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 № 1

по курсу «Основы научных исследований»

Построение и анализ эмпирической математической модели

Вариант 35

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

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

Москва – 2016

Оглавление

Введение 3
1 Теоретические основы 4
1.1 Постановка задачи 4
1.2 Выбор функциональной зависимости 4
1.3 Определение параметров эмпирической модели 6
1.4 Статистический анализ эмпирической модели 7
2 Практическая часть 9
2.1 Задание 9
2.2 Определение эмпирической зависимости 9
2.3 Результаты расчетов 12
3 Выводы 15

Введение

Цель работы – овладение методами выбора, построения и анализа эмпирических математических моделей литейных процессов, закрепление навыков работы с ЭВМ.

Содержание работы:

1. Ознакомиться с теоретической частью работы по настоящему указанию.
2. На основании экспериментальных данных, соответствующих указанному преподавателем варианту, произвести выбор типа математической модели.
3. Методом наименьших квадратов определить параметры математической модели.
4. Проверить адекватность математической модели экспериментальным данным, произведя необходимые расчеты на ЭВМ.
5. Оформить отчет.

1 Теоретические основы

1.1 Постановка задачи

При изучении многих процессов ввиду сложности или из-за недостаточно ясного их понимания не представляется возможным получить математическое описание на основе анализа механизма явлений. В этом случае пользуются построением эмпирических моделей процессов и явлений. На практике часто именно эмпирическая модель является единственным доступным инструментом анализа процессов.

Пусть, например, при изучении функциональной зависимости $y = f(x)$ с помощью эксперимента получен ряд измерений величин x и y: $\{x_i, y_i\}$, $i = 1, 2,..., n$. Задача построения эмпирической модели состоит в том, чтобы выбрать такую формулу: $\hat{y} = \hat{f}(x)$, значения которой при $x = x_i$ возможно мало отличались бы от опытных данных y_i. Геометрически задача построения эмпирической формулы состоит в проведении кривой, возможно «ближе» примыкающей к системе точек $\{x_i, y_i\}$.

Построение эмпирической модели включает следующие этапы:

- выбор формы функциональной зависимости;
- определение ее наилучших параметров;
- статистический анализ полученной математической модели.

1.2 Выбор функциональной зависимости

При выборе функциональной зависимости предпочтение отдается простым формулам. Универсального способа решения этой задачи не существует. Если экспериментальные данные имеют характер монотонной зависимости (возрастающей или убывающей), то процедура выбора вида эмпирической формулы может быть формализована. Рассмотрим эту процедуру для следующих наиболее часто применяемых зависимостей:

$$y = a \cdot x + b; \tag{1}$$

$$y = b \cdot x^a; \tag{2}$$

$$y = b \cdot a^x; \tag{3}$$

$$y = \frac{a}{x} + b; \tag{4}$$

$$y = \frac{1}{a \cdot x + b}; \tag{5}$$

$$y = \frac{x}{a + b \cdot x}; \tag{6}$$

$$y = a \cdot \lg x + b. \tag{7}$$

Особенностью моделей (1)–(7) является то, что с помощью элементарных преобразований, представленных в табл. 1, все они могут быть сведены к линейной зависимости.

Таблица 1 – Расчетные параметры для моделей (1)–(7)

