

Санкт-Петербургский государственный университет
Кафедра математической теории игр и статистических решений

Заковряшин Егор Михайлович

Магистерская диссертация

**Анализ взаимосвязи волнения морской
поверхности и характеристик радиосигнала
методами прикладной статистики**

Направление 01.04.02

Прикладная математика и информатика

Магистерская программа

Исследование операций и системный анализ

Научный руководитель,
кандидат физ.-мат. наук,
старший преподаватель
Кумачева С. Ш.

Санкт-Петербург

2018

Содержание

Введение	3
Актуальность	4
Постановка задачи	6
Глава 1. Анализ эмпирических данных по волнению морской поверхности	7
1.1 Первичный анализ данных	7
1.2 Проверка гипотез статистической однородности на выборках волнения	10
1.3 Проверка гипотез согласия на выборках волнения	11
1.4 Удаление выбросов	14
1.5 Выводы из анализа эмпирических данных по волнению морской поверхности	17
Глава 2. Анализ эмпирических данных по замиранию амплитуды радиосигнала	18
2.1 Первичная обработка и анализ данных	18
2.2 Гипотезы согласия выборок амплитуды	22
2.3 Классификация результатов. Выводы из анализа эмпирических данных по замиранию амплитуды радиосигнала	28
Глава 3. Альтернативные методы исследования эмпирических данных	30
3.1 Исследование выборок кривыми Пирсона	30
3.1.1 Обзор кривых Пирсона	30
3.1.2 Классификация данных по семействам кривых Пирсона . .	31
3.1.3 Исследование сдвигов	35
3.1.4 Выводы из анализа с помощью кривых Пирсона	39
3.2 Исследование подвыборок	40
3.2.1 Описание подходов	40
3.2.2 Исследование последовательных подвыборок	40
3.3.3 Исследование случайных подвыборок	46
3.3.4 Выводы из анализа подвыборок	50
Глава 4. Взаимосвязь волнения морской поверхности и амплитуды радиосигнала	51

4.1 Корреляция волнения и сглаженных значений амплитуды	51
4.2 Корреляция волнения и остаточных значений амплитуды	52
4.3 Алгоритм исследования	54
4.4 Выводы из анализа взаимосвязи волнения морской поверхности и амплитуды радиосигнала	55
Заключение	56
Список литературы	58
Приложение	61

Введение

На сегодняшний день на планете господствует эра информации. Нередко говорят о том, что ключевое умение человека в наш век - умение работать с ней, а сегодняшний мир опутан паутиной электрических сетей, обеспечивающих почти мгновенную связь на далекие расстояния. И, пожалуй, сейчас поиск новых средств обработки, передачи и сохранения информации имеет ключевое значение для последующей эволюции человечества.

Достижения в области беспроводных информационно-коммуникационных систем обеспечили поистине безграничные возможности. Процесс революционных изменений в области беспроводной связи затрагивает широкие и перспективные сферы не только коммерции, но и государственной обороны и научных исследований.

Допустим, что с целью улучшения качества беспроводного сигнала, осуществляется некоторая последовательность операций по преобразованию входящих/исходящих сигналов. Следуют закономерные вопросы: когда это желательно, какие будут ограничения и какие преимущества это даст? Ответы на них можно найти, обратившись к теории связи, значение которой значительно возросло с развитием вычислительной техники, позволившей проводить более сложные преобразования сигналов.

В теории связи выделяются два различных научных направления: статистическая теория помехоустойчивого приема сигнала и теория помехоустойчивого кодирования сигнала. Идеи, относящиеся к первому направлению, получили широкое развитие в исследованиях Котельникова [1] и нашли широкое применение, ввиду доступности математического аппарата и многообразия поддающихся анализу ситуаций. По-другому обстоит дело со вторым направлением. Развивающие его теоретические исследования К. Шеннона [2] актуальны и привлекательны с позиции разработанного инструментария, но, к сожалению, их практическое приложение было затруднено отсутствием подходящего математического аппарата. К счастью, вскоре последовала разработка реальных схем кодирования, аппроксимирующих идеальные схемы Шеннона, теория которых сочетала в себе статистические и алгебраические методы.

Статистическая теория неспроста занимает значимое место в исследованиях теории связи, поскольку в основе ее описания лежит фактор случайности. Необходимости в осуществление связи не возникло бы, если бы мы заранее знали, что нам хочет сообщить передающий, будь то записывающий прибор или человек. Поэтому в теории связи предполагается, что передатчик ассоциирован с некоторой случайной величиной, которую невозможно точно предсказать на приемнике. Кроме того, как уже говорилось выше, передаваемый сигнал искажается во время распространения, т. е. снижается точность его передачи. Наглядным примером является распространение радиоволн через ионосферу. Радиоволны определенной частоты преломляются, когда проходят через ионизированный газ. Степень преломления зависит от спектра различных факторов. Учесть все результирующие изменения и произвести точный расчет является невозможной задачей. Таким образом, изменение напряжения на выходах антенны приемника непредсказуемо в деталях, и мы говорим, что оно изменяется случайно. Однако принимаемая мощность - статистически устойчива, т. е. изменяется незначительно при усреднении по различным интервалам времени (секунды, минуты, месяцы и т.д.), что побуждает нас строить адекватные статистические модели для изучения таких явлений.

Основные аспекты связи в наши дни имеют прочную научную основу и обширную методологию. Однако горизонты для исследований как в области теории, так и в области практики еще продолжают формироваться, расширяя границы научных исследований.

Актуальность

В данной работе рассматриваются исследования, необходимые для моделирования канала связи приема сигнала со спутника на исследовательскую платформу и передачи его в обратном направлении. Схематичное изображение модели канала связи [3] может быть представлено на Рис. 1. На нем выходной сигнал источника случайно порождается совокупностью возможных дис-

кретных сообщений M , после чего каждому сообщению сопоставляется сигнал соответствующей формы, который впоследствии посылается передатчиком. Переданный сигнал затем проходит через канал, и его искаженный вариант поступает на вход приемника.

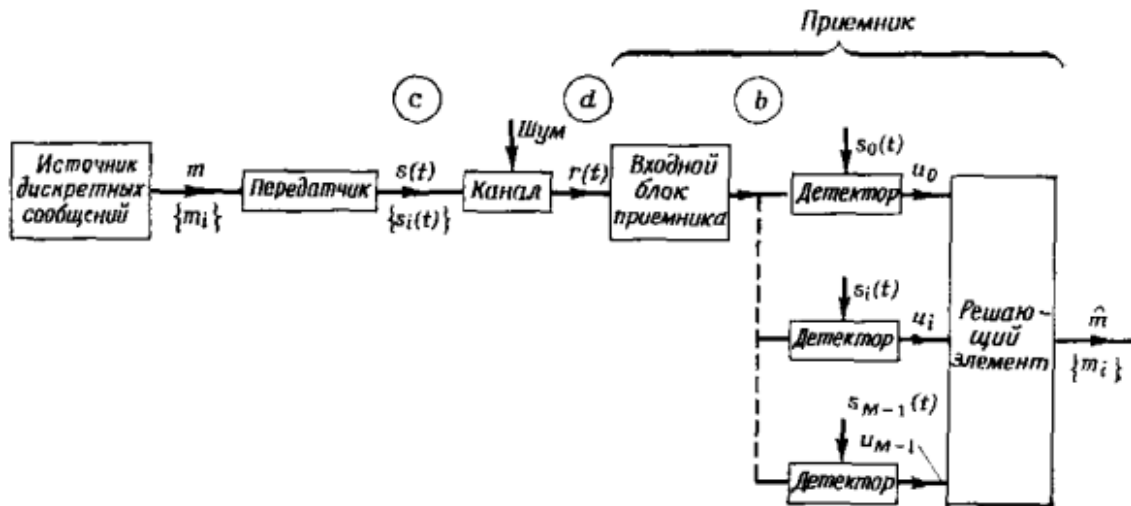


Рис. 1: Модель канала связи [3].

Задача приемника состоит в том, чтобы получить оценку сообщения, порожденного источником, при помощи ряда детекторов и решающего элемента. Детекторы сравнивают напряжение полученного сигнала с напряжением элементов из совокупности исходных сигналов M и передают результаты сравнения решающему элементу.

Необходимо выявить закон распределения выходных сигналов источника, а также рассмотреть влияние различных факторов, искажающих сигнал. В частности, этим фактором может являться волнение морской поверхности, так как приемник находится на плавучей платформе.

Фундаментальной целью исследования являлось получение аналитических функций распределения для имеющегося набора данных, а также выявление зависимости между характеристиками радиосигнала и волнением морской поверхности. На основе полученных функций распределения в дальнейшем будет смоделирован канал связи.

Постановка задачи

В рамках проведенного исследования были поставлены следующие задачи:

- Произвести предварительную обработку данных;
- На основании предварительного изучения данных составить алгоритм дальнейшего исследования предоставленных выборок;
- Получить оценки параметров распределений и эмпирические функции распределения (эмпирические характеристики распределений);
- Получить аналитические функции представления волнения морской поверхности и замирания радиосигнала;
- Установить или опровергнуть влияние волнения морской поверхности на амплитуду радиосигнала;
- В том случае, если таковое влияние наблюдается, выявить закономерности влияния.

Глава 1. Анализ эмпирических данных по волнению морской поверхности

Данные для исследования были предоставлены АО «Концерн «ЦНИИ «Электроприбор» [4]. Данный институт занимается выполнением полного цикла работ от фундаментально-поисковых исследований, до производства и поддержки продукции в эксплуатации. Одним из ведущих направлений организации является разработка средств связи и автоматизированных комплексов радиосвязи. Натурный эксперимент проводился с 31 октября по 7 ноября 2016 года. Данные содержат снятые при помощи 3-х волнографов показания волнения морской поверхности, а также характеристики принятого и переданного радиосигнала при различных углах места.

Предоставленные данные, описывающие волнение морской поверхности, содержат в себе 3 выборки показаний различных волнографов, измеренных в метрах над уровнем моря. Волнографы расположены в различных частях исследовательской платформы. Период снятия показаний волнения волнографами: со 2-го ноября по 7-е ноября 2016 года, частота снятия показаний — 4 показания в секунду. Каждая из выборок была представлена в виде подвыборок, которые соответствовали дню снятия показаний и подлежали объединению. Приблизительный объем каждой выборки в объединенном виде — 2 миллиона элементов.

В рекомендации по работе с данными было указано, что одна из выборок нежелательна к проведению исследования, т.к. волнограф, снимавший показания, искажал данные.

1.1 Первичный анализ данных

Исследуемые выборки имеют названия `waveHeight` и `waveHeight2`, которых мы будем придерживаться в дальнейшем. Анализ выборок проводился при помощи прикладного программного пакета R [5, 6]. Описательная статистика данных выборок продемонстрирована в таблице 1.

Таблица 1: Описательная статистика выборок волнения морской поверхности.

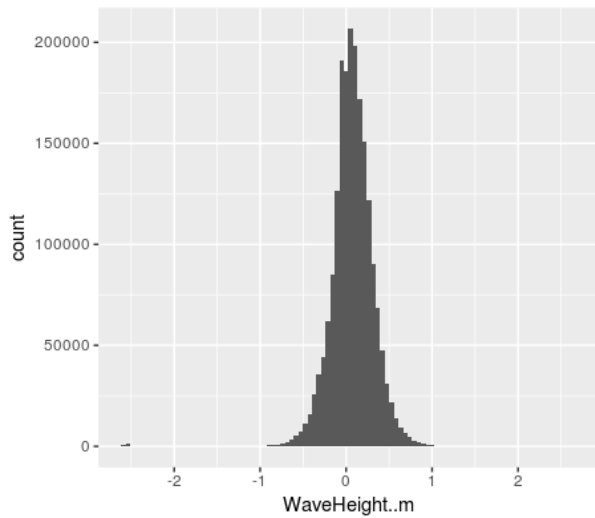
	Выборка "waveHeight"	Выборка "waveHeight2"
Количество элементов	1957988	1957988
Арифметическое среднее	0.069	-0.372
Стандартное отклонение	0.25	0.24
Медиана	0.07	-0.365
Среднее абсолютное отклонение	0.2	0.2
Минимум	-2.57	-4.09
Максимум	2.66	3.83
Размах	5.23	7.91
Асимметрия	-1.55	-0.87
Эксцесс	16.87	12.06

Как можно видеть из таблицы, объем выборок приближен к 2-м миллионам экземпляров в каждой. Несмотря на то, что данные фиксировались волнографами, находившимися на одной платформе, в один промежуток времени, в характеристиках выборок наблюдаются некоторые различия. В частности, минимальные и максимальные значения волнения, зафиксированные вторым волнографом, несколько больше значений первого по модулю, что делает размах второй выборки в целом шире. Наблюдаются также различия в средних, асимметричности и эксцессе выборок. Стандартное отклонение, однако, приблизительно одинаковое.

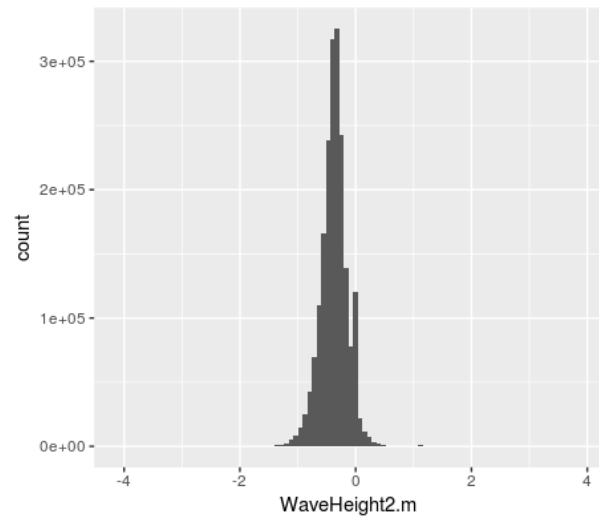
В выборках присутствует большое количество повторяющихся значений. Вызвано это тем, что точность записи волнения включает в себя четыре знака после запятой. К сожалению, при таких объемах выборки и сравнительно небольшом размахе совпадающие значения неизбежны.

На рис.2 представлены гистограммы и графики эмпирических плотностей для рассматриваемых выборок.

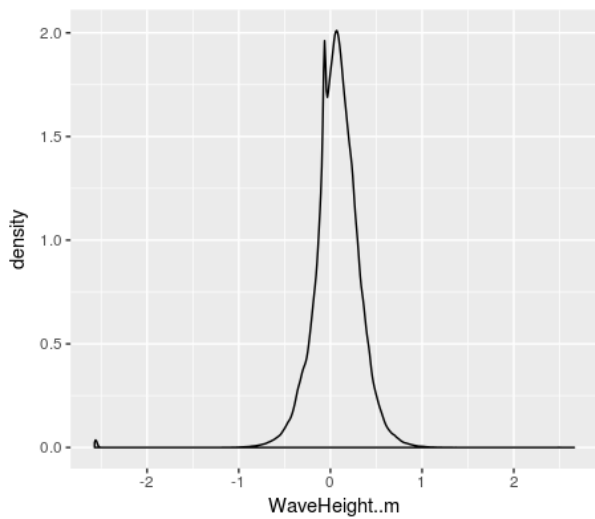
Внешний вид, а также небольшие разности значений среднего арифметического и медианы обеих выборок наводят на подозрения о том, что эти выборки извлечены из нормальной генеральной совокупности. Однако, если



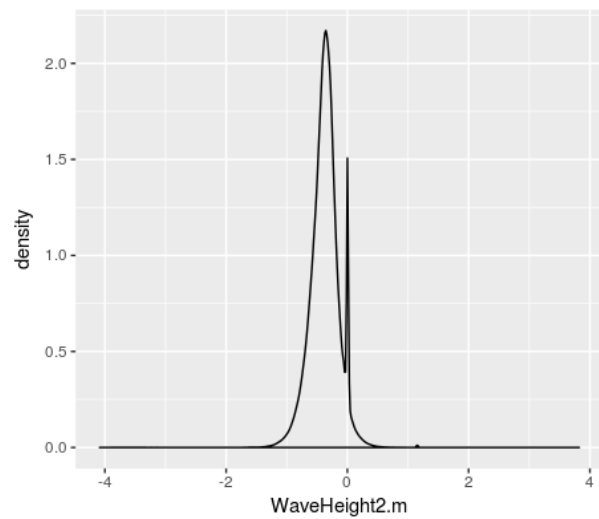
a)



b)



c)



d)

Рис. 2: Графическое представление выборок: а) гистограмма `waveHeight`, б) гистограмма `waveHeight2`, в) эмпирическая плотность `waveHeight`, д) эмпирическая плотность `waveHeight2`.

