Math 20340: Statistics for Life Sciences
Fall 2009 Mid-semester Exam 3 — Solutions to practice problems

1. Achievement test scores of all high school seniors in a certain state have mean 60. A random sample of 100 students from one large high school had a mean score of 58, with sample standard deviation 8. A researcher want to see if this data provides sufficient evidence to suggest that this high school is inferior.

   a): State the null and alternative hypotheses for this test.

   Solution: Let μ be the average score for the entire high school.

   $H_0 : \mu = 60$

   $H_a : \mu < 60$

   b): Calculate the value of the test statistic.

   Solution: Test statistic is $\frac{\bar{x}-60}{8/\sqrt{100}} = -2.5$

   c): Calculate the p-value of the test statistic.

   Solution: p-value is $P(z \leq -2.5) = .0061$.

   d): Calculate the critical value for rejecting the null at 5% significance.

   Solution: It's a one-sided test, so critical value is -1.645.

   e): Is there sufficient evidence to suggest that this high school is inferior, at 5% significance?

   Solution: Yes; p-value is smaller that $.05$, test statistic is more negative than -1.645.

   f): Is there sufficient evidence to suggest that this high school is inferior, at 1% significance?

   Solution: Yes; p-value is smaller that $.01$.

2. A coin is called *fair* if, when tossed repeatedly, the proportion of times it comes up Heads is in the long run .5. Charlie has a theory that the coin used by the umpire at the start of Saturday's game against Syracuse was not fair. To test his theory, he tosses the coin 100 times. It comes up Heads 61 times. At 5% significance, is there enough evidence to support Charlies's feeling that the coin is not fair? State clearly your null and alternative hypothesis.

   Solution: Let p be the proportion of times that this particular coin comes up Heads.

   $H_0 : p = .5$

   $H_a : p \neq .5$

   Test statistic: $\frac{\hat{p}-.5}{\sqrt{.5*.5/100}} = 2.2 \ (\hat{p} = .61)$

   p-value: $P(z > 2.2 \text{ or } z < -2.2) = .0278$.

   There's evidence to accept Charlie's claim at 5% significance, but not at 1% significance.

3. Salmon grown at a commercial hatchery have weights that are normally distributed. When the "Norwegian method" is used, the adult salmon have mean weight 7.6 pounds. A hatchery claims that a modification to the method that they are using increases the average weight. Suppose a random sample of 16 fish grown using the new method yielded an average weight of 8.1 pounds with a sample standard deviation of 1.2 pounds.

a): State the null and alternative hypotheses for this test.

Solution: Let μ be average weight of salmon using new process.

$H_0 : \mu = 7.6$

$H_a : \mu > 7.6$

b): Write down the test statistic.

Solution: Test statistic: $\frac{8.1-7.6}{1.2/\sqrt{16}} = 1.666...$

c): What is the distribution of the test statistic?

Solution: It's a t distribution with 15 degrees of freedom

d): What is the critical value for rejecting the null at 5% significance?

Solution: From the t table, $t_{.05} = 1.753$ (it's a one-sided test)

e): Is there strong enough evidence to accept the hatchery's claim at 5% level of significance?

Solution: No, because the test statistic is not as large as the p-value.

4. **a)** A certain null hypothesis H_0 is tested against an alternative H_1, and accepted at 5% significance. With the same data, will H_0 be accepted at 2% significance?

   Solution: Since we accept H_0 at 5%, the p-value is bigger than .05; so we also accept at 2%.

   b) Explain what is meant by the p-value of a test statistic. (Your answer should not include the words "critical value" or "significance level".)

   Solution: The p-value of a test statistic is the probability of observing a value of the test statistic as extreme or more extreme as actually observed, *if the null hypothesis is true*.

   c) What is the *power* of a hypothesis test? How does the power change as the true value of the mean changes?

   Solution: The power of a test is the probability that the test will accept the alternative, when the alternative is true. The power of the test increases as the difference between the true mean and the null-hypothesized mean increases.

5. Professor G. believes that students who skip Friday afternoon classes don't perform as well on exams as those who show up diligently. His students (especially those who like to start the weekend watching a Friday afternoon soccer game) are skeptical, and ask for proof. So Prof. G. points out that on the last midterm in the large calculus class that he teaches, the scores of 30 randomly selected students who always attend class on Friday had a mean of 78 and a sample variance of 40, while the scores of 30 randomly selected students who never attend class on Friday had a mean of 75 and a sample variance of 45. Has the Professor proved his point?

