Применение специальных биолого-математических методов в агрономии

Любая разность между средними арифметическими двух генеральных совокупностей всегда существенна. В опытной работе обычно сравнивают средние величины не генеральных, а выборочных совокупностей, поэтому приходится доказывать методами математической статистики достоверность различий в действии изучаемых факторов и на основании этого делать выводы и предложения для производства.

Использование статистических методов или критериев проверки гипотез – основа принятия решений при изучении явлений, обусловленных случайной вариацией.

Статистическая гипотеза – это научное предположение о тех или иных статистических законах распределения рассматриваемых случайных величин, которое может быть проверено на основе выборки. Почти всегда задача заключается в проверке нулевой гипотезы (H_0) – предположении об отсутствии реального различия между фактическими и ожидаемыми наблюдениями. Если в результате проверки различия между фактическими и теоретическими наблюдениями близки к нулю или находятся в области допустимых значений, то нулевая гипотеза не опровергается (подтверждается). Если же различия оказываются в критической для данного статистического критерия области, невозможны при нашей гипотезе и поэтому несовместимы с ней, то гипотеза опровергается.

Принятие H_0 гипотезы означает, что данные не противоречат предположению об отсутствии различий между фактическими и теоретическими или двумя рядами фактических распределений. Отбрасывание гипотезы означает, что эмпирические данные несовместимы с H_0 и верна другая, альтернативная гипотеза.

Справедливость H_0 проверяется вычислением статистических критериев проверки для определенного уровня значимости. Для проверки H_0 используют критерии двух видов: параметрические и непараметрические.

Параметрические критерии – это критерии, основанные на предположении, что распределение признака в совокупности подчиняется известном закону (например, закону нормального распределения). К таким показателям относятся критерии t и F, использование которых предполагает вычисление оценок параметров распределения.

Непараметрические критерии – это критерии, использование которых не требует вычисления оценок неизвестных параметров распределения и приближенного значения распределения признака. Этими критериями пользуются только в предварительных исследованиях.

Статистические показатели выборочной совокупности (выборки) являются приближенными оценками показателей (параметров) генеральной совокупности. Оценка может быть точечная (представлена одним числом) или интервальная.

Через выборочную среднюю $\bar{x}$ и ее ошибку $S_{\bar{x}}$ можно дать точечную оценку генеральной средней в виде $\bar{x} \pm S_{\bar{x}}$. Это означает, что $\bar{x}$ – оценка генеральной средней μ с ошибкой, равной $S_{\bar{x}}$.

Интервальная оценка характеризуется двумя числами – концами интервала, покрывающего оцениваемый параметр.

Доверительный интервал – это такой интервал, который с заданной вероятностью покрывает оцениваемый параметр. Центр доверительного интервала – выборочная оценка точки, а пределы, или доверительные границы, определяются средней ошибкой оценки и уровнем вероятности.

Математически доверительный интервал для генеральной средней записывается так:

$$\bar{x} \pm t \cdot S_{\bar{x}},$$

где $t \cdot S_{\bar{x}}$ – предельная ошибка выборочной средней при данном числе свободы и принятом уровне значимости (t_{05} или t_{01}).

Крайние точки интервала (начало – $\bar{x} - t \cdot S_{\bar{x}}$; конец – $\bar{x} + t \cdot S_{\bar{x}}$) называют доверительными границами. Интервальную оценку используют для статистической проверки гипотез при сравнении выборочных средних. Например, при $n = 30$ были получены такие выборочные средние и ошибки средних:

$$\bar{x}_1 \pm S_{\bar{x}_1} = 37{,}4 \pm 0{,}3;$$
$$\bar{x}_2 \pm S_{\bar{x}_2} = 34{,}8 \pm 0{,}5.$$

