



BME 2200: BME Biostatistics and Research Methods

Lecture 7: ANOVA, t-test



Lucas C. Parra
Biomedical Engineering Department
City College of New York



parra@ccny.cuny.edu



Content, Schedule

1. Scientific literature:

- Literature search
- Structure biomedical papers, engineering papers, technical reports
- Experimental design, correlation, causality.

2. Presentation skills:

- Report – Written report on literature search (individual)
- Talk – Oral presentation on biomedical implant (individual and group)

3. Graphical representation of data:

- Introduction to MATLAB
- Plot formats: line, scatter, polar, surface, contour, bar-graph, error bars. etc.
- Labeling: title, label, grid, legend, etc.
- Statistics: histogram, percentile, mean, variance, standard error, box plot

4. Biostatistics:

- Basics of probability
- t-Test, ANOVA
- Linear regression, cross-validation
- Error analysis
- Test power, sensitivity, specificity, ROC analysis

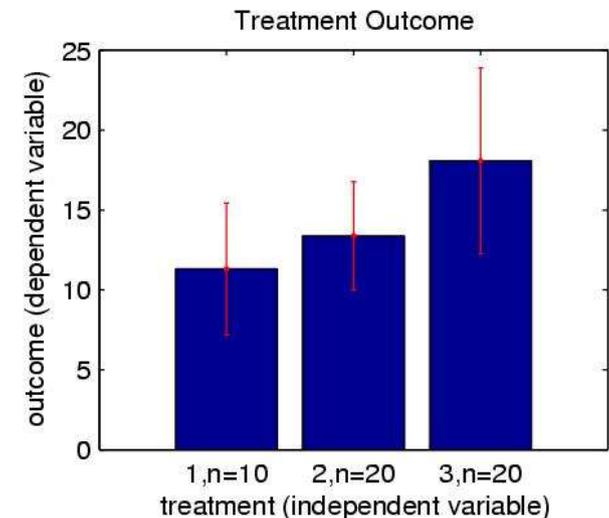
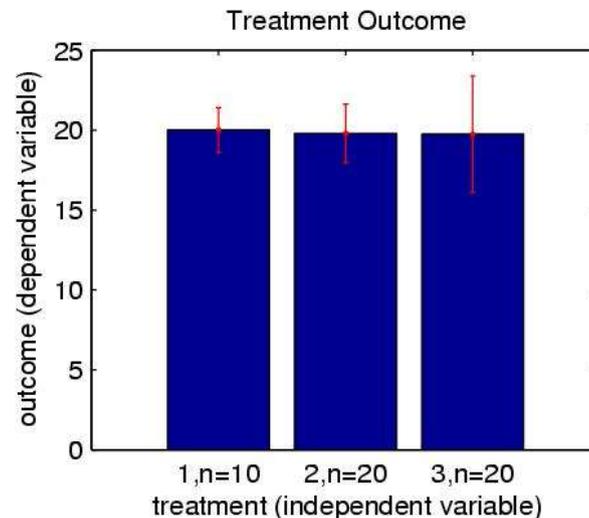
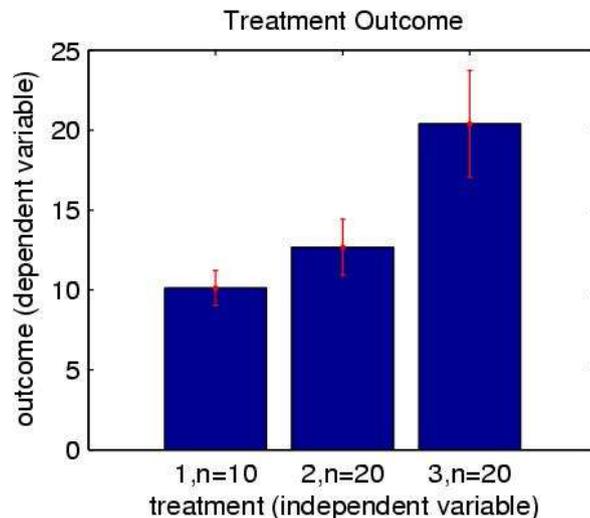


When are the means different?

