

Смешанная модель дисперсионного анализа в пакете Statistica на примере данных Sergey.Pustylnikov (черновик)

Нохрин Д.Ю. (*nokh на forum.disser.ru*)

Описание данных – см.

<http://forum.disser.ru/index.php?showtopic=2518&pid=10184&st=0&#entry10184>

сами данные продублированы в Приложении.

«Мне нужно ответить на вопросы:

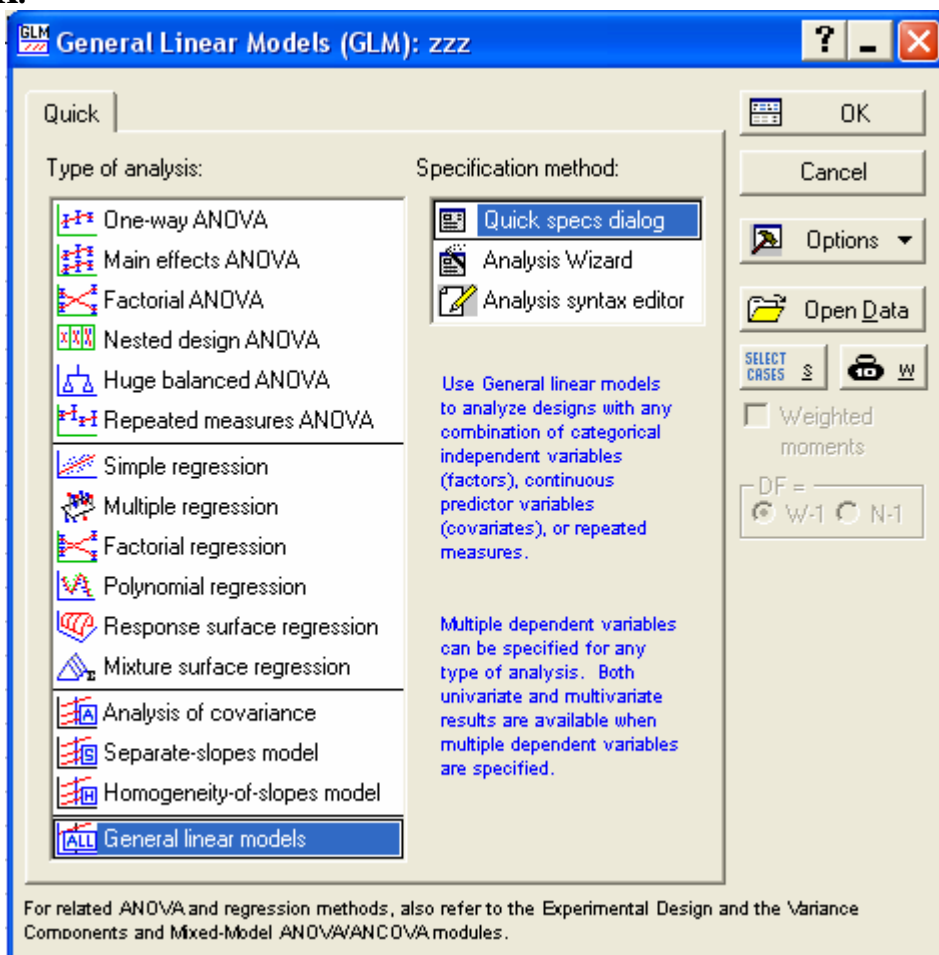
- а) При каждой конкретной концентрации есть ли достоверное отличие между опытом (2 повтора по 3 лунки) и контролем (2 повтора по 3 лунки)? Каков уровень стат. значимости?
- б) При каждой конкретной концентрации есть ли достоверное отличие между опытом с веществом 1 и веществом 2 (2 повтора по 3 лунки)? Каков уровень стат. значимости?
- в) В каких случаях (при каких концентрациях и каких веществах) было достоверным (и было ли) влияние дополнительного фактора? Каков уровень стат. значимости?»

Этап 1.

Дисперсионный анализ в черновом варианте

Имеем четырёхфакторный дисперсионный анализ. Факторы: Дополнительный фактор, Вещество и Концентрация – фиксированные (контролируются исследователем). Фактор Повторность – случайный, т.к. проведённые две повторности - лишь одни из бесконечного множества возможных повторных реализаций эксперимента. Модель дисперсионного анализа (ДА) сочетающая фиксированные и случайные эффекты называется смешанной (Mixed design ANOVA). Для учёта случайного фактора анализ проводится в модуле общих линейных моделей - General Linear Models (GLM).

Statistics – Advancad Linear/Nonlinear Models – General Linear Models – General Linear Models – ОК.



Форма GLM

Закладка Quick. Variables. Dependent: Data, Categorical predictor: Repl (Повторность), Factor (Дополнительный фактор), Subs (Вещество), Conc (Концентрация).

!!! Закладка Options. **Random Factors: Repl**, Sum of squares: Type III (т.к. дисперсионный комплекс равномерный), Галочка в No intercept (Чтобы ненужная информация не выдавалась) – ОК.

Форма GLM Results

Закладка Quick. All effects. Получаем таблицу черного варианта дисперсионного анализа. Как к черновому к нему следует относиться, т.к., возможно, для окончательного варианта потребуется дополнительно преобразовывать данные. Выделяем статистически значимые эффекты.

Таблица 1. Результаты черного дисперсионного анализа всего массива данных

Источник изменчивости	Effect	SS	df	MS	Den.Syn. Error df	Den.Syn. Error MS	F	p
Главные эффекты:								
Repl	Random	0,004108	1	0,004108	2,381429	0,036802	0,1116	0,765488
Factor	Fixed	0,026120	1	0,026120	1,000000	0,018572	1,4064	0,445987
Subs	Fixed	0,395476	1	0,395476	1,000000	0,012418	31,8457	0,111653
Conc	Fixed	5,299404	6	0,883234	6,000000	0,007739	114,1245	0,000006
Двухфакторные взаимодействия:								
Repl*Factor	Random	0,018572	1	0,018572	0,000000			
Repl*Subs	Random	0,012418	1	0,012418	0,000000			
Factor*Subs	Fixed	0,033566	1	0,033566	1,000000	0,000440	76,3647	0,072535
Repl*Conc	Random	0,046435	6	0,007739	0,037290	0,001488	5,1998	0,906846
Factor*Conc	Fixed	0,113573	6	0,018929	6,000000	0,004680	4,0445	0,056575
Subs*Conc	Fixed	1,638478	6	0,273080	6,000000	0,011238	24,2988	0,000582
Трёхфакторные взаимодействия:								
Repl*Factor*Subs	Random	0,000440	1	0,000440	6,000000	0,014430	0,0305	0,867189
Repl*Factor*Conc	Random	0,028081	6	0,004680	6,000000	0,014430	0,3243	0,901788
Repl*Subs*Conc	Random	0,067430	6	0,011238	6,000000	0,014430	0,7788	0,615379
Factor*Subs*Conc	Fixed	0,022261	6	0,003710	6,000000	0,014430	0,2571	0,938543
Четырёхфакторное взаимодействие:								
Repl*Factor*Subs*Conc	Random	0,086581	6	0,014430	0,000000	0,003164	4,5607	0,000355
Ошибка		0,354413	112	0,003164				

Значимы: различия между концентрациями, двухфакторное взаимодействие «Вещество x Концентрация» и четырёхфакторное взаимодействие. Также проявились тенденция к взаимодействиям «Фактор x Вещество» и «Фактор x Концентрация». Пустые ячейки – эффекты, статистическую значимость которых не удалось оценить, т.е. её нельзя оценить ни напрямую, ни косвенно - через декомпозицию сумм квадратов по формуле Сэттерсвейта (см. Монтгомери).

Четырёхфакторное взаимодействие - проблема, т.к. оно указывает на то, что мы, строго говоря, не вправе рассматривать эффекты меньшего порядка изолированно друг от друга, т.е. «всё связано со всем». Сами по себе взаимодействия интересны, т.к. позволяют выявлять не только различия между переменными, но и влияние на эти различия других факторов. Однако с методологической точки зрения взаимодействие высокого порядка – плохо, т.к. указывает на плохую контролируемость условий эксперимента и, соответственно, на сложность прогнозирования результата. Поступить в этой ситуации можно по-разному; я обозначу тот путь, как вёл бы анализ сам в данном конкретном случае.

Взаимодействие факторов может быть артефактом неправильного выбора шкалы и в этом случае прибегают к преобразованию исходных данных. Различают 3 типа

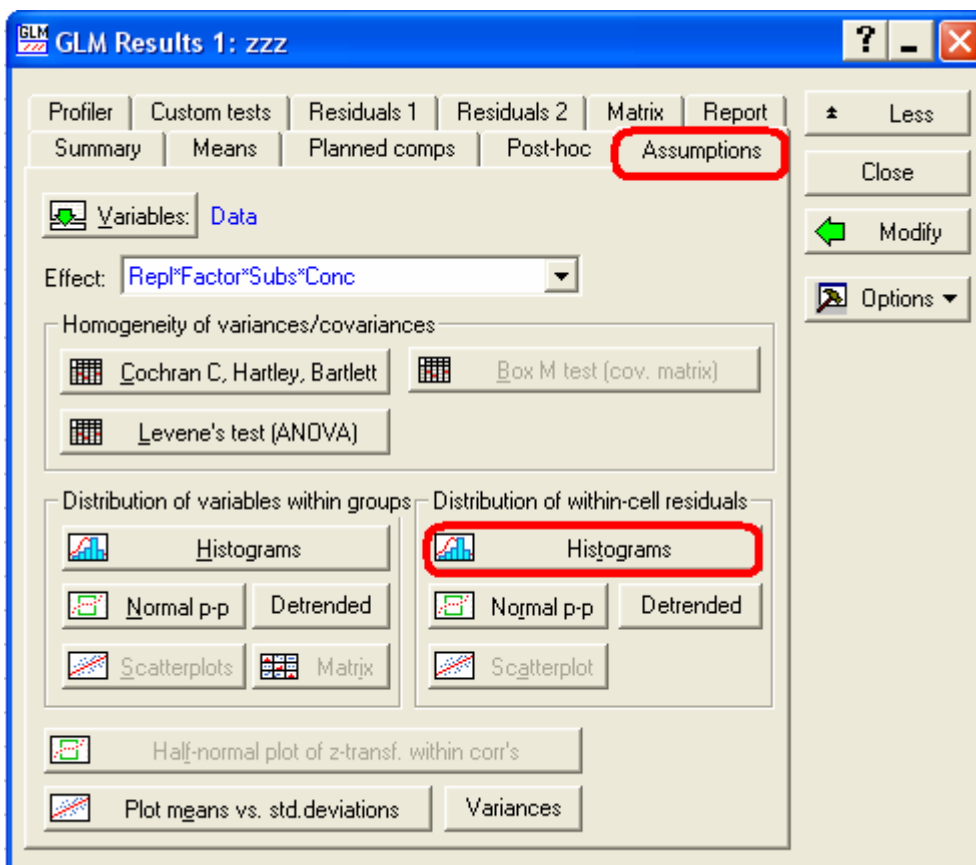
преобразований: (1) ведущие к нормализации распределения ошибки, (2) ведущие к устранению неоднородности дисперсий (гетероскедастичности) и (3) ведущие к аддитивности. В нашем случае нет именно аддитивности, т.е. при действии факторов их совместный эффект не суммируется, а проявляется мультипликативно, нелинейно, что и составляет взаимодействием факторов. Часто выбор адекватного преобразования решает сразу все три проблемы, а наиболее частой причиной проблем является асимметричность распределения показателя. Поэтому первым шагом после чернового анализа будет проверка нормальности распределения ошибки анализа. Заодно полезно оценить и выполнение второго требования дисперсионного анализа (ДА) – однородность дисперсий в ячейках комплекса. Следует сразу отметить, что многофакторные модели ДА достаточно устойчивы к отклонениям от данных требований, т.е. робастны. Тем не менее, мы должны постараться сделать всё возможное, чтобы приблизить данные к требованиям модели.