Необходимо определить, существенно ли различаются эти выборочные средние при 5%-ном уровне значимости, т. е. проверить нулевую гипотезу H_0:

$$\mu_1 - \mu_2 = d = 0.$$

Для $\nu = n - 1 = 30 - 1 = 29$ степеней свободы $t_{05} = 2{,}05$, и при 95%-ном уровне вероятности доверительные интервалы составят:

$$\bar{x}_1 \pm t_1 \cdot S_{\bar{x}_1} = 37{,}4 \pm 2{,}05 \cdot 0{,}3 = 37{,}4 \pm 0{,}615 = 36{,}785 \div 38{,}015;$$
$$\bar{x}_2 \pm t_2 \cdot S_{\bar{x}_2} = 34{,}8 \pm 2{,}05 \cdot 0{,}5 = 34{,}8 \pm 1{,}025 = 33{,}775 \div 35{,}825.$$

Доверительные интервалы для генеральных средних перекрывают друг друга. Поэтому разность между выборочными средними $d = \bar{x}_1 - \bar{x}_2 = 2{,}6$ нельзя переносить на генеральные средние μ_1 и μ_2, так как генеральная разность между ними может быть равна нулю или быть отрицательной величиной, когда $\mu_2 > \mu_1$. Поэтому $H_0 : d = 0$ не отвергается.

Нулевую гипотезу об отсутствии существенных различий методом интервальной оценки можно определить другим способом – по формуле

$$S_d = \sqrt{S_{\bar{x}_1}{}^2 + S_{\bar{x}_2}{}^2},$$

где вначале вычисляется ошибка разности средних (S_d), а затем рассмотренным выше способом рассчитываются доверительные интервалы для генеральной разности средних $d \pm t \cdot S_1$.

Величина, указывающая границу предельным случайным отклонениям, называется наименьшей существенной разностью (НСР) и определяется по соотношению

$$\text{НСР} = t \cdot S_d.$$

Если фактическая разность между выборочными средними $d \geq$ НСР, то нулевая гипотеза отвергается, а если $d <$ НСР – не отвергается.

Наименьшая существенная разность используется при построении доверительных интервалов и проверке статистических гипотез. Доверительный интервал для разности генеральных средних определяется по отношению $d \pm$ НСР, где НСР $= t \cdot S_d$ – предельная ошибка разности выборочных средних при данном числе степеней свободы $\nu = n_1 + n_2 - 2$ и принятом уровне значимости.

Если по величине $\bar{x} \pm t \cdot S_{\bar{x}}$ оцениваются доверительные границы, в которых находится средняя арифметическая, то по величине стандартного отклонения S оценивается интервал для отдельного значения X и всей совокупности $\bar{x} \pm t \cdot S$. Внутри этого интервала будут находиться с 95%-ным или 99%-ным уровнем вероятности значения генеральной средней μ и все индивидуальные значения варьирующей величины, где величина $t \cdot S$ – область разброса индивидуальных значений.

При сравнении средних необходимо учитывать два момента:

1) средние двух независимых выборок сравнивают, когда единицы наблюдения первой и второй выборок не связаны общим условием;

2) сравнивают те две сопряженные выборки, единицы наблюдения одной из которых сопряжены каким-то общим условием с единицами наблюдения второй.

При сравнении средних двух независимых выборок по t-критерию Стъюдента оценивается существенность разности средних ($d = \bar{x}_1 - \bar{x}_2$), а при сравнении двух сопряженных выборок – существенность средней разности ($d = \sum d : n$).

Оценка разности средних независимых выборок по *t*-критерию и НСР$_{05}$

В теории статистики ошибка разности или суммы средних арифметических выборок при одинаковом числе наблюдений ($n_1 = n_2$) определяется по формуле

$$S_d = \sqrt{S_{\bar{x}_1}{}^2 + S_{\bar{x}_2}{}^2},$$

где S_d – ошибка разности или суммы;

$S_{\bar{x}_1}$ и $S_{\bar{x}_2}$ – ошибка сравниваемых средних арифметических $\bar{x}_1$ и $\bar{x}_2$.

