Дисперсионный анализ

Основы биостатистики, осень 2022

Марина Варфоломеева

- Однофакторный дисперсионный анализ
- Условия применимости дисперсионного анализа
- Запланированные сравнения
- Пост хок тесты
- Фиксированные и случайные факторы
- Ошибка измерения, воспроизводимость

Пример: Мелатонин как регулятор циркадного ритма

После джет-лага режим выравнивается, когда глаза привыкают к ритму освещения на новом месте.

Мелатонин:

- регулятор суточного ритма
- вырабатывается ночью
- свет снижает выработку

Проверка сомнительного результата

Одно исследование показало, что продукцию мелатонина можно регулировать, освещая внутреннюю сторону коленей (Campbell, Murphy, 1998).

Проверка (Wright, Czeisler, 2002):

3 группы людей (22 чел.) будили ночью и 3 часа

- освещали ярким светом глаза
- освещали ярким светом колени
- не освещали ничего (контроль)

По уровню продукции мелатонина регистрировали сдвиг циркадного ритма (в часах).

Влияет ли экспериментальная процедура (группа) на сдвиг циркадного ритма?

Проверка сомнительного результата

Одно исследование показало, что продукцию мелатонина можно регулировать, освещая внутреннюю сторону коленей (Campbell, Murphy, 1998).

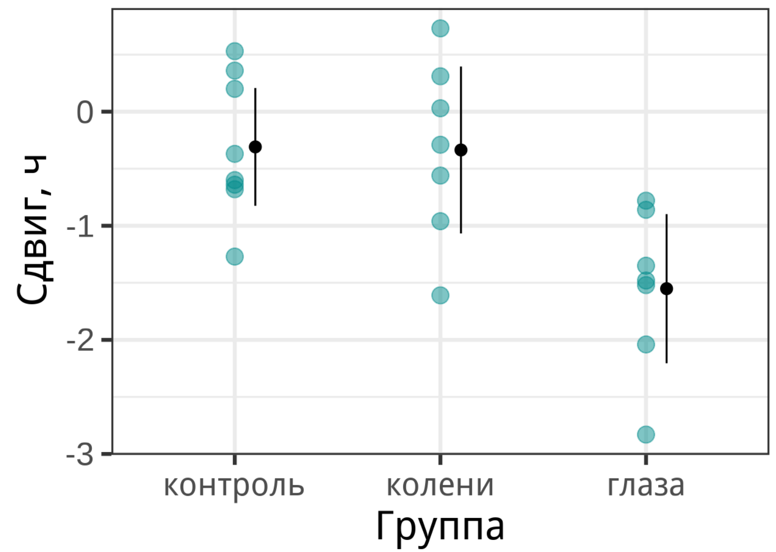
Проверка (Wright, Czeisler, 2002):

3 группы людей (22 чел.) будили ночью и 3 часа

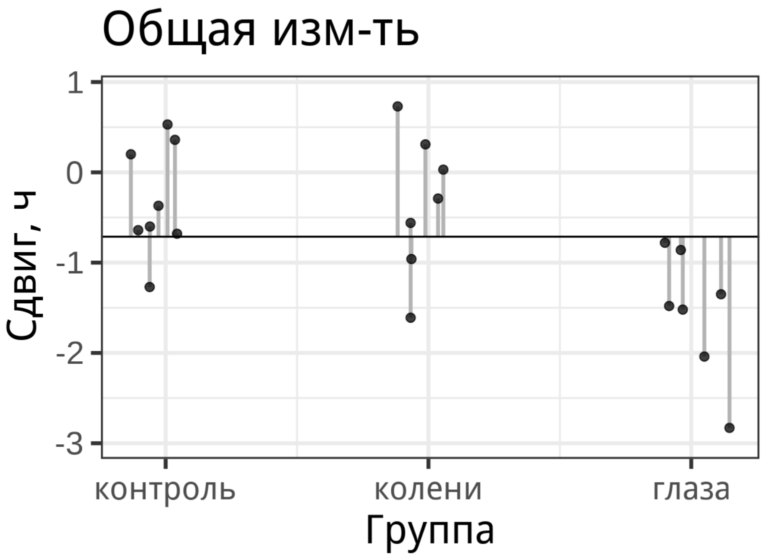
- освещали ярким светом глаза
- освещали ярким светом колени
- не освещали ничего (контроль)

По уровню продукции мелатонина регистрировали сдвиг циркадного ритма (в часах).

Влияет ли экспериментальная процедура (группа) на сдвиг циркадного ритма?

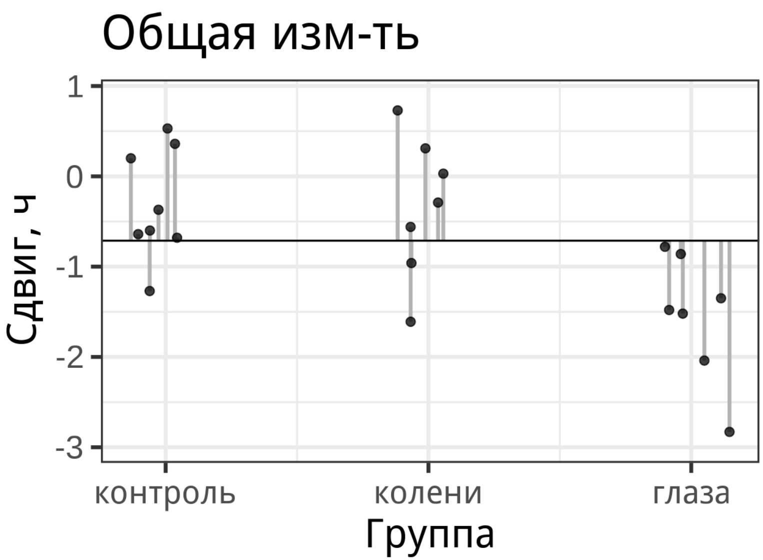


Дисперсионный анализ



Дисперсионный анализ (analysis of variance, ANOVA) — метод одновременной проверки гипотез о равенстве средних значений в нескольких группах.

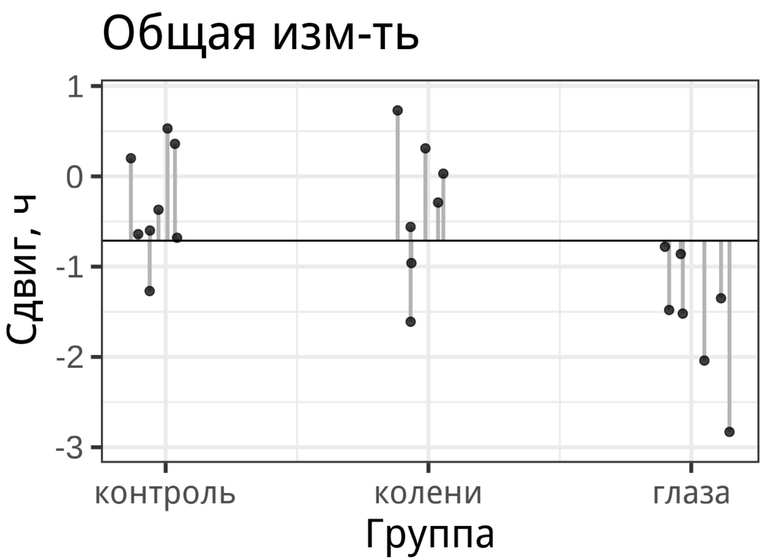
Дисперсионный анализ



Дисперсионный анализ (analysis of variance, ANOVA) — метод одновременной проверки гипотез о равенстве средних значений в нескольких группах.

Насколько наблюдения из одной группы более похожи друг на друга, чем из разных групп?

