Writing the Rules of Heredity

In the mid 1800's, an Augustinian friar named Gregor Mendel formalized quantitative observations on heredity in the pea plant. He undertook hybridization experiments that utilized purebred or true breeding plants with specific qualities over many generations to observe the passage of these traits. Some of these physical traits included: seed shape, flower color, plant height and pod shape.

The pea plant (*Pisum sativum*) offered a great advantage of being able to control the fertilization process and having large quantities of offspring in a short period of time. In a simple experiment of tracking the passage of a single trait (monohybrid cross) like flower color through multiple generations he was able to formulate rules of heredity. In this case, pea plants either produced white flowers or purple flowers for many generations (true breeding purple flower or true breeding white flower). These true breeding plants are referred to as the Parental Generation (P). By removing the male parts of the pea flower (anthers containing pollen), Mendel was able to control for self-pollination. The hybridization came from applying the pollen from one true breeding plant to the female part (the pistil) of the opposite true breeding plant. The subsequent offspring are referred to as the First Filial Generation (F₁). In the first generation, all flowers are purple. Permitting self-pollination generates a Second Filial Generation (F₂). This generation sees the re-emergence of the white flowered plants in an approximate ratio of 3 purple flowered to 1 white flowered plants.

Male and female parts of flowers. Mendel removed the anthers containing pollen to prohibit self-pollination and selectively applied the pollen to stigmas in order to control the "hybridization". Credit: LadyofHats Mariana Ruiz [Public domain]
The loss of one variant on the trait in the F₁ plants with the re-emergence in the F₂ prompted Mendel to propose that each individual contained 2 hereditary particles where each offspring would inherit 1 of these particles from each parent. Furthermore, the loss of one of the variants in the F₁ was explained by one variant masking the other, as he explained as being dominant. The re-emergence of the masked variation, or recessive trait in the next generation was due to the both particles being of the masked variety. We now refer to these hereditary particles as genes and the variants of the traits as alleles.

### Mendel's Rules of Segregation and Dominance

The observations and conclusions that Mendel made from the monohybrid cross identified that inheritance of a single trait could be described as passage of genes (particles) from parents to offspring. Each individual normally contained two particles and these particles would separate during production of gametes. During sexual reproduction, each parent would contribute one of these particles to reconstitute offspring with 2 particles. In the modern language, we refer to the genetic make-up of the two "particles" (in this case, alleles) as the genotype and the physical manifestation of the traits as the phenotype. Therefore, Mendel's first rules of inheritance are as follows:

1. **Law of Segregation**
   - During gamete formation, the alleles for each gene segregate from each other so that each gamete carries only one allele for each gene

2. **Law of Dominance**
   - An organism with at least one dominant allele will have the phenotype of the dominant allele.
   - The recessive phenotype will only appear when the genotype contains 2 recessive alleles. This is referred to as homozygous recessive
   - The dominant phenotype will occur when the genotype contains either 2 dominant alleles (homozygous dominant) or on dominant and one recessive (heterozygous)
The Punnett Square is a tool devised to make predictions about the probability of traits observed in the offspring in the F$_2$ generation and illustrate the segregation during gamete formation.

The Single Trait Cross (Monohybrid Cross)

Monohybrid cross (one trait cross) observing the pod shape of peas

Monohybrid cross (on trait cross) observing the pod color of peas.
Corn Coloration in an F₂ Population (activity)

A corn cob contains hundreds of kernels. Each kernel is a seed that represents an individual organism. In the cob, we can easily see kernel color as a phenotype.

