

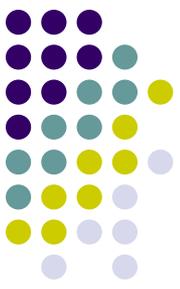
Тема: Численное моделирование. Метод Монте-Карло.

Содержание лекции:

1. История
 - 1.1 Алгоритм Буффона для определения числа Пи
 - 1.2 Связь стохастических процессов и дифференциальных уравнений
 - 1.3 Рождение метода Монте-Карло в Лос-Аламосе
 - 1.4 Дальнейшее развитие и современность
2. Интегрирование методом Монте-Карло
 - 2.1 Обычный алгоритм Монте-Карло интегрирования
 - 2.2 Использование выборки по значимости
3. Оптимизация
4. Применение в физике
 - 4.1 Алгоритм Метрополиса
 - 4.2 Динамическая (кинетическая) формулировка
 - 4.3 Прямое моделирование методом Монте-Карло
 - 4.4 Квантовый метод Монте-Карло



Метод Монте-Карло – это ...



Мéтод Мóнте-Кáрло (методы Монте-Карло) — общее название группы численных методов, основанных на получении большого числа реализаций стохастического (случайного) процесса, который формируется таким образом, чтобы его вероятностные характеристики совпадали с аналогичными величинами решаемой задачи.

Используется для решения задач в областях физики, математики, экономики, оптимизации, теории управления и др.

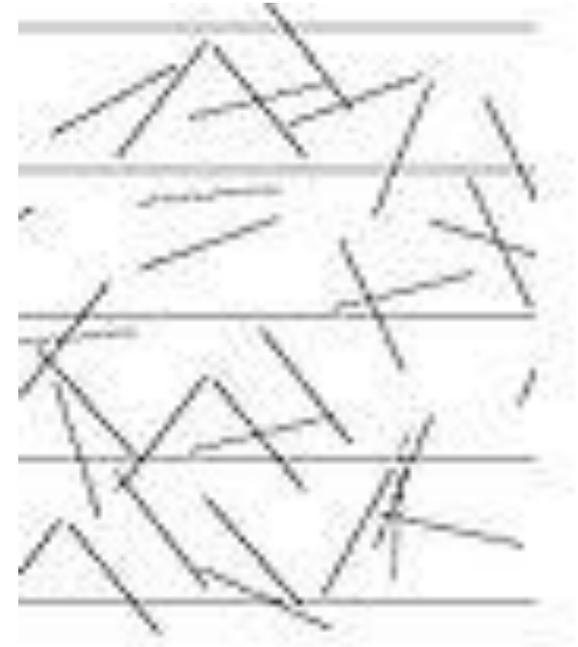
Алгоритм Буффона для определения числа Пи



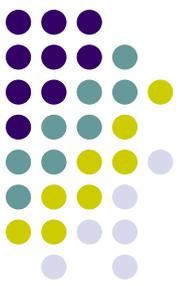
Случайные величины использовались для решения различных прикладных задач достаточно давно.

Примером может служить способ определения числа Пи, который был предложен Буффоном еще в 1777 году.

Суть метода была в бросании иглы длиной L на плоскость, расчерченную параллельными прямыми, расположенными на расстоянии r друг от друга.



Алгоритм Буффона для определения числа Пи



Вероятность (как видно из дальнейшего контекста, речь идёт не о вероятности, а о математическом ожидании количества пересечений за один опыт; вероятностью это становится лишь при условии, что $r > L$) того, что отрезок пересечет прямую, связана с числом Пи:

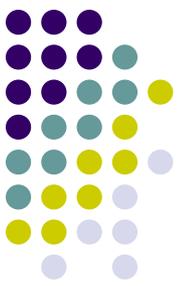
$$p = \int_0^{\pi} \int_0^{l \sin \theta} \frac{1}{r\pi} dA d\theta$$

где

A — расстояние от начала иглы до ближайшей к ней прямой;

θ — угол иглы относительно прямых.

Алгоритм Буффона для определения числа Пи



Этот интеграл просто взять:

$$p = \frac{2L}{r\pi}$$

(при условии, что $r > L$), поэтому подсчитав долю отрезков, пересекающих прямые, можно приближенно определить это число. При увеличении количества попыток точность получаемого результата будет увеличиваться.

Алгоритм Буффона для определения числа Пи



В 1864 году капитан Фокс, выздоравливая после ранения, чтобы как-то занять себя, реализовал эксперимент по бросанию иглы. Результаты представлены в следующей таблице:

	Число бросаний	Число пересечений	Длина иглы	Расстояние между прямыми	Вращение	Зн-ие Пи
Первая попытка	500	236	3	4	отсутствует	3.1780
Вторая попытка	530	253	3	4	присутствует	3.1423
Третья попытка	590	939	5	2	присутствует	3.1416

Комментарии:

- Вращение плоскости применялось (и как показывают результаты — успешно) для того, чтобы уменьшить систематическую ошибку.
- В третьей попытке длина иглы была больше расстояния между линиями, что позволило не увеличивая числа бросаний эффективно увеличить число событий и повысить точность.



Создание математического аппарата стохастических методов началось в конце 19-го века. В 1899 году лорд Релей показал, что одномерное случайное блуждание на бесконечной решётке может давать приближенное решение параболического дифференциального уравнения.

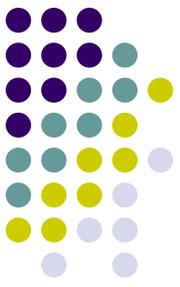
Колмогоров в 1931 году дал большой толчок развитию стохастических подходов к решению различных математических задач, поскольку он сумел доказать, что цепи Маркова связаны с некоторыми интегро-дифференциальными уравнениями.



В 1933 году Петровский показал, что случайное блуждание, образующее Марковскую цепь, асимптотически связано с решением эллиптического дифференциального уравнения в частных производных.

После этих открытий стало понятно, что стохастические процессы можно описывать дифференциальными уравнениями и, соответственно, исследовать при помощи хорошо на тот момент разработанных математических методов решения этих уравнений.

Рождение метода Монте-Карло в Лос-Аламосе



Сначала Энрико Ферми в 1930-х годах в Италии, а затем Джон фон Нейман и Станислав Улам в 1940-х в Лос-Аламосе предположили, что можно использовать связь между стохастическими процессами и дифференциальными уравнениями «в обратную сторону». Они предложили использовать стохастический подход для аппроксимации многомерных интегралов в уравнениях переноса, возникших в связи с задачей о движении нейтрона в изотропной среде.

Рождение метода Монте-Карло в Лос-Аламосе



Идея была развита Станиславом Уламом, который, по иронии судьбы, также как и Фокс боролся с вынужденным безделием во время выздоровления после болезни, и, раскладывая пасьянсы, задался вопросом, какова вероятность того, что пасьянс «сложится».

Ему в голову пришла идея, что вместо того, чтобы использовать обычные для подобных задач соображения комбинаторики, можно просто поставить «эксперимент» большое число раз и, таким образом, подсчитав число удачных исходов, оценить их вероятность. Он же предложил использовать компьютеры для расчётов методом Монте-Карло.

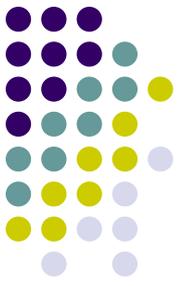
Рождение метода Монте-Карло в Лос-Аламосе



Появление первых электронных компьютеров, которые могли с большой скоростью генерировать псевдослучайные числа, резко расширило круг задач, для решения которых стохастический подход оказался более эффективным, чем другие математические методы.

После этого произошёл большой прорыв и метод Монте-Карло применялся во многих задачах, однако его использование не всегда было оправдано из-за большого количества вычислений, необходимых для получения ответа с заданной точностью.

Рождение метода Монте-Карло в Лос-Аламосе



Годом рождения метода Монте-Карло считается 1949 год, когда в свет выходит статья Метрополиса и Улама «Метод Монте-Карло».

Название метода происходит от названия города в княжестве Монако, широко известного своими многочисленными казино, поскольку именно рулетка является одним из самых широко известных генераторов случайных чисел.

Станислав Улам пишет в своей автобиографии «Приключения математика», что название было предложено Метрополисом в честь его дяди, который был азартным игроком.

Дальнейшее развитие и современность

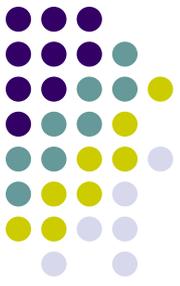


В 1950-х годах метод использовался для расчётов при разработке водородной бомбы. Основные заслуги в развитии метода в это время принадлежат сотрудникам лабораторий ВВС США и корпорации RAND.

В 1970-х годах в новой области математики — теории вычислительной сложности было показано, что существует класс задач, сложность (количество вычислений, необходимых для получения точного ответа) которых растёт с размерностью задачи экспоненциально. Иногда можно, пожертвовав точностью, найти алгоритм, сложность которого растёт медленнее, но есть большое количество задач, для которого этого нельзя сделать (например, задача определения объёма выпуклого тела в n -мерном евклидовом пространстве) и метод Монте-Карло является единственной возможностью для получения достаточно точного ответа за приемлемое время.

