

ДИСПЕРСИОННЫЙ АНАЛИЗ КАК СРЕДСТВО ОЦЕНКИ РЕЗУЛЬТАТОВ

Н. ЛАЗАРЕВА, канд. с.-х. наук, член ВНАП, ООО «Коудайс МКорма»

В предыдущей статье «Оценка результатов испытаний в условиях птицефабрики», опубликованной в №5—2016, обсуждался вопрос оценки достоверности влияния различных факторов на рост и развитие цыплят. При этом речь шла о сравнении двух выборок или двух групп, например, птичников — контрольных и опытных.

Но на практике мы зачастую имеем дело с несколькими выборками одновременно, когда исследуемый фактор имеет разные градации как количественные, так и качественные. Например, это могут быть разные дозы лекарства (количественные градации) или различные лекарства (качественная градация). Выборки можно объединить в единый статистический комплекс при следующих условиях: во-первых, факторы воздействия должны быть независимыми друг от друга, во-вторых, выборки должны производиться способом случайного отбора из нормально распределяющейся совокупности. В дисперсионном анализе принято называть признаки, которые меняются под действием факторов, результативными. В птицеводстве это, как правило, показатели выращивания (зоотехнические и ветеринарные): сохранность, выбраковка, прирост живой массы, конверсия корма и др. При этом результативные признаки могут быть выражены в различных условных единицах — в процентах, индексах и т.д.

С математической точки зрения суть дисперсионного анализа заключается в разложении общей дисперсии всего статистического комплекса на составляющие и сравнении их друг с другом. Если на результативный признак действует один фактор в нескольких градациях (однофакторный дисперсионный анализ), составляющих общей дисперсии будет две: внутригрупповая (остаточная) и межгрупповая (факториальная). Для их расчета применяют так называемые девиаты (D), зависимость между ними выражается следующим равенством:

$$D_y = D_x + D_e,$$

где D_x — межгрупповая девиата, или сумма квадратов отклонений групповых средних от общей средней комплекса;
 D_e — внутригрупповая девиата, или сумма из сумм квадратов отклонений отдельных вариантов от их групповых средних;

D_y — общая девиата, или сумма квадратов отклонений вариант от общей средней.

Отношение девиаты к соответствующему числу степеней свободы и будет составлять выборочную дисперсию — факториальную, остаточную или общую.

Девиаты рассчитываются по формулам:

$$D_y = \sum x^2 - H; D_x = \frac{\sum(\sum x^2)}{n} - H; D_e = D_y - D_x,$$

$$\text{где } H = \frac{[\sum(\sum x)]^2}{N}.$$

Рассмотрим пример. На птицефабрике N испытывали разные дозировки антибиотика, который должен способствовать лучшей сохранности цыплят. Всего было 4 дозировки препарата, поэтому можно считать, что на птицу влиял регулируемый фактор в четырех градациях. В каждом варианте опыта (градации) было 5 птичников, то есть общее количество наблюдений равнялось 20 (объем дисперсионного комплекса). В таблице 1 приведены результаты по сохранности.

Таблица 1. Результаты применения кормового антибиотика: сохранность бройлеров за период выращивания

Дозы препарата (фактор А), г/т	Сохранность по птичкам, %					Средняя сохранность, %
	1	2	3	4	5	
50	93	94	92	94	96	93,8
65	95	96	95	97	96	95,8
80	95	95,5	96,5	96	94,5	95,5
100	95,5	97	97,5	94,5	95,5	96,0

Обозначения в таблице 2:

a — количество градаций фактора, действие которого оценивается;

n — количество повторов (в нашем примере — количество птичников);

$\sum x$ — сумма значений (вариант) по каждой градации (варианту опыта);

$(\sum x)^2$ — квадрат этой суммы;

$\sum x^2$ — сумма квадратов каждого значения, входящего в данную градацию (вариант опыта);

N — общее число наблюдений или объем дисперсионного комплекса.

Для удобства расчетов от каждого значения сохранности отнимаем число 90 (на конечный результат это не повлияет), и эти данные располагаем по вертикали. Рассчитываем

величины $\sum x$, $(\sum x)^2$ и $\sum x^2$ по каждой градации, а также суммы этих величин по всем четырем градациям.