№ п/п	x_s	y_s	Уравнение	Преобразование к линейной модели $Y = AX + B$
1	$0{,}5 \cdot (x_1 + x_n)$	$0{,}5 \cdot (y_1 + y_n)$	$y = a \cdot x + b$	$Y = y;\ X = x;\ A = a;\ B = b$
2	$\sqrt{x_1 \cdot x_n}$	$\sqrt{y_1 \cdot y_n}$	$y = b \cdot x^a$	$Y = \lg y;\ X = \lg x;\ A = a;\ B = \lg b$
3	$0{,}5 \cdot (x_1 + x_n)$	$\sqrt{y_1 \cdot y_n}$	$y = b \cdot a^x$	$Y = \lg y;\ X = x;\ A = \lg a;\ B = \lg b$
4	$\frac{2 \cdot x_1 \cdot x_n}{x_1 + x_n}$	$0{,}5 \cdot (y_1 + y_n)$	$y = \frac{a}{x} + b$	$Y = y;\ X = \frac{1}{x};\ A = a;\ B = b$
5	$0{,}5 \cdot (x_1 + x_n)$	$\frac{2 \cdot y_1 \cdot y_n}{y_1 + y_n}$	$y = \frac{1}{a \cdot x + b}$	$Y = \frac{1}{y};\ X = x;\ A = a;\ B = b$
6	$\frac{2 \cdot x_1 \cdot x_n}{x_1 + x_n}$	$\frac{2 \cdot y_1 \cdot y_n}{y_1 + y_n}$	$y = \frac{x}{a + b \cdot x}$	$Y = \frac{1}{y};\ X = \frac{1}{x};\ A = a;\ B = b$
7	$\sqrt{x_1 \cdot x_n}$	$0{,}5 \cdot (y_1 + y_n)$	$y = a \cdot \lg x + b$	$Y = y;\ X = \lg x;\ A = a;\ B = b$

Выбор вида эмпирической модели производят следующим образом.

Используя данные эксперимента для крайних точек (x_1, y_1) и (x_n, y_n), по табл. 1 определяют x_s и y_s для всех семи зависимостей. Далее значение y_s сравнивают с экспериментально наблюдаемым. Конечно, может оказаться, что среди экспериментальных величин отсутствуют измерения при $x = x_s$ – тогда соответствующее значение для величины y определяют с помощью линейной интерполяции по формуле:

$$\hat{y}_s = y_i + \frac{y_{i+1} - y_i}{x_{i+1} - x_i} \cdot \left(x_s - x_i \right), \tag{8}$$

где x_i и x_{i+1} – ближайшие промежуточные значения, между которыми содержится x_s ($x_i < x_s < x_{i+1}$);

y_i и y_{i+1} – соответствующие им средние значения функции отклика.

По результатам сравнения выбирается такая эмпирическая модель, для которой величина $\left| y_s - \hat{y}_s \right|$ минимальна.

1.3 Определение параметров эмпирической модели

После выбора типа эмпирической модели определяют ее параметры, которыми являются для рассматриваемых зависимостей коэффициенты a и b. Для определения наилучших значений коэффициентов a и b используется метод наименьших квадратов (МНК). Поскольку каждая из рассматриваемых моделей может быть сведена к линейной, МНК применяется для получения линейной зависимости:

$$Y = A \cdot X + B. \tag{9}$$

Зависимости между параметрами линейной модели (обозначены заглавными буквами) и параметрами исходной модели (строчные буквы) приведены в табл. 1.

Сущность метода наименьших квадратов применительно к получению наилучшей линейной модели состоит в следующем. Наилучшими коэффициентами A и B модели (9) считаются те, для которых будет минимальной сумма квадратов отклонений, вычисляемая следующим образом:

$$S\left(A, B\right) = \sum_{i=1}^{n} \left(AX_i + B - \overline{Y}_i \right)^2 \cdot p_i, \tag{10}$$

где n – число опытов;

$\overline{Y}_i$ – среднее значение функции отклика в i-ом опыте;

p_i – число параллельных измерений в i-ом опыте.

Отсюда, используя необходимые условия экстремума функции двух переменных $S = S(A, B)$, получаем систему для определения коэффициентов A и B:

$$\begin{aligned} \frac{\partial S}{\partial A} &= 0; \\ \frac{\partial S}{\partial B} &= 0. \end{aligned} \qquad (11)$$

После преобразований с учетом (10) получим формулы для определения коэффициентов A и B:

$$\begin{aligned} A &= \frac{\left(\sum_{i=1}^{n} p_i X_i\right)\left(\sum_{i=1}^{n} p_i \overline{Y}_i\right) - \left(\sum_{i=1}^{n} p_i X_i \overline{Y}_i\right)\left(\sum_{i=1}^{n} p_i\right)}{\left(\sum_{i=1}^{n} p_i X_i\right)^2 - \left(\sum_{i=1}^{n} p_i X_i^2\right)\left(\sum_{i=1}^{n} p_i\right)}; \\ B &= \frac{\left(\sum_{i=1}^{n} p_i X_i \overline{Y}_i\right)\left(\sum_{i=1}^{n} p_i X_i\right) - \left(\sum_{i=1}^{n} p_i \overline{Y}_i\right)\left(\sum_{i=1}^{n} p_i X_i^2\right)}{\left(\sum_{i=1}^{n} p_i X_i\right)^2 - \left(\sum_{i=1}^{n} p_i X_i^2\right)\left(\sum_{i=1}^{n} p_i\right)}. \end{aligned} \qquad (12)$$

