

BIOL 701 Likelihood Methods in Biology: Homework #2
Due Monday, February 18

1. Radioactive decay is a classic “Poisson process”. The number of alpha particles emitted from a sample of radioactive material during a fixed unit of time follows the Poisson distribution:

$$\Pr(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}$$

where X is the number of particles emitted (the random variable), k is a particular value of X , and λ is the mean emission rate (which depends on the amount and type of radioactive material). We use an alpha particle detector to investigate an unknown material. The number of particles detected per minute is recorded for 20 distinct, 1-minute time intervals (see Table 1). Use the likelihood ratio method to test the hypothesis that the mean emission rate, $\lambda = 2$.

Interval I.D.	Particles detected	Interval I.D.	Particles detected
1	3	11	1
2	2	12	3
3	2	13	1
4	5	14	4
5	4	15	5
6	6	16	4
7	1	17	1
8	1	18	2
9	3	19	0
10	0	20	4

2. You are interested in estimating the selfing rate for an annual plant. You have a molecular marker with two alleles ('A' and 'a'). You have access to seeds that you can assume are in Hardy-Weinberg equilibrium (let's say that your seed bank was created by artificially pollinating plants selected at random from within a large population)¹. You grow the seeds in a field, allow the resulting plants to either self or outcross according to their own inclination, and then you genotype the offspring. The genotypic numbers that you observe in the offspring are $n_{AA} = 186$, $n_{Aa} = 228$, and $n_{aa} = 86$.

You would like to estimate the frequency of selfing in your experimental population (s). But you do not know the frequency of the A-allele (p_A).

Part 1: Write down the log likelihood as a function of the variables $n_{AA}, n_{Aa}, n_{aa}, s$, and p_A (don't put the actual numbers in, use the variables).

Hint: You should assume Hardy-Weinberg genotypic frequencies for the parental generation, but the offspring are only expected to be in Hardy-Weinberg genotypic frequencies if the selfing rate is 0.

Normally, we would try to find generic equations for the maximum likelihood estimates of both \hat{s} and \hat{p}_A . The math is very tedious in this case. So, I am going to tell you that the global value of $\hat{p}_A = 0.6$ for this example.

Part 2: Solve for the MLE of \hat{s} given this knowledge. Start by substituting the values from the real data and the value of $\hat{p}_A = 0.6$ into your equation for the log-likelihood. Then find the maximum likelihood value of \hat{s} . This will involve a bit of tedious math, and using the quadratic equation.

Suppose you are interested in testing the hypothesis that the selfing rate is 0.

Part 3: Solve for the MLE of p given the real data and the value of $s = 0$.

Part 4: Calculate the log likelihoods for the hypothesis that $s = 0$ and for the global maximum likelihood point.

Part 5: Can you reject a null hypothesis that $s = 0$?

¹Recall that Hardy-Weinberg equilibrium assumes random mating and predicts genotypic frequencies to be p^2 for the homozygote genotype of an allele that occurs at frequency p . Heterozygotes containing that allele should occur at frequency $2p(1 - p)$.