

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ В ПОЛИКОМПОНЕНТНОЙ СИСТЕМЕ ЦИФРОВОГО ОБРАЗОВАНИЯ

Е. Н. ДИК, С. А. АРСЛАНБЕКОВА, Э. Ф. МУРЗИНА

Башкирский государственный аграрный университет

Уфа, Россия

В статье рассматривается построение математической модели эксплуатационного процесса работы двигателя, а именно: определяется связь смазочного материала и мощности двигателя на основе методов теории корреляции и регрессионного анализа. Известно, эти методы математически формализуют связи переменных количественных признаков, далее изучаются свойства модели и степень адекватности построенной корреляционной модели. Приводится лабораторный пример построения линейной модели о связи расхода моторного масла в зависимости от мощности двигателя автомобиля.

Техническое обеспечение современного производства в АПК подтверждает многопрофильную подготовку кадров инженерной специализации в высшем учебном заведении. Инженер-исследователь, он же генератор идей, осваивает функции проектирования, моделирования, конструирования. Инженер-производственник эксплуатирует, внедряет технические комплексы в производственную сферу. В целом, будущий инженер должен овладеть общепрофессиональными компетенциями – инженерным ядром.

В статье рассматривается роль инженера-исследователя, способного построить, адаптировать модель технического процесса и прогнозировать изменения модели под конкретные условия. В частности, создана математическая модель линейной корреляции для несгруппированных данных, проверена достоверность и надежность на соответствие реальным данным.

Следующая задача освещает тему исследования. Имеются данные, приведенные в табл. 1, нормы расхода моторных масел на угар и замену Y л/100 л. т. в зависимости от максимальной мощности двигателя автомобиля X л. с.

Выбор факторного признака определил X – максимальную мощность двигателя автомобиля, результативный признак принадлежит переменной Y – расходу моторных масел на угар и замену.

Табл. 1

X	39	42	53	70	75	90	110	115	150	170
Y	1,3	1,3	0,9	2,2	2,2	2,2	2,8	2,4	2,5	2,6

Предполагаем, что зависимость расхода моторного масла от мощности двигателя автомобиля носит линейный характер. Поэтому математическая модель, представляющая уравнение регрессии, построена методом наименьших квадратов.

Для расчета вспомогательных статистик $\bar{x}$, $\bar{y}$, S_x , S_y , r потребуются данные табл. 2.

Табл. 2

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	y_i	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	x^2	xy
39	-52,4	2745,76	1,3	-0,74	0,5476	1521	50,7
42	-49,4	2440,36	1,3	-0,74	0,5476	1764	54,6
53	-38,4	1474,56	0,9	-1,14	1,2996	2809	47,7
70	-21,4	457,96	2,2	0,16	0,0256	4900	154
75	-16,4	268,96	2,2	0,16	0,0256	5625	165
90	-1,4	1,96	2,2	0,16	0,0256	8100	198
110	18,6	345,96	2,8	0,76	0,5776	12100	308
115	23,6	556,96	2,4	0,36	0,1296	13225	276
150	58,6	3433,96	2,5	0,46	0,2116	22500	375
170	78,6	6177,96	2,6	0,56	0,3136	28900	442
914		17904,4	20,4		3,704	101444	2071

Рассчитаем выборочные средние квадратические отклонения, среднее произведений исследуемых признаков и коэффициент корреляции:

$$S_x = \sqrt{S_x^2} = \sqrt{\frac{1}{n-1} \cdot \sum (x_i - \bar{x})^2} = \sqrt{\frac{1}{10-1} \cdot 17904,4} = 44,6;$$

$$S_y = \sqrt{S_y^2} = \sqrt{\frac{1}{n-1} \cdot \sum (y_i - \bar{y})^2} = \sqrt{\frac{1}{9} \cdot 3,704} = 0,64;$$

$$\overline{xy} = \frac{1}{n} \cdot \sum_{i=1}^n x_i y_i = \frac{1}{10} \cdot 2071 = 207,1;$$

$$r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{S_x \cdot S_y} = \frac{207,1 - 91,4 \cdot 2,04}{44,6 \cdot 0,64} = 0,72.$$

Расчетный коэффициент по опытным данным позволяет оценить коэффициент корреляции

$$t_p = \frac{|r| \cdot \sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0,72 \cdot \sqrt{10-2}}{\sqrt{1-0,72^2}} = 2,93.$$

Значение теоретического коэффициента выбираем из таблицы критических точек распределения Стьюдента. Сравним t_p и t_T : $t_p = 2,93$; $t_T = 2,31$. Если $t_p > t_T$, т. е. $2,93 > 2,31$, то корреляция X и Y значима, существует линейная связь этих признаков.

