# ГЕНЕРАТОР СЛУЧАЙНЫХ ЧИСЕЛ

Авторы:

Михайлов Дмитрий Владимирович

Соколов Алексей Николаевич

Научный руководитель:

Батова Марина Михайловна

## Отчёт

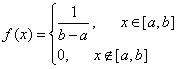
Мы представляем генератор случайных чисел, который по заданным входным параметрам, зависящим от вида распределения, генерирует выбранное пользователем распределение.

Нами реализовано приложение с пользовательским интерфейсом, в котором пользователю сначала необходимо выбрать название распределения, а затем указать числовые значения параметров выбранного распределения.

Приложение написано на языке программирования Python при помощи нескольких модулей.

Среди них наиболее важную роль играет модуль PyQt5, который обеспечивает взаимодествие пользователя с программой.

Нашим приложением генерируются следующие распределения:

1) Равномерное распределение

Функция плотности вероятности:

2) Нормальное распределение

Функция плотности вероятности:

3)Логнормальное распределение

Функция плотности вероятности:

4) Распределение Пуассона

Функция вероятности:

5) Экспоненциальное распределение

Функция плотности вероятности:

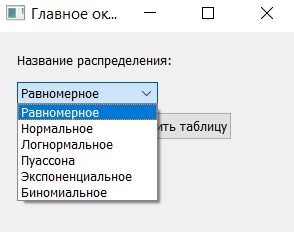
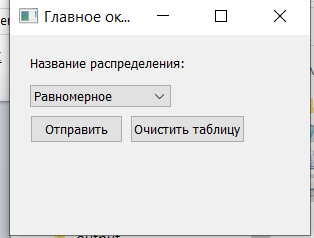
6) Биномиальное распределение

Функция вероятности:

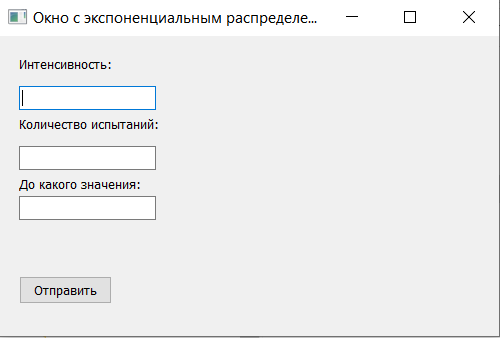
Все распределения генерируются при помощи модуля random

**1) Взаимодействие с пользовательским интерфейсом**

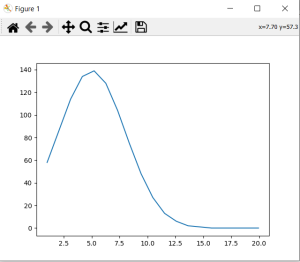
При запуске приложения пользователю предлагается выбрать распределение:



После нажатия кнопки «Отправить» пользователю необходимо ввести соответствующие данные параматров выбранного распределения. Например для экспоненциального распределения необходимо ввести интенсивность, количество испытаний и правую границу значений:



После нажатия кнопки «Отправить» появится график полученного распределения следующего вида:



Области графика можно увеличивать, уменьшать, при наведении мыши на график можно смотреть точные значения.

В это же время (сразу после нажатия кнопки «Отправить») все данные занесутся в таблицу excel в лист с соответствующим распределением:

**2) Описание и корректность работы программы**

Все параметры сгенерированного распределения совпадают с введёнными пользователем параметрами, необходимыми для генерации распределения. В этом можно убедиться с помощью графика и сгенерированного в таблице распределения. Но как на программном уровне достигается генерация распределения, отвечающая заданным параметрам?

Для этого воспользуемся определением определённого интеграла Римана.

Определение интеграла Римана.

Пусть на отрезке [a, b] определена некоторая функция. f.  Будем считать a < b.

**Разбиением (неразмеченным) отрезка** [a, b] назовём конечное множество точек отрезка [a, b] , в которое входят точки  a и b. Как видно из определения, в разбиение всегда входят хотя бы две точки. Точки разбиения можно расположить по возрастанию.