Дисперсионный анализ

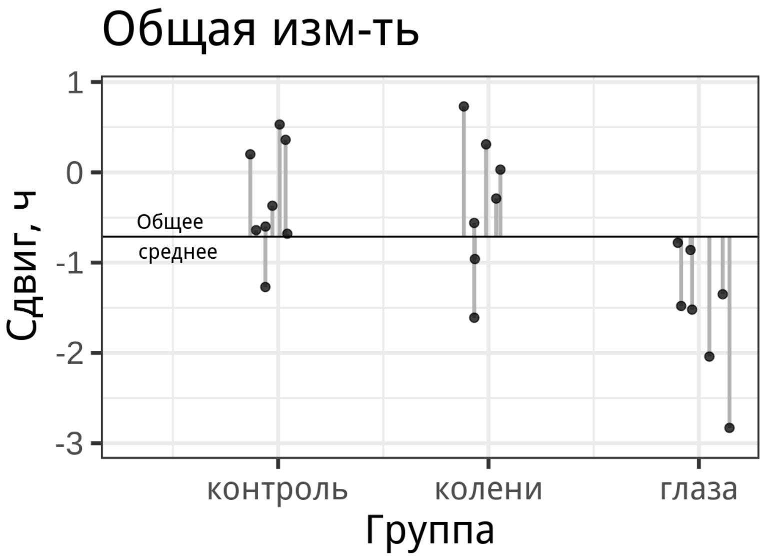


Дисперсионный анализ (analysis of variance, ANOVA) — метод одновременной проверки гипотез о равенстве средних значений в нескольких группах.

Насколько наблюдения из одной группы более похожи друг на друга, чем из разных групп?

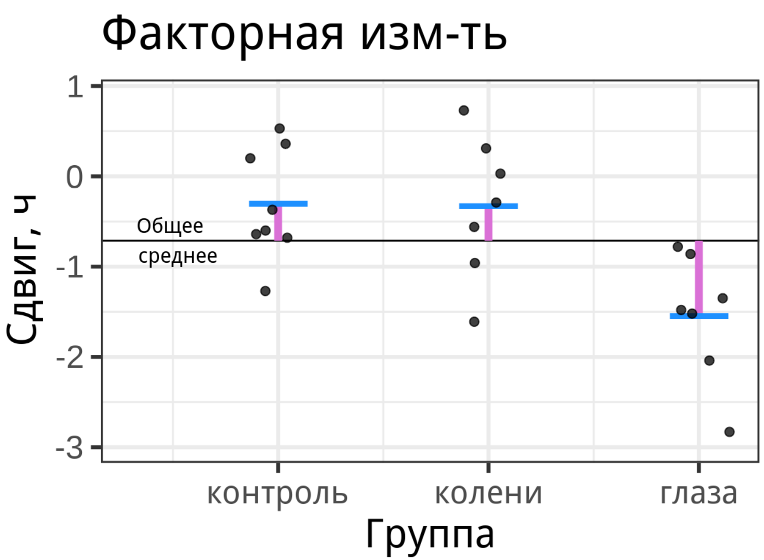
- однофакторный (как в примере)
- многофакторный (деление на группы сразу по нескольким факторам).

Общая изменчивость



Общая изменчивость SS_t — это сумма квадратов отклонений наблюдаемых значений y_i от общего среднего $\bar{y}$

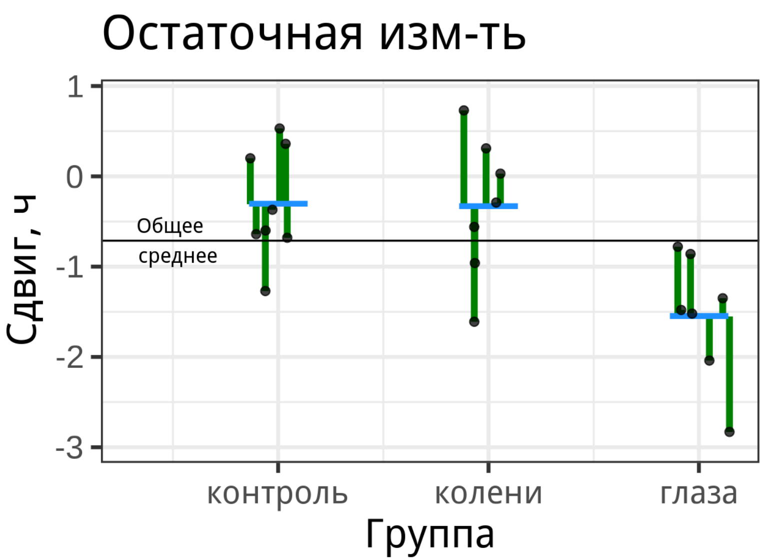
Факторная (межгрупповая) изменчивость



Отклонения внутригрупповых средних от общего среднего в генеральной совокупности — это эффект фактора $\alpha_j = \mu_j - \mu$, где $j = 1, 2, \ldots, p$ — это одна из p групп.

Мы оцениваем эффект фактора по реальным данным $\bar{y}_j - \bar{y}$

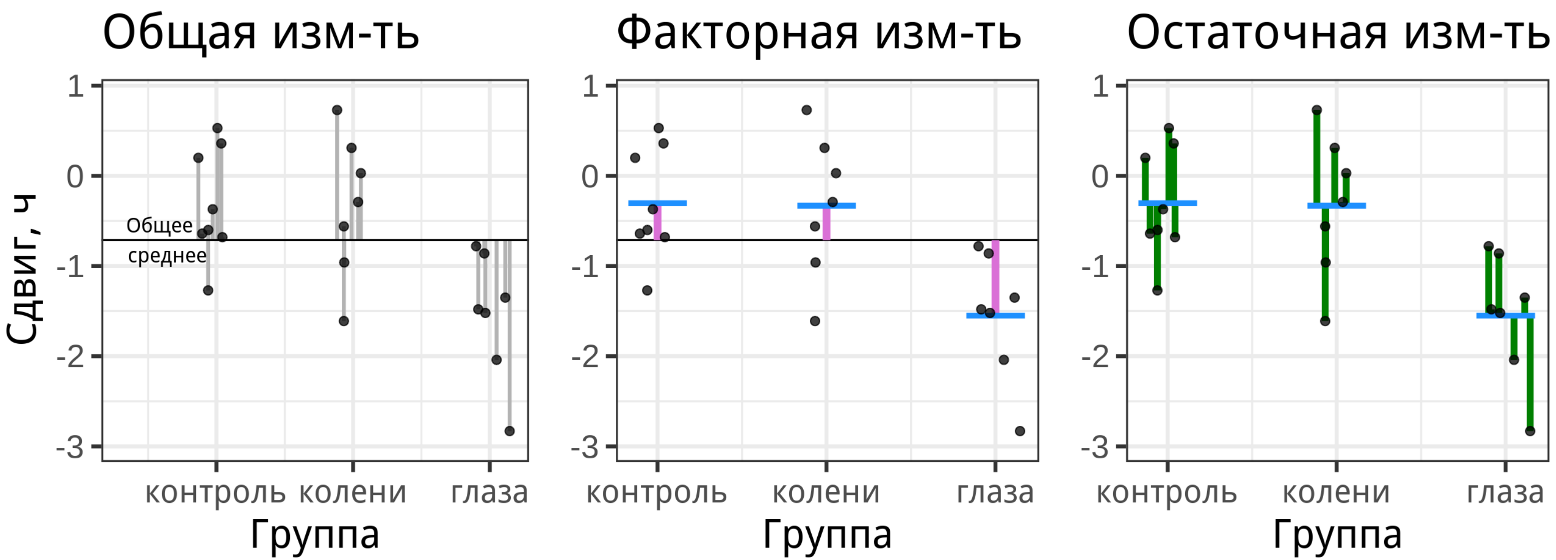
Остаточная (внутригрупповая) изменчивость



Отклонения значений от средних внутри групп (**остатки**) — это изменчивость, которую не может объяснить группировка по фактору. Ещё её называют случайной изменчивостью.