Этап 2.

Проверка выполнения предположений (assumptions) о характере данных, лежащих в основе дисперсионного анализа

Форма GLM Results

Закладка Resids (сокращение от residuals – остатки). Остатки – то, что остаётся от данных после снятия всех эффектов главных факторов и их взаимодействий, т.е. ошибка в модели ДА. Можно кое-что посмотреть здесь. Кнопка «More results» Закладка Asumptiouns. Distribution of within-cell residuals. Histograms.



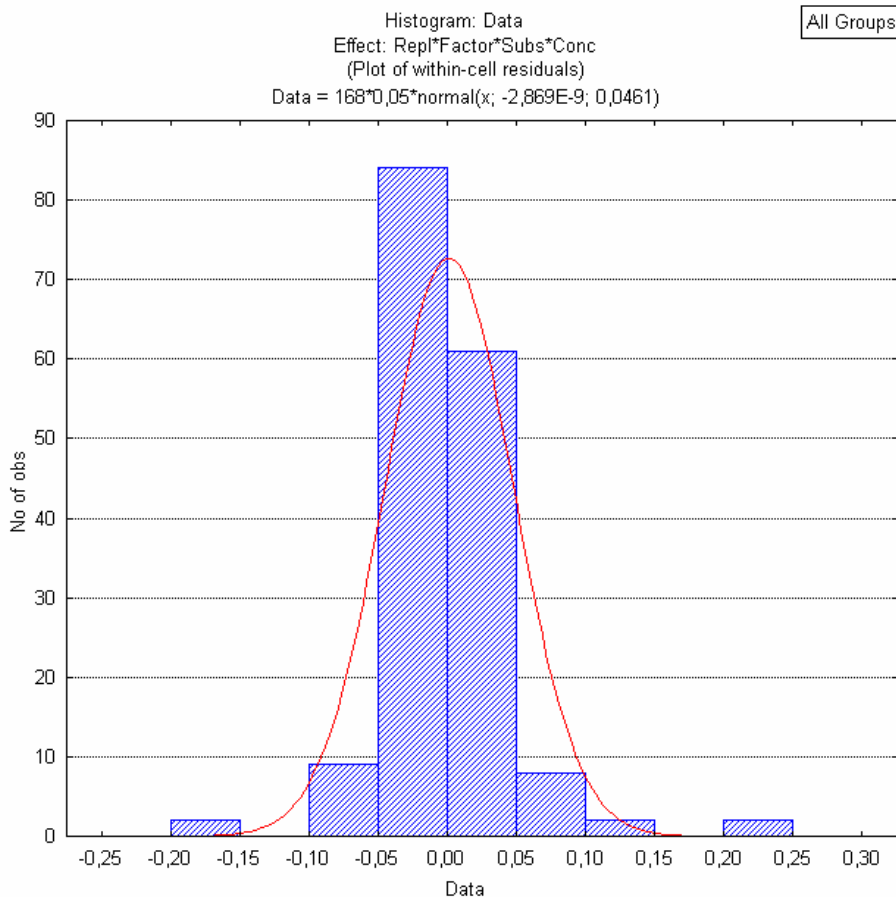


Рис. 1. Гистограмма распределения остатков

Распределение выглядит более-менее симметричным, следовательно исходные данные не нуждаются в преобразовании. Тем не менее, избыток значений в центральной части распределения указывает на наличие положительного эксцесса, что лучше чем асимметрия, но с чем сложнее бороться. Вероятнее всего эксцесс здесь возникает в результате смеси двух близких к нормальному распределений с близкими средними значениями (около нуля), но с различной дисперсией. В принципе, с причинами эксцесса можно разбираться. Например, можно предположить, что эти два распределения ошибок порождены различиями в точности первой и второй повторностей эксперимента. Тогда если провести аналогичные дисперсионные анализы отдельно для каждой повторности распределения ошибки в каждой из них должно быть нормальным. Или два распределения порождает действие дополнительного фактора, тогда нужно провести два анализа только для каждой градации этого фактора в отдельности. Мне это делать лень, к тому же возможны и другие причины такого отклонения от нормальности - здесь нужны частные знания о конкретном методе исследования. Важно, что данные не требуют преобразования и что четырёхфакторное взаимодействие скорее всего реально, а не является артефактом эффекта шкалы.

В данном модуле отсутствуют опции для проверки распределения остатков должным образом. Поэтому лучше сохранить остатки отдельно и исследовать их адекватными методами. Кнопка «Predicted and residuls». Копируем колонку Data resides и сохраняем в отдельный файл – по умолчанию просто Var1. Работаем с этим файлом.

Statistics – Basic statistics/Tables – Descriptive statistics. Variables: Var1. Закладка Normality. Близкое к оптимальному количество интервалов (Number of intervals) определяем по эмпирической формуле Sturges (Закс, 1976. Статистическое оценивание):

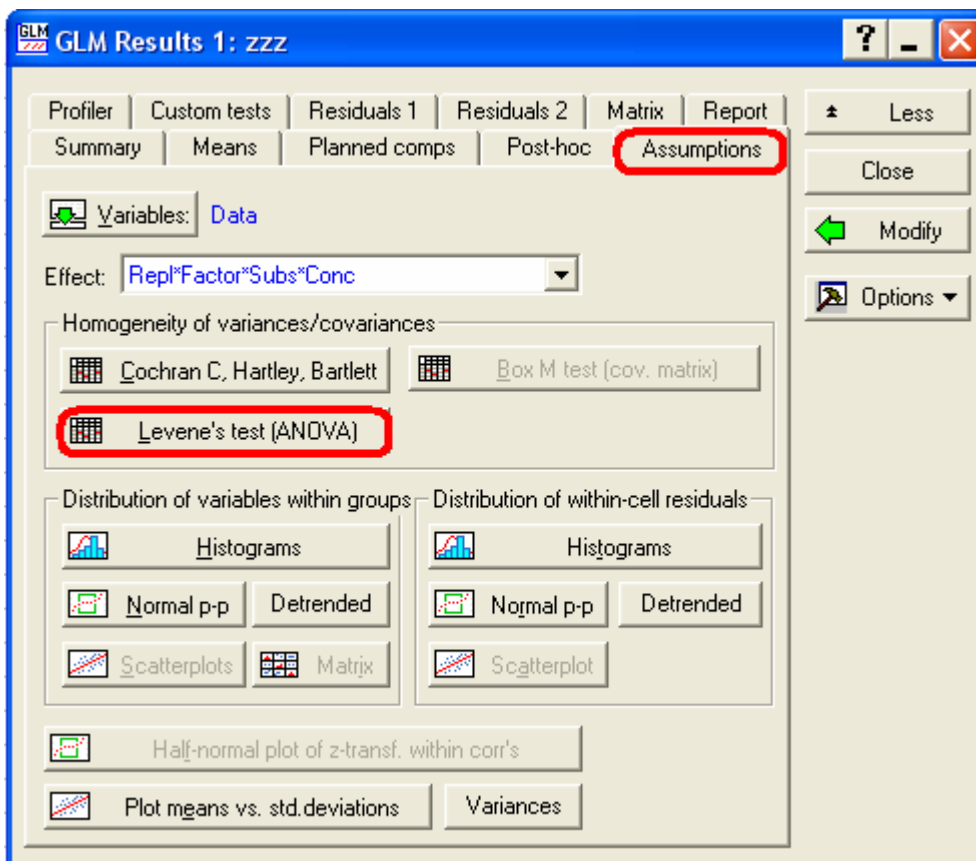
$$k \approx 1 + 3,32 \lg n, \text{ где } \lg n \text{ – десятичный логарифм числа наблюдений. В нашем случае } k \approx 1 + 3,32 \lg 168 \approx 8,39 \text{ т.е. 8-9 классов, лучше оставим нечётное 9.}$$

Ставим галочку в Shapiro-Wilk W test. OK.

Результат: K-S d=,11303, p<,05 ; Lilliefors p<,01; **Shapiro-Wilk W=,86706, p=,00000**

Наиболее мощный тест на нормальность из доступных трёх – Шапиро-Уилка, смотрим на него в первую очередь. Если есть желание убедиться, что ненормальность – результат именно эксцесса – здесь же в Descriptive statistics закладка Advanced и галочки в эксцессе и его стандартной ошибке (Kurtosis и Std.er., Kurtosis). Рассчитать значимость придётся вручную по формулам из учебников.

Разобравшись с нормальностью/ненормальностью проверяем второе требование метода ДА – однородность дисперсий. Возвращаемся к таблице результатов в Assumptions (Форма GLM Results, Кнопка «More results», Закладка Assumptions). Тесты Кохрена, Хартли и Бартлетта уступают в мощности тесту Ливена, смотрим только его:



Тест Ливена на однородность дисперсий:

	MS	MS	F	p
Data	0,002588	0,000533	4,853047	0,000000

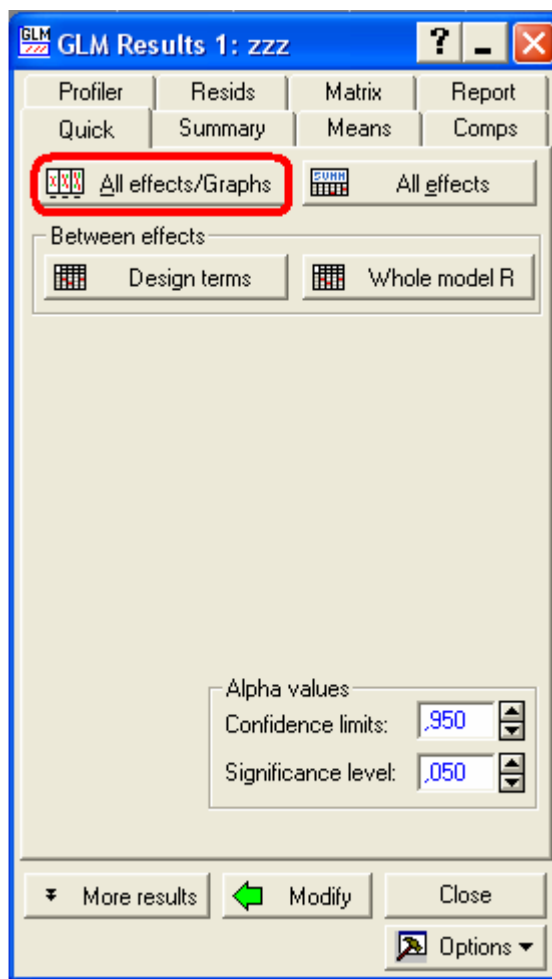
Налицо высоко статистически значимая неоднородность дисперсий в ячейках дисперсионного комплекса (гетероскедастичность).

