

Федеральное государственное бюджетное образовательное  
учреждение высшего образования «Сибирский государственный  
университет телекоммуникаций и информатики» (СибГУТИ)

Кафедра Вычислительных систем (ВС)

Лабораторная работа №0  
по дисциплине «Моделирование»

Выполнил:

студент гр. ИВ-622

Гайдук П.А

Работу проверила:

Ассистент кафедры ВС

Петухова Я.В

Новосибирск 2020

## Оглавление

Постановка задачи .....	3
Теоретические сведения .....	3
Выполнение и оценка результатов эксперимента .....	5
Заключение.....	9
Приложение .....	10

## Постановка задачи

Взять готовую реализацию генератора псевдослучайных чисел на отрезке  $[0,1]$  и убедиться в его равномерном распределении, используя такие параметры как  $\chi^2$  и автокорреляцию.

## Теоретические сведения

$\chi^2$  - метод, который позволяет оценить значимость различий между фактическим (выявленным в результате исследования) количеством исходов или качественных характеристик выборки, попадающих в каждую категорию, и теоретическим количеством, которое можно ожидать в изучаемых группах при справедливости нулевой гипотезы.

Процедура проверки гипотез с использованием критериев типа  $\chi^2$  предусматривает группирование наблюдений. Область определения случайной величины разбивают на  $k$  непересекающихся интервалов граничными точками  $(x_0, x_1, x_2, \dots, x_k)$ , где  $x_0$  — нижняя грань области определения случайной величины;  $x_k$  — верхняя грань. В соответствии с заданным разбиением подсчитывают число  $n_i$  выборочных значений, попавших в  $i$ -й интервал, и вероятности попадания в интервал, соответствующие теоретическому закону с функцией распределения.

Рассчитывается критерий «хи-квадрат» по формуле:

$$\chi^2 = \frac{(n_1 - p_1 N)^2}{p_1 N} + \frac{(n_2 - p_2 N)^2}{p_2 N} + \dots + \frac{(n_k - p_k N)^2}{p_k N}$$

, где

- $p_i$  - теоретическая вероятность попадания чисел в  $i$ -ый интервал
- $k$  - количество интервалов
- $N$  - общее количество сгенерированных чисел
- $n_i$  - количество попавших чисел в интервал
- $\chi^2$ - критерий, который позволяет определить, удовлетворяет ли генератор случайных чисел требованиям равномерного распределения или нет

Если  $\chi_{\text{эксп}} \leq \chi_{\text{таб}}$ , то гипотеза не противоречит опытным данным, иначе - отвергается.

Условие применения  $\chi^2$  является наличие в каждом интервале не менее пяти наблюдений. И следует помнить, что для равномерного распределения все значения вероятности одинаковы. Далее надо провести корреляционный анализ. Корреляционный анализ – популярный метод статистического исследования, который используется для выявления степени зависимости одного показателя от другого.

Автокорреляция - статистическая взаимосвязь между последовательностями величин одного ряда, взятыми со сдвигом, например, для случайного процесса — со сдвигом по времени.

$$a(\tau) = \frac{\sum_{i=1}^n (x_i - E(x)) * (x_{i+\tau} - E(x))}{(N - \tau) S^2(x)}$$

, где

$E(x)$  - математическое ожидание — среднее значение случайной величины при стремлении количества выборок или количества её измерений

$S^2(x)$  - выборочная дисперсия случайной величины — это оценка теоретической дисперсии распределения, рассчитанная на основе данных выборки:

$$S^2(x) = \frac{1}{N} \sum_{i=1}^N (x_i - x)^2 = \frac{1}{N} \sum_{i=1}^N x_i^2 - (x)^2$$

$x_{i+\tau}$  - множество значений другой случайной величины (полученной из значений прошлой случайной величины, но с некоторым смещением)

$n$  – мощность множества случайных величин

$\tau$ - смещение последовательности

Генераторы случайных чисел:

- Стандартный генератор случайных чисел, с улучшенным интерфейсом(<https://github.com/effolkronium/random>) – добавлена потоковая безопасность, реализованная в методе `random_thread_local`, также упрощено получение случайных чисел методом `get` и `shuffle`.

- Генератор псевдослучайных чисел из библиотеки Росо – генератор случайных чисел с нелинейной аддитивной обратной связью.
- Генератор псевдослучайных чисел из библиотеки Boost mt19937. В документации этой библиотеки генераторы характеризуются длиной цикла (приблизительная оценка качества генерируемой последовательности) и скоростью генерации числа.

### Выполнение и оценка результатов эксперимента

При запуске программы выводится заданное количество псевдослучайных чисел. Для того, чтобы вычислить  $\chi^2$  необходимо разбить эту последовательность на  $k$  равных интервалов, для каждого интервала поставить в соответствие количество попавших в него чисел и подставить в приведенную выше формулу.

