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»

Дисциплина: «Моделирование»

Лабораторная работа №1

Тема работы: «Исследование псевдослучайных чисел»

Студент: Левушкин И. К.
Группа: ИУ7-72Б
Преподаватель: Рудаков И. В.

Москва, 2020 г.

Задание

Реализовать критерий оценки случайности последовательности.

Сравнить результаты работы данного критерия на одноразрядных, двухразрядных и трехразрядных последовательностях псевдослучайных целых чисел.

Последовательности получать алгоритмическим способом и табличным способом.

Формализация

Реализация последовательностей случайных последовательностей

В данной лабораторной работе генерация псевдослучайных чисел осуществляется с помощью вихря Мерсенна [1], табличные значения берутся из трёх файла.

Критерий оценки случайности последовательности

В качестве критерия проверки случайности последовательности был выбран критерий равномерности (критерий частот), использующий χ^2-критерий [2].

Кратко описать данный критерий можно следующим образом:

- Пусть k – количество всех возможных принимаемых значений,
- p_s - вероятность того, что наблюдение относится к категории s,
- n - количество проведенных экспериментов,
- Y_s - число экспериментов относящихся к категории s.

Поскольку все возможные принимаемые значения должны быть равновероятны, то $p_s = \frac{1}{k}$

Тогда получим статистику V, также известную как расстояние Пирсона D:

$$V = \frac{k}{n} \sum_{s=1}^{k} \left(Y_s^2 \right) - n$$

Если V равно нулю, то распределение абсолютно равномерно, то есть все возможные принимаемые значения входят в анализируемую последовательность по одному разу.

Для определения «приемлемого» значения V используются значения из χ^2-распределения с $\nu = k - 1$ степенями свободы:

	p = 1%	p = 5%	p = 25%	p = 50%	p = 75%	p = 95%	p = 99%
$\nu = 9$	2.088	3.325	5.899	8.343	11.39	16.92	21.67
$\nu = 89$	60.93	68.25	79.68	88.33	97.60	112.02	122.94
$\nu = 899$	803.31	830.41	870.05	898.33	927.23	969.86	1000.57

Таблица 1

Если в таблице выбрать число x, стоящее на ν-ой строке и в столбце p, то это означает, что вероятность того, что значение V будет меньше либо равно x, равно x, приближенно равна p, если n достаточно велико.

Чтобы определить насколько большим должно быть n, воспользуемся эмпирическим правилом:

Необходимо взять n настолько большим, чтобы все значения величин np_s *были больше или равны пяти.*

В нашем случае p_s равно:

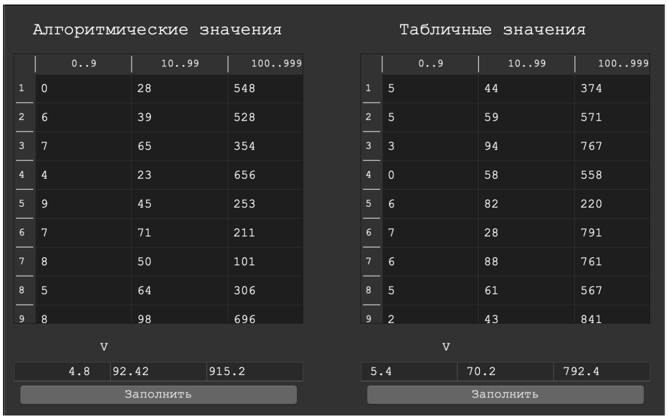
- $\frac{1}{10} = 0.1$ при $k = 10$
- $\frac{1}{90} = 0.0(1)$ при $k = 90$
- $\frac{1}{900} = 0.00(1)$ при $k = 900$

В соответствии с сформулированным эмпирическим правилом необходимо провести $n >= 50$, $n >= 450$, $n >= 4500$ наблюдений для одноразрядных, двухразрядных и трехразрядных чисел, соответственно.

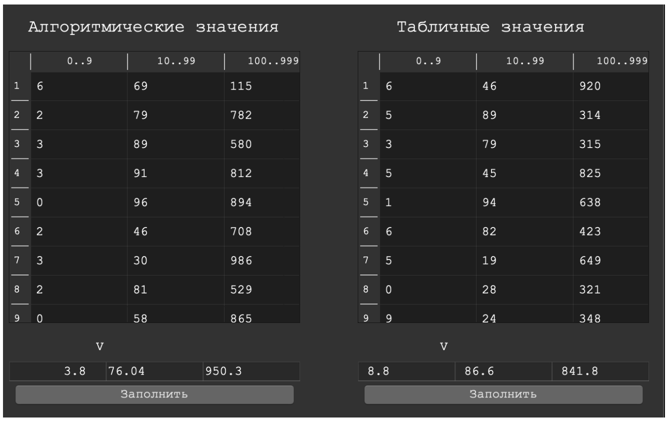
В данной лабораторной работе используется $n = 100$, $n = 1000$, $n = 10000$ наблюдений для одноразрядных, двухразрядных и трехразрядных чисел, соответственно.

