***Пименов Сергей Юрьевич,***

*инженер-конструктор,*

*АО «НИИП имени В.В. Тихомирова»,*

*г. Жуковский, Московская область;*

***Тинаев Василий Владимирович,***

*инженер-конструктор,*

*АО «Государственный Рязанский приборный завод»,*

*г. Рязань*

# ПРИМЕНЕНИЕ ЦЕНТРАЛЬНОЙ ПРЕДЕЛЬНОЙ ТЕОРЕМЫ ДЛЯ КОМПЬЮТЕРНОГО МОДЕЛИРОВАНИЯ СЛУЧАЙНЫХ СИГНАЛОВ

Существует несколько определений центральной предельной теоремы. Во всех источниках сказано, что сумма достаточно большого количества случайных чисел будет распределена по нормальному закону, однако нигде не указано, какое конкретно количество случайных чисел необходимо. В этой статье мы попытались разобраться, как же все-таки влияет число слагаемых на качество результата.

Нормальной (или гауссовой) называется случайная величина X, определенная на всей числовой оси (− ∞, ∞), имеющая плотность распределения вероятности f(x) [3]:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{1}$$

где μ – математическое ожидание,

σ – среднеквадратическое отклонение.

Центральная предельная теорема утверждает, что сумма достаточно большого количества слабо зависимых случайных величин, имеющих примерно одинаковые масштабы (ни одно из слагаемых не доминирует, не вносит в сумму определяющего вклада), имеет распределение, близкое к нормальному.

Для исследования возьмем несколько источников случайной величины, каждый из которых представляет собой сумму случайных величин, распределенных равномерно. Источники отличаются между собой количеством слагаемых. Случайные числа с равномерным распределением формируются с помощью встроенных средств языка программирования Java. Выберем $\mu = 0$ и $\sigma = 1$. Такое распределение называется стандартным нормальным распределением [1].

Соответствие полученных последовательностей нормальному закону распределения будем оценивать с помощью критериев Колмогорова и $\chi^2$, а также качественно по графическому изображению плотности вероятности.

Критерий Колмогорова проверяет гипотезу о принадлежности выборки заданному закону распределения, соответствие реального распределения теоретическому. Статистика критерия вычисляется по формуле:

$$D_n = \sup\left|F_n(x) - F(x)\right| \tag{2}$$

где $F_n(x)$ – эмпирическое значение функции распределения выборки случайных чисел для значения x,

F(x) – теоретическое значение функции распределения случайной величины с нормальным законом распределения для значения x,

sup – точная верхняя грань.

Чем меньше значение $D_n$, тем ближе эмпирическое распределение к теоретическому.

Критерий согласия $\chi^2$ применяется для проверки гипотез о согласии наблюдаемых выборочных данных с предполагаемым законом распределения случайной величины. Статистика критерия $\chi^2$ определяется формулой:

$$\chi^2_{k-1} = \sum_{j=0}^{k} \frac{(n_j - E_j)^2}{E_j} \tag{3}$$

где $\chi^2_{k-1}$ - распределение $\chi^2$ с (k-1) степенями свободы,

$n_j$ – эмпирическая частота появления случайного числа в заданном поддиапазоне,

$E_j$ – теоретическая частота появления случайного числа в заданном поддиапазоне.

Вообще сфера применения критериев согласия чрезвычайно широка. Это задачи, возникающие в различных областях техники, в массовом производстве, в финансах и экономике, при обработке результатов измерений, при контроле качества и испытаниях продукции, при обработке экспериментальных наблюдений в научных исследованиях [2].

В эксперименте использовались 10 источников с различным числом слагаемых, распределенных равномерно. С помощью каждого источника сгенерированы последовательности длиной 5.000.000, посчитаны статистики критериев Колмогорова и $\chi^2$. Результаты приведены в таблице 1.

Таблица 1 – Статистические данные по критериям Колмогорова и χ2

| Число слагаемых | $D_n$ | $\chi^2$ |
|---|---|---|
| 2 | 0,01756 | 0,02564 |
| 3 | 0,009036 | 0,00758 |
| 5 | 0,006441 | 0,002489 |
| 10 | 0,003431 | 0,0007118 |
| 20 | 0,002342 | 0,0002963 |
| 30 | 0,001604 | 0,0002363 |
| 40 | 0,001594 | 0,0002154 |
| 50 | 0,001583 | 0,0002068 |
| 100 | 0,001578 | 0,0001859 |
| 500 | 0,001572 | 0,0001801 |

По данным таблицы 1 видно, что при увеличении количества слагаемых величины $D_n$ и $\chi^2$ уменьшаются, следовательно характеристики эмпирических последовательностям приближаются к теоретическим. При количестве слагаемых более 30 снижение $D_n$ и $\chi^2$ незначительно.

На рисунке 1 даны графики плотностей вероятности полученных последовательностей. Точками обозначены значения для эмпирических выборок, сплошной линией – график теоретической плотности вероятности для стандартного нормального закона распределения, полученный по формуле (1). По рисункам видно, что при увеличения числа слагаемых эмпирическая гистограмма приближается к теоретической кривой. Графики для числа слагаемых 40, 50, 100 и 500 не представлены, так как при данном масштабе не отличаются от графика для 30 слагаемых.

