ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САМАРСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИМЕНИ АКАДЕМИКА С.П. КОРОЛЕВА»

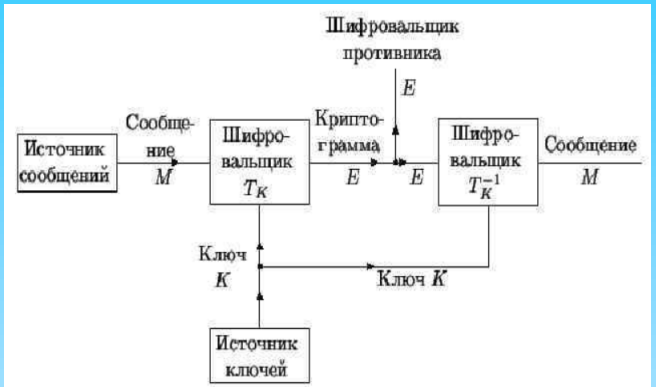
Автоматизированная система исследования качества генераторов случайных и псевдослучайных чисел

Обучающийся: Андрей Юрьевич Тышкун, гр. 6231-090401D

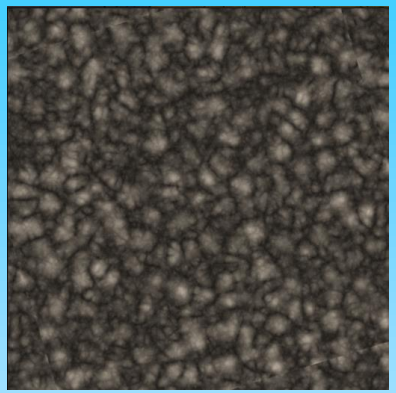
Руководитель: Константин Евгеньевич Климентьев, доцент кафедры ИСТ, к.т.н., доцент

Самара, 2023

Введение

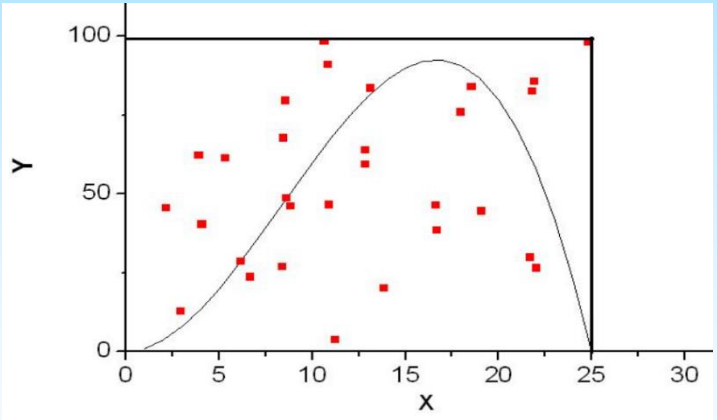


Генерация ключей для шифрования

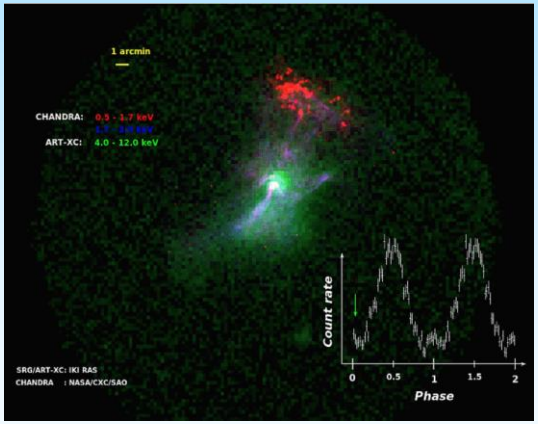


Генерация текстур

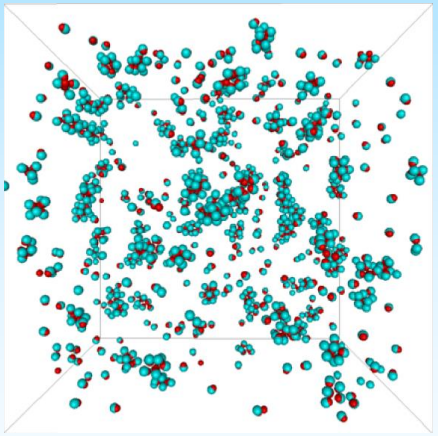
Проведение лотереи



Вычисление площади под графиком методом Монте-Карло



Исследование периодических сигналов с космического источника



Моделирование слипания частиц

Введение

Существуют различные типы генераторов случайных чисел:

- генераторы истинно случайных последовательностей (ГСП);
- генераторы псевдослучайных последовательностей (ГПСП);
- генераторы на основе комбинированного метода.

Критерии качества генераторов:

- простота и компактность;
- быстродействие;
- переносимость;
- возможность их распараллеливания для генерирования последовательностей.

Критерии качества генерируемых числовых последовательностей:

- стохастичность поведения;
- равномерность распределения;
- отсутствие скрытых зависимостей;
- длинный период.

Цели и задачи

Цель работы – создание программного обеспечения, обладающего возможностью проверки гипотезы о случайности последовательностей чисел, генерируемых различными генераторами, и выявление их отклонений от случайности, если такое присутствует, с помощью батареи тестов, включающей в себя различные статистические и графические тесты.

Задачи:

- Провести анализ предметной области
- Обзор систем-аналогов
- Сформулировать требования к системе оценки качества случайных чисел
- Выполнить проектирование системы
- Разработать систему
- Провести исследования генераторов

Описание предметной области

Наиболее распространённые виды генераторов ПСП:

- линейные конгруэнтные генераторы: $x_{n+1} = (ax_n + c)\%m$;
- мультипликативные линейные конгруэнтные генераторы: $x_{n+1} = (ax_n)\%m$;
- генераторы LFSR на основе регистров сдвига;
- генераторы Вихрь Мерсенна на основе матриц сдвиговых регистров: $x_{k+n} = x_{k+m} \oplus \left(x_k^u | x_{k+1}^l\right)A$;
- модифицированные и комбинированные генераторы, такие как KISS, Ran2 и RANECU.
- специальные типы генераторов, например, криптографически стойкие генераторы.

M	A
$2^{31}-1$ = 2147483647	16807
	48271
	69621
	39373
	742938285
	1754050460
$2^{31}-85$ =2147483563	40014
$2^{31}-249$ =2147483399	40692
$2^{32}-5$ = 4294967291	1223106847
$2^{61}-1$=2305843009213693951	1070922063159934167
$2^{64}-59$=18446744073709551557	2227057010910366687

