In genetics, statistics can be particularly helpful when trying to figure out if your genes are behaving in the way you think they should be!

**The Chi- Squared Test**

Bio12 AP **Genetics, Stats & Math… Reunited at Last! ;-)**  Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_

Date: \_\_\_\_\_\_\_\_\_\_\_\_\_\_

**AP Science Practice 2: The student can use mathematics appropriately**

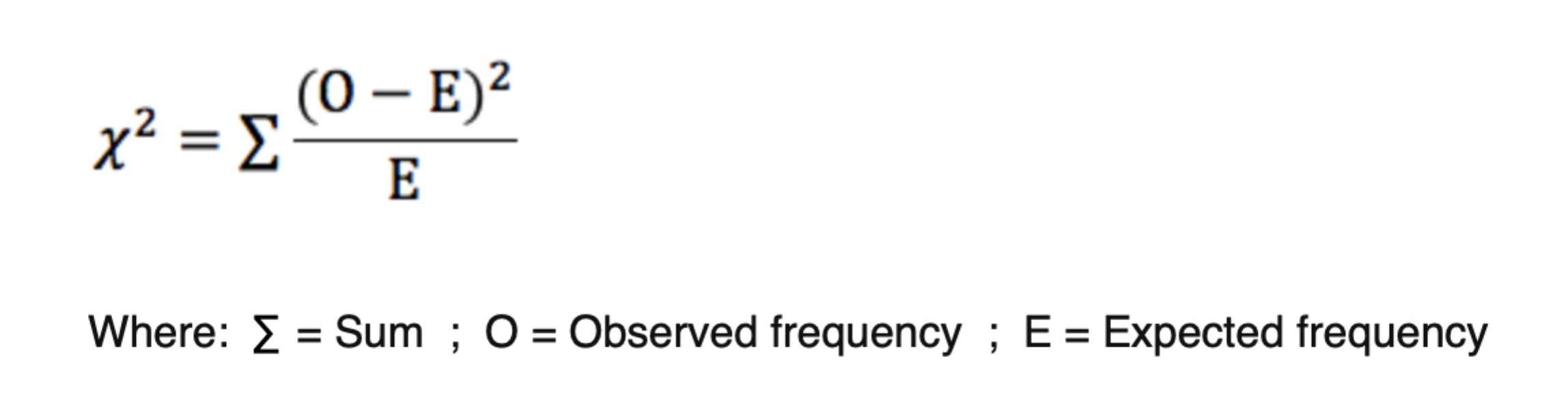
*-justify the selection of a mathematical routine to solve problems*

*-apply mathematical routines to quantities that describe natural phenomena*

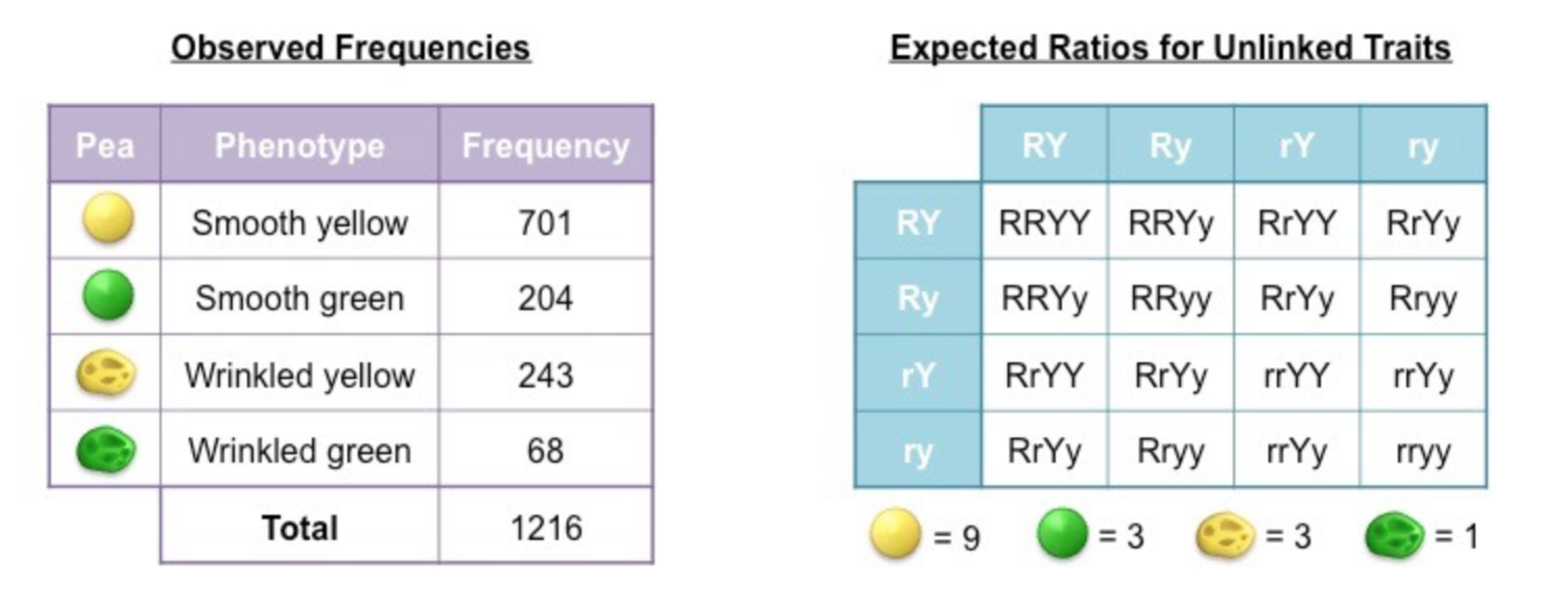
*-estimate numerically quantities that describe natural phenomena*

**Chi-squared (χ2) tests** can be applied to data from dihybrid crosses and are super useful to determine whether the difference between an observed and expected frequency distribution is statistically significant 🡪 ie is the predicted 9:3:3:1 ratio being upheld? .