A standard prospective study asks whether different treatment regimens will affected subjects differently, e.g. one group of subjects may received drug 1 to lower cholesterol, another received drug 2 to lower cholesterol, and a third received a placebo as a control.

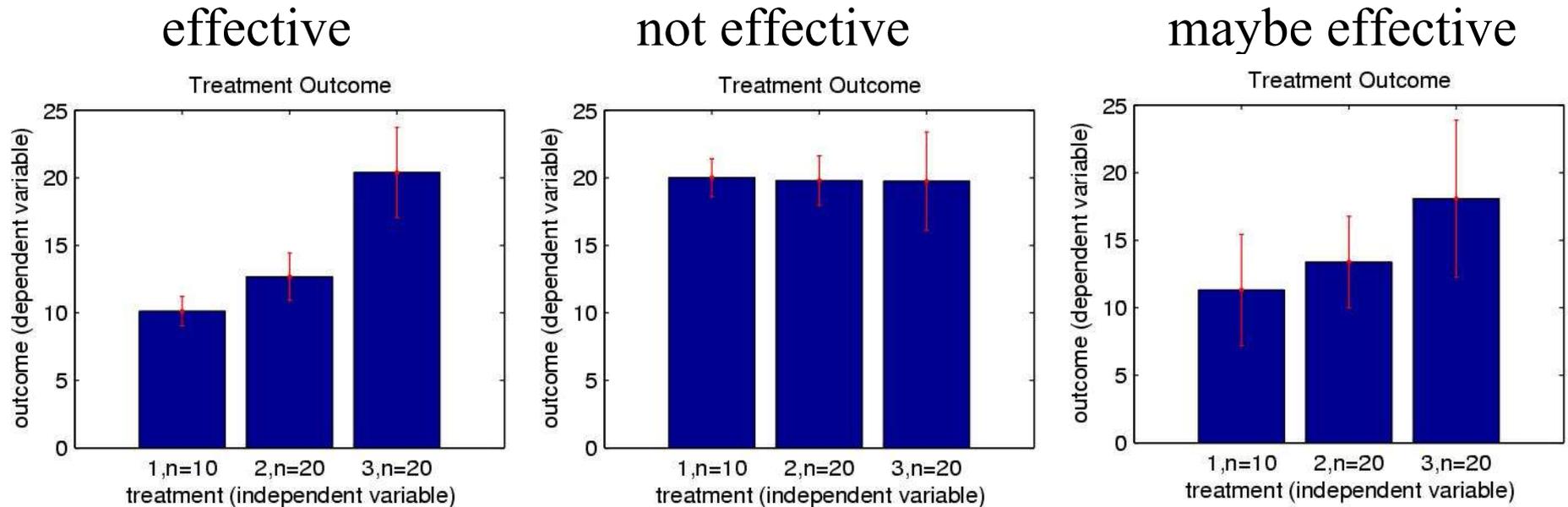
When can we assert that changing an independent variable (the drug) will have a **significant** effect on the dependent variable (cholesterol)?

We could look at the **mean** and **standard deviation** (std) of the cholesterol in the different groups. Here are the possible outcomes:





When are the means different?



- Big difference in the mean indicates 'significant' effect.
- Big variability within a group indicates less 'significant' effect.

An intuitive measure of significance is then

$$\frac{\text{variability between means}}{\text{variability within groups}} \longrightarrow \frac{\text{variance of means}}{\text{mean of variances}}$$



Definitions of mean and variance (again)

We have m groups with n_i data points in the i -th group.