обратить внимание на существенное отличие коэффициентов асимметрии и эксцесса от нуля, данное предположение может быть отвергнуто [7]. Обе выборки являются отрицательно асимметричными.

Видно, что первая выборка расположена симметрично относительно нуля, в то время, как вторая смещена в сторону отрицательной полуоси. При этом, график плотности второй выборки выглядит «уже», чем график первой выборки. С физической точки зрения данные различия можно объяснить

тем, что волнографы были расположены в разных частях платформы.

Тем не менее, можно отвергнуть предположение о сходстве обеих выборок с нормальной совокупностью. Рассмотрим законы распределения Райса и Накагами [8, 9]. Данные распределения наиболее часто используются для исследования характеристик радиосигнала [10]. Их выбор для исследования волнения морской поверхности обоснован большой долей зрительного сходства с ними эмпирических плотностей исследуемых выборок. Закон распределения Райса является обобщением распределения Релея, которое, в свою очередь, опирается на нормальное распределение. Распределение Накагами является обобщением гамма-распределения. Данное предположение будет рассмотрено чуть ниже.

1.2 Проверка гипотез статистической однородности на выборках волнения

С целью обобщить показания волнографов, был произведен тест на однородность данных выборок волнения. Принятие гипотезы однородности, в данном случае, позволит использовать только одну выборку из двух при дальнейших исследованиях взаимосвязи с амплитудой радиосигнала. В качестве используемых критериев однородности был выбран критерий однородности Колмогорова-Смирнова [11, 12, 13].

Согласно критерию Колмогорова-Смирнова, нулевая гипотеза о том, что данные выборки извлечены из одной генеральной совокупности, должна быть отклонена.

Тест однородности Колмогорова-Смирнова:

$$D = 0.7121, p\text{-value} < 2.2e-16$$

Опираясь на предположение о том, что данные выборки извлечены из распределений Райса или Накагами, было решено сместить их, руководствуясь следующими соображениями:

- минимальные значения выборок должны быть не меньше нуля;

- средние выборок должны примерно совпадать.

Вообще говоря, такое горизонтальное смещение никаким образом не влияет на свойства выборок, за исключением математического ожидания. Убедиться в этом можно, обратив внимания на таблицу 2. Тем не менее, несмотря на произведенное смещение, критерий однородности Колмогорова-Смирнова дает тот же результат - гипотезу об однородности выборок следует отвергнуть:

Тест однородности Колмогорова-Смирнова:

$$D = 0.021871, p\text{-value} < 2.2e-16$$

Таблица 2: Описательная статистика сдвинутых выборок волнения морской поверхности.

	Сдвинутая выборка "waveHeight"	Сдвинутая выборка "waveHeight2"
Количество элементов	1957988	1957988
Арифметическое среднее	3.715	3.715
Стандартное отклонение	0.25	0.24
Медиана	3.716	3.722
Среднее абсолютное отклонение	0.2	0.2
Минимум	1.073	0
Максимум	6.306	7.913
Размах	5.23	7.91
Асимметрия	-1.55	-0.87
Эксцесс	16.87	12.06

Следует заметить, что в данном случае величина статистики Колмогорова-Смирнова уменьшилась примерно в 35 раз по сравнению с предыдущим, что, однако, не повлияло на результат, несмотря на малое значение статистики.

1.3 Проверка гипотез согласия на выборках волнения

Согласно выше приведенным предположениям о теоретических распределениях выборочных совокупностей, были проверены тесты гипотез согласия о

том, что данные выборки извлечены из распределений Райса или Накагами. Гипотезы проверялись при помощи теста Колмогорова.

Оценка параметров распределения Накагами производилась по методу Гринвуда-Дюранда, который показал себя как наиболее результативный среди прочих существующих, согласно [8], а также по методу моментов, как наиболее распространенному и простому в реализации. Оценки параметров распределений Райса были получены при помощи метода моментов по тем же причинам, несмотря на результаты полученные в [9], ввиду того, что оценочный метод максимального правдоподобия в контексте данного теоретического распределения является не в меру громоздким и сложен в реализации.

По результатам теста Колмогорова гипотезу согласия для выборки waveHeight и распределения Накагами по обоим методикам оценки необходимо отклонить.

Тест Колмогорова на согласие с распределением Накагами, аппроксимация методом Гринвуда-Дюранда:

$$D = 0.068924, p\text{-value} < 2.2e-16$$

Тест Колмогорова на согласие с распределением Накагами, аппроксимация методом моментов:

$$D = 0.048588, p\text{-value} < 2.2e-16$$

Результаты теста и график сравнения эмпирической и теоретической плотностей представлены на рис. 3. Несмотря на явную зрительную схожесть, строго принять гипотезу не удалось. При этом оценка параметров распределения по методу моментов, как зрительно, так и по статистике критерия дает большее сходство с теоретическим распределением, нежели оценка методом Гринвуда-Дюранда. Согласно выводам, полученным в [8], должно быть наоборот, но только в том случае, если выборка действительно извлечена из теоретического распределения Накагами, что ставит наше предположение о законе распределения под сомнение.

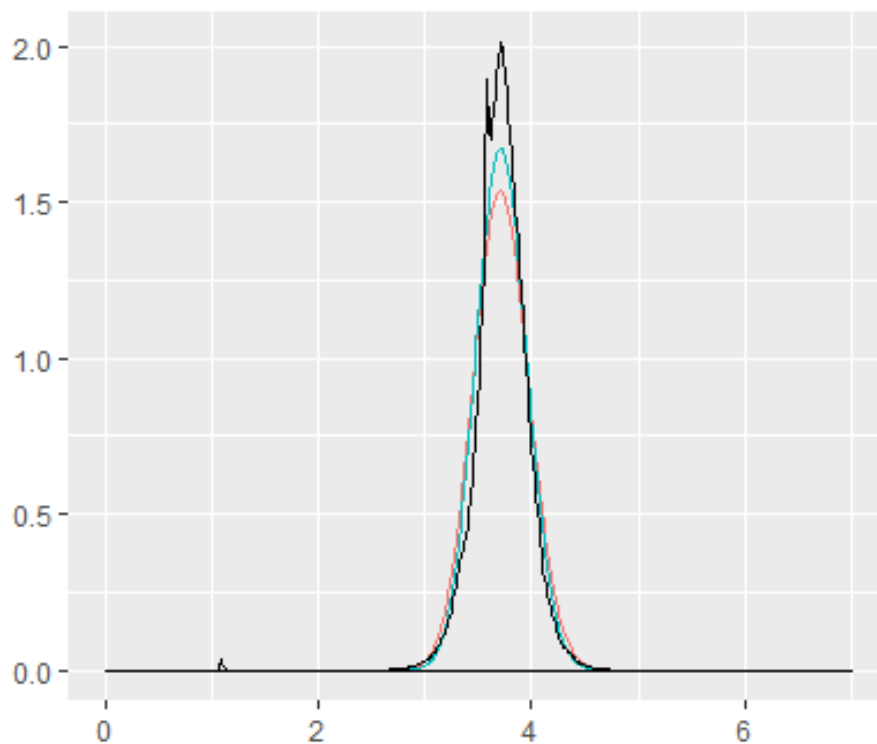


Рис. 3: График сравнения теоретических и эмпирической плотностей. Красным - оценка методом Гринвуда-Дюранда, синим - методом моментов. Черным - эмпирическая плотность.

По результатам теста Колмогорова гипотезу согласия выборки `waveHeight` и распределения Райса необходимо отклонить.

Тест Колмогорова на согласие с распределением Райса:

$$D = 0.040095, p\text{-value} < 2.2e-16$$

Результаты теста и график сравнения эмпирической и теоретической плотностей представлены на рис. 4. Как по статистике критерия, так и зрительно, теоретическая плотность распределения Райса лучше описывает эмпирические данные.

Тем не менее, результаты обоих тестов, следует отклонить, несмотря на маленькие, в сравнении с аналогичными статистиками для выборок меньших величин, значения результирующих статистик.

Результаты проверки гипотез согласия по выборке `waveHeight2` не представлены, т.к. полностью повторяют вышеприведенные.

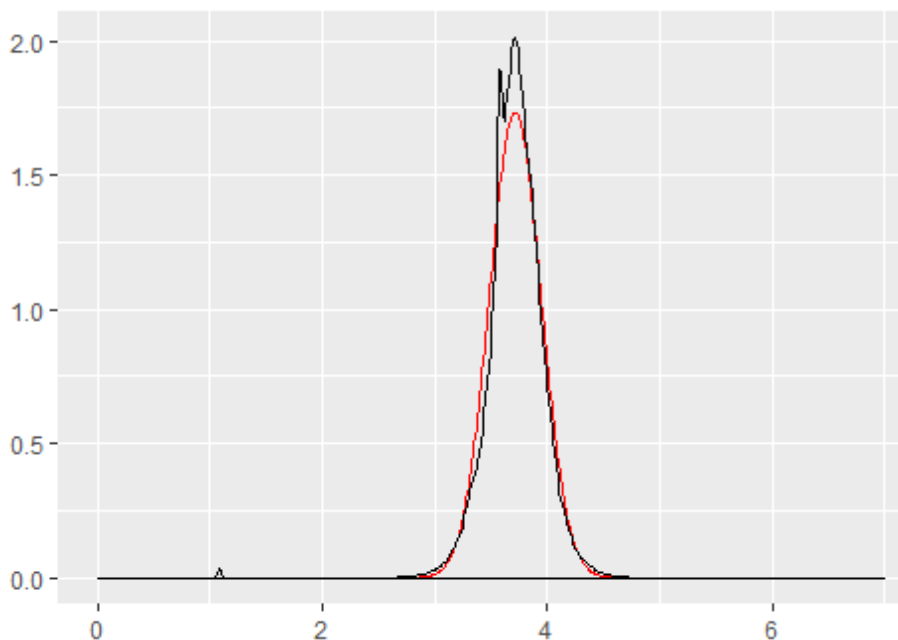


Рис. 4: График сравнения теоретических и эмпирической плотностей. Красным - оценка методом моментов, черным - эмпирическая плотность.

1.4 Удаление выбросов

При рассмотрении графиков эмпирических плотностей было сделано предположение о том, что в выборках возможно присутствие так называемых выбросов (пики слева медианы `waveHeight` и справа медианы `waveHeight2`, см. рис. 2). Было принято решение удалить данные пики из рассмотрения при условии, что их доля в выборке незначительна, и провести тесты на гипотезы согласия еще раз.

Рассмотрим пошаговое описание удаления выбросов для выборки `waveHeight`. Аналогичный алгоритм может быть применен ко второй выборке `waveHeight2`.

1. Зрительно установить интервал значений, соответствующий выбросу;
2. Установить общее количество элементов выборки на этом интервале;
3. Установить частоты повторяющихся элементов выборки на этом интервале;
4. Для аппроксимации количества удаляемых значений рассмотреть сим-

метричный, относительно медианы, интервал значений и установить общее количество элементов в нем;

5. Рассчитать среднее количество элементов, которое необходимо вычесть из частот повторяющихся элементов рассматриваемого интервала;
6. Произвести вычитание, если это возможно.

Зрительно было установлено, что выброс приходится на значения от 3,55 до 3,60. Из выборки были удалены все элементы из данного диапазона с целью выяснить их количество. Уникальных элементов среди них (напомним, что в выборке присутствуют повторяющиеся значения) было 180. Частота их повторения варьировалась от 1 до 1466. Количество удаленных элементов равняется 171923. Сравнительная гистограмма представлена на рис. 5.

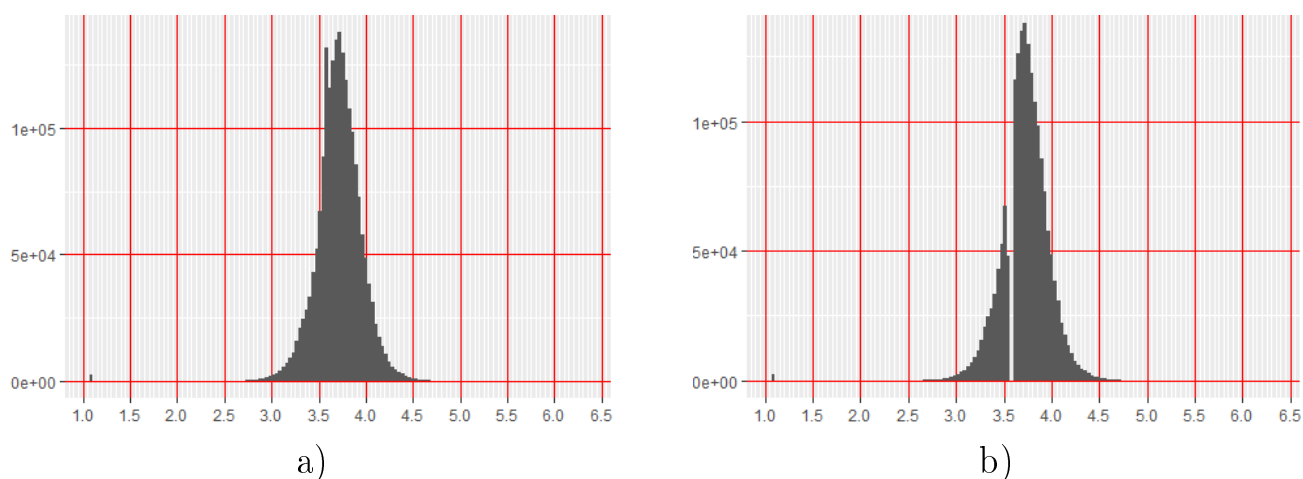


Рис. 5: Сравнение гистограмм: а) изначальная гистограмма waveHeight, б) модифицированная гистограмма waveHeight.

Далее, для аппроксимации количества элементов, необходимых к удалению, аналогичным образом была рассмотрена совокупность элементов выборки, принадлежащих интервалу, симметричному относительно медианы выборки, т. е. те элементы выборки, что принадлежат интервалу от 3,85 до 3,90. Количество элементов выборки в данном интервале равняется 134653.

Таким образом, разница в количестве симметричных относительно медианты интервалов составляет 37270, т.е. для удаления выброса из выборки

необходимо удалить 37270 элементов выборки из интервала от 3,55 до 3,60. Данное количество составляет менее 2% от общего объема выборки или, иначе говоря, площадь данного пика составляет примерно 0,019, следовательно, ей можно пренебречь.