**5 Steps to completing a chi-squared test:**

* Identify hypotheses (null versus alternative)
* Construct a table of frequencies (compare observed versus expected phenotypes)
* Apply the chi-squared formula 
* Determine the degree of freedom (df)
* Identify the p value (should be <0.05 for statistical significance) & compare calculated x2  to the critical x2 value at p<0.05

**Let’s Try It!**

****

**Step 1: Identify the Hypotheses**

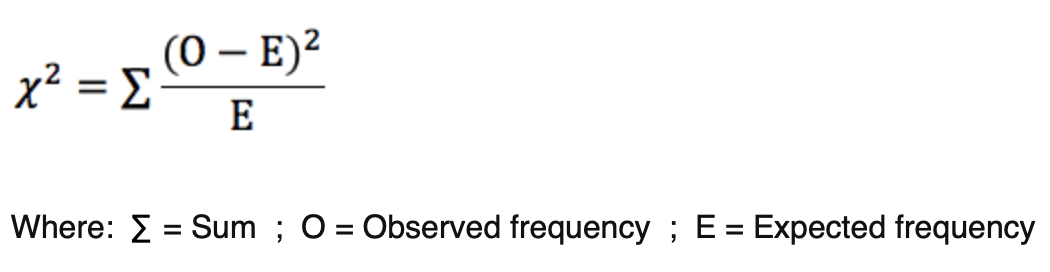
Null Hypothesis (HO): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Alternate Hypothesis (HA): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Step 2: Construct a table of frequencies**

A table must be constructed that compares observed and expected frequencies for each possible phenotype. Expected frequencies are calculated by first determining the expected ratios and then multiplying against the observed total

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Smooth Yellow** | **Smooth Green** | **Wrinkled Yellow** | **Wrinkled Green** | **Total** |
| **Observed** | 701 | 204 | 243 | 68 | 1216 |
| **Expected** | 1216 x (9/16)  =684 |  |  |  |  |

****

**Step 3: Apply the chi-squared formula**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Smooth Yellow** | **Smooth Green** | **Wrinkled Yellow** | **Wrinkled Green** |
| **Observed (O)** | 701 |  |  |  |
| **Expected (E)** | 684 |  |  |  |
| **(O – E)** | 17 |  |  |  |
| **(O-E)2**  **E** | 0.42 |  |  |  |

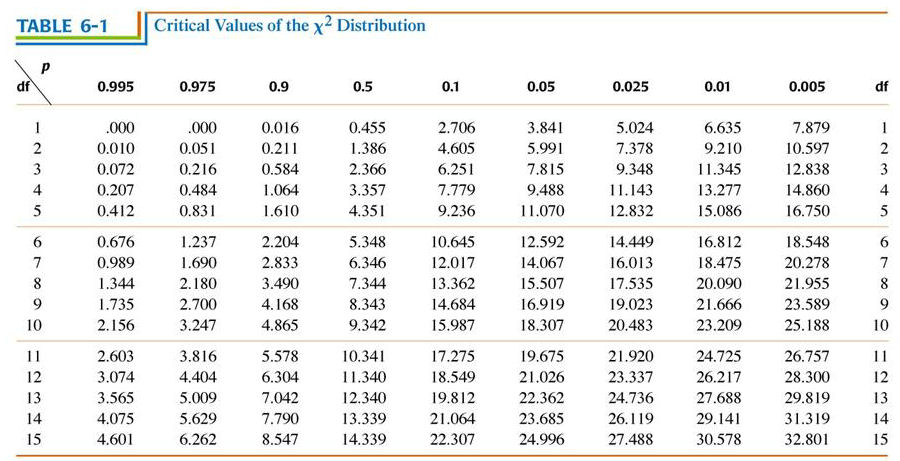
**X2= the sum of the last row**

**Step 4: Determine the Degrees of Freedom (df)-** Degrees of freedom take into account the number of different classes represented by the cross, in this case = 4 (since 4 different phenotypes). The more classes you have, the greater the variation you are likely to see from any expected numbers. ***To calculate degrees of freedom, you subtract 1 from the number of classes.* *Our example df=***

**Step 5: Identify the critical value at p=0.05 & compare –**  Find the critical x2 value in the table for the df determined at p = 0.05, (or less) this level gives us 95 % confidence! **If our calculated x2 number is HIGHER than the critical value at p= 0.05 then we REJECT the null hypothesis**. It means there is something aside from just chance accounting for our results. **If our calculated x2 is below the critical p value then we ACCEPT the null hypothesis.**



**Apply the table. What can we conclude from our example?**



**Practice-**

**1)**  In sweet pea plants, the trait for purple flowers (P) is dominant to red flowers (p). Similarly, the trait for long pollen (L) is dominant to round pollen (l).

Two heterozygotes are crossed, yielding the following frequencies for the F1 generation:296 purple long pollen plants, 19 purple round pollen plants, 27 red long pollen plants, 85 red round pollen plants.

Use the chi-squared test to determine if these results are due to independent assortment. Include a null and alternate hypothesis and a table to show all calculations