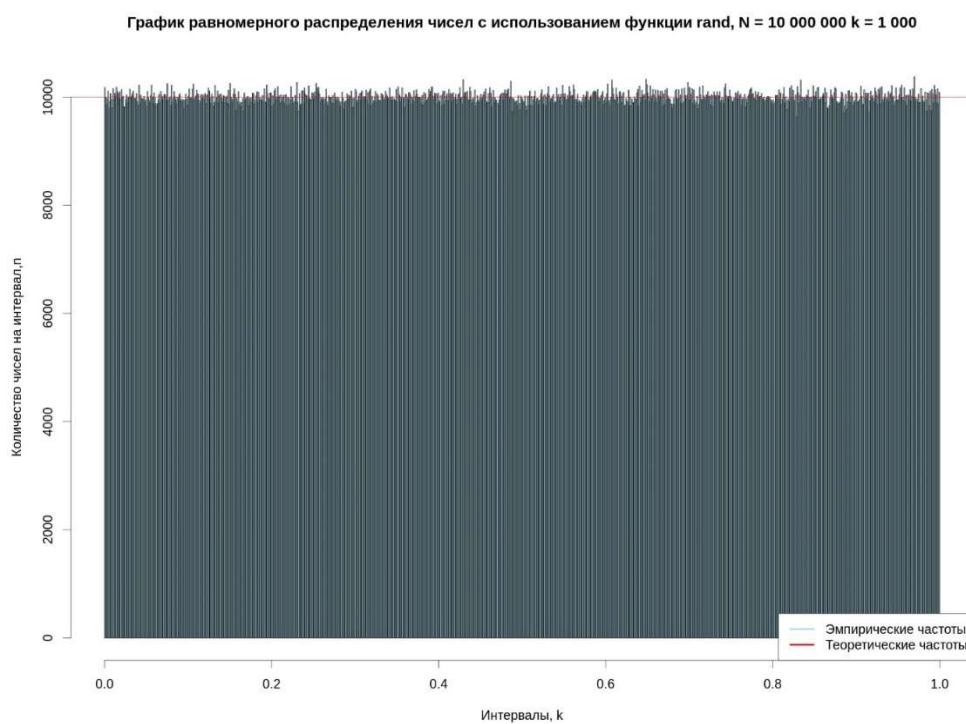
Листинг программы для построения гистограммы и вычисления  $\chi^2$  и автокорреляции, путем смещения интервала на  $\tau$

```
res_n10e7_1000.R
d1 = read.csv("../outPoco.csv", header=FALSE)
h <- hist(d1$V1, breaks=1000)
png("task0.png", width = 1920, height = 1440, units = "px", pointsize = 25)
count <- 0
for (val in 1:1000){
  count = count + ((h$counts[val] - 10000) * (h$counts[val] - 10000) / 10000)
}
print (count)
cor (d1$V1[-1],d1$V1[-length(d1$V1)])
```

Результат для  $N = 10^7$ ,  $k = 1000$  (rand)

