Name: _____ Date: _____ Period: _____

<u>Chi-Square (X^2) Tutorial and Practice</u>
AP Biology

When do I use a Chi square (X^2) test?

The **Chi-square test** is a statistical method that makes a comparison between the data collected in an experiment (i.e., the observed results) versus the data an investigator expected to find. The Chi-square test is a way determine <mark>whether the difference between observed and expected results is due to random chance alone or if some other factor is involved</mark>. For instance, if you toss a coin 10 times, although your **expected results** are 5 heads and 5 tails, you will often get a different result (ex: 6 heads and 4 tails), without anything being unusual with your coins or the way you tossed them. Therefore, the difference between the observed and expected results for the coin toss may be due to random chance alone.

Ultimately, the Chi square test allows us to determine a probability value (p value) for our data. A <mark>p value of 0.05</mark> means that there is a 5% chance that the difference between the observed and expected results is due to random chance alone. This means that there is a 95% chance that the difference between the observed and expected results is real and significant, meaning some factor other than chance caused the difference.

If the Chi square test yields a <mark>p value above 0.05</mark>, **scientists fail to reject the null hypothesis.** This means that there may not be a statistically significant difference between the observed and expected values. Let's say the p value is 0.3 (which is greater than 0.05). This means that there is a 30% chance that the difference between the observed and expected results is due to random chance alone and only a 70% chance that the difference is caused by another real and significant factor. As such, it's pretty likely that the difference is due to random chance and that there may not be a statistically significant difference between the observed and expected results. <mark>**Again, this means that we fail to reject the null hypothesis.**</mark>

If the Chi square test yields a <mark>p value below or equal to 0.05, scientists reject the null hypothesis.</mark> This means that there may be a statistically significant difference between the observed and expected values. Let's say the p value is 0.02 (below 0.05). This means that there is only a 2% chance that the difference between the observed and expected results is due to random chance alone and a 98% chance that the difference is caused by another real and significant factor. As such, it's pretty likely that the difference is due to another factor and that there may be a statistically significant difference between the observed and expected results. <mark>**As such, this means that we reject the null hypothesis.**</mark>

1. **Example #1**

For example, let's say that we are doing genetic testing on a population of elephants. (This takes us back to when we learned how to complete Punnett squares in pre-AP biology). The initial population has elephants that are all heterozygous (Aa) for the trait of trunk length. Trunk length in elephants is controlled by a single gene, where the dominant allele (A) codes for the long trunk phenotype and the recessive allele (a) codes for the short trunk phenotype. Researchers sampled 100 offspring from this initial population and found the following phenotype frequencies. These are the OBSERVED results.

***Phenotype = physical trait

Observed Results

Phenotype	Number of Offspring <mark>Observed</mark> with this Phenotype
Long Trunks	80
Short Trunks	20

We can use or knowledge of genetics to predict our expected results. If we know that all parents in the initial population are heterozygous (Aa) for the trait of trunk length, we can predict the expected offspring phenotype frequencies using a Punnett square (see below).

	A	A
A	AA (Long Trunk)	Aa (Long Trunk)
a	Aa (Long Trunk)	aa (Short Trunk)

According to this Punnett square, we would predict that 75% of the offspring would have long trunks and 25% of the offspring would have short trunks. If we know the number of offspring sampled (100), we can use this and the predicted phenotype frequencies to predict the number of offspring elephants out of the sample with long trunks and short trunks. These are the <mark>**EXPECTED**</mark> results.

Expected Results

Phenotype	Predicted Frequencies	Number of Offspring Predicted with this Phenotype
Long Trunks	75% or 0.75	100 x 0.75 = 75
Short Trunks	25% or 0.25	100 x 0.25 = 25

Using the Chi square test, we would be able to determine if there is a statistically significant difference between the observed results (80 long trunks / 20 short trunks) and the expected results (75 long trunks / 25 short trunks)

2. **Example #2**

Suppose we were trying to determine if there is a statistically significant difference between the number of turtles with brown shells and the number of turtles with green shells in a population. We sampled 200 turtles from the population and found that 92 of them were brown-shelled and 108 of them were green-shelled.

