

Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное учреждение высшего
профессионального образования
«Московский государственный технический университет им. Н. Э. Баумана»

Факультет «Машиностроительные технологии»
Кафедра «Литейные технологии»

Отчет по лабораторной работе № 4
по курсу «Основы научных исследований»

Моделирование случайных процессов методом Монте-Карло

Вариант ...

За основу этого отчета взят пример № 1 из методических указаний

Выполнил: студент гр. МТ5–17
Карпенко Д. Н.

Проверил: асс. каф. МТ5
Карпенко Д. Н.

Москва – 2016

Оглавление

Введение	3
1 Теоретическая часть	4
1.1 Моделирование в условиях неопределенности.....	4
1.2 Метод Монте-Карло	5
1.3 Генерация случайных чисел	8
1.4 Обработка результатов вычислительного эксперимента и анализ качества.....	9
2 Практическая часть.....	11
2.1 Задание	11
2.2 Предварительный этап	11
2.3 Моделирование методом Монте-Карло.....	12
2.4 Оценка результатов моделирования	14
3 Выводы	15

Введение

Цель работы – освоение теоретико-вероятностного подхода к моделированию процессов литейного производства с использованием метода Монте-Карло, закрепление навыков автоматизации вычислений с помощью ЭВМ.

Содержание работы:

1. Ознакомиться с теоретической частью работы по настоящему указанию.
2. На основании исходных данных определить значения функции отклика или параметры работы системы (в зависимости от варианта задания) без учета случайного разброса по факторам.
3. Провести вычислительный эксперимент с использованием метода Монте-Карло с учетом случайных значений факторов. Оценить статистические характеристики функции отклика, при необходимости построить гистограмму.
4. Сравнить полученные результаты с результатом расчета по детерминистической модели. Оценить показатели качества изучаемого процесса и потери от нестабильности факторов (процент брака, коэффициент потерь и т. п.).
5. Сделать выводы.
6. Оформить отчет по работе.

1 Теоретическая часть

1.1 Моделирование в условиях неопределенности

В общем случае математическую модель можно представить как некий оператор \mathbb{F} , который осуществляет однозначное отображение множества входных величин $\Omega_X = \{x_1, x_2, \dots, x_n\}$ на множество выходных $\Omega_Y = \{y_1, y_2, \dots, y_m\}$:

$$\mathbb{F}: \Omega_X \rightarrow \Omega_Y. \quad (1)$$

В дальнейшем мы будем рассматривать математические модели, имеющие только один выходной параметр y , который однозначно определяется из соотношения:

$$y = F(x_1, x_2, \dots, x_n), \quad (2)$$

где F также следует понимать как оператор, в качестве которого может выступать как функция (функциональный оператор), так и алгоритм, включающий в себя, например, логические условия (алгоритмический оператор).

В реальных прикладных исследованиях в качестве y довольно часто выступает некая величина, характеризующая качество изделия или процесса. В этом случае параметр y может служить критерием для оценки пригодности того или иного технологического решения.

Главной особенностью операторов (1) и (2) является то, что они детерминистичны, то есть при одних и тех же значениях входных величин мы всегда будем получать одни и те же значения выходных величин. Однако в реальном производстве параметры технологического процесса подвержены случайным возмущениям, которые приводят к отклонению результирующих показателей от ожидаемых значений. Такая неопределенность в параметрах модели называется стохастической.

Если параметр y является показателем качества, было бы желательно на основании модели (2) уметь предсказывать величину подобных откло-

нений. Разумеется, при этом нужно учитывать, что факторы x_1, x_2, \dots, x_n в общем случае являются стохастическими случайными величинами, каждая из которых подчинена своему закону распределения: $P_1(x_1), P_2(x_2)$ и т. д. Эти законы распределения должны быть установлены для каждого из факторов заранее на основании экспериментальной оценки.

в прикладных исследованиях для моделирования случайных процессов чаще всего используют методы приближенной оценки распределения выходной величины $P(y)$. Одним из самых широко распространенных методов такого рода является метод Монте-Карло.

1.2 Метод Монте-Карло

Получить приблизительное представление о распределении выходной величины y можно с помощью вычислительного эксперимента. В ходе этого эксперимента должно быть получено большое количество реализаций стохастического процесса, описываемого соотношением (2) и набором распределений $\{P_i(x_i)\}, i = 1 \dots n$, после чего производится обработка результатов с помощью обычных методов математической статистики. В этом и состоит суть метода Монте-Карло.