Рассчитаем доверительный интервал r с надежностью 0,95 и объемом выборки $n = 10 < 50$:

$$r - t_\gamma \cdot \sigma_r \leq \hat{r} \leq r + t_\gamma \cdot \sigma_r.$$

При надежности $\gamma = 0,95$ коэффициент, входящий в предыдущее неравенство, равен $t_\gamma = 2,26$ (по таблице функции Лапласа).

Вычисляем погрешность коэффициента корреляции:

$$\sigma = \frac{1-r^2}{\sqrt{n-2}} = \frac{1-0,72^2}{\sqrt{10-2}} = 0,17;$$

$$0,72 - 2,26 \cdot 0,17 \leq \hat{r} \leq 0,72 + 2,26 \cdot 0,17.$$

Таким образом, размах значений коэффициента корреляции с вероятностью 0,95 заключен в интервале $r \in [0,34; 1]$. То есть в 34 % случаев и более прогнозируется влияние факторного признака на результативный. Следует изучить влияние и других факторов, возможно увеличивающих долю исследуемых связей.

Строим линейную модель – эмпирическую линию регрессии (рис. 1), используя данные табл. 2:

$$y_x = y(x) = \bar{y} + r \cdot \frac{s_y}{s_x} (x - \bar{x}); \quad y_x = 2,04 + 0,0103x - 0,9414;$$

$$y_x = 0,0103x + 1,0986.$$

Из уравнения $y_x = 0,0103x + 1,0986$ следует, что при увеличении максимальной мощности двигателя автомобиля на 100 л. с. расход моторных масел на угар и замену возрастет на 1,03 л/100 л. т.

Исследование нормы расхода смазочного материала построенной линейной зависимостью объяснимо в 52 % случаев и 48 % изменений обусловлено неучтенными в модели факторами. Например, нагрузками на транспорт или отказ технического узла конструкции автомобиля. Значит, адекватность модели по исследованию результативного признака требует дополнительного цикла изучения.

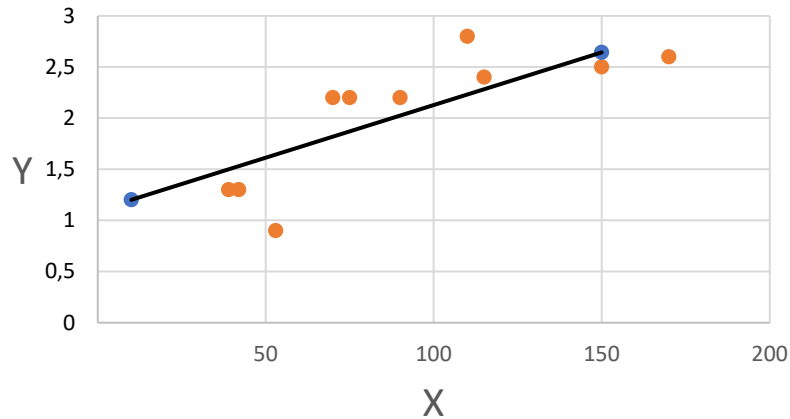


Рис. 1. Линейная регрессия Y на X (прямая с уравнением $y_x = 0,0103x + 1,0986$)

Проверка соответствия исследуемых величин их реально существующим значениям осуществляется по критерию Фишера – Снедекора. Для расчета статистики $F_H = \frac{R^2(n-2)}{1-R^2}$; $R^2 = 1 - \frac{\sum(y_i - \bar{y}_{xi})^2}{\sum y_i - \bar{y}}$ применим вспомогательные данные табл. 3.

