Порядок выполнения лабораторной работы №1
«Получение случайных чисел с заданным законом распределения»

**Задание** (в качестве примера): получить числа с равномерным законом распределения в интервале [2,7].

1. **Выбор метода получения случайных чисел** (среди методов, приведенных в методических указаниях к лабораторной работе).

В задании указано, что случайные числа описываются равномерным законом распределения в интервале [a=2,b=7]. Плотность распределения для таких чисел задается формулой

$$f(y) = \begin{cases} 1/(b-a), \text{ если } y€[a,b] \\ 0, \text{ вне } y€[a,b] \end{cases}$$

| Ситуация выбора метода получения случайных чисел | Решение |
|---|---|
| Задана аналитически плотность распределения случайной величины f(y), можно <u>аналитически</u> в общем случае решить уравнение вида, взяв интеграл от f(y) $$\int_{-\infty}^{y_j} f(x)dx = \xi_j$$ | **Метод обратной функции (прямой метод)** |
| Получение случайных чисел с нормальным законом распределения.<br>Получение случайных чисел, распределённых по закону Пуассона. | Получение случайных чисел с использованием предельных теорем теории вероятностей |
| Пусть требуется получить последовательность случайных чисел $\{y_j\}$ с функцией плотности f(y), возможные значения которой лежат в интервале (c, d). Функция плотности распределения f(y) представляется в виде кусочно-постоянных функций, т.е. интервал (c, d) разбивается на $\ell$ подинтервалов и на каждом подинтервале f(y) считается постоянной | Приближенный универсальный способ получения случайных чисел (метод кусочной аппроксимации функции плотности) |

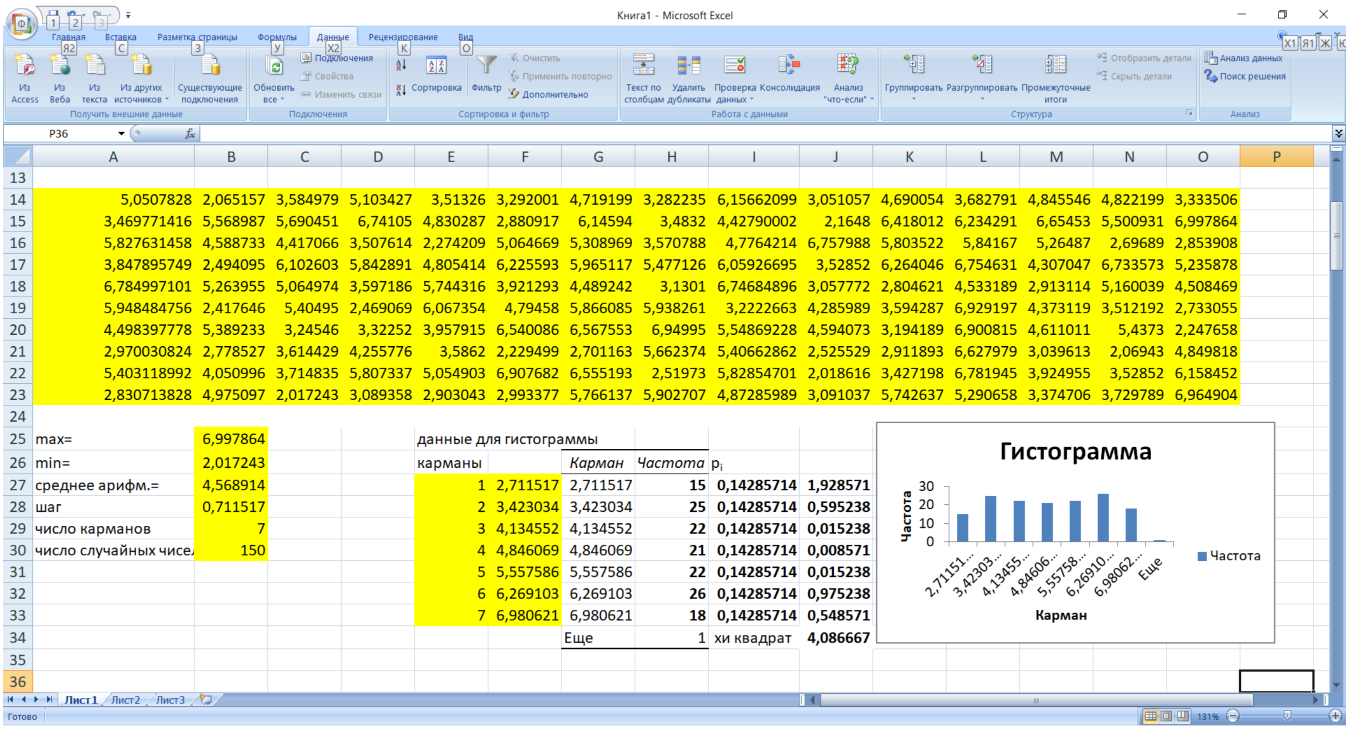
2. **Решить уравнение вида** $\int_{-\infty}^{y_j} f(x)dx = \xi_j$, где $\xi_j$ – случайное число с равномерным законом распределения в интервале [0,1], которую программно можно определить, используя функцию генерации случайных чисел с равномерным законом распределения, имеющуюся в библиотеке используемого программного инструментария, например, =random(); $y_j$ – <u>неизвестная искомая случайная величина</u>, распределенная по требуемому в задании закону.

