**EVOLUTION OF TUSKLESSNESS IN AFRICAN ELEPHANTS**

**AUTHORS**

Jennifer Buntz, Central New Mexico Community College, Albuquerque, NM, USA

Jessica Ross, Midwood High School, Brooklyn, NY, USA

Tony Weisstein, Truman State University, Kirksville, MO, USA

Heather Zimbler-DeLorenzo, Georgia State University Perimeter College, Decatur, GA, USA

**ABSTRACT**

**LEARNING OBJECTIVES**

1. Be able to explain how natural selection produces change in populations within the context of tuskless elephants.
2. Calculate allele frequencies in a population given phenotypic data.
3. Calculate expected genotype frequencies given the allele frequencies.
4. Use a chi-squared test to determine whether a population is in Hardy-Weinberg Equilibrium.
5. Calculate allele frequencies for a population using sex linked traits.
6. Be able to explain how sex-linked traits affect evolution of characteristics differently.

**TEACHING NOTES**

This activity is divided into sections that can be taught together as one large class/lab activity or divided into sections that can be selected based on the level of the students or interest of the class.

We suggest that this activity be taught using Part 1 (Hardy Weinberg Equilibrium) and Part 2 (Chi-Square). Part 3 takes the second video of the Pre-Class Activity and has students apply the Hardy Weinberg Equilibrium using the newly published genetic information about tusklessness. It begins by starting with a sample example using hemophilia and moving on to tusklessness.

**ADDITIONAL MATERIALS:**

Hardy Weinberg PPT

Hardy Weinberg Schematic Diagram

**Pre-Class Activity:**

The Pre-Class Activity is provided to give the students biological context of the problem of elephant tusklessness. It contains two introductory videos that can either be assigned out of class as a pre-class activity or shown in class. The first video introduces the declining elephant population and how researchers are trying to find estimate population size and includes questions to make sure students understand the important points. The second video introduces students to the phenomenon of tuskless in elephants.

**In-Class Activity**

**Part I:**

This section presents students with the basics of the Hardy Weinberg Equilibrium. We have tried to explain the relationships between the two equations in a manner so that students can understand. We also have included an example that you can take the students through. To help with teaching of these equations and providing visual representation, you can use the PowerPoint and schematic drawing provided.

We also provided two methods to teach students how to calculate allelic and genotypic frequencies. You can choose whether you want to present both methods or just one, depending on the kind of students performing this activity.

After learning about Hardy Weinberg, we present the case of the tuskless elephants and *simulated* data using reports from Zambia. Students are asked to now go through and calculate allelic and genotypic frequencies using the data provided.

This section ends with a biological application of the Hardy Weinberg and also is leading students to what will be presented in Part 3.

**Part II:**

Most discussions of the chi-squared test do not explain the formula for each genotype's proportional deviation, **(O – E)2 / E**. We have tried to explain how this formula was obtained by thinking about what each part means: for example, to measure how different the observed and expected numbers are for each genotype, we begin with the term **O – E**. This value will be positive for some genotypes and negative for others, so we then square this term to prevent these values from canceling out. To explain the final division by **E**, it can be useful to frame the discussion in terms of two samples of very different size. For example, if you flipped a fair coin 1000 times, you would get an average of 500 heads and 500 tails. However, for a sample this large, you wouldn't be terribly surprised if you got heads 9 more times than average. By contrast, if you had flipped the same coin only 20 times, you would be very surprised if you got heads 9 more times than average. This is because the proportional deviation is only 0.45% (= 9 / 500) in the first case, but 90% (= 9 / 10) in the latter case.

Degrees of freedom basically indicate the number of different ways that a data set can meet a specified list of constraints. For example, when choosing a point on the (x,y) plane, we would normally have two degrees of freedom (the x- and y-coordinates). However, under the added constraint that the x- and y-coordinates be equal, knowing either coordinate would immediately tell us the other coordinate, so we would have only one remaining degree of freedom. For the chi-squaredtest of Hardy-Weinberg Equilibrium, our data set starts with three degrees of freedom (the numbers of AA, Aa, and aa individuals), then adds two constraints (the sample size N and the allele frequency p), leaving us with one degree of freedom. Another way to think about this issue is that, for specified values of N and p, knowing the observed number of any one genotype is enough to calculate the observed number of the other two genotypes.

**Part III:**

This section presents students with extending the concept of Hardy Weinberg Equilibrium to the case of X-linked. We have tried to explain how X-linked allelic frequencies need to be addressed in a manner so that students can understand. We also have included an example that you can take the students through using hemophilia.

After learning about Hardy Weinberg in the case of hemophilia, we present the case of the tuskless elephants and *simulated* data using reports from Zambia. Students are asked to now go through and calculate allelic and genotypic frequencies using the data provided, but this time for each type of individual.

**ANSWER KEY**

**Pre-Class Activity**

**Video 1 Questions**

1. What is the Great Elephant Census and why is this project important?

The Great Elephant Census is the first-ever pan-African survey of savanna elephants using standardized data collection and validation methods. Results will provide a baseline that governments and wildlife conservation organizations can use to coordinate conservation efforts.

1. What is the difference between a total count and sample count?

A total count is when the entire study area is surveyed, while a sample count is when parts of the study area are surveyed and the results from those parts are applied to the entire study area to estimate the total population size.