1.4 Статистический анализ эмпирической модели

Процедура проверки гипотезы о том, что полученная эмпирическая модель удовлетворительно или адекватно описывает экспериментальные данные, состоит в следующем. По результатам эксперимента вычисляется дисперсия ошибок $S_{ош}^2$, определяющая меру рассеивания, вызванного экспериментальной ошибкой:

$$S_{ош}^2 = \frac{\sum_{i=1}^{n}\sum_{j=1}^{p_i}\left(Y_{ij} - \overline{Y}_i\right)^2}{\sum_{i=1}^{n} p_i - n}. \qquad (13)$$

Дисперсия ошибок $S_{ош}^2$ сравнивается с дисперсией адекватности $S_{ад}^2$:

$$S_{ад}^2 = \frac{1}{n-2}\sum_{i=1}^{n} p_i \left(\overline{Y}_i - \hat{Y}_i\right)^2, \qquad (14)$$

где $\hat{Y}_i = A \cdot X_i + B$ – значение величины Y, предсказанное полученной эмпирической моделью для каждого из значений X_i.

С использованием этих двух дисперсий рассчитывается значение статистического критерия Фишера, с использованием которого и оценивается адекватность полученной модели:

$$F_p = \frac{S_{ад}^2}{S_{ош}^2}. \qquad (15)$$

Расчетное значение критерия Фишера сравнивается с табличным F_m, которое выбирается в соответствии с доверительной вероятностью P и числом степеней свободы числителя $\nu_{ад}$ и знаменателя $\nu_{ош}$. Величина доверительной вероятности в данной работе принята равной 0,95. Числа степеней свободы вычисляются по следующим формулам:

$$\nu_{ад} = n - 2; \qquad (16)$$

$$\nu_{ош} = \sum_{i=1}^{n} p_i - n. \qquad (17)$$

2 Практическая часть

2.1 Задание

На основании экспериментальных данных, представленных в табл. 2, определить зависимость предела прочности $\sigma_в$ немодифицированного алюминиевого сплава АЛ2 от массивности тела отливки, определяемой ее диаметром $d_{отл}$:

$$\sigma_в = f\left(d_{отл}\right). \qquad (18)$$

Таблица 2 – Результаты эксперимента

i	$d_{отл}$, мм	$\sigma_в$, МПа				$\bar{\sigma}_в$, МПа
1	15	201	206	210	207	206
2	30	162	168	170	172	168
3	45	150	156	159	159	156
4	60	130	134	135	137	134

2.2 Определение эмпирической зависимости

Экспериментальные данные свидетельствуют о том, что зависимость (18) является монотонной, так что будем искать ее в виде одной из формул (1)–(7). При этом x будет означать диаметр отливки $d_{отл}$, а y – предел прочности $\sigma_в$.

Выбор вида функциональной зависимости

Для выбора вида функциональной зависимости необходимо вычислить координаты промежуточной точки (x_s, y_s) по формулам из табл. 1, определить значения функции отклика $\hat{y}_s$ при x_s, используя при необходимости формулу (8), и определить разность $\left|y_s - \hat{y}_s\right|$. Результаты расчетов сведены в табл. 3

Таблица 3 – Расчетные параметры моделей

№ п/п	Уравнение	x_s	y_s	$\hat{y}_s$	$\lvert y_s - \hat{y}_s \rvert$
1	$y = a \cdot x + b$	37,5	170,00000	162,0	8,00000
2	$y = b \cdot x^a$	30,0	166,14452	168,0	1,85548
3	$y = b \cdot a^x$	37,5	166,14452	162,0	4,14542
4	$y = \frac{a}{x} + b$	24,0	170,00000	183,2	13,20000
5	$y = \frac{1}{a \cdot x + b}$	37,5	162,37647	162,0	0,37647
6	$y = \frac{x}{a + b \cdot x}$	24,0	162,37647	183,2	20,82353
7	$y = a \cdot \lg x + b$	30,0	170,00000	168,0	2,00000

Как видно из табл. 3, наиболее подходящей моделью является уравнение (5).

