

Приклади розв'язування задач на перевірку статистичних гіпотез

1. Підприємство виготовляє однакові деталі двома способами. Першим способом виготовлено 10 деталей, витрати сировини були такими 1,4;1,6;1,2;1,5;1,4;1,6;1,5;1,8;1.1;1,4

Другим способом виготовлено 6 деталей, витрати сировини були такими 1,8;1,7;1,9;1,3;1,6;1,5. Припускаючи, що дисперсія витрат сировини однакова, при рівні значущості $\alpha = 0.02$ перевірити гіпотезу

$H_0 : \sigma_1 = \sigma_2 = 0$ при альтернативній гіпотезі

$H_1 : \sigma_1 \neq \sigma_2 \neq 0$.

Розв'язування

1-й спосіб. Треба перевірити гіпотезу про рівність математичних сподівань двох нормальних генеральних сукупностей. Згідно з гіпотезою H_0 , критична область буде двобічною і визначається умовою

$$P\left(v > v_{\frac{\alpha}{2}}\right) = P\left(v < -v_{\frac{\alpha}{2}}\right) = \frac{\alpha}{2}$$

де статистична характеристика v визначена формулою (13.1). Ступінь волі дорівнює $n_1 + n_2 - 2 = 10 + 6 - 2 = 14$. З таблиці критичних значень α для 14 ступенів волі одержимо

$$v_{\frac{\alpha}{2}} = 2,62$$

За даними вибірки треба знайти \bar{x}_1, S_1^2 і \bar{x}_2, S_2^2 , Для цього на лист Excel занесемо до діапазону A2.:J2 дані першої вибірки, а до діапазону A1:F1 дані другої вибірки. Далі, послідовно викликаючи функції СРЗНАЧ(A2:J2), ДИСП(A2.:J2), СРЗНАЧ (A1:F1), fMOI(A1:F1), отримаємо

$$\bar{x}_1 = 1,45; \quad S_1^2 = 0,04; \quad \bar{x}_2 = 1,63; \quad S_2^2 = 0,05.$$

Тепер за формулою (13.1) обчислюємо

$$v_{cp} = 1,717.$$

Отже, $1,72 < 2,62$, тому гіпотеза H_0 може бути прийнята.

2-й спосіб. Те ж саме можна зробити за допомогою пакета аналізу, що входить до складу Microsoft Excel. Для цього треба вибрати команду **Анализ данных** в меню **Сервис**. Далі вибрати розділ **Двухвыборочный t-тест с одинаковыми дисперсиями** і виконувати команди в режимі діалогу. Результати розрахунків наведені в табл. 13.1.

Таблиця 13.1

Величина	Позначення в пакеті аналізу	Змінна 2	Змінна 1
\bar{x}_1	Среднее	1.63	1.45
S_1^2	Дисперсия	0,05	0,04
n_1	Наблюдения	6	10
	Объединенная дисперсия	0,0427	

	Гипотетическая разность средних	0	
Ступінь волі	Df	14	
V_{crit}	t-статистика	1.717	
	P(T<=0 одностороннее)	0,0539	
	t критическое одностороннее	2.2637	
a/2	t критическое двухстороннее	0,1079	
Критичне значення	t критическое двухстороннее	2,6244	

1. Є дані про вартість (в тис. гривень) проданих трьох видів виробів певним магазином в окремі дні тижня

Таблиця 13.2

вівторок	середа	четвер	п'ятниця	субота
10,2	10,8	10,7	13,0	12,0
11,5	9,8	11,5	13,2	11,5
12,0	12,1	12,0	11,5	11,8

Припускаючи нормальний закон розподілу одержаної суми кожного дня та рівність дисперсій, перевірити гіпотезу $H_0 : a_1 = a_2 = a_3 = a_4 = a_5$ при рівні значущості $\alpha = 0.05$.

Розв'язування

Умови прикладу дозволяють застосувати до розв'язання задачі критерій дисперсійного аналізу. Рішення знайдемо за допомогою пакета аналізу, що входить до складу Microsoft Excel. Для цього треба перенести таблицю 13.2 на лист Excel і вибрати команду **Анализ данных** в меню **Сервис**. Далі вибрати розділ **Однофакторный дисперсионный анализ** і виконувати команди в режимі діалогу. Результати розрахунків наведені в табл. 13.3. (збережено російську мову оригіналу)

Таблиця 13.3

Группы	Счет	Сумма	Среднее	Дисперсия
Столбец 1	3	33,7	11,23333333	0,863333333
Столбец 2	3	32,7	10,9	1,33
Столбец 3	3	34,2	11,4	0,43
Столбец 4	3	37,7	12,56666667	0,863333333
Столбец 5	3	35,3	11,76666667	0,063333333