Для удаления данных элементов было найдено среднее число элементов, которое необходимо вычесть из частоты уникальных элементов интервала, т.е. число элементов, необходимых к удалению, было разделено на количество уникальных элементов интервала. Таким образом, из частоты каждого уникального элемента интервала было вычтено 207 элементов при условии, что это можно сделать. Результирующий график эмпирической плотности и гистограмма выборки представлены на рис. 6.

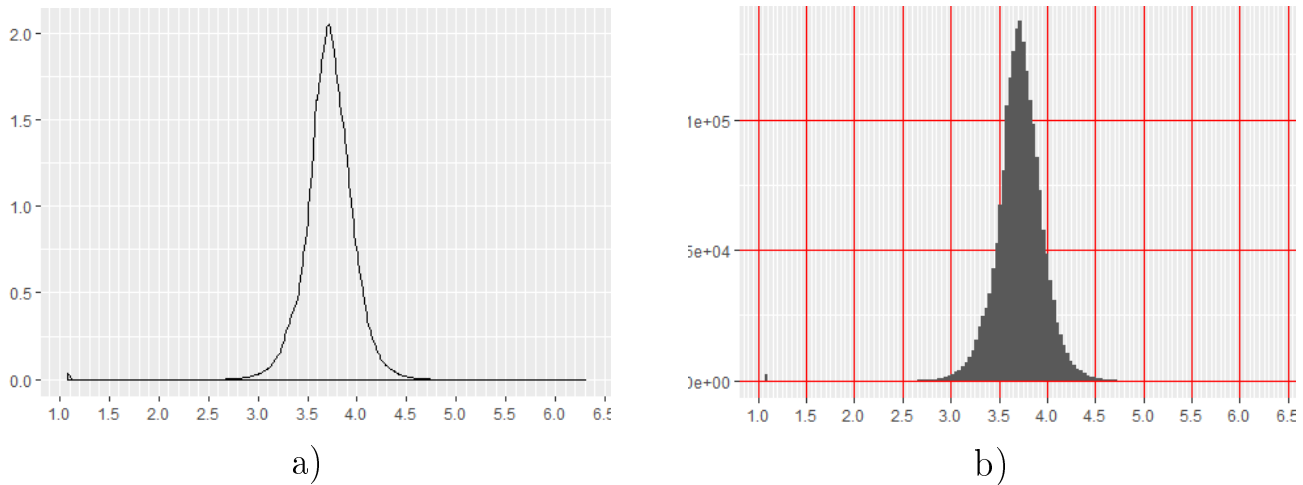


Рис. 6: Выборка waveHeight без выбросов: а) эмпирическая плотность, б) гистограмма.

Как можно видеть, выброс успешно устранен из выборки, что, однако, не изменило ситуацию с критериями согласия. Каждый из них по-прежнему следует отклонять. При этом, значения оценок параметров распределений изменяются незначительно, относительно их значений без удаления выбросов. То же следует сказать и об описательной статистике (см. сравнение в таблице 3).

Можно сделать вывод о том, что удаление выбросов не несет в себе значимой пользы в данном случае. Выводы касательно выборки waveHeight, приведенные в данном разделе, могут быть применены и к выборке waveHeight2.

Таблица 3: Сравнение значений описательных характеристик для выборки waveHeight до и после удаления выбросов.

	Сдвинутая выборка "waveHeight" с выбросами	Сдвинутая выборка "waveHeight" без выбросов
Количество элементов	1957988	1923626
Арифметическое среднее	3.715	3.718
Стандартное отклонение	0.25	0.25
Медиана	3.716	3.721
Среднее абсолютное отклонение	0.2	0.2
Минимум	1.073	1.074
Максимум	6.306	6.306
Размах	5.23	5.23
Асимметрия	-1.55	-1.57
Экцесс	16.87	16.80

1.5 Выводы из анализа эмпирических данных по волнению морской поверхности

Из приведенных выше результатов можно сделать вывод о том, что, хотя данные и взяты с одной платформы, факт того, что каждую из выборок писал разный волнограф, нельзя игнорировать. Т. е., при работе с данными по амплитуде радиосигнала нельзя ограничиваться работой с одной из выборок волнения, необходимо проводить исследование со всеми существующими, в нашем случае — с обеими.

Также можно сказать, что выборки волнения морской поверхности не извлечены из распределений Накагами или Райса, хотя визуально они очень похожи. Скорее всего, отклонение гипотез при таких маленьких статистиках Колмогорова вызваны большими объемами выборок.

Удаление выбросов из выборок также не дало существенных результатов, как в изменении параметров оценки и описательных характеристик выборок, так и в принятии гипотез согласия.

Глава 2. Анализ эмпирических данных по замиранию амплитуды радиосигнала

Предоставленные данные, описывающие замирание амплитуды радиосигнала, содержат в себе 43 выборки. Показания замерялись в децибелах, частота снятия — 100 показаний в секунду. Каждой выборке соответствуют определенная специфика сигнала (прием или передача) и угол места. Показания снимались в различные промежутки времени в период с 31-го октября по 6-е ноября. Объемы выборок также варьировались от нескольких тысяч до нескольких миллионов элементов. Данные по флуктуации амплитуды представлены в логарифмическом масштабе.

Формат записи названий выборок будет иметь следующий вид: *"специфика сигнала : угол места : идентификатор выборки"*. Например, запись вида *"VBAU_T : 10 : 30dB_10grad_ch2"* обозначает, что снятие показаний замирания амплитуды производилось во время передачи сигнала при угле места в 10 градусов. Идентификаторы выборок необходимы по причине того, что отдельно взятым специфике сигнала и углу места могут соответствовать сразу несколько наборов данных, разнесенных по времени.

В рекомендациях по работе с данными было указано, что некоторые из выборок подлежат объединению в том случае, если они принадлежат одной специфике сигнала и имеют один угол места (в дальнейшем такое сочетание будем именовать классом выборки). Кроме того, временной промежуток между ними должен быть достаточно мал для осуществления объединения. В этом случае составлен новый идентификатор выборки на основе старых.

2.1 Первичная обработка и анализ данных

Перед проведением первичного исследования данных замирания амплитуды, все временные метки представленных выборок были проанализированы на предмет возможного объединения выборок одного класса. По окончании данной операции успешно удалось объединить 9 выборок в 3. Таким образом, к

исследованию представлено 37 выборок различных классов.

Ввиду большой громоздкости представления результатов исследования всех 37-ми выборок, мы представим лишь некоторые из них, которые являются наиболее показательными.

Составим описательные статистики для нескольких выборок. Их представление можно найти в таблице 4.

Таблица 4: Описательная статистика выборок различных классов:

класс А: $VBAU_R : 10 : 20dB_ch2$,

класс В: $VBAU_R : 17 : 35dB_ch2$,

класс С: $VBAU_T : 10 : 40dB_10grad_ch2$,

класс D: $VBAU_R : 13 : 50dB_ch2_1$.

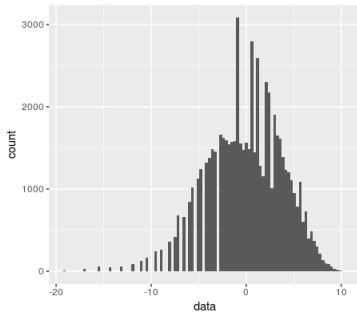
	Класс А	Класс В	Класс С	Класс D
Количество элементов	63005	56188	295308	10338
Арифметическое среднее	0.117	5.197	6.622	-8.127
Стандартное отклонение	3.828	7.029	4.473	5.701
Медиана	0.240	3.524	7.0461	-9.54
Среднее абсолютное отклонение	4.057	7.569	4.401	3.705
Минимум	-19.082	-21.581	-21.581	-21.581
Максимум	10.46	18.925	18.925	16.261
Размах	29.542	40.506	40.506	37.842
Асимметрия	-0.386	0.167	-0.775	1.635
Эксцесс	0.236	-0.966	1.487	3.734

Как можно видеть из таблицы, объемы выборок варьируются от нескольких тысяч до нескольких сотен тысяч. Обобщая данное наблюдение, можно сказать, что среди всех 37-ми выборок, минимальный объем - 4374 элемента, максимальный - 3651143. Можно также сказать, что средние рассматриваемых выборок, как и их стандартные отклонения, разительно отличаются между собой. Данный факт ставит под сомнение предположение о том, что все выборки были извлечены из одной генеральной совокупности. Размах выборок приблизительно одинаков, за исключением первой рассматриваемой выборки. У нее он значительно меньше остальных.

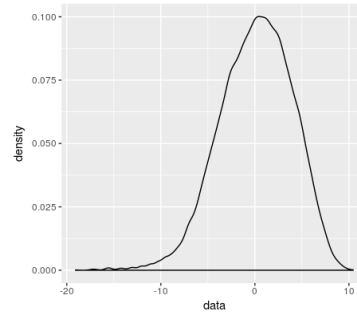
Незначительное отличие от нуля коэффициентов асимметричности и эксцесса у первой выборки, позволяет выдвинуть предположение о том, что данная выборка извлечена из нормальной совокупности. К сожалению, подобные предположения не представляются правомочными в контексте вида оставшихся 3-ех выборок.

Гистограммы и графики эмпирических плотностей для рассматриваемых наборов данных можно найти на рис. 7. Как можно видеть по графикам, внешний вид выборки *VBAU_R : 10 : 20dB_ch2* подтверждает наши подозрения касательно того, что она может быть извлечена из нормальной генеральной совокупности. Тем не менее частоты значений, находящихся чуть правее медианы, значительно превосходят симметричные им слева от медианы. Внешний вид выборки *VBAU_R : 17 : 35dB_ch2* является иллюстрацией того, что ряд исследуемых выборок может является смесью распределений. Выборка *VBAU_T : 10 : 40dB_10grad_ch2* призвана продемонстрировать волнообразную структуру некоторых выборок. Данная волнообразность вызвана тем, что некоторые исследуемые выборки имеют сильно дискретную структуру в связи с разреженными повторяющимися значениями (как и в ситуации с волнением морской поверхности, точность измерений представляется недостаточно высокой для таких объемов данных — всего 4 знака после запятой). Выборка *VBAU_R : 13 : 50dB_ch2_1* демонстрирует факт того, что среди всех исследуемых наборов данных существуют и те, которые не поддаются какой-либо классификации и обобщению.

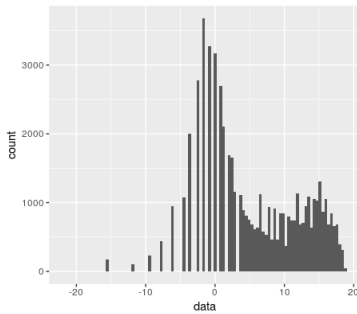
Таким образом, можно говорить о том, что множество рассматриваемых наборов данных по замиранию амплитуды не может быть обобщено в одну генеральную совокупность по причине их большой разобщенности. Тем не менее, большую часть данного набора составляют выборки, которые в результате визуального анализа можно отнести к какой-либо генеральной совокупности, что впоследствии и будет проверено.



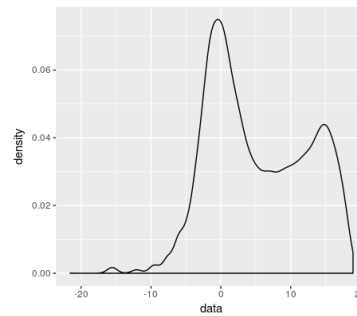
VBAU_R : 10 : 20dB_ch2,
гистограмма



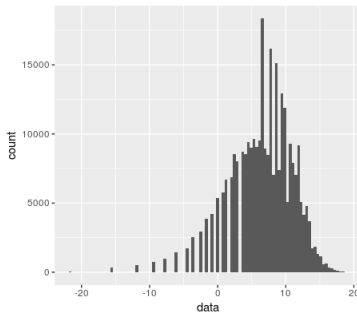
VBAU_R : 10 : 20dB_ch2,
ПЛОТНОСТЬ



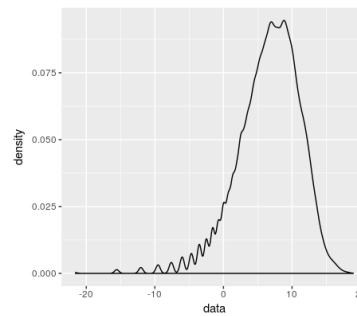
VBAU_R : 17 : 35dB_ch2,
гистограмма



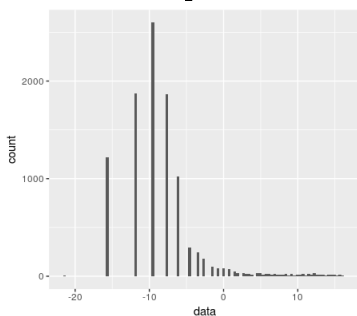
VBAU_R : 17 : 35dB_ch2,
ПЛОТНОСТЬ



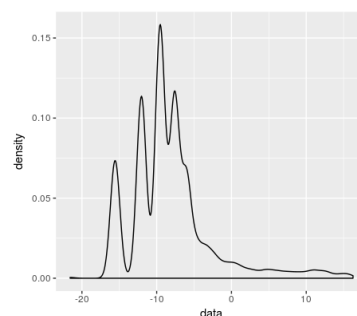
VBAU_T : 10 : 40dB_10grad_ch2,
гистограмма



VBAU_T : 10 : 40dB_10grad_ch2,
ПЛОТНОСТЬ



VBAU_R : 13 : 50dB_ch2_1,
гистограмма



VBAU_R : 13 : 50dB_ch2_1,
ПЛОТНОСТЬ

Рис. 7: Графики гистограмм и плотностей.

2.2 Гипотезы согласия выборок амплитуды

Прежде чем приступить к проверке критериев согласия, необходимо определиться с теоретическими распределениями, к которым мы будем относить наборы данных замирания амплитуды. Согласно [14], наиболее часто характеристики радиосигналов моделируются при помощи теоретических распределений Райса и Накагами, о которых уже шла речь выше. Также к этому списку можно добавить и нормальное распределение, по следующим соображениям:

- распределения Райса и Накагами базируются на основе распределения Гаусса;
- некоторые выборки, как зрительно, так и согласно их центральным моментам, по результатам первичного анализа могут быть отнесены к классу нормального распределения.

Условимся о том, что будем отдавать предпочтение на проверку критериев согласия теоретическим распределениям Райса и Накагами. Однако, будем проводить дополнительные проверки критериев согласия с другими теоретическими распределениями (к примеру на согласие с нормальным распределением) в том случае, если, как зрительно, так и по характеристикам, выборку гипотетически можно отнести к ним.

Согласно определению [8, 9], случайные величины, извлекаемые из теоретических распределений Райса или Накагами, могут иметь значения строго большие нуля, поэтому для работы с ними необходимо перейти из логарифмической шкалы к стандартной. В проводимом исследовании все статистические характеристики выборок, за исключением среднего, минимального и максимального значений, инвариантны относительно произведенного сдвига.

Будем применять критерии оценки параметров распределений Райса и Накагами, которые применялись при исследовании волнения, т.е. методы моментов и Гринвуда-Дюранда для распределения Накагами, метод моментов для распределения Райса. Параметры других теоретических распределений,

выявленных в ходе исследования, также будем исследовать методом моментов.

Как и в случае с описательной статисткой и предварительной обработкой данных, мы не будем представлять результаты исследования всех 37-ми выборок, ввиду их большой громоздкости, а ограничимся теми выборками, которые были представлены в предыдущем подразделе.