В данном случае уравнение имеет вид $\int_{a}^{y_j} (1/(b-a))\, dx = \xi_j$, если x€[a,b], вне данного интервала функция плотности распределения случайной величины равна f(x)=0. Решение имеет вид

$y_j = a+(b-a)* \xi_j$. Для интервала, указанного в задании, решение имеет вид $y_j = 2+5* \xi_j$. Это выражение и определяет формулу для получения случайного числа с равномерным законом распределения в интервале [2,7].

3. **Сгенерируем 200 случайных чисел, построим гистограмму для полученной выборки и рассчитаем среднее значение и дисперсию.**
   Используем для простоты реализации MS Excel.

   На рис.1 представлены основные результаты расчетов. Ниже приведены использованные формулы и пояснения к расчетам. При выполнении данного задания результаты изменятся, поскольку будут сгенерированы другие случайные числа. Это не влияет на порядок выполнения работы.



| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 13 | | | | | | | | | | | | | | | |
| 14 | 5,0507828 | 2,065157 | 3,584979 | 5,103427 | 3,51326 | 3,292001 | 4,719199 | 3,282235 | 6,15662099 | 3,051057 | 4,690054 | 3,682791 | 4,845546 | 4,822199 | 3,333506 |
| 15 | 3,469771416 | 5,568987 | 5,690451 | 6,74105 | 4,830287 | 2,880917 | 6,14594 | 3,4832 | 4,42790002 | 2,1648 | 6,418012 | 6,234291 | 6,65453 | 5,500931 | 6,997864 |
| 16 | 5,827631458 | 4,588733 | 4,417066 | 3,507614 | 2,274209 | 5,064669 | 5,308969 | 3,570788 | 4,7764214 | 6,757988 | 5,803522 | 5,84167 | 5,26487 | 2,69689 | 2,853908 |
| 17 | 3,847895749 | 2,494095 | 6,102603 | 5,842891 | 4,805414 | 6,225593 | 5,965117 | 5,477126 | 6,05926695 | 3,52852 | 6,264046 | 6,754631 | 4,307047 | 6,733573 | 5,235878 |
| 18 | 6,784997101 | 5,263955 | 5,064974 | 3,597186 | 5,744316 | 3,921293 | 4,489242 | 3,1301 | 6,74684896 | 3,057772 | 2,804621 | 4,533189 | 2,913114 | 5,160039 | 4,508469 |
| 19 | 5,948484756 | 2,417646 | 5,40495 | 2,469069 | 6,067354 | 4,79458 | 5,866085 | 5,938261 | 3,2222663 | 4,285989 | 3,594287 | 6,929197 | 4,373119 | 3,512192 | 2,733055 |
| 20 | 4,498397778 | 5,389233 | 3,24546 | 3,32252 | 3,957915 | 6,540086 | 6,567553 | 6,94995 | 5,54869228 | 4,594073 | 3,194189 | 6,900815 | 4,611011 | 5,4373 | 2,247658 |
| 21 | 2,970030824 | 2,778527 | 3,614429 | 4,255776 | 3,5862 | 2,229499 | 2,701163 | 5,662374 | 5,40662862 | 2,525529 | 2,911893 | 6,627979 | 3,039613 | 2,06943 | 4,849818 |
| 22 | 5,403118992 | 4,050996 | 3,714835 | 5,807337 | 5,054903 | 6,907682 | 6,555193 | 2,51973 | 5,82854701 | 2,018616 | 3,427198 | 6,781945 | 3,924955 | 3,52852 | 6,158452 |
