Skill Drills: Chi-Squared

Chi-Square

$$\chi^2 = \sum \frac{(o-e)^2}{e}$$

o = observed results

e = expected results

Degrees of freedom are equal to the number of distinct possible outcomes minus one.

Chi-Square Table

p value	Degrees of Freedom							
	1	2	3	4	5	6	7	8
0.05	3.84	5.99	7.82	9.49	11.07	12.59	14.07	15.51
0.01	6.64	9.21	11.34	13.28	15.09	16.81	18.48	20.09

Steps to solve Chi -Squared Problems:

1. List out your observed results
2. Calculate your expected results
    a. Equal expectations? (divide by the number of outcomes)
    b. Unequal expectations? (multiply by the proportions)
3. Subtract the expected from the observed
4. Square the number from step 3
5. Divide the number from step 4 by the expected result
6. Repeat steps 3-5 for all data points/outcomes
7. Add up all the values **(this is your chi-squared value)**
8. Calculate the degrees of freedom and find your critical value using the table
9. Compare your chi-squared value to the critical value
    a. If chi-squared is > crit value then you should reject the null hypothesis
    b. If chi-squared is < crit value then you should fail to reject the null hypothesis

obs #	exp #	o-e	$(o\text{-}e)^2$	$(o\text{-}e)^2/e$
18	26.8	-8.8	78.0	2.9
48	39.2	8.8	78.0	2.0
32	23.2	8.8	78.0	3.4
25	33.8	-8.8	78.0	2.3
Sum of column 5 = X^2 =				10.6

Practice Problems:

1. AP Biology students collected data while studying the common isopod (“Roly Poly”). They created three chambers for the insect to explore: no cover (their control), wet paper towel, leaf litter (old leaves from outside). They used 12 isopods.Results (after 10 minutes of observations): 2 in the control, 4 under the paper towel and 6under the leaf litter.Was this statistically significant? Did the roly poly prefer the leaf litter or was this result due just to chance? Apply your chi square knowledge to find out!

   # of Outcomes: ___________________ degrees of freedom (df):_________________

   Null Hypothesis(H_O): There is no difference between the control and the other habitats as the isopods move between different habitats. (expect only 4 bugs in each habitat….ie EQUAL)

   Alternate Hypothesis (H_A): There is a difference between the control and other habitats as the isopods move between different habitats.

Location	Observed (o)	Expected (e)	(o-e)	$(o-e)^2$	$(o-e)^2/e$
Control					
Wet Paper Towel					
Leaf Litter					

chi square =__________ critical value =__________

Conclusion:

2. Two students conducted the photosynthesis floating disk assay with spinach leaves. The students wanted to see if there was a difference between having carbon dioxide (baking soda dissolved in water) or not in the cups.
This is what they tested:
1)Control: 1 tablespoon of baking soda
2) ½ tablespoon baking soda
3)1 teaspoon baking soda
4)½ teaspoon baking soda
5)Just a sprinkle of baking soda
6)No baking soda

They expected that the control would have the best results and that all 10 disks would float in the 15 minute time frame. They expected 5 disks would float in cup #2, cup #3 would have 4 disks, 3 disks would float in cup 4 and cup 5 would have 2 disks floating. They expected none to float in the “no baking soda” cup.

Results: 10 disks floated in cups #1 and 2. 9 disks floated in cup #3. 4 floated in cup 4. 1 floated in #5 and none floated in cup #6.

Did this student get what they expected? Apply your chi square knowledge to find out!

Location	Observed (o)	Expected (e)	(o-e)	$(o-e)^2$	$(o-e)^2/e$
Control					

chi square =__________ critical value =__________

Conclusion:

3. A student performed a cross between a brown toad and a green toad. When born, 146 toads that were produced from the cross matured into 87 toads with brown skin and 59 toads with green skin.
   a. Perform a chi-square test for the null hypothesis that both parents are heterozygous (consider how many of the 146 toads born would be brown vs. green based on the results of a Punnett square if both parents are heterozygous) for the skin-color gene.
   b. Explain whether your hypothesis is supported by the chi-square test and justify your explanation.

