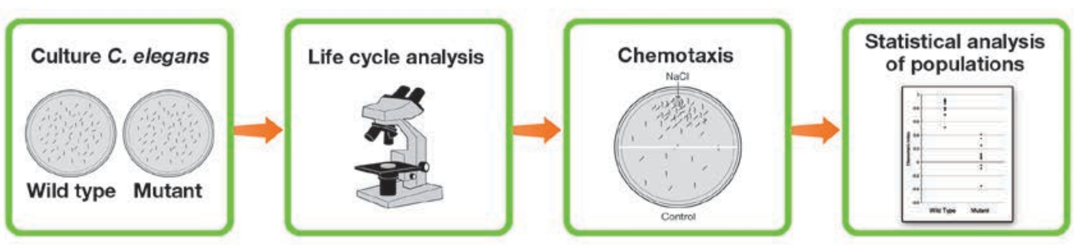
Biotechnology Explorer™

C. elegans Behavior Kit

Chi Square Analysis Supplement

explorer.bio-rad.com

Catalog #166-5120EDU



This kit contains temperature-sensitive reagents.
Open immediately and see individual components for storage temperature.

Please see redemption instructions on how to receive your *C. elegans*.

Duplication of any part of this document is permitted for classroom use only.

Please visit explorer.bio-rad.com to access our selection of language translations for Biotechnology Explorer kit curricula.

Chi Square (χ^2) Analysis of Distribution

Chi square (χ^2) is a statistical test that compares observed data collected in the course of an experiment with data expected to result from the experiment based on a hypothesis. The chi square test is a method to determine whether differences between observed and expected results are due to random chance, or if an unforeseen factor is influencing the outcome of the experiment. In other words, the chi square is a test to help answer the question: "Do the results of my experiment support the hypothesis?" To illustrate how to use chi square analysis let's look at an example that uses M&M'S color distribution as a model.

Example of Chi Square Analysis

Suppose you are given a 12.60 ounce pack of milk chocolate M&M'S and you are asked to predict how many candies of each color are within the bag. M&M'S come in 6 colors; cyan blue, orange, green, bright yellow, red, and brown. Having no additional information you hypothesize that each pack of milk chocolate M&M'S would have an equal distribution of each color. Let's do an experiment step by step to determine if your hypothesis is correct.

Step 1: State the Hypothesis

First, we need to come up with a null hypothesis for the data and an alternative hypothesis of what is going on if the null hypothesis is rejected. In general form, this can look like the following:

H_0 (null hypothesis): The data are consistent with an expected result.

H_a (alternative hypothesis): The data are **not** consistent with an expected result.

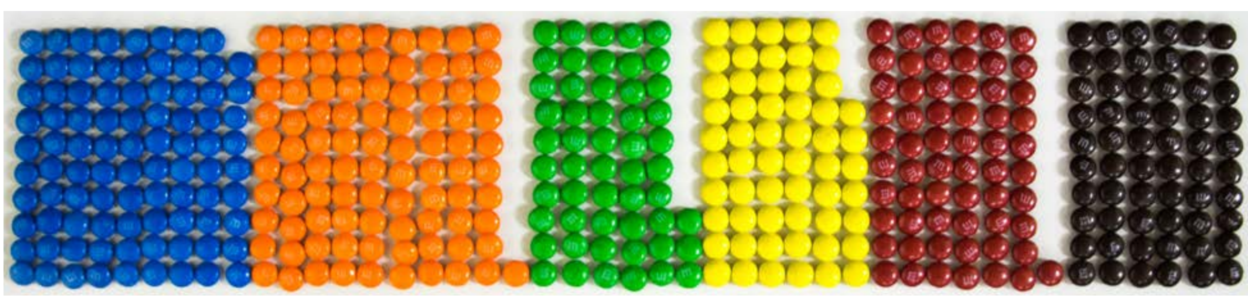
You hypothesize that all the colors of M&M'S are equally distributed. This is the null hypothesis. The alternative hypothesis would be that not all colors of M&M'S are equally distributed. Therefore the hypotheses for this experiment are as follows:

H_0 (null hypothesis): All colors of M&M'S are equally distributed.

H_a (alternative hypothesis): All colors of M&M'S are **not** equally distributed.

