + 1}{2} - \left| R_i - \frac{n_1 + n_2 + 1}{2} \right| \right\}, \qquad (5.16)$$

где R_i – ранги элементов первой выборки в общем вариационном ряду. Проверяемая гипотеза не отклоняется при $S_{\alpha/2} < S < S_{1-\alpha/2}$. Критические значения статистики для $n_1, n_2 \le 10$ доступны в таблице, приведенной, например, в [109], а для больших значений n_i соответствующая таблица может быть элементарно расширена методами статистического моделирования.

В случае принадлежности выборок случайных величин одному и тому же закону распределение $G\left(S\middle|H_0\right)$ статистики (5.16) при справедливости проверяемой гипотезы H_0 не зависит от вида этого закона.

Математическое ожидание и дисперсия статистики (5.16) имеют вид

$$E[S]=\begin{cases}\dfrac{n_1(n_1+n_2+2)}{4} & \text{при четном } (n_1+n_2),\\[2ex] \dfrac{n_1(n_1+n_2+1)^2}{4(n_1+n_2)} & \text{при нечетном } (n_1+n_2);\end{cases}$$

$$D[S]=\begin{cases}\dfrac{n_1n_2(n_1+n_2-2)(n_1+n_2+2)}{48(n_1+n_2-1)} & \text{при четном } (n_1+n_2),\\[2ex] \dfrac{n_1n_2(n_1+n_2+1)\left[(n_1+n_2)^2+3\right]}{48(n_1+n_2)^2} & \text{при нечетном } (n_1+n_2).\end{cases}$$

При объемах выборок $n_1,\ n_2>10$ дискретное распределение нормированной статистики

$$S^*=(S-E[S])/\sqrt{D[S]} \tag{5.17}$$

достаточно хорошо приближается стандартным нормальным законом. В этом случае проверяемая гипотеза не отклоняется при $N^*_{\alpha/2}<S^*<N^*_{1-\alpha/2}$, где N^*_α – соответствующая квантиль стандартного нормального закона. Дискретностью распределений статистик (5.16) и (5.17) практически можно пренебречь, начиная с $n_1,\ n_2>40$.

5.2.2. Критерий Муда

Статистика критерия имеет вид [74, 85]

$$M=\sum_{i=1}^{n_1}\left(R_i-\frac{n_1+n_2+1}{2}\right)^2, \tag{5.18}$$

где R_i – ранги элементов первой выборки в общем вариационном ряду двух выборок. Проверяемая гипотеза не отклоняется при $M_{\alpha/2} < M < M_{1-\alpha/2}$. Критические значения данной статистики для $n_1, n_2 \le 10$ доступны в [109], а для больших значений n_i таблица может быть легко расширена методами статистического моделирования.

При $n_1, n_2 > 10$ распределение нормированной статистики

$$M^* = \frac{\left(M - E[M] + \frac{1}{2}\right)}{\sqrt{D[M]}}, \tag{5.19}$$

где

$$E[M] = \frac{n_1(n_1 + n_2 + 1)(n_1 + n_2 - 1)}{12},$$

$$D[M] = \frac{n_1 n_2 \left(n_1 + n_2 + 1\right)(n_1 + n_2 + 2)(n_1 + n_2 - 2)}{180},$$

хорошо приближается стандартным нормальным законом [40], а при $n_1, n_2 > 20$, как показали исследования, дискретностью распределений статистик (5.18), (5.19) вообще можно пренебречь. При использовании статистики (5.19) проверяемая гипотеза не отклоняется при $N^*_{\alpha/2} < M^* < N^*_{1-\alpha/2}$.

Проблемы с дискретностью статистики (5.19) существенно меньше, чем со статистикой (5.17) критерия Ансари–Бредли.

Критерий показывает мощность несколько выше по сравнению с критерием Ансари–Бредли.

5.2.3. Критерий Сижела–Тьюки

Статистика критерия строится следующим образом [84]. Вариационный ряд, построенный по объединенной выборке, $x_1 \le x_2 \le \ldots \le x_n$, где $n = n_1 + n_2$, преобразуется в последовательность вида

$$x_1, x_n, x_{n-1}, x_2, x_3, x_{n-2}, x_{n-3}, x_4, x_5, \ldots,$$

т. е. оставшийся ряд «переворачивается» каждый раз после приписывания рангов паре крайних значений. В качестве статистики критерия используется сумма рангов элементов первой выборки.

Статистика критерия имеет вид

$$R = \sum_{i=1}^{n_1} R_i \ . \tag{5.20}$$

Проверяемая гипотеза не отклоняется при $R_{\alpha/2} < R < R_{1-\alpha/2}$. Статистика критерия Сижела–Тьюки является аналогом критерия Манна–Уитни, но предназначенным для проверки гипотез об однородности параметров масштаба. Поэтому при проверке гипотезы могут использоваться квантили распределения статистики Манна–Уитни [109].

При $n_1, n_2 > 10$ распределение нормированной статистики

$$R^* = (R - E[R]) / \sqrt{D[R]}\,, \tag{5.21}$$

где

$$E[R] = \frac{n_1(n_1 + n_2 + 1)}{2}, \quad D[R] = \frac{n_1 n_2 \left(n_1 + n_2 + 1\right)}{12},$$

достаточно хорошо приближается стандартным нормальным законом. В этом случае проверяемая гипотеза не отклоняется при $N^*_{\alpha/2} < R^* < N^*_{1-\alpha/2}$. При этом дискретностью распределения статистики можно практически пренебречь с $n_1, n_2 > 30$.

Можно заметить, что статистика (5.21 более дискретна по сравнению со статистикой (5.19) Муда, но менее – по сравнению со статистикой (5.17) критерия Ансари–Бредли.

По мощности критерий эквивалентен критерию Ансари–Бредли.

5.2.4. Критерий Клотца

Статистика критерия имеет вид [35]

$$L = \sum_{i=1}^{n_1} u^2_{R_i/(n_1+n_2+1)} \ , \tag{5.22}$$

где n_1 и n_2 – объемы сравниваемых выборок; R_i – ранг i-го элемента первой выборки в общем упорядоченном по возрастанию ряду (n_1+n_2) значений объединенной выборки; u_γ – γ-квантиль стандартного нормального распределения. Значения $u^2_{i/(N+1)}$, называемые метками критерия, которые несложно вычислить, приведены, например, в [109].

Гипотеза о равенстве параметров масштаба не отклоняется с достоверностью α, если $L_{\alpha/2} < L < L_{1-\alpha/2}$, где критические значения $L_{\alpha/2}$, $L_{1-\alpha/2}$ можно найти в [109].

При $n_1, n_2 > 10$ нормализованная статистика

$$L^* = \frac{L - E[L]}{\sqrt{D[L]}}, \tag{5.23}$$

где

$$E[L] = \frac{n_1}{n_1+n_2} \sum_{i=1}^{n_1+n_2} u^2_{\frac{i}{n_1+n_2+1}},$$

$$D[L] = \frac{n_1 n_2}{(n_1+n_2)(n_1+n_2-1)} \sum_{i=1}^{n_1+n_2} u^4_{\frac{i}{n_1+n_2+1}} -$$

$$-\frac{n_2}{n_1(n_1+n_2-1)} \left[\frac{n_1}{n_1+n_2} \sum_{i=1}^{n_1+n_2} u^2_{\frac{i}{n_1+n_2+1}} \right]^2,$$

хорошо аппроксимируется стандартным нормальным законом. Гипотеза равенства параметров масштаба не отклоняется с достоверностью α, если $N^*_{\alpha/2} < L^* < N^*_{1-\alpha/2}$, где N^*_α – соответствующая квантиль стандартного нормального закона.

Дискретностью распределений статистик критерия можно пренебречь при $n_1, n_2 > 10$. При таких же объёмах выборок можно считать несущественным отклонение распределения $G(L^* | H_0)$ статистики (5.23) от стандартного нормального закона.

Критерий Клотца демонстрирует мощность, превышающую мощность критерия Муда, который в свою очередь имеет преимущество перед критериями Ансари–Бредли и Сижела–Тьюки.

5.2.5. Критерий Кейпена

Статистика критерия Кейпена [7] задается соотношением

$$K = \sum_{i=1}^{n_1} a_{n_1+n_2}\left(R_i\right), \tag{5.24}$$

где n_1 и n_2 – объемы сравниваемых выборок; R_i – ранг i-го элемента первой (меньшей по объёму) выборки в общем упорядоченном по возрастанию ряду $n = n_1 + n_2$ значений объединенной выборки. Метки $a_n\left(i\right)$ представляют собой квадрат математического ожидания i-й порядковой статистики в выборке объема n из стандартного нормального закона и приближённо определяются соотношением $a_n\left(i\right) = u^2_{(i-3/8)/(n+1/4)}$, где u_γ – γ-квантиль стандартного нормального распределения. Значения $a_n\left(i\right)$ для некоторых n приведены, например, в [109].

Гипотеза о равенстве параметров масштаба не отклоняется с достоверностью α, если $K_{\alpha/2} < K < K_{1-\alpha/2}$, где критические значения $K_{\alpha/2}$, $K_{1-\alpha/2}$ также можно найти в [109].

При $n_1, n_2 > 10$ можно использовать нормализованную статистику

$$K^* = \frac{K - E[K]}{\sqrt{D[K]}}, \tag{5.25}$$

где

$$E[K] = n_1, \quad D[K] = \frac{n_1 n_2}{\left(n_1 + n_2\right)\left(n_1 + n_2 - 1\right)} \sum_{i=1}^{n_1+n_2} a^2_{n_1+n_2}(i) - \frac{n_1 n_2}{n_1 + n_2 - 1}.$$

Гипотеза равенства параметров масштаба не отклоняется с достоверностью α, если $N^*_{\alpha/2} < K^* < N^*_{1-\alpha/2}$.

По своим свойствам и мощности критерий эквивалентен критерию Клотца.

5.2.6. k-выборочный критерий Флайне–Киллина

Модифицированный в [12] непараметрический критерий Флайне–Киллина [20] предназначен для проверки однородности дисперсий $k \ge 2$ выборок с объёмами n_i, $i = \overline{1,k}$. Статистика критерия формируется следующим образом.