1. Retrieve an F₂ corn cob
2. Count a total of 100 kernels
   1. Tally the number of Yellow Kernels within that 100 (in the dried state, anything yellow or honey colored counts as yellow)
   2. Tally the number of Purple Kernels within that 100 (in the dried state, purple colored kernels may appear brown)
   3. Ignore any speckled kernels that may have yellow and purple within them
3. Compare numbers with the class as a whole
4. From the numbers:
   1. Is there a dominant color?
   2. Which is dominant, if there is?
   3. Create a Punnet square to illustrate the expected number of each color in a simple dominant:recessive paradigm.
The Two Trait Cross (Dihybrid Cross)

Mendel continued his experimentation where he looked at two traits. These two trait crosses are called **dihybrid crosses**. While the monohybrid cross would yield 3:1 ratios of the phenotypes, the dihybrid crosses would yield 9:3:3:1 ratios of all the combinations of each phenotype.

![Dihybrid Cross Diagram](diagram.png)

**Mendel's Rule of Independent Assortment**

The dihybrid cross revealed another law of inheritance to Mendel. By observing the 9:3:3:1 ratio, Mendel concluded that traits were not tied to each other. That is to say, if a pea pod was yellow, it could still be either smooth or wrinkled in texture. This lack of linkage between genes yielding different characteristics was dubbed the **Law of Independent Assortment**. Genes for different traits can segregate independently during the formation of gametes.
Kernel Coloration and Texture in an F<sub>2</sub> Population (activity)

1. Retrieve a dihybrid F<sub>2</sub> corn cob
2. Count a total of 200 kernels
   1. Tally the number of Yellow Kernels that are rounded and smooth in texture
   2. Tally the number of Yellow Kernels that are shriveled and wrinkly in texture (honey colored)
   3. Tally the number of Purple Kernels within that rounded and smooth in texture
   4. Tally the number of Purple Kernels within that are shriveled and wrinkly in texture
   5. Ignore any speckled kernels that may have yellow and purple within them
3. Compare numbers with the class as a whole
4. Each kernel constitutes an individual organism (a seed that can give rise to a whole new plant).
   From the numbers:
   1. Is there a dominant texture (smooth or shriveled)?
   2. Which is dominant, if there is?
   3. Is there a color that always pairs with a texture or do these characteristics <em>assort independently</em>?
   4. Create a Punnet square to illustrate the expected number of each color/texture combination in a simple dominant:recessive paradigm.
**Genetics leaves a bad taste in my mouth... or not**

Some of our personal preferences arise from the way we were brought up. Culture plays a role in our likes and dislikes. Likewise, our experiences play a role in how we respond to certain stimuli. Another major factor that plays a role into our preferences comes wired in our genome. The DNA in our cells is the instruction manual for who we are. We are programmed to seek out things of a nutritive values in order to acquire raw materials like carbohydrates, proteins and lipids. In our search for nutritive compounds we have learned to avoid things that don't taste good. Bitter things have a tendency to be associated with toxic compounds in nature. When eating a food item for the first time, molecules hit our tongue and stimulate multiple sensations: sweet, sour, salty, savory and bitter. Attributed to these multiple taste types are a diverse family of receptors that bind to the molecules that result in our perception of these sensations. Something bitter might make us learn to avoid this food item in the future.

One type of bitter receptor senses the presence of a chemical called phenylthiocarbamide (PTC). This chemical chemically resembles toxic compounds found in plants but is non-toxic. The ability to taste PTC is comes from the gene called TAS2R38. This gene encodes a protein that on our tongues that communicates the bitterness of this chemical. There are two common alleles of this gene with at least five more uncommon variants. Within the two common forms, a single nucleotide polymorphism (SNP) is responsible for changing one amino acid in the receptor. It's this difference of one amino acid that results in the ability of the receptor to either respond or not respond to PTC. We inherit one copy of the gene from our father and one copy from our mother. Based on how our parents gametes were formed and what alleles we received during the fertilization event determines how we respond to this chemical. Because we each have 2 copies of this gene, we can utilize simple Mendelian genetics to understand which allele is dominant or recessive.