Структура общей изменчивости

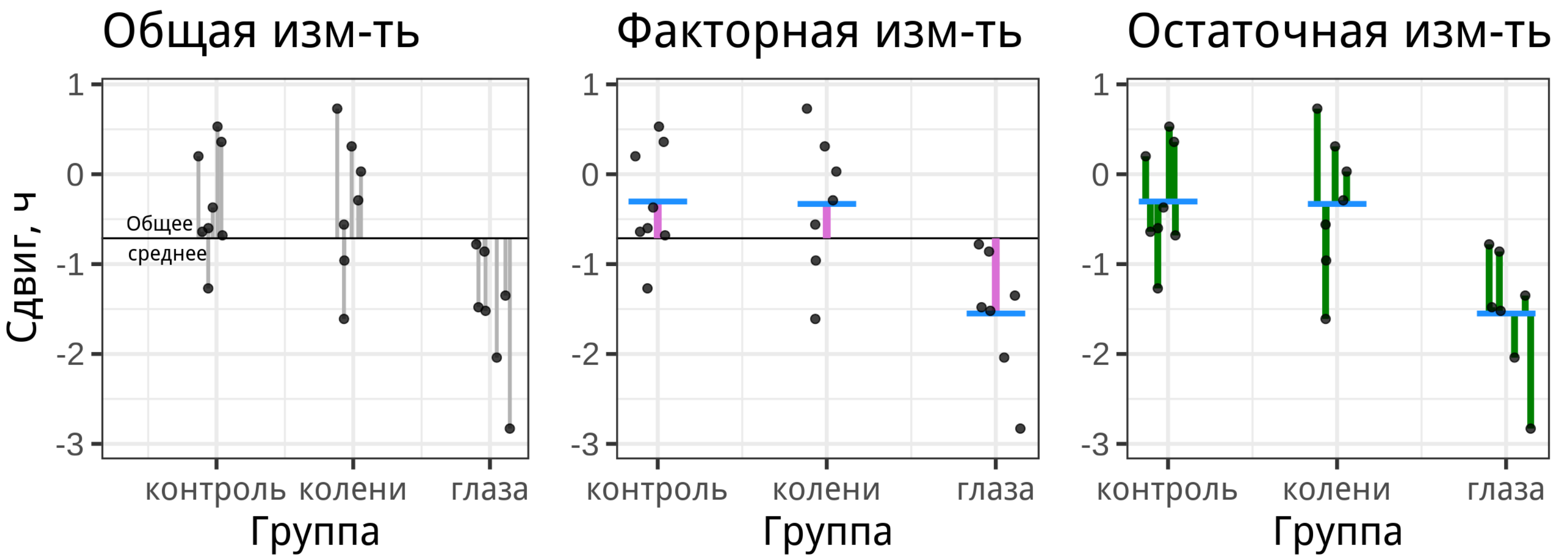
$$SS_t = SS_x + SS_e$$



Общая изменчивость	Факторная изменчивость	Остаточная изменчивость
...	...	...
$SS_t = \sum\sum(\bar{y} - y_{ij})^2$	$SS_x = \sum n_j(\bar{y}_j - \bar{y})^2$	$SS_e = \sum\sum(\bar{y}_j - y_{ij})^2$
$df_t = n - 1$	$df_x = p - 1$	$df_e = n - p$

От изменчивостей к дисперсиям

$$SS_t = SS_x + SS_e \qquad MS_t \neq MS_x + MS_e$$



Общая дисперсия	Факторная дисперсия	Остаточная дисперсия
$MS_t = \frac{SS_t}{df_t}$	$MS_x = \frac{SS_x}{df_x}$	$MS_e = \frac{SS_e}{df_e}$
$SS_t = \sum\sum(\bar{y} - y_{ij})^2$	$SS_x = \sum n_j(\bar{y}_j - \bar{y})^2$	$SS_e = \sum\sum(\bar{y}_j - y_{ij})^2$
$df_t = n - 1$	$df_x = p - 1$	$df_e = n - p$

MSx и MSe помогают тестировать значимость фактора

Если зависимости нет, то $\mu_1 = \ldots = \mu_p$ — средние равны во всех p группах, и $MS_x \sim MS_e$

при условии, что

- дисперсии остатков в группах равны
- фактор имеет фиксированное число градаций

MSx и MSe помогают тестировать значимость фактора

Если зависимости нет, то $\mu_1 = \ldots = \mu_p$ — средние равны во всех p группах, и $MS_x \sim MS_e$

при условии, что

- дисперсии остатков в группах равны
- фактор имеет фиксированное число градаций

- $H_0: \mu_1 = \ldots = \mu_p$ — средние во всех p группах равны.
- $H_A: \exists\, i, j: \mu_i \neq \mu_j$ — **хотя бы одно** среднее отличается от общего среднего.

MSx и MSe помогают тестировать значимость фактора

Если зависимости нет, то $\mu_1 = \ldots = \mu_p$ — средние равны во всех p группах, и $MS_x \sim MS_e$

при условии, что

- дисперсии остатков в группах равны
- фактор имеет фиксированное число градаций

- $H_0: \mu_1 = \ldots = \mu_p$ — средние во всех p группах равны.
- $H_A: \exists\, i, j: \mu_i \neq \mu_j$ — **хотя бы одно** среднее отличается от общего среднего.

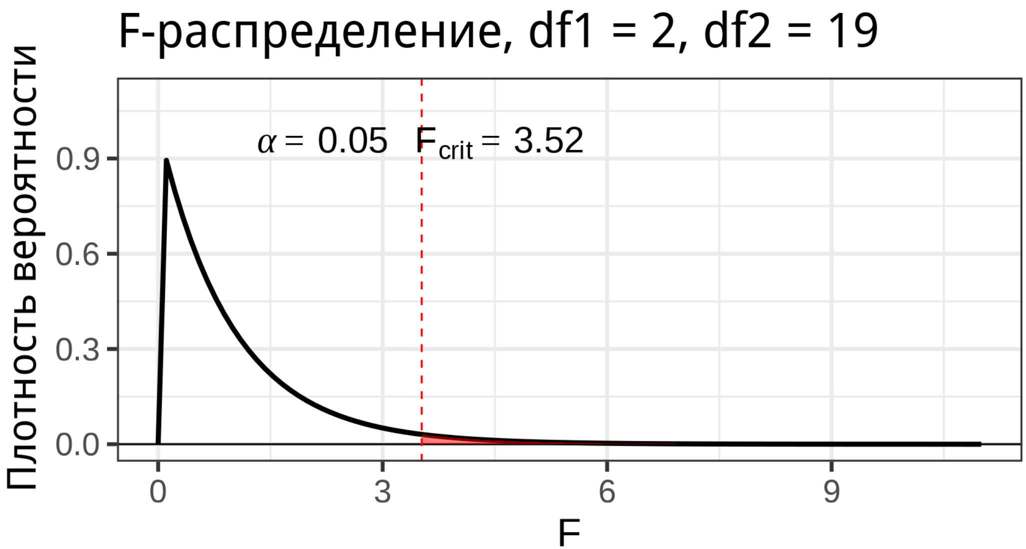
$$F_{df_x, df_e} = \frac{MS_x}{MS_e}$$

Тестирование значимости фактора при помощи F-критерия

- $H_0 : \mu_1 = \ldots = \mu_p$ — средние во всех p группах равны.
- $H_A : \exists\, i, j : \mu_i \neq \mu_j$ — **хотя бы одно** среднее отличается от общего среднего.

$$F_{df_x, df_e} = \frac{MS_x}{MS_e}$$

В однофакторном дисперсионном анализе $df_x = p - 1$ и $df_e = n - p$.