Total number of data points:

$$N = \sum_{i=1}^m n_i$$

Sample mean of the i -th treatment group:

$$\langle x_i \rangle = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij}$$

Sample variance of the i -th treatment group:

$$s_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (x_{ij} - \langle x_i \rangle)^2$$



The F statistics

This ratio is called the F statistics

$$F = \frac{s_{between}^2}{s_{within}^2}$$

The within class variance is the mean of the variances of each individual treatment group:

$$s_{within}^2 = \langle s_x^2 \rangle = \frac{1}{m} \sum_{i=1}^m s_i^2$$

The between class variance is the variance of the means times the number of samples (assuming all groups have the same $n=n_i$):

$$s_{between}^2 = n s_{\langle x \rangle}^2 = \frac{n}{(m-1)} \sum_{i=1}^m (\langle x_i \rangle - \langle x \rangle)^2$$



Groups of different sizes

To define the variances for the case that we have different numbers of samples in each class introduce the term *variation* v for the un-normalized variance s^2 .

$$v_{total} = v_{between} + v_{within}$$

Define the total sample variance

$$s_{total}^2 = \frac{1}{(N-1)} v_{total} = \frac{1}{(N-1)} \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \langle x \rangle)^2$$

The between class variance is the variance of the means times the number of samples

$$s_{between}^2 = \frac{1}{(m-1)} v_{between} = \frac{1}{(m-1)} \sum_{i=1}^m n_i (\langle x_i \rangle - \langle x \rangle)^2$$

The within class variance is can always be computed from:

$$s_{within}^2 = \frac{1}{(N-m)} v_{within}$$



The F statistics

- When F is large there is a significant difference in the mean outcomes as compared to the variability in the data.
- When there is no significant difference between the means as compared to the variability in the data then F should be small.

What is a large or small F ?

Null Hypothesis: Lets assume there is NO DIFFERENCE between the groups, i.e. treatment had no effect. We assume that all samples are drawn independently from the same normal distribution.

With this assumption we expect the sample means and sample variances of the different groups to be about the same. We also expect the variance of the mean to be n times the sample variance. Therefore we **expect under the Null Hypothesis:**

$$F \approx 1$$

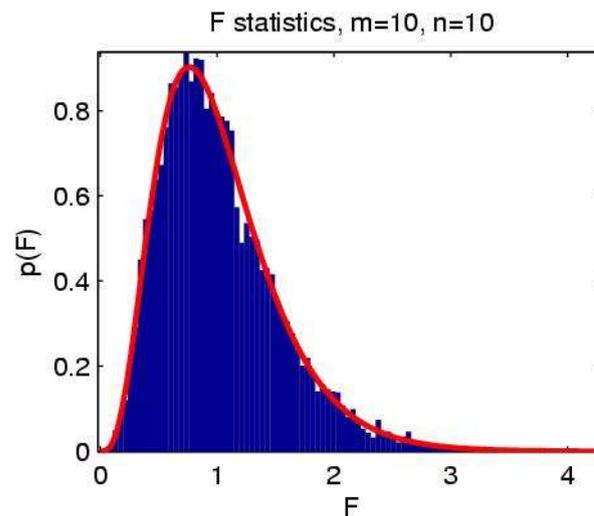


The F statistics

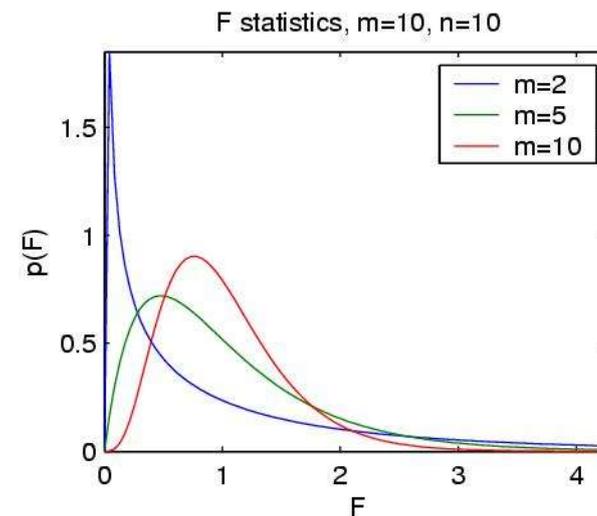
- F is a statistic that is estimated from random samples. It is therefore itself a random number.
- F could **by chance** take on values smaller or much larger than 1.
- Its probability density function $p(F)$ is given e.g. Schaum's ch. 4.