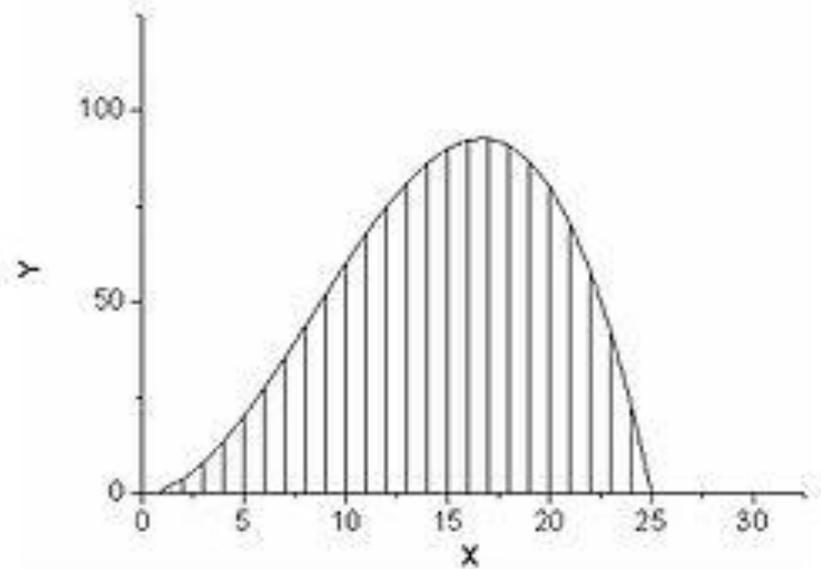
В настоящее время основные усилия исследователей направлены на создание эффективных Монте-Карло алгоритмов различных физических, химических и социальных процессов для параллельных вычислительных систем.

Интегрирование методом Монте-Карло



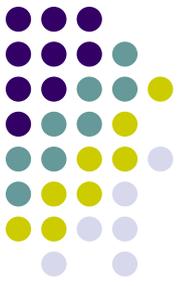
Предположим, необходимо взять интеграл от некоторой функции. Воспользуемся неформальным геометрическим описанием интеграла и будем понимать его как площадь под графиком этой функции.

Для определения этой площади можно воспользоваться одним из обычных численных методов интегрирования: разбить отрезок на подотрезки, подсчитать площадь под графиком функции на каждом из них и сложить.



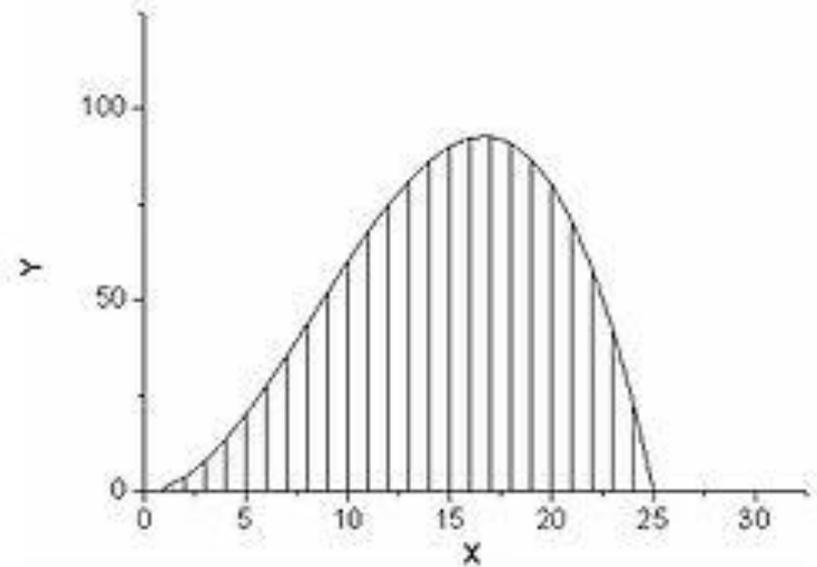
Численное интегрирование функции
детерминистическим методом

Интегрирование методом Монте-Карло



Предположим, что для функции, представленной на рисунке, достаточно разбиения на 25 отрезков и, следовательно, вычисления 25 значений функции. Представим теперь, мы имеем дело с n -мерной функцией. Тогда нам необходимо 25^n отрезков и столько же вычислений значения функции. При размерности функции больше 10 задача становится огромной.

Поскольку пространства большой размерности встречаются, в частности, в задачах теории струн, а также многих других физических задачах, где имеются системы со многими степенями свободы, необходимо иметь метод решения, вычислительная сложность которого бы не столь сильно зависела от размерности. Именно таким свойством обладает **метод Монте-Карло**.



Численное интегрирование функции
детерминистическим методом

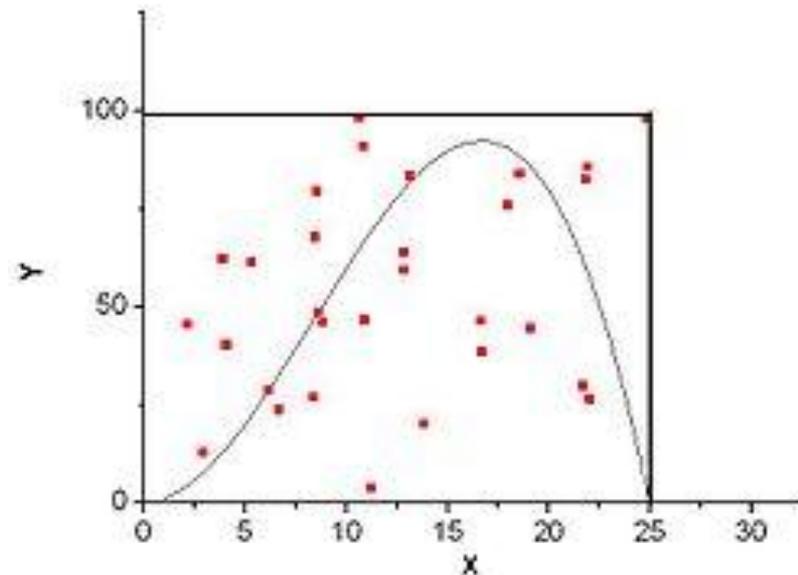
Обычный алгоритм Монте-Карло интегрирования



Для определения площади под графиком функции можно использовать следующий стохастический алгоритм:

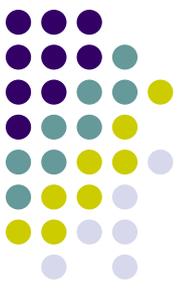
- ограничим функцию прямоугольником (n -мерным параллелепипедом в случае многих измерений), площадь которого S_{par} можно легко вычислить;
- «набросаем» в этот прямоугольник (параллелепипед) некоторое количество точек (N штук), координаты которых будем выбирать случайным образом;
- определим число точек (K штук), которые попадут под график функции;
- площадь области, ограниченной функцией и осями координат, S даётся следующим выражением

$$S = S_{par} \frac{K}{N}$$



численное интегрирование функции
методом Монте-Карло

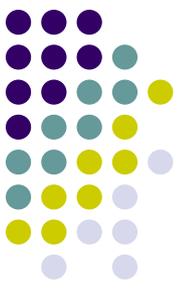
Обычный алгоритм Монте-Карло интегрирования



Для малого числа измерений интегрируемой функции производительность Монте-Карло интегрирования гораздо ниже, чем производительность детерминистических методов.

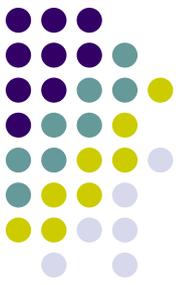
Тем не менее, в некоторых случаях, когда функция задана неявно, а необходимо определить область, заданную в виде сложных неравенств, стохастический метод может оказаться более предпочтительным.

Использование выборки по значимости



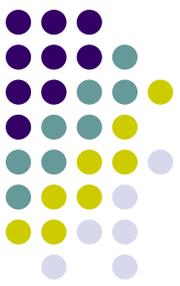
Очевидно, что точность вычислений можно увеличить, если область, ограничивающая искомую функцию, будет максимально к ней приближена. Для этого необходимо использовать случайные с распределением, форма которого максимально близка к форме интегрируемой функции. На этом основан один из методов улучшения сходимости в вычислениях методом Монте-Карло: выборка по значимости.

Применение в физике



Компьютерное моделирование играет в современной физике важную роль и метод **Монте-Карло** является одним из самых распространённых во многих областях от квантовой физики до физики твёрдого тела, физики плазмы и астрофизики.

Алгоритм Метрополиса

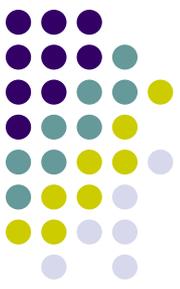


Традиционно метод Монте-Карло применялся для определения различных физических параметров систем, находящихся в состоянии термодинамического равновесия. Предположим имеется набор $W(S)$ возможных состояний физической системы S . Для определения среднего значения \bar{A} некоторой величины A необходимо рассчитать

$$\bar{A} = \sum_S A(S)P(S)$$

где суммирование производится по всем состояниям S из $W(S)$, $P(S)$ — вероятность состояния S .

Прямое моделирование методом Монте-Карло

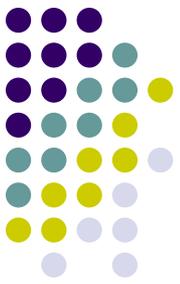


Прямое моделирование методом Монте-Карло какого-либо физического процесса подразумевает моделирование поведения отдельных элементарных частей физической системы.

По сути это прямое моделирование близко к решению задачи из первых принципов, однако обычно для ускорения расчётов допускается применение каких-либо физических приближений.