1. How does the sample count work?

A sample count involves dividing the survey areas into regions, called strata, of varying sizes. Surveyors then fly along the transect lines to estimate the number in that area, or stratum. Using this method, only elephants seen in the sections of the survey area are counted and used to determine elephant density in that area. These numbers are used to estimate the overall number of elephants based on a mathematical algorithm.

1. What is a transect?

A transect is a line across a habitat or part of a habitat. It can be as simple as a string or rope placed in a line on the ground. The number of organisms of each species along a transect can be observed and recorded at regular intervals.

1. How do they measure the strip?

Aerial sampling count consists of flying along parallel transects randomly or systematically distributed across the study area. Elephants are counted bt observers into strip samples (= sampling units). The strip samples are materialized by two streamers fixed perpendicularly to wing struts and parallel to the aircraft’s fuselage. Only animals seen between the streamers will be counted. Aircraft must fly at a constant height during the survey.

1. What are the 3 things used to calculate the elephant census?

The researchers use specific, premeasured areas as their sample, they use cameras to determine what animals are inside the area, and they use laser altimeters to calculate their height above the ground, which the counting strip depends on.

1. Why do we need accurate census data?

An accurate count of the African elephant population using up-to-date scientific techniques is a vital step in managing conservation efforts, identifying poaching hotspots and guiding law enforcement interventions, and assessing the impact of threats, such as habitat loss.

**Video 2 Questions**

1. What is Joyce Poole studying?

Dr. Joyce Poole, a scientist who has studied elephants for many years. She is Co-Founder and Co-Director of ElephantVoices. She has a Ph.D. in elephant behavior from Cambridge University, and has studied the social behavior and communication of elephants for over 40 years, dedicating her life to their conservation and welfare. Her contributions to science include the discovery of musth in male African elephants, the description of the contextual use of elephant vocalizations, including those below the level of human hearing, and the discovery of vocal imitation.

1. There was a civil war from 1977 to 1992, what happened during those years to the elephants?

Mozambique’s civil war from 1977 to 1992 had a grim outcome for elephants: During that time, some 90% were killed for the ivory in their tusks, which were sold to finance the war. Now, researchers report this intense hunting dramatically altered a major elephant population there, favoring female elephants born without tusks.

1. What is the interesting trait that elephants have that is being studied?

The adaptation of tuskless elephants, however, has an associated genetic mutation that kills male elephants before they’re born; even though this emerging trait may have helped save the elephant population.

1. What did Poole notice when she came to Gorongosa?

Dr. Poole noticed a high proportion of tuskless female elephants in Gorongosa.

1. What do males need their tusks for?

Males need tusks to fight other males for females (intrasexual competition). Without tusks, males are less likely to survive these fights and reproduce. So there is stronger selective pressure for males to have tusks than for females to have tusks.

1. What happened to the tusked elephants during the civil war?

Only about 200 of an estimated 2,500 elephants living there survived the ravages of the 15-year-long war during which poachers targeted tusked elephants for ivory.

1. Who is currently at a selective advantage, why?

Poaching was the selective pressure that reduced the proportion of tusked females, which increased the proportion of tuskless females.

**Part 1.**

1. Using the simulated elephant data provided in the table above, calculate the allele frequencies for tusked and tusklessness for the population - p and q - for the different time periods.

*p = Tusks*

*q = Tusklessness (only female are tuskless; there are no tuskless males)*

*1960: q2 = 12500/250000*

*q= .224*

*1-q=p*

*p=.776*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **1960s** | **1970s** | **1980s** | **1990s** | **2000s** |
| p | .766 | .708 | .613 | .564 | .730 |
| q | .224 | .292 | .387 | .436 | .270 |
| Total Frequency | 1 | 1 | 1 | 1 | 1 |

1. How do the frequencies change over time?

*The frequency of tusklessness begins to increase in females once the ivory wars begin. After the ivory ban, we began to see a decrease in tusklessness.*

1. Are there differences between males and females?

*Males are not tuskless. Therefore, only females can carry the gene of tusklessness. If they have both copies, then they display the characteristics. What students do not know yet is that males with the gene do not survive (don’t tell them yet).*

1. What is the ecological impact on the environment with the changes in the elephant populations?

 *African elephants play an important role in balancing natural ecosystems. They trample*

*forests and dense grasslands, making room for smaller species to co-exist. Elephants*

*also create water holes used by other wildlife as they dig dry riverbeds when rainfall is low. Herds travel over vast rangelands, and they disperse seeds in their dung, which helps generate new green growth. If the populations of elephants decrease over time, then so will their ecological impact. Elephants without tusks may be less able to push over trees or dig up riverbeds.*

1. What can be some reasons for why there are no tuskless males found in the population?

 *Students may come up with a variety of answers.*

**Part 2.**

1. Populations of pocket mice living in Mexico and the southwestern United States contain a mixture of light-colored and dark-colored individuals. Coloration is mainly due to a single gene: *DD* and *Dd* individuals are dark-colored, while *dd* individuals are light-colored. Researchers studying a population that lives on light-colored outcrops counted 9 *DD* mice, 34 *Dd* mice, and 34 *dd* mice.
2. Based on these numbers, calculate the frequencies of the *D* and *d* alleles in this population. Use these frequencies to calculate the expected number of *DD*, *Dd*, and *dd* mice, given a population size of 77.

