

РАЗДЕЛ 2. ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ

Лекция 7. Принципы имитационного моделирования

Вернемся к рассмотренной задаче о встрече, которую мы решали аналитически.

Задача 1. Пусть два объекта (например, пешеход и велосипедист) движутся друг другу навстречу (**рис. 7.1**) со скоростями V_1 и V_2 соответственно. Необходимо узнать: когда и где встретятся эти объекты?

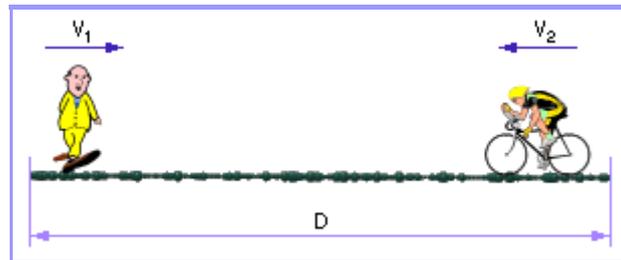


Рис. 7.1. Задача о встрече

Имитационный способ представления задачи 1

При имитационном способе решения обязательным является наличие некоего счётчика, который позволяет моделировать процесс по шагам или по деталям процесса.

И м и т а ц и о н н ы й а л г о р и т м и ч е с к и й с п о с о б

Повторяя пошагово расчёт в цикле, на каждом этапе алгоритма будем имитировать течение процесса (**рис. 1.19**). Обратите внимание, что процесс берётся не в целом, а в деталях, по шагам. Переменная t является координатой, а значит, отслеживается счётчиком с шагом h . Идея имитации — продвигать пешехода и велосипедиста на величину $V \cdot h$ на каждом такте, где h — достаточно малая величина. Поскольку мы рассматриваем множество актов движения по отдельности, можно по ходу менять все переменные модели, например, V . Если путь пройден большой (S_1), то можно устроить привал ($V = 0$) на некоторое время. Остановка процесса имитации определяется суммой путей, пройденных велосипедистом и пешеходом навстречу друг другу, и сравнением её с расстоянием D .

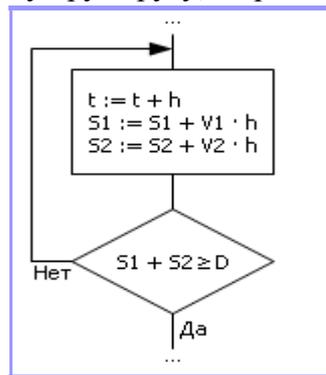


Рис. 7.2. Блок-схема решения задачи о встрече (имитационный алгоритмический способ)

На формально-математическом языке алгоритм выглядит так, как показано ниже.

```

t := t + h * e
S1 := S1 + V1 * h * e
S2 := S2 + V2 * h * e
e := ed(D - (S1 + S2))
f := not(e)
stop(f)
    
```

e – вспомогательный флаг;
 f – флаг, показывающий, был ли пройден к текущему моменту t весь путь или нет;
 $ed(x)$ – единичная функция: $ed(x) = 1$ при $x \geq 0$, иначе $ed(x) = 0$;
 $stop(z)$ – функция останова вычислений при $z > 0$.

Имитационный геометрический способ

Решение может быть найдено геометрически. Для этого в осях (t, S) схемой, показанной на **рис. 7.3**, строятся траектории движения объектов.

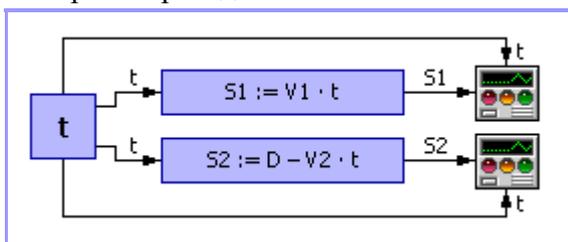


Рис. 7.3. Схема решения задачи о встрече (имитационный геометрический способ)

На **рис. 7.4** вы видите картину, образованную двумя осциллограммами. Точка, в которой пересекаются осциллограммы, является предполагаемой точкой встречи двух объектов.

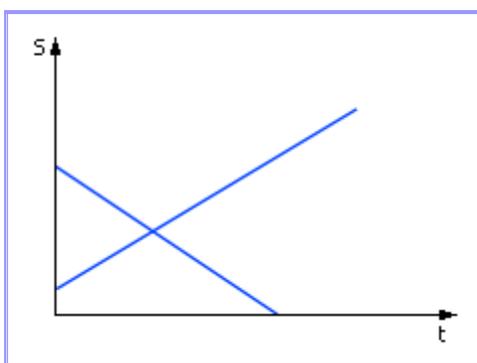


Рис. 7.4. Вид решения задачи о встрече (имитационный геометрический способ)

Имитационная статистическая постановка задачи

Главное отличие имитационных моделей от аналитических, которые мы рассмотрели выше, состоит в том, что имитационную модель можно постепенно последовательно усложнять. При этом результативность модели не падает.

Усложним задачу 1, введя в неё дополнительное условие. Представим, что на пути первого и/или второго объекта встретится помеха – пусть это будет участок железной дороги со шлагбаумом, который работает по случайному закону. Если шлагбаум открыт, то объект может переходить железную дорогу, в противном случае он не имеет права этого делать.

Промоделировать случайную работу шлагбаума можно с помощью генератора случайных чисел (ГСЧ). В различные моменты времени ГСЧ будет выдавать случайное число $r = 0$ или $r = 1$, это будет означать, что шлагбаум закрыт или, соответственно, открыт (см. **рис. 7.5**).



Рис. 7.5. Вид функции случайных помех (к задаче о встрече)

Частоту открывания шлагбаума можно контролировать, увеличивая или, наоборот, уменьшая число q , пересчитав случайное число r в z по формуле: $z := ed(q - r)$.

На **рис. 7.6** дана иллюстрация усложнённой задачи 1.

