Name ______________________

**Final Exam**

Notes: You may use your calculator for this exam, but it is necessary to SHOW YOUR WORK for each answer question where it is necessary. Solutions that appear out of nowhere will not be given full credit. You have 120 minutes to complete the exam. If you have any questions about the questions themselves (e.g. what they mean), please feel free to ask me. GOOD LUCK!

1. (10 pts.) Suppose the scores on an exam approximately follow a Normal\((77.2, 5.4)\) distribution.

   a) Use the 68-95-99.7 Rule to determine the proportion of students who scored less than 88 on the exam.

   By the 68-95-99.7 rule (Ch. 3), 95% of the students score within 2 standard deviations of the mean. The boundaries are \(77.2 - 2(5.4) = 66.4\) and \(77.2 + 2(5.4) = 88\). Therefore, 95% lie between 66.4 and 88. The remaining 5% are divided evenly between the left and right tails, leaving 2.5% less than 66.4. All of this together means that \(2.5 + 95 = 97.5\%\) of the students score less than 88.

   b) What is the proportion of students who scored between 73 and 82?

   By normal computations (Ch. 3), we are seeking \(P(73 \leq X \leq 82)\).

   \[
P(73 \leq X \leq 82) = P\left(\frac{73 - 77.2}{5.4} \leq \frac{X - 77.2}{5.4} \leq \frac{82 - 77.2}{5.4}\right)
   = P(-0.78 \leq Z \leq 0.89)
   \]

   From Table A (Z tables), this is equal to \(.8133 - .2177 = .5956\)

2. (10 pts.) Suppose the number of parking tickets that George gets in a year follows a normal distribution with mean 1.7 and standard deviation 0.4.

   a) What is the probability that in a particular year, George gets 2 or more parking tickets?

   Again using the normal methods of Chapter 3, we seek \(P(X \geq 2)\).

   \[
P(X \geq 2) = P\left(\frac{X - 1.7}{0.4} \geq \frac{2 - 1.7}{0.4}\right) = P(Z \geq 0.75)
   \]

   From Table A (Z tables), this is equal to \(1 - .7734 = .2266\)
b) Suppose we observe George for 10 years. What is the probability that the AVERAGE number of tickets George gets per year over that span is greater than 2?

By the Central Limit Theorem (Ch. 10), we know that \( \bar{x} \sim N(1.7, 0.4/\sqrt{10}) \)

Using the normal methods of Chapter 3, we seek \( P(\bar{x} \geq 2) \).

\[
P(\bar{x} \geq 2) = P\left( \frac{\bar{x} - 1.7}{0.4/\sqrt{10}} \geq \frac{2 - 1.7}{0.4/\sqrt{10}} \right) = P(Z \geq 2.37)
\]

From Table A (Z tables), this is equal to 1 - .9911 = \textbf{.0089}

3. (10 pts.) Suppose we want to study the relationship between ACT scores and GPA.

For GPA, we observed: \( \bar{y} = 3.068 \) and \( s_y = 0.627 \)

For ACT scores, we observed: \( \bar{x} = 19.72 \) and \( s_x = 4.39 \)

The correlation between student scores and time spent studying is \( r = 0.572 \)

a) Find the equation of the least-squares regression line for predicting GPA from ACT scores.

From Chapter 5, the least squares regression line has the form:

\[
y = a + bx \quad \text{where} \quad b = r \frac{s_y}{s_x} \quad \text{and} \quad a = \bar{y} - b\bar{x}
\]

Thus, \( b = 0.0572 \cdot \frac{0.627}{4.39} = 0.0817 \) and \( a = 3.068 - 0.0817(19.72) = 1.457 \)

This gives the line \( \hat{y} = 1.457 + 0.0817x \).

b) What percent of the observed variation in student GPA can be explained by the linear relationship between their GPA and ACT scores?

\( r^2 = 0.327184 \) so \textbf{32.7184} % of the variation is explained by the linear relationship.

c) One student scores has a GPA of 2.723 and an ACT score of 22. What is the predicted GPA for a student who has an ACT score of 22? What is the residual for this particular student?