Step 2: Create a table to analyze your data and calculate the chi square value

You open a bag of M&M'S containing 411 candies. If your hypothesis is correct, then the bag will contain 68.5 candies of each color. With the hypothesis set, you sort and count each color and log your data in the chart below:



Cyan Blue Orange Green Bright Yellow Red Brown

Expected and observed data table

Color	Cyan Blue	Orange	Green	Bright Yellow	Red	Brown	Total M&M'S
Expected Quantity	68.5	68.5	68.5	68.5	68.5	68.5	411
Actual Quantity	89	91	53	57	61	60	411

Do the data fit your hypothesis? Looking at the data, cyan blue and orange candies are the highest represented colors, while green is the least represented color in the bag. It is unlikely that each bag of M&M'S will contain all colors in the same exact ratio. But do the observed results differ enough from the expected results to say that your null hypothesis is rejected? To determine whether your null hypothesis can be statistically rejected we will use chi square analysis.

The chi square value can be calculated using the following equation:

$\chi^2 = \Sigma(o_i - e_i)^2/e_i$, where:

χ^2 = chi square value

o_i = observed/measured value i, where i = 1, 2, etc. up to the number of different conditions

e_i = expected value i, where i = 1, 2, etc. up to the number of different conditions

Σ = sum of all the values from i = 1 to number of different conditions

To help up to calculate the χ^2 value, we will create a table that includes the data that we collected and calculate $(o - e)^2/e$ for each condition. Once all values have been calculated we will add up the $(o - e)^2/e$ values for each condition. The sum of all of the $(o - e)^2/e$ values is our χ^2 value.

Chi Square Table

H_0 (null hypothesis): All colors of M&M'S are equally represented in the bag. H_a (alternate hypothesis): All colors of M&M'S are **not** equally represented in the bag.			
	Observed, o	Expected, e	$(o - e)^2/e$
Condition 1, Cyan blue	89	68.5	6.14
Condition 2, Orange	91	68.5	7.39
Condition 3, Green	53	68.5	3.51
Condition 4, Bright yellow	57	68.5	1.93
Condition 5, Red	61	68.5	0.82
Condition 6, Brown	60	68.5	1.05
Σ, Sum	411	411	20.81 (χ^2)

In this example, the χ^2 value is 20.84.

Step 3: Calculate Degrees of Freedom

Next you will calculate the degrees of freedom for your experiment. The degree of freedom of each of these experiments is the number of conditions minus 1.

DF (degrees of freedom) = # of conditions – 1

In this case there are 6 conditions: cyan blue, orange, green, bright yellow, red, and brown. Therefore, the degree of freedom = 6 – 1 = 5

Step 4: Determine Significance Level

The significance level (p-value) is the probability that a χ^2 rejection of your null hypothesis will be incorrect. For example, if the significance level is 0.05 and your chi square test shows that your null hypothesis is correct (your data statistically match the expected results), then there is a 5% chance that the null hypothesis is not, in fact, correct. The lower the significance level is set, the less your chance of making a statistically incorrect conclusion. Typically, scientists set the significance level to **0.05**.

Step 5: Use a Chi Square Distribution Table to Determine the χ^2 Critical Value

We have determined that the degrees of freedom in our example are **5** and that we want a significance level of 0.05. Using these values we can determine the χ^2 critical value using the chi square distribution table below. A χ^2 value greater than the χ^2 critical value indicates that there is a statistically significant difference between the observed and expected values. Therefore, if our calculated χ^2 value is greater than the χ^2 critical value, then we can negate the null hypothesis. On the other hand, a χ^2 value less than the χ^2 critical value indicates that there is not a significant difference between observed and expected values. Therefore, if our calculated χ^2 value is less than the χ^2 critical value, then we cannot negate the null hypothesis. According to the table, the χ^2 critical value for our experiment is 11.07.

Chi Square Distribution Table

Degrees of freedom	p-value	1	2	3	4	5	6	7	8
Significance level	0.05	3.84	5.99	7.82	9.49	11.07	12.59	14.07	15.51

In the case of our data, χ^2 = 20.84, which is greater than 11.07 so the null hypothesis can be rejected. Therefore, the hypothesis that all colors of M&M'S are equally distributed is considered untrue.

