Лабораторная работа 3.
Критерии значимости

Теоретические сведения

Основные задачи математической статистики разделяют на две категории, тесно связанные между собой, но отличающиеся постановкой задач: оценивание параметров и проверка статистических гипотез. Основной задачей оценивания параметров является получение по выборке оценок, наилучших в том или ином смысле. При проверке гипотез задача ставится иначе: требуется по выборке принять или отвергнуть некоторое предположение о распределении генеральной совокупности, из которой извлечена выборка.

Статистической гипотезой называется любое предположение о виде (***непараметрическая гипотеза***) или параметрах (***параметрическая гипотеза***) неизвестного распределения.

Одну из гипотез выделяют в качестве ***основной*** (или ***нулевой***) H_0, а другую, являющуюся логическим отрицанием H_0, – в качестве ***конкурирующей*** (или ***альтернативной***) гипотезы $\bar{H}$.

Правило, по которому принимается решение принять или отклонить проверяемую гипотезу, называется ***критерием проверки статистической гипотезы*** (***статистическим критерием***). При этом заранее выбирают допустимое значение ошибки вывода, которое называется ***уровнем значимости*** статистического критерия и обозначается α (это вероятность отвергнуть нулевую гипотезу, когда она верна).

Статистические критерии, с помощью которых проверяются гипотезы о значениях параметров распределения или о соотношениях между ними в предположении, что тип распределения известен, называются ***критериями значимости,*** или ***параметрическими критериями***.

Рассмотрим критерии значимости, предназначенные для проверки гипотез о значениях параметров *в случае выборок из нормального распределения*, которое на практике встречается наиболее часто. (В случае нарушения предположения о нормальном распределении выборок необходимо использовать другие критерии.)

Нормальное распределение имеет два параметра: математическое ожидание a и дисперсию σ^2, которые оцениваются с помощью выборочного среднего $\bar{x}$ и выборочной дисперсии (исправленной выборочной дисперсии s^2) соответственно. Выборочное среднее является оценкой для среднего значения измеряемой величины и может служить оценкой того или иного показателя качества. Дисперсия характеризует разброс экспериментальных значений, а следовательно, служит мерой точности. Например, если произведено несколько измерений одной и той же величины, то дисперсия может характеризовать точность прибора, метода измерения и т. д.

1. Проверка гипотезы о равенстве математического ожидания нормального распределения заданному значению.

Нулевая гипотеза $H_0:\ a = a_0$.

Альтернативная гипотеза $\bar{H}:\ a \neq a_0$.

Требуется по выборке объема n проверить гипотезу H_0 при заданном уровне значимости α. При этом предполагается, что выборка взята из нормально распределенной генеральной совокупности.

1 случай. Если дисперсия σ^2 *известна, то гипотеза* H_0 *принимается* (т. е. согласуется с результатами наблюдений) *при условии, что*

$$u_{\text{расч}} = \frac{|\,\overline{x} - a_0\,|}{\sqrt{\sigma^2 / n}} < u_{\text{табл}} = u_\alpha, \qquad (1)$$

где квантиль u_α удовлетворяет соотношению $\Phi(u_\alpha) = \dfrac{1-\alpha}{2}$.

2 случай. Если дисперсия σ^2 *неизвестна, то гипотеза* H_0 *принимается при*

$$t_{\text{расч}} = \frac{|\,\overline{x} - a_0\,|}{\sqrt{s^2 / n}} < t_{\text{табл}} = t_{\alpha;\, n-1}, \qquad (2)$$

где квантиль $t_{\alpha;\, n-1}$ определяется по таблице распределения Стьюдента.

2. Проверка гипотезы о равенстве заданному значению дисперсии нормального распределения.

Нулевая гипотеза $H_0 : \sigma^2 = \sigma_0^2$.

Альтернативная гипотеза $\overline{H} : \sigma^2 \neq \sigma_0^2$.

Гипотеза H_0 при заданном уровне значимости α *принимается, если*

$$\chi^2_{1-\alpha/2;\, n-1} < \chi^2_{\text{расч}} = \frac{(n-1)s^2}{\sigma_0^2} < \chi^2_{\alpha/2;\, n-1}, \qquad (3)$$

где квантили $\chi^2_{1-\alpha/2;\, n-1}$ и $\chi^2_{\alpha/2;\, n-1}$ определяются по таблице распределения χ^2.

3. Сравнение двух дисперсий нормально распределенных признаков. Такая задача возникает, если требуется сравнить точность приборов, инструментов, методов измерения. Лучшим будет тот прибор, инструмент, метод, который дает меньший разброс результатов, т. е. меньшую дисперсию.

Нулевая гипотеза $H_0 : \sigma_1^2 = \sigma_2^2$.

Альтернативная гипотеза $\overline{H} : \sigma_1^2 \neq \sigma_2^2$.

Пусть для первой дисперсии по выборке объема n_1 найдена несмещенная оценка s_1^2, для второй – по выборке объема n_2 оценка s_2^2.

Гипотеза H_0 *при заданном уровне значимости* α *принимается, если*

$$F_{\text{расч}} = \frac{s^2_{\max}}{s^2_{\min}} < F_{\text{табл}} = F_{\alpha/2;\, f_1;\, f_2}. \qquad (4)$$

Здесь $F_{\text{расч}}$ равно отношению *большей* несмещенной оценки дисперсии к *меньшей*, квантиль $F_{\alpha/2;\, f_1;\, f_2}$ определяется по таблице распределения Фишера, причем f_1 и f_2 – числа степеней свободы *соответственно* числителя и знаменателя, т. е. *большей* и *меньшей* оценок дисперсий.

Если гипотеза о равенстве дисперсий принимается, то эти дисперсии считаются ***однородными***. (Термин «однородные» в статистике означает «являющиеся оценкой одного и того же параметра».)

Замечание 1. Критерий Фишера (4) может использоваться также для проверки гипотезы о равенстве дисперсии заданному значению σ_0^2. В этом случае число степеней свободы известной дисперсии принимается равным ∞: $f_{\sigma_0^2} = \infty$.

Замечание 2. Критерий Фишера (4) может применяться также для проверки гипотезы о равенстве нескольких дисперсий нормально распределенных признаков. В этом случае проверяют гипотезу о равенстве наибольшей и наименьшей из сравниваемых дисперсий. Если они признаются однородными, то можно принять гипотезу о равенстве всех сравниваемых дисперсий.

4. Сравнение двух средних в случае независимых нормально распределенных признаков.

Нулевая гипотеза $H_0: \ a_1 = a_2$.

Альтернативная гипотеза $\overline{H}: a_1 \neq a_2$.

Требуется по выборкам объемов n_1 и n_2 проверить гипотезу H_0 при заданном уровне значимости α.

1 случай. Если дисперсии σ_1^2 и σ_2^2 *известны*, то *гипотеза* H_0 *принимается при условии, что*

$$u_{\text{расч}} = \frac{|\,\overline{x}_1 - \overline{x}_2\,|}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} < u_{\text{табл}} = u_\alpha, \tag{5}$$

где квантиль u_α удовлетворяет соотношению $\Phi(u_\alpha) = \dfrac{1-\alpha}{2}$.

2 случай. Если дисперсии σ_1^2 и σ_2^2 не известны, но на основании проверки соответствующей гипотезы по критерию Фишера признаны *однородными*, то *гипотеза* H_0 *принимается при*

$$t_{\text{расч}} = \frac{|\,\overline{x}_1 - \overline{x}_2\,|}{\sqrt{s^2\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}} < t_{\text{табл}} = t_{\alpha;\,f}, \tag{6}$$

где общая средневзвешенная дисперсия s^2 вычисляется по формуле

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

и имеет число степеней свободы $f = n_1 + n_2 - 2$; значение $t_{\alpha;f}$ определяется по таблице квантилей распределения Стьюдента.