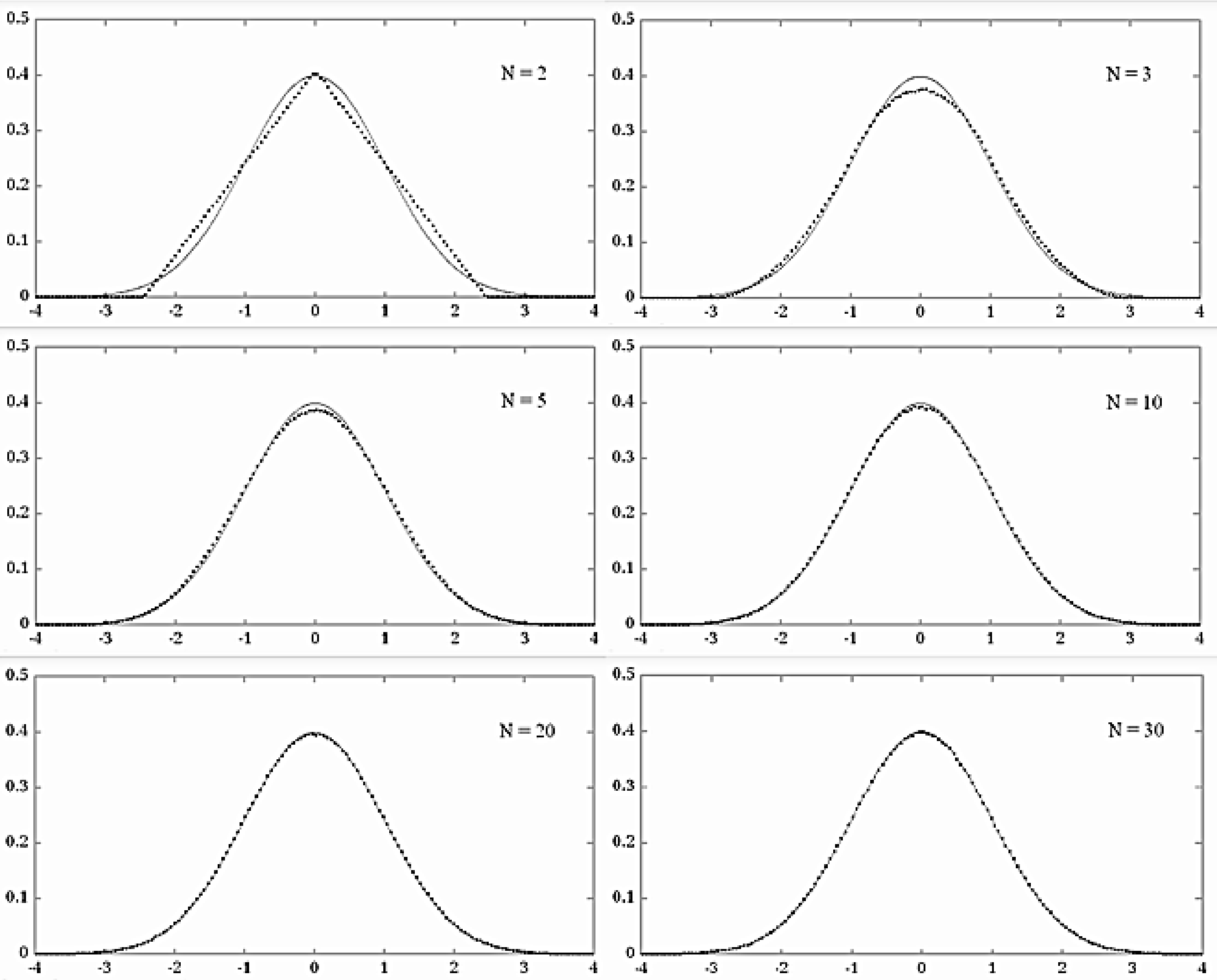


Рисунок 1– Графики плотностей вероятности случайных последовательностей

По итогам проведенной работы можно сделать следующие выводы:

Степень соответствия плотности распределения последовательностей случайных чисел, полученных по центральной предельной теореме, зависит от количества слагаемых. Можно выбирать различную точность под конкретные решаемые задачи.

Недостаток использования теоремы – необходимость подбора дисперсии случайных чисел с равномерным распределением при получении дисперсии стандартного нормального распределения в зависимости от числа слагаемых.

Вычисление случайной величины по центральной предельной теореме при большом количестве слагаемых требует значительных временных затрат. При использовании 3 слагаемых график плотности вероятности начинает приближаться к нормальному; при 30 и более слагаемых график плотности вероятности близок к теоретическому. Дальнейшее увеличение не приводит к существенным изменениям характеристик, однако увеличивает время генерации одного отсчета.

*СПИСОК ЛИТЕРАТУРЫ*

1. *Гельгор А.Л. Методы моделирования случайных величин и случайных процессов. – СПб.: Издательство политехнического университета, 2012.*

2. *Денисов В.И., Лемешко Б.Ю., Постовалов С.Н. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим: Методические рекомендации. – Новосибирск: Новосибирский государственный технический университет, 1998.*

3. *Дорофеев Н.В. Компьютерное и имитационное моделирование. Методические указания к лабораторным работам. – Муром: ФГБО УВПО «Владимирский государственный университет имени Столетовых», 2014.*