Гарантией надежности вывода о существенности или несущественности различий между $\bar{x}_1$ и $\bar{x}_2$ служит отношение разницы к ее ошибке. Это отношение получило название критерия существенности разности:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{S_{\bar{x}_1}{}^2 + S_{\bar{x}_2}{}^2}} = \frac{d}{S_d}.$$

Если $t_{\text{факт}} > t_{\text{теор}}$, то нулевая гипотеза об отсутствии существенности различий между средними опровергается, а если различия находятся в пределах случайных колебаний для принятого уровня значимости (т. е. $H_0 : \mathrm{d} = 0$), – не опровергается.

Теоретическое значение критерия t находят по таблице, зная, что число степеней свободы $\nu = n_1 + n_2 - 2$ по принятому уровню значимости.

Проверить H_0 можно и по величине НСР, выраженной в единицах варьирующего признака. Когда разность между средними ($d >$ НСР) попадает в критическую область существенных различий, она признается значимой и H_0 опровергается. Когда d < НСР и лежит в области случайных колебаний, то H_0 не опровергается.

Например, в борьбе с аскохитозом гороха было проведено протравливание семян 80%-ным ТМТД (4 кг/т) и был подсчитан процент поражения растений. Необходимо определить 5%-ные доверительные интервалы и проверить значимость действия препарата.

В связи с тем, что пробные площадки для учета поражения растений аскохитозом не связаны общим условием, полученные данные следует обрабатывать по типу несопряженных выборок. Последовательность вычислений при сравнении средних несопряженных выборок приведена в табл. 1.

Таблица 1. **Последовательность вычислений несопряженной выборки**

Номер пробной площадки	Контроль (без протравливания)			При обработке семян		
	Процент поражения аскохитозом	$\bar{x}_1 = X - A_1$	$\bar{x}_1^{\,2}$	Процент поражения аскохитозом	$\bar{x}_2 = X - A_2$	$\bar{x}_2^{\,2}$
1	41	$\bar{x}_1 = 41 - 34 = 7$	49	10	$\bar{x}_2 = 10 - 10 = 0$	0
2	29	$\bar{x}_1 = 29 - 34 = -5$	25	11	$\bar{x}_2 = 11 - 10 = 1$	1
3	37	$\bar{x}_1 = 37 - 34 = 3$	9	15	$\bar{x}_2 = 15 - 10 = 5$	25
4	29	$\bar{x}_1 = 29 - 34 = -5$	25	13	$\bar{x}_2 = 13 - 10 = 3$	9
5	32	$\bar{x}_1 = 32 - 34 = -2$	4	8	$\bar{x}_2 = 8 - 10 = -2$	4
6	25	$\bar{x}_1 = 25 - 34 = -9$	81	11	$\bar{x}_2 = 11 - 10 = 1$	1
7	26	$\bar{x}_1 = 26 - 34 = -8$	64	8	$\bar{x}_2 = 8 - 10 = -2$	4
8	27	$\bar{x}_1 = 27 - 34 = -7$	49	9	$\bar{x}_2 = 9 - 10 = -1$	1
9	45	$\bar{x}_1 = 45 - 34 = 11$	121	10	$\bar{x}_2 = 10 - 10 = 0$	0
10	51	$\bar{x}_1 = 51 - 34 = 17$	289	5	$\bar{x}_2 = 5 - 10 = -5$	25
	$\sum X = 342$	$\sum \bar{x}_1 = 2$	$\sum \bar{x}_1^{\,2} = 716$	$\sum X = 100$	$\sum \bar{x}_2 = -5$	$\sum \bar{x}_2^{\,2} = 70$
	$\bar{x}_1 = \frac{\sum X}{n} = \frac{342}{10} = 34{,}2$		$A_1 = 34$	$\bar{x}_1 = \frac{\sum X}{n} = \frac{100}{10} = 10$		$A_1 = 10$

Расчеты по каждой выборке проводят в следующем порядке:

1) сумма квадратов

$$\sum (X - \bar{x}_1)^2 = \sum \bar{x}_1^{\,2} - \frac{(\sum \bar{x}_1)^2}{n} = 716 - \frac{4}{10} = 715{,}6;$$

$$\sum (X - \bar{x}_2)^2 = \sum \bar{x}_2^{\,2} - \frac{(\sum \bar{x}_2)^2}{n} = 70 - \frac{25}{10} = 67{,}5;$$