Таблица 2. Расчет дисперсионного комплекса

Параметры	Варианты опыта (градации фактора воздействия)				Суммы
	50 г/т	65 г/т	80 г/т	100 г/т	
Сохранность по птичникам, % ($X - 90$)	3	5	5	5,5	$a = 4$
	4	6	5,5	7	
	2	5	6,5	7,5	
	4	7	6	4,5	
	6	6	4,5	5,5	
N	5	5	5	5	$N = 20$
$\sum x$	19	29	27,5	30	105,5
$(\sum x)^2$	361	841	756,25	900	2858,25
$\sum x^2$	81	171	153,75	186	591,75

Далее рассчитываем девиаты:

$$D_y = 591,75 - 105,52/20 = 591,75 - 556,51 = 35,24;$$

$$D_A = 2858,25/5 - 556,5125 = 571,65 - 556,51 = 15,14;$$

$$D_e = 35,24 - 15,14 = 20,1$$

и дисперсии:

$$S_A^2 = \frac{D_A}{a-1} = \frac{15,14}{3} = 5,0458;$$

$$S_e^2 = \frac{D_e}{N-n} = \frac{20,1}{16} = 1,25625.$$

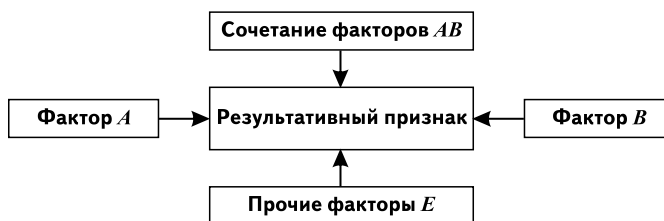
Результаты сводим в заключительную таблицу 3 (см. Приложение 1) и определяем соотношения дисперсий:

$$F_A = S_A^2/S_e^2 = 5,0458/1,25625 = 4,02.$$

Для того чтобы сделать вывод о том, повлияла или нет доза антибиотика на сохранность цыплят, сравниваем полученную величину соотношения факториальной и остаточной дисперсий 4,02 с рассчитанными для двух уровней значимости (берем из приложения 3) с учетом объема выборки. Как видим, полученное в ходе расчетов число больше значения для 95%-го уровня доверительной вероятности, но меньше для 99%-го. Поэтому можно сделать вывод, что дозировка антибиотика повлияла на увеличение сохранности цыплят с вероятностью 95%.

Очень часто на практике мы сталкиваемся с действием на организм не одного, а сразу нескольких факторов. В таком случае стоит вопрос не только о достоверности влияния

каждого из факторов, но и о доле влияния каждого фактора на результирующий признак. Алгоритм дисперсного двухфакторного анализа в принципе не отличается от однофакторного, но более громоздкий, поскольку приходится оценивать не один, а два фактора, а также их совместное действие. Это можно изобразить в виде схемы:



Общая сумма квадратов отклонений D_y содержит четыре компонента варьирования: $D_y = D_a + D_b + D_{ab} + D_e$, а общая факториальная сумма квадратов отклонений D_e состоит из трех членов: $D_x = D_a + D_b + D_{ab}$.

Дисперсионный анализ двухфакторных равномерных комплексов отличается от анализа однофакторных тем, что к расчету девиат D_y , D_x и D_e добавляется расчет факториальных девиат D_a , D_b и девиаты совместного действия факторов D_{ab} , которая вычисляется как: $D_{ab} = D_x - D_a - D_b$. Также необходимо определить степени свободы для факториальных дисперсий A , B и дисперсии совместного действия A и B .

Пример 2. Изучали действие двух иммуномодуляторов на показатели естественной резистентности цыплят-бройлеров. Птица получала препарат двумя путями: через корм и через воду. Кровь отбирали при убое шести голов для каждого варианта опыта и определяли, в частности, уровень лизоцима в сыворотке крови. Результаты приведены в таблице 4.

Для удобства расчетов отнимаем от каждого значения уровня лизоцима число 42, умножаем на 10 (чтобы избавиться от дробей) и заносим данные в таблицу 5 — по каждой градации каждого фактора отдельно. В данном расчете, как видим, требуется определить $\sum x$, $(\sum x)^2$, $\sum x^2$ и их суммы не только по всему дисперсионному комплексу, но и по каждому фактору отдельно: $\sum xA$, $\sum xB$ и т.д. Также в данную таблицу можно занести значения $H = (\sum x)^2/n$.