По исходным выборкам вычисляются абсолютные значения $z_{ji} = \left| x_{ji} - \tilde{x}_i \right|$, где $\tilde{x}_i$ – выборочная медиана i-й выборки, $i = \overline{1,k}$. Далее строится вариационный ряд объединённой выборки z_{ji}, $j = \overline{1,n_i}$, $i = \overline{1,k}$. Для элементов i-й выборки на основании рангов R_{ji} её элементов z_{ji} в объединённой выборке строятся метки

$$a_{n,R_{ji}} = \Phi^{-1}\left(\frac{1 + R_{ji} / (n+1)}{2} \right), \; j = \overline{1,n_i},$$

где $n = \sum_{i=1}^{k} n_i$, находятся $\bar{A}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} a_{n,R_{ji}}$. Статистика критерия имеет вид:

$$\chi_0^2 = \frac{\sum_{i=1}^{k} n_i \left(\bar{A}_i - \bar{a} \right)^2}{V^2}, \tag{5.26}$$

где $\bar{a} = \frac{1}{n} \sum_{j=1}^{n} a_{n,j}$, $V^2 = \frac{1}{n-1} \sum_{j=1}^{n} \left(a_{n,j} - \bar{a} \right)^2$.

Асимптотическим распределением статистики (5.26) при справедливости H_0 и больших объёмах выборок является χ^2_{k-1}-распределение.

Отклонением реального распределения статистики (5.26) от χ^2_{k-1}-распределения в ситуации принадлежности выборок нормальному закону можно пренебречь при объёмах анализируемых выборок порядка 100. С ростом числа сравниваемых выборок k сходимость к соответствующему χ^2_{k-1}-распределению не улучшается, но и сильно не ухудшается.

Отметим, что при ограниченных объёмах выборок распределения $G(\chi^2_0|H_0)$ статистики (5.26) зависят от закона, которому принадлежат выборки.

Исходный критерий Флайне–Киллина [20] был построен в предположении о принадлежности анализируемых выборок симметричным законам. Именно в такой ситуации асимптотическим распределением статистики (5.26) оказывается χ^2_{k-1}-распределение. Отметим, что сходимость распределений статистики (5.26) к χ^2_{k-1}-распределению при симметричных законах, отличных от нормального, более медленная, чем при нормальном.

5.3. Сравнительный анализ мощности критериев однородности дисперсий

В связи с применением критериев однородности дисперсий в приложениях, как правило, специалистов волнуют две связанные проблемы. Первая заключается в крайней неустойчивости большей части существующих параметрических критериев однородности дисперсий, а вторая касается оценки мощности критериев.

Вопросам анализа мощности критериев однородности дисперсий, среди множества публикаций, были посвящены работы [42, 12, 67]. Эти же вопросы рассматривались в [45, 48, 114, 115, 116, 117, 80, 119, 49, 50, 51, 52]. Общие итоги сравнительного анализа множества параметрических и непараметрических критериев однородности дисперсий, подведенные в [134], кратко можно сформулировать следующим образом.

При $k=2$ параметрические критерии Бартлетта, Кокрена, Фишера, Хартли, Неймана–Пирсона и Z–критерий Оверолла–Вудворда являются эквивалентными, а при $k>2$ преимущество оказывается за критерием Кокрена. При этом мощность параметрических критериев существенно выше непараметрических аналогов.

Явное преимущество в мощности параметрических критериев заставляет рассмотреть возможность их применения в условиях нарушения классического предположения о нормальности (в условиях принадлежности выборок различным законам [115, 49, 22]).

Другую группу практически эквивалентных по мощности критериев образуют критерии Лайарда, Миллера, О'Брайена и модифицированный Z–критерий. Некоторое преимущество имеет пара критериев Лайарда, Миллера. Различие в мощности модифицированного Z–критерия и критерия О'Брайена заметно только на достаточно далёких конкурирующих гипотезах. При этом эта группа имеет преимущество в мощности по сравнению с критерием Левене. Считается, что, как и последний, эти критерии достаточно устойчивы к нарушению стандартного предположения о нормальности анализируемых выборок.

Критерий Ньюмана с ростом объёмов выборок всё заметнее уступает в мощности критерию Левене. В то же время он имеет явное преимущество в мощности по сравнению с критериями Блисса–Кокрена–Тьюки, Кадуэлла–Лесли–Брауна и Линка. Три последних эквивалентны по мощности.

Следует отметить, что критерии О'Брайена, Левене и модифицированный Z–критерий, относящиеся к группе "устойчивых" критериев, при малых объёмах выборок уступают в мощности критериям Ньюмана, Линка, Блисса–Кокрена–Тьюки и Кадуэлла–Лесли–Брауна, но с ростом n имеет явное преимущество перед последними, а также перед непараметрическими критериями.

Непараметрические критерии существенно уступают в мощности большинству из рассмотренных параметрических критериев.

Исключение составляют параметрические критерии Ньюмана, Блисса–Кокрена–Тьюки, Кадуэлла–Лесли–Брауна и Линка, имеющие некоторое преимущество в мощности над непараметрическими критериями лишь при малых объёмах выборок ($n_i = 10 \div 20$), а при увеличении n_i заметно уступают всем непараметрическим критериям.

В группе непараметрических критериев результаты анализа показывают заметное преимущество в мощности критерия Клотца. Затем следуют критерий Флайне-Киллина и критерий Муда. Ещё меньшую мощность и практическую эквивалентность демонстрируют критерии Ансари–Бредли и Сижела–Тьюки. Некоторый «разнобой» в оценках мощности критериев Муда, Ансари–Бредли и Сижела–Тьюки при объемах выборок $n = 10$ и $n = 20$ объясняется различной степенью дискретности распределений статистик этих критериев.

Критерии Бартлетта, Кокрена, Хартли, Левене, Неймана–Пирсона, Лаарда, Миллера, О'Брайена, Блисса–Кокрена–Тьюки, Кадуэлла–Лесли–Брауна, Z–критерий Оверолла–Вудворда, модифицированный Z–критерий могут применяться при числе выборок $k > 2$. При $k > 2$ критерии Бартлетта, Кокрена, Хартли, Неймана–Пирсона и Z–критерий Оверолла–Вудворда уже не образуют группу эквивалентных критериев с одинаковой мощностью. Исключение составляют лишь критерии Бартлетта и Неймана–Пирсона, которые остаются практически эквивалентными по мощности, а также пара критериев Миллера и Лайарда.

Среди многовыборочных критериев (при числе выборок $k > 2$) на первой позиции с явным преимуществом, как и было показано в [48, 114], находится критерий Кокрена. На второй оказывается критерий О'Брайена, однако в случае анализа 3-х выборок и близких конкурирующих гипотез он не имеет заметного преимущества по сравнению с Z–критерием Оверолла–Вудворда, Неймана-Пирсона и Бартлетта. В то же время критерий О'Брайена явно мощнее модифицированного Z–критерия, критериев Хартли, Миллера, Лайарда и критерия Левене. В этой связи, нелишне напомнить, что критерий О'Брайена, как и критерий Левене, является устойчивыми к нарушению стандартного предположения о нормальности.

В условиях нарушения стандартного предположения о нормальности можно констатировать, что критерии Бартлетта, Кокрена, Хартли, Фишера, Неймана–Пирсона и Z–критерий Оверолла-Вудворда остаются эквивалентными по мощности в ситуации нарушения стандартного предположения о нормальности и принадлежности 2-х анализируемых выборок некоторому симметричному закону.

Аналогично, эквивалентной по мощности остаётся группа критериев Блисса–Кокрена–Тьюки, Кадуэлла–Лесли–Брауна и Линка.

В случае принадлежности выборок законам с более “лёгкими хвостами” (по сравнению с нормальным законом) критерии упорядочиваются по мощности практически так же, как и при нормальном законе.

При (симметричных) законах с более “тяжелыми хвостами” по сравнению с нормальным законом порядок предпочтения меняется следующим образом:

Флайне–Киллина $\succ$ *Клотца* $\succ$ *Муда* $\succ$ *Левене* $\succ$ *Сижела–Тьюки* $\sim$ *Ансари–Бредли* $\succ$ *Миллера* $\succ$ *О’Брайена* $\succ$ *Лайарда* $\succ$ *Модифицированный Z–критерий* $\succ$ ***группа критериев*** *(Бартлетта, Кокрена, Хартли, Фишера, Неймана–Пирсона, Z–критерий Оверолла–Вудворда)* $\succ$ *Ньюмана* $\succ$ ***группа критериев*** *(Блисса–Кокрена–Тьюки, Кадуэлла–Лесли–Брауна, Линка)*

5.4. Пример применения критериев однородности дисперсий для проверки гипотез об отсутствии тренда

Продемонстрируем применение критериев однородности дисперсий для проверки гипотезы об отсутствии тренда в характеристиках рассеяния. Как и ранее, сделаем это на примере реализации такой возможности в программной системе [148].

Процедура загрузки анализируемой выборки и разбиение её на части, с возможно отличающимися закономерностями (вызванными наличием или отсутствием тренда) и перспективные для целей обнаружения тренда, не отличается от описанной в разделе 3.8 и показанной на рис.3.1 и 3.2. В данном случае анализируемая выборка (см. рис. 5.1) разбивается на 4 равные части.

После выбора раздела меню “Проверить однородность дисперсий” открывается вкладка (см. рис. 5.2), позволяющая выбирать выделенные для анализа части временного ряда и применяемые для этого критерии однородности дисперсий.

Выделив сравниваемые части (анализируемой выборки) и отметив используемые критерии, запускаем проверку. Результаты проверки отражаются на вкладке, показанной на рис.5.3.

