

Задача №2. Тестирование последовательностей псевдослучайных чисел.

Датчики псевдослучайных чисел разрабатываются так, чтобы генерируемые ими последовательности можно было считать реализациями независимых случайных величин, равномерно распределённых на единичном отрезке. Ваше задание — реализовать такой датчик и проверить генерируемую последовательность на равномерность и независимость.

Выполните следующие шаги:

1. Рассчитайте 100 псевдослучайных чисел методом, соответствующим вашему варианту. Описание методов дано на второй и третьей страницах.
 2. Приведите первые 10 чисел этой последовательности.
 3. Постройте гистограмму с 10 столбцами для полученной последовательности.
 4. Проверьте гипотезу о том, что последовательность имеет распределение $R(0, 1)$ критерием хи-квадрат, разбив интервал $[0; 1)$ на десять равных интервалов.
 5. Повторите шаги 3 и 4 для последовательности длиной в 10000 чисел.
 6. Изучите тест перестановок (он описан на третьей странице) и проверьте этим тестом первые 9999 чисел вашей последовательности, разбив их на тройки.
- Используйте уровень значимости 5%.

Методы генерации псевдослучайных чисел и проверяющие по вариантам:

Вариант 1. Метод серединных квадратов, $z_1 = 1661$.

Проверяет Романов А.Д., adromanov_4@edu.hse.ru

Вариант 2. Метод серединных произведений, $z_1 = 8731$, $z_2 = 1617$.

Проверяет Фурманов К.К., furmach@inbox.ru.

Вариант 3. Линейный конгруэнтный метод, $z_1 = 2456$, $a = 113$, $b = 2$, $m = 10000$.

Проверяет Румянцева Е.В., evrumyantseva.2006@yandex.ru.

Вариант 4. Степенной остаточный №1, $z_1 = 1237$.

Проверяет Нуртдинов Р.А., ranurtdinov@edu.hse.ru

Вариант 5. Линейный конгруэнтный метод, $z_1 = 10^9$, $a = 16807$, $b = 0$, $m = 2147483647$.

Проверяет Коломенцов А.М., amkolomentsov@edu.hse.ru.

Вариант 6. Метод серединных квадратов, $z_1 = 1687$.

Проверяет Дубновицкая А.А., n.dubnovitskaya@gmail.com

Вариант 7. Метод серединных произведений, $z_1 = 3308$, $z_2 = 1949$.

Проверяет Дубновицкая А.А., n.dubnovitskaya@gmail.com

Вариант 8. Степенной остаточный №2, $z_1 = 7724$.

Проверяет Чернышёва И.К., riina_mustika@mail.ru.

Вариант 9. Линейный конгруэнтный метод, $z_1 = 2456$, $a = 2119$, $b = 0$, $m = 32767$.

Проверяет Коломенцов А.М., amkolomentsov@edu.hse.ru.

Вариант 10. Степенной остаточный №3, $z_1 = 4508$.

Проверяет Серебrenников А.Д., adserebrennikov@edu.hse.ru

Срок сдачи: до 21 марта включительно.

Штрафы за опоздание: как и раньше.

Описание методов генерации

Метод серединных квадратов.

Допустим, мы хотим генерировать числа с четырьмя знаками после запятой. Для этого берётся некое начальное четырёхзначное число z_1 и возводится в квадрат — получается восьмизначное число. Средние четыре цифры этого восьмизначного числа задают следующий элемент этой последовательности — z_2 . Затем z_2 возводится в квадрат и берутся средние четыре цифры от квадрата и т. д.

В конце укладываем все полученные числа в пределы от нуля до единицы простым способом: $x_i = \frac{z_i}{10000}$.

Если при возведении в квадрат получается число меньше чем из восьми знаков, считаем, что оно всё же восьмизначное, просто у него в начале стоят нули.

Пример генерируемой последовательности: 0.1111, 0.2343, 0.4896, ...

Метод серединных произведений.