For the purposes of our Chi square test, we will call these our <mark>**OBSERVED**</mark> results.

Observed Results

Turtle Shell Color	Number of Turtles
Brown Shells	92
Green Shells	108

If we had predicted that we would find equal numbers of brown-shelled turtles and green-shelled turtles (i.e. 50% of each), then our **EXPECTED** results are as follows…
("There is no statistically significant difference between the number of turtles on this island with brown and green shells").

Expected Results

Turtle Shell Color	Predicted Frequencies	Number of Turtles
Brown Shells	50% or 0.50	200 x 0.50 = 100
Green Shells	50% or 0.50	200 x 0.50 = 100

Using the Chi square test, we would be able to determine if there is a statistically significant difference between the observed results (92 brown shells / 108 green shells) and the expected results (100 brown shells / 100 green shells).

How do I perform a Chi square test?

1. **State the null hypothesis**

    -   This is a negative statement, basically saying that <mark>there is no statistically significant difference between our observed and expected results</mark>

        For Example #1 given above, our null hypothesis would be… "There is no statistically significant difference between the number of long-trunked and short-trunked offspring observed in the population and the number of long-trunked and short-trunked offspring expected based on our Punnett square."

        For Example #2 given above, our null hypothesis would be… "There is no statistically significant difference between the number of brown-shelled and green-shelled turtles observed in the population and the number of brown-shelled and green-shelled turtles expected." Because we expect to have equal numbers of each shell color, we could write this null hypothesis more simply as… "<mark>There is no statistically significant difference between the number of brown-shelled and green-shelled turtles in the population.</mark>"

2. **Determine your expected values**

    -   The way you calculate your expected values will be different for each situation
    -   We have already done this for Example #1 and Example #2 given above.
    -   Expected and observed values are always whole numbers. This is why we converted our expected frequencies/percentages (ex: 75% long trunks and 25% short trunks for Example #1) to whole numbers of elephants (ex: 75 with long trunks and 25 with short trunks).

3. **Calculate X^2**

    -   The formula is: $X^2 = \sum \frac{(o-e)^2}{e}$ *(this formula is on the AP Bio Formula Sheet)*
    -   Where o = observed value, e = expected value, and $\sum$ = the sum of

- So you would need to calculate $\frac{(o-e)^2}{e}$ separately for each value (ex: each phenotype from Example #1) and then add the results together. See a sample calculation below for Example #1.

Phenotype	Observed	Expected	(o – e)²/e
Long Trunks	80	75	$(80-75)^2/75 = 0.33$
Short Trunks	20	25	$(20-25)^2/25 = 1$
	Σ (this Greek letter sigma means sum)	Total (this is your X² value!)	1.33

4. You will also need to know the degrees of freedom.

- This is calculated using the formula (n-1)
- where n = the number of sets of results, or the number of categories we're looking at (ex: the number of possible phenotypes from Example #1)
- For Example #1… degrees of freedom = n-1 = 2-1 = 1

5. Compare the X² value against a table of critical values.

- On the table below, refer to the row that corresponds to the correct number of degrees of freedom for your data set
- Look up the critical number at the intersection of the correct degrees of freedom and the p = 0.05 column. "p" stands for probability level. **Scientists almost always use a 0.05 probability level.**
- For Example #1, the critical value (aka critical number) is 3.84 (see circled value on the chart below).

