Lesson 3-9: Drawing Conclusions

Learning Goals:
#1: How do we write conclusions for a hypothesis test?
#2: How do we use our graphic display calculator to complete a Chi-Squared hypothesis test?

Warm-Up:
1. A study compared noncombat mortality rates for U.S. military personnel who were deployed in combat situations to those not deployed. The results of a random sample of 1580 military personnel:

<table>
<thead>
<tr>
<th>Cause of Death</th>
<th>Unintentional Injury</th>
<th>Illness</th>
<th>Homicide</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deployed</td>
<td>183</td>
<td>30</td>
<td>11</td>
<td>224</td>
</tr>
<tr>
<td>Not deployed</td>
<td>784</td>
<td>264</td>
<td>308</td>
<td>1356</td>
</tr>
<tr>
<td>Total</td>
<td>967</td>
<td>294</td>
<td>319</td>
<td>1580</td>
</tr>
</tbody>
</table>

   a. State the null hypothesis

   \[ H_0: \text{Cause of Death and deployment are independent} \]

   b. Calculate the degrees of freedom.

   \[ df = (\text{# of rows}-1)(\text{# of columns}-1) = (2-1)(3-1) = 3 \]

   c. Show the expected value for being not deployed and dying due to an illness is 252.

   \[ E_{ij} = \frac{(\text{Row total})(\text{Column total})}{\text{Grand Total}} = \frac{(1356)(294)}{1580} = 252.318873 \approx 252 \]

   Running a Chi-Squared Test:

   Calculator Steps:

   Use this to get the Expected Values, \( X^2 \), p-value, and degrees of freedom.

   1. Enter the table into a MATRIX on the calculator
      2ND \( \rightarrow \) MATRIX (\( x^{-1} \)) \( \rightarrow \) EDIT \( \rightarrow \) [A]
   2. Enter the number of rows and columns for the given table
      row X column
   3. Calculate the expected value table
      STAT \( \rightarrow \) CALC \( \rightarrow \) TESTS \( \rightarrow \) C: \( x^2 \) Test...
      (Observed = [A], Expected = [B]) \( \rightarrow \) CALCULATE

   \[ X^2 = 5.811550483 \]
   \[ p = 0.1211474583 \]
   \[ df = 3 \]

   NEVER INCLUDE THE "TOTAL" CELLS WHEN TYPING INTO THE CALCULATOR

   BACK TO THE WARM-UP:

   1. Run a Chi-Squared using the data in the table above and state the \( X^2 \) and p-value.

   \[ X^2 = 51.90497331 \]
**IB Math Studies Yr 1**

**Practice Finding these Statistics:**

1. One hundred people were interviewed outside a chocolate shop to find out which flavor of ice cream they preferred. The results are given in the table below.

<table>
<thead>
<tr>
<th></th>
<th>Strawberry</th>
<th>Coffee</th>
<th>Orange</th>
<th>Vanilla</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>23</td>
<td>18</td>
<td>8</td>
<td>8</td>
<td>57</td>
</tr>
<tr>
<td>Female</td>
<td>15</td>
<td>6</td>
<td>12</td>
<td>10</td>
<td>43</td>
</tr>
<tr>
<td>Totals</td>
<td>38</td>
<td>24</td>
<td>20</td>
<td>18</td>
<td>100</td>
</tr>
</tbody>
</table>

a. Run a chi-squared test at a 5% significance level to determine the value of $\chi^2$.

\[ \chi^2 = 6.881306353 \]

b. Calculate the p-value.

\[ p = .0757787228 \]

2. The producer of a TV dancing show asked a group of 150 viewers their age and the type of Latin dance they preferred. The types of Latin dances in the show were Argentine tango, Samba, Rumba and Cha-cha-cha. The data obtained were organized in the following table.

<table>
<thead>
<tr>
<th>Dance</th>
<th>Argentine tango</th>
<th>Samba</th>
<th>Rumba</th>
<th>Cha-cha-cha</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 years old and younger</td>
<td>35</td>
<td>23</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>Older than 20 years old</td>
<td>20</td>
<td>17</td>
<td>18</td>
<td>15</td>
</tr>
</tbody>
</table>

a. Calculate the value of $\chi^2$, after a chi-squared test has been performed using the data above.

\[ \chi^2 = 6.553368506 \]

b. State the corresponding p-value

\[ p = .0875813773 \]
FINAL STEP: Drawing Conclusions
After you run your chi-squared test in the calculator, you must draw a conclusion about your null hypothesis.

Write your conclusion *in context*
- **CHI-SQUARED Critical Value**...A critical value is a number that separates the "reject the null hypothesis" statement from the "fail to reject the null."
- A **significance level**...is "our level of doubt". We also call it alpha (α). We usually pick a significance level of 10%, 5% or 1%.