Результаты дисперсионного анализа часто представляют в виде таблицы

Источник изменчивости	SS	df	MS	F
Фактор	$SS_x = \sum n_j(\bar{y}_j - \bar{y})^2$	$df_x = p - 1$	$MS_x = \frac{SS_x}{df_x}$	$F_{df_x df_e} = \frac{MS_x}{MS_e}$
Случайная	$SS_e = \sum\sum(\bar{y}_j - y_{ij})^2$	$df_e = n - p$	$MS_e = \frac{SS_e}{df_e}$	
Общая	$SS_t = \sum\sum(\bar{y} - y_{ij})^2$	$df_t = n - 1$		

Минимальное описание результатов в тексте должно содержать F_{df_x,df_e} и p.

Дисперсионный анализ данных о мелатонине

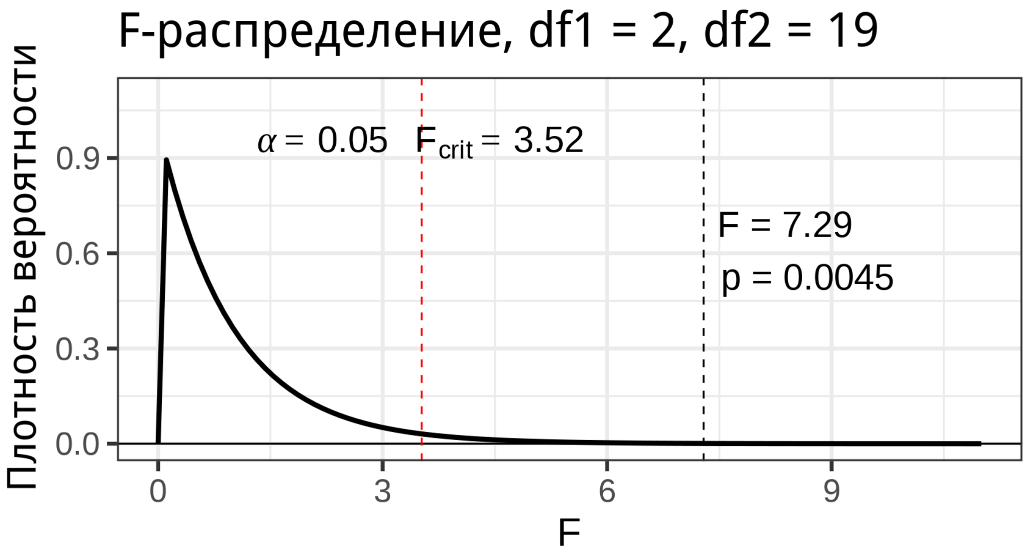
- $H_0 : \mu_{\text{контроль}} = \mu_{\text{колени}} = \mu_{\text{глаза}}$ — средние во всех 3 группах равны.
- H_A : **хотя бы одно** среднее отличается от общего среднего.

	SS	df	MS	F	P
Группа	7.22	2	3.612	7.29	0.004
Случайная	9.41	19	0.496		
Общая	16.64	21			

Дисперсионный анализ данных о мелатонине

- $H_0: \mu_{\text{контроль}} = \mu_{\text{колени}} = \mu_{\text{глаза}}$ — средние во всех 3 группах равны.
- H_A : **хотя бы одно** среднее отличается от общего среднего.

	SS	df	MS	F	P
Группа	7.22	2	3.612	7.29	0.004
Случайная	9.41	19	0.496		
Общая	16.64	21			



Коэффициент детерминации — мера объясненной изменчивости

$$SS_t = SS_x + SS_e$$

- SS_t — общая изменчивость
- SS_x — объясненная фактором изменчивость
- SS_e — остаточная изменчивость

Как оценить, какую долю от всей изменчивости зависимой переменной объясняет фактор?

Коэффициент детерминации — мера объясненной изменчивости

$$SS_t = SS_x + SS_e$$

- SS_t — общая изменчивость
- SS_x — объясненная фактором изменчивость
- SS_e — остаточная изменчивость

Как оценить, какую долю от всей изменчивости зависимой переменной объясняет фактор?

Коэффициент детерминации:

$$R^2 = \frac{SS_x}{SS_t}$$

$0 < R^2 < 1$

Можно записать в долях или процентах.

Какую долю изменчивости объясняет фактор в примере про мелатонин?

	SS	df	MS	F	P
Группа	7.22	2	3.612	7.29	0.004
Случайная	9.41	19	0.496		
Общая	16.64	21			

Какую долю изменчивости объясняет фактор в примере про мелатонин?

	SS	df	MS	F	P
Группа	7.22	2	3.612	7.29	0.004
Случайная	9.41	19	0.496		
Общая	16.64	21			

$$R^2 = \frac{7.224}{16.64} = 0.434$$

или $R^2 = 43.4~\%$

Условия применимости однофакторного дисперсионного анализа

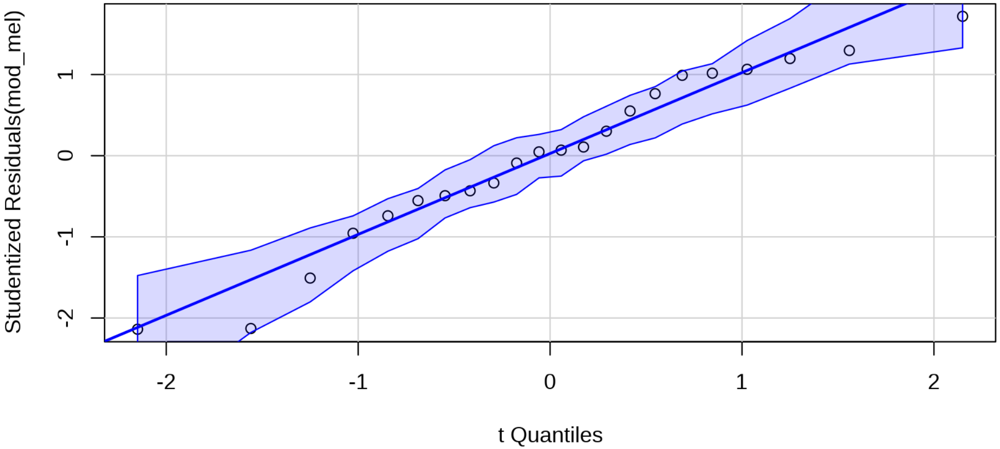
Условия применимости однофакторного дисперсионного анализа

F-тесту можно верить, если выполняются условия применимости:

- Случайность и независимость наблюдений внутри групп
- Нормальное распределение остатков
- Равенство дисперсий остатков в группах по фактору