Табл. 3

y_i	$\bar{y}_x$	$y_i - \bar{y}_{xi}$	$(y_i - \bar{y}_{xi})^2$
1,3	1,5003	-0,2003	0,0401
1,3	1,5312	-0,2312	0,0535
0,9	1,6445	-0,7445	0,5543
2,2	1,8196	0,3804	0,1447
2,2	1,8711	0,3289	0,1082
2,2	2,0256	0,1744	0,0304
2,8	2,2316	0,5684	0,3231
2,4	2,2831	0,1169	0,0137
2,5	2,6436	-0,1436	0,0206
2,6	2,8496	-0,2496	0,0623
			1,3508

Выбираем значения последних столбцов $\sum(y_i - \bar{y})^2 = 3,7040$, $\sum(y_i - \bar{y}_{xi})^2 = 1,3508$. Далее $R^2 = 1 - \frac{1,3508}{3,7040} = 0,635$; $F_H = \frac{0,635 \cdot (10-2)}{1-0,635} = 13,918$.

При уровне значимости $\alpha = 0,05$ и числа степеней свободы $k_1 = 1$, $k_2 = n - 2 = 10 - 2 = 8$ по таблице критических точек распределения Фишера – Снедекора находим $F_T = F_{\alpha; k_1; k_2} = F_{0,05; 1; 8} = 5,32$. Имеем $F_H = 13,918 > 5,32$, следовательно, эмпирическая зависимость, полученная методами математической статистики, позволяет планировать эксперимент.

Относительная погрешность эмпирического уравнения регрессии исследована по формуле

$$\delta = \frac{\sigma_u}{\bar{y}} \cdot 100 \%; \quad \sigma_u = \sqrt{D_u} = \sqrt{\frac{\sum(u_i - \bar{u})^2}{n-2}}; \quad u_i = y_i - \bar{y}_x; \quad \bar{u} = \frac{1}{n} \sum (y_i - \bar{y}_x)^2.$$

Рассчитаем $\sigma_u = \sqrt{\frac{1,5308}{8}} = 0,437$; $\delta = \frac{0,437}{2,04} \cdot 100 \% = 21 \%$. Относительная погрешность получила допустимое значение, точность оценки теоретически полученных и экспериментальных данных очевидна.

Исследование коэффициентов $a_0 = 1,0986$, $a_1 = 0,0103$ через расчет средних квадратических значений:

$$S_{a0} = S_{y/x} \cdot \sqrt{\frac{[x^2]}{n[x^2] - ([x])^2}}; \quad S_{a1} = S_{y/x} \sqrt{\frac{n}{n[x^2] - ([x])^2}}; \quad S_{y/x} = S_y \sqrt{1 - r^2}.$$

Выбираем из табл. 2 $[x] = 914$, $[x^2] = 101444$. Учитывая, что $n = 10$, $r^2 = 0,52$ и $S_y = 0,64$, находим:

$$S_{y/x} = 0,64 \cdot \sqrt{1 - 0,52} = 0,443;$$

$$S_{a0} = 0,443 \cdot \sqrt{\frac{101444}{10 \cdot 101444 - (914)^2}} = 0,14;$$

$$S_{a1} = 0,443 \cdot \sqrt{\frac{10}{10 \cdot 101444 - (914)^2}} = 0,0033.$$

$\frac{S_{a0}}{|a_0|} = \frac{0,14}{1,0986} = 0,13 < 0,5$ и $\frac{S_{a1}}{|a_1|} = \frac{0,0033}{0,0103} = 0,32 < 0,5$, значения коэффициентов удовлетворяют граничным значениям неравенств, а потому значимы для планирования эксперимента по найденной эмпирической зависимости (см. рис. 1).

Окончательно, уравнение регрессии $y_x = 0,0103x + 1,0986$, описывающее зависимость расхода моторных масел на угар и замену от максимальной мощности двигателя автомобиля, значимо описывает опытные данные для анализа и прогнозирования технических величин.

Таким образом, роль инженера-исследователя представлена, как часть многопрофильной подготовки инженерных кадров, на примере знания законов специальных математических дисциплин, лежащих в основе инженерной специализации.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. **Губин, В. И.** Статистические методы обработки экспериментальных данных: учеб. пособие / В. И. Губин, В. Н. Осташков. – Тюмень: ТюмГНГУ, 2007. – 202 с.