3 случай. Если дисперсии σ_1^2 и σ_2^2 не известны и на основании проверки по критерию Фишера признаны *неоднородными*, то проверка также проводится по критерию Стьюдента, однако этот критерий является приближенным. В этом случае *гипотеза* H_0 *принимается, если*

$$t_{\text{расч}} = \frac{|\,\overline{x}_1 - \overline{x}_2\,|}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} < t_{\text{табл}} = t_{\alpha;f}, \tag{7}$$

где квантиль $t_{\alpha;f}$ определяется по таблице распределения Стьюдента при

$$f \approx \frac{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^2}{\dfrac{\left(\dfrac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \dfrac{\left(\dfrac{s_2^2}{n_2}\right)^2}{n_2 - 1}}.$$

Отметим, что при сравнении двух средних в случае неизвестных дисперсий возникает необходимость проверки двух различных гипотез по одним и тем же данным. Сперва проверяют гипотезу о равенстве дисперсий, а затем гипотезу о равенстве средних.

5. Сравнение нескольких средних в случае независимых нормально распределенных признаков. Для сравнения нескольких средних в случае независимых нормально распределенных признаков используется специальная статистическая процедура, которая называется дисперсионным анализом. Однако можно сделать вывод и на основании критерия Стьюдента, проверив гипотезу о равенстве наибольшего и наименьшего средних.

Опишем процедуру ***однофакторного дисперсионного анализа.***

Пусть имеется N независимых выборок объемов $n_1, n_2, \ldots, n_N$ соответственно и задан уровень значимости α. Обозначим через $\overline{x}_i, s_i^2$ несмещенные оценки математического ожидания и дисперсии, полученные по i-й выборке, $f_i = n_i - 1$.

Нулевая гипотеза $H_0 : a_1 = a_2 = \ldots = a_N$.

Альтернативная гипотеза $\overline{H}$: не все эти математические ожидания равны между собой.

Условием применимости метода дисперсионного анализа является, помимо нормальности выборок, однородность дисперсий. Следовательно, как и в случае двух выборок, процедуре сравнения средних должно предшествовать сравнение дисперсий.

Идея однофакторного дисперсионного анализа заключается в разбиении общей дисперсии, которая получается при объединении всех наблюдений в одну выборку, на два независимых слагаемых – *факторную (межгрупповую)* дисперсию $s^2_{\text{факт}}$, порождаемую различием между группами (выборками), и *остаточную (внутригрупповую)* дисперсию $s^2_{\text{ост}}$, обусловленную случайными помехами и неучтенными факторами: $s^2_{\text{общ}} = s^2_{\text{факт}} + s^2_{\text{ост}}$. Дисперсионный анализ был первоначально предложен Р. Фишером и определен им как метод «отделения дисперсии, приписываемой одной группе причин, от дисперсии, приписываемой другим группам».

Межгрупповая дисперсия рассчитывается по формуле

$$s^2_{\text{факт}} = \frac{1}{N-1} \sum_{i=1}^{N} (\overline{x}_i - \overline{\overline{x}})^2 n_i,$$

где $\overline{\overline{x}}$ – выборочное среднее, рассчитанное по объединенной выборке; число степеней межгрупповой дисперсии равно $f_{\text{факт}} = N - 1$. Остаточная дисперсия представляет собой взвешенное среднее оценок дисперсий и рассчитывается по формуле

$$s^2_{\text{ост}} = \frac{f_1 s_1^2 + f_2 s_2^2 + \ldots + f_N s_N^2}{f_1 + f_2 + \ldots + f_N}, \qquad f_{\text{ост}} = f_1 + f_2 + \ldots + f_N.$$

Гипотеза H_0 *при заданном уровне значимости* α *принимается (не противоречит экспериментальным данным), если*

$$F_{\text{расч}} = \frac{s^2_{\text{факт}}}{s^2_{\text{ост}}} < F_{\text{табл}} = F_{\alpha;\, f_{\text{факт}};\, f_{\text{ост}}},$$

где $F_{\alpha;\, f_{\text{факт}};\, f_{\text{ост}}}$ определяется по таблице квантилей распределения Фишера.

6. Сравнение двух средних в случае зависимых нормально распределенных признаков. Такая задача возникает, если две выборки взаимосвязаны. Например, проводятся измерения одних и тех же величин на одних и тех же объектах двумя разными методами и требуется определить, одинаковы ли результаты использования двух методов измерения. Либо если проводятся измерения какой-то характеристики для одних и тех же объектов до и после некоторого воздействия и требуется определить, влияет ли это воздействие на значение характеристики.

В этом случае имеются две выборки одинакового объема n :

$$x_{11}, \quad x_{12}, \quad \ldots, \quad x_{1n};$$

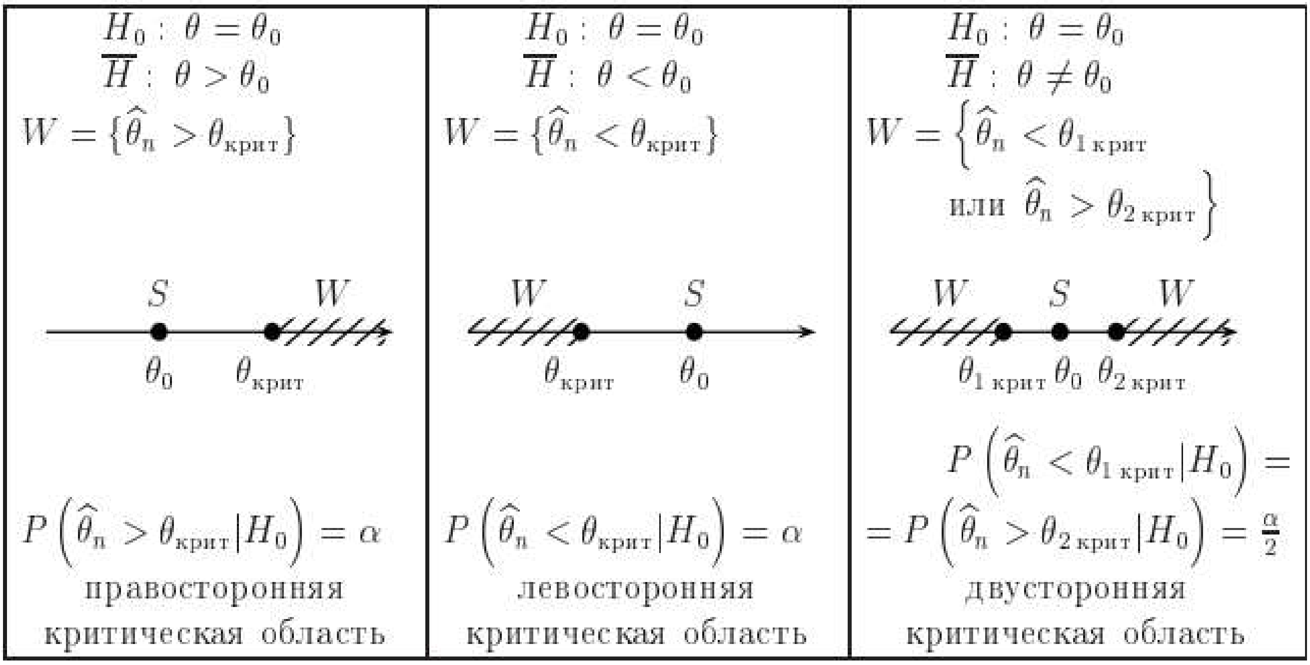
$$x_{21}, \quad x_{22}, \quad \ldots, \quad x_{2n}.$$

Поскольку значения в каждой паре x_{1i}, x_{2i} связаны (например, измерены на одном и том же объекте), то получим новую выборку с элементами $\Delta x_i = x_{1i} - x_{2i}$.

Задача сводится к проверке гипотезы о равенстве нулю среднего значения новой выборки, т. е. $H_0 : a_{\Delta x} = 0$. Эта проверка проводится по критерию (2).