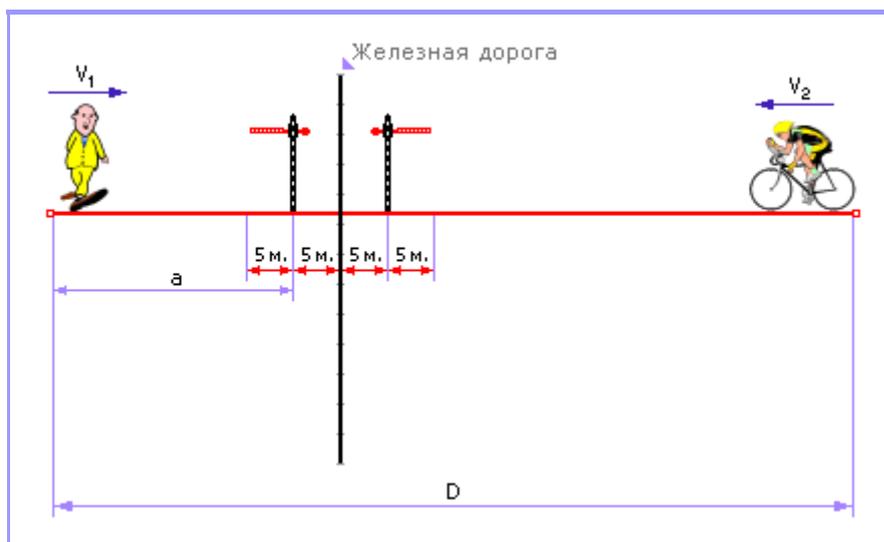


Рис. 7.6. Иллюстрация к усложнённой задаче о встрече

На рис. 7.7 представлена алгоритмическая схема задачи.

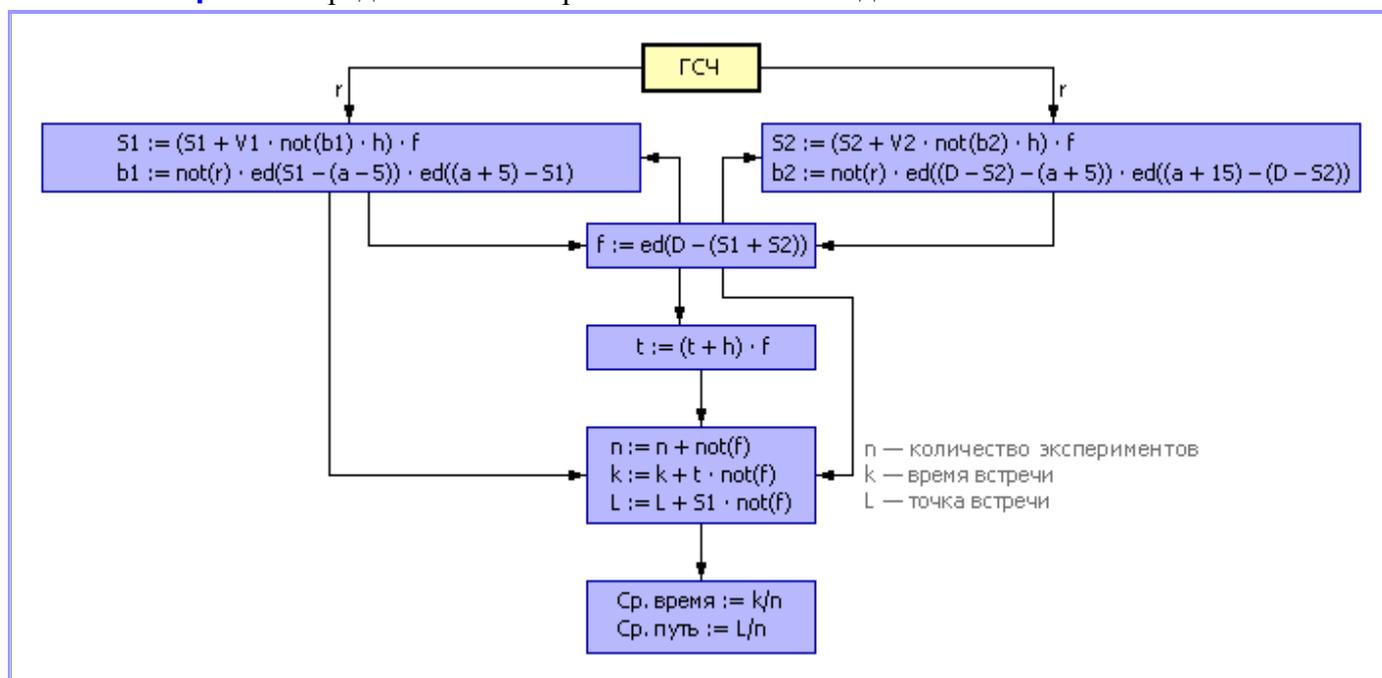


Рис. 7.7. Схема решения задачи о встрече (имитационный статистический способ)

Условия $b1$ и $b2$ контролируют, находится ли первый и/или второй объект менее чем за 5 метров от шлагбаума, когда тот закрыт. Условия $b1 = 1$ ($b2 = 1$) – это условия «не двигаться», если объект находится в зоне шлагбаума и шлагбаум закрыт; a – место нахождения шлагбаума, расстояние до шлагбаума от нуля; f – флаг встречи. Если $f = 0$, то встреча произошла и моделирование начинается снова с $t = 0$, $S1 = 0$, $S2 = 0$, а к статистическим счётчикам необходимо прибавить итоги эксперимента — номер эксперимента, время встречи, место встречи.

Поскольку алгоритм использует случайные числа в качестве исходных данных, придётся сделать несколько экспериментов и найти средние значения выходных величин. Результат одного эксперимента случаен и ни о чем не говорит. Среднее значение более информативно. Ещё более информативны сведения о первом и втором моменте – среднем и разбросе значений вокруг него (дисперсии) и так далее.

И м и т а ц и о н н ы й к р и т е р и а л ь н ы й с п о с о б

Этот способ – ещё один шаг к усложнению модели и приближению её к реальным условиям. Если во всех предыдущих случаях скорость V была известна, то теперь она –

непредсказуема, даже статистически, что чаще всего и наблюдается в реальных условиях. Скорость V объекта принимается из некоторых дополнительных моделей. Причём эти модели могут быть с обратными связями. Например, скорость может зависеть от обстановки, степени достижения цели, плана, целесообразности поведения, знания местности, желания двигаться (психологический мотив). Такая постановка ведёт к адаптивным системам и системам искусственного интеллекта.

Например, если объект (пешеход) имеет возможность осмотреть местность с обзорной башни, то он может заранее просчитать свой путь и затем следовать ему. Без обзорной башни пешеход может зайти в безвыходный тупик или бесконечно долго перебирать варианты пути. Для адаптивных систем вводится критерий для оценки перспективности выбора направления движения.

Статистическое моделирование – базовый метод моделирования, заключающийся в том, что модель испытывается множеством случайных сигналов с заданной плотностью вероятности. Целью является статистическое определение выходных результатов. В основе статистического моделирования лежит *метод Монте-Карло*. Напомним, что имитацию используют тогда, когда другие методы применить невозможно.

Метод Монте-Карло

Рассмотрим метод Монте-Карло на примере вычисления интеграла, значение которого аналитическим способом найти не удастся.

