Inference for numerical data - exercises

E. Pastucha

October 2020

Here are 5 exercises to solidify your knowledge about t-tests. For each exercise go slowly through the steps:

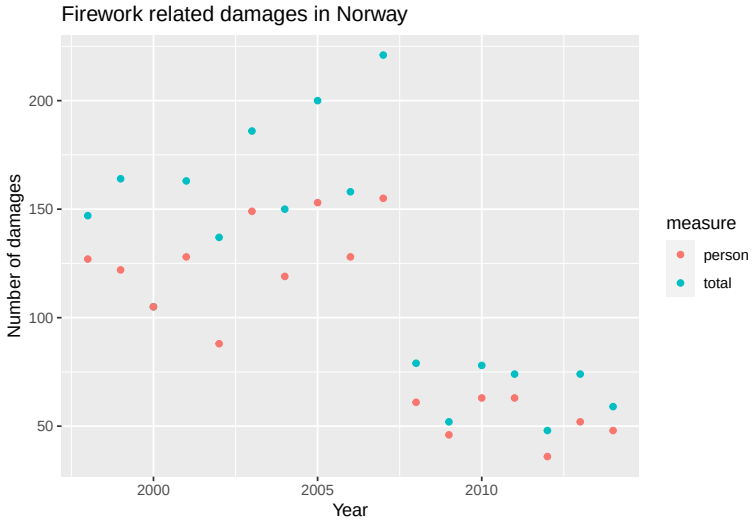
1. Get aquinted with the data - get some summary statistics, some plots.
2. Decide on the test (one sample, difference of means, paired data)
3. Comment whether the data is fulfilling CLT conditions.
4. Set up hypothesis.
5. Conduct the test.
6. Form conclussions.

1. Feeding chickens

File 'chickenweights.csv' contains information about 36 chickens, their weight and primary food. Compare the two weight distributions and examine whether there is a significant difference in their average weight.

2. Fireworks

The file *fireworks.csv* contains data about fireworks damages in Norway.



Use a t.test to compare the total number of damages per year for the two groups: group 1: years prior to and including 2007, group 2: years after 2007.

3. Plant growth

The file: *plantgrowth.csv* contains results of an experiment where plants were treated with different chemical compunds. All you're provided is category: treatment or control group, and weight of a plant. Choose a suitable statistical test and use that to determine whether there is a significant difference in the mean plant weight for the control group *ctrl* and the first treatment group *trt1*.

4. Major League Baseball players

The file *major_league_baseball_players.csv* contains information about major league baseball players.

a) For all player types (positions), calculate the average and standard deviation of the weight. Sort the list according to the average weight.

b) Use t-test to compare the average weight of a *Designated_Hitter* and a *Third_Baseman*

5. Nobel prizes

Data: *nobel_laureates.csv*

Choose a suitable statistical test and use that to determine whether there is a significant difference in the mean number of people sharing the award for Chemistry and the mean number of people sharing the award for Peace (for all available years).