# Lesson Plan

This lesson was developed by Taylor Hamilton of James Campbell High School.

**Grade Level: 11th & 12th**

**Standard(s): AP Statistics Standards**

**Length of time to implement the lesson/activity: 1 class period**

**Lesson Plan:**

LESSON PLAN — Mr. Hamilton
AP Stats & Trig/Pre-Calculus — **DATE:** Wednesday 3/29/17

***Period 1: AP Statistics***

**OBJECTIVE:** Today ALL students will learn about Chi-Squared Goodness of Fit through the CRB

**OPENING (10 min):** Teacher will make announcements for the day. Students will listen to the announcements.

*Announcements:*
*Read Bulletin
-STAR testing will be next week Monday
-Today we will learn about one of the chi-squared tests on goodness of fit
-The month of April will be a crazy month. Make sure you are prepared and know what's going on with testing and myself being out for two day towards the end of the month for a conference.
-Any questions?

**EXAMPLES (30 min):** Teacher will conduct a scaffolded example problem with decreasing guidance and built in CFU's. Students will take notes on the example problem, ask and answer questions, and reason through the problem.

*Answer Key Attached
-Lecture on Chi-Squared Goodness of Fit
Give students an intro about the Coconut Rhinoceros Beetle
-Go over two Chi-Squared Goodness of Fit example on the CRB

*CFUs:*
-What are the similarities between chi-squared and other tests we have done?
-When do you use chi-squared?
-How will you make sure you can differentiate between chi-squared and one/two prop z-tests?
-Why is it important to look at this example on the CRB?
-Can someone explain the entire process we went through to solve this problem?

**ACTIVITY/GROUP WORK (35 min):** Teacher will let students complete the second problem on the crossing of CRB species and facilitate quick presentations on the problem. Students will complete the problem, construct a presentation on chi-squared, and share with the class.

-Walk around and ask them follow up CFU questions
-Break up the significance test and have each POD share one section of the four-step process due to time

**CLOSING:** Teacher will conclude the lesson. Students will clean up work spaces and prepare for next class.
*Homework:* See week sheet

1

**Assessment and Rubric**

CHAPTER 11: INFERENCE FOR DISTRIBUTIONS OF CATEGORICAL DATA Name: _____
AP Statistics Period: ___ Date: _____

**11-1 Chi Square Goodness of Fit Test:**

(1)_____ distributions were used for _____.

(2) ____ distributions were used for _____.

(3)_____ distributions will be used for _____.

Picture:

The **chi-square distributions** are a family of distributions that take only _____ values and are skewed to the _____. A specific chi-square distribution is specified by one parameter, called the _____. df = _____ where n is _____.

Properties:

- The area under a chi-square curve is equal to _____.
- Each chi-square curve begins at _____, increases to a peak, and then approaches the horizontal axis asymptotically from above.
- As the degrees of freedom increases, the distribution becomes _____.
- The mean of a particular chi-square distribution is equal to _____.
- The mode (peak) of the chi-square density curve is at _____.

Conditions that must be met in order to use the chi-square distributions

1. **Random:** The data come from a _____ sample or a _____ experiment.
2. **Large Sample Size:** All expected counts are at least ____.
3. **Independent:** Individual observations are _____. When sampling without replacement, we must also check the _____ condition.

Two warnings:

1. The chi-square test statistic compares observed and expected _____ and **NOT PROPORTIONS!**
2. When checking the large sample size condition, be sure to look at _____ counts and **NOT OBSERVED COUNTS!**

Example 1: Local scientists believe that the Coconut Rhinoceros Beetle is becoming a huge problem in the State of Hawaii, especially Oahu as it is destroying some of our plants here in our community. After careful qualitative research, they believe that certain communities on Oahu have a big influence on how destructive the Coconut Rhinoceros Beetle (CRB) can be. To see if certain communities are related to the success the CRB is having, a random sample of 80 CRB's were selected and their location was recorded. Overall, 32 were found in Ewa Beach, 20 in Kapolei, 16 in Waianae, and 12 in Pearl City. Does the data provide convincing evidence that CRB success in certain communities are not uniformly distributed throughout the island?

Follow up analysis

When a chi-square test gives a P-value that is significant, we must perform a follow up analysis. To do this, we find the largest _____ of the chi-square test statistic.

