**Лабораторная работа 1**

**Моделирование случайных величин**

 1 Моделирование одной случайной величины

При моделировании случайных величин широко используется метод Монте-Карло. Этот метод объединяет группу методов численного решения различных математических задач при помощи моделирования случайных величин.

 Название “Монте-Карло” произошло от города Монте-Карло, известного своими казино, так как простейшим прибором для генерирования случайных чисел служит игральная рулетка. Возникновение метода связывают с именами Дж. Неймана, С. Улама, Н. Метрополиса, которые в 40-х годах прошлого века работали в Лос-Аламосе, США. Развитию данного метода способствовало бурное развитие ЭВМ.

 При решении задач методом Монте-Карло нужно:

1. Выбрать случайную величину (СВ) для решения конкретной задачи
2. Найти значение произвольной СВ.

 Необходимыми элементами для моделирования являются *датчики случайных чисел*, в качестве которых используются физические устройства*.* Для их реализации используют шумящие радиоэлектронные приборы. Они применяются довольно редко, так как нет возможности повторно воспроизвести выборочную последовательность для повторения расчетов.
 *Псевдослучайные числа:* пригодность случайных чисел определяется не процессом их получения, а тем, что они обладают интересующими нас свойствами независимых случайных величин.
Псевдослучайными числами называются такие, которые вычисляются по заданной формуле и могут быть использованы вместо случайных чисел при решении задач численным методом. Из сказанного следует, что оказываются тождественными те свойства случайных и псевдослучайных чисел, которые требуются для моделирования широкого круга задач. По отношению к этим задачам разницы между физически генерируемыми "случайными" числами и псевдослучайным числами практически нет.
 К преимуществам псевдослучайных чисел можно отнести:
 1. Небольшие затраты машинного времени для их получения.
 2.Возможность повторных воспроизведений последовательности чисел.
 3. Необходимость однократного тестирования алгоритмов вычисления псевдослучайных чисел.

 Для моделирования непрерывной случайной величины нужно учесть, что СВ *X*, удовлетворяющая уравнению *W(x) = α* имеет плотность вероятности *w(x).* Таким образом розыгрыш значения непрерывной СВ *X* с заданной функцией распределения (ФР) *W(x)* сводится к следующей процедуре:

#### нужно получить случайное число  и в качестве значения *x* взять , где введено обозначение для обратной функции по отношению к ФР *W(x).*

При моделировании СВ воспользуемся *методом обратной функции*, сущность которого заключается в том, что по ФР входной переменной путем розыгрыша находится значение этого признака. На рис.1 показаны графики плотности вероятности (ПВ - слева) и ФР (справа) некоторой величины, подчиняющейся нормальному закону.



Рис. 1 Графики плотности вероятности (слева) и функции распределения (справа)

 Как известно из теории вероятности, вся площадь под кривой ПВ равна единице, а ФР определяется как накопленная сумма (интеграл) от ПВ, поэтому на правой границе диапазона изменения переменной ФР асимптотически стремится к единице. На рис.1 показана ситуация, когда заштрихованная область под кривой ПВ равна приблизительно 0,75, т.е. вероятность в этом случае составляет 75%. На графике ФР (справа) эта же вероятность отложена на вертикальной оси, а определенное при этом значение аргумента совпадает с аргументом на горизонтальной оси левого графика. Ось ординат ФР (правого графика) имеет диапазон от 0 до 1 (100%), поскольку определяет собой вероятность любого события. Вследствие этого для моделирования надо взять любое случайное число из интервала [0,1] и отложить его на вертикальной оси, далее следует из этой точки провести горизонтальную линию до пересечения с кривой ФР, а из точки пересечения опустить перпендикуляр на горизонтальную ось. Таким образом, на этой оси получаем первое разыгранное значение некоторого признака. В этом заключается сущность метода обратной функции.

 Разыграв определенное количество случайных чисел (разумеется, уже не с помощью рулетки), получим то же количество значений моделируемого признака. Подобная процедура повторяется для всех разыгрываемых переменных, и в итоге получаем таблицу смоделированных значений, состоящую из *n* строк (количество розыгрышей) и  *m* столбцов (число переменных).

 Для нормально распределенных величин методом обратной функции значение СВ *Х*  рассчитывается по формуле

 ,

где  - функция Лапласа;  - среднее значение и среднеквадратичное отклонение;  *a* - разыгранная величина из интервала [0,1].

2 Моделирование системы случайных величин

 В ситуации, когда необходимо смоделировать систему случайных нормально распределенных величин (*X,Y*), следует поступать следующим образом.

 Пусть система случайных величин (*X,Y*) имеет двумерное нормальное распределение с параметрами  (*ρ* - коэффициент корреляции между *х* и *у*). В этом случае компонент *Х*  имеет нормальное распределение с параметрами , а компонент *Y* - нормальное распределение с такими параметрами:

* Математическое ожидание ;
* Стандартное отклонение .

 Математическое ожидание *Y* зависит от того, какое значение *х* принял компонент *Х*.

 Следовательно, моделирование заданной нормальной системы двух СВ сводится к имитации нормальной СВ *Х* с параметрами  и нормальной СВ *Y* с параметрами  и .

3 Работа на компьютере

3.1. *Excel*

 **Пример 1**.Разыграть методом Монте-Карло10 значений величин, которые распределены по нормальному закону со средним значением, равным 3, и дисперсией, равной 1.