Примером могут служить расчёты различных процессов методом молекулярной динамики: с одной стороны система описывается через поведение её элементарных составных частей, с другой стороны, используемый потенциал взаимодействия зачастую является эмпирическим.

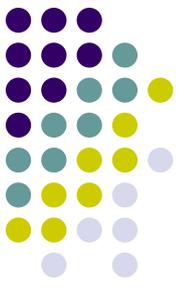
Прямое моделирование методом Монте-Карло



Примеры прямого моделирования методом Монте-Карло:

- **Моделирование облучения твёрдых тел ионами в приближении бинарных столкновений.**
- **Прямое Монте-Карло моделирование разреженных газов.**
- **Большинство кинетических Монте-Карло моделей относятся к числу прямых (в частности, исследование молекулярно-пучковой эпитаксии).**

Квантовый метод Монте-Карло



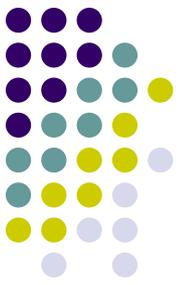
Квантовый метод Монте-Карло широко применяется для исследования сложных молекул и твёрдых тел. Это название объединяет несколько разных методов.

Первый из них это **Вариационный метод Монте-Карло**, который по сути является численным интегрированием многомерных интегралов, возникающих при решении уравнения Шрёдингера.

Для решения задачи, в которой участвует 1000 электронов, необходимо взятие 3000-мерных интегралов, и при решении таких задач метод Монте-Карло имеет огромное преимущество в производительности по сравнению с другими численными методами интегрирования.

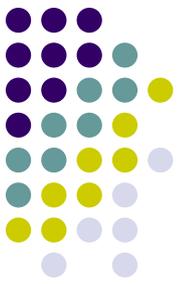
Другая разновидность метода Монте-Карло — это **диффузионный метод Монте-Карло**.





Метод Монте–Карло — это численный метод решения математических задач при помощи моделирования случайных величин.

Происхождение метода Монте-Карло



Датой рождения метода Монте-Карло принято считать 1949 г., когда появилась статья под названием «The Monte Carlo method». Создателями этого метода считают американских математиков Дж. Неймана и С. Улама.

Любопытно, что теоретическая основа метода была известна давно.

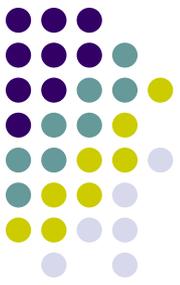
Более того, некоторые задачи статистики рассчитывались иногда с помощью случайных выборок, т. е. фактически методом Монте-Карло.

Однако до появления компьютеров этот метод не мог найти сколько-нибудь широкого применения, ибо моделировать случайные величины вручную — очень трудоемкая работа.

Таким образом, возникновение метода Монте-Карло как весьма универсального численного метода стало возможным только благодаря появлению компьютеров.

Само название «Монте-Карло» происходит от города Монте-Карло в княжестве Монако, знаменитого своими казино. Дело в том, что одним из простейших механических приборов для получения случайных величин является рулетка.

Происхождение метода Монте-Карло



Для того чтобы понять, о чем пойдет речь, рассмотрим простой пример. Предположим, что нам нужно вычислить площадь плоской фигуры S . Это может быть произвольная фигура с криволинейной границей, заданная графически или аналитически, связная или состоящая из нескольких кусков. Пусть это будет фигура, изображенная на рис.1, и предположим, что она вся расположена внутри единичного квадрата.

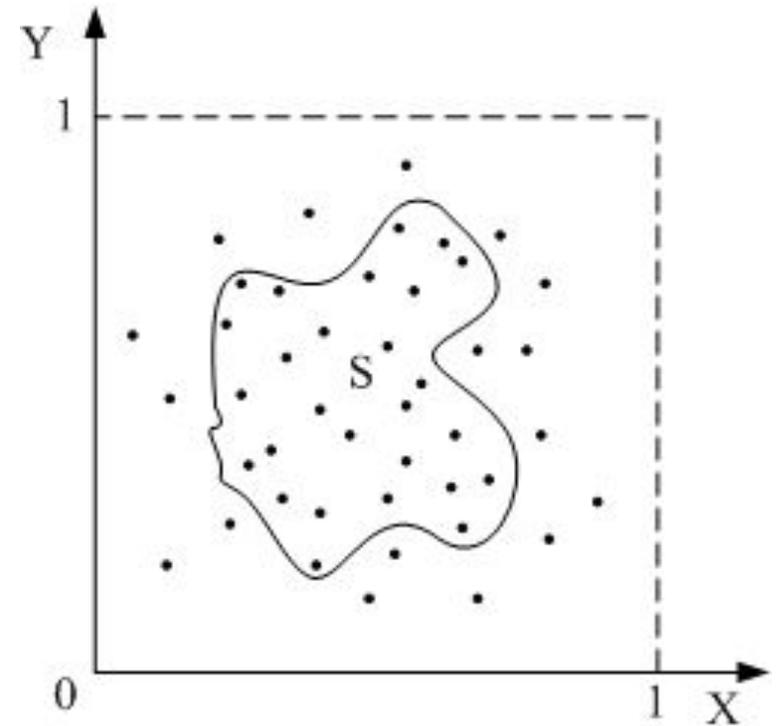
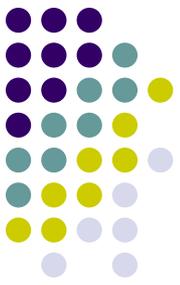


Рис.1 Вычисление площади фигуры

Квантовый метод Монте-Карло



Выберем внутри квадрата множество случайных точек N .

Обозначим через N' число точек, попавших при этом внутрь S . Геометрически очевидно, что площадь S приближенно равна отношению N'/N .

Чем больше N , тем больше точность этой оценки.

Например, выберем случайным образом 44 точки.

Пусть 13 из них оказались внутри S .

Отношение N'/N равно $13/44=0,30$, в то время как истинная площадь S равна $0,35$.

Две особенности метода Монте-Карло



Первая особенность метода - простая структура вычислительного алгоритма.

Как правило, составляется программа для осуществления одного случайного испытания (надо выбрать случайную точку в квадрате и проверить, принадлежит ли она S).

Затем это испытание повторяется N раз, причем каждый опыт не зависит от всех остальных, и результаты всех опытов усредняются.

Поэтому иногда метод Монте-Карло называют *методом статистических испытаний*.

Две особенности метода Монте-Карло



Вторая особенность метода - погрешность вычислений, как правило, пропорциональна $\sqrt{D/N}$, где D - некоторая постоянная. N - число испытаний.

Отсюда видно, что для того, чтобы уменьшить погрешность в 10 раз (иначе говоря, чтобы получить в ответе еще один верный десятичный знак), нужно увеличить N (т. е. объем работы) в 100 раз.

Однако одну и ту же задачу можно решать различными вариантами метода Монте-Карло, которым отвечают различные значения D .

Во многих задачах удастся значительно увеличить точность, выбрав способ расчета, которому соответствует значительно меньшее значение D .

Задачи, решаемые методом Монте-Карло



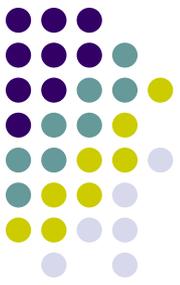
Во-первых, метод Монте-Карло позволяет моделировать любой процесс, на протекание которого влияют случайные факторы.

Во-вторых, для многих математических задач, не связанных с какими-либо случайностями, можно искусственно придумать вероятностную модель (и даже не одну), позволяющую решать эти задачи (например, для рис.1).

Таким образом, можно говорить о методе Монте-Карло как об универсальном методе решения математических задач.

Особенно интересно, что в некоторых случаях выгодно отказаться от моделирования истинного случайного процесса и вместо этого использовать искусственную модель.

Задачи, решаемые методом Монте-Карло



Вернемся к примеру (рис.1).

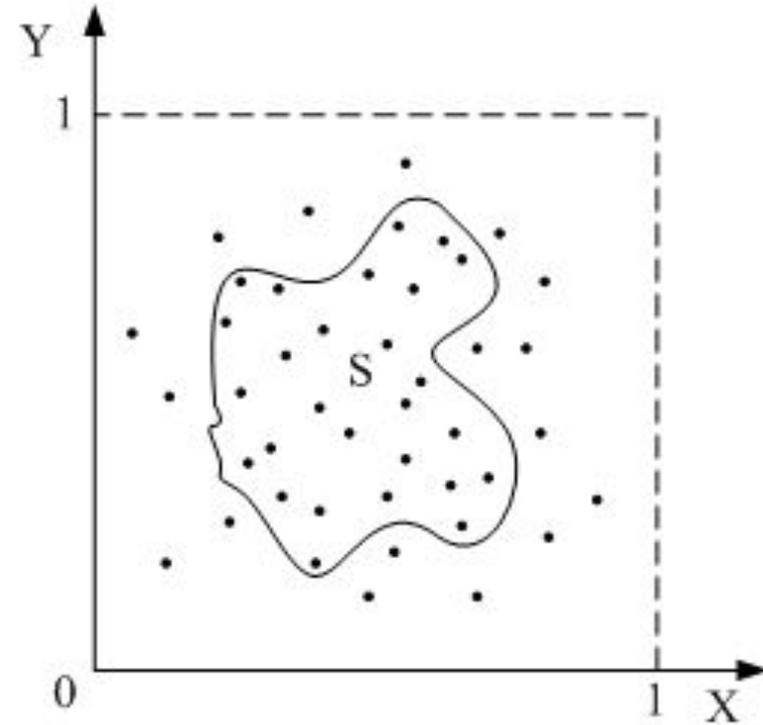
Для расчета нам нужно было выбирать случайные точки в единичном квадрате. Как это сделать физически?