Degrees of Freedom	Probability				
	0.9	0.5	0.1	0.05	0.01
1	0.02	0.46	2.71	3.84	6.64
2	0.21	1.39	4.61	5.99	9.21
3	0.58	2.37	6.25	7.82	11.35
4	1.06	3.36	7.78	9.49	13.28
5	1.61	4.35	9.24	11.07	15.09

6. Make a conclusion

- **If the X^2 value that you calculated in Step 3 is higher than the critical value at the p = 0.05 level then you can reject the null hypothesis.** In other words, there may be a *statistically significant difference between the observed and expected results.* (i.e. the observed results do not "match" the expected results)

  If the X^2 value is higher than the critical value at a p value of 0.05, that means that the true p value for your data is lower than 0.05. (A high X^2 value corresponds with a low p value.) Remember from the information on the first page of this packet that a p value lower than 0.05 means that it is likely that there is some real and significant factor causing the differences between the observed and expected values (the differences are statistically significant and caused by that factor!).

- **If the X^2 value is less than the critical number then you fail to reject the null hypothesis.** In other words, there may not be a statistically significant difference between the observed and expected results. (The observed results "match" the expected results)

  If the X^2 value is lower than the critical number at a p value of 0.05, that means that the true p value for your data is higher than 0.05. (A low X^2 value corresponds with a high p value.) Remember from the information on the first page of this packet that a p value higher than 0.05 means that it is likely that any differences between the observed and expected values are due to random chance alone.

  Note: A low X^2 value corresponds with a high p value (above 0.05)

  For Example #1, the calculated Chi square value **(1.33) is lower than the critical value (3.84)**, **so we fail to reject the null hypothesis**. This means there is not a statistically significant difference between the observed and expected results. In other words, there is not a statistically significant difference between the number of offspring observed with each phenotype and the number expected to have each phenotype based on the Punnett square, and so our null hypothesis holds.

Now… please complete the two problems on the following pages to practice using Chi square analysis

Problem #1

Naked mole rats are a burrowing rodent native to parts of East Africa. They have a complex social structure in which only one female (the queen) and one to three males reproduce, while the rest of the members of the colony function as workers. Mammal ecologists suspected that they had an unusual male to female ratio. They counted the numbers of each sex in one colony.

Sex	Number of animals
Female	52
Male	34

State the Null hypothesis

Calculate the expected results

Calculate the chi-squared value

Sex	Observed	Expected	$(o - e)^2/e$
Female	52		
Male	34		
	Σ	Total (this is your X^2 value!)	

What are the degrees of freedom?

DF =

Compare the Chi square (X^2) value with the critical value/number from the chart below

Degrees of Freedom	Probability				
	0.9	0.5	0.1	0.05	0.01
1	0.02	0.46	2.71	3.84	6.64
2	0.21	1.39	4.61	5.99	9.21
3	0.58	2.37	6.25	7.82	11.35
4	1.06	3.36	7.78	9.49	13.28
5	1.61	4.35	9.24	11.07	15.09

Make a conclusion (Do you reject or fail to reject your null hypothesis? What does that mean for THIS scenario?)

Problem #2

You have been wandering about on a seashore and you have noticed that a small snail (the flat periwinkle) seems to live only on seaweeds of various kinds. You decide to investigate whether the animals prefer certain kinds of seaweed by counting numbers of animals on different species. You end up with the following data

Type of Seaweed	Number of animals on each kind of seaweed
serrated wrack	45
bladder wrack	38
egg wrack	10
spiral wrack	5
other algae	2
Total	100

State the Null hypothesis

Calculate the expected results

Calculate the chi-squared value

Seaweed	Observed	Expected	$(o - e)^2/e$
serrated wrack	45		
bladder wrack	38		
egg wrack	10		
spiral wrack	5		
other algae	2		
	Σ	Total (this is your X^2 value!)	

What are the degrees of freedom?

DF =

Compare the calculated value with the critical value/number from the chart below

Degrees of Freedom	Probability				
	0.9	0.5	0.1	0.05	0.01
1	0.02	0.46	2.71	3.84	6.64
2	0.21	1.39	4.61	5.99	9.21
3	0.58	2.37	6.25	7.82	11.35
4	1.06	3.36	7.78	9.49	13.28
5	1.61	4.35	9.24	11.07	15.09

Make a conclusion (Do you reject or fail to reject your null hypothesis? What does that mean for THIS scenario?)