

Statistica

Одновременное сравнение
нескольких выборок.

*Однофакторный ANOVA и критерий
Краскела-Уоллиса*

Анализ единственной выборки

1. Определение шкалы измерения.
2. Проверка «нормальности» распределения для количественных данных / определение типа распределения.
3. Попытки преобразования «ненормально» распределенных переменных / анализ и поиск источников гетерогенности в многомодальных распределениях.
4. Проверка предположения об отличии значения в выборки от какой-либо константы.
5. Подгонка, т.е. определение типа распределения.

Выбор статистического теста при сравнении распределений (сравнении центральных тенденций)

Задача	Количественная шкала, нормальное распределение	Порядковая шкала или отклонение от нормального распределения	Номинальная шкала
Сравнить одну группу с гипотетическим значением	t-тест Стьюдента для одной выборки	Тест Вилкоксона	Тест хи-квадрат
Сравнить две не связанные совокупности	t-тест Стьюдента для не связанных совокупностей	Тест Манна-Уитни	Тест Фишера (тест хи-квадрат)
Сравнить две связанные совокупности	t-тест Стьюдента для связанных совокупностей	Тест Вилкоксона	Тест Мак-Неймера
Сравнить более двух не связанных совокупностей	Однофакторный дисперсионный анализ	Тест Краскела-Уоллиса	Тест хи-квадрат
Сравнить более двух связанных совокупностей	Дисперсионный анализ с повторными измерениями	Тест Фридмана	Тест Кохрана

Почему **НЕЛЬЗЯ** использовать комбинации парных сравнений для случаев множественных сравнений?

Популяция	Популяция 1	Популяция 2	Популяция 3	Популяция 4	Популяция 5
Популяция 1		Нет различий	Есть различия (P=0.0354)	Есть различия (P=0.0154)	Есть различия (P=0.0054)
Популяция 2			Нет различий	Есть различия (P=0.0237)	Есть различия (P=0.058)
Популяция 3				Нет различий	Нет различий
Популяция 4					Нет различий
Популяция 5					

Итого: 10 парных сравнений. На уровне $\alpha=0,05$ воображимо 0,5 случая **АБСОЛЮТНО СЛУЧАЙНЫХ** различий, связанных только с логикой принятия гипотез.

гипотез.

Почему **НЕЛЬЗЯ** использовать комбинации парных сравнений для случаев множественных сравнений?

		Принята гипотеза	
		H_0	H_1
Верна гипотеза	H_0		Ошибка 1 рода: вероятность найти различия, где их нет. Вероятность ошибки первого рода – это уровень значимости (α или P).
	H_1	Ошибка 2 рода: вероятность не увидеть различий, где они есть. Это «близорукость», или «слепота» критерия, вред от неё не очень большой.	

Почему НЕЛЬЗЯ использовать комбинации парных сравнений для случаев множественных сравнений: **разность гипотез**

Популяция	Популяция 1	Популяция 2	Популяция 3	Популяция 4	Популяция 5
Популяция 1		Нет различий	Нет различий	Нет различий	Нет различий
Популяция 2	Нет различий		Нет различий	Нет различий	Нет различий
Популяция 3	Нет различий	Есть различия		Нет различий	Нет различий
Популяция 4	Нет различий	Нет различий	Нет различий		Нет различий
Популяция 5	Нет различий	Нет различий	Нет различий	Нет различий	

На самом деле тут, прежде всего, есть только **ОДНО СРАВНЕНИЕ:**

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$$

H_1 : хотя бы одно равенство не выполняется

Почему НЕЛЬЗЯ использовать комбинации парных сравнений для случаев множественных сравнений: **пример своими руками**

- Создаем в EXCEL (*.xls) файл;
- 15 переменных (var1...var15);
- 25 строк, которые заполняем [=СЛЧИС()];
- Экспортируем в виртуальную машину и в STATISTICA;
- Считаем t-критерием (независимым для переменных) попарные различия всех переменных со всеми;
- Считаем количество значимых ($\alpha=0,05$) различий.