Представим такой эксперимент.

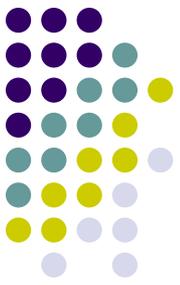
Рис. 1 с фигурой S и квадратом повешен на стену в качестве мишени.

Стрелок, находящийся на некотором расстоянии от стены, стреляет N раз, целясь в центр квадрата. Конечно, все пули не будут ложиться точно в центр: они пробьют на мишени N случайных точек.

Можно ли по этим точкам оценить площадь S ?



Задачи, решаемые методом Монте-Карло



Результат такого опыта изображен на рис. 2.

В этом опыте $N = 44$, $N' = 27$ и $N'/N = 0,61$, что превышает истинное значение площади (0,35) почти в 2 раза.

Впрочем, и так ясно, что при очень высокой квалификации стрелка результат опыта будет очень плохим, так как почти все пули будут ложиться вблизи центра и попадут в S .

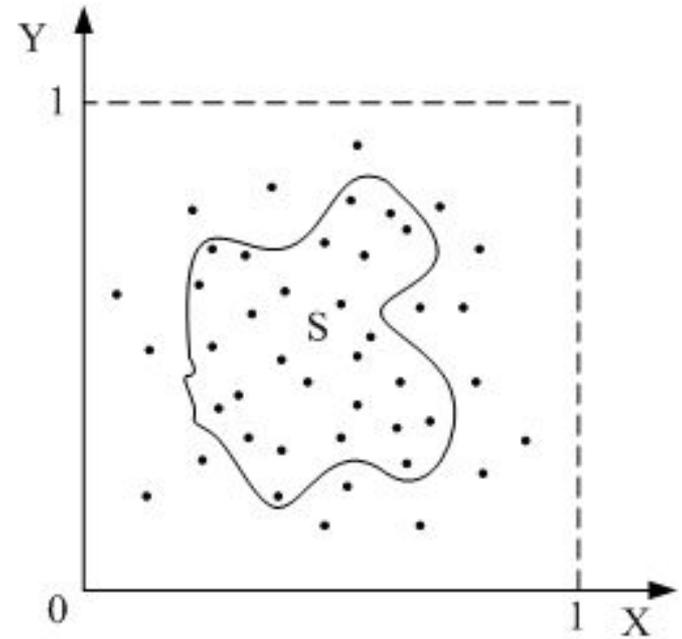
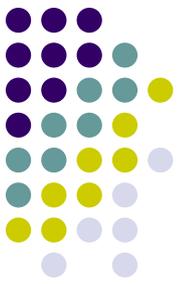


Рис. 2. Неравномерное распределение точек

Задачи, решаемые методом Монте-Карло



Нетрудно понять, что наш метод вычисления площади будет справедлив только тогда, когда случайные точки будут не «просто случайными», а еще и «равномерно разбросанными» по всему квадрату. Чтобы придать этим словам точный смысл, необходимо познакомиться с определением случайных величин и некоторыми их свойствами.

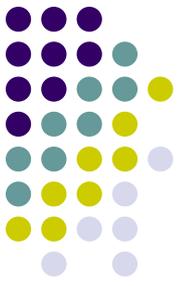
Вообще, при математическом описании модели любого сколько-нибудь сложного процесса или явления всегда возникает задача учета случайности.

В течение долгого периода исследователи ограничивались рассмотрением разного рода игр, особенно игр в кости, поскольку их изучение позволяет ограничиваться простыми и прозрачными математическими моделями.

Однако следует заметить, что многие отлично понимали то, что позднее (в 1657 году) было прекрасно сформулировано Х. Гюйгенсом: *"...Я полагаю, что при внимательном изучении предмета читатель заметит, что имеет дело не только с игрой, но здесь закладываются основы очень интересной и глубокой теории"*.

И эта теория, изучающая количественные закономерности случайных явлений, была создана и названа теорией вероятностей.

Случайные величины



За последние десятилетия неизмеримо выросла роль, которую играет теория вероятностей в современном естествознании. После того, как молекулярные представления о строении вещества получили всеобщее признание, стало неизбежным широкое использование теории вероятностей в физике и химии.

Заметим, что с точки зрения молекулярной физики каждое вещество состоит из огромного числа малых частиц, находящихся в непрерывном движении и в процессе этого движения взаимодействующих друг с другом. Естественно, что при таких условиях обычные для физических теорий методы математических исследований становятся бессмысленными.

Теория вероятностей развилась из потребностей практики: в абстрактной форме она отражает закономерности, присущие случайным событиям массового характера. Эти закономерности играют исключительно важную роль в физике, химии, различных технических дисциплинах, экономике, военном деле и так далее.

Случайные величины

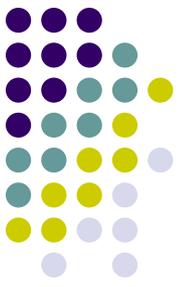


До некоторого времени теория вероятностей представляла собой еще не сложившуюся математическую науку, в которой основные понятия были недостаточно корректно определены. Сказанное относится к так называемой классической теории вероятностей.

В 1900 году на II Международном математическом конгрессе в Париже немецкий ученый Д. Гильберт сделал доклад, в котором указал 23 важнейшие проблемы, стоящие перед математикой.

Одна из них — аксиоматическое построение теории вероятностей. В настоящее время эта проблема успешно решена. Наиболее удачный подход к построению аксиоматической теории вероятностей предложил А. Н. Колмогоров. Этот подход тесно связывает теорию вероятностей с современной теорией меры, а также теорией множеств.

Некоторые определения

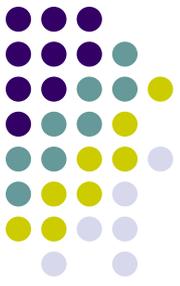


Опыт — это осуществление какого-либо комплекса условий, который может быть воспроизведен много раз.

Под ***событием*** будем понимать результат опыта или наблюдения.

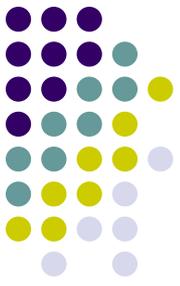
События могут быть ***элементарными*** (неразложимыми) и ***составными*** (разложимыми).

Элементарное событие — это событие, которое происходит в результате единичного опыта.



Составное событие — это совокупность элементарных событий.

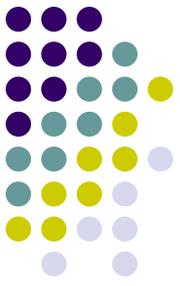
Пример событий: Игральный кубик подбрасывается два раза. Пусть составное событие определено следующим образом: "сумма выпавших цифр равна 6". Тогда элементарные события будут: " $5 + 1$ ", " $4 + 2$ ", " $3 + 3$ ", " $2 + 4$ " и " $1 + 5$ ". Любые другие сочетания не относятся к рассматриваемому составному событию.



Генеральной совокупностью называют совокупность событий, которые могут быть реализованы в результате бесконечного числа однотипных опытов.

Выборочной совокупностью, или просто ***выборкой***, называют совокупность случайно отобранных событий из генеральной совокупности.

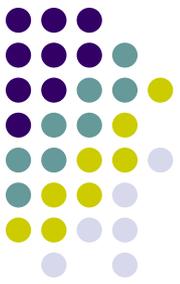
Объемом совокупности называют число событий этой совокупности.



Случайной величиной будем называть величину, которая в результате опыта (наблюдения, испытания) может принять одно из возможных значений, но какое именно — неизвестно.

Вероятность некоторого события — это мера его "благоприятствия".

Некоторые определения



События будем называть **равновозможными**, если мера их "благоприятствия" одинакова.

В реальных условиях, когда количество опытов конечно, мера "благоприятствия" определяется не вероятностью, а **частотью**.

Пусть событие A наблюдалось в m опытах из n опытов (испытаний).

Тогда частота события A - $W(A)$ определяется формулой:

$$\boxed{W(A) = m/n.} \quad (1)$$

Когда n достаточно велико, то "работает" одна из предельных теорем (закон больших чисел — теорема Бернулли), и может быть записано приближенное равенство:

$$\boxed{P(A) \approx W(A),} \quad (2)$$

где $P(A)$ — вероятность события A .

Некоторые определения

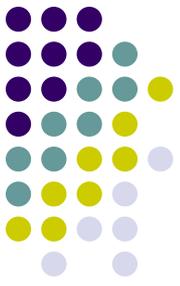


Отметим, что по принятому в теории вероятностей соглашению вероятность произвольного события полагается изменяющейся от 0 до 1.

При этом нулевая вероятность соответствует невозможному событию (которое никогда произойти не может), а единичная – достоверному событию (которое обязательно произойдет).

Слова «случайная величина» в обыденном смысле употребляют тогда, когда хотят подчеркнуть, что неизвестно, каким будет конкретное значение этой величины, причем иногда за этими словами скрывается просто незнание, какова эта величина.

Некоторые определения



Однако математик употребляет эти же слова, вкладывая в них вполне определенное содержание. Действительно, говорит он, мы не знаем, какое значение примет эта величина в данном конкретном случае, но мы знаем, какие значения она может принимать, и знаем, каковы вероятности тех или иных значений.

