**CEE 574: Advanced Hydrology**

**Homework 1: Hypothesis testing and trend analysis using hydrologic data.**

Due date: April 18. 2013.

***Question 1:*** Annual peak flow data for the Elkhorn river in Nebraska is provided to you in an excel spreadsheet (Sheet1 of HW1_data.xlsx). Answer the questions below using the data.

(a) For a river restoration project in 1970, flood frequencies were needed. Assume you were the engineer in charge of the project in 1970, and estimate the magnitude of 5-, 10-, and 25-year floods first using the data from 1920 to 1970 and plot flood magnitudes (Q) as a function of their corresponding return periods (RP).
Note that an X year flood has an exceedance probability of 1 in X years. If you are not familiar with calculating exceedance probabilities please see the handout for flood duration curves in the homework folder.

(b) Now estimate the 5-, 10-, and 25-year floods again using the full length of the data this time. Also plot the Q-RP relationship for the whole data on the same figure you did for (a). Report how the two curves are different from each other.

(c) The state claims that ***the flood frequencies have changed in the river and the flood magnitudes calculated in the past should be re-evaluated.*** Using the data, develop and test the hypothesis for the probability of exceeding $Q_5$, $Q_{10}$, and $Q_{25}$ for both $\alpha$=0.05 and $\alpha$=0.1. Your null hypothesis should be that the probability of exceedances have not changed: $P(Q > Q_{RP}) = P_{RP}$, where $Q_{RP}$ and $P_{RP}$ are magnitude and exceedance probability of a flood discharge with a known return period calculated previously.

(d) Do you get consistent results for different $\alpha$ values? If not, what does this indicate?

Note: In hypothesis testing $\alpha$ is the level of significance or the exceedance probability of the test statistic, while p is the cumulative probability [P(T<t)] of the test statistic. In a two-tail test we want the test statistic to have a cumulative probability less than p=1- $\alpha$ /2 (upper bound) and higher than p=$\alpha$ /2 (lower bound) to remain in the mid portion of the pdf where 95% ($\alpha$=0.05) or 90% ($\alpha$=0.1) confidence interval of the data is.

***Question 2:*** Annual precipitation and runoff data for the Niobrara river near Sparks, NE is given in HW1_data.xlsx. In this region, an irrigation dam was constructed in 1964. The Nebraska wildlife service claims, based on the annual runoff records, that the dam resulted in reduced flows and therefore pose threat to fish. The irrigation district, on the other hand attributes the reduced runoff amounts to climate change, and argues that the irrigation project has not altered the natural course of the river flows with at any level of significance.
This is a typical example for comparisons between two independent data groups (before and after 1964 annual runoff) (Table 17.3.2). Your assignment is to do the analysis to

suggest whether the changes in runoff was due to climate change (largely due to precipitation) or potentially due to land use change.

(a) Use the two-sample t-test to test the null hypothesis, Ho: the means of the two groups are equal ($\alpha$=0.05). Both for runoff and precipitation in both time periods before and after 1964 (before and after the irrigation project).
(b) Use the rank-sum test to test for $\alpha$=0.05, to test the null hypothesis: Ho: the medians of the two groups are equal (again do this for both annual runoff and precipitation). Evaluate the consistence of your results and draw a conclusion. Please read the note below.

Note: When drawing your conclusion for the tests you should report (1) the hypothesis tested. (2) the calculated value of the statistic (t, Z, $X^2$, etc.), and its corresponding p value. (3) the $\alpha$ level of the test significance ($\alpha$=0.05, or 0.1), and a final statement: "The null hypothesis is Accepted/Rejected at the $\alpha$ significance level, indicating that *all work and no play makes Jack a dull boy*".

***Question 3***: The data for this question is also given in HW1_data.xlsx, where you will find annual mean discharge and peak flow of a river.

a) calculate the Kendall's correlation coefficient for annual mean discharge with time. Test the significance of this correlation using the non-parametric Kendall (or often known as Mann-Kendall) for $\alpha$=0.05. Your null hypothesis is Ho: the distribution of annual mean discharge does not change as a function of time.
b) Repeat (a) for peak discharge in this river.
c) Compare and summarize your conclusions in a paragraph. For both cases report $\tau$, Z and p of the data.

***Question 4***: In the excel spreadsheet you will find annual precipitation and runoff data for **four** basins. These basins are located in the same geography and climate. For a regional water resources development project you are given limited amount of funding to study one representative basin among the four basins, because your client claims that these basins behave similarly and therefore studying one representative basin in detail would be sufficient. Before you accept the client's claim, you would like test if the basins have an identical mean behavior. For this purpose, calculate the annual runoff ratio of the basins (runoff/precipitation) as a surrogate index for annual hydrologic behavior, and apply the ANOVA procedure in MATLAB for annual runoff. For test level use $\alpha$=0.05.