1. Place a piece of “Control” paper on the tongue and indicate if there is a taste
2. Place a piece of “PTC” paper on the tongue and indicate if there is a taste and the taste severity
3. Fill out the table for the class to identify how many non-tasters, tasters or super-tasters there are.
4. Indicate if you believe the trait is dominant or recessive (ability to taste or not taste)
5. Assign a descriptor allele for the dominant (a capital letter) or the recessive (a lowercase letter) and draw a Punnet square for the F2 generation of 2 Heterozygous parents.
6. Compare the class tally of tasters and non-tasters in the class and discuss with your instructor if there is a clear dominance of this trait.

**Table: PTC Tasting Tally**

<table>
<thead>
<tr>
<th>Phenotypes</th>
<th>Number</th>
<th>% Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PTC Tasters</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Dominant or Recessive)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PTC Non-Tasters</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Dominant or Recessive)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Questions:**

1. How do you explain the presence of those who can't taste PTC, those who can taste it and those who really can't stand the taste of it?
2. This chemical is non-toxic and doesn't exist in nature. Do you think there is a selective pressure that confers an advantage to those who do taste it?
**Exercise: Coding Bitterness**

Prior to this exercise, review the [Central Dogma](http://openlab.citytech.cuny.edu/bio-oer/genetics/).

The full coding sequence of **TAS2R38** is 1,002 bases (334 amino acids) long. A segment of the gene is shown below where the SNP (in red) occurs. Variant 1 is the version of the gene that encodes for the ability to taste PTC. Variant 2 is the version of the gene that is unable to bind to PTC. This SNP mutation is called a **missense** mutation because it changes the amino acid. Some mutations cause the insertion of a premature stop codon. This **nonsense mutation** results in a truncated protein and can be disastrous to the function. We already know that the simple substitution of one nucleotide translates to a change in one amino acid and determines the ability to taste PTC. Imagine if a large group of amino acids from the protein was missing.

With template strand (“Complement”) information:

1. Write the sequence of the coding strand.
2. Write the sequence of the mRNA.
3. Use the Genetic Code Chart to translate the amino acid sequence.

**Variant 1**
- **Coding Strand**: 5'-
- **Complement**: 3'-TTC TCC GTC CGT GAC TCG-5'
- **mRNA**: 5'-
- **Amino Acid**:

**Variant 2**
- **Coding Strand**: 5'-
- **Complement**: 3'-TTC TCC GTC GGT GAC TCG-5'
- **mRNA**: 5'-
- **Amino Acid**:

---

Jeremy Seto jseto@citytech.cuny.edu

Credit: [Mouagip](http://开放资源/)[Public domain]
Sex-Linked Genes

For the most part, mammals have gender determined by the presence of the Y chromosome. This chromosome is gene poor and a specific area called sex determining region on Y (SRY) is responsible for the initiation of the male sex determination. The X-chromosome is rich in genes while the Y-chromosome is a gene desert. The presence of an X-chromosome is absolutely necessary to produce a viable life form and the default gender of mammals is traditionally female.

Chromosomal painting techniques can reveal the gender origin of mammalian cells. By using fluorescent marker sequences that can hybridize specifically to X or Y chromosomes through Fluorescence In Situ Hybridization (FISH), gender can be identified in cells.

The male cells have an X and a Y while the female cells have X and X combination.

Credit: Janice Y Ahn, Jeannie T Lee [CC BY 2.0]
The genes encoding photoreceptor proteins for the long wave-length (reds) and middle wave-lengths (greens) reside on the X chromosome at Xq25. Since the Y-chromosome is not homologous, any mutation to either of these genes that render them non-functional results in an inability to perceive either of those colors. Men are more susceptible to the condition of red-green colorblindness since they are hemizygous. This means that there is no corresponding gene that could complement a deficient red or green photoreceptor gene.

Dr. Shinobu Ishihara published his test for color perception in 1917 and this test is widely used to detect deficits in color perception. Below are examples of Ishihara plates. Record the number that you perceive in each plate and discuss with the rest of the class.