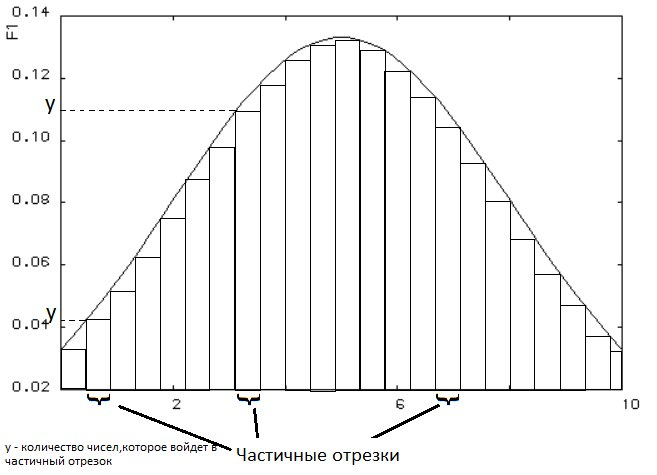
Точки разбиения, между которыми нет других точек разбиения, называются **соседними**. Отрезок, концами которого являются соседние точки разбиения, называется **частичным отрезком разбиения**.

Длина наибольшего из отрезков называется **диаметром разбиения**.

**Интегралом Римана функции f в пределах от  a до b** называется предел интегральных сумм Римана функции f на разбиениях отрезка [a, b] при диаметре разбиения, стремящемуся к нулю.

Основная идея программы очень похожа на определение определённого интеграла: заданный пользователем диапазон разбивается на маленькие части – частичные отрезки. Для каждого частичного отрезка вычисляется количество чисел сгенерированного распределения, которое будет лежать в нём. Это количество зависит от длины диапазона и количества чисел – параметров, заданных пользователем. Примерная модель программы представлена на рисунке:

**Генерация разбиений на примере нормального распределения**



Для дальнейшего объяснения необходимо погрузиться в детали реализации программы.

Введём обозначения:

*quantity* – количество чисел,

*begin* – начало диапазона,

*end* – конец диапазона,

Суть построения всех распределений состоит в том что заданный диапазон разбивается на равные части — частичные отрезки, длина которых вычисляется по формуле:

(конец диапазона – начало диапазона + 1)/length,

где *length* – некоторая переменная, полученная в результате следующих вычислений:

*prefix* – дробная часть длины диапазона

введём переменную gcd:

**gcd = НОД(prefix, 10^(len(prefix))) \***len - длина

Тогда

**length = ((10^(len(prefix)))/gcd)\*(end – begin)**

Но если length получается **очень** большое, то получится слишком много разбиений и вычислительной мощности компьютера не хватит для таких вычислений. Поэтому воспользуемся формулой Стёрджесса, чтобы сократить длину настолько, чтобы вычислительной мощности было достаточно и распределение аппроксимировалось к идеальному:

**length = length/(1 + 3.322\*lg(length))**

Для каждого частичного отрезка высчитывается величина *height*, равная количеству чисел, принадлежащих отрезку, которое войдёт в итоговую выборку. В зависимости от вида распределения высчитываться это число будет по-разному.

Как было сказано выше, изначально числа генерируются с помощью модуля random, программа лишь решает для каждого следующего числа, включать ли его в итоговую выборку или нет. Когда количество чисел в некотором частичном отрезке, включённых в выборку, достигает максимума, то есть будет равна height, то для данного частичного отрезка числа более генерироваться не будут.

Можно заключить, что сгенерированные распределения отвечают заданным параметрам благодаря способу высчитывания частичных отрезков и высоты столбиков — количества чисел, которое войдёт в частичный отрезок.

**3) Виды распределений и особенности их генерирования**

**1) Равномерное распределение**

Для данного распределения пользователю необходимо указать количество чисел и диапазон чисел, в котором будет генерироваться распределение. В равномерном распределении в каждом разбиении должно лежать количество чисел, равное quantity/length. Это достигается с помощью программы. Как только все разбиения полностью заполняются, программа завершается.