После заполнения такой таблицы можно приступить к расчету девиат, не забывая о том, что мы умножали все цифры на 10, чтобы избавиться от дробей, поэтому каждую девиату надо будет разделить на 100:

Таблица 3. Результаты анализа дисперсий

Варьирование	k	D	S^2	$F_{\text{факт}}$	Уровень значимости	
					0,05	0,01
По фактору A	3	15,14	5,04583	4,02	2,85	4,44
Остаточное	16	20,1	1,25625	—	—	—
Общее	19	35,24	1,8547	—	—	—

Таблица 4. Уровень лизоцима в сыворотке крови цыплят, %

Способ дачи препарата — фактор B	Иммуномодуляторы — фактор A						Среднее по B
	A_1			A_2			
	Результаты определения уровня лизоцима						
Через воду B_1	48,5	48,3	47,5	46,6	46,5	45,8	46,87
	47,1	47,4	46,4	46,5	45,7	46,1	
Через корм B_2	47,4	46,5	45,7	46,1	45,4	46,5	46,18
	46,7	47,5	46,2	45,2	46,0	44,9	
Среднее по A	47,1			45,94			

Таблица 5. Расчет двухфакторного дисперсионного комплекса

Параметры	A_1		A_2		Сумма
	B_1	B_2	B_1	B_2	
X (преобразованные показатели лизоцима)	65 63 55 51 54 44	54 45 37 47 55 42	46 45 38 45 37 41	41 34 45 32 40 29	$a = 2$ $b = 2$
n	6	6	6	6	$N = 24$
$\sum x$	332	280	252	221	1085
$(\sum x)^2$	110 224	78 400	63 504	48 841	—
$(\sum x)^2/n$	18 672	13 308	10 660	8 327	50 967
$\sum x^2$	18370,67	13 066,67	10 584	8 140,17	50 161,51
nA	12		12		24
$\sum xA$	332 + 280 = 612		252 + 221 = 473		—
$(\sum xA)^2/nA$	31 212		18 644,0833		49 856,0833
nB	12		12		24
$\sum xB$	332 + 252 = 584		280 + 221 = 501		—
$(\sum xB)^2/nB$	28 421,33		20 916,75		49 338,08

Таблица 6. Результаты двухфакторного дисперсионного анализа

Варьирование	k	D	S^2	$F_{\text{факт}}$	Уровень значимости	
					0,05	0,01
По фактору A	1	8,05	8,05	19,98	4,35	8,10
По фактору B	1	2,87	2,87	7,122	—	—
Совместно AB	1	0,18	0,18	0,447	—	—
Остаточное	20	8,06	0,403	—	—	—
Общее	23	19,16	0,833	—	—	—

$$D_y (\text{общая}) = (50\,967 - 1085^2/24)/100 = (50\,967 - 49\,051,042)/100 = 1915,958/100 = 19,16;$$

$$D_x (\text{межгрупповая}) = (50\,161,51 - 1085^2/24)/100 = (50\,161,51 - 49\,051,042)/100 = 1109,468/100 = 11,10;$$

$$D_e (\text{остаточная}) = 19,16 - 11,10 = 8,06;$$

$$D_A (\text{по фактору } A) = (49\,856,0833 - 49\,051,042)/100 = 805,0413/100 = 8,05;$$

$$D_B (\text{по фактору } B) = (49\,338,08 - 49\,051,042)/100 = 287,038/100 = 2,87;$$

$$D_{AB} (\text{совместная по } A, B) = D_x - D_A - D_B = 11,10 - 8,05 - 2,87 = 0,18.$$

Далее рассчитываем дисперсии варьирования (с учетом степеней свобод) фактора A , фактора B , совместного AB и остаточного:

$$S_A^2 = 8,05/1 = 8,05;$$

$$S_B^2 = 2,87/1 = 2,87;$$

$$S_{AB}^2 = 0,18/1 = 0,18;$$

$$S_e^2 = 8,06/20 = 0,403;$$

$$S_y^2 = 19,16/23 = 0,833.$$

Результаты сводим в заключительную таблицу 6 (см. Приложение 2) и определяем соотношение дисперсий:

$$F_A = S_A^2/S_e^2 = 8,05/0,403 = 19,98;$$

$$F_B = S_B^2/S_e^2 = 2,87/0,403 = 7,122;$$

$$F_{AB} = S_{AB}^2/S_e^2 = 0,18/0,403 = 0,447.$$

Критерий Фишера (F) для фактора A больше критического значения 8,10 при значимости 0,01, а для фактора B — больше критического значения 4,35 при уровне значимости 0,05. Из этого следует, что и вид иммуномодулятора, и способ его дачи достоверно усиливают резистентность цыплят, что выражается в увеличении уровня лизоцима в

