# Лекция 6: Статистическое моделирование

1. Общие сведения о статистическом моделировании.
2. Методы генерирования случайной величины.
3. Марковские процессы.

# 1 Общие сведения о статистическом моделировании

***Основным отличием*** статистических методов является построение генеральной совокупности:

***последовательность вариантов исходных данных, поступающих на вход системы, определяется не самим исследователем в зависимости от плана эксперимента, а генерируются с помощью датчика случайных чисел на компьютере.***

Далее реакция проверяется не на реальном объекте исследований, а на модели.

Таким образом, основное место при использовании статистических методов занимает компьютер.

***В качестве моделей***, на которых проверяется возможная реакция системы, применяются:

- ***вероятностные аналитические модели***

  (влияние случайных факторов учитывается с помощью задания вероятностных характеристик случайных процессов. Это приводит к усложнению вычислительной задачи и ограничивает применение данных моделей сравнительно простыми системами);

- ***имитационные модели***

  (введение случайных возмущений не вносит принципиальных усложнений, что делает их наиболее часто применяемыми).

Исследование сложных процессов и систем, подверженных случайным возмущениям, с помощью имитационного моделирования принято называть ***статистическим моделированием***.

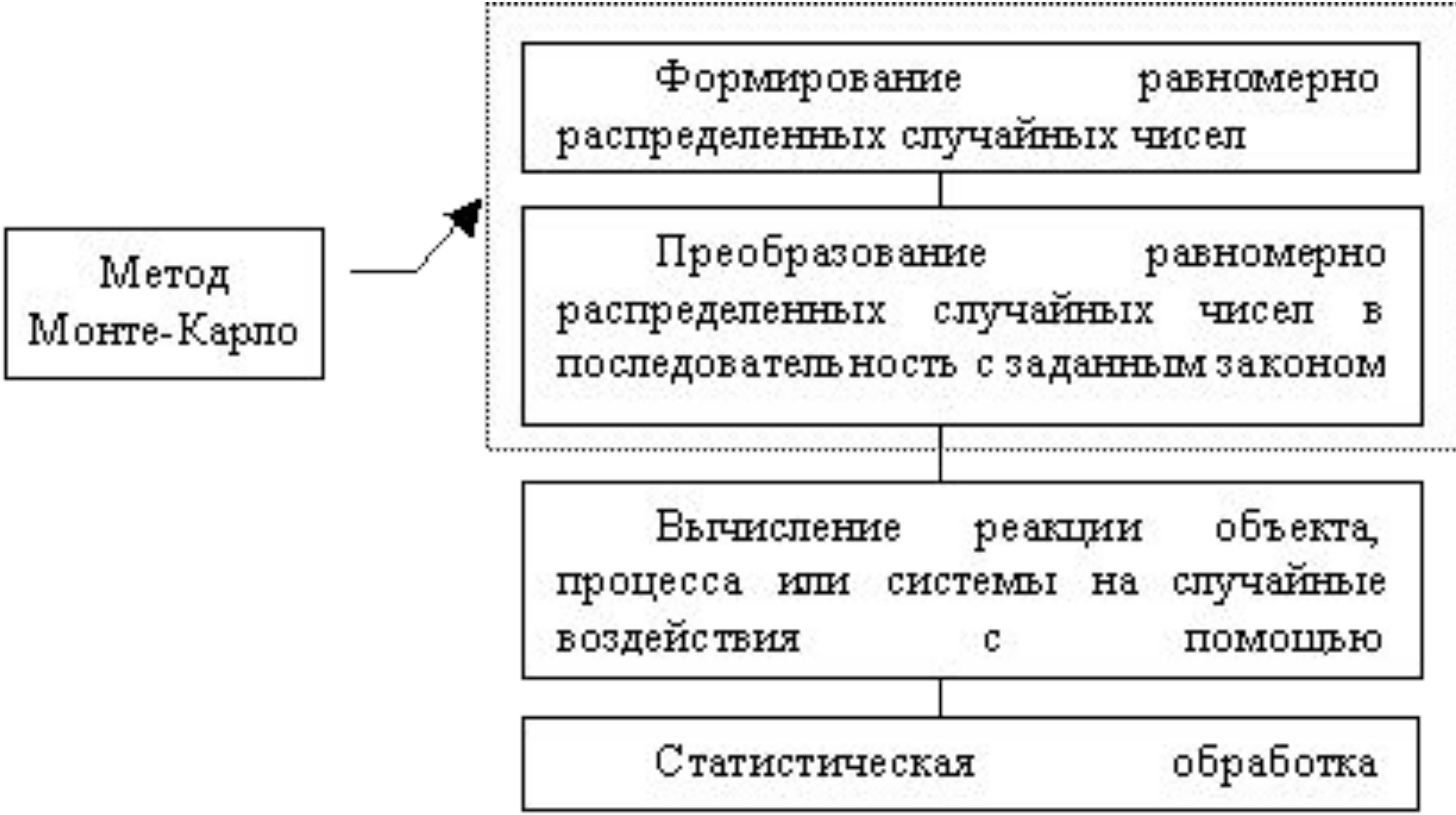
***Статистическая модель случайного процесса*** - это алгоритм, с помощью которого имитируют работу сложной системы, подверженной случайным возмущениям, причем полагается, что взаимодействие элементов системы носит вероятностный характер.