Таблица параметров мультипликативного генератора

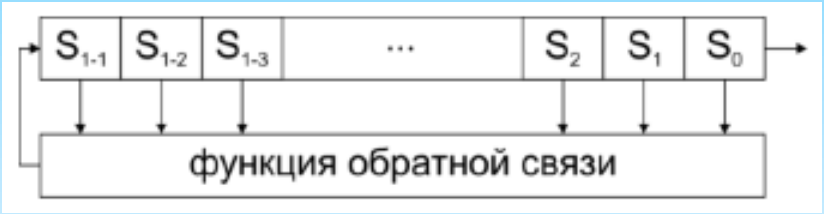


Схема генератора LFSR

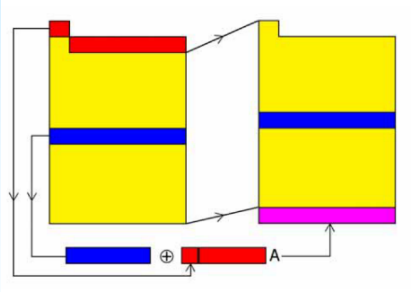


Схема генератора Вихрь Мерсенна

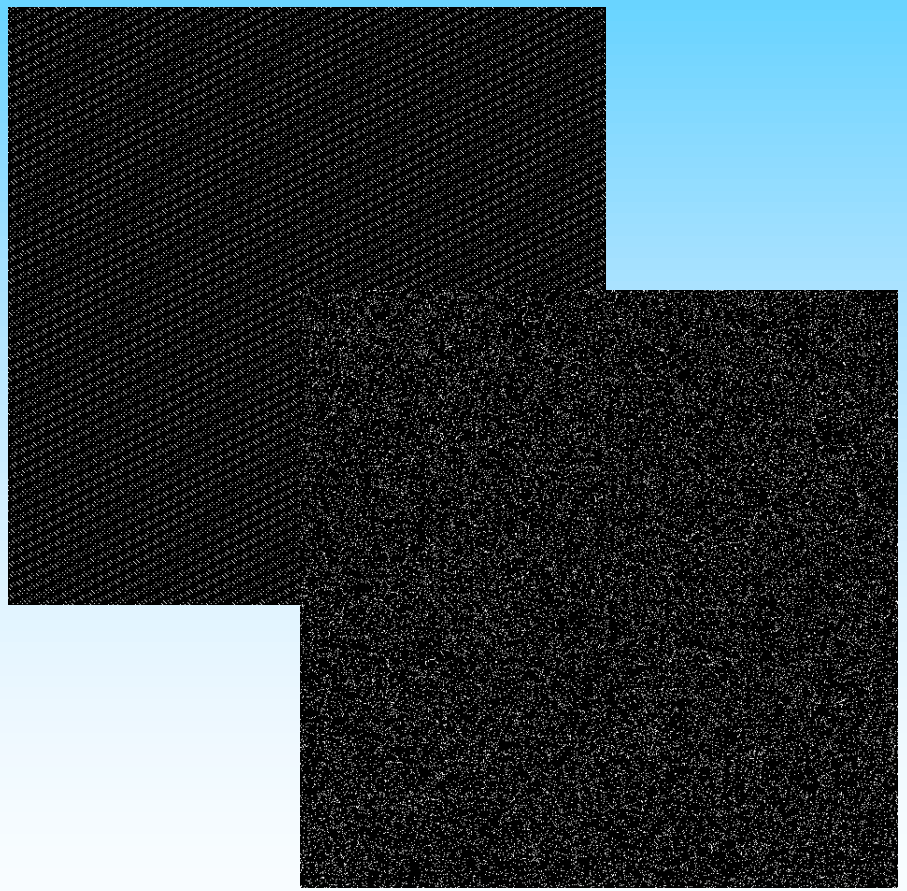
Источник	m	A	c
MS Visual Basic	2^{24}	1140671485	12820163
Glibc, ANSI C, Watcom C/C++	2^{32}	1103515245	12345
Numerical. Recipes , ГОСТ 28640–2012	2^{32}	1664525	1013904223
Borland C/C++	2^{32}	22695477	1
Borland Delphi	2^{32}	134775813	1
MS Visual C/C++	2^{32}	214013	2531011
Super-Duper от Дж. Марсальи	2^{32}	69069	любое нечетное
Java, GCC	2^{48}	25214903917	11
Д. Кнут	2^{64}	6364136223846793005	1442695040888963407
Super-Duper от Дж. Марсальи	2^{64}	6906969069	любое нечетное

Таблица параметров линейный конгруэнтных генераторов

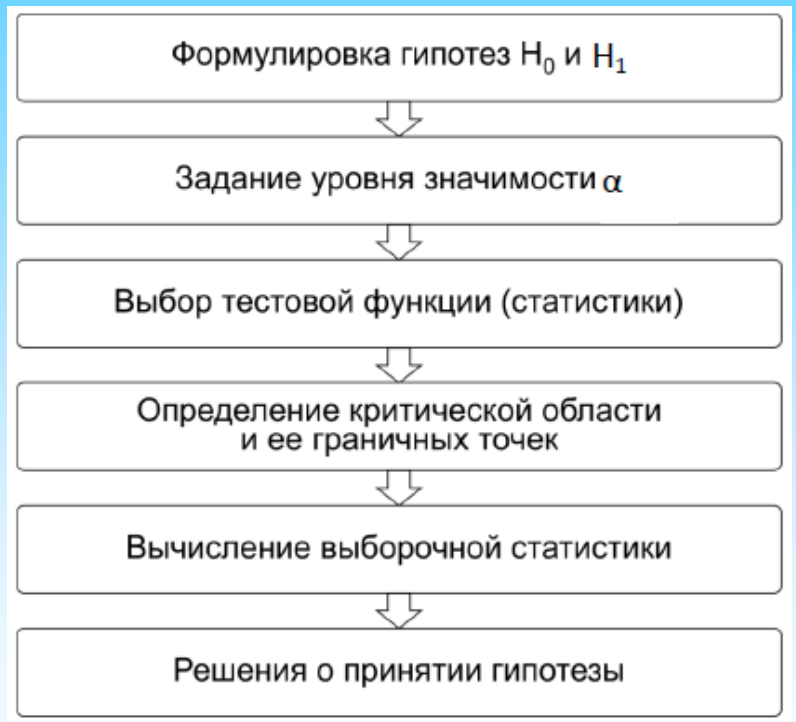
Описание предметной области

Для анализа уровня случайности последовательностей используются:

- Статистические тесты
- Графические тесты

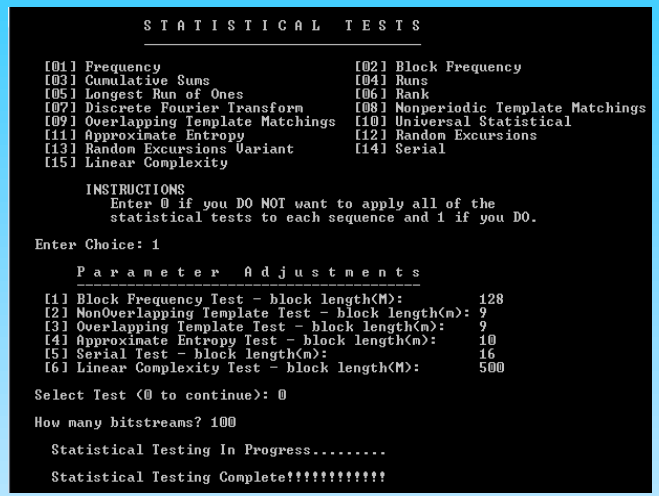


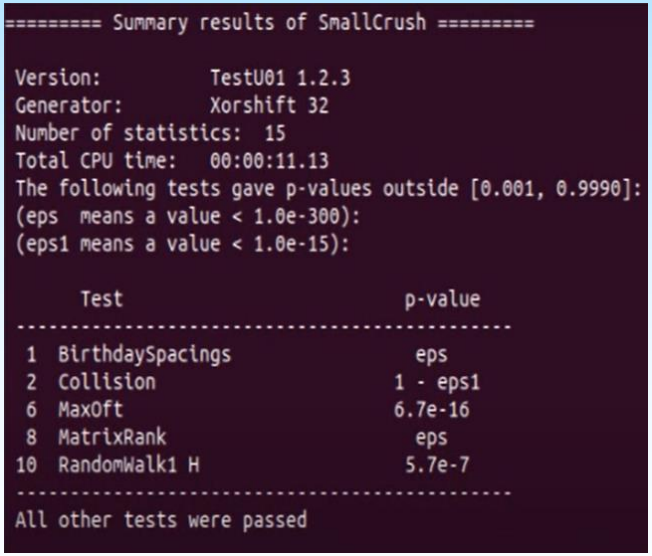
Результаты графического тестирования



Обобщенный алгоритм статистического тестирования

Системы-аналоги





	DIEHARD	TestU01	NIST SP 800-22
Качественные тесты	+	+	+
Встроенные генераторы	+	+	+
Битовые тесты	-	-	+
Скорость	-	-	+
Универсальность	-	-	-
Графические тесты	-	-	-
Переносимость	-	-	+
Интуитивно понятный интерфейс	-	-	-
Русский язык	-	-	+
База данных тестирования	-	-	-
Эмпирические критерии	+	+	+

	DIEHARD	TestU01	NIST SP 800-22
Теоретические критерии	-	-	-
Градация тестов по их силе	-	+	-
Определение области тестирования	-	-	+
Настройки параметров тестов/тестирования	-	+	-
Возможность получения интегральной оценки	-	+	+
Возможность тестирования нескольких последовательностей	-	+	+
Отчет по каждому тесту	+	+	+

Требования к системе оценки качества случайных чисел

Требования к системе:

- наличие встроенных генераторов псевдослучайных последовательностей;
- наличие эмпирических и теоретических критериев;
- наличие оценочных и графических тестов различного качества;
- определение области тестирования;
- настройки параметров тестов;
- настройки параметров тестирования;
- база данных тестирования;
- возможность получения интегральной оценки;
- возможность тестирования нескольких последовательностей;
- отчет по каждому тесту.