Рассмотри выборку $VBAU_R : 10 : 20dB_ch2$. Данная выборка, как по зрительному наблюдению, так и по исследованию ее описательных характеристик, может быть извлечена из нормальной генеральной совокупности, помимо предполагаемых распределений Накагами и Райса. Гипотезы согласия проверялись при помощи критериев Колмогорова [11, 12] и критерия Пирсона [11, 12]. Начальное количество интервалов для критерия Пирсона бралось согласно формуле Стерджесса [11]. Результаты гипотез и итоговые статистики представлены ниже.

Тест Колмогорова на согласие с нормальным распределением:

$$D = 0.0339171, p\text{-value} < 2.22e-16$$

Тест Колмогорова на согласие с распределением Накагами, аппроксимация методом Гринвуда-Дюранда:

$$D = 0.0478445, p\text{-value} < 2.22e-16$$

Тест Колмогорова на согласие с распределением Накагами, аппроксимация методом моментов:

$$D = 0.0457108, p\text{-value} < 2.22e-16$$

Тест Колмогорова на согласие с распределением Райса:

$$D = 0.0345198, p\text{-value} < 2.22e-16$$

Тест Пирсона на согласие с нормальным распределением:

$$X^2 = 2777.08, p\text{-value} < 2.22e-16$$

Тест Пирсона на согласие с распределением Накагами, аппроксимация методом Гринвуда-Дюранда:

$$X^2 = 117434, p\text{-value} < 2.22e-16$$

*Тест Пирсона на согласие с распределением Накагами,
аппроксимация методом моментов:*

$$X^2 = 1097510, p\text{-value} < 2.22e-16$$

Тест Пирсона на согласие с распределением Райса:

$$X^2 = 3358.76, p\text{-value} < 2.22e-16$$

Как видно из представленных выводов, все результирующие значения статистик X^2 очень велики, поэтому данные гипотезы необходимо отклонить. Все результирующие значения статистик Колмогорова, напротив, относительно малы. Однако несмотря на это, все гипотезы согласия необходимо отвергнуть. Графическое сравнение эмпирической и теоретической плотностей можно увидеть на рис. 8.

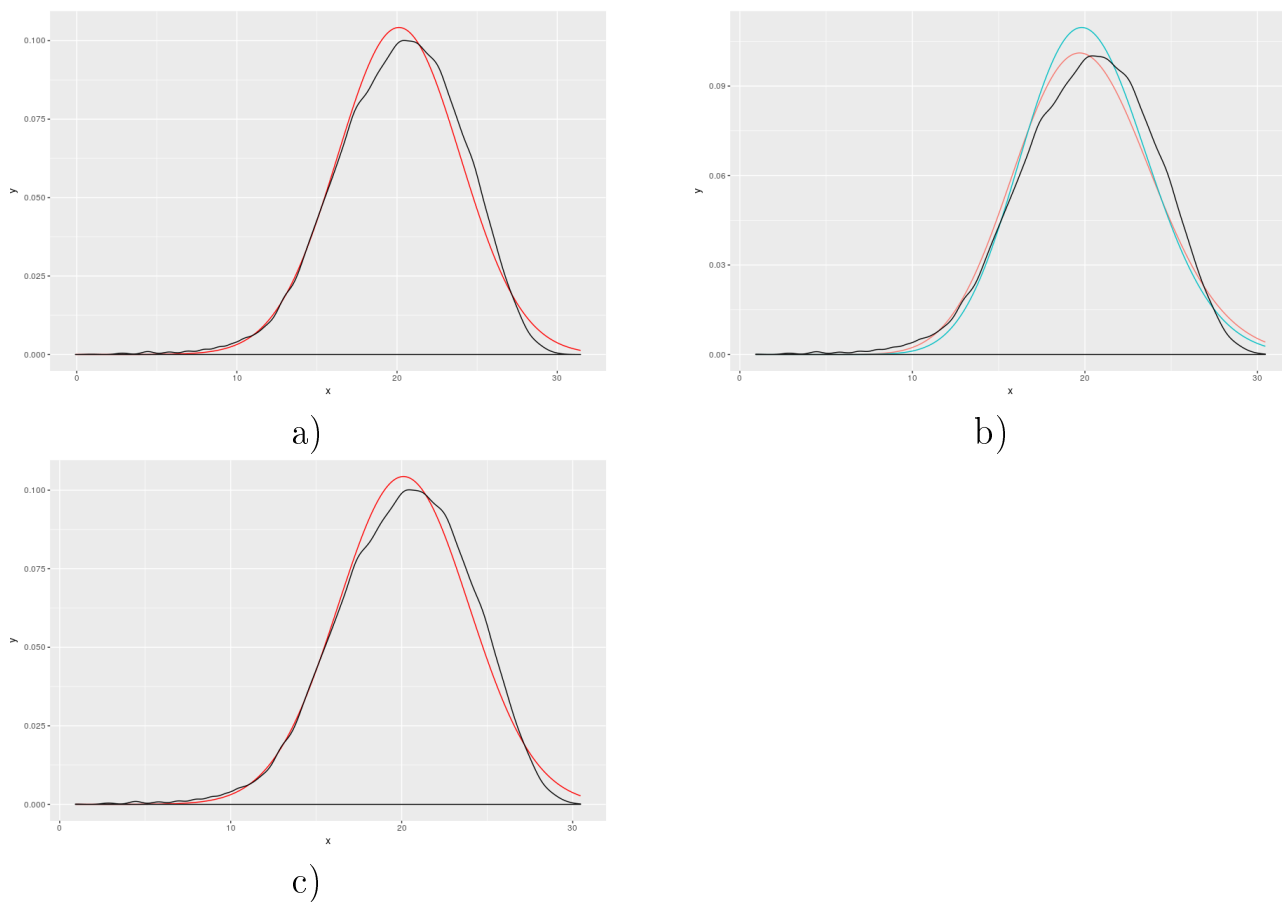


Рис. 8: Сравнительные графики для $VBAU_R : 10 : 20dB_ch2$:
а) нормальное распределение, б) распределение Накагами, красным - оценка по методу Гринвуда-Дюранда, синим - моментов, с) распределение Райса.

Аналогично ситуации с исследованием волнения морской поверхности, результат визуального сравнения отличается от результатов проверок. Зрительно исследуемая выборка хорошо укладывается во все проверяемые теоретические распределения.

Далее рассмотрим выборку $VBAU_T : 10 : 40dB_10grad_ch2$. Данная выборка по внешнему поведению схожа с предыдущей. Результаты применения критериев согласия следующие:

Тест Колмогорова на согласие с нормальным распределением:

$$D = 0.0565421, p\text{-value} < 2.22e-16$$

Тест Колмогорова на согласие с распределением Накагами, аппроксимация методом Гринвуда-Дюранда:

$$D = 0.0731366, p\text{-value} < 2.22e-16$$

Тест Колмогорова на согласие с распределением Накагами, аппроксимация методом моментов:

$$D = 0.0628264, p\text{-value} < 2.22e-16$$

Тест Колмогорова на согласие с распределением Райса:

$$D = 0.0570408, p\text{-value} < 2.22e-16$$

Тест Пирсона на согласие с нормальным распределением:

$$X^2 = 44386.6, p\text{-value} < 2.22e-16$$

Тест Пирсона на согласие с распределением Накагами, аппроксимация методом Гринвуда-Дюранда:

$$X^2 = 535009, p\text{-value} < 2.22e-16$$

Тест Пирсона на согласие с распределением Накагами, аппроксимация методом моментов:

$$X^2 = 10284900, p\text{-value} < 2.22e-16$$

Тест Пирсона на согласие с распределением Райса:

$$X^2 = 64768.4, p\text{-value} < 2.22e-16$$

В данном случае все проверяемые гипотезы необходимо отклонить по

аналогичным причинам. Относительно значений статистик данная выборка хуже описывается рассматриваемыми теоретическими распределениями, нежели чем предыдущая. На рис. 9 представлены графики сравнения эмпирической и теоретической плотностей рассматриваемой выборки.

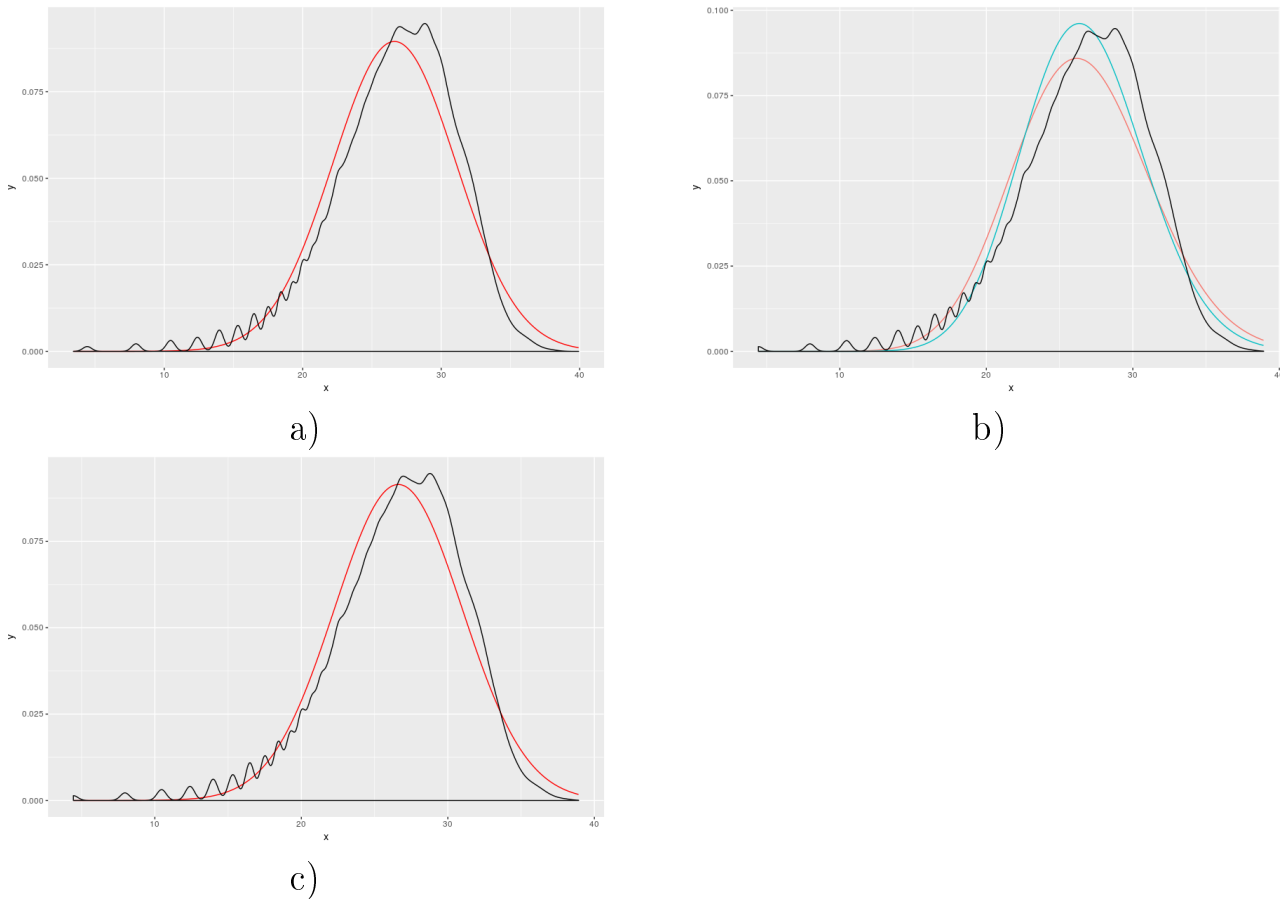


Рис. 9: Сравнительные графики для $VBAU_T : 10 : 40dB_10grad_ch2$: а) нормальное распределение, б) распределение Накагами, красным - оценка по методу Гринвуда-Дюранда, синим - моментов, с) распределение Райса.

Из рисунка видно, что все теоретические кривые плотностей находятся чуть левее относительно пика эмпирических данных. Этим можно объяснить увеличение численных значений статистик в сравнении с предыдущим случаем, где вершины совпадали.

Наконец рассмотрим выборку $VBAU_R : 13 : 50dB_ch2_1$. К данной выборке довольно проблематично подобрать подходящий закон распределения. Тем не менее, для полноты исследования, все подобные ей выборки были

исследованы на согласие с теоретическими распределениями Райса и Накагами. Результаты тестов на согласия приведены ниже:

Тест Колмогорова на согласие с распределением Накагами, аппроксимация методом Гринвуда-Дюранда:

$$D = 0.170593, p\text{-value} < 2.22e-16$$

Тест Колмогорова на согласие с распределением Накагами, аппроксимация методом моментов:

$$D = 0.242264, p\text{-value} < 2.22e-16$$

Тест Колмогорова на согласие с распределением Райса:

$$D = 0.177435, p\text{-value} < 2.22e-16$$

Тест Пирсона на согласие с распределением Накагами, аппроксимация методом Гринвуда-Дюранда:

$$\chi^2 = 5719.58, p\text{-value} < 2.22e-16$$

Тест Пирсона на согласие с распределением Накагами, аппроксимация методом моментов:

$$\chi^2 = 5675.01, p\text{-value} < 2.22e-16$$

Тест Пирсона на согласие с распределением Райса:

$$\chi^2 = 26362.3, p\text{-value} < 2.22e-16$$

Аналогично с предыдущими проверками, все гипотезы согласия были отклонены, что вполне закономерно. Иллюстрирующие сравнительные графики представлены на рис. 10.

Аналогичным образом были проверены гипотезы согласия для оставшихся 34-х совокупностей данных по замиранию амплитуды. Ввиду большой громоздкости, результаты проверок оставшихся гипотез согласия не будут приведены здесь в полной мере. Вместо этого в следующем подразделе будут представлены таблица, содержащая в себе классификацию выборок по результатам исследования и общие выводы по полученным результатам.

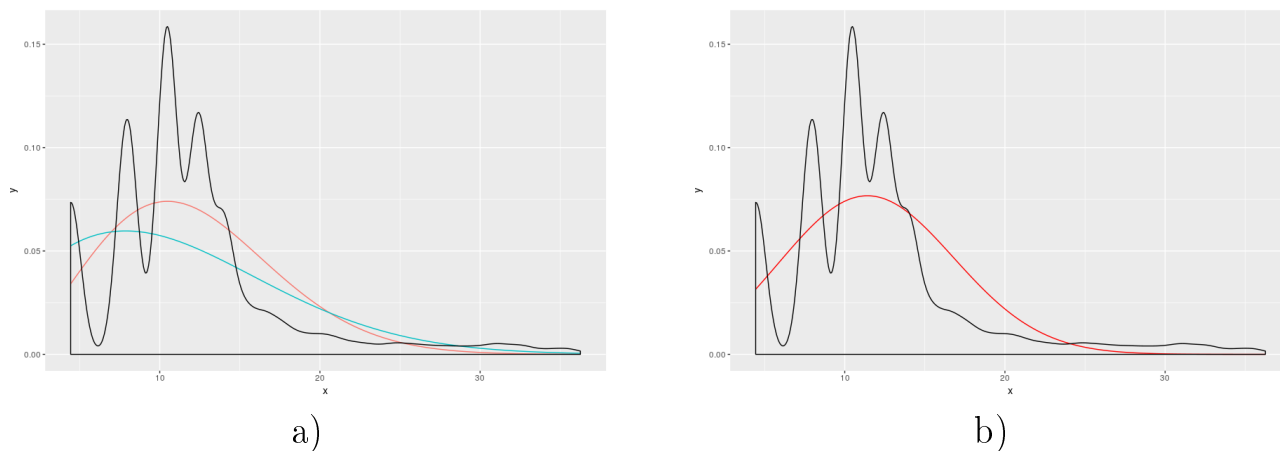


Рис. 10: Сравнительные графики для $VBAU_R : 13 : 50dB_ch2_1$:
 а) распределение Накагами, красным - оценка по методу Гринвуда-Дюранда, синим - моментов, б) распределение Райса.