Skill Drills: Standard Deviation & SEM

Standard Deviatic Standard Error of the Mean*

$$S = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

$$SE_{\bar{x}} = \frac{s}{\sqrt{n}}$$

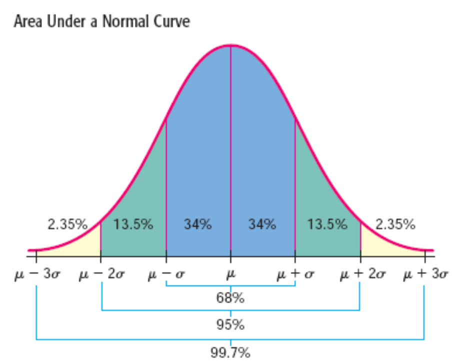
$\bar{x}$ = sample mean

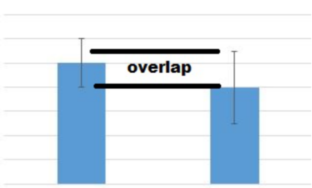
n = size of the sample

s = sample standard deviation (i.e., the sample-based estimate of the standard deviation of the population)

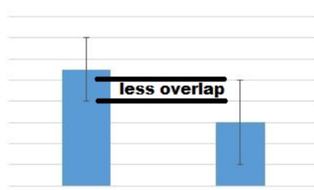
Why are these used?:

- Standard Deviation
  - A better measure of variability
  - Tells how much your data spreads from the mean
- Standard Error of the Mean (SEM)
  - Use SEM bars on graphs to show variability from the population mean
  - Helps us make an inference about the validity of our data
  - Shows how well the mean of that sample matches the entire population's mean

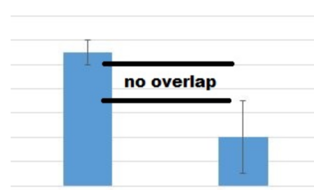




When standard deviation error bars overlap quite a bit, it's a clue that the difference is not statistically significant.



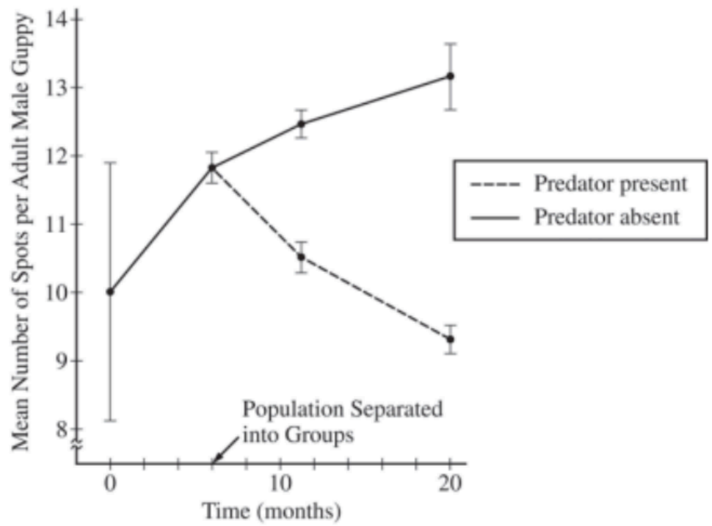
When standard deviation error bars overlap even less, it's a clue that the difference is probably not statistically significant.