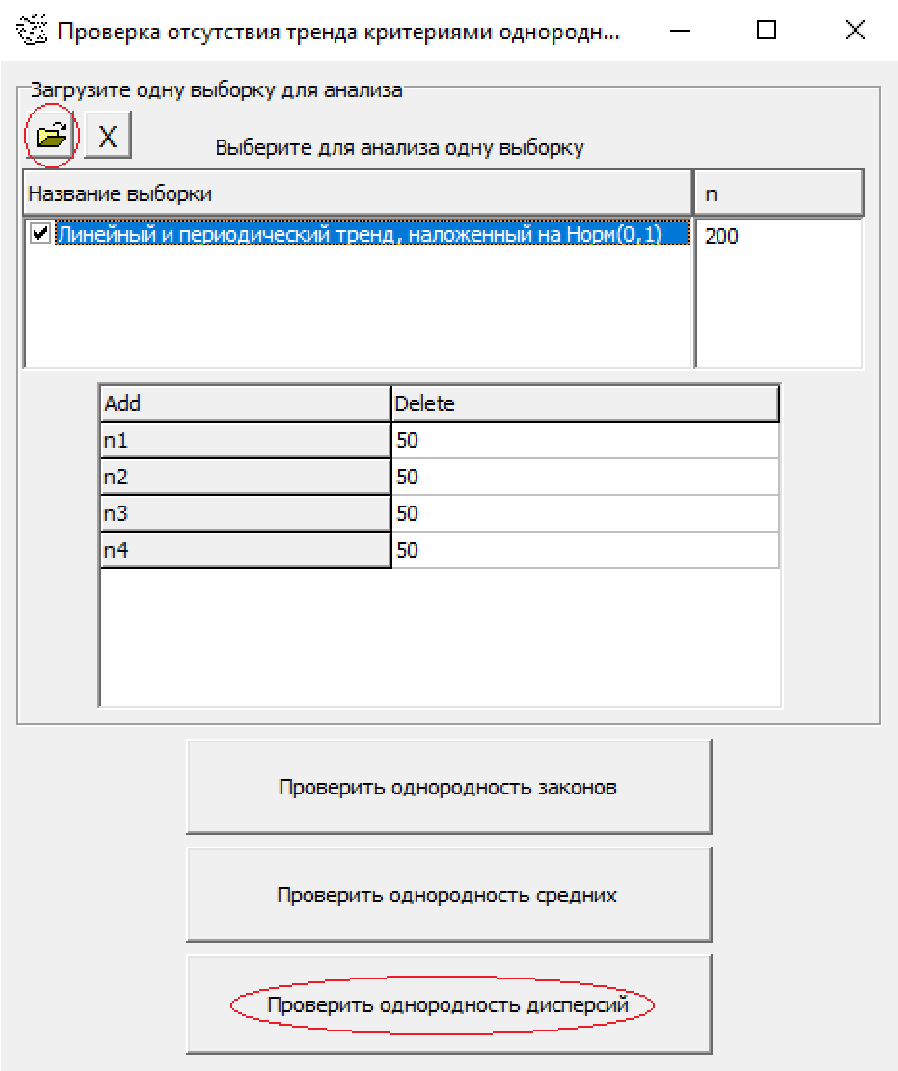


Рис. 5.1. Загрузка выборки с временным рядом измерений
и разбиение её на части

Применение критериев однородности дисперсий для проверки гипотезы об отсутствии тренда в дисперсии продемонстрируем на примере временного ряда, показанного на рис. 5.4.

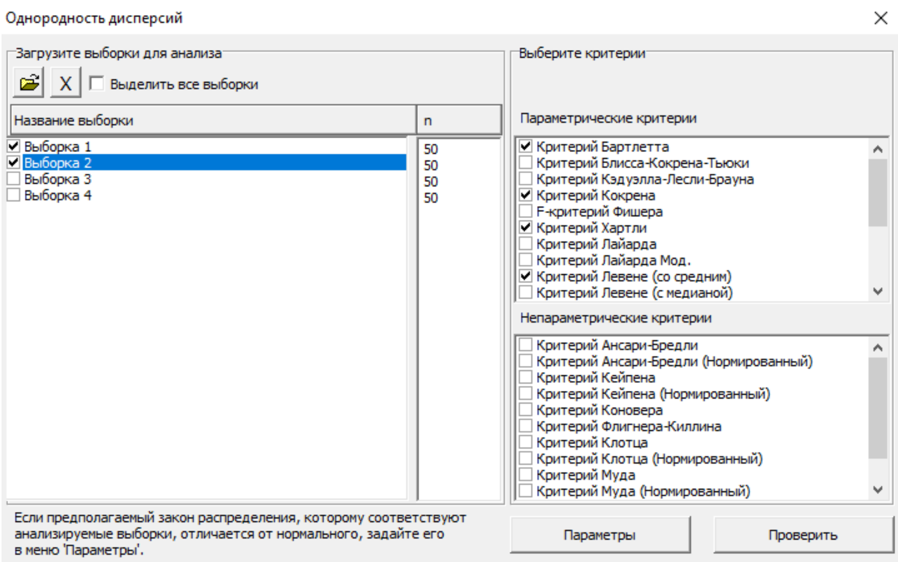


Рис. 5.2. Выбор анализируемых частей выборки используемых критериев

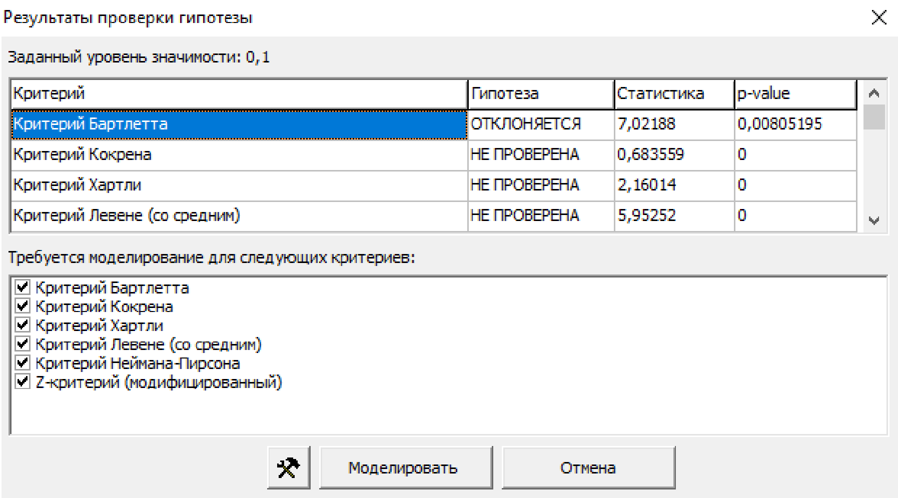


Рис. 5.3. Результаты проверки по критериям с известными асимптотическими распределениями статистик

Временной ряд представляет собой последовательность из двухсот псевдослучайных величин ξ_i, подчиняющихся стандартному нормальному закону, на который наложена комбинация линейного и периодического тренда вида $X_i = \xi_i(1+0.6t_i + 0.8\sin(2k\pi t_i))$ где $k=4$, $t_i=(i-1)\Delta t$, шаг $\Delta t = 1/n$ связан с объемом выборки n (см. раздел 1.2).

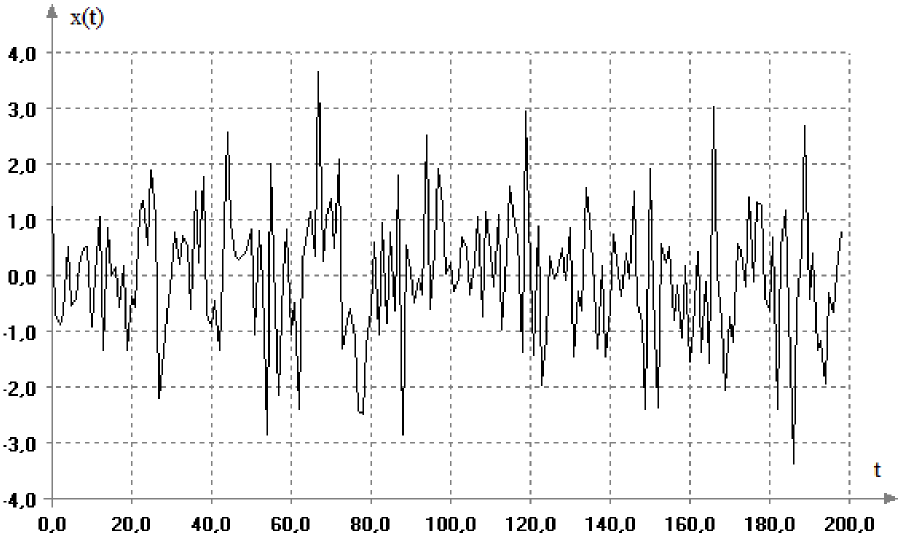


Рис. 5.4. Временной ряд измерений с линейным и периодическим трендом в дисперсии

В данном случае временной ряд разбит на 4 равные выборки.

В таблице 5.1 приведены результаты проверки однородности (по некоторым из критериев) первых двух частей ряда по 50 измерений. Оценка p_{value} для критерия Бартлетта вычислена по известному асимптотическому распределению. Оценки p_{value} для остальных критериев таблице найдены в результате интерактивного моделирования при числе статистических испытаний метода Монте–Карло $N=10^5$. В данном случае, судя по оценкам p_{value}, гипотеза об отсутствии тренда **отклоняется** по всем применяемым критериям при заданной вероятности ошибки 1-го рода $\alpha > 0.0175$.

Следует обратить внимание на то, что в случае анализа 2-х выборок критерии Бартлетта, Кокрена, Хартли и Неймана–Пирсона дают одинаковый результат. В данном случае они эквивалентны, так как для критерия Бартлетта моделирование также даёт значение $p_{value} = 0.0078$.

Таблица 5.1

Результаты проверки однородности дисперсий первых 2-х частей

№ п/п	Критерий	Статистика	p_{value}
1	Критерий Бартлетта	7.02188	0.0080
2	Критерий Кокрена	0.683559	0.0078
3	Критерий Хартли	2.16014	0.0078
4	Критерий Левене (с оценкой среднего)	5.95252	0.0174
5	Критерий Неймана–Пирсона	1.07507	0.0078
6	Z-критерий Овероллла–Вудворда	6.14013	0.0137

В таблице 5.2 при использовании несколько другого перечня критериев приведены полученные результаты проверки однородности последних 3-х частей. В данном случае гипотеза об отсутствии тренда **отклоняется** по всем применяемым критериям при заданной вероятности ошибки 1-го рода $\alpha > 0.097$.

Таблица 5.2

Результаты проверки однородности дисперсий последних 3-х частей

№ п/п	Критерий	Статистика	p_{value}
1	Критерий Бартлетта	7.64139	0.0219
2	Критерий Кокрена	0.43829	0.0969
3	Критерий Лайарда	7.35034	0.0403
4	Критерий Левене (с оценкой среднего)	3.28535	0.0416
5	Критерий О`Брайена	2.73921	0.0649
6	Z-критерий Овероллла–Вудворда (мод)	2.72528	0.0625

В таблице 5.3 приведены результаты проверки однородности всех

4-х частей по тем же, что и ранее, критериям.

И в этом случае гипотеза об отсутствии тренда **отклоняется** по всем применяемым критериям при задании вероятности ошибки 1-го рода $\alpha > 0.0295$.

Как видим, критерии однородности дисперсий являются эффективным средством обнаружения тренда в характеристиках рассеяния временных рядов.