*Total # allele copies = 2 × 77 = 154.*

*# D alleles = 2 × 9 (DD mice) + 1 × 34 (Dd mice) = 52.*

 *Frequency (p) of the D allele = 52 / 154 = 0.338.*

*# d alleles = 2 × 34 (dd mice) + 1 × 34 (Dd mice) = 102.*

 *Frequency (q) of the d allele = 102 / 154 = 0.662.*

*Expected # of DD mice = p2 × N = (0.338)2 × 77 = 8.779.* ***Do not round!***

*Expected # of Dd mice = 2pq × N = 2(0.338)(0.662) × 77 = 34.442.*

*Expected # of dd mice = q2 × N = (0.662)2 × 77 = 33.779.*

1. Compare the observed and expected numbers of each genotype. **Based solely on these numbers**, would you predict that this population is in Hardy-Weinberg Equilibrium or not? Explain your reasoning.

*The observed numbers for each genotype are very close to the numbers that would be expected under Hardy-Weinberg Equilibrium. We would therefore predict that the population will be in Hardy-Weinberg Equilibrium, though we still need to conduct a chi-squared test to be sure.*

1. Now conduct a chi-squared test on these data. Summarize your results in a sentence of the form "This population [does / does not] appear to be in Hardy-Weinberg Equilibrium for the D gene." List the value of your chi-squared statistic and the number of degrees of freedom for your chi-squared test.

*Calculating (O – E)2 / E for each genotype:*

 *DD: (9 – 8.779)2 / 8.779 = 0.006.*

 *Dd: (34 – 34.442)2 / 34.442 = 0.006.*

 *dd: (34 – 33.779)2 / 33.779 = 0.001.*

*Summing these to find the chi-squared value yields χ2 = 0.013. This test has one degree of freedom (see Teaching Notes), so we will compare our calculated χ2 value to the corresponding critical value of 3.841 (see table of critical values in main text). Doing so tells us that this population does appear to be in Hardy-Weinberg Equilibrium for the* D *gene (χ2 = 0.013, 1 df).*

1. While studying the Zambian elephant population, researchers took tissue samples so that they could determine each elephant's genotype. Once the tusklessness gene was discovered, the researchers could re-analyze these tissue samples for that gene. They found that, of the 500 female elephants sampled during the 1980s, 186 were *TT*, 164 were *Tt*, and 150 were *tt*. (Recall that elephants with at least one *T* allele have tusks, while those with two *t* alleles do not.)
2. Based on these numbers, determine whether the female elephants in the Zambian population were in Hardy-Weinberg Equilibrium during this sampling period. Summarize your results in a single sentence and state the value of your chi-squared statistic and the number of degrees of freedom.

*Total # allele copies = 2 × 500 = 100.*

*Frequency (p) of the T allele = (2 × 186 + 1 × 164) / 1000 = 0.536.*

*Frequency (q) of the t allele = (2 × 150 + 1 × 164) / 1000 = 0.464.*

*Expected numbers: 143.648 TT, 248.704 Tt, 107.648 tt.*

*(O – E)2 / E for each genotype: TT = 12.487, Tt = 28.849, tt = 16.663. The sum is 57.998, which is much larger than the critical value for 1 df.*

*This population does not appear to be in Hardy-Weinberg Equilibrium for the* T *gene (χ2 = 57.998, 1 df).*

1. Look at the **(O – E)2 / E** column of your chi-squared table. Which of the three genotypes has the largest proportional deviation? Does the observed population contain more or fewer individuals with this genotype than would be expected under Hardy-Weinberg Equilibrium?

*At a value of 28.849, the heterozygous Tt genotype has the largest proportional deviation. The observed population contains many fewer Tt individuals (164) than would be expected under Hardy-Weinberg Equilibrium (248.704).*

1. What specific evolutionary processes might cause the deviation you noted in part (b)? For example, if a lot of Zambia's elephants *migrated* to other areas due to local poaching, would that explain why certain genotypes were more or less common than expected? Could natural or artificial *selection* be favoring certain genotypes? Explain your reasoning.

*Many different evolutionary processes could cause the lower-than-expected number of heterozygotes. It seems unlikely that Tt elephants would be much likelier than other genotypes to migrate away from Zambia. On the other hand, migration from nearby elephant populations into Zambia could explain these results if those populations had low frequencies of Tt elephants. The results could also be the result of selection for tusklessness. In particular, if poachers selectively kill tusked elephants, they would increase the frequency of the t (tuskless) allele but would not correspondingly increase the frequency of Tt heterozygotes. The chi-squared test would detect this result as a deviation from Hardy-Weinberg Equilibrium caused by too few heterozygotes.*

**Part 3.**

1. Using the simulated elephant data provided in the table above, calculate the allele frequencies for tusked and tusklessness for the males as well as females; p and q for the different time periods, assuming that this gene is X-linked recessive.

Think about which organisms should be q? (Tuskless)

p = Tusks

q = Tusklessness (only female are tuskless; there are no tuskless males)

1960: q2 = 12500/125,000

q= .1

1-q=p

p=.683

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| males | **1960s** | **1970s** | **1980s** | **1990s** | **2000s** |
| p | 0 | 0 | 0 | 0 | 0 |
| q | 1 | 1 | 1 | 1 | 1 |
| Total Frequency | 1 | 1 | 1 | 1 | 1 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| females | **1960s** | **1970s** | **1980s** | **1990s** | **2000s** |
| p | .316 | .412 | .548 | .616 | .382 |
| q | .683 | .588 | .452 | .383 | .618 |
| Total Frequency | 1 | 1 | 1 | 1 | 1 |

1. How do the frequencies change over time for males and females?

The frequency of the recessive allele (for tusklessness) increases over time as poaching increases. It begins to decrease in the 2000’s when the civil war and poaching are less prevalent.

