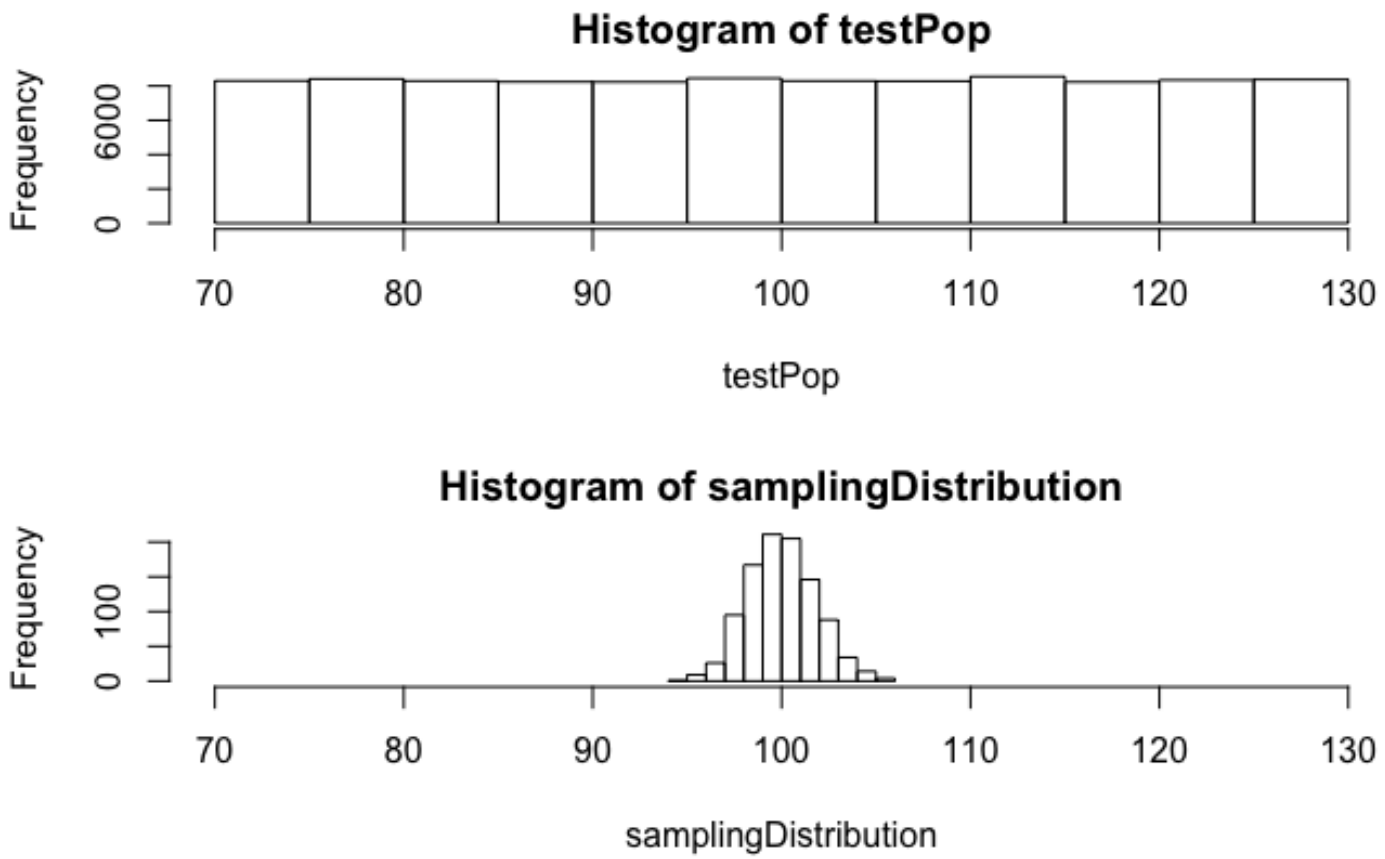
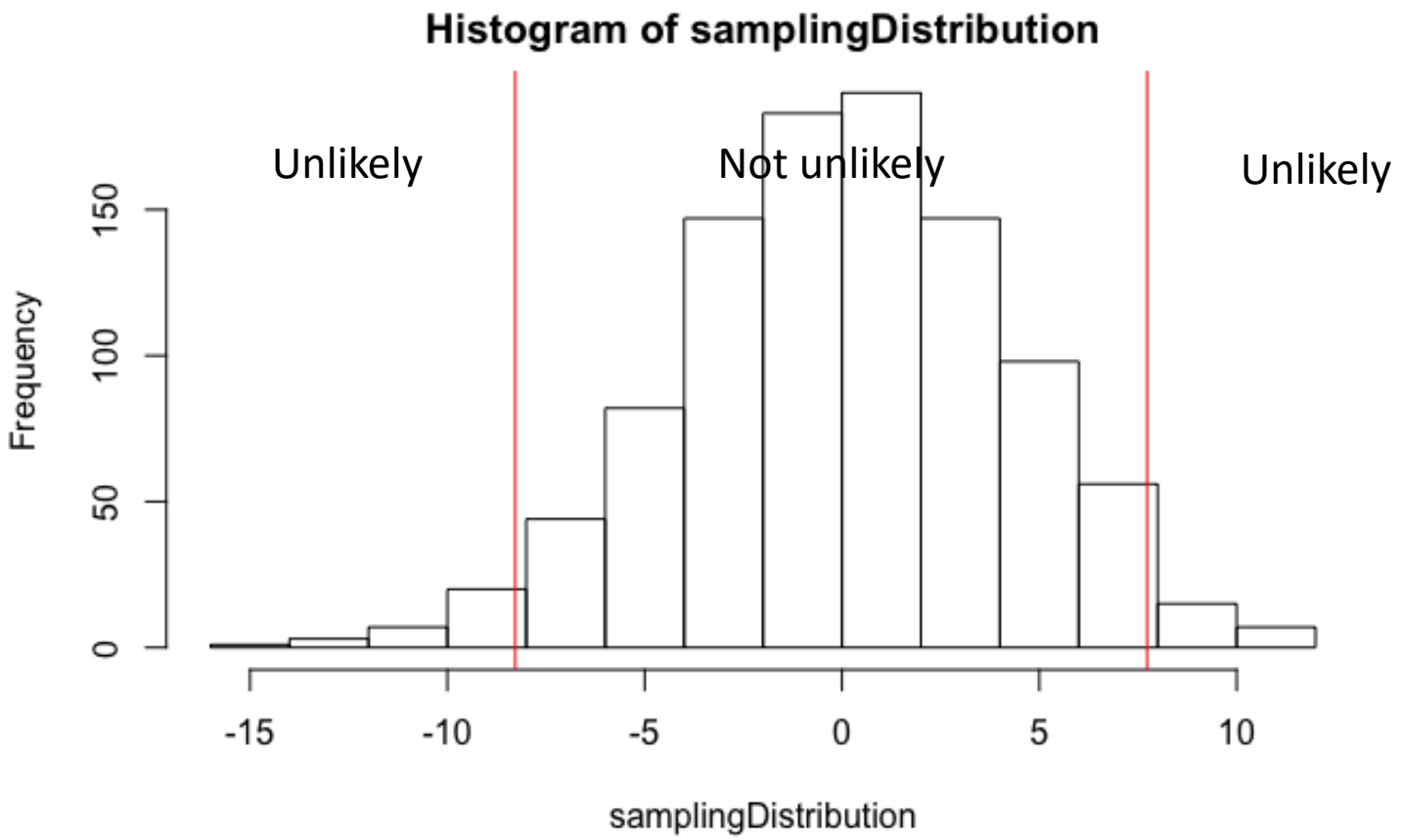
All handouts for this class: **https://tinyurl.com/IST772crowston**

