Stat 5102 (Geyer) Spring 2012
Homework Assignment 5
Due Wednesday, February 29, 2012

Solve each problem. Explain your reasoning. No credit for answers with no explanation. If the problem is a proof, then you need words as well as formulas. Explain why your formulas follow one from another.

5-1. In each of the situations below explain whether a one-tailed or a two-tailed test is more appropriate. If you can't tell, give arguments for both sides.

(a) Two groups are considered independent samples from the population of interest. One is given a treatment for pain, the other is given a placebo, which looks and feels like the treatment, but does not contain the active ingredient and hence should not work. The response is a quantitative measure of each individuals report of pain relief. The response will be considered approximately normal and a two-sample t test done.

(b) Same as above, except there is no placebo. Two active treatments are compared. One is aspirin; the other is a new medication being considered for approval.

(c) Same as above, except the two active treatments being compared are both new.

5-2. Suppose the following data

$$\begin{array}{rrrrrrr} -0.33 & 3.52 & -0.08 & 3.36 & 7.26 & 3.56 & 7.46 \\ 7.14 & 3.73 & 4.64 & 0.05 & 1.06 & 5.81 & 3.04 \end{array}$$

are differences between treatment and control in a paired comparison experiment. We wish to test the null hypothesis that unknown true difference of population means is zero using an upper-tailed test.

(a) Perform a t test and interpret the P-value.

(b) Perform a sign test and interpret the P-value.

(c) Perform a signed rank test and interpret the P-value.

If you want to use R on this problem, the data are in the file

`http://www.stat.umn.edu/geyer/5102/data/prob5-2.txt`

5-3. Two groups in physics lab have been measuring the density of aluminum at room temperature (20° C). They got the following summary statistics

	n	$\overline{X}_n$	S_n
Group I	10	2.792	0.241
Group II	8	2.538	0.313

(Units are grams per cubic centimeter.) Assume the measurements for group I are IID $\mathcal{N}(\mu_1, \sigma_1^2)$ and the measurements for group II are IID $\mathcal{N}(\mu_2, \sigma_2^2)$. We want to perform a test of $H_0 : \mu_1 = \mu_2$ versus $H_1 : \mu_1 \neq \mu_2$. Perform Welch's approximate test, come as close as you can to the P-value.

If you want to use R on this problem, the data are in the file

```
http://www.stat.umn.edu/geyer/5102/data/prob5-3.txt
```

5-4. Develop the hypothesis test dual to the two-independent-sample confidence interval for binomial data developed in problem 4-5. Suppose we have two independent samples of size m and n from two different populations. We are interested in the fractions p and q of the populations that have a certain property. We estimate these proportions by the sample proportions $\hat{p}_m$ and $\hat{q}_n$ which are the fractions of the first and second samples having the property. We know from problem 4-5 that

$$\frac{(\hat{p}_m - \hat{q}_n) - (p - q)}{\sqrt{\frac{\hat{p}_m(1-\hat{p}_m)}{m} + \frac{\hat{q}_n(1-\hat{q}_n)}{n}}}$$

is asymptotically standard normal. Describe how to do one and two tailed tests concerning the parameter of interest $p - q$.

5-5. Develop a hypothesis test for the difference of two exponential distributions. Suppose we have two independent samples of size m and n both of which are assumed IID exponential but with possibly different rates λ_1 and λ_2.

(a) Develop an asymptotically pivotal quantity which is a function of the parameter of interest $\lambda_1 - \lambda_2$ and the sample means $\overline{X}_m$ and $\overline{Y}_n$ of the two samples.

(b) Describe how to do one and two tailed tests concerning the parameter of interest $\lambda_1 - \lambda_2$.

(c) Suppose the data are $m = 30$, $n = 40$, $\overline{X}_m = 21.7$, and $\overline{Y}_n = 33.3$. Compute the P-value for a two-tailed test of the null hypothesis $H_0 : \lambda_1 = \lambda_2$ and interpret the P-value.

5-6. Suppose I have taken a random sample of size 100 of ears of corn from a field. My sample has mean ear length of 6.13 inches and standard deviation 1.44 inches. This gives me a 95% confidence interval for the true mean ear length all the corn in the field of 6.13 ± 0.28 inches.

Suppose I want a more accurate 95% confidence interval with a half-width (plus-or-minus) of 0.10 inches. What sample size would I need to get that?

5-7. Suppose I intend to collect data about the effect of coaching on SAT scores. The data will be SAT scores for individuals before and after taking a cram course. Suppose the test-retest variability without coaching is known to be about 50 points (assume this is the population standard deviation).

(a) How large a sample size do I need to have a power of 0.95 of detecting a true mean difference due to coaching as small as 10 points (the null hypothesis being no difference) at the 0.05 significance level? The test will be an upper-tailed test, since we expect that coaching cannot hurt. The sample size will be large, so use the power calculation based on the asymptotic normal distribution for this part.

(b) How large a sample size do I need to have a power of 0.95 of detecting a true mean difference due to coaching as small as 40 points (the null hypothesis being no difference) at the 0.05 significance level? Again, the test will be an upper-tailed test. The sample size will be small, so assume before-after differences of scores are normally distributed and use the power calculation based on the noncentral t distribution for this part.

5-8. In the following table

1.13	−1.12	−1.30	1.16	−0.19
−1.18	0.12	0.02	−1.11	0.35
−0.49	−0.11	−0.45	−0.17	−1.66
2.70**	0.03	0.14	−1.64	0.61
−0.35	1.80*	2.65**	−0.73	−1.32

* $P < 0.05$, ** $P < 0.01$

the asterisks are just footnote symbols, but tables like this are so common that no one familiar with the literature needs to look at the footnote. One star means "significant" (statistically significant at the 0.05 level), and two stars means "highly significant" (statistically significant at the 0.01 level). The stars are supposed to indicate the interesting results.

In this table suppose all twenty-five numbers are z-scores for different one-tailed, upper-tailed tests. The stars in the table do not reflect any correction for multiple testing. That is a z-score is declared "significant" (gets a star) if $z \geq 1.645$ and is declared "highly significant" (gets two stars) if $z \geq 2.326$. Here 1.645 and 2.326 are the one tailed 0.05 and 0.01 z critical values.

(a) What critical values should replace 1.645 and 2.326 in order to apply a Bonferroni correction to this multiple testing situation?

(b) What would the result of the Bonferroni correction be in terms of stars?

5-9. For the data in problem 5-2, which is also in the URL

`http://www.stat.umn.edu/geyer/5102/data/prob5-2.txt`

compute a nonparametric confidence interval for the median difference. Note that this is not the same as the difference of medians — the data given are $Z_i = X_i - Y_i$ and we want a confidence interval for the median of the Z_i, which is not the same as the difference of the medians of the X_i and Y_i.

(a) First assume only that the distribution of the differences is continuous and calculate a confidence interval associated with the sign test.

(b) Then assume only that the distribution of the differences is continuous and symmetric, in which case the median is also the center of symmetry, and calculate a confidence interval associated with the signed rank test.

In each case calculate a conventional two-sided interval that has confidence level at least 95% but no larger than 95% than it has to be. In each case state the exact confidence level that is achieved.

5-10. Suppose we are doing an upper-tailed sign test with null hypothesis with sample size 20.

(a) Suppose the test statistic is $T = 16$, that is, sixteen of the data values are above the parameter value hypothesized under H_0, compute the conventional P-value and interpret it.

(b) Same except compute the fuzzy P-value and interpret it.