Иногда возникает необходимость сравнения гипотезы $H_0 : \theta = \theta_0$ с ***односторонней*** альтернативой $\bar{H}_1 : \theta > \theta_0$ или $\bar{H}_2 : \theta < \theta_0$. Например, если известно, что неравенство $\theta < \theta_0$ невозможно, то в качестве альтернативной рассматривается гипотеза $\bar{H} : \theta > \theta_0$.

Вид критической области W и области S принятия гипотезы зависит от вида альтернативной гипотезы.



Таким образом, в зависимости от вида альтернативной гипотезы $\bar{H}$ выбирают *правостороннюю*, *левостороннюю* или *двустороннюю* критическую область.

Например, при проверке гипотезы $H_0 : a = a_0$ против альтернативы $\bar{H}_1 : a > a_0$ требуется выяснить, соответствует ли выборочное среднее значение норме или превосходит ее. Пусть дисперсия σ^2 известна. Оценкой для параметра a является $\bar{x}$. Ясно, что если $\bar{x} < a_0$, то гипотезу H_0 следует предпочесть альтернативе $\bar{H}_1$. Если же $\bar{x} > a_0$, то *гипотеза* H_0 *принимается на уровне значимости* α, *если выполняется условие* (1) *с* $u_\text{табл} = u_{2\alpha}$, т. е. табличное значение определяется для удвоенного уровня значимости.

Аналогично с удвоенным уровнем значимости определяются табличные значения при использовании критериев (2), (5), (6), (7) в случае односторонних альтернатив.

Критерий (3) используется следующим образом. В случае альтернативы $\bar{H}_1 : \sigma^2 > \sigma_0^2$ *гипотеза* $H_0 : \sigma^2 = \sigma_0^2$ *при заданном уровне значимости* α *принимается, если*

$$\chi^2_\text{расч} = \frac{(n-1)s^2}{\sigma_0^2} < \chi^2_{\alpha;\, n-1};$$

в случае альтернативы $\overline{H}_2 : \sigma^2 < \sigma_0^2$ гипотеза H_0 принимается, если

$$\chi^2_{\text{расч}} = \frac{(n-1)s^2}{\sigma_0^2} > \chi^2_{1-\alpha;\, n-1}.$$

Контрольные вопросы

1. В чем заключается основная задача при проверке статистической гипотезы?
2. Что называется статистической гипотезой?
3. В каком случае статистическая гипотеза называется простой? сложной?
4. В чем разница между нулевой и альтернативной гипотезами?
5. В каком случае статистическая гипотеза называется параметрической? непараметрической?
6. Что называется критерием значимости? Что называется критерием согласия?
7. Что называется уровнем значимости статистического критерия?
8. Как видоизменяется критерий проверки гипотезы в случае односторонней альтернативы?
9. Что характеризует выборочное среднее? Что характеризует выборочная дисперсия?
10. Как рассчитать выборочное среднее?
11. Как рассчитать несмещенную оценку дисперсии?
12. Какие критерии используются для проверки гипотез о математических ожиданиях одной и двух независимых нормальных выборок?
13. Какие критерии используются для проверки гипотез о дисперсиях одной и двух независимых нормальных выборок?
14. Что такое однородность дисперсий и как она проверяется?
15. Для проверки каких гипотез используется критерий Фишера?
16. Как используется критерий Фишера для проверки однородности нескольких дисперсий?
17. Для проверки каких гипотез используется критерий χ^2?
18. Чем отличается процедура проверки гипотезы о равенстве математического ожидания заданному значению для случаев известной и неизвестной дисперсии?
19. Для проверки каких гипотез используется критерий Стьюдента?
20. Чем отличается процедура проверки гипотезы о равенстве средних двух зависимых и независимых нормальных выборок?
21. В чем заключается процедура проверки гипотезы о равенстве средних в случае парных (зависимых) выборок?
22. Как учитывается предположение о равенстве дисперсий при сравнении средних?

Пример и методические указания
по выполнению лабораторной работы в Excel

Решить задачи, используя критерии значимости, предназначенные для проверки гипотез о значениях параметров нормального распределения. Уровень значимости принять $\alpha = 0{,}05$.

Во всех задачах считать, что исследуемые признаки имеют нормальное распределение.

Задача 1. Проводилось исследование предела прочности на разрыв различных по химической структуре твердых смол. При этом требовалось выяснить, вносят ли действия оператора какое-нибудь смещение в результаты наблюдений. Было взято по два образца каждой смолы. Двум операторам А и В предложили испытать по одному образцу смол каждого типа. Можно ли говорить, что наблюдаются различия между результатами наблюдений двух операторов?

Смола	127	135	138	139	146	152
Оператор А	5250	4975	5050	5075	4795	5190
Оператор В	5230	4980	5020	5085	4750	5120

Решение. Исследуемый признак – предел прочности на разрыв смол. Нужно определить, одинаковы ли значения результатов наблюдений у двух операторов. Поскольку два оператора испытывали одни и те же образцы смол, имеем задачу сравнения средних в случае зависимых выборок.

Проверяем гипотезу H_0 о том, что в среднем результаты наблюдений одинаковы (критерии подразумевают, что нулевая гипотеза всегда выдвигается о равенстве параметров):

$H_0 : a_{\Delta x} = 0$

при альтернативе $\overline{H}$, согласно которой есть различия в результатах наблюдений:

$\overline{H} : a_{\Delta x} \neq 0$.

По формуле (2) (дисперсия σ^2 неизвестна и будет оцениваться по выборке), гипотеза H_0 при альтернативе $\overline{H}$ на уровне значимости α принимается (не противоречит результатам наблюдений, нет оснований отвергнуть гипотезу), если

$$t_{\text{расч}} = \frac{|\overline{\Delta x}|}{\sqrt{s^2 / n}} < t_{\text{табл}} = t_{\alpha;\, n-1}.$$

Рассчитаем выборку значений $\Delta x_i = x_{1i} - x_{2i}$ разностей результатов наблюдений двух операторов.

Смола	127	135	138	139	146	152
$\Delta x_i = x_{1i} - x_{2i}$	20	-5	30	-10	45	70

Объем выборки $n = 6$. Рассчитаем несмещенные оценки среднего и дисперсии:

$$\overline{\Delta x} = \frac{1}{6} \cdot (20 - 5 + 30 - 10 + 45 + 70) = 25;$$

$$D_{\text{в}} = \frac{1}{6} \cdot ((20-25)^2 + (-5-25)^2 + (30-25)^2 + (-10-25)^2 + (45-25)^2 + (70-25)^2) =$$

$$= \frac{1}{6} \cdot (25 + 900 + 25 + 1225 + 400 + 2025) = \frac{1}{6} \cdot 4600 = \frac{2300}{3};$$