Задача. Найти значение интеграла:

$$y = \int_{x_1}^{x_2} f(x) dx$$

На **рис. 7.8** представлен график функции $f(x)$. Вычислить значение интеграла этой функции – значит, найти площадь под этим графиком.

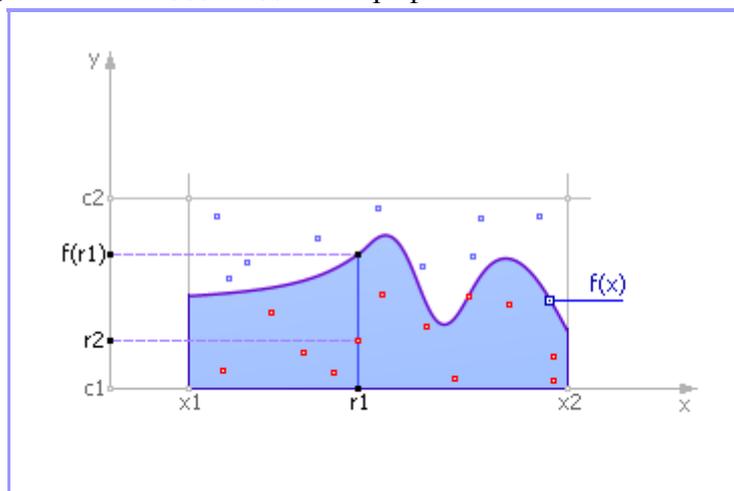


Рис. 7.8. Определение значения интеграла методом Монте-Карло

Ограничиваем кривую сверху, справа и слева. Случайным образом распределяем точки в прямоугольнике поиска. Обозначим через N_1 количество точек, принятых для испытаний (то есть попавших в прямоугольник, эти точки изображены на **рис. 7.8** красным и синим цветом), и через N_2 – количество точек под кривой, то есть попавших в закрашенную площадь под функцией (эти точки изображены на **рис. 7.8** красным цветом). Тогда естественно предположить, что количество точек, попавших под кривую по отношению к общему числу точек пропорционально площади под кривой (величине интеграла) по отношению к площади испытываемого прямоугольника. Математически это выражается так:

$$\frac{N_2}{N_1} = \frac{y}{(x_2 - x_1)(c_2 - c_1)}$$

Рассуждения эти, конечно, статистические и тем более верны, чем большее число испытываемых точек мы возьмем. Фрагмент алгоритма метода Монте-Карло в виде блок-

схемы выглядит так, как показано на [рис. 7.9](#).

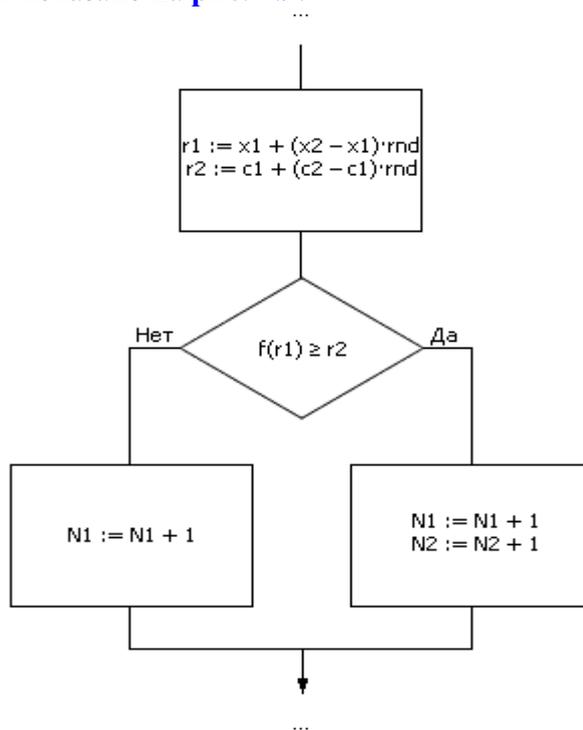


Рис. 7.9. Фрагмент алгоритма реализации метода Монте-Карло

Значения r_1 и r_2 на [рис. 7.9](#) являются равномерно распределенными случайными числами из интервалов $(x_1; x_2)$ и $(c_1; c_2)$ соответственно.

Метод Монте-Карло чрезвычайно эффективен, прост, но необходим «хороший» генератор случайных чисел. Вторая проблема применения метода заключается в определении объема выборки, то есть количества точек, необходимых для обеспечения решения с заданной точностью. Эксперименты показывают: чтобы увеличить точность в 10 раз, объем выборки нужно увеличить в 100 раз; то есть точность примерно пропорциональна корню квадратному из объема выборки:

$$\text{точность} \cong \sqrt{\text{объем выборки}}$$

Схема использования метода Монте-Карло при исследовании систем со случайными параметрами

Построив модель системы со случайными параметрами, на ее вход подают входные сигналы от генератора случайных чисел (ГСЧ), как показано на [рис. 7.10](#). ГСЧ устроен так, что он выдает *равномерно распределенные* случайные числа r_{pp} из интервала $[0; 1]$.

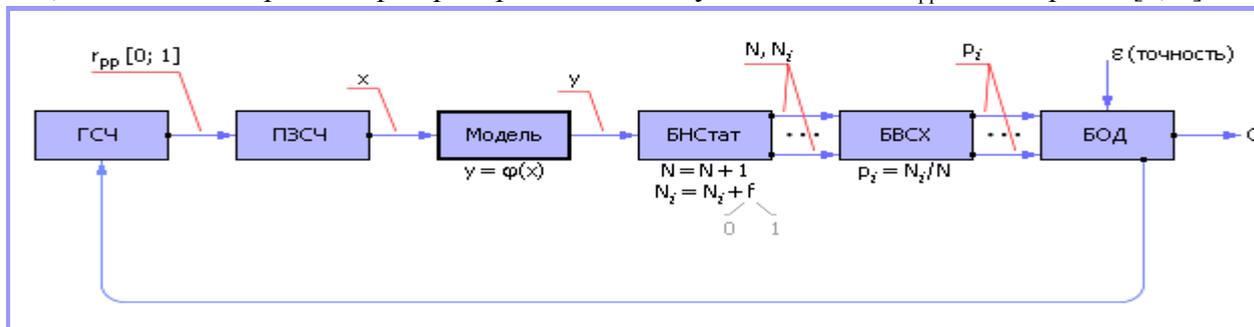


Рис. 7.10. Общая схема метода статистического моделирования

Так как одни события могут быть более вероятными, другие — менее вероятными, то равномерно распределенные случайные числа от генератора подают на преобразователь закона случайных чисел (ПЗСЧ), который преобразует их в *заданный* пользователем закон распределения вероятности, например, в нормальный или экспоненциальный закон. Эти преобразованные случайные числа x подают на вход модели. Модель обрабатывает

входной сигнал x по некоторому закону $y = \varphi(x)$ и получает выходной сигнал y , который также является случайным.