Приложение 1. Сводная таблица однофакторного дисперсионного анализа

Варьирование	Числа степеней свободы	Суммы квадратов отклонений, или девиаты D	Средние квадраты отклонений, или дисперсии S^2	Дисперсионное отношение $F_{факт}$ (критерий Фишера)
По фактору A	$k_A = a - 1$	D_A	$S_A^2 = D_A/k_A$	$F_{факт.} = S_A^2/S_e^2$
Остаточное	$K_e = N - a$	D_e	$S_e^2 = D_e/k_e$	—
Общее	$k_y = N - 1$	D_y	$S_y^2 = D_y/k_y$	—

Приложение 2. Сводная таблица двухфакторного дисперсионного анализа

Варьирование	Числа степеней свободы	Девиаты	Дисперсии	$F_{факт}$
По фактору A	$k_A = a - 1$	D_A	$S_A^2 = D_A/k_A$	S_A^2/S_e^2
По фактору B	$k_B = b - 1$	D_B	$S_B^2 = D_B/k_B$	S_B^2/S_e^2
Совместно AB	$k_{AB} = (a - 1)(b - 1)$	D_{AB}	$S_{AB}^2 = D_{AB}/k_{AB}$	S_{AB}^2/S_e^2
Остаточное	$K_e = N - ab$	D_e	$S_e^2 = D_e/k_e$	—
Общее	$K_y = N - 1$	D_y	—	—

Приложение 3. Значения F-критерия Фишера при разных уровнях значимости — 0,05 (верхняя строка) и 0,01 (нижняя строка)

k	k_1 — степени свободы для большей дисперсии												
	1	2	3	4	5	6	7	8	9	10	20	30	∞
1	161	200	216	225	230	234	237	239	241	242	248	250	254
	4052	4999	5403	5625	5764	5859	5928	5982	6022	6056	6209	6261	6366
2	18,51	19	19,16	19,25	19,30	19,33	19,36	19,37	19,38	19,39	19,44	19,46	19,50
	98,49	99,01	99,17	99,25	99,30	99,33	99,34	99,36	99,38	99,40	99,45	99,47	99,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,88	8,84	8,81	8,78	8,66	8,62	8,53
	34,12	30,81	29,46	28,71	28,24	27,91	27,67	27,49	27,34	27,23	26,69	26,50	26,12
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,80	5,75	5,63
	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,54	14,02	13,83	13,46
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,78	4,74	4,56	4,50	4,36
	16,26	13,27	12,06	11,39	10,97	10,67	10,45	10,27	10,15	10,05	9,55	9,38	9,02
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	3,87	3,81	3,67
	13,74	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,39	7,23	6,88
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,63	3,44	3,38	3,23
	12,25	9,55	8,45	7,85	7,46	7,19	7,00	6,84	6,71	6,62	6,15	5,98	5,65
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,34	3,15	3,08	2,93
	11,26	8,65	7,59	7,01	6,63	6,37	6,19	6,03	5,91	5,82	5,36	5,20	4,86
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	2,94	2,86	2,71
	10,56	8,02	6,99	6,42	6,06	5,80	5,62	5,47	5,35	5,26	4,80	4,64	4,31
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,77	2,70	2,54
	10,04	7,56	6,55	5,99	5,64	5,39	5,21	5,06	4,95	4,85	4,41	4,25	3,91
20	4,35	3,49	3,10	2,87	2,71	2,60	2,52	2,45	2,40	2,35	1,12	2,04	1,84
	8,10	5,85	4,94	4,43	4,10	3,87	3,71	3,56	3,45	3,37	2,94	2,77	2,42
30	4,17	3,32	2,92	2,69	2,53	2,42	2,34	2,27	2,21	2,16	1,93	1,84	1,62
	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,06	2,98	2,55	2,39	2,01
∞	3,64	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,57	1,46	1,00
	6,63	4,61	3,78	3,32	3,02	2,80	2,64	2,51	2,41	2,32	1,88	1,70	1,00

сыворотке крови. Фактор «способ дачи препарата» имеет меньшее влияние по сравнению с фактором «тип препарата». Практический вывод из полученных результатов: для усиления естественной резистентности цыплят оптимально применять иммуномодулятор A_1 путем его выпойки.

Таким образом, для оценки влияния тех или иных факторов на изучаемый объект можно использовать дисперсионный анализ. Применение специальных статистических компьютерных программ существенно облегчает расчеты при большом количестве данных. ■