Example 2: Biologists come up with a crazy and mad idea to cross pairs of Coconut Rhinoceros Beetles (CRB) having genetic makeup Bb, indicating that each CRB has one dominant gene (B) and one recessive gene (b) for its rhino horn. Each offspring CRB will receive one gene for the rhino horn from each parent. The following table, often called a *Punnett square*, shows the possible combinations of gene received by the offspring:

| | | Parent 2 passes on: **B** | **b** |
|---|---|---|---|
| Parent 1 passes on: | **B** | BB | Bb |
| | **b** | Bb | bb |

The Punnet square suggests that the expected ratio of rhino horn (BB) to stunted rhino horn (Bb) to no rhino horn (bb) CRB should be 1:2:1. In other words, the biologists predict that 25% of the offspring will have the rhino horn, 50% will have a stunted rhino horn, and 25% will have no rhino horn.

To test their hypothesis about the distribution of offspring, the biologist's mate 84 randomly selected pairs of CRB with and without the rhino horn. Of the 84 offspring, 23 had the rhino horn, 50 had the stunted rhino horn, and 11 had no rhino horn. Do these data differ significantly from what the biologists have predicted?

## A.P. Statistic Chapter and Unit Test Rubrics
## Campbell High School – Hamilton and Tong

| *50% Multiple Choice* | *50% Free Response (E, P, I)* |
|---|---|
| • 10+ = 4 | • All Exemplary = 4 |
| • 8 – 9 = 3.5 | • Mostly Exemplary/Some Partial = 3.5 |
| • 7 = 3 | • Mostly Partial/Some Exemplary = 3 |
| • 5 – 6 = 2.5 | • Some Partial/Some Incomplete = 2.5 |
| • 4 – = 2 | • Mostly Incomplete = 2 |

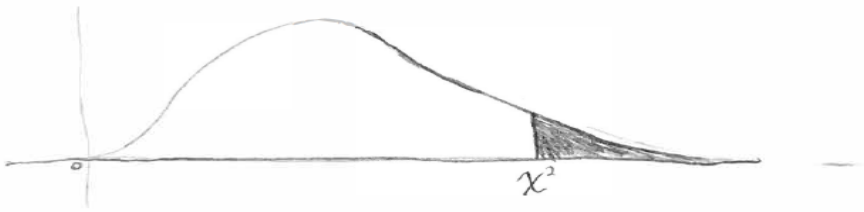
| ***AP Exam Calculator*** | ***1*** | ***2*** | ***3*** | ***4*** |
|---|---|---|---|---|
| **Q1** | 1.875 | 3.750 | 5.625 | 7.500 |
| **Q2** | 1.875 | 3.750 | 5.625 | 7.500 |
| **Q3** | 1.875 | 3.750 | 5.625 | 7.500 |
| **Q4** | 1.875 | 3.750 | 5.625 | 7.500 |
| **Q5** | 1.875 | 3.750 | 5.625 | 7.500 |
| **Q6** | 3.125 | 6.250 | 9.375 | 12.500 |

# Student Sample

**CHAPTER 11: INFERENCE FOR DISTRIBUTIONS OF CATEGORICAL DATA**
**AP Statistics** **Period: ___ Date: _____**
**11-1 Chi Square Goodness of Fit Test:**

(1) normal distributions were used for 1 & 2 prop z test.
(2) t distributions were used for 1 & 2 samp t test.
(3) chi-square distributions will be used for $\chi^2$ G.o.f, homogeneity, association.

Picture:




The **chi-square distributions** are a family of distributions that take only positive values and are skewed to the right. A specific chi-square distribution is specified by one parameter, called the degrees of freedom (df). df = n-1 where n is number of categories.

Properties:

- The area under a chi-square curve is equal to 1.
- Each chi-square curve begins at 0, increases to a peak, and then approaches the horizontal axis asymptotically from above.
- As the degrees of freedom increases, the distribution becomes more normal.
- The mean of a particular chi-square distribution is equal to df.
- The mode (peak) of the chi-square density curve is at df - 2.

Conditions that must be met in order to use the chi-square distributions

1. **Random:** The data come from a random sample or a randomized experiment.
2. **Large Sample Size:** All expected counts are at least 5.
3. **Independent:** Individual observations are independant. When sampling without replacement, we must also check the 10% condition.

Two warnings:

1. The chi-square test statistic compares observed and expected counts and **NOT PROPORTIONS!**
2. When checking the large sample size condition, be sure to look at expected counts and **NOT OBSERVED COUNTS!**

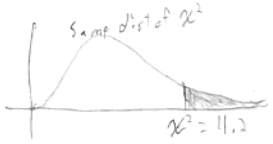
**CHAPTER 11: INFERENCE FOR DISTRIBUTIONS OF CATEGORICAL DATA** **Name:** _____
**AP Statistics** **Period:** ___ **Date:** _____

Example 1: Local scientists believe that the Coconut Rhinoceros Beetle is becoming a huge problem in the State of Hawaii, especially Oahu as it is destroying some of our plants here in our community. After careful qualitative research, they believe that certain communities on Oahu have a big influence on how destructive the Coconut Rhinoceros Beetle (CRB) can be. To see if certain communities are related to the success the CRB is having, a random sample of 80 CRB's were selected and their location was recorded. Overall, 32 were found in Ewa Beach, 20 in Kapolei, 16 in Waianae, and 12 in Pearl City. Does the data provide convincing evidence that CRB success in certain communities are not uniformly distributed throughout the island?