Demonstration

```
for i=1:N,
    x=rand(m,n);
    F(i) = n*std(mean(x)).^2/mean(std(x).^2);
end
```



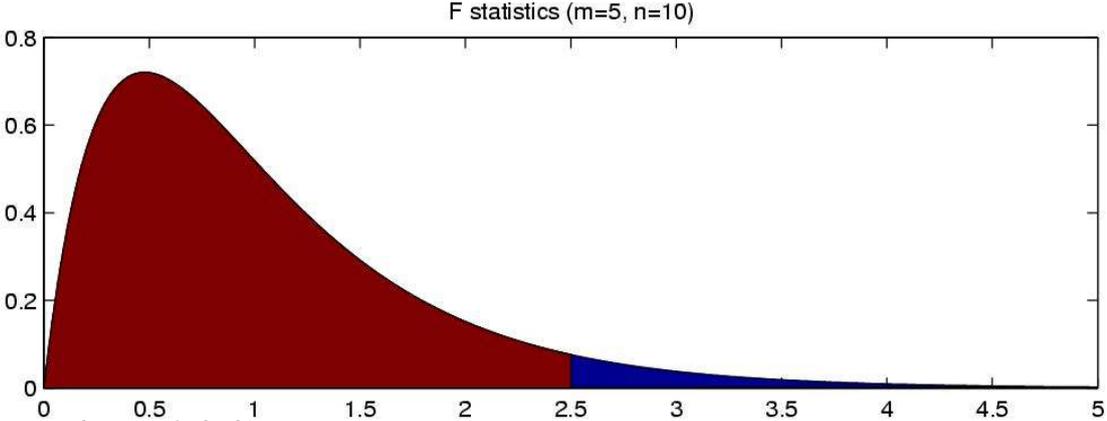
`hist(F)`



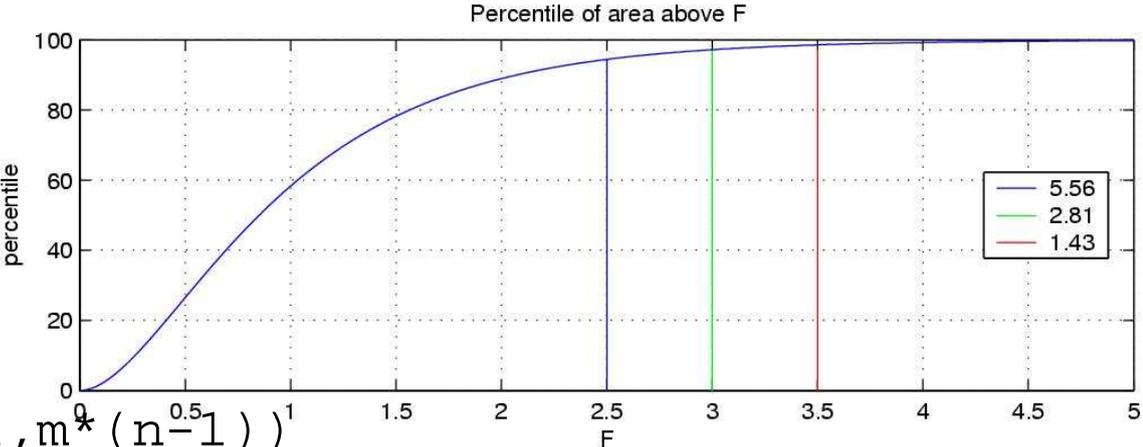
`fpdf(f, m-1, m*(n-1))` 9

Statistical significance – P value

Even when the null hypothesis is true we get occasionally a large value for F . A **statistically significant value** for F is a value that we get with low likelihood, e.g. $p < 0.05$ or $p < 0.01$.