2) ошибка выборочной средней

$$S_{\bar{x}} = \sqrt{\frac{\sum (X - \bar{x})^2}{n \cdot (n-1)}};$$

$$S_{\bar{x}_1} = \sqrt{\frac{715{,}6}{10 \cdot (10-1)}} = 2{,}82;$$

$$S_{\bar{x}_2} = \sqrt{\frac{67{,}5}{10 \cdot (10-1)}} = 0{,}87;$$

относительная ошибка средней

$$S_{\bar{x}}\% = \frac{S_{\bar{x}}}{\bar{x}} \cdot 100\ \%;$$

3) доверительный интервал для генеральной средней

$$\bar{x} \pm t_{05} \cdot S_{\bar{x}}.$$

При степенях свободы $\nu = n - 1 = 10 - 1 = 9$ по прил. 1 находим t_{05}, которое равно 2,26. Далее получим:

$$\bar{x}_1 \pm t_{05} \cdot S_{\bar{x}_1} = 34{,}2 \pm 2{,}26 \cdot 2{,}82 = 34{,}2 \pm 6{,}4 = 27{,}8 \div 40{,}6;$$
$$\bar{x}_2 \pm t_{05} \cdot S_{\bar{x}_2} = 10 \pm 2{,}26 \cdot 0{,}87 = 10{,}0 \pm 2{,}0 = 8{,}0 \div 12{,}0;$$

4) критерий существенности

$$t_{\text{факт}} = \frac{d}{S_d} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{S_{\bar{x}_1}{}^2 + S_{\bar{x}_2}{}^2}} = \frac{34{,}2 - 10{,}0}{\sqrt{2{,}82^2 + 0{,}87^2}} = \frac{24{,}2}{\sqrt{7{,}95 + 0{,}76}} = 8{,}2.$$

Полученный критерий $t_{\text{факт}}$ сравнивается с теоретическим, который находится по прил. 1 при числе степеней свободы $\nu = n_1 + n_2 - 2 = 10 + + 10 - 2 = 18$, $t_{\text{теор}} = 2{,}1$ на 5%-ном уровне значимости. Сопоставляя фактическое значение с теоретическим, приходим к выводу, что $t_{\text{факт}} > t_{\text{теор}}$.

Вывод. Разность существенна при 5%-ном уровне значимости.

Задание. Дать точечную и интервальную оценку генеральной совокупности по выборочной средней $\bar{x}$, ошибке среднего арифметического $S_{\bar{x}}$ и среднеквадратическому отклонению S.

Определить, существенно ли различаются эти выборочные средние на 1 %-ном и 5 %-ном уровнях значимости (т. е. проверить нулевую гипотезу) по t-критерию и НСР для независимых выборок.

Исходные данные приведены в табл. 2.

Таблица 2. **Исходные данные к заданию**

№ п. п.	$\bar{x}_1$	$S_{\bar{x}_1}$	S_1	n_1	$\bar{x}_2$	$S_{\bar{x}_2}$	S_2	n_2
1	73,41	0,56	5,03	15	76,76	0,70	6,30	20
2	81,46	0,83	7,44	20	86,28	0,84	7,53	25
3	1,88	0,09	0,77	25	1,95	0,09	0,78	15
4	6,25	0,26	2,35	30	5,93	0,23	2,07	10
5	4,84	0,14	1,22	35	4,58	0,14	1,28	25
6	19,20	0,91	8,16	40	16,40	0,79	7,10	45
7	16,10	0,64	5,74	45	13,64	0,57	5,06	40
8	79,64	3,94	35,43	50	68,26	3,40	30,43	55
9	68,59	2,96	26,48	55	58,21	2,55	22,81	50
10	80,26	0,63	5,65	60	74,74	0,65	5,82	25
11	87,75	0,95	8,52	10	85,49	0,96	8,60	15
12	1,98	0,09	0,83	65	1,93	0,09	0,82	60
13	5,04	0,23	2,08	15	5,16	0,25	2,24	70
14	3,83	0,13	1,14	35	3,80	0,10	0,89	15
15	13,71	0,64	5,74	30	14,79	0,76	6,77	25
16	11,48	0,42	3,79	15	12,08	0,43	3,83	20
17	59,23	2,80	25,04	10	65,24	3,33	29,78	15
18	51,36	2,01	17,98	25	55,36	2,11	18,91	20
19	68,75	1,09	8,42	10	73,53	1,06	8,19	25
20	71,60	1,06	8,18	15	74,85	1,06	8,22	10

