

ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

В группе пациентов реакция на лечение препаратом А лучше, чем препаратом В. Будут ли наблюдаться подобные различия во всей совокупности подобных пациентов или они являются случайными?

Ответ - с помощью критерия значимости.

Три возможных причины наблюдаемых различий:

1. Препарат А действительно лучше препарата В.
2. На различия мог повлиять какой-то другой фактор (например, возраст пациентов).
3. Различия вызваны случайным разбросом данных

Рисунок 1

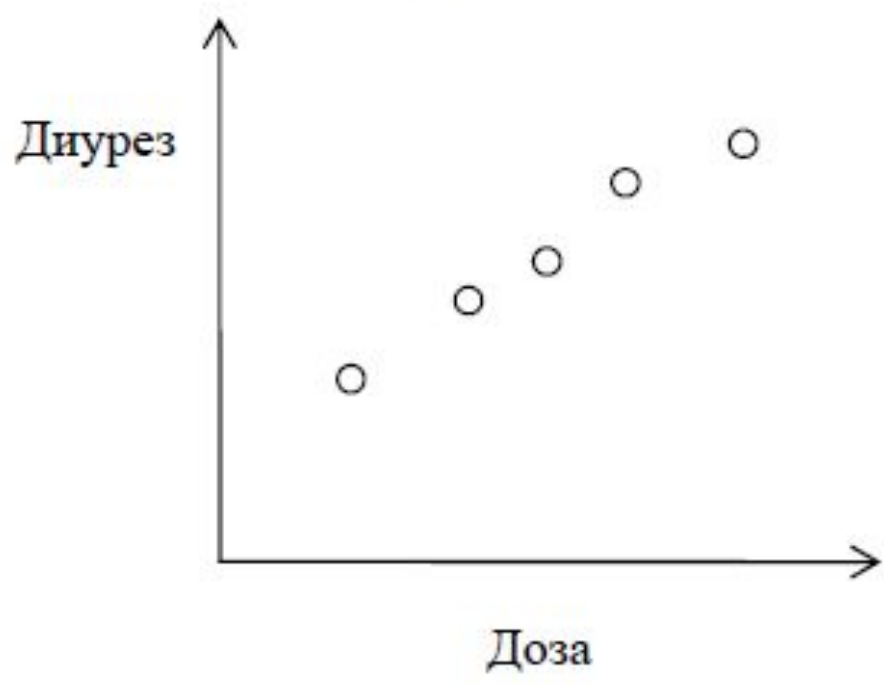
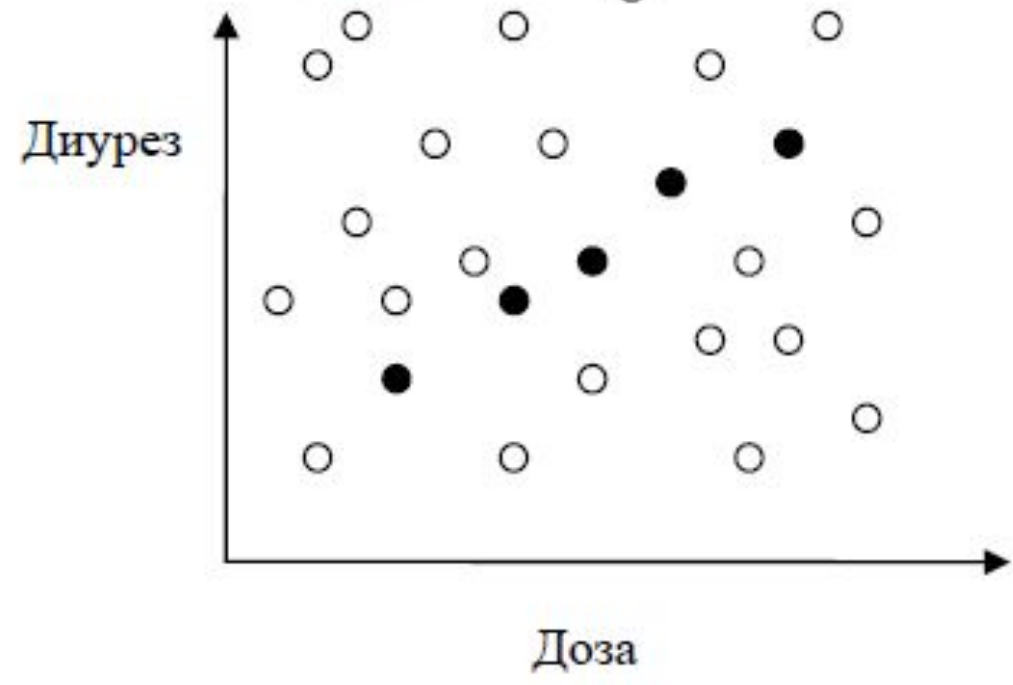