1. As you go through the plates above, note the number that you see (if any).
2. The genes for the Red and Green receptors are on the X-chromosome, who are most affected by mutation? Create a Punnet Square to illustrate how this works.
3. Can women be color-blind for red/green?
4. Humans have 3 color light receptors and have trichromatic vision. Some women are described as possibly having tetrachromatic vision (seeing 4 colors) and being able to discriminate colors invisible to the rest of us. Describe a mechanism for why this could happen. Why is there a possible gender bias?
The case of Queen Victoria

Hemophilia literally translates to blood loving. This is a description of a series of disorders where an individual has an inability to clot blood after a cut. In modern times, clotting factors may be administered to an afflicted individual, but a prior treatment involved blood transfusions. A very famous family had a genetic predisposition to hemophilia and due to the proliferative nature of this family, we have some statistical power to verify predictions on the probabilities of passing the disease state. Below is a partial pedigree for Victoria, Queen of the United Kingdom of Great Britain and Ireland and Empress of India. The filled in shapes represent individuals who suffered from hemophilia.

1. From the pedigree above, what can you say about this form of hemophilia with respect to dominance?
2. From this pedigree, can you comment on the probable chromosome where the deficiency occurs?
3. Assign genotypes for Prince Albert and Queen Victoria and perform a Punnet Square to illustrate if their offspring reflect your statements on dominance and chromosome location.
4. Albert and Victoria were 1st cousins. Do you believe this had anything to do with the propagation of this disease? What does your Punnet Square tell you?
5. Highlight the definitive carriers of the disease gene in the pedigree above.

Additional Resources
- Full case study can be acquired at the National Center for Case Study Teaching in Science.
- Human Factor IX mRNA sequence
X-inactivation
The mammalian X-chromosome contains significantly more genetic information than the Y-chromosome. This gene dosage is controlled for in females through a process called X-inactivation where one of the X-chromosomes is shut down and highly condensed into a Barr body. Inactivation of the X-chromosome occurs in a stochastic manner that results in females being cellular mosaics where a group of cells have inactivated the paternal X-chromosome and other patches of cells have inactivated the maternal X-chromosome. The most striking example of mosaicism is the calico cat. A calico cat (tortoise shell cat) is always a female. One of the genes that encodes coat color in cats resides on the X-chromosome and exist as either orange or black alleles. Due to the stochastic inactivation, the patterning of orange and black fur is a distinctive quality of calicos.

While the genetic information for the the orange or black coat color exists in all cells, they are not equally expressed. This type of heritable trait in spite of the presence of the genetic material (DNA) is called epigenetic to imply that it is "above" (epi) genetics.
Drosophila: Thomas Hunt Morgan

Around 1908, Thomas Hunt Morgan began to explore the genetics of what was to become a model organism, *Drosophila melanogaster* (Fruit fly). This small organism had a relatively short life cycle, great fecundity and was easily managed. From these flies that normally have red eye coloring, he and his students found white-eyed mutants. The lab noted that white-eyed flies were almost exclusively male. This gender imbalance lead Morgan to believe that the trait was sex-linked. In 1911, Morgan published a paper that described the inheritance patterns of 5 eye-colors in *Drosophila* (Morgan, 1911).

While DNA was not yet known as the source of genetic information, Morgan's studies revealed that the location of genes most likely resided on the chromosomes. By cataloging many mutations in the lab, he was able to construct a map of gene locations. His 1922 paper specifically stated that some traits were sex-linked and therefore residing on the sex chromosome. When performing crosses of white-eyed males to wild-type females, he continued to find white-eyed trait only in males. However, in the subsequent cross of females from that generation with white-eyed males, the presence of white-eyed males and females were revealed. This indicated that the white-eyed trait was recessive and resided on the X chromosome.

Morgan received the Nobel Prize in Physiology or Medicine in 1933 for his inference of chromosomes being a physical mechanism for packaging genetic information in the cells.

- **Morgan** TH. *The Origin of Five Mutations in Eye Color in Drosophila and Their Modes of Inheritance*. *Science*. 1911 Apr 7;33(849):534-7
Non-Mendelian Genetics

Co-Dominance and multiple alleles

Co-dominance is said to occur when there is an expression of two dominant alleles. The prototypical case for this is the human ABO blood grouping.