$$\chi^2 = 1061.437$$

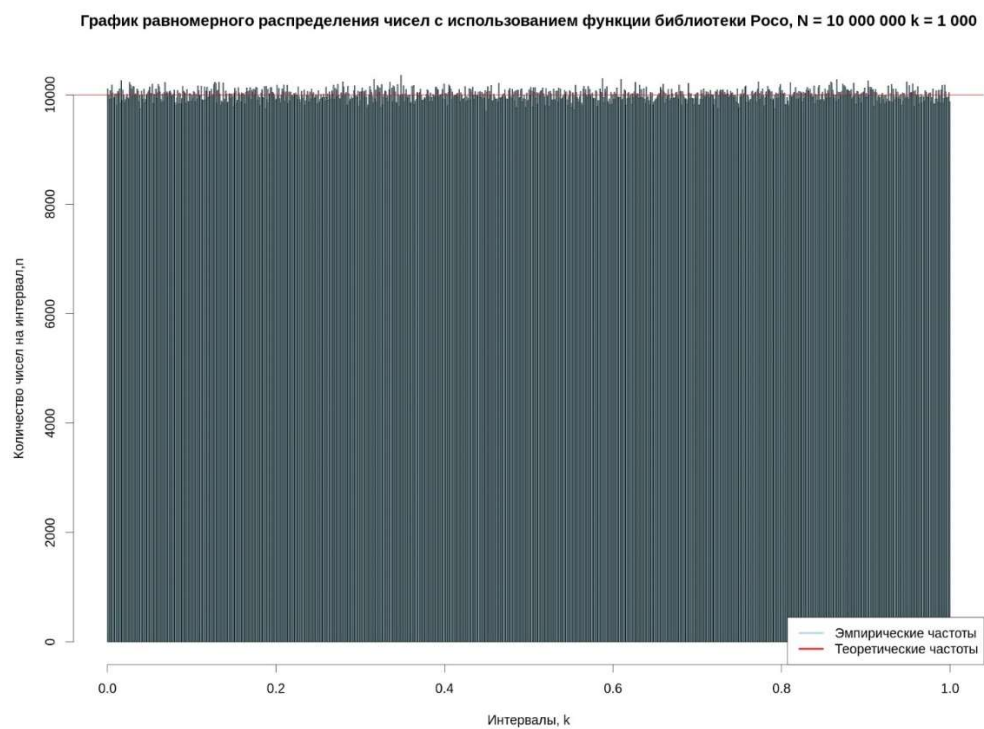
$$a(\tau) = 0.000301628$$



Результат для  $N = 10^7$ ,  $k = 1000$  (Roco)

$$\chi^2 = 1051.617$$

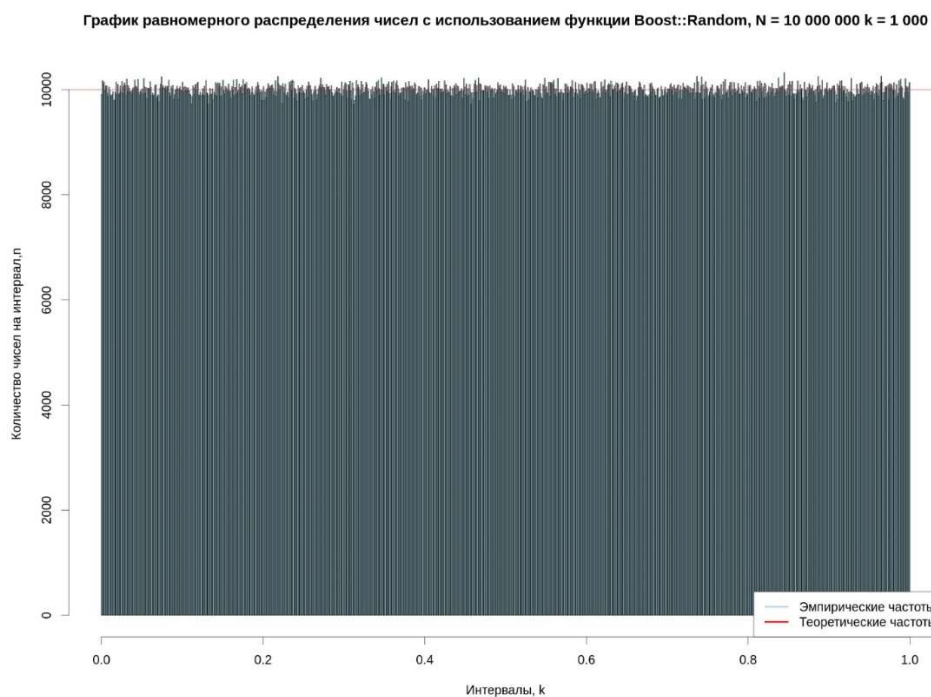
$$a(\tau) = 0.0001881426$$



Результат для  $N = 10^7$ ,  $k = 1000$  (Boost mt19937)

$$\chi^2 = 1021.457$$

$$a(\tau) = 0.0002098877$$



<p> N - количество  чисел,  k—количество  интервалов </p>	<p> N = 10000  k = 1000 </p>	<p> N = 10000000  k = 100 </p>	<p> N = 10000000  k = 1000 </p>
ГСЧ			
rand	$\chi^2 = 957.2$	$\chi^2 = 114.2601$	$\chi^2 = 1061.437$
Росо::Random	$\chi^2 = 984$	$\chi^2 = 86.5736$	$\chi^2 = 1051.617$
Boost::Random mt19937	$\chi^2 = 1025.6$	$\chi^2 = 107.2408$	$\chi^2 = 1021.457$

```

pgaiduk@pgaiduk-VirtualBox:~/Documents/univer/sem8/modeling/lab0/r$ Rscript autocorrelation.R
[1] "τ = 1 0.000187932728960825"
[1] "τ = 2 -0.000444151057183172"
[1] "τ = 3 -0.000168248687077047"
[1] "τ = 4 0.000176680012636754"
[1] "τ = 5 -0.000425491682380691"
[1] "τ = 6 -0.000122738664324294"
[1] "τ = 7 -2.51330174007403e-05"
[1] "τ = 8 0.000322211108549141"
[1] "τ = 9 7.37876665091732e-05"
pgaiduk@pgaiduk-VirtualBox:~/Documents/univer/sem8/modeling/lab0/r$

