PRACTICE EXERCISES
INF 397C

Introduction to Research in Information Studies
School of Information
University of Texas at Austin

Dr. Philip Doty

Version 3.5
1. Define the following terms and symbols:

- \( n \)
- \( \sum x \)
- \( x \)
- \( \bar{x} \)
- \( s \)
- \( s^2 \)
- \( N \)
- \( \mu \)
- \( \sigma \)
- \( \sigma^2 \)
- coefficient of variation (CV)
- mode
- median
- arithmetic mean
- range
- variance
- interquartile range (IQR)
- standard deviation
- sample
- statistic
- parameter
- frequency distribution

2. A sample of the variable \( x \) assumes the following values:

9  11  13  3  7  2  8  9  6  10

Compute:
(a) \( n \)
(b) \( \sum x \)
(c) \( \bar{x} \)
(d) \( s \)
(e) \( s^2 \)
(f) median
(g) mode
(h) range
(i) CV
3. A sample of the variable $x$ assumes the following values:

\begin{center}
57  51  58  52  50  59  57  51  59  56  \\
50  53  54  50  57  51  53  55  52  54
\end{center}

Generate a frequency distribution indicating $x$, frequency of $x$, cumulative frequency of $x$, relative frequency of $x$, and cumulative relative frequency of $x$.

4. For the frequency distribution in problem 3, compute:

(a) $n$
(b) $\sum x$
(c) $\bar{x}$
(d) $s$
(e) $s^2$
(f) median
(g) mode
(h) range
(i) CV

5. Generate a histogram for the data in problem 3.

6. Generate a frequency polygon for the data in problem 3.

7. Define the following terms:

skewness
ordinate
abscissa
central tendency
bimodal
ordered pair
Cartesian plane
stem-and-leaf plot
outlier
reliability
dispersion or variability
negatively skewed
positively skewed
validity
box plot
whiskers
8. What is the relationship between or among the terms?

(a) sample/population
(b) $\bar{x}/\mu$
(c) $s/\sigma$
(d) variance/standard deviation
(e) $n/N$
(f) statistic/parameter
(g) mean, median, mode of the normal curve
(h) coefficient of variation/IQR

9. Graph the following sample distributions (histogram and frequency polygon) using three different pairs of axes.

<table>
<thead>
<tr>
<th>Score</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>4</td>
</tr>
<tr>
<td>26</td>
<td>10</td>
</tr>
<tr>
<td>27</td>
<td>6</td>
</tr>
<tr>
<td>28</td>
<td>3</td>
</tr>
<tr>
<td>29</td>
<td>3</td>
</tr>
<tr>
<td>30</td>
<td>1</td>
</tr>
<tr>
<td>31</td>
<td>4</td>
</tr>
<tr>
<td>32</td>
<td>6</td>
</tr>
<tr>
<td>33</td>
<td>5</td>
</tr>
<tr>
<td>34</td>
<td>4</td>
</tr>
<tr>
<td>35</td>
<td>8</td>
</tr>
<tr>
<td>36</td>
<td>5</td>
</tr>
<tr>
<td>37</td>
<td>3</td>
</tr>
<tr>
<td>38</td>
<td>2</td>
</tr>
<tr>
<td>39</td>
<td>1</td>
</tr>
<tr>
<td>40</td>
<td>10</td>
</tr>
<tr>
<td>41</td>
<td>12</td>
</tr>
<tr>
<td>42</td>
<td>8</td>
</tr>
<tr>
<td>43</td>
<td>10</td>
</tr>
<tr>
<td>44</td>
<td>13</td>
</tr>
<tr>
<td>45</td>
<td>24</td>
</tr>
<tr>
<td>46</td>
<td>14</td>
</tr>
<tr>
<td>47</td>
<td>20</td>
</tr>
</tbody>
</table>

10. For each of the distributions in problem 9, answer the following questions.

(a) Is the curve of the distribution positively or negatively skewed?
(b) What is n?
(c) Is the mode > median? Compute the answer and also answer it graphically, i.e., label the position of the mode and the median on the curve.
(d) Is the mean > mode? Compute the answer and also answer it graphically as in part (c) of this question.
(e) What is the variance of the distribution?
(f) What is the standard deviation of the distribution?
(g) What is the range of the distribution?
(h) What is the coefficient of variation of the distribution?
(i) Which measure of central tendency, mode, median, or (arithmetic) mean, is the fairest and clearest description of the distribution? Why?
11. Define:
  error model
  quartile
  percentile
  $Q_1$
  $Q_2$
  $Q_3$
  freq (x)
  cf (x)
  rel freq (x)
  cum rel freq (x)
  PR
  z-scores
  $\chi$
  deviation score
  $s$ of z-scores
  $\bar{x}$ of z-scores
  $\sum x$
  centile
  Interquartile Range (IQR)
  non-response bias
  self-selection