Оценка параметров модели осуществляется с помощью статистических методов: метода максимального правдоподобия, метода наименьших квадратов, метода моментов.

***Этапы методики статистического моделирования:***

1. Моделирование на компьютере псевдослучайных последовательностей с заданной корреляцией и законом распределения вероятностей (метод Монте-Карло), имитирующих случайные значения параметров при каждом испытании.
2. Преобразование полученных числовых последовательностей на имитационных математических моделях в генеральную совокупность.
3. Статистическая обработка результатов моделирования.

# *Обобщенный алгоритм метода статистических испытаний*



***Две области применения метода*** статистического моделирования:

- для изучения стохастических систем;
- для решения детерминированных задач.

В детерминированных системах предсказываемые значения могут быть вычислены точно, а в стохастических – лишь с некоторой долей вероятности.

Основная идея *для решения детерминированных задач*: замена детерминированной задачи эквивалентной схемой некоторой стохастической системы, выходные характеристики последней совпадают с результатом решения детерминированной задачи.

***Достоинства***:

- уменьшение погрешности с ростом числа испытаний (статистическая устойчивость результатов);
- возможность получения сведений о поведении реального объекта или процесса в произвольные моменты времени.

***Основная сложность - учет стохастических воздействий:***

- точность получаемых оценок зависит от размера совокупности случайных чисел, генерируемых системой, что приводит к *росту вычислительных затрат*, обусловленных созданием данной совокупности;
- качество получаемых на основе статистических моделей результатов, их точность и достоверность определяются исходными (базовыми) последовательностями случайных чисел. Это приводит к *необходимости разработки простых и экономичных способов формирования последовательностей случайных чисел* требуемого качества.

# 2 Методы генерирования случайной величины

***Методы***, используемые для получения случайных числовых последовательностей с заданными вероятностными характеристиками, ***различаются видом распределения*** случайной величины на заданном интервале $(a,b)$:

- ***равномерным*** $f(x)=\begin{cases}1/(b-a), & a\le x\le b\\ 0, & x<a, x>b\end{cases}$ ;

- ***нормальным***;

- ***распределением Бернулли*** (случайная величина принимает значение 1 с вероятностью $p$ и 0 с вероятностью 1=1-$p$ ;

- ***биномальным*** $P_n(m)=C_n^m p^m (1-p)^{n-m}$ ($n$ – общее число испытаний; $m$ – число успешных опытов);

- ***Пуассона*** (вероятность реализации случайной величины со значением m и параметром распределения $\lambda$:

$$P(m)=\frac{\lambda^m}{m!}\exp(-\lambda)$$

Численный метод, моделирующий случайные величины, равномерно распределенные на интервале (0,1), получил название "***метод статистических испытаний***" или "***метод Монте-Карло***".