Расчет коэффициентов модели

Для определения коэффициентов преобразуем модель (5) к линейному виду. Для этого преобразуем данные табл. 2 в соответствии с формулами табл. 1: $X = x$, $Y = 1/y$.

Таблица 4 – Преобразованные данные эксперимента

i	X	Y				$\bar{Y}$
1	15	0,004975	0,004854	0,004762	0,004831	0,00485558
2	30	0,006173	0,005952	0,005882	0,005814	0,00595538
3	45	0,006667	0,006410	0,006289	0,006289	0,00641388
4	60	0,007692	0,007463	0,007407	0,007299	0,00746542

Для расчета коэффициентов по формулам (12) предварительно вычислим входящие в них суммы. Поскольку $p_i = p$ в каждом из опытов одинаковое, суммы можно преобразовать к более простому виду.

$$\sum_{i=1}^{n} p_i = p \cdot n = 4 \cdot 4 = 16;$$

$$\sum_{i=1}^{n} p_i \cdot X_i = p \cdot \sum_{i=1}^{n} X_i = 4 \cdot (15 + 30 + 45 + 60) = 600;$$

$$\sum_{i=1}^{n} p_i \cdot X_i^2 = p \cdot \sum_{i=1}^{n} X_i^2 = 4 \cdot \left(15^2 + 30^2 + 45^2 + 60^2\right) = 27000;$$

$$\sum_{i=1}^{n} p_i \cdot \overline{Y}_i = p \cdot \sum_{i=1}^{n} \overline{Y}_i = 4 \cdot (0,00485558 + 0,00595538 + 0,00641388 +$$
$$+ 0,00746542) = 0,0987611;$$

$$\sum_{i=1}^{n} p_i \cdot X_i \cdot \overline{Y}_i = p \cdot \sum_{i=1}^{n} X_i \cdot \overline{Y}_i = 4 \cdot (15 \cdot 0,00485558 + 30 \cdot 0,00595538 +$$
$$+ 45 \cdot 0,00641388 + 60 \cdot 0,00746542) = 3,95218;$$

Подставим полученные значения в формулы (12) и вычислим коэффициенты линейной модели:

$$A = \frac{600 \cdot 0,0987611 - 3,95218 \cdot 16}{600^2 - 27000 \cdot 16} = 0,0000552535;$$

$$B = \frac{3,95218 \cdot 600 - 0,0987611 \cdot 27000}{600^2 - 27000 \cdot 16} = 0,00410056.$$

Таким образом, линейная эмпирическая модель выглядит следующим образом:

$$Y = 0,0000552535 \cdot X + 0,00410056. \qquad (19)$$

В соответствии с табл. 1: $a = A$ и $b = B$, так что мы можем перейти к нелинейной модели в исходных переменных:

$$y = \frac{1}{0,0000552535 \cdot x + 0,00410056} = \frac{100000}{5,52535 \cdot x + 410,056}. \qquad (20)$$

Для удобства последующего использования мы преобразовали полученную модель к более «читаемому» виду.

Проверка модели на адекватность

Дисперсию ошибок вычислим по формуле (13) с использованием данных табл. 4:

$$S_{ош}^2 = \frac{1}{16-4}\cdot\left(\left(0,004975-0,00485558\right)^2+\ldots+\left(0,007299-0,00786542\right)\right)$$
$$=2,28096\times10^{-8}.$$

В соответствии с формулой (14) для получения значения дисперсии адекватности прежде всего нужно вычислить значения полученной функции в экспериментальных точках:

$$\hat{Y}_1 = 0,0000552535\cdot15+0,00410056=0,00492936;$$
$$\hat{Y}_2 = 0,0000552535\cdot30+0,00410056=0,00575816;$$
$$\hat{Y}_3 = 0,0000552535\cdot45+0,00410056=0,00658697;$$
$$\hat{Y}_4 = 0,0000552535\cdot60+0,00410056=0,00741577.$$