<table>
<thead>
<tr>
<th>Red blood cell type</th>
<th>Group A</th>
<th>Group B</th>
<th>Group AB</th>
<th>Group O</th>
</tr>
</thead>
<tbody>
<tr>
<td>A antigen</td>
<td>A</td>
<td>B</td>
<td>AB</td>
<td>O</td>
</tr>
<tr>
<td>Antibodies in Plasma</td>
<td>Anti-B</td>
<td>Anti-A</td>
<td>None</td>
<td>Anti-A and Anti-B</td>
</tr>
<tr>
<td>Antigens in Red Blood Cell</td>
<td>A antigen</td>
<td>B antigen</td>
<td>A and B antigens</td>
<td>None</td>
</tr>
</tbody>
</table>

Three alleles exist in the ABO system: A, B and O. This results in four blood types: A, B, O and the blended AB.
Incomplete Dominance

During Mendel's time, people believed in a concept of blending inheritance whereby offspring demonstrated intermediate phenotypes between those of the parental generation. This was refuted by Mendel's pea experiments that illustrated a Law of Dominance. Despite this, non-Mendelian inheritance can be observed in sex-linkage and co-dominance where the expected ratios of phenotypes are not observed clearly. **Incomplete dominance** superficially resembles the idea of blending inheritance, but can still be explained using Mendel's laws with modification. In this case, alleles do not exert full dominance and the offspring resemble a mixture of the two phenotypes.

The most obvious case of a two allele system that exhibits incomplete dominance is in the snapdragon flower. The alleles that give rise to flower coloration (Red or White) both express and the heterozygous genotype yields pink flowers. There are different ways to denote this. In this case, the superscripts of R or W refer to the red or white alleles, respectively. Since no clear dominance is in effect, using a shared letter to denote the common trait with the superscripts (or subscripts) permit for a clearer denotation of the ultimate genotype to phenotype translations.

Problem: Incomplete Dominance

If pink flowers arose from blending inheritance, then subsequent crosses of pink flowers with either parental strain would continue to dilute the phenotype. Using a Punnet Square, perform a test cross between a heterozygous plant and a parental to predict the phenotypes of the offspring.
Epistasis and Modifier Genes

Interplay of multiple enzymes in a biochemical pathway will alter the phenotype. Some genes will modify the actions of another gene.

Epistasis refers to the event where a gene at one locus is dependent on the expression of a gene at another genomic locus. Stated another way, one genetic locus acts as a modifier to another. This can be visualized easily in the case of labrador retriever coloration where three primary coat coloration schemes exist: black lab, chocolate lab and yellow lab.

Genes do not exist in isolation and the gene products often interact in some way. Epistasis refers to the event where a gene at one locus is dependent on the expression of a gene at another genomic locus. Stated another way, one genetic locus acts as a modifier to another. This can be visualized easily in the case of labrador retriever coloration where three primary coat coloration schemes exist: black lab, chocolate lab and yellow lab.

Chocolate lab (top), Black lab (middle), Yellow lab (bottom) coat colorations arise from the interaction of 2 gene loci, each with 2 alleles.

Credit: Erikeltic [CC-BY-SA 3.0]

Black lab (BB or Bb) and Chocolate lab (bb)
Credit: deposliff [CC BY-SA 3.0]

Black lab (EE or Ee) and Yellow lab (ee)
Credit: [Public Domain]
Two genes are involved in the coloration of labradors. The first is a gene for a protein called TYRP1, which is localized to the melanosomes (pigment storing organelles). Three mutant alleles of this gene have been identified that reduce the function of the protein and yield lighter coloration. These three alleles can be noted as "b" while the functioning allele is called "B". A heterozygous (Bb) or a homozygous dominant individual will be black coated while a homozygous recessive (bb) individual will be brown.

