

*STATISTICA*TM

Двух(много)факторный одномерный дисперсионный анализ

*Многомерный дисперсионный
анализ*

*Дисперсионный анализ с
повторными измерениями*

Ковариационный анализ

Выбор статистического теста при сравнении распределений (сравнении центральных тенденций)

Задача	Количественная шкала, нормальное распределение	Порядковая шкала или отклонение от нормального распределения	Номинальная шкала
Сравнить более двух связанных совокупностей в зависимости от одного фактора	Дисперсионный анализ с повторными измерениями	Тест Фридмана	Тест Кохрана
Сравнить две или более несвязанных совокупностей в зависимости от двух и большего числа факторов	Двух(много)-факторный дисперсионный анализ	?	?
Сравнить две или более несвязанных совокупностей ОДНОВРЕМЕННО в зависимости от двух и большего числа факторов, часть из которых – качественные (дискретные), а часть – количественные (непрерывные)	Ковариационный анализ	?	?

Вспомним: идея однофакторного ANOVA

- Один признак;
- Много групп объектов (выборок);
- Сравнение дисперсии признака между группами (межгрупповая изменчивость) и внутри групп (внутригрупповая изменчивость);

Идея двухфакторного ДА

Группа	Популяция 1 - самцы	Популяция 1 - самки	Популяция 2 – самцы	Популяция 2 - самки
Популяция 1 - самцы		P=???	P=???	P=???
Популяция 1 - самки			P=???	P=???
Популяция 2 – самцы				P=???
Популяция 2 - самки				

6 попарных сравнений? Нет! Всего три сравнения:

$$(1) H_0: \mu_{\text{самцы}} = \mu_{\text{самки}}; H_1: \mu_{\text{самцы}} \neq \mu_{\text{самки}};$$

$$(2) H_0: \mu_{\text{популяц1}} = \mu_{\text{популяц2}}; H_1: \mu_{\text{популяц1}} \neq \mu_{\text{популяц2}};$$

(3) Гипотезы относятся к взаимодействию «пол x популяция»

Все же

**попробуем
попарные**

сравнения...

- Две гипотетические популяции насекомых
- В каждой популяции СЛУЧАЙНО извлечены самцы и самки
- Измерены массы особей и длины надкрылий
- Вопросы:
 - 1) есть ли различия между популяциями?
 - 2) есть ли различия между полами?
 - 3) есть ли различия между одноименными полами в разных популяциях?
 - 4) однотипны ли различий у самцов и самок между популяциями, если различия между популяциями вообще

	A	B	C	D	E
1	N	Популяция	Пол	Масса г	Длина mm
17	16	1	самцы	3,5	16,3
18	17	1	самцы	3,3	20,7
19	18	1	самцы	3,3	22,2
20	19	1	самцы	3,6	14,8
21	20	1	самцы	3,5	22,2
22	21	2	самки	6,1	20,3
23	22	2	самки	7,0	22,6
24	23	2	самки	6,2	19,4
25	24	2	самки	5,2	15,0
26	25	2	самки	7,6	25,1
27	26	2	самки	5,3	15,4
28	27	2	самки	5,9	18,2
29	28	2	самки	5,4	15,8
30	29	2	самки	6,2	22,4
31	30	2	самцы	3,9	21,7
32	31	2	самцы	4,4	15,4
33	32	2	самцы	3,9	19,2
34	33	2	самцы	3,2	16,7

Идея двухфакторного ДА

Популяция	Пол	
	Самцы	Самки
Популяция 1	Масса = ?? (n=??)	Масса = ?? (n=??)
Популяция 2	Масса = ?? (n=??)	Масса = ?? (n=??)