$$f_{pdf}(f, m-1, m^*(n-1))$$



p-value

$$f_{cdf}(f, m-1, m^*(n-1))$$



The F statistics

Shape of the pdf of F is determined by the **degrees of freedom**. Degrees of freedom counts how many random number go into the computation of the variance:

$$F = \frac{S_{between}^2}{S_{within}^2} = n \frac{\frac{1}{(m-1)} \sum_{i=1}^m (\langle x_i \rangle - \langle x \rangle)^2}{\frac{1}{m(n-1)} \sum_{i=1}^m \sum_{j=1}^n (x_{ij} - \langle x_i \rangle)^2}$$

- The nominator uses m random numbers (the group means) minus one mean that reduces the estimated variability.

$$\nu_b = m - 1$$

- The denominator uses $N=mn$ random numbers (all the samples) minus m means that reduce the estimated variability.

$$\nu_w = N - m$$



Single factor ANOVA

This method of judging a statistically significant deviation from the null hypothesis based on variances is called **analysis of variance**.

Example: Determine if there is a statistically significant difference in the height of your classmates depending on their origine.

- What is the between group variance and the within group variance?
- What is the F value?
- What is the p value of the null-hypothesis and the likelihood that there is in fact a difference between the groups.

We have only looked at **single factor**: the origine. If one were to look at additional factors such as gender we would need to perform a **multiple factor ANOVA**.

To see the problem just think what would happen if we analyzed 100 different factors by applying single factor ANOVA 100 times.



Pairwise comparison – t test

In the case of only two groups the F statistics reduces to the t statistic (easy to show that $t^2=F$)

$$t = \frac{\langle x_1 \rangle - \langle x_2 \rangle}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$$

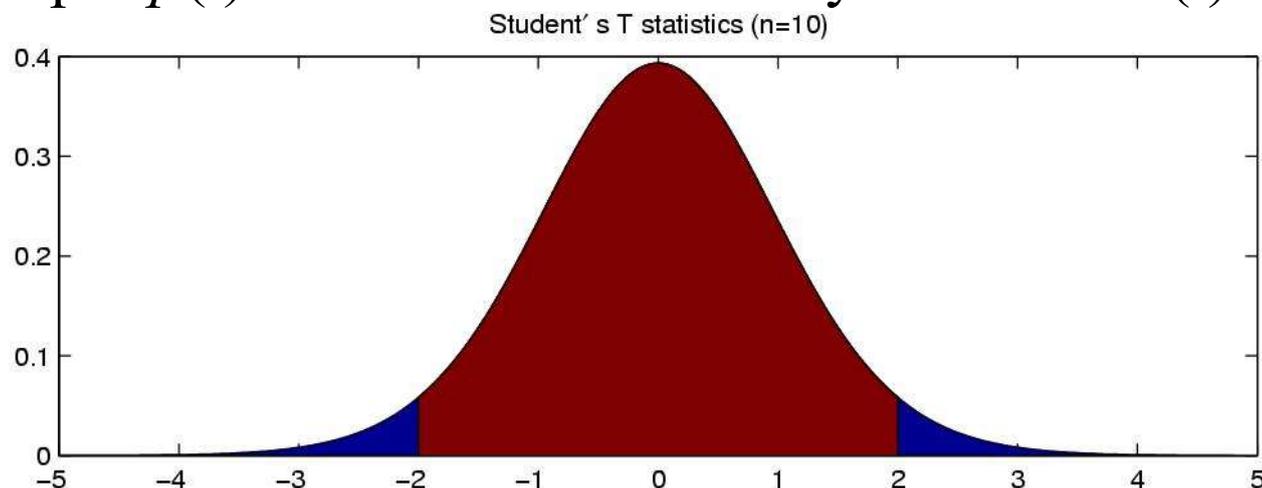
The only difference is that t can be positive or negative indicating that one group has a larger or smaller mean than the other group.

Since *a priory* we do not know which group is larger we have to consider both cases – two tailed t test!