Задачу ***моделирования случайных чисел*** с нормальным законом распределения решают в несколько ***этапов***:

1. Вначале имитируют равномерное распределение и получают последовательность псевдослучайных чисел, равномерно распределенных на интервале (0,1).
2. Затем, используя равномерно распределенную псевдослучайную величину, получают последовательность псевдослучайных чисел с нормальным законом распределения (чаще всего в нормированном виде, т.е. $M_x = 0,\ \sigma = 1$).

***Основные способа формирования последовательности нормально распределенных случайных величин***:

1. ***Прямое преобразование*** псевдослучайного числа $y$ являющегося реализацией случайной величины $Y$, равномерно распределенной на интервале [0,1], с помощью некоторой функции $W$ в число $x$, которое может рассматриваться как реализация случайной величины $X$, имеющей нормальный закон распределения.
2. **Отсеивание псевдослучайных чисел** из первоначальной последовательности $Y$ равномерно распределенной на интервале [0,1], таким образом, чтобы оставшиеся числа были распределены по нормальному закону.
3. ***Моделирование условий***, соответствующих центральной предельной теореме теории вероятности.

***Методы моделирования нормально распределенной случайной величины***:

- ***полярных координат***

  (первый способ получения. Вычисляет две независимые нормально распределенные случайные величины $x_1$ и $x_2$ с $M_x = 0$ и $\sigma = 1$ по двум заданным независимым равномерно распределенным случайным числам $y_1$ и $y_2$;

- ***метод, основанный на центральной предельной теореме***

  (третий способ получения. Основан на приближенном воспроизводстве условий, при которых справедлива центральная предельная теорема теории вероятности)

# 3 Марковские цепи

Под ***марковским процессом*** понимается случайный процесс, эволюция которого после любого заданного значения временного параметра $t$ не зависит от эволюции, предшествовавшей $t$, при условии, что значение процесса в этот момент фиксировано.

***«Будущее» процесса не зависит от «прошлого» при известном «настоящем».***

Понятие введено в 1907г А.А. Марковым.

Направление известно под названием ***теории цепей Маркова или «динамики вероятностей»***.

Основы общей теории марковских процессов с непрерывным временем были заложены Колмогоровым.

***По существу*** марковские цепи аналогичны методу динамического программирования.

***Отличие***: на каждом шаге учитывается вероятность попадания системы в то или иное состояние. В связи с этим этот метод называют ***стохастическим динамическим программированием***.

***Область применения***: исследование операций и теория принятия оптимальных решений.

***Основаны*** на понятии случайной функции и относятся к частным случаям случайных процессов.

Если аргументом случайной функции является время или какой-то другой аргумент, то такой ***процесс*** называют ***случайным***.

Случайные процессы могут быть *с дискретным* или *непрерывным состоянием или временем*.

***Важное свойство случайных процессов - вероятностная связь между состояниями случайного процесса***.

(Если в случайном процессе вероятность перехода системы в каждое последующее состояние зависит только от предыдущего состояния, то такой процесс называется ***процессом без последействия – сложная цепь)***. Обычно применяют так называемый процесс укрупнения состояний путем математических преобразований, объединяя предшествующие состояния в одно.

Марковский процесс удобно задавать графом переходов из состояния в состояние.

Два варианта описания марковских процессов – **с дискретным и непрерывным временем**:

*В первом случае переход* из одного состояния в другое происходит *в заранее известные моменты времени* – такты (1, 2, 3, 4, …). Переход осуществляется на каждом такте, то есть исследователя интересует только последовательность состояний, которую проходит случайный процесс в своем развитии, и не интересует, когда конкретно происходил каждый из переходов.