For an ACT score of 22, \( \hat{y}(22) = 1.457 + 0.0817(22) = 3.254 \).
Thus, the residual is $y - \hat{y}(22) = 2.723 - 3.254 = -0.531$.

4. (10 pts.) Suppose that in order to stay on scholarship, Aaron needs to score a B or better in at least 4 of the 5 classes he is taking. Suppose that scores in each class are independent and the probability of scoring a B or better in each course is 0.89.

a) What is the probability that Aaron gets a B or better in all 5 classes?

This is a binomial question (Ch. 12, note the probability of success is given). We seek:

$$P(X = 5) = \binom{5}{5}(.89)^5(.11)^0 = .5584.$$  

b) What is the probability that Aaron maintains his scholarship?

Aaron maintains his scholarship if he scores a B or higher in at least 4 classes, so we seek

$$P(X \geq 4) = P(X = 4) + P(X = 5) = \binom{5}{4}(.89)^4(.11)^1 + .5584 \text{ (from part a)}$$

$$= .3451 + .5584 = .9035.$$  

You also could have used the TI-83 Plus and plugged in $1 - \text{binomcdf}(5, 0.89, 3)$ as this will subtract off the probability that he scores 0, 1, 2 or 3 B’s or higher and leave you with the probability he scores 4 or 5 B’s or higher.

5. (10 pts.) Suppose people can be assigned to seven numbered groups. The probabilities of being assigned to each group are as follows:

<table>
<thead>
<tr>
<th>Group X</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.232</td>
<td>0.187</td>
<td>0.098</td>
<td>0.113</td>
<td>0.101</td>
<td>0.188</td>
<td>0.081</td>
</tr>
</tbody>
</table>

(Ch. 9)

a) Is this a legitimate assignment of probabilities? Why or why not?

Yes. Note that $.232 + .187 + .098 + .113 + .101 + .188 + .081 = 1$ so the probabilities add to one. Also note that each probability is between 0 and 1.

b) What is the probability that someone is assigned to a group numbered less than 3?

$$P(X < 3) = .232 + .187 = .419.$$  

c) What is the probability that someone is assigned to a group with an even number?

$$P(\text{even}) = .187 + .113 + .188 = .488.$$
d) What is the probability that someone is assigned to a group numbered less than 3 or with an even number?

\[ P(X<3 \text{ or even}) = 0.232 + 0.187 + 0.113 + 0.188 = 0.720. \] Note that groups 1, 2, 4 or 6 satisfy at least one of the conditions and thus are included.

e) Why is your answer for part (d) not the sum of the answers for parts (b) and (c)?

The events of a group less than 3 and an even numbered group are not disjoint because 2 falls in both categories. If we were to add our answers from parts b and c, we would have counted group 2 twice!

6. (10 pts.) A study is held to study student habits. A group of 365 students at the U. were asked how many credit hours they were taking this semester. The average response was 14.8 with standard deviation 2.43.

a. Give a 95% confidence interval for the mean number of credit hours for all U. students this semester.

(Ch. 16) A 95% confidence interval has the form:

\[ \bar{x} \pm t^{*} \frac{s}{\sqrt{n}} \]

In our case, this becomes \[14.8 \pm 1.984 \frac{2.43}{\sqrt{365}}.\] This becomes \[14.8 \pm 0.252 = [14.458, 15.052].\]

Notice that we used \( t^{*} \) instead of \( z^{*} \) because the standard deviation of 2.43 describes the sample of 365 students. The degrees of freedom are equal to the sample size – 1.

b. Does this give evidence that the mean number of credit hours taken by students is not 16 at level \( \alpha = 0.05 \)? State \( H_{0} \) and \( H_{A} \) and base the test on the confidence interval from part (a).

\[ H_{0} : \mu = 16 \]
\[ H_{A} : \mu \neq 16 \]