Вывод по этапу 2. Требования дисперсионного анализа нарушены. Возможно они нарушаются вследствие одного и того же явления – смеси двух распределений ошибок с нулевым средним, но существенно разными дисперсиями. К результатам анализа для эффектов «на грани статистической значимости» следует относиться с осторожностью, а для множественных межгрупповых сравнений в рамках дисперсионного комплекса – предпочесть процедуру Games & Howell. С другой стороны, исходные данные не нуждаются в преобразовании и «черновой» варианта анализа будем рассматривать как «чистовой».

Этап 3.

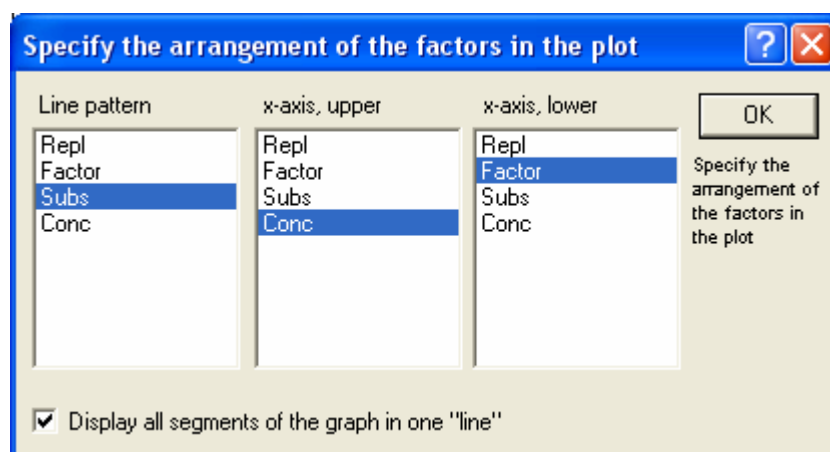
Исследуем статистически значимое взаимодействие самого высокого порядка

Т.к. четырёхфакторное взаимодействие оказалось значимым и мы не сможем от него избавиться с помощью преобразований – придётся исследовать в чём оно заключается.

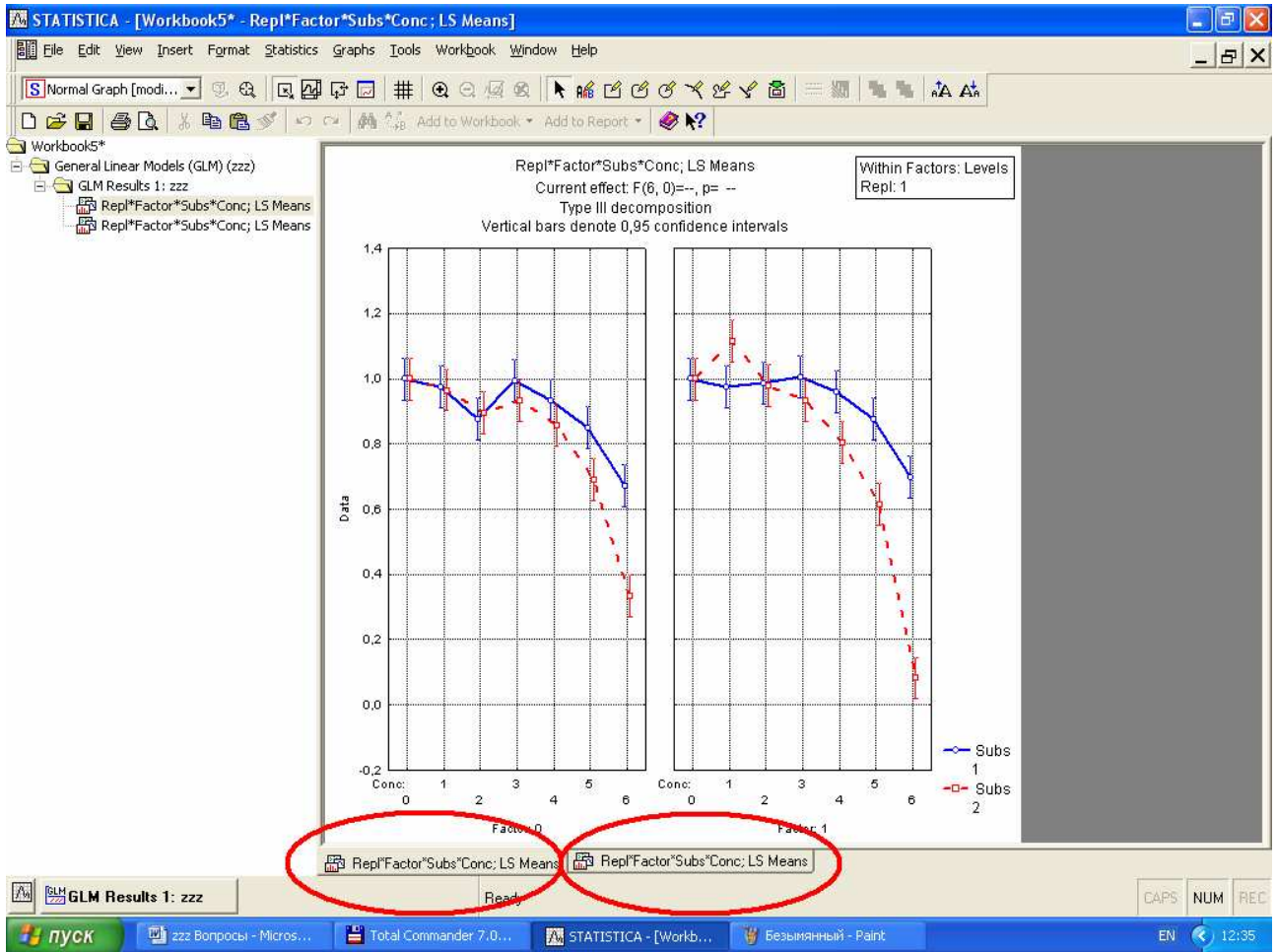


Выбираем четырехфакторное взаимодействие (самое нижнее) – ОК.

Теперь нужно верно расставить выделения, чтобы рисунки были наглядными (можно методом проб-ошибок):

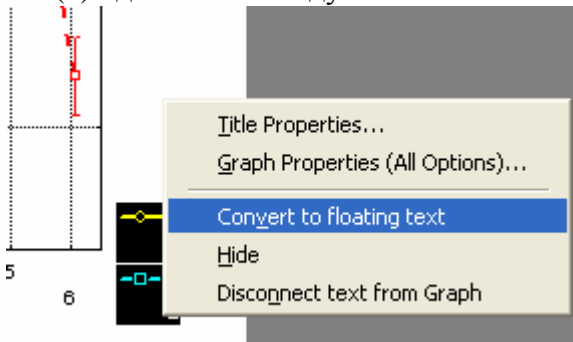


При таком расположении графики для фактора Repl будут выданы на отдельных закладках:

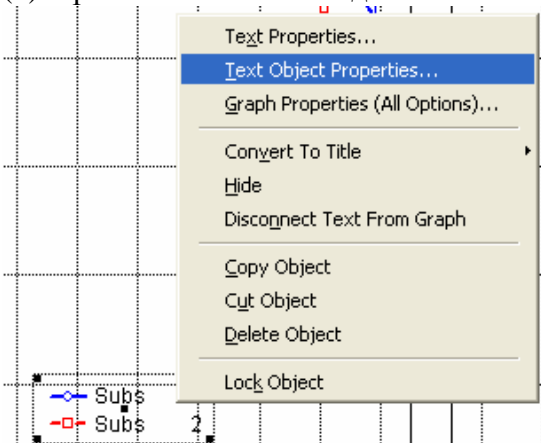


Доводим рисунки до приемлемого вида:

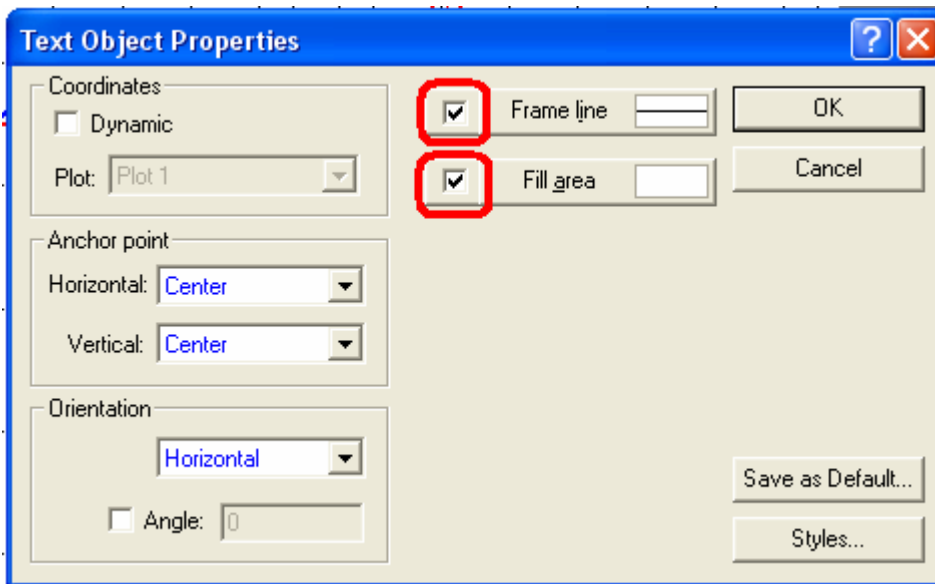
(1) сдвигаем легенду:



(2) Правый клик на легенде:



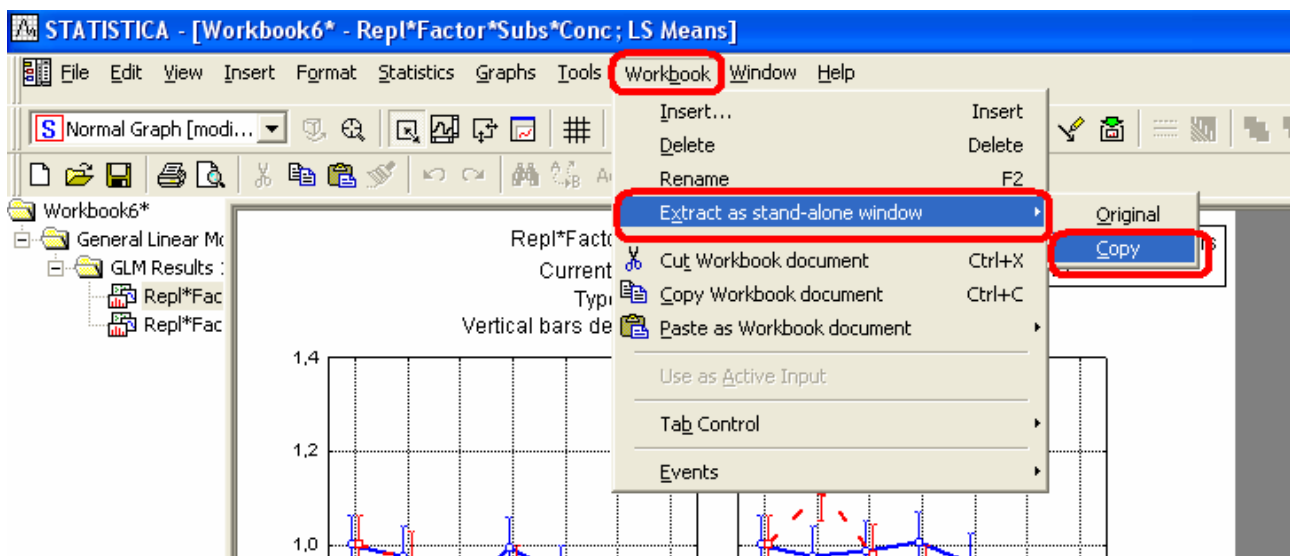
(3) Ставим галочки



(4) Удаляем лишний текст, переводим названия осей, изменяем минимум-максимум шкалы у, изменяем размеры рисунков и т.д. и т.п. («методом тыка» или см. книгу Боровикова).