Почему **НЕЛЬЗЯ** использовать комбинации парных сравнений для случаев множественных сравнений?

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»

Идея ANOVA (Analysis of variation)

	Группа 1	Группа 2
Наблюдение 1	2	6
Наблюдение 2	3	7
Наблюдение 3	1	5
Среднее	2	6
Сумма квадратов (<i>SS</i>)	2	2
Общее среднее	4	
Общая сумма квадратов	28	

SS ошибки

~Внутригрупповая дисперсия

Сумма SS

~Общая

Квадрат (*SS*) эффекта

(межгрупповых различий)

$$24$$

$$(28 - (2 + 2))$$

Идея ANOVA (Analysis of variation)

$$SS \rightarrow SS/(n - 1) = MS$$

$$F = \frac{\text{оценка дисперсии **между** группами}}{\text{оценка дисперсии **внутри** групп}}$$

статистика:

$$F = \frac{MS_B}{MS_W}$$

Преимущества ANOVA

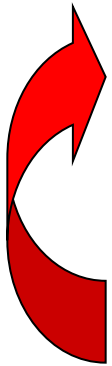
- Возможность сложных, более чем парных сравнений;
- Возможность больших, более чем однофакторных сравнений;
- Возможность оценки ВЗАИМОДЕЙСТВИЯ между факторами;
- Устойчивость к малым объемам выборок;

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.



Где лежит «ANOVA»? «Куда нажимать?»

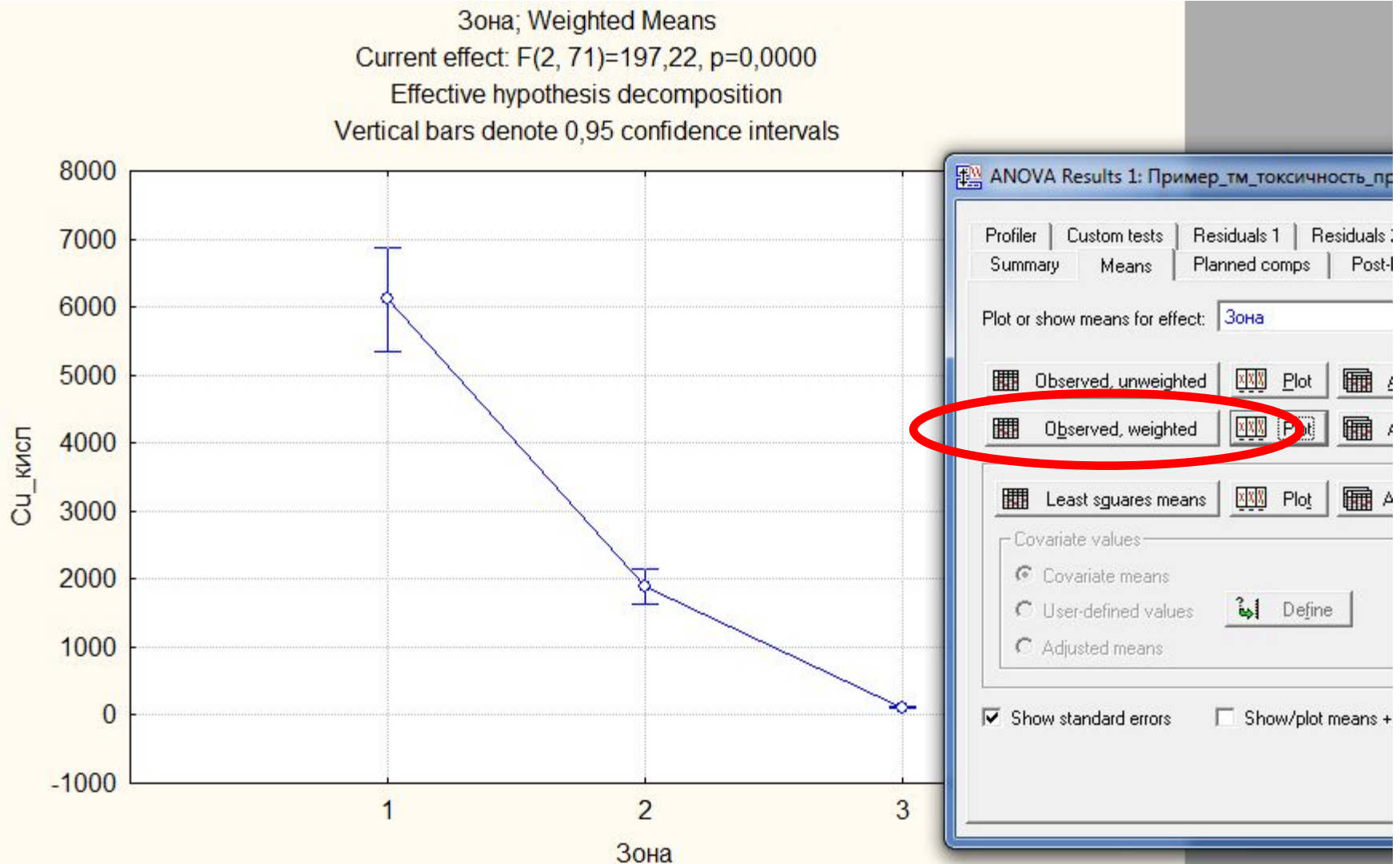
- Пример «Пример_тм_токсичность_проверка.xls»;
- Организация файла;
- Зависимая – независимая переменная;
- Интерпретация общего результата;
- Построение графиков средних;
- Проверка предположений;
- Анализ запланированных контрастов;
- Анализ незапланированных сравнений – апостериорные сравнения;