Таблица 5.3

Результаты проверки однородности дисперсий всех 4-х выборок

№ п/п	Критерий	Статистика	p_{value}
1	Критерий Бартлетта	11.7583	0.0085
2	Критерий Кокрена	0.36436	0.0294
3	Критерий Лайарда	10.9585	0.0233
4	Критерий Левене (с оценкой среднего)	3.24762	0.0249
5	Критерий О`Брайена	3.15569	0.0242
6	Z-критерий Овероллла–Вудворда (мод)	3.03056	0.0281

Подчеркнём, что в системе [148] представленное там множество параметрических критериев проверки однородности дисперсий можно применять и в условиях нарушения стандартного предположения о нормальности. Это реализуется за счет интерактивного моделирования распределений статистик при различных параметрических моделях законов распределения.

6. О ПРИМЕНЕНИИ КРИТЕРИЕВ В УСЛОВИЯХ ОКРУГЛЕНИЯ ДАННЫХ

В последнее время всё чаще сталкиваются с необходимостью анализа очень больших последовательностей данных. При попытках применения критериев проверки статистических гипотез для анализа больших выборок, как правило, обнаруживаются проблемы, препятствующие использованию классических результатов.

Проблемы, возникающие в связи с применением для анализа больших выборок множества критериев согласия, были рассмотрены в работах [138, 58, 139]. Оказалось, что основной причиной некорректности выводов при анализе больших данных является их ограниченная точность при "неограниченных" объёмах выборок: выборки большие, но данные в них округлены. Поэтому в "больших выборках" оказывается много повторяющихся наблюдений, в результате чего с ростом n распределения статистик начинают удаляться от предельных (асимптотических) распределений этих статистик, что исключает возможность классических результатов.

В случае Big Data эта проблема решается достаточно просто. При анализе Big Data с использованием соответствующего критерия проверки гипотезы, статистика должна вычисляться не по всему большому массиву, а по выборкам, извлекаемым из "генеральной совокупности", роль которой в данном случае играет анализируемый большой массив данных. Объём извлекаемой выборки должен не превышать некоторой величины $n_{\max}$, при которой (при данной точности) распределение статистики $G(S_{n_{\max}}|H_0)$ критерия при справедливости проверяемой гипотезы H_0 ещё реально не отличается от предельного распределения $G(S|H_0)$ этой статистики.

Реальные измерения всегда являются округлёнными с некоторой погрешностью Δ. Наличие округлений может влиять на свойства статистических критериев.

Когда ошибка округления $\Delta << \sigma$, где σ – среднеквадратическое отклонение наблюдаемой величины (ошибки измерения), с ростом объёмов выборок n распределение статистики $G(S_n|H_0)$ критерия (быстро) сходится к асимптотическому (предельному) $G(S|H_0)$ и только при больших n начинает отклоняться от него.

В случае соизмеримости Δ и σ, обозначим это как $\Delta \sim \sigma$, распределение статистики $G(S_n|H_0)$ может вообще не сходиться к асимптотическому $G(S|H_0)$, а с ростом n будет лишь отдаляться от него. Так происходит с распределениями статистик одновыборочных критериев: непараметрических критериев согласия, критериев нормальности, критериев экспоненциальности [140, 141, 59, 60] и других, таким же образом в подобной ситуации меняются распределения статистик проверки гипотез об отсутствии тренда.

Влияние ошибок округления на распределения статистик критериев однородности покажем на примере двухвыборочного критерия Андерсона–Дарлинга–Петита со статистикой (3.7).

Влияние погрешности округления на распределения статистик критериев однородности законов при справедливости H_0 без потери общности рассмотрим в случае принадлежности анализируемых выборок стандартному нормальному закону.

При одинаковых ошибках округления анализируемых выборок $\Delta_1 = \Delta_2$ и объёмах $n_i = 50$ и при $\Delta_i \le 0.01\sigma$ распределения $G(A^2|H_0)$ практически не отклоняются от распределения $a2(S)$, но отклоняются при неравных Δ_i. На рис. 6.1 распределения $G(A^2|H_0)$ статистики критерия при $n_i = 50$ показаны в зависимости от Δ_2 при $\Delta_1 = 0.01\sigma$. В этом случае отклонение распределения $G(A^2|H_0)$ от $a2(S)$ при $\Delta_1 = 0.1\sigma$ еще практического значения не имеет. При тех же Δ_i с ростом n_i отклонения $G(A^2|H_0)$ от $a2(S)$ увеличивается.

Таким же образом ошибки округления сказываются на распределениях статистик других двухвыборочных (Смирнова, Лемана–Розенблатта) и k-выборочных (Андерсона–Дарлинга, Жанга и других) критериев однородности законов.

Исследование распределений статистик 2-х и k-выборочных параметрических критериев однородности средних показало, что ошибки округления результатов измерений не оказывают на них какого-либо заметного влияния.

Напротив, на распределения статистик параметрических критериев, используемых для проверки аналогичных гипотез относительно дисперсий, ошибки округления могут оказывать существенное влияние.

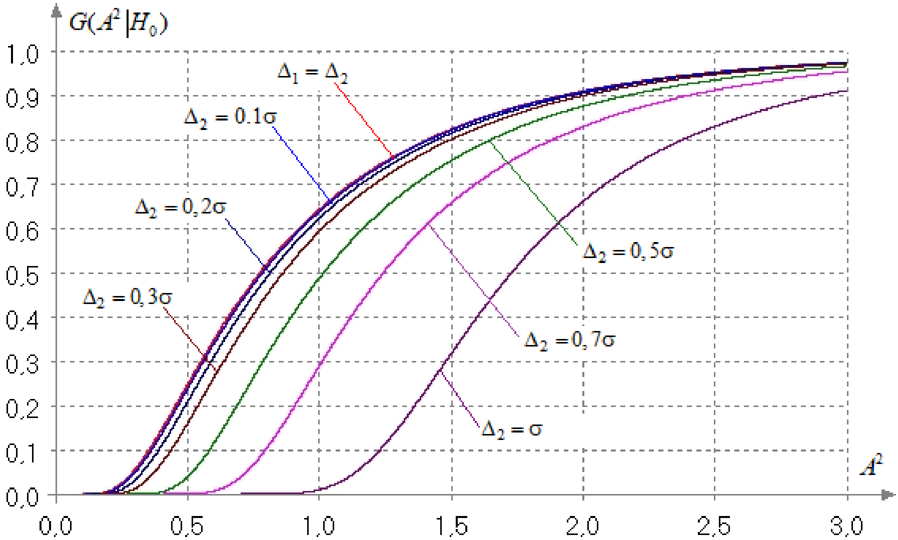


Рис. 6.1. Зависимость распределений статистики (3.7) критерия Андерсона–Дарлинга–Петита от Δ_2 (при $\Delta_1 = 0.01\sigma$)

Покажем это на критерии Бартлетта со статистикой (5.1). Асимптотическим распределением статистики критерия Бартлетта при выполнении стандартного предположения о нормальности и числе сравниваемых выборок k является χ^2_{k-1}-распределение. При наличии округления и равных Δ_i реальные распределения $G(\chi^2|H_0)$ статистики критерия не отклоняются от χ^2_{k-1}-распределения. При существенных и различающихся Δ_i реальные распределения $G(\chi^2|H_0)$ отклоняются от χ^2_{k-1}-распределения. Рис. 6.2 иллюстрирует зависимость распределения $G(\chi^2|H_0)$ статистики (5.1) критерия Бартлетта от погрешности округления Δ_2 наблюдений во второй выборке при $\Delta_1 = 0.01\sigma$ и при объёмах выборок $n_i = 1000$ в случае $k = 2$.

Как можно видеть, при существенно отличающихся Δ_1 и Δ_2 распределения $G(\chi^2 | H_0)$ значительно отклоняются от χ_1^2-распределения. Заметим, что при меньших объёмах выборок, например, при $n_i = 100$ отклонением реального распределения от асимптотического при $\Delta_2 = 0.5\sigma$ ещё можно пренебречь.

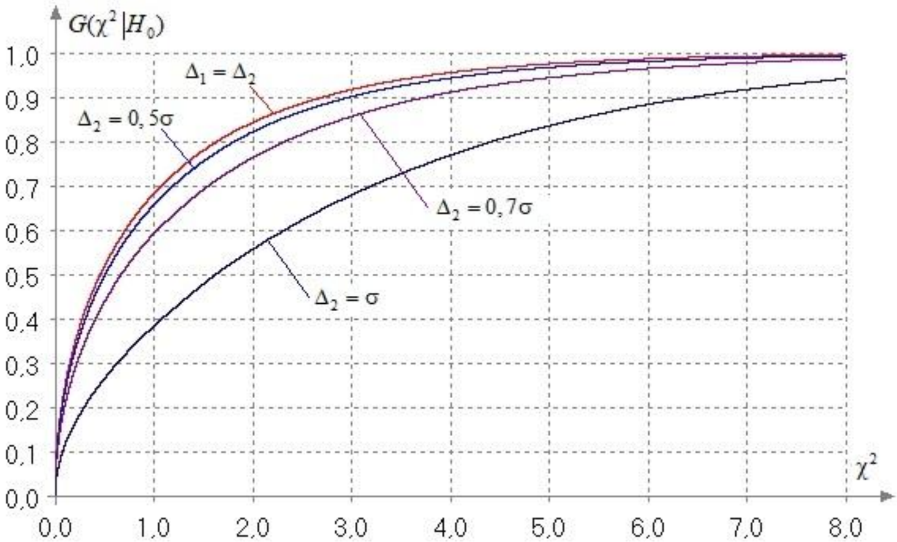


Рис. 6.2. Зависимость распределения статистики (5.1) критерия Бартлетта от Δ_2 (при $\Delta_1 = 0.01\sigma$ и $n_i = 1000$)

Если реальные распределения $G(\chi^2 | H_0)$ статистики при наличии округления и равных Δ_i не отклоняются от асимптотических, то при справедливости конкурирующих гипотез H_i распределения $G(\chi^2 | H_i)$ зависят от погрешности округления. При этом мощность критериев с ростом Δ_i снижается даже при равных Δ_i [58].

Исследования показали, что аналогичным образом при неравных Δ_i с ростом ошибок округления меняются распределения статистик других параметрических критериев однородности дисперсий, рассмотренных в [134].

Примечание. При использовании критериев однородности для проверки гипотез об отсутствии тренда сравниваются части единой анализируемой последовательности, в которой, возможно, измерения представлены с одной и той же ошибкой округления Δ. В такой ситуации распределения $G(S|H_0)$ статистик критериев однородности не должны зависеть от Δ.

В случае соизмеримости Δ и σ могут изменяться и распределения $G(S|H_0)$ статистик специальных критериев, рассмотренных в разделе 2 и ориентированных именно на проверку гипотез об отсутствии тренда (и случайности). При этом характер зависимости различается для разных критериев.