Требования к тестам:

- Тесты набора должны находить все возможные виды отклонений от случайности и находить скрытые зависимости.
- Тесты должны быть независимыми. Не следует использовать тестов больше, чем это необходимо.
- Тесты должны быть универсальными. Многие существующие тесты рассчитаны на определенный вид последовательностей.

Состав тестов

Название теста	Описание
Статистические тесты	
Тест серий	Проверяет неравномерность распределения m-битных слов.
Проверка гипотезы равномерного распределения случайной величины с помощью критерия хи-квадрат	Исследует последовательности на равномерность распределения в ней сгенерированных чисел.
Оценка математического ожидания каждой выборки случайных чисел	Позволяет эффективнее определить, насколько выборка соответствует равномерному распределению. Такой подход считается более надежным, поскольку он учитывает не только значения выборки, но и их отклонения от ожидаемых значений.
Проверка кумулятивных сумм	Вычисляется максимальное отклонение накопленной суммы элементов последовательности от начальной точки отсчета. Определяемый дефект – слишком много нулей или единиц в начале последовательности.
Частотный тест	Проверяет соотношение распределения нулей и единиц. Определяет слишком ли много нулей или единиц в последовательности
Проверка случайных отклонений	Определение отклонений от ожидаемого числа посещений различных состояний при произвольном обходе

Название теста	Описание
Статистические тесты	
Тест на равномерность битов	Определяет количество непрерывных серий одинаковых битов на всей длине последовательности. Определяемый дефект – колебание потока бит слишком быстрое или медленное.
Тест «блоков» в подпоследовательностях	В данном тесте определяется самая длинная серия единиц внутри блока длиной m бит. Показывает отклонение максимальных длин серий единиц от теоретического закона распределения.
Спектральный тест (Spectral test)	Предназначен для исследования периодических свойств последовательности битов на основе высот выбросов преобразования Фурье.
Проверка аппроксимированной энтропии	Определяет меру согласования наблюдаемого значения энтропии, исследуемой СП с ожидаемым значением. Исследует неравномерность распределения m-битных пересекающихся серий.
Проверка линейной сложности	Определяем минимальный регистр сдвига для генерации выделенной подпоследовательности. Исследует сложность данной последовательности, регистр не должен быть слишком коротким.

Состав тестов

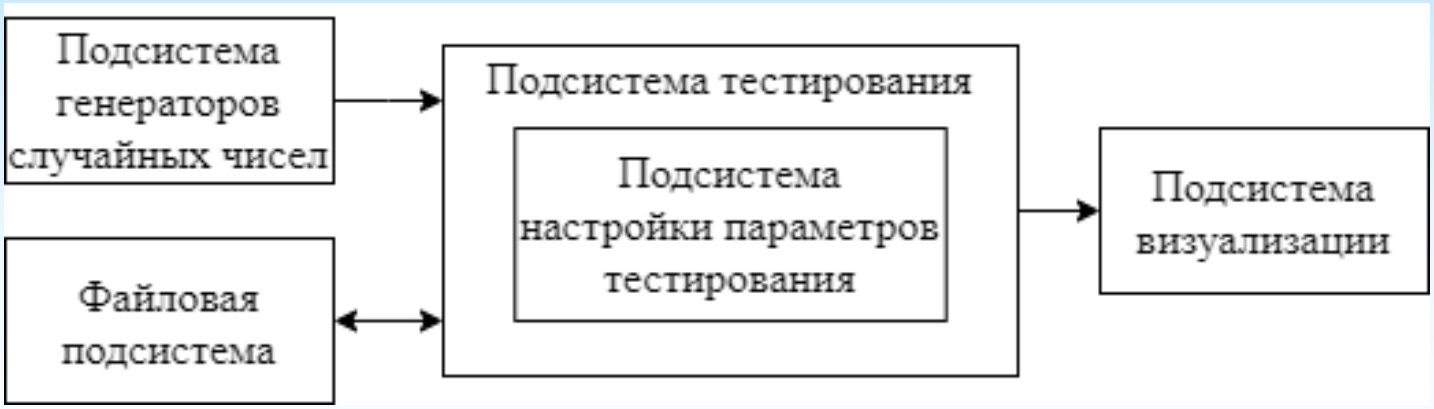
Название теста	Описание
Статистические тесты	
Проверка рангов матриц	Определяется ранг матрицы, тем самым исследуется линейная независимость подстрок фиксированной длины, составляющих первоначальную последовательность.
Частотный тест в подпоследовательностях	Тест определяет количество единиц внутри блока длиной M бит. Определяет действительно ли частота повторения единиц в блоке длиной M бит приблизительно равна M÷2.
Универсальный тест Маурера	Определяется число бит между одинаковыми шаблонами в исходной последовательности. Исследуется сжимаемость последовательности.
Проверка непересекающихся шаблонов	В данном тесте подсчитывается количество заранее определённых шаблонов, найденных в исходной последовательности. Исследует частоту встречи определенных шаблонов.
Покер тест	Проверяет равномерность распределения символов в исследуемой последовательности, анализируя различные комбинации чисел в подпоследовательностях.
Тест «стопка книг»	Проверяет равномерность распределения числе в исследуемой последовательности, на основе алгоритма «перемести на передний план» преобразования данных для сжатия.

Название теста	Описание
Графические тесты	
Гистограмма распределения элементов последовательности	Обеспечивает возможность оценить равномерность распределения чисел в последовательности и выявить частоту появления определенных чисел.
Распределение на плоскости (с фильтрующей процедурой и без)	Данный тест предназначен для определения зависимостей между элементами исследуемой последовательности. Если зависимость между элементами отсутствует, точки должны располагаться хаотично, иначе наблюдаются узоры.
Проверка серий	Основанный на анализе частоты появления нулей, единиц и серий, состоящих из k бит. Данный тест предназначен для оценки равномерности распределения символов в исследуемой последовательности.
Проверка на монотонность	Данный тест оценивает равномерность распределения символов в последовательности, используя анализ длины участков, на которых элементы последовательности не возрастают и не убывают.
Автокорреляционный тест	Основной целью данного теста является определение корреляции между сдвинутыми копиями последовательности с целью выявления возможной зависимости между подпоследовательностями, составляющими анализируемую последовательность.
Спектральный тест	Основная цель данного теста заключается в проверке равномерности распределения 0 и 1 в анализируемой последовательности. Для этого используется анализ высоты выбросов, полученных в результате преобразования Фурье.