В блоке накопления статистики (БНСтат) установлены фильтры и счетчики. Фильтр (некоторое логическое условие) определяет по значению y , реализовалось ли в конкретном опыте некоторое событие (выполнилось условие, $f = 1$) или нет (условие не выполнилось, $f = 0$). Если событие реализовалось, то счетчик события увеличивается на единицу. Если событие не реализовалось, то значение счетчика не меняется. Если требуется следить за несколькими разными типами событий, то для статистического моделирования понадобится несколько фильтров и счетчиков N_i . Всегда ведется счетчик количества экспериментов – N .

Далее отношение N_i к N , рассчитываемое в блоке вычисления статистических характеристик (БВСХ) по методу Монте-Карло, дает оценку вероятности p_i появления события i , то есть указывает на частоту его выпадения в серии из N опытов. Это позволяет сделать выводы о статистических свойствах моделируемого объекта.

Например, событие А совершилось в результате проведенных 200 экспериментов 50 раз. Это означает, согласно методу Монте-Карло, что вероятность совершения события равна: $p_A = 50/200 = 0.25$. Вероятность того, что событие не совершится, равна, соответственно, $1 - 0.25 = 0.75$.

Обратите внимание: когда говорят о вероятности, полученной экспериментально, то ее называют **частотью**; слово **вероятность** употребляют, когда хотят подчеркнуть, что речь идет о теоретическом понятии. При большом количестве опытов N частота появления события, полученная экспериментальным путем, стремится к значению теоретической вероятности появления события.

В блоке оценки достоверности (БОД) анализируют степень достоверности статистических экспериментальных данных, снятых с модели (принимая во внимание точность результата ϵ , заданную пользователем) и определяют необходимое для этого количество статистических испытаний. Если колебания значений частоты появления событий относительно теоретической вероятности меньше заданной точности, то экспериментальную частоту принимают в качестве ответа, иначе генерацию случайных входных воздействий продолжают, и процесс моделирования повторяется. При малом числе испытаний результат может оказаться недостоверным. Но чем более испытаний, тем точнее ответ, согласно центральной предельной теореме.

Заметим, что оценивание ведут по худшей из частот. Это обеспечивает достоверный результат сразу по всем снимаемым характеристикам модели.

Пример 1. Решим простую задачу. Какова вероятность выпадения монеты орлом кверху при падении ее с высоты случайным образом?

Начнем подбрасывать монетку и фиксировать результаты бросков (табл. 7.1).

Таблица 7.1.

Результаты испытаний бросания монеты

Количество опытов N	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Значение счетчика выпадения орла N_o	0	0	1	1	2	3	4
Значение счетчика выпадения решки N_p	1	2	2	3	3	3	3
Частость выпадения орла $P_o = N_o/N$	0	0	0.33	0.25	0.4	0.5	0.57
Частость выпадения решки $P_p = N_p/N$	1	1	0.66	0.75	0.6	0.5	0.43

Будем подсчитывать частость выпадения орла как отношение количества случаев выпадения орла к общему числу наблюдений. Посмотрите в табл. 7.1. случаи для $N = 1$,

$N = 2$, $N = 3$ – сначала значения частоты нельзя назвать достоверными. Попробуем построить график зависимости P_o от N – и посмотрим, как меняется частота выпадения орла в зависимости от количества проведенных опытов. Разумеется, при различных экспериментах будут получаться разные таблицы и, следовательно, разные графики. На **рис. 7.11** показан один из вариантов.

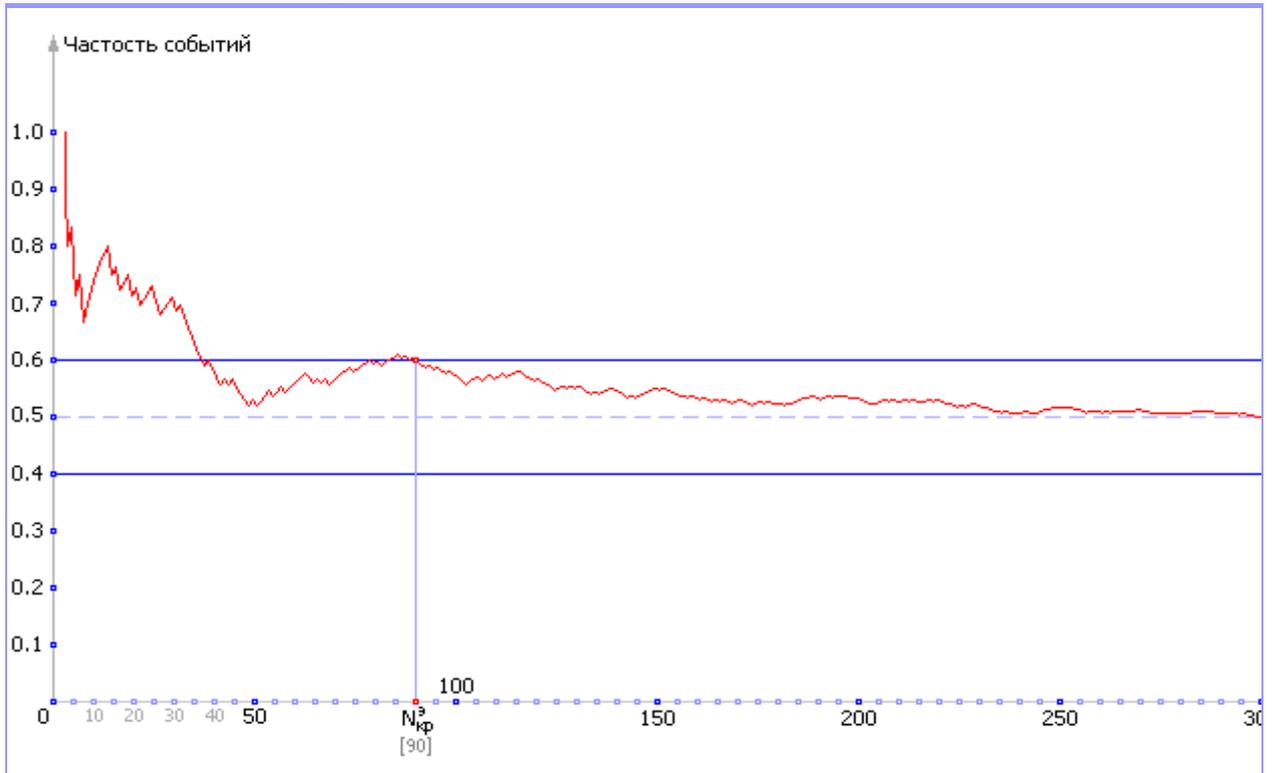


Рис. 7.11. Экспериментальная зависимость частоты появления случайного события от количества наблюдений и ее стремление к теоретической вероятности

Сделаем некоторые выводы.

