Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет
имени Н. Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н. Э. Баумана)

ФАКУЛЬТЕТ «Информатика и системы управления»

КАФЕДРА «Программное обеспечение ЭВМ и информационные технологии»

ОТЧЁТ

По лабораторной работе № 1

По курсу: «Моделирование»

Тема: «Генераторы случайных чисел»

Студент: Керимов А. Ш.

Группа: ИУ7-74Б

Оценка (баллы): ____________

Преподаватель: Рудаков И. В.

Москва

2020

Оглавление

1 Теоретическая часть

1.1 Вихрь Мерсенна

В качестве генератора псевдослучайных чисел был выбран вихрь Мерсенна, а именно вариант MT19937 с периодом $2^{19937} - 1$, поставляемый стандартной библиотекой языка C++.

Алгоритм основывается на свойствах простых чисел Мерсенна и обеспечивает быструю генерацию высококачественных по критерию случайности псевдослучайных чисел.

Вихрь Мерсенна лишён многих недостатков, присущих другим ГПСЧ, таких как малый период, предсказуемость, легко выявляемые статистические закономерности.

1.2 Табличный метод

Табличные генераторы в качестве источника случайных чисел используют специальным образом составленные таблицы, содержащие проверенные некоррелированные цифры.

Табличный генератор случайных чисел в лабораторной работе использует таблицу из книги «A Million Random Digits with 100,000 Normal Deviates».

1.3 Критерий равномерности

В качестве критерия равномерности был выбран критерий серий.

Пусть имеется последовательность целых чисел $\langle X_{2n} \rangle = X_0, X_1, \ldots, X_{2n-1}$, элементы которого, как предполагается, независимы и распределены между 0 и $d - 1$. Требование, предъявляемое к этой последовательности, состоит в следующем: пары последовательных чисел должны быть распределены независимо и равномерно.

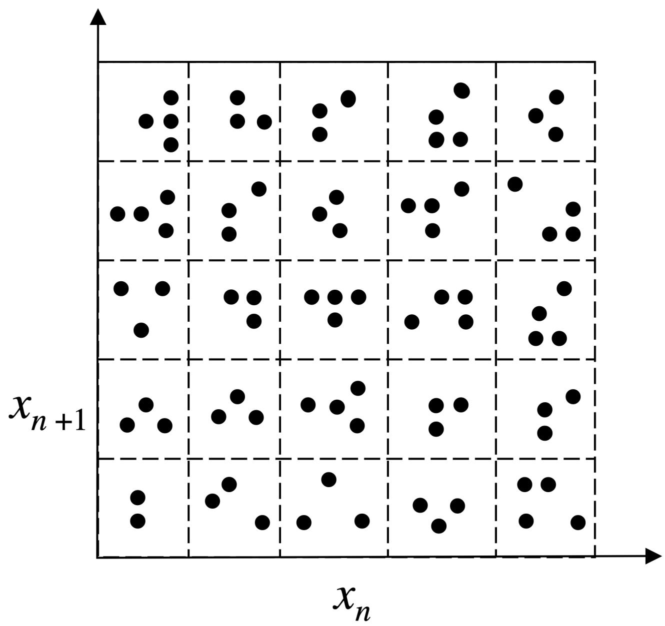


Рис. 1.1: Геометрический смысл критерия серий

Для каждой категории чисел (q, r), где $0 \leqslant q, r < d$, подсчитываем количество случаев, когда пара $(X_{2j}, X_{2j+1}) = (q, r)$ для $j = \overline{0, n-1}$. Затем применяем критерий хи-квадрат к этим $k = d^2$ категориям, с вероятностью $p = 1/d^2$ отнесения пары чисел к каждой из категорий.

В критерии хи-квадрат вычисляется статистика

$$\chi^2 = \sum_{i=0}^{k-1} \frac{(Y_i - np_i)^2}{np_i}, \tag{1.1}$$

где n — число независимых наблюдений, p_i — вероятность того, что наблюдение относится к категории i, Y_i — число наблюдений, которые относятся к категории i, при этом:

$$\sum_{i=0}^{k-1} Y_i = n, \quad \sum_{i=0}^{k-1} p_i = 1. \tag{1.2}$$

С учётом также $p_0 = p_1 = \ldots = p_{k-1} = p$, формулу (1.1) можно преобразовать:

$$\chi^2 = \frac{p}{n} \sum_{i=0}^{k-1} Y_i^2 - n. \tag{1.3}$$

Число степеней свободы ν статистики хи-квадрат на единицу меньше числа категорий k.