| 23 | 2,830713828 | 4,975097 | 2,017243 | 3,089358 | 2,903043 | 2,993377 | 5,766137 | 5,902707 | 4,87285989 | 3,091037 | 5,742637 | 5,290658 | 3,374706 | 3,729789 | 6,964904 |
| 24 | | | | | | | | | | | | | | | |
| 25 | max= | 6,997864 | | | данные для гистограммы | | | | | | | | | | |
| 26 | min= | 2,017243 | | | карманы | | *Карман* | *Частота* | $p_i$ | | | | | | |
| 27 | среднее арифм.= | 4,568914 | | | 1 | 2,711517 | 2,711517 | **15** | **0,14285714** | **1,928571** | | | | | |
| 28 | шаг | 0,711517 | | | 2 | 3,423034 | 3,423034 | **25** | **0,14285714** | **0,595238** | | | | | |
| 29 | число карманов | 7 | | | 3 | 4,134552 | 4,134552 | **22** | **0,14285714** | **0,015238** | | | | | |
| 30 | число случайных чисел | 150 | | | 4 | 4,846069 | 4,846069 | **21** | **0,14285714** | **0,008571** | | | | | |
| 31 | | | | | 5 | 5,557586 | 5,557586 | **22** | **0,14285714** | **0,015238** | | | | | |
| 32 | | | | | 6 | 6,269103 | 6,269103 | **26** | **0,14285714** | **0,975238** | | | | | |
| 33 | | | | | 7 | 6,980621 | 6,980621 | **18** | **0,14285714** | **0,548571** | | | | | |
| 34 | | | | | | | Еще | 1 | хи квадрат | **4,086667** | | | | | |



Рис.1 Результаты расчета величины критерия согласия Пирсона (хи квадрат) для задания

Рассмотрим порядок расчетов, представленных на рис.1 по шагам в MS Excel. Можно взять любой программный инструмент для выполнения расчетов.

3.1. Предварительно необходимо сгенерировать 150 (число выбрано произвольно) чисел с равномерным законом распределения в интервале [0,1].
Выполнить команду ДАННЫЕ/АНАЛИЗ_ДАННЫХ/ГЕНЕРАЦИЯ_СЛУЧАЙНЫХ_ЧИСЕЛ. В открывшемся окне «Генерация случайных чисел» ввести исходные данные:

- В поле ввода «Число переменных» ввести в данном случае число 15. Это число определяет число ячеек по горизонтали, т.е. число столбцов, области ячеек, которая отводится под случайные числа.
- В поле ввода «Число случайных чисел» ввести в данном случае число 10. Это число определяет число ячеек по вертикали, т.е. число строк, области ячеек, которая отводится под случайные числа.
- Выбрать в раскрывающемся списке «Распределение» вариант «Равномерное». Автоматически в окне диалога «Генерация случайных чисел» появятся параметры равномерного распределения; интервал [0,1].
- Включить радиокнопку «Выходной интервал» и ввести в ставшее активным поле ввода адрес левой верхней ячейки области размером 15*10, в которой программа разместит все двести случайных чисел.