The second gene is tied to the gene for Melanocortin 1 Receptor (MC1R) and influences if the eumelanin pigment is expressed in the fur. This gene has the alleles denoted "E" or "e". A yellow labrador will have a genotype of either Bbee or bbee.

The interplay between these genes can be described by the following diagram:
Mendel's Observations

<table>
<thead>
<tr>
<th>Parental Cross</th>
<th>F₁ Phenotypes</th>
<th>F₂ Phenotypes</th>
<th>F₂ Ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td>Round x Wrinkled Seeds</td>
<td>All Round</td>
<td>5474 Round: 1850 Wrinkled</td>
<td>2.96:1</td>
</tr>
<tr>
<td>Yellow x Green Seeds</td>
<td>All Yellow</td>
<td>6022 Yellow: 2001 Green</td>
<td>3.01:1</td>
</tr>
<tr>
<td>Purple x White Flowers</td>
<td>All Purple</td>
<td>705 Purple: 224 White</td>
<td>3.15:1</td>
</tr>
<tr>
<td>Tall x Dwarf Plants</td>
<td>All Tall</td>
<td>787 Tall: 227 Dwarf</td>
<td>2.84:1</td>
</tr>
</tbody>
</table>

Probability: Past Punnett Squares

Punnett Squares are convenient for predicting the outcome of monohybrid or dihybrid crosses. The expectation of two heterozygous parents is 3:1 in a single trait cross or 9:3:3:1 in a two-trait cross. Performing a three or four trait cross becomes very messy. In these instances, it is better to follow the rules of probability. **Probability** is the chance that an event will occur as a fraction or percentage. In the case of a monohybrid cross, 3:1 ratio means there is a 3/4 (0.75) chance of the dominant phenotype with a 1/4 (0.25) chance of a recessive phenotype.

\[
\frac{1}{6} \times \frac{1}{6} = \frac{1}{36}
\]

1/6 x 1/6 = 1/36

1/6  x  1/6  =  1/36 chance of both being three
A single die has a 1 in 6 chance of being a specific value. In this case, there is a 1/6 probability of rolling a 3. It is understood that rolling a second die simultaneously is not influenced by the first and is therefore independent. This second die also has a 1/6 chance of being a 3.

We can understand these rules of probability by applying them to the dihybrid cross and realizing we come to the same outcome as the 2 monohybrid Punnett Squares as with the single dihybrid Punnett Square.

This forked line method of calculating probability of offspring with various genotypes and phenotypes can be scaled and applied to more characteristics.
The Chi-Square Test

The $\chi^2$ statistic is used in genetics to illustrate if there are deviations from the expected outcomes of the alleles in a population. The general assumption of any statistical test is that there are no significant deviations between the measured results and the predicted ones. This lack of deviation is called the null hypothesis ($H_0$). $\chi^2$ statistic uses a distribution table to compare results against at varying levels of probabilities or critical values. If the $\chi^2$ value is greater than the value at a specific probability, then the null hypothesis has been rejected and a significant deviation from predicted values was observed. Using Mendel’s laws, we can count phenotypes after a cross to compare against those predicted by probabilities (or a Punnett Square).

<table>
<thead>
<tr>
<th>Degrees of Freedom (DF)</th>
<th>p-Value for $\chi^2$ test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.1</td>
</tr>
<tr>
<td>1</td>
<td>2.706</td>
</tr>
<tr>
<td>2</td>
<td>4.605</td>
</tr>
<tr>
<td>3</td>
<td>6.251</td>
</tr>
<tr>
<td>4</td>
<td>7.779</td>
</tr>
<tr>
<td>5</td>
<td>9.236</td>
</tr>
<tr>
<td>6</td>
<td>10.65</td>
</tr>
<tr>
<td>7</td>
<td>12.02</td>
</tr>
<tr>
<td>8</td>
<td>13.36</td>
</tr>
<tr>
<td>9</td>
<td>14.68</td>
</tr>
<tr>
<td>10</td>
<td>15.99</td>
</tr>
</tbody>
</table>