Сохраняем оба рисунка в форматах: **stg** (векторный, для возможности правки в будущем) и **bmp** (растровый, для вставки в текст или презентацию).

!!! Для этого их нужно предварительно вырезать из рабочего пространства:



Располагаем оба рисунка на одной странице, РАСПЕЧАТЫВАЕМ её и анализируем графически:

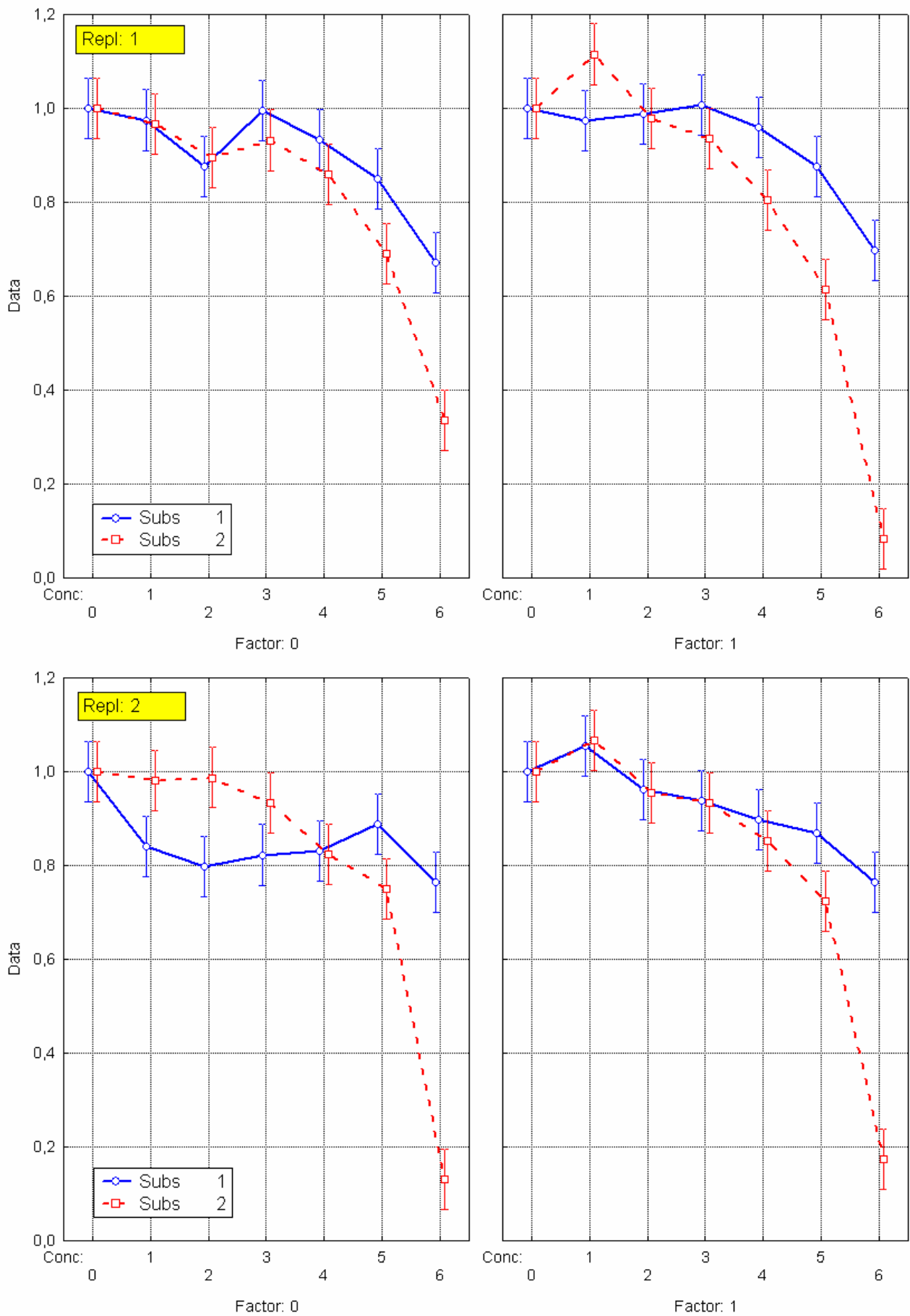


Рис. 2. Четырёхфакторное взаимодействие

Грамотная интерпретация взаимодействий столь высоких порядков может представлять сложности для новичка. Нужно за что-то «зацепиться». Полагаю, что в данном случае такой зацепкой является Дополнительный фактор. Как видно из результатов таблицы 1, он дважды проявил себя тенденцией к статистической значимости во взаимодействиях «Фактор x Вещество» и «Фактор x Концентрация», но «не дотянул» до значимости. Однако оба эти взаимодействия соединившись в эффектом повторности стали высоко статистически значимыми. Т.е. ключевым в этом четырёхфакторном взаимодействии является именно дополнительный фактор. Трактовать его вмешательство должен специалист по данному методу исследования. Как неспециалист могу только заметить, что результаты анализа в присутствии доп. фактора выглядят более стабильными (насколько это можно заключить только по двум повторностям).

Вывод по этапу 3 («черновой», вывод неспециалиста). В отсутствие действия дополнительного фактора динамика изменения концентраций веществ 1 и 2 изменяется в повторностях менее предсказуемо, чем в его присутствии.

Будем считать, что с дополнительным фактором и четырёхфакторным взаимодействием разобрались. Теперь нужно принять внестатистическое решение: проводить интересующие сравнения по данному варианту анализа или исключить из анализа вариант с отсутствием дополнительного фактора и провести аналогичный трёхфакторный ДА? Я бы пошёл вторым путём, т.е. разделил бы исследование на 2 части: методологическую и собственно исследовательскую. В методологической части описал бы всё что описал выше, а в окончательный анализ включил бы только вариант с дополнительным фактором. Но у специалиста могут быть иные соображения + демонстрация в учебных целях проведения запланированных сравнений одинакова + для скорости написания этого текста: сделаем интересующие сравнения в контексте данного анализа.

Этап 4.

Межгрупповые сравнения внутри дисперсионного комплекса

Тема множественных сравнений обширна: масса методов и литературы. В целом, смотреть различия между интересующими ячейками дисперсионного комплекса или их группами можно двумя способами:

- (1) Запланированные ортогональные сравнения (Planned Comparisons)
- (2) Запланированные неортогональные или незапланированные сравнения (Post-hoc Comparisons).

Первый способ наиболее грамотный, т.к. во-первых, у исследователя всегда есть ключевые вопросы заранее, ещё до проведения эксперимента (см. выделение зелёным в самом начале текста), а во-вторых такой анализ является более мощным, т.е. возможны ситуации, когда интересующий эффект в запланированных сравнениях оказывается статистически значимым, а в Post-hoc – нет. Поэтому всегда следует пытаться в первую очередь использовать Запланированные сравнения. Доступно прочитать о методике составления контрастов и их проверке на ортогональность можно в Закс Л. Статистическое оценивание (1976). Т.к. теме контрастов в отечественной литературе вообще не уделяется внимания, а она важнее Post-hoc – рассмотрим её подробнее.

Согласно рекомендациям Sokal & Rohlf (Biometry..., 1995) выбор способа сравнений средних в ячейках дисперсионного комплекса зависит от числа групп a и числа сравнений k . (см. Табл. 2). В многофакторных экспериментах число групп всегда велико, поэтому под a будем рассматривать число групп для фактора с наибольшим числом градаций – т.е. Концентрацию ($a=6$).

Таблица 2. Рекомендации по межгрупповым сравнениям внутри дисперсионного комплекса (Sokal & Rohlf, Biometry..., 1995)

Тип сравнений	Число сравнений k	Метод	Условия
Запланированные ортогональные	$k \leq (a - 1)$	Ортогональное разложение MS между группами	
Запланированные неортогональные и незапланированные	$k < \frac{a(a-1)}{2}$	метод Bonferroni метод Dunn-Šidák последовательная техника Bonferroni методом Dunn-Šidák 9(менее консервативная)	
	Все возможные сравнения: $k = \frac{a(a-1)}{2}$	метод T метод T' метод Hochberg's GT2 метод Tukey-Kramer метод Welsch	Равные выборки Примерно равные Неравные выборки Неравные выборки Равные выборки
	Все возможные контрасты: $k > \frac{a(a-1)}{2}$	метод Scheffé метод T метод Hochberg's GT2	Малые a и df (Примерно) равные Неравные выборки
	Все возможные ячейки комплекса: $k ? \frac{a(a-1)}{2}$	метод SS-STP (Simultaneous Sums of Squares Test Procedure)	
В случае неравных дисперсий используется метод Games and Howell			

Следует заметить, что ориентировочно и для скорости интерпретировать результаты ДА можно и по графикам интересующих эффектов и их взаимодействий. При этом если 95%-ные доверительные интервалы не перекрываются – межгрупповые различия считают статистически значимыми. Однако такой способ обладает минимальной мощностью – даже меньше чем метод SS-STP, т.к при построении доверительных интервалов для среднего какой-либо ячейки используется информация только о наблюдениях внутри этой ячейки, а не всего дисперсионного комплекса.

Тем не менее начинать анализ нужно всё равно с графика. То, как первоначальные вопросы сформулированы исследователем (см. выделение зелёным в начале текста) и с учётом статистической значимости взаимодействия «Концентрация x Вещество», требует рассмотрения именно этого двухфакторного взаимодействия (считаем, что с дополнительным фактором разобрались в методологической части исследования). Поэтому строим рисунок для него. Но если бы статистически значимым было не данное как раз интересующее нас взаимодействие, а скажем только эффект концентрации, то рисунок следовало бы строить для него, признав отсутствие статистически значимых различий между веществами.