1. How do these frequencies change compared to in part I?

**ACTIVITY**

***Pre-Class Activity***

***Video 1: Great Elephant Census***

According to data collected by the World Wildlife Fund (WWF), just under 100 years ago there were over 10 million African elephants. But decades of poaching and conflict have since decimated African elephant populations. In 2016, experts estimated that Africa’s elephant population had dropped by 111,000 elephants in the span of a decade. Today, there are just 415,000 elephants across Africa. While elephant poaching is trending downward, with significant declines in East Africa, poaching continues to steer the species dangerously nearer to extinction.

Watch the Great Elephant Census provided by HHMI Biointeractive (<https://www.biointeractive.org/classroom-resources/great-elephant-census>).

Video Questions:

1. What is the Great Elephant Census and why is this project important?
2. What is the difference between a total count and sample count?
3. How does the sample count work?
4. What is a transect?
5. How do they measure the strip?
6. What are the 3 things used to calculate the elephant census?
7. Why do we need accurate census data?

***Video 2: Selection of Tuskless Elephants***

Tusks are teeth—upper incisors to be exact. During the first year of life, a baby elephant’s tusks will replace his set of milk teeth, extending from a socket in the skull. Prying bark from trees, digging for water or roots, fighting other bulls—elephants’ tusks perform a variety of functions. Tusks continue to grow throughout an elephant’s life, becoming longer and thicker generally with age. Male tusks can grow to be seven times the weight of female tusks as they age.

There is also genetic variability with regard to tusk length and thickness, with some older males having smaller tusks. However, heavy poaching of large-tusked elephants can influence the biology of future generations of elephants because poachers are selecting tusk size, therefore will tend to kill older males with very large tusks. This results in removing breeding-aged males who happen to have very big tasks out of the population. Those males then no longer pass on their genes for large tusks.

The normal level of tusklessness in an elephant population is somewhere between 3-4% (CITE 2008 African Journal of Ecology). The number of tuskless female elephants in Zambia’s South Luangwa National Park and Lupande Game Management Area has increased from 10.5% in 1969 to 38.2% in 1989. As this coincided with the ivory wars, tusklessness may be the result of illegal hunting for ivory (Jachmann et al. 1995). A 1991 elephant conservation plan in Uganda reported a higher-than-normal percentage of tuskless elephants in Queen Elizabeth National Park and singled out poaching as the main cause. In a 1989 survey of Queen Elizabeth National Park revealed tusklessness in the elephant population to be between 9 and 25%.

Watch the Selection of Tuskless Elephants

(<https://www.biointeractive.org/classroom-resources/selection-tuskless-elephants>)

Video Questions:

1. What is Joyce Poole studying?
2. There was a civil war from 1977 to 1992, what happened during those years to the elephants?
3. What is the interesting trait that elephants have that is being studied?
4. What did Poole notice when she came to Gorongosa?
5. What do males need their tusks for?
6. What happened to the tusked elephants during the civil war?
7. Who is currently at a selective advantage, why?

***In-Class Activity***

**Part I: Hardy-Weinberg Equilibrium**

Hardy-Weinberg Equilibrium, also referred to as the Hardy-Weinberg principle, is used to compare allele frequencies in a given population over a period of time. A population must meet five rules in order to be considered “in equilibrium”.

1) No gene mutations may occur and therefore allele changes do not occur.

2) There must be no migration of individuals either into or out of the population.

3) Random mating must occur, meaning individuals mate by chance.

4) No genetic drift, a random change due to change in allele frequency, may occur.

5) No natural selection, a change in allele frequency due to the environment, may occur.

Hardy-Weinberg Equilibrium never occurs in nature because there is always at least one rule being violated. Hardy-Weinberg Equilibrium is an ideal state that provides a baseline against which scientists measure gene evolution in a given population. When the assumptions aren’t met, we may get different genotype frequencies. The Hardy-Weinberg equations can be used for any population; the population does not need to be in equilibrium.

There are two equations necessary to solve a Hardy-Weinberg Equilibrium question:

1. The first Hardy-Weinberg equation (p + q = 1) concerns estimating the frequency of **ALLELES** in a population. Each gene usually has two alleles (diploid organism), one from each parent. These alleles are denoted as the dominant (A) and recessive (a) forms. These are represented as ‘p‘ and ‘q‘ is the equation. In a population, the combined frequency of both the alleles must equal 1 (or 100%). Therefore, if the frequency of one allele is known, it is possible to calculate the frequency of the other allele simply by rearranging the equation.
2. The Hardy-Weinberg equation used to determine **GENOTYPE** frequencies is: p2 + 2pq + q2 = 1. Since a genotype contains two alleles the equation is based on (p + q) X (p + q) = 1. You can simplify this equation using the FOIL rule (first, outside, inside, last) = p2 + pq + pq +q2 = 1. Simplified, this equation is p2 + 2pq + q2 = 1.

Here 𝑝² is the frequency of individuals with the homozygous dominant (AA) genotype, 2𝑝𝑞 is the frequency of individuals with the heterozygous (Aa) genotype, and 𝑞² is the frequency of individuals with the homozygous recessive (aa) genotype. Again, if one genotype frequency is known, it is possible to use the Hardy-Weinberg equations to work out the others.