  На рис.2 приведены все исходные данные в окне «Генерация случайных чисел».

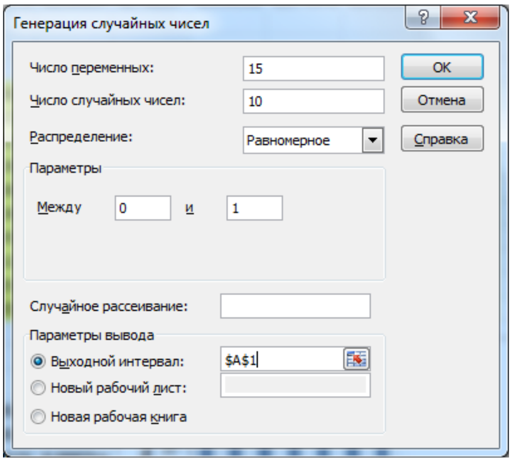
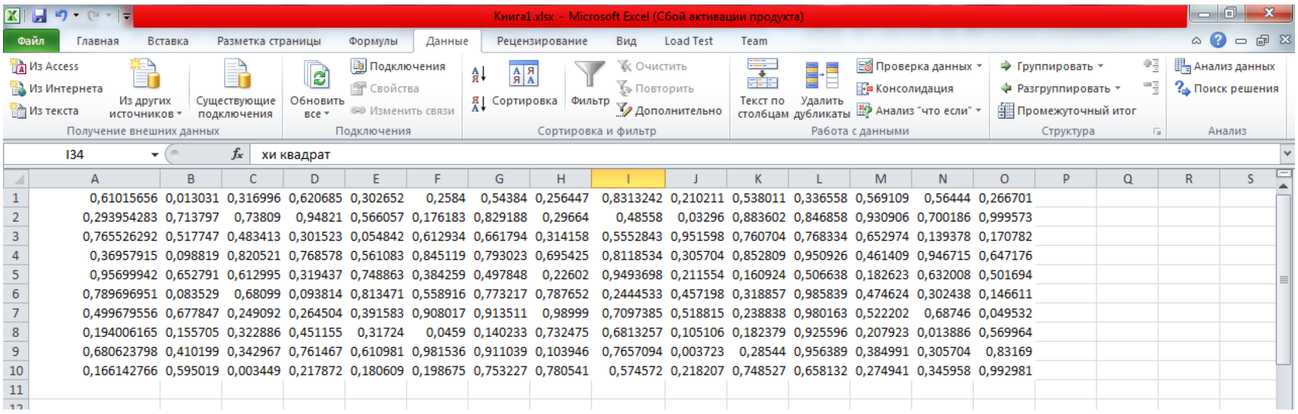



Рис.2. Исходные данные в окне диалога «Генерация случайных чисел»

- Нажать командную кнопку «ОК». Результат представлен на рис.3.



| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0,61015656 | 0,013031 | 0,316996 | 0,620685 | 0,302652 | 0,2584 | 0,54384 | 0,256447 | 0,8313242 | 0,210211 | 0,538011 | 0,336558 | 0,569109 | 0,56444 | 0,266701 |
| 2 | 0,293954283 | 0,713797 | 0,73809 | 0,94821 | 0,566057 | 0,176183 | 0,829188 | 0,29664 | 0,48558 | 0,03296 | 0,883602 | 0,846858 | 0,930906 | 0,700186 | 0,999573 |
| 3 | 0,765526292 | 0,517747 | 0,483413 | 0,301523 | 0,054842 | 0,612934 | 0,661794 | 0,314158 | 0,5552843 | 0,951598 | 0,760704 | 0,768334 | 0,652974 | 0,139378 | 0,170782 |
| 4 | 0,36957915 | 0,098819 | 0,820521 | 0,768578 | 0,561083 | 0,845119 | 0,793023 | 0,695425 | 0,8118534 | 0,305704 | 0,852809 | 0,950926 | 0,461409 | 0,946715 | 0,647176 |
| 5 | 0,95699942 | 0,652791 | 0,612995 | 0,319437 | 0,748863 | 0,384259 | 0,497848 | 0,22602 | 0,9493698 | 0,211554 | 0,160924 | 0,506638 | 0,182623 | 0,632008 | 0,501694 |
| 6 | 0,789696951 | 0,083529 | 0,68099 | 0,093814 | 0,813471 | 0,558916 | 0,773217 | 0,787652 | 0,2444533 | 0,457198 | 0,318857 | 0,985839 | 0,474624 | 0,302438 | 0,146611 |
| 7 | 0,499679556 | 0,677847 | 0,249092 | 0,264504 | 0,391583 | 0,908017 | 0,913511 | 0,98999 | 0,7097385 | 0,518815 | 0,238838 | 0,980163 | 0,522202 | 0,68746 | 0,049532 |
| 8 | 0,194006165 | 0,155705 | 0,322886 | 0,451155 | 0,31724 | 0,0459 | 0,140233 | 0,732475 | 0,6813257 | 0,105106 | 0,182379 | 0,925596 | 0,207923 | 0,013886 | 0,569964 |
| 9 | 0,680623798 | 0,410199 | 0,342967 | 0,761467 | 0,610981 | 0,981536 | 0,911039 | 0,103946 | 0,7657094 | 0,003723 | 0,28544 | 0,956389 | 0,384991 | 0,305704 | 0,83169 |
| 10 | 0,166142766 | 0,595019 | 0,003449 | 0,217872 | 0,180609 | 0,198675 | 0,753227 | 0,780541 | 0,574572 | 0,218207 | 0,748527 | 0,658132 | 0,274941 | 0,345958 | 0,992981 |

Рис. 3 Область ячеек A1:O10, в которой размещены случайные числа

3.2. Сгенерируем случайные числа с равномерным законом распределения в интервале [2,7]. Формула, с помощью которой рассчитываются случайные числа $y_j$, имеет вид $y_j = a+(b-a)* \xi_j$, $j=1,2,\ldots,150$, где $a=2$, $b=7$, $\xi_j$ – случайное число с равномерным законом распределения в интервале [0,1]:

- Ввести в ячейку A14 формулу =2+5*A1.
- Скопировать эту формулу в область ячеек A14:O23 (см. рис.1).

3.3. Рассчитать шаг для определения карманов при построении гистограммы. Карманы определяют точки, отмечаемые на оси абсцисс гистограммы.

- Определить максимальное случайное число в области A14:O23. Ввести в ячейку B25 формулу =МАКС(A14:O23).
- Сравнить полученное значение с максимальным значением случайных чисел с равномерным законом распределения в интервале [2,7], которое равно 7. Сравниваемые значения близки по величине, что свидетельствует о качестве сгенерированных чисел по данной выборке.
- Определить минимальное случайное число в области A14:O23. Ввести в ячейку B26 формулу =МИН(A14:O23).
- Сравнить полученное значение с минимальным значением случайных чисел с равномерным законом распределения в интервале [2,7], которое равно 2. Сравниваемые

значения близки по величине, что свидетельствует о качестве сгенерированных чисел по данной выборке.

- Определить среднее арифметическое число в области A14:O23. Ввести в ячейку B27 формулу =(B25-B26)/$B$29.
- Сравнить полученное значение с математическим ожиданием случайных чисел с равномерным законом распределения в интервале [2,7], которое определяется по формуле М=(а+в)/2=(7-2)/2=4,5. Полученные значения близки по величине, что свидетельствует о качестве сгенерированных чисел по данной выборке.
- Задать число карманов для построения гистограммы. Это число карманов (подинтервалов), на которое будет разделена область изменения сгенерированных случайных чисел. Ввести в ячейку B29 число 7. В общем случае рекомендуется делить на 5-20 карманов. Этот диапазон определен опытным путем.
- Определить величину шага, с которым будут определяться карманы на гистограмме по оси абсцисс. Ввести в ячейку B28 формулу =(B25-B26)/$B$29.
- Задать количество сгенерированных случайных чисел. Ввести в ячейку B30 число 150.
- В ячейки B25:B30, E25:E26? I34 введены комментарии к использованным данным и выполненным расчетам.