Итак - график:

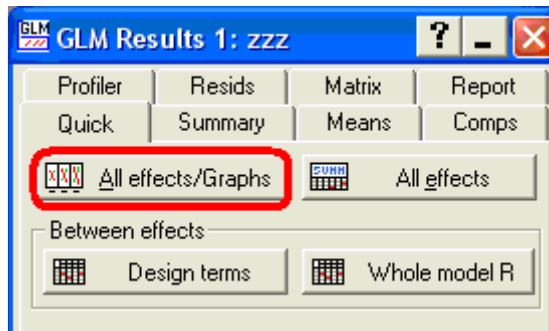


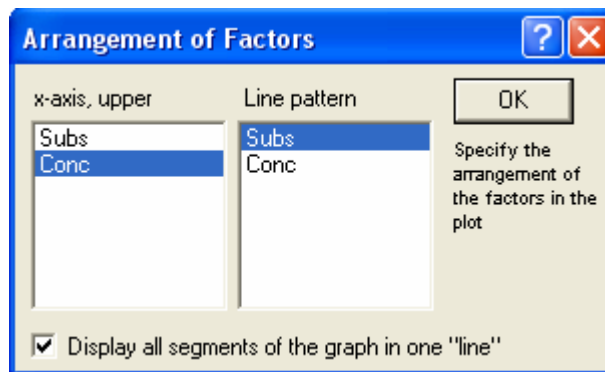
Table of All Effects: zzz

Over-parameterized model
Type III decomposition

Effect	Deg...	MS	Den.Syn...	Den.Sy...	F	p
Repl	1	,004	2,381	,037	,1	,765
Factor	1	,026	1,000	,019	1,4	,446
Subs	1	,395	1,000	,012	31,8	,112
Conc	6	,883	6,000	,008	114,1	,000*
Repl*Factor	1	,019	0,000	---	---	---
Repl*Subs	1	,012	0,000	---	---	---
Factor*Subs	1	,034	1,000	,000	76,4	,073
Repl*Conc	6	,008	,037	,001	5,2	,907
Factor*Conc	6	,019	6,000	,005	4,0	,057
Subs*Conc	6	,273	6,000	,011	24,3	,001*
Repl*Factor*Subs	1	,000	6,000	,014	,0	,867
Repl*Factor*Conc	6	,005	6,000	,014	,3	,902
Repl*Subs*Conc	6	,011	6,000	,014	,8	,615
Factor*Subs*Conc	6	,004	6,000	,014	,3	,939

Double-click on an effect to produce a graph or a Spreadsheet of means.

Options: Close dialog on OK, Graph, Spreadsheet, Means: Unweighted, Weighted, Least squares, Compute standard errors.



Результат после всех доводок (сохраняем рисунок в форматах stg и bmp):

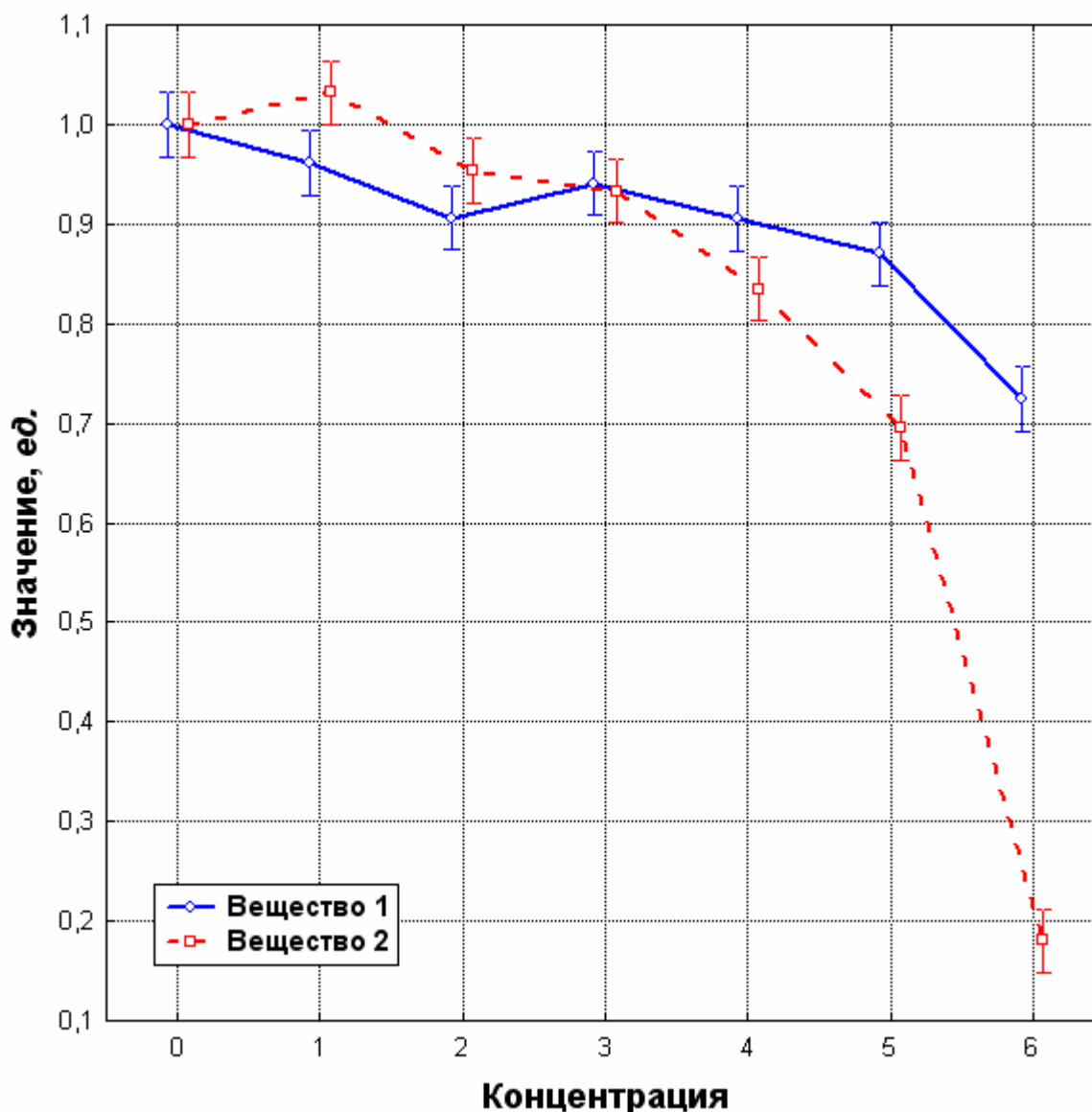


Рис. 3. Двухфакторное взаимодействие Концентрация x Вещество

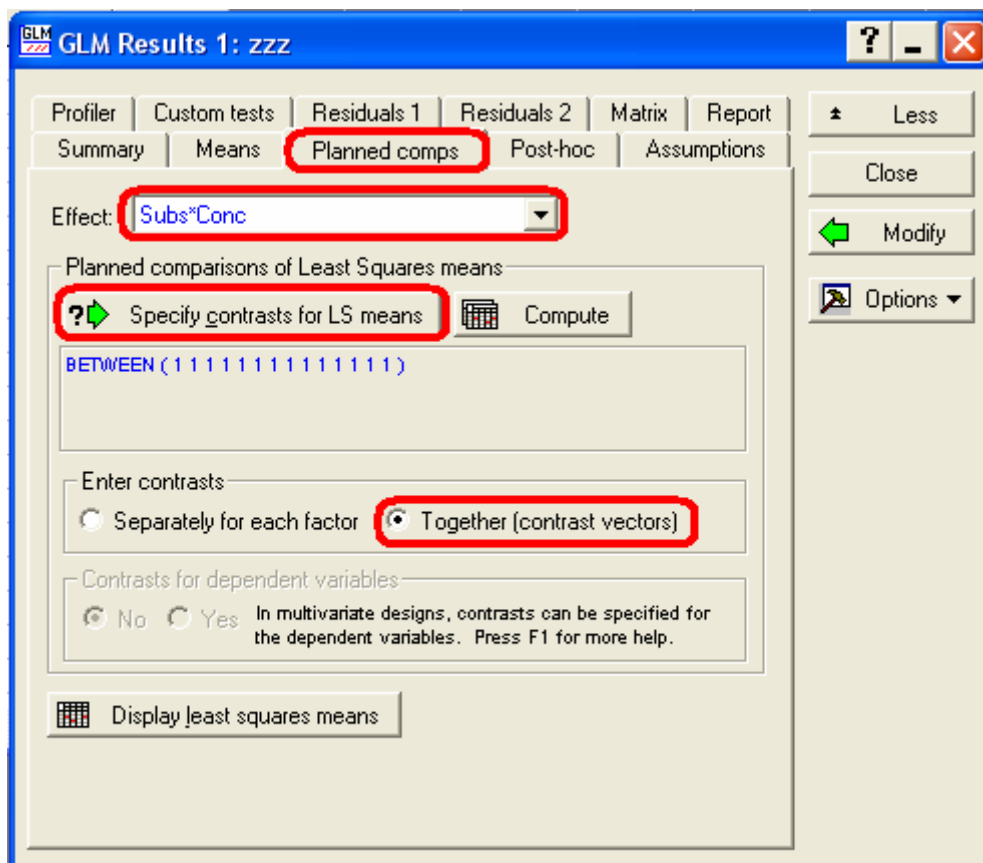
Вариант 1. Интерпретация через 95%-ные ДИ.

Согласно этому **очень консервативному** подходу для Концентраций 0, 2 и 3 различий между веществами нет, в Концентрации 1 значение больше для Вещества 2, в Концентрациях 4-6 – для вещества 1. В целом по анализу это так. Однако если посмотреть на графики Рис. 2, то видно, что устойчивые различия между веществами наблюдаются только в концентрациях 5-6 (меньшие значения для вещества 2).

Вариант 2. Запланированные ортогональные сравнения

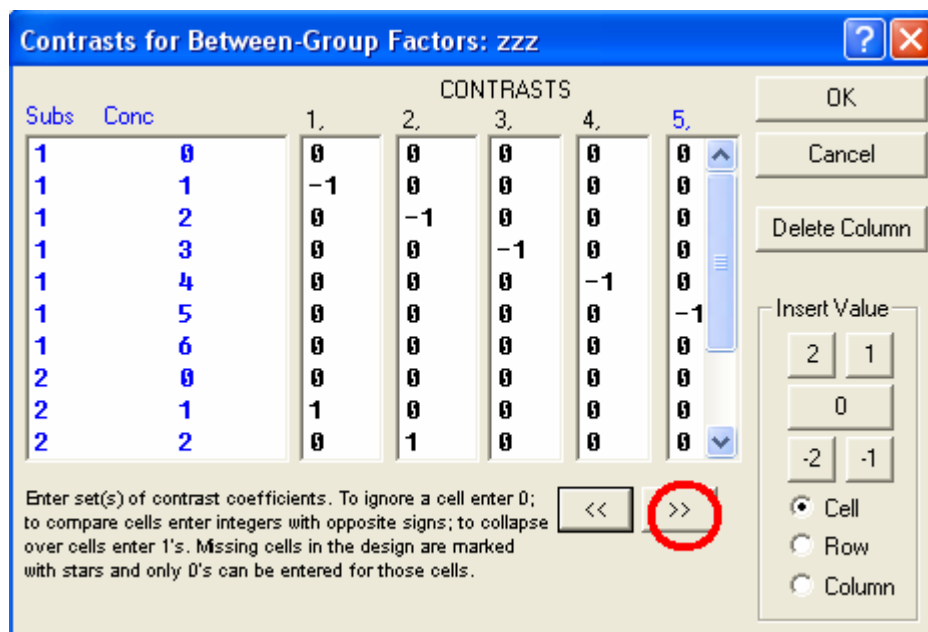
Т.к. вопросов у исследователя много больше, чем возможно ответить в ходе запланированных сравнений, т.е. $k > (a-1)$, данный вариант даётся в методических целях. Предположим, что различия в нулевой концентрации не представляли интереса, а нужно было бы найти различия между веществами в отличных от нуля концентрациях, т.е. это урезанный вопрос «Б»: «**б) При каждой конкретной концентрации есть ли достоверное отличие между опытом с веществом 1 и веществом 2 (2 повтора по 3 лунки)?**». Ответим на него, построив линейные контрасты по Шеффе (Scheffé).