Рисунок 2



Принцип построения критериев значимости

- *Нулевая гипотеза* (H_0): исследуемые факторы не оказывают никакого влияния на исследуемый параметр, и между сравниваемыми группами нет различий (наблюдаемые различия являются результатом случайного разброса данных);
- *Альтернативная гипотеза* (H_1): выбор определяется формулировкой задач.
- Критерий значимости определяет правило, по которому принимается решение принять или отвергнуть нулевую гипотезу

Пример

Исследователь хочет узнать, влияет ли препарат на температуру тела. Схема эксперимента: взять две группы, одной дать препарат, другой плацебо, измерить температуру и вычислить для обеих групп среднюю температуру и стандартные отклонения. Даже если препарат не оказывает никакого действия, средние значения в группах вряд ли совпадут. Насколько вероятно, что наблюдаемые различия случайны?

Нулевая гипотеза: препарат не оказывает влияния на температуру.

Альтернативная гипотеза: препарат влияет на температуру

- Рассчитывается **статистика критерия**, которая сравнивается с **критической величиной**, полученной из набора статистических таблиц.
- Если рассчитанное значение (статистика критерия) **больше** табличного, нулевая гипотеза отклоняется, и различия признаются **статистически значимыми**.

<i>Действия</i> <i>Состояния</i>	Нулевая гипотеза принимается	Нулевая гипотеза отвергается
Нулевая гипотеза верна	Верное решение	Ошибка первого рода α
Нулевая гипотеза неверна	Ошибка второго рода β	Верное решение

- α - *уровень значимости критерия*
- **максимальная приемлемая вероятность в тестах на статистическую значимость - 5% ($\alpha=0,05$)**
- $(1-\beta)$ – *мощность критерия*

Параметрические критерии. Критерий Стьюдента

Разработан **Уильямом Госсетом** для оценки качества пива в компании Гиннесс. В связи с обязательствами перед компанией по неразглашению коммерческой тайны, статья Госсета вышла в 1908 году в журнале «Биометрика» под псевдонимом «Student»