Проверка нормальности распределения остатков

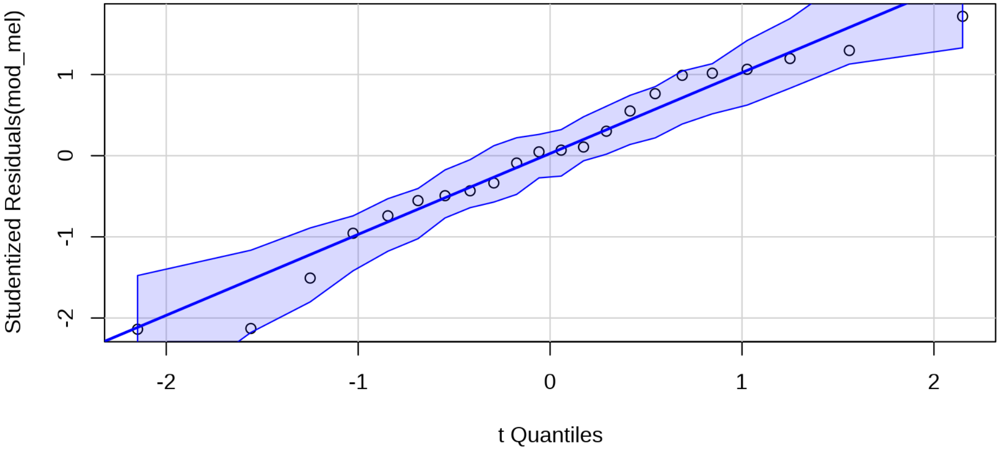
Квантильный график остатков



Знакомый график

Проверка нормальности распределения остатков

Квантильный график остатков

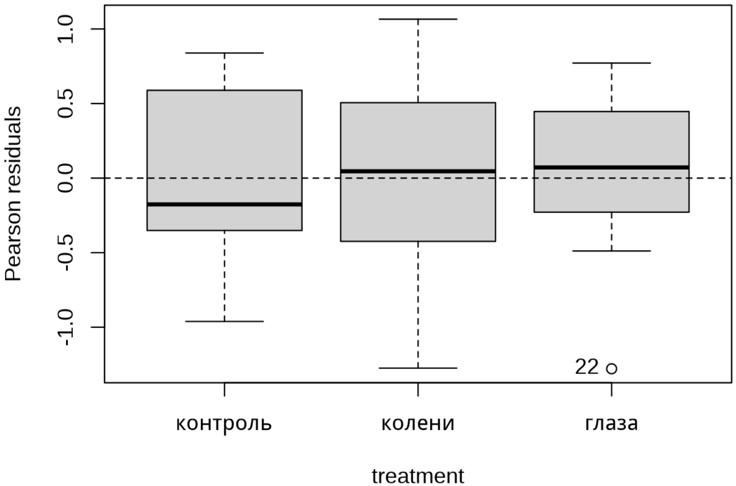


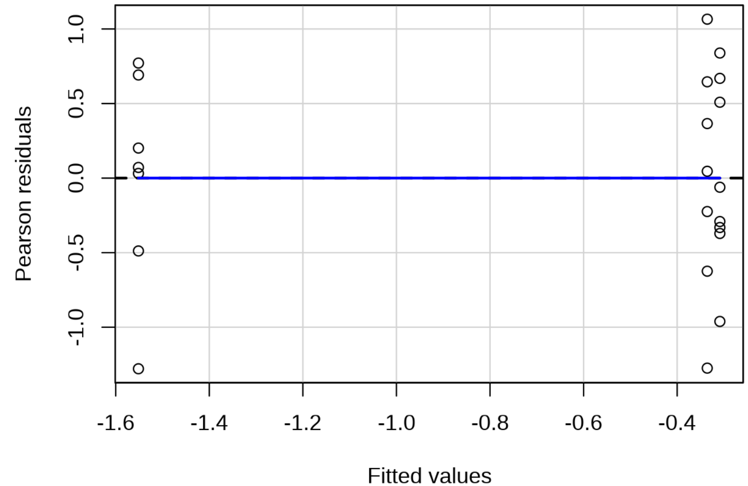
Знакомый график

- Точки должны лежать на одной прямой, если квантили наблюдаемого распределения остатков соответствуют квантилям теоретического распределения.

Проверка равенства дисперсий остатков в группах

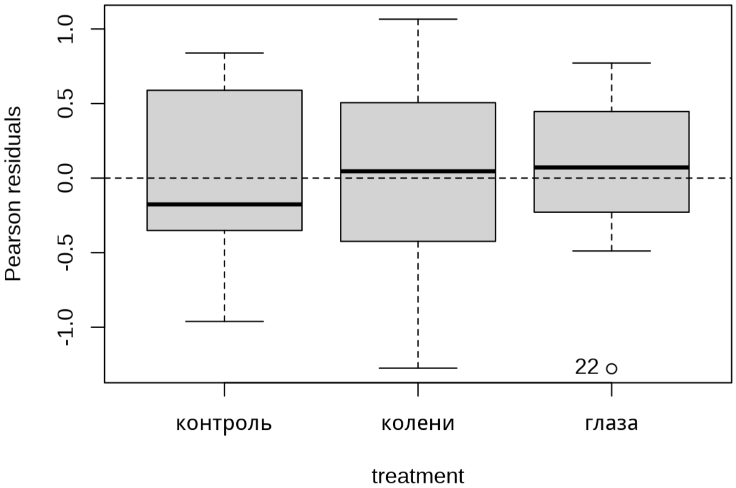
- остатки в группах по фактору
- остатки в зависимости от предсказанных значений

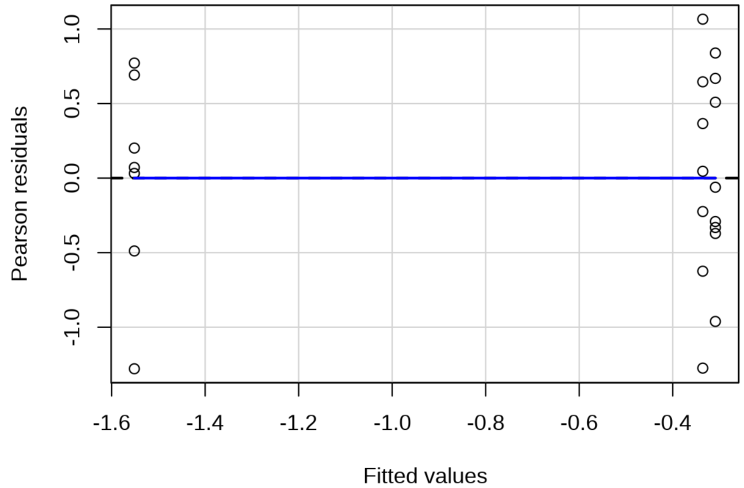




Проверка равенства дисперсий остатков в группах

- остатки в группах по фактору
- остатки в зависимости от предсказанных значений





- Разброс остатков должен быть одинаков
  - в группах
  - вне зависимости от предсказанных значений

Устойчивость дисперсионного анализа

Устойчивость (robustness) — свойство статистического метода, описывающее устойчивость его результатов при нарушении условий применимости.

Устойчивость дисперсионного анализа

Устойчивость (robustness) — свойство статистического метода, описывающее устойчивость его результатов при нарушении условий применимости.

Дисперсионный анализ устойчив к отклонениям от условий применимости, если

- размеры групп примерно одинаковы
- равные объемы групп
- большие выборки

Устойчивость дисперсионного анализа

Устойчивость (robustness) — свойство статистического метода, описывающее устойчивость его результатов при нарушении условий применимости.

Дисперсионный анализ устойчив к отклонениям от условий применимости, если

- размеры групп примерно одинаковы
- равные объемы групп
- большие выборки

Более устойчив к отклонениям от нормальности распределения остатков

Менее устойчив к неравенству дисперсий

Если условия применимости нарушены

- Трансформация данных
- Обобщенная линейная модель (например, для счетных данных или долей)
- Непараметрический тест Краскала-Уоллиса (Kruskal-Wallis test)
- Тест, основанный на пермутациях

Какие именно группы различаются?

Запланированные сравнения

Как понять, какие именно группы различаются?

Дисперсионный анализ говорит только, есть ли влияние фактора.

Два способа понять, какие группы различаются

Два способа понять, какие группы различаются

Запланированные сравнения
(= planned comparisons,
= linear contrasts)

- Можно сравнить выбранные группы.
- Набор гипотез (и сравнений) должен быть определен заранее.
- Делать можно вне зависимости от результатов дисперсионного анализа.