When standard deviation error bars do not overlap, it's a clue that the difference may be significant

Practice Problems:

1. Adult male guppies (*Poecilia reticulata*) exhibit genetically determined spots, while juvenile and adult female guppies lack spots. In a study of selection, male and female guppies from genetically diverse populations were collected from different mountain streams and placed together in an isolated environment containing no predators. The study population was maintained for several generations in the isolated area before being separated into two groups. One group was moved to an artificial pond containing a fish predator, while a second group was moved to an artificial pond containing no predators. The two groups went through several generations in their new environments. At different times during the experiment, the mean number of spots per adult male guppy was determined as shown in the figure below. Vertical bars in the figure represent two standard errors of the mean (SEM).



   a. Identify the independent variable in this experiment.

   b. Identify the dependent variable in this experiment.

   c. Propose a title for this graph..

   d. Identify the SEM at time 0 months.

   e. Identify the SEM at time 6 months.

   f. Notice the change in the length of SEM bars between time 0 and time 6. Explain what the change in error bars between these times signifies with respect to the mean number of spots in the adult guppy population at the beginning of the investigation and after 6 months.

2. **In a second trial, the data listed in the table below was collected.**
    a. Complete the table below

Time (months)	Guppy 1	Guppy 2	Guppy 3	Guppy 4	Mean	Standard Deviation	1 SEM (68% confidence)	2 SEM (95% confidence)
0	9	7	12	13				
6	11	12	11	12				
20	13	15	12	11				

    b. A male guppy at time 6 months has 15 spots. Justify if this guppy falls into the 95% confidence range for mean number of spots per adult guppy.

3. The unicellular green alga Chlorella is often used as a model organism to study the effects of various substances on the growth of photosynthetic organisms. Researchers studying the detoxification of certain metals by Chlorella first collected data on the growth dynamics of the alga. A small number of Chlorella were added to 1,500 mL of culture medium that contained all of the inorganic nutrients needed for growth. Every five days for 30 days, the researchers performed multiple counts to determine the concentration of Chlorella in the culture (Table 1).

Table 1. The growth dynamics of a culture of *Chlorella* over a period of 30 days

Days Since Inoculation	Concentration of *Chlorella* ($\times 10^6$ per mL)
0	0.01 (0.005)
5	0.08 (0.01)
10	0.55 (0.1)
15	1.9 (0.2)
20	2.8 (0.4)
25	3.2 (0.25)
30	3.2 (0.3)

$\pm 2SE_{\bar{x}}$ values are shown in parentheses.

Using the template, construct an appropriately labeled graph to represent the data in Table 1. Based on the data, determine whether the concentration of Chlorella on day 20 is statistically different from the concentration on day 15 and from the concentration on day 25.

Skill Drills: Water Potential

Water Potential (Ψ)

$\Psi = \Psi_P + \Psi_S$

Ψ_P = pressure potential

Ψ_S = solute potential

The water potential will be equal to the solute potential of a solution in an open container because the pressure potential of the solution in an open container is zero.

The Solute Potential of a Solution

$\Psi_S = -iCRT$

i = ionization constant (this is 1.0 for sucrose because sucrose does not ionize in water)

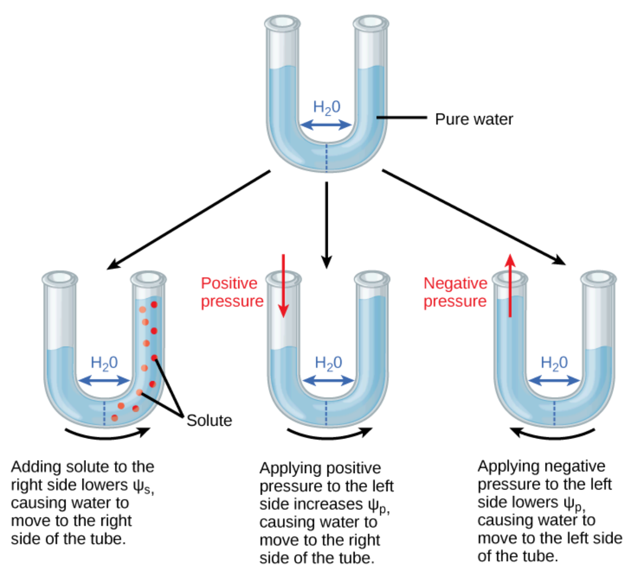
C = molar concentration

R = pressure constant (R = 0.0831 liter bars/mole K)

T = temperature in Kelvin (ºC + 273)

What is it used for?

