MATH 476 Statistics

Fred J. Hickernell **Spring 2006**
Final Exam **Thursday, May 11**

Instructions:

i. *This exam consists of FOUR questions. Answer all of them.*

ii. *The data and situations portrayed in this test are all realistic. Some are fictitious.*

iii. *This exam is closed book, but you may use 2 double-sided letter-size sheets of notes, the statistical tables from the text, and JMP.*

iv. *Write down test statistic values, p-values, or other key output from JMP where appropriate.*

v. *Show all your work to justify your answers. Answers without adequate justification will not receive credit.*

1. (25 marks)
We Serve Inc. provides network servers to businesses in Chicago. They want to demonstrate that the mean time to failure of their servers is at least 100 hours. They perform an experiment where they measure the first n network server failure times. Assume that these are i.i.d. random variables $X_1, \ldots, X_n \sim \text{Exponential}(\mu)$, and let $Y_n = X_1 + \cdots + X_n$ be the total time until the n^{th} server failure. We Serve Inc. sets a null hypothesis of $H_0 : \mu \leq 100$, an alternative hypothesis of $H_a : \mu > 100$, and a rejection region for Y_5 of $[1000, \infty)$.

   a) Why did We Serve Inc. set $H_0 : \mu \leq 100, H_a : \mu > 100$ rather than $H_0 : Y_5 \leq 100, H_a : Y_5 > 100$?

   Answer: Hypothesis tests must be about population parameters, not sample quantities.

   b) Why did We Serve Inc. set $H_0 : \mu \leq 100, H_a : \mu > 100$ rather than $H_0 : \mu > 100, H_a : \mu \leq 100$?

   Answer: They wish to reject H_0, which is a stronger statement than not rejecting H_0.

   c) What is the significance level of this test?

   Answer: Since the X_i are exponential random variables, it follows that $Y_n/\mu \sim \text{Gamma}(n)$. Thus, the significance level for this test is

   $$\alpha = \Pr(Y_5 \geq 100 | H_0 \textit{ is true}) = \Pr(Y_5 \geq 1000 | \mu = 100) = 1 - F_{\text{Gamma}}(10, 5) = 0.0293$$

   where $F_{\text{Gamma}}(y, n)$ is the probability distribution function for a Gamma(n) *random variable. See* `ExponentialPowerFinal06.jmp` *for the calculation.*

   d) Keeping in mind that that this test rejects the null hypothesis if $Y_5 \geq 1000$, sketch its power function as a function of μ.

   e) Sketch the power function for the test that rejects the null hypothesis if $Y_5 \geq 2000$. What are the advantages and disadvantages of this test over the first test?

f) Sketch the power function for the test that rejects the null hypothesis if $Y_{10} \geq 2000$. What are the advantages and disadvantages of this test over the first test?

Answer: The three power functions are plotted using JMP. The second test is more conservative by setting a stricter criterion for rejection. Thus, a false rejection is less likely, but it is also harder to reject the null hypothesis even if it is false. The third test is better than the first since the chance of a false rejection is less likely, but also the chance of rejection when the null hypothesis is false is greater. The disadvantage is that more data are needed.

2. (25 marks)
Windy City Co. produces cooling fans to prevent network servers from heating up and hopefully reducing the chance of crashes. Windy City Co. tests their fans with fifteen different makes and models of computer servers. For each model of computer tested Windy City Co. measures the time to failure of two identical servers, one fitted with their cooling fan and one without:

No Fan	221	187	312	85	246	152	109	81	12	40	60	445	101	19	135
Fan	390	230	545	137	248	245	425	94	56	232	91	408	133	69	200

The in-house statistician at Windy City performs a two-sample t-test and arrives at a p-value of 0.0894, which is too large to reject the null hypothesis that there is no difference in mean failure times for the servers with and without fans. You are hired as a consulting statistician to see what can or should be done.

a) Is this the correct hypothesis test and conclusion? Explain why or why not.

Answer: This is not the correct test. This is a problem of matched pairs since the two servers in each pair are identical except for the fan.

b) If this is not the correct hypothesis test and conclusion, then perform the proper test and arrive at the correct conclusion.

Answer: The correct matched pairs test is shown in `ServerFailCoolFan.jmp`*. Now the p-value is* 0.0040 *for the two-sided alternative. One may reject the null hypothesis that that there is no difference in favor of the alternative that the fan lengthens the time failure.*

c) What assumptions were made in using a t-test? Can those assumptions be weakened and still arrive at the same conclusion?

Answer: The t-test assumes that the difference of failure times is normal. In fact, this assumption appears to be violated (see `ServerFailCoolFan.jmp`*), but the Wilcoxon Sign-Rank test gives the same conclusion of rejecting the null hypothesis that that there is no difference in favor of the alternative that the fan lengthens the time to failure.*

3. (25 marks)
Teachers would like parents to be involved, but not too much or too little. Teachers from several classes in grades 1, 6, and 9, were asked to estimate the numbers of parental involvement in their children's education

	Too Little	About Right	Too Much
Grade 1	3	87	14
Grade 6	16	72	18
Grade 9	35	30	28

a) Is there evidence of a relationship between the level of parental involvement and the level of study of the children? Are there any features of the data that might affect your conclusion.