   Be sure to state what your null and alternative hypotheses are, and at what significance level you are drawing your conclusion.

   Solution: Let μ_1 be average score among students who attend class of Fridays, and μ_2 the average among those who do not.

   $H_0 : \mu_1 = \mu_2$

   $H_a : \mu_1 > \mu_2$

Test statistic: $\frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_1^2}{n_1}}} = \frac{78 - 75}{\sqrt{\frac{40}{30} + \frac{45}{30}}} = 1.78....$

p-value: $P(z > 1.78) = .0375$.

There's evidence to accept Prof. G.'s claim at 5%, but not at 1% significance

6. How many hours a night does the average Notre Dame senior sleep during finals week? I want to estimate this by taking a random sample. Back in May I polled 18 randomly selected seniors, and got the following answers to the question "how many hours did you sleep last night?": 1 person said 4 hours, 2 said 4.5 hours, 4 said 5 hours, 7 said 5.5 hours, 3 said 6 hours and 1 said 6.5 hours.

a) Is there reason to believe from this data that the population is close to being normally distributed?

Solution: A histogram of the data clearly suggests a distribution that is roughly normal.

b) Calculate $\bar{x}$ and s for this data.

Solution: $\bar{x} = 5.333...$ and $s = .618....$

c) Using an appropriate t-distribution, construct a 90% confidence interval for the true average number of hours that a senior sleeps per night during finals week.

Solution: Since there are 18 sample points we will use a t-distribution with $df = 17$. A 90% confidence interval is
$$(\bar{x} - t_{.05}s/\sqrt{n}, \bar{x} + t_{.05}s/\sqrt{n}) = (5.079..., 5.586...).$$

7. There are two processes available to produce a certain key ingredient for a drug. The length of time needed to produce 1mg of this ingredient is in each case known to be normally distributed. A researcher wants to see if there is evidence to suggest a difference in the mean production times. Because of cost constraints, she can only study limited samples. In eight runs of process A, she found the average time to be 11.2 minutes, with sample standard deviation 2.4. In 12 tries of process B, she found the average time to be 9.6 minutes, with sample standard deviation 2.1.

a) Do a test at 5% significance to see if there is a difference between the mean production times for the two processes.

Solution: First we need the pooled estimator for the variance; it is

$$s^2 = \frac{7 * (2.4)^2 + 11 * (2.1)^2}{18} = 4.935.$$

Letting μ_A be the true average for process A and μ_B the true average for process B, we are testing $H_0 : \mu_A = \mu_B$ against $H_a : \mu_A \neq \mu_B$. The (standardized) test statistic for this test is

$$\frac{\bar{x}_A - \bar{x}_B}{\sqrt{\frac{s^2}{n_1} + \frac{s^2}{n_2}}} = \frac{11.2 - 9.6}{\sqrt{\frac{4.935}{8} + \frac{4.935}{12}}} = 1.577....$$

From a t-table we see that the 10% critical value for a t-distribution with 18 degrees of freedom is 1.33 and the 5% critical value is 1.734, so the p-value for this test is between .1 and .2; this leads us to accept the null.

b) What assumption do you have to make about the population variances to make this test valid?

Solution: We assumed that the two population variances were equal (a reasonable assumption given the sample variances).

3

8. Do women live longer than men? Here's data on the life expectancy (at birth) of 7 randomly selected UN countries, broken down by men and women (the data is from the CIA World Factbook, 2008):

	Yemen	Tanzania	Tonga	Syria	Brunei	Burundi	Andorra
Men	60.61	49.41	67.6	69.27	73.12	50.48	80.35
Women	64.54	52.04	72.76	72.02	77.59	52.12	85.14

Do a paired-difference test on this data to see if it provides evidence for women having a greater life expectancy than men.

Solution: The seven data points for the difference (Women - Men) are

$$3.93 \quad 2.63 \quad 5.16 \quad 2.75 \quad 4.47 \quad 1.64 \quad 4.79$$

with sample mean 3.63 and sample standard deviation 1.30. Doing a t-test with $df = 6$ to test the alternative hypothesis that the true average difference is greater than 0 we have a test statistic of $3.63/(1.3/\sqrt{7}) = 7.34$. Since $t_{.005} = 3.707$, this is very highly significant; we definitely reject the null.