```

Результат расчета автокорреляции для генератора из библиотеки Росо

```

pgaiduk@pgaiduk-VirtualBox:~/Documents/univer/sem8/modeling/lab0/r$ Rscript autocorrelation.R
[1] "τ = 1 -0.000301635410657703"
[1] "τ = 2 -0.000113198681342385"
[1] "τ = 3 -0.000149941650239367"
[1] "τ = 4 0.000597184480618657"
[1] "τ = 5 -0.000538733981490953"
[1] "τ = 6 0.000945098175690499"
[1] "τ = 7 0.000214980476787693"
[1] "τ = 8 0.000441443823162285"
[1] "τ = 9 -6.43470006326753e-05"

```

Результат расчета автокорреляции для стандартного генератора C

```

pgaiduk@pgaiduk-VirtualBox:~/Documents/univer/sem8/modeling/lab0/r$ Rscript autocorrelation.R
[1] "τ = 1 0.000209855583059798"
[1] "τ = 2 9.13057539779654e-05"
[1] "τ = 3 0.000571793099219385"
[1] "τ = 4 -0.000322829021151534"
[1] "τ = 5 3.87362280225995e-05"
[1] "τ = 6 -0.000192000929634325"
[1] "τ = 7 -0.000281071893240437"
[1] "τ = 8 -0.000391764978881831"
[1] "τ = 9 -9.8929492662744e-05"

```

Результат расчета автокорреляции для генератора Boost::Random::mt19937



## Заключение

В данной лабораторной работе провели ряд экспериментов по исследованию равномерного распределения в трех различных генераторах псевдослучайных чисел в языке программирования C++.

По результатам экспериментов видно, что критерий  $\chi^2$ , вычисленный на основе эмпирических данных ( $\chi^2 = 1061.437$ ,  $\chi^2 = 1051.617$  и  $\chi^2 = 1021.457$ ) удовлетворяет условию  $\chi_{\text{эксп}}^2 < \chi_{\text{табл}}^2(1142.848)$ , из чего можно сделать вывод о том, что гипотезы о равномерном распределении в генераторах псевдослучайных чисел применяются.

В каждом из проведенных экспериментов автокорреляционная функция при изменении параметра  $\tau$  (смещение в последовательности от 1 до половины наших интервалов) приближена к нулю, что говорит нам об очень слабой корреляции. Эта величина показывает, что зависимость между данными крайне мала.

Уменьшая количество интервалов, при неизменном количестве генерируемых чисел, заметим, что коэффициент автокорреляции будет стремиться к нулевому значению. При увеличении количества псевдослучайных чисел коэффициент автокорреляции также будет стремиться к 0.

Таким образом мы изучили распределение псевдослучайных чисел используя параметры  $\chi^2$  и автокорреляцию.

## Приложение

### Generate.cpp

```
#include <Poco/Random.h>
#include <iostream>
#include "../random/include/effolkronium/random.hpp"
#include <limits.h>
#include <Poco/RandomStream.h>

#include <boost/random.hpp>
#include <ctime>
#include <cstdint>

using Poco::Random;
using Poco::RandomInputStream;
using Rand = effolkronium::random_static;
using namespace std;

int main(){
    // ##### POCO
    // Random rnd;
    // rnd.seed();
    // cout << "numbers" << endl;
    // for (int i = 0; i < 10000000; i++)
    //     std::cout << (rnd.next() + 0.0) / (INT_MAX + 0.0) << std::endl;
    // RandomInputStream ri;
    // std::string rs;
    // ri >> rs;
    // ##### POCO END

    // ##### crand
    /*for (int i = 0; i < 10000000; i++){
        auto val = (Rand::get<Rand::common>(0, INT_MAX) + 0.0) / (INT_MAX
+ 0.0);
        cout << val << endl;
    }*/
    // ##### crand end

    std::time_t now = std::time(0);
    boost::random::mt19937 gen{static_cast<std::uint32_t>(now)};
    for (int i = 0; i < 10000000; i++){
        std::cout << (gen() + 0.0) / (gen.max() + 0.0) << '\n';
    }

    return 0;
}
```

### Autocorrelation.R

```
#d1 =
read.csv("/home/pgaiduk/Documents/univer/sem8/modeling/lab0/outPoco.csv", header=FALSE)
#cor (d1$V1[-1],d1$V1[-length(d1$V1)])

#d1 =
read.csv("/home/pgaiduk/Documents/univer/sem8/modeling/lab0/outPoco.csv", header=FALSE)
```

```

#d1 =
read.csv("/home/pgaiduk/Documents/univer/sem8/modeling/lab0/effolkroni
um.csv", header=FALSE)
d1 =
read.csv("/home/pgaiduk/Documents/univer/sem8/modeling/lab0/outBoost.c
sv", header=FALSE)

shifter <- function(x, n = 1) { if (n == 0) x else c(tail(x, -n),
head(x, n)) }

for (i in 1:9){
  print(paste("τ =", i, cor(d1$V1, shifter(d1$V1, i))))
}

```