Оценка существенности разности средних сопряженных выборок по *t*-критерию и НСР

Для оценки существенности разности средних сопряженных выборок по *t*-критерию и НСР расчеты проводят в следующем порядке:

1) вычисляют S_d по формуле

$$S_{\bar{d}} = \sqrt{\frac{\sum d^2 - (\sum d)^2 : n}{n\,(n-1)}}\ ;$$

2) вычисляют $t_{\text{факт}}$ по формуле

$$t_{\text{факт}} = \frac{\bar{d}}{S_{\bar{d}}};$$

3) число степеней свободы ν находят по равенству

$$\nu = n - 1,$$

где n – число сопряженных пар;

4) вычисляют НСР по формуле

$$\text{НСР}_{05} = t_{05} \cdot S_d;$$

5) на основании решения задачи конкретного варианта делают письменные выводы.

Все расчеты удобнее проводить по форме табл. 3.

Таблица 3. **Последовательность вычисления сопряженных выборок**

Номер пары наблюдений	X_1	X_2	Разность $d = X_1 - X_2$	Квадрат разности d^2
1 2 … n				
Сумма	$\sum X_1$	$\sum X_1$	$\sum d$	$\sum d^2$
Среднее $\bar{x}$	$\bar{x}_1 = \frac{\sum X_1}{n}$;	$\bar{x}_2 = \frac{\sum X_2}{n}$		

Задание. Дать точечную и интервальную оценку генеральной совокупности по выборочной средней $\bar{x}$, ошибке среднего арифметического $S_{\bar{x}}$ и стандартному отклонению S.

Определить, существенно ли различаются эти выборочные средние на 1 %-ном и 5 %-ном уровнях значимости (т. е. проверить нулевую гипотезу) по *t*-критерию и НСР для независимых выборок.

Исходные данные приведены в табл. 4.

Таблица 4. **Исходные данные к заданию**

№ п. п.	$\bar{x}_1$	$S_{\bar{x}_1}$	S_1	$n_1 = n_2$	$\bar{x}_2$	$S_{\bar{x}_2}$	S_2
1	77,10	0,87	6,75	25	63,32	0,96	7,42
2	79,10	0,88	6,85	30	66,67	1,04	8,03
3	5,22	0,29	2,22	15	5,65	0,26	2,00
4	4,08	0,12	0,96	35	4,27	0,11	0,86
5	15,48	0,65	5,02	10	18,05	0,83	6,45
6	13,65	0,40	3,11	40	15,32	0,53	4,14
7	58,97	2,24	17,39	15	65,73	3,03	23,45
8	53,13	1,70	13,16	55	50,52	2,29	17,76
9	37,60	0,81	3,63	20	34,85	0,98	4,38
10	48,75	2,19	9,79	60	43,65	1,45	6,47
11	2,30	0,19	0,86	25	2,35	0,23	1,04
12	6,20	0,73	3,27	50	6,15	0,80	3,56
13	3,25	0,29	1,29	30	2,90	0,27	1,21
14	21,65	2,61	11,66	45	21,35	2,81	12,58
15	12,10	1,19	5,33	35	10,10	0,95	4,24
16	61,25	0,46	6,16	40	59,05	0,42	5,64
17	63,19	0,59	7,97	25	60,19	0,45	5,98
18	4,37	0,19	2,61	10	4,97	0,21	2,84
19	5,11	0,27	3,58	30	4,92	0,23	3,13
20	18,18	0,57	7,69	15	13,75	0,69	9,22
21	16,40	0,69	9,20	20	14,61	0,75	10,02
22	3,14	0,14	1,82	45	2,06	0,08	1,04
23	8,38	0,41	5,54	35	4,66	0,22	2,94
24	10,82	0,49	6,58	15	6,48	0,25	3,33
25	28,32	1,40	18,76	10	13,30	0,65	8,70

