1. The Oliver Twist Foundation offered a twelve-session behavior-management training package for workers at juvenile residential facilities in another state. You randomly selected a sample of \( n = 48 \) workers from workers throughout the state who participated in the classes last year. You selected these workers with the restriction that no member of the sample could have attended the same twelve session series. Members of the sample attended an average of \( \bar{X} = 9 \) sessions. At the conclusion of the training each worker participated in three role-play exercises designed to give the opportunity to demonstrate behavior management skills. Possible total scores for all three exercises ranged from \( Y = 0 \) to \( Y = 12 \). The mean score for sample members was \( \bar{Y} = 8 \). The correlation between number of sessions attended and skill score on the role-plays was \( r = +0.275 \).

1.01. Calculate the proportion of unexplained variance
\[
1 - r^2 = ________
\]
(Round \( 1 - r^2 \) to two decimal places for this answer, but use the unrounded \( 1 - r^2 \) for further calculations.)

1.02. Calculate \( t_{obt} \). Show your work and round to three decimal places.
\[
t_{obt} = r \sqrt{\frac{n - 2}{1 - r^2}} = ________
\]

1.03. What are the degrees of freedom for \( t \)?
\[
df = ________
\]

1.04. Find \( t_{crit} \) at a non-directional \( \alpha = .05 \) from the table of critical values.
\[
t_{crit} = t_{(\alpha/2, df)} = ________
\]

1.05. Compare the values for \( t_{obt} \) and \( t_{crit} \). At a non-directional \( \alpha = .05 \), should you reject the Null Hypothesis that attendance (\( X \)) and skill score (\( Y \)) are uncorrelated? Why or why not?
2. Table 1 shows the relationship between intervention and outcome found in a study of \( n = 40 \) individuals.

<table>
<thead>
<tr>
<th>Intervention Status (X)</th>
<th>Outcome (Y)</th>
<th>Not Improved</th>
<th>Improved</th>
<th>Row Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intervention</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td></td>
<td>( f_A = 2 )</td>
<td>( f_B = 18 )</td>
<td>( f_A + f_B = 20 )</td>
</tr>
<tr>
<td><strong>No Intervention</strong></td>
<td></td>
<td>( f_C = 8 )</td>
<td>( f_D = 12 )</td>
<td>( f_C + f_D = 20 )</td>
</tr>
<tr>
<td><strong>Column Total</strong></td>
<td></td>
<td>( f_A + f_C = 10 )</td>
<td>( f_B + f_D = 30 )</td>
<td>( n = 40 )</td>
</tr>
</tbody>
</table>

The correlation between intervention status and outcome was \( r_\Phi = +0.346 \). Use a chi-square test to evaluate this phi coefficient.

\[
\chi^2_{\text{obt}} = \frac{n \left[ (f_B \cdot f_C) - (f_A \cdot f_D) \right]^2}{(f_A + f_B)(f_C + f_D)(f_A + f_C)(f_B + f_D)}
\]

2.01. \( f_B \cdot f_C \)

2.02. \( f_A \cdot f_D \)

2.03. \( \left( f_B \cdot f_C - f_A \cdot f_D \right)^2 \) 

2.04. \( \left( f_B \cdot f_C - f_A \cdot f_D \right)^2 \) 

2.05. \( \frac{n \left[ (f_B \cdot f_C) - (f_A \cdot f_D) \right]^2}{(f_A + f_B)(f_C + f_D)(f_A + f_C)(f_B + f_D)} \)

2.06. \( (f_A + f_B) \cdot (f_C + f_D) \cdot (f_A + f_C) \cdot (f_B + f_D) \)

2.07. Calculate \( \chi^2_{\text{obt}} \). Show your work and round to three decimal places. \( \chi^2_{\text{obt}} = \) ________.

2.08. What are the degrees of freedom for this \( \chi^2 \)? \( df = \) ________.

2.09. Find \( \chi^2_{\text{crit}} \) at \( \alpha = .05 \) from the table of critical values. \( \chi^2_{\text{crit}} = \chi^2_{\alpha; df} = \) ________.

2.10. Compare the values for \( \chi^2_{\text{obt}} \) and \( \chi^2_{\text{crit}} \). At \( \alpha = .05 \), should you reject the Null Hypothesis that intervention status (X) and outcome (Y) are uncorrelated? Why or why not?

3. We wish to see if amount of recent in-service training is correlated with the frequency of service provision mistakes made by protective services workers.

The independent variable refers to the amount of in-service training (\( X \) – measured by number of hours) a worker attended in the previous calendar year.

The dependent measure refers to the number of cases out of 15 randomly selected cases from each worker’s caseload (\( Y \)) where there were service provision mistakes. Mistakes refer to cases where the worker did not provide an appropriate service or inadequately provided an appropriate service. We shall limit the cases selected for review to those opened in the January – March quarter immediately following training.

Based upon data from a pilot study, we have reason to believe that the number of hours of in service training and number of service provision mistakes have normal distributions.

To identify the appropriate test, we must
- identify tests appropriate to the research design used.
- identify tests appropriate to the level at which the dependent (outcome) variable is being measured and, if relevant, to the distribution of dependent measure scores.
3.01. Appropriate Test = ________________

We want to be able to detect a *medium to large* effect size of \( \rho = 0.4 \).

