Instructor: E. JACKIEWICZ

Print Your Name: ___________________________________________

CLASS TIME: _______________________________

Honor Statement:

I have neither given nor received information regarding this exam, and I will not do so until the examination period is over.

Signed ___________________________________________

DIRECTIONS:
This is a closed book examination. You may use a calculator and 8x11 page with hand written notes, no completely solved problems are allowed. Turn in the notes with your exam.
There are 10 questions and 5 True-False statements for 5 points extra credit. 7 multiple choice questions are worth 5 points each. Place letter answers (A-E as appropriate) in the table below. For problems 1-3 (65 points total) provide complete and well-organized full answers with all work showing, include requested sketches.

Relax and good luck!

Place answers to multiple choice questions below, use letters A-E as appropriate

<table>
<thead>
<tr>
<th>Question number</th>
<th>#4</th>
<th>#5</th>
<th>#6</th>
<th>#7</th>
<th>#8</th>
<th>#9</th>
<th>#10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Answer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Place answers to extra credit true-false statements below, circle T or F as appropriate

<table>
<thead>
<tr>
<th>Statement number</th>
<th>#1</th>
<th>#2</th>
<th>#3</th>
<th>#4</th>
<th>#5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Answer (circle one)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Part ONE: Show all work on problems 1 - 3**

**Question 1 Volunteer Work of College Students.** Researcher collected information about the number of hours ASU students are spending weekly in volunteer service. Table below gives summary statistics for two independent samples of male and female students. **Do data present evidence at 5% significance level that mean weekly time spent in volunteering is higher for females than for males?**

<table>
<thead>
<tr>
<th></th>
<th>Female</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample mean</td>
<td>4.8</td>
<td>3.5</td>
</tr>
<tr>
<td>Sample standard deviation</td>
<td>1.87</td>
<td>2.48</td>
</tr>
<tr>
<td>Sample size</td>
<td>13</td>
<td>17</td>
</tr>
</tbody>
</table>

Test appropriate hypothesis using **nonpooled t-test procedure.** Assume normal populations.

a. (5 points) Let $\mu_1$ = mean weekly volunteer hours for all female ASU students $\mu_2$ = mean weekly volunteer hours for all male ASU students. Use above notation to formulate appropriate null and alternative hypothesis.

b. (5 points) Compute appropriate test statistics.

c. (5 points) Sketch the rejection region for your test, use $\alpha = 0.05$ and 27 degrees of freedom. Clearly mark critical value(s), label rejection and nonrejection parts on your sketch.

d. (5 points) Decide if null hypothesis is rejected or not at $\alpha = 0.05$, explain your decision (circle one answer from selections below) and clearly answer question posed in the problem, use a full sentence.

A) $H_0$ rejected, we have evidence for $H_a$
B) $H_0$ not rejected, we have no evidence for $H_a$
C) $H_0$ rejected, we have no evidence for $H_a$
D) $H_0$ not rejected, we have evidence for $H_a$

Answer question posed using a complete sentence:
**Question #2 Combating Midday Drowsiness.** A researcher examined different methods office workers at ASU use to combat midday drowsiness. Table below shows the distribution of 5 such methods for a random sample of 70 office workers at ASU.

<table>
<thead>
<tr>
<th>Method</th>
<th>Beverage</th>
<th>Nap</th>
<th>Walk</th>
<th>Snack</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>observed frequency</td>
<td>20</td>
<td>16</td>
<td>10</td>
<td>9</td>
<td>15</td>
</tr>
</tbody>
</table>

At 10% significance level is there sufficient evidence to conclude that some methods selected by office workers to combat drowsiness are preferred over others? Test hypotheses listed below by means of **Chi-square Goodness-of-Fit test**, use $\alpha = 0.10$

$H_0$: All methods are equally preferred  
$H_a$: Some methods are preferred over others

**Use the following steps:**

a) (5 points) Compute the expected counts under assumption that $H_0$ is true.

