



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»

Лабораторная работа № 3

Тема Исследование псевдослучайных чисел

Студент Сукочева А.

Группа ИУ7-736

Оценка (баллы) _____

Преподаватель Рудаков И.В.

Москва.
2021 г

Задание на лабораторную работу

Изучить методы генерирования псевдослучайных чисел, а также критерии оценки случайности последовательности. Реализовать критерий оценки случайной последовательности. Сравнить результаты работы данного критерия на одноразрядных, двухразрядных и трехразрядных последовательностях целых чисел. Последовательности получать алгоритмическим и табличным способами.

Теоретическая часть

Для выполнения работы был выбран критерий «хи-квадрат». Это один из самых известных статистических критериев, также это основной метод, используемый в сочетании с другими критериями.

С помощью этого критерия можно узнать, удовлетворяет ли генератор случайных чисел требованию равномерного распределения или нет.

Для оценки по этому критерию необходимо вычислить статистику V по формуле:

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

где n – количество независимых испытаний, k – количество категорий, Y_s — число наблюдений, которые действительно относятся к категории S , p_s — вероятность того, что каждое наблюдение относится к категории s .

Значение V является значением критерия «хи-квадрат» для экспериментальных данных. Приемлемое значение этого критерия можно определить по таблице 1. Для этого используем строку с $v = k-1$, где $k = 10, 90, 900$ для задания лабораторной. P в этой таблице — это вероятность того, что экспериментальное значение $V_{\text{эксп.}}$ будет меньше табулированного (теоретического) $V_{\text{теор.}}$ или равно ему. Ее также можно рассматривать как доверительную вероятность.

Если вычисленное V окажется меньше 1%-й точки или больше 99%-й точки, можно сделать вывод, что эти числа недостаточно случайные. Если V лежит между 1% и 5% точками или между 95% и 99% точками, то эти числа «подозрительны». Если V лежит между 5% и 10% точками или 90%-95% точками, то числа можно считать «почти подозрительными». Обычно необходимо произвести проверку три раза и более с разными данными. Если по крайней мере два из трех результатов оказываются подозрительными, то числа рассматриваются как недостаточно случайные.

	$p = 1\%$	$p = 5\%$	$p = 25\%$	$p = 50\%$	$p = 75\%$	$p = 95\%$	$p = 99\%$
$\nu = 1$	0.00016	0.00393	0.1015	0.4549	1.323	3.841	6.635
$\nu = 2$	0.02010	0.1026	0.5754	1.386	2.773	5.991	9.210
$\nu = 3$	0.1148	0.3518	1.213	2.366	4.108	7.815	11.34
$\nu = 4$	0.2971	0.7107	1.923	3.357	5.385	9.488	13.28
$\nu = 5$	0.5543	1.1455	2.675	4.351	6.626	11.07	15.09
$\nu = 6$	0.8721	1.635	3.455	5.348	7.841	12.59	16.81
$\nu = 7$	1.239	2.167	4.255	6.346	9.037	14.07	18.48
$\nu = 8$	1.646	2.733	5.071	7.344	10.22	15.51	20.09
$\nu = 9$	2.088	3.325	5.899	8.343	11.39	16.92	21.67
$\nu = 10$	2.558	3.940	6.737	9.342	12.55	18.31	23.21
$\nu = 11$	3.053	4.575	7.584	10.34	13.70	19.68	24.72
$\nu = 12$	3.571	5.226	8.438	11.34	14.85	21.03	26.22
$\nu = 15$	5.229	7.261	11.04	14.34	18.25	25.00	30.58
$\nu = 20$	8.260	10.85	15.45	19.34	23.83	31.41	37.57
$\nu = 30$	14.95	18.49	24.48	29.34	34.80	43.77	50.89
$\nu = 50$	29.71	34.76	42.94	49.33	56.33	67.50	76.15
$\nu > 30$	$\nu + \sqrt{2\nu}x_p + \frac{2}{3}x_p^2 - \frac{2}{3} + O(1/\sqrt{\nu})$						
$x_p =$	-2.33	-1.64	-.674	0.00	0.674	1.64	2.33

Таблица 1.