1. Видно, что при малых значениях N , например, $N = 1$, $N = 2$, $N = 3$ ответу вообще доверять нельзя. Например, $P_o = 0$ при $N = 1$, то есть вероятность выпадения орла при одном броске равна нулю! Хотя всем хорошо известно, что это не так. То есть пока мы получили очень грубый ответ. Однако, посмотрите на график: в процессе **накопления** информации ответ медленно, но верно приближается к правильному (он выделен пунктирной линией). К счастью, в данном конкретном случае правильный ответ нам известен: в идеале, вероятность выпадения орла равна 0.5 (в других, более сложных задачах, ответ нам, конечно, будет неизвестен). Допустим, что ответ нам надо знать с точностью $\varepsilon = 0.1$. Проведем две параллельные линии, отстоящие от правильного ответа 0.5 на расстояние 0.1 (см. **рис. 7.11**). Ширина образовавшегося коридора будет равна 0.2. Как только кривая $P_o(N)$ войдет в этот коридор так, что уже никогда его не покинет, можно остановиться и посмотреть, для какого значения N это произошло. Это и есть **экспериментально вычисленное критическое значение** необходимого количества опытов $N_{кр}^{\varepsilon}$ для определения ответа с точностью $\varepsilon = 0.1$; ε -окрестность в наших рассуждениях играет роль своеобразной трубки точности. Заметьте, что ответы $P_o(91)$, $P_o(92)$ и так далее уже не меняют сильно своих значений (см. **рис. 7.11**); по крайней мере, у них не изменяется первая цифра после запятой, которой мы обязаны доверять по условиям задачи.

2. Причиной такого поведения кривой является действие **центральной предельной теоремы**. Пока здесь мы сформулируем ее в самом простом варианте «Сумма случайных величин есть величина неслучайная». Мы использовали среднюю величину P_o , которая несет в себе информацию о сумме опытов, и поэтому постепенно эта величина становится все более достоверной.

3. Если проделать еще раз этот опыт сначала, то, конечно, его результатом будет

другой вид случайной кривой. И ответ будет другим, хотя примерно таким же. Проведем целую серию таких экспериментов (см. **рис. 7.12**). Такая серия называется **ансамблем реализаций**. Какому же ответу в итоге следует верить? Ведь они, хоть и являются близкими, все же разнятся. На практике поступают по-разному. Первый вариант – вычислить среднее значение ответов за несколько реализаций (см. табл. 7.2).

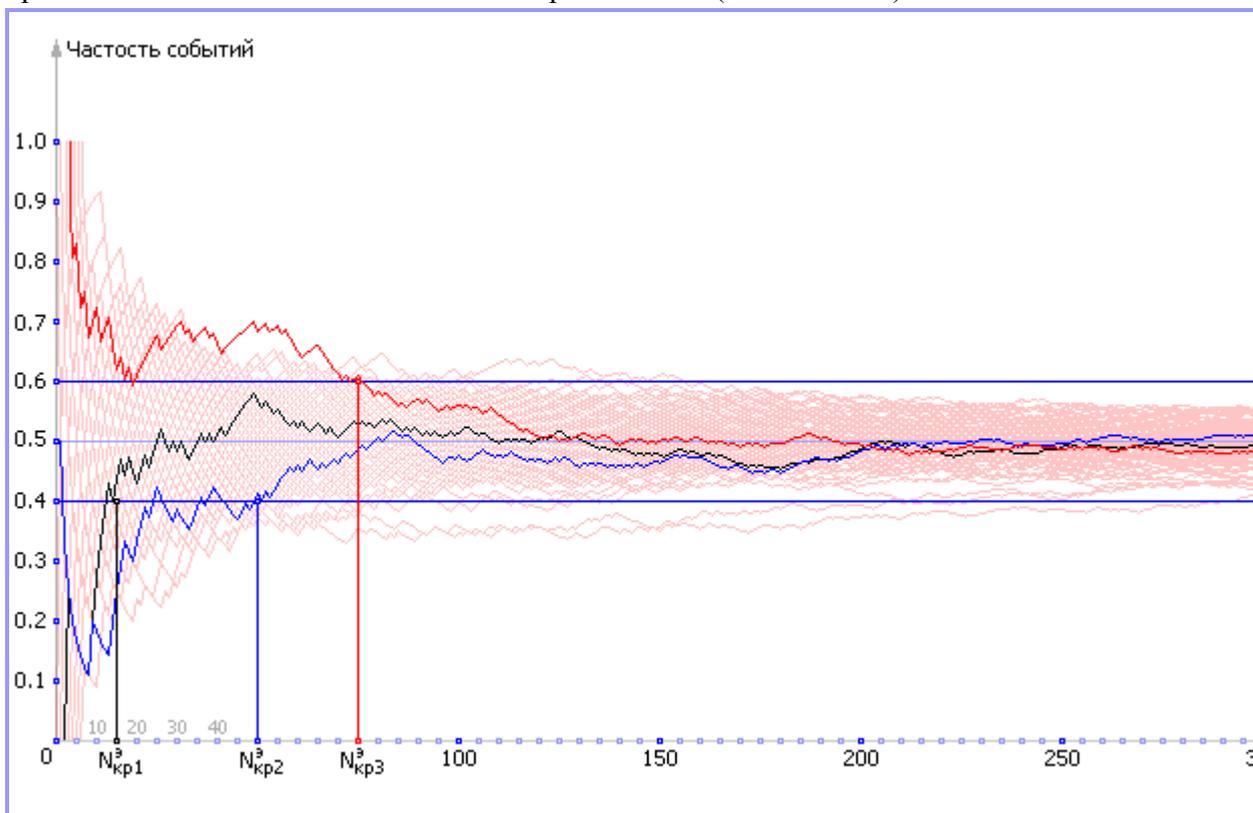


Рис. 7.12. Экспериментально снятый ансамбль случайных зависимостей частоты появления случайного события от количества наблюдений

Мы поставили несколько экспериментов и определяли каждый раз, сколько необходимо было сделать опытов, то есть $N_{кр}^3$. Было проделано 10 экспериментов, результаты которых были сведены в табл. 7.2. По результатам 10-ти экспериментов было вычислено среднее значение $N_{кр}^3$.