In order to use the table, one must determine the stringency of the test. The lower the p-value, the more stringent the statistics. Degrees of Freedom (DF) are also calculated to determine which value on the table to use. Degrees of Freedom are the number of classes or categories there are in the observations minus 1. DF=n-1

In the example of corn kernel color and texture, there are 4 classes: Purple & Smooth, Purple & Wrinkled, Yellow & Smooth, Yellow & Wrinkled. Therefore, DF = 4 - 1 = 3 and choosing p < 0.05 to be the threshold for significance (rejection of the null hypothesis), the $\chi^2$ must be greater than 7.82 in order to be significantly deviating from what is expected. With this dihybrid cross example, we expect a ratio of 9:3:3:1 in phenotypes where 1/16th of the population are recessive for both texture and color while 9/16th of the population display both color and texture as the dominant. 3/16th will be dominant for one phenotype while recessive for the other and the remaining 3/16th will be the opposite combination.
With this in mind, we can predict or have expected outcomes using these ratios. Taking a total count of 200 events in a population, 9/16(200)=112.5 and so forth. Formally, the $\chi^2$ value is generated by summing all combinations of:

\[(\text{observed-expected})^2/\text{expected}\]

**Chi-Square Test: Is this coin fair or weighted? (activity)**

1. Everyone in the class should flip a coin 2x and record the result (assumes class is 24)
2. Fair coins are expected to land 50% heads and 50% tails
   - 50% of 48 results should be 24
   - 24 heads and 24 tails is already written in the "Expected" column
3. As a class, compile the results in the "Observed" column (total of 48 coin flips)
4. In the last column, subtract the expected heads from the observed heads and square it, then divide by the number of expected heads
5. In the last column, subtract the expected tails from the observed tails and square it, then divide by the number of expected tails
6. Add the values together from the last column to generate the $\chi^2$ value
7. Compare the value with the value at 0.05 with DF=1
   - there are 2 classes or categories (head or tail), so DF = 2 - 1 = 1
   - Were the coin flips fair (not significantly deviating from 50:50)?

<table>
<thead>
<tr>
<th>Heads or Tails</th>
<th>Expected Number</th>
<th>Observed Number</th>
<th>$\frac{(\text{Observed-Expected})^2}{\text{Expected}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heads</td>
<td>24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tails</td>
<td>24</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[\chi^2 = \sum \frac{(\text{Observed-Expected})^2}{\text{Expected}}\]

Let's say that the coin tosses yielded 26 Heads and 22 Tails. Can we assume that the coin was unfair? If we toss a coin an odd number of times (eg. 51), then we would expect that the results would yield 25.5 (50%) Heads and 25.5 (50%) Tails. But this isn't a possibility. This is when the $\chi^2$ test is important as it delineates whether 26:25 or 30:21 etc. are within the probability for a fair coin.
Chi-Square Test of Kernel Coloration and Texture in an F$_2$ Population (activity)

1. From the counts, one can assume which phenotypes are dominant and recessive
2. Fill in the "Observed" category with the appropriate counts
3. Fill in the "Expected Ratio" with either 9/16, 3/16 or 1/16
4. The total number of counted event was 200, so multiply the "Expected Ratio" x 200 to generate the "Expected Number" fields
5. Calculate the $\frac{(\text{Observed}-\text{Expected})^2}{\text{Expected}}$ for each phenotype combination
6. Add all $\frac{(\text{Observed}-\text{Expected})^2}{\text{Expected}}$ values together to generate the $\chi^2$ value and compare with the value on the table where DF=3
7. Do we reject the Null Hypothesis or were the observed numbers as we expected as roughly 9:3:3:1?
   - What would it mean if the Null Hypothesis was rejected? Can you explain a case in which we have observed values that are significantly altered from what is expected?