На основании этих данных мы не можем точно предсказать результат одного испытания, связанного с этой случайной величиной, но можем весьма надежно предсказать совокупность результатов большого числа испытаний. Чем больше испытаний (как говорят, чем больше статистика), тем точнее будут наши предсказания.

Итак, чтобы задать случайную величину, надо указать, какие значения она может принимать и каковы вероятности этих значений.

Дискретные случайные величины



Характеристики дискретных случайных величин

Случайная величина y называется **дискретной**, если она может принимать дискретное множество значений y_1, y_2, \dots, y_n .

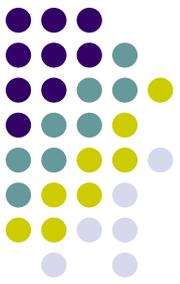
В теории вероятностей рассматриваются также дискретные случайные величины, которые могут принимать бесконечное число значений.

Дискретная случайная величина Y определяется таблицей

$$y = \begin{pmatrix} y_1 & y_2 \dots y_n \\ p_1 & p_2 \dots p_n \end{pmatrix} \quad (3)$$

где y_1, y_2, \dots, y_n — возможные значения величины y , а p_1, p_2, \dots, p_n — соответствующие им вероятности.

Дискретные случайные величины



Точнее говоря, вероятность того, что случайная величина y примет значение y_i (обозначим ее через $P\{y = y_i\}$), равна p_i :

$$P\{y = y_i\} = p_i$$

Таблица (3) называется **распределением случайной величины**.

Числа x_1, x_2, \dots, x_n могут быть, вообще говоря, любыми. Однако вероятности p_1, p_2, \dots, p_n должны удовлетворять двум условиям:

1) все p_i положительны:

$$p_i > 0; \tag{4}$$

2) сумма всех p_i равна 1:

$$p_1 + p_2 + \dots + p_n = 1 \tag{5}$$

(это условие означает, что y обязана в каждом случае принять одно из значений y_1, y_2, \dots, y_n).



Математическим ожиданием дискретной случайной величины y называется сумма произведений всех ее возможных значений на их вероятности:

$$\langle y \rangle = M(y) = \sum_{i=1}^n y_i p_i$$

(6)

Чтобы выяснить физический смысл этой величины, запишем ее в следующем виде:

$$M(y) = \frac{\sum_{i=1}^n y_i p_i}{\sum_{i=1}^n p_i}$$

Отсюда видно, что $M(y)$ - это (генеральное) среднее значение величины y , причем более вероятные значения y входят в сумму с большими весами.



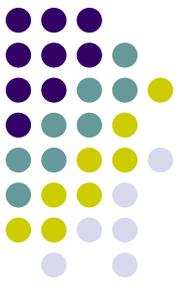
Усреднение с весами очень часто встречается в самых различных областях науки.

Например, в механике: если в точках x_1, x_2, \dots, x_n (на оси Ox) расположены массы m_1, m_2, \dots, m_n , то абсцисса центра тяжести этой системы вычисляется по формуле

$$\langle x \rangle = \sum_{i=1}^n m_i p_i \left(\sum_{i=1}^n m_i \right)^{-1}$$

Конечно, в этом случае сумма всех масс не обязана равняться единице.

Математическое ожидание, мода, медиана



Отметим основные свойства математического ожидания.

Если c - какая-нибудь не случайная величина, то

$$M(y + c) = M(y) + c \quad (7)$$

$$M(c y) = cM(y) \quad (8)$$

Если y, z — любые случайные величины, то

$$M(y + z) = M(y) + M(z). \quad (9)$$

Математическое ожидание — важная, но не единственная характеристика "усредненного" значения случайной величины.

"Усредненное" значение может характеризоваться также **модой** и **медианой**.



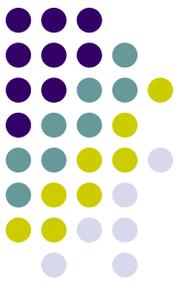
Модой случайной величины называется ее наиболее вероятное значение (то, для которого вероятность достигает максимума).

Медианой случайной величины называется такое значение m из области определения случайной величины, для которого

$$P\{y < m\} \approx P\{y > m\}.$$

Из этого следует, что для случайной величины принятие значения, меньшего медианы, равновероятно принятию значения, большего медианы.

Математическое ожидание, мода, медиана

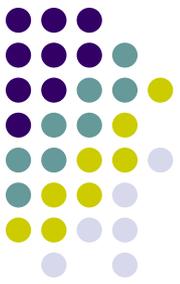


Математическое ожидание, мода и медиана случайной величины по-разному характеризуют усредненное значение случайной величины. Эти различия становятся понятными из следующего примера.

Если некоторая фирма хочет оценить необходимое количество денег для выдачи зарплаты, ее интересует средняя зарплата сотрудника, то есть математическое ожидание зарплаты сотрудника, рассматриваемой как случайная величина.

При устройстве на работу будущего сотрудника должна интересовать не средняя зарплата, а ее мода — заработная плата, получаемая большинством сотрудников.

Сопоставление собственной заработной платы с медианой заработной платы говорит сотруднику о том, принадлежит ли он к хорошо или плохо оплачиваемой части сослуживцев.



Дисперсией дискретной случайной величины называется число

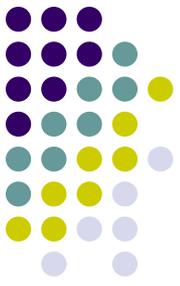
$$D(y) = M(y - M(y))^2 = \langle (y - \langle y \rangle)^2 \rangle \quad (10)$$

Следовательно, дисперсия $D(y)$ - это математическое ожидание квадрата отклонения случайной величины от ее среднего значения $M(y)$.

Очевидно, всегда $D(y) > 0$.

Математическое ожидание и дисперсия - важнейшие числовые характеристики случайной величины y . Каково их практическое значение?

Дисперсия



Если мы будем наблюдать величину y много раз и получим значения y_1, y_2, \dots, y_n , то среднее арифметическое от этих значений будет близко к $M(y)$:

$$\frac{y_1 + y_2 + \dots + y_N}{N} = M(y) \quad (11)$$

Дисперсия $D(y)$ характеризует разброс этих значений около среднего $M(y)$.

Формулу (10) для дисперсии можно преобразовать с помощью формул (7)-(9) :

$$D(y) = M[y^2 - 2M(y) \cdot y + M(y)^2] = M(y^2) - 2M(y)M(y) + M(y)^2,$$

Откуда

$$D(y) = M(y^2) - M(y)^2. \quad (12)$$

Вычислять дисперсию по этой формуле обычно проще, чем по формуле (10) .

Дисперсия



Отметим основные свойства дисперсии. Если c — какая-нибудь не случайная величина, то

$$D(y + c) = D(y), \quad (13)$$

$$D(cy) = c^2 D(y). \quad (14)$$

Важную роль в теории вероятностей играет понятие **независимости случайных величин**.

В действительности это довольно сложное понятие, но в простейших случаях оно очевидно: допустим, что кроме случайной величины y мы наблюдаем еще за случайной величиной z . Если распределение величины y не меняется от того, что нам уже известно значение, которое приняла величина z , то естественно считать, что y от z не зависит.

Дисперсия

Для независимых случайных величин y и z справедливы следующие соотношения:

$$M(yz) = M(y)M(z), \quad (15)$$

$$D(y+z) = D(y) + D(z). \quad (16)$$

Рассмотрим случайную величину χ с распределением

$$\chi \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}$$

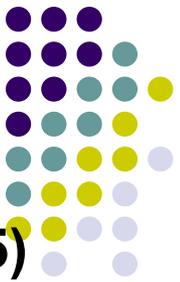
Очевидно, что реализацией этой величины можно считать число очков, выпадающих на игральной кости: любое значение одинаково вероятно.

Вычислим математическое ожидание χ по формуле (6):

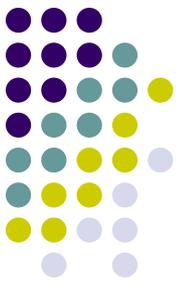
$$M(\chi) = 1 \cdot \frac{1}{6} + \dots + 6 \cdot \frac{1}{6} = 3,5;$$

дисперсию χ по формуле (12):

$$D(\chi) = M(\chi^2) - (M(\chi))^2 = 1^2 \cdot \frac{1}{6} + \dots + 6^2 \cdot \frac{1}{6} - (3,5)^2 = 2,917$$



Дисперсия



Рассмотрим еще один пример. Возьмем случайную величину θ

$$\theta = \begin{pmatrix} 3 & 4 \\ 1/2 & 1/2 \end{pmatrix}$$

Реализацией этой величины можно считать игру в орлянку с условием, что, например, орел - это 3 очка, а решетка - 4 очка. В этом случае

$$M(\theta) = 0,5 \cdot 3 + 0,5 \cdot 4 = 3,5;$$

$$D(\theta) = 0,5 (3^2 + 4^2) - (3,5)^2 = 0,25.$$

Мы видим, что $M(\theta) = M(\chi)$, но $D(\theta) < D(\chi)$. Последнее можно было легко предвидеть, так как значения θ могут отличаться от 3,5 только на $\pm 0,5$, в то время как в значениях χ разброс может достигать $\pm 2,5$.

