ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ
ВОЛГОГРАДСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
КАФЕДРА «СОПРОТИВЛЕНИЕ МАТЕРИАЛОВ»

СТАТИСТИЧЕСКАЯ ОБРАБОТКА РЕЗУЛЬТАТОВ ИСПЫТАНИЙ НА ИЗНАШИВАНИЕ

Методические указания к лабораторной работе

Волгоград
2010

УДК 620.178.16:519.25 (075)

Р е ц е н з е н т

канд. техн. наук, доцент *А. А. Белов*

Издается по решению редакционно-издательского совета
Волгоградского государственного технического университета

Статистическая обработка результатов испытаний на изнашивание : метод. указания к лабораторной работе / сост. С. Н. Паршев, А. С. Столярчук. – Волгоград : Изд-во ИУНЛ ВолгГТУ, 2010. – 15 с.

Составлены для организации и проведения студентами учебно-исследовательской лабораторной работы по дисциплине «Прикладная физика (триботехника)». Приводятся методики расчетов, а также основные справочные сведения, что позволяет не прибегать к дополнительной литературе. Предполагается знание студентами основных положений раздела высшей математики: теории вероятностей и математической статистики.

Цель работы. Методами математической статистики установить зависимость относительной износостойкости среднеуглеродистой стали при граничном трении от микротвердости поверхностного слоя.

1. ОСНОВНЫЕ ТЕОРЕТИЧЕСКИЕ СВЕДЕНИЯ

1.1. Корреляционный анализ. Коэффициент корреляции при нормальном законе распределения случайных величин. Проверка нулевой гипотезы

При функциональной (**детерминированной**) зависимости между переменными величинами каждому значению аргумента x соответствует определенное значение функции y, то есть $y = f(x)$. Для экспериментальных величин такого однозначного соответствия между переменными нет. При этом необходимо выявлять зависимости особого вида (стохастические), при наличии которых одна величина реагирует на изменение другой изменением некоторых параметров своего закона распределения (например, закона Гаусса при нормальном распределении).

Изменение одной экспериментальной величины (например, y_i) в связи с изменением другой экспериментальной величины (x_i) может быть описано в виде двух составляющих: **стохастической** (обусловлена действием общих факторов, имеющих определенную физическую природу и формирующих в среднем зависимость рассматриваемых величин друг от друга) и **случайной** (зависит от индивидуальных случайных факторов). Очевидно, если случайная составляющая равна нулю, между опытными величинами существует детерминированная зависимость $y = f(x)$, если наоборот – стохастическая составляющая отсутствует, то рассматриваемые величины являются независимыми, то есть в абсолютном смысле случайными.

На практике обычно отличны от нуля обе составляющие. В этом случае тесноту взаимосвязей для генеральной совокупности оценивают **генеральным коэффициентом корреляции** ρ (для генеральной совокупности число опытов стремится к бесконечности, что возможно только теоретически). Генеральный коэффициент корреляции теоретически характеризует тесноту связи между экспериментальными величинами, рассматриваемыми как случайные. Коэффициент корреляции изменяется в пределах $-1 \le \rho \le 1$. Для независимых случайных величин, если он подчиняется закону Гаусса, $\rho = 0$. При положительном значении коэффициента корреляции) с возрастанием одной случайной величины в среднем возрастает и другая. При $\rho < 0$ с возрастанием одной величины другая убывает. Коэффициент корреляции может быть близок к нулю и для случая коррелированных величин, если зависимость между ними нелинейная, т.е. коэффициента корреляции в этом случае недостаточно для оценки их взаимосвязи.

Принято считать [1], что в случае линейной зависимости между исследуемыми экспериментальными случайными величинами их взаимосвязь определяется значениями ρ (таблица 1.1), то есть “случайными” они являются не в полном смысле этого слова (имеют и стохастическую и случайную компоненты). Тем не менее, в дальнейшем в тексте сохранен термин “случайные”, что продиктовано только сложившимися к настоящему времени традициями в научно-технической литературе.

