

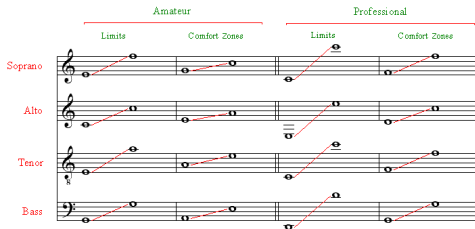
Прикладная статистика. Занятие 5. Дисперсионный анализ.

15 марта 2011 г.

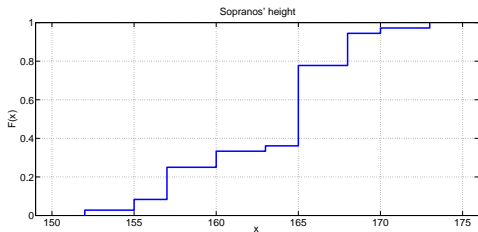
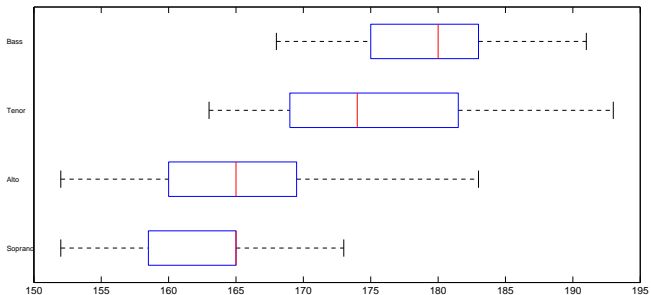
Рост певцов хора

В 1979 году 130 участников Нью-Йоркской ассоциации хорового пения сообщили данные своего роста; для каждого известен также регистр голоса. Есть ли связь между ростом и регистром?

Vocal Ranges



Рост певцов хора



Рост певцов хора

H_0 : рост и регистр голоса не связаны.

H_1 : для каких-то видов регистра голоса средний рост отличается.

Source	SS	df	MS	F	Prob>F
Groups	6901.4	3	2300.47	55.73	5.34718e-023
Error	5201.1	126	41.28		
Total	12102.5	129			

SS — сумма квадратов отклонений, df — число степеней свободы, MS — дисперсия, F — статистика критерия;

строка Groups — оценки по выборочным средним, строка Error — оценки по выборочным дисперсиям.

Предположения ANOVA:

- 1 значения признака во всех группах нормально распределены;
- 2 дисперсия значений признака во всех группах одинакова;
- 3 все наблюдения независимы.

Метод устойчив к нарушению первых двух предположений.

Рост певцов хора

Критерий Стьюдента для проверки гипотезы равенства роста певцов с альтом и сопрано: $p = 0.2460$ — против двусторонней альтернативы, $p = 0.1230$ — против односторонней альтернативы.

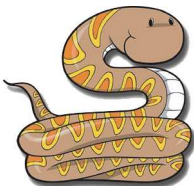
Критерий Стьюдента для проверки гипотезы равенства роста певцов с тенором и басом: $p = 0.0597$ — против двусторонней альтернативы, $p = 0.0298$ — против односторонней альтернативы.

Критерий Джонкхиера для проверки наличия тренда (увеличение роста с понижением регистра голоса): $p < 0.00001$.

Скорость обживания клеток у гремучих змей

Place, Abramson, Habituation of the rattle response in western diamondback rattlesnakes, *Crotalus atrox* (2008): испытания проводились в течение четырёх дней. Каждый день гремучая змея помещалась в клетку, крышка которой автоматически открывалась и закрывалась каждые 5 минут. Первое время при этом змея начинала греметь, но со временем обживалась и переставала реагировать. Для 6 змей известен номер открытия крышки, при котором впервые змея не начинала греметь.

Требуется проверить, отличается ли скорость обживания клетки для различных змей и для различных дней проведения испытаний.