Два способа понять, какие группы различаются

Запланированные сравнения
(= planned comparisons,
= linear contrasts)

- Можно сравнить выбранные группы.
- Набор гипотез (и сравнений) должен быть определен заранее.
- Делать можно вне зависимости от результатов дисперсионного анализа.

Post hoc тесты

- Сравниваются все возможные группы.
- Нет четких заранее сформулированных гипотез.
- Делать можно, только если влияние соответствующего фактора оказалось значимым.

Запланированные сравнения средних

$d = \bar{Y}_2 - \bar{Y}_1$ — разница средних значений в двух группах

Запланированные сравнения средних

$d = \bar{Y}_2 - \bar{Y}_1$ — разница средних значений в двух группах

Стандартная ошибка этой разницы:

$$SE_d = \sqrt{MS_e\Big(\frac{1}{n_1} + \frac{1}{n_2}\Big)}$$

Запланированные сравнения средних

$d = \bar{Y}_2 - \bar{Y}_1$ — разница средних значений в двух группах

Стандартная ошибка этой разницы:

$$SE_d = \sqrt{MS_e \Big(\frac{1}{n_1} + \frac{1}{n_2}\Big)}$$

Доверительный интервал для этой разницы будет накрывать истинное значение $\mu_2 - \mu_1$ в заданном проценте повторных выборок:

$d - |t_{\alpha,df}| \cdot SE_d \le \mu_2 - \mu_1 \le d + |t_{\alpha,df}| \cdot SE_d$

$df = N - p$

Сравним контроль и опыт

Сравним время сдвига циркадного ритма в группе, где освещали колени, и в контроле.

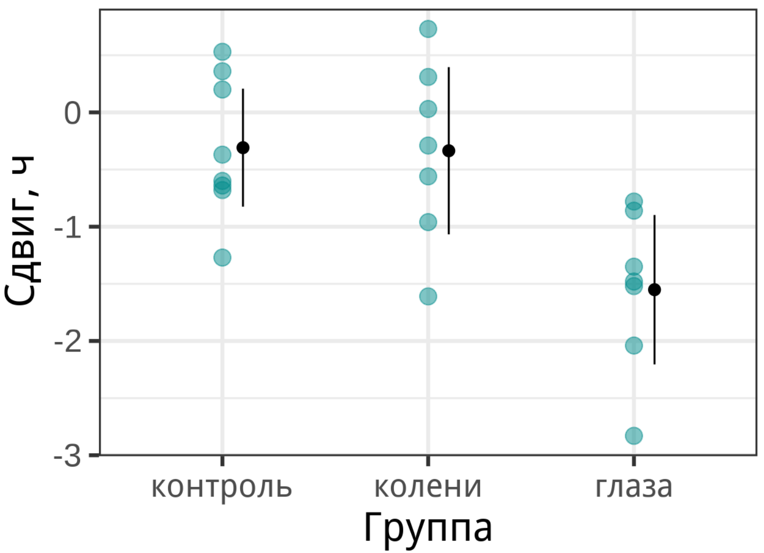
$d = \bar{Y}_2 - \bar{Y}_1 = -0.336 - (-0.309) = -0.027$

$MS_e = 0.496$

$SE_d = \sqrt{0.496\left(\frac{1}{8} + \frac{1}{7}\right)} = 0.364$

$df = 22 - 3 = 19$

$|t_{0.05,19}| = 2.093$



Сравним контроль и опыт

Сравним время сдвига циркадного ритма в группе, где освещали колени, и в контроле.

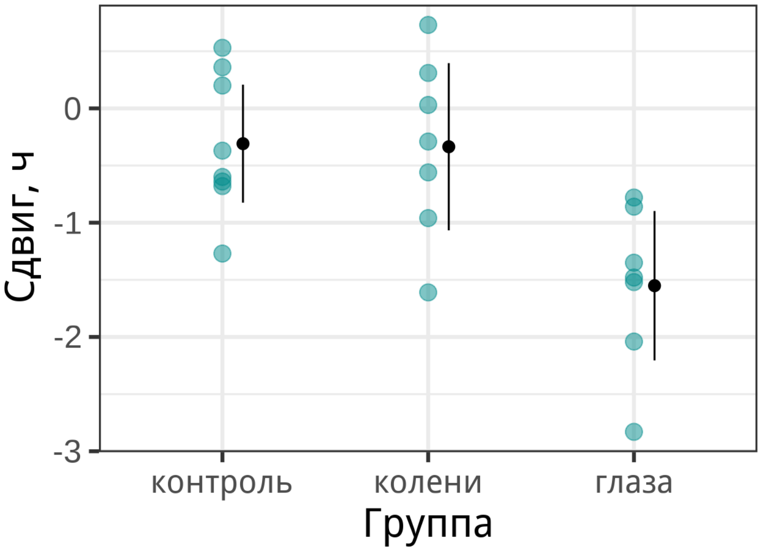
$d = \bar{Y}_2 - \bar{Y}_1 = -0.336 - (-0.309) = -0.027$

$MS_e = 0.496$

$SE_d = \sqrt{0.496\left(\frac{1}{8} + \frac{1}{7}\right)} = 0.364$

$df = 22 - 3 = 19$

$|t_{0.05,19}| = 2.093$



Доверительный интервал :

$-0.027 - 2.093 \cdot 0.364 \le \mu_2 - \mu_1 \le -0.027 + 2.093 \cdot 0.364$

$-0.79 \le \mu_2 - \mu_1 \le 0.736$ или -0.027 ± 0.763

Сравним контроль и опыт

Сравним время сдвига циркадного ритма в группе, где освещали колени, и в контроле.

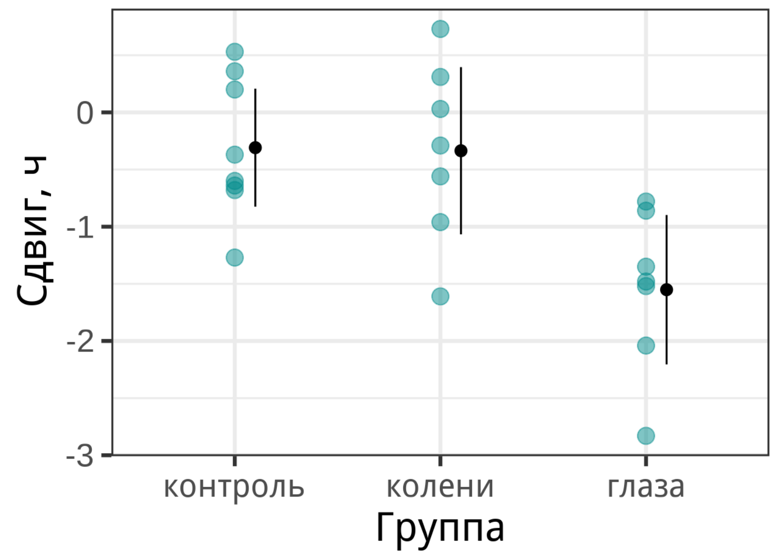
$d = \bar{Y}_2 - \bar{Y}_1 = -0.336 - (-0.309) = -0.027$

$MS_e = 0.496$

$SE_d = \sqrt{0.496\left(\frac{1}{8} + \frac{1}{7}\right)} = 0.364$

$df = 22 - 3 = 19$

$|t_{0.05,19}| = 2.093$



Доверительный интервал :

$-0.027 - 2.093 \cdot 0.364 \leq \mu_2 - \mu_1 \leq -0.027 + 2.093 \cdot 0.364$

$-0.79 \leq \mu_2 - \mu_1 \leq 0.736$ или -0.027 ± 0.763

Он включает 0, значит нет статистически-значимой разницы времени сдвига циркадного ритма в контроле и в группе людей, которым освещали колени.