Непрерывные случайные величины

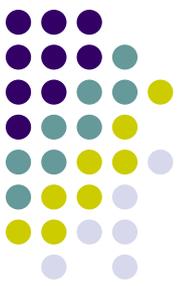


Предположим, что на плоскости в начале координат расположено некоторое количество радия. При распаде каждого атома радия из него вылетает α -частица. Направление ее будем характеризовать углом ϕ .

Так как и теоретически, и практически возможны любые направления вылета, то эта случайная величина может принимать любое значение от 0 до 2π .

Мы будем называть случайную величину ξ **непрерывной**, если она может принимать любое значение из некоторого интервала (a, b) .

Непрерывные случайные величины



Непрерывная случайная величина ξ определяется заданием интервала (a, b) , содержащего возможные значения этой величины, и функции $p(x)$, которая называется **плотностью вероятностей случайной величины ξ** (или **плотностью распределения ξ**).

Физический смысл $p(x)$ следующий: пусть (a', b') - произвольный интервал, содержащийся в (a, b) . Тогда вероятность того, что ξ окажется в интервале (a', b') , равна интегралу

$$P\{a' < \xi < b'\} = \int_{a'}^{b'} p(x) dx \quad (17)$$

Непрерывные случайные величины



На рис. 3 заштрихованная площадь равна значению этого интеграла.

Множество значений ξ может быть любым интервалом. Возможен даже случай $a = -\infty$, а также $b = \infty$.

Однако плотность $p(x)$ должна удовлетворять двум условиям, аналогичным (4)-(5), для дискретных величин:

1. плотность $p(x)$ положительна:
$$p(x) > 0 \quad (18)$$

2. интеграл от плотности $p(x)$ по всему интервалу (a, b) равен 1:

$$\int_a^b p(x) dx = 1 \quad (19)$$

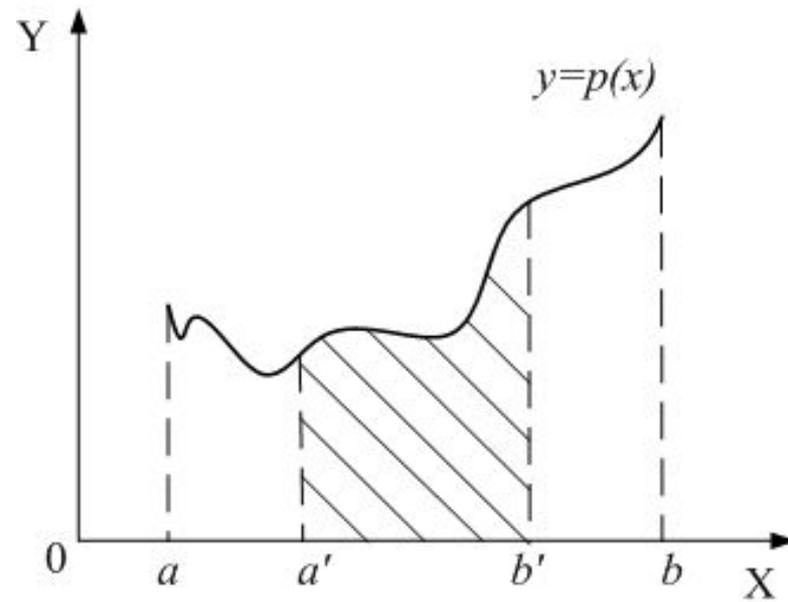


Рис. 3. Плотность вероятности непрерывной случайной величины

Непрерывные случайные величины



Математическим ожиданием непрерывной случайной величины называется число

$$M(\xi) = \int_a^b xp(x)dx$$

(20)

Смысл этой характеристики такой же, как в случае дискретной случайной величины.

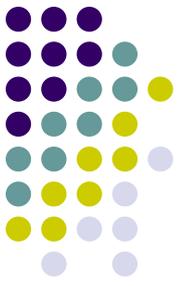
Отметим, что для любой непрерывной случайной величины ξ при каждом x имеем

$$P\{\xi = x\} = 0.$$

Физический смысл имеет не вероятность попадания ξ в заданную точку x , а вероятность попадания в сколь угодно малый интервал:

$$P\{x \leq \xi < x+dx\} = p(x)dx.$$

Непрерывные случайные величины



Рассмотрим пример.

Случайная величина γ , определенная в интервале $(0,1)$ и имеющая плотность $p(x)=1$, называется **равномерно распределенной** в $(0,1)$.

В самом деле, какой бы интервал (a', b') внутри $(0, 1)$ мы ни взяли, вероятность того, что γ попадет в (a', b') , равна

$$\int_{a'}^{b'} p(x)dx = b' - a'$$

т.е. длине этого интервала.

В частности, если мы разделим интервал $(0,1)$ на любое число интервалов равной длины, то вероятность попадания γ в любой из них будет одна и та же.

Легко вычислить, что

$$M(\gamma) = \int_0^1 xp(x)dx = \int_0^1 xdx = \frac{1}{2} \quad D(\gamma) = \int_0^1 x^2 p(x)dx - (M\gamma)^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$$



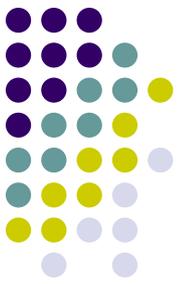
Нормальной (или **гауссовской**) случайной величиной называется случайная величина ζ , определенная на всей оси $(-\infty, \infty)$ и имеющая плотность

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\}, \quad (21)$$

где a и $\sigma > 0$ — числовые параметры.

Параметр a не влияет на форму кривой $y=p(x)$: изменение его приводит лишь к сдвигу кривой вдоль оси x . Однако при изменении σ форма кривой меняется.

Нормальные случайные величины

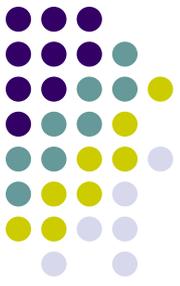


В самом деле, легко видеть, что

$$\max p(x) = p(a) = \frac{1}{\sigma \sqrt{2\pi}} \quad (22)$$

Если уменьшать σ , то $\max p(x)$ будет возрастать. Но вся площадь под кривой $y = p(x)$ по условию (19) равна 1. Поэтому кривая будет вытягиваться вверх в окрестности $x = a$, но убывать при всех достаточно больших значениях x .

Нормальные случайные величины



На рис. 4 построены две нормальные плотности, соответствующие $a = 0$, $\sigma = 1$ и $a=0$, $\sigma = 0,5$. Можно доказать, что

$$M(\zeta) = a, \quad D(\zeta) = \sigma^2. \quad (23)$$

Любые вероятности вида $P \{x' < S < x''\}$ легко вычисляются с помощью таблицы, в которой приведены значения функции

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt, \quad (24)$$

называемой **интегралом вероятностей**.

Нормальные случайные величины очень часто встречаются при исследовании самых различных по своей природе вопросов.

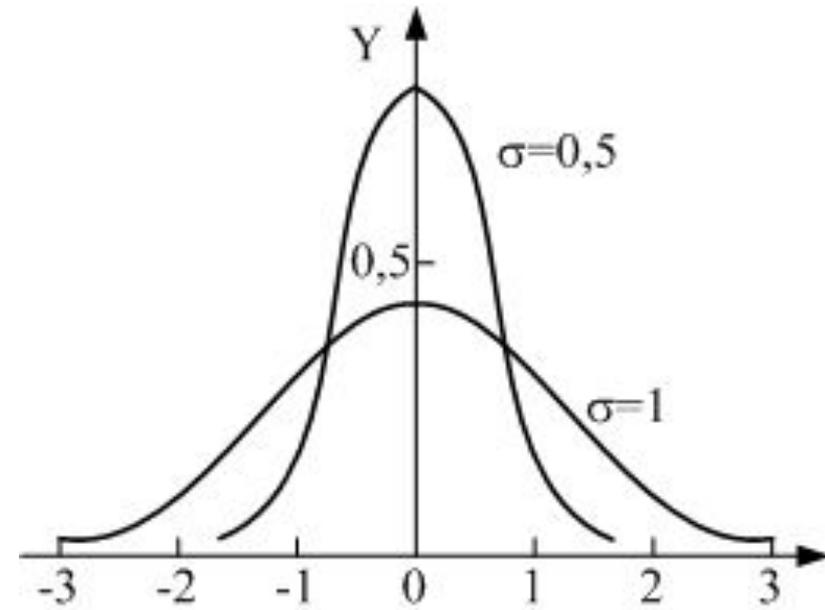
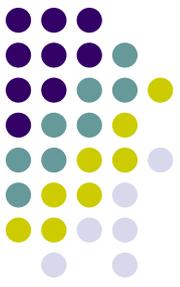


Рис 4. Нормальные распределения

Правило «трех сигм»



Вероятность попадания реализаций нормально распределенной случайной величины в интервал $a \pm 3\sigma$ составляет более 99,7%:

$$P\{a - 3\sigma < \zeta < a + 3\sigma\} = \Phi(3) = 0,997.$$

Вероятность 0,997 настолько близка к 1, что иногда последнюю формулу интерпретируют так:

при одном испытании практически невозможно получить значение ζ , отличающееся от $M(\zeta)$ больше, чем на 3σ .

Центральная предельная теорема



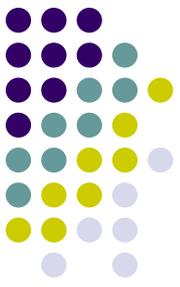
Эта замечательная теорема была впервые сформулирована **П. Лапласом**. Обобщением этой теоремы занимались многие выдающиеся математики, в том числе **П.Л.Чебышев, А.А.Марков, А.М.Ляпунов**. Доказательство ее достаточно сложно.