- Determine where water will move in a system when looking at osmosis
    - Water travels from HIGH POTENTIAL to LOW POTENTIAL
- Water potential is impacted by solute concentration and pressure of the system
    - In pure water, $\Psi_s = 0$
    - In an open beaker, $\Psi_P = 0$
- If you add more solute you will be lowering the water potential



Practice Problems:

1. If a plant cell's Ψ_P = 2 bars and its Ψ_S = -3.5 bars, what is the resulting Ψ?

2. The plant cell from question #1 is placed in a beaker of sugar water with Ψ_S = -4.0 bars. In which direction will the net flow of water be?

3. The original cell from question # 1 is placed in a beaker of sugar water with Ψ_S = -0.15 MPa (megapascals). We know that 1 MPa = 10 bars. In which direction will the net flow of water be?

4. The value for Ψ in root tissue was found to be -3.3 bars. If you place the root tissue in a 0.1 M solution of sucrose at 20°C in an open beaker, what is the Ψ of the solution, and in which direction would the net flow of water be?

5. NaCl dissociates into 2 particles in water: Na^+ and Cl^-. If the solution in question 4 contained 0.1 M NaCl instead of 0.1 M sucrose, what is the Ψ of the solution, and in which direction would the net flow of water be?

6. At 20°C, a plant cell containing 0.6 M glucose is in equilibrium with its surrounding solution containing 0.5 M glucose in an open container. What is the cell's Ψ_P?

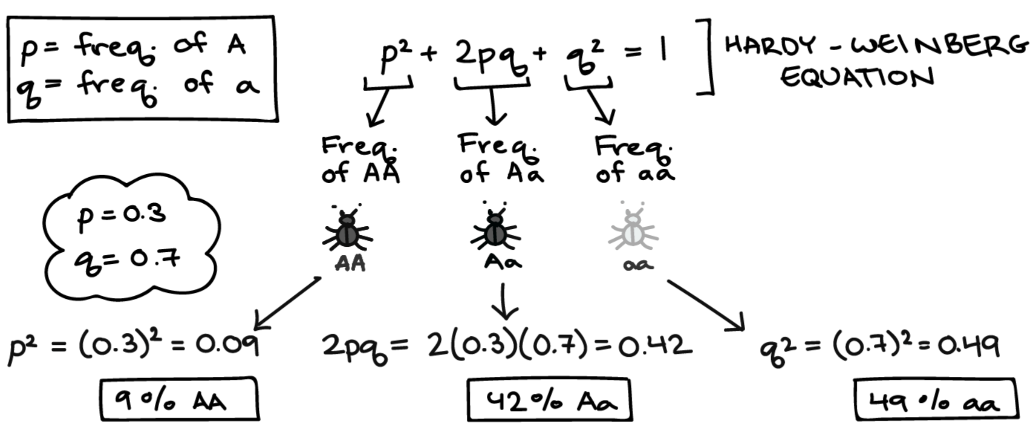
Skill Drills: Hardy-Weinberg Equilibrium

Hardy-Weinberg Equations

$p^2 + 2pq + q^2 = 1$	p = frequency of the dominant allele in a population
$p + q = 1$	q = frequency of the recessive allele in a population

What is it used for?