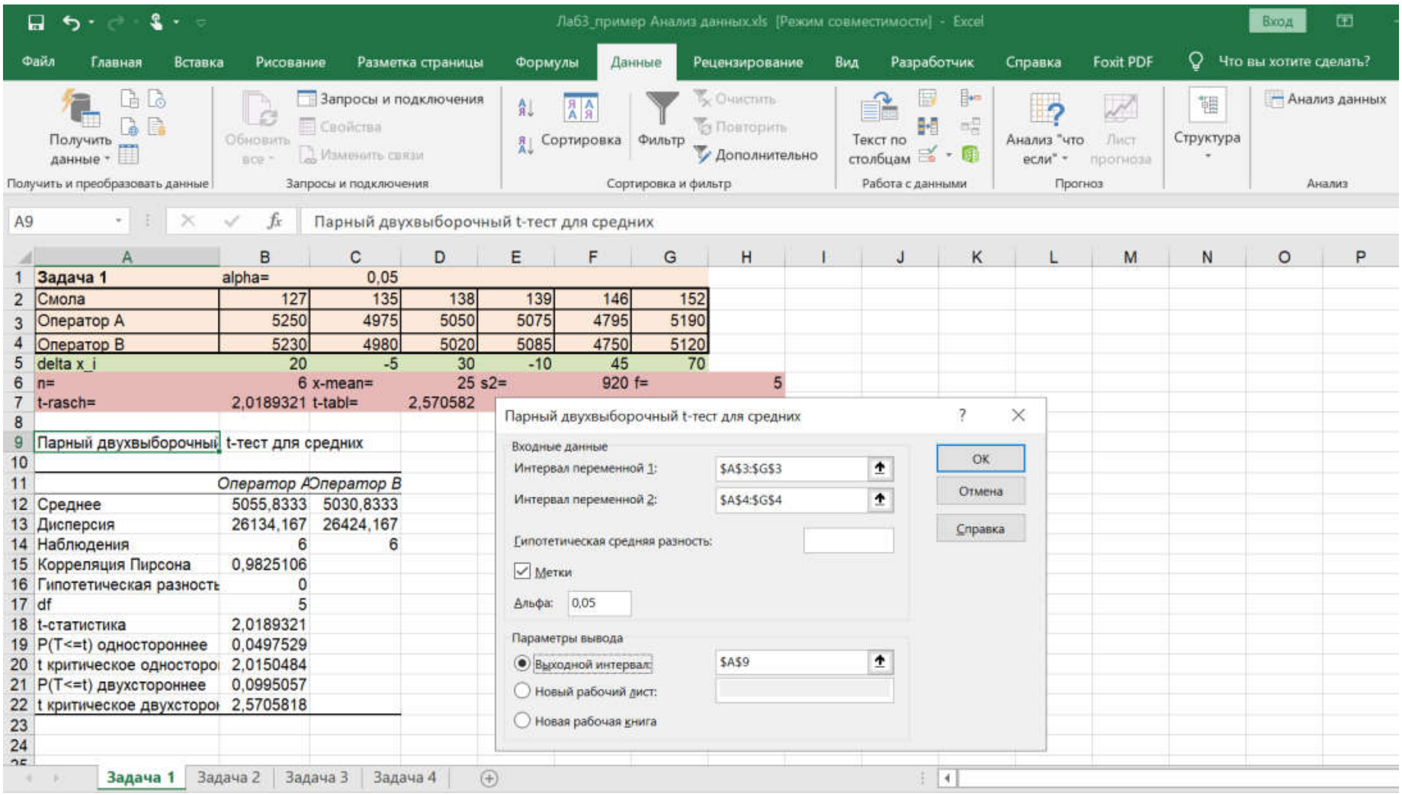
$$s^2 = \frac{6}{5} \cdot \frac{4600}{6} = 920.$$

Поскольку $t_{\text{расч}} = \dfrac{25}{\sqrt{920/6}} \approx 2{,}02 < t_{\text{табл}} = t_{0{,}05;\,5} = 2{,}57$, то делаем вывод: на уровне значимости 0,05 можно утверждать, что гипотеза H_0 не противоречит экспериментальным данным, т. е. различия между результатами измерений у двух операторов следует признать незначительными.

Ниже приведен фрагмент рабочего листа и даны рекомендации по выполнению расчетов для решения задачи в Excel с помощью встроенных статистических функций и с помощью Пакета анализа.

Пакет анализа – набор средств анализа данных в Microsoft Excel, предназначенный для решения сложных статистических и инженерных задач. Средства, которые включены в Пакет анализа данных, доступны через команду меню **Данные→Анализ данных**.

Если этой команды нет в меню, необходимо загрузить надстройку **Пакет анализа**: 1) открыть вкладку **Файл→Параметры→Надстройки**; 2) в раскрывающемся списке **Управление** выбрать пункт **Надстройки Excel** и нажать кнопку **Перейти**; 3) в диалоговом окне **Надстройки** установить флажок **Пакет анализа** и нажать кнопку ОК.



Методические указания по использованию EXCEL

E X C E L

Использование статистических функций

Функция СЧЁТ() подсчитывает количество ячеек, содержащих числа;
функция СРЗНАЧ() вычисляет выборочное среднее;
функция ДИСП.В() (в Excel2007 и более ранних версиях - ДИСП()) рассчитывает несмещенную оценку дисперсии;
функция СТЬЮДЕНТ.ОБР.2Х() (в Excel2007 и более ранних версиях СТЬЮДРАСПОБР()) возвращает квантиль распределения Стьюдента: *например*, в ячейке D7 записана формула =СТЬЮДЕНТ.ОБР.2Х(C1;H6)

Использование Пакета анализа

Данные→Анализ данных→Парный двухвыборочный t-тест для средних

Сравните результаты:

t-статистика = $t_{расч}$;

t критическое двухстороннее = $t_{табл}$ (для случая односторонней альтернативы имеется t критическое одностороннее);

P(T<=t) двухстороннее – это пороговое значение уровня значимости, показывающее, что при $\alpha > 0{,}0995$ гипотеза H_0 при альтернативе $\overline{H}$ отвергается.

Задача 2. Для того чтобы определить, сокращается ли время сварки на отливках, если при литье вместо сырой формовочной смеси использовать сухую смесь, было проведено специальное исследование. Стоимость литья в случае сухой формовочной смеси выше, но есть мнение, что это может быть оправдано, если время сварки значимо уменьшится. В таблице приведены значения времени сварки в минутах. Можно ли сказать, что имеет место значимое уменьшение времени сварки при использовании сухой формовочной смеси?

	1	2	3	4	5
Сырая смесь	19	28	14	29	15
Сухая смесь	21	15	11	12	21

Решение. В таблице приведены значения времени сварки и требуется проверить, имеет ли место значимое уменьшение времени сварки при использовании второй технологии или время сварки в среднем одинаковое для двух технологий.

Имеем задачу сравнения средних в случае независимых выборок. Проверяем при уровне значимости $\alpha = 0{,}05$ нулевую гипотезу H_0 о том, что в среднем время сварки одинаковое для двух технологий (критерии подразумевают, что нулевая гипотеза всегда выдвигается о равенстве параметров):

$H_0:\ a_1 = a_2$

при альтернативе $\overline{H}$, согласно которой для второй технологии время в среднем сокращается (односторонняя альтернатива):

$\overline{H}: a_1 > a_2$.

При сравнении двух средних нужно учитывать, однородны ли дисперсии двух выборок, поэтому предварительно проверяется вспомогательная гипотеза о равенстве дисперсий:

$H_0^{/}: \sigma_1^2 = \sigma_2^2;$

$\overline{H}^{/}: \sigma_1^2 \neq \sigma_2^2.$

Проверка проводится по критерию Фишера. Гипотеза $H_0^{/}$ при альтернативе $\overline{H}^{/}$ на уровне значимости α принимается, если

$$F_{\text{расч}} = \frac{s_{\max}^2}{s_{\min}^2} < F_{\text{табл}} = F_{\alpha/2; f_1; f_2},$$

где f_1 и f_2 – числа степеней свободы большей и меньшей оценок дисперсий соответственно.

Рассчитаем статистические характеристики выборок.

	n_i	$\overline{x_i}$	s_i^2	f_i
Сырая смесь	5	21	50,5	4
Сухая смесь	5	16	23	4

Объемы выборок $n_1 = n_2 = 5$. Число степеней свободы каждой оценки дисперсии равно числу наблюдений, по которым она рассчитана, минус 1:

$$f_1 = n_1 - 1 = 5 - 1 = 4; \qquad f_2 = n_2 - 1 = 5 - 1 = 4.$$

Расчетное значение критерия Фишера (нужно разделить большую оценку дисперсии на меньшую) равно

$$F_{\text{расч}} = \frac{50,5}{23} \approx 2,2;$$

табличное значение $F_{\text{табл}} = F_{0,025;\,4;\,4} = 9,6$. Поскольку $F_{\text{расч}} = 2,2 < F_{\text{табл}} = 9,6$, то делаем вывод: на уровне значимости 0,05 можно считать дисперсии однородными.

