# Lab 3

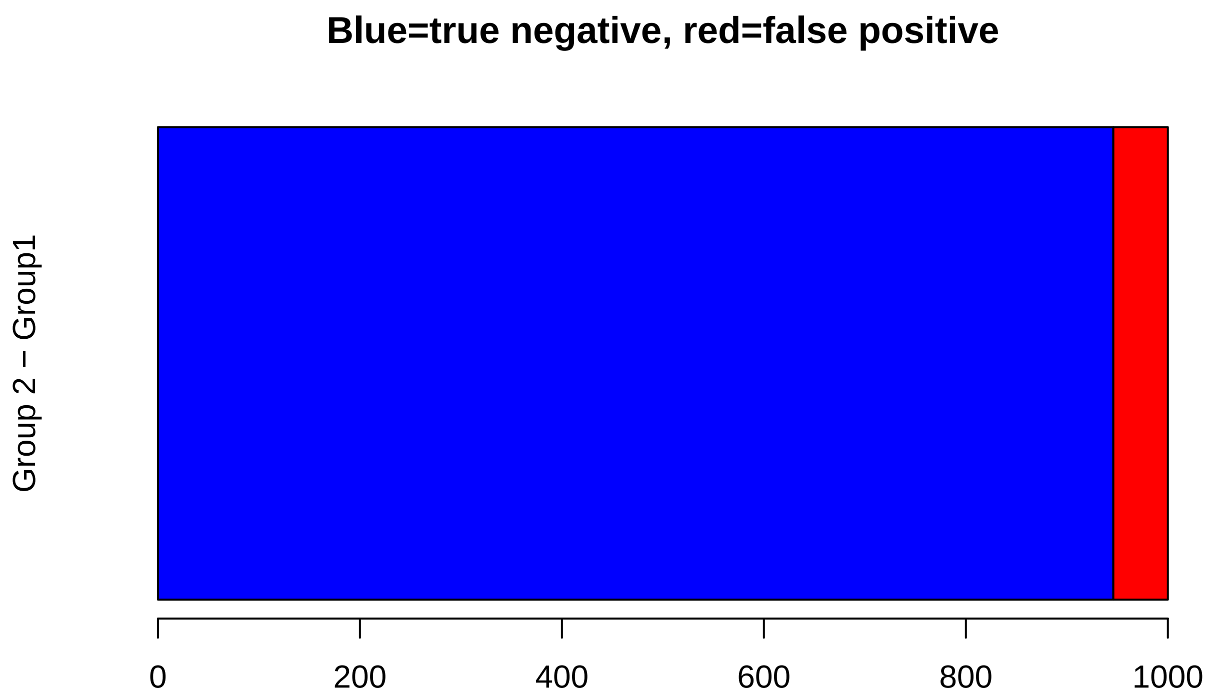
*Wiktor Soral*

*March 3rd 2017*

## Dangers of multiple testing

- Suppose you have repeatedly performed 100 experiments, where you compared means of two groups of people. All the procedures were the same, but each time you have drawn a new sample of participants.
- Suppose that you somehow know, that the 'true' difference of means equals 0.
- I you use t-test and treat results with p-values less than 0.05 as statistically significant, how often across 100 replications will you find significant results (false positives)?