**2) Нормальное и логнормальное распределения**

Для данных распределений пользователю необходимо указать математическое ожидание, среднеквадратическое отклонение, количество чисел и диапазон чисел, в котором будет генерироваться распределение. В каждом разбиении должно лежать количество чисел, равное интегралу от функции плотности вероятности с границами, равными границам разбиения, умноженному на количество чисел в распределении, которое задавалось пользователем.

**3) Экспоненциальное распределение**

Пользователю необходимо ввести интенсивность (величина обратная математическому ожиданию), количество чисел и диапазон (от 0 до указанного числа). Программа использует метод обратного преобразования. Пусть U – последовательность равномерно распределенных величин на отрезке (0;1), Х – искомая выборка, тогда



Итоговая формула:

**4) Распределение Пуассона**

Пользователь вводит математическое ожидание и количество испытаний. Для создания выборки используется свойство: ,где Е – распределенные экспоненциально с интенсивностью 1 случайные величины.

**5) Биномиальное распределение**

Пользователь вводит вероятность успеха, число испытаний и необходимое количество чисел. Для создания используются испытания Бернулли.

## 4) Вывод в файл

Все созданные распределения сохраняются в excel файл под названием «Distributions.xlsx». Для каждого вида распределения существует свой лист, каждое распределение добавляется в новый столбец листа.

Для удобства была создана кнопка для очистки файла, при нажатии на которую содержимое файла бесследно удаляется.

Данные удобно импортировать в различные системы автоматизированного проектирования.

**5) Недостатки приложения**

Существует 3 аспекта, замедляющие работу программы:

1) Программа написана на интерпретируемом языке(python);

2) Алгоритм осуществляет проверку на возможность вхождения сгенерированного числа в конечную выборку;

3) Программа однопоточна.

В результате возможна генерация чисел только в определённых пределах и в определённых количествах. Теоретически высчитанно, что если брать диапазон длины 100. то возможна генерация порядка миллиона целых чисел или 15000 дробных чисел.

**6) Область применения**

Область применения программы включает:

1) Исследование параметров личности человека в психологии и психиатрии;

2) Вероятностную модель в финансах;

3) Расчёт вероятности ковенантов по кредитной линии в экономике;

4) Карты контроля качества;

5) Теория массового обслуживания;

6) Телекоммуникация;

7) Медицинская статистика.

Например, для исследования параметров личности человека обычно необходимо нормальное распределение, а для построения финансовой модели нужно воспользоваться равномерным распределением. Распределение Пуассона используется в картах контроля качества, теории массового обслуживания, телекоммуникации, медицинской статистике.

Если говорить непосредственно о нашем приложении, то данные, сгенерированные в файл excel и графики удобно импортировать для дальнейшей работы в различные системы автоматизированного проектирования(matlab, mathcad и др.).

**7) Отличия от аналогов**

Реализация представленного проекта уникальна своим подходом к генерации чисел. Для того, чтобы параметры итогового распределения совпадали с введёнными пользователем значениями, использовалась концепция определения интеграла Римана, а для того, чтобы

программа работала быстро и в то же время не терялась точность итогового распределения, использовалась формула Стёрджесса. Всё проекты, представленные ниже, представляют собой либо аппаратную реализацию с применением теории по схемотехнике, либо логическую, либо программную. Но даже в проектах с программной реализацией не использовались перечисленные выше концепции.

**Преимущества нашей программы над аналогами с программной реализацией:**

1) В некоторых аналогах числа генерируются в полуавтоматическом режиме, то есть тратятся ресурсы на высчитывание числа, попадающего в итоговую выборку и выборка получается не случайной, а значит, высок шанс того, что выборки будут повторяться.

2) Многие программные продукты затрагивают аппаратную часть и требуют дополнительной аппаратуры, а наше приложение мобильно и независимо, его можно развёртывать на любом устройстве.