IST772
Confidence Intervals (Week 4)

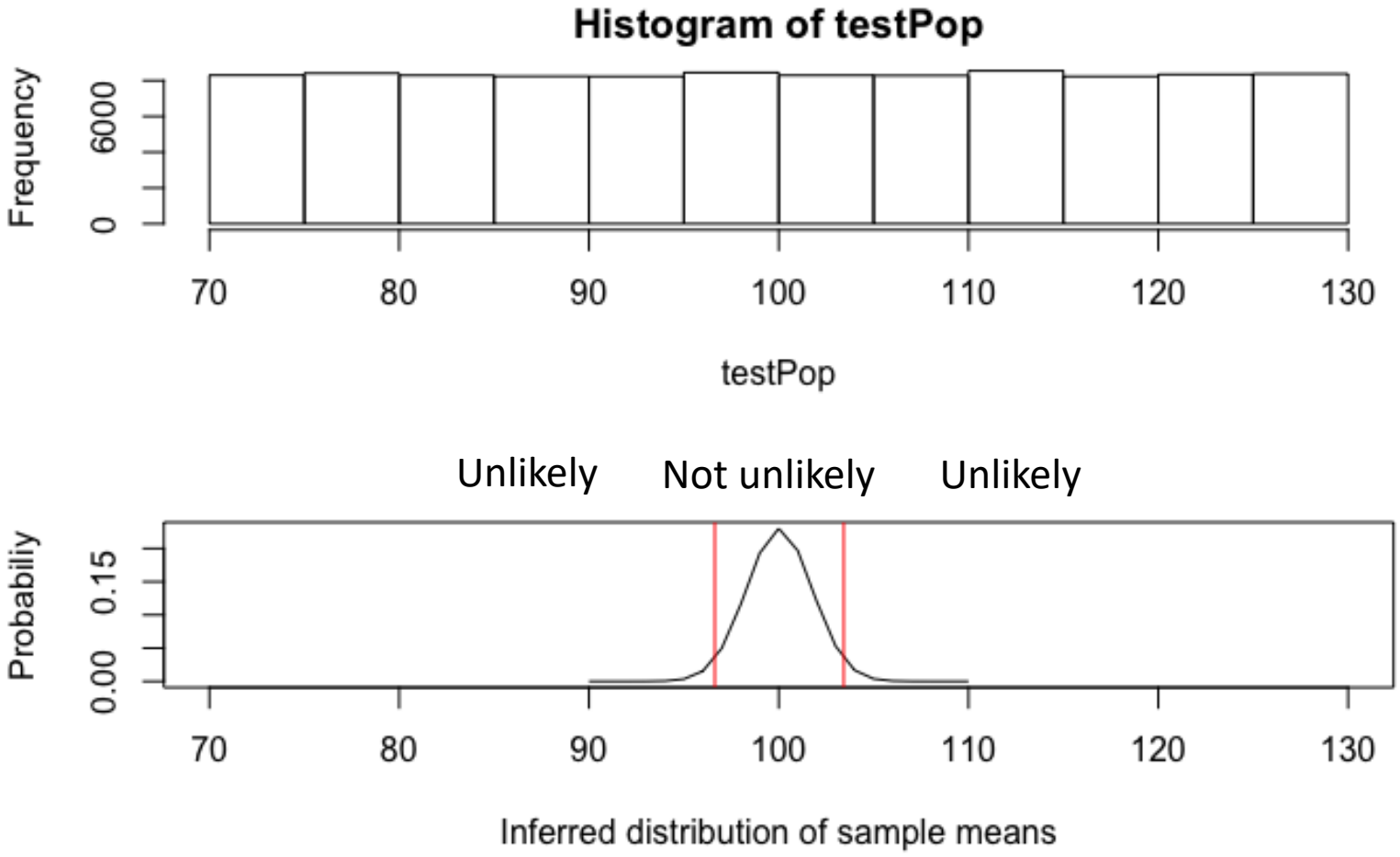
Distribution of sample means ≠ distribution of raw data