2.3 Классификация результатов. Выводы из анализа эмпирических данных по замиранию амплитуды радиосигнала

Ввиду громоздкости представления, условная классификация выборок по результатам исследования расположена в Приложении 1. Классификация разбита на две таблицы, в одной из которых содержатся выборки, которые при зрительном анализе могут быть отнесены к какому-либо теоретическому распределению, во второй — выборки для которых этого сделать, вообще говоря, нельзя.

Как видно из таблиц, больше половины выборок достаточно хорошо описываются рассматриваемыми теоретическими распределениями. К сожалению, ни одна из гипотез согласия не была принята, несмотря на довольно сильное зрительное сходство теоретических и эмпирических плотностей.

Наибольшее количество «уникальных» по внешнему виду выборок наблюдается при приеме. Уникальные в данном случае означает то, что зрительно выборка выделяется из ряда остальных и не может быть отнесена к какому-либо классу теоретических распределений. При этом определить какую-то зависимость между градусом приема/передачи сигнала и возможно-

стью отнести данную выборку к какому-либо теоретическому распределению не представляется возможным, т.к. не прослеживается какой-либо закономерности.

Если говорить о том, какое из теоретических распределений наиболее качественно описывает данные выборки, то все распределения показывают себя примерно одинаково в том случае, если выборка описывается хорошо. Нельзя, однако, сказать, что какое-то из теоретических распределений предпочтительно, т.к. строго принять ни одну из гипотез согласия ни на одной из выборок не удалось. Нельзя однозначно указать причину этого, так как:

- Зрительно многие выборки нельзя отнести к одной категории (многие из них выглядят уникально), и не все из них могут быть описаны представленными выше распределениями (если говорить шире, то в целом каким-либо теоретическим распределением);
- Те из выборок, которые выглядят похожими друг на друга, и хорошо описываются (по крайней мере, зрительно) представленными выше распределениями, содержат в себе большое количество элементов. Данный факт сильно влияет на статистики критериев согласия, что не позволяет строго принять гипотезы согласия.

Распределение Райса как качественно (по значению статистики критерия), так и зрительно, оно, пусть и ненамного, превосходит остальные два по степени близости к эмпирическим данным.

По причине того, что выборки имеют большой объем было решено искать новые подходы к исследованию данных, которые будут рассмотрены в главе 3.

Глава 3. Альтернативные методы исследования эмпирических данных

По причине того, что все гипотезы согласия, проверенные на выборках замирания амплитуды, пришлось отклонить, было решено обратить внимание на другие методы анализа экспериментальных данных.

3.1 Исследование выборок кривыми Пирсона

3.1.1 Обзор кривых Пирсона

Были изучены несколько литературных источников [15, 16, 17], в которых рассматриваются методы аппроксимации экспериментальных данных. Согласно [15], использование рядов Тейлора и Фурье для класса данных задач затруднительно, поскольку они не обладают необходимыми свойствами плотности распределения. Ряд Грамма-Шарлье также не может быть использован в данном исследовании, поскольку он близок к вариациям нормального распределения. При использовании данного ряда с другими видами распределений мы можем столкнуться с его неконсистентным поведением. Рекомендации [15] уточняют, что использовать его желательно только в том случае, если коэффициент асимметрии незначительно отличается от нулевого. То же можно сказать и относительно кривых Джонсона.

Наиболее пригодными в контексте нашего исследования являются кривые Пирсона [16]. Данные кривые описываются дифференциальным уравнением вида:

$$\frac{dy}{y} = -\frac{x-M}{b_2x^2+b_1x+b_0}dx,$$

где началом отсчета x служит середина ряда, M - мода выборки. Данное дифференциальное уравнение выражает общие свойства функции распределения. Постоянные коэффициенты b_2, b_1, b_0 вычисляются на основе начальных и центральных моментов исследуемой выборки.

Вообще говоря, кривые Пирсона делятся на 7 семейств кривых, каждое из которых включает в себя ряд теоретических распределений. Согласно [17], распределения, которые были рассмотрены нами ранее, т.е. нормальное, Райса и Накагами, относятся к I семейству кривых Пирсона. Поэтому в данном исследовании мы сосредоточимся именно на нем.

Классификация на семейства производится при помощи переменной, названной *каппой Пирсона*:

$$k = -\frac{b_1^2}{4b_2b_0}$$

Кривая может быть отнесена к I семейству, в том случае, если значение $k < 0$. Общее уравнение кривых I семейства выглядит следующим образом:

$$y = \begin{cases} 0, & \text{если } x < a_1, \\ y_0\left(1 + \frac{x}{a_1}\right)^{m_1}\left(1 - \frac{x}{a_2}\right)^{m_2}, & \text{если } a_1 < x < a_2, \\ 0, & \text{если } x > a_2, \end{cases}$$

где коэффициенты a_1, a_2, m_1, m_2 также вычисляются на основе начальных и центральных моментов исследуемой выборки.

При равенстве $k = 0$ кривая относится к семейству нормального распределения.

3.1.2 Классификация данных по семействам кривых Пирсона

По аналогии с предыдущими этапами исследований, мы не будем приводить результаты классификации для всех выборок, ввиду их большого объема. Поэтому ограничимся несколькими иллюстрационными примерами и приведем вывод по результатам исследования в целом.

Рассмотрим выборку *VBAU_R : 10 : 15dB_ch2*. По результатам классификации данную выборку можно отнести к I семейству кривых Пирсона, т.к. значение каппы Пирсона меньше нуля. Найденные коэффициенты функции следующие:

- $a_1 = 15.79$;
- $a_2 = 44.56$;
- $m_1 = 24.81$;
- $m_2 = 68.27$;
- $y_0 = 0.144$.

На рис. 11 представлен график сравнения эмпирической и теоретической (распределение Райса) плотностей данной выборки, а также наложена кривая Пирсона, которая получилась в результате аппроксимации эмпирических данных.

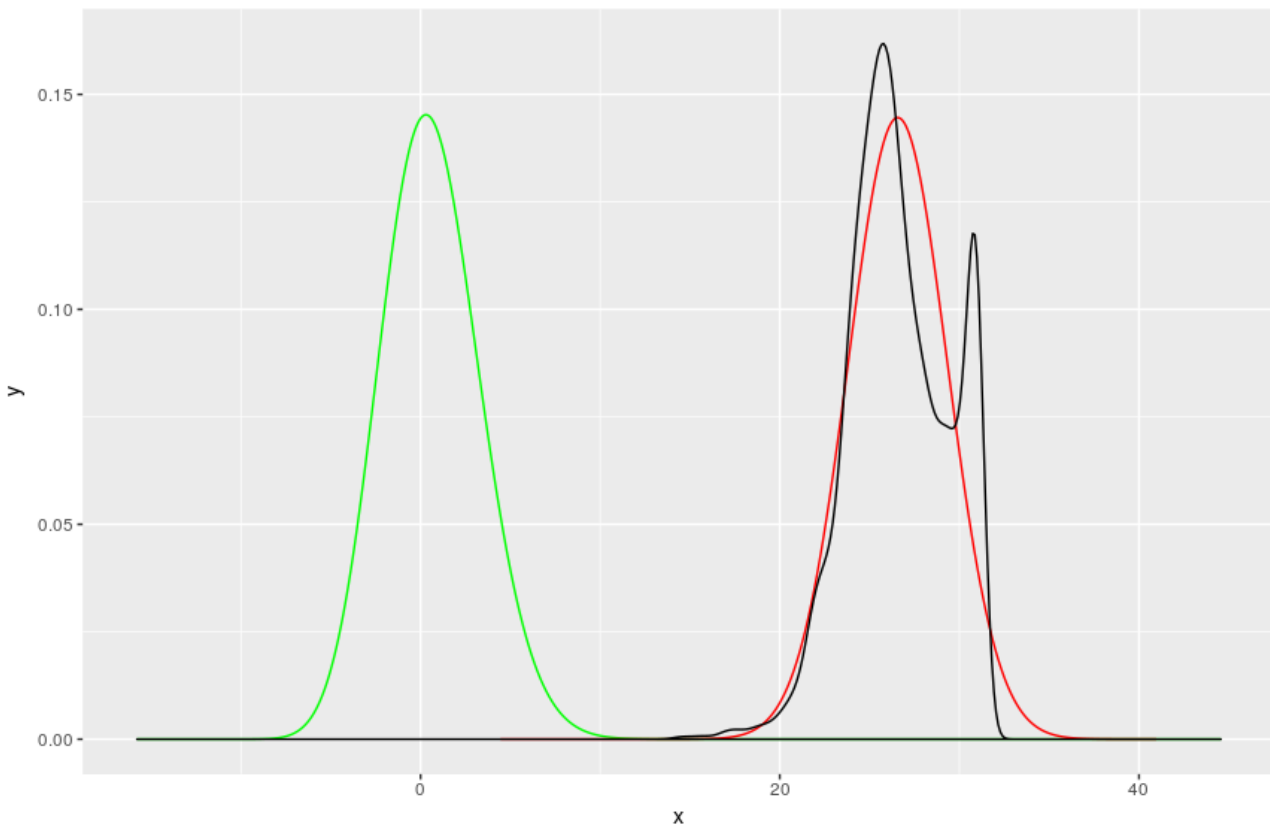


Рис. 11: Сравнительный график для $VBAU_R : 10 : 15dB_ch2$, черным цветом представлена эмпирическая плотность, красным - теоретическая плотность Райса, зеленым - кривая Пирсона.

Следует помнить, что все исследуемые выборки были сдвинуты на 20 dB вправо с целью иметь возможность исследовать их на принадлежность распределениям Райса или Накагами. На рис. 11 можно наблюдать факт того, что центр аппроксимированной нами кривой находится левее, чем центры эмпирической и теоретической плотностей. Расстояние между ними, однако, составляет примерно 26 dB (установлено эмпирически). Таким образом можно утверждать, что построенная нами кривая Пирсона находится левее, чем плотность на 6 dB. Однозначно определить причину этого на данном этапе исследования не удалось. Тем не менее, если совместить теоретическую плотность Райса для данной выборки с построенной кривой Пирсона, можно наблюдать практически полное их совпадение (см рис. 12).

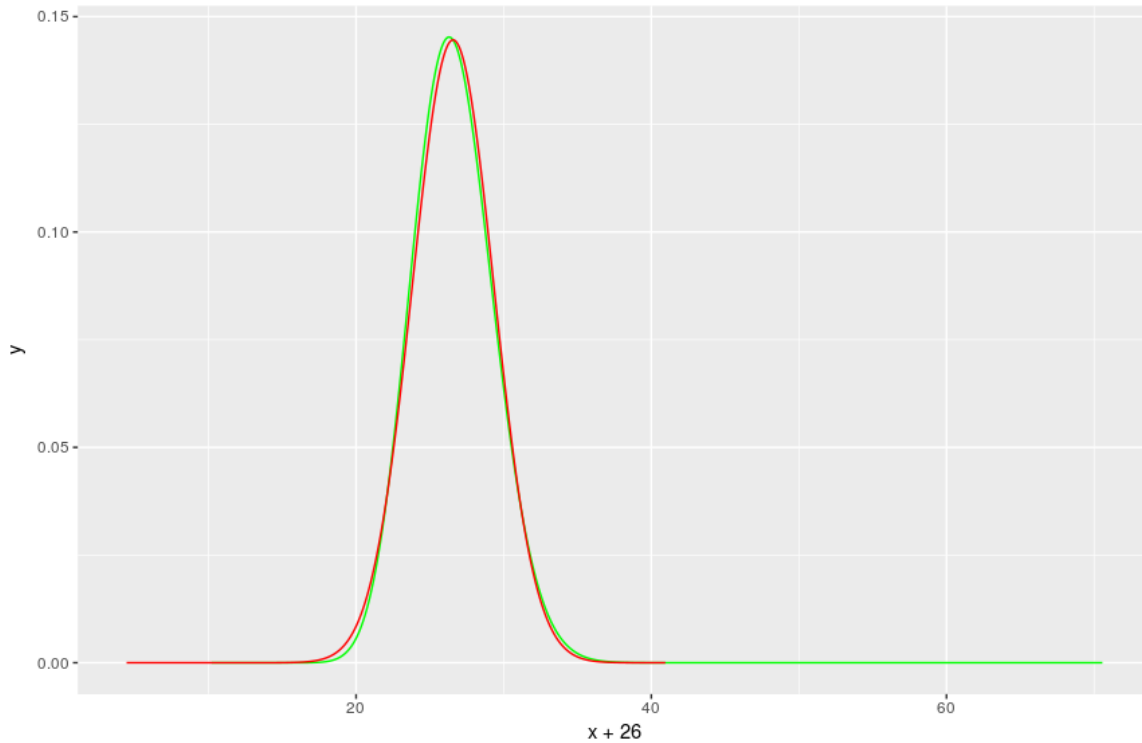


Рис. 12: Сравнительный график для $VBAU_R : 10 : 15dB_ch2$, красным цветом представлена теоретическая плотность Райса, зеленым - кривая Пирсона, сдвинутая на 26 dB вправо.

К сожалению, не все построенные кривые Пирсона давали похожую степень совпадения с теоретическими плотностями распределений. Можно, однако, отметить еще пару кривых, степень схожести которых с соответствующей

теоретической плотностью была такой же. На рис. 13 представлены аналогичные графики сравнений для выборок $VIU_T : 10 : 50dB_20grad_ch2$ и $VIU_T : 10 : 35dB_ch2$. В обоих случаях наблюдался подобный сдвиг, для первой выборки он составил 31 дВ, для второй - 22 дВ.

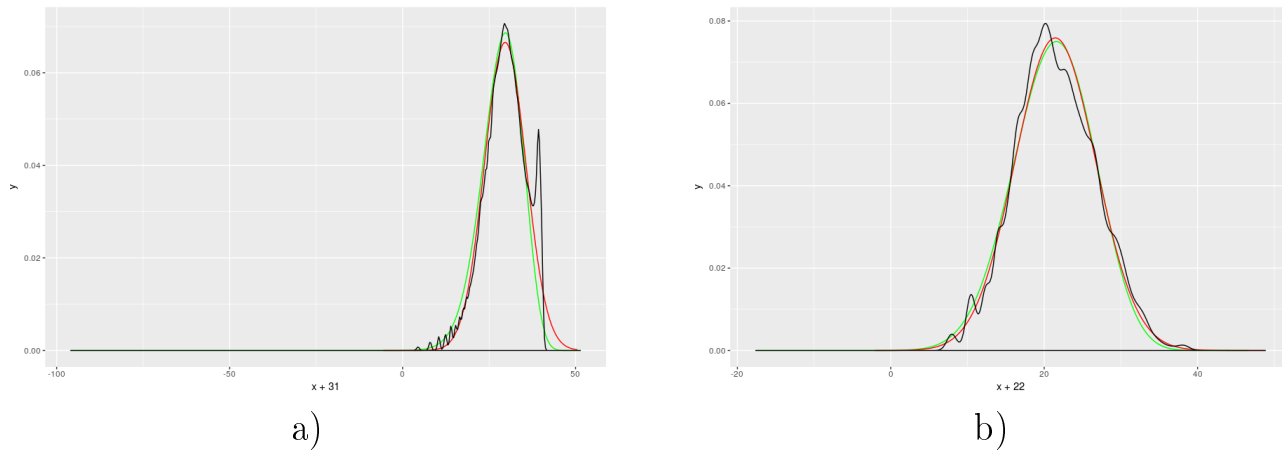


Рис. 13: Сравнительные графики для
 а) $VIU_T : 10 : 50dB_20grad_ch2$,
 б) $VIU_T : 10 : 35dB_ch2$.