Если дисперсии однородны, то гипотеза H_0 при *односторонней* альтернативе $\overline{H}$ на уровне значимости α принимается, если $\bar{x}_1 < \bar{x}_2$ (среднее время для первой технологии оказалось меньше среднего времени для второй) либо

$$t_{\text{расч}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}} < t_{\text{табл}} = t_{2\alpha;\,f},$$

где $s^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$, $f = n_1 + n_2 - 2$.

Вычисляем общую средневзвешенную оценку дисперсии и ее число степеней свободы:

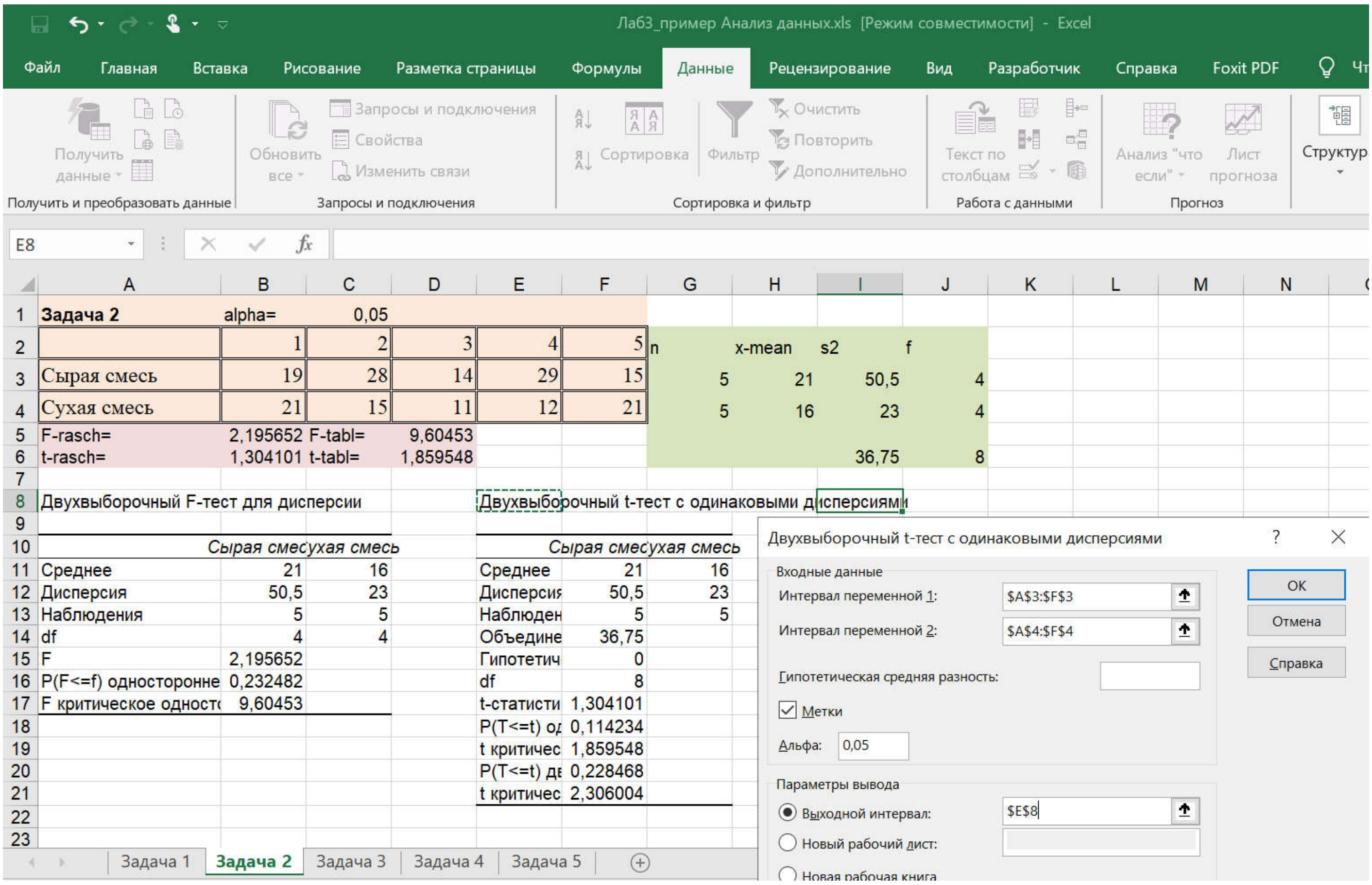
$$s^2 = \frac{(5-1)\cdot 50,5 + (5-1)\cdot 23}{5+5-2} = 36,75; \qquad f = 5 + 5 - 2 = 8;$$

сравниваем расчетное и критическое значения критерия Стьюдента:

$$t_{\text{расч}} = \frac{21 - 16}{\sqrt{36,75\cdot\left(\dfrac{1}{5} + \dfrac{1}{5}\right)}} \approx 1,3 < t_{\text{табл}} = t_{0,1;\,8} = 1,86,$$

делаем вывод: на уровне значимости 0,05 можно утверждать, что гипотеза H_0 не противоречит экспериментальным данным, т. е. нет оснований утверждать, что при использовании сухой смеси время литья значимо уменьшится.

Ниже приведен фрагмент рабочего листа и даны рекомендации по выполнению расчетов для решения задачи в Excel с помощью встроенных статистических функций и с помощью Пакета анализа.



Использование статистических функций

Функция F.ОБР.ПХ() (в более ранних версиях FРАСПОБР()) вычисляет квантиль распределения Фишера. *Например*, если в ячейке C1 записан заданный уровень значимости α, а в ячейках J3 и J4 стоят числа степеней свободы соответственно большей и меньшей оценок дисперсий, то $F_{\text{табл}}$ определяется по формуле =F.ОБР.ПХ(C1/2;J3;J4) (соответственно, =FРАСПОБР(C1/2;J3;J4)).

Использование Пакета анализа

Данные→Анализ данных→Двухвыборочный F-тест для дисперсии

Сравните результаты: средние, дисперсии, объемы выборок (Наблюдения), числа степеней свободы дисперсий;

F = $F_{\text{расч}}$;

F критическое одностороннее = $F_{\text{табл}}$;

P(F<=f) одностороннее – это пороговое значение уровня значимости, показывающее, что при $\alpha > 0{,}23$ нулевая гипотеза об однородности дисперсий отвергается.

Данные→Анализ данных→Двухвыборочный t-тест с одинаковыми дисперсиями.

Сравните результаты: средние, дисперсии, объемы выборок (Наблюдения);

Объединенная дисперсия= s^2;

df = f ;

t-статистика = $t_{\text{расч}}$;

t критическое одностороннее = $t_{\text{табл}}$ (для случая двухсторонней альтернативы имеется t критическое двухстороннее);

P(T<=t) одностороннее – это пороговое значение уровня значимости, показывающее, что при $\alpha > 0{,}114$ гипотеза H_0 при альтернативе $\overline{H}$ отвергается.

Задача 3. При производстве синтетического волокна для уменьшения последующей усадки продукция, движущаяся непрерывным потоком, подвергается термической обработке. Даны результаты измерений величины усадки в процентах для волокон после обработки при двух температурах: $120^\circ C$ и $140^\circ C$. До начала эксперимента предполагалось, что дисперсии усадки при рассмотренных температурах не равны между собой. Требуется проверить, будет ли усадка при $140^\circ C$ *больше*, чем при $120^\circ C$.

$120^\circ C$	3,45	3,62	3,6	3,49	3,64	3,56	3,52	3,53	3,57	3,44	3,56	3,43
$140^\circ C$	3,72	4,01	3,54	3,67	4,03	3,4	3,96	3,6	3,76	3,91		

Решение. В таблице приведены значения величины усадки и требуется проверить, будут ли значения усадки во второй выборке в среднем больше, чем в первой, или усадка в среднем одинакова в обеих выборках.

