Ma 320 Test 2 – Section 2

S. Sawyer — Washington University — March 23, 2001

Calculators and a single 8½ by 11 cribsheet can be used. All problems are 15 points unless indicated otherwise.

1. An oil services company wants to estimate the number of ships per day that come within two miles of a new oil platform. On 60 randomly selected days, the numbers of different ships that come that close to the platform have a sample mean of 7.2 and sample standard deviation of 2.9. Find a 95% confidence interval for the average number of ships per day that are expected to come within two miles of that site. (*Hint*: You can assume a Student-t distribution is the same as a normal if the number of degrees of freedom $d \geq 50$.)

Answer: The formula for a 95% normal confidence interval is $\overline{X} \pm 1.960\sigma/\sqrt{n}$. Here $\overline{X} = 7.2$ and we can assume $\sigma = s = 2.9$. The confidence interval is $7.2 \pm 1.960(2.9)/\sqrt{60} = (6.466,\ 7.934)$.

2. A company advertises that the average output voltage of a certain device that it sells is 130 volts with a standard deviation of 2.1 volts. In a shipment of 40 such devices, the average output voltage was 128.6 volts. Is this evidence that the true average output voltage is less than 130 volts at level of significance 0.05, assuming that the advertised value of 2.1 is correct for the standard deviation? What is the P-value?

Answer: Here $H_0 : \mu = 130$ and $H_1 : \mu < 130$. The population standard deviation $\sigma = 2.1$ is given. The level of confidence is $\alpha = 0.05$. The observed test statistic is $T = (128.6 - 130)/(2.1/\sqrt{40}) = -4.2163$. T has a standard normal distribution given H_0. Here $z(\alpha) = z(0.05) = -1.645$, so we reject H_0 and conclude that the true average voltage is significantly lower. The P-value is $P = P(T \leq -4.2163) = 2.48 \times 10^{-5}$.

3. (i) Scientists believe that the amount of the organic matter at the bottom of a particular lake is important for the health of the lake environment. Thicknesses were measured in centimeters at eight different sites in the lake in 1992 and again in 1998. Based on the data in the following table, is there evidence for a significant change in the amount of organic matter at these sites? What is the P-value? Assume that the data are normally distributed.

Site	1992	1998
1	42	43
2	51	38
3	53	53
4	65	56
5	71	62
6	36	33
7	32	35
8	33	32

(ii) What test statistic did you use? What is its distribution under hypothesis H_0?

Answers: (i) This is a paired design, so that the basic data are the eight paired differences $\{D_i\} = \{1, -13, 0, -9, -9, -3, 3, -12\}$. Here $H_0 : \mu_D = 0$ (or $\mu_A = \mu_B$) and $H_1 : \mu_D \neq 0$ (or $\mu_A \neq \mu_B$). The level of significance is $\alpha = 0.05$. By calculation $\overline{D} = -3.875$ and $s_D = 5.743$. The observed test statistic $T_D = \overline{D}/(s_D/\sqrt{8}) = -1.908$. T_D has a Student's-t distribution with 7 degrees of freedom given H_0. Here $t(7, 0.975) = 2.3646$ so we accept H_0 at $\alpha = 0.05$. The P-value is $P = P(|t(7)| \geq 1.908) = 2P(t(7) \geq 1.908) = 0.09805$.

(ii) The test statistic I used is $T_D = \overline{D}/(s_D/\sqrt{n})$. T_D has a Student's-t distribution with $n-1$ degrees of freedom given H_0.

4. (i) A factory guarantees that a particular machine part is sufficiently uniform that the standard deviation of its diameter is no more than 20 millimeters. A sample of 10 parts had a sample standard deviation of 30 millimeters. Is this evidence that the true standard deviation is greater than 20mm? What would you conclude at level of significance 0.05? What is the P-value?

(ii) What test statistic did you use? What is its distribution under hypothesis H_0?

Answers: (i) Here $H_0 : \sigma = 20$ and $H_1 : \sigma > 20$. The observed test statistic $T = 9 \times 30^2/20^2 = 20.25$. T has a chi-square distribution with 9 degrees of freedom given H_0. Here $\chi^2(9, 0.95) = 16.92$ so that we reject H_0 and conclude that the standard deviation is greater than 20mm at level of significance $\alpha = 0.05$. The P-value is $P = P(\chi^2(9) \geq 20.25) = 0.0164$.

(ii) The test statistic I used is $T = (n-1)s^2/\sigma^2$. T has a chi-square distribution with $n-1$ degrees of freedom given H_0.

5. The magnitudes of earthquakes in California that can be detected at a particular listening station can be modeled as having an exponential distribution with mean 3.2, as measured on the Richter scale. Find the probability that three or more of the next ten earthquakes will have magnitude 7.0 or larger. (*Hint*: This problem involves both an exponential and a binomial distribution. Recall that the *mean* of an exponential is one over its *rate*.)

Answer: The magnitude L for one earthquake satisfies $P(L > t) = \exp(-t/\mu)$ so that $P(L > 7.0) = \exp(-7.0/3.2) = 0.1122$. Now find $P(T \geq 3)$ where T has a binomial distribution with $p = 0.1122$ and $n = 10$. Using a TI-83 calculator, $P(T \leq 2) =$ binomcdf(10,0.1122,2) $= 0.9073$ so that $P(T \geq 3) = 1 - P(T \leq 2) = 0.0927$ since T takes on integer values only.

6. Essay (5 points) Explain the following terms: P-value, level of significance, and power. Given an example to clarify your statements. What is the relationship between the P-value and the level of significance?

7. (Extra Credit) Six different species of parrots are known to live in both inhabited and uninhabited forests in a tropical country. It is thought that parrots living in inhabited areas might tend to be larger. The average weights of parrots of these six species in the two types of forest are:

Parrot Species	Inhabited Forest	Uninhabited Forest
1	168	148
2	85	88
3	66	73
4	67	68
5	62	53
6	75	69

Is there is significant different in the weights of these parrots between inhabited and uninhabited areas? Carry out a Wilcoxon Signed Rank test to find out. What is the sum of the signed ranks? What is the (two-sided) P-value?

Answer: This is a paired design, so that the basic data are the six paired differences $\{D_i\} = \{20, -3, -7, -1, 9, 6\}$. Here $H_0 : \mu_D = 0$ (or $\mu_A = \mu_B$) and $H_1 : \mu_D \neq 0$ (or $\mu_A \neq \mu_B$), which we test using the signed ranks of the D_i instead of the D_i themselves. The level of significance is $\alpha = 0.05$. (This is understood if the level of significance isn't stated explicitly.)

The absolute values are $\{20, 3, 7, 1, 9, 6\}$ and the corresponding ranks are $\{6, 2, 4, 1, 5, 3\}$. The signed ranks are the ranks with the original signs $\{R_i\} = \{6, -2, -4, -1, 5, 3\}$. By calculation, $\overline{R} = 1.167$ and $s_R = 4.07$. The observed test statistic for the signed ranks is $T_R = 1.167/(4.07/\sqrt{6}) = 0.702$. T_R is assumed to have a Student's-t distribution with 5 degrees of freedom given H_0. Here $t(5, 0.975) = 2.5702$, so that we accept H_0 and conclude that there is no evidence for a difference in parrot weights. The P-value is $P = 2P(t(5) \geq 0.702) = 0.5140$.