МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РФ

ФГОУ ВПО «БРЯНСКАЯ ГОСУДАРСТВЕННАЯ СЕЛЬСКОХОЗЯЙСТВЕННАЯ АКАДЕМИЯ»

ИНЖЕНЕРНО-ТЕХНОЛОГИЧЕСКИЙ ФАКУЛЬТЕТ

Кафедра эксплуатации машинно-тракторного парка

Оценка дисперсии воспроизводимости опыта, адекватности модели и значимости коэффициентов

Методические указания для выполнения практической работы по дисциплине «Основы научных исследований и патентоведение»

студентам инженерно-технологического факультета обучающихся по профилям:

110800-01 «Технические системы в агробизнесе»

110800-03 «Машины и оборудование для хранения и переработки сельскохозяйственной продукции»

190100-01 «Машины и оборудование природообустройства и дорожного строительства»

Брянск 2013

УДК 631.3.072:629.3(076)
ББК 39.3:40.72
Д 93

Дьяченко, А.В. Оценка дисперсии воспроизводимости опыта, адекватности модели и значимости коэффициентов : методические указания / А.В. Дьяченко. - Брянск: Издательство Брянской ГСХА, 2013. - 12 с.

Методические указания предназначены для выполнения практической работы по дисциплине «Основы научных исследований и патентоведение» для студентов инженерно-технологического факультета обучающихся по профилям 110800-01 «Технические системы в агробизнесе», 110800-03 «Машины и оборудование для хранения и переработки сельскохозяйственной продукции», 190100-01 «Машины и оборудование природообустройства и дорожного строительства. Целью выполнения практической работы является получение навыков анализа полученных экспериментальных данных и аппроксимирующей зависимости.

Рецензент: к.э.н., доцент каф. ТОЖиПП Исаев Х.М.

Рекомендовано к изданию методической комиссией инженерно-технологического факультета Брянской государственной сельскохозяйственной академии, протокол № 1от 27 сентября2013 г.

Для полученных экспериментальных данных и принятой модели применяют следующие статистические оценки: оценка дисперсии воспроизводимости опыта, оценка адекватности модели и оценка значимости ее коэффициентов.

Оценка дисперсии воспроизводимости (погрешности опыта) определяется на основании данных параллельных опытов и характеризует равноточность измерений во всех опытах.

Для проверки нулевой гипотезы, состоящей в том, что дисперсии во всех опытах равны между собой, т.е. проверка, значимо или незначимо отличаются оценки дисперсии в каждом опыте, более предпочтительно использование *критерия Кохрена,* который представляй собой отношение максимальной дисперсии $D_{y\,\max}$ к сумме всех дисперсий в N опытных точках:

$$G = D_{y\max} \Big/ \sum_{i=1}^{N} D_{y_i} \;, \tag{1}$$

где $D_{y_i} = \dfrac{1}{m-1} \sum_{j=1}^{m} \left(y_{ij} - \bar{y}_i \right)^2 ;$

$$\bar{y}_i = \frac{1}{m} \sum_{j=1}^{m} y_{ij} \;;$$

m - число измерений в i-м опыте, N - число опытов.

Вычисленное по формуле (1) значение критерия Кохрена G при принятом уровне значимости α (чаще всего α = 0,05) сравнивается с табличными G_T (табл. 1), которое является функцией числа степеней свободы m - 1 и N.

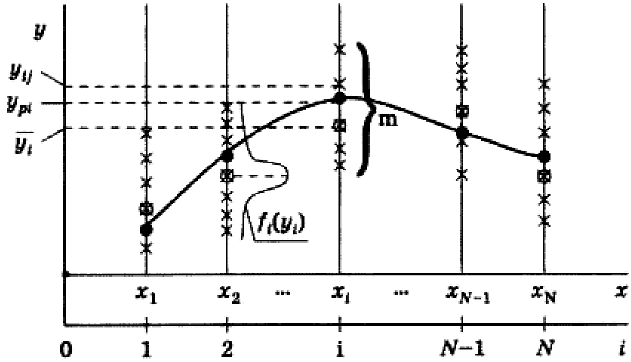

Рисунок 1 - Графическая интерпретация результатов эксперимента.

Если $G < G_T$, то гипотеза о равноточности не отвергается. Тогда погрешность опыта, оцениваемая средней квадратической погрешностью при определения среднего значения $\bar{y}_i$.

$$\sigma^2_{\bar{y}_0} = D_{\bar{y}_0} = \frac{1}{mN}\sum_{i=1}^{N} D_{y_i} , \qquad (2)$$

где $mN = n$ – общее количество измерений.