Practice Exercise

The makers of M&M'S, Mars Chocolate North America, were contacted to find out what color distribution they include in their M&M'S candies. They replied to our inquiry with an email stating that milk chocolate M&M'S are made up of 24% cyan blue, 20% orange, 16% green, 14% bright yellow, 13% red, and 13% brown candies. However, in their response they also state, "Each large production batch is blended to those ratios and mixed thoroughly. However, since the individual packages are filled by weight on high-speed equipment, and not by count, it is possible to have an unusual color distribution."

Use the percentages provided by Mars Chocolate North America to set your expected values, and use the observed distribution of colors provided in the above experiment as your observed values. Fill out the table below and determine whether the M&M'S color distribution observed is statistically in line with the expected proportions set by the company.

H_0 (null hypothesis): M&M'S are represented in the bag in the following proportions: 24% cyan blue, 20% orange, 16% green, 14% bright yellow, 13% red, and 13% brown.

H_a (alternate hypothesis): M&M'S are **not** represented in the bag in the following proportions: 24% cyan blue, 20% orange, 16% green, 14% bright yellow, 13% red, and 13% brown.

	Observed, o	Expected, e	$(o - e)^2/e$
Condition 1, Cyan blue	89		0.94
Condition 2, Orange	91		0.94
Condition 3, Green	53		2.48
Condition 4, Bright yellow	57		0.01
Condition 5, Red	61		1.07
Condition 6, Brown	60		0.81
Σ, Sum	411		χ^2= ______

Degrees of freedom:	Significance Level = **0.05**

Does χ^2 analysis reject the null hypothesis?

Yes _____ No _____

Practice Exercise Solution

H_0 (null hypothesis): M&M'S are represented in the bag in the following proportions: 24% cyan blue, 20% orange, 16% green, 14% bright yellow, 13% red and 13% brown.

H_a (alternate hypothesis): M&M'S are **not** represented in the bag in the following proportions: 24% cyan blue, 20% orange, 16% green, 14% bright yellow, 13% red and 13% brown.

	Observed, o	Expected, e	$(o - e)^2/e$
Condition 1, Cyan blue	89	98.64	0.94
Condition 2, Orange	91	82.2	0.94
Condition 3, Green	53	65.76	2.48
Condition 4, Bright yellow	57	57.54	0.01
Condition 5, Red	61	53.43	1.07
Condition 6, Brown	60	53.43	0.81
Σ, Sum	411	411	χ^2= 6.25

Degrees of freedom: 5

Significance Level = 0.05

Does χ^2 analysis reject the null hypothesis?

Yes_____ No __X__

In the practice exercise, the expected values for each color are determined by multiplying the total number of candies by the percent representation as specified by Mars Chocolate North America. For example, Mars Chocolate North America states that 24% of candies are cyan blue. Therefore the expected value for cyan blue is 411 x 0.24 = 98.64.

Once all expected values are calculated, the χ^2 value is calculated. In this example χ^2 = **6.25**. We determine the degrees of freedom to be **5**. This has not changed from the previous example as there are still 6 conditions (candy colors) in each bag. We subtract 1 from our number of conditions to arrive at the degrees of freedom. Next we look up the χ^2 critical value in the chi square distribution table. Because the degrees of freedom and the significance level have not changed from the previous example, the χ^2 critical value is still 11.07.

The calculated χ^2 value, 6.25, is lower than the χ^2 critical value of 11.07. This means that we **cannot** reject the null hypothesis. Therefore, we can conclude that Mars Chocolate North America did provide the color distribution of candies that they said they would.

Chi Square Analysis of your *C. elegans* experiment

The experiment performed in class culminated in counting *C. elegans* that migrated to the NaCl side of the plate and *C. elegans* that migrated to the control side of the plate. *C. elegans* migration toward NaCl is used as a measure of associative learning behavior since only a population of *C. elegans* that is capable of associating NaCl with favorable feeding environments migrates towards NaCl. Using the data collected, a chemotaxis index (CI) was calculated that allows you to graph and visually determine differences in behavior between the wild-type and mutant *C. elegans* populations. However, how do the observed results compare to our hypothesis that *C. elegans* will migrate toward NaCl after conditioning with food (*E. coli*)? Do the wild-type or mutant results accept or reject this null hypothesis?