В частности, в условиях соизмеримости Δ и σ влияния величины Δ на распределения $G(S|H_0)$ статистик никак не сказывается для критериев: автокорреляции (2.2), Морана (2.4), Люнга–Бокса (2.5), Дюффа–Роя (2.6), модификация критерия автокорреляции (2.8), Вальда–Вольфовица (2.10), рангового с метками Клотца (2.30), Хсу с Н-статистикой (2.27).

Влияние на распределениях $G(S|H_0)$ сказывается незначительно и практически не влияет на результаты статистических выводов для критериев: Вальда–Вольфовица (рангового (2.12)), Бартелса (2.17), Холлина (2.19) и (2.20).

В условиях соизмеримости Δ и σ существенно меняются, например, распределения статистик следующих критериев: кумулятивной суммы (2.18), Фостера–Стюарта (2.23) и (2.24), Кокса–Стюарта (2.26), рангового с метками Сэвиджа (2.31), инверсий (2.32), сериального Вальда–Вольфовица (2.33).

В качестве примера на рис. 6.3 демонстрируется изменение распределения $G(I^*|H_0)$ статистики (2.31) критерия инверсий в зависимости от Δ при объеме выборок $n=100$.

На рис. 6.4 аналогичная картина представлена для распределений $G(S^*|H_0)$ статистики (2.26) критерия Кокса–Стюарта, используемого для проверки гипотезы об отсутствии тренда в математическом ожидании.

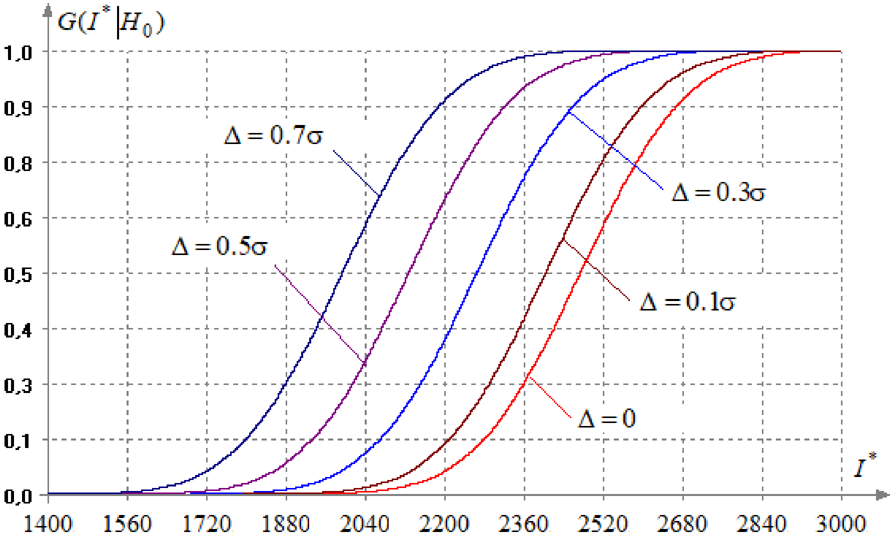


Рис. 6.3. Зависимость от Δ распределения статистики (2.32) критерия инверсий (при $n = 100$)

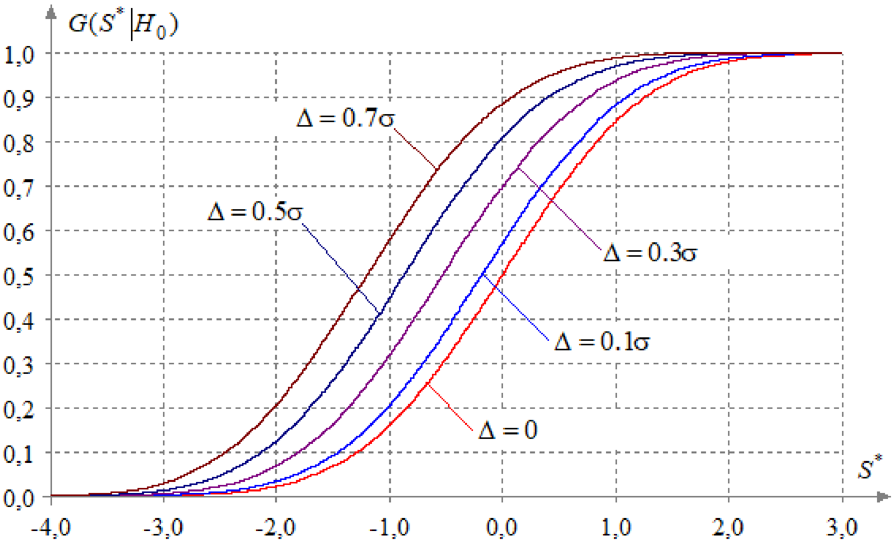


Рис. 6.4. Зависимость от Δ распределения статистики (2.26) критерия Кокса–Стюарта (при $n = 100$)

Показанная картина подчеркивает, что вследствие ошибок округления могут существенно меняться распределения статистик критериев, соответствующие справедливости проверяемой гипотезы H_0. Если пренебрегать этим фактом, то при проверке будут существенно увеличиваться вероятности ошибок 1-го рода.

При использовании конкретных критериев в приложениях, а также в автоматизированных системах обработки данных необходимо учитывать возможное влияние погрешностей округления на распределения статистик критериев.

Сигналом к осторожности в использовании классических результатов относительно применяемых критериев является наличие в анализируемых выборках слишком большого количества повторяющихся значений. Если этого нет, можно опираться на классические результаты.

Изменение свойств критерия под влиянием погрешностей округления не исключает возможность его корректного применения. Для этого достаточно смоделировать неизвестное распределение статистики $G(S|H_0)$ критерия при соответствующих Δ_i и n_i, и в результате N имитационных экспериментов найти эмпирическое распределение статистики $G_N(S_n|H_0)$. Далее, имея распределение $G_N(S_n|H_0)$, по вычисленному значению S^* статистики критерия можно найти оценку $p_{value} = 1 - G_N(S^*|H_0)$.

Такой подход к применению рассматриваемых в руководстве критериев в условиях наличия ошибок округления реализован в системе [148], с использованием которой проведены настоящие исследования. Реализация подобной программной поддержки в специализированных системах обработки данных не представляет принципиальных трудностей.

ЗАКЛЮЧЕНИЕ

Настоящее руководство позволяет ориентироваться в совокупности критериев, которые могут использоваться для проверки гипотез об отсутствии в анализируемой последовательности каких-либо случайных или неслучайных закономерностей, противоречащих предположению о том, что мы имеем дело с выборкой независимых, одинаково распределенных случайных величин.

Поводом для отклонения проверяемой гипотезы могут быть различные причины, в том числе наличие в последовательности некоторого тренда.

Представленные в руководстве описания специальных критериев, предназначенных для проверки гипотезы об отсутствии тренда, с указанием их преимуществ и недостатков, оценки мощности критериев относительно некоторых конкурирующих гипотез дают возможность специалистам, решающим задачи статистического анализа в конкретной прикладной области, осознанно подходить к выбору критериев, не останавливаясь на использовании какого-то одного.

В то же время, нельзя дать однозначного ответа на вопрос, какой критерий лучше всего использовать в целях проверки данного вида гипотез. Относительно разного вида конкурирующих гипотез критерии по мощности могут упорядочиваться различным образом.

Приходится отметить, что применение совокупности критериев при анализе одной и той же последовательности может приводить к противоположным выводам о результатах проверки. Это естественно, так как статистики критериев опираются на различные меры отклонения анализируемой последовательности от проверяемой гипотезы H_0. А причиной “разнобоя” в выводах кроется в низкой мощности критериев и в разной способности различать конкретные альтернативы.

Для отслеживания изменения закономерностей в анализируемой последовательности можно сравнивать различные части последовательности, применяя для этого критерии однородности законов, критерии однородности математических ожиданий или однородности дисперсий. В этих целях в данном руководстве рассмотрено множество критериев однородности, свойства которых подробно изложены в [134]. Практика показывает, что использование критериев однородности для анализа частей последовательности (при проверке гипотез об отсутствии тренда) обычно позволяет прийти к однозначному ответу в ситуациях, когда выводы о справедливости H_0 по специальным критериям оказываются противоположными.

Очевидно, что в настоящем руководстве представлено не всё множество критериев, которые могут использоваться для проверки рассмотренных гипотез. Что-то упущено из существующих критериев и, в то же время, предлагаются новые критерии. Однако приведенные в руководстве описания критериев и их свойств должны способствовать корректному применению данных критериев в приложениях и, следовательно, способствовать повышению качества статистических выводов.

Использование в процессе проверки гипотез процентных точек уже не соответствует современным требованиям к качеству статистических выводов. Более серьёзные основания для принятия того или иного решения о результатах проверки гипотезы даёт оценка достигнутого уровня значимости. Однако оценка p_{value} требует знания распределения статистики при справедливости проверяемой гипотезы H_0, которое не всегда известно. В этом плане можно отметить возрастающую роль компьютерных технологий и статистического моделирования и, позволяющих в интерактивном режиме исследовать распределения статистик и оценивать значение p_{value}. Подобным образом оценки p_{value} находятся в программной системе [148], которой можно воспользоваться при необходимости. Такие технологии позволяют применять статистические критерии, в том числе, в условиях нарушения стандартных предположений, когда исключается возможность использования классических результатов.