Таблица 7.2.

Экспериментальные данные необходимого количества бросков монеты для достижения точности $\varepsilon = 0.1$ при вычислении вероятности выпадения орла

Опыт	$N_{кр}^3$
1	288
2	95
3	50
4	29
5	113
6	210
7	30
8	42
9	39
10	48
Среднее $N_{кр}^3$	94

Таким образом, проведя 10 реализаций разной длины, мы определили, что достаточно **в среднем** было сделать 1 реализацию длиной в 94 броска монеты.

Еще один важный факт. Внимательно рассмотрите график на **рис. 7.12**. На нем нарисовано 100 реализаций – 100 красных линий. Отметьте на нем абсциссу $N = 94$ вертикальной чертой. Есть какой-то процент красных линий, которые не успели пересечь ε -окрестность, то есть $(P^{\text{эксп}} - \varepsilon \leq P^{\text{теор}} \leq P^{\text{эксп}} + \varepsilon)$, и войти в коридор точности до момента $N = 94$. Обратите внимание, таких линий 5. Это значит, что 95 из 100, то есть 95%, линий достоверно вошли в обозначенный интервал.

Таким образом, проведя 100 реализаций, мы добились примерно 95%-ного доверия к полученной экспериментально величине вероятности выпадения орла, определив ее с точностью 0.1. Для сравнения полученного результата вычислим теоретическое значение $N_{\text{кр}}^T$ теоретически. Однако для этого придется ввести понятие доверительной вероятности Q_F , которая показывает, насколько мы готовы верить ответу. Например, при $Q_F = 0.95$ мы готовы верить ответу в 95% случаев из 100. Формула теоретического расчета числа экспериментов имеет вид: $N_{\text{кр}}^T = k(Q_F) \cdot p \cdot (1 - p) / \varepsilon^2$, где $k(Q_F)$ – коэффициент Лапласа, p – вероятность выпадения орла, ε – точность (доверительный интервал). В табл. 7.3 показаны значения теоретической величины количества необходимых опытов при разных Q_F (для точности $\varepsilon = 0.1$ и вероятности $p = 0.5$).

Таблица 21.3.

Теоретический расчет необходимого количества бросков монеты для достижения точности $\varepsilon = 0.1$ при вычислении вероятности выпадения орла

Доверительная вероятность Q_F	Коэффициент Лапласа $k(Q_F)$	Требуемое число опытов $N_{\text{кр}}^T = k(Q_F) \cdot p \cdot (1 - p) / \varepsilon^2$
0.90	2.72	68
0.95	3.84	96
0.99	6.66	167

Как видите, полученная нами оценка длины реализации, равная 94 опытам, очень близка к теоретической, равной 96. Некоторое несовпадение объясняется тем, что, видимо, 10 реализаций недостаточно для точного вычисления $N_{\text{кр}}^T$. Если вы решите, что вам нужен результат, которому следует доверять больше, то измените значение доверительной вероятности. Например, теория говорит нам, что если опытов будет 167, то всего 1-2 линии из ансамбля не войдут в предложенную трубку точности. Но имейте в виду, количество экспериментов с ростом точности и достоверности растет очень быстро.

Второй вариант, используемый на практике – провести **одну** реализацию и **увеличить полученное для нее $N_{\text{кр}}^T$ в 2 раза**. Это считают хорошей гарантией точности ответа (см. **рис. 7.13**). Если присмотреться к ансамблю случайных реализаций, то можно обнаружить, что сходимости частоты к значению теоретической вероятности происходит по кривой, соответствующей обратной квадратичной зависимости от числа экспериментов (см. **рис. 7.14**). Это действительно так получается и теоретически. Если изменять задаваемую точность ε и исследовать количество экспериментов, требуемых для обеспечения каждой из них, то получится табл. 7.4.

Таблица 7.4.

Теоретическая зависимость количества экспериментов, необходимых для обеспечения заданной точности при $Q_F = 0.95$

Точность ε	Критическое число экспериментов $N_{\text{кр}}^T$
0.1	96
0.01	9600
0.001	960000