12. Define the relationship(s) between or among the terms:
  $Q_2$/median
  median/N or n
  z-score/raw score
  z-score/deviation score/standard deviation
  $Q_1$/ $Q_2$/ $Q_3$
  z-score/$\chi$/$s$ or $\sigma$
  median/fifth centile
  cf (x)/freq (x)/PR
  Literary Digest poll/bias

13. The observations of the values of variable $x$ can be summarized in the population frequency distribution below.

<table>
<thead>
<tr>
<th>$x$</th>
<th>freq (x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
</tr>
</tbody>
</table>
For this distribution of \( x \), calculate:

(a) Cumulative frequency, relative frequency, and cumulative relative frequency for each value
(b) \( N \)
(c) the range
(d) median
(e) mode
(f) \( \mu \)
(g) \( \sigma \)
(h) \( Q_1, Q_2, \) and \( Q_3 \)
(i) CV (coefficient of variation)
(j) IQR
(k) the percentile rank of \( x = 8, x = 2, x = 3 \)
(l) \( z \)-scores for \( x = 6, x = 8, x = 2, x = 3, x = 9 \)


15. Generate a box plot for the data in problem 3.

16. For a normally distributed distribution of variable \( x \), where \( \mu = 50 \) and \( \sigma = 2.5 \) [\( ND(50, 2.5) \)], calculate:

(a) the percentile rank of \( x = 45 \)
(b) the \( z \)-score of \( x = 52.6 \)
(c) the percentile rank of \( x = 58 \)
(d) the 29.12th percentile
(e) the 89.74th percentile
(f) the \( z \)-score of \( x = 45 \)
(g) the percentile rank of \( x = 49 \)

17. Define:
\[ \alpha \]
- sampling distribution
- Central Limit Theorem
- Standard Error (SE)
- \( ND(\mu, \sigma) \)
- decile
- descriptive statistics
- inferential statistics
- effect size
- confidence interval (C.I.) on \( \mu \)
- Student’s \( t \)
- degrees of freedom (df)
- random sampling
- stratified random sample
18. Define the relationship(s) between or among the terms:

\[ E(\bar{x}) / \mu \]
\[ \alpha / \text{df} / t \]
Central Limit Theorem/C.I. on \( \mu \)
\[ \alpha / \text{C.I. on } \mu \text{ when } \sigma \text{ is known} \]
\[ \alpha / \text{C.I. on } \mu \text{ when } \sigma \text{ is not known} \]
\[ z / \text{confidence interval on } \mu \]
\[ t / \text{C.I. on } \mu \]
\[ t / z \]

19. The following values indicate the number of microcomputer applications available to a sample of 10 computer users.

2, 5, 9, 5, 3, 6, 6, 3, 1, 13

For the population from which the sample was drawn, \( \mu = 4.1 \) and \( \sigma = 2.93 \).

Calculate:

a. The expected value of the mean of the sampling distribution of means
b. The standard deviation of the sampling distribution of means.

20. From previous research, we know that the standard deviation of the ages of public library users is 3.9 years. If the "average" age of a sample of 90 public library users is 20.3 years, construct:

a. a 95% C.I. around \( \mu \)

b. a 90% C.I. around \( \mu \)

c. a 99% C.I. around \( \mu \).

What does a 95% interval around \( \mu \) mean?

What is our best estimate of \( \mu \)?

21. The sample size in problem 20 was increased to 145, while the "average" age of the sample remained at 20.3 years. Construct three confidence intervals around \( \mu \) with the same levels of confidence as in Question 20.

22. Determine \( t \) for a C.I. of 95% around \( \mu \) when \( n \) equals 20, 9, \( \infty \), and 1.

23. Determine \( t \) on the same values of \( n \) (20, 9, \( \infty \), and 1) as in Question 22, but for a C.I. of 99% around \( \mu \). Should this new interval be narrower or wider than a 95% confidence interval on \( \mu \)? Why? Answer the question both conceptually and algebraically.
24. The following frequency distribution gives the values for variable $x$ in a sample drawn from a larger population.