- States that genetic variation in a population will remain constant from generation to generation
- A population (or an allele) is not evolving if it is in Hardy-Weinberg Equilibrium
- Conditions that have to be met for a population to be in Hardy-Weinberg Equilibrium:
  - No genetic drift (bottleneck or founder effect)
  - No gene flow (immigration or emigration)
  - No mutations
  - No nonrandom mating (no selection of mates)
  - No natural selection



Practice Problems:

1. You have sampled a population in which you know that the percentage of the homozygous recessive genotype (aa) is 36%. Using that 36%, calculate the following:
    a. The frequency of the "aa" genotype (q^2)

    b. The frequency of the "a" allele (q).

    c. The frequency of the "A" allele (p).

    d. The frequencies of the genotypes "AA" (p^2) and "Aa" (2pq)

2. Sickle-cell anemia is a genetic disease. You can be SS (normal), Ss (heterozygous), or ss (sickle cell anemia). Individuals with the homozygous recessive genotype (ss, about 9%) often die prematurely due to the disease. Use the Hardy-Weinberg equations to calculate the following:
    a. The frequency of the recessive allele in the population.

    b. The frequency of the dominant allele in the population.

    c. The frequency of homozygous dominant individuals in the population.

    d. The frequency of heterozygous individuals in the population.

3. A farmer purchased 600 sheep. 150 sheep ended up having an economically undesirable feature: crinkly-hair, caused by the recessive allele cr.
    a. What is the frequency of this undesirable allele in the herd?

    b. What percent of the herd is likely to be heterozygous?

4. Sixty flowering plants are planted in a flowerbed. Forty of the plants are red-flowering homozygous dominant (AA). Twenty of the plants are white-flowering homozygous recessive (aa).
   a. Find q and p:

   b. What are the genotype frequencies for the original generation?
      p^2= 2pq= q^2=

   c. The plants naturally pollinate and reseed themselves for several years. In a subsequent year, 178 red-flowered plants, 190 pink-flowered plants, and 52 white-flowered plants were found in the flowerbed. Use a chi-squared analysis to determine if the population is in Hardy-Weinberg equilibrium. Use the frequencies calculated earlier to find your "expected" values.

      HINTS: Remember these things!!!
      1) You must compare equal populations with chi squared
      2) Hardy-Weinberg is a null hypothesis. If a population is not evolving, the frequencies stay the same in every generation. So your null in this case is that the population is in Hardy-Weinberg equilibrium.

Show your work here:

	Observed	Expected
p^2	178	
2pq	190	
q^2	52	
Total	420	420

P-value = 0.05 Degrees of freedom = 2 Critical value = 5.99 Chi-squared value =

   d. What should you do with the null hypothesis in this case?

   e. Is the population in HWE?

   f. Is the population evolving?

Skill Drills: Population Growth

Rate and Growth

Rate

$$\frac{dY}{dt}$$

Population Growth

$$\frac{dN}{dt} = B - D$$

Exponential Growth

$$\frac{dN}{dt} = r_{max}N$$

Logistic Growth

$$\frac{dN}{dt} = r_{max}N\left(\frac{K-N}{K}\right)$$

dY = amount of change

dt = change in time

B = birth rate

D = death rate

N = population size

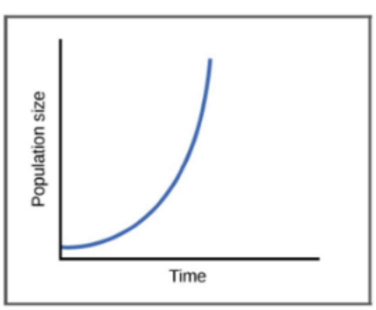
K = carrying capacity

r_{max} = maximum per capita growth rate of population

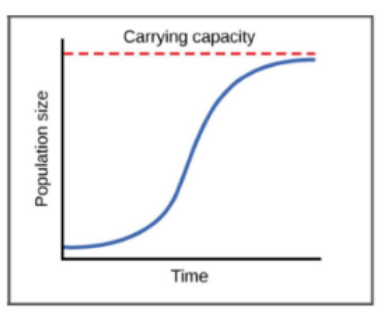
What is it used for?