Интерпретация общего результата

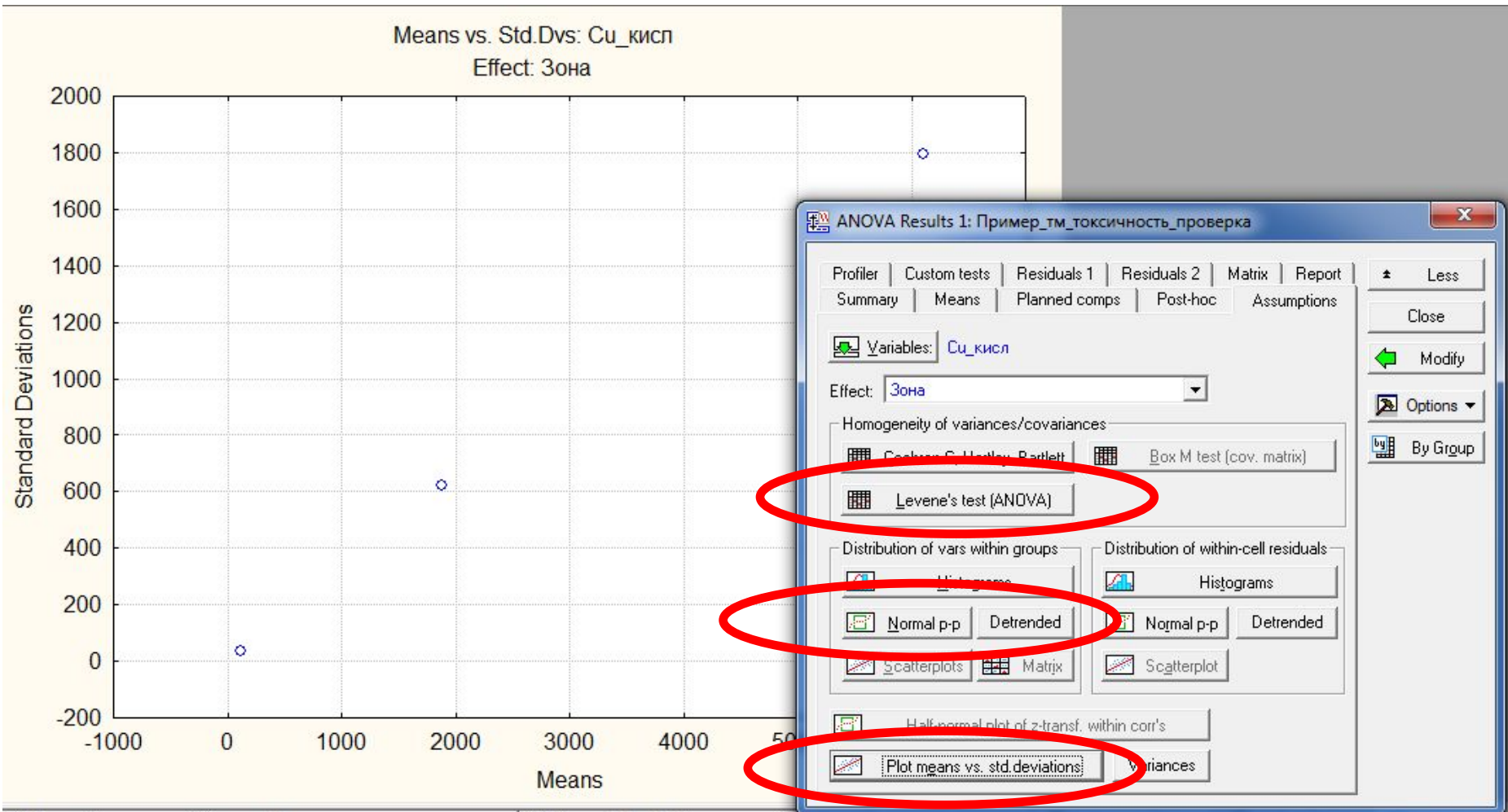
Univariate Tests of Significance for Су_кисл (Пример) Sigma-restricted parameterization Effective hypothesis decomposition					
Effect	SS	Degr. of Freedom	MS	F	p
Intercept	540259964	1	540259964	460,8808	0,00
Зона	462375478	2	231187739	197,2198	0,00
Error	83228588	71	1172234		