<table>
<thead>
<tr>
<th>Method</th>
<th>Beverage</th>
<th>Nap</th>
<th>Walk</th>
<th>Snack</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected frequency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

b. (5 points) Partial test statistics (without category “snack”) equals 4.07. Compute complete test statistics by adding a missing part, show your work.

\[ \chi^2 = 4.07 + \_\_\_ = \_\_\_ \]

c. (5 points) Give p-value for your test, you may compute it directly using an appropriate calculator function or estimate it from tables. Use appropriate degrees of freedom and include a sketch explaining your answer.

Degrees of freedom: _______ P-value:_________

Sketch:

d. (5 points) Decide if null hypothesis is rejected or not at $\alpha = 0.10$, explain your decision and clearly answer question posed in the problem.

$H_0$ rejected $H_0$ not rejected (circle one)

Answer question posed (circle one of answers below)

A) There is evidence at $\alpha = 0.10$ that some methods are preferred over others
B) There is no evidence at $\alpha = 0.10$ that some methods are preferred over others
Question#3

Researcher investigated a relationship between length (X) and weight (Y) of snakes. He observed a linear trend and obtained following least squares regression line:

\[ \hat{y} = 6.89x - 303.07 \]

Length was measured in centimeters, weight in grams. Longest snake in the sample was 1.8 meters (180 cm) long. SST= 9811 and SSE=1194

a) (5 points) Interpret the slope of this regression line in the context of the problem, be very specific.

b) (10 points) Compute coefficient of determination and linear correlation coefficient and give percentage of total variability among the weight of the snakes (Y) that is explained by the regression line. Round answers to 4 decimal places.

Coefficient of determination  

Linear correlation coefficient  

Percentage explained

(c) (5 points) Use the given equation to predict the weight of a snake that is 1.5 meters (150 cm) long.

d) (5 points) Explain briefly why predicting a weight of a snake that is 2.2 meters (220cm) long may give unreasonable results.
Part TWO: Multiple choice questions, 5 points each

Use following information for questions 4-5

Age of supporters in local election in 2012. Suppose we want to find out if there is a difference between average age of supporters of two candidates for State Senat in AZ, one is a Democrat, the other a Republican. Let $\mu_1$ = mean age of the supporters of a Democrat and $\mu_2$ = mean age of the supporters of a Republican We collect a two independent random samples of supporters for each candidate, data is given below:

Democrat: $\bar{x}_1=35, s_1=4.5, n_1=23$  \hspace{1cm} Republican: $\bar{x}_2=47, s_2=3.1, n_2=25$

Question 4: Assuming ages of each population of supporters have normal distribution compute Margin of error in 95% Confidence Interval for $(\mu_1 - \mu_2)$, use non-pooled T interval procedure with 38 degrees of freedom. Give 2 decimal places.

(A) $E=1.27$  \hspace{1cm} (B) $E=2.28$  \hspace{1cm} (C) $E=1.13$  \hspace{1cm} (D) $E=1.90$  \hspace{1cm} (E) none of these

Question 5: Suppose our 95% confidence interval is roughly (-14, -10). Use that interval to decide if there is convincing evidence at $\alpha = 0.05$ that there is a difference between two population means. Select one of the following responses:

(A) Yes, CI does not contain zero  \hspace{1cm} (B) No, CI does not contain zero  \hspace{1cm} (C) No, this CI can't even be correct, because it has negative endpoints  \hspace{1cm} (D) not enough information

Use following information for Questions 6-8

Suppose you are testing hypotheses: $H_0: \ p = 0.8$ , where $p=$true proportion of smokers in some population

Question 6: Suppose you have following alternative hypothesis $H_a: \ p > 0.8$ and your test statistics is $z=1.57$. Compute the p-value for your test (4 decimal places)

(A) 0.0582  \hspace{1cm} (B) 0.9418  \hspace{1cm} (C) 0.0291  \hspace{1cm} (D) 0.1164  \hspace{1cm} (E) none of these

Question 7: Suppose the p-value for your right-tailed test was 0.037, what is the p-value if we change alternative hypothesis to $H_a: \ p \neq 0.8$ ?