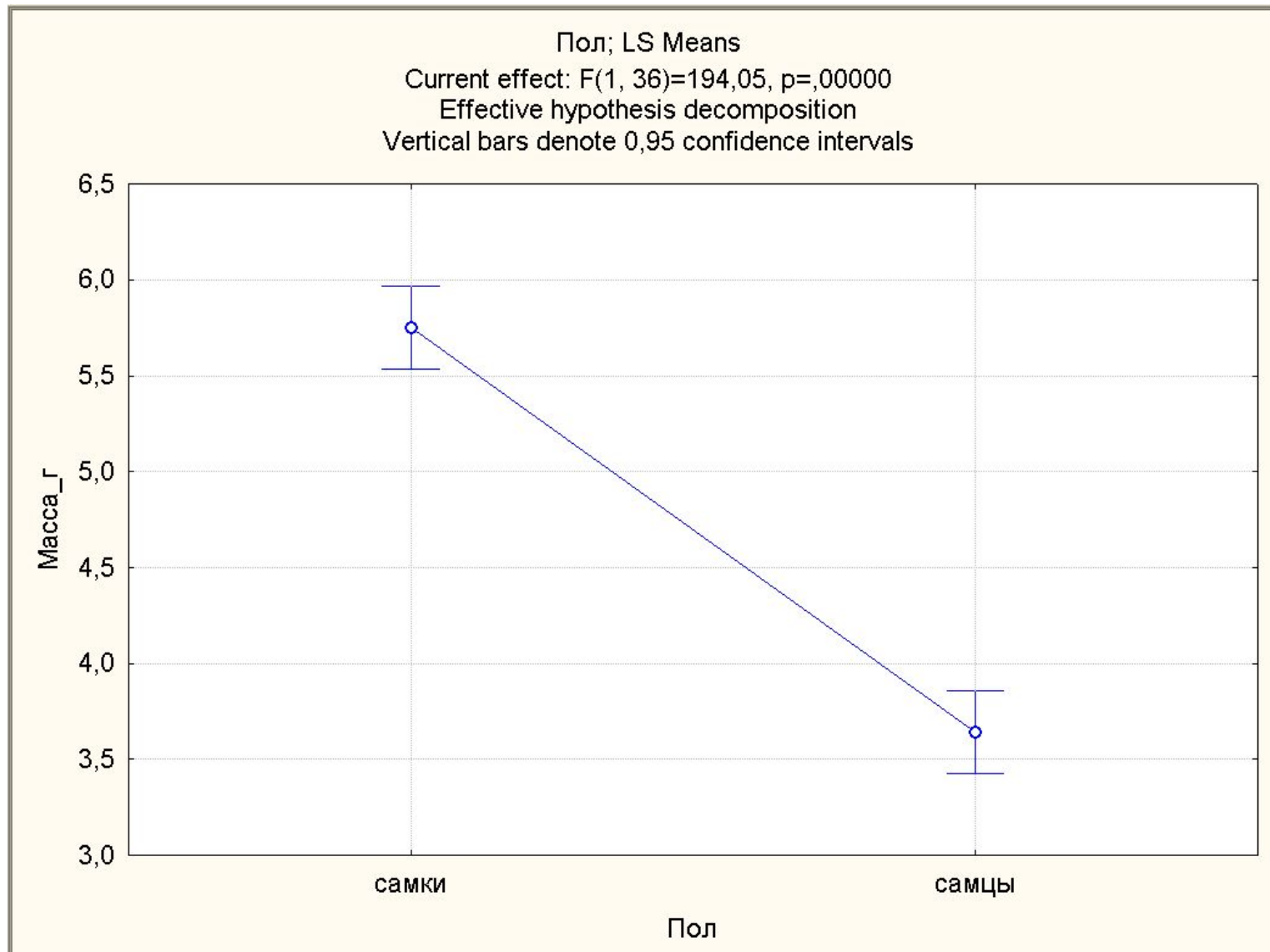
Сравнение 2: между популяциями.

Сравнение 1: между полами.