Черным цветом представлена эмпирическая плотность, красным - теоретическая плотность Райса, зеленым - кривая Пирсона.

Как видно из рисунков, построенные нами кривые для этих 3-х выборок достаточно хорошо совпадают как с эмпирическими, так и с теоретическими плотностями Райса. К сожалению, такого нельзя сказать обо всем семействе данных по замиранию амплитуды. Чаще всего построенная кривая Пирсона достаточно хорошо повторяла внешний вид только эмпирической плотности. Этого не происходило в тех случаях, когда эмпирическая плотность выборки имела внешний вид, схожий с представленным на рис. 10. А также в тех случаях, когда рассчитать аналитическое представление кривой Пирсона не представлялось возможным ввиду того, что вычисляемые для нее коэффициенты были слишком велики для представления при помощи пакета R.

На рис. 14 продемонстрирован пример построенной кривой Пирсона, внешний вид которой типичен для оставшихся выборок амплитуды. Данная кривая построена для выборки $VBAU_T : 19 : 40dB_20grad_ch2$. Сдвиг пика построенной кривой Пирсона от пика эмпирической плотности данной выборки

ки также присутствует и составляет 33 dB.

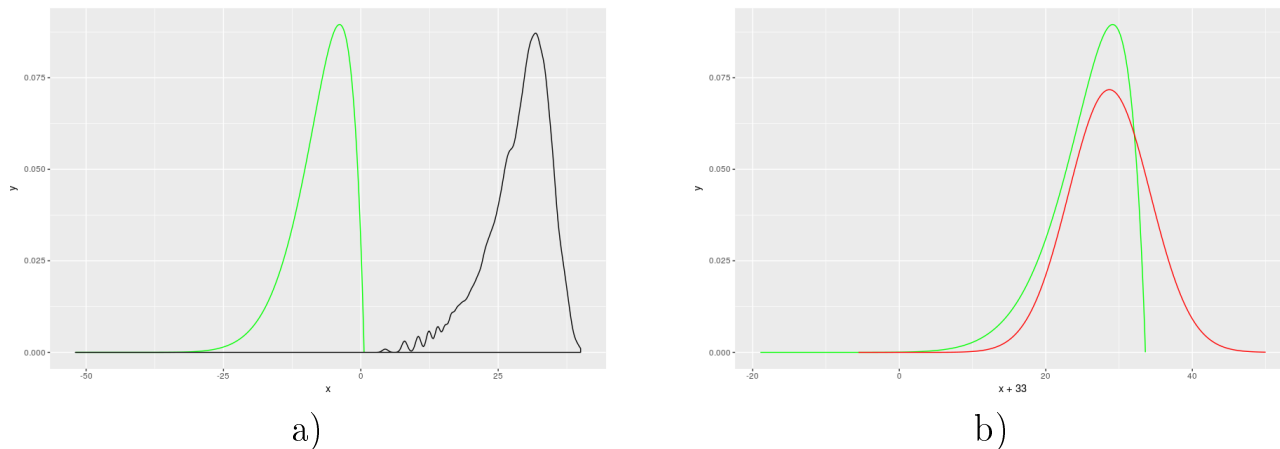


Рис. 14: Сравнительные графики для $VIU_T : 10 : 50dB_80m$:

- a) сравнение с эмпирической плотностью,
- b) сравнение с теоретической плотностью Райса.

Черным цветом представлена эмпирическая плотность, красным — теоретическая плотность Райса, зеленым — кривая Пирсона.

В целом, в результате аппроксимации выборок замирания амплитуды при помощи кривых Пирсона, мы получили ряд аналитических функций, которыми можно представлять плотности исследуемых наборов данных в зависимости от их класса. К минусам можно отнести тот факт, что не для всех выборок удалось найти данную функцию ввиду того, что получаемые коэффициенты кривых технически невозможно вычислить в \mathbb{R} .

К тому же, невыясненной осталась проблема со сдвигами получаемых функций и плотности изначальных данных относительно друг друга.

3.1.3 Исследование сдвигов

Возможной причиной наличия обнаруженных сдвигов может являться влияние волнения морской поверхности. Для того, чтобы проверить данную гипотезу, было необходимо вычислить средний показатель волнения на момент записи показаний замирания амплитуды, после чего произвести регрессионный анализ зависимости величины сдвига от показателя волнения. Сопоставление величин сдвига и волнения можно наблюдать в таблице 5.

Таблица 5: Таблица зависимости величины сдвига от показателя волнения

Идентификатор выборки	Сдвиг (dB)	Средний показатель волнения по выборке "waveHeight"	Средний показатель волнения по выборке "waveHeight2"
<i>VBAU_T : 10 : 30dB_10grad</i>	31	0.13	-0.343
<i>VBAU_T : 19 : 40dB_80m</i>	33	0.11	-0.342
<i>VBAU_T : 10 : 40dB_40grad</i>	36	0.11	-0.341
<i>VBAU_R : 10 : 5dB_ch2</i>	23.5	-0.024	0.0016
<i>VBAU_R : 10 : 15dB_ch2</i>	26	-0.013	0.0014
<i>VBAU_R : 10 : 65dB_ch2</i>	35	-0.034	-0.39
<i>VBAU_R : 32 : 80dB_30grad_ch2</i>	33	0.027	-0.44
<i>VIU_T : 15 : 40dB_100m_ch2</i>	33	0.115	-0.347
<i>VIU_T : 17 : 40dB_90m_ch2</i>	36	-0.0044	-0.48
<i>VIU_T : 29 : 50dB_20grad_ch2</i>	31	0.11	-0.38
<i>VIU_T : 35 : 45dB_40m_ch2</i>	33	-0.064	-0.42
<i>VIU_T : 35 : 50dB_40m_ch2</i>	31	-0.134	-0.47

В данной таблице представлены только те выборки замираний амплитуды, для которых имелось возможным построить кривую Пирсона I семейства и имелись соответствующие по времени показатели волнения (т.к. некоторые показания замираний амплитуды снимались раньше снятия показаний волнения морской поверхности).

По имеющимся данным было произведено два линейных регрессионных анализа для первой и второй выборок волнения соответственно.

Результаты анализа для первой выборки следующие:

$$Sh = 31.5406 + 9.1669 \cdot wH,$$

где Sh - это сдвиг, wH - средняя высота волны, записанной с показаний первого волнографа.

Полученная модель должна удовлетворять основным постулатам регрессионного анализа, согласно [18, 19, 22]. Рассмотрим доверительные интервалы для полученных коэффициентов в таблице 6.

Таблица 6: Доверительные интервалы коэффициентов 1-ой модели

Доверительные интервалы			
	2.5%	97.5%	Значимость
Свободный коэффициент	28.926864	34.154430	Значим
wH	-20.827393	39.161128	Не значим

Как можно видеть, коэффициент при высоте волны содержит ноль в своем доверительном интервале, что может означать тот факт, что он не значим. Более того, скорректированный коэффициент R^2 равен -0.051. Поэтому данную модель можно признать неадекватной.

Рассмотрим результаты анализа для второй модели:

$$Sh = 25.3487 - 19.5786 \cdot wH_2,$$

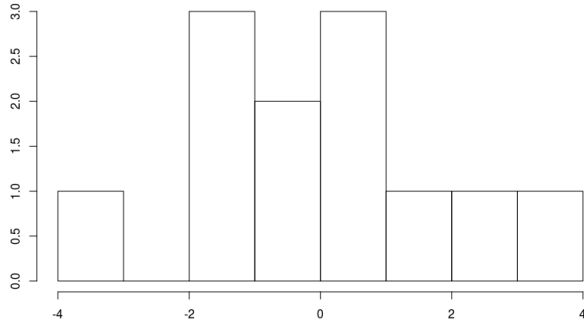
где Sh - это сдвиг, wH_2 - средняя высота волны, записанная с показаний второго волнографа. Доверительные интервалы для коэффициентов данной модели представлены в таблице 7.

Таблица 7: Доверительные интервалы коэффициентов 2-ой модели

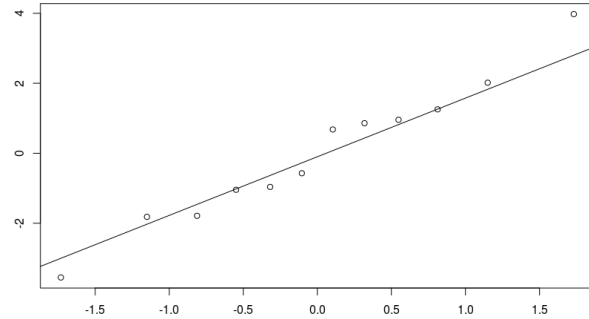
Доверительные интервалы			
	2.5%	97.5%	Значимость
Свободный коэффициент	22.149198	28.548176	Значим
wH_2	-28.372603	-10.784531	Значим

Можно видеть, что оба коэффициента данной модели являются значимыми. Скорректированный коэффициент R^2 для данной модели равен 0,682, что является приемлемым уровнем. На рис. 15 рассмотрены гистограмма распределения остатков модели и график квантилей.

По внешнему виду графиков можно сделать предположение о том, что остатки модели распределены нормально. Проверим данную гипотезу при помощи тестов Шапиро-Уилка, Колмогорова и Пирсона [23, 24, 25].



a)



b)

Рис. 15: Представление остатков регрессионной модели: а) гистограмма, б) квантили.

Тест Шапиро-Уилка на согласие с нормальным распределением:

$$W = 0.977452, p\text{-value} = 0.97142$$

Тест Колмогорова на согласие с нормальным распределением:

$$D = 0.131404, p\text{-value} = 0.96857$$

Тест Пирсона на согласие с нормальным распределением:

$$P = 3, p\text{-value} = 0.39163$$

Нулевая гипотеза не отвергается по результатам всех тестов. Более того, проверка гипотезы с помощью данных тестов также помогает убедиться в том факте, что среднее арифметическое остатков равняется нулю, а в модели отсутствуют систематические ошибки.

Далее проверим гомоскедастичность остатков модели при помощи тестов Гольфельда-Квандта и Уайта [5, 26].

Тест Гольфельда-Квандта на гомоскедастичность остатков:

$$W = 0.977452, p\text{-value} = 0.97142$$

Тест Уайта на гомоскедастичность остатков:

$$GQ = 0.924763, p\text{-value} = 0.5293$$

Нулевая гипотеза о гомоскедастичности остатков также не отвергается. Наконец проверим гипотезы о том, что остатки модели независимы. Для проверки данной гипотезы будем использовать тесты Дарбина-Уотсона и Льюинга-

Бокса [5].

Тест Дарбина-Уотсона на независимость остатков:

$$DW = 1.92546, p\text{-value} = 0.38215$$

Тест Льюинга-Бокса на независимость остатков:

$$X^2 = 0.158526, p\text{-value} = 0.69052$$

По проведенному исследованию можно сказать, что построенная модель зависимости сдвига от волнения морской поверхности является адекватной и пригодна к использованию. Таким образом, уровень волнения влияет на сдвиг обратно пропорционально. Однако абсолютные значения остатков данной модели могут достигать 4 dB, что может не являться пригодным результатом, поскольку такой уровень точности может быть непригоден.

3.1.4 Выводы из анализа с помощью кривых Пирсона

Проведенные исследования дали нам возможность представлять законы распределения, из которых извлечены некоторые из исследуемых выборок, в виде алгебраических функций. К сожалению, не для всех выборок удалось построить аппроксимирующую кривую Пирсона по причине того, что вычисление коэффициентов для них невозможно с программной точки зрения. В случае трех из исследованных выборок, можно сказать, что аналитические кривые Пирсона, построенные для них, и теоретические плотности распределения Райса, вид которых мы получили в гл. 2, зрительно практически совпадают.

Факт наличия сдвигов между аналитическими кривыми Пирсона и эмпирическими плотностями выборок не удалось строго объяснить аналитически. Несмотря на это, была исследована зависимость данных сдвигов от волнения морской поверхности, показания которого снимались одновременно со снятием показаний замирания амплитуды. Было построено две модели, т.к. в нашем распоряжении имелось две выборки волнения. Из них адекватной признана только одна, что может указывать на один из двух выводов. Первый - положение волнографа, фиксировавшего показания использованной в модели выборки волнения, является важным с точки зрения корректировки

сдвигов. Второй - результат данного исследования может являться удачным стечением обстоятельств. Т.е. гипотеза о том, что рассматриваемые сдвиги зависят от волнения не верна, а полученная адекватная модель — неудачное совпадение.

3.2 Исследование подвыборок

3.2.1 Описание подходов

Ввиду того, что объемы исследуемых выборок замираний амплитуды достаточно большие, было решено обозначить такие подходы к исследованию, при которых возможно сокращение количества элементов выборок, т.е. извлекать подвыборки [7, 19]. Условно, принятые нами подходы можно разделить на два направления:

- имитировать получение в качестве исходных данных большее количество выборок меньшего объема, т.е. разбивать исходный массив данных на подмассивы определенного объема, соблюдая при этом порядок элементов выборки;
- рассматривать исходную выборку в качестве генеральной совокупности и извлекать подвыборки случайным образом.

Остановимся на каждом из подходов подробнее.

3.2.2 Исследование последовательных подвыборок

Будем оперировать исходными выборками следующим образом:

1. Исходную выборку последовательно разобьем на подвыборки объема n . Подобное разбиение будет имитировать подобие другого натурального эксперимента, при котором мы получаем большее количество выборок меньшего объема.

2. Проверим гипотезы согласия на выбранные теоретические распределения для каждой из полученных подвыборок.
3. Проанализируем общий процент принятых гипотез согласия с каждым из выбранных теоретических распределений для всех подвыборок в контексте одной исходной выборки.
4. Рассмотрим статистические характеристики полученных оценок параметров теоретических распределений для подвыборок.
5. Повторим эти шаги для всех имеющихся выборок замирания амплитуды.

Как и в случае с исследованием на предмет согласия с выбранными теоретическими распределениями исходных выборок, мы не будем приводить результаты исследования для всех выборок, ввиду их громоздкости, а ограничимся несколькими примерами.

Рассмотрим выборку амплитуды радиосигнала $VBAU_R : 10 : 15dB_ch2$ как выборку с минимальным объемом среди всех прочих (4.374 элементов).

Последовательно разобьем эту выборку на подвыборки объемом 200 элементов. Для каждой из полученных подвыборок проверим гипотезы согласия на соответствие теоретическим распределениям Райса и Накагами, а также нормальному распределению. Гипотезы согласия будем проверять при помощи критерия Колмогорова и критерия Пирсона. Затем, аналогичным образом, повторим проверку для подвыборок объемом в 400 и 600 элементов. Подобные разбиения соответствуют снятию показаний замираний амплитуды радиосигнала в течение 2-ух, 4-ех и 6-ти секунд соответственно.

Результаты проверки гипотез согласия для подвыборок данной выборки представлены в таблице 8.