3.4. Построим гистограмму для полученной в области A14:O23 выборки случайных чисел:

- Ввести в ячейки E27:E33 порядковые номера карманов от 1 до 7.
- Определить точки окончания каждого кармана на оси абсцисс. Ввести в ячейку F27 формулу =2+$B$28*E27. Скопировать эту формулу в область ячеек F28:F33.
- Выполнить команду ДАННЫЕ/АНАЛИЗ_ДАННЫХ.
- В открывшемся окне «Анализ данных» в списке «Инструменты анализа» выбрать команду «Гистограмма» нажать командную кнопку «ОК».
- В открывшемся окне «Гистограмма» ввести исходные данные:
    - В поле ввода «Входной интервал» ввести адрес области A14:O23;
    - В поле ввода «Интервал карманов» ввести адрес области E28:E33;
    - Включить радиокнопку «Выходной интервал» и ввести адрес верхней левой ячейки, начиная с которой будут выведены результаты расчетов для построения гистограммы, G26.
    - Включить переключатель «Вывод графика». Результат представлен на рис.4

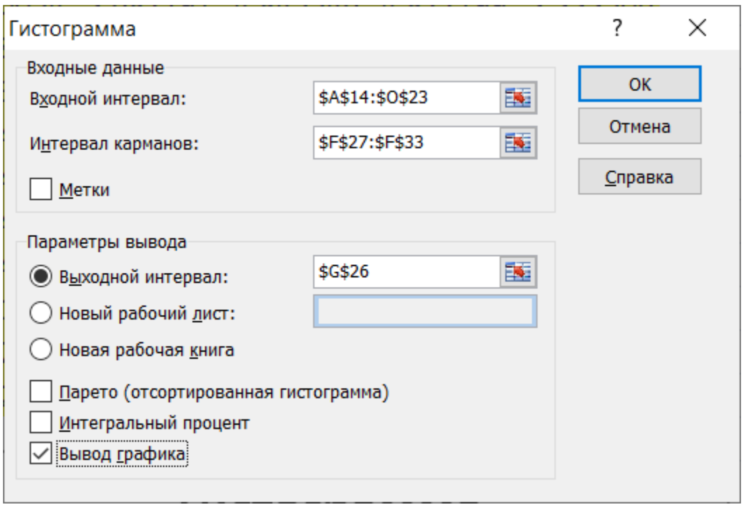



Рис.4. Исходные данные для ввода при построении гистограммы

- Нажать командную кнопку «ОК». Гистограмма построена. В области ячеек H27:H33 выведены соответственно количество сгенерированных случайных чисел, попавших по величине в соответствующий карман (частота).

3.5. Выбрать и рассчитать критерий согласия для проверки гипотезы о соответствии сгенерированных случайных чисел заданному закону распределения (вопрос качества построенного датчика случайных чисел).

- В таблице представлены типичные ситуации выбора критерия согласия. Для выполнения задания выберем третий вариант – критерий согласия Пирсона.

| Типовая ситуация выбора критерия согласия | Решение |
|---|---|
| Имеется несколько групп статистических данных, собранных в разных условиях. Вопрос – можно ли объединить эти данные в одну группу? | Критерий Смирнова (определяет степень принадлежности статистических данных одной генеральной совокупности) |
| Задан теоретически закон распределения случайной величины, а также заданы количественно параметры закона распределения | Критерий Колмогорова (используется функция распределения случайных величин с заданными теоретическими значениями количественных параметров закона распределения) |
| Имеются статистические данные об объекте | Критерий Пирсона или хи квадрат (используется функция плотности распределения случайных величин с рассчитанными на основе статистических данных значениями количественных параметров закона распределения) |

Критерий рассчитывается по формуле

$$\chi^2 = \sum_{i=1}^{k} \frac{(m_i - NP_i)^2}{NP_i},$$

где $m_i$ – количество значений случайной величины, попавших в i-й карман (частота в области ячеек H27:H33); $P_i$ – теоретическая вероятность попадания случайной величины в i-й карман, вычисленная по теоретическому распределению (область ячеек I27:I3, вероятности одинаковы для всех карманов 1/7, 7 – число карманов); k – количество карманов, на которые разбивается интервал изменения случайной величины; N=150 – количество сгенерированных случайных чисел.