Имеем задачу сравнения средних в случае независимых выборок. Проверяем при уровне значимости $\alpha = 0{,}05$ нулевую гипотезу H_0 о том, что в среднем усадка одинакова в обеих выборках (критерии подразумевают, что нулевая гипотеза всегда выдвигается о равенстве параметров):

$H_0:\ a_1 = a_2$

при альтернативе $\overline{H}$, согласно которой усадка при $140^\circ C$ больше, чем при $120^\circ C$ (односторонняя альтернатива):

$\overline{H} : a_1 < a_2.$

При сравнении двух средних нужно учитывать, однородны ли дисперсии двух выборок. Предварительно проверяется вспомогательная гипотеза о равенстве дисперсий:

$H_0^{/} : \sigma_1^2 = \sigma_2^2;$

$\overline{H}^{/} : \sigma_1^2 \neq \sigma_2^2.$

Проверка проводится по критерию Фишера. Гипотеза $H_0^{/}$ при альтернативе $\overline{H}^{/}$ на уровне значимости α принимается, если

$$F_{\text{расч}} = \frac{s_{\max}^2}{s_{\min}^2} < F_{\text{табл}} = F_{\alpha/2;\,f_1;\,f_2},$$

где f_1 и f_2 – числа степеней свободы большей и меньшей оценок дисперсий соответственно.

Рассчитаем статистические характеристики выборок.

	n_i	$\overline{x_i}$	s_i^2	f_i
$120^\circ C$	12	3,53	0,005	11

$140^\circ C$	10	3,76	0,046	9

Объемы выборок $n_1 = 12$; $n_2 = 10$. Число степеней свободы каждой оценки дисперсии равно числу наблюдений, по которым она рассчитана, минус 1:

$$f_1 = n_1 - 1 = 12 - 1 = 11; \quad f_2 = n_2 - 1 = 10 - 1 = 9.$$

Расчетное значение критерия Фишера (нужно разделить большую оценку дисперсии на меньшую) равно

$$F_{\text{расч}} = \frac{0,046}{0,005} \approx 9,2\,;$$

табличное значение $F_{\text{табл}} = F_{0,025;\,9;\,11} = 3,59$ (уровень значимости делится на 2 в соответствии с формулой; числа степеней свободы берутся в порядке, соответствующем порядку оценок дисперсий – сначала число степеней большей оценки дисперсии, затем меньшей). Поскольку $F_{\text{расч}} = 9,2 > F_{\text{табл}} = 3,59$, то делаем вывод: на уровне значимости 0,05 дисперсии следует признать неоднородными.

Если дисперсии неоднородны, то гипотеза H_0 при *односторонней* альтернативе $\overline{H}$ на уровне значимости α принимается, если $\bar{x}_1 > \bar{x}_2$ (среднее первой выборки оказалось больше второго среднего) либо

$$t_{\text{расч}} = \frac{\bar{x}_2 - \bar{x}_1}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} < t_{\text{табл}} = t_{2\alpha;\,f}\,,$$

где

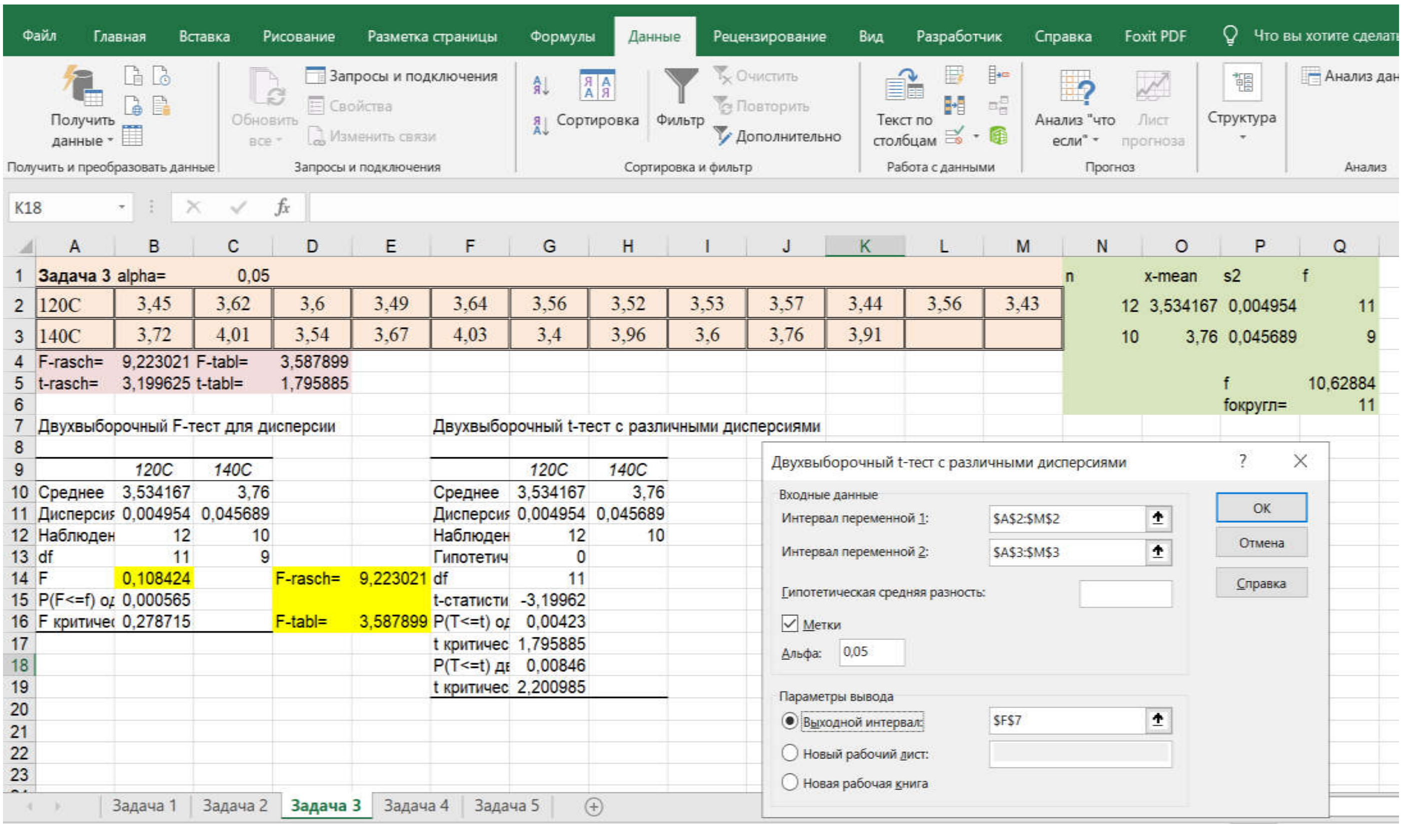
$$f \approx \frac{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^2}{\dfrac{\left(\dfrac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \dfrac{\left(\dfrac{s_2^2}{n_2}\right)^2}{n_2 - 1}} = \frac{\left(\dfrac{0,005}{12} + \dfrac{0,046}{10}\right)^2}{\dfrac{\left(\dfrac{0,005}{12}\right)^2}{11} + \dfrac{\left(\dfrac{0,046}{10}\right)^2}{9}} \approx 11.$$

Сравнивая расчетное и критическое значения критерия Стьюдента:

$$t_{\text{расч}} = \frac{3,76 - 3,53}{\sqrt{\left(\dfrac{0,005}{12} + \dfrac{0,046}{10}\right)}} \approx 3,2 > t_{\text{табл}} = t_{0,1;\,11} = 1,8,$$

делаем вывод: на уровне значимости 0,05 гипотеза H_0 отвергается, т. е. усадка при $140^\circ C$ больше, чем при $120^\circ C$.

