Use the information in the following setting to answer questions 1 - 8
In order to assess the risk of impaired neurocognitive performance due to repeated concussive events investigators at the University of North Carolina at Chapel Hill interviewed 91 randomly selected soccer athletes, 96 randomly selected non-soccer athletes, and 53 non-athletes. This last group served as the control group. The main interest for the investigative team was whether there was compelling evidence for the existence of chronic neuropsychological dysfunction in the population of collegiate soccer players. Of interest here is whether or not the rates for concussive events are distributed evenly for the soccer athletes, non-soccer athletes and the controls. Table 1 provides the cross tabulated results for this study. Let $\alpha = 0.05$.


<table>
<thead>
<tr>
<th>Athlete Category</th>
<th>Number of Concussions</th>
<th>0 Concussions</th>
<th>1 Concussion</th>
<th>2 Concussions</th>
<th>3 or more Concussions</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soccer</td>
<td></td>
<td>45</td>
<td>25</td>
<td>11</td>
<td>10</td>
<td>91</td>
</tr>
<tr>
<td></td>
<td>$\chi^2_{11} = 3.71$</td>
<td>$\chi^2_{12} = 3.695$</td>
<td>$\chi^2_{13} = 0.85$</td>
<td>$\chi^2_{14} = 3.27$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Soccer</td>
<td></td>
<td>68</td>
<td>15</td>
<td>8</td>
<td>5</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>$\chi^2_{21} = 0.37$</td>
<td>$\chi^2_{22} = 0.073$</td>
<td>$\chi^2_{23} = 0.17$</td>
<td>$\chi^2_{24} = 0.17$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Athletes</td>
<td></td>
<td>45</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td>$\chi^2_{31} = 2.93$</td>
<td>$\chi^2_{32} = 2.45$</td>
<td>$\chi^2_{33} = 0.71$</td>
<td>$\chi^2_{34} = 3.3$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td></td>
<td>158</td>
<td>45</td>
<td>22</td>
<td>15</td>
<td>240</td>
</tr>
</tbody>
</table>

$Exp_{ij}$ = is the expected frequency for cell i,j.

$\chi^2_{ij}$ is the $\chi^2$ contribution for cell i,j.
Relationships between categorical variables: Concussion Events versus Sports Participation - Chi Square Analysis

1) Compute the expected cell frequencies and corresponding \( \chi^2 \) contributions for each cell. Use 2-decimal precision when computing and entering your expected values and 3-decimal precision for the \( \chi^2 \) contributions.

\[
E_{ij} = \frac{(O_{i.})(O_{.j})}{O_{..}}
\]

for example:

\[
E_{1a} = \frac{(91.45)(240)}{17.06} = 125.0
\]

2) The appropriate Ho/Ha pair for this setting is:

- \( H_0 \): no assoc. between prop. of concussions and sport categories
- \( H_a \): they are associated

3) The number of degrees of freedom for this study is:

\[(r-1)(c-1) = (4-1)(3-1) = 3 \times 2 = 6\]

4) The critical \( \chi^2 \) value is:

look at what value gives a 0.05 right-tail

\[ \chi^2_{0.05} = 12.59 \]

5) The decision rule is:

\[
\text{reject } H_0 \text{ if } \chi^2_{\text{test}} > \chi^2_{\text{crit}}
\]

6) The test statistic value, \( \chi^2_{\text{test}} \), is:

add up all the \( \chi^2 \)

\[ \chi^2_{\text{test}} = 22.1 \]

7) The statistical decision is: Reject Ho FTR Ho

\[ 22.1 > 12.59 \]

8) The English interpretation for this study is:

at \( \alpha = 0.05 \) we conclude that there is a significant association between the proportion of concussive events and the sport participation categories.
Relationships between categorical variables: Concussion Events versus Sports Participation - Chi Square Analysis

Use the information in the following setting to answer questions 9 through 11

Suppose we want to find a 90% confidence interval for the true difference in proportion suffering from 3 or more concussions for soccer athletes vs non-soccer athletes.

9) Find the critical value $z_{crit}$:

$z_{crit} = 1.645$

10) The margin of error is:

$$ME = z^* \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} = 1.645 \sqrt{\frac{10(\frac{91}{q_1})}{91} + \frac{5}{96} \left(\frac{91}{q_6}\right)} = 0.06557$$

11) Find the 90% CI:

$$\left(p_1 - p_2\right) \pm ME = \left(\frac{10}{91} - \frac{5}{96}\right) \pm 0.0656 = (-0.0078, 0.1234)$$

Use the information in the following setting to answer questions 12 through 15

Now we want to perform a hypothesis test to see if a greater proportion of non-athletes suffer 0 concussions than the proportion of soccer players that suffer 0 concussions. Perform the test at the $\alpha = .01$ level.

12) The appropriate Ho/Ha pair for this setting is:

$$H_0: \pi_{soccer} - \pi_{Non-Ath} \geq 0$$

$$H_a: "" < 0$$

13) The test statistic $z_{test}$ is:

$$\bar{p} = \frac{n_1p_1 + n_2p_2}{n_1 + n_2} = \frac{53 \cdot \frac{45}{53} + 91 \cdot \frac{45}{91}}{53 + 91} = \frac{90}{144} = 0.625$$

then:

$$z_{test} = \frac{(p_1 - p_2)}{\sqrt{\frac{\bar{p}(1-\bar{p})}{n_1} + \frac{\bar{p}(1-\bar{p})}{n_2}}} = -4.24$$

14) The p-value is:

$$0.0000$$

15) The English interpretation for this study is:

at $\alpha = 0.01$ we conclude the proportion of soccer athletes suffering 0 concussions is less than the proportion of non-athletes suffering 0 concussions.