Likelihood of an observation



Confidence intervals

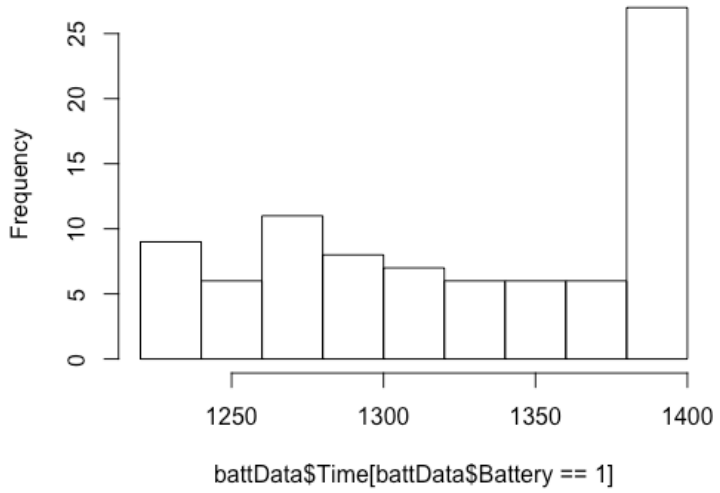


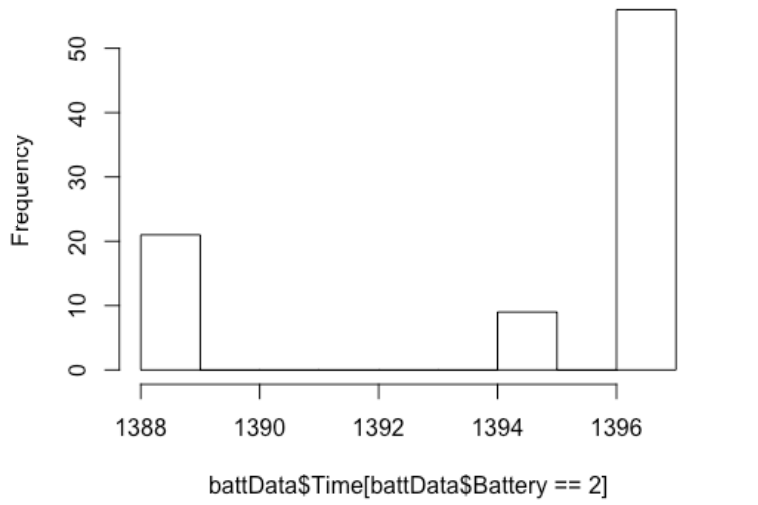
Week 4 practice exam

- Open the file `week4practiceexam.docx` on Blackboard and answer the questions in the document
- You will have 15 minutes to work on the exam
- We will debrief afterwards
- Points are given for taking the exam
- You can upload the exam on Blackboard if you want feedback on your answers; otherwise just submit “I do not need feedback”

Comparing battery life (time to discharge) for two types of batteries

Goal: Analyze the difference in group means, construct a confidence interval around the group mean difference, interpret the confidence interval correctly.