Вычислим теперь дисперсию адекватности с использованием данных табл. 4 и полученных выше значений:

$$S_{ад}^2 = \frac{4}{4-2}\cdot\left(\left(0,00485558-0,00492936\right)^2+\ldots+\left(0,00746542-0,00741577\right)^2\right)=$$
$$=1,53522\times10^{-7}.$$

Следовательно, расчетное значение критерия Фишера будет равно:

$$F_p = \frac{1,53522\times10^{-7}}{2,28096\times10^{-8}} = 6,73059\,.$$

По таблице значений критерия Фишера с учетом доверительной вероятности P = 0,95 и числа степеней свободы числителя $\nu_{ад}$ = 2 и знаменателя $\nu_{ош}$ = 12 (см. формулы (16) и (17)) найдем соответствующее значение: F_m = 3,89. Поскольку расчетное значение критерия Фишера получилось больше табличного, можно говорить о том, что полученная модель (20) не адекватна экспериментальным данным.

2.3 Результаты расчетов

Результаты расчетов представлены в табл. 5. На рис. 1 и 2 показаны графики линейной и нелинейной эмпирических моделей. Таблица и графики сформированы учебной программой для лабораторной работы № 1 автоматически.

Таблица 5 – Результаты расчетов

	Результаты Программы	Результаты Пользователя
1. Выбрана модель y = 1/(a*x + b)		
2. Разница среднегеометрического модели и эксперимента	0.3765	0.3765
2. Коэффициент A	0.00005525	0.00005525
Коэффициент B	0.004101	0.004101
3. Дисперсия ошибок Sош^2	0.00000002281	0.00000002281
Дисперсия адекватности Sад^2	0.0000001535	0.0000001535
4. Отношение дисперсий Sад^2/Sош^2	6.7306	6.7306
5. Модель не адекватно описывает данные эксперимента		