---

**Drawing Conclusions In Context**

We can use the \( \chi^2 \) value or the p-value.
(These values always come from our calculator after running the Chi-Squared Test)

- **Compare the calculated \( \chi^2 \) to the CHI-SQUARED CRITICAL VALUE**
- **Compare the p-value to the SIGNIFICANCE LEVEL** (α). → always a decimal

<table>
<thead>
<tr>
<th>Greek Letters</th>
<th>&lt;</th>
<th>English Letters</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \chi^2 )</td>
<td>&lt;</td>
<td>Critical Value</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>&lt;</td>
<td>p-value</td>
</tr>
</tbody>
</table>

**ANYTIME YOU GET A LESS THAN SIGN:**

ACCEPT \( H_0 \)

---

**Let's try some examples:**

1) Determine if you would accept or reject the null hypothesis in each of the following cases:
   a) CV = 3.456 and \( \chi^2 = 3.654 \)
      \[ \chi^2 = 3.654 \geq \text{C.V., } = 3.456 \]
      Reject \( H_0 \)

   b) a 5% significance level and p-value = 0.032
      \[ \alpha = 0.05 \quad \square \quad p = 0.032 \]
      Reject \( H_0 \)

---

**Steps:**

1. Identify what values you are given
2. Identify which values can be compared to one another
3. Put the Greek letters on the left, and English on the right.
4. Draw your conclusion.
You Try some!
Draw a Final Conclusion based on the given statistics:

<table>
<thead>
<tr>
<th>CV = 8.731 and $\chi^2 = 7.921$</th>
<th>P-value = 0.23 and $\chi^2 = .01$, alpha = 0.02</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi^2 = 7.921 \quad \Box \quad \text{C.V.} = 8.731$</td>
<td>$\alpha = .02 \quad \Box \quad \text{p} = .23$</td>
</tr>
<tr>
<td>Accept $H_0$</td>
<td>Accept $H_0$</td>
</tr>
<tr>
<td>$\chi_{calc}^2 = 7.23$ and $\chi_{crit}^2 = 6.05 \quad \Box \quad p-value = 7.05$</td>
<td>$\alpha = .05 \quad \Box \quad \chi^2 = 6.416, \text{P-value} = .0231$</td>
</tr>
<tr>
<td>$\chi^2 = 7.23 \quad \Box \quad \text{C.V.} = 6.05$</td>
<td>$\alpha = .05 \quad \Box \quad \text{p} = .0231$</td>
</tr>
<tr>
<td>Reject $H_0$</td>
<td>Reject $H_0$</td>
</tr>
</tbody>
</table>

PUTTING ALL TOGETHER!!!!

2) The local park is used for walking dogs. The sizes of the dogs are observed at different times of the day. The table below shows the numbers of dogs present, classified by size, at three different times last Sunday.

<table>
<thead>
<tr>
<th></th>
<th>Small</th>
<th>Medium</th>
<th>Large</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morning</td>
<td>9</td>
<td>18</td>
<td>2</td>
<td>29</td>
</tr>
<tr>
<td>Afternoon</td>
<td>11</td>
<td>6</td>
<td>13</td>
<td>30</td>
</tr>
<tr>
<td>Evening</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>24</td>
</tr>
<tr>
<td>Total</td>
<td>27</td>
<td>32</td>
<td>24</td>
<td>83</td>
</tr>
</tbody>
</table>

a) The critical value, at the 5% level of significance, is 9.488. What conclusion can be drawn from this test? **Give a reason for your answer.**

\[ \chi^2 = 14.39958578 \quad \Box \quad \text{C.V.} = 9.488 \]

Since we are given both we can use either $\text{p}$ or $\chi^2$.

\[ \alpha = .05 \quad \Box \quad \text{p-value} = .0049199837 \]

Reject $H_0$. The size of dog and time of walk are dependent because $\chi^2 > \text{critical value.}$
3) The veterinarian has gathered the following data about the weight of dogs and the weight of their puppies.

<table>
<thead>
<tr>
<th>Puppy</th>
<th>Dog</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Heavy</td>
<td>Light</td>
</tr>
<tr>
<td>Heavy</td>
<td>36</td>
<td>27</td>
</tr>
<tr>
<td>Light</td>
<td>22</td>
<td>35</td>
</tr>
<tr>
<td>Total</td>
<td>58</td>
<td>62</td>
</tr>
</tbody>
</table>

The veterinarian wishes to test the following hypotheses.

