Genetics Lab 4
Mendelian Genetics
Analysis of Data

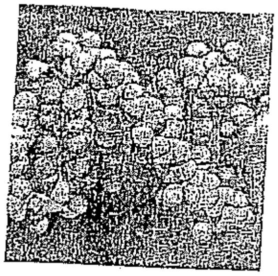
In our last lab exercise we collected raw data from F2 ears of corn from monohybrid and dihybrid crosses as well as from the offspring of a testcross. Many groups collected data that closely resembled the ratios that Mendel saw in his crosses involving pea plants, and we *assumed* that our observations could be explained according to Mendel's 4 postulates and accepted rules of inheritance. However, we were not expected to go beyond a "first-look" type of analysis.

Today we will test the "goodness of fit" of our data with a statistical test called the Chi-Square (X^2) test. One can apply this test when one has a theoretical or expected outcome to an experiment. The ultimate goal of the test is to make a reasonable statement as to whether any deviation from the expected outcome may be due to chance. (See detail introduction handout)

Before we return to our data from last week we will perform another lab exercise to show the general application of the X^2 test . *Then* we will use our own data collected from the F2 ears of corn to test it's fit with observations of Mendel.

<u>Exercise 1—BEAN COUNTERS</u>
Follow the instructions for Part 1A in your handout. Do the exercise in class and answer the questions that follow. **Include in your answers, <u>and in your lab report</u> discussion of this exercise the null hypothesis for this exercise and whether you would reject or fail to reject it. (The null hypothesis will be based on limited information about, and your initial observations of the beans)**

<u>Exercise 2, 3, 4</u>
Apply X^2 analysis to the data you collected last week from the monohybrid, dihybrid and test crosses. For each cross state the null hypothesis and decide whether to reject or fail to reject it.



INVESTIGATION 4

The Chi-Square Test

The purpose of the chi-square (χ^2) test is to determine whether experimentally obtained data constitute a good fit to, or a satisfactory approximation of, a theoretical, expected ratio; that is, the χ^2 test enables one to determine whether it is reasonable to attribute deviations from a perfect fit to chance. Obviously, if deviations are small, they can be more reasonably attributed to chance than if they are large. The question we try to answer with the χ^2 test is, "How small must the deviations be to be attributed to chance alone?" The formula for χ^2 is as follows:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

where O = the observed number of individuals in a particular phenotype, E = the expected number in that phenotype, and Σ = the summation of all possible values of $(O - E)^2/E$ for the various phenotypic categories.

Before you apply the χ^2 test to any of the data you have collected in Investigations 1, 2, and 3, use the following example for practice: In a cross of tall tomato plants to dwarf ones, the F_1 consisted entirely of tall plants and the F_2 consisted of 102 tall and 44 dwarf plants. Do these F_2 data fit a ratio of 3:1? To answer this question, χ^2 was calculated; the calculations are summarized in Table 4.1.

The calculated χ^2 value is 2.0548, but what does this χ^2 value mean? If the observed numbers (O) were exactly equal to the corresponding theoretically expected numbers (E), the fit would be perfect, and χ^2 would be zero. Thus, a small value for χ^2 indicates that the observed and expected ratios are in close agreement, whereas a large value indicates marked deviation from the expected ratio. Because chance deviations from the theoretical values are expected to occur, the question to be answered is, "Are the observed deviations within the limits expected by chance?" Generally, statisticians have agreed on the arbitrary limits of odds of 1 chance in 20 (probability = .05) for drawing the line between acceptance and rejection of the hypothesis as a satisfactory explanation of the data tested.

A χ^2 value of 3.841 for a two-term ratio corresponds to a probability of 1 in 20, or .05. One would obtain a χ^2 value of 3.841 resulting from chance alone in only about 5% of similar trials if the hypothesis is true. When χ^2 exceeds 3.841 for a two-term ratio, the probability that the deviations can be attributed to chance alone is less than 1 in 20. The hypothesis of compatibility between the observed

TABLE 4.1. Summary of the Calculation of χ^2 in a Hypothetical Illustration

Phenotype	Genotype	O	E	(O − E)	$(O - E)^2$	$(O - E)^2/E$
Tall	*T*–	102	109.5	−7.5	56.25	0.5137
Dwarf	*tt*	44	36.5	7.5	56.25	1.5411
Totals		146	146.0			$\chi^2 = 2.0548$

TABLE 4.2. Table of χ^2 Values*

df	*P* = .99	.95	.80	.50	.20	.05	.01
1	.000157	.00393	.0642	.455	1.642	3.841	6.635
2	.0201	.103	.446	1.386	3.219	5.991	9.210
3	.115	.352	1.005	2.366	4.642	7.815	11.345
4	.297	.711	1.649	3.357	5.989	9.488	13.277
5	.554	1.145	2.343	4.351	7.289	11.070	15.086
6	.872	1.635	3.070	5.348	8.558	12.592	16.812
7	1.239	2.167	3.822	6.346	9.803	14.067	18.475
8	1.646	2.733	4.594	7.344	11.030	15.507	20.090
9	2.088	3.325	5.380	8.343	12.242	16.919	21.666
10	2.558	3.940	6.179	9.342	13.442	18.307	23.209
15	5.229	7.261	10.307	14.339	19.311	24.996	30.578
20	8.260	10.851	14.578	19.337	25.038	31.410	37.566
25	11.524	14.611	18.940	24.337	30.675	37.652	44.314
30	14.953	18.493	23.364	29.336	36.250	43.773	50.892

* Selected data from R.A. Fisher and F. Yates, *Statistical tables for biological, agricultural and medical research* (London: Oliver and Boyd, 1943).

and expected ratios is thus rejected. In the practice example, $\chi^2 = 2.0548$, which is considerably less than 3.841, the maximum allowable value for a two-term ratio associated with the probability of 1 in 20. Therefore, you can attribute the deviations to chance alone and accept the hypothesis that the data fit a 3:1 ratio.