Рассмотрим N одинаковых независимых случайных величин $\xi_1, \xi_2, \dots, \xi_n$, так что распределения вероятностей этих величин совпадают. Следовательно, их математические ожидания и дисперсии также совпадают (предполагается, что они конечны). Величины эти могут быть как непрерывными, так и дискретными.

Обозначим

$$M(\xi_1) = M(\xi_2) = \dots = M(\xi_n) = m,$$
$$M(\xi_1) = M(\xi_2) = \dots = M(\xi_n) = b^2.$$

Центральная предельная теорема



Рассмотрим теперь нормальную случайную величину ε с такими же параметрами:

$$a = Nm, \quad \sigma = b\sqrt{N} .$$

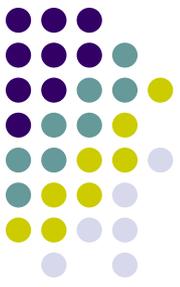
В центральной предельной теореме утверждается, что *для любого интервала (a', b') при больших N*

$$P\{a' < \rho_N < b'\} \approx \int_{a'}^{b'} p_{\zeta_N}(x) dx \quad (25)$$

Физический смысл этой теоремы очевиден: сумма ρ_N большого числа одинаковых случайных величин приближенно нормальна.

На самом деле эта теорема справедлива при гораздо более широких условиях: все слагаемые не обязаны быть одинаковыми и независимыми; существенно только, чтобы отдельные слагаемые не играли слишком большой роли в сумме.

Центральная предельная теорема



Именно эта теорема объясняет, почему нормальные случайные величины так часто встречаются в природе.

В самом деле, каждый раз, когда мы сталкиваемся с суммарным воздействием большого числа незначительных случайных факторов, результирующая случайная величина оказывается нормальной.

Например, отклонение артиллерийского снаряда от цели почти всегда оказывается нормальной случайной величиной, так как оно зависит и от метеорологических условий на разных участках траекторий, и от многих других факторов.

Общая схема метода Монте-Карло



Допустим, что нам требуется вычислить какую-то неизвестную величину m .

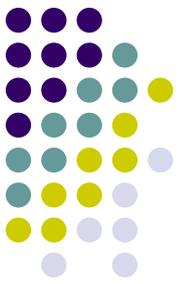
Попытаемся придумать такую случайную величину ξ , чтобы $M(\xi) = m$.

Пусть при этом $D(\xi) = b^2$.

Рассмотрим N независимых случайных величин $\xi_1, \xi_2, \dots, \xi_N$, распределения которых совпадают с распределением ξ .

Если N достаточно велико, то согласно центральной предельной теореме распределение суммы $\rho_N = \xi_1 + \xi_2 + \dots + \xi_N$ будет приблизительно нормальным с параметрами

$$a = Nm, \quad \sigma = b\sqrt{N}.$$



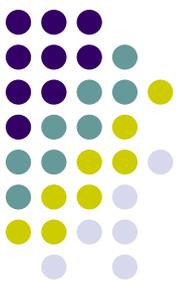
Из (25) следует, что

$$P\{Nm - 3b\sqrt{N} < \rho_N < Nm + 3b\sqrt{N}\} \approx 0,997$$

Если мы разделим выражение, стоящее в фигурной скобке, на N , то получим эквивалентное неравенство, и вероятность его останется такой же:

$$P\{m - 3b/\sqrt{N} < \rho_N/N < m + 3b/\sqrt{N}\} \approx 0,997$$

Общая схема метода Монте-Карло



Последнее соотношение перепишем в несколько ином виде:

$$P \left\{ \left| \frac{1}{N} \sum_{j=1}^N \xi_j - m \right| < \frac{3b}{\sqrt{N}} \right\} \approx 0.997 \quad (26)$$

Это - чрезвычайно важное для метода Монте-Карло соотношение. Оно дает нам и метод расчета m , и оценку погрешности.

В самом деле, найдем N значений случайной величины ξ (все равно, находить ли один раз по одному значению каждой из величин $\xi_1, \xi_2, \dots, \xi_n$ или найти N значений величины ξ , так как все эти случайные величины имеют одно и то же распределение и предполагаются независимыми). Из (26) видно, что среднее арифметическое этих значений будет приближенно равно m .

С большой вероятностью погрешность такого приближения не превосходит величины $\frac{3b\sqrt{N}}{N}$. Очевидно, эта погрешность стремится к нулю с ростом N .

Общая схема метода Монте-Карло



На практике очень часто предпочитают ориентироваться не на оценку сверху $3b\sqrt{N}$,

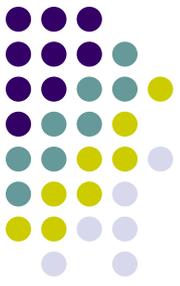
а на *вероятную ошибку*, которая приблизительно равна

$$r_N = 0,6745b/\sqrt{N}$$

и фактическая погрешности расчета в этом случае равна

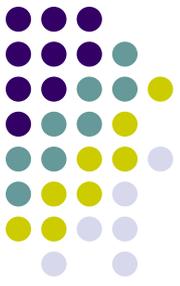
$$\left| \frac{1}{N} \sum_{j=1}^N \xi_j - m \right|$$

Прохождение нейтронов сквозь пластинку



Вероятностные законы взаимодействия отдельной элементарной частицы (нейтрона, фотона, мезона и др.) с веществом известны. Обычно требуется найти макроскопические характеристики процессов, в которых участвует огромное количество таких частиц: плотности, потоки и т. п.

Прохождение нейтронов сквозь пластинку



Рассмотрим простейший вариант задачи о прохождении нейтронов сквозь пластинку.

Пусть на однородную бесконечную пластинку толщиной h падает поток нейтронов с энергией E_0 ; угол падения 90° .

При столкновении с атомами вещества, из которого состоит пластинка, нейтроны могут упруго рассеиваться или поглощаться.

Предположим для простоты, что энергия нейтрона при рассеянии не меняется и любое направление «отскока» нейтрона от атома одинаково вероятно (последнее иногда справедливо в веществах с тяжелыми атомами).

Прохождение нейтронов сквозь пластинку

На рис. 5 изображены различные варианты взаимодействия нейтронов с пластинкой:

a — нейтрон проходит сквозь пластинку,

б — нейтрон поглощается в пластинке,

в — нейтрон отражается пластинкой.

Требуется вычислить вероятность прохождения нейтрона сквозь пластинку p^+ ; вероятность отражения нейтрона пластин p^- и вероятность поглощения нейтрона в пластинке p^0 .

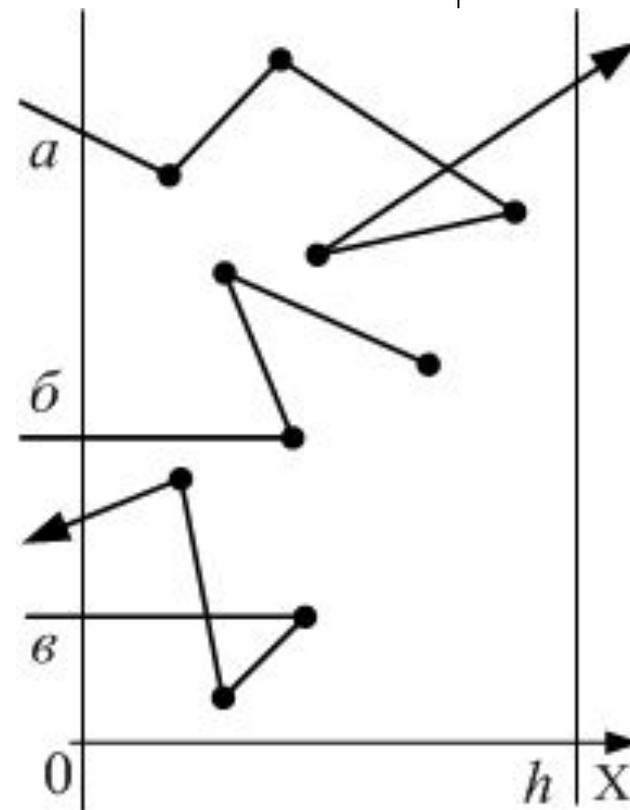
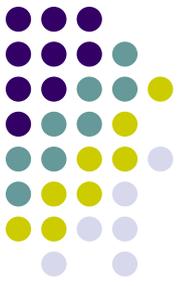


Рис. 5. Различные случаи «судьбы» нейтронов, падающих на пластинку

Прохождение нейтронов сквозь пластинку

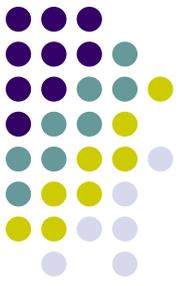


Взаимодействие нейтронов с веществом характеризуется в рассматриваемом случае двумя постоянными Σ_c и Σ_s , которые называются **сечением поглощения и сечением рассеяния** (индексы c и s —это первые буквы английских слов capture -- захват и scattering -- рассеяние).

Сумма этих сечений $\Sigma = \Sigma_c + \Sigma_s$ называется **полным сечением**.

Физический смысл сечений следующий:
при столкновении нейтрона с атомом вещества вероятность поглощения равна Σ_c / Σ , а вероятность рассеяния равна Σ_s / Σ .