Ниже приведен фрагмент рабочего листа Excel и отмечены нюансы использования инструмента **Двухвыборочный F-тест для дисперсии** из Пакета анализа.



Методические указания по использованию EXCEL

Использование Пакета анализа

Данные→Анализ данных→Двухвыборочный F-тест для дисперсии

Сравним результаты: средние, дисперсии, объемы выборок (Наблюдения), числа степеней свободы дисперсий рассчитаны корректно.

Получили значение F<1, поскольку в Пакете анализа для расчета $F_{расч}$ запрограммировано деление первой дисперсии на вторую. Рассчитаем в столбце E величины, обратные к F и F критическое одностороннее, чтобы получить $F_{расч}$ и $F_{табл}$ соответственно: E14=1/B14; E16=1/B16.

Данные→Анализ данных→Двухвыборочный t-тест с различными дисперсиями

Сравним результаты: средние, дисперсии, объемы выборок (Наблюдения);

df = f (округленное значение);

t-статистика = $-t_{расч}$ (в Пакете анализа t-статистика рассчитывается без модуля);

t критическое одностороннее = $t_{табл}$ (для случая двухсторонней альтернативы имеется t критическое двухстороннее);

P(T<=t) одностороннее – это пороговое значение уровня значимости, показывающее, что при $\alpha > 0{,}0423$ гипотеза H_0 при альтернативе $\overline{H}$ отвергается.

Задача 4. Для проверки точности двух станков проведены измерения некоторого признака выпускаемых ими однотипных деталей. Можно ли на основании этих данных при уровне значимости $\alpha = 0{,}05$ сделать вывод о том, что точность первого станка выше точности второго?

Станок	$\overline{x_i}$	s_i^2	n_i
1-й	120,3	12,25	25
2-й	118,9	28,4	18

Решение. Точность станка характеризуется разбросом значений признака выпускаемых изделий, произведенных станком: чем меньше разброс значений (т. е. дисперсия значений измеренного признака выпускаемых изделий), тем выше точность станка.

Имеем задачу сравнения двух дисперсий: проверяем при уровне значимости $\alpha = 0{,}05$ нулевую гипотезу H_0 о том, что разброс значений измеренного признака выпускаемых изделий одинаков для двух станков (критерии подразумевают, что нулевая гипотеза всегда выдвигается о равенстве параметров):

$H_0: \sigma_1^2 = \sigma_2^2$

при альтернативе $\overline{H}$, согласно которой для первого станка разброс значений меньше, чем для второго (одностороння альтернатива):

$\overline{H}: \sigma_1^2 < \sigma_2^2.$

Проверка проводится по критерию Фишера. Гипотеза H_0 при *односторонней* альтернативе $\overline{H}$ на уровне значимости α принимается, если $s_1^2 > s_2^2$ (оценка дисперсии для первого станка оказалась меньше второй оценки дисперсии) либо

$$F_{\text{расч}} = \frac{s_{\max}^2}{s_{\min}^2} < F_{\text{табл}} = F_{2\cdot\alpha/2;f_1;f_2},$$

где f_1 и f_2 – числа степеней свободы большей и меньшей оценок дисперсий соответственно.

Объемы выборок $n_1 = 25$; $n_2 = 18$. Число степеней свободы каждой оценки дисперсии равно числу наблюдений, по которым она рассчитана, минус 1:

$$f_1 = n_1 - 1 = 25 - 1 = 24; \quad f_2 = n_2 - 1 = 18 - 1 = 17.$$

Расчетное значение критерия Фишера (нужно разделить большую оценку дисперсии на меньшую) равно

$$F_{\text{расч}} = \frac{28{,}4}{12{,}25} \approx 2{,}32;$$

табличное значение $F_{\text{табл}} = F_{0{,}05;\,17;\,24} = 2{,}07$ (уровень значимости делится на 2 в соответствии с формулой критерия Фишера, но удваивается в случае односторонней альтернативы; числа степеней свободы берутся в порядке, соответствующем порядку оценок дисперсий – сначала число степеней большей оценки дисперсии, затем меньшей).

Поскольку $F_{\text{расч}} = 2{,}32 > F_{\text{табл}} = 2{,}07$, то на уровне значимости 0,05 гипотеза H_0 о равенстве дисперсий должна быть отвергнута. Таким образом, по результатам экспериментальных данных на уровне значимости 0,05 точность первого станка признаем выше точности второго.

Задача 5. Измерялось сопротивление проволок трех типов. Утверждается, что между проволоками разных типов в среднем нет различий. Можно ли принять гипотезу об одинаковом среднем значении сопротивления для проволок трех типов?

A	121	120	124	121	120	124	126	120
B	120	119	126	128	126	124	122	127
C	129	132	137	139	130	132	137	136

Решение.

1 способ. Имеем задачу сравнения нескольких средних в случае независимых выборок. Решим задачу методом **однофакторного дисперсионного анализа.**

Проверяется гипотеза о равенстве средних трех выборок:

$H_0: a_1 = a_2 = a_3$

при альтернативе $\overline{H}:$ не все средние равны между собой.

Процедуре сравнения средних предшествует проверка однородности дисперсий:

$H_0^{/}: \sigma_1^2 = \sigma_2^2 = \sigma_3^2;$

$\overline{H}^{/}:$ не все дисперсии равны между собой.

Проверку однородности дисперсий можно провести по критерию Фишера, сравнив наибольшую и наименьшую из оценок дисперсий.

Рассчитаем статистические характеристики выборок.

	n_i	$\overline{x}_i$	s_i^2	f_i
A	8	122	5,43	7
B	8	124	11,14	7
C	8	134	13,71	7

Сравнивая расчетное значение критерия Фишера (нужно разделить большую оценку дисперсии на меньшую) с табличным

$$F_{\text{расч}} = \frac{13,71}{5,43} \approx 2,53 < F_{\text{табл}} = F_{0,025;\,7;\,7} = 4,99,$$

делаем вывод, что на уровне значимости 0,05 дисперсии можно считать однородными. Это позволяет использовать дисперсионный анализ для проверки однородности средних.

Имеем $N = 3$ выборки объемами $n_1 = n_2 = n_3 = 8$. Объединяя все выборки в одну, вычисляем среднее объединенной выборки

$$\overline{\overline{x}} = \frac{1}{24} \cdot (121 + 120 + 124 + ... + 137 + 136) \approx 126,67.$$

Рассчитаем межгрупповую дисперсию по формуле

$$s_{\text{факт}}^2 = \frac{1}{N-1} \sum_{i=1}^{N} (\overline{x}_i - \overline{\overline{x}})^2 n_i =$$

$$= \frac{1}{3} \cdot ((122 - 126,67)^2 \cdot 8 + (124 - 126,67)^2 \cdot 8 + (134 - 126,67)^2 \cdot 8) = 330,67,$$