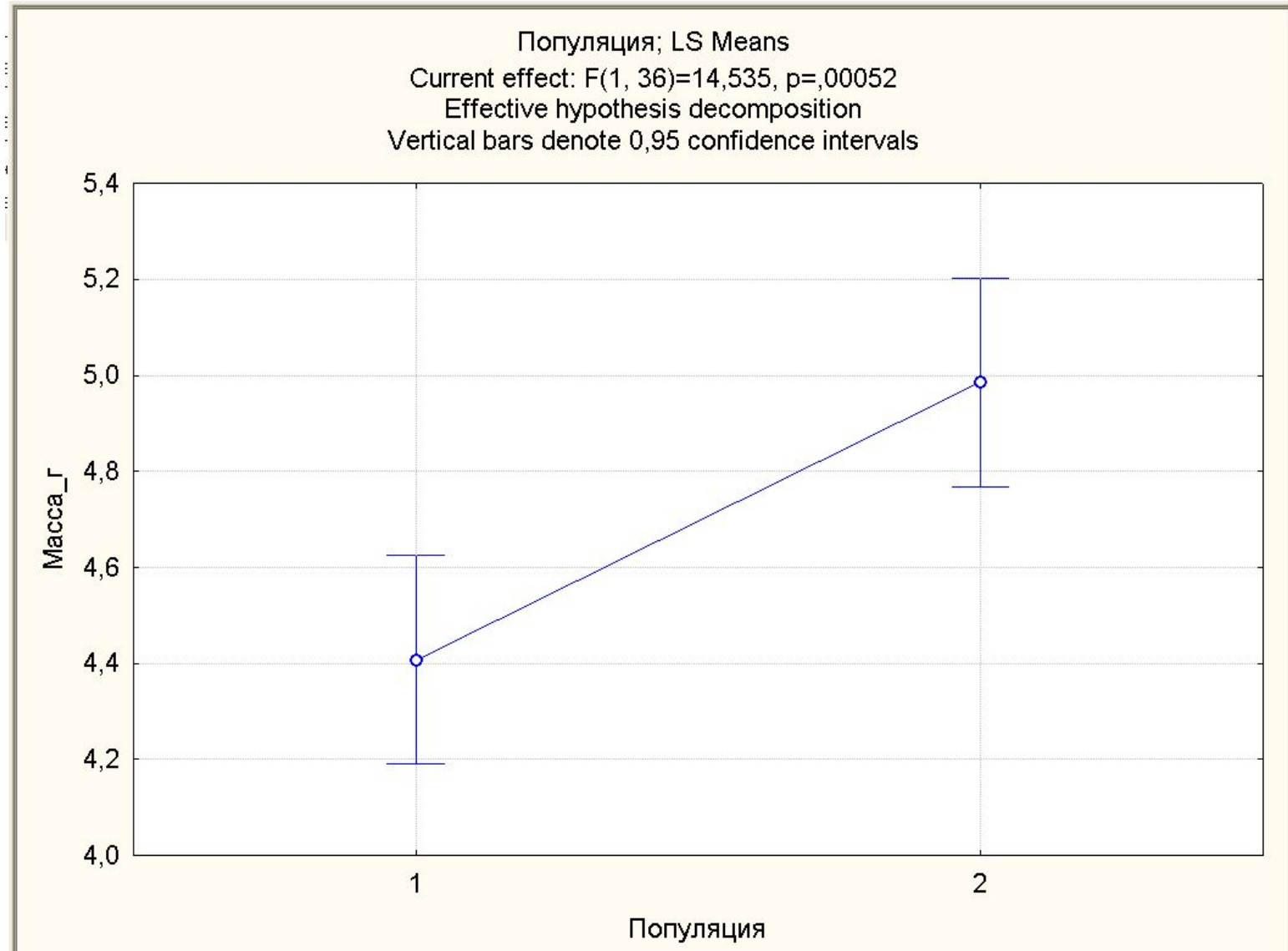
Общие результаты 2-ф ANOVA

Univariate Tests of Significance for Масса_г (Две_популяции) Sigma-ограниченная параметризация Effective hypothesis decomposition						
Эффект	SS	Degr. of Freedom	MS	F	p	
ОТРЕЗОК	873,3859	1	873,3859	3841,085	0,000000	
Популяция	3,3049	1	3,3049	14,535	0,000519	
Пол	44,1222	1	44,1222	194,046	0,000000	
Популяция*Пол	0,1570	1	0,1570	0,690	0,411499	
Ошибка	8,1857	36	0,2274			

