STAT 8004, Homework 1

Group # ... (Replace this)
Members: ... (Replace this)

Jan. 30, 2013

This homework is due Thu., 2013/02/06, 5:30pm.

Instructions: Generate a PDF file from it and submit the PDF file to blackboard. Each group should submit one file with file names **hw[number]-[groupnumber].pdf**. For exmaple, "hw01-1.pdf" for homeowrk 1 and group 1. Please also include your R code in the appendix.

Problem 1. (20 points) Suppose you lived in a world mimicing the movie "Pacific Rim". The scientists found that the time interval between two seperate Kaiju (a term in the movie, meaning "Giant Monster") invasion events is decreasing. For successive events, the time interval between two events is $y_1, \ldots, y_n$. We assume that $y_1, \ldots, y_n$ are independent random variables from exponential distribution with the expectation $\lambda, \lambda/2, \ldots, \lambda/n$ respectively. The scientists would like to gain information about λ so that they can help decide how many Jaeger's (a term in the movie, meaning "Mechanic Hunters") they need to fight Kaiju's.

a) Formulate a likelihood ratio test for testing $\mathrm{H}_0 : \lambda = 1$ versus $\mathrm{H}_1 : \lambda < 1$. For $(y_1, y_2) = (1.2, 0.4)$, would such test with size 0.20 accept or reject H_0?

b) Describe a procedure for forming a level 0.95 one-sided confidence interval of the form $(0, \theta_u)$ [you do not need to come up with a closed form expression and can express that you would need to calculate the quantiles of certain distributions and do a numerical search to form the confidence interval]. Use your procedure to find (approximately) a realized confidence interval of the form $(0, \theta_u)$ for the sample $(y_1, y_2) = (1.2, 0.4)$ (you may want to write a computer program for this).

Problem 2. (40 points) Jaffe, Parker and Wilson have investigated the concentration of several hydrophobic organic substances (such as hexachlorobenzene, chlordane, heptachlor, aldrin, dieldrin, endrin) in the Wolf River in Tennessee. Measurements were taken downstream of an abandoned dump site that had previously been used by the pesticide industry to dispose of its waste products.

It was expected that these hydrophic substances might have a nonhomogeneous vertical distribution in the river because of differences in density between these compounds and water and because of the adsorption of these compounds on sediments, which could lead

to higher concentrations on the bottom. It is important to check this hypothesis because the standard procedure of sampling at six-tenths of the depth could miss the bulk of these pollutants if the distribution were not uniform.

Grab samples were taken with a La Motte-Vandorn water sampler of 1 litre capacity at various depths of the river. This sampler consists of a horizontal plexiglas tube of 7 centimetres diameter and a plunger of each side which shuts the sampler when the sampler is at the desired depth. Ten surface, 10 mid-depth and 10 bottom samples were collected, all within a relatively short period. Until they were analysed the samples were stored in 1-quart mason jars at low temperature.

In the analysis of the samples, a 250-millilitre water sample was taken from each mason jar and was extracted with 1 millilitre of either hexanes or petroleum ether. A sample of the extract was then injected into a gas chromatograph and the output was compared against standards of known concentrations. The test procedure was repeated two more times, injecting different samples of the extract in the gas chromatograph. The average aldrin and hexachlorobenzene (HCB) concentrations (in nanograms per liter) in these 30 samples are given in the data.

a). (10 points) Load the data into R. Investigate the function *boxplot* and draw a box plot of average Aldrin and HCB concentrations at different depth of the river. Explain the plots.

b) (10 points). Build up a one-way ANOVA model to compare the average Aldrin concentrations at different depth of the river. Explain what the parameter means in your model. Also estimate the parameters in your model (You can use R).

c) (10 points). Repeat the same procedure in b) to analyze the average HCB concentrations at different depth. Please use a model different from b).

d). (10 points) List ANOVA tables to test whether the average Aldrin and HCB concentrations are the same across different depth of the river.