<table>
<thead>
<tr>
<th>$x$</th>
<th>freq($x$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>43</td>
<td>6</td>
</tr>
<tr>
<td>42</td>
<td>11</td>
</tr>
<tr>
<td>39</td>
<td>3</td>
</tr>
<tr>
<td>36</td>
<td>7</td>
</tr>
<tr>
<td>35</td>
<td>7</td>
</tr>
<tr>
<td>34</td>
<td>15</td>
</tr>
</tbody>
</table>

Calculate:

(a) $\bar{x}$
(b) $s$
(c) $SE_\mu = \sigma_\bar{x}$
(d) $E(\bar{x})$
(e) $Q_1$, $Q_2$, and $Q_3$
(f) mode
(g) CV
(h) IQR
(i) a 95% C.I. on $\mu$
(j) the width of the confidence interval in part (i)
(k) a 99% C.I. on $\mu$
(l) the width of the confidence interval in part (k)
(m) PR (percentile rank) of $x = 42$
(n) our best estimates of $\mu$ and $\sigma$ from the data.

25. Generate a box plot for the data in problem 24.

26. Construct a stem-and-leaf plot for the following data set; indicate $Q_1$, $Q_2$, and $Q_3$ on the plot; and generate the six-figure summary. Be sure that you are able to identify the stems and the leaves and to identify their units of measurement.

The heights of members of an extended family were measured in inches. The observations were: 62, 48, 56, 37, 37, 26, 74, 66, 29, 49, 72, 77, 69, 62, and 64.
27. \( H_0 \): There is no relationship between computer expertise and minutes spent doing known-item searches in an OPAC at \( \alpha = 0.10 \).

Should we reject the \( H_0 \) given the following data? Remember that the acceptable error rate is 0.10.

**TIME (MINS)**

<table>
<thead>
<tr>
<th>EXPERTISE</th>
<th>( \leq 5 )</th>
<th>( &gt; 5, \leq 10 )</th>
<th>( &gt; 10 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novice</td>
<td>14</td>
<td>20</td>
<td>19</td>
</tr>
<tr>
<td>Intermediate</td>
<td>15</td>
<td>16</td>
<td>9</td>
</tr>
<tr>
<td>Expert</td>
<td>22</td>
<td>11</td>
<td>2</td>
</tr>
</tbody>
</table>

28. Answer Question 27 at an acceptable error rate of 0.05.

29. Define:

- statistical hypothesis
- \( H_0 \)
- \( H_1 \)
- \( \alpha \)
- Type I error
- Type II error
- \( \chi^2 \)
- nonparametric
- contingency table
- statistically significant
- \( E \) (expected value) in \( \chi^2 \)
- \( O \) (observed value) in \( \chi^2 \)

30. Discuss the relationship(s) between or among the terms:

- \( \alpha / \chi^2 \)
- \( df / R / C \) [in \( \chi^2 \) situation]
- \( H_0 / H_1 \)
- \( \chi^2 / \alpha / df \)
- \( \alpha / \text{Type I error} \)
- \( E / O / \chi^2 \)
- Type I / Type II error
SELECTED ANSWERS 3.5

2. (a) \( n = 10 \)

(b) \( \sum x = 78 \)

(c) \( \bar{x} = \frac{\sum x}{n} = \frac{78}{10} = 7.8 \)

(d) \( s = \sqrt{\frac{\sum x^2 - n\bar{x}^2}{n-1}} = \sqrt{\frac{714 - 608.4}{9}} = \sqrt{11.73} = 3.43 \)

(e) \( s^2 = 11.73 \)

(f) \( P(\text{med}) = \frac{n+1}{2} \text{ th score} = \frac{10+1}{2} \text{ th score} = 5.5^{th} \text{ score} \)

\[ \text{med} = 5.5^{th} \text{ score} = \frac{5\text{th score} + 6\text{th score}}{2} = \frac{8+9}{2} = 8.5 \]

(g) \( \text{mode} = 9 \)

(h) \( \text{range} = \text{highest value} - \text{lowest value} = 13 - 2 = 11 \)

(i) \( \text{coefficient of variation (CV)} = \frac{s}{\bar{x}} = \frac{3.43}{7.8} = 0.44 \)

4. (a) \( n = 20 \)

(b) \( \sum x = 1079 \)