Using non-directional \( \alpha = 0.05 \) and a power level of \( 1-\beta = 0.80 \), how large a sample of protective services workers should we select?

3.02. Total Sample Size = __________

4. We want to evaluate the relative effectiveness of three training techniques (workbook, video and interactive computer program) on accuracy of risk assessment in child maltreatment cases.

The subjects were taught to use a child maltreatment risk protocol to evaluate risk of maltreatment within twelve components of the situation. Based upon the protocol, subjects would classify maltreatment risk within five categories: (1) minimal to no risk, (2) slight risk, (3) moderate risk, (4) strong risk and (5) extreme risk.

After completion of training, we had each subject evaluate twenty case vignettes using the risk protocol. The dependent measure was number of correct evaluations of risk.

We could conduct three statistical tests of difference in assessment accuracy between pairs of training techniques:

i. workbook versus video training.

iii. workbook versus interactive computer program.

iii. video training versus interactive computer program.

4.01. If we conducted each of the three statistical tests at comparison-wise \( \alpha = .05 \), what would be the experiment-wise \( \alpha \)? Round your answer to two decimal places.

4.02. Excluding Ryan’s approximation, under what circumstance may we conduct the three statistical tests between pairs of training techniques and maintain an experiment-wise \( \alpha = .05 \)?

5. We randomly selected \( n = 36 \) fifth grade students from those students referred for disruptive behavior problems in the Thatcher school district. The Thatcher school district has 28 elementary schools. When we randomly selected students, we did so with the restriction that no student could be in the same class as another student selected for participation in the study.

We randomly assigned the students to either a reward condition (received additional free play time for each full day without disruptive behavior) or a punishment condition (lost free play time for disruptive behavior).

We compared the students from the two conditions on number of instances of disruptive behavior during the most recent six-week grading period following onset of intervention.

At post-test, the frequency of disruptive behavior for students in the reward condition had a mean of \( \bar{Y} = 3.4 \), a median of \( Y_{50} = 1 \), and a mode of \( Y_{Mode} = 0 \). The minimum and maximum disruptive behavior scores were \( Y_{Min} = 0 \) and \( Y_{Max} = 15 \).

For students in the punishment condition, the frequency of disruptive behavior had a mean of \( \bar{Y} = 3.0 \), a median of \( Y_{50} = 1 \), and two modes – \( Y_{Mode} = 1 \) and 4. The minimum and maximum disruptive behavior scores were \( Y_{Min} = 0 \) and \( Y_{Max} = 10 \).

5.01 The independent variable in this study refers to which of the following.

- disruptive behavior
- (1) disruptive, (2) mildly disruptive, (3) not disruptive
- frequency of disruptive behavior
- playtime condition
- referral status
- (1) reward, (2) punishment
- (1) school, (2) school district
- sixth grade students
- unit of analysis

5.02 The levels of the independent variable in this study refers to which of the following.

- disruptive behavior
- (1) disruptive, (2) mildly disruptive, (3) not disruptive
- frequency of disruptive behavior
- playtime condition
- referral status
- (1) reward, (2) punishment
- (1) school, (2) school district
- sixth grade students
- unit of analysis
5.03 The dependent variable in this study refers to which of the following.
- disruptive behavior
- (1) disruptive, (2) mildly disruptive, (3) not disruptive
- frequency of disruptive behavior
- playtime condition

5.04 The dependent measure in this study refers to which of the following.
- disruptive behavior
- (1) disruptive, (2) mildly disruptive, (3) not disruptive
- frequency of disruptive behavior
- playtime condition

5.05 Identify the level of measurement (Nominal, Ordinal, Interval, or Ratio) for the dependent measure in this study.

5.06 What type of research design did this study use?
- single sample comparison
  [Uses a single sample. Evaluates difference between sample dependent measure values and known or theoretical population dependent measure values.]
- dependent samples comparison:
  matched-pairs
  [Uses two samples of matched sampling units (pairs). Evaluates difference between dependent measure values for the matched pairs of sampling units.]
- j=2 independent samples comparison:
  [Uses samples formed by assigning sampling units to two different levels of the independent variable or uses samples selected from two different populations. Evaluates difference between dependent measure values for the two samples.]
- two variable correlation:
  [Uses a single sample. Evaluates correlation between ordered pairs of dependent measure values (X,Y).]

5.07 Which, if any, statistical test(s) would be appropriate?
- chi-square test for the phi correlation
- 2xk chi-square test of independence
- dependent samples Student t test
- exact probability test or large sample Student t approximation for the Spearman correlation
- goodness-of-fit chi-square test
- independent samples Student t test
- McNemar change test
- single sample Kolmogorov-Smirnov test
- single sample Student t test
- Student t test for the Pearson correlation
- Wilcoxon/Mann-Whitney U test
- Wilcoxon T (matched pair, signed ranks) test
- none of these tests are appropriate.

5.08 Please explain your conclusion. Describe how this study meets or does not meet each of the two most important assumptions for statistical hypothesis testing. If you chose a specific test, describe how the study meets the additional assumptions for the specific test chosen.

Your discussion of the assumptions must support your conclusion about which, if any, test is appropriate. Your entire answer should be less than 200 words.