Чем это отличается от обычного доверительного интервала?

Запланированные сравнения дают доверительный интервал:

$-0.79 \leq \mu_2 - \mu_1 \leq 0.736$ или -0.027 ± 0.763

Чем это отличается от обычного доверительного интервала?

Запланированные сравнения дают доверительный интервал:

$-0.79 \leq \mu_2 - \mu_1 \leq 0.736$ или -0.027 ± 0.763

Обычный доверительный интервал к разнице средних был бы шире:

$-0.813 \leq \mu_2 - \mu_1 \leq 0.759$ или -0.027 ± 0.786

Чем это отличается от обычного доверительного интервала?

Запланированные сравнения дают доверительный интервал:

$-0.79 \leq \mu_2 - \mu_1 \leq 0.736$ или -0.027 ± 0.763

Обычный доверительный интервал к разнице средних был бы шире:

$-0.813 \leq \mu_2 - \mu_1 \leq 0.759$ или -0.027 ± 0.786

Т.е. запланированные сравнения — более мощный метод.

Условия применимости запланированных сравнений

Те же, что и у дисперсионного анализа:

- независимость наблюдений
- нормальное распределение остатков
- одинаковые дисперсии в группах

Условия применимости запланированных сравнений

Те же, что и у дисперсионного анализа:

- независимость наблюдений
- нормальное распределение остатков
- одинаковые дисперсии в группах

Менее устойчивы к отклонениям от условий применимости.

Какие именно группы различаются?

Пост хок тесты

Пост хок тесты

Пост хок тесты (post hoc tests) — позволяют узнать, какие именно группы различаются.

Пост хок тесты

Пост хок тесты (post hoc tests) — позволяют узнать, какие именно группы различаются.

- Делать можно, только если влияние соответствующего фактора оказалось значимым.
- Сравниваются все возможные группы.

Пост хок тесты

Пост хок тесты (post hoc tests) — позволяют узнать, какие именно группы различаются.

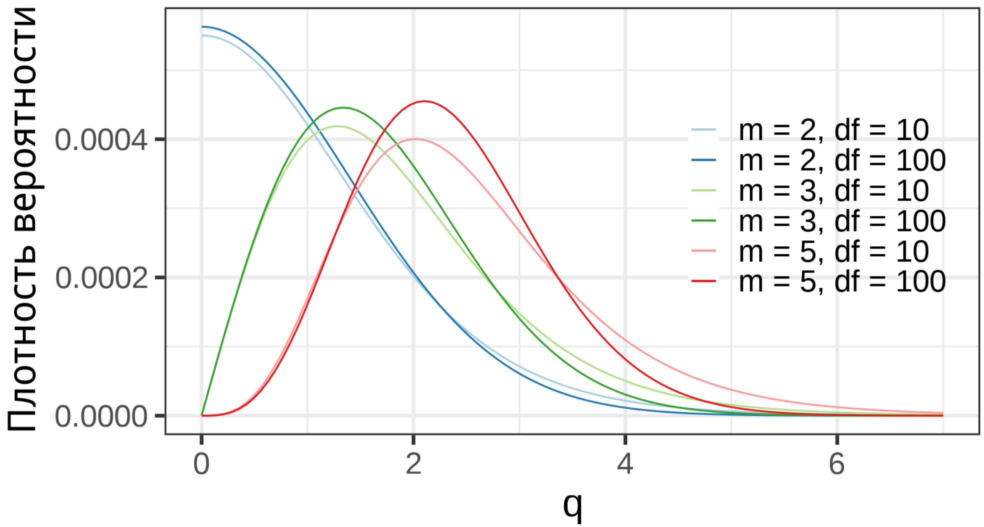
- Делать можно, только если влияние соответствующего фактора оказалось значимым.
- Сравниваются все возможные группы.

- Тест Тьюки (Tukey's Honest Significant Difference, HSD) — пост хок тест, основанный на распределении стьюдентизированного размаха

Распределение стьюдентизированного размаха

(studentized range distribution)

- Аналог t-распределения для любого числа выборок.
- Стандартизованная разница минимального и максимального средних из нескольких выборок.
- Форма зависит от df и от числа выборок m.



Функция распределения для случая равных дисперсий и разных объемов групп:

$$q = \frac{\bar{y}_{max} - \bar{y}_{min}}{\sqrt{s^2 \frac{1}{2}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}}$$

Пост хок Тест Тьюки

(Tukey's Honest Significant Difference)

= стьюдентизированный t-критерий

$$q = \frac{\bar{y}_i - \bar{y}_j}{\sqrt{MS_e \frac{1}{2}(\frac{1}{n_i} + \frac{1}{n_j})}}$$

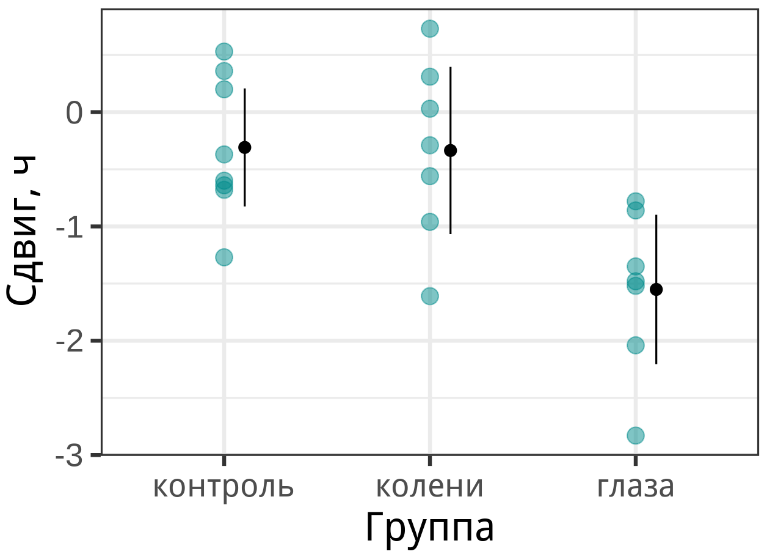
Подчиняется распределению стьюдентизированного размаха
с параметрами $df = df_e = n - p$ и $m = p$ (число групп).

Требуется равенство дисперсий.

Сделаем пост хок тест

Сравним средние во всех группах

	Разность	Н.гр.	В.гр.	p
колени-контроль	-0.027	-0.953	0.899	0.997
глаза-контроль	-1.243	-2.168	-0.317	0.008
глаза-колени	-1.216	-2.172	-0.260	0.012



Фиксированные и случайные факторы

Фиксированные факторы

	Свойства фиксированных факторов
Уровни фактора	заранее определенные и воспроизводимые уровни
Гипотезы	о средних значениях отклика для уровней фактора $H_0 : \mu_1 = \mu_2 = \ldots = \mu_i = \mu$
Экстраполяция	только на уровни из анализа
Число уровней фактора	Если много уровней, нужно много наблюдений

- Разные группы в клинических испытаниях
- Горизонты на литорали
- Разные возрастные группы

Случайные факторы

	Свойства случайных факторов
Уровни фактора	случайная выборка из возможных уровней
Гипотезы	о дисперсии отклика между уровнями фактора $H_0 : \sigma^2_{rand.fact.} = 0$
Экстраполяция	на все возможные уровни
Число уровней фактора	Лучше больше 5 уровней

- Пациенты в клинических испытаниях (если несколько измерений на одном человеке)
- Семьи или выводки (если наблюдения на нескольких членах семьи)

Дисперсионный анализ со случайным фактором

Дисперсионный анализ со случайным фактором

Главная задачане сравнение средних, а определение **компонентов дисперсии** (variance components) случайных факторов.