Проектирование системы

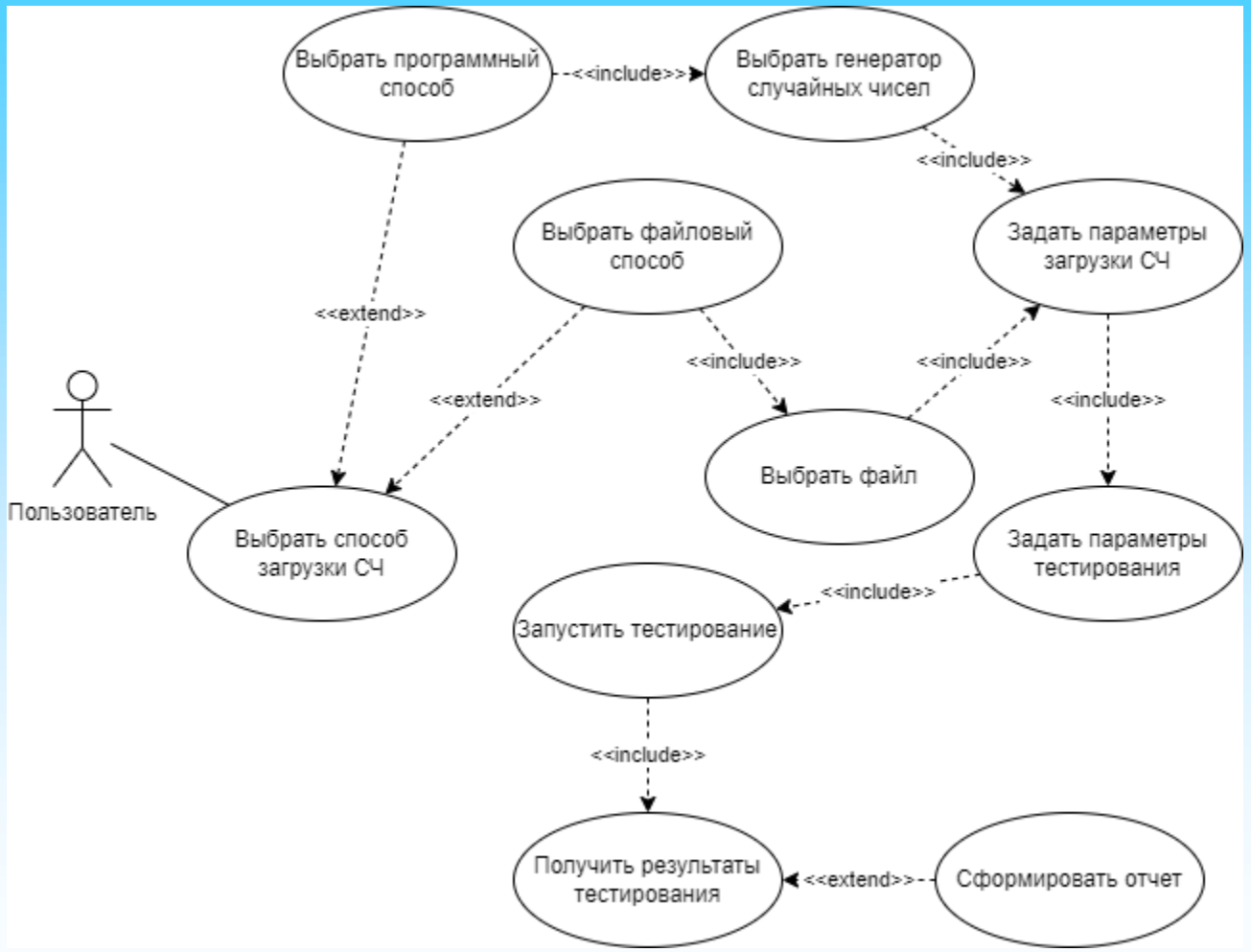
В состав структурной схемы входят следующие подсистемы:

1) Подсистема генераторов случайных чисел, которая отвечает за генерирование случайных чисел.
2) Файловая подсистема, которая позволяет загружать случайные числа в систему для тестирования в формате .txt или сохранять результаты тестирования.
3) Подсистема тестирования для тестирования случайных чисел, в её состав входит подсистема настройки параметров тестирования.
4) Подсистема визуализации, которая отвечает за графические отображения результатов тестирования.