(A) 0.037  \hspace{1cm} (B) 0.0185  \hspace{1cm} (C) 0.074  \hspace{1cm} (D) Not enough information

Question 8: Suppose you did not reject $H_0$ in favor of $H_a: \ p \neq 0.8$ (at 5% significance level) and then you computed 95% CI for $p$, would that interval contain 0.8?

(A) Yes  \hspace{1cm} (B) No  \hspace{1cm} (C) Not enough information
Use following Information for Questions 9-10

**Recreational Reading and Gender.** A book store owner wished to determine if there is a relationship between gender and type of books selected by his customers. He collected a random sample of customers and classified them according to gender (Male, Female) and type of books they selected (Mystery, Romance, Self-Help, Other). He used Chi-square test of independence and received test statistics $\chi^2 = 12.3$. Does he have sufficient evidence at $\alpha = 0.05$ that gender and type of books selected are associated?

**Question 9:** Formulate null and alternative hypotheses:

A) $H_0$: Variables are not associated  
$H_a$: Variables are associated

B) $H_0$: Variables are associated  
$H_a$: Variables are not associated

C) $H_0$: $\mu_1 = \mu_2$,  
$H_a$: $\mu_1 \neq \mu_2$

D) $H_0$: $p = 0$,  
$H_a$: $p \neq 0$

**Questions 10:** Use tables to give the rejection region for your test, use appropriate degrees of freedom.

(A) $\chi^2 \geq 9.488$  
(B) $\chi^2 \geq 13.277$  
(C) $\chi^2 \geq 7.815$  
(D) $\chi^2 \geq 12.592$

---

**Extra Credit Questions (1 point each)** Answer True or False for each:

E1: P-value for the hypothesis test can be negative.

**True**  
**False**

E2: Linear correlation coefficient $r = 0.79$ indicates stronger linear trend than $r = -0.79$.

**True**  
**False**

E3: In linear regression analysis it is appropriate to fit the Least Squares regression line to the data even if the plot of our data clearly does not show a linear trend.

**True**  
**False**

E4: In testing hypothesis using critical value approach, the area over the rejection region is equal to the significance level ($\alpha$) we selected for our test.

**True**  
**False**

E5: Degrees of freedom for chi-square tests depend on the size of the sample(s).

**True**  
**False**
Key.

Q1
a) \( H_0 : \mu_1 = \mu_2 \quad H_a : \mu_1 > \mu_2 \)
b) ____t= 1.64____
c) \( C_v=1.703 \) Sketch represents t-curve, center at 0, rejection region marked right of 1.703, nonrejection region left of the cv.
d) Select B.
We have evidence at 5% significance level that mean weekly time spent volunteering is higher for females than for males

Q2
a) \( E=14 \) for all
b) + __1.79__ = ___5.86___
c) Degrees of freedom: _____4_________ P-value:___0.2098___
Sketch should represent a Chi-square curve, 5.86 is marked and region right of it is shaded indicating that the shaded area=p-value

d) Select \( H_0 \) not rejected, and Select B

Q3
a) For every cm increase in length, the weight of a snake increases by 6.89 grams.

b)Coefficient of determination_______0.8783__________
Linear correlation coefficient_______0.9372__________
Percentage explained___________87.83%__________

c) 150*6.89-303.07=730.43 grams

d) Length of 2.2 meters is outside the range of our data. This would be an extrapolation and it may not give reasonable results, since we can't be certain that past 1.8 meters the relationship between length and weight will remain the same.

<table>
<thead>
<tr>
<th>Question number</th>
<th>#4</th>
<th>#5</th>
<th>#6</th>
<th>#7</th>
<th>#8</th>
<th>#9</th>
<th>#10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Answer:</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>C</td>
</tr>
</tbody>
</table>

Place answers to extra credit true-false statements below, circle T or F as appropriate

<table>
<thead>
<tr>
<th>Statement number</th>
<th>#1</th>
<th>#2</th>
<th>#3</th>
<th>#4</th>
<th>#5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Answer (circle one)</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>T</td>
<td>F</td>
</tr>
</tbody>
</table>