Примеры использования:

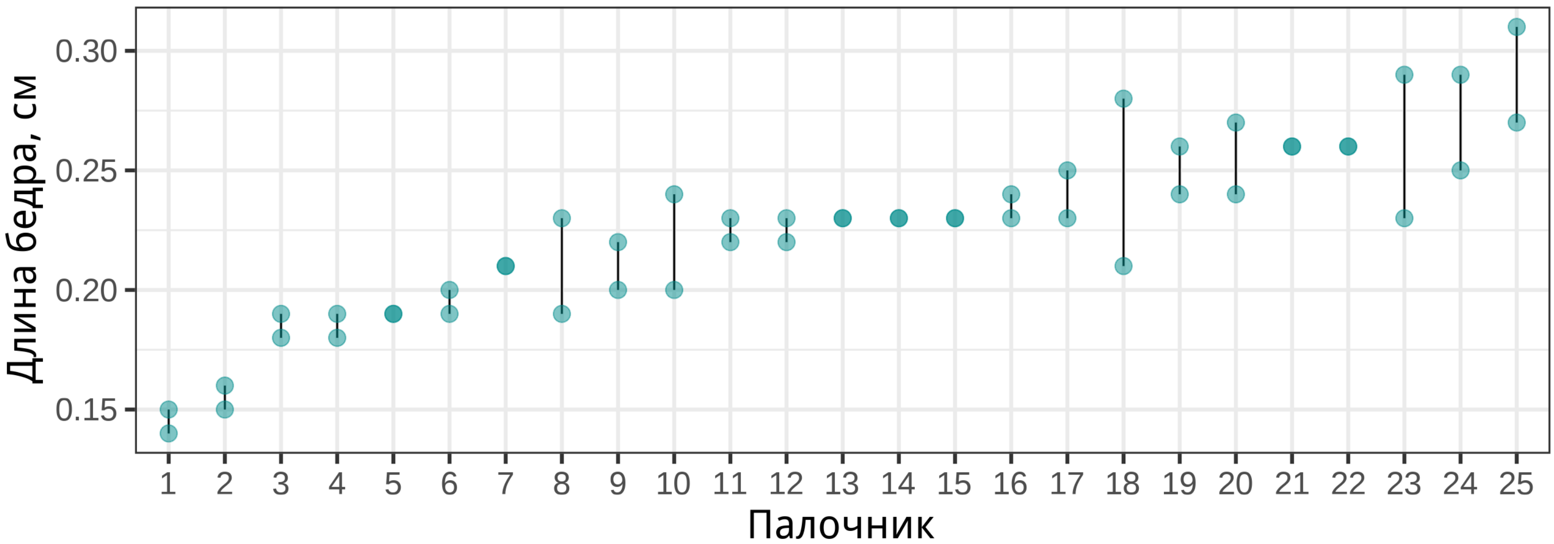
- определение вклада генов и среды в изменчивость признака при селекции
- определение ошибки измерения

Ошибка измерения

Палочник Timema cristinae живет в чаппарале в Калифорнии. В исследовании морфологических адаптаций палочников к разным видам растений учитывали ошибку измерения (Nosil, Crespi, 2006).

Каждого палочника фотографировали и измеряли. Затем повторяли всю процедуру.

Насколько велика ошибка измерений по сравнению с изменчивостью признака?



Дисперсионный анализ со случайным фактором

Если только один случайный фактор, то вычисления такие же, как для модели с фиксированным фактором.

	SS	df	MS
Группа (палочник)	0.059132	24	0.002464
Случайная	0.008900	25	0.000356
Общая	0.068032	49	

Дисперсионный анализ со случайным фактором

Если только один случайный фактор, то вычисления такие же, как для модели с фиксированным фактором.

	SS	df	MS
Группа (палочник)	0.059132	24	0.002464
Случайная	0.008900	25	0.000356
Общая	0.068032	49	

Задача оценить компоненты дисперсии, поэтому F-тест не нужен.

Два уровня изменчивости

	SS	df	MS
Группа (палочник)	0.059132	24	0.002464
Случайная	0.008900	25	0.000356
Общая	0.068032	49	

Два уровня изменчивости

	SS	df	MS
Группа (палочник)	0.059132	24	0.002464
Случайная	0.008900	25	0.000356
Общая	0.068032	49	

Внутри групп

σ^2 и ее оценка MS_e — дисперсия между измерениями на одном и том же объекте.

Два уровня изменчивости

	SS	df	MS
Группа (палочник)	0.059132	24	0.002464
Случайная	0.008900	25	0.000356
Общая	0.068032	49	

Внутри групп

σ^2 и ее оценка MS_e — дисперсия между измерениями на одном и том же объекте.

Между группами

Средние в группах $\mu_i \sim N(\mu_A, \sigma_A^2)$

σ_A^2 и ее оценка — s_A^2 — дисперсия между средними в группах (т.е. на разных объектах).

$$s_A^2 = \frac{MS_{\text{группа}} - MS_e}{n_{\text{группа}}}$$

Два уровня изменчивости

	SS	df	MS
Группа (палочник)	0.059132	24	0.002464
Случайная	0.008900	25	0.000356
Общая	0.068032	49	

Внутри групп

σ^2 и ее оценка MS_e — дисперсия между измерениями на одном и том же объекте.

Между группами

Средние в группах $\mu_i \sim N(\mu_A, \sigma_A^2)$

σ_A^2 и ее оценка — s_A^2 — дисперсия между средними в группах (т.е. на разных объектах).

$$s_A^2 = \frac{MS_{\text{группа}} - MS_e}{n_{\text{группа}}}$$

В примере

$MS_e = 0.000356$

Два уровня изменчивости

	SS	df	MS
Группа (палочник)	0.059132	24	0.002464
Случайная	0.008900	25	0.000356
Общая	0.068032	49	

Внутри групп

σ^2 и ее оценка MS_e — дисперсия между измерениями на одном и том же объекте.

Между группами

Средние в группах $\mu_i \sim N(\mu_A, \sigma_A^2)$

σ_A^2 и ее оценка — s_A^2 — дисперсия между средними в группах (т.е. на разных объектах).

$$s_A^2 = \frac{MS_{\text{группа}} - MS_e}{n_{\text{группа}}}$$

В примере

$$MS_e = 0.000356$$

$$s_A^2 = \frac{0.002464 - 0.000356}{2} = 0.00105$$

Воспроизводимость

Воспроизводимость (repeatability) — способ оценить долю изменчивости признака по отношению ко всей изменчивости.

$$Repeatability = \frac{s_A^2}{s_A^2 - MS_e}$$

Воспроизводимость

Воспроизводимость (repeatability) — способ оценить долю изменчивости признака по отношению ко всей изменчивости.

$$Repeatability = \frac{s_A^2}{s_A^2 - MS_e}$$

$$Repeatability = \frac{0.00105}{0.00105 - 0.000356} = 0.747503$$

Воспроизводимость

Воспроизводимость (repeatability) — способ оценить долю изменчивости признака по отношению ко всей изменчивости.

$$Repeatability = \frac{s_A^2}{s_A^2 - MS_e}$$

$$Repeatability = \frac{0.00105}{0.00105 - 0.000356} = 0.747503$$

Т.е. 75% общей изменчивости объясняется изменчивостью признака,
а 25% — ошибкой измерения.

Условия применимости дисперсионного анализа со случайным фактором

Те же самые + два новых

- Уровни фактора (группы) выбраны слуайно из возможных уровней.
- Средние в группах нормально распределены.

Summary

Summary

Что почитать

- Quinn, Keough, 2002, pp. 173-207
- Logan, 2010, pp. 254 - 282
- Open Intro to Statistics
- Sokal, Rohlf, 1995, pp. 179-260
- Zar, 2010, pp. 189-207