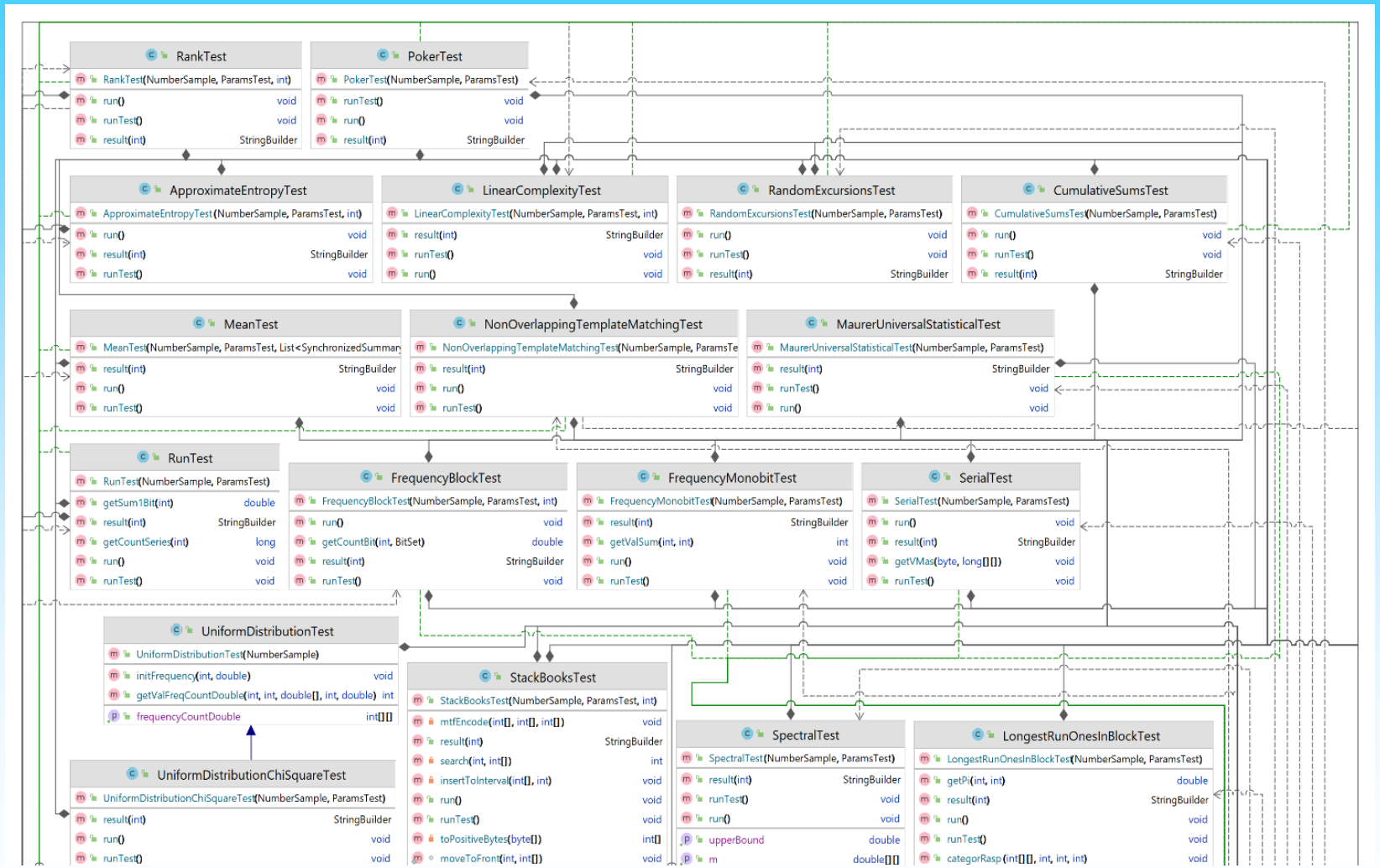
Проектирование системы

Диаграмма вариантов использования



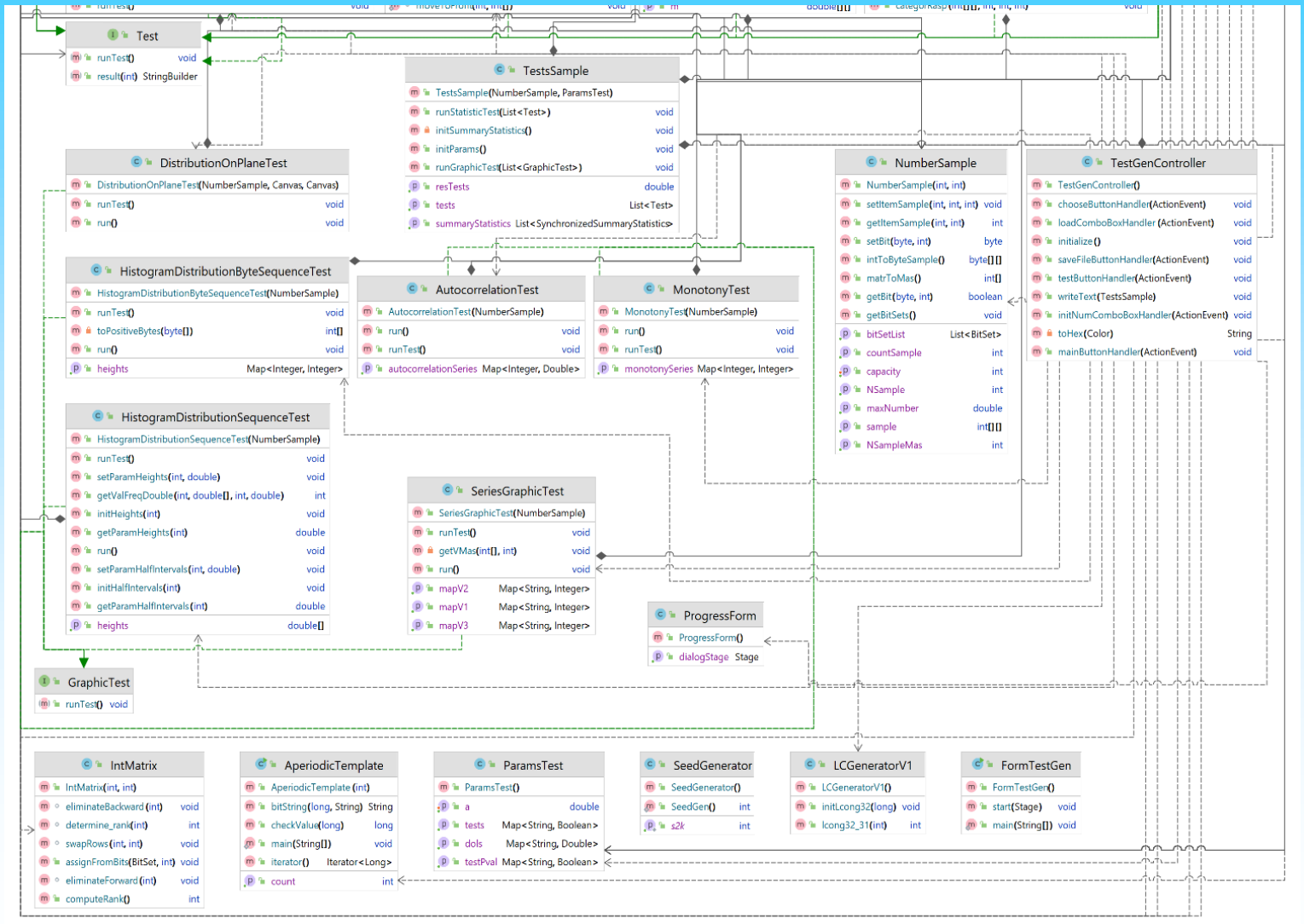
Проектирование системы

Диаграмма классов



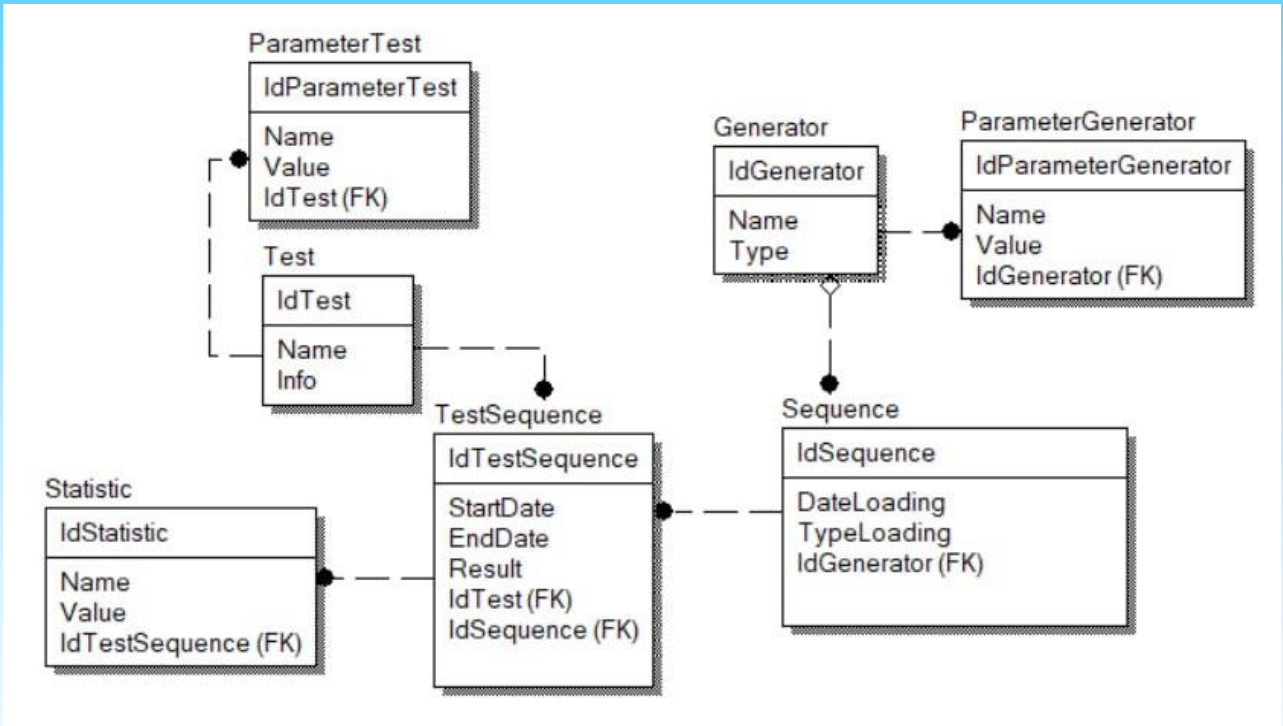
Проектирование системы

Диаграмма классов(продолжение)