Проверить, включен ли пакет Анализа данных. При выделении опции Данные должно появиться окно Анализа данных.



Если этого нет, то через Главная / Правой клавишей мыши (ПКМ)/ Настройка панели быстрого доступа/ Надстройки /Пакет анализа/ Перейти.



 Для решения воспользуемся табличным процессором Excel. Через *Сервис – Надстройки* включить Пакет анализа (отметить соответствующее окно) (рис.2).



Рисунок 2 - Включение пакета анализа

Через *Сервис –Анализ данных* перейти к окну "Анализ данных ", в котором выделить опцию "Генерация случайных чисел" (рис.3).



Рисунок 3 - Открытие окна "Анализ данных "

 В появившемся окне выполнить следующие операции (рис.4):

* В строке числа переменных поставить 1 (одна переменная разыгрывается).
* В строке числа случайных чисел поставить 10, так как по условиям примера необходимо получить 10 чисел.
* В следующей строке выбрать вид распределения: нормальное.
* В окне параметры указать данные в примере значения: среднее = 3, а стандартное отклонение (то же, что среднеквадратичное отклонение - СКО) равно квадратному корню из заданной дисперсии и составляет 1.
* В строке параметров вывода указать столбец, куда запишутся разыгранные значения.



Рисунок 4 - Окно "Генерация случайных чисел"

 Разыгранные 10 значений нормально распределенных величин показаны на рис.5.



Рисунок 5 - Разыгранные значения

 При необходимости можно уменьшить разрядность полученных данных, нажав соответствующую кнопку . Например, после четырехкратного нажатия выделенной кнопки разыгранные значения становятся равными (рис.6).



Рисунок 6 - Разыгранные значения с уменьшенной разрядностью

 **Пример 2.** Сформировать последовательность из 100 реализаций (*xi, yi*) системы (*X,Y*), имеющей двумерное нормальное распределение с параметрами = 5; = 1,25; =7; =2; ρ=0,8.

 Здесь необходимо смоделировать 2 величины:

1. СВ *Х* с параметрами = 5; = 1,25.

2. СВ *Y*  с параметрами 

 

Введем в таблицу Excel исходные данные (рис.7), где в диапазоне В1:В6 находятся параметры моделируемого распределения.



Рис.7 - Параметры распределения

В столбце Е смоделируем 100 значений СВ *Х* с параметрами = 5; = 1,25 с помощью операции "Генерация случайных величин".

В ячейку D2 вставим формулу расчета величины  затем скопируем ее до ячейки D101 (рис.8).



Рисунок 8 - Расчет СВ 

В ячейку F2 ввести формулу =НОРМОБР(СЛЧИС();D2;1,2), которая позволяет найти случайное число из нормального распределения, и скопировать ее содержимое до ячейки F101 (рис.9).



Рисунок 9 - Расчет СВ *Y*

На рис.10 показана диаграмма рассеяния между *X* и *Y* , по которой можно судить о наличии положительной корреляционной связи.



Рисунок 10 - Диаграмма рассеяния между *X* и *Y*

Расчетное значение коэффициента корреляции составляет



 0,83 , что весьма близко к начальному значению 0,8.

 Задание

В этой работе 4 варианта. Первый по списку студент выбирает вариант 1, второй - вариант 2 и т.д. Затем процедура повторяется.

1. Разыграть случайные величины, подчиняющиеся нормальному закону при следующих оценках среднего значения (СЗ) и среднеквадратичного отклонения (СКО):

 Вариант 1: СЗ =10; СКО = 2; объем выборки: 20.

 Вариант 2: СЗ =20; СКО = 3; объем выборки: 20.

 Вариант 3: СЗ =70; СКО = 5; объем выборки: 20.

 Вариант 4: СЗ =40; СКО = 4; объем выборки: 20.

 По разыгранным данным построить гистограмму.

 2.По разыгранным значениям проверить, в самом ли деле СЗ и СКО близки к задаваемым характеристикам. Для этого через опцию "Описательные статистики" для каждого варианта оценить параметры выборки. Эта опция находится в окне *Анализа данных* (рис.11).



Рисунок 11– Включение опции "Описательные статистики"

 На рис.12 показано окно описательных статистик



Рисунок 12 – Окно описательных статистик

 3. Провести моделирование случайных величин, распределенных по равномерному закону. Объем выборки - 20.

 Вариант 1: от 1 до 10.

 Вариант 2: от 10 до 20.

 Вариант 3: от 20 до 40.

 Вариант 4: от 40 до 50.

 Построить гистограммы.

 4. Для всех вариантов.

Сформировать последовательность из 20 реализаций (*xi, yi*) системы (*X,Y*), имеющей двумерное нормальное распределение с параметрами = 2; = 1; =5; =2; ρ=0,9.

Вопросы к защите работы:

1.В чем сущность метода Монте-Карло?

2. При розыгрыше случайных величин (СВ) методом обратной функции используется плотность вероятности или функция распределения?

3. В чем разница между плотностью вероятности и функцией распределения?

4. Можно ли использовать метод Монте-Карло для моделирования СВ, не подчиняющихся нормальному закону?

5. Что показывает диаграмма рассеяния?

6. Как будет выглядеть диаграмма рассеяния при коэффициенте корреляции, близком к 1; близком к (-1); близком к нулю?

7. В чем отличие корреляции от ковариации?