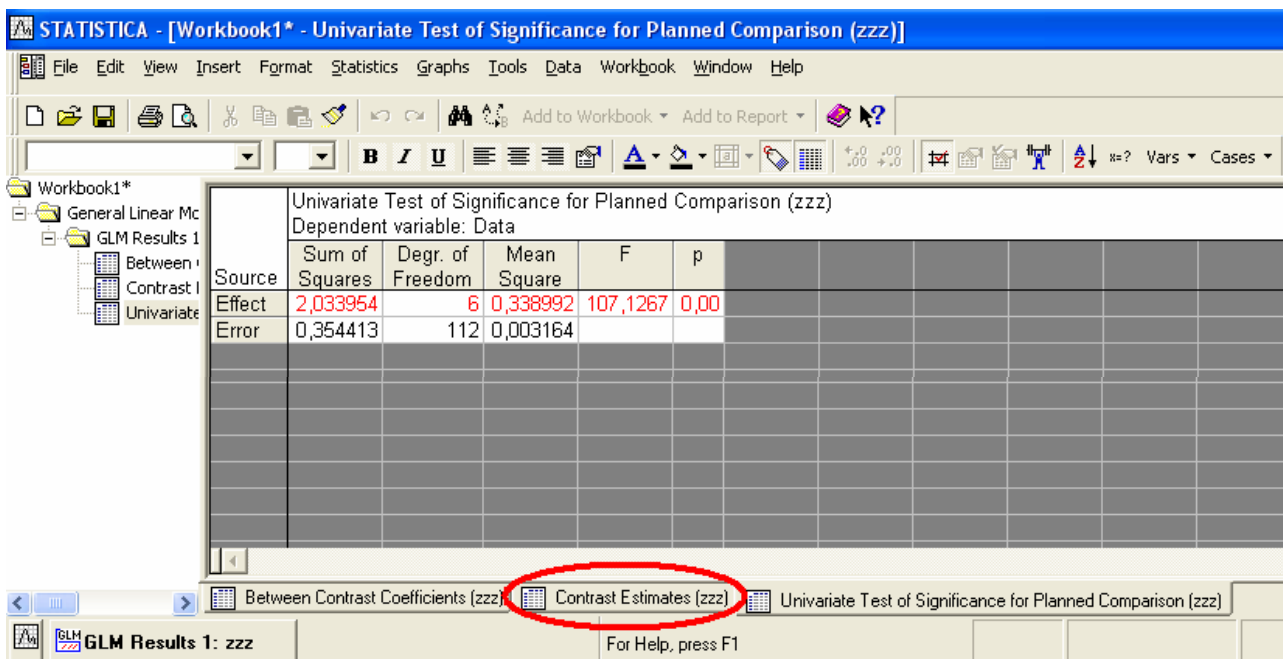
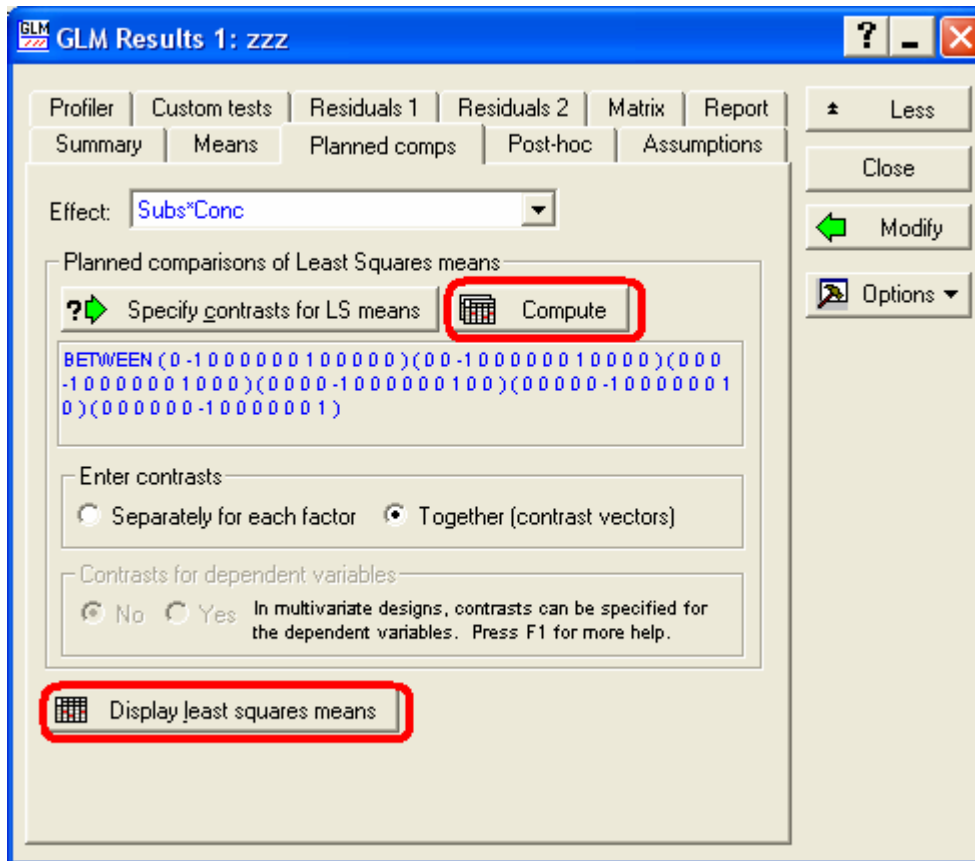
Форма **GLM Results**, Кнопка «More results», Закладка **Planned comps.** Выбираем нужный эффект, т.е. **Subs*Conc** и переходим в окно задания контрастов.



Т.к. $k \leq (a-1)$, то при $a=6$ имеем возможность задания только 5 контрастов.

Контраст 1: Вещество 1 в Концентрации 1 (ставим метку -1) против Вещества 2 в Концентрации 1 (ставим метку 1). Т.к. остальные вещества и концентрации в этом контрасте не задействованы – в остальные ячейки вносим 0. Аналогично далее Subs 1, Cons 2 vs Subs 2, Cons 2 и т.д. по пятый контраст (различия веществ в концентрации 6).





Результаты:

	Estimate	Std.Err	t	p	-95,00%	+95,00%
CNTRST1	0,070516	0,022965	3,0706	0,002681	0,025014	0,116019
CNTRST2	0,048059	0,022965	2,0927	0,038635	0,002556	0,093562
CNTRST3	-0,007848	0,022965	-0,3417	0,733201	-0,053350	0,037655
CNTRST4	-0,070847	0,022965	-3,0850	0,002565	-0,116350	-0,025344
CNTRST5	-0,175109	0,022965	-7,6250	0,000000	-0,220611	-0,129606
CNTRST6	-0,544028	0,022965	-23,6892	0,000000	-0,589530	-0,498525

Т.о. в данном случае незначимым оказался только контраст 3, т.е. различия между веществами в концентрации 3. По сравнению с этим, грубый метод доверительных интервалов не позволил обнаружить разницу в концентрации 2 (значение больше для Вещества 2).

Средние:

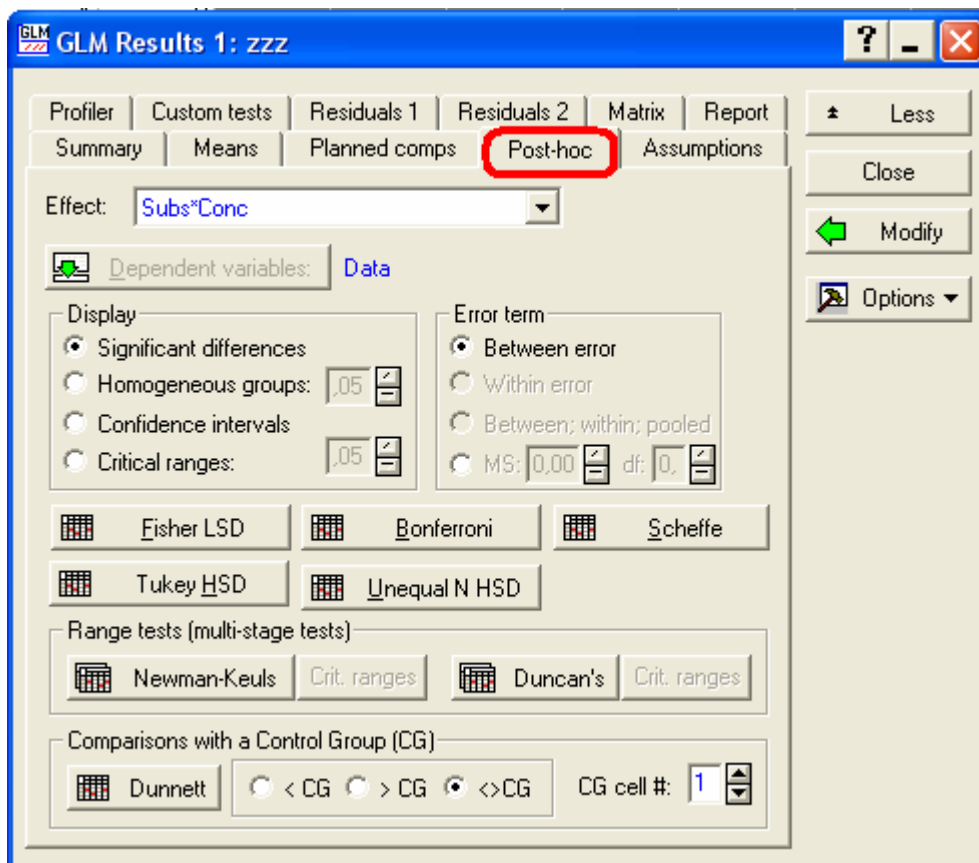
Subs*Conc; LS Means (zss)							
Current effect: F(6, --)=24,299, p=,00058							
Type III decomposition							
Cell No.	Subs	Conc	Data Mean	Data Std.Err.	Data -95,00%	Data +95,00%	N
1	1	0	1,000000	0,016239	0,967825	1,032175	12
2	1	1	0,961459	0,016239	0,929284	0,993634	12
3	1	2	0,905872	0,016239	0,873697	0,938047	12
4	1	3	0,940970	0,016239	0,908794	0,973145	12
5	1	4	0,905517	0,016239	0,873342	0,937692	12
6	1	5	0,870256	0,016239	0,838080	0,902431	12
7	1	6	0,724597	0,016239	0,692422	0,756772	12
8	2	0	1,000000	0,016239	0,967825	1,032175	12
9	2	1	1,031975	0,016239	0,999800	1,064151	12
10	2	2	0,953931	0,016239	0,921756	0,986106	12
11	2	3	0,933122	0,016239	0,900947	0,965297	12
12	2	4	0,834670	0,016239	0,802495	0,866845	12
13	2	5	0,695147	0,016239	0,662972	0,727322	12
14	2	6	0,180569	0,016239	0,148394	0,212744	12

Контрасты могут быть и куда более сложными – см. Закса, Монтгомери.