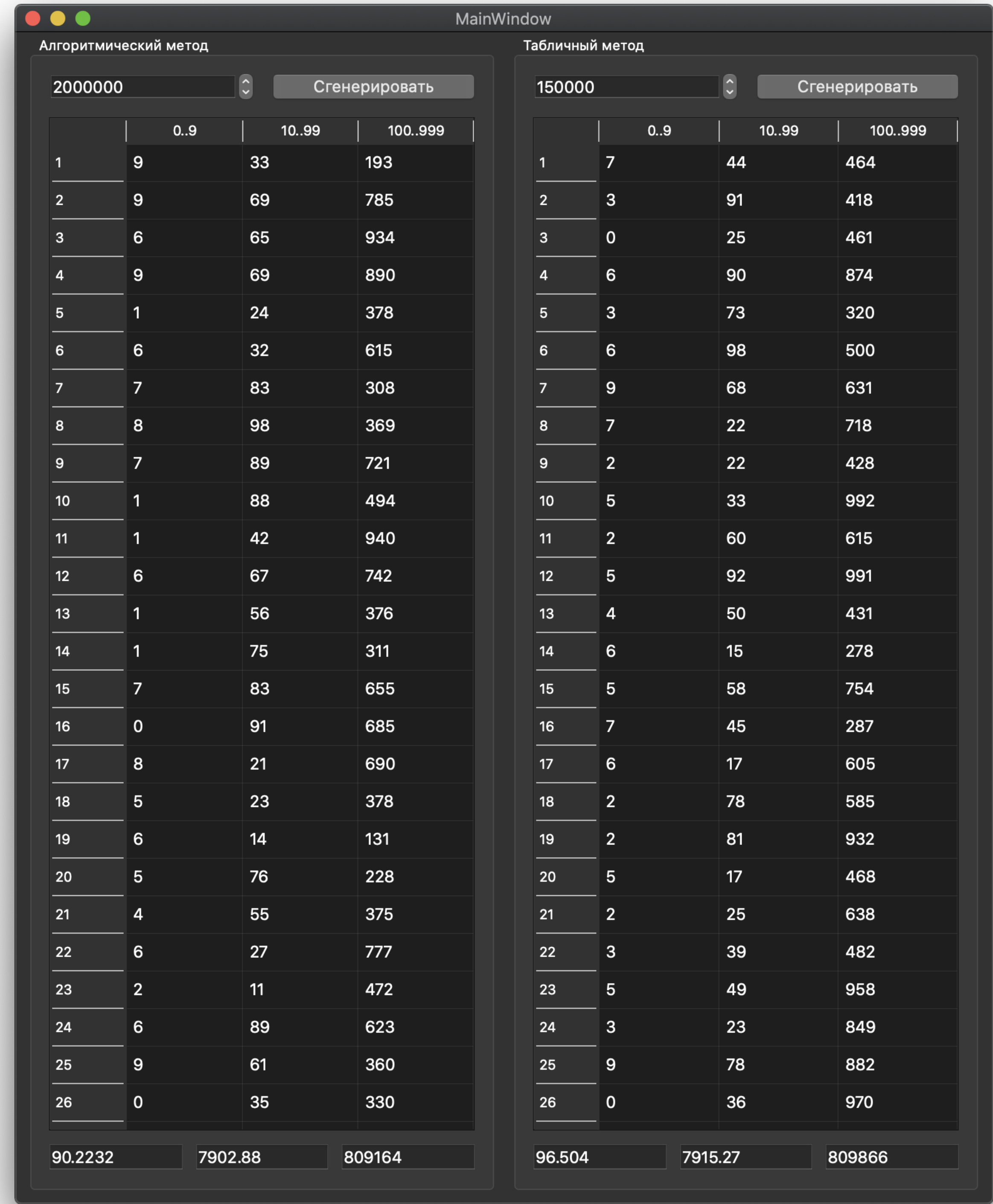
Далее значение статистики сравнивается с приемлемым (табличным). Если значение χ^2 много больше или много меньше табличного, то гипотеза о равномерности случайной величины не выполняется, т. к. разброс чисел слишком велик или мал соответственно. Если значение χ^2 лежит между теоретическими значениями двух рядом стоящих столбцов — гипотеза о равномерности случайной величины выполняется с вероятностью p, которая, в идеале, должна стремиться к 50 %, но хорошим считается результат 5 — 95 %.

