**Chi Square Analysis Exercise:**  **Name:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Ap Biology Hr.:\_\_\_\_\_\_\_\_\_**

**Background Information:**

Statistics can be used to determine if differences among groups are significant, or simply the result of predictable error. The test most often used to determine if data is good is the Chi-square test. This test can determine if deviations from the expected values are due to chance alone or to some other circumstance.

**Formula:** **o** = Observed number of individuals

 **e** = Expected number of individuals **X2 =** $Σ$ **(o-e)2**

$Σ$= The sum of the values **e**

 (in this case, the difference, squared & divided by the expected number.)

**Example 1: Corn Seedlings**. An F1 cross between Heterozygous parents ( Gg) for color is made. A punnett square of the cross (Gg X Gg) would predict that the expected ratio of green : albino seedlings would be 3:1 (GG, Gg, Gg, gg).

**# Observed (o)** = Number from the collected data. From the above cross 72 green seedling and 12 albino seedlings were produced. Enter those values into the data table below. Add the two values to get the total observed.

**# Expected (e)=** Change the ratio into a fraction (3= ¾, 1= ¼) and multiply by total observed (84). The expected # of green seedlings is .75 x 84 = 63. Enter this in the data table below and calculate the # for albino seedlings.

* There is a small difference between the observed and the expected results, but is this data close enough that the difference can be explained by random chance or variation in the sample?

To determine if the observed data falls within the acceptable limits, a Chi-square analysis is preformed to test the validity of a **null hypothesis**; that there is no statistically significant differences between the observed and the expected data. If the Chi-square analysis indicates that the data varies too much from the expected an alternative hypothesis is accepted.

* Complete the Chi-square analysis to examine the null hypothesis, to see if the data from the cross above will be expected to fit the 3;1 ratio. Follow the directions given in the chart.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Phynotype/Genotype** | **# Observed (o)** | **# Expected (e)** | **(o-e)** | **(o-e)2** | **(o-e)2****e** |
| **Green** |  |  |  |  |  |
| **Albino** |  |  |  |  |  |
| **Total** |  |   **X2 =** $Σ$ **(o-e)2/e**  |  |

|  |  |
| --- | --- |
| **Probability****(p)** | **Degrees of Freedom (df)** |
| **1** | **2** | **3** | **4** | **5** |
| **0.05** | **3.84** | **5.99** | **7.82** | **9.49** | **11.1** |
| **0.01** | **6.64** | **9.21** | **11.3** | **13.2** | **15.1** |
| **0.001** | **10.8** | **13.8** | **16.3** | **18.5** | **20.5** |

**Critical Values Table:**

**1 . Determine the degrees of freedom (df)**.

Take the number of phenotypes and subtract 1. Since there are 2 possible (green & albino) in this experiment df =1

**2. Find the p value.**

Under the df column, find the critical value in the probability p=0.05 row. The value is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_. If the Chi-square is greater than or equal to the critical value then the null hypothesis is rejected. This means there is no statistically significant difference between the observed & expected data. In other words chance alone cannot explain the deviations observed and there is a reason to doubt the original hypothesis. The minimum probability for rejecting a null hypothesis is generally 0.05.

* These results are said to be significant at the probability of p=0.05. This means that only 5% of the time would you expect to see similar data if the null hypothesis were correct; you are 95% sure that the data does not fit the 3:1 ratio.
* Since the data does not fit the 3;1 ratio. Reasons must be considered for the variation. Additional experimentation would be necessary. Perhaps the testing size was too small or errors were made in the data collection.

**Problem 1:** In a study of incomplete dominance in tobacco seedlings, the following counts were made from a cross between two heterozygous (Gg) plants:

1. Complete a Punnett square for the cross to determine the expected ratio.

Show your work and write the ratio.

1. Fill in the data below and complete the analysis.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Phynotype/Genotype** | **# Observed (o)** | **# Expected (e)** | **(o-e)** | **(o-e)2** | **(o-e)2****e** |
| **Green (GG)** | **22** |  |  |  |  |
| **Yellow-Green (Gg)** | **50** |  |  |  |  |
| **Albino (gg)** | **12** |  |  |  |  |
| **Total** | **84** |   **X2 =** $Σ$ **(o-e)2/e**  |  |

1. How many degrees of freedom are there?\_\_\_\_\_\_\_\_
2. Do you accept or reject the null hypothesis? Why?

**Problem 2:** When pure-breeding long-wing ***Drosophila*** are mated with pure-breeding short-winged flies, the F1 offspring have an intermediate wing length. When several intermediate wing-length flies are allowed to interbreed the following results are obtained; 230 long-wings, 510 intermediate length wings, 260 short wings.

1. Write the genotype for the F1 intermediate wing-length flies. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
2. Write a hypothesis describing the mode of inheritance of wing length in the ***Drosophila*** (this is your null hypothesis)
3. Complete the following table:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Phynotype/Genotype** | **# Observed (o)** | **# Expected (e)** | **(o-e)** | **(o-e)2** | **(o-e)2****e** |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
| **Total** |  |   **X2 =** $Σ$ **(o-e)2/e**  |  |

1. How many degrees of freedom are there? \_\_\_\_\_\_\_\_

1. E. What is the probability value for this data?\_\_\_\_\_\_\_
2. Do you accept or reject the null hypothesis? Why?