Таблиця 13.4

Источники вариации	ss	df	MS	F	P-Значение	F критическое
Между группами	4,869333	4	1,217333333	0,2228281	3,47804984	
и внутри групп	333	12	1,714553991	12	8	
	7,1	16	0,71			

Итого 11,96933 1
333 4

Де через SS позначено обчислення сум, що входять до чисельника (рядок **Между группами**) і знаменника (рядок **Внутри групп**) формули (13.4). Через df позначено величини $N - 1$ та $n - N$ ступенів волі (у цьому випадку маємо: $N = 5$, $n = 15$). Через MS позначено чисельник (рядок **Между группами**) і знаменник (рядок **Внутри групп**) формули (13.4). Через F позначено результат обчислень за формулою (13.4). Через **F критическое** позначено критичну точку розподілу Фішера з 4,10 ступенів волі та рівнем значущості $\alpha = 0,05$. Оскільки

$$F = 1,714 < F_{кр} = 3,48,$$

то гіпотеза H_0 може бути прийнята.

3. За рівнем значущості $\alpha = 0,05$ перевірити гіпотезу про нормальний розподіл генеральної сукупності, якщо відомі емпіричні та теоретичні частоти

n_k	6	13	38	74	106	85	30	14
n'_k	3	14	42	82	99	76	37	13

Розв'язування

У даному випадку теоретичні частоти n'_k задані, кількість варіант вибірки $m=8$. Тому за формулою (13.5) знаходимо ступінь волі $k = 8 - 3 = 5$.

З таблиці критичних точок розподілу $\chi^2(\alpha, k)$ для $\alpha = 0,05$ та $k = 5$ знаходимо $\chi^2_{кр} = 11,1$.

Для обчислення $\chi^2_{сн}$ за формулою (13.6) використаємо розрахункову таблицю

n_k	n'_k	$n_k - n'_k$	$(n_k - n'_k)^2$	$(n_k - n'_k)^2 / n'_k$
6	3	-3	9	3
13	14	-1	1	0,07
38	42	-4	16	0,38
74	82	-8	64	0,78
106	99	7	49	0,49
85	76	9	81	1,07
30	37	-7	49	1,32
14	13	1	1	0,08
				$\chi^2_{сн} = 7,19$

Таким чином, $\chi^2_{сн} = 7,19 < 11,1 = \chi^2_{кр} >$ тому за правилом Пірсона гіпотезу H_0 слід прийняти. Отже, дані вибірки узгоджуються з гіпотезою H_0 , тому що розбіжність емпіричних та теоретичних частот незначна.

Завдання для самостійної роботи

1. За рівнем значущості $\alpha = 0,01$ перевірити гіпотезу про нормальний розподіл генеральної сукупності, якщо відомі емпіричні та теоретичні частоти

n_k	8	16	40	72	36	18	10
n^*_k	6	18	36	76	39	18	7

$$\{ \chi^2_{\text{ен}} = 3,061 < 13,3 = \chi^2_{\text{кр}} \}$$

2. Хімічна лабораторія перевіряє м'ясо курчат на вміст діоксину. В першій партії проаналізовано вісім проб, у таблиці наведено вміст речовини в міліграмах на кілограм

15	20	16	22	24	14	18	20
----	----	----	----	----	----	----	----

В другій партії проаналізовано 10 проб

17	21	15	26	13	11	10	18	21	22
----	----	----	----	----	----	----	----	----	----

Припускаючи, що дисперсія вимірів однакова, за рівнем значущості $\alpha = 0,02$ перевірити гіпотезу про рівність середнього вмісту діоксину. $\{v_{\text{ен}} = 0,570 < 2,583 = v_{\text{кр}}\}$

3. Для перевірки впливу матеріального стимулювання шести групам з п'яти чоловік були запропоновані задачі однакової складності. Чим більше номер групи, тим вище рівень стимулювання. У таблиці наведено кількість задач, що розв'язана членами кожної групи. Використовуючи однофакторний дисперсійний аналіз перевірити гіпотезу про відсутність ефектів впливу матеріального стимулювання.

Група 1	Група 2	Група 3	Група 4	Група 5	Група 6
10	8	12	12	24	19
11	10	17	15	16	18
9	16	14	16	22	27
13	13	9	16	18	25
7	12	16	19	20	24

$$\{F = 12,637 > F_{\text{об}} = 2,621\}$$