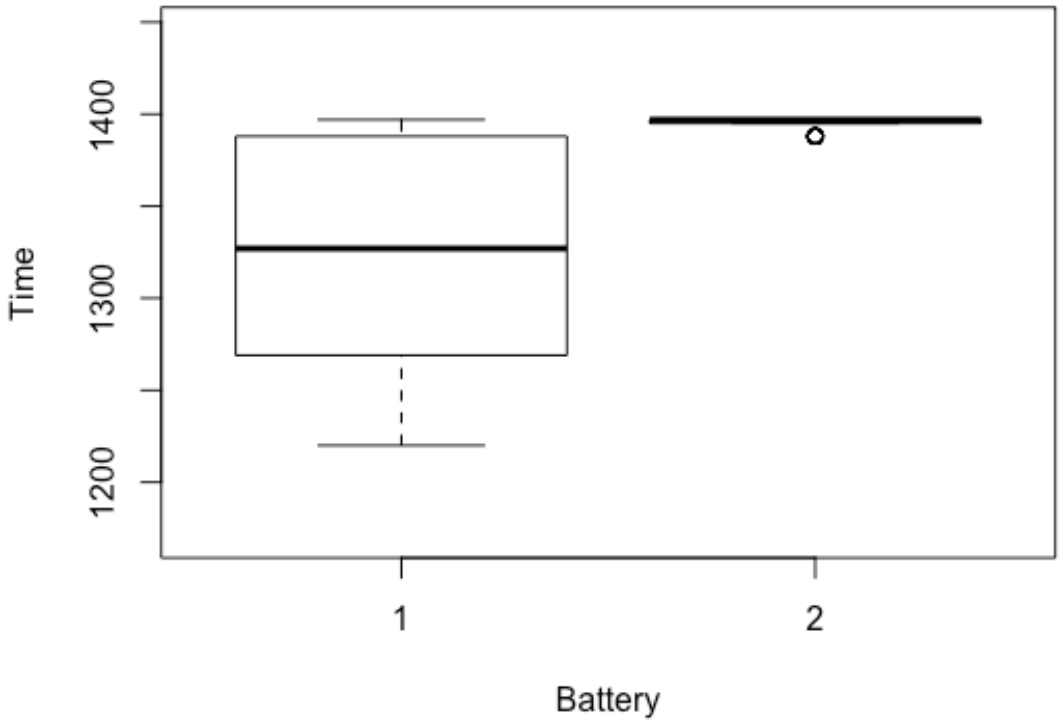
Breakout 1 – Constructing and Interpreting a Confidence Interval

- Open notebook `1. week4rstudio1.Rmd`
- Import “batterydata.r.csv” into R
- Run descriptive statistics and graphical diagnostics on the whole data set as well as the two groups separately
- Run a t-test to obtain the confidence interval
- Write a comment with a correct interpretation of the CI
- Share your code on https://codeshare.io/aJDyRX

Results of the t-test: Mean Diff = -71 minutes

Welch Two Sample t-test

data: Time by Battery
t = -11.252, df = 85.728, p-value < 2.2e-16
alternative hypothesis: true difference
in means is not equal to 0
95 percent confidence interval:
-83.49022 -58.41675
sample estimates:
mean in group 1 mean in group 2
1323.640 1394.593