Проектирование системы

Логическая модель базы данных

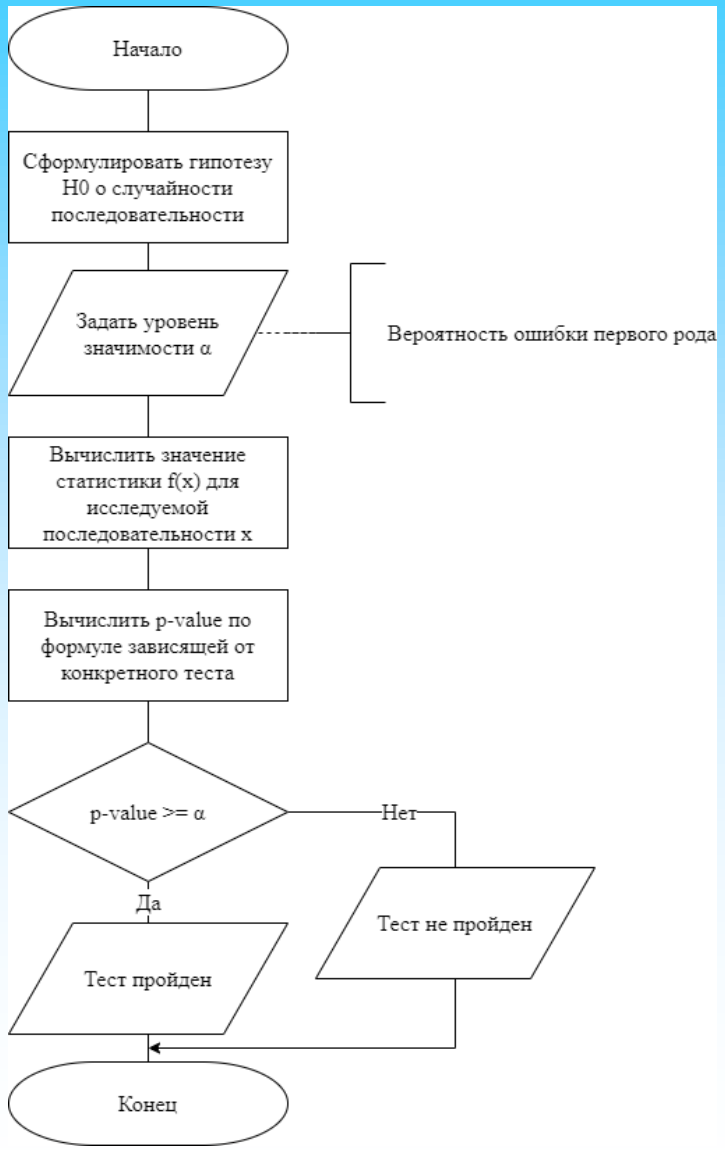


Проектирование системы

Алгоритм тестирования генераторов случайных чисел

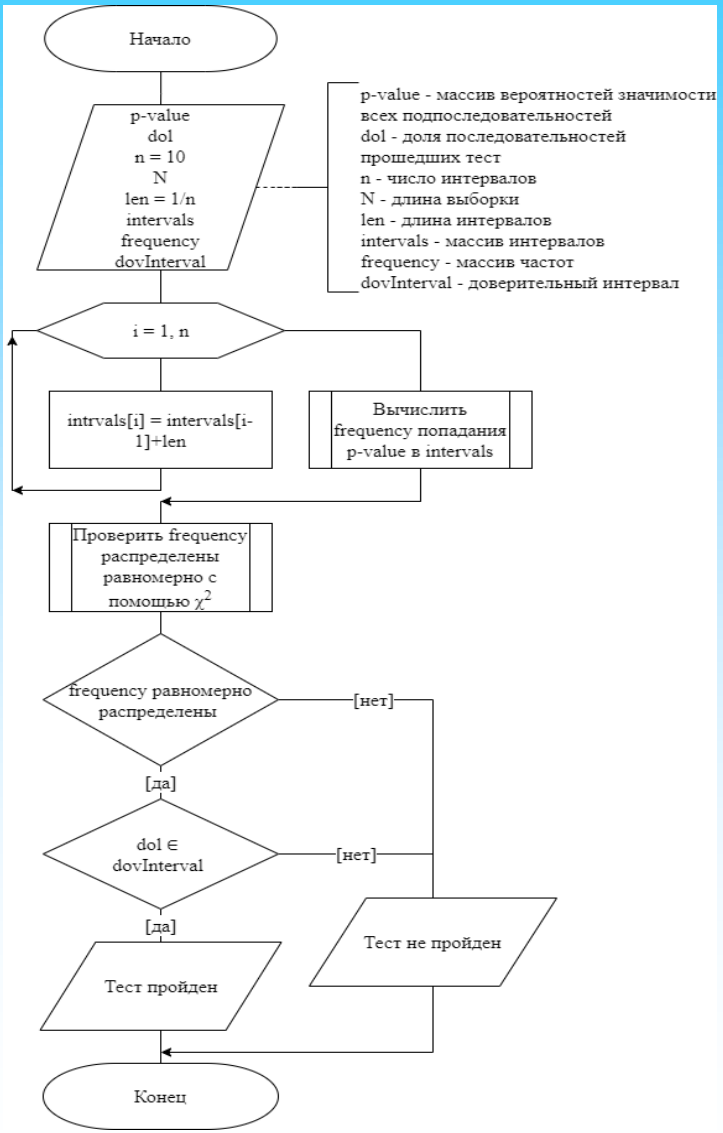
Проектирование системы

Алгоритм работы статистического теста для одной подпоследовательности



Проектирование системы

Алгоритм вывода о прохождении теста подпоследовательностями

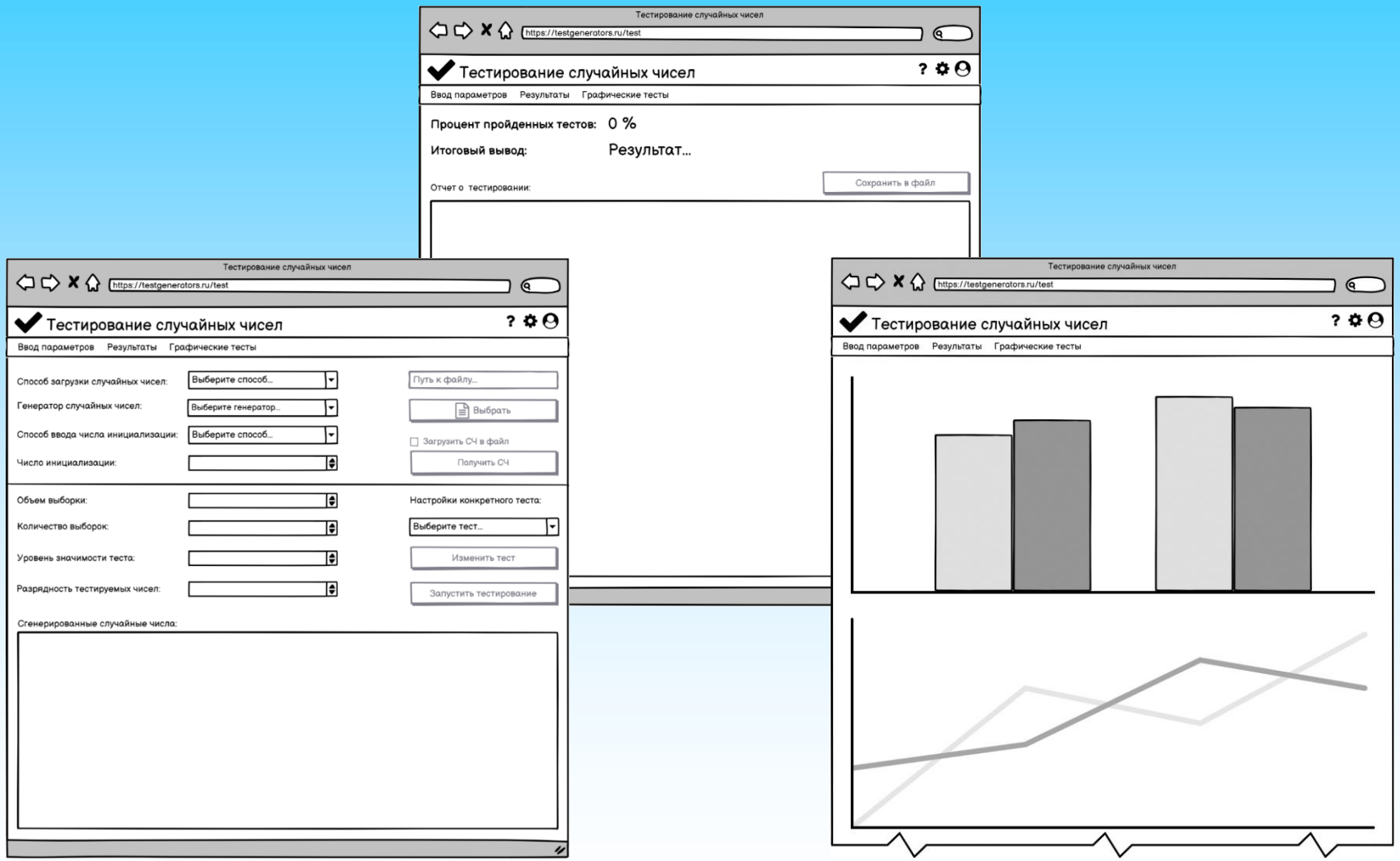


Критерии вычисляющие и использующие p-value для проверки соответствия эмпирических распределений теоретическим:

- критерий χ2 Пирсона;
- критерий Колмогорова-Смирнова.

Реализация системы

Прототип пользовательского интерфейса



Контрольный пример

Входные данные

- линейный конгруэнтный генератор хорошего качества:

$x_{n+1} = (ax_n + c)\%m$, где a = 1664525, c = 1, m = 2^{32}, которые представлены в ГОСТ 28640-2012 для генерирования хороших значений;

- SecureRandom генератор:

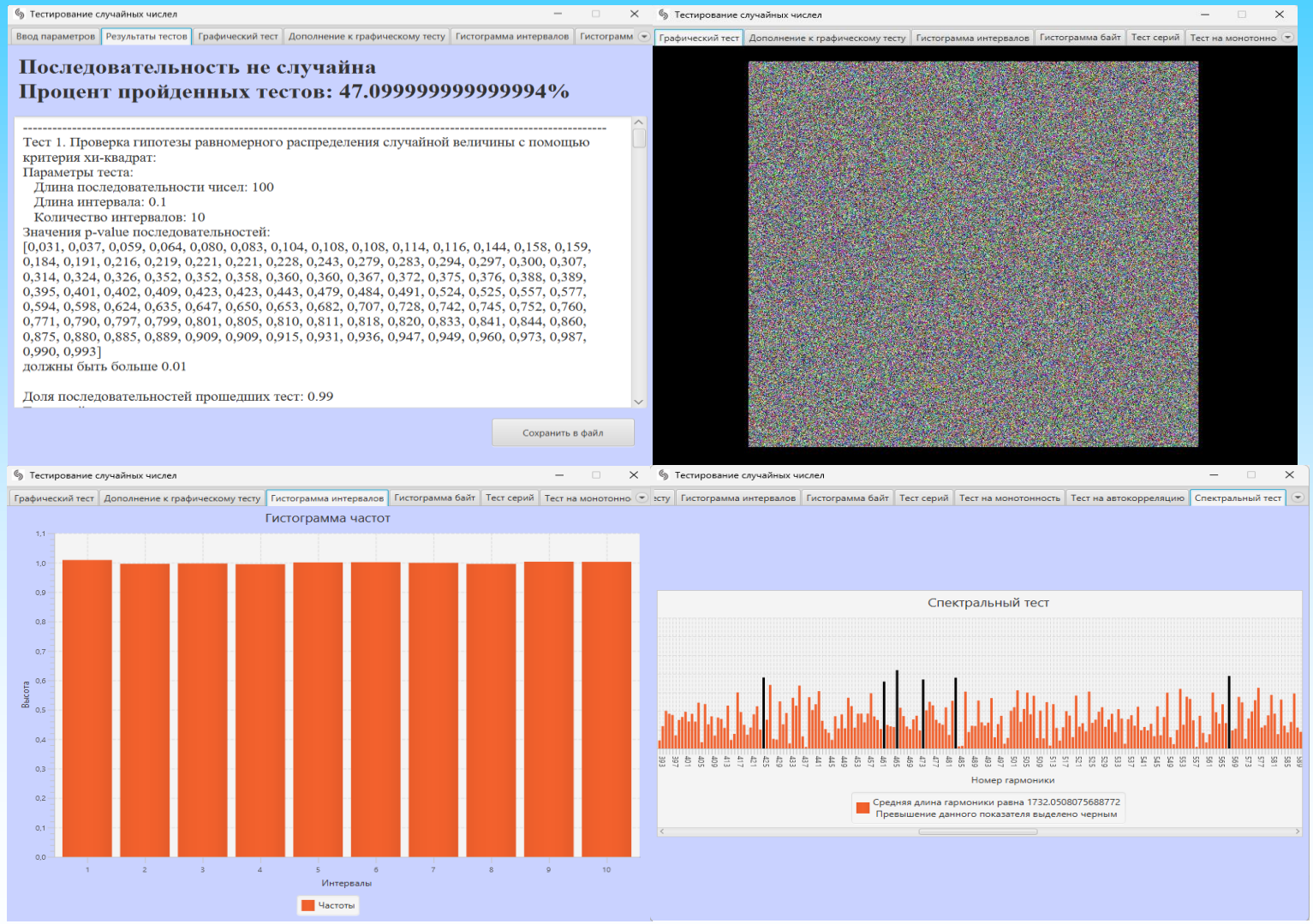
Широко используется для генерации криптографически сильных случайных последовательностей. Использует детерминированный алгоритм на основе хеш-функции для формирования псевдослучайной последовательности из истинно случайного seed.

Проверим следующие свойства взятых генераторов:

- свойства линейных конгруэнтных генераторов, у которых младшие биты генерируемых чисел, часто далеки от случайности;
- Качество SecureRandom генератора.

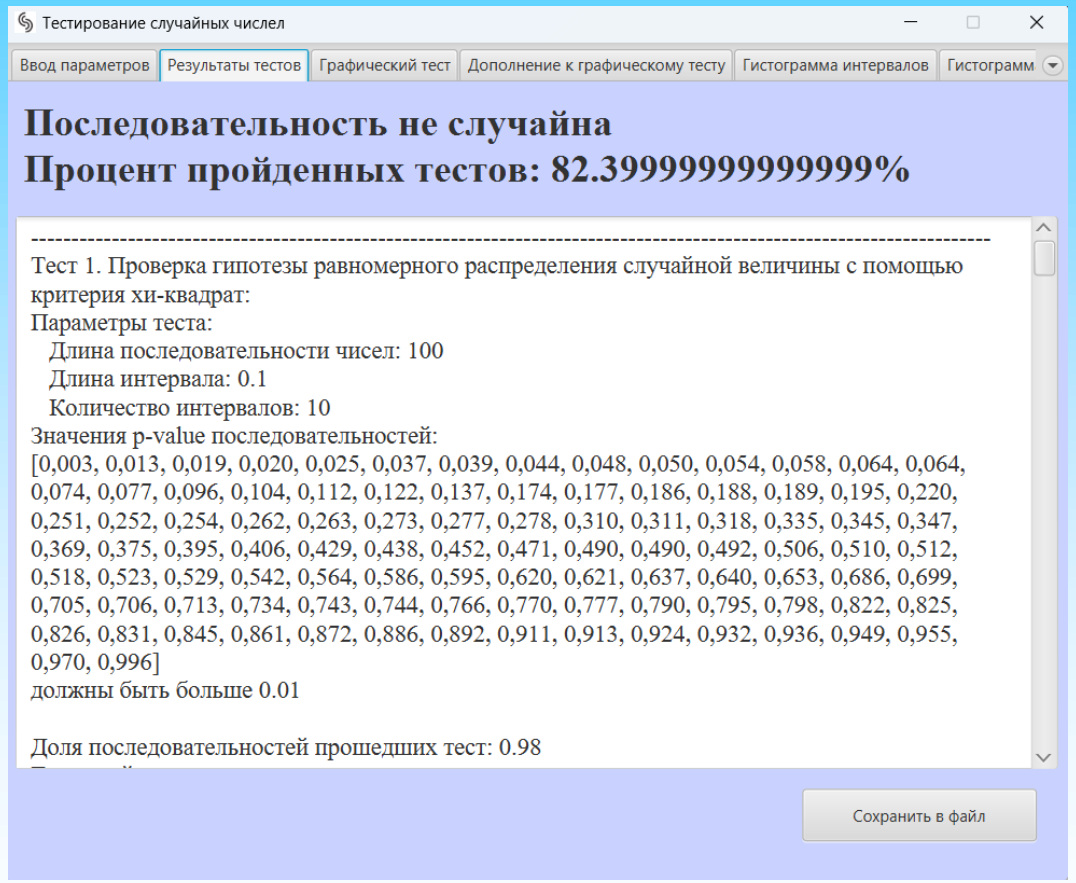
Контрольный пример

LCG генератор хорошего качества. Разрядность 31.



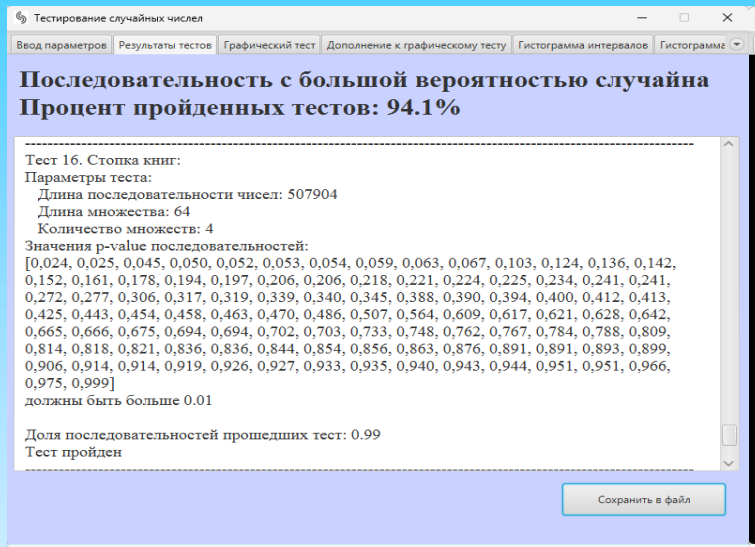
Контрольный пример

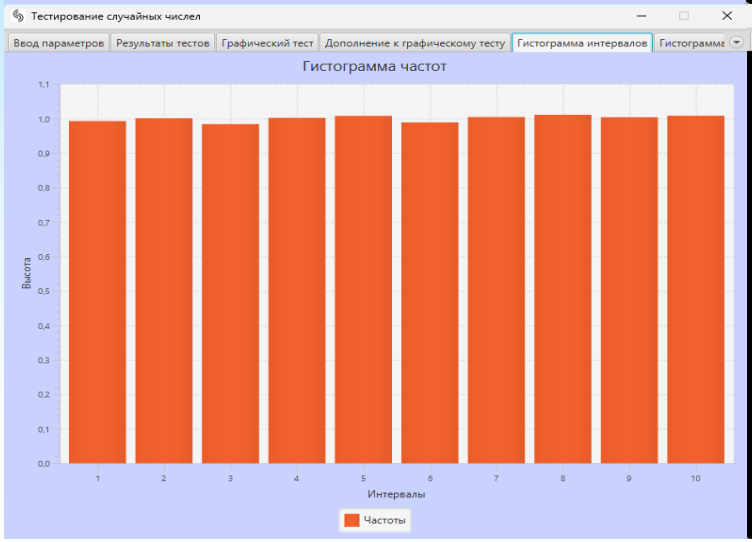
LCG генератор хорошего качества. Разрядность 16.