Таблица 1.1: Некоторые процентные точки χ^2-распределения

	$\nu = 99$	$\nu = 8099$	$\nu = 809999$
$p = 0{,}999$	61,137	7711,395	806071,478
$p = 0{,}995$	66,510	7774,928	806724,263
$p = 0{,}99$	69,230	7805,867	807040,986
$p = 0{,}975$	73,361	7851,452	807506,270
$p = 0{,}95$	77,046	7890,800	807906,582
$p = 0{,}75$	89,181	8012,797	809140,152
$p = 0{,}5$	98,334	8098,333	809998,333
$p = 0{,}25$	108,093	8184,476	810857,121
$p = 0{,}05$	123,225	8309,474	812093,692
$p = 0{,}025$	128,422	8350,336	812495,519
$p = 0{,}01$	134,642	8398,015	812962,897
$p = 0{,}005$	138,987	8430,585	813281,250
$p = 0{,}001$	148,230	8498,004	813937,922

2 Результат работы

Для исследования генераторов псевдослучайных чисел алгоритмическим методом было сгенерировано по 2 000 000 одно-, двух- и трёх-разрядных чисел, табличным методом — по 150 000 чисел.



Алгоритмический метод:

	0..9	10..99	100..999
1	9	33	193
2	9	69	785
3	6	65	934
4	9	69	890
5	1	24	378
6	6	32	615
7	7	83	308
8	8	98	369
9	7	89	721
10	1	88	494
11	1	42	940
12	6	67	742
13	1	56	376
14	1	75	311
15	7	83	655
16	0	91	685
17	8	21	690
18	5	23	378
19	6	14	131
20	5	76	228
21	4	55	375
22	6	27	777
23	2	11	472
24	6	89	623
25	9	61	360
26	0	35	330
	90.2232	7902.88	809164

Табличный метод:

	0..9	10..99	100..999
1	7	44	464
2	3	91	418
3	0	25	461
4	6	90	874
5	3	73	320
6	6	98	500
7	9	68	631
8	7	22	718
9	2	22	428
10	5	33	992
11	2	60	615
12	5	92	991
13	4	50	431
14	6	15	278
15	5	58	754
16	7	45	287
17	6	17	605
18	2	78	585
19	2	81	932
20	5	17	468
21	2	25	638
22	3	39	482
23	5	49	958
24	3	23	849
25	9	78	882
26	0	36	970
	96.504	7915.27	809866

Рис. 2.1: Результат работы программы

Вывод

Рассмотрены алгоритмический (MT19937) и табличный генераторы псевдослучайных чисел. Оба метода удовлетворяют критерию серий — получают оценку 5 — 95 %, а именно:

Таблица 2.1: Оценка ГПСЧ

	0..9	**10..99**	**100..999**
Алгоритмический метод	$p = 72{,}422$ %	$p = 93{,}923$ %	$p = 74{,}400$ %
Табличный метод	$p = 55{,}227$ %	$p = 92{,}639$ %	$p = 54{,}141$ %

Достоинством данных методов является быстродействие и генерация качественных псевдослучайных чисел. Хороший результат табличного ГПСЧ обусловлен тем, что в таблице содержатся проверенные некоррелированные цифры. В то же время, для хранения большого количества цифр, от которого зависит период генератора, требуется много памяти.

Алгоритм MT19937 этих недостатков лишён, его период составляет практически недостижимые $2^{19937} - 1$, однако по качеству генерируемых чисел немного уступает табличному методу.