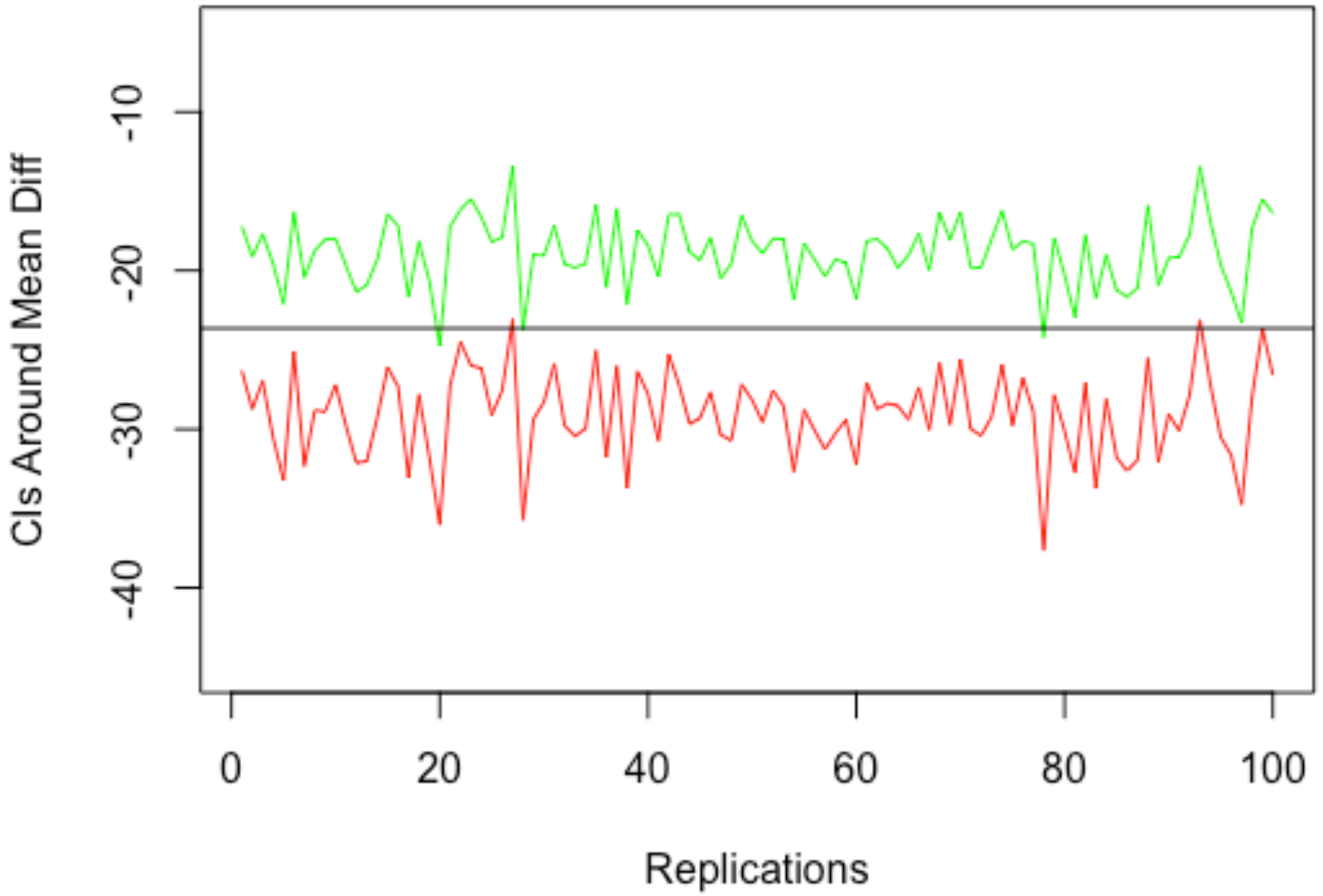
A Correct Interpretation

- We analyzed the difference in discharge time between n=86 NiCad batteries and n=86 Li-Ion batteries. Results showed a mean difference of -71 minutes, indicating that on average NiCad batteries reached a discharged state 71 minutes sooner than Li-Ion batteries.
- We constructed a 95% confidence interval around this mean difference, which ranged from -83.5 minutes to -58.4 minutes. **Note that this confidence interval may or may not contain the true population value.** The width of the confidence band, about plus or minus 12.5 minutes, gives some indication of amount of uncertainty around the point estimate of -71 minutes. To reduce this uncertainty, we would have to increase sample size, reduce variability in discharge times within groups, or both.

Breakout 2 – Repeated CIs from the Population

- Open 2. week4rstudio2.Rmd
- Import a whole population of battery test data
- Calculate the true population mean difference between the two types of batteries
- Construct a custom function to sample n=172 batteries and calculate a CI around the mean difference
- Replicate that function 100 times, saving the CIs
- Plot and interpret the results

Out of 100 replications, five CIs cross the line



The mean of this population, -23.6 minutes, is outside of the original CI you calculated.

Paper of the Week – Morey et al., 2016

- Discusses common misinterpretations of the confidence interval, including the "Fundamental Confidence Fallacy:"
  - *If the probability that a random interval contains the true value is* X%*, then the plausibility or probability that* ***a particular observed interval*** *contains the true value is also* X%*; or, alternatively, we can have* X% *confidence that the observed interval contains the true value.*


Psychon Bull Rev (2016) 23:103–123
DOI 10.3758/s13423-015-0947-8

THEORETICAL REVIEW

Note: This statement is false, for reasons discussed in the article.

The fallacy of placing confidence in confidence intervals

Richard D. Morey[1] · Rink Hoekstra[2] · Jeffrey N. Rouder[3] · Michael D. Lee[4] · Eric-Jan Wagenmakers[5]

Homework and Practice Exam

- Make sure you are using the updated syllabus that I distributed at the beginning of the semester (on the wall and in the handouts folder).
- The homework for week four is based on exercises 7 - 10 on page 66 (but be sure to answer the questions in the supplied notebook).