Таблица 1.1

Уровни взаимосвязи между случайными величинами

Модуль коэффициента корреляции	Взаимосвязь между исследуемыми величинами
$0{,}00 \le /\rho/ < 0{,}20$	Связи практически нет
$0{,}20 \le /\rho/ < 0{,}50$	Существует слабая связь
$0{,}50 \le /\rho/ < 0{,}75$	Существует средняя связь
$0{,}75 \le /\rho/ < 0{,}95$	Существует сильная связь
$0{,}95 \le /\rho/ \le 1{,}00$	Связь практически функциональная

При анализе результатов механических испытаний в случае малой выборки (число испытаний, т. е. выборка $n \le 50$) при линейной зависимости между нормально распределенными (по закону Гаусса) случайными величинами в качестве количественной оценки тесноты связи между этими величинами используют [2] **выборочный коэффициент корреляции r_0**

$$r_0 = \frac{m}{S_x S_y} \tag{1.1}$$

Здесь выборочный смешанный центральный момент подсчитывается по формуле

$$m = \frac{1}{n}\sum_{1}^{n}(x_i - \bar{x})(y_i - \bar{y}), \tag{1.2}$$

где $\bar{x}$, $\bar{y}$ – выборочные средние случайных величин x_i, y_i:

$$\bar{x} = \frac{1}{n}\sum_{1}^{n} x_i,$$

$$\bar{y} = \frac{1}{n}\sum_{1}^{n} y_i, \tag{1.3}$$

Выборочные средние квадратические отклонения:

$$S_x = \sqrt{\frac{\sum_{1}^{n}(x_i - \bar{x})^2}{n-1}},$$

$$S_y = \sqrt{\frac{\sum_{1}^{n}(y_i - \bar{y})^2}{n-1}}. \tag{1.4}$$

При малом объеме выборки n значение выборочного коэффициента корреляции (1.1) целесообразно корректировать [3] по формуле

$$r = r_0\left[1 + \frac{1 - r_0^2}{2\,(n-3)}\right]. \tag{1.5}$$

Выборочный коэффициент корреляции, как и другие выборочные характеристики, является случайным параметром (в виду ограниченности n) и может принимать различные значения, но при увеличении числа опытов до бесконечности его средняя величина в пределе стремится к генеральному коэффициенту корреляции ρ.

При рассмотрении независимых величин, для которых генеральный коэффициент корреляции ρ равен нулю, выборочный коэффициент r может отличаться от нуля. В связи с этим возникает важная практическая задача:

***проверка нулевой гипотезы**, то есть гипотезы о равенстве нулю генерального коэффициента корреляции на основании эмпирических данных, что эквивалентно признанию изучаемых величин независимыми для генеральной совокупности (только по результатам исследований ограниченной выборки). Для решения поставленной задачи необходимо, прежде всего, установить закон распределения выборочного коэффициента корреляции, потому что при ограниченных объемах n распределение выборочного коэффициента в частных случаях может существенно отличаться от нормального закона.*

Для проверки гипотезы об отсутствии корреляции между исследуемыми величинами x_i, y_i (нулевой гипотезы) используют преобразование Фишера, который показал, что распределение случайной величины

$$u = \frac{1}{2}\ell n\frac{1+r}{1-r} \tag{1.6}$$

хорошо аппроксимируется нормальным законом с математическим ожиданием

$$a_u \cong \frac{1}{2}\ell n\frac{1+\rho}{1-\rho} \tag{1.7}$$

и дисперсией

$$\sigma_u{}^2 = \frac{1}{n-3}. \tag{1.8}$$

Проверка нулевой гипотезы, т. е. гипотезы: $\rho = 0$ (при альтернативной $\rho \neq 0$), заключается в вычислении по формулам (1.6), (1.8) значений u, σ_u и сопоставлении выборочного значения u с критическим, найденным по таблице 1.2 для вероятности $P = 1-\alpha/2$ (α – принимаемый исследователем, уровень значимости). Если выполняется условие

$$|u| \le z_p\sigma_u\,, \tag{1.9}$$

нулевую гипотезу ($\rho = 0$) принимают. В том случае, когда $|u| > z_p\sigma_u$, нулевую гипотезу отвергают. Здесь Z_p – квантиль нормированного нормального распределения; определяется по табл. 1.2.