Некоторые процентные точки χ^2 -распределения.
(Источник: Кнут Д. Э. «Искусство программирования»)

Таким образом, процедура проверки следующая:

1. Выделяем k категорий. В нашем случае это количество возможных полученных значений: 10, 90 и 900 для одноразрядных, двухразрядных и трехразрядных.
2. Запускаем генератор случайных чисел N раз.
3. Определяем количество случайных чисел, попавших в каждую категорию.
4. Вычисляем значение V по формуле (1).
5. Сравниваем полученное значение с теоретическими значениями в таблице, определяем к какому интервалу оно относится.
6. Делаем выводы о случайности величины, возможны три случая:
 1. Если $V_{\text{эксп}}$ лежит между 1% и 99% точками, то генератор удовлетворителен. (Однако необходимо учитывать «подозрительные результаты», о которых написано выше)
 2. Если $V_{\text{эксп}}$ меньше 1% точки, то генератор не удовлетворителен, так как разброс чисел слишком мал, чтобы быть случайным.

3. Если $V_{\text{эксп}}$ больше 99% точки, то генератор не удовлетворителен, так как разброс чисел слишком велик, чтобы быть случайным.

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

Таблица 2. Значения $V_{\text{теор}}$ для количества степеней свободы по заданию.

По таблице 2 можно будет сделать выводы о полученных в программе значениях.

Результаты работы программы

В качестве алгоритмического метода был взят **линейный конгруэнтный метод** генерации псевдослучайных чисел.

Программа, реализованная в лабораторной работе, выводит на экран таблицу из 7 столбцов и 12 строк. 10 строк представлены для того, чтобы можно было пронаблюдать, какие числа возвращает генератор случайных чисел. Для каждого из реализованных методов в таблице есть по три столбца для чисел с разным количеством разрядов.

В последнем столбце выводится значение V , подсчитанное для каждого столбца. ($N = 10000$)

```
lab_03 [main] python3 lab1.py
```

--Табличный метод--				--Алгоритмический метод--		
№	1 разряд	2 разряд	3 разряд	1 разряд	2 разряд	3 разряд
0	9	99	956	3	51	977
1	4	53	487	6	92	500
2	3	89	760	7	31	223
3	2	30	758	8	84	122
4	9	74	107	5	19	673
5	1	79	802	4	20	180
6	3	67	923	7	19	695
7	2	11	992	6	74	986
8	3	44	621	5	99	125
9	6	90	599	8	34	212
коэф	10.02599999999984	71.46800000000076	863.0	7.727999999999156	74.36599999999999	899.1800000000003

Рис 1. Первый запуск программы

```
lab_03 [main] python3 lab1.py
```

--Табличный метод--				--Алгоритмический метод--		
№	1 разряд	2 разряд	3 разряд	1 разряд	2 разряд	3 разряд
0	8	47	792	3	75	189
1	6	27	736	6	38	416
2	6	17	771	7	97	211
3	0	82	597	8	60	742
4	0	58	477	9	71	297
5	6	34	300	8	18	712
6	8	73	244	1	69	211
7	2	76	420	0	14	318
8	3	68	487	5	49	933
9	6	76	720	2	72	820
коэф	8.479999999999563	61.495999999999185	835.1000000000004	9.067999999999302	84.01399999999921	923.1200000000008

Рис 2. Второй запуск программы

```
lab_03 [main] ⚡ python3 lab1.py
```

--Табличный метод--				--Алгоритмический метод--		
№	1 разряд	2 разряд	3 разряд	1 разряд	2 разряд	3 разряд
0	7	84	671	5	61	249
1	6	74	107	8	74	332
2	9	13	885	9	73	199
3	2	25	118	8	36	462
4	5	67	906	3	33	821
5	8	84	342	4	68	344
6	0	23	164	5	29	679
7	8	18	246	2	44	498
8	8	56	379	5	89	841
9	6	55	506	4	72	580
коэф	13.22400000000016	86.22799999999916	863.3600000000006	6.9979999999995925	82.15999999999985	854.5400000000009

Рис 3. Третий запуск программы

Для правильной оценки случайности методов было проведено 3 испытания.

Сравним полученные данные с таблицей 2. Получим следующий результат:

	Табличный метод			Алгоритмический метод		
№ эксперимента	1 разряд	2 разряд	3 разряд	1 разряд	2 разряд	3 разряд
1	50%-75%	5%-25%	5%-25%	25%-50%	5%-25%	50%-75%
2	50%-75%	1%-5%	5%-25%	50%-75%	25%-50%	50%-75%
3	75%-95%	25%-50%	5%-25%	25%-50%	25%-50%	5%-25%

Таблица 3. Оценка полученных результатов.

Из таблицы 3 видно, что в некоторых случаях при применении табличного метода значения оказываются «подозрительным», однако это не критично и результаты работы генераторов можно признать удовлетворительными. Для алгоритмического метода полученные значения v находятся в рамках 5%-75%, поэтому можно признать и этот метод удовлетворительным.