Библиографический список

1. *Ansari A.R.* Rank-tests for dispersions / A. R. Ansari, R. A. Bradley // AMS. – 1960. – Vol. 31, № 4. – P. 1174–1189.
2. *Bartlett M.S.* Properties of sufficiency of statistical tests / M. S. Bartlett // Proc. Roy. Soc. – 1937. – A 160. – P. 268–287.
3. *Bartels R.* The rank version of von Neumann's ratio test for randomness // JASA. 1982. V. 77, №377. P. 40-46.
4. *Behrens W.U.* Ein Beitrag Zur Fehlerberechnung bei wenigen Beobachtungen // Landw. Jb., 1929. B. 68, – S. 807-837.
5. *Bliss C.I., Cochran W.G., Tukey J.W.* A rejection criterion based upon the range // Biometrika. 1956. Vol. 43. No. 3/4. – P. 418-422.
6. *Brown M.B.* Robust Tests for Equality of Variances / M. B. Brown, A. B. Forsythe // J. Amer. Statist. Assoc. – 1974. – Vol. 69. – P. 364–367.
7. *Capon J.* Asymptotic efficiency of certain locally most powerful rank tests / J. Capon // AMS. – 1961. – Vol. 32, № 1. – P. 88–100.
8. *Chen S., Pokojovy M.* Modern and classical *k*-sample omnibus tests // Wiley Online Library, 2017. DOI: 10.1002/wics.1418
9. *Cochran W.G.* The distribution of the largest of a set of estimated variances as a fraction of their total / W. G. Cochran // Annals of Eugenics. – 1941. – Vol. 11. – P. 47–52.
10. *Conover W.J.* Several k-sample Kolmogorov-Smirnov tests // T*he Annals of Mathematical Statistics.* – 1965. – Vol. 36, No. 3. – P.1019-1026.
11. *Conover W.J.* Practical Nonparametric Statistics / W. J. Conover. – 3d ed. – Wiley, 1999. – 584 p.
12. *Conover W.J., Johnson M.E., Johnson M.M.* A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data // Technometrics. – 1981. – Vol. 23, No. 4. – P. 351-361.
13. *Cox D.R., Stuart A.* Quick sign tests for trend in location and dispersion // Biometrika. 1955. – V.42. – P.80-95.
14. *Cox M.G.* The evaluation of key comparison data // Metrologia. 2002. Vol. 39. No. 6. – P.589-585. DOI: 10.1088/0026-1394/39/6/10
15. *Cox M.G.* The evaluation of key comparison data: determining the largest consistent subset // Metrologia. 2007. Vol. 44. No. 3. – P.187-200. DOI: 10.1088/0026-1394/44/3/005
16. *Dufor J.–M., Roy R.* Some robust exact results on sample for randomness // J. of Econometrics. 1985. – V. 29. – P. 257-273.
17. *Fisher R. A.* The fiducial argument in statistical inference // Annals of Eugenics, Lond., 1935. Vol. **6**, No. 4. – P. 391-398.

18. *Fisher R. A.* The asymptotic approach to Behrens's integral, with further tables for the *d* test of significance // Annals of Eugenics, Lond., 1941. Vol. **11**. – P. 141-172.

19. *Fisher R.A., Yates F.* Statistical tables for biological, agricultural and medical research. London & Edinburgh: Oliver and Boyd, 1948.

20. *Fligner M.A., Killeen T.J.* Distribution-Free Two-Sample Tests for Scale // Journal of American Statistical Association. 1976. Vol. 71. No. 353. – P.210-213.

21. *Foster F.G., Stuart A.* Distribution-free tests in time series dated on the breaking of records // JRSS. 1954. – V. B16, №1. – P.1-22.

22. *Gorbunova A.A., Lemeshko B.Yu.* Application of Parametric Homogeneity of Variances Tests under Violation of Classical Assumption // Proceedings, 2nd Stochastic Modeling Techniques and Data Analysis International Conference. 5 - 8 June 2012, Chania, Crete, Greece. P.253-260. (http://www.smtda.net/images/1_SMTDA2012_Proceedings_D-J_119-338.pdf)

23. *Hartley H. O.* The maximum F-ratio as a short-cut test of heterogeneity of variance / H. O. Hartley // Biometrika. – 1950. – Vol. 37. – P. 308–312.

24. *Hines W.* Increased power with modified forms of the Levene (Med) test for heterogeneity of variance / W. Hines, R. Hines // Biometrics. – 2000. – Vol. 56. – P. 451–454.

25. *Hoeffding W.* Optimum non-parametric test // Proc.11 th Berkeley Symp., 1950. – P.83-82.

26. *Hollander M.* Non-parametric Statistical Methods / M. Hollander, D. A. Wolfe. – 2nd ed. – New York : Wiley, 1999.

27. *Hollin M., Ingeubleek J.-F., Puri M.L.* Linear serial rank tests for randomness against ARMA alternatives // Annal. Statist. 1985. V. 13. – P. 1156-1181.

28. *Hollin M., Ingeubleek J.-F., Puri M.L.* Linear and quadratic serial rank tests for randomness against social dependence // J. Of Time Serial Annal. 1987. V. 8. – P. 409-424.

29. *Hollin M., Merald G.* Rank-tests for randomness against first order serial dependence // JASA. 1988. V. 83. – P.1117-1129.

30. *Hollin M., Puri M.L.* Optimal rank-based procedures for time series analysis: Testing on ARMA model against other ARMA models // Annal. Statist. 1988. V. 16. – P. 402-432.

31. *Hollin M., Laforet A., Merald G.* Distribution-free tests against dependence: signed or unsigned ranks? // J. of Stat. Planning and Inference. 1990. V. 24. – P. 151-165.

32. *Hsieh H.K.* Nonparametric tests for scale shift at a unknown time point // Communications in Statistics - Theory and Methods, 1984. – V.13. № 11. – P.1335-1355.

33. *Hsu D.A.* Test for variance shift at an unknown time point / D. A. Hsu // Appl. Statist., 1977. – V.26, № 3. – P.279-284.

34. *Kiefer J.* K-Sample Analogues of the Kolmogorov-Smirnov and Cramer-v. Mises Tests // Annals of Mathematical Statistics**,** 1959. Vol. 30. No. 2. – P. 420-447.

35. *Klotz J.* Nonparametric tests for scale / J. Klotz // AMS. – 1962. – Vol. 33. – P. 498–512.

36. *Knoke J.D.* Testing for randomness against autocorrelation: The parametric case // Biometrica. 1975. – V.62. – P. 571-575.

37. *Knoke J.D.* Testing for randomness against autocorrelation: Alternative tests // Biometrica. 1977. – V. 64, №3. – P. 523-529.

38. *Kruskal W. H.* Use of ranks in one-criterion variance analysis / W. H. Kruskal, W. A. Wallis // J. Amer. Statist. Assoc. – 1952. – Vol. 47. – P. 583–621.

39. *Kruskal W. H.* Use of ranks in one-criterion variance analysis / W. H. Kruskal, W. A. Wallis // J. Amer. Statist. Assoc. – 1953. – Vol. 48. – P. 907–911.

40. *Laubsher N. F.* Exact critical Values for Mood`s distribution-free test statistic for dispersion and its normal approximation / N. F. Laubsher, F. E. Steffens, E. M. De Lange // Technometrics. – 1968. – Vol. 10, № 3. – P. 497–508.

41. *Layard M.W.J.* Robust large-sample tests for homogeneity of variances // Journal of the American Statistical Association. – 1973. – Vol. 68. – P. 195–198.

42. *Lee H.B., Katz G.S., Restori A.F.* A Monte Carlo Study of Seven Homogeneity of Variance Tests // Journal of Mathematics and Statistics. 2010. Vol. 6, No. 3. P. 359-366.

43. *Lehmann E.L.* Consistency and unbiasedness of certain nonparametric tests / E. L. Lehmann // Ann. Math. Statist. – 1951. – Vol. 22, № 1. – P. 165–179.

44. *Lemeshko B. Y.* Application of Homogeneity Tests: Problems and Solution / B. Y. Lemeshko, I. V. Veretelnikova, S. B. Lemeshko, A. Y. Novikova // In: Rykov V., Singpurwalla N., Zubkov A. (eds) Analytical and Computational Methods in Probability Theory. ACMPT 2017. Lecture Notes in Computer Science. : monograph. - Cham : Springer, 2017. - 10684. - P. 461-475.

45. *Lemeshko B.* Bartlett and Cochran tests in measurements with probability laws different from normal / B. Lemeshko, E. Mirkin // Measurement Techniques. – 2004. – Vol. 47, № 10. – P. 960–968.

46. *Lemeshko B. Yu.* Power and robustness of criteria used to verify the homogeneity of means / B.Yu. Lemeshko, S.B. Lemeshko // Measurement Techniques. – 2008. – Vol. 51, № 9. – P. 950–959.

47. *Lemeshko B. Yu.* Statistical distribution convergence and homogeneity test power for Smirnov and Lehmann–Rosenblatt tests / B. Yu. Lemeshko, S. B. Lemeshko // Measurement Techniques – 2005. – Vol. 48, № 12. – P. 1159–1166.

48. *Lemeshko B.Yu.* Application and power of criteria for testing the homogeneity of variances. Part I. Parametric criteria / B. Yu. Lemeshko, S. B. Lemeshko, A. A. Gorbunova // Measurement Techniques. – 2010. – Vol. 53, № 3. – P. 237–246.

49. *Lemeshko B.Yu., Lemeshko S.B., and A.A. Gorbunova*. Application and power of criteria for testing the homogeneity of variances. Part II. Nonparametric criteria // Measurement Techniques, Vol. 53, No. 5, 2010. P.476-486. DOI: 10.1007/s11018-010-9530-x

50. *Lemeshko B.Y., Sataeva T.S.* Application and Power of Parametric Criteria for Testing the Homogeneity of Variances. Part III // Measurement Techniques, 2017. Vol. 60. No. 1. – P. 7-14. DOI: 10.1007/s11018-017-1141-3

51. *Lemeshko B.Y., Sataeva T.S.* Application and Power of Parametric Criteria for Testing the Homogeneity of Variances. Part IV // Measurement Techniques, 2017. Vol. 60. No. 5. – P. 425-431. DOI: 10.1007/s11018-017-1213-4

52. *Lemeshko B.Yu., Sataeva T.S.* On the Properties and Application of Tests for Homogeneity of Variances in the Problems of Metrology and Control // Advances in Intelligent Systems and Computing, Vol. 543, 2017. – P. 784-798. DOI: 10.1007/978-3-319-48923-0_84

53. *Lemeshko B.Yu., Gorbunova A.A., Lemeshko S.B., Rogozhnikov A.P.* Application of Nonparametric Goodness-of-fit tests for Composite Hypotheses in Case of Unknown Distributions of Statistics // Proceedings of the International Workshop “Applied Methods of Statistical Analysis. Applications in Survival Analysis, Reliability and Quality Control” – AMSA’2013, Novosibirsk, Russia, 25-27 September, 2013. P. 8-24.

54. *Lemeshko B.Yu., Lemeshko S.B., Rogozhnikov A.P.* Real-Time Studying of Statistic Distributions of Non-Parametric Goodness-of-Fit Tests when Testing Complex Hypotheses // Proceedings of the International Workshop “Applied Methods of Statistical Analysis. Simulations and Statistical Inference” – AMSA’2011, Novosibirsk, Russia, 20-22 September, 2011. P. 19-27.

55. *Lemeshko B.Yu.* Application of nonparametric Kuiper and Watson tests of goodness-of-fit for composite hypotheses / B. Yu. Lemeshko, A. A. Gorbunova // Measurement Techniques. – 2013. – Vol. 56. – № 9. – P.965-973.