Алгоритм метода Монте-Карло состоит из следующих этапов:

1. Задаются начальные условия: математическая модель вида (2) и распределения для случайных факторов $\{P_i(x_i)\}, i = 1 \dots n$. Эти распределения могут быть как непрерывными (равномерное, нормальное и т. п.), так и задаваться гистограммой, построенной по результатам измерения. При этом некоторые из факторов могут моделироваться как детерминистические – тогда они задаются как постоянная величина.
2. Задаются точностью (величиной ε -окрестности) итогового распределения и доверительной вероятностью $p(\varepsilon)$, на основании которых вычисляют минимальное количество опытов N_{\min} .

3. На основании распределений $\{P_i(x_i)\}$ по каждому из факторов формируется случайная выборка объемом $N > N_{\min}$. Таким образом, мы получаем нечто вроде матрицы:

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1N} \\ x_{21} & x_{22} & \cdots & x_{2N} \\ \vdots & \vdots & x_{ij} & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nN} \end{pmatrix},$$

где x_{ij} означает j -ю реализацию i -го фактора: $i = 1 \dots n$; $j = 1 \dots N$.

4. Полученные случайные значения факторов подставляются в модель (2) и вычисляется N значений выходной величины y , которую в данном случае будем называть функцией отклика. Элементы матрицы X подставляются в модель по столбцам:

$$y_j = F(x_{1j}, x_{2j}, \dots, x_{nj}).$$

5. Полученная выборка $\{y_j\}$, $j = 1 \dots N$, обрабатывается посредством обычных методов математической статистики, например: определяется среднее значение и дисперсия, строится гистограмма распределения, по возможности делается предположение о законе распределения, которому подвержены значения функции отклика y . Полезно также определить максимальное и минимальное значения в выборке. Результаты моделирования заносятся в таблицу вида таблицы 1.

На основании статистической обработки делаются выводы о показателях моделируемого процесса.

При генерации значений случайных факторов в некоторых случаях нужно следить за тем, чтобы они не выходили за определенные границы. Ограничения могут накладываться физическим смыслом процесса, либо исследуемая модель (например, эмпирическая формула) справедлива только на ограниченных интервалах значений факторов.

Как и всякий метод численного моделирования, метод Монте-Карло является приближительным. Точность полученного закона распределения зависит от числа опытов N , следствием чего является необходимость проведения большого числа экспериментов.

Точность результата в методе Монте-Карло можно оценить через вероятность попадания оцениваемой постоянной вероятности P в окрестность полученных в эксперименте значений относительной частоты с некоторой наперед заданной точностью ε . Из интегральной теоремы Муавра – Лапласа можно получить следующую оценку вероятности:

$$p\left(\left|\frac{m}{N} - P\right| \leq \varepsilon\right) = 2 \cdot \Phi\left(\varepsilon \cdot \sqrt{\frac{N}{P(1-P)}}\right), \quad (3)$$

где $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt$ – функция Лапласа.

Полезно знать соотношения между функцией Лапласа Φ и другими часто используемыми статистическими функциями:

$$\Phi(x) = \frac{1}{2} \operatorname{erf} \frac{x}{\sqrt{2}} = F_{N(0,1)}(x) - \frac{1}{2}, \quad (4)$$

где erf – так называемая функция ошибок;

$F_{N(0,1)}$ – интегральная функция стандартного нормального распределения.

Анализ функции (3) показывает, что она имеет минимум при $P = 0,5$ (при прочих постоянных параметрах). Это значит, что для оценки точности достаточно вычислить $p(\varepsilon)$ при этом значении вероятности. С учетом этого, а также соотношений (4) можно записать выражение для вероятности (3) через две другие функции:

$$p(\varepsilon) = 2 \cdot \Phi(2\varepsilon \cdot \sqrt{N}) = \operatorname{erf}(\varepsilon \cdot \sqrt{2N}) = 2 \cdot F_{N(0,1)}(2\varepsilon \cdot \sqrt{N}) - 1. \quad (5)$$

Из соотношений (5) можно также получить оценку минимального числа опытов, которые нужно провести, чтобы получить требуемое распределение, лежащее в ε -окрестности с заданной вероятностью $p(\varepsilon)$:

$$N_{\min} = \left(\frac{\Phi^{-1}(p(\varepsilon)/2)}{2\varepsilon} \right)^2 = \frac{1}{2} \left(\frac{\operatorname{erf}^{-1} p(\varepsilon)}{\varepsilon} \right)^2 = \left(\frac{F_{N(0,1)}^{-1} \left(\frac{p(\varepsilon)+1}{2} \right)}{2\varepsilon} \right)^2, \quad (6)$$

где Φ^{-1} , erf^{-1} и $F_{N(0,1)}^{-1}$ – обратные функции для функции Лапласа, функции ошибок и интегральной функции стандартного нормального распределения соответственно.