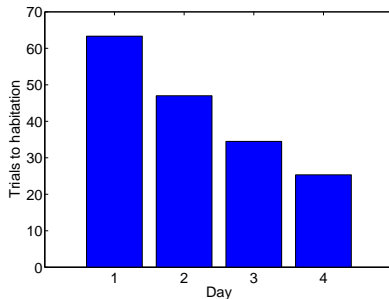


Скорость обживания клеток у гремучих змей

H_{01} : скорость обживания одинакова во все дни проведения испытания.

H_{02} : скорость обживания одинакова для всех змей.

Source	SS	df	MS	F	Prob>F
Days	4877.79	3	1625.93	3.32	0.0487
Snakes	3042.21	5	608.44	1.24	0.3382
Error	7345.96	15	489.73		
Total	15265.96	23			



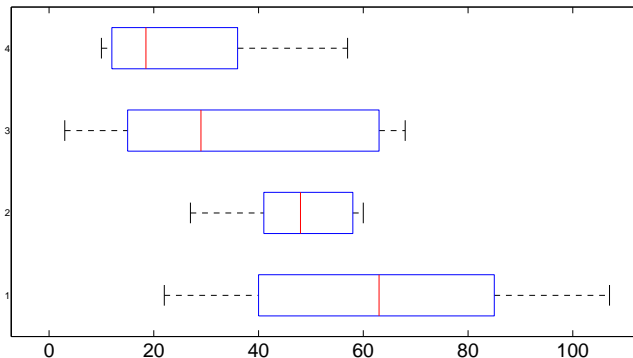
Скорость обживания клеток у гремучих змей

Критерий Фридмана убирает влияние одного из факторов, оценивает значимость оставшегося:

- если нейтрализовать влияние особи, получаем $p = 0.0384$;
- если нейтрализовать влияние дня, получаем $p = 0.1643$.

Скорость обживания клеток у гремучих змей

Однофакторный дисперсионный анализ с учётом только дня: $p = 0.0485$.



Критерий Джонкхиера для проверки наличия тренда (ускорение обживания): $p = 0.0037$.

Марихуана и скорость реакции

Изучалось воздействие марихуаны на скорость реакции. В качестве испытуемых были выбраны по 12 человек из каждой категории:

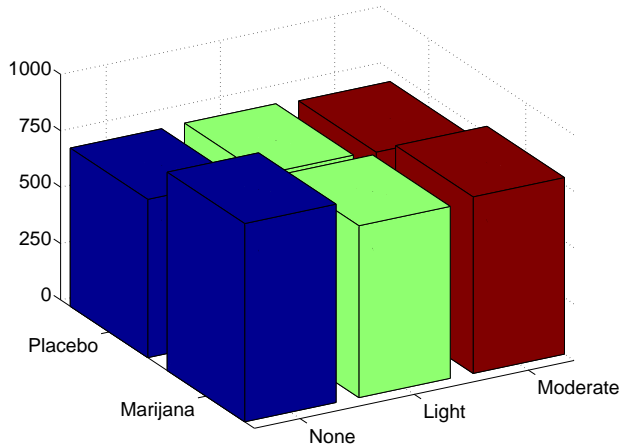
- никогда не пробовали марихуану;
- иногда употребляют марихуану;
- регулярно употребляют марихуану.

Испытуемые были разделены на две равные группы; половине из них дали выкурить две сигареты с марихуаной, вторая половина выкурила две обычные сигареты с запахом и вкусом марихуаны. Сразу после этого все испытуемые прошли тест на скорость реакции.

Требуется оценить влияние марихуаны на скорость реакции, учитывая фактор предыдущего опыта употребления.