- There are two different models of population growth; exponential and logistic
    - Exponential growth model: a population living under ideal conditions (ie easy access to food, abundant food, free to reproduce, etc)
        - Living in an ideal situation
        - Population grows rapidly
    - Logistic growth model: the per capita rate of increase approaches zero as the population size nears its carrying capacity
        - There are actually limits to the population size
        - The density of individuals exceeds the system's resource availability
        - Adds in carrying capacity

Exponential Growth



Logistic Growth



Practice Problems:

1. Here is data for a given population:
   Population Size = 50
   Births = 10
   Deaths = 4
   a. What is the growth rate for this population?

   b. Using this growth rate, how many individuals will be added to the population after one generation?

2. A population is in logistic growth. The population size is 50, the growth rate is the same from the previous question, and carrying capacity is 400.
   a. How many individuals are added to the population after one generation?

   b. What will the total population size be in three generations?

3. The following population, C, has no limits on food resources or space:
   Population size = 500
   Births = 240
   Deaths = 170
   a. Calculate the r for this population if it is experiencing exponential growth.

   b. How many individuals will be in the population at the start of the second generation?

   c. How many individuals will be in the population at the start of the third generation?

Skill Drills: Simpson's Diversity Index

Simpson's Diversity Index

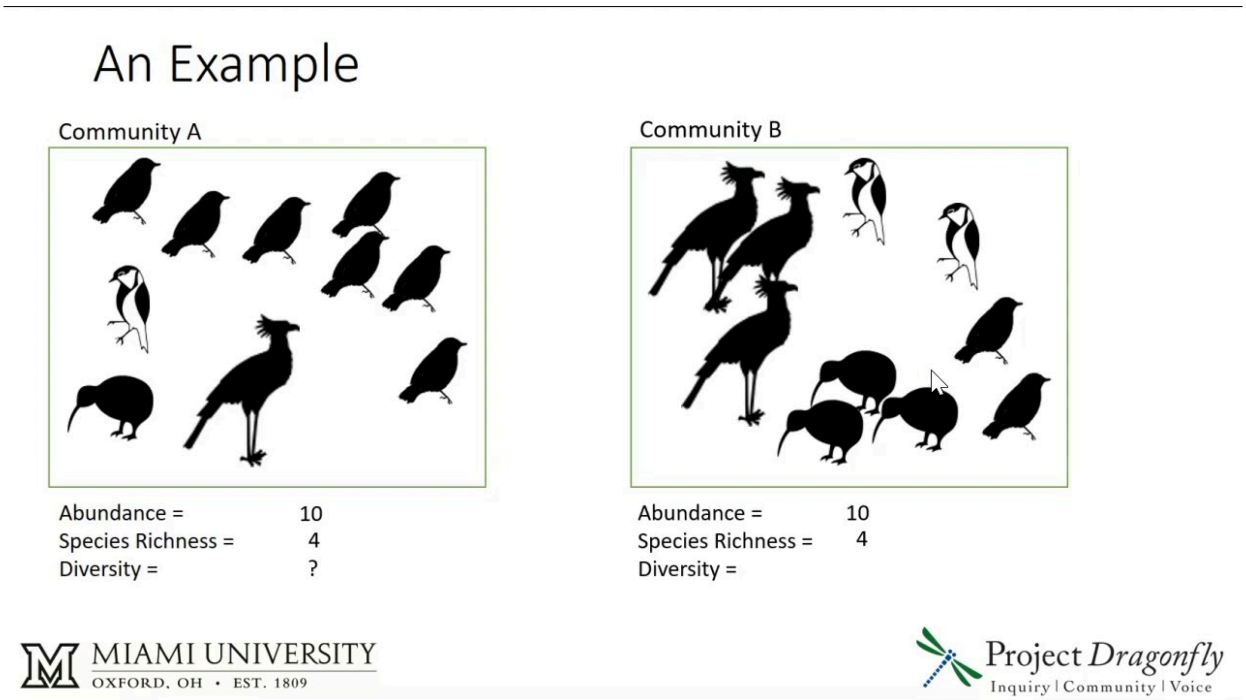
$$\text{Diversity Index} = 1 - \sum\left(\frac{n}{N}\right)^2$$