Оценка соответствия между наблюдаемыми (фактическими) и ожидаемыми (теоретическими) распределениями по критерию χ^2 (хи-квадрат)

Критерий χ^2 применяют в том случае, когда необходимо установить соответствие двух сравниваемых рядов распределения – эмпирического (фактического) и теоретического или двух эмпирических. Особенно широко критерий соответствия используется в генетическом анализе, когда необходимо убедиться в том, является ли обнаруженное отклонение от теоретически ожидаемого расщепления (1:1; 1:2:1; 3:1; 9:3:4; 9:3:3:1; 9:7 и т. д.) отклонением закономерным или оно лежит в пределах возможных случайных колебаний. Этот показатель используется при изучении качественных признаков для оценки соответствия эмпирических данных теоретической предпосылке, нулевой гипотезе (H_0). Гипотеза отвергается, если $\chi^2_{\text{факт}} > \chi^2_{\text{теор}}$, и не отвергается, если $\chi^2_{\text{факт}} < \chi^2_{\text{теор}}$ или когда фактические и теоретические ожидаемые частоты совпадают ($\chi^2 = 0$).

Теоретически ожидаемые частоты обозначают через M, а опытные, эмпирически полученные, – через m. Б. А. Доспехов данные частоты обозначает соответственно через F и f, что неудобно, так как через F обозначают в генетике и селекции соответствующее поколение гибрида, а в дисперсионном анализе F – это критерий Фишера. Общей мерой отклонения фактических данных от теоретических, т. е. критерия соответствия χ^2, будет сумма отношений квадратов разностей между частотами эмпирического и теоретического распределений к частотам теоретического распределения для данной группы:

$$\chi^2 = \frac{(m_1 - M_1)^2}{M_1} + \frac{(m_2 - M_2)^2}{M_2} + \ldots + \frac{(m_n - M_n)^2}{M_n} = \sum \frac{(m - M)^2}{M}.$$

Критерий χ^2 широко используется в генетическом анализе соответствия расщепления гибридов теоретически ожидаемому, для оценки независимости (или сопряженности), в распределении объектов выборки (наиболее широко применяют при решении токсикологических и других задач в защите растений), определения степени соответствия фактического распределения изучаемого признака нормальному и оценки соответствия двух эмпирических распределений.

Следует помнить, что в формулу для определения χ^2 должны подставляться только частоты. Величина χ^2 зависит от числа степеней свободы. В наиболее типичных случаях число степеней свободы ν определяется по формуле

$$\nu = (c - 1) \cdot (k - 1),$$

где c – число строк;

k – число колонок в анализируемой таблице.

Пример. В результате расщепления во втором поколении гибридов люпина по алкалоидности были получены следующие результаты: алкалоидных семян – 679 шт., безалкалоидных – 483 шт. Необходимо установить, соответствует ли эмпирическое распределение частей (m) теоретически ожидаемому наследованию в соотношении 9 : 7.

Расчет осуществляем согласно табл. 5.

Таблица 5. **Последовательность расчета критерия χ^2**

Показатели	Расщепление в F_2		Сумма
	Алкалоидные	Безалкалоидные	
Ожидаемое расщепление H_0	9	7	16
Наблюдаемые частоты m	679	483	1162
Ожидаемые частоты M	654	508	1162
Разность $m - M$	+25	–25	–
Квадрат разности $(m - M)^2$	625	625	–
Соотношение $\frac{(m-M)^2}{M}$	0,96	1,23	2,19 = χ^2

Ожидаемые частоты определяем умножением теоретически ожидаемой доли растений на общее число наблюдений. Так, доля алкалоидных растений должна быть равна 9/16 и, следовательно,

$$M_1 = 9/16 \cdot 1162 = 654,$$
$$M_2 = 7/16 \cdot 1162 = 508.$$
$$M_1 + M_2 = 654 + 508 = 1162.$$

