STA 114: Statistics

HW 9

Due Wed Nov 16 2011

1. Smile durations (in seconds) of a 8 week old baby are modeled as independent observations from $\mathsf{Uniform}(0, \theta)$, $\theta > 0$.
   (a) Describe the critical region of a size α ML test for $H_0 : \theta = \theta_0$ against $H_1 : \theta \neq \theta_0$ for some given number $\theta_0 > 0$.
   (b) Describe the critical region of a size α ML test for $H_0 : \theta \geq \theta_0$ against $H_1 : \theta < \theta_0$ for some given number $\theta_0 > 0$.
   (c) Calculate the p-value based on ML tests for testing $H_0 : \theta = 30$ against $H_1 : \theta \neq 30$ when observed data are: 10.4, 19.6, 12.8, 14.8, 1.3, 0.7, 5.8, 6.9, 8.9, 9.4.
   (d) Calculate the p-value based on ML tests for testing $H_0 : \theta \geq 30$ against $H_1 : \theta < 30$ with the same observations as in part (d).

2. Again consider the uniform model: $X_1, \cdots, X_n \overset{\text{IID}}{\sim} \mathsf{Uniform}(0, \theta)$, $\theta > 0$.
   (a) Show that, for any $\theta_0 > 0$, all ML tests of $H_0 : \theta \leq \theta_0$ against $H_1 : \theta > \theta_0$ are identical and have size 0.
   (b) Is $\theta \leq \theta_0$ a bad candidate to be labelled a null hypothesis? Explain.

3. Suppose the time interval (in minutes) between two successive eruptions of a geyser follows an $\mathsf{Exponential}(\lambda)$ distribution. The data available consists of n observed intervals $X_1, \cdots, X_n$. Find the (approximate) p-value based on ML tests for testing $H_0 : \lambda \leq 1/60$ against $H_1 : \lambda > 1/60$ (the null hypothesis says that the expected waiting time, $1/\lambda$, is at least an hour). Perform the ML tests (give accept/reject H_0 verdicts) of size 1%, 5% and 10%.

4. Consider data $X = (X_1, \cdots, X_n)$ and $Y = (Y_1, \cdots, Y_m)$ modeled as $X_i \overset{\text{IID}}{\sim} \mathsf{Normal}(\mu_1, \sigma^2)$, $Y_j \overset{\text{IID}}{\sim} \mathsf{Normal}(\mu_2, \sigma^2)$ with σ fixed.
   (a) Describe critical region of a size α test for $H_0 : \mu_1 = \mu_2$ against $\mu_1 \neq \mu_2$ based on the medians x_{med} and y_{med}.
   (b) Get the p-value based on these tests when X_i's and Y_j's, respectively, are bodyweight gains of 2 groups of rats put to a high and a low protein diet, with $\sigma = 15$ and observed data:

Diet	Observations
High	134 146 104 119 124 161 107 83 113 129 97 123
Low	70 118 101 85 107 132 94

5. We want to compare Duke students' weekly expenditure on food between years 2010 and 2011. We have data $X = (X_1, \cdots, X_n)$ from 2010 on n students, and data $Y = (Y_1, \cdots, Y_m)$ from 2011 on another m students. Consider the model $X_i \overset{\text{IID}}{\sim} \mathsf{Normal}(\mu_1, \sigma_1^2)$, $Y_j \sim \mathsf{Normal}(\mu_2, \sigma_2^2)$. The observed data are

Year	Expendtures
2010	139 210 75 131 184 140 140 127 100 129 145 145 120 237
2011	125 140 200 200 190 100 140 250 125 180 110 125 120 130 140 150 120 100 95 195 95 130

(a) Assume $\sigma_1 = \sigma_2$ (but the common value is unknown) and calculate the p-value based on the ML tests for testing $H_0 : \mu_1 = \mu_2$ against $H_1 : \mu_1 \neq \mu_2$. Perform the ML tests (give accept/reject H_0 verdicts) of size 1%, 5% and 10%.

(b) Repeat part (a) without assuming equal variances. Use Welch's tests.