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Expected Ratio (9/16, 3/16, 1/16)</th>
<th>Expected Number (Total # x Ratio)</th>
<th>Observed Number</th>
<th>$\frac{(\text{Observed}-\text{Expected})^2}{\text{Expected}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purple &amp; Smooth</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Purple &amp; Wrinkly</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yellow &amp; Smooth</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yellow &amp; Wrinkly</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ \chi^2 = \sum \frac{(\text{Observed}-\text{Expected})^2}{\text{Expected}} \]
Hardy-Weinberg Principle
The Hardy-Weinberg principle is a mathematical model used to describe the equilibrium of two alleles in a population in the absence of evolutionary forces. This model was derived independently by G.H. Hardy and Wilhelm Weinberg. It states that the allele and genotype frequencies across a population will remain constant across generations in the absence of evolutionary forces. This equilibrium makes several assumptions in order to be true:
1. An infinitely large population size
2. The organism involved is diploid
3. The organism only reproduces sexually
4. There are no overlapping generations
5. Mating is random
6. Allele frequencies equal in both genders
7. Absence of migration, mutation or selection

As we can see, many items in the list above can not be controlled for but it allows for us to make a comparison in situations where expected evolutionary forces come into play (selection etc.).

Hardy-Weinberg Equilibrium
The alleles in the equation are defined as the following:
- Genotype frequency is calculated by the following:
  \[
genotype\ frequency = \frac{\text{# individuals of given genotype}}{\text{total # individuals in population}}
\]
- Allele frequency is calculated by the following:
  \[
  \text{allele frequency} = \frac{\text{# of copies of an allele in a population}}{\text{total # of alleles in population}}
  \]
- In a two allele system with dominant/recessive, we designate the frequency of one as \( p \) and the other as \( q \) and standardize to:
  \[p = \text{Dominant allele frequency}\]
  \[q = \text{recessive allele frequency}\]
- Therefore the total frequency of all alleles in this system equal 100% (or 1)
  \[p + q = 1\]
- Likewise, the total frequency of all genotypes is expressed by the following quadratic where it also equals 1:
  \[p^2 + 2pq + q^2 = 1\]
- This equation is the Hardy-Weinberg theorem that states that there are no evolutionary forces at play that are altering the gene frequencies.
Calculating Hardy-Weinberg Equilibrium (activity)

This exercise refers to the PTC tasting exercise. One can test for selection for one allele within the population using this example. Though the class size is small, pooling results from multiple section can enhance the exercise. Remember to surmise the dominant/recessive traits from the class counts.

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Genotype</th>
<th>Genotype Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Super Taster</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taster</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-taster</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. What is the recessive phenotype and how can we represent the genotype?
2. What is the dominant phenotype and how can we represent the genotypes?
3. What is the frequency of recessive genotype? \( q^2 \)
4. What is the frequency of the recessive allele? \( q \)
5. What is the frequency of the dominant allele? \( p = 1 - q \)
6. Use Hardy-Weinberg to calculate the frequency of heterozygotes in the class. \((2pq)\)
7. Use Hardy-Weinberg to calculate the frequency of homozygotes in the class. \( (p^2) \)
8. Using an aggregate of multiple section, compare the local allelic and genotypic frequencies with what the Hardy-Weinberg would predict.
9. With this small number in mind, we can see that there are problems with the assumptions required for this principle. The instructor will perform the following simulation in class to illustrate the effects on multiple populations with the effects of selection and/or population limitations. A coefficient of fitness can be applied to illustrate a selective pressure against an allele.
   - Population Genetics Simulation of Alleles

10. In the case of a selective pressure, a fitness coefficient \( (w) \) can be introduced. A research article http://www.jci.org/articles/view/64240 has shown that the Tas2R38 receptor aids in the immune response against Pseudomonas. Imagine a situation where there is an epidemic of antibiotic resistant Pseudomonas. This would show that the dominant allele will have a selective advantage.
   - Modify the fitness coefficient in the Population Genetics Simulator and describe the effects this would have over many successive generations.

Jeremy Seto  jseto@citytech.cuny.edu