Answer: This is a contingency table. Entering the data and performing the analysis in JMP by `ParentalInovlvement.jmp` *gives a p-value* < 0.0001*, indicating strong evidence for a relationship between the level of parental involvement and the level of study of the children. The number of observations in the upper left corner is smaller than one would like.*

b) Consider the parents of Grade 9 students. Construct a 95% confidence interval for the proportion of parents that have about the right amount of involvement. Someone claims that less than a third of all ninth grade parents have about the right amount of involvement. Can you refute this claim or not?

Answer: In JMP we may choose `Analyze` $\rightarrow$ `Distribution`*, and then* `Confidence interval`*, to get a confidence interval of* $[0.236248, 0.422987]$*. Since* 33% *lies in this confidence interval, you cannot refute the claim.*

c) Why do you think that there would be a higher number of ninth grade parents with extreme levels of involvement than for the lower grades?

Answer: Some parents may get busy and so become less involved, while others are concerned for their children's college education and so become more involved.

4. (25 marks)
The tastes of thirty cheese samples are rated by a panel of experts, whose scores are combined into a single score `http://lib.stat.cmu.edu/DASL/Stories/CheddarCheeseTaste.html`. The cheese manufacturing company would like to discover the relationship between taste and certain elements of the chemical composition, namely, the concentrations of acetic acid, hydrogen sulfide and lactic acid.

Taste	Acetic Acid	Hydrogen Sulfide	Lactic Acid
12.3	4.543	3.135	0.86
20.9	5.159	5.043	1.53
39	5.366	5.438	1.57
47.9	5.759	7.496	1.81
5.6	4.663	3.807	0.99
25.9	5.697	7.601	1.09
37.3	5.892	8.726	1.29
21.9	6.078	7.966	1.78
18.1	4.898	3.85	1.29
21	5.242	4.174	1.58
34.9	5.74	6.142	1.68
57.2	6.446	7.908	1.9
0.7	4.477	2.996	1.06
25.9	5.236	4.942	1.3
54.9	6.151	6.752	1.52
40.9	6.365	9.588	1.74
15.9	4.787	3.912	1.16
6.4	5.412	4.7	1.49
18	5.247	6.174	1.63
38.9	5.438	9.064	1.99
14	4.564	4.949	1.15
15.2	5.298	5.22	1.33
32	5.455	9.242	1.44
56.7	5.855	10.199	2.01
16.8	5.366	3.664	1.31
11.6	6.043	3.219	1.46
26.5	6.458	6.962	1.72
0.7	5.328	3.912	1.25
13.4	5.802	6.685	1.08
5.5	6.176	4.787	1.25

a) What is the best regression model for taste as a function of the other variables? How do you determine which variables to keep in the model and which to leave out?

Answer: First, we run `Fit Model` *with all explanatory variables. The output gives* $p < 0.0001$ *for the ANOVA table, but* $p = 0.9420$ *for the t-test of the Acetic Acid term. Leaving out this variable, we run* `Fit Model` *again. This gives* $p < 0.0001$ *for the ANOVA table, but now all the p-values for the for the t-tests of each term are smaller than 2%. The model is*

$$\textit{Taste} \approx -27.6 + 2.95 \times \textit{Hydrogen Sulfide} + 19.9 \times \textit{Lactic Acid}$$

b) In the output of your regression model there is an Analysis of Variance (ANOVA) table. Explain the meaning of the entries in that table. If the p-value is small enough what is your conclusion: the model is very likely correct, perhaps correct, or very likely wrong?

Answer: Let n *be the number of data and* $k-1$ *the number of terms in the model excluding the intercept. Let* y_i *be the* i^{th} *observation of taste,* $\hat{y}_i$ *the* i^{th} *fitted value of taste, and* $\bar{y}$

the sample average of the y_i. The degrees of freedom for the model, error and total are $k-1$, $n-k$, and $n-1$, respectively. The sums of squares of the model, error and total are

$$\text{SSTot} = \sum_{i=1}^{n}(y_i - \bar{y})^2, \qquad \text{SSMod} = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2, \qquad \text{SSErr} = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2,$$

respectively. The mean square model and error are the sum of squares divided by their respective degrees of freedom, and the F Ratio is the ratio of the mean square model over mean square error. This F Ratio follows an F distribution, and the p-value is given under the null hypothesis that the explanatory variables are unrelated to the response. If the p-value is small enough, then the model is perhaps correct, but one must test for individual terms.

c) In the output of your regression model there is also a table of t-tests for individual regression coefficients. What are the meanings of the p-values here?

Answer: These p-values test whether individual terms in the model are significant. If the p-values are small enough, say $< 5\%$, then the terms are significant.

d) What kind of pattern does one desire in the plot of residuals versus predicted values?

Answer: There should be no discernible pattern in the plot of residuals versus predicted values. A pattern indicates that further terms are needed in the model or that a transformation of one or more variables is needed.

e) What chemical composition makes the best tasting cheese?

Answer: Increasing the concentrations of hydrogen sulfide and lactic acid improve the taste of cheese, although this could probably not be taken to the extreme.