Yes, it does give evidence against the null hypothesis because 16 is not in a \( 1 - \alpha = 95\% \) confidence interval for the mean (we computed the confidence interval in part a). This fact comes near the end of Chapter 14.
7. (10 pts.) A sales manager is hoping that profits are growing for her stores. To test this, suppose she takes a SRS of 41 stores, for which the mean monthly profit (in thousands of dollars) is \( \bar{x}_1 = 3.25 \) and \( s_1 = 35.4 \). Suppose a SRS last year of size 53 had mean monthly profit \( \bar{x}_2 = 5.23 \) and \( s_2 = 98.3 \). Test the hypothesis that the mean number of sales grew in the last year at level \( \alpha = 0.05 \).

a. State \( H_0 \) and \( H_A \).

\[ H_0 : \mu = 0 \quad \text{or} \quad H_0 : \mu_1 = \mu_2 \]
\[ H_A : \mu > 0 \quad \text{or} \quad H_A : \mu_1 > \mu_2 \]

b. Carry out the test and give the P-value.

This is a 2-sample t procedure (Ch. 17, note the two samples with different sizes). We base the test on:

\[
\begin{align*}
\bar{x}_1 - \bar{x}_2 & = \frac{25.3 - 23.5}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{1.8}{0.872} = 2.064. \\
\end{align*}
\]

The degrees of freedom is equal to the smaller sample size – 1 = 41 – 1 = 40.

Using Table C (t tables), the p value for 40 degrees of freedom is between 0.02 and 0.025.

c. State your conclusion.

Since the p-value is less than \( \alpha = 0.05 \), we reject the null hypothesis. Thus, profits are growing.

d. Give a 90\% confidence interval for the difference of the mean monthly profits.

A 90\% confidence interval has the form:

\[
(\bar{x}_1 - \bar{x}_2) \pm t_{.05} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}
\]

In our case, this becomes \((25.3 - 23.5) \pm 1.684(0.872)\) which becomes \(1.8 \pm 1.468 = [0.332, 3.268]\).
8. (10 pts.) Sleeping patterns for adults are believed to vary normally with mean 7.3 hours and standard deviation 1.2 hours. Suppose for our class (n = 180), the average number of hours slept by students was 7.2 hours. Is this an indication that our class has different sleeping patterns than other adults?

a. State $H_0$ and $H_A$.

$H_0 : \mu = 7.3$

$H_A : \mu \neq 7.3$

Notice it is two-sided because we are seeking evidence if our class is different, which doesn’t specify a direction.

b. Carry out the test and give the P-value.

This is a $z$ test situation, because the standard deviation of 1.2 describes the population of adults.

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{7.2 - 7.3}{1.2 / \sqrt{180}} = -1.11.$$  

Because our alternative is 2-sided, the p-value is $2 \cdot P(Z<-1.11)$.

From Table A (Z tables), this is equal to $2 \cdot (0.1335) = 0.2670$.

c. State your conclusion.

**The p-value is large (0.2670), so we can not reject the null hypothesis.** We do not have enough evidence to show that our class has different sleeping patterns than the rest of adults.

d. Give a 99% confidence interval for the number of hours slept by students in our class.

A 99% confidence interval has the form:

$$\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}}$$

In our case, this is $7.2 \pm 2.576 \frac{1.2}{\sqrt{180}}$.

This becomes $7.2 \pm 0.230 = [6.97, 7.43]$. 
9. (10 pts.) Suppose an incumbent governor’s approval rating is in question. Suppose in a poll of 321 voters, 183 were satisfied with the governor’s performance. Is this an indication that the governor’s approval rating is greater than 50%? Test at level \( \alpha = 0.05 \).

a. State \( H_0 \) and \( H_A \).

\[
H_0 : \ p = 0.5 \\
H_A : \ p \geq 0.5
\]

b. Carry out the test and give the P-value.

This is a one-sample proportion situation (Ch. 18, note we are asking about percentages). We base the test on:

\[
z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}.
\]

First we must compute the sample proportion. \( \hat{p} = \frac{183}{321} = 0.570 \).