 This equation is the binomial expansion of the first equation (p + q = 1).

 (p + q = 1)2 OR (p + q) x (p + q) = 1 x 1

In the simplest possible situation we have a single gene with only two alleles. These alleles might be A and a. In a real world population, we can only see phenotypes, not genotypes or alleles. AA and Aa are different genotypes but both produce the same phenotype. Therefore, the observed frequency of allele A equals the sum of all of the AA genotype plus half of the Aa genotype (the A half). The observed frequency of allele a is therefore half of the Aa individuals (the a half) plus all of aa individuals. If you know one value, you can of course just subtract it from 1 (100%) to get the value of the other. In other words, the observed frequency of A = 100%(AA) + 50%(Aa) and a = 50%(Aa) and 100%(aa).

—--------------------------------------------------------------------------------------------------------------------

**EXAMPLE**: An allele W, for white wool, is dominant over allele w, for black wool. In a sample of 900 sheep, 891 are white and 9 are black. Calculate the allele frequencies within this population, assuming that the population is in H-W equilibrium. There are two approaches to finding the allele frequencies.

By convention, we use the dominant phenotype to name the alleles. In this case, white is dominant over black, so we'll use 'W' for white fur and 'w' for black fur. Assign 'p' to be the frequency of W, the white wool allele. Assign 'q' to be the frequency of w, the black wool allele.

 𝑝² + 2𝑝𝑞 + 𝑞² = 1

 (WW) (Ww) (ww)

p = allelic frequency of W q = allelic frequency of w

total number of individuals = 900

 black wool sheep (ww) = 9 individuals

 white wool sheep (WW and Ww) = 891 individuals

Method 1: If you know that your population is in Hardy Weinberg Equilibrium, you can use this short cut technique.

**Step 1:** We start by trying to determine the number of homozygous recessive alleles because this is easily known from the information. However, the number of homozygous recessive individuals is q2, NOT q (which is the allele frequency). To find q, take the square root of q2.

 q2 = 9 black wool sheep = 0.01 therefore q = 0.1

 900 total sheep

**Step 3:** Once you have q, finding p is easy! The sum of the frequencies of both alleles must equal 1.

Now, p + q = 1 thus p = 1 – q or q = 1- p.

p = 1 – 0.1 = 0.9

q = 1 – 0.9 (p) = 0.1

**Step 4:** Use p and q to calculate the remaining genotypes. It is suggested that you calculate q2 even though that's what you started with. Realize that rounding will give you slightly different values and all 3 genotype frequencies may not EXACTLY equal 1 because of this. It's fine.

 p2 = 0.9 X 0.9 = 0.81 **HOMOZYGOUS DOMINANT**

 2pq = 2 x (0.9) x (0.1) = 0.18 **HETEROZYGOUS**

 q2 = 0.1 x 0.1 = 0.01 **HOMOZYGOUS RECESSIVE**

**Step 5:** Check your math. This is easy because the equation should equal 1.

Method 2: If we do not know whether the population is in Hardy Weinberg equilibrium, we can do the same calculations using the numbers of sheep with each genotype.

**Step 1:** We must be provided with the numbers of sheep with each genotype.

WW (HOMOZYGOUS DOMINANT for white wool) = 729 sheep

 Ww (HETEROZYGOUS and displaying white wool) = 162 sheep

 ww (HOMOZYGOUS RECESSIVE for black wool) = 9 sheep

**Step 2**: Calculate genotypic frequency of the homozygous dominant. To do this you must remember that every individual has two alleles for each gene. A white homozygous dominant sheep would have two copies of the W allele. You would then divide this by the number of total alleles, which is equal to the number of individuals times 2 (since we all have 2 alleles per gene).

 Genotypic frequency of WW = W + W = 729 + 729 = 0.81(p2)

Total alleles = 900 + 900

**Step 3**: Calculate genotypic frequency of the heterozygous. A white heterozygous sheep would have one copy of the W allele and one copy of the w allele. You would then divide this by the number of total alleles, which is equal to the number of individuals times 2 (since we all have 2 alleles per gene).

 Genotypic frequency of Ww = W + w = 729 + 162 = 0.18 (2pq)

 Total alleles = 900 + 900

**Step 4**: Calculate genotypic frequency of the homozygous recessive. A black sheep would have two copies of the w allele. You would then divide this by the number of total alleles, which is equal to the number of individuals times 2 (since we all have 2 alleles per gene).

Genotypic frequency of ww = w + w = 9 + 9 = 0.18 (q2)

Total alleles = 900 + 900

**Step 5**: Now we can calculate allelic frequencies using the genotypes. For each genotype, there are two alleles. W allele comes from both the homozygous dominant and heterozygous genotypes. w allele comes from the homozygous recessive and heterozygous genotypes. Then, divide the total of the allele by the total number of possible alleles. Remember that p + q will equal 1. If you calculate one of the allele frequencies, you can check your math by subtracting from 1.