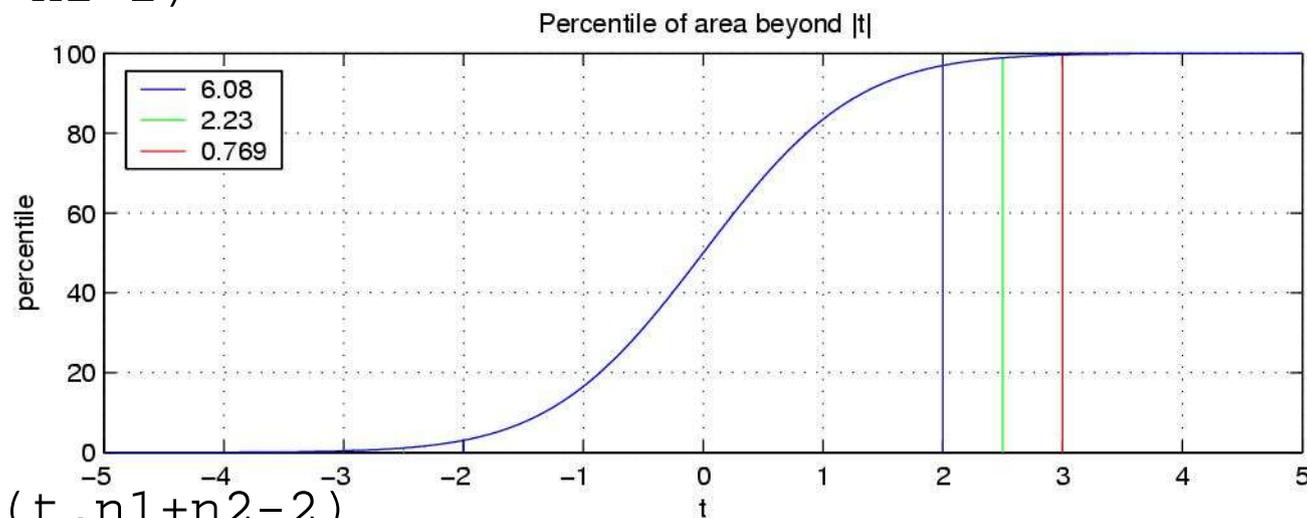


Pairwise comparison – t test

Under the null hypothesis t follows Student's T probability density with degrees of freedom $\nu = n_1 + n_2 - 2$. See Schaum's ch.4 for the formula of the pdf $p(t)$ and cumulative density function $Pr(t)$



$$t_{pdf}(t, n_1 + n_2 - 2)$$



$$Pr = tcdf(t, n_1 + n_2 - 2)$$



p-value

If $Pr(t, \nu)$ is the cumulative density function of t with ν degrees of freedom, then the likelihood that we observe $|t|$ by chance is

$$p = 2 \left(1 - Pr(|t|, \nu) \right)$$

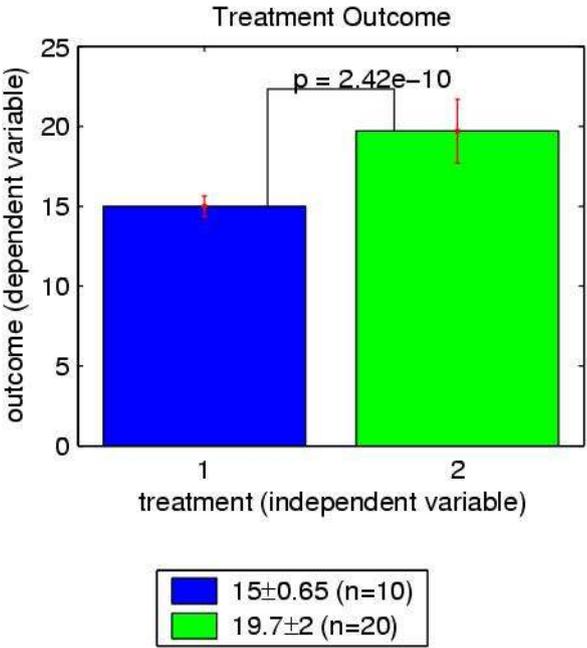
This likelihood is typically just called **p value**, and often quoted as a boundary. For instance, if we obtain a p value of 0.0089 we would simply write $p < 0.01$.