Результаты работы

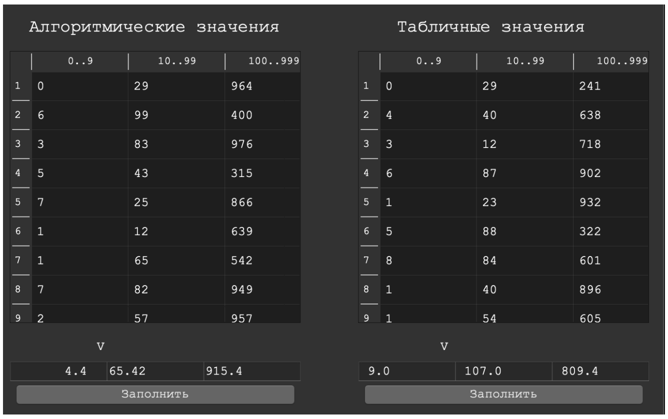
Ниже приведены результаты работы программы. В таблицах представлены первые 10 полученных чисел.



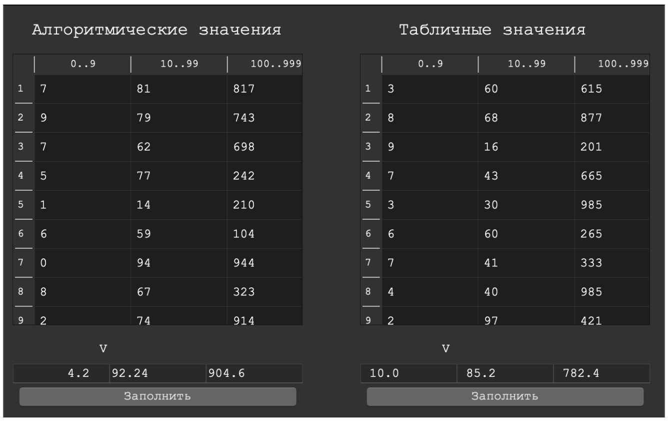
Пример №1 работы программы



Пример №2 работы программы



Пример №3 работы программы



Пример №4 работы программы

Вывод

Исходя из приведенных результатов работы программы следует, что:

№	V (алг.) 0..9	V (алг.) 10..99	V (алг.) 100..999	V (табл.) 0..9	V (табл.) 10..99	V (табл.) 100..999
1	4.8	92.42	915.2	5.4	70.2	792.4
2	3.8	76.04	950.3	8.8	86.6	793.5
3	4.4	97.1	915.4	9.0	107.0	809.4
4	4.2	92.24	904.6	10.0	85.2	782.4

Таблица 2

Сравнивая эти значения с значениями V из таблицы 1, можно прийти к выводам, что вероятность p при котором величина V принимает указанные значения:

№	p (алг.) 0..9	p (алг.) 10..99	p (алг.) 100..999	p (табл.) 0..9	p (табл.) 10..99	p (табл.) 100..999
1	<= 25%	<= 75%	<= 75%	<= 25%	<= 25%	<= 1%
2	<= 25%	<= 50%	<= 95%	<= 75%	<= 50%	<= 1%
3	<= 25%	<= 75%	<= 75%	<= 75%	<= 95%	<= 5%
4	<= 25%	<= 75%	<= 75%	<= 75%	<= 50%	<= 1%

Таблица 3

Видно, что при получении трехразрядных последовательностей с помощью таблицы случайных чисел при большом n (10000) V принимает такие малые значения, что вероятность p стремится к 0. Другими словами, наблюдаемые значения очень близки к ожидаемым, что неудивительно, поскольку таблица случайных чисел по определению подразумевает, что это набор цифр такой, что вероятность возникновения любой цифры от 0 до 9 одна и та же ($p_s = \frac{1}{k}$).

Также, заметим, что значения, сгенерированные алгоритмически, как и ожидалось, являются удовлетворительно случайными по отношению к χ^2-критерию (в среднем 25% $<= p <=$ 75%). Другими словами, V достаточно большое, чтобы не считать результаты искусственными как в случае с табличным заполнением, но в то же время достаточно маленькое, чтобы считать результаты равномерно распределенными.

Таким образом, полученные результаты полностью соответствуют ожидаемым.

Список литературы

[1] Mersenne Twister: A 623-Dimensionally Equidistributed Uniform Pseudo-Random Number Generator [Электронный ресурс]. – Режим доступа: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.315.6296&rep=rep1&type=
свободный – (10.10.2020)

[2] Случайные числа [Электронный ресурс]. – Режим доступа: https://vk.com/doc10903696_194274866?hash=449e2b0813b76d81ce&dl=db99c067f1a
свободный – (10.10.2020)