*Во втором случае* исследователя интересует и цепочка меняющих друг друга состояний, и моменты времени, в которые происходили такие переходы.

Если вероятность перехода не зависит от времени, то ***марковскую цепь*** называют ***однородной***.

Рассмотрим *численный пример*, в котором имитируется стрельба из пушки по цели.

Определим следующие три состояния: $S_0$ — цель не повреждена; $S_1$ — цель повреждена; $S_2$ — цель разрушена.

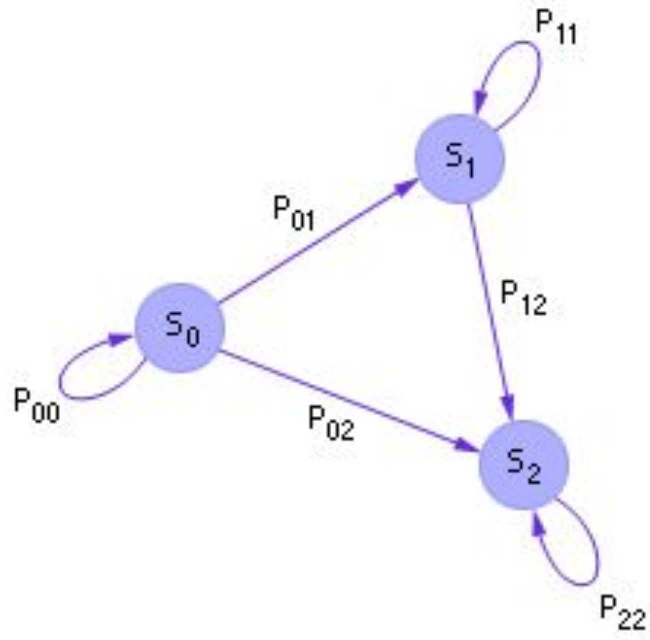
Таблица 2 – Вектор начальных вероятностей

| | $S_0$ | $S_1$ | $S_2$ |
|---|---|---|---|
| $P_0$ | 0,8 | 0,2 | 0 |

Таблица 3 – Матрица вероятностей перехода дискретного марковского процесса

| | В $S_0$ | В $S_1$ | В $S_2$ | Сумма вероятностей переходов |
|---|---|---|---|---|
| Из $S_0$ | 0.45 | 0.40 | 0.15 | 0.45+0.40+0.15=1 |
| Из $S_1$ | 0 | 0.45 | 0.55 | 0+0.45+0.55=1 |
| Из $S_2$ | 0 | 0 | 1 | 0+0+1=1 |

## *Представление процесса в виде марковской цепи*




Проимитируем, используя таблицу случайных чисел, процесс стрельбы.