1.3 Генерация случайных чисел

Главной проблемой метода Монте-Карло является генерация больших массивов случайных чисел по заданным законам распределения. Существует два основных способа получения случайных чисел:

- аппаратные генераторы случайных чисел (ГСЧ);
- генераторы псевдослучайных чисел (ГПСЧ).

Применение аппаратных ГСЧ ограничено по ряду причин: некоторые из них необходимо интегрировать с компьютером, их свойства могут быть нестабильными, генерируемые последовательности являются невоспроизводимыми, а генерация больших последовательностей может занимать значительное время. Поэтому для решения задач моделирования случайных процессов в большинстве случаев применяются генераторы псевдослучайных чисел.

Генератор псевдослучайных чисел – это алгоритм, позволяющий получать последовательности чисел, которые по своим свойствам мало отличаются от последовательностей истинно случайных чисел. Во многих случаях ГПСЧ представляет собой рекуррентное соотношение, формирующее регулярные последовательности псевдослучайных чисел в неко-

тором диапазоне. Наиболее известным и распространенным является линейный конгруэнтный метод, при котором последовательность генерируется в соответствии с соотношением:

$$r_k = (a \cdot r_{k-1} + c) \bmod m, r_0 \leq m, \quad (7)$$

где a, c, m – определенные константы: $0 \leq a < m, 0 \leq c < m, m \leq 2$;

r_0 – начальное число последовательности;

\bmod – операция вычисления остатка от деления.

Начальное значение r_0 либо определяется случайно (например, с помощью аппаратного ГСЧ), либо назначается исследователем. Константы a, c и m подбираются специальным образом, поскольку при неправильном выборе генерируемая последовательность не будет удовлетворять требованиям, которые накладываются на истинно случайные числа.

Из последовательности, генерируемой соотношением (7), легко получить последовательность псевдослучайных чисел, равномерно распределенную на интервале от 0 до 1:

$$R_k(0, 1) = r_k / m. \quad (8)$$

Последовательность (8) может быть использована для получения последовательности псевдослучайных чисел с любым наперед заданным непрерывным распределением. Поэтому случайные величины, равномерно распределенные на интервале от 0 до 1, еще называют базовыми случайными величинами.

Большинство существующих математических программ позволяют генерировать большие последовательности псевдослучайных величин с заданным законом распределения.

1.4 Обработка результатов вычислительного эксперимента и анализ качества

Результатом вычислительного эксперимента, проведенного с помощью метода Монте-Карло на модели (2) является набор значений функ-

ции отклика $\{y_j\}, j = 1 \dots N$, который, по сути, моделирует случайный разброс выходного параметра y . Первичная обработка полученных значений проводится стандартными статистическими методами. Так, можно определить среднее значение функции отклика

$$\bar{y} = \frac{1}{N} \sum_{j=1}^N y_j \quad (9)$$

и дисперсию значений функции отклика

$$s^2 = \frac{1}{N-1} \sum_{j=1}^N (y_j - \bar{y})^2. \quad (10)$$

В данном случае определяется несмещенная дисперсия, поскольку величина выборки ограничена.

Кроме этого, может быть построена гистограмма распределения величины y , по которой можно судить о виде закона распределения. Гистограмма – это столбчатая диаграмма, характеризующая частоту попадания случайной величины в различные интервалы значений. Гистограмму распределения по выборке предпочтительно строить в относительных частотах. Границы интервалов следует выбирать с учетом минимального и максимального значений таким образом, чтобы в крайние интервалы попадало как можно меньше значений функции отклика.

После первичной обработки значений функции отклика необходимо оценить показатели, которые будут определять качество исследуемого объекта или процесса.

Часто функция отклика y – это критерий качества, для которого задано некое критическое значение $y_{кр}$, служащее границей, отделяющей качественное изделие от бракованного (при этом само $y_{кр}$ в ряде случаев может быть случайной величиной). В этом случае можно посчитать, какой процент значений функции отклика удовлетворяет критерию качества $y_{кр}$ (например, больше или меньше его), а исходя из этого мы можем определить ожидаемый процент брака.