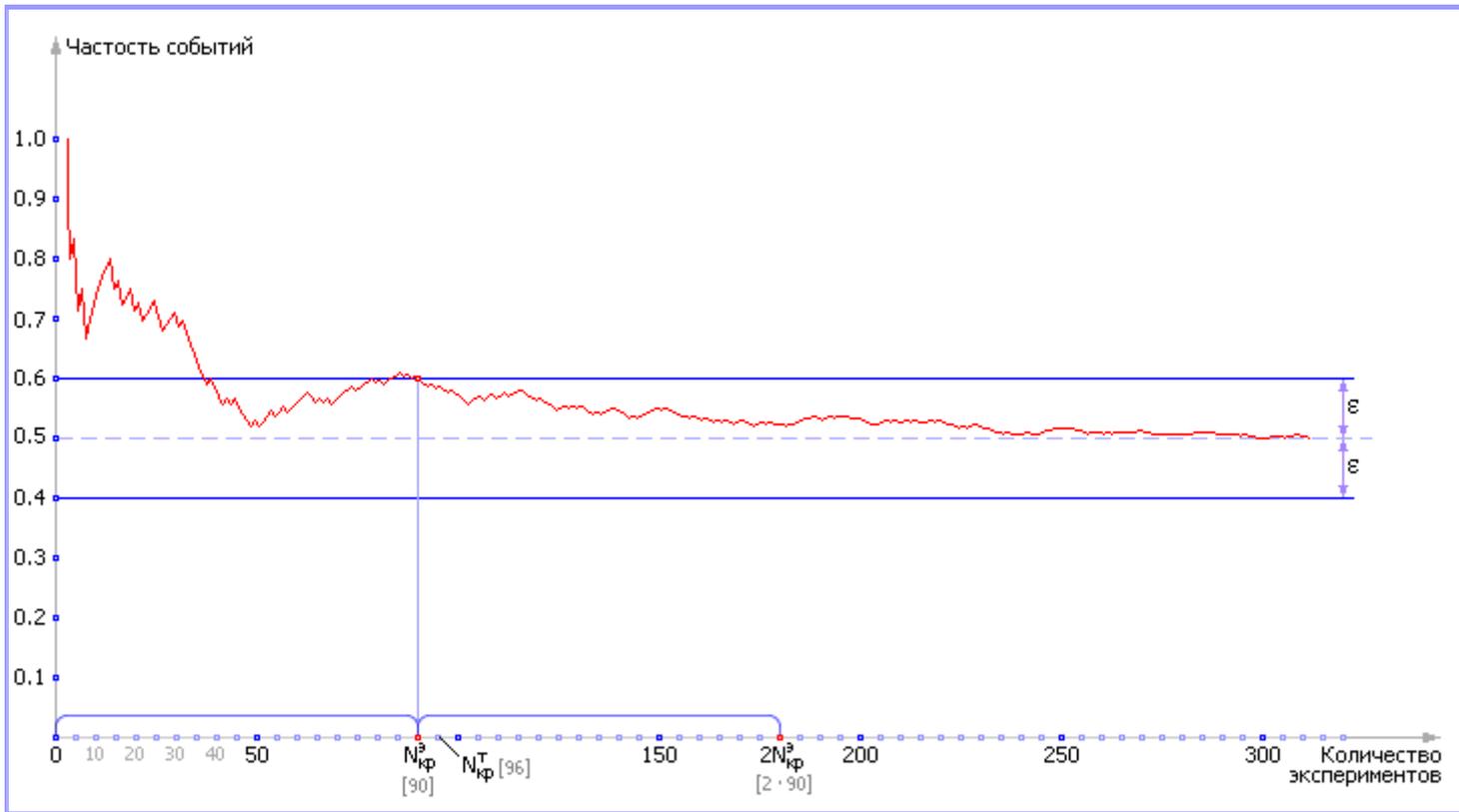


Рис. 7.13. Иллюстрация экспериментального определения $N_{кр}^э$ по правилу «умножь на два»

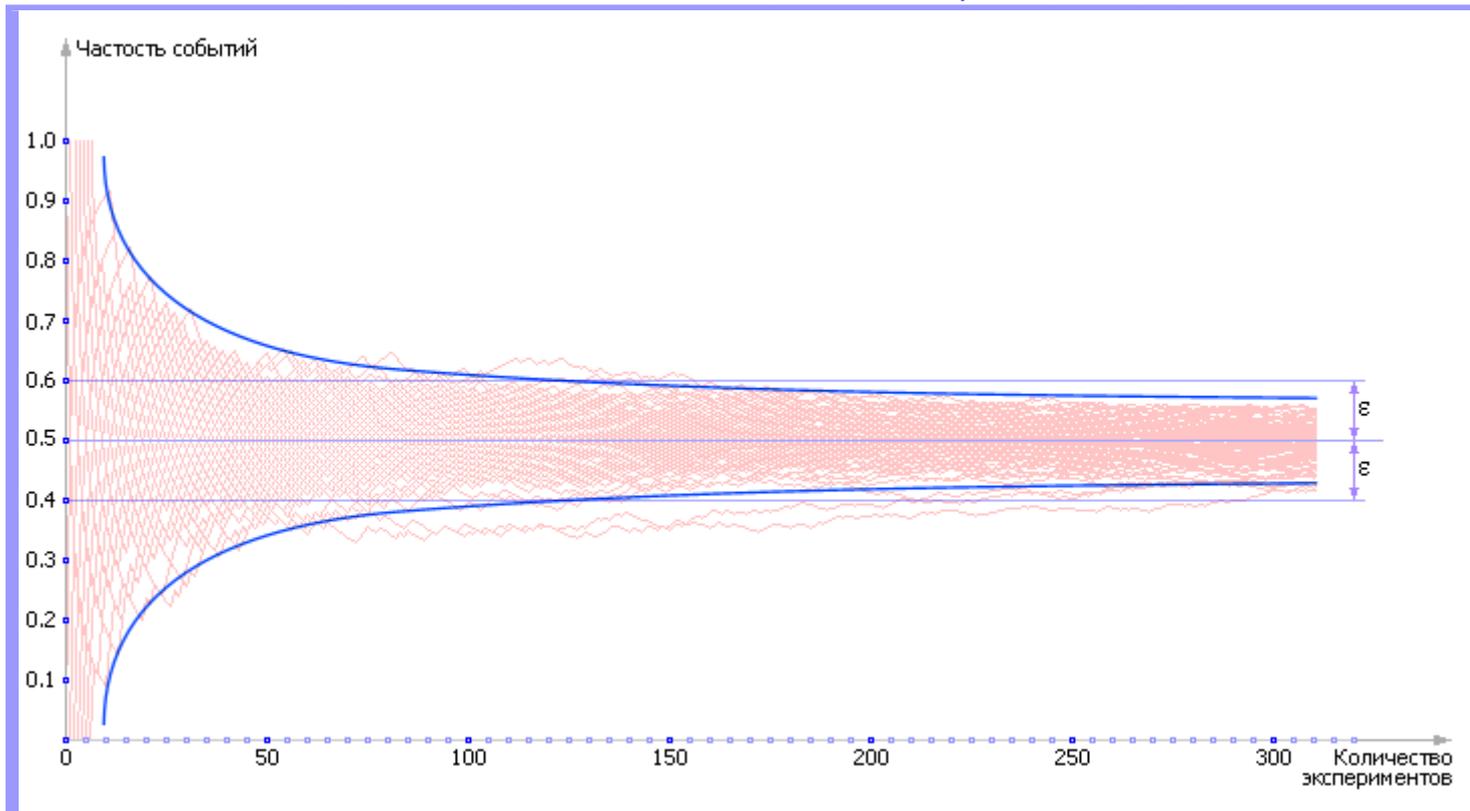


Рис. 7.14. Иллюстрация скорости схождения экспериментально получаемой частоты к теоретической вероятности

Построим по табл. 7.4 график зависимости $N_{кр}^э(\varepsilon)$ (см. [рис. 7.15](#)). Графики подтверждают приведенную выше оценку:

$$\text{точность} \cong \sqrt{\text{объем выборки}}$$

Заметим, что оценок точности может быть несколько. Некоторые из них будут еще обсуждаться.

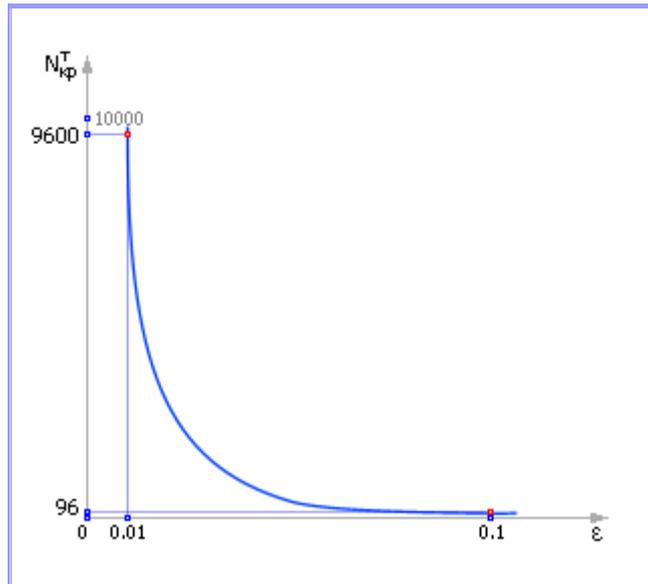


Рис. 7.15. Зависимость числа экспериментов, требуемых для достижения заданной точности ϵ при фиксированном $Q_F = 0.95$

Пример 2. Нахождение площади фигуры методом Монте-Карло. Определим методом Монте-Карло площадь пятиугольника с координатами углов $(0, 0)$, $(0, 10)$, $(5, 20)$, $(10, 10)$, $(7, 0)$. Нарисуем в двухмерных координатах заданный пятиугольник, вписав его в прямоугольник, чья площадь составляет $(10 - 0) \cdot (20 - 0) = 200$ (см. [рис. 7.16](#)).