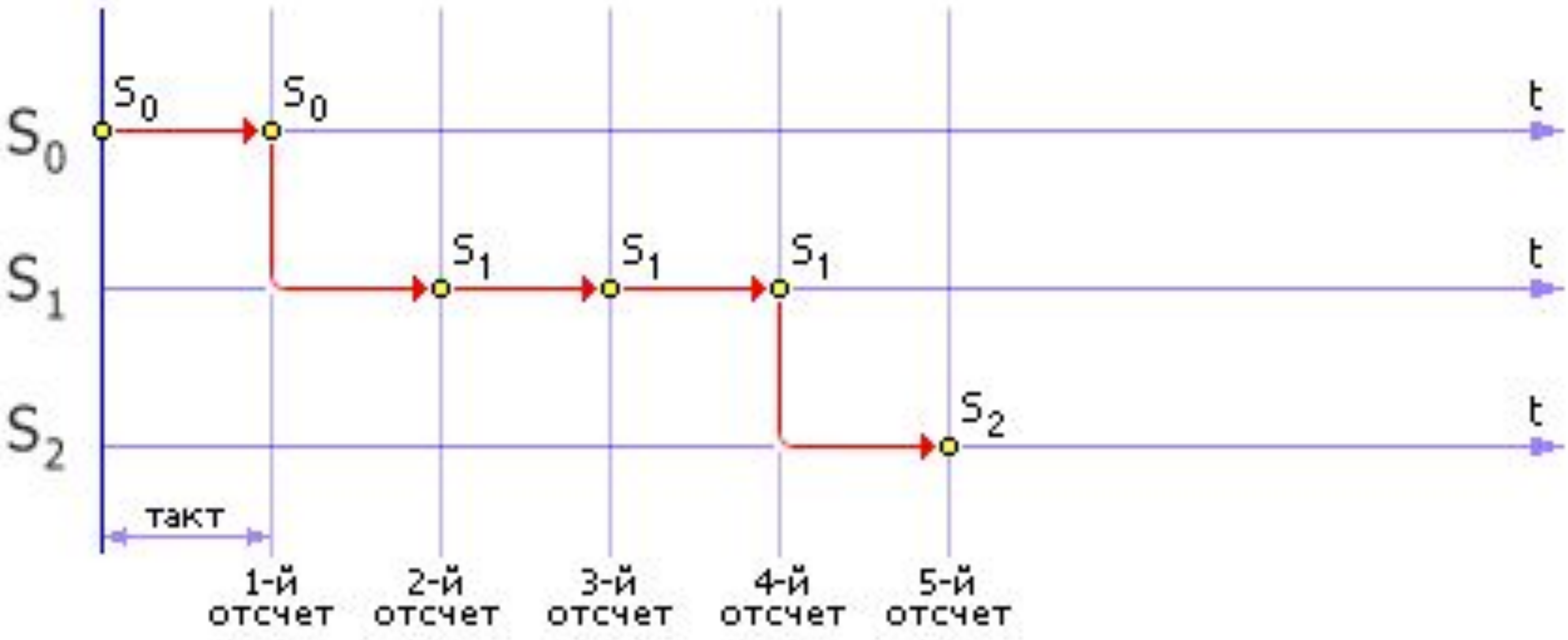
Пусть начальное состояние будет $S_0$. Возьмем последовательность из таблицы случайных чисел: 0.31, 0.53, 0.23, 0.42, 0.63, 0.21, ….

Проимитируем, используя таблицу случайных чисел, процесс стрельбы.