Подставляя эмпирические и теоретические ожидаемые частоты в формулу для определения χ^2, получают:

$$\chi^2 = \sum \frac{(m-M)^2}{M} = \frac{(679-654)^2}{654} + \frac{(483-508)^2}{508} = 2{,}19.$$

Вывод. При $k - 1 = 1$ теоретическое значение $\chi^2 = 3{,}84$. Так как $\chi^2_{\text{факт}} < \chi^2_{\text{теор}}$, то нулевая гипотеза не отвергается.

Пример. Из 50 имаго колорадского жука было обработано актарой 30 имаго (опытные насекомые), а 20 не обрабатывались (контроль). В опытной группе погибло 28 жуков, а в контрольной – 4.

Доказывают ли результаты опыта токсическое действие актары или гибель жуков зависит от случайных причин?

Обработка результатов опыта и расчеты приведены в табл. 6 и 7. Напоминаем, что m и M – соответственно фактически наблюдаемое и теоретически ожидаемое количество насекомых.

Таблица 6. **Токсичность актары и вычисление ожидаемого количества насекомых по группам (схема 2×2)**

Группа насекомых	Количество насекомых				Σ	%
	погибших		здоровых			
	m_1	M	m_2	M		
Опытная	28	19,2	2	10,8	30	60
Контрольная	4	12,8	16	7,2	20	40
Σ	32	32	18	18	50	100
%	64	–	36	–	–	100

Таблица 7 **Различия между фактическими и ожидаемыми количествами насекомых в группах (*m* – *M*)**

Группа насекомых	Количество насекомых		Сумма
	погибших	здоровых	
Опытная	8,8	–8,8	0
Контрольная	–8,8	8,8	0
$\sum$	0	0	0

Количество насекомых по группам распределилось следующим образом:
– погибших в опытной группе

$$M_1 = \frac{60 \cdot 32}{100} = 19{,}2;$$

– здоровых в опытной группе

$$M_2 = \frac{60 \cdot 18}{100} = 10{,}8;$$

– погибших в контрольной группе:

$$M_3 = \frac{40 \cdot 32}{100} = 12{,}8;$$

– здоровых в контрольной группе:

$$M_4 = \frac{40 \cdot 18}{100} = 7{,}2;$$

$$\chi^2 = \sum \frac{(m - M)^2}{M} = \frac{8{,}8^2}{19{,}2} + \frac{(-8{,}8)^2}{10{,}8} + \frac{(-8{,}8)^2}{12{,}8} + \frac{8{,}8^2}{7{,}2} =$$

$$= 4{,}03 + 7{,}17 + 6{,}05 + 10{,}76 = 28{,}01.$$

При $n = (c - 1) \cdot (k - 1) = (2 - 1) \cdot (2 - 1) = 1$ теоретическое значение $\chi^2_{05} = 3{,}84$ (прил. 2)

Вывод. Наблюдается существенное увеличение гибели имаго колорадского жука после обработки их актарой ($\chi^2_{\text{факт}} > \chi^2_{05}$). Нулевая гипотеза о независимости гибели вредителя от актары отвергается.

Задание. Соответствуют ли эмпирические расщепления частот *m* теоретически ожидаемому наследованию в соотношении 9 : 4 : 3.

Исходные данные приведены в табл. 8.

Таблица 8. **Исходные данные к заданию**

№ п. п.	Семена		
	серые	белые	мраморные
	9	4	3
1	657	284	219
2	261	124	97
3	467	215	156
4	384	162	124
5	291	134	103
6	262	162	–
7	521	243	186
8	571	276	207
9	662	324	238
10	344	168	126
11	372	189	149
12	440	217	169
13	419	222	163
14	653	303	238
15	591	283	212
16	679	237	314
17	654	320	232
18	426	224	166
19	422	215	157
20	398	207	157
21	361	187	132
22	242	117	87
23	300	148	112
24	353	168	129
25	415	193	143