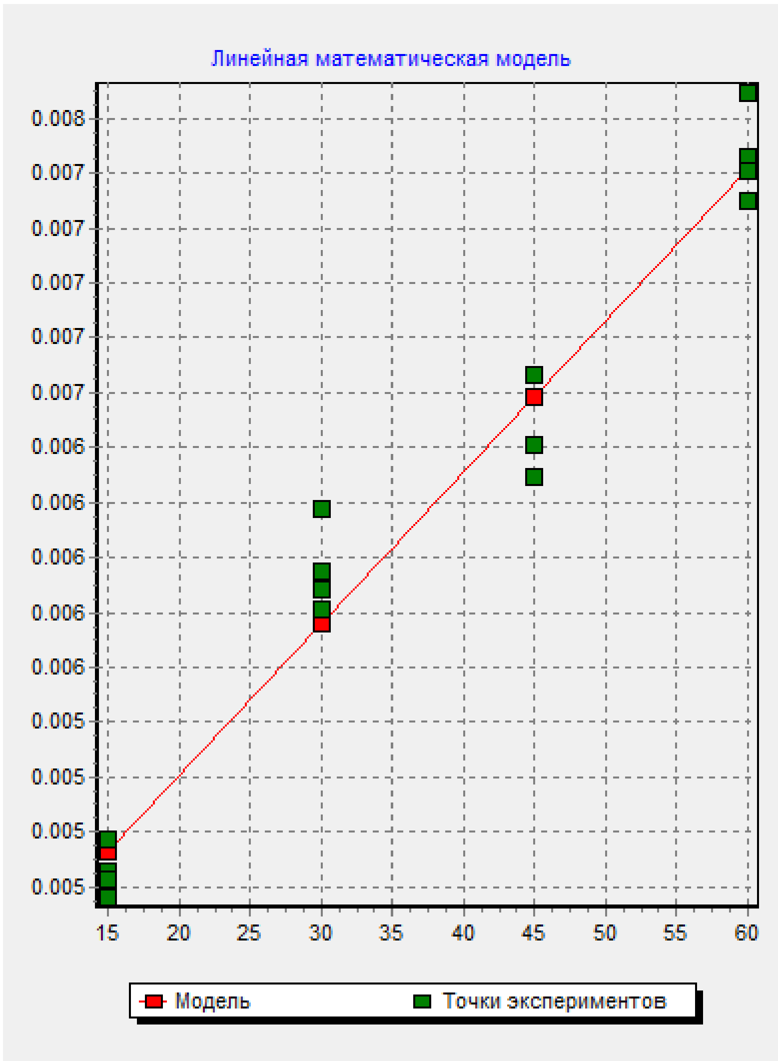

Рис. 1 – Линейная эмпирическая модель

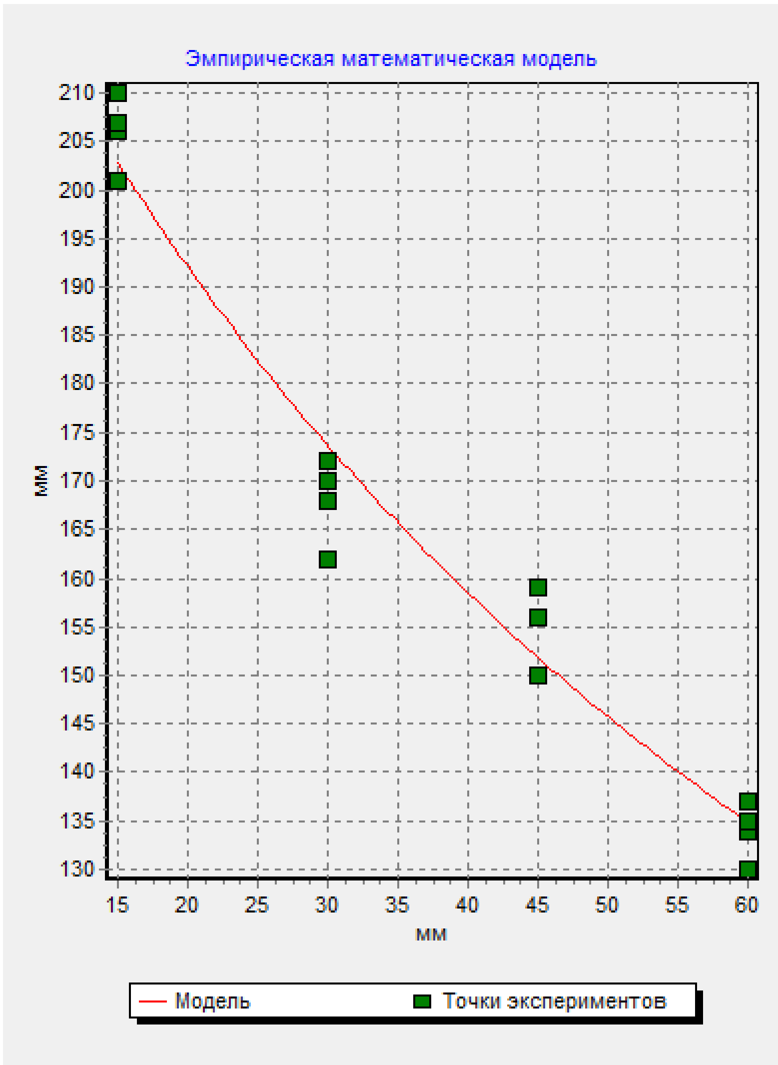

Рис. 2 – Нелинейная эмпирическая модель

3 Выводы

1. В результате обработки экспериментальных данных была получена следующая эмпирическая модель:

$$\sigma_в = \frac{100000}{5{,}52535 \cdot d_{отл} + 410{,}056},$$

где $\sigma_в$ – предел прочности немодифицированного сплава АЛ2, МПа;

$d_{отл}$ – диаметр отливки, мм.

2. Статистическая проверка показала, что полученная модель не адекватна экспериментальным данным, следовательно, ее нельзя использовать в практических целях. Для получения адекватной модели без проведения дополнительных экспериментов, необходимо исследовать другие зависимости, не входящие в использованный в данной работе набор из семи моделей.