n = total number of organisms of a particular species

N = total number of organisms of all species

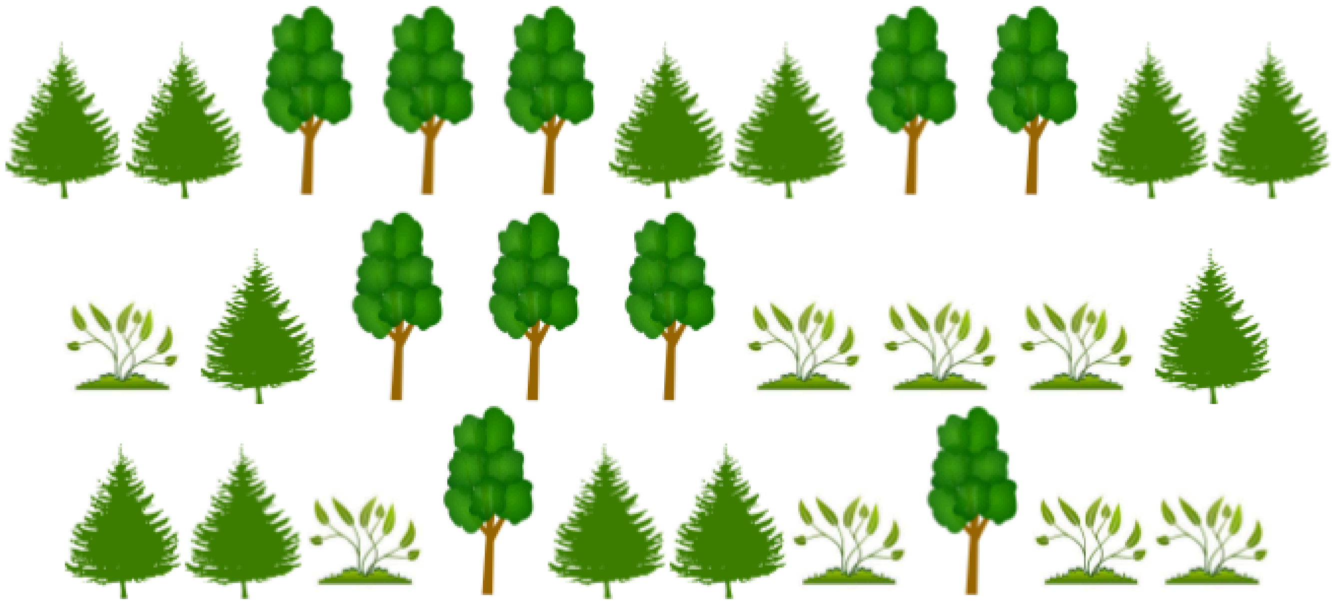
What is it used for?

- Allows us to calculate the diversity in the ecosystem that is being studied
    - Takes into account the species richness and relative abundance
- High Index = high diversity
- Low Indes = low diversity



Practice Problems:

1. Instructions: Use the community picture to complete the problems below.



   a. Calculate species richness.

   b. Calculate the relative abundance.

   c. Calculate the Simpson's diversity index

2. Looking at two different tins of nuts (mixed nuts and deluxe mixed nuts) students wanted to find out if there was a difference between the two. Using each type of nut as a species and treating each tin as their own community they collected the data in the table below.

Species	**Number of each species in “Mixed Nuts”**	**Number of each species in “Deluxe Mixed Nuts”**
Brazil Nut	1	8
Cashew	8	15
Pecan	1	7
Almond	15	18
Hazelnut	5	36
Peanut	85	31
Total	**115**	**115**

a. What is the species richness for the two tins of nuts? Which one is higher?

b. What is the relative abundance for the two tins of nuts? Which one is higher?

c. What is the simpson’s diversity index for the two tins of nuts? Which on is higher

3. Why is having a high diversity index genetically important?