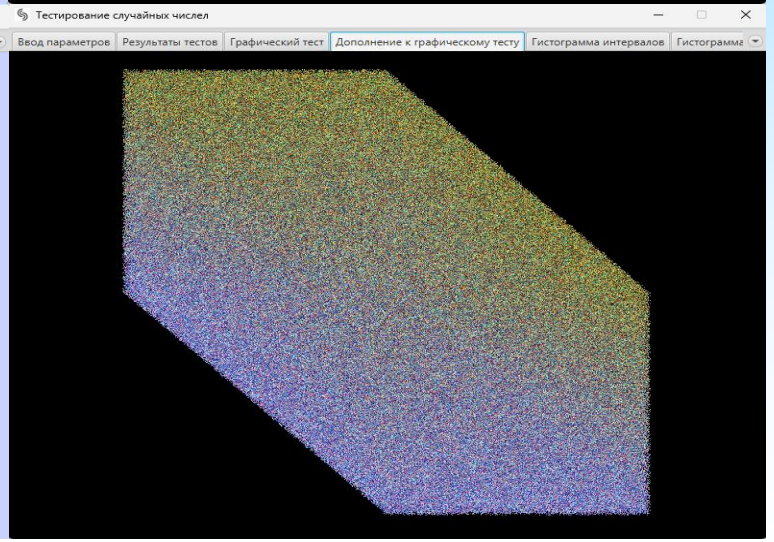


Контрольный пример

SecureRandom генератор. Разрядность 31.





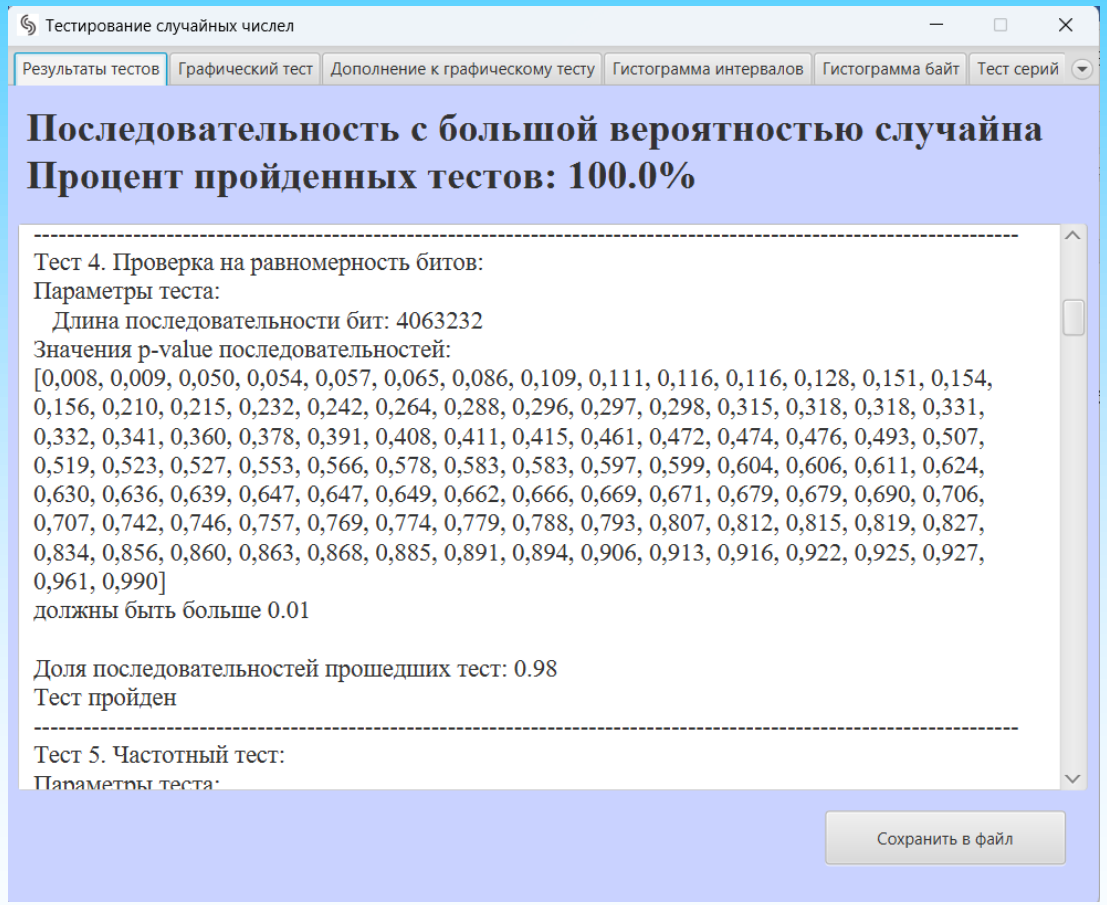


Контрольный пример

SecureRandom генератор. Разрядность 31.

Контрольный пример

SecureRandom генератор. Разрядность 16.



Результаты исследования генераторов

Генератор	Процент пройденных тестов %	Не пройденные тесты
ГОСТ Р ИСО 24153-2012	94,1	Проверка непересекающихся шаблонов
GFSR генератор с 3 параметрами (p = 1279, q = 418, w = 32)	88,2	Проверка линейной сложности, стопка книг
GFSR генератор с 5 параметрами (p = 521, q1 = 86, q2 = 197, q3 = 447, w = 32)	88,2	Проверка кумулятивных сумм, проверка случайных отклонений
Генератор Таусворта (на основе комбинации трех последовательностей параметров (31, 13, 12), (29, 2, 4) и (28, 3, 17))	82,4	Проверка гипотезы равномерного распределения случайной величины с помощью критерия хи-квадрат, проверка случайных отклонений, проверка непересекающихся шаблонов
Генератор Твистера (с параметрами (624, 397, 31, 32, 0x9908b0df, 11, 7, 15, 18, 0x9d2c5680, 0xefc60000))	100	-
LCD v1 (a = 1664525, c = 1, m = 2^{32})	94,1	Тест серий битов
LCD v2 (a = 2100005341, c = 15, m = 2^{32})	88,2	Тест серий битов, проверка непересекающихся шаблонов
KISS генератор (на основе двух GFSR и 1 мультипликативного генератора)	100	-
Java Random генератор	94,1	Проверка непересекающихся шаблонов
Java SecureRandom генератор	100	-
Java SplittableRandom генератор	88,2	Стопка книг, проверка непересекающихся шаблонов

Генератор	Процент пройденных тестов %	Не пройденные тесты
Мультипликативный генератор v1 (a = 0.011, c= π, m = 2^{32})	23,5	Проверка гипотезы равномерного распределения случайной величины с помощью критерия хи-квадрат, оценка математического ожидания каждой выборки случайных чисел, проверка кумулятивных сумм, проверка на равномерность битов, частотный тест, тест серий битов, спектральный тест, проверка аппроксимированной энтропии, частотный тест в подпоследовательностях, универсальный статистический тест Маурера, проверка непересекающихся шаблонов, стопка книг, покер-тест
Мультипликативный генератор v2 (a = 11, c= π, m = 2^{32})	35,3	Проверка гипотезы равномерного распределения случайной величины с помощью критерия хи-квадрат, оценка математического ожидания каждой выборки случайных чисел, проверка кумулятивных сумм, проверка на равномерность битов, частотный тест, тест серий битов, проверка аппроксимированной энтропии, проверка линейной сложности, частотный тест в подпоследовательностях, проверка непересекающихся шаблонов, стопка книг
Мультипликативный генератор v3 (a = 37, m = 2^{32})	41,2	Аналогично мультипликативному генератору v2, кроме теста «стопка книг»
Мультипликативный генератор v4 (a = 214013, c= 2531011, m = 232)	58,8	Проверка кумулятивных сумм, частотный тест, спектральный тест, проверка аппроксимированной энтропии, тест «блоков» в подпоследовательностях, проверка непересекающихся шаблонов, стопка книг
Мультипликативный генератор v5 (a = 16807, m = 2^{31})	94,1	Тест «блоков» в подпоследовательностях

Заключение

- Проведен анализ предметной области
- Выполнен обзор систем-аналогов
- Сформулировать требования к системе оценки качества случайных чисел
- Выполнено проектирование системы
- Выполнена разработка системы
- Проведены исследования генераторов