M1D7: Statistical methods and finalize data analysis

save as pdf

10/04/19

1. Statistics lecture
2. Statistical analysis on your data
3. Plan/work on data summary
4. Postlab: Recap Mod1

<u>Announcements</u>

on print menu

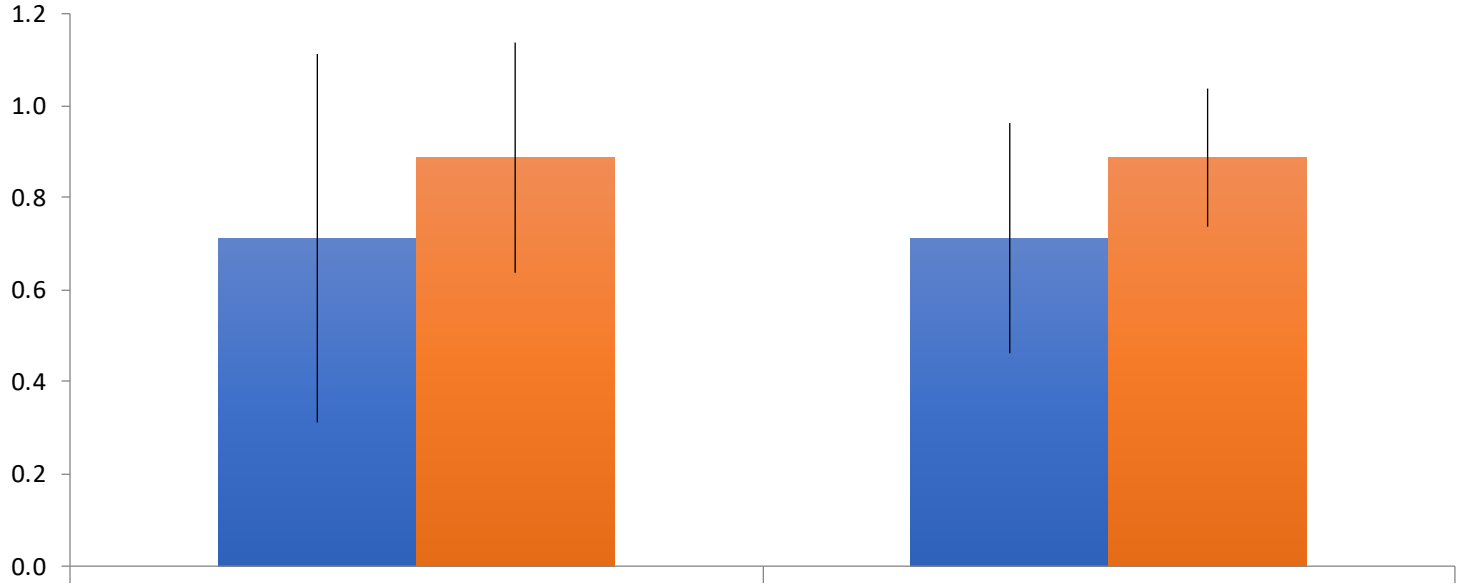
- Notebook due 10/5, 10pm
  - Graded in detail: **M1D4**, must post a pdf to Stellar
- Extra office hours—Instructors:
  - Sat. (10/12) 10a-4:30p, 56-302
- Office hours– Prof. Engelward:
  - Fri. (10/11) 8a-12p
- Data Summary draft due 10pm, Mon. 10/14

Regime change for Mod 2 & Mod 3!

- <u>ALL Future Assignments</u> should be posted to Stellar
  - I will still print them and return them in paper form
- When posting to Stellar– make sure your <u>name</u> is in your homework document

Confidence intervals show the variance in the data set

- At 95% confidence interval, there is a 95% chance that the true mean is within the defined range



Error bars depict: 95% CI st. dev.