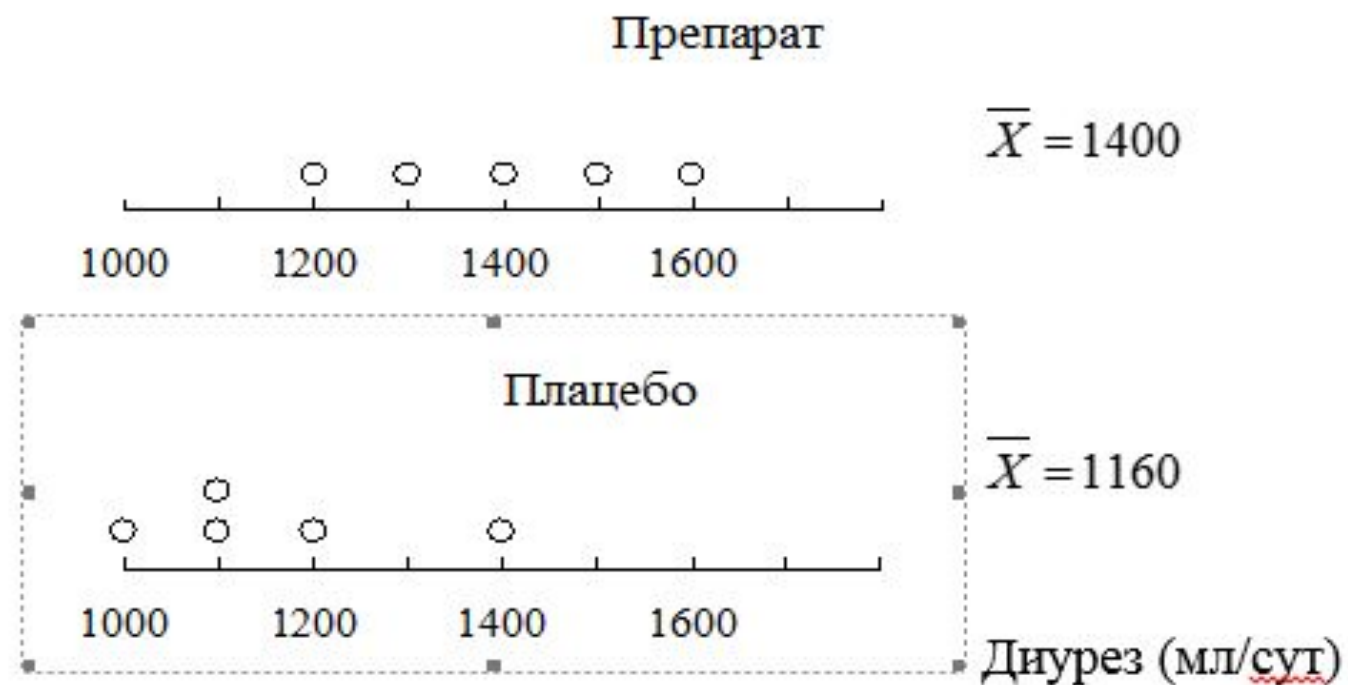


Условия применения:

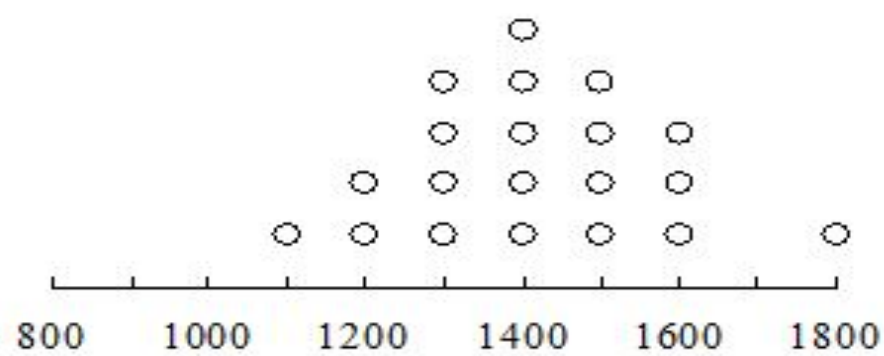
Предназначен для сравнения средних в двух группах

1. Исходная совокупность имеет нормальное распределение
2. Дисперсии в двух сравниваемых выборках равны
3. Выборки извлечены случайным и независимым образом

Изучение действия диуретического препарата

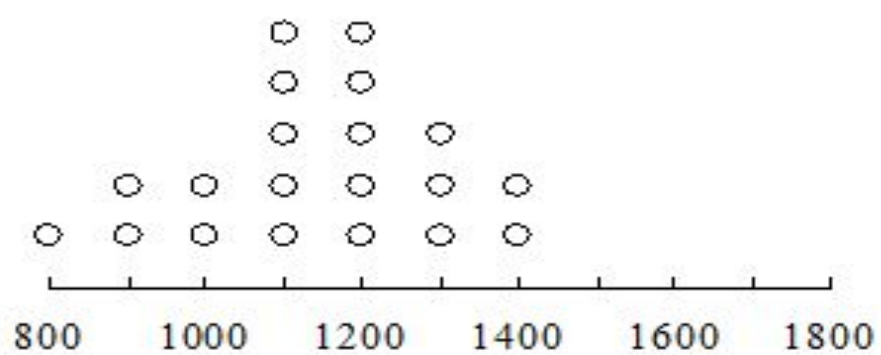


Препарат



$$\bar{X} = 1400$$

Плацебо



$$\bar{X} = 1160$$

Диурез (мл/сут)

Величина критерия Стьюдента t представляет собой отношение разности средних к стандартной ошибке разности средних:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{\bar{X}_1 - \bar{X}_2}}$$

Процедура применения критерия Стьюдента

1. Проверяем соответствие распределения нормальному закону.
2. Проверяем равенство выборочных дисперсий с помощью **критерия Фишера**:

Находим отношение большей и меньшей по величине дисперсий. Полученное значение F сравниваем с критическим табличным значением в зависимости от числа степеней свободы

$$v_1 = v_{\text{числ}} = n_{\text{числ}} - 1; \quad v_2 = v_{\text{знам}} = n_{\text{знам}} - 1.$$

Если рассчитанное значение критерия меньше табличного, то между дисперсиями нет значимых различий и их можно считать равными.