(c) \( \bar{x} = \frac{\sum x}{n} = \frac{1079}{20} = 53.95 \)

(d) \( s = \sqrt{\frac{\sum x^2 - n\bar{x}^2}{n-1}} = \sqrt{\frac{58395 - 58212}{19}} = \sqrt{9.63} = 3.1 \)

(e) \( s^2 = 9.63 \)

(f) \( P(\text{med}) = \frac{n+1}{2} \text{ th score} = \frac{20+1}{2} \text{ th score} = 10.5^{th} \text{ score} \)

\[ \text{median} = \frac{53 + 54}{2} = 53.5 \]
(g) mode = 50, 51, 57 (trimodal)

(h) range = 59 - 50 = 9

(i) \( CV = \frac{s}{\bar{x}} = \frac{3.1}{53.95} = 0.06 \)

10. (a) Pos, Pos, Neg

(b) \( n = 28, 221, 217 \)

(c) mode = 26, 5, 12

\[ P(\text{median}) = 14.5\text{th}, 111\text{th}, 109\text{th observations} \]
\[ \text{median} = 26.5, 4, 10 \]

\[ \text{mode}_1 < \text{median}_1; \text{mode}_2 > \text{median}_2; \text{mode}_3 > \text{median}_3 \]

(d) mean = \( \bar{x} = \frac{\sum x}{n} = \frac{776 + 907 + 2058}{28} = 27.7, 4.1, 9.5 \)

\[ \text{mean}_1 > \text{mode}_1; \text{mode}_2 > \text{mean}_2; \text{mode}_3 > \text{mean}_3 \]

(e) \( s_1^2 = \frac{\sum x^2 - n\bar{x}^2}{n - 1} = \frac{22218 - (28)(27.7)^2}{28 - 1} = \frac{733.88}{27} = 27.18 \)

\[ s_2^2 = \frac{10887 - (221)(4.1)^2}{220} = \frac{7171.99}{220} = 32.60 \]

\[ s_3^2 = \frac{21948 - (217)(9.5)^2}{216} = \frac{2363.75}{216} = 10.94 \]

(f) \( s_1, s_2, s_3 = \sqrt{s_1^2} = \sqrt{27.18}, \sqrt{32.60}, \sqrt{10.94} = 5.21, 5.71, 3.31 \)

(g) Range = highest observation - lowest observation = H - L = 53 -25 = 28;
\[ 85 - 1 = 84; 14 - 2 = 12 \]

(h) \( CV = \frac{s}{\bar{x}} = \frac{5.21}{27.7}, \frac{5.71}{4.1}, \frac{3.31}{9.5} = 0.19, 1.39, 0.35 \)
13. (a) $x$ | freq $x$ | cum freq $x$ | rel freq $x$ | cum rel freq $x$
---|---|---|---|---
9  | 3  | 25  | 0.12  | 1.00  
8  | 9  | 22  | 0.36  | 0.88  
6  | 5  | 13  | 0.20  | 0.52  
3  | 2  | 8   | 0.08  | 0.32  
2  | 6  | 6   | 0.24  | 0.24  
25 |     |     | 1.00  |       

(b) $N = 25$

(c) Range = $H - L = 9 - 2 = 7$

(d) $P(\text{median}) = \frac{N + 1}{2} = \frac{13}{2}$

(e) mode = 8

(f) $\mu = \frac{\sum x}{N} = \frac{147}{25} = 5.88$

(g) $\sigma = \sqrt{\frac{\sum x^2 - N\mu^2}{N}} = \sqrt{\frac{1041 - (25)(5.88)^2}{25}} = \sqrt{\frac{174.64}{25}} = \sqrt{7.07} = 2.66$

(h) $P(Q_2) = 13\text{th observation}; Q_2 = 6$

\[
P(Q_1) = \frac{1 + P(Q_2)}{2} = \frac{1 + 13}{2} = 7\text{th observation from the beginning}; Q_1 = 3
\]

\[
P(Q_3) = \frac{1 + P(Q_2)}{2} = \frac{1 + 13}{2} = 7\text{th observation from the end}; Q_3 = 8
\]