3) Многие продукты реализовывают какое-то конкретное распределение. Наше приложение может генерировать 6 различных видов распределений.

4) Пользовательский интерфейс: пользователю предоставлен пользовательский интерфейс для генерации распределений, также предоставлена возможность убедиться в правильности разбиения с помощью графика и возможность хранения данных в таблице. Преимущество записи данных в excel состоит в том, что в этой утилите реализованы функции, отображающие параметры распределения.

## Патенты аналоги

| Документ | Номер документа | Дата публикации | Название |
| --- | --- | --- | --- |
| [RU2211481C2\_20030827](https://searchplatform.rospatent.gov.ru/doc/RU2211481C2_20030827) | 2211481 | 2003.08.27 | ГЕНЕРАТОР СЛУЧАЙНЫХ ЧИСЕЛ |

|  |  |  |  |
| --- | --- | --- | --- |
| [RU2002114316A\_20031227](https://searchplatform.rospatent.gov.ru/doc/RU2002114316A_20031227) | 2002114316 | 2003.12.27 | Генератор псевдослучайных последовательностей |

|  |  |  |  |
| --- | --- | --- | --- |
| [RU2001127314A\_20030810](https://searchplatform.rospatent.gov.ru/doc/RU2001127314A_20030810) | 2001127314 | 2003.08.10 | Генератор случайных чисел |

|  |  |  |  |
| --- | --- | --- | --- |
| [RU2223593C1\_20040210](https://searchplatform.rospatent.gov.ru/doc/RU2223593C1_20040210) | 2223593 | 2004.02.10 | ГЕНЕРАТОР ПСЕВДОСЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ |

|  |  |  |  |
| --- | --- | --- | --- |
| [RU99672U1\_20101120](https://searchplatform.rospatent.gov.ru/doc/RU99672U1_20101120) | 99672 | 2010.11.20 | ГЕНЕРАТОР СЛУЧАЙНЫХ ЧИСЕЛ |

|  |  |  |  |
| --- | --- | --- | --- |
| [RU2294559C1\_20070227](https://searchplatform.rospatent.gov.ru/doc/RU2294559C1_20070227) | 2294559 | 2007.02.27 | УСТРОЙСТВО ДЛЯ ГЕНЕРАЦИИ ПСЕВДОСЛУЧАЙНОЙ ПОСЛЕДОВАТЕЛЬНОСТИ ДВОИЧНЫХ ЧИСЕЛ С ИСПОЛЬЗОВАНИЕМ ЭЛЛИПТИЧЕСКИХ КРИВЫХ |

|  |  |  |  |
| --- | --- | --- | --- |
| [RU2427886C2\_20110827](https://searchplatform.rospatent.gov.ru/doc/RU2427886C2_20110827) | 2427886 | 2011.08.27 | ГЕНЕРАТОР ПСЕВДОСЛУЧАЙНЫХ БИНАРНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ |

|  |  |  |  |
| --- | --- | --- | --- |
| [SU1758645A2\_19920830](https://searchplatform.rospatent.gov.ru/doc/SU1758645A2_19920830) | 1758645 | 1992.08.30 | Генератор случайных чисел |

|  |  |  |  |
| --- | --- | --- | --- |
| [RU2092892C1\_19971010](https://searchplatform.rospatent.gov.ru/doc/RU2092892C1_19971010) | 2092892 | 1997.10.10 | ГЕНЕРАТОР РАВНОМЕРНО РАСПРЕДЕЛЕННЫХ СЛУЧАЙНЫХ ЧИСЕЛ |

|  |  |  |  |
| --- | --- | --- | --- |
| [RU2081450C1\_19970610](https://searchplatform.rospatent.gov.ru/doc/RU2081450C1_19970610) | 2081450 | 1997.06.10 | ГЕНЕРАТОР N-ЗНАЧНОЙ ПСЕВДОСЛУЧАЙНОЙ ПОСЛЕДОВАТЕЛЬНОСТИ |