Вариант 3. Запланированные неортогональные сравнения

Т.к. интересующих сравнений много, наиболее грамотный вариант для обчёта результатов этого эксперимента - запланированные неортогональные сравнения. В принципе они не дифференцируются от незапланированных, а поэтому попадают в ту же категорию Post-hoc сравнений. Как указывалось выше, в случае неравных дисперсий в ячейках комплекса (а тест Ливена обнаружил её – см. выше) используется метод Games and Howell - относительно либеральная модификация Tukey's HSD. Пакет Statistica 6.0 не позволяет им воспользоваться. Посмотрим, а чем можно воспользоваться? Переходим на закладку Post-hoc.

Bonferroni отмечаем как в высшей степени консервативный метод для случая большого числа сравнений (благоприятствует ошибке второго типа – необнаружению различий когда они на самом деле есть), менее консервативной последовательной техники методом Дана-Шидака в пакете нет. Unequal N HSD – тоже отмечаем, т.к. у нас равные объёмы выборок N в ячейках комплекса. Тест Newman-Keuls хорош лишь в определённых ситуациях, я его не использую. Duncan's test – вообще некорректный для множественных сравнений, т.к. не учитывает собственно множественности сравнений©. Тест множественных сравнений Scheffé достаточно консервативен и обладает малой мощностью, особенно в нашем случае - когда и групп много, и число степеней свободы велико.



В данном пакете нам оставляют только:

Fisher LSD (The least significant difference test) – разновидность t-критерия, самый либеральный из Post-hoc. Используется только после обнаружения различий F-критерием.

Tukey HSD (Tukey honestly significant difference test) – наиболее строгий (консервативный) метод из Post-hoc, основанный на статистике стьюдентизированного размаха.

Т.к. эти два метода есть во всех статпакетах часто применяют их связку: значимые в HSD различия признаются значимыми, незначимые в LSD различия признаются незначимыми, по значимым в LSD но незначимым в HSD различиям решение не принимается и оставляется потомкам.

Проводить анализ полученных таблиц не представляет сложностей и для новичка. Для сокращения размеров таблиц можно отформатировать столбцы до трёх знаков после запятой (деципроцентного уровня значимости хватает для принятия решений даже в строгих случаях).

LSD test; variable Data (zzz) Probabilities for Post Hoc Tests Error: Between MS = ,00316, df = 112,00													
Cell No.	{2}	{3}	{4}	{5}	{6}	{7}	{8}	{9}	{10}	{11}	{12}	{13}	{14}
1	,96146	,90587	,94097	,90552	,87026	,72460	1,0000	1,0320	,95393	,93312	,83467	,69515	,18057
2	0,096	0,000	0,011	0,000	0,000	0,000	1,000	0,167	0,047	0,004	0,000	0,000	0,000
3		0,017	0,374	0,016	0,000	0,000	0,096	0,003	0,744	0,220	0,000	0,000	0,000
4			0,129	0,988	0,124	0,000	0,000	0,000	0,039	0,238	0,002	0,000	0,000
5				0,125	0,003	0,000	0,011	0,000	0,574	0,733	0,000	0,000	0,000
6					0,127	0,000	0,000	0,000	0,037	0,232	0,003	0,000	0,000
7						0,000	0,000	0,000	0,000	0,007	0,124	0,000	0,000
8							0,000	0,000	0,000	0,000	0,000	0,202	0,000
9								0,167	0,047	0,004	0,000	0,000	0,000
10									0,001	0,000	0,000	0,000	0,000
11										0,367	0,000	0,000	0,000
12											0,000	0,000	0,000
13												0,000	0,000
14													0,000

Tukey HSD test; variable Data (zzz)													
Probabilities for Post Hoc Tests													
Error: Between MS = ,00316, df = 112,00													
Cell No.	{2}	{3}	{4}	{5}	{6}	{7}	{8}	{9}	{10}	{11}	{12}	{13}	{14}
	,96146	,90587	,94097	,90552	,87026	,72460	1,0000	1,0320	,95393	,93312	,83467	,69515	,18057
1	0,919	0,006	0,371	0,006	0,000	0,000	1,000	0,981	0,761	0,188	0,000	0,000	0,000
2		0,473	1,000	0,462	0,009	0,000	0,919	0,130	1,000	0,994	0,000	0,000	0,000
3	0,473		0,960	1,000	0,955	0,000	0,006	0,000	0,705	0,996	0,121	0,000	0,000
4	1,000	0,960		0,956	0,127	0,000	0,371	0,009	1,000	1,000	0,001	0,000	0,000
5	0,462	1,000	0,956		0,958	0,000	0,006	0,000	0,694	0,995	0,126	0,000	0,000
6	0,009	0,955	0,127	0,958		0,000	0,000	0,000	0,026	0,272	0,955	0,000	0,000
7	0,000	0,000	0,000	0,000	0,000		0,000	0,000	0,000	0,000	0,001	0,991	0,000
8	0,919	0,006	0,371	0,006	0,000	0,000		0,981	0,761	0,188	0,000	0,000	0,000
9	0,130	0,000	0,009	0,000	0,000	0,000	0,981		0,055	0,003	0,000	0,000	0,000
10	1,000	0,705	1,000	0,694	0,026	0,000	0,761	0,055		1,000	0,000	0,000	0,000
11	0,994	0,996	1,000	0,995	0,272	0,000	0,188	0,003	1,000		0,003	0,000	0,000
12	0,000	0,121	0,001	0,126	0,955	0,001	0,000	0,000	0,000	0,003		0,000	0,000
13	0,000	0,000	0,000	0,000	0,000	0,991	0,000	0,000	0,000	0,000	0,000		0,000
14	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	

В этих таблицах – ответы на все оставшиеся **зелёные** вопросы.

PS

То, что данных (3 лунки планшета на самую маленькую ячейку дисперсионного комплекса) вовсе не мало можно показать следующим способом. Отбросим каждое третье наблюдение, как будто в анализе было только 2 лунки -эти данные также добавил в приложение. Повторим анализ и убедимся, что точно к таким же выводам можно было придти **сократив затраты на реактивы и расход культуры клеток на треть (!)**.

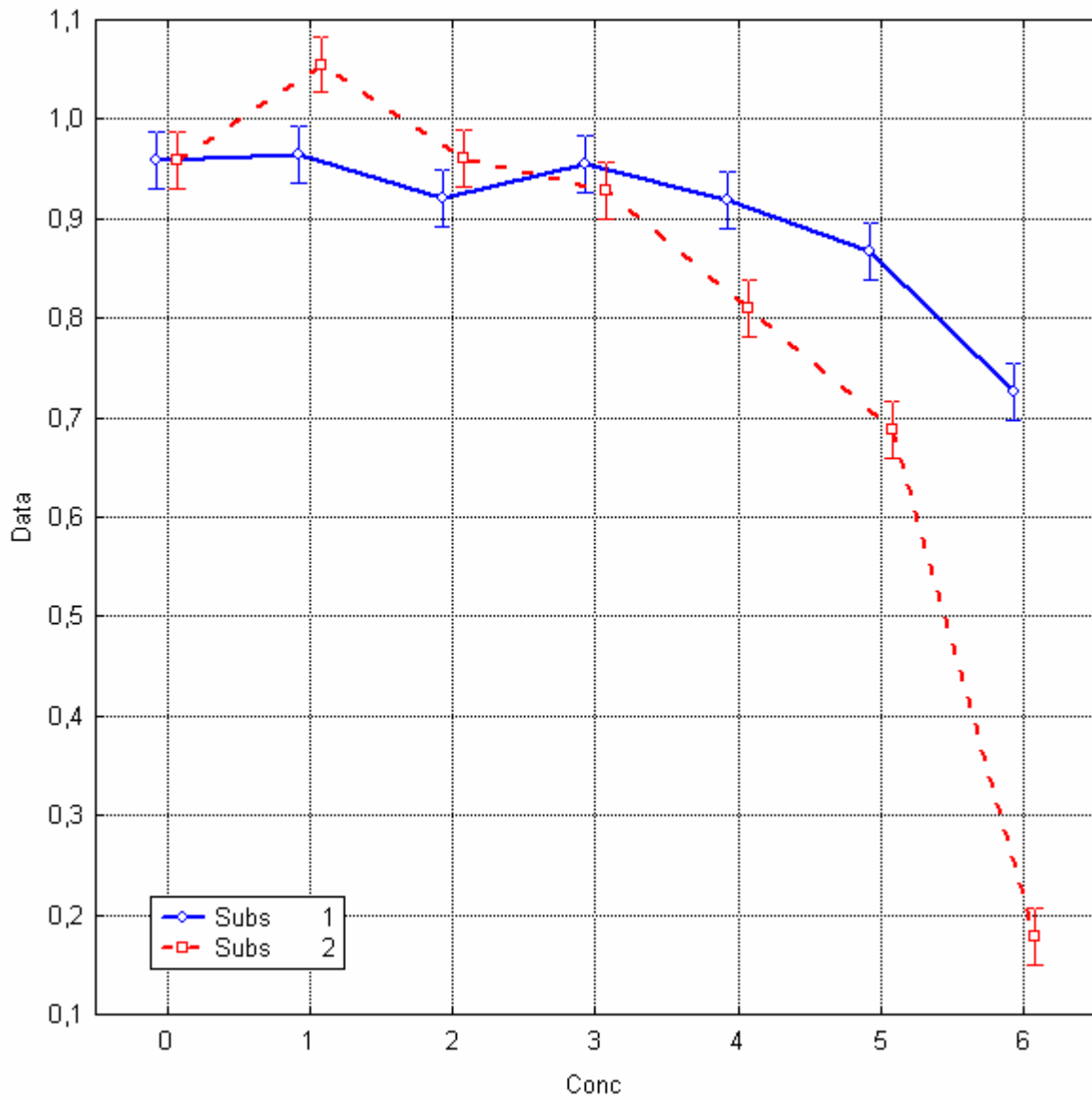


Рис. 4. Двухфакторное взаимодействие Концентрация x Вещество по сокращённым на треть данным

Приложение. Данные.

