Math 2200  Exam 3  Spring 2009

Put your answers on the answer card.
The answers offered have been rounded off; take that into account when looking for the correct answer.
The first 6 questions are worth 2 points each, the next 3 are 3 points each, the remaining questions are 10 points each: total 101.

1. Suppose you do a hypothesis test at the significance level $\alpha = .05$ and fail to accept (i.e., you reject) the null hypothesis. If instead you had selected the significance level $\alpha = .01$ you would also have failed to accept the null hypothesis.
   a. Definitely yes.
   b. Definitely no.
   c. Possibly but not definitely.

2. When doing a hypothesis test you would like to set things up so that the probability of Type I error and of Type II error are quite small. Although these two goals are generally in conflict with each other you can make progress on both goals if you can increase the sample size, $n$.
   a. True.
   b. False.

3. Hypothesis tests for equality of means are unreliable when the two sample sizes differ by a factor of 5 or more.
   a. True.
   b. False.
4. If sampling to estimate a proportion is done without replacement then observations in the sample are not independent and hence the normal approximation to the sampling distribution of \( \hat{p} \) should not be used.
   a. True.
   b. False.

5. You suspect a certain proportion has the value \( p = .6 \). You could collect data, construct a 95% confidence interval for \( p \) and look to see if .6 is in that interval. Alternatively you could, with \( \alpha = .05 \), test the hypothesis \( H_0 : p = .6 \) against the two sided alternative. Although these two approaches are not the same, they produce very similar types of insight into the accuracy of your suspicions.
   a. True.
   b. False.

6. The \( P \)-value for a chi squared test of homogeneity is always computed using a single tail; i.e., \( P \)-value = \( P(\chi^2 > \text{calculated value}) \).
   a. True.
   b. False.
7. You test the null hypothesis that the value of a proportion is \( p = .5 \) against the two sided alternative using a significance level \( \alpha = .05 \). You use the normal approximation for the sampling distribution of \( \hat{p} \) and compute the value \( z = 2 \). Should you conclude that
   a. There is an (approximately) 5% chance that the null hypothesis is correct.
   b. There is an (approximately) 5% chance that the null hypothesis is not correct.
   c. There is an (approximately) 2.5% chance that the null hypothesis is correct.
   d. There is an (approximately) 2.5% chance that the null hypothesis is not correct.
   e. There is an (approximately) 95% chance that the alternative hypothesis is correct.
   f. There is an (approximately) 97.5% chance that the alternative hypothesis is correct.
   g. The computed value is too close to the critical value to allow a decision.
   h. Both b and e.
   i. Both d and f.
   j. none of the above.

8. According to the book’s guidelines, using the normal approximation when describing the sampling distribution of an estimate for a proportion would be appropriate for which of the following situations?
   1. \( H_0 : p = .1 \), sample size \( n = 70 \)
   2. \( H_0 : p = .2 \), sample size \( n = 60 \)
   3. \( H_0 : p = .3 \), sample size \( n = 50 \)
   4. \( H_0 : p = .4 \), sample size \( n = 40 \)
   5. \( H_0 : p = .5 \), sample size \( n = 30 \)
   6. \( H_0 : p = .6 \), sample size \( n = 20 \)
   7. \( H_0 : p = .7 \), sample size \( n = 10 \)
   a. All.
   b. None.
   c. Only the first two.
   d. Only the last two.
   e. Only the first four.
   f. Only the last four.
   g. Only numbers 2, 3, 4, and 5.
   h. All except the first and last.
9. You are interested in knowing the fraction, \( p \), of drivers on Highway 70 who obey the speed limit and plan to test the null hypothesis \( H_0 : p = 0.8 \) against the one-sided alternative \( H_1 : p > 0.8 \) using a significance level \( \alpha = 0.05 \). You collect data and compute the test statistic \( z = 1.22 \). What do you conclude?
   a. 1.22 is smaller than the critical value 1.96 and hence we accept the null hypothesis.
   b. 1.22 is smaller than the critical value 1.96 and hence we do not accept the null hypothesis.
   c. 1.22 is smaller than the critical value 1.645 and hence we accept the null hypothesis.
   d. 1.22 is smaller than the critical value 1.645 and hence we do not accept the null hypothesis.

10. A Gallup Poll in March of this year asked 500 people "Generally speaking, do you think nuclear power plants are safe or not safe?" 56% said "Safe", the remainder answered "Not Safe" or "Unsure". What is the 95% confidence interval for estimating the fraction of the population who believe nuclear power plants are safe?
   a. (.50, .59)
   b. (.51, .59)
   c. (.50, .62)
   d. (.52, .60)
   e. (.52, .59)
   f. (.52, .61)
   g. (.53, .61)
   h. (.53, .62)
11. The Center for Disease Control and Prevention reports a survey result that 411 of 1012 men over age 64 suffered from some form of arthritis. You use this data to test the null hypothesis that 43% of men in that age group suffer. Use a two sided alternative. What is the $P$ value produced by the test?