 W allele = 729 + 729 + 162 = 1620

(Homozygous W) (Heterozygous)

 w allele = 9 + 9 + 162 = 180

(Homozygous w) (Heterozygous)

Total alleles = 900 + 900 = 1800

Allelic Frequencies:

 W allele = 1620/1800 = 0.9

 w allele = 180/1800 = 0.1

—---------------------------------------------------------------------------------------------------------------------

Simulated data is provided for elephants in Zambia at different time periods. The time periods represent before the ivory wars, during the Ivory wars (1970s and 1980s), and after the ivory wars. The ivory ban was put in place in 1989. Tusklessness is a recessive allele (q).

|  | **1960s** | **1970s** | **1980s** | **1990s** | **2000s** |
| --- | --- | --- | --- | --- | --- |
| **Tusked Males** | 125,000 | 9,000 | 500 | 5,000 | 7,000 |
| **Tuskless Males** | 0 | 0 | 0 | 0 | 0 |
| **Tusked Females** | 112,500 | 7,470 | 350 | 3,100 | 4,980 |
| **Tuskless Females** | 12,500 | 1,530 | 150 | 1,900 | 1,020 |
| **Totals** | **250,000** | **18,000** | **1,000** | **10,000** | **14,000** |

1. Using the simulated elephant data provided in the table above, calculate the allele frequencies for tusked and tusklessness for the population - p and q - for the different time periods.
2. How do the frequencies change over time?
3. Are there differences between males and females?

Elephant tusks evolved from teeth, giving the species an evolutionary advantage. They serve a variety of purposes: digging, lifting objects, gathering food, stripping bark from trees to eat, and defense. The tusks also protect the trunk—another valuable tool for drinking, breathing, and eating, among other uses.

1. What is the ecological impact on the environment with the changes in the elephant populations?
2. What can be some reasons for why there are no tuskless males found in the population?

**Part II: The chi-squared test**

As described earlier, Hardy-Weinberg Equilibrium predicts that the genotypes within a population will be in the ratio *p*2 : 2*pq* : *q*2. To determine how well a particular population fits this ratio, biologists use a statistical analysis called a ***chi-squared* *test*.**

To set up a chi-squared test, we start by determining the number of individuals with each genotype (AA, Aa, and aa). We use these values to calculate the frequency of each allele, and then the number of individuals expected under Hardy-Weinberg Equilibrium (obtained by multiplying the expected genotype frequencies p2, 2pq, and q2 by the total population size N). For example, in the wool color example from Part I, we found that p = 0.9 and q = 0.1 in a population of 900 sheep. We therefore find that p2 = 0.81: in other words, under Hardy-Weinberg Equilibrium, 81% of the sheep would be WW. We then take 81% of 900 to find an expected number of 729 WW sheep. Similar calculations show that the population should contain 162 Ww sheep and 9 ww sheep if it is at Hardy-Weinberg Equilibrium.

—--------------------------------------------------------------------------------------------------------------------

EXAMPLE:The data below are from a study of the MN blood group in humans. Similar to the more familiar ABO blood group, *M* and *N* are alleles that code for a protein expressed on the surface of blood cells. People with the *MM* genotype produce only the *M* form of the protein, those with the *NN* genotype produce only the *N* form, and those with the *MN* genotype produce both.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Genotype** | **# observed** | **# expected (under HWE)** | **# Obs –****# Exp** | **(O – E)2** | **(O – E)2/E** |
| ***MM*** | 1101 |  |  |  |  |
| ***MN*** | 1496 |  |  |  |  |
| ***NN*** | 503 |  |  |  |  |
| **TOTAL** | 3100 |  |  |  |  |

**Step 1:** Based on the observed numbers of each genotype, calculate the frequency of the *M* and *N* alleles. (You may use *p* to represent the former and *q* for the latter.)

**Step 2:** Use these frequencies to calculate the expected frequency of each genotype if this population is in Hardy-Weinberg Equilibrium.

**Step 3:** Finally, multiply the frequency of each genotype by the **total** population size to calculate the number of people that you would expect to show each genotype if the population is in Hardy-Weinberg Equilibrium. Keep at least two decimal places to minimize rounding error. Record these values in the table above.

Next, we will measure how different the observed and expected numbers are for each genotype. Mathematically, we do this by subtracting the expected number from the observed number. Next, we square each value to prevent positive and negative values from canceling each other out.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Genotype** | **# observed** | **# expected (under HWE)** | **# Obs –****# Exp** | **(O – E)2** | **(O – E)2/E** |
| ***MM*** | 1101 | 1102.84 |  |  |  |
| ***MN*** | 1496 | 1492.32 |  |  |  |
| ***NN*** | 503 | 504.84 |  |  |  |
| **TOTAL** | 3100 | 3100.00 |  |  |  |

**Step 4**: For each genotype, calculate the difference between the observed and expected numbers and square the result. Record these values in the table above. Note that, if your calculations are correct, the entries in column 3 **(# Obs – # Exp)** should add up to zero.

We then divide by the expected number of each genotype to determine the *proportional* deviation. Finally, we add these proportional deviations across all genotypes to calculate the value of the $χ^{2}$(chi-squared) statistic.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Genotype** | **# observed** | **# expected (under HWE)** | **# Obs –****# Exp** | **(O – E)2** | **(O – E)2/E** |
| **AA** | 1101 | 1102.84 | -1.84 | 3.38 |  |
| **Aa** | 1496 | 1492.32 | 3.68 | 15.53 |  |
| **aa** | 503 | 504.84 | -1.84 | 3.38 |  |
| **TOTAL** | 3100 | 3100.00 |  |  | $χ^{2}$ **=**  |

**Step 5**: Calculate the proportional deviation **(O – E)2 / E** for each genotype, and record your results in the table above. Finally, add up these values for all three genotypes and record the result as the value of the $χ^{2}$statistic for this data set.