With small sample sizes, 95% CI can be more reflective of sample variance

Calculating Confidence interval in excel

α

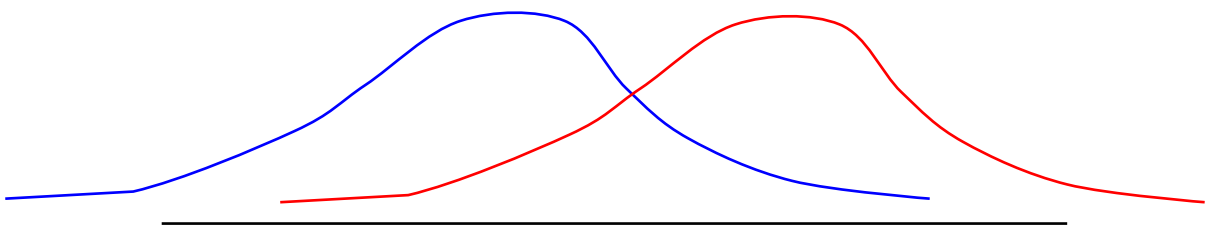
= CONFIDENCE.t(confidence level, standard dev., size)

0.05 =STDEV n

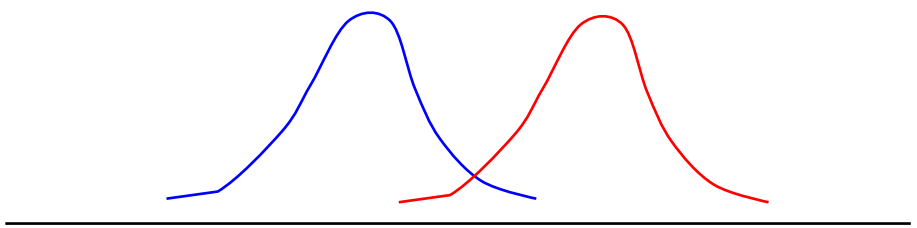
Need to calculate standard deviation in separate cell.

Student's *t*-test used to determine if populations are significantly different

- Assume data follows *t*-distribution
- At $p < 0.05$, there is less than a 5% chance that populations are the same (95% chance that populations are different)
- Examines ratio of signal (means):noise (variance)



high noise



low noise

Calculating Student's *t* in excel

NT MMS

p = TTEST (array1, array2,2,3)

Use the fewest assumptions:

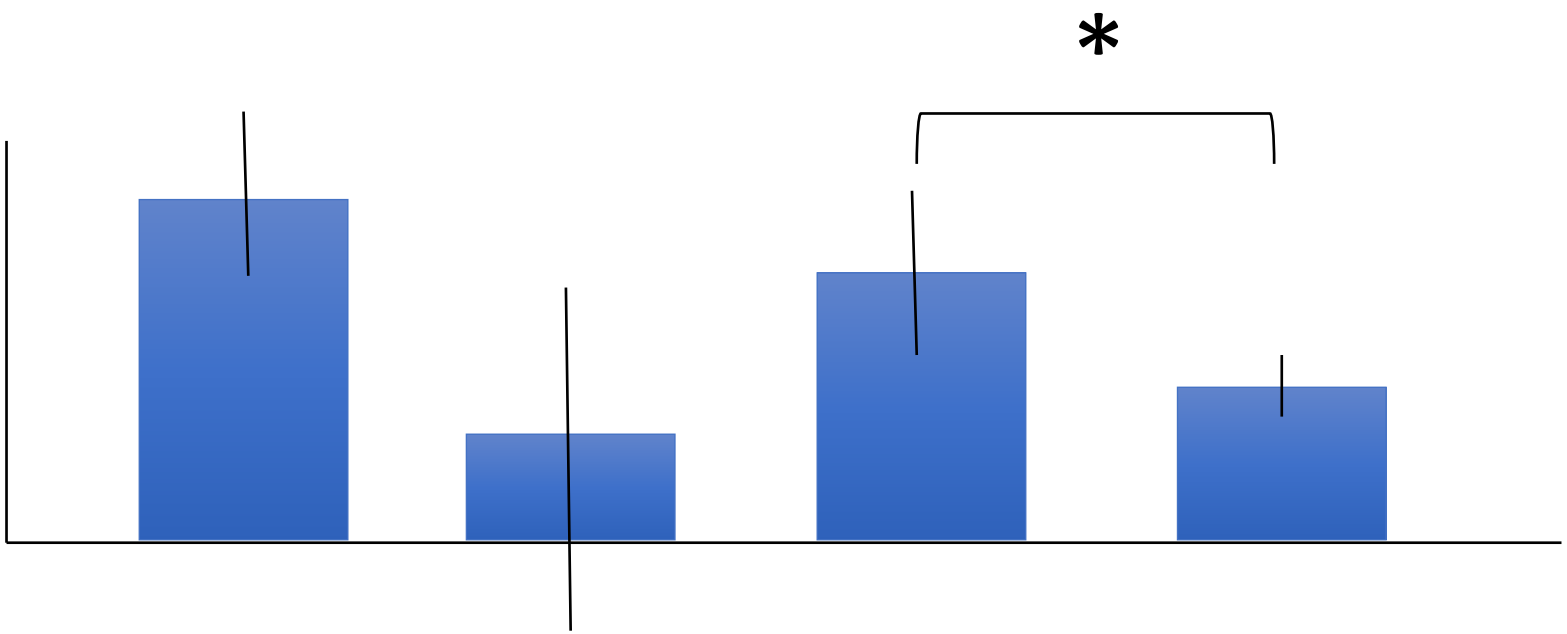
two-tailed

unequal variance

Can only compare two data sets at a time

*Make sure it is clear on your plots/writing which conditions are being compared

How will you use statistics in your data analysis?



What if the data are not statistically significant?