(i) $CV = \frac{\sigma}{\mu} = \frac{2.66}{5.88} = 0.45$

(j) $IQR = Q_3 - Q_1 = 8 - 3 = 5$

(k) $PR = \% \text{ lower} + 1/2 \% \text{ of that score}$

\[
PR(8) = 0.52 + \frac{(0.36)}{2} = 0.70
\]

\[
PR(2) = 0 + \frac{(0.24)}{2} = 0.12
\]

\[
PR(3) = 0.24 + \frac{(0.08)}{2} = 0.28
\]
(l) $z = \frac{x - \mu}{\sigma}$ (for populations)

$$z_6 = \frac{x - \mu}{\sigma} = \frac{6 - 5.88}{2.66} = 0.05; z_8 = \frac{8 - 5.88}{2.66} = 0.80; z_2 = \frac{2 - 5.88}{2.66} = -1.46$$

$$z_3 = \frac{3 - 5.88}{2.66} = -1.08; z_9 = \frac{9 - 5.88}{2.66} = 1.17$$

16. ND(50, 2.5)

(a) PR(45); $z_{45} = \frac{x - \mu}{\sigma} = \frac{45 - 50}{2.5} = -2.00 \Rightarrow -0.4772$

$$0.5000 - 0.4772 = 0.0228, 2.28\%, 2.28\text{th percentile}$$

(b) $z_{52.6} = \frac{52.6 - 50}{2.5} = 1.04$

(c) PR(58); $z_{58} = \frac{58 - 50}{2.5} = 3.2 \Rightarrow 0.4993$

$$0.5000 + 0.4993 = 0.9993, 99.93\%, 99.93\text{rd percentile}$$

(d) find the 29.12th percentile

$$0.5000 - 0.2912 = 0.2088 \Rightarrow z = -0.55$$

$$-0.55 = \frac{x - 50}{2.5}$$

$$x - 50 = -1.375$$

$$x = 48.63$$

(e) find the 89.74th percentile

$$0.8974 - 0.5000 = 0.3974 \Rightarrow z = 1.27$$

$$z = 1.27 = \frac{x - 50}{2.5}$$

$$x = 53.175$$

(f) $z_{45} = \frac{45 - 50}{2.5} = -2.00$
(g) PR(49); $z_{49} = \frac{49 - 50}{2.5} = -\frac{1}{2.5} = -0.4 \Rightarrow -0.1554$

$0.5000 - 0.1554 = 0.3446$, 34.46%, 34.46$^{\text{th}}$ percentile

19. 2, 5, 9, 5, 3, 6, 6, 3, 1, 13

(a) $E(\bar{x}) = \mu = 4.1$

(b) $SE_{\mu} = \sigma = \frac{\sigma}{\sqrt{n}} = \frac{2.93}{\sqrt{10}} = 0.93$

20. $\sigma = 3.9$ years, $n = 90$, $\bar{x} = 20.3$ years

(a) 95% C.I. on $\mu$

$\bar{x} - zSE_{\mu} \leq \mu \leq \bar{x} + zSE_{\mu}$

$SE_{\mu} = \sigma = \frac{\sigma}{\sqrt{n}} = \frac{3.9}{\sqrt{90}} = 0.41$

$z_{95\%} = 1.96$ [Remember that $\frac{0.95}{2} = 0.4750 \Rightarrow 1.96 = z$]

$\bar{x} - zSE_{\mu} \leq \mu \leq \bar{x} + zSE_{\mu}$

$20.3 - (1.96)(0.41) \leq \mu \leq 20.3 + (1.96)(0.41)$

$19.5 \leq \mu \leq 21.1$  

interval width = 1.6

(b) 90% C.I. on $\mu$  

$z_{90\%} = 1.65$

$\bar{x} - zSE_{\mu} \leq \mu \leq \bar{x} + zSE_{\mu}$

$20.3 - (1.65)(0.41) \leq \mu \leq 20.3 + (1.65)(0.41)$

$19.62 \leq \mu \leq 20.98$  

interval width = 1.36

(c) 99% C.I. on $\mu$  

$z_{99\%} = 2.58$

$\bar{x} - zSE_{\mu} \leq \mu \leq \bar{x} + zSE_{\mu}$

$20.3 - (2.58)(0.41) \leq \mu \leq 20.3 + (2.58)(0.41)$

$19.25 \leq \mu \leq 21.35$  

interval width = 2.1

Our best estimate of $\mu$ is $\bar{x}$, 20.3 years
21. \( N = 145, \bar{x} = 20.3 \) years, \( \sigma = 3.9 \) years