Находим объединенную оценку дисперсии:

1) для выборок одинакового объема $n_1 = n_2$:

$$s^2 = \frac{s_1^2 + s_2^2}{2}$$

2) для выборок разного объема:

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

Вычисляем t по формуле:

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s^2}{n_1} + \frac{s^2}{n_2}}}$$

Сравниваем полученное значение t с критическим значением из таблицы в зависимости от числа степеней свободы $v = n_1 + n_2 - 2$ и уровня значимости α .

Если рассчитанное значение t больше табличного, отвергаем нулевую гипотезу и утверждаем, что с уровнем значимости α существуют различия между двумя группами

Пример

Изучить, как влияют два типа музыки на способность студентов решать задачи, требующие умственной концентрации. Для решения задачи была отобрана однородная группа из 30 студентов. Эта группа была случайным образом разделена на две подгруппы по 15 человек. Члены каждой из групп выполняли серию из 40 задач под музыку разных типов. В результате измерялось количество задач, решенных каждым студентом за отведенное время.

Группа 1 (тип музыки 1)	Группа 2 (тип музыки 2)
26 22 23 26 20	18 23 21 20 20
22 26 25 24 21	28 20 16 20 26
23 23 19 29 22	21 25 17 18 19

Есть ли различия в количестве решенных задач?

$$\overline{X_1} = 23,4; \quad s_1 = 2,64.$$

$$\overline{X_2} = 20,8; \quad s_2 = 3,38.$$

Для первой группы медиана равна 23, три моды: 22, 23 и 26; 25-ый процентиль равен 22, 75-ый процентиль - 26.

Для второй группы медиана равна 20, мода – также 20, 25-ый процентиль равен 18, 75-ый процентиль – 23.

Интервалы \pm два стандартных отклонения для обеих групп включают практически все выборочные значения. Таким образом, обе выборки можно считать извлеченными из нормально распределенной совокупности.

Критерий Фишера:

$$F = \frac{3,38^2}{2,64^2} = 1,64$$

Число степеней свободы :

$$V_{\text{числ}} = 15 - 1 = 14;$$

$$V_{\text{знам}} = 15 - 1 = 14$$

		<i>Число степеней свободы для числителя</i>							
		<i>($\nu_{\text{меж}}$)</i>							
		1	2	3	...	12	13	14	
<i>Число степеней свободы для знаменателя</i>	<i>($\nu_{\text{внутр}}$)</i>	1	161	199	216	...			
	2								
	...								
	14	4.60	3.74	3.34	...		2.51	2.48	
		8.86	6.51	5.56	...		3.75	3.70	
	...								

Рассчитанное значение **меньше** критического, поэтому между выборочными дисперсиями **нет статистически значимых различий** – можно применить критерий Стьюдента.

$$s^2 = \frac{s_1^2 + s_2^2}{2} = \frac{2,64^2 + 3,38^2}{2} = 9,21$$

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s^2}{n_1} + \frac{s^2}{n_2}}} = \frac{23,4 - 20,08}{\sqrt{\frac{9,21}{15} + \frac{9,21}{15}}} = 2,35$$

$$v = n_1 + n_2 - 2 = 15 + 15 - 2 = 28$$

		<i>Уровень значимости α</i>					
v	---	0.05	0.02	0.01	---	0.001	
28	---	2.048	2.467	2.763	---	3.674	

Группы отличаются статистически значимо с уровнем значимости 0,05.

Студенты, слушавшие музыку первого типа, лучше справились с предложенными задачами. Вероятность ошибки нашего заключения менее 5%

- Применение критерия Стьюдента в классическом варианте правомерно лишь при соблюдении указанных условий применения. По данным ряда авторов, в медицинских исследованиях лишь около 20% количественных признаков удовлетворяют указанным условиям. Таким образом, в большинстве случаев использование t-критерия неправомерно

- www.biometrica.tomsk.ru

Если схема эксперимента предполагает не две, а три и более групп?



Попарные сравнения групп –
проблема множественных сравнений!!!