Построение графиков средних (**невзвешанных**)

Построение графиков средних (**взвешанных**)

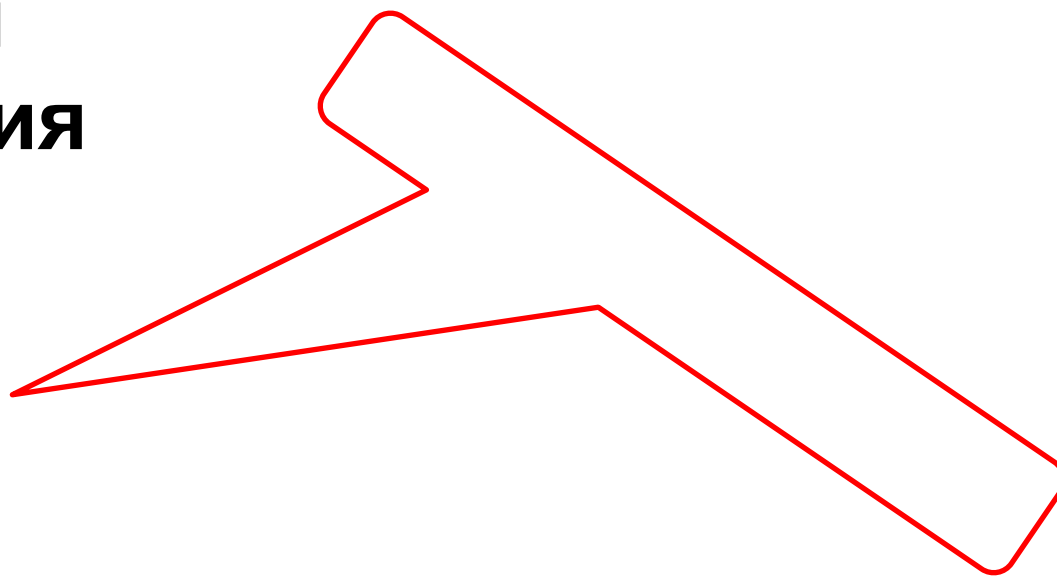


Проверка предположений. Какой путь избрать далее?