—----------------------------------------------------------------------------------------------------------------

The $χ^{2}$statistic measures the overall *mismatch* between the observed and expected numbers of each genotype. A $χ^{2}$value of zero means that the numbers correspond perfectly, meaning that the observed data are perfectly consistent with the hypothesis being tested (here, that the population is in Hardy-Weinberg Equilibrium). By contrast, we will obtain a small but positive $χ^{2}$value if the observed and expected numbers for each genotype are slightly different, or a larger $χ^{2}$value if the numbers for each genotype are substantially different. To determine how well our data fit our hypothesis, we must therefore determine whether our $χ^{2}$value is considered "small" or "large."

Statisticians have constructed tables of critical values ($χ\_{crit}^{2}$) for the $χ^{2}$ test. Each critical value can be obtained by generating thousands of random data sets and plotting the $χ^{2}$ value for each. In most cases, we want to identify a $χ\_{crit}^{2}$ such that 95% of the random data sets will yield $χ^{2}$ < $χ\_{crit}^{2}$. (You may have encountered this idea before in your statistics classes, where you used a confidence level of α = 0.05 to determine statistical significance.) Any $χ^{2}$ greater than $χ\_{crit}^{2}$ is unlikely to happen by chance and is therefore considered sufficient evidence to reject the hypothesis being tested.

To use a critical value table, we must also know the degrees of freedom (d.f.) for our data set. When testing Hardy-Weinberg Equilibrium, we will always have one degree of freedom, though the details are beyond the scope of this activity.

Once we have determined our data set's $χ^{2}$value and degrees of freedom, we can look up the appropriate $χ\_{crit}^{2}$value in a table like the one at the right. (A larger version of this table, with values up to 100 degrees of freedom at different levels of confidence, can be found at <https://www.itl.nist.gov/div898/handbook/eda/section3/eda3674.htm> .) In this case, our calculated $χ^{2}$value of 0.019 is less than the critical value for 1 degree of freedom. Because our $χ^{2}$value is so small, we conclude that our data fit Hardy-Weinberg Equilibrium quite well (more precisely, our data do not justify rejecting the hypothesis that the population is in Hardy-Weinberg Equilibrium).

**Sample problems:**

1. Populations of pocket mice living in Mexico and the southwestern United States contain a mixture of light-colored and dark-colored individuals. Coloration is mainly due to a single gene: *DD* and *Dd* individuals are dark-colored, while *dd* individuals are light-colored. Researchers studying a population that lives on light-colored outcrops counted 9 *DD* mice, 34 *Dd* mice, and 34 *dd* mice.
2. Based on these numbers, calculate the frequencies of the *D* and *d* alleles in this population. Use these frequencies to calculate the expected number of *DD*, *Dd*, and *dd* mice, given a population size of 77.
3. Compare the observed and expected numbers of each genotype. **Based solely on these numbers**, would you predict that this population is in Hardy-Weinberg Equilibrium or not? Explain your reasoning.
4. Now conduct a chi-squared test on these data. Summarize your results in a sentence of the form "This population [does / does not] appear to be in Hardy-Weinberg Equilibrium for the D gene." List the value of your chi-squared statistic and the number of degrees of freedom for your chi-squared test.
5. While studying the Zambian elephant population, researchers took tissue samples so that they could determine each elephant's genotype. Once the tusklessness gene was discovered, the researchers could re-analyze these tissue samples for that gene. They found that, of the 500 female elephants sampled during the 1980s, 186 were *TT*, 164 were *Tt*, and 150 were *tt*. (Recall that elephants with at least one *T* allele have tusks, while those with two *t* alleles do not.)
6. Based on these numbers, determine whether the female elephants in the Zambian population were in Hardy-Weinberg Equilibrium during this sampling period. Summarize your results in a single sentence and state the value of your chi-squared statistic and the number of degrees of freedom.
7. Look at the **(O – E)2 / E** column of your chi-squared table. Which of the three genotypes has the largest proportional deviation? Does the observed population contain more or fewer individuals with this genotype than would be expected under Hardy-Weinberg Equilibrium?
8. What specific evolutionary processes might cause the deviation you noted in part (b)? For example, if a lot of Zambia's elephants *migrated* to other areas due to local poaching, would that explain why certain genotypes were more or less common than expected? Could natural or artificial *selection* be favoring certain genotypes? Explain your reasoning.

**Part III: Hardy Weinberg with Sex Linked Traits**

Elephants are a diploid organism with X-Y sex determination (females are XX ; males are XY). The inheritance of a trait (phenotype) that is determined by a gene located on one of the sex chromosomes is called sex linked inheritance. The expectations of sex-linked inheritance in any species depend on how the chromosomes determine sex. For example, in humans, males are heterogametic, with one X chromosome and one Y chromosome. But females are homogametic, with two X chromosomes. In human males, the entire X chromosome is active. In general terms, traits determined by genes on sex chromosomes are not different from traits determined by autosomal genes. Sex-linked traits are distinguishable by their mode of transmission through successive generations of a family. In humans it is called X-linked or Y-linked inheritance.

If allele frequencies differ between the sexes, it takes two generations of random mating to attain Hardy-Weinberg equilibrium. Sex-linked loci require multiple generations to attain equilibrium because one sex has two copies of the gene and the other sex has only one.

