SUTD ESD

Statistics 2017

Homework Assignment 4

Due: 5pm, Monday 24 April.

You may submit this assignment on *eDimension*, or in the homework box on level 7, Building 1.

Refer to the *Excel* spreadsheet for the data. Show working.

Question 1. In a study on the effect of vitamin B on learning, 12 matched pairs of children were randomly divided into two groups. One child in each pair received a vitamin B tablet every day and the other received a placebo tablet. The spreadsheet shows their gains in the scores on a standardized test over the course of the study. Let $\alpha = 0.05$.

(a) Find the p-value using the *sign test*, to determine if vitamin B improves scores on the test.

(b) Repeat part (a) using the *signed rank test*. Discuss any discrepancy in your results.

Question 2. In ANOVA for single factor experiments, define MST = SST/$(N-1)$. Is it possible that MST = MSA + MSE? Fully justify your answer.

Question 3. Water salinity measurements at three sites are given in the spreadsheet, and an ANOVA table is produced. However, one entry (marked by X) has been accidentally deleted.

(a) What can you conclude from the ANOVA F?

(b) Find X with help from the ANOVA table. (Hint: you can do this using any method you like, but check your answer.)

Question 4. Refer again to the sugar content example given in the spreadsheet. Use the Bonferroni method to determine which shelves have significantly different mean sugar content. Use $\alpha = 0.05$.

Question 5. Refer to the spreadsheet 'IQ', which records some adopted children's IQ with the socioeconomic status of their biological parents as well as adoptive parents.

(a) Sketch a line chart for the cell means, and from it comment on whether there is any significant interaction.

(b) Construct an ANOVA table for this two-factor experiment. What conclusions can you draw from it?

Question 6. Let $x_1, x_2, \ldots, x_{2n}, x_{2n+1}$ be iid observations from a *double exponential* distribution, with probability density function

$$f(x|\theta) = \frac{1}{2}\exp\big(-|x-\theta|\big).$$

Find the maximum likelihood estimate of θ. (Hint: Week 2.)

Question 7. In a particular computer game, each player controls a 'hero' that can be used to deal damage against other heroes. One hero is notorious for having an attribute which can be described as 'critical strike', namely, 15% of the time she deals much more than her usual damage. Another hero possesses an attribute known as 'bash', that is, 17% of the time he can disable the target as well as dealing extra damage.

Due to humans' lack of intuition for what is random, some of the game's attributes are *not* determined using a truly random process. To illustrate this, normal strikes and critical strikes (respectively, bashes) in a game are recorded in the spreadsheet: numbers indicate runs of normal strikes, while each critical strike (respectively, bash) is indicated by an X (respectively, B).

(a) Give a 95% confidence interval for the proportion of critical strikes, and for the proportion of bashes. Are they consistent with the purported rates?

(b) If critical strikes (respectively, bashes) occur truly randomly (and hence independently), what is the probability of (at least) two critical strikes (respectively, bashes) in a row? Find the expected number of two critical strikes (respectively, bashes) in a row, then briefly discuss whether the data seems consistent with this. (Hint: you may need to work with random variables in a non-trivial way.)

(c) Perform a Wald-Wolfowitz *runs test* with $\alpha = 0.05$ to determine if critical strikes (respectively, bashes) occur independently.