2 Практическая часть

2.1 Задание

С помощью метода Монте-Карло определить статистические характеристики и ожидаемый процент брака по прочности на разрыв σ_B при производстве серого чугуна, если установлена эмпирическая зависимость между пределом прочности (МПа) и химическим составом чугуна (%):

$$\sigma_B = 477,335 - 82,025 \cdot C + 34,675 \cdot Si - 4,475 \cdot Mn, \quad (11)$$

где C , Si , Mn – процентное содержание углерода, кремния и марганца соответственно.

Известно, что при производстве чугуна используют следующий химический состав: $C = 3,4 \%$, $Si = 1,8 \%$, $Mn = 0,7 \%$. Экспериментальным путем установлено, что химический состав из-за неточности дозирования и других случайных возмущений подвержен случайным отклонениям, хорошо согласующимся с нормальным распределением. Стандартные отклонения для каждого из компонентов: $s_C = 0,5 \%$; $s_{Si} = 0,3 \%$; $s_{Mn} = 0,2 \%$.

Пригодными считаются образцы, чья прочность на разрыв превышает $\sigma_{B_{кр}} = 240$ МПа.

Количество опытов должно обеспечить попадание полученного распределения в ε -окрестность действительной вероятности при точности $\varepsilon = \pm 5 \%$ с вероятностью $p(\varepsilon) = 0,97$.

2.2 Предварительный этап

Прежде всего, оценим прочность на разрыв по детерминистической модели (11) при принятом содержании компонентов сплава, без учета возможных отклонений:

$$\sigma_B^0 = 477,335 - 82,025 \cdot 3,4 + 34,675 \cdot 1,8 - 4,475 \cdot 0,7 = 257,733 \text{ (МПа)}.$$

Видно, что данная прочность удовлетворяет заявленным требованиям с запасом и обеспечивает качество сплава по данному показателю.

Моделирование методом Монте-Карло осуществляется в соответствии с алгоритмом, изложенным в подразд. 1.2.

Для моделирования с учетом случайных отклонений по химическому составу необходимо задаться законами распределения, которым будут подвержены факторы. На основании исходных данных можно принять, что содержание каждого из компонентов есть величина случайная, подчиненная нормальному закону распределения с математическим ожиданием, равным нормативному содержанию, используемому в производстве, и среднеквадратическим отклонением, равным стандартному отклонению по каждому из факторов. Характеристики распределений по каждому из факторов приведены в таблице 1.

Таблица 1 – Характеристики распределений для случайных факторов

Характеристики	Факторы		
	<i>C</i>	<i>Si</i>	<i>Mn</i>
Вид распределения	нормальное		
Математическое ожидание μ	3,4	1,8	0,7
Среднеквадратическое отклонение σ	0,5	0,3	0,2

Минимальное количество опытов определяем по формуле (6), а именно с использованием последнего выражения в этой формуле:

$$N_{\min} = \left(\frac{F_{N(0,1)}^{-1} \left(\frac{p(\varepsilon) + 1}{2} \right)}{2\varepsilon} \right)^2 = \left(\frac{F_{N(0,1)}^{-1} \left(\frac{0,97 + 1}{2} \right)}{2 \cdot 0,05} \right)^2 \approx 471.$$

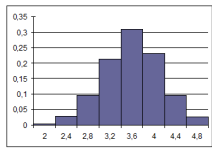
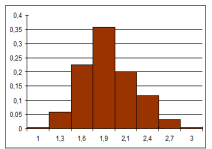
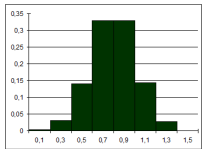
Для вычисления значения функции в числителе мы воспользовались функцией **НОРМСТОБР** из программы *Microsoft Office Excel*. Примем количество опытов $N = 500$.

2.3 Моделирование методом Монте-Карло

В соответствии с таблицей 1 в программе *Microsoft Office Excel* формируем три массива псевдослучайных чисел объемом 500 значений каждый. Для каждой выборки проводим статистическую обработку по фор-

мулам (9) и (10) и строим гистограмму. Полученные случайные значения факторов подставляем в формулу (11) и определяем значение функции отклика, в качестве которой здесь выступает прочность на разрыв. Данные по основному этапу метода Монте-Карло сведены в таблицу 2.