Таблица 1 – Значения критерия Кохрена

N	$m-1$								
	1	2	3	4	5	6	8	16	∞
2	0,999	0,975	0,939	0,906	0,853	0,816	0,734	0,660	0,500
3	0,967	0,871	0,798	0,746	0,677	0,633	0,547	0,475	0,333
4	0,907	0,768	0,684	0,629	0,560	0,518	0,437	0,372	0,250
6	0,781	0,666	0,532	0,480	0,418	0,382	0,314	0,261	0,167
8	0,680	0,516	0,438	0,391	0,336	0,304	0,246	0,202	0,125
10	0,602	0,445	0,373	0,331	0,282	0,254	0,203	0,166	0,100
15	0,471	0,334	0,276	0,242	0,203	0,182	0,143	0,114	0,067
20	0,389	0,271	0,221	0,192	0,160	0,142	0,111	0,088	0,050
30	0,293	0,198	0,159	0,138	0,114	0,100	0,077	0,060	0,033
120	0,100	0,063	0,049	0,042	0,034	0,029	0,022	0,017	0,008

Оценка адекватности аппроксимирующей зависимости исследуемому объекту обычно производится с помощью *критерия Фишера*, который в данном случае определяется как отношение дисперсии адекватности D_{y_a} к дисперсии опыта $D_{\bar{y}_0}$, определенной по формуле (2). Дисперсия адекватности, характеризующая рассеивание данных эксперимента $\bar{y}_i$ вокруг аппроксимирующей зависимости (кривая на рис. 1), определяется по формуле

$$D_{y_a} = \frac{1}{N-k} \sum_{i=1}^{N} \left(y_{pi} - \bar{y}_i \right)^2 , \tag{3}$$

где k - число коэффициентов в принятой математической модели; $k = s + 1$; s - порядок принятого полинома;

y_{pi} - расчетное значение функции в i-й точке при аппроксимации ее зависимостью вида $Y = f(X)$.

Тогда *F*-критерий запишется в виде

$$F = D_{y_a} / D_{\bar{y}_0} \tag{4}$$

Полученное в соответствии с (4) значение F сравнивается с табличным $F_T = f\,[N - k;\, N(m-1)]$ (см. табл. 2). Если $F < F_T$ то гипотеза об адекватности зависимости $Y = f(X)$ исследуемому объекту не отвергается. Подчеркнем, что вычисление D_{y_a}, а следовательно, и проверка адекватности с помощью *F*-критерия возможна только при $N > k$. Заметим, что если погрешность опыта, оцениваемая дисперсией, известна априори, то при $D_{y_a} \le D_{\bar{y}_0}$ математическая модель адекватна объекту.

Таблица 2 – Значения критерия Фишера

N(m-1)	*N-k*									
	1	2	3	4	5	6	8	12	24	∞
3	10,13	9,55	9,28	9,12	9,01	8,94	8,84	8,74	8,64	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,57	3,41	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,90	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54
12	4,75	3,88	3,49	3,26	3,11	3,00	2,85	2,69	2,50	2,30
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,35	2,13
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08	1,84
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62
∞	3,84	2,99	2,60	2,37	2,21	2,09	1,94	1,75	1,52	1,00

* Здесь *k* — количество искомых параметров (коэффициентов).

Пример

Предположим, задачи, выполняемые оператором, разбиты по сложности на три категории: $x_i; i=1\ldots N; N=3$. Среднее время, затрачиваемое на решение задачи y_i:

x_i	1	2	3
y_i	2	3	5

Получена аппроксимирующая зависимость $y_p = 0{,}33 + 1{,}5x$. Дисперсия $D_{\bar{y}_0} = 0{,}05$. Произвести проверку адекватности.

Решение. В нашем случае $\boldsymbol{N}$ = 3; k = 2; $\bar{y}_i = y_i$.

$$D_{y_a} = \frac{1}{N-1}\sum_{i=1}^{N}\left(y_{pi} - y_i\right)^2 = \sum_{i=1}^{3}\left(0{,}33 + 1{,}5x_i - y_i\right)^2 = 0{,}167 \cdot$$

По формуле (4) $F = 0{,}167/0{,}05 = 3{,}34$. По табл. 2 при $N - k$ =1 и $N(m-1) \to \infty$ (так как $D_{\bar{y}_0}$ задано, то подразумевается, что число повторений опыта достаточно велико) $\boldsymbol{F_T}$ = 3,84 и, следовательно, $\boldsymbol{F}$ = 3,34 < $\boldsymbol{F_T}$ = 3,84, и гипотеза об адекватности полученной линейной зависимости исследуемому объекту не отвергается.