—--------------------------------------------------------------------------------------------------------------------

EXAMPLE: We can use the example of hemophilia to demonstrate how to calculate HW frequencies. Hemophilia is a bleeding disorder in which blood does not clot properly. Blood contains many proteins, called clotting factors, that can help to stop bleeding after injury or surgery. People with hemophilia have low amounts of either factor VIII (eight) or factor IX (nine), which are key factors responsible for normal blood clotting. Since there are no genes for clotting factors on the Y chromosome, this means that males only have one allele for factor VIII and one allele for factor IX. Thus, if a male has a hemophilia allele on his only X chromosome, he will display the disorder. Possible male genotypes are: XHY or XhY. Remember that male genotypes are no different in the Hardy Weinberg equation, except that p2 or q2 are not possible in males.

A female inherits two copies of the factor VIII or factor IX gene, one from her mother and one from her father. A female with a hemophilia allele on one X chromosome usually has a normal allele on her other X chromosome that can produce normal clotting factor, so she has some protection against having hemophilia (although is a carrier of the disorder). Possible female genotypes are: XHXH, XHXh, or XhXh. Females therefore fit the Hardy Weinberg genotypic frequencies (p2 and 2pq and q2) since they are diploid for the X chromosome.

We can determine the allele frequency and heterozygotic carrier frequency in a population for which the frequency of the trait is known.

If the X-linked recessive frequency is 1/5000 and there are 1/2500 Female carriers, we can do the calculation as follows:

 **Step 1:** Identify the dominant and recessive allele frequencies in sex chromosomes.

 p = regular clotting

 q = hemophilia

 **Step 2:** Using the information from above, calculate the frequencies.

 q = 1/5000 = 0.002

 p = 1- q = 1 - 0.002 = 0.998

 **Step 3:** Using p and q, calculate the frequencies of the genotypes.

 *FEMALES:*

p2 = 0.998 x 0.998 = **HOMOZYGOUS DOMINANT (normal)**

 2pq = 2 x (0.998) x (0.002) = **HETEROZYGOUS (carriers)**

 q2 = 0.002 x 0.002 = **HOMOZYGOUS RECESSIVE**

*MALES:*

q = 0.002 = males afflicted

p = 0.998 = normal males

—-------------------------------------------------------------------------------------------------------

In order to calculate allele frequencies in X-linked disorders, the following things need to be kept in mind:

q = Males are hemizygous for the X-chromosome, therefore frequency of affected males

is the frequency of the mutant allele

2q = for rare X-linked disorders, frequency of heterozygous carrier females is twice the

frequency of affected males

q = the frequency of homozygous females is very low

We want to study the evolution of tusklessness, which is a sex-linked trait with two alleles on the X-chromosome (with no counterpart on the Y-chromosome). However, this presence of this dominant trait in a male is lethal (due to the lack of a second sex chromosome to mask it). Two genes associated with tooth development in mammals are responsible for the tuskless elephant phenomenon. One of these genes is connected to the X chromosome and is lethal to males, while humans who have the same gene mutation exhibit similar teeth defects. If a female tuskless elephant were to become pregnant with a male, she would spontaneously abort the fetus half of the time. Her daughters would have a 50% chance of being tuskless (Jachmann, 1995).

Thinking about an X-linked recessive disorder, what would happen if the trait was lethal in males? There would be no males with the disorder. We see this case in elephants. Here is the data from the elephants from the past 50 years.

|  | **1960s** | **1970s** | **1980s** | **1990s** | **2000s** |
| --- | --- | --- | --- | --- | --- |
| **Tusked Males** | 125,000 | 9,000 | 500 | 5,000 | 7,000 |
| **Tuskless Males** | 0 | 0 | 0 | 0 | 0 |
| **Tusked Females** | 112,500 | 7,470 | 350 | 3,100 | 4,980 |
| **Tuskless Females** | 12,500 | 1,530 | 150 | 1,900 | 1,020 |
| **Totals** | **250,000** | **18,000** | **1,000** | **10,000** | **14,000** |

1. Using the simulated elephant data provided in the table above, calculate the allele frequencies for tusked and tusklessness for the males as well as females; p and q for the different time periods, assuming that this gene is X-linked recessive.
2. How do the frequencies change over time for males and females?
3. How do these frequencies change compared to in part I?

**REFERENCES**

Med.umich.edu. 2022. [online] Available at: <https://www.med.umich.edu/lrc/coursepages/m1/humangenetics/pdfs/Lec4-5PopulationGenetics.pdf> [Accessed 24 February 2022].

Centers for Disease Control and Prevention. 2022. *How Hemophilia is Inherited | Hemophilia | NCBDDD | CDC*. [online] Available at: <https://www.cdc.gov/ncbddd/hemophilia/inheritance-pattern.html> [Accessed 24 February 2022].

Jachmann H, Berry PSM, Imae H. 1995. Tusklessness in African elephants: a future trend. African Journal of Ecology 33: 230-235.

Princeton University. 2022. *Princeton scientists identify genes behind tusklessness in African elephants facing poaching pressure*. [online] Available at: <https://www.princeton.edu/news/2021/11/01/princeton-scientists-identify-genes-behind-tusklessness-african-elephants-facing> [Accessed 24 February 2022].

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

This material is based upon work supported by the National Science Foundation under Grant No.1919613. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.