Exam 5

Instructions:

- Please turn off and put away all cell phones and any listening devices.
- Each question is weighted the same, and graded using a 5 point rubric. The exam is worth 200 points.
- Work should be clearly labeled and easy to follow.
- Partial credit will only be given for work showing significant progress toward the correct answer.
- You may use a graphing calculator on this exam.
- Work in pen will not be accepted.
- You may use one 8.5" x 11", or smaller handwritten sheet of notes. Both sides may be used, but photocopies of any kind are not allowed.
- Do not hesitate to ask the instructor if you do not understand what a question is asking, of if you are showing enough work.

Round all probabilities to three significant digits

Date: ________  Name: ________  Class(circle one):  4:20  6:00
1. The results of a random sample of students by the location of school and the number of those students achieving basic skill levels in three subjects is listed in the table below. Test the claim that achieving a basic skill level in a subject is independent of school location.

<table>
<thead>
<tr>
<th>Location of School</th>
<th>Subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban</td>
<td>Reading 43</td>
</tr>
<tr>
<td></td>
<td>Math 42</td>
</tr>
<tr>
<td></td>
<td>Science 38</td>
</tr>
<tr>
<td>Suburban</td>
<td>Reading 63</td>
</tr>
<tr>
<td></td>
<td>Math 66</td>
</tr>
<tr>
<td></td>
<td>Science 65</td>
</tr>
</tbody>
</table>

Ho: Claim: Subject is independent of school location

HA: If false! They are dependent.

Matrix A 2x3

\[ \chi^2 - \text{Test} \]

(Calc. 80.18734827) p-Value = .802 > .05

Fail to reject Ho.

There is not sufficient evidence to warrant rejection of the claim that subject is independent of school location.

2. A simple random sample of front-seat occupants involved in car crashes is obtained. Among 2823 occupants not wearing seat belts, 1.1% were killed. Among 7765 occupants wearing seat belts 0.21% were killed. Construct a 90% confidence interval estimate of the difference between fatality rates for those not wearing and those wearing seat belts. What can you conclude from the result?

\[ \hat{p}_1 - \hat{p}_2 \]

\[ 2 - \text{Prop} Z\text{Int} \]

\[ \text{NW} \]

\[ \sum X_1 = 0.011(2823) = 31.053 \approx (3) \]

\[ n_1 = 2823 \]

\[ \sum X_2 = 0.021(7765) = 163.065 \approx (16) \]

\[ n_2 = 7765 \]

\[ \hat{p}_1 = 0.004 \]

\[ \hat{p}_2 = 0.002 \]

\[ Z_{0.05} = 1.645 \]

\[ \text{CI} (0.00559, 0.0123) \]

Since 0.00 is not contained in the CI, we have evidence there is a significant difference in the fatality rates of the two groups. The values are all positive, so those not wearing have higher fatality rates.
3. An experiment was conducted at a university to compare the mean number of study hours per week by student athletes and by non-athletes. The mean for 55 student athletes was 20.6 hours with a standard deviation of 4.5 hours. For 200 non-athletes, the mean was 23.5 hours with a standard deviation of 3.8 hours. Test the claim that the mean number of study hours per week for the non-athletes is greater than the mean for the athletes.

\[ \text{NA A} \]

Claim: \( \mu_1 > \mu_2 \)

If false: \( \mu_1 \leq \mu_2 \)

\[ \text{H}_0: \mu_1 = \mu_2 \]

(claim) \( \text{H}_A: \mu_1 > \mu_2 \)

There is sufficient evidence to support the claim that the mean number of study hours for non-athletes is greater than for athletes.

\[ \text{P-value} = .0000193 < .05 \]

\[ \text{reject } \text{H}_0 \]

(calc 1.9369305x10^-5)

4. A study of people who refused to answer survey questions provided the randomly selected sample data shown below. At the 0.01 significance level, test the claim that the cooperation of the subject (response or refusal) is independent of the age category. Does any particular age group appear to be particularly uncooperative?

<table>
<thead>
<tr>
<th>Age</th>
<th>18-21</th>
<th>22-29</th>
<th>30-39</th>
<th>40-49</th>
<th>50-59</th>
<th>60 and over</th>
</tr>
</thead>
<tbody>
<tr>
<td>Responded</td>
<td>73</td>
<td>255</td>
<td>245</td>
<td>136</td>
<td>138</td>
<td>202</td>
</tr>
<tr>
<td>Refused</td>
<td>11</td>
<td>20</td>
<td>33</td>
<td>16</td>
<td>27</td>
<td>49</td>
</tr>
</tbody>
</table>

\[ \text{H}_0: \text{Claim: cooperation is independent of age category.} \]

\[ \text{H}_A: \text{If false: they are dependent.} \]

\[ \chi^2 \text{Test} \]

\[ \text{P-value} = .00111 < .01 \]

\[ \text{reject } \text{H}_0 \]

(calc .0011114514)

Age 60 and over seem to be the most uncooperative.

There is enough evidence to warrant rejection of the claim that cooperation is independent of age category.

(c dependent)
5. As part of the National Health Survey, data were collected on the weights of men in two different age brackets. Test the claim that the mean weight for the younger group is different than that of the older group by constructing an appropriate confidence interval. Use α = 0.05. Interpret the results.

\[ H_0: \mu_1 = \mu_2 \]

\[ H_A: \mu_1 \neq \mu_2 \]

\[ CL = 1 - 0.05 = 0.95 \]

\[ 2\text{-Samp T-Int}(\bar{x}, s_1, \bar{x}, s_2) \]

\[ x_1 = 176.3 \text{ lb} \quad \bar{x}_2 = 178.2 \text{ lb} \]

\[ s_1 = 35.2 \text{ lb} \quad s_2 = 29.9 \text{ lb} \]

\[ n_1 = 751 \quad n_2 = 1031 \]

\[ CL = 0.95 \]

\[ CI(-5.0, 1.2) \]

\[ CI_{(-5.0, 1.2)} \]

Since zero is contained in the CI, we have evidence that there is no significant difference in the two groups' mean weight.