Therefore, \( z = \frac{0.570 - 0.5}{\sqrt{0.5(1-0.5)/321}} = \frac{0.07}{0.033} = 2.13 \).

The p-value is \( P(Z > 2.13) \). From Table A (Z tables), this is \( 1 - .9834 = .0166 \).

c. State your conclusion.

Since the p-value is less than \( \alpha = 0.05 \), we reject the null hypothesis. Thus, the governor’s approval rating IS higher than 50%.

d. Give a 95% confidence interval for the proportion of voters who approve the governor’s performance.

A 95% confidence interval has the form:

\[
\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}.
\]

In our case, this becomes \( .570 \pm 1.96 \sqrt{\frac{.57(1-.57)}{321}} \).

This in turn becomes \( .570 \pm .054 = [.5159, .6243] \).
10. (10 pts.) Suppose a study is held to see if smoking habits differ between men and women. Suppose of a sample of 55 women, 32 were smokers. Suppose of a sample of 50 men, 26 were smokers. Is this an indication that there is a difference between the proportions of men and women who smoke?

a. State \( H_0 \) and \( H_A \).

\[ H_0 : p_1 = p_2 \]
\[ H_A : p_1 \neq p_2 \]

Note the alternative is two-sided because we are seeking evidence for a difference in proportions, which does not specify direction.

b. Carry out the test and give the P-value.

This is a two-sample proportion test (Ch. 19, note two independent samples of men and women). We base our test on:

\[
z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1 - \hat{p}) \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}
\]

We need to compute some proportions to proceed.

\[
\hat{p}_1 = \frac{32}{55} = 0.582, \quad \hat{p}_2 = \frac{26}{50} = 0.52 \quad \text{and the pooled proportion} \quad \hat{p} = \frac{32 + 26}{55 + 50} = \frac{58}{105} = 0.552.
\]

Now,
\[
z = \frac{0.582 - 0.52}{\sqrt{0.552(1 - 0.552) \left( \frac{1}{55} + \frac{1}{50} \right)}} = \frac{0.062}{0.097} = 0.64.
\]

Since our alternative is 2-sided, the p-value is \(2 \cdot P(Z > 0.64)\). From Table A (Z tables), this is \(2 \cdot (1 - 0.7389) = 2 \cdot 0.2611 = 0.5222\).

c. State your conclusion.

Since the p-value is so large (.5222), we can not reject the null hypothesis. Thus, there is no evidence to support a difference in the proportions of men and women smokers.

d. Give a 95% confidence interval for the difference of proportions of men and women smokers.

A 95% confidence interval has the form:

\[
(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}}.
\]
In our case, this becomes \((.582 - .52) \pm 1.96 \sqrt{\frac{.582(1 -.582)}{55} + \frac{.52(1 -.52)}{50}}\).

This in turn becomes \(0.062 \pm 0.190 = [-.128, .252]\).

Formulas

Least squares regression line: \(\hat{y} = a + bx\) where \(b = r \frac{s_y}{s_x}\) and \(a = \bar{y} - b \bar{x}\)

Binomial probabilities: \(P(X = k) = \left(\begin{array}{c}n \\ k\end{array}\right) p^k (1-p)^{n-k}\) where \(\left(\begin{array}{c}n \\ k\end{array}\right) = \frac{n!}{k!(n-k)!}\)

Confidence intervals:

Mean of a normal population: \(\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}}\)

One-sample t confidence interval: \(\bar{x} \pm t^* \frac{s}{\sqrt{n}}\)

Two-sample t confidence interval: \((\bar{x}_1 - \bar{x}_2) \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\)

Confidence interval for a population proportion: \(\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}\)

Confidence interval for 2 population proportions: \((\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}}\)

Test statistics:

Z test for population mean: \(z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}\)

One sample t test for population mean: \(t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}\)

Two sample t test for population mean: \(t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}\)

Z test for one proportion: \(z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}\)

Z test for two proportions: \(z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1 - \hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}\)