H0: A puppy's weight is independent of its parent's weight.
H1: A puppy's weight is dependent of the weight of its parent.

a) Perform a chi-squared test at a 1% significance level. Should the null hypothesis be rejected? Justify your reasoning.

\[ \alpha = 0.01 \]

\[ \chi^2 = 0.43 \] \[ p = 0.52 \]

Accept H_0. The weight of a puppy and weight of the parents are independent because \( \alpha \) is less than the p-value.

Practice!

4) The eye colour and gender of 500 students are noted and the results are indicated in the table below.

<table>
<thead>
<tr>
<th></th>
<th>Blue</th>
<th>Brown</th>
<th>Green</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>18</td>
<td>152</td>
<td>50</td>
<td>220</td>
</tr>
<tr>
<td>Female</td>
<td>40</td>
<td>180</td>
<td>60</td>
<td>280</td>
</tr>
<tr>
<td>Total</td>
<td>58</td>
<td>332</td>
<td>110</td>
<td>500</td>
</tr>
</tbody>
</table>

It is believed that eye colour is related to gender in a school in Banff. It is decided to test this hypothesis by using a \( \chi^2 \) test at the 5% level of significance.

b) Determine the expected number of brown eyed females.

\[ E.V. = \frac{(280)(332)}{500} = 185.92 \]

c) Use your graphics display calculator to determine the p-value of this test.

\[ p = 0.1064651864 \]

d) Should the null hypothesis be rejected? Justify your reasoning.

\[ \alpha = 0.05 \] \[ p = 0.1064651864 \]

No, it should be accepted because \( \alpha \) is less than the p-value.
Lesson 3-8: Homework

1) The following contingency table displays a poll taken of 50 random people who like or dislike football, and also records their gender.

<table>
<thead>
<tr>
<th></th>
<th>Likes football</th>
<th>Dislikes football</th>
<th>sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>21</td>
<td>5</td>
<td>26</td>
</tr>
<tr>
<td>Female</td>
<td>7</td>
<td>17</td>
<td>24</td>
</tr>
<tr>
<td>sum</td>
<td>28</td>
<td>22</td>
<td>50</td>
</tr>
</tbody>
</table>

a. Write a null and alternative hypothesis

\[ H_0 : \text{Gender and liking football are independent} \]
\[ H_1 : \text{Gender and liking football are dependent} \]

b. Calculate a \( \chi^2 \) value.

\[ \chi^2 = 13.4870338 \]

c. Draw a conclusion, if the chi-squared critical value is 12. Justify your reasoning.

\[ \chi^2 = 13.4870338 \geq C.V. = 12 \]

Reject \( H_0 \), gender and liking football are dependent because \( \chi^2 > C.V. \)

2) State your conclusion based on the following calculator printout. Assume the test is being conducted with a 5% level of significance.

\[ X^2 = 5.811550483 \]
\[ P = .1211474583 \]
\[ df = 3 \]

\( \alpha = .05 \)

\( \alpha = .05 \leq P = .1211474583 \)

Accept \( H_0 \)!
3) The table below shows the scores for 12 golfers for their first two rounds in a local golf tournament.

<table>
<thead>
<tr>
<th>Round 1 (x)</th>
<th>71</th>
<th>79</th>
<th>66</th>
<th>73</th>
<th>69</th>
<th>76</th>
<th>68</th>
<th>75</th>
<th>82</th>
<th>67</th>
<th>69</th>
<th>74</th>
</tr>
</thead>
<tbody>
<tr>
<td>Round 2 (y)</td>
<td>73</td>
<td>81</td>
<td>68</td>
<td>75</td>
<td>70</td>
<td>79</td>
<td>69</td>
<td>77</td>
<td>83</td>
<td>68</td>
<td>72</td>
<td>76</td>
</tr>
</tbody>
</table>

(i) Write down the mean score in Round 1.

\[ \bar{x} = 72.41606667 \]

(ii) Write down the standard deviation in Round 1.

\[ s = 74.25 \]

a. Write down the correlation coefficient, \( r \).

\[ r = .990147403 \]

b. Write down the equation of the regression line of \( y \) on \( x \).

\[ y = 1.014045802 x + .8161832061 \]

c. Another golfer scored 70 in Round 1. Calculate an estimate of his score in Round 2.

\[ y = 1.014045802(70) + .8161832061 \]

\[ y = 71.79938935 \]

d. Another golfer scored 89 in Round 2. Determine whether you can use the equation of the regression line to estimate his score in Round 1. Give a reason for your answer.

\[ 89 = 1.014045802 x + .8161832061 \]

\[ -.8161832061 \]

\[ 88.18381679 = 1.014045802 x \]

\[ \frac{88.18381679}{1.014045802} \]

\[ x = 86.96236069 \]