6. Listed below are incomes (in thousands of dollars) and the percents of the incomes donated to charity for seven families.

<table>
<thead>
<tr>
<th></th>
<th>Income  ($)</th>
<th>50</th>
<th>65</th>
<th>48</th>
<th>42</th>
<th>59</th>
<th>72</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>x_1</td>
<td>Percent Donated</td>
<td>4</td>
<td>8</td>
<td>5</td>
<td>5</td>
<td>10</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>x_2</td>
<td>Percent Donated</td>
<td>4</td>
<td>8</td>
<td>5</td>
<td>5</td>
<td>10</td>
<td>7</td>
<td>6</td>
</tr>
</tbody>
</table>

a. Construct a scatterplot. Label each axis.

b. Assume there is no linear correlation between the amount of income and the percent donated to charity. Find the best predicted percent donated for a family that has an income of 55 thousand dollars.

\[ 2\text{-Var Stats, L_1, L_2} \]

\[ \bar{y} = 6.428571429 \]

\[ \text{so} \quad \bar{y} \approx 69 \% \]
7. A manufacturer claims that athletes can increase their vertical jump heights by using the company's new Strength Shoes. Eight athletes are randomly selected and their vertical jump heights are measured before using the new shoes, and then again after they have used them for 8 months. Test the manufacturer's claim using $\alpha = .10$. (Dependent)

<table>
<thead>
<tr>
<th>Vertical Jump Heights</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
</tr>
<tr>
<td>After</td>
</tr>
</tbody>
</table>

$\bar{L}_3 = \bar{L}_2 - \bar{L}_1$

Claim: $\mu_d > 0$
If true: $\mu_d \leq 0$

$H_0: \mu_d = 0$
$H_a: \mu_d > 0$

There is sufficient evidence to reject the null hypothesis.

$T$-Test ($Data$)

$\mu_0 = 0$
$\bar{L}_3 = \bar{L}$

$\mu_d > \mu_0$

$P$-Value = .0262 < .10

Reject $H_0$

There is sufficient evidence to support the claim that they can increase their vertical jump heights by using the new shoes.

8. In a 1993 survey of 500 college students, 171 said that they used illegal drugs during the previous year. In a recent survey of 720 college students, 263 said they used illegal drugs during the previous year. Use a 0.05 significance level to test the claim that the rate of college students using illegal drugs in 1993 was greater than or equal to the current rate.

Claim: $p_1 \geq p_2$
If false: $p_1 < p_2$

$H_0: p_1 = p_2$
$H_a: p_1 < p_2$

There is sufficient evidence to warrant rejection of the claim that the rate of illegal drug use in 1993 was greater than or equal to the current rate.

2-Prop $Z$-Test

$\left\{ \begin{array}{l} X_1 = 171 \\ N_1 = 500 \end{array} \right.$

$X_2 = 263$ (Calc .03.335102)

$P$-value = .0123 < .05

Reject $H_0$
9. A tax preparation company claims that the tax preparation methods of adults are distributed as shown in the table below. You randomly select 300 adults and ask each which type of tax preparation he or she prefers. Using $\alpha = 0.01$, test the claim that the observed frequencies agree with the claimed distribution.

<table>
<thead>
<tr>
<th>Claimed Dist.</th>
<th>Accountant</th>
<th>By Hand</th>
<th>Computer Software</th>
<th>Friend/Family</th>
<th>Tax Prep Service</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survey Results</td>
<td>25%</td>
<td>20%</td>
<td>35%</td>
<td>5%</td>
<td>15%</td>
</tr>
<tr>
<td>n = 300</td>
<td>71</td>
<td>40</td>
<td>101</td>
<td>35</td>
<td>53</td>
</tr>
</tbody>
</table>

$H_o$: Claim: Observed freq. agree with claimed distribution

$H_a$: $\neq$ false: At least 1 is different.

$X^2 = \frac{(25)(300) + (20)(300) + (35)(300) - (25)(300)}{300} = 45$

$X^2_{cdf}(35.121, 10.99, 4)$

$P$-Value $= 0.001 < 0.01$

There is sufficient evidence to reject $H_o$.

(Calc $4.386937246 \times 10^{-7}$)

10. The ages (in years) of five children and the number of words in their vocabulary are listed in the table below.

<table>
<thead>
<tr>
<th>Age</th>
<th>1</th>
<th>5</th>
<th>2</th>
<th>4</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary Size</td>
<td>5</td>
<td>1900</td>
<td>1490</td>
<td>2400</td>
<td></td>
</tr>
</tbody>
</table>

(a) Use the $P$-value and the linear correlation coefficient to determine whether there is a correlation between the two variables.

LinReg Test $P$-Value $= .0000196 < .05$

Linear correlation $r = .999$

(b) Find the linear regression equation.

$y = -4.95 + 4.83x$

$\hat{y} = -45.73488372$

$\hat{b} = 4.834302326$

(c) Find the best predicted number of words if a child's age is 3.

$X = 3 \Rightarrow \hat{y} = -4.95 + 4.83(3)$

$\hat{y} = 95.4$ words