56. *Lemeshko B.Yu.* Interactive investigation of statistical regularities in testing composite hypotheses of goodness of fit / B. Yu. Lemeshko, S. B. Lemeshko, A. P. Rogozhnikov // Statistical Models and Methods for Reliability and Survival Analysis : monograph. – Wiley-ISTE, 2013. – Chap. 5. – P. 61–76.

57. *Lemeshko B.Yu.* Solving problems of using some nonparametric goodness-of-fit tests / B.Yu. Lemeshko, A.A. Gorbunova, S.B. Lemeshko, A.P. Rogozhnikov // Optoelectronics, Instrumentation and Data Processing. – 2014. – Vol. 50, № 1. – P.21-35.

58. *Lemeshko B., Lemeshko S., Semenova M. Features* of testing statistical hypotheses under big data analysis // Applied Methods of Statistical Analysis. Statistical Computation and Simulation - AMSA'2019, Novosibirsk, Russia, 18-20 September, 2019: Proceedings of the International Workshop. - Novosibirsk: NSTU publisher, 2019. - P.122-137.

59. *Lemeshko B.Yu., Lemeshko S.B.* Effect of the Roundoff on the Properties of Criteria for Testing Statistical Hypotheses // Optoelectronics, Instrumentation and Data Processing. 2020. – Vol. 56. – No. 3. – P.35-45. DOI: 10.3103/S8756699020030103

60. *Lemeshko B.Y., Lemeshko S.B.* About the effect of rounding on the properties of tests for testing statistical hypotheses // Journal of Physics: Conference Series. – 2021. – Vol. 1715. 012063. DOI: 10.1088/1742-6596/1715/1/012063

61. *Lemeshko B.Yu., Veretel'nikova I.V.* Power of k-sample tests aimed at checking the homogeneity of laws // Measurement Techniques, Vol. 61, No. 7, October, 2018. – P. 647-654. DOI: 10.1007/s11018-018-1479-1

62. *Lemeshko B.Yu., Veretelnikova I.V.* On Some New k-Samples Tests for Testing the Homogeneity of Distribution Laws // 2018 14th International scientific-technical conference on actual problems of electronic instrument engineering (APEIE) – 44894 Proceedings. Vol.1, Part 4, Novosibirsk, 2018. – P.153-157. IEEE Catalog Number CFP18471-PRT. **DOI:** 10.1109/APEIE.2018.8545256

63. *Lemeshko B., Veretelnikova I.* On application of k-samples homogeneity tests // Applied Methods of Statistical Analysis. Statistical Computation and Simulation - AMSA'2019, Novosibirsk, Russia, 18-20 September, 2019: Proceedings of the International Workshop. - Novosibirsk: NSTU publisher, 2019. - P.138-151.

64. *Leslie R.T., Brown B.M.* Use of range in testing heterogeneity of variance // Biometrika. 1966. Vol. 53. No.1/2. – P. 221-227.

65. *Levene H.* Robust tests for equality of variances // Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling. – 1960. – P. 278-292.

66. *Levene* Test for Equality of Variances [Электронный ресурс] // e-Handbook of Statistical Methods. – Режим доступа : http://www.itl.nist.gov/div898/handbook/ eda/section3/eda35a.htm. –Загл. с экрана.

67. *Lim T.-S., Loh W.-Y.* A comparison of tests of equality of variances // Computational Statistics & Data Analysis. 1996. Vol. 22, No. 5. P. 287-301.

68. *Link R.F.* The sampling distribution of the ratio of two ranges from independent samples // The annals of mathematical statistics. 1950. Vol. 21, No. 1. – P. 112-116.

69. Ljung G. M., Box G. E. P. On a measure of lack of fit in time series models // Biometrika. 1978. – V. 65. – P. 297-303.

70. *Mann H. B.* On a test of whether one of two random variables is stochastically larger than the other / H. B. Mann, D. R. Whitney // Ann. Math. Statist. – 1947. – Vol. 18. – P. 50–60.

71. Mc Gielchrist C.A., Woodyer K.D. Note on a distribution-free CISIM technique // Technometrics. 1975. V. 17, №3. – P. 321-325.

72. *Miller R.G.* Jackknifing variances // The Annals of Mathematical Statistics – 1968. – Vol. 39. – P. 567–582.

73. *Milton R. C.* An extended table of critical values for the Mann –Whitney (Wilcoxon) two-sample statistic / R. C. Milton // J. Amer. Statist. Ass. – 1964. – Vol. 59. – P. 925–934.

74. *Mood A.* On the asymptotic efficiency of certain nonparametric tests / A. Mood // AMS. – 1954. – Vol. 25. No. 3. – P. 514–522.

75. Moran P. A. P. Some theoremson time series 2: The siginificance of th serial correlations coefficient // Biometrika. 1948. – V. 35. – P. 255-260.

76. *Neel J.H.* A Monte Carlo Study of Levene`s Test of Homogeneity of Variance: Empirical Frequencies of Type I Error in Normal Distributions Paper : presented at the Annual Meeting of the American Educational Research Association Convention / J.H. Neel, W.M. Stallings. – Chicago, Illinois, 1974. – April.

77. *Newman D.* The distribution of range in samples from a normal population, expressed in terms of an independent estimate of standard deviation // Biometrika. 1939. Vol. 31. No.1/2. – P. 20-30.

78. *Pettitt A.N.* A two-sample Anderson-Darling rank statistic // Biometrika. 1976. Vol. 63. No.1. P. 161-168.

79. *Rosenblatt M.* Limit theorems associated with variants of the von Mises statistic / M. Rosenblatt // Ann. Math. Statist. – 1952. – Vol. 23. – P. 617–623.

80. *Sataeva T. S., Lemeshko B.Yu.* About properties and power of classical tests of homogeneity of variances // Proceedings 2016 11th International Forum on Strategic Technology (IFOST), June 1-3, 2016, Novosibirsk, Russia. Part 1. – P. 350-354. DOI: 10.1109/IFOST.2016.7884125

81. *Scheffe H.* On Solutions of the Behrens-Fisher Problem, Based on the t-Distribution // The Annals of Mathematical Statistics. 1943. Vol. 14, No. 1. – P. 35-44.

82. *Scheffe H.* Practical Solutions of the Behrens-Fisher Problem // Journal of the American Statistical Association. 1970. Vol. 65. No. 332. – P. 1501-1508.

83. *Scholz F.W., Stephens M.A.* K-Sample Anderson–Darling Tests // Journal of the American Statistical Association. 1987. Vol. 82. No. 399. – P. 918-924.

84. *Siegel S.* A nonparametric sum of rank procedure for relative spread in unpaired samples / S. Siegel, J.W. Tukey // JASA. – 1960. – Vol. 55, № 291. – P. 429–445.

85. *Sukhatme B. V.* On certain Two-sample nonparametric tests for variances / B. V. Sukhatme // AMS. – 1957. – Vol. 28, № 1. – P. 188–194.

86. *Terry M.E.* Some rank order test which are most powerful against specific parametric alternatives // The Annals of Mathematical Statistics. 1952. Vol. 23. – P.346-366.

87. *O'Brien R.G.* Robust techniques for testing heterogeneity of variance effects in factorial designs // Psychometrika. 1978. Vol. 43, No. 3. P. 327-342.

88. *Overall J.E., Woodward J.A.* A simple test for heterogeneity of variance in complex factorial design // Psychometrika. 1974. Vol. 39. No. 3. – P. 311-318.

89. *Overall J.E., Woodward J.A.* A robust and powerful test for heterogeneity of variance // University of Texas Medical Branch Psychometric Laboratory. 1976.

90. *Parra-Frutos I.* The behaviour of the modified Levene's test when data are not normally distributed // Computational Statistics. – 2009. – Vol. 24. – P. 671–693.

91. *Veretel`nikova I.V., Lemeshko B.Yu.* The analytical review of tests for randomness and the absence of a trend // 2014 12th International conference on actual problems of electronics instrument engineering (APEIE) 34006 Proceedings. Vol. 1. Novosibirsk, 2014. – P.532-539.

92. *Wald A., Wolfowitz J.* On a test whether two samples are from the same population / A. Wald, J. Wolfowitz // Ann. Math Statist. 1940. Vol. 11. – P. 147-162.

93. *Welch B. L.* The Significance of the Difference Between Two Means when the Population Variances are Unequal // Biometrika. 1938. Vol. 29, No. 3/4. – P. 350-362.

94. *Welch B. L.* The generalization of "Student's" problem when several different population variances are involved // Biometrika. 1947. Vol. 34, No. 1/2. – P. 28-35.

95. *Wilcoxon F.* Individual comparisons by ranking methods / F. Wilcoxon // Biometrics Bulletin. – 1945. – № 1. – P. 80–83.

96. Woodward R.H., Goldsmith P.L. Cumulative sum techniques. I.C.I. Monograph. №3, Oliver and Boyd, 1964.

97. *Zhang J.* Powerful goodness-of-fit and multi-sample tests / J. Zhang // PhD Thesis. York University, Toronto. 2001. – 113 p. URL: http://www.collectionscanada.gc.ca/obj/s4/f2/dsk3/ftp05/NQ66371.pdf (дата обращения 28.01.2013).

98. *Zhang J.* Powerful Two-Sample Tests Based on the Likelihood Ratio / J. Zhang // Technometrics. – 2006. – V. 48. – No. 1. – P.95-103. DOI 10.1198/004017005000000328

99. *Zhang J., Wu Y.* k-Sample tests based on the likelihood ratio // Computational Statistics & Data Analysis. – 2007. – V. 51. – No. 9. – P. 4682-4691.

100. *Андерсон Т.* Статистический анализ временных рядов. – М.: Мир, 1976. – 756 с.

101. *Большев Л. Н.* Таблицы математической статистики / Л. Н. Большев, Н. В. Смирнов. – М. : Наука, 1983. – 416 с.

102. *Ван дер Варден Б.Л.* Математическая статистика. – М.: Иностранная литература, 1960. – 435 с.

103. *Веретельникова И.В., Лемешко Б.Ю.* Аналитический обзор критериев проверки случайности и отсутствия тренда // Труды XII международной

конференции “Актуальные проблемы электронного приборостроения” АПЭП-2014. Т.6, Новосибирск, 2014. – С.16-23.

104. *Веретельникова И.В., Лемешко Б.Ю.* К вопросам применения критериев проверки случайности и отсутствия тренда // Материалы Российской НТК “Обработка информации и математическое моделирование”, Новосибирск. 2014. – С.25-28.