Bio-Rad scientists have performed hundreds of NaCl chemotaxis experiments on *C. elegans* and thus expected values are well established. According to previous experiments, 91% of wild-type *C. elegans* typically migrate toward NaCl side of the plate while the remaining 9% of *C. elegans* will migrate toward the control side of the plate.

Use the data described above to set your expected value, and use the data collected in the course of your experiment as the observed values.

Fill out the worksheet below and determine whether your experimental results support the following null hypothesis:

H_0 (null hypothesis): *C. elegans* will migrate towards NaCl after conditioning with food.

You can calculate the percent migration as follows:

$$\text{Percent } C.\ elegans \text{ on NaCl side} = \frac{\#\ C.\ elegans \text{ on NaCl side}}{\text{Total } C.\ elegans}$$

$$\text{Percent } C.\ elegans \text{ on Control side} = \frac{\#\ \text{C. elegans on Control side}}{\text{Total } C.\ elegans}$$

Wild-Type *C. elegans* Data Analysis

H_0 (null hypothesis): The data are consistent with an expected result.

H_a (alternative hypothesis): The data are **not** consistent with an expected result.

H_0 (null hypothesis):____

____

H_a (alternative hypothesis):____

____

Wild-Type *C. elegans* Chi Square Analysis

H_0 (null hypothesis): H_a (alternate hypothesis):			
	Observed, o	Expected, e	$(o - e)^2/e$
Percent *C. elegans* on NaCl side		0.91	
Percent *C. elegans* on control side		0.09	
		Σ, Sum	$\chi^2 =$ ______
Degrees of freedom: ______________		Significance Level = **0.05**	
Does χ^2 analysis reject the null hypothesis? Yes _____ No _____			

Mutant *C. elegans* Data Analysis

H_0 (null hypothesis): The data are consistent with an expected result.

H_a (alternative hypothesis): The data are **not** consistent with an expected result.

H_0 (null hypothesis):____

____

H_a (alternative hypothesis):____

____

Mutant *C. elegans* Chi Square Analysis

H_0 (null hypothesis): H_a (alternate hypothesis):			
	Observed, o	Expected, e	$(o - e)^2/e$
Percent *C. elegans* on NaCl side		0.91	
Percent *C. elegans* on control side		0.09	
		Σ, Sum	$\chi^2 =$ ______
Degrees of freedom: ____________		Significance Level = **0.05**	
Does χ^2 analysis reject the null hypothesis? Yes _____ No _____			

Bio-Rad
Laboratories, Inc.

Life Science
Group

Web site *www.bio-rad.com* ***USA*** *800 424 6723* ***Australia*** *61 2 9914 2800* ***Austria*** *01 877 89 01* ***Belgium*** *09 385 55 11* ***Brazil*** *55 11 5044 5699*
Canada *905 364 3435* ***China*** *86 21 6169 8500* ***Czech Republic*** *420 241 430 532* ***Denmark*** *44 52 10 00* ***Finland*** *09 804 22 00*
France *01 47 95 69 65* ***Germany*** *089 31 884 0* ***Greece*** *30 210 9532 220* ***Hong Kong*** *852 2789 3300* ***Hungary*** *36 1 459 6100* ***India*** *91 124 4029300*
Israel *03 963 6050* ***Italy*** *39 02 216091* ***Japan*** *03 6361 7000* ***Korea*** *82 2 3473 4460* ***Mexico*** *52 555 488 7670* ***The Netherlands*** *0318 540666*
New Zealand *64 9 415 2280* ***Norway*** *23 38 41 30* ***Poland*** *48 22 331 99 99* ***Portugal*** *351 21 472 7700* ***Russia*** *7 495 721 14 04*
Singapore *65 6415 3188* ***South Africa*** *27 861 246 723* ***Spain*** *34 91 590 5200* ***Sweden*** *08 555 12700* ***Switzerland*** *026 674 55 05*
Taiwan *886 2 2578 7189* ***Thailand*** *800 88 22 88* ***United Kingdom*** *020 8328 2000*

10040890 Rev A US/EG *Sig 1212*