Как видно из таблицы, наибольшему проценту принятых гипотез согласия по критерию Колмогорова соответствует разбиение исходной выборки на подвыборки объемом в 200 элементов. В этом случае доля принятых гипотез для каждого теоретического распределения примерно одинакова. Гипотезы

Таблица 8: Результаты исследования последовательных подвыборок с различным объемом. Исходная выборка – $VBAU_R : 10 : 15dB_ch2$

Доля принятых гипотез	Объем подвыборок		
	200 эл-в	400 эл-в	600 эл-в
<i>По критерию Колмогорова</i>			
Нормальное распр.	0.43	0.1	0
Распр. Накагами (метод Гринвуда)	0.38	0.1	0
Распр. Накагами (метод моментов)	0.43	0.3	0
Распр. Райса	0.43	0	0
<i>По критерию Пирсона</i>			
Нормальное распр.	0.24	0.5	0.43
Распр. Накагами (метод Гринвуда)	0.23	0.3	0.43
Распр. Накагами (метод моментов)	0.19	0.3	0.43
Распр. Райса	0.28	0.5	0.57

согласия по критерию Пирсона ведут себя несколько иным образом. С ростом объема подвыборок растет доля принятых гипотез. При этом видно, что доля принятых гипотез для распределения Райса больше, чем для остальных теоретических распределений. Однако, если мы разобьем выборку на подвыборки объемом в 1000 элементов, что составляет четверть от объема выборки, процент принятых по критерию Пирсона гипотез упадет до нуля. Таким образом, можно сказать, что прирост доли принятых гипотез для данной выборки ограничен.

На рисунке 16 представлена гистограмма одной из подвыборок, для которой гипотезы согласия с каждым из теоретических распределений были отклонены. Видно, что в данной подвыборке большое количество повторяющихся значений сосредоточено в интервале от 31 до 31.25, что является причиной отклонения гипотезы. Данный результат может быть обобщен для остальных подвыборок данной выборки, гипотезы согласия для которых были отклонены.

Рассмотренная нами выборка является минимальной среди всех прочих. И даже при задании небольших объемов подвыборок, их общее коли-

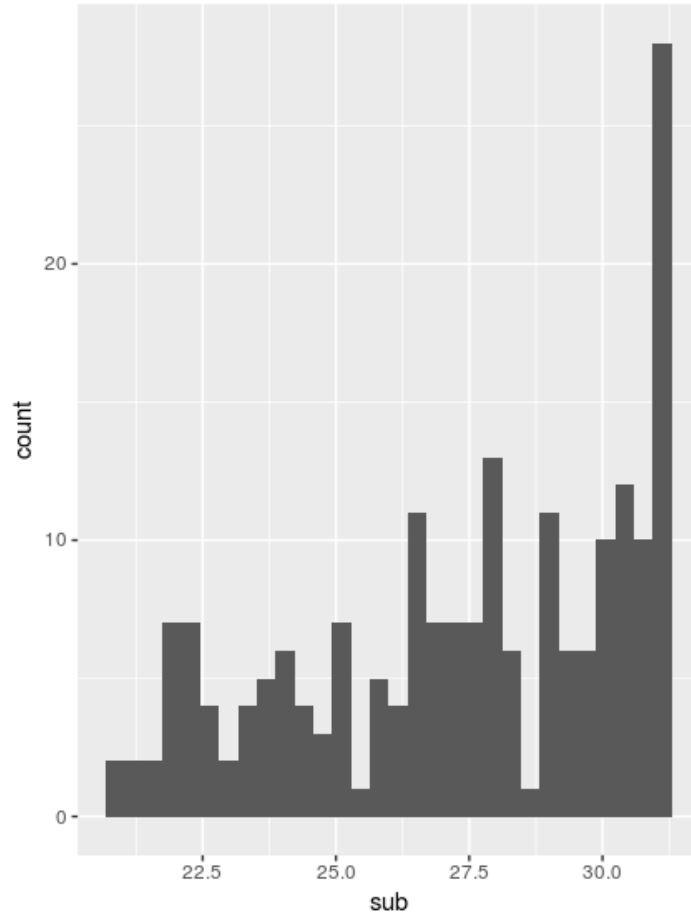


Рис. 16: Гистограмма подвыборки.

чество является небольшим. Поэтому рассмотрим выборку $VBAU_T : 15 : 35dB_100m_ch2$ как выборку со средним объемом (711.538 элементов). Результаты разбиения и проверки гипотез согласия представлены в таблице 9.

Аналогично с результатами для выборки $VBAU_R : 10 : 15dB_ch2$, наибольший процент принятых гипотез по критерию Колмогорова наблюдается при разбиении исходной выборки на подвыборки объемом в 200 элементов, что справедливо и для гипотез согласия Пирсона в контексте данной выборки. Также видно, что наибольшая доля принятых гипотез согласия приходится на распределения Райса и нормальное распределение.

Аналогично предыдущей выборке, гипотезы согласия для подвыборок не принимались в том случае, если на краях подвыборки присутствовало большое количество повторяющихся элементов.

В таблице 10 представлены основные статистические характеристики

Таблица 9: Результаты исследования последовательных подвыборок с различным объемом. Исходная выборка $VBAU_T : 15 : 35dB_100m_ch2$

Доля принятых гипотез	Объем подвыборок		
	200 эл-в	400 эл-в	600 эл-в
<i>По критерию Колмогорова</i>			
Нормальное распр.	0.18	0.09	0.05
Распр. Накагами (метод Гринвуда)	0.16	0.07	0.03
Распр. Накагами (метод моментов)	0.17	0.07	0.04
Распр. Райса	0.19	0.1	0.05
<i>По критерию Пирсона</i>			
Нормальное распр.	0.27	0.21	0.15
Распр. Накагами (метод Гринвуда)	0.24	0.17	0.12
Распр. Накагами (метод моментов)	0.21	0.13	0.07
Распр. Райса	0.27	0.21	0.15

(среднее арифметическое и стандартное отклонение) оценок параметров теоретических распределений для последовательных подвыборок объемом в 200 элементов выборки $VBAU_T : 15 : 35dB_100m_ch2$. К рассмотрению принимались только те подвыборки, для которых удалось принять гипотезу согласия с соответствующим распределением.

Из таблицы видно, что стандартное отклонение совокупности оценок распределения Накагами для обоих методов оценивания очень велико, из чего можно сделать вывод о большой разнице между оценками для различных подвыборок. Стандартное отклонение совокупности оценок нормального распределения и распределение Райса, напротив, невелико. Соответственно, все подвыборки имеют приблизительно одинаковые оценки для данных двух теоретических распределений.

Таблица 10: Статистические характеристики оценок параметров распределений последовательных подвыборок выборки $VBAU_T : 15 : 35dB_100m_ch2$.

Параметр	Среднее арифметическое	Стандартное отклонение
<i>Нормальное распределение</i>		
μ	26.9	3.01
σ	3.70	1.34
<i>Распределение Накагами (метод Гринвуда)</i>		
μ	23.9	24.13
ω	760.23	156.07
<i>Распределение Накагами (метод моментов)</i>		
μ	24.18	23.75
ω	759.57	157.16
<i>Распределение Райса</i>		
ν	26.61	3.14
σ	3.77	1.38

Таким образом, можно сделать следующие выводы:

- Гипотезы согласия на подвыборках отвергаются по причине большого количества повторяющихся значений на их краях. Данный факт может иметь воздействие на принятие гипотез согласия на исходных выборках.
- Наибольшая доля принятых гипотез достигается при объеме подвыборок в 200 элементов. С увеличением объема доля принятых гипотез, как правило, снижается.
- Доля принятых гипотез на согласие с распределением Райса и нормальным распределением больше, чем для распределения Накагами.
- Подвыборочные оценки параметров нормального распределения и распределения Райса лежат компактнее, нежели оценки параметров распределения Накагами.

Данные выводы можно обобщить на всю совокупность выборок замираний амплитуды радиосигнала.

3.3.3 Исследование случайных подвыборок

Извлечение случайных подвыборок из выборки несет в себе больший смысл, нежели разбиение выборки на последовательные подвыборки, с точки зрения статистической значимости. Поэтому построим продолжение нашего исследования следующим образом:

1. Из исходной выборки случайным образом будем извлекать подвыборку объема n . Извлечение случайной подвыборки будем производить при помощи встроенных средств в прикладном программном пакете R.
2. Проверим гипотезы согласия с выбранными теоретическими распределениями для извлеченной подвыборки.
3. Повторим первые два пункта m раз.
4. Проанализируем общий процент принятых гипотез согласия с каждым из выбранных теоретических распределений для m случайно извлеченных подвыборок в контексте одной исходной выборки.
5. Рассмотрим статистические характеристики полученных оценок параметров теоретических распределений, а также проверим гипотезы однородности для извлеченных подвыборок.
6. Повторим описанные шаги для всех имеющихся выборок замирания амплитуды.

Как и в случае с предыдущим пунктом, мы рассмотрим пару выборок в качестве примера и обобщим выводы, полученные в результате этой части исследования.

Рассмотрим выборку из предыдущего пункта: *VBAU_T : 15 : 35dB_100m_ch2*. Случайным образом будем извлекать из неё подвыборки объемом в 200, 400 и 600 элементов и проверять на них гипотезы согласия с выбранными теоретическими распределениями. Будем проводить 5.000 повторений теста на

согласие для исследования доли принятых гипотез на каждое сочетание объема извлекаемых подвыборок и теоретического распределения. Результаты расчетов представлены в таблице 11.

Таблица 11: Результаты исследования случайных подвыборок с различным объемом. Исходная выборка $VBAU_T : 15 : 35dB_100m_ch2$

Доля принятых гипотез	Объем подвыборок		
	200 эл-в	400 эл-в	600 эл-в
<i>По критерию Колмогорова</i>			
Нормальное распр.	0.86	0.43	0.12
Распр. Накагами (метод Гринвуда)	0.53	0.06	0.002
Распр. Накагами (метод моментов)	0.78	0.26	0.04
Распр. Райса	0.87	0.43	0.11
<i>По критерию Пирсона</i>			
Нормальное распр.	0.79	0.39	0.18
Распр. Накагами (метод Гринвуда)	0.55	0.14	0.03
Распр. Накагами (метод моментов)	0.27	0.03	0.003
Распр. Райса	0.73	0.34	0.13

Аналогично ситуации с последовательными выборками, наибольшая доля принятых гипотез наблюдается при объеме выборок в 200 элементов. При дальнейшем увеличении объема случайных подвыборок доля успешно принятых гипотез стремится к нулю. Стоит заметить, что, как и в случае с исследованием последовательных подвыборок, доля принятых гипотез согласия с нормальным распределением и распределением Райса на порядок выше, чем доля принятых гипотез согласия с распределением Накагами. При этом между собой доли принятых гипотез для нормального распределения и распределения Райса отличаются незначительно.

Аналогичным образом исследуем выборку $VIU_T : 35 : 0dB_0dB_5dB_20dB$, которая является результатом соединения 4-х выборок. Результаты исследования представлены в таблице 12.

Из таблицы видно, что для случайных подвыборок данной выборки, объем которых равен 200 элементам, практически всегда принимаются гипо-

Таблица 12: Результаты исследования случайных подвыборок с различным объемом. Исходная выборка – $VIU_T : 35 : 0dB_0dB_5dB_20dB$

Доля принятых гипотез	Объем подвыборок		
	200 эл-в	400 эл-в	600 эл-в
<i>По критерию Колмогорова</i>			
Нормальное распр.	0.99	0.95	0.84
Распр. Накагами (метод Гринвуда)	0.92	0.59	0.24
Распр. Накагами (метод моментов)	0.95	0.86	0.63
Распр. Райса	0.99	0.94	0.84
<i>По критерию Пирсона</i>			
Нормальное распр.	0.95	0.89	0.83
Распр. Накагами (метод Гринвуда)	0.85	0.59	0.42
Распр. Накагами (метод моментов)	0.81	0.53	0.32
Распр. Райса	0.92	0.88	0.80

тезы согласия с нормальным распределением и распределением Райса, разница между долями которых незначительна. Доля принятых гипотез согласия для распределения Накагами также велика, однако она уступает двум вышеупомянутым. Как и в предыдущих случаях, увеличение объема подвыборок приводит к снижению доли принятых гипотез.

В таблице 13 рассмотрены основные статистические характеристики (среднее арифметическое и стандартное отклонение) оценок параметров теоретических распределений для случайных подвыборок объемом в 200 элементов исходной выборки $VIU_T : 35 : 0dB_0dB_5dB_20dB$. К рассмотрению принимались только те подвыборки, для которых удалось принять гипотезу согласия с соответствующим распределением в рамках 5.000 повторных извлечений.

Из таблицы видно, что оценочные параметры всех теоретических распределений лежат довольно плотно, в отличие от последовательных подвыборок. Это говорит о том, что данные параметры можно использовать в качестве оценочных параметров и для исходных выборок. Однако, если рассматривать гипотезы статистической однородности для данных подвыборок, то в сред-

Таблица 13: Статистические характеристики оценок параметров распределений случайных подвыборок выборки $VIU_T : 35 : 0dB_0dB_5dB_20dB$.

Параметр	Среднее арифметическое	Стандартное отклонение
<i>Нормальное распределение</i>		
μ	24.78	0.38
σ	5.34	0.27
<i>Распределение Накагами (метод Гринвуда)</i>		
μ	5.41	0.63
ω	643.03	18.49
<i>Распределение Накагами (метод моментов)</i>		
μ	6.02	0.64
ω	643.19	18.82
<i>Распределение Райса</i>		
ν	24.17	0.4
σ	5.38	0.29

нем каждая подвыборка однородна с 20% остальных подвыборок, что может свидетельствовать о слабой однородности данной совокупности подвыборок.

Данные результаты можно обобщить на большинство из исследуемых выборок замираний амплитуды. Но нельзя сказать, что каждая из них так же хорошо описывается нормальным распределением и распределением Райса при извлечении случайных подвыборок. К примеру, у выборки $VBAU_R : 21 : 85dB_20grad_ch2$ доля принятых гипотез согласия критерием согласия Колмогорова с нормальным распределением и распределением Райса составляет всего 0.37 для обеих гипотез (для критерия Пирсона — 0.34 и 0.25 соответственно) при объеме подвыборок в 200 элементов. К сожалению, какой-либо закономерности доли принятых гипотез от класса выборки амплитуды радиосигнала обнаружить не удалось.

Таким образом, можно сделать следующие выводы по данной части исследования:

- Наибольшая доля принятых гипотез достигается при объеме подвыборок в 200 элементов. С увеличением объема доля принятых гипотез сни-

жается.

- Доля принятых гипотез согласия с распределением Райса и нормальным распределением больше, чем с распределением Накагами. Стоит отметить, что для некоторых выборок доля принятых гипотез на случайных подвыборках была близка к единице.
- Подвыборочные оценки параметров в целом лежат компактнее, по сравнению с предыдущим методом исследования – извлечением последовательных подвыборок.

3.3.4 Выводы из анализа подвыборок

Несмотря на то, что в результате вышеприведенных исследований, доля принятых гипотез согласия с распределением Райса и нормальным распределением приблизительно равна для подвыборок, которые извлекаются различными методами, мы не можем игнорировать результаты аппроксимации кривыми Пирсона. Согласно их результатам, ни одна из исследуемых выборок не попала в семейство кривых Пирсона, которое соответствует нормальному распределению. Таким образом, можно сделать вывод о том, что большинство из исследуемых нами выборок замираний амплитуды радиосигнала распределено по закону Райса, оценки параметров которого рассчитывались по методу моментов.

Глава 4. Взаимосвязь волнения морской поверхности и амплитуды радиосигнала

Задачей данного этапа исследования являлось выявление корреляции между морским волнением и замираниями амплитуды сигнала. И, если таковая будет выявлена, то следующим этапом является построение регрессионных моделей.