Таблица 1.2

Значения квантили Z_p нормированного нормального распределения величины *u* в зависимости от вероятности *P* [2, с. 207]

P	Тысячные доли *P*				
	0	2	4	6	8
0,80	0,842	0,849	0,856	0,863	0,871
0,81	0,879	0,885	0,893	0,900	0,908
0,82	0,915	0,922	0,931	0,938	0,946
0,83	0,954	0,962	0,970	0,978	0,986
0,84	0,994	1,003	1,011	1,019	1,028
0,85	1,036	1,045	1,054	1,063	1,071
0,86	1,080	1,089	1,098	1,108	1,117
0,87	1,126	1,136	1,146	1,155	1,165
0,88	1,175	1,185	1,195	1,206	1,216
0,89	1,227	1,237	1,248	1,259	1,270
0,90	1,282	1,293	1,305	1,317	1,329
0,91	1,341	1,353	1,366	1,379	1,392
0,92	1,405	1,419	1,433	1,447	1,461
0,93	1,476	1,491	1,506	1,522	1,538
0,94	1,555	1,572	1,589	1,607	1,626
0,95	1,645	1,665	1,685	1,706	1,728
0,96	1,751	1,774	1,799	1,825	1,852
0,97	1,881	1,911	1,943	1,977	2,014
0,98	2,054	2,097	2,144	2,197	2,257
0,99	2,326	2,409	2,512	2,652	2,878

1.2. Линейный регрессионный анализ. Эмпирическая линия регрессии. Нахождение вероятного расположения (доверительного интервала) теоретической линии регрессии и оценка ошибок ее определения

Если исследуемые случайные величины подчиняются закону Гаусса и, кроме того, являются зависимыми (т. е. между ними имеются стохастические связи), то с изменением одной величины, в общем случае, могут меняться все статистические параметры другой случайной величины. В частном случае, когда рассматриваются только два параметра из этих статистик, можно записать

$$a_{y/x} = f_1(x), \tag{1.10}$$

$$\sigma^2_{y/x} = f_2(x) . \tag{1.11}$$

Первую зависимость называют уравнением **теоретической линии регрессии**, а вторую – скедастической зависимостью (здесь и далее для упрощения рассматривается только регрессия «*y*» по «*x*»).

Регрессионный анализ результатов испытаний включает оценку параметров уравнения **эмпирической линии регрессии** и ее графическое построение с учетом скедастической зависимости, а также проверку гипотезы о соответствии выбранной функции (1.10) данным опыта, то есть гипотезы адекватности выбранной математической модели (в нашем случае принимается простейшая математическая модель – линейная функция). Эмпирическая линия регрессии при этом, естественно, служит лишь некоторым приближением к теоретической линии регрессии (тем лучшим, чем больше объем эмпирической выборки). Для разных выборок, т. е. в различных сериях экспериментов, параметры выбранной функции, очевидно, будут отличаться. Таким образом, в опытах можно получить множество эмпирических линий (число которых равно числу серий экспериментов), образующих некоторую область “вокруг” теоретической линии регрессии, которую мы не можем точно установить вследствие ограниченного объема выборки.

Обычно регрессионному анализу предшествует корреляционный, на основании которого производят оценку средних значений изучаемых величин, а также их выборочных дисперсий и выборочного коэффициента корреляции ($\bar{x}$, $\bar{y}$, S_x^2, S_y^2, r). Тогда уравнение эмпирической линии регрессии, являющейся, как указано, лишь некоторым экспериментальным случайным приближением (зависимым от объема выборки и количества выборок, т.е. серий экспериментов) к теоретической линии регрессии (1.10), записывают [2] для принятой линейной модели в виде

$$Y = \bar{y} + r\,\frac{S_y}{S_x}\,(x_i - \bar{x}). \qquad (1.12)$$

При малом объеме выборки (число пар экспериментальных величин $n \le 50$) для упрощения анализа можно принять, что дисперсия случайной величины Y не зависит от x_i, т. е. скедастическая зависимость (1.11) имеет вид

$$\sigma^2_{y/x} = Const\,. \tag{1.13}$$

В этом случае параметры уравнения (1.12) могут быть определены по формулам (1.1)–(1.5). Кроме того, если имеется *n* пар экспериментальных величин $(x_1, y_1), (x_2, y_2) \ldots (x_n, y_n)$, то в качестве оценки дисперсии величины *Y* вместо $\sigma^2_{y/x}$ может быть использована выборочная дисперсия

$$S^2_{y/x} = \frac{1}{n-2}\sum_1^n (y_i - Y_i)^2 \cong S^2_y (1 - r_0^2)\frac{n-1}{n-2}\,. \tag{1.14}$$

Величина
$$\delta_y = \sqrt{S^2_{y/x}} \tag{1.15}$$

используется в качестве меры индивидуального рассеяния экспериментальных значений *Y* “вокруг” линии регрессии, то есть в качестве первой (но не единственной) меры **ошибки** определения эмпирической линии регрессии по уравнению (1.12).

