

CE 385S – Stochastic Hydrology – Fall 2016

Homework Assignment #2 – Due Thursday 09/22/2016

All the data sets used in this homework assignment are posted on Canvas on Files -> Homework
Report all your test results both in terms of the test statistics as well as in terms of the p-value.

Problem 1 – Paired comparison

Given are values of data for counts of thunderstorms reported in the northeastern United States (x) and the Great Lakes states (y) for 21 years. We want to test whether or not the thunderstorm patterns in the two regions are similar.

- (a) Plot a box plot of x and y
- (b) Plot x and y on a scatterplot.
- (c) What can you say about similarities/differences among the two data sets based on (a) and (b)?
- (d) Test the hypothesis that x and y come from the same population with the signed rank test with a significance level 5%.
- (e) Test the same hypothesis with a paired t-test with a significance level 5%.

Problem 2 – Comparison of independent samples

Given are unit well yields, in gallons per minute per foot of water-bearing material for wells within valleys containing fracturing versus valleys with no fracturing. We want to test whether fracturing is associated with higher mean unit well yield with:

- (a) Rank sum test with a significance level 5%.
- (b) Two sample t-test with a significance level 5%.

Problem 3 – Comparison of groups

Discharge from pulp liquor waste may have contaminated shallow groundwater with caustic, high pH effluent. We want to determine whether the pH of samples taken from three sets of piezometers are all identical.

- (a) Plot the box plot of the three groups. Are there clear differences among them?
- (b) Test the similarity of the three groups with an ANOVA test and significance level 5%.
- (c) Test the similarity of the three groups with a Kruskal-Wallis test, significance level 5%.

Problem 3 - Data Analysis – Step 2!

You will find in your first hw some comments on the analysis you are planning for this class. This portion of the homework will be used to advance your analysis. If I had specific questions for you in the previous homework, please answer them as part of this homework. In the previous homework I asked for some general ideas on what you would like to analyze. I would like to narrow down some more specific tasks now. Try to list initial steps you could take in your analysis and if you already have seen some tools (hypothesis testing) to address those questions, start looking into that and perform your analysis. You can also use this homework for asking me any specific question you have on your data.