Where did the value of 3.841 come from? Statisticians have published extensive tables listing χ^2 values (see Table 4.2). Notice that across the top of Table 4.2 are probability (*P*) values; down the left side are "degrees of freedom" values (*df* = 1, 2 ... 30). The number of degrees of freedom is generally one less than the number of terms in the ratio. In the practice example, with two terms in the ratio (3:1), you have one degree of freedom in interpreting χ^2. Thus, in the one-degree-of-freedom row and under the .05 column you find the χ^2 value of 3.841, the maximum value of χ^2 that we are willing to accept and still attribute the deviations observed to chance alone.

In the hypothetical example, χ^2 was calculated to be 2.0548. Looking in the one-degree-of-freedom row, you see that this value of χ^2 falls between the .05 ($\chi^2 = 3.841$) and the .20 ($\chi^2 = 1.642$) probability columns. This means that the probability that the deviations observed may be attributed to chance alone is between 5% and 20%; that is, if you were to do this same experiment 100 times, you would expect to observe deviations as large as you obtained or larger in between 5 and 20 of the 100 trials owing to chance alone. With this background in the meaning, calculation, and interpretation of χ^2, you can apply the test to some data you have collected yourself.

OBJECTIVES OF THE INVESTIGATION

Upon completion of this investigation, the student should be able to

1. **calculate** χ^2 to determine whether a given set of data approximates a theoretically expected ratio and
2. **interpret** a calculated χ^2 value, given the appropriate number of degrees of freedom and a table of χ^2 values.

Materials needed for this investigation:

container of equal ~~quantities~~ weight of colored and white beans
calculator

I. APPLICATION OF THE CHI-SQUARE TEST

A. Application of Chi-Square to New Data

Colored and white beans were carefully selected, ~~for equality of size and uniformity of shape~~. *Equal* ~~quantities~~ of each color were thoroughly mixed and placed in a single container. 2

Remove a random sample (one petri dish cover or plastic cup full) of this mixture. Such a sample is shown in Figure 4.1. Segregate and count the beans of the different colors. Record your data in Table 4.3 and then calculate the expected numbers based on the size of the sample and the known ratio of colored to white beans in the entire population. Complete Table 4.3 and calculate χ^2.

TABLE 4.3. Calculation of χ^2 for a Sample Removed from a Large Population Consisting of Equal Numbers of Colored and White Beans

Classes (Phenotypes)	Observed (O)	Expected (E)	Deviations (O − E)	$(O - E)^2$	$(O - E)^2/E$
Colored	______	______	______	______	______
White	______	______	______	______	______
Totals	______	______	______	χ^2 =	______

FIGURE 4.1. Sample from a large container of equal quantities of colored and white beans. Petri dish lid that held the sample is shown adjacent to beans.

1. How many degrees of freedom do you have in the interpretation of the χ^2 value?

____

2. Using Table 4.2, determine what χ^2 values lie on either side of the χ^2 calculated in Table 4.3.

   Record these values here. ______________________________

3. What probability values are associated with these table values of χ^2?

____

4. In the space provided, write a brief interpretation of the χ^2 value you have just calculated.

____

____

____

____

5. Place the χ^2 value for your sample in the proper column (under the appropriate P value) of the table on the chalkboard that is prepared in the style of Table 4.4.

TABLE 4.4. χ^2 Values Obtained by Various Class Members for Individual Samples Removed from a Large Population Consisting of Equal Numbers of Colored and White Beans

Degrees of Freedom	P = .99	.95	.80	.50	.20	.05	.01

Note that this table is in the same form as Table 4.2, but space has been provided in the body of the table for recording the χ^2 values obtained by all members of the class. The actual distribution of the χ^2 values from all of the experiments conducted in the class can be observed in the completed table. Thus, although the population is known to be composed of equal numbers of colored and white beans, random samples drawn from it may vary considerably in how closely they approximate the theoretical 1:1 ratio. Nevertheless, in only about 5% of the samples drawn will the χ^2 values be expected to equal or exceed 3.841.[1]

B. Application of Chi-Square to Data from Investigation 1 — Corn DATA

Transfer the F_2 data obtained in Investigation 1 (see Table 1.2) to Table 4.5. Calculate χ^2 for the total of males and females based on the hypothesis that the classical monohybrid F_2 ratio has been obtained.

1. How many degrees of freedom do you have in the interpretation of this χ^2 value?

____

[1] Included in the *Instructor's manual* is a set of data from 21 students that demonstrates this principle.

Calculation of χ^2 on Data from Monohybrid cross (Color)

Phenotypes	O	E	O – E	$(O - E)^2$	$(O - E)^2/E$
Totals				$\chi^2 =$	

df =
p =

Calculation of χ^2 on Data from Monohybrid cross (texture)

Phenotypes	O	E	O – E	$(O - E)^2$	$(O - E)^2/E$
Totals				$\chi^2 =$	

df =
p =

Calculation of χ^2 on Data from Dihybrid cross

Phenotypes	O	E	O − E	$(O - E)^2$	$(O - E)^2/E$
Totals				$\chi^2 =$	

df =
P =

Calculation of χ^2 on Data from test cross

Phenotypes	O	E	O − E	$(O - E)^2$	$(O - E)^2/E$
Totals				$\chi^2 =$	

df =
P =