Более точный подход заключается в оценке зоны вероятного расположения теоретической линии регрессии. Для этого, с принятой вероятностью, которая определяется выбранным или заданным уровнем значимости α, строится ***доверительный интервал****. Процедура такого построения заключается в следующем. Для ряда значений* x_i *по формуле (1.12), в случае линейной гипотезы, находят величину Y, а также ее дисперсию*

$$S^2_{Y/x} = S^2_{y/x}\left[\frac{1}{n} + \frac{(x_i - \bar{x})^2}{(n-1)\,S^2_x}\right]. \tag{1.16}$$

Далее составляют доверительный интервал для $a_{y/x}$:

$$Y - S_{Y/x}\cdot t_{\alpha,k} \le a_{y/x} \le Y + S_{Y/x}\cdot t_{\alpha,k}\,, \tag{1.17}$$

где $t_{\alpha,k}$ – значение квантили статистики t распределения Стьюдента для вероятности P. Значения $t_{\alpha,k}$ приведены в табл. 1.3, где число степеней свободы определяется по формуле: k = n – 2. Для дискретных значений x, задаваемых с выбранным шагом Δx, по формуле (1.17) строят нижнюю и верхнюю границы доверительного интервала. Очевидно, что такое приближенное построение будет тем точнее, чем меньше выбранный шаг.

Таблица 1.3

Значения α-пределов $t_{\alpha,k}$ распределения Стьюдента [2, с. 208]
в зависимости от k (при $k > 30$ $t_{\alpha,k} = Z_p$; (см. табл. 1.2)

k	α			
	0,100	0,050	0,025	0,010
1	6,314	12,706	25,452	63,657
2	2,920	4,303	6,205	9,925
3	2,353	3,182	4,177	5,841
4	2,132	2,776	3,495	4,604
5	2,015	2,571	3,163	4,032
6	1,943	2,447	2,969	3,707
7	1,895	2,365	2,841	3,499
8	1,860	2,306	2,752	3,355
9	1,833	2,262	2,685	3,250
10	1,812	2,228	2,634	3,169
12	1,782	2,179	2,550	3,055
14	1,761	2,145	2,510	2,977
16	1,746	2,120	2,473	2,921
18	1,734	2,101	2,445	2,878
20	1,725	2,086	2,423	2,845
22	1,717	2,074	2,405	2,819
24	1,711	2,064	2,391	2,797
26	1,706	2,056	2,379	2,779
28	1,701	2,048	2,369	2,763
30	1,697	2,042	2,360	2,750

2. ОБЩИЕ РЕКОМЕНДАЦИИ ПО ВЫПОЛНЕНИЮ И ОФОРМЛЕНИЮ РАБОТЫ

2.1. Проведение испытаний на изнашивание и использование корреляционного анализа

В сводной табл. 2.1 приведены результаты экспериментального определения микротвердости и относительной износостойкости стали 40 в различном структурном состоянии, определяемом режимами термообработки.

Таблица 2.1

Экспериментальные результаты исследования влияния микротвердости $H_{0,98}$ (x_i) на износостойкость (y_i) и их статистическая обработка

№ № п. п.	x_i	y_i	$\left(x_i-\overline{x}\right)$	$\left(y_i-\overline{y}\right)$	$\left(x_i-\overline{x}\right)\cdot\left(y_i-\overline{y}\right)$	$\left(x_i-\overline{x}\right)^2$	$\left(y-\overline{y}\right)^2$
01							
02							
03							
04							
05							
06							
07							
08							
09							
10							
11							
12							
13							
14							
15							
16							

Нанесите экспериментальные значения, полученные в опытах (табл. 2.1), на графики, принимая за аргумент x – микротвердость; за функцию y – относительную износостойкость. Визуально оцените, в каких случаях в принятых осях координат результаты можно аппроксимировать линейными зависимостями, в каких – нет. Приведите в отчете поля разброса экспериментальных точек на полученных таким образом графиках (поля регрессии).

С целью определения выборочных коэффициентов корреляции прочностных параметров с микротвердостью используйте формулы (1.1)–(1.5) из п. 1.1.

На основании проведенного корреляционного анализа сделайте выводы об уровнях взаимосвязи (см. табл. 1.1) между микротвердостью поверхностного слоя и относительной износостойкостью выбранной марки среднеуглеродистой стали.