State: $H_0$: CRB destruction is uniformly distributed throughout the island  $H_\alpha$: CRB destruction isn't uniformly distributed throughout the island  $\alpha = 0.05$

plan: $\chi^2$ GoF test
random - yes, random sample of 80 CRB
large sample size - yes, $80(.25) = 20 \geq 5$ The samp. dist. is approximately norm because all expected counts are $\geq 5$
independant - yes, $80 \leq (\frac{1}{10})$(thousands of CRB), so we can assume independance

do:




$df = n - 1 = 4 - 1 = 3$

$\chi^2 = \sum \frac{(O-E)^2}{E} = \frac{(32-20)^2}{20} + \frac{(20-20)^2}{20} + \frac{(16-20)^2}{20} + \frac{(12-20)^2}{20}$

$\chi^2 = 7.2 + 0 + 0.8 + 3.2 = 11.2$

$\chi^2$ CDF → p-value = 0.010692  table C → p-value ≈ 0.01

conclude: Assuming $H_0$ is true (CRB destruction is uniformly distributed), there is a 0.010692 probabil of getting $\chi^2$ of 11.2 or greater purely by chance. This provides strong evidence against $H_0$ at $\alpha = 0.05$ level. Therefore we reject $H_0$ and conclude that CRB destruction aren't uniformly distributed. Because this is significant, the largest component of $\chi^2$ is 7.2 because Ewa Beach had much more destruction than expected

Follow up analysis
When a chi-square test gives a P-value that is significant, we must perform a follow up analysis. To do this, we find the largest component of the chi-square test statistic.

**CHAPTER 11: INFERENCE FOR DISTRIBUTIONS OF CATEGORICAL DATA** **Name:** _____
**AP Statistics** **Period:** ___ **Date:** _____

Example 2: Biologists come up with a crazy and mad idea to cross pairs of Coconut Rhinoceros Beetles (CRB) having genetic makeup Bb, indicating that each CRB has one dominant gene (B) and one recessive gene (b) for its rhino horn. Each offspring CRB will receive one gene for the rhino horn from each parent. The following table, often called a *Punnett square*, shows the possible combinations of gene received by the offspring:

| | | Parent 2 passes on: **B** | **b** |
|---|---|---|---|
| Parent 1 passes on: | **B** | BB | Bb |
| | **b** | Bb | bb |

The Punnet square suggests that the expected ratio of rhino horn (BB) to stunted rhino horn (Bb) to no rhino horn (bb) CRB should be 1:2:1. In other words, the biologists predict that 25% of the offspring will have the rhino horn, 50% will have a stunted rhino horn, and 25% will have no rhino horn.

To test their hypothesis about the distribution of offspring, the biologist's mate 84 randomly selected pairs of CRB with and without the rhino horn. Of the 84 offspring, 23 had the rhino horn, 50 had the stunted rhino horn, and 11 had no rhino horn. Do these data differ significantly from what the biologists have predicted?

State: $H_0$: the proportion of CRB offspring with a horn is 0.25, stunted horn is 0.5, and no horn is 0.25 $H_a$: The proportions for each category don't match the predictions
$\alpha = 0.05$ $df = n-1 = 2$

plan: $\chi^2$ GoF test - check conditions
random - yes, "84 randomly selected pairs"
large sample size - yes, $84(.25) = 21$ $84(.5) = 42$ $\geq 5$ The sampling distribution is approximately $\chi^2$ because all expected values are $\geq 5$
independant - yes, $84 \leq \frac{1}{10}$ (thousands of CRB), so we assume independance

do: samp dist of $\chi^2$


$\chi^2 = \sum \frac{(o-E)^2}{E} = \frac{(23-21)^2}{21} + \frac{(50-42)^2}{42} + \frac{(11-21)^2}{21} = 0.19 + 1.52 + 4.76 = 6.47$

p-value = 0.03936

conclude: Assuming $H_0$ (predicted proportions are true) there is a 0.03936 probability of obtaining a $\chi^2$ value of 6.47 or greater purely by chance. This provides strong evidence against $H_0$ at $\alpha = 0.05$ level. Therefore, we reject $H_0$ and conclude that the proportion of CRB horns don't match the predicted 1:2:1 ratio. Because this is significant, the largest component of $\chi^2$ is 4.76 since the no horn count was less than expected