Таблица 2 – План реализации метода Монте-Карло

Эксперименты	Факторы			Функция отклика
	$C, \%$	$Si, \%$	$Mn, \%$	$\sigma_B, \text{МПа}$
j				
1	2,6147	1,5534	0,3859	315,0018
2	3,3448	1,4818	0,6779	251,3292
3	3,4536	2,6799	0,7214	283,7511
•	•	•	•	•
•	•	•	•	•
•	•	•	•	•
500	4,1135	2,0606	0,9854	206,9642
Статистические показатели выборки				
Среднее значение	3,3920	1,7852	0,6968	257,8936
Стандартное отклонение	0,5140	0,3180	0,2056	44,8368
Минимальное значение	1,8689	0,9387	0,0876	149,6086
Максимальное значение	4,6495	2,7946	1,1998	381,3668
Гистограмма				см. рис. 1

Видно, что статистические характеристики сделанных выборок незначительно отличаются от заданных значений, что вполне ожидаемо. Среднее значение прочности мало отличается от значения, рассчитанного по детерминистической модели, и так же удовлетворяет критерию качества. Гистограмма для значений прочности представлена на рис. 1.

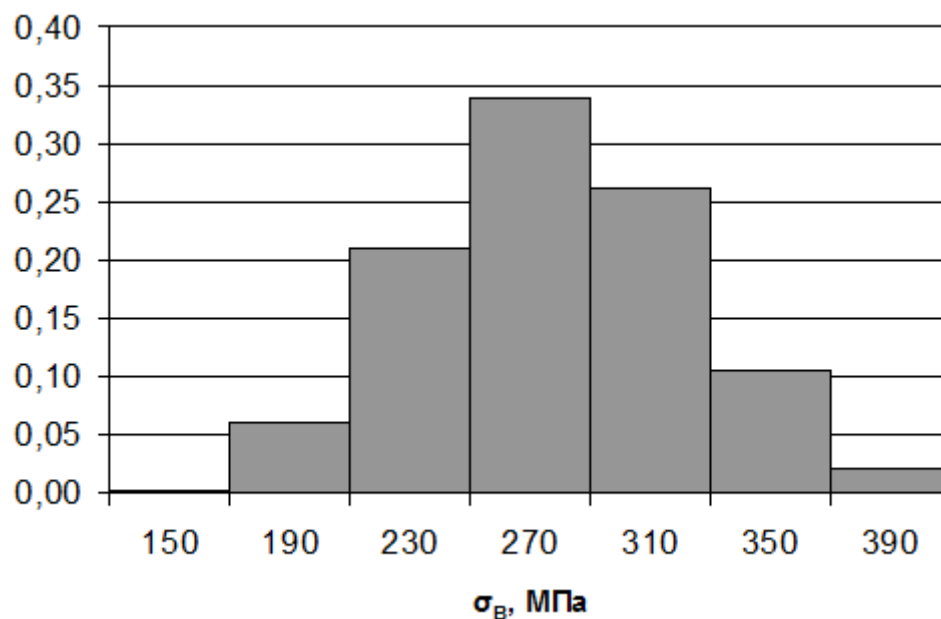


Рис. 1 – Гистограмма для значений прочности на разрыв

2.4 Оценка результатов моделирования

Из гистограммы на рис. 1 видно, что довольно большая доля значений лежит ниже критического значения. Согласно оценке, проведенной средствами программы *Microsoft Office Excel*, доля значений меньше 240 МПа составляет 35,8 %.

3 Выводы

1. Расчет по детерминистической модели (11) показывает, что прочность сплава с запасом превышает требуемое значение, а используемый химический состав может быть пригодным для использования.

2. Полученная в результате моделирования методом Монте-Карло средняя прочность слабо отличается от прочности, рассчитанной по детерминистической модели, и также удовлетворяет требованиям качества. Тем не менее, разброс значений прочности, вызванный случайными колебаниями химического состава, может привести к существенной величине брака (35,8 % по результатам моделирования).

3. Снижения столь высокого процента брака можно добиться одним из двух способов:

- поскольку закон распределения значений прочности близок к симметричному (см. рис. 1) можно изменить исходный химический состав с тем, чтобы повысить среднюю прочность;
- повысить точность дозирования исходных компонентов, тем самым уменьшив итоговую дисперсию функции отклика.

Эффективность каждого из методов может быть оценена моделированием, проведенным для новых параметров исходных данных.