Как записать результаты использования ANOVA?

Значени
я
средних



- $F(2, 71) = 1176,00; P << 0,0001$

Степени свободы для «фактора» и
«ошибки»

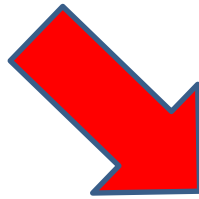
Значение
критерия

Значимос
ть

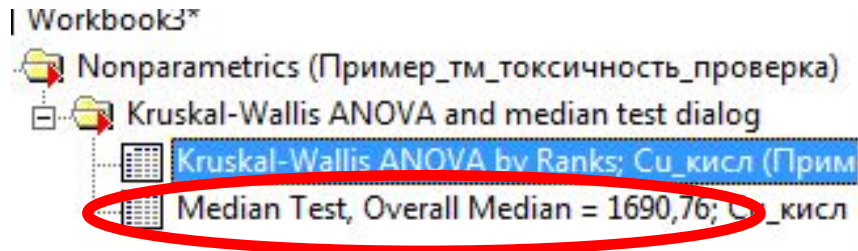
Выбор статистического теста при сравнении распределений (сравнении центральных тенденций)

Задача	Количественная шкала, нормальное распределение	Порядковая шкала или отклонение от нормального распределения	Номинальная шкала
Сравнить одну группу с гипотетическим значением	t-тест Стьюдента для одной выборки	Тест Вилкоксона	Тест хи-квадрат
Сравнить две не связанные совокупности	t-тест Стьюдента для не связанных совокупностей	Тест Манна-Уитни	Тест Фишера (тест хи-квадрат)
Сравнить две связанные совокупности	t-тест Стьюдента для связанных совокупностей	Тест Вилкоксона	Тест Мак-Неймера
Сравнить более двух не связанных совокупностей	Однофакторный дисперсионный анализ	Тест Краскела-Уоллиса	Тест хи-квадрат
Сравнить более двух связанных совокупностей	Дисперсионный анализ с повторными измерениями	Тест Фридмана	Тест Кохрана

**Критерий Краскела-Уоллиса –
непараметрический аналог однофакторного
ANOVA: что такое ранги?**



Критерий Краскела-Уоллиса – непараметрический аналог однофакторного ANOVA



Kruskal-Wallis ANOVA by Ranks: Cu_кисл (Пример_Independent (grouping) variable: Зона
Kruskal-Wallis test: $H(2, N=74) = 64,89189$ $p = ,0000$

Depend.:	Code	Valid N	Sum of Ranks				
Cu_кисл							
1	1	24	1500,000				
2	2	25	950,000				
3	3	25	325,000				

Медианный тест
- грубая версия
Краскела-
Уоллиса.

Удельная активность ^{90}Sr в костной ткани *Apodemus uralensis*

Сравнение трех выборок
(критерий Краскела-Уолиса) $H(2, N=55)=8.5$, $p=0.01$

Попарное сравнение выборок
(критерий Манна-Уитни):

Линия 1 – линия 2: $p=0.04$

Линия 1 – линия 3: $p=0.01$

Линия 2 – линия 3: $p=0.74$

N=26

N=13

N=16

Удельная активность ^{90}Sr в костной
ткани малой лесной мыши составляет
 29 ± 2 Бк/г (указано среднее значение и
стандартная ошибка средней, N=55)

Удельная активность ^{90}Sr в костной ткани *Microtus oeconomus*

Сравнение четырех выборок (линии 1- 4) (критерий Краскела-Уолиса) $H(3, N=64)=4.5, p=0.21$.

Попарное сравнение выборок (тест Манна-Уитни) также не выявило статистически значимых ($p<0.05$) различий между ними.

Различия между выборками с линий 1-4 и 5 не значимы (тест Манна-Уитни) $p=0.13$, что может быть связано с малым размером выборки с линии 5.

N = 23 9 16 16 4