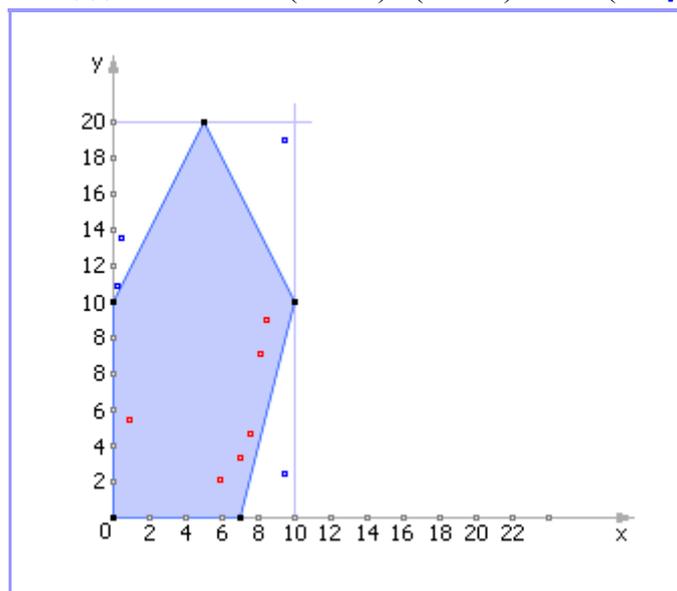


Рис. 7.16. Иллюстрация к решению задачи о площади фигуры методом Монте-Карло

Используем таблицу случайных чисел для генерации пар чисел R, G , равномерно распределенных в интервале от 0 до 1. Число R будет имитировать координату X ($0 \leq X \leq 10$), следовательно, $X = 10 \cdot R$. Число G будет имитировать координату Y ($0 \leq Y \leq 20$), следовательно, $Y = 20 \cdot G$. Сгенерируем по 10 чисел R и G и отобразим 10 точек $(X; Y)$ на [рис. 7.16](#) и в табл. 7.5. Статистическая гипотеза заключается в том, что количество точек, попавших в контур фигуры, пропорционально площади фигуры: $6:10 = S:200$. То есть, по формуле метода Монте-Карло, получаем, что площадь S пятиугольника равна: $200 \cdot 6/10 = 120$. Проследим, как менялась величина S от опыта к опыту (табл. 7.6). Поскольку в ответе все еще меняется значение второго разряда, то возможная неточность составляет пока больше 10%. Точность расчета может быть увеличена с ростом числа испытаний (см. [рис. 7.17](#)).

Таблица 7.5. Решение задачи методом Монте-Карло

Номер точки	R	G	X	Y	Точка (X; Y) попала в прямоугольник?	Точка (X; Y) попала в пятиугольник?
1	0.8109	0.3557	8.109	7.114	Да	Да
2	0.0333	0.5370	0.333	10.740	Да	Нет
3	0.1958	0.2748	1.958	5.496	Да	Да
4	0.6982	0.1652	6.982	3.304	Да	Да
5	0.9499	0.1090	9.499	2.180	Да	Нет
6	0.7644	0.2194	7.644	4.388	Да	Да
7	0.8395	0.4510	8.395	9.020	Да	Да
8	0.0415	0.6855	0.415	13.710	Да	Нет
9	0.5997	0.1140	5.997	2.280	Да	Да
10	0.9595	0.9595	9.595	19.190	Да	Нет
Всего:					10	6

Таблица 7.6. Оценка точности ответа

Количество испытаний N	Оценка вероятности попадания случайной точки в испытываемую область	Оценка площади S методом Монте-Карло
1	$1/1 = 1.00$	200
2	$1/2 = 0.50$	100
3	$2/3 = 0.67$	133
4	$3/4 = 0.75$	150
5	$3/5 = 0.60$	120
6	$4/6 = 0.67$	133
7	$5/7 = 0.71$	143
8	$5/8 = 0.63$	125
9	$6/9 = 0.67$	133
10	$6/10 = 0.60$	120

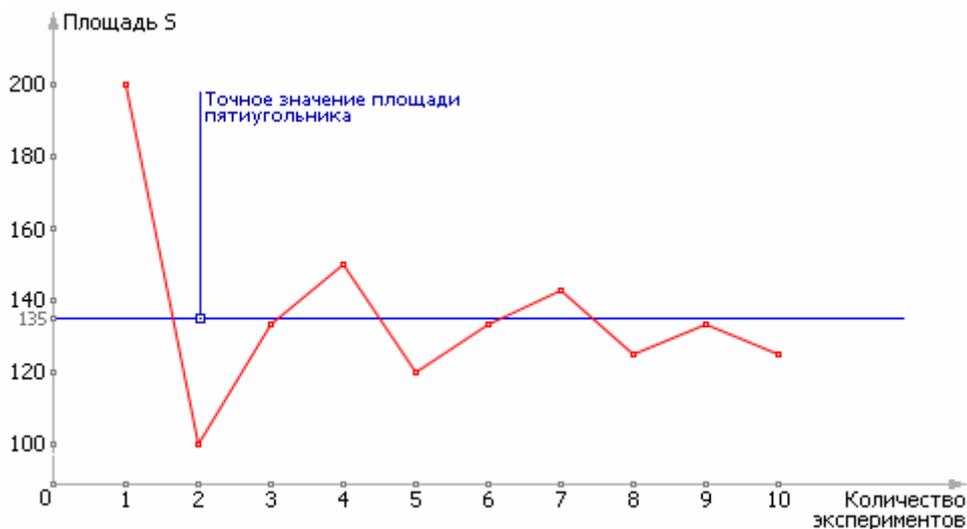


Рис. 7.17. Иллюстрация процесса сходимости экспериментально определяемого ответа к теоретическому результату

на начало лекции
на содержание документа