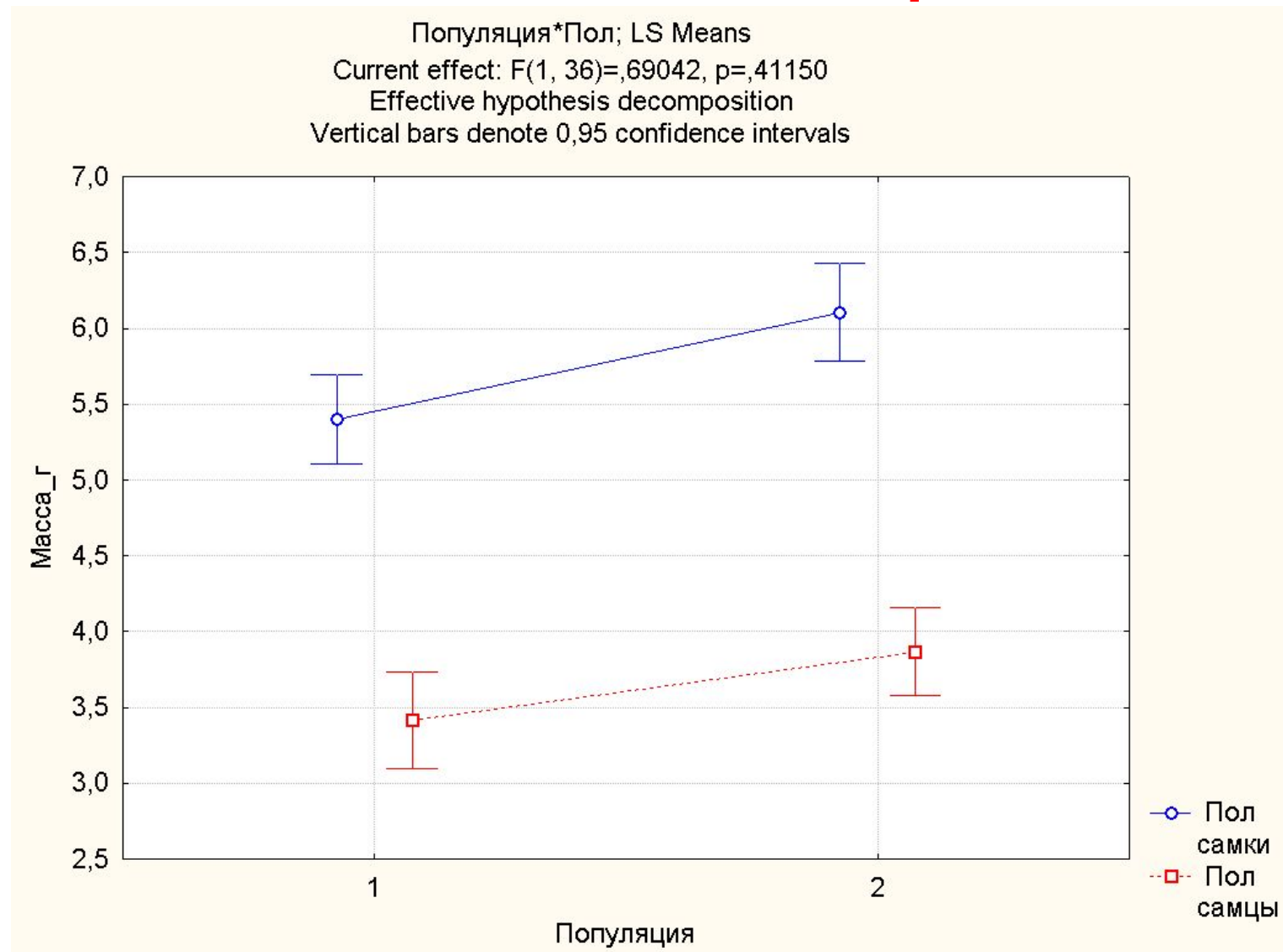
Главные эффекты: половые различия



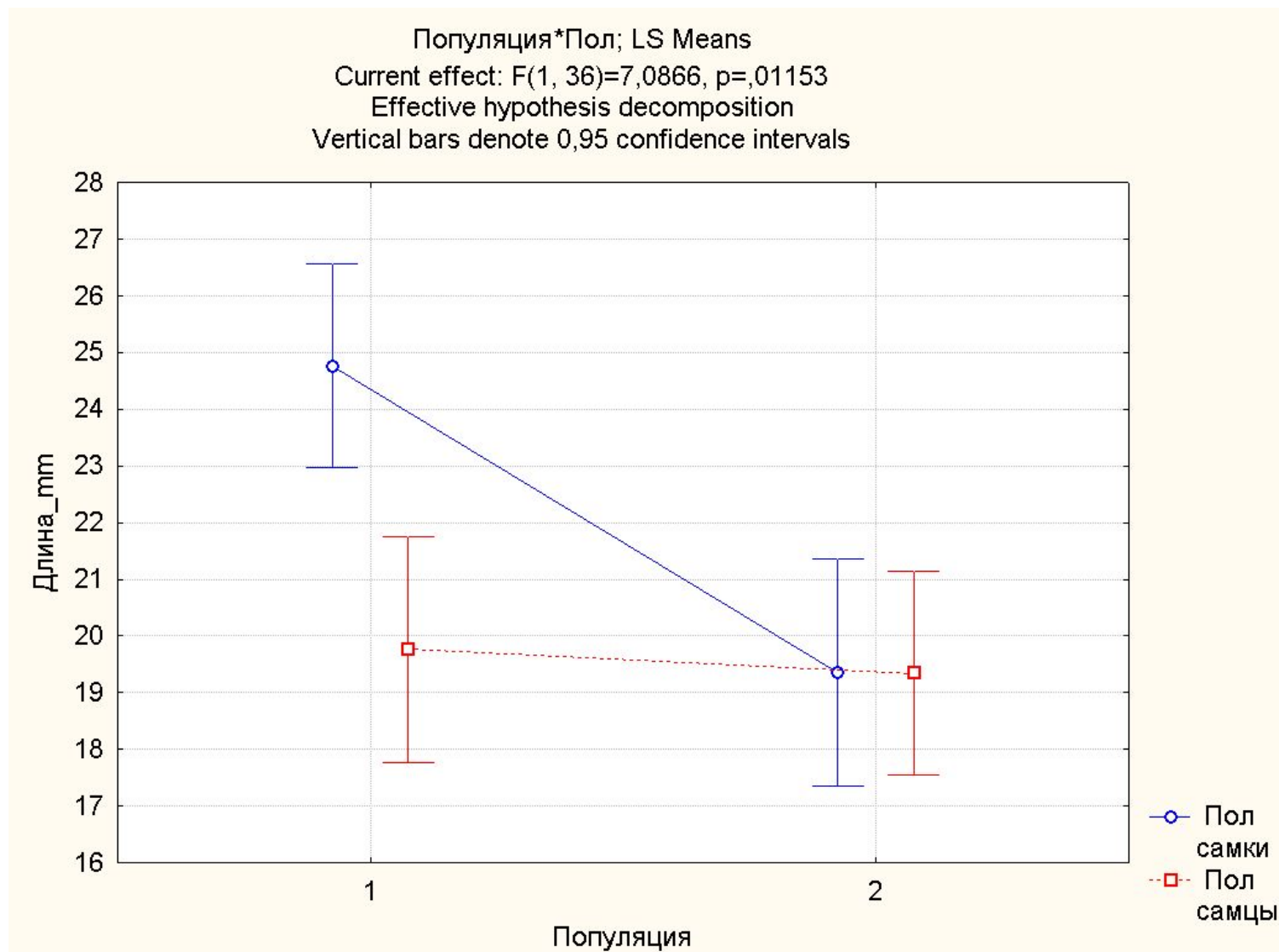
Главные эффекты: популяционные различия



Взаимодействие факторов: одновременная оценка половых и популяционных различий: сопоставление дисперсий