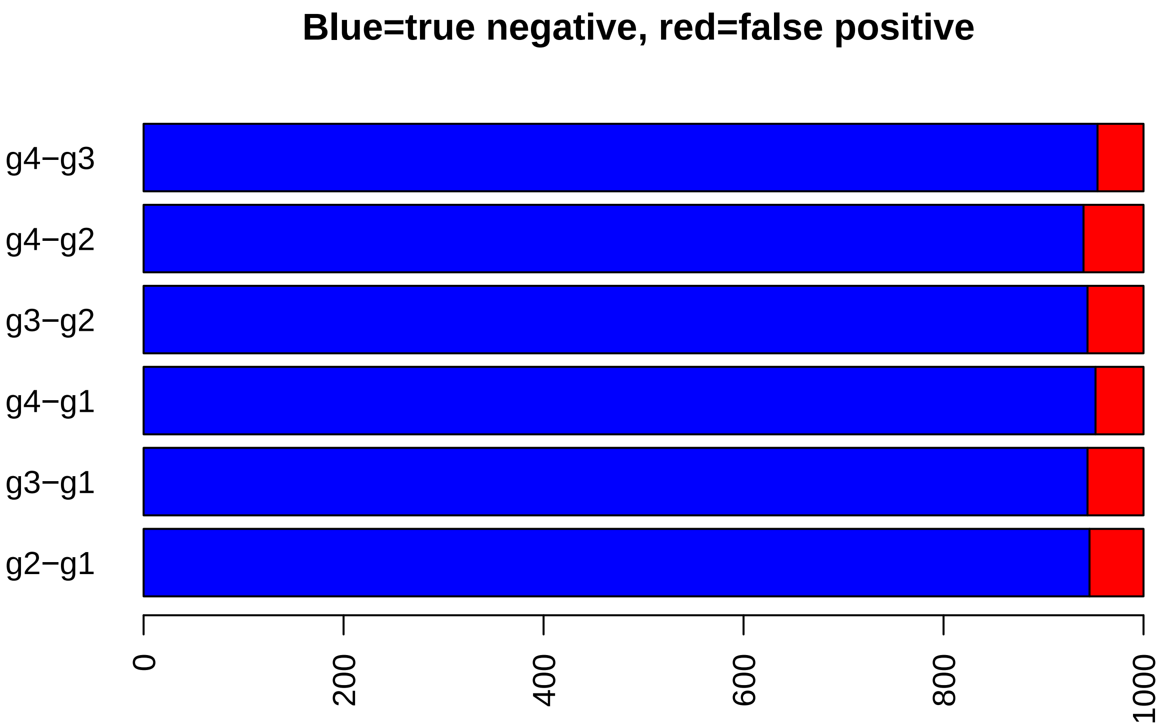
## Dangers of multiple testing




## Dangers of multiple testing

- Now suppose that you once again repeatedly performed 100 experiments, but you compared means of four groups of people.
- How many between group comparisons you can make?
- Recall binomial coefficient: $\frac{n!}{k!(n-k)!}$ - number of ways $k$ objects can be choosen from among $n$ objects, and $x!$ means factorial, e.g. $4! = 4 * 3 * 2 * 1$
- I you use t-test and treat results with p-values less than 0.05 as statistically significant, how often across 100 replications will you find at least one significant result (false positive)?