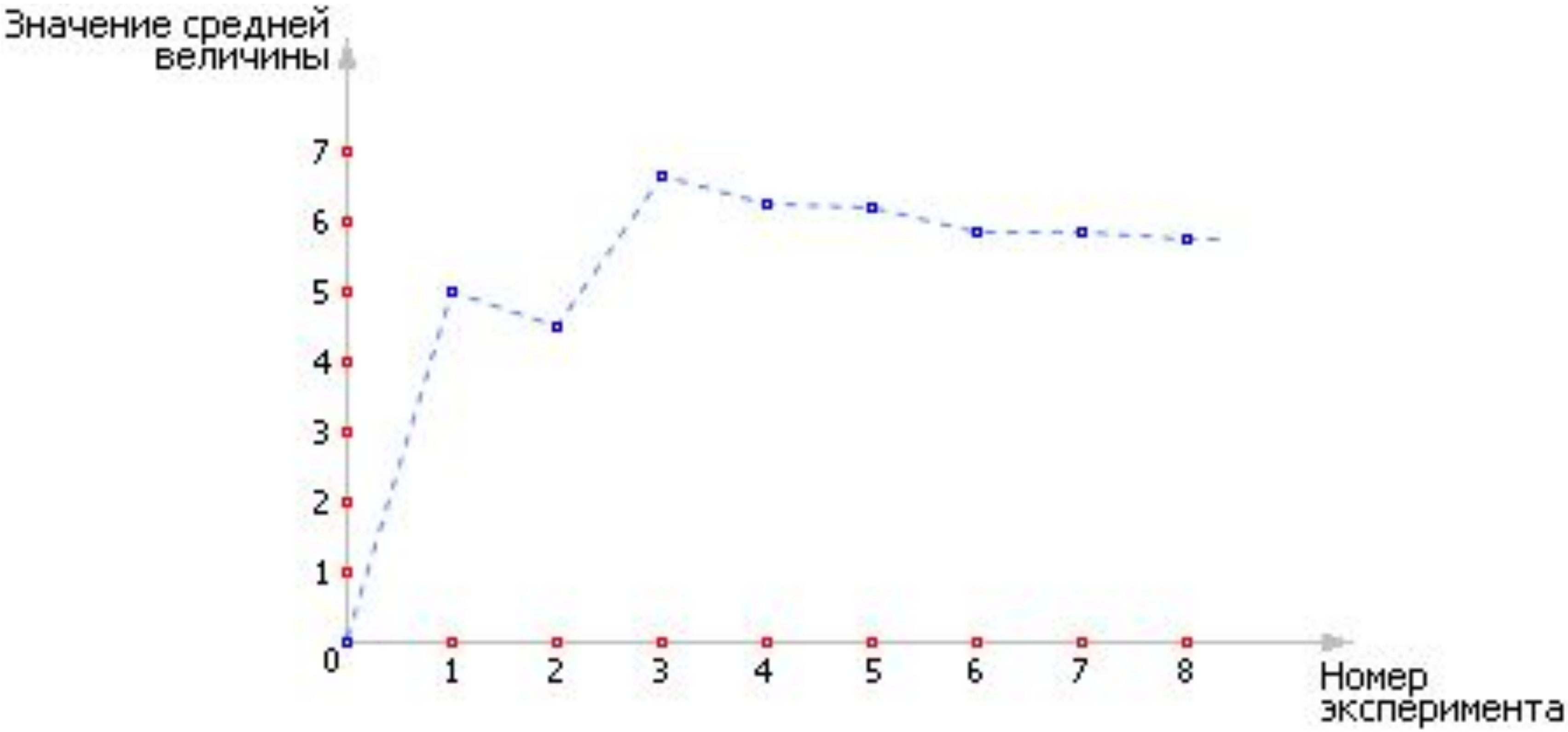
Пусть начальное состояние будет $S_0$. Возьмем последовательность из таблицы случайных чисел: 0.31, 0.53, 0.23, 0.42, 0.63, 0.21, ….

**0.31**: цель находится в состоянии $S_0$ и остается в состоянии $S_0$, так как $0 < \mathbf{0.31} < 0.45$; **0.53**: цель находится в состоянии $S_0$ и переходит в состояние $S_1$, так как $0.45 < \mathbf{0.53} < 0.45 + 0.40$;
**0.23**: цель находится в состоянии $S_1$ и остается в состоянии $S_1$, так как $0 < \mathbf{0.23} < 0.45$;
**0.42**: цель находится в состоянии $S_1$ и остается в состоянии $S_1$, так как $0 < \mathbf{0.42} < 0.45$;
**0.63**: цель находится в состоянии $S_1$ и переходит в состояние $S_2$, так как $0.45 < \mathbf{0.63} < 0.45 + 0.55$.
Так как достигнуто состояние $S_2$ (далее цель переходит из $S_2$ в состояние $S_2$ с вероятностью 1), то цель поражена. Для этого в данном эксперименте потребовалось ***5 снарядов***.

# *Временная диаграмма, получаемая во время процесса моделирования*



Повторяя циклы моделирования случайных процессов, получаем статистику:




Ряд сходится к некоторой величине, которая и является ответом. В нашем случае – это 6. Именно столько снарядов в среднем рекомендуется иметь в боевом запасе пушки для уничтожения цели при таких вероятностях попаданий.