105. *Веретельникова И.В. Лемешко Б.Ю.* О мощности критериев случайности и отсутствия тренда в характеристиках рассеяния // Труды XIII международной конференции “Актуальные проблемы электронного приборостроения” АПЭП-2016. Т.8, Новосибирск, 2016. – С.113-118.

106. *Веретельникова И.В., Лемешко Б.Ю.* О критериях проверки отсутствия тренда в характеристиках рассеяния // Материалы Российской НТК “Обработка информации и математическое моделирование”, Новосибирск. 2015. – С. 42-53.

107. *Гаек Я., Шидак З.* Теория ранговых критериев. – Главная редакция физико-математической литературы изд-ва «Наука», 1971. – 376 с.

108. *Закс Л.* Статистическое оценивание / Л. Закс. – М. : Статистика, 1976. – 598 с.

109. *Кобзарь А.И.* Прикладная математическая статистика. Для инженеров и научных работников / А.И. Кобзарь. – М. : Физматлит, 2006. – 816 с.

110. *Крамер Г.* Математические методы статистики. – М.: Мир, 1975. – 648 с.

111. *Лемешко Б.Ю.* О сходимости распределений статистик и мощности критериев однородности Смирнова и Лемана–Розенблатта / Б. Ю. Лемешко, С. Б. Лемешко // Измерительная техника. – 2005. – № 12. – С. 9–14.

112. *Лемешко Б.Ю.* О применении критериев проверки однородности законов распределения / Б.Ю. Лемешко, С.Б. Лемешко, И.В. Веретельникова // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2017. – № 41. – С. 24-31.

113. *Лемешко Б.Ю.* Об устойчивости и мощности критериев проверки однородности средних / Б.Ю. Лемешко, С.Б. Лемешко // Измерительная техника. – 2008. – № 9. – С. 23–28.

114. *Лемешко Б.Ю.* О применении и мощности критериев проверки однородности дисперсий. Ч. I. Параметрические критерии / Б.Ю. Лемешко, С.Б. Лемешко, А.А. Горбунова // Измерительная техника. – 2010. – № 3. – С. 10–16.

115. *Лемешко Б.Ю.* О применении и мощности критериев проверки однородности дисперсий. Ч. II. Непараметрические критерии / Б.Ю. Лемешко, С.Б. Лемешко, А.А. Горбунова // Измерительная техника. – 2010. – № 5. – С. 11–18.

116. *Лемешко Б.Ю., Сатаева Т.С.* Применение и мощность параметрических критериев проверки однородности дисперсий. Ч. 4 // Измерительная техника. – 2017. – № 5. – С. 12-17.

117. *Лемешко Б.Ю., Сатаева Т.С.* Применение и мощность параметрических критериев проверки однородности дисперсий. Ч. 4 // Измерительная техника. – 2017. – № 5. – С. 12-17.

118. *Лемешко Б.Ю., Лемешко С.Б., Постовалов С.Н., Чимитова Е.В.* Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: Монография. Новосибирск : Изд-во НГТУ, 2011. 888 с.

119. *Лемешко Б. Ю.* Критерии Бартлетта и Кокрена в измерительных задачах при вероятностных законах, отличающихся от нормального / Б.Ю. Лемешко, Е.П. Миркин // Измерительная техника. – 2004. – № 10. – С. 10–16.

120. *Лемешко Б.Ю.* Корреляционный анализ наблюдений многомерных случайных величин при нарушении предположений о нормальности / Б.Ю. Лемешко, С.С. Помадин // Сибирский журнал индустриальной математики. 2002. – Т. 5, № 3. – С. 115–130.

121. *Лемешко Б.Ю., Комиссарова А.С., Щеглов А.Е.* Применение некоторых критериев проверки гипотез случайности и отсутствия тренда // Метрология. 2010. № 12. – С. 3-25.

122. *Лемешко Б.Ю., Комиссарова А.С., Щеглов А.Е.* Свойства и мощность некоторых критериев случайности и отсутствия тренда // Научный вестник НГТУ. – 2012. – № 1(46). – С. 53-66.

123. *Лемешко Б.Ю.* Исследование распределений статистик корреляционного анализа при отклонении многомерного закона от нормального / Б.Ю. Лемешко, С.С. Помадин // Тр. V международной конференции "Актуальные проблемы электронного приборостроения" АПЭП-2000. Новосибирск, 2000. - Т. 7. - С. 184-187.

124. *Лемешко Б.Ю.* Корреляционный анализ наблюдений многомерных случайных величин при нарушении предположений о нормальности / Б.Ю. Лемешко, С.С. Помадин // Сибирский журнал индустриальной математики. 2002. - Т.5. - № 3. - С.115-130.

125. *Лемешко Б.Ю.* Корреляционный анализ многомерных случайных величин при нарушении предположений о нормальности / Б.Ю. Лемешко, С.С. Помадин // Труды 10-го Юбилейного Международного Симпозиума по непараметрическим и робастным методам в кибернетике. Томск. – 2004. – С. 114-128.

126. *Лемешко С.Б.* Критерий независимости Аббе при нарушении предположений нормальности / С. Б. Лемешко // Измерительная техника, 2006. № 10. - С.9-14.

127. *Лемешко Б.Ю., Помадин С.С.* Проверка гипотез о математических ожиданиях и дисперсиях в задачах метрологии и контроля качества при вероятностных законах, отличающихся от нормального // Метрология. 2004. – № 3. – С.3-15.

128. *Лемешко Б.Ю., Постовалов С.Н.* Компьютерные технологии анализа данных и исследования статистических закономерностей: Учеб. пособие. – Новосибирск: Изд-во НГТУ, 2004. – 120 с.

129. *Лемешко Б.Ю.* О решении проблем применения некоторых непараметрических критериев согласия / Б.Ю. Лемешко, А.А. Горбунова, С.Б. Лемешко, А.П. Рогожников // Автометрия. 2014. – **Т.**50. – № 1. – С.26-43.

130. *Лемешко Б.Ю.* Непараметрические критерии согласия: Руководство по применению: Монография / Б.Ю. Лемешко.– М.: ИНФРА-М, 2014. – 163 с.

131. *Лемешко Б.Ю., Блинов П.Ю.* Критерии проверки отклонения распределения от равномерного закона. Руководство по применению: Монография / Б.Ю. Лемешко. П.Ю. Блинов. – М.: НИЦ ИНФРА-М. 2015. – 183 с. – (Научная мысль).

132. *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).

133. *Лемешко Б.Ю.* Критерии проверки отклонения от экспоненциального закона. Руководство по применению : монография / Б.Ю. Лемешко, П.Ю. Блинов. – Москва : ИНФРА-М, 2021. – 352 с. – (Научная мысль).

134. *Лемешко Б.Ю.* Критерии проверки гипотез об однородности. Руководство по применению : монография / Б.Ю. Лемешко. – 2-е изд., перераб. и доп. – Москва : ИНФРА-М, 2021. – 248 с. – (Научная мысль).

135. *Лемешко Б.Ю.* Мощность k-выборочных критериев проверки однородности законов /Б.Ю. Лемешко, И.В. Веретельникова // Измерительная техника. – 2018. – № 7. – С. 3-7. DOI: *10.32446/0368-1025it-2018-7-3-7*

136. *Лемешко Б.Ю., Лемешко С.Б., Веретельникова И.В.* Расширение области применения критериев однородности // Обработка информации и математическое моделирование : материалы Рос. на-уч.-техн. конф. [Новосибирск, 23–24 апр. 2020 г.]. – Новосибирск : СибГУТИ, 2020. – С. 100-105.

137. *Лемешко Б.Ю.* О применении критериев проверки однородности средних / Б.Ю. Лемешко, С.Б. Лемешко, И.В. Веретельникова, А.Ю. Новикова // Вестник СибГУТИ. 2018. – № 1. – С. 41-55.

138. *Лемешко Б.Ю. Лемешко С.Б., Семёнова М.А.* К вопросу статистического анализа больших данных // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2018. – № 44. – С. 40-49. DOI: 10.17223/19988605/44/5

139. *Лемешко Б.Ю., Лемешко С.Б., Веретельникова И.В., Блинов П.Ю.* Критерии проверки статистических гипотез при анализе больших выборок: проблемы и их решение // Марчуковские научные чтения. 2019 : Труды

Международной конференции "Актуальные проблемы вычислительной и прикладной математики" / Ин-т вычислительной математики и матем. геофизики СО РАН. Новосибирск, 1-5 июля 2019 г. / Новосибирск : ИПЦ НГУ. - С. 277-283.

140. *Лемешко Б.Ю., Лемешко С.Б.* Влияние округления на свойства критериев проверки статистических гипотез // Автометрия. 2020. – Т. 56, № 3. – С. 35-45. DOI: 10.15372/AUT20200305

141. *Лемешко Б.Ю., Лемешко С.Б.* О влиянии ошибок округления на распределения статистик критериев согласия // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2020. – № 53. – С. 47-60. DOI: 10.17223/19988605/53/5

142. *Ликеш И., Ляга Й*. Основные таблицы математической статистики. – М.: Финансы и статистика, 1985.

143. *Миттаг Х.-Й.* Статистические методы обеспечения качества / Х.-Й. Миттаг, Х. Ринне. – М. : Машиностроение. 1995. – 600 с.

144. *Орлов А.И*. О проверке однородности двух независимых выборок / А. И. Орлов // Завод. лаб. – 2003. – Т. 69, №. 1. – С. 55–60.

145. Р 50.1.033–2001. Рекомендации по стандартизации. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Ч. I. Критерии типа хи-квадрат. – М. : Изд-во стандартов, 2002. – 87 с.

146. Р 50.1.037–2002. Рекомендации по стандартизации. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Ч. II. Непараметрические критерии. – М. : Изд-во стандартов, 2002. – 64 с.

147. *Смирнов Н.В.* Оценка расхождения между эмпирическими кривыми распределения в двух независимых выборках / Н.В. Смирнов // Бюл. МГУ, Серия А. – 1939. – Т. 2, № 2. – С. 3–14.

148. Статистический анализ интервальных наблюдений одномерных непрерывных случайных величин "Интервальная статистика 5.4" / Лемешко Б.Ю., Лемешко С.Б., Блинов П.Ю., Веретельникова И.В., Новикова А.Ю. // Свидетельство о регистрации программы для ЭВМ RU 2018666213, 13.12.2018. Заявка № 2018663206 от 22.11.2018. https://ami.nstu.ru/~headrd/ISW.htm. (дата обр. 24.01.2021)

149. *Химмельблау Д*. Анализ процессов статистическими методами. – М.: Мир, 1973.

150. *Хальд А*. Математическая статистика с техническими приложениями. – М.: ИЛ, 1956.