Для проверки нулевой гипотезы (см. п. 1.1) используйте преобразования Фишера и его анализ согласно выражениям (1.6)–(1.9). Уровень значимости (α) *задается преподавателем. Обоснуйте правомерность своих выводов и дайте им вероятностную трактовку (на основании проверки нулевой гипотезы).*

2.2. Применение регрессионного анализа

Для построения эмпирической линии регрессии (после визуальной проверки линейности ожидаемой зависимости) воспользуйтесь уравнением (1.12). Значения параметров $\bar{x}$, $\bar{y}$, S_x, S_y, r вычислены ранее. На основании полученных результатов постройте эмпирическую линию регрессии в координатах x, y.

Дисперсию и индивидуальное рассеяние результатов наблюдений «вокруг» линии регрессии, т. е. меру основной ошибки определения эмпирической линии регрессии рассчитайте по формулам (1.14), (1.15).

На основании построенной эмпирической линии регрессии установите выборочные (случайные) границы изменения относительной износостойкости, соответствующие интервалу допустимых значений микротвердости.

Для расчета и построения на графиках доверительных интервалов возможного расположения теоретической линии регрессии используйте выражения (1.16), (1.17) (см. п. 1.2). Уровень значимости (α) задается преподавателем. Результаты расчетов удобно свести в таблицу (рекомендуемая форма приведена в виде табл. 2.2). При расчетах использовать табл. 1.3.

Таблица 2.2

Номер	X_i (микротвердость)	Y	$S_{Y/x}$	$S_{Y/x} \cdot t_{\alpha,k}$	$Y - S_{Y/x} \cdot t_{\alpha,k}$	$Y + S_{Y/x} \cdot t_{\alpha,k}$
1 2 3 4 5 … 16						

Границы доверительных областей нанести пунктирными линиями на графики.

3. ВОПРОСЫ ДЛЯ САМОКОНТРОЛЯ ПРИ ПОДГОТОВКЕ К ОТЧЕТУ

1. Чем обусловлено применение корреляционного и регрессионного методов анализа в практике обработки результатов испытаний на изнашивание.
2. Сформулируйте основные задачи данного исследования.
3. Что определяет генеральный коэффициент корреляции? В каких пределах он изменяется?
4. Приведите характерные уровни взаимосвязи между случайными величинами и генеральные коэффициенты корреляции, соответствующие этим уровням.
5. В каких случаях и для чего используют выборочный коэффициент корреляции?
6. В чем заключается проверка нулевой гипотезы?
7. Какие зависимости называют уравнением линии регрессии и скедастической зависимостью?
8. Какие задачи решает регрессионный анализ?
9. Какой параметр используют в качестве основной ошибки определения положения линии регрессии?
10. Как оценивается интервал вероятного нахождения теоретической линии регрессии?

СПИСОК РЕКОМЕНДОВАННЫХ ИСТОЧНИКОВ

1. *Шиндовский, Э. О*. Статистические методы управления качеством. Контрольные карты и планы контроля : пер. с нем. / Э. О. Шиндовский, О. Шюрц. – М.: Мир, 1976. – 597 с.

2. Степнов, М. Н. Статистические методы обработки результатов механических испытаний : справ / М. Н. Степнов. – М.: Машиностроение, 1985. – 232 с.

3. Зако, Л. Статистические оценивание : пер. с нем. / Л. Зако ; под ред. Ю. П. Адлера, Б. Г. Горского. – М.: Статистика, 1976. – 598 с.

У ч е б н о е   и з д а н и е

Сергей Николаевич **Паршев**
Александр Станиславович **Столярчук**

СТАТИСТИЧЕСКАЯ ОБРАБОТКА РЕЗУЛЬТАТОВ ИСПЫТАНИЙ НА ИЗНАШИВАНИЕ

Методические указания к лабораторной работе

Темплан 2010 г. (учебно-методическая литература). Поз. № 138.
Подписано в печать 26.05.2010. Формат 60х84 1/16. Бумага офсетная.
Гарнитура Times. Печать офсетная. Усл. печ. л. 0,93.
Тираж 10 экз. Заказ

Волгоградский государственный технический университет.
400131, г. Волгоград, пр. им. В. И. Ленина, 28, корп. 1.

Отпечатано в типографии ИУНЛ ВолгГТУ.
400131, г. Волгоград, пр. им. В. И. Ленина, 28, корп. 7.