## Dangers of multiple testing

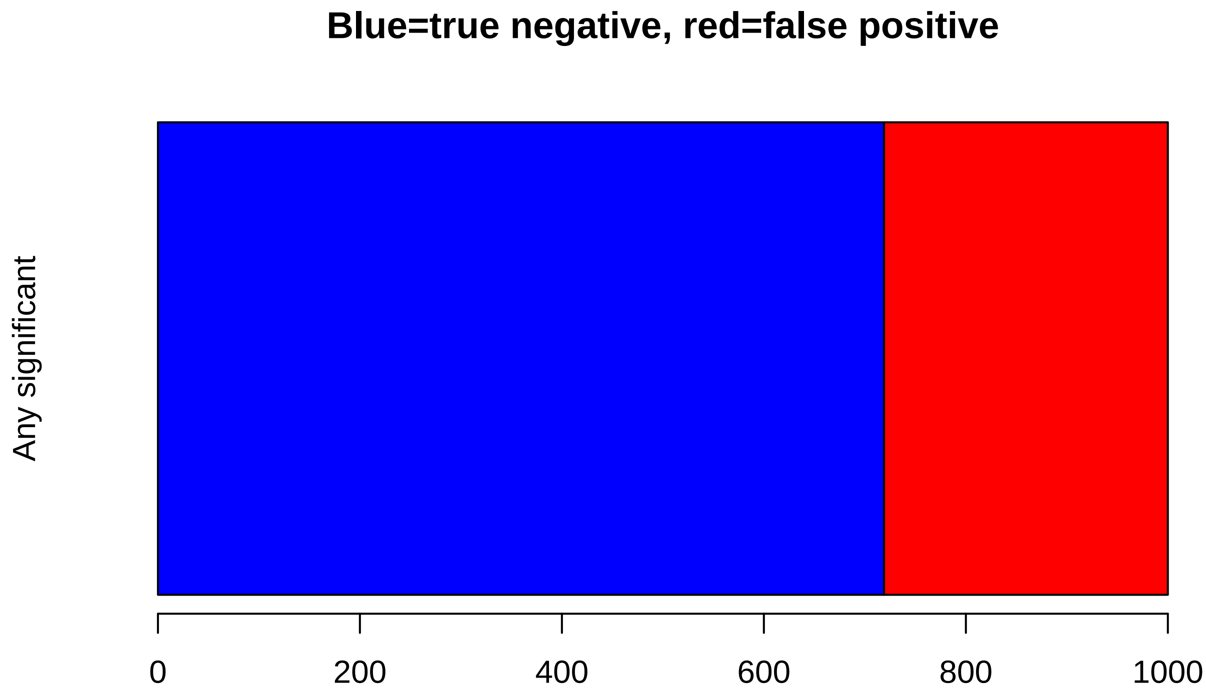



## Dangers of multiple testing

- Although you would expect to find significant differences in only 5 of 100 studies (when H0 is true), the probability of finding at least one significant result is in fact far more greater

## Dangers of multiple testing

- With 4 groups (6 comparisons) it is equal to $0.05^6 = 0.2649$!!!




### Post-hoc tests

- When you obtain significant F-value, you usually will like to check which group means differ
- One way to do it is by using Fisher's Least Significant Difference (LSD) test
- $LSD = t * \sqrt{2 * MS_R/n^*}$
- where t is the critical, tabled value of the t-distribution with the df associated with $MS_R$, and $n^*$ is the number of scores used to calculate the means of interest
- Simple t-value calculator
- LSD = minimum difference between a pair of means necessary for statistical significance
- LSD does not account for making multiple comparisons!

### Post-hoc tests

- One way to account for making multiple comparisons is by using Tukey's Honest Significant Difference (HSD)
- $HSD = q * \sqrt{MS_R/n^*}$
- where q is the relevant critical value of the studentized range statistic, and all the rest is the same
- In order to obtain critical q-value, in the studentized range statistic table look up the q value for and $\alpha = 0.05$, $df = \nu$ - df related to $MS_R$, and $k = p = r$ - number of groups
- Simple q-value calculator

### Post-hoc tests

- There is a number of other post-hoc procedures available
- Particularly popular is Scheffe's test due to it's conservatism (the greatest penalty on multiple comparisons), however Scheffe's test is likely to lead to Type II errors

- During this lab Tukey's test will be procedure of choice
- However in special design situations, other post-hoc procedures may also be preferable and should be explored as alternatives

## Planned comparisons

- Recall $y_{ij} = \mu + \alpha_j + \epsilon_{ij}$
- With 3 groups we have $\alpha_1, \alpha_2, \alpha_3$, we can assign to and multiply each coefficient by a value, like $-1, 0, 1, 2$, so that the sum of assigned values would be equal to 0
- The assigned values are known as contrasts coefficients, and are used to model specific relations between group means
- If $L$ is a matrix of different contrast, and $B$ is a vector of group mean coefficients, we can test whether, $L * B = 0$