	Data	Repl	Factor	Subs	Conc
1	0,838951311	1	0	1	0
2	0,926966292	1	0	1	0
3	1,2340824	1	0	1	0
4	0,948864437	1	0	1	1
5	0,978191254	1	0	1	1
6	0,998057808	1	0	1	1
7	0,887372724	1	0	1	2
8	0,865614118	1	0	1	2
9	0,875074381	1	0	1	2
10	1,0084641	1	0	1	3
11	1,02076244	1	0	1	3
12	0,957378674	1	0	1	3
13	0,963054833	1	0	1	4
14	0,945080332	1	0	1	4
15	0,894940935	1	0	1	4
16	0,823988958	1	0	1	5
17	0,843855512	1	0	1	5
18	0,880750539	1	0	1	5
19	0,708573742	1	0	1	6
20	0,649920108	1	0	1	6
21	0,656542293	1	0	1	6
22	0,838951311	1	0	2	0
23	0,926966292	1	0	2	0
24	1,2340824	1	0	2	0
25	1,00467999	1	0	2	1
26	0,963054833	1	0	2	1
27	0,92994391	1	0	2	1
28	0,91953762	1	0	2	2
29	0,890210803	1	0	2	2
30	0,878858487	1	0	2	2
31	0,927105831	1	0	2	3
32	0,926159805	1	0	2	3
33	0,943188279	1	0	2	3
34	0,829665116	1	0	2	4
35	0,840071406	1	0	2	4
36	0,906293251	1	0	2	4
37	0,678300899	1	0	2	5
38	0,6962754	1	0	2	5
39	0,695329373	1	0	2	5
40	0,319756909	1	0	2	6
41	0,333001278	1	0	2	6
42	0,351921805	1	0	2	6
43	0,951490357	1	1	1	0
44	0,987142022	1	1	1	0
45	1,06136762	1	1	1	0
46	0,935125658	1	1	1	1
47	1,01227352	1	1	1	1
48	0,977206312	1	1	1	1
49	1,07071888	1	1	1	2
50	0,971361777	1	1	1	2
51	0,92226768	1	1	1	2
52	1,02162478	1	1	1	3
53	1,00058445	1	1	1	3
54	0,999415546	1	1	1	3
55	1,00642899	1	1	1	4
56	0,971361777	1	1	1	4
57	0,902396259	1	1	1	4
58	0,849795441	1	1	1	5

59	0,862653419	1	1	1	5
60	0,915254237	1	1	1	5
61	0,720046756	1	1	1	6
62	0,679135009	1	1	1	6
63	0,696668615	1	1	1	6
64	0,951490357	1	1	2	0
65	0,987142022	1	1	2	0
66	1,06136762	1	1	2	0
67	1,20748101	1	1	2	1
68	1,09292811	1	1	2	1
69	1,04617183	1	1	2	1
70	1,00642899	1	1	2	2
71	0,986557569	1	1	2	2
72	0,939801286	1	1	2	2
73	0,977206312	1	1	2	3
74	0,902396259	1	1	2	3
75	0,925774401	1	1	2	3
76	0,789012274	1	1	2	4
77	0,749269433	1	1	2	4
78	0,87434249	1	1	2	4
79	0,599649328	1	1	2	5
80	0,603156049	1	1	2	5
81	0,641729982	1	1	2	5
82	0,0958503799	1	1	2	6
83	0,0829924021	1	1	2	6
84	0,0689655172	1	1	2	6
85	0,980513729	2	0	1	0
86	1,03542958	2	0	1	0
87	0,984056687	2	0	1	0
88	0,800708592	2	0	1	1
89	0,897254207	2	0	1	1
90	0,824623561	2	0	1	1
91	0,709477414	2	0	1	2
92	0,905225864	2	0	1	2
93	0,775907883	2	0	1	2
94	0,796279894	2	0	1	3
95	0,887511072	2	0	1	3
96	0,783879539	2	0	1	3
97	0,853852967	2	0	1	4
98	0,785651019	2	0	1	4
99	0,852967228	2	0	1	4
100	0,862710363	2	0	1	5
101	0,930912312	2	0	1	5
102	0,86891054	2	0	1	5
103	0,746678477	2	0	1	6
104	0,766164748	2	0	1	6
105	0,779450841	2	0	1	6
106	0,980513729	2	0	2	0
107	1,03542958	2	0	2	0
108	0,984056687	2	0	2	0
109	1,09565988	2	0	2	1
110	0,953941541	2	0	2	1
111	0,89105403	2	0	2	1
112	0,997342781	2	0	2	2
113	0,976085031	2	0	2	2
114	0,987599646	2	0	2	2
115	0,93445527	2	0	2	3
116	0,939769708	2	0	2	3
117	0,923826395	2	0	2	3
118	0,774136404	2	0	2	4
119	0,852967228	2	0	2	4
120	0,842338353	2	0	2	4

121	0,73870682	2	0	2	5
122	0,76173605	2	0	2	5
123	0,751992914	2	0	2	5
124	0,125775022	2	0	2	6
125	0,125775022	2	0	2	6
126	0,139946856	2	0	2	6
127	0,968532526	2	1	1	0
128	0,983963691	2	1	1	0
129	1,04750378	2	1	1	0
130	1,10287443	2	1	1	1
131	1,03479576	2	1	1	1
132	1,02753404	2	1	1	1
133	1,00030257	2	1	1	2
134	0,954916793	2	1	1	2
135	0,932223903	2	1	1	2
136	0,972163389	2	1	1	3
137	0,931316188	2	1	1	3
138	0,91225416	2	1	1	3
139	0,936762481	2	1	1	4
140	0,890468986	2	1	1	4
141	0,863237519	2	1	1	4
142	0,865960666	2	1	1	5
143	0,892284418	2	1	1	5
144	0,845990923	2	1	1	5
145	0,776096823	2	1	1	6
146	0,763388805	2	1	1	6
147	0,752496218	2	1	1	6
148	0,968532526	2	1	2	0
149	0,983963691	2	1	2	0
150	1,04750378	2	1	2	0
151	1,05476551	2	1	2	1
152	1,06838124	2	1	2	1
153	1,07564297	2	1	2	1
154	0,95128593	2	1	2	2
155	0,953101362	2	1	2	2
156	0,960363086	2	1	2	2
157	0,937670197	2	1	2	3
158	0,879576399	2	1	2	3
159	0,980332829	2	1	2	3
160	0,820574887	2	1	2	4
161	0,819667171	2	1	2	4
162	0,917700454	2	1	2	4
163	0,726172466	2	1	2	5
164	0,704387292	2	1	2	5
165	0,744326778	2	1	2	5
166	0,165204236	2	1	2	6
167	0,184266263	2	1	2	6
168	0,173373676	2	1	2	6

Сокращённые на треть данные

	Data	Repl	Factor	Subs	Conc
1	0,838951311	1	0	1	0
2	0,926966292	1	0	1	0
3	0,948864437	1	0	1	1
4	0,978191254	1	0	1	1
5	0,887372724	1	0	1	2
6	0,865614118	1	0	1	2
7	1,0084641	1	0	1	3
8	1,02076244	1	0	1	3

9	0,963054833	1	0	1	4
10	0,945080332	1	0	1	4
11	0,823988958	1	0	1	5
12	0,843855512	1	0	1	5
13	0,708573742	1	0	1	6
14	0,649920108	1	0	1	6
15	0,838951311	1	0	2	0
16	0,926966292	1	0	2	0
17	1,00467999	1	0	2	1
18	0,963054833	1	0	2	1
19	0,91953762	1	0	2	2
20	0,890210803	1	0	2	2
21	0,927105831	1	0	2	3
22	0,926159805	1	0	2	3
23	0,829665116	1	0	2	4
24	0,840071406	1	0	2	4
25	0,678300899	1	0	2	5
26	0,6962754	1	0	2	5
27	0,319756909	1	0	2	6
28	0,333001278	1	0	2	6
29	0,951490357	1	1	1	0
30	0,987142022	1	1	1	0
31	0,935125658	1	1	1	1
32	1,01227352	1	1	1	1
33	1,07071888	1	1	1	2
34	0,971361777	1	1	1	2
35	1,02162478	1	1	1	3
36	1,00058445	1	1	1	3
37	1,00642899	1	1	1	4
38	0,971361777	1	1	1	4
39	0,849795441	1	1	1	5
40	0,862653419	1	1	1	5
41	0,720046756	1	1	1	6
42	0,679135009	1	1	1	6
43	0,951490357	1	1	2	0
44	0,987142022	1	1	2	0
45	1,20748101	1	1	2	1
46	1,09292811	1	1	2	1
47	1,00642899	1	1	2	2
48	0,986557569	1	1	2	2
49	0,977206312	1	1	2	3
50	0,902396259	1	1	2	3
51	0,789012274	1	1	2	4
52	0,749269433	1	1	2	4
53	0,599649328	1	1	2	5
54	0,603156049	1	1	2	5
55	0,0958503799	1	1	2	6
56	0,0829924021	1	1	2	6
57	0,980513729	2	0	1	0
58	1,03542958	2	0	1	0
59	0,800708592	2	0	1	1
60	0,897254207	2	0	1	1
61	0,709477414	2	0	1	2
62	0,905225864	2	0	1	2
63	0,796279894	2	0	1	3
64	0,887511072	2	0	1	3
65	0,853852967	2	0	1	4
66	0,785651019	2	0	1	4
67	0,862710363	2	0	1	5
68	0,930912312	2	0	1	5
69	0,746678477	2	0	1	6
70	0,766164748	2	0	1	6

71	0,980513729	2	0	2	0
72	1,03542958	2	0	2	0
73	1,09565988	2	0	2	1
74	0,953941541	2	0	2	1
75	0,997342781	2	0	2	2
76	0,976085031	2	0	2	2
77	0,93445527	2	0	2	3
78	0,939769708	2	0	2	3
79	0,774136404	2	0	2	4
80	0,852967228	2	0	2	4
81	0,73870682	2	0	2	5
82	0,76173605	2	0	2	5
83	0,125775022	2	0	2	6
84	0,125775022	2	0	2	6
85	0,968532526	2	1	1	0
86	0,983963691	2	1	1	0
87	1,10287443	2	1	1	1
88	1,03479576	2	1	1	1
89	1,00030257	2	1	1	2
90	0,954916793	2	1	1	2
91	0,972163389	2	1	1	3
92	0,931316188	2	1	1	3
93	0,936762481	2	1	1	4
94	0,890468986	2	1	1	4
95	0,865960666	2	1	1	5
96	0,892284418	2	1	1	5
97	0,776096823	2	1	1	6
98	0,763388805	2	1	1	6
99	0,968532526	2	1	2	0
100	0,983963691	2	1	2	0
101	1,05476551	2	1	2	1
102	1,06838124	2	1	2	1
103	0,95128593	2	1	2	2
104	0,953101362	2	1	2	2
105	0,937670197	2	1	2	3
106	0,879576399	2	1	2	3
107	0,820574887	2	1	2	4
108	0,819667171	2	1	2	4
109	0,726172466	2	1	2	5
110	0,704387292	2	1	2	5
111	0,165204236	2	1	2	6
112	0,184266263	2	1	2	6