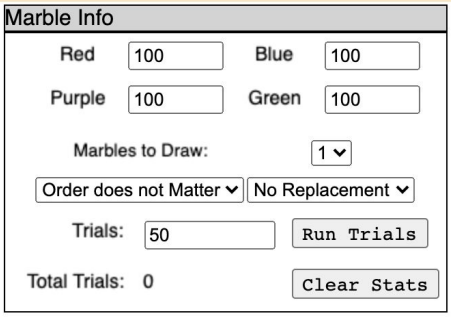
You've lost your marbles!
An Introduction to Chi-Square Analysis



Part I: In this activity, we will use a simulated bag of marbles to determine if a random draw matches statistical expectations. Go to this marble drawing simulator and set up the conditions following the picture to the right. You are filling a bag with 400 marbles (100 of each of four colors) and will pull out 50, one at a time, and will not return the pulled ones to the bag. Remember to use actual numbers of marbles, NOT percentages, in all your calculations.

Begin by writing a **null hypothesis.** What do you expect to happen when you pull 50 marbles from this bag?

I would expect 12 of them to be red, 12 of them to be blue, 12 of them to be purple, and 13 of them to be green.

Data:

	observed	**expected**	**(o-e)**	**$(o-e)^2/e$**
Red	13	12	13 - 12 = 1	1/12 = 0.083
Blue	8	12	8 - 12 = -4	16/12 = 1.333
Purple	16	12	16 - 12 = 4	16/12 = 1.333
Green	13	13	13 - 13 = 0	0/12 = 0.000
			$X^2 = \Sigma(o-e)^2/e$	2.750

Write a conclusion for your experiment in the table below. Include the following elements in every summary of a chi-square calculation that you complete in AP Biology.

Conclusion:

a. Identify the critical value at p = 0.05	b. Identify the degrees of freedom (n-1)	c. State the X^2 sum that you calculated.
7.81 is the critical value	3 degrees of freedom	2.750 is my chi-squared sum

d. Explain your results. How does your X^2 sum compare to the critical value, and what does this allow you to conclude about the significance of the variation in your data? Do you reject, or fail to reject, the null hypothesis?

The chi-squared sum for this experiment (at 2.75) was far below the critical threshold (at 7.81). The variation from our expected values and actual values was not due to any phenomena, it was the result of mere chance (there was more than a 25% chance that these results would occur even with our null hypothesis being true). Thus, I failed to reject my null hypothesis.

Part II: Design your own! Now, make up an experiment of your own. You must still draw one marble at a time, and the marbles cannot be returned to the bag (no replacement). However, you can change the number of colors (set quantities to 0 for colors you want to omit), can change the numbers of marbles (they don't have to be equal), and/or change the number of marbles drawn. Explain how you will set up your experiment here:

I will use 100 purple marbles, 100 green marbles, and 200 blue marbles (no red marbles this time). I will still, as required, draw one marble at a time and refrain from returning marbles to the bag (no replacement).

What is the null hypothesis?

Out of 50 marbles drawn, I expect to draw 25 blue marbles, 12 purple marbles, and 13 green marbles.

Data:

	observed	**expected**	**(o-e)**	**$(o-e)^2/e$**
Red				
Blue	33	25	33 - 25 = 8	64/25 = 2.560
Purple	12	12	12 - 12 = 0	0/12 = 0.000
Green	5	13	5 - 13 = -8	64/13 = 4.923
			$X^2 = \Sigma(o-e)^2/e$	7.483

Conclusion:

a. Identify the critical value at p = 0.05

5.99 is the critical value

b. Identify the degrees of freedom (n-1)

2 degrees of freedom

c. State the X^2 sum that you calculated.

7.483 is my chi-squared sum

d. Explain your results. How does your X^2 sum compare to the critical value, and what does this allow you to conclude about the significance of the variation in your data? Do you reject, or fail to reject, the null hypothesis?

The chi-squared sum for this experiment (at 7.483) was much greater than the critical threshold (at 5.99). The variation from our expected values and actual values *was* due to a phenomena (there was less than a 5% chance that these results would occur with our null hypothesis being true). Thus, I have rejected my null hypothesis.