   (a) 95\% C.I. on \( \mu \) \[ 19.67 \leq \mu \leq 20.93 \] interval width = 1.26

22. 95\% C.I. on \( \mu \), \( \therefore \alpha = 0.05 \)

\[ \begin{align*}
\text{df} & = n - 1 \\
\text{n} & = 20, 9, \infty, 1 \\
t & = 2.093, 2.306, 1.960, ?
\end{align*} \]

23. 99\% C.I. on \( \mu \), \( \therefore \alpha = 0.01 \)

\[ \begin{align*}
\text{df} & = n - 1 \\
t & = 2.861, 3.355, 2.576, ?
\end{align*} \]

24. (a) \( \bar{x} = \frac{\sum x}{n} = \frac{1844}{49} = 37.63 \)

   (b) \( s = \sqrt{\frac{\sum x^2 - n\bar{x}^2}{n-1}} = \sqrt{\frac{70048 - 69385}{48}} = 3.72 \)

   (c) \( SE_\mu = \sigma_x = \frac{s}{\sqrt{n-1}} = \frac{3.72}{6.93} = 0.537 \)

   (d) \( E(\bar{x}) = \mu \)

   (e) \( P(Q_2) = \frac{n+1}{2} \) th position = 25\textsuperscript{th} observation; \( Q_2 = 36 \)

\[ \begin{align*}
P(Q_1) = \frac{1 + \downarrow P(Q_1)}{2} & = \frac{1 + 25}{2} = 13\text{th observation from the beginning}; Q_1 = 34 \\
P(Q_3) & = \frac{1 + \downarrow P(Q_2)}{2} = \frac{1 + 25}{2} = 13\text{th observation from the end}; Q_3 = 42
\end{align*} \]

(f) mode = 34

(g) \( CV = \frac{s}{\bar{x}} = \frac{3.72}{37.63} = 0.10 \)

(h) IQR = \( Q_3 - Q_1 = 42 - 34 = 8 \)
(i) 95% C.I. on $\mu$, $\therefore \alpha = 0.05$

$SE_\mu = 0.537$ (see (c) above)

$df = 48$, $t = 2.021$

$\bar{x} - zSE_\mu \leq \mu \leq \bar{x} + zSE_\mu$

$37.63 - (2.021)(0.537) \leq \mu \leq 37.63 + (2.021)(0.537)$

$37.63 - 1.085 \leq \mu \leq 37.63 + 1.085$

$36.5 \leq \mu \leq 38.7$

(m) $PR_{42} = \%$ below $+ 1/2$ % of that score $= 0.65 + \frac{(0.22)}{2} = 0.76$
27.

<table>
<thead>
<tr>
<th>TIME (MINS)</th>
<th>≤5</th>
<th>&gt;5, ≤10</th>
<th>&gt;10</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXPERTISE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Novice</td>
<td>14 (21.12)</td>
<td>20 (19.46)</td>
<td>19 (12.42)</td>
</tr>
<tr>
<td>Intermediate</td>
<td>15 (15.94)</td>
<td>16 (14.69)</td>
<td>9 (9.38)</td>
</tr>
<tr>
<td>Expert</td>
<td>22 (13.95)</td>
<td>11 (12.85)</td>
<td>2 (8.20)</td>
</tr>
</tbody>
</table>

\[ \chi^2 = \frac{(O - E)^2}{E} \]

\[ E = \frac{R \times C}{n} \]

\[ \chi^2 = \frac{(14 - 21.12)^2}{21.12} + \frac{(15 - 15.94)^2}{15.94} + \frac{(22 - 13.95)^2}{13.95} + \frac{(20 - 19.46)^2}{19.46} + \frac{(16 - 14.69)^2}{14.69} \]

\[ + \frac{(11 - 12.85)^2}{12.85} + \frac{(19 - 12.42)^2}{12.42} + \frac{(9 - 9.38)^2}{9.38} + \frac{(2 - 8.20)^2}{8.20} \]

\[ \chi^2 = 2.40 + 0.06 + 4.65 + 0.01 + 0.12 + 0.27 + 3.49 + 0.02 + 4.69 = 15.71 \]

\[ \text{df} = (R - 1)(C - 1) = 2 \times 2 = 4 \]

\[ \chi^2 (4, 0.10) = 7.78 \]

15.71 > 7.78, reject \( H_0 \)

28. \[ \chi^2 (4, 0.05) = 9.49 \]

15.71 > 9.49, reject \( H_0 \)