Взаимодействие факторов: одновременная оценка половых и популяционных различий: сопоставление дисперсий



Преимущества ANOVA

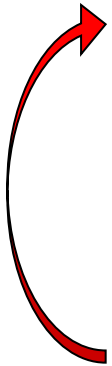
- Возможность сложных, более чем парных сравнений;
- **Возможность сложных, более чем однофакторных сравнений;**
- **Возможность оценки ВЗАИМОДЕЙСТВИЯ между факторами;**
- Устойчивость к малым объемам выборок;

Допущения ANOVA

Требования	Строгость	Примечание
Переменные – в количественных шкалах	Абсолютная	
Объем выборок	Не большая (можно оперировать 3–5 наблюдениями в ячейке плана)	При малых n снижается мощность
Нормальность распределения в выборках	Значительная но не абсолютная	При больших n (>15-25 в группе) нарушениями нормальности можно пренебречь; небольшими нарушениями нормальности можно пренебречь почти всегда.
Однородность дисперсии в сравниваемых группах	Значительная но не абсолютная	ANOVA устойчив относительно небольших нарушений однородности дисперсий
Отсутствие корреляции между средним в группе и дисперсией	Видимо, довольно значительная.	

Оптимальные действия с выборками / переменными для подготовки к анализу ANOVA

- Оптимальное планирование усилий; обеспечение рандомизации; манипулирование объемами наблюдений на этапе сбора данных/группировки;
- Проверка нормальности (хотя бы относительной симметричности распределений);
- Преобразование для сильно отклоняющихся переменных;
- Анализ равенства дисперсий и скоррелированности «средние – дисперсия» в ходе выполнения ANOVA.



**Для сильно строгих пользователей:
если ничто не помогает хотя бы приблизиться к
нормальному распределению
и равенству дисперсий:**

**Критерий Шейрера-Рэя-Хара
(непараметрический аналог
многофакторному ДА)**

!!!

Что можно/нужно проанализировать?

- Интерпретация общего результата;
- Построение графиков средних;
- Проверка предположений;
- Анализ незапланированных сравнений – апостериорные сравнения.

Про ТМ

- **Файл «Пример_тм_токсичность_2ФДА.xls»;**
- **Лист «для_2фДА»;**
- **Проанализируем 2ф ANOVA «длину корня» и «долю некрозов» в зависимости от двух факторов: зоны нагрузки и типа эксперимента; оценим взаимодействие факторов;**

**Неполные планы двухфакторного ДА:
анализ главных эффектов,
но надо быть недоверчивым и
перепроверяться в однофакторных
сравнениях!**

Популяция	Пол	
	Самцы	Самки
Популяция 1	Масса = ?? (n=??)	Масса = ?? (n=??)
Популяция 2	Нет измерений (мы поздно пришли, лёт самок уже закончился)	Масса = ?? (n=??)

Специальные схемы: ДА с повторными измерениями

- Файл «Пример_тм_токсичность_2ФДА.xls»;
- Лист «данные»;
- Проанализируем 2ф ANOVA «длину корня» и «долю некрозов» в двух разных экспериментах как повторные измерения и оценим эффект между ними;

Специальные схемы: **МНОГОМЕРНЫЙ** ДА

- Файл «Пример_тм_токсичность_2ФДА.xls»;
- Лист «данные»;
- Проанализируем **МНОГОМЕРНЫМ** 2ф ANOVA концентрации металлов в пробах в зависимости от зоны нагрузки;

Специальные схемы: **КОВАРИАЦИОННЫЙ** анализ

- Файл «Пример_тм_токсичность_2ФДА.xls»;
- Лист «данные»;
- Проанализируем **КОВАРИАЦИОННЫМ** анализом длину корня в первом эксперименте в зависимости от:
 - 1) «зоны нагрузки» (категориальная переменная) и «рН» (непрерывная переменная) или
 - 2) «зоны нагрузки» (категориальная переменная) и «логарифма токсичности» (непрерывная переменная)

???

- ???