4.1 Корреляция волнения и сглаженных значений амплитуды

Необходимым действием перед данным этапом исследования является объединение выборок замирания амплитуды сигнала и волнения морской поверхности в двухмерную выборку. Каждую представленную выборку амплитуды мы объединим с обеими выборками волнения. Принципом объединения выборок является время снятия показаний.

После произведенного соединения, следует отметить два факта:

- не каждую выборку показаний амплитуды удалось объединить с выборками волнения по причине того, что некоторые из выборок показаний амплитуды не имеют общего временного показателя с выборками волнения;
- частота снятия показаний замирания амплитуды сигнала в 4-ре раза превышает частоту снятия показаний волнения морской поверхности (100 показаний в секунду против 4-х).

Таким образом, при объединении выборок игнорируется 96% данных замирания амплитуды. Для того, чтобы сгладить влияние этого факта, мы будем рассматривать корреляцию волнения морской поверхности со сглаженными скользящим средним значениями амплитуды.

Так как нам не удалось установить факт того, что рассматриваемые выборки извлечены из нормального распределения, в качестве коэффициента корреляции мы будем рассматривать коэффициент Спирмена [19]. Ис-

пользование коэффициента ранговой корреляции Спирмена здесь оправдано тем, что после объединения выборок мы получили две двухмерные выборки сопоставления волнения морской поверхности с амплитудой радиосигнала для каждой выборки амплитуды радиосигнала. Таким образом, от количественных шкал мы можем перейти к порядковым, чтобы выявить существование номинальной зависимости, что позволяет нам использовать коэффициент Спирмена [7, 20].

По итогам расчета коэффициента корреляции Спирмена было установлено, что корреляция между волнением морской поверхности и сглаженными значениями амплитуды сигнала, как таковая, отсутствует. На подавляющем большинстве выборок коэффициент корреляции не превышает по модулю 0.05. Исключение составляют две выборки амплитуды:

- *VIU_T : 17 : 40dB_90m_ch2* - коэффициенты равны -0.239 и -0.235 для выборок волнения *waveHeight* и *waveHeight2* соответственно;
- *VBAU_R : 10 : 75dB_ch2* - коэффициенты равны -0.003 и -0.202 для выборок волнения *waveHeight* и *waveheight2* соответственно.

Согласно шкале Чеддока [27], данная связь между волнением морской поверхности и амплитудой радиосигнала может быть классифицирована как *слабая*.

Таким образом, можно установить, что зависимости между значениями замирания амплитуды сигнала и волнением морской поверхности нет для исследуемых данных.

4.2 Корреляция волнения и остаточных значений амплитуды

Так как в результате исследования не обнаружено зависимости между сглаженными значениями амплитуды сигнала и волнением морской поверхности, было решено изменить подход.

Будем рассматривать выборки значений амплитуды как временной ряд. Зависимость будем искать между значениями волнения и значениями временных рядов амплитуды за вычетом трендовых значений. Выделение тренда осуществлялось при помощи метода Хольта [28].

Метод Хольта был выбран как классический метод прогнозирования временных рядов, способный учитывать тенденцию к росту или падению значений временного ряда (т.е. тренда). Метод позволяет, помимо оценки текущего уровня временного ряда, оценивать также его наклон. Текущий уровень и наклон ряда уточняются на каждом шаге исследования временного ряда.

Рассмотрим, в качестве примера, расчет корреляции для выборки амплитуды *VBAU_T : 10 : 40dB_10grad_1_ch2*. Согласно методу Хольта, тренд данного временного ряда амплитуды имеет следующий вид:

$$tr = 13.371 - 0.019 \cdot x,$$

где x - элементы ряда амплитуды. Вычитая таким образом полученный тренд из значений ряда, вычисляем коэффициенты корреляции Спирмена между выборками волнения и рядом амплитуды. Полученные значения корреляций: -0.0108 для выборки *waveHeight* и 0.0313 для выборки *waveHeight2*. Можно видеть, что для данного временного ряда амплитуды корреляция с волнением отсутствует.

Данный результат обобщается на все доступное множество временных рядов амплитуды сигнала. Все полученные значения корреляций, аналогично с предыдущим подходом к вычислению, не превышают значения в 0.05 по модулю, без исключений.

Таким образом, можно сделать вывод о том, что и в этом случае отсутствует зависимость между значениями амплитуды сигнала за вычетом тренда и волнением морской поверхности. При отсутствии статистически значимой корреляции, необходимость в проведении регрессионного анализа отсутствует.

4.3 Алгоритм исследования

В результате данного исследования был разработан алгоритм проведения анализа данных по волнению морской поверхности и амплитудной флуктуации, который может быть использован в контексте подобных исследований.

Алгоритм данного исследования состоял из следующих этапов:

1. Произвести предварительную обработку и анализ данных, включающие в себя:
 - (а) Объединение данных одного рода (волнение/амплитуда) по времени, если таковая необходимость имеется;
 - (б) Описание наборов данных базовыми статистическими характеристиками;
 - (с) Построение гистограммы и графика эмпирической плотности данных;
 - (d) Формулировка и выдвижение изначальных гипотез о виде распределения данных, сделанное на основе результатов предыдущих этапов;
 - (е) Выполнение дополнительных операций над данными, если таковая необходимость имеется. В нашем случае мы переводили данные о замирании амплитуды из логарифмической шкалы методом сдвига.
2. Отобрать методы оценивания параметров для выбранных теоретических распределений.
3. Произвести проверки гипотез согласия при помощи различных критериев. Параметры распределений рассчитываются согласно отобранным методам оценки.
4. По результатам критериев согласия сделать выводы о принадлежности совокупностей данных рассматриваемым теоретическим распределениям.

5. Оценить степень корреляции между волнением морской поверхности и замиранием амплитуды сигнала с целью выяснить целесообразность проведения регрессионного анализа для установления взаимосвязи.
6. В случае значимой корреляции произвести регрессионный анализ данных амплитуды и волнения.
7. В случае исследования данных, описывающих волнение морской поверхности, дополнительно рассмотреть гипотезы однородности для исследуемых наборов данных.
8. В случае отвержения гипотез согласия для предполагаемых теоретических соединений:
 - (a) Построить кривые Пирсона, с целью получения аналитических функций, аппроксимирующих плотности распределений рассматриваемых выборок;
 - (b) Извлечь подвыборки из исходных данных, при условии, что исходный объем данных позволяет это сделать, и исследовать их на предмет соответствия выбранным теоретическим распределениям.

4.4 Выводы из анализа взаимосвязи волнения морской поверхности и амплитуды радиосигнала

В результате проведенного исследования было установлено, что корреляция между волнением морской поверхности и замиранием амплитуды радиосигнала незначительна, как в случае сглаженных значений амплитуды, так и в случае значений амплитуды за вычетом тренда. Ввиду этого, было решено не производить дальнейший регрессионный анализ.

Заключение

Данная работа является отправной точкой в исследовании, направленном на моделирование канала связи приема сигнала со спутника на исследовательскую платформу и передачу его в обратном направлении. Важность данной задачи продиктована стремлением улучшить существующие методы обработки сигналов, поступающих и исходящих с морских судов.

В рамках работы был разработан подход к исследованию задач подобного рода, который включает в себя ряд стандартных статистических операций, таких как выдвижение и проверка гипотез согласия и поиск корреляции между наборами данных, а также применение методов аппроксимации плотности распределений полиномами высших порядков.

Несмотря на то, что в результате представленного исследования не удалось строго получить теоретические законы распределения, которым соответствуют исследуемые наборы данных волнения морской поверхности и замирания амплитуды радиосигнала, нам удалось аппроксимировать поведение плотности выборок для некоторых из наборов данных. Вообще говоря, нельзя говорить о каком-либо обобщении полученных результатов для всех 37-ми рассматриваемых выборок амплитуды. Причиной тому является сильная разобщенность данных в зависимости от класса выборки. Однако согласно результатам анализа подвыборок исследуемых выборок, предположение о том, что большинство из них распределены по закону Райса, можно считать справедливым. Оценки параметров распределения Райса в данном случае были найдены при помощи метода моментов. Данные оценки параметров могут различаться в зависимости от класса выборки.

В результате поиска взаимосвязи между волнением морской поверхности и амплитудой принятого и переданного радиосигнала было установлено, что взаимосвязи между ними не наблюдается в рамках данного набора данных.

Можно с уверенностью сказать, что задача моделирования канала связи не является тривиальной. На каждом этапе исследования наблюдались нюансы и специфические аспекты алгоритмов, существенно зависящих от структу-

ры данных, которая менялась от выборки к выборке. Проведенная в данном исследовании работа является значимой базой для будущих этапов исследования.

Список литературы

- [1] Котельников В. А. Теория потенциальной помехоустойчивости. // М. 1956 г. 152 с.
- [2] Shannon C. A Mathematical Theory of Communication. // Bell System Technical Journal. 1948. vol. 27, p. 379–423, 623–656.
- [3] Возенкрафт Дж., Джекобс И. Теоретические основы техники связи. Пер. с англ. / Под редакцией Добрушина Р. Л. // М. 1969 г. 640 с.
- [4] АО «Концерн «ЦНИИ «Электроприбор» // <http://www.elektropribor.spb.ru/index.html>, 24.04.2018.
- [5] Буре В. М., Парилина Е. М., Седаков А. А., Шевкопляс Е. В. Методы прикладной статистики в R и Excel. // СПб. 2016 г. 152 с.
- [6] R manual. // <https://cran.r-project.org/manuals.html>, 24.04.2018.
- [7] Рао С. Р. Линейные статистические методы и их применения. Пер с англ. / Под редакцией Линника Ю. В. // М. 1968 г. 548 с.
- [8] Kolar R., Jirik R., Jan J. Estimator comparison of the Nakagami-m parameter and its application in echocardiography. // Radioengineering, 2004, vol. 13(1) p. 8-12.
- [9] Яковлева Т. В. Условия применимости статистической модели Райса и расчет параметров райсовского сигнала методом максимума правдоподобия // Компьютерные исследования и моделирование. 2014 г., Т.6, № 1, с.13-25.
- [10] Миддлтон Д. Введение в статистическую теорию связи. // М. 1962 г. 832 с.
- [11] Айвазян С. А., Енюков И. С, Мешалкин Л. Д. Прикладная статистика. // М. 1983 г. 472 с.

- [12] Кобзарь А. И. Прикладная математическая статистика. Для инженеров и научных работников. // М. 2006 г. 816 с.
- [13] Lilliefors H. On the Kolmogorov-Smirnov test for normality with mean and variance unknown // Journal of the American Statistical Association. 1967. vol. 62, no. 318. p. 309-402.
- [14] Перов А. И. Статистическая теория радиотехнических систем. // М. 2003 г. 400с.
- [15] Заикин П. В., Погореловский М. А., Микшина В. С. Аппроксимация эмпирических функций полиномами высших порядков. // Вестник кибернетики. 2015 г. № 4. с. 129—134
- [16] Беренгард Ю., Беренгард С. Выравнивание статистических распределений. Кривые Пирсона. // http://www.simumath.net/library/book.html?code=Treat_Exper_Pearson_distr, 24.04.2018
- [17] Голик Ф. В. Аппроксимация кривыми Пирсона плотности распределения суммы независимых одинаково распределенных случайных величин // Кибернетика и программирование. 2017 г. № 2. с. 17-41.
- [18] Ивченко Г. И., Медведев Ю. И. Математическая статистика: учебное пособие для вузов. // М. 1984 г. 248 с.
- [19] Кендал М., Стьюарт А. Статистические выводы и связи. Пер. с англ. / Под редакцией Колмогорова А. Н. // М. 1973 г. 899 с.
- [20] Макаров А. А., Тюрин Ю. Н. Анализ данных на компьютере. // М. 2002 г. 528 с.
- [21] Кендал М., Стьюарт А. Теория Распределений. Пер. с англ. / Под редакцией Колмогорова А. Н. // М. 1966 г. 588 с.

- [22] Eyduran E., Ozdemir T., Alarслан E. Importance of diagnostics in multiple regression analysis // Journal of Applied Sciences. 2005. no. 5, p. 1792-1796.
- [23] Буре В. М., Парилина Е. М. Теория вероятностей и математическая статистика. // СПб. 2013 г. 416 с.
- [24] Лемешко Б. Ю, Лемешко С. Б, Постовалов С. Н. Мощность критериев согласия при близких альтернативах. // Измерительная техника. 2007 г. № 2. с. 22-27.
- [25] Razali N., Yар. B.W. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests // Journal of Statistical Modeling and Analysis. 2011. vol. 2, no. 1. p. 21-33.
- [26] Breusch T. S., Pagan A. R. A simple test for heteroscedasticity and random coefficient variation // Econometrica. 1979. vol. 47, no. 5. p. 1287-1294
- [27] Спирип А. А., Башина О. Э. Общая теория статистики: Статистическая методология в изучении коммерческой деятельности. // М. 1996 г. 296 с.
- [28] Montgomery D., Jennings C., Kulahci M. Introduction to time series analysis and forecasting. // Wiley-Interscience, 2008 y., p. 472

Приложение

Приложение 1

Таблица 1. Выборки замираний амплитуды, которые могут быть отнесены к теоретическим распределениям.

Класс выборки	Пики теоретического и эмпирического распределений находятся на одной горизонтальной прямой	Пики теоретического и эмпирического распределений незначительно расходятся
VBAU_T : 10	<i>30dB_10grad_ch2, 40dB_10grad_ch2_1</i>	<i>40dB_10grad_ch2</i>
VBAU_T : 15		<i>30dB_100m, 35dB_100m</i>
VBAU_T : 19		<i>35dB_80m</i>
VBAU_R : 10	<i>5dB_ch2, 20dB_ch2</i>	<i>15dB_ch2, 25dB_ch2, 65dB_ch2, 75dB_ch2</i>
VBAU_R : 21		<i>75dB_20grad_ch2, 85dB_20grad_ch2</i>
VBAU_R : 32		<i>80dB_30grad_ch2, 85dB_30grad_ch2</i>
VIU_T : 15		<i>35dB_100m_ch2, 40dB_100m_ch2</i>
VIU_T : 17	<i>40dB_90m_ch2</i>	
VIU_T : 29	<i>50dB_20grad_ch2</i>	
VIU_T : 32	<i>60dB_ch2</i>	
VIU_T : 35	<i>Соединенные выборки (0dB_ch2_1, 0dB_ch2_2, 5dB_ch2, 20dB_ch2), 35dB_ch2</i>	<i>45dB_40m_ch2, 50dB_40m_ch2</i>

Таблица 2. Выборки замираний амплитуды, которые не могут быть отнесены к теоретическим распределениям.

Класс выборки	Пик плотности в начале промежутка	Пик плотность в конце промежутка	Остальные
VBAU_T : 19			<i>40dB_80m</i>
VBAU_T : 65			<i>40dB_40grad</i>
VBAU_R : 10		<i>Соединенные выборки (30dB_ch2, 45dB_ch2)</i>	
VBAU_R : 13	<i>50dB_ch2_1</i>	<i>Соединенные выборки (45dB_ch2, 45dB_ch2_1, 55dB_ch2)</i>	<i>50dB_ch2_2</i>
VBAU_R : 15			<i>NAdB_ch2, 35dB_ch2</i>
VIU_T : 29		<i>55dB_20grad_ch2</i>	
VIU_T : 32			<i>50dB_30grad_ch2</i>
VIU_T : 35		<i>0db_ch2_2, 20dB_ch2</i>	