- Для расчета критерия согласия Пирсона ввести в ячейку J27 формулу =(H27-$B$30*I27) ^2/($B$30*I27).
- Скопировать формулу из ячейки J27 в область ячеек J28:J33 (см. рис.1).
- В ячейку J34 ввести формулу =СУММ(J27:J33). Расчет значения критерия согласия Пирсона выполнен.
- Далее проверяется гипотеза о соответствии сгенерированных случайных чисел теоретическому распределению (в данном случае равномерному закону распределения в интервале [2,7]).

При $N \to \infty$ закон распределения величины $U=\chi 2$ зависит только от числа карманов и приближается к закону распределения $\chi 2$.

Вычисляется U и определяется число степеней свободы n =k – r – 1, где k– число карманов, r – количество явных параметром теоретического распределения (для равномерного распределения r=0, для экспоненциального распределения r=1). В задании n = 7 – 0– 1=6. Это строка входа в таблицу χ2- распределения.

Затем по таблицам χ2- распределения (рис.5) определяют P(χ2 γ ≥ χ2). Если эта вероятность превышает некоторой уровень значимости y, то считается, что гипотеза о соответствии сгенерированных случайных чисел теоретическому распределению не опровергается.


| n-1 | .995 | .990 | .975 | .950 | .900 | .750 | .500 | .250 | .100 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.00004 | 0.00016 | 0.00098 | 0.00393 | 0.01579 | 0.10153 | [illegible] | 1.32330 | 2.70554 |
| 2 | 0.01003 | 0.02010 | 0.05064 | 0.10259 | 0.21072 | 0.57536 | 1.38629 | 2.77259 | 4.60517 |
| 3 | 0.07172 | 0.11483 | 0.21580 | 0.35185 | 0.58437 | 1.21253 | 2.36597 | 4.10834 | 6.25139 |
| 4 | 0.20699 | 0.29711 | 0.48442 | 0.71072 | 1.06362 | 1.92256 | 3.35669 | 5.38527 | 7.77944 |
| 5 | 0.41174 | 0.55430 | 0.83121 | 1.14548 | 1.61031 | 2.67460 | 4.35146 | 6.62568 | 9.23636 |
| 6 | 0.67[illegible] | [illegible] | [illegible] | [illegible] | [illegible].20413 | 3.45460 | 5.34812 | 7.84080 | 10.64464 |
| 7 | 0.98926 | 1.23904 | 1.68987 | 2.16735 | 2.83311 | 4.25485 | 6.34581 | 9.03715 | 12.01704 |
| 8 | 1.34441 | 1.64650 | 2.17973 | 2.73264 | 3.48954 | 5.07064 | 7.34412 | 10.21885 | 13.36157 |
| 9 | 1.73493 | 2.08790 | 2.70039 | 3.32511 | 4.16816 | 5.89883 | 8.34283 | 11.38875 | 14.68366 |
| 10 | 2.15586 | 2.55821 | 3.24697 | 3.94030 | 4.86518 | 6.73720 | 9.34182 | 12.54886 | 15.98718 |

Рис. 5 Фрагмент таблицы χ2- распределения (URL: https://math.semestr.ru/group/xixi.php)

По таблице в строке с n=6 находим χ2=5,34, ближайшее большее к вычисленному значению U=χ2=4,09. Далее по столбцу вверх определяем вероятность y=0,5, с которой гипотеза о соответствии сгенерированных случайных чисел теоретическому распределению (в данном случае равномерному закону распределения в интервале [2,7]) не опровергается. Этот вывод свидетельствует о качестве построенного датчика случайных чисел с заданным законом определения. Данный датчик можно использовать для имитационного моделирования.