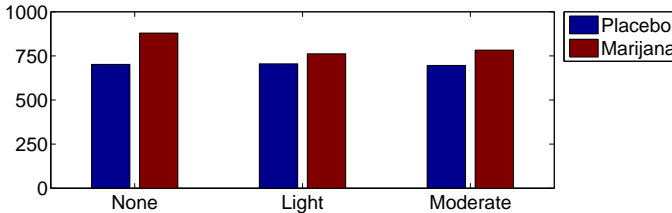
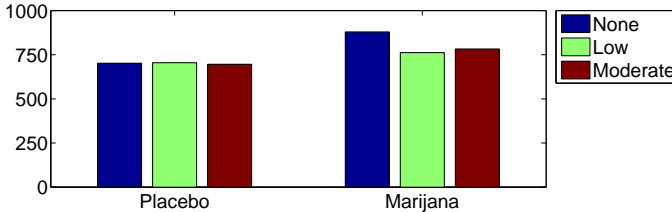
Марихуана и скорость реакции

Плохой график:



Марихуана и скорость реакции

Хорошие графики:



Марихуана и скорость реакции

H_{01} : средняя скорость реакции одинакова при употреблении и марихуаны, и сигарет.

H_{02} : средняя скорость реакции не зависит от предыдущего опыта употребления марихуаны.

H_{03} : средняя скорость реакции не зависит от комбинации факторов.

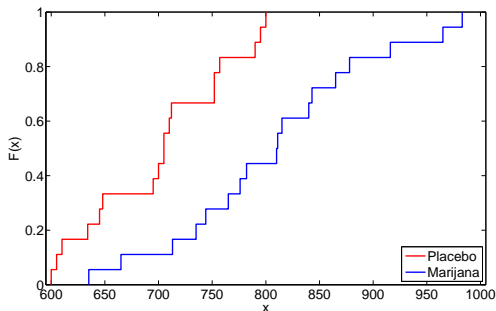
Source	SS	df	MS	F	Prob>F
Group	103041	1	103041	17.58	0.0002
Past use	23634.5	2	11817.2	2.02	0.1508
Interaction	23642.2	2	11821.1	2.02	0.1507
Error	175796.3	30	5859.9		
Total	326114	35			

Марихуана и скорость реакции

Вывод: гипотеза о том, что предыдущий опыт употребления влияет на скорость реакции, не принимается \Rightarrow данные по группам можно объединить.

Для объединённых данных:

- однофакторный дисперсионный анализ: $p = 0.00036$;
- критерий Уилкоксона, двусторонняя альтернатива: $p = 0.000596$;
- критерий Стьюдента, односторонняя альтернатива: $p = 0.00018$, $ci = (61.3, \infty)$;



Лечение гипертонии

72 пациента проходили лечение от гипертонии. Для лечения использовались три вида лекарств, при этом их эффект изучался как при использовании специальной диеты, так и в её отсутствие; кроме того, в ряде случаев применялась психотерапия. Данные — артериальное давление пациента по окончании лечения.

Требуется сравнить эффективность методов для лечения гипертонии.

Дизайн $[3 \times 2 \times 2]$.

Лечение гипертонии

Трёхфакторный дисперсионный анализ, все взаимодействия:

Source	SS	df	MS	F	Prob>F
Therapy	2048	1	2048	13.07	0.0006
Diet	5202	1	5202	33.2	0
Drug	3675	2	1837.5	11.73	0.0001
Therapy*Diet	32	1	32	0.2	0.6529
Therapy*Drug	259	2	129.5	0.83	0.4425
Diet*Drug	903	2	451.5	2.88	0.0638
Therapy*Diet*Drug	1075	2	537.5	3.43	0.0388
Error	9400	60	156.67		
Total	22594	71			

Лечение гипертонии

Значимость многофакторных взаимодействий:

- Diet*Drug: воздействие диеты различно при различных применяемых препаратах (или наоборот, действие препаратов зависит от диеты);
- Therapy*Diet*Drug: воздействие одного из факторов различно при различных комбинациях двух других. Хотя эффект Therapy*Drug незначим в целом, значимость Therapy*Diet*Drug говорит о том, что влияние Therapy*Drug необходимо оценивать отдельно для пациентов, использующих и не использующих диету.

Прикладная статистика
Семинар 5. Дисперсионный анализ.

Рябенко Евгений
riabenko.e@gmail.com