a. 0.055
b. 0.065
c. 0.075
d. 0.085
e. 0.095
f. 0.105
g. 0.115
h. 0.125
i. 0.135

12. Denote by $p_M$ and $p_W$ the fractions of men and of women in favor of increasing the state income tax. A survey finds 17 of 41 men in favor and 22 of 38 women in favor. Construct a 95% confidence interval for the difference $p_M - p_W$. The endpoints of the interval are

a. $-0.164 \pm 0.124$
b. $-0.164 \pm 0.218$
c. $-0.164 \pm 0.315$
d. $-0.164 \pm 0.088$
e. $-0.142 \pm 0.218$
f. $-0.142 \pm 0.146$
g. $-0.142 \pm 0.082$
h. $-0.142 \pm 0.124$
i. $-0.108 \pm 0.124$
j. $-0.108 \pm 0.112$
13. Suppose the high temperatures in St. Louis on the next three Mondays are 60°, 66° and 66°. Use that data to construct a 90% confidence interval for the quantity \( T = \) average daily high temperature in St. Louis at this time of year.
   a. (58.2, 69.8)
   b. (58.0, 70.0)
   c. (57.8, 69.8)
   d. (58.0, 69.6)
   e. (58.6, 70.2)
   f. (57.6, 70.2)
   g. (57.8, 69.6)
   h. (58.2, 69.2)

14. You wish to estimate a proportion and to have the 90% confidence interval for your estimate have total length .1. What is the smallest sample size, \( n \), which will insure this? (In doing the computation use the worst case (the "cautious approach") for \( \hat{p} \), as is done in the book.)
   a. 126
   b. 147
   c. 163
   d. 198
   e. 216
   f. 257
   g. 271
   h. 292
   i. 311
   j. 323
15. You want to test the null hypothesis that a four-sided die is fair using a $\chi^2$ test with significance level $\alpha = .05$. You roll the die 40 times and obtain

\[
\begin{array}{c|cccc}
\text{Face} & 1 & 2 & 3 & 4 \\
\hline
\text{Number of times} & 15 & 12 & 7 & 6 \\
\end{array}
\]

What are the computed value of $\chi^2$, the appropriate number of degrees of freedom, and the outcome of the test?

a. $\chi^2 = 5.4, df = 5$, accept $H_0$.
b. $\chi^2 = 5.4, df = 4$, reject $H_0$.
c. $\chi^2 = 5.4, df = 3$, accept $H_0$.
d. $\chi^2 = 5.4, df = 2$, reject $H_0$.
e. $\chi^2 = 4.2, df = 4$, accept $H_0$.
f. $\chi^2 = 4.2, df = 4$, reject $H_0$.
g. $\chi^2 = 4.2, df = 3$, accept $H_0$.
h. $\chi^2 = 4.2, df = 2$, reject $H_0$.

16. Perform a hypothesis test for the equality of $p_{old}$, the proportion of older people who snore, and $p_{young}$, the proportion of younger people who snore, using the survey data in which 13 of 32 young people snore and 23 of 51 older people snore. Use the a two sided alternative. What is the $P$-value you obtain.

a. .689
b. .565
c. .412
d. .326
e. .295
f. .205
g. .142
h. .102
i. .086
j. .056
17. Note: the computed and critical values given below are correct.

You want to know if students' thoughts about going to graduate school change from year to year. You select 20 students from each class, freshman, sophomore, junior, and senior, and ask them their plans. The data is

<table>
<thead>
<tr>
<th>Planing to go to grad school</th>
<th>1st year</th>
<th>2nd year</th>
<th>3rd year</th>
<th>4th year</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>8</td>
<td>13</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Not planning to go to grad school</td>
<td>14</td>
<td>12</td>
<td>7</td>
<td>7</td>
</tr>
</tbody>
</table>

You want to test the hypothesis that the actual fraction of students planning to go to grad school is the same for all the years at a significance level $\alpha = .05$. You compute $\chi^2 = 7.60$. Do you

a. Reject the hypothesis because 7.60 is greater than 5.99, the 5% critical value for $\chi^2$ with $df = 2$.

b. Do not reject the hypothesis because 7.60 is less than 7.81, the 5% critical value for $\chi^2$ with $df = 3$.

c. Do not reject the hypothesis because 7.60 is less than 9.49, the 5% critical value for $\chi^2$ with $df = 4$.

d. Change to $\alpha = .10$ and then reject the hypothesis because 7.60 is greater than 6.251, the 10% cutoff for $\chi^2$ with $df = 3$.

e. Do not reject the hypothesis because 7.60 is much less than 14.1 the 5% critical value for $\chi^2$ with $df = 7$.

f. Notice that the "at least 10 observations in each cell" requirement was not met and decide to use a paired t-test instead.

g. Notice that the "at least 10 observations in each cell" requirement was not met and combine the data from the first two columns and the data from the last two making a two by two table and then redo the analysis.

h. Notice that the "at least 10 observations in each cell" requirement was not met and instead do a regression of percent planning to go to grad school against year to see if the regression line has a positive slope.

i. None of the above.