## Planned comparisons

- Lets say $\alpha_1 = -1.2, \alpha_2 = 1.2, \alpha_3 = 2.3$, if we multply each coefficient by values of $-2, 1, 1$, then…
- $(-2) * (-1.2) + (1) * (1.2) + (1) * (2.3) = 0$, if $H_0$ if true, and if we move some value to the right side…
- $(-2) * (-1.2) = (1) * (1.2) + (1) * (2.3)$, we can also move -2 to the right, and for clarity we can multiply both sides by -1…
- $1.2 = \frac{1.2+2.3}{2}$
- In other words we test, whether mean of group 1 is equal to averaged means of group 2 and group 3

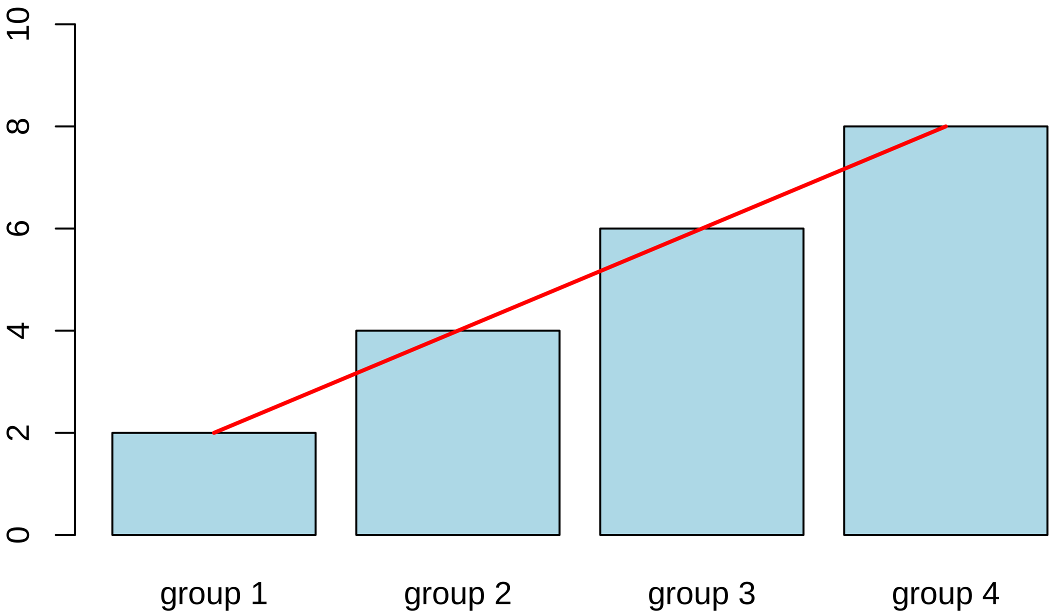
## Planned comparisons

- In our analyses, we can use as many contrasts as we want to test specific hypotheses…
- However, if we want to properly divide the model sum of squares, it necessary to have contrasts that are orthogonal
- For a given analysis you can find a set of up to $k - 1$ orthogonal contrasts, where $k$ is number of groups - try and check at home

| | $\alpha_1$ | $\alpha_2$ | $\alpha_3$ |
|---|---|---|---|
| | -2 | 1 | 1 |
| $\times$ | 0 | -1 | 1 |
| | 0 | -1 | 1 |