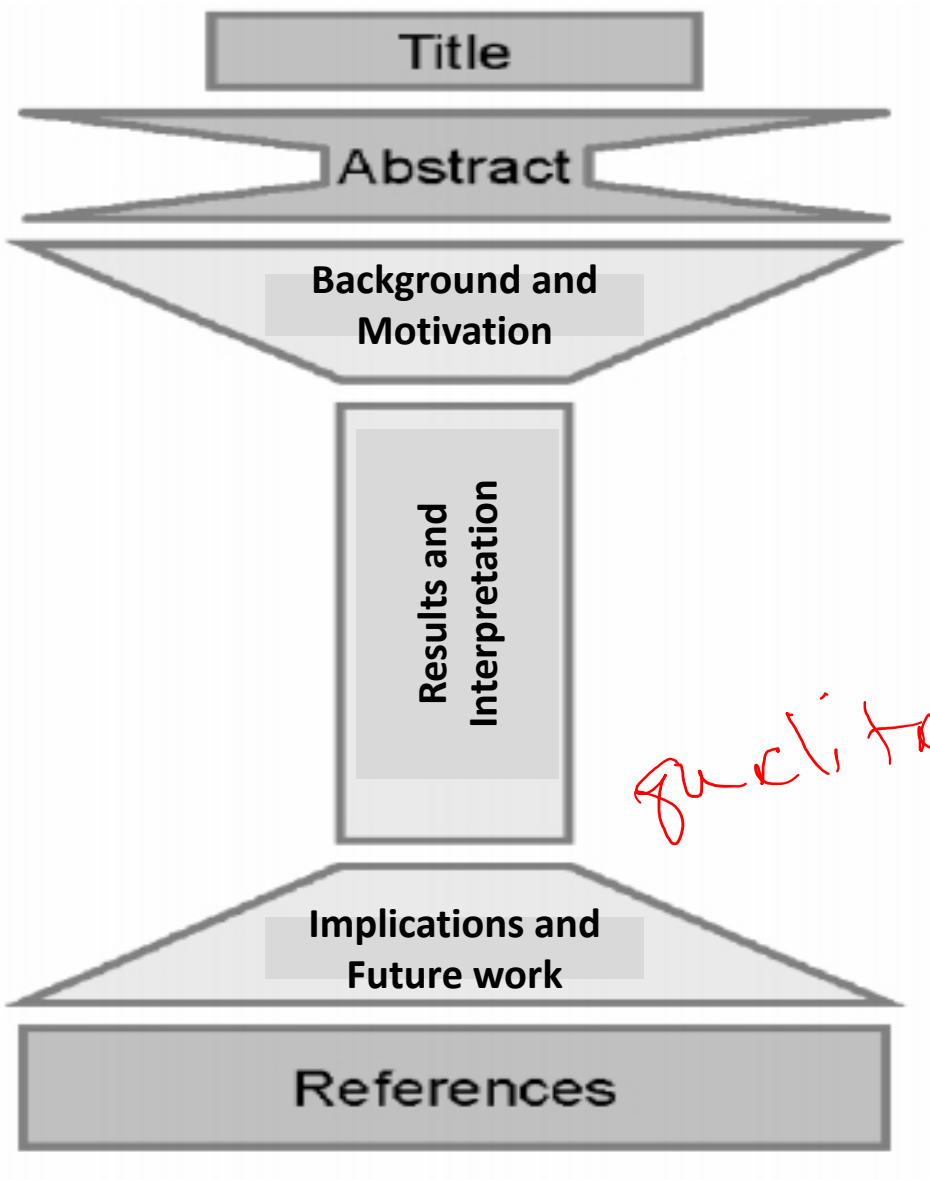
p = 0.1

p = 0.08

p = 0.055

Module 1: Measuring Genomic Stability Recap

M1 Data Summary



Title: take-away message

Abstract: the only section ***not*** in bullet points

ALL bullet points:

-background and motivation (include references) BCR

- schematics

-Results and interpretation

- Cell loading
- Comet Chip assay (Trevigen analysis) (TR purple team)
- H2AX assay (Matlab and ImageJ analysis)
  - H2AX foci and EdU staining analysis
    - Compare foci count and signal intensity

qualitative

Implications and future work (include references)

References (*see wiki for format suggestions*)

Background & Motivation

- Impact statement
  - general background
  - describe previous work in the field
- Specific background (e.g. BER,NMDA/MMS, Arsenic, CometChip, H2AX) Citations!
  - introduce topics, pathways and specific technologies necessary to understand the experimental approach
  - Reference schematic figure
  - narrow focus to the specific question addressed in your study
- Knowledge gap/statement of problem
  - what is unknown, therefore motivating your study
- Hypothesis
  - what do you propose will be the outcome of your study?
- A brief preview of your findings
  - Here we show…
  - end with broad implications of the study

Results & Interpretation

- Figures and captions
  - ***Decide on the figures first*** 1-3 panels
  - Use figure subpanels (label with letters)
  - Text: limited on figure, explicit in caption
  - reasonable size
  - descriptive title
  - Intro/purpose at beginning of in caption
  - caption descriptive of image, very light on methods

Do Not use Wiki Images
Don't forget stats

- Results and Interpretation (each page needs subtitle below figure caption)
  - Goal / intent / purpose of experiment = intro topic bullet
  - What you did: experiments and expectations, describe controls
  - What you found: quantitatively describe your result, referring to the figure ("Figure 1a shows…"
  - What does this indicate: interpret your result, what does it mean?
  - What does this motivate you to do next: **transition** to next experiment

Implications & Future Work

- Start with a very similar bullet to the last bullet in your Background/Motivation
  - Restate major results and broad implications
- Follow same order as in Figures/Results
  - Describe your conclusions from your data
    - If necessary describe caveats of experiment and suggest improvements
  - Identify unknowns and speculate (within reason)
    - Don't make huge generalizations or overreach
- Propose future experiments, identify new questions that arise
  - *What was Prof. Engelward's request for implications (hint: Arsenic)?*
- Come back to the big picture/impact statement topic introduced in background

Same order

comparable to environment? 10 μg/L