Прохождение нейтронов сквозь пластинку



Длина свободного пробега нейтрона λ (т. е. длина пути от столкновения до столкновения) — это случайная величина. Она может принимать любые положительные значения с плотностью вероятностей

$$p(x) = \Sigma e^{-\Sigma x}.$$

Мы можем сразу написать выражение для средней длины свободного пробега:

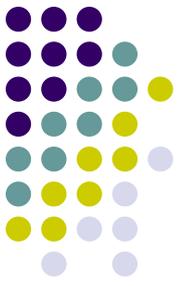
$$\langle \lambda \rangle = 1/\Sigma,$$

а также формулу для розыгрыша λ :

$$\lambda = -(1/\Sigma) \ln \gamma,$$

где γ – случайная величина на интервале (0,1).

Прохождение нейтронов сквозь пластинку



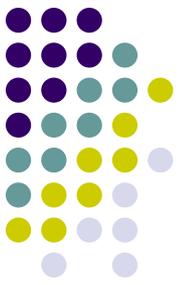
Остается еще выяснить, как выбирать случайное направление нейтрона после рассеяния.

Так как задача симметрична относительно оси x , то направление вполне определяется одним углом ϕ между направлением скорости нейтрона и осью Ox .

Можно доказать, что требование равной вероятности любого направления в этом случае равносильно требованию, чтобы косинус этого угла $\mu = \cos\phi$ был равномерно распределен в интервале $(-1,1)$ и

$$\mu = 2\gamma - 1.$$

Схема моделирования траекторий нейтронов



Предположим, что нейтрон испытал k -е рассеяние внутри пластинки в точке с абсциссой x_k и после этого начал двигаться в направлении μ_k .

Разыграем длину свободного пробега

$$\lambda = -(1/\Sigma)\ln\gamma$$

и вычислим абсциссу следующего столкновения (рис.6)

$$x_{k+1} = x_k + \lambda \mu_k.$$

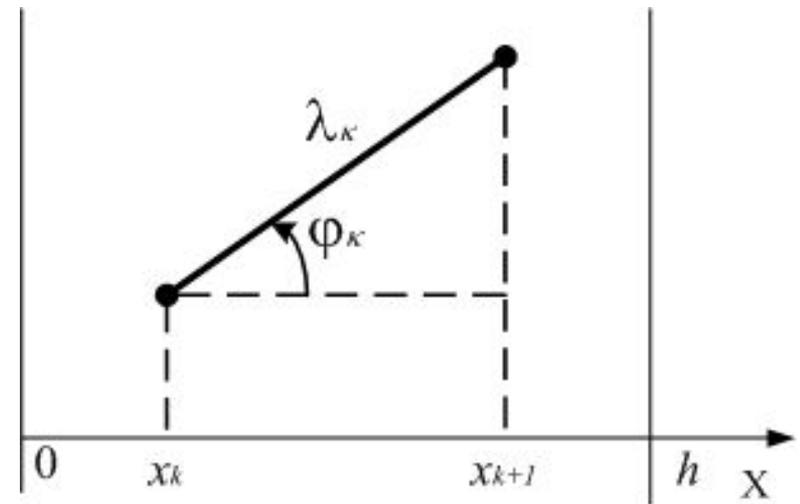
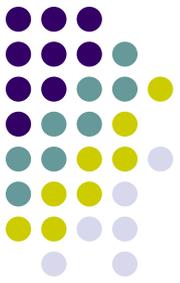


Схема моделирования траекторий нейтронов



Проверим условие прохождения сквозь пластинку:

$$x_{k+1} > h.$$

Если это условие выполнено, то счет траектории нейтрона заканчивается и добавляется единица к счетчику прошедших частиц.

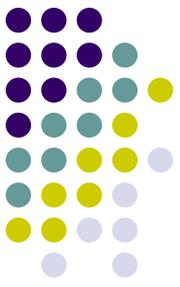
В противном случае проверяем условие отражения:

$$x_{k+1} < 0.$$

Если это условие выполнено, то счет траектории заканчивается и добавляется единица к счетчику отраженных частиц.

Если же и это условие не выполнено, т. е. $0 \leq x_{k+1} \leq h$, то, значит, нейтрон испытал $k+1$ -е столкновение внутри пластинки, и надо разыграть «судьбу» нейтрона при столкновении.

Схема моделирования траекторий нейтронов



Выбираем очередное значение γ и проверяем условие поглощения:

$$\gamma < \Sigma_c / \Sigma$$

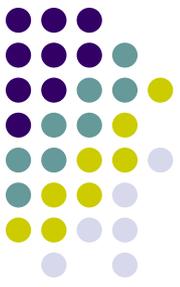
Если последнее неравенство выполнено, то счет траектории заканчивается и добавляется единица к счетчику поглощенных частиц.

В противном случае мы считаем, что нейтрон испытал рассеяние в точке с абсциссой x_{k+1} . Тогда разыгрываем новое направление скорости нейтрона

$$\mu_{k+1} = 2\gamma - 1$$

и затем повторяем весь цикл снова (но, конечно, уже с другими значениями γ).

Схема моделирования траекторий нейтронов

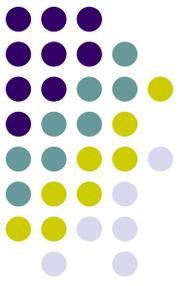


Все γ написаны без индексов, так как имеется в виду, что каждое значение γ используется всего один раз. Для расчета одного звена траектории нужны три значения γ . Начальные значения для каждой траектории:

$$x_0 = 0, \mu_0 = 1$$

После того как будут сосчитаны N траекторий, окажется, что N^+ нейтронов прошли сквозь пластинку, N^- нейтронов отразились от нее, а N^0 нейтронов были поглощены в ней.

Схема моделирования траекторий нейтронов



Очевидно, искомые вероятности приближенно равны отношениям:

$$p^+ \approx \frac{N^+}{N}, p^- \approx \frac{N^-}{N}, p^0 \approx \frac{N^0}{N}.$$

На рис. 7 приведена блок-схема программы для расчета этой задачи.

Индекс j — это номер траектории, индекс k — номер столкновения (вдоль траектории).

Отметим только, что метод Монте-Карло позволяет решать гораздо более сложные задачи об элементарных частицах: исследуемая среда может состоять из различных веществ и иметь любую геометрическую структуру; энергия частицы при каждом столкновении может меняться. Можно учитывать много других ядерных процессов (например, возможность деления атома при столкновении с нейтроном и образование при этом новых нейтронов). Можно рассчитать условие возникновения и поддержания цепной реакции и т. д.

Схема моделирования траекторий нейтронов

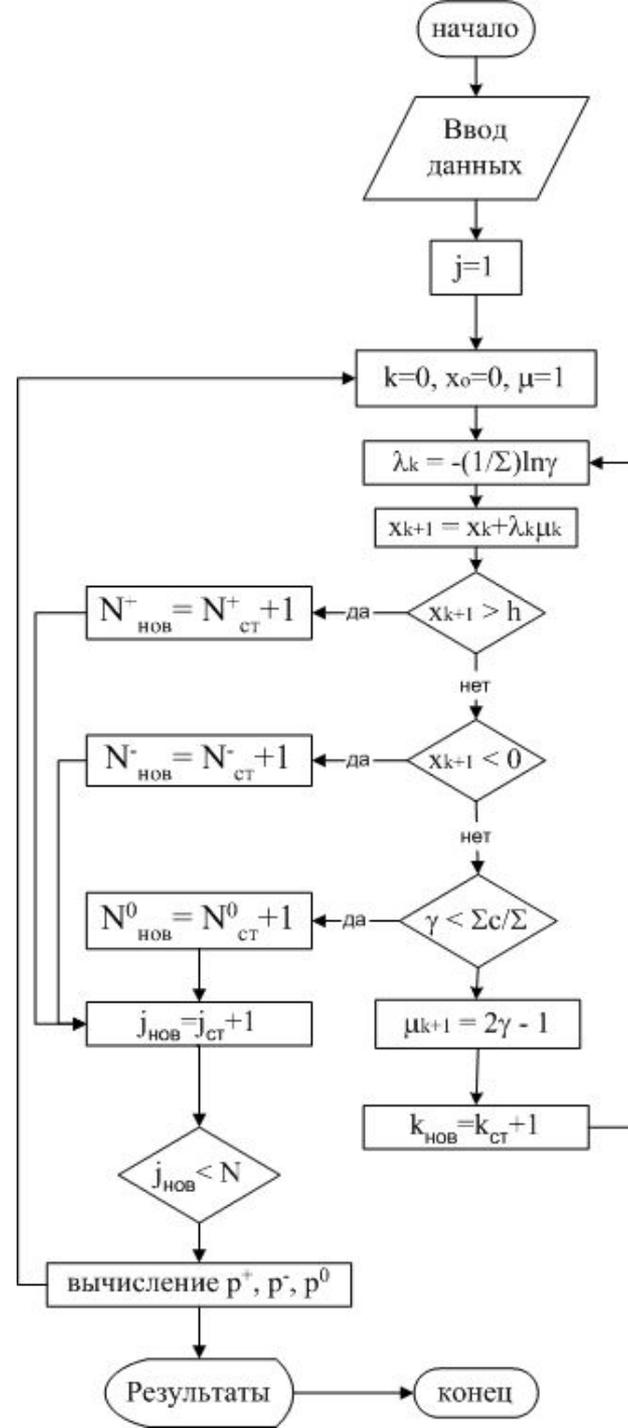


Рис. 7. Блок-схема программы расчета прохождения нейтронов в пластине