Допустим, мы хотим генерировать числа с четырьмя знаками после запятой. Для этого берём два четырёхзначных числа z_1 и z_2 и перемножаем их — получается восьмизначное число. Средние четыре цифры этого восьмизначного числа задают следующий элемент последовательности — z_3 . Затем умножаем z_2 на z_3 и берём средние четыре цифры из этого произведения — получаем z_4 , которое потом умножается на z_3 и т. д.

В конце укладываем все полученные числа в пределы от нуля до единицы простым способом: $x_i = \frac{z_i}{10000}$.

Если при возведении в квадрат получается число меньше чем из восьми знаков, считаем, что оно всё же восьмизначное, просто у него в начале стоят нули.

Пример генерируемой последовательности (первые два числа — взятые с потолка начальные значения): 0.1111, 0.5509, 0.1204, 0.6328, ...

Линейный конгруэнтный метод.

Волшебным образом назначаются три параметра генерируемой последовательности a , b и m и берётся начальное число $z_1 < m$. Следующее число рассчитывается по формуле $z_i = (az_{i-1} + b) \bmod m$, \bmod — остаток от деления.

Полученная последовательность укладывается в пределы от 0 до 1 так: $x_i = \frac{z_i}{m}$.

Степенной остаточный №1.

Назначаем начальное число $z_1 < 10000$. Последующие числа получаем из соотношения $z_i = ((z_{i-1}^{2.5}) \div 100) \bmod 10000$, \div — целая часть от деления, \bmod — остаток от деления.

Полученная последовательность укладывается в пределы от 0 до 1 так: $x_i = \frac{z_i}{10000}$.

Степенной остаточный №2.

Назначаем начальное число $z_1 < 10000$. Последующие числа получаем из соотношения $z_i = ((z_{i-1} + 17)^{2.2} \div 100) \bmod 10000$, \div — целая часть от деления, \bmod — остаток от деления.

Полученная последовательность укладывается в пределы от 0 до 1 так: $x_i = \frac{z_i}{10000}$.

Степенной остаточный №3.

Назначаем начальное число $z_1 < 10000$. Последующие числа получаем из соотношения $z_i = ((z_{i-1} + 58)^{1.6} \text{ div } 10) \bmod 10000$, div — целая часть от деления, \bmod — остаток от деления.

Полученная последовательность укладывается в пределы от 0 до 1 так: $x_i = \frac{z_i}{10000}$.

Тест перестановок (permutation test)

Это один из критериев, с помощью которых проверяют независимость последовательности случайных чисел. Возьмём для примера такую последовательность:

0.278 0.490 0.001 0.977 0.404 0.551 0.303 0.322 0.880

Разобьём её на три тройки:

Первая тройка: 0.278 0.490 0.001

Вторая тройка: 0.977 0.404 0.551

Третья тройка: 0.303 0.322 0.880

Элементам каждой тройки поставим в соответствие *ранг*: 1 — наименьшему элементу, 2 — среднему, 3 — наибольшему.

В наших тройках ранги распределятся так:

Первая тройка: 2 3 1.

Вторая тройка: 3 1 2.

Третья тройка: 1 2 3.

Всего существует $3!=6$ возможных перестановок рангов, которые в случае независимых элементов последовательности должны возникать равновероятно. Тест перестановок заключается в расчёте частот каждой из шести перестановок в тестируемой последовательности и применении критерия согласия хи-квадрат для проверки основной гипотезы о равновероятности этих перестановок. Последовательность проходит тест, если основная гипотеза не отвергается на выбранном уровне значимости.

Тест может применяться и к подпоследовательностям из четырёх чисел, тогда число возможных перестановок будет $4!=24$. Но в этом задании нужно сделать тест для троек.

Примечание. В некоторых вариантах встречаются тройки с повторяющимися элементами вроде {0.6875, 0.7945, 0.6875} — в идеальных датчиках такого нет и тест на такое не рассчитан. В учебных целях давайте считать, что из одинаковых чисел ранг больше у того, которое встречается позже. Так, тройке {0.6875, 0.7945, 0.6875} будут соответствовать ранги {1, 3, 2}. Это бессмысленный произвол, надо же как-то считать.