

ESM 206A Problem Set 1

A: *t*-test practice

For each of the following questions, (a) state the appropriate null and alternate hypotheses; (b) state which statistical procedure should be used to test the null hypothesis; (c) run the test, reporting the values of t , degrees of freedom, and P ; (d) draw conclusions based on $\alpha = 0.1$, $\alpha = 0.05$, $\alpha = 0.01$, and $\alpha = 0.001$.

1. A species of marine arthropod lives in seawater that contains calcium in a concentration of 30 mmole/kg of water. Thirteen of the animals are collected and the calcium concentrations in their coelomic fluid are found to be: 28, 27, 29, 29, 30, 30, 31, 30, 33, 27, 30, 32, and 31 mmole/kg. Test the appropriate hypothesis to conclude whether members of this species maintain a coelomic calcium concentration less than that of their environment.
2. Twenty-five intertidal crabs were placed in air at 24.3°C, and their body temperatures were measured as: 25.8, 24.6, 26.1, 22.9, 25.1, 27.3, 24.0, 24.5, 23.9, 26.2, 24.3, 24.6, 23.3, 25.5, 28.1, 24.8, 23.5, 26.3, 25.4, 25.5, 23.9, 27.0, 24.8, 22.9, 25.4. Test the appropriate hypothesis to conclude whether members of this species maintain their bodies at a different temperature than the air.
3. The menstrual cycles in a random sample of fifteen women, measured in days, were found to be: 26, 24, 29, 33, 25, 26, 23, 30, 31, 30, 28, 27, 29, 28. Test the appropriate hypothesis to determine whether the average human menstrual cycle equals a lunar month (29.5 days).
4. Over eight repeated experiments, the dissolving time (in seconds) of a drug in gastric juices was found to be: 42.7, 43.4, 44.6, 45.1, 45.6, 45.9, 46.8, 47.6. Test the null hypothesis that the mean dissolving time is no more than 45 seconds.
5. Using the data from the previous question, test the null hypothesis that the mean dissolving time is no *less* than 45 seconds.
6. Two different blood-clotting agents were tested. Each patient was given one of the drugs and the blood-clotting time (in seconds) measured. Drug B: 8.8, 8.4, 7.9, 8.7, 9.1, 9.6. Drug G: 9.9, 9.0, 11.1, 9.6, 8.7, 10.4, 9.5. Test the appropriate hypothesis to determine whether the drugs differ in their effectiveness.
7. A new fertilizer is claimed (by its manufacturer) to increase the growth rate of plants. Individual plants were given one of two fertilizers at random, and their heights (in cm) measured after one month. Present fertilizer: 48.2, 54.6, 58.3, 47.8, 51.4, 52.0, 55.2, 49.1, 49.9, 52.6. New fertilizer: 52.3, 57.4, 55.6, 53.2, 61.3, 58.0, 59.8, 54.8. Test the manufacturer's claim.

8. Using the following data, determine whether male and female turtles have the same mean serum cholesterol concentrations (measured in mg/100 ml).

<i>Male</i>	<i>Female</i>
220.1	223.4
218.6	221.5
229.6	230.2
228.8	224.3
222.0	223.8
224.1	230.8
226.5	

9. The file **Creekflow.xls** contains data on annual peak flows in CFS for two local creeks. Determine whether, on average, peak streamflow differs between Atascadero and Carpenteria Creeks. The two worksheets in the file have the same data arranged in different forms; which one you use will depend on which test you use.

B: Assessing normality

For each of the problems in part A, answer the following questions (remember that if your data are in multiple groups, you want to look at distributions within each group separately):

- Do you think the data are a sample from a normally distributed population?
- If your answer in (a) is No, what transformation do you think might make the data more normal?
- Try your transformation from (b). Are you satisfied that the new variable is a random sample from a normally distributed population?

C: Integrative problems

1. An industrial site has been found to be contaminated with a toxic chemical called TcCB, and the company responsible has performed a cleanup operation. The EPA has determined that concentrations in the soil above 6 parts per million (ppm) are unsafe.

Your job is to determine whether the cleanup has been successful. The company has taken a set of soil samples from the site and sent them to an independent lab for analysis. The results (in ppm) are contained in the file **tccb.xls** on the class website.

- Determine appropriate null and alternate hypotheses to make the decision. Explain why you chose these.
- If you make a type I error, what will be the consequence? If you make a type II error, what will be the consequence?
- What statistical procedure will you use to test the null hypothesis? Justify your choice in words.
- Run the test, report the results, and state your conclusions.

- e. Examine distributions of the data. Are you concerned that the assumptions of the test might be violated? Why or why not? If so, how might you fix it?
2. The Channel Islands are in a transition zone between two “biogeographic provinces” that are largely defined by average water temperature. You want to understand whether this transition has any effect on giant kelp, which is important habitat for many species. The dataset **kelp1.xls** on the course web page contains data on kelp density measured at two sites in the Channel Islands (a site at Anacapa Island and a site at Santa Rosa Island, spanning the range of the transition zone) at 14 times during 1998 and 1999. In each case, the density estimate is the number of stems within a randomly located 1 x 100 meter transect.

Your job is to determine whether the true kelp density differs between the two sites.

- a. State your null and alternate hypotheses.
- b. What statistical procedure will you use to test the null hypothesis? Justify your choice in words.
- c. Run the test, report the results, and state your conclusions.
- d. Examine distributions of the data. Are you concerned that the assumptions of the test might be violated? Why or why not? If so, how might you fix it?