ее число степеней свободы равно $f_{\text{факт}} = N - 1 = 2$. Остаточная дисперсия представляет собой взвешенное среднее оценок дисперсий и рассчитывается по формуле

$$s_{\text{ост}}^2 = \frac{f_1 s_1^2 + f_2 s_2^2 + f_3 s_3^2}{f_1 + f_2 + f_3} = \frac{7 \cdot 5,43 + 7 \cdot 11,14 + 7 \cdot 13,71}{7 + 7 + 7} = 10,095,$$

$$f_{\text{ост}} = f_1 + f_2 + f_3 = 21.$$

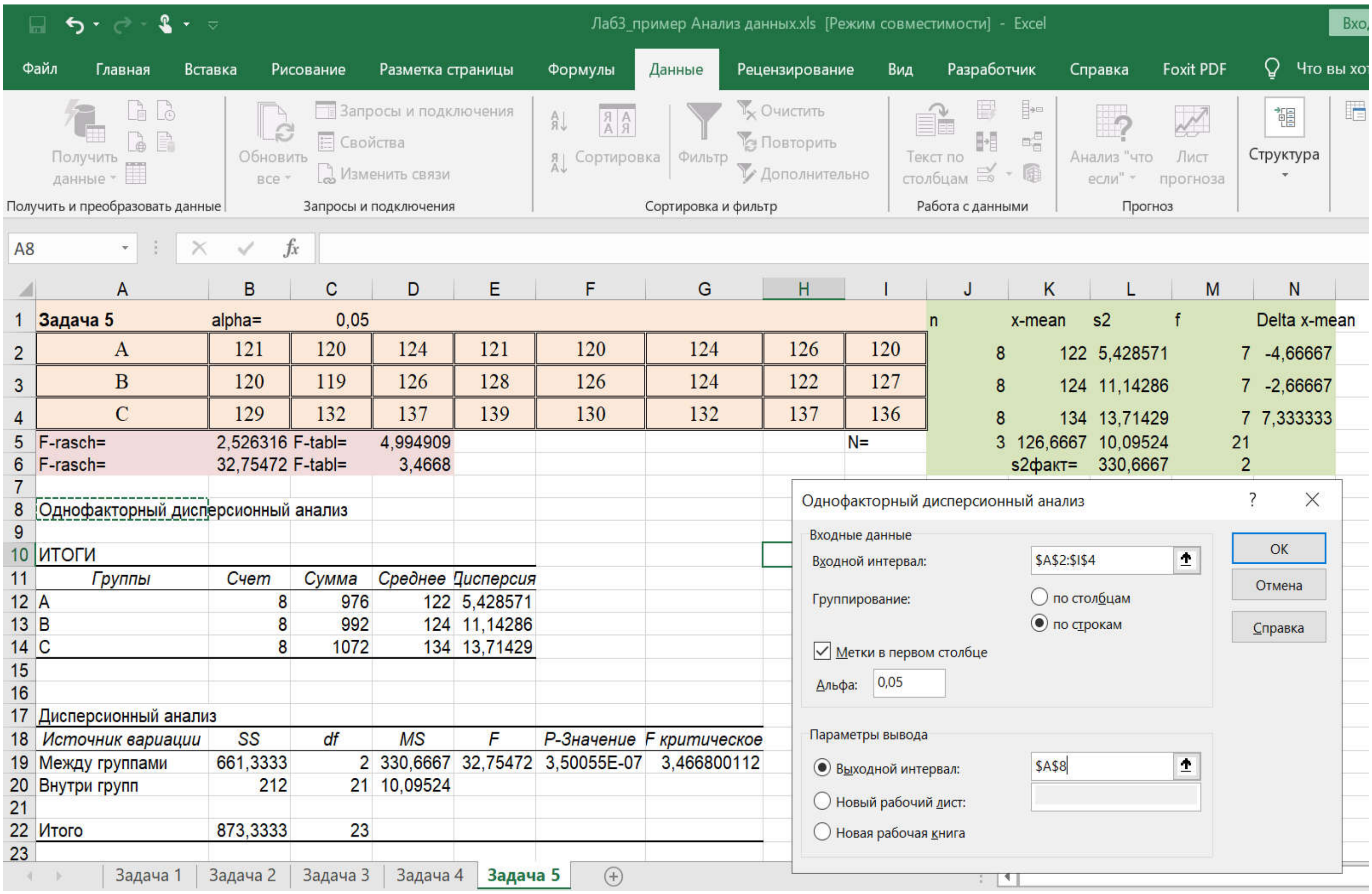
Гипотеза H_0 при заданном уровне значимости α принимается (не противоречит экспериментальным данным), если

$$F_{\text{расч}} = \frac{s_{\text{факт}}^2}{s_{\text{ост}}^2} < F_{\text{табл}} = F_{\alpha;\, f_{\text{факт}};\, f_{\text{ост}}},$$

где $F_{\alpha;\, f_{\text{факт}};\, f_{\text{ост}}}$ определяется по таблице квантилей распределения Фишера.

Поскольку $F_{\text{расч}} = \dfrac{330,67}{10,095} \approx 32,75 > F_{\text{табл}} = F_{0,05;\, 2;\, 21} = 3,47$, то делаем вывод: на уровне значимости 0,05 гипотеза о равенстве трех средних отклоняется.

На фрагменте рабочего листа Excel представлены расчеты по решению задачи методом однофакторного дисперсионного анализа с помощью встроенных статистических функций и с помощью Пакета анализа.



Методические указания по использованию EXCEL

E *Использование Пакета анализа*

Данные→Анализ данных→Однофакторный дисперсионный анализ

Сравните результаты.

Предварительные итоги по выборкам: объем выборки (Счет), сумма элементов в выборке (Сумма), среднее, дисперсия.

Результаты дисперсионного анализа:

df = f (числа степеней свободы факторной и остаточной дисперсий);

MS – факторная и остаточная дисперсии;

SS – суммы (MS=SS/df);

F = $F_{\text{расч}}$;

F критическое = $F_{\text{табл}}$;

P-Значение – это пороговое значение уровня значимости, показывающее, что при $\alpha > 3{,}5 \cdot 10^{-7}$ гипотеза H_0 при альтернативе $\overline{H}$ отвергается.

<u>*2 способ.*</u> Имеем задачу сравнения нескольких средних в случае <u>независимых</u> выборок: проверяется гипотеза о равенстве средних трех выборок:

$H_0: \ a_1 = a_2 = a_3$

при альтернативе $\overline{H}$: не все средние равны между собой.

Решим задачу с помощью критерия Стьюдента, проверив гипотезу о равенстве наибольшего и наименьшего средних.

Рассчитаем статистические характеристики выборок.

	n_i	$\overline{x_i}$	s_i^2	f_i
A	8	122	5,43	7
B	8	124	11,14	7
C	8	134	13,71	7

Итак, проверим при уровне значимости $\alpha = 0{,}05$ нулевую гипотезу $H_0^{/}$ о равенстве наибольшего и наименьшего средних (средних выборок А и С):

$H_0^{/}: \ a_1 = a_3$

$\overline{H}^{/}: a_1 \neq a_3.$

При сравнении двух средних нужно учитывать, однородны ли дисперсии двух выборок, поэтому предварительно проверяется вспомогательная гипотеза о равенстве дисперсий:

$H_0^{//}: \sigma_1^2 = \sigma_3^2;$

$\overline{H}^{//}: \sigma_1^2 \neq \sigma_3^2.$

Проверка проводится по критерию Фишера: так как

$$F_{\text{расч}} = \frac{13{,}71}{5{,}43} \approx 2{,}53 < \ F_{\text{табл}} = F_{0{,}025;\,7;\,7} = 4{,}99,$$

то на уровне значимости 0,05 дисперсии признаем однородными.

Если дисперсии однородны, то гипотеза $H_0^{/}$ при альтернативе $\overline{H}^{/}$ на уровне значимости α принимается, если

$$t_{\text{расч}} = \frac{|\,\overline{x}_1 - \overline{x}_3\,|}{\sqrt{s^2\left(\dfrac{1}{n_1} + \dfrac{1}{n_3}\right)}} < t_{\text{табл}} = t_{\alpha;\,f},$$

где

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_3 - 1)s_3^2}{n_1 + n_3 - 2} = \frac{7 \cdot 5,43 + 7 \cdot 13,71}{8 + 8 - 2} \approx 9,57; \qquad f = n_1 + n_3 - 2 = 8 + 8 - 2 = 14.$$

Сравнивая расчетное и критическое значения критерия Стьюдента:

$$t_{\text{расч}} = \frac{|122 - 134|}{\sqrt{9,57 \cdot \left(\dfrac{1}{8} + \dfrac{1}{8}\right)}} \approx 7,76 > t_{\text{табл}} = t_{0,05;\,14} = 2,14,$$

заключаем: на уровне значимости 0,05 можно утверждать, что гипотеза $H_0^{/}$ о равенстве наименьшего и наибольшего средних отвергается. Следовательно, гипотеза о равенстве трех средних также отвергается.