This means that there is a less than 1% chance of being wrong if we assert that the means are really different.

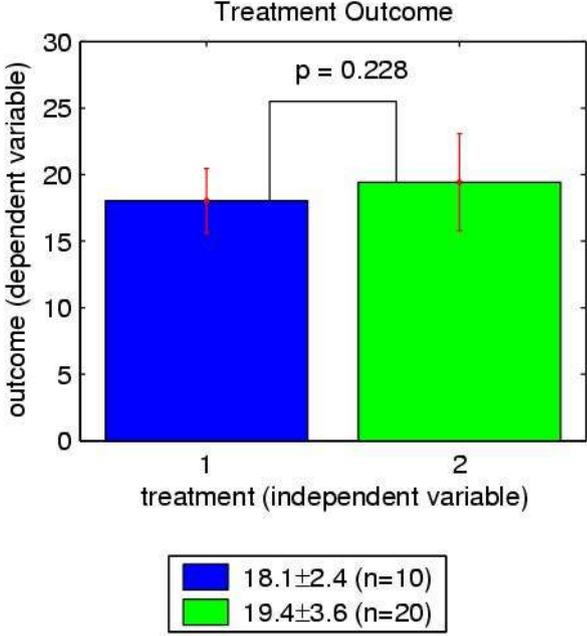
Pairwise comparison – t test

Examples for possible treatment outcomes:

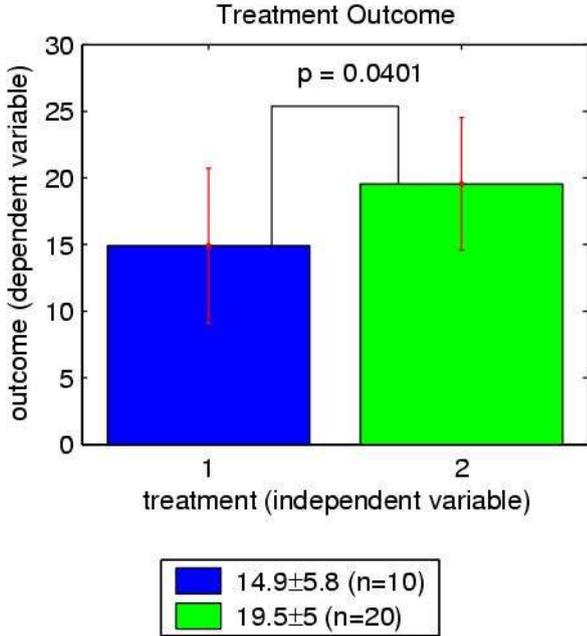
very significant



not significant



significant (p<0.05)





Pairwise comparison – t test

Assignment 7:

Determine if there is a statistically significant difference in the pulse rate between males and females in this class. Give the significance value for your finding.

Determine if there is a statistically significant difference in height between males and females. Give the significance value for your finding.

Read chapter 3 and 4 in Glantz.

Voluntary: For a more in-depth study read Schaum's ch. 7 and 9.



Multiple comparisons

With multiple groups always first perform ANOVA. Once the null hypothesis is rejected use the t test to detect the group that differs by doing pairwise comparisons with **adjusted significance values!**

Problem example: With 3 groups there are 3 possible pairs. At a $p=0.05$ each test has a 5% chance of confirming a difference of two groups when actually the null hypothesis is true. The total chance of erroneously concluding that one group is significantly different is then approximately 15%.

A simple conservative correction to this problem is the

Bonferroni t test (good for small m):

If there are k possible pairwise comparisons apply the conventional t test on each pair using p/k for a desired significance value of p .

For large m use the sequential **Holm t test**.