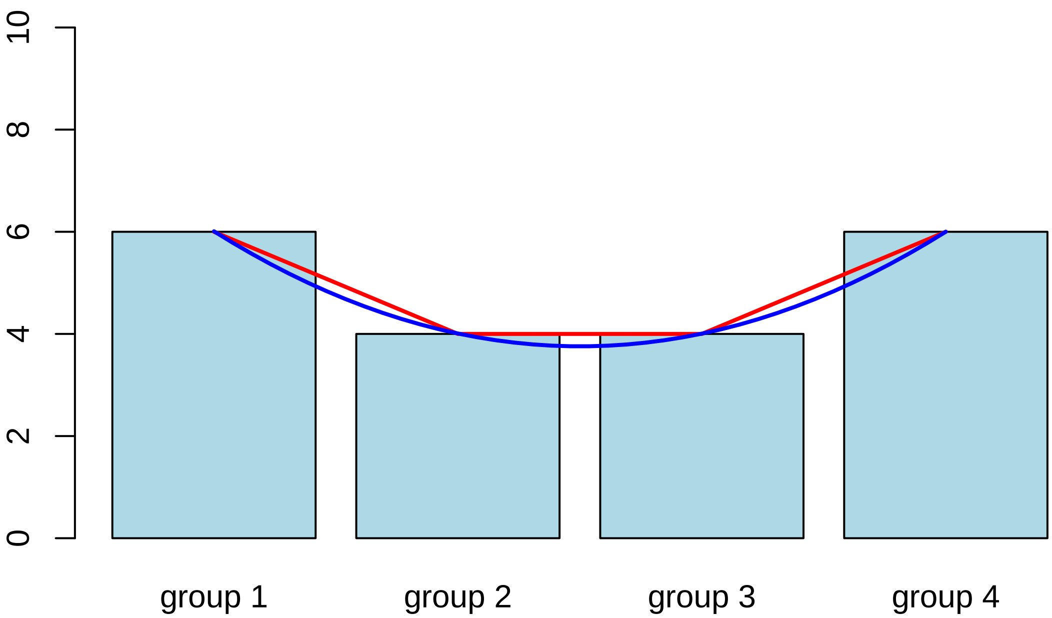
## Analysis of polynomial trends

- Linear trend



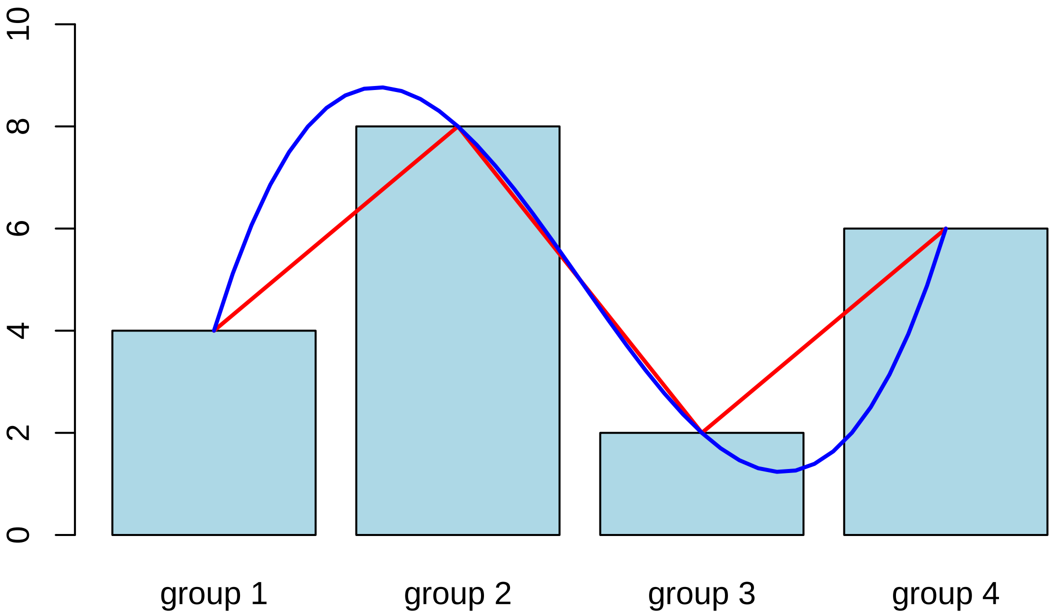

#### Analysis of polynomial trends

- Quadratic trend




#### Analysis of polynomial trends

- Cubic trend



10
8
6
4
2
0
group 1
group 2
group 3
group 4