Hardy-Weinberg Equilibrium in R (swirl lesson)

Author: Abigail E. Cahill, Albion College, acahill@albion.edu

Course Information

Department: **Biology**
Level: **Upper Undergraduate**
Course type: **Lecture with computer lab component**
Students: **Majors**
Number of Students: **16**

Focus: Students will use swirl to understand Hardy-Weinberg equilibrium. The lesson starts with observed numbers of individuals for each genotype, and students will work through a number of steps to assess whether or not the population is in equilibrium.

Overview: This lesson will allow students to test a population to see if it meets Hardy Weinberg expectations starting with observed numbers of individuals in each genotype within the population. It will require them to calculate observed allele frequencies, expected genotype frequencies, and use a chi-square test to get (and then interpret) a p-value.

Learning objectives:

By the end of this swirl lesson, students will be able to

1. Test a population for Hardy-Weinberg equilibrium when given observed genotype data (numbers of individuals of each genotype)
2. Use the chisq.test() function in R and interpret a p-value

Lesson sequence:

1. In-class swirl intro (see pre-lesson activities)
2. In-class Hardy Weinberg content (see pre-lesson activities)
3. Hardy Weinberg swirl lesson

Pre-lesson activities: Prior to this lesson, students had completed a short in-class exercise in R, using the first module in the R Programming course by Team swirl (http://swirlstats.com/scn/rprog.html).

Students had also completed about 3 hours of course content on the basics of Hardy-Weinberg equilibrium, introducing what it is, how to calculate values, and why it matters. Prior to taking this course, they had also been exposed to the very basics of HWE in an introductory ecology and evolution course.

The lesson expects that students will know what alleles are and understand allele and genotype frequencies. It also expects that they will know the HWE formulas ($p+q=1$) and ($p^2+2pq+q^2=1$).

Although the lesson has students calculate a chi-square test, students do not need to have an in-depth understanding of the test itself. They should know how to interpret a p-value.

Post-lesson activities: Students completed other Hardy-Weinberg problems both in an R-markdown in-class assignment (currently available from Abigail Cahill by request, acahill@albion.edu) and on an exam. These follow-up problems did not break HWE into small steps in the way the swirl lesson does, so hopefully students used the swirl lesson to understand each step of the calculation along the way and why the steps are necessary.

Implementation notes: This exercise was completed by students during a 45-minute class period. It was a part of an Ecological Genetics class, used relatively early in the semester. The students completed the lesson in RStudio Cloud; although a cloud-based delivery is not necessary, it was very helpful when we moved to online instruction during the COVID-19 pandemic.