Оценка значимости коэффициентов аппроксимирующей зависимости, взятой в виде алгебраического полинома (уравнения регрессии), в смысле отличия значений этих коэффициентов от нуля обычно проводится отдельно для каждого коэффициента a_l, где $l = 0 \ldots s$, с помощью *критерия Стьюдента*

$$t_l = \frac{|a_l|}{\sigma_{a_l}}, \tag{5}$$

где $\sigma_{a_l} = \sqrt{D_{a_l}}$; D_{a_l} - дисперсия коэффициента регрессии a_l.

Величина D_{a_l} определяется следующим образом. Решается система нормальных уравнений относительно коэффициентов a_l.

$$
\begin{aligned}
&a_0 \sum_i x_i^0 + a_1 \sum_i x_i^1 + a_2 \sum_i x_i^2 + \ldots + a_s \sum_i x_i^s = \sum_i x_i^0 y_i;\\
&a_0 \sum_i x_i^1 + a_1 \sum_i x_i^2 + a_2 \sum_i x_i^3 + \ldots + a_s \sum_i x_i^{s+1} = \sum_i x_i^1 y_i;\\
&\quad \ldots \quad \ldots \quad \ldots \quad \ldots\\
&a_0 \sum_i x_i^s + a_1 \sum_i x_i^{s+1} + a_2 \sum_i x_i^{s+2} + \ldots + a_s \sum_i x_i^{2s} = \sum_i x_i^s y_i.
\end{aligned}
$$

При этом правые части уравнений $\nu_l = \sum_{i=1}^{N} x_i^l y_i$ не заменяются их численными значениями. В результате решения системы для коэффициентов a_l находят линейные зависимости от величин ν_l. Если в эти зависимости подставить численные значения ν_l то получим численные значения коэффициентов a_l. Если же в них подставить вместо ν_l единицу, а вместо остальных ν нули, то можно получить для каждого a_l значение M_l с помощью которого и находят

$$D_{a_l} = M_l D_{\bar{y}_0} \tag{6}$$

где $D_{\bar{y}_0}$ - вычислено по выражению (2).

В частности, при линейной зависимости $Y = a_0 + a_1 x$ и при $\sum_{i=1}^{N} x_i = 0$ зависимость $a_l(\nu_l)$ имеет вид

$$a_0 = \nu_0 N^{-1};\; a_1 = \nu_1 \left(\sum_{i=1}^{N} x_i^2 \right)^{-1}.$$

Тогда при $\nu_l = 1$ имеем $M_0 = N^{-1};\; M_1 = \left(\sum_{i=1}^{N} x_i^2 \right)^{-1}.$

Воспользовавшись (6), запишем:

$$D_{a_0} = D_{\bar{y}_0} N^{-1};\ D_{a_1} = \left(\sum_{i=1}^{N} x_1^2 \right)^{-1} D_{\bar{y}_0}.$$

Таблица 3 – Значения коэффициента Стьюдента

v	t	v	t	v	t
1	12,71	6	2,45	20	2,09
2	4,30	8	2,31	30	2,04
3	3,18	10	2,23	120	1,98
4	2,78	15	2,13	00	1,96

Значение t_l, определенное по формуле (5), сравнивают с табличным t_T, найденным для числа степеней свободы $v = N(m - 1)$ при принятом уровне значимости, обычно 0,05 (табл. 3). Если $t_l < t_T$ коэффициент a_l считается незначимым (т.е. можно принять a_l = 0) и соответствующее слагаемое исключается из уравнения регрессии. Заметим, что при $m = 1$ имеем $v = 0$ и рассмотренный метод оценки неприменим. В этом случае оценка значимости коэффициента может быть произведена путем сравнения дисперсий адекватности D_{y_a} при наличии члена аппроксимирующего полинома с коэффициентом a_l и при его отсутствии ($a_l = 0$). Если дисперсия адекватности для второго варианта близка к дисперсии для первого, то рассматриваемый коэффициент можно считать незначимым.

Литература

1. Методы исследований и организация экспериментов / под ред. проф. К. П. Власова — Х.: Издательство «Гуманитарный Центр», 2002.— 256 с.

Для заметок

Учебное издание

Дьяченко Антон Вячеславович

Оценка дисперсии воспроизводимости опыта, адекватности модели и значимости коэффициентов

Подписано к печати .. 2013 г. Формат 60 х 84. 1/16.
Бумага печатная. Усл. п .л. 0,75. Тираж 50 экз. Изд. № .

Издательство Брянской государственной сельскохозяйственной академии
243365 Брянская область, Выгоничский район, с. Кокино, Брянская ГСХА