1. A population contains individuals with two different genotypes, 35% with GENOTYPE 1 and 65% with GENOTYPE 2. Individuals with GENOTYPE 1 have a Poisson number of offspring with mean $\mu = 2.3$ while individuals with GENOTYPE 2 are slightly more fit and have a Poisson number of offspring with mean $\mu = 3.1$. An individual is selected at random from the population. Given that this individual has four offspring, what is the probability that the selected individual has of GENOTYPE 1? *(Hint: Recall that for the Poisson distribution that $P(X = k) = \mathrm{e}^{-\mu}\mu^k/k!$.)*

Solution:

If we let $Y$ be the number of offspring, then the conditional distribution of $Y$ given the genotype of the individual is Poisson where the mean depends on the genotype. The information we are given is $\mathbb{P}(G_1) = 0.35$, $\mathbb{P}(G_2) = 0.65$, $\mathbb{P}(Y = 4 \mid G_1) = \mathrm{e}^{-2.3}(2.3)^4/4! \doteq 0.1169$, and $\mathbb{P}(Y = 4 \mid G_2) = \mathrm{e}^{-3.1}(3.1)^4/4! \doteq 0.1733$. We are asked to compute $\mathbb{P}(G_1 \mid Y = 4)$. Bayes' theorem is needed as we are trying to find a conditional probability that is the inverse of our given information. You can write a short probability tree to guide the calculation, but it suffices to simply recall the definition of conditional probability.

$$\begin{aligned} \mathbb{P}(G_1 \mid Y = 4) &= \frac{\mathbb{P}(G_1 \text{ AND } Y = 4)}{\mathbb{P}(Y = 4)} \\ &= \frac{(0.35)(0.1169)}{(0.35)(0.1169) + (0.65)(0.1733)} \\ &\doteq 0.266 \end{aligned}$$

2. Consider a simple gene $A$ with alleles $a_1$ and $a_2$. The three possible genotypes are $a_1a_1$, $a_1a_2$, and $a_2a_2$. The probability distribution of the genotype of a single offspring for some of the possible crosses are as follows.

| | Offspring Genotype | | |
|---|---|---|---|
| Parent Genotypes | $a_1a_1$ | $a_1a_2$ | $a_2a_2$ |
| $a_1a_1 \times a_1a_1$ | 1 | 0 | 0 |
| $a_1a_1 \times a_1a_2$ | 1/2 | 1/2 | 0 |

Suppose that there is a trait $T$ that is partially determined by the single gene $A$ and partially determined by environmental conditions. Specifically, suppose that the probability that an individual exhibits trait $T$ depends on the genotype as follows: $\mathbb{P}(T \mid a_1a_1) = 0.8$, $\mathbb{P}(T \mid a_1a_2) = 0.4$, and $\mathbb{P}(T \mid a_2a_2) = 0.1$. Furthermore, suppose that incidence of traits are conditionally independent given the genotypes among any group of individuals.

In a certain population, the proportions of individuals with genotypes $a_1a_1$ and $a_1a_2$ are respectively 0.6 and 0.4. Now consider a cross between PARENT 1, selected at random from the population, with PARENT 2, an individual with genotype $a_1a_1$.

(a) What is the probability that the first offspring of this cross has the trait $T$?

(b) Given that the first offspring has trait $T$, what is the probability that PARENT 1 has genotype $a_1a_1$?

(c) Given that the first offspring has trait $T$, what is the probability that the second offspring from a cross of the same parents also has trait $T$?

Solution: We begin again by creating a notation for the events and random variables in the problem, say with $P_1$, $O_1$, and $O_2$ representing Parent 1 and the first and second offspring.

(a) The first question is to find $\mathbb{P}(O_1 = T)$. To find this probability, we would want to condition on the offspring genotype, and to find this, we need to condition on the parent genotype. You can create a tree with three levels to describe this situation and to guide the calculations.

With equations, we are computing using the Law of Total Probability, or simply summing all of the ways that the first offspring can have the trait.

$$\begin{aligned}
\mathbb{P}(O_1 = T) &= \mathbb{P}(P_1 = a_1a_1 \text{ AND } O_1 = a_1a_1 \text{ AND } O_1 = T) \\
&\quad + \mathbb{P}(P_1 = a_1a_2 \text{ AND } O_1 = a_1a_1 \text{ AND } O_1 = T) \\
&\quad + \mathbb{P}(P_1 = a_1a_2 \text{ AND } O_1 = a_1a_2 \text{ AND } O_1 = T) \\
&= (0.6)(1)(0.8) + (0.4)(0.5)(0.8) + (0.4)(0.5)(0.4) \\
&= 0.72
\end{aligned}$$

(b) Now we are asked to find $\mathbb{P}(P_1 = a_1a_1 \mid O_1 = T)$. This is a Bayes' theorem problem and you can use the tree from the previous problem. With equations, we can just apply the rule for conditional probabilities.

$$\begin{aligned}
\mathbb{P}(P_1 = a_1a_1 \mid O_1 = T) &= \frac{\mathbb{P}(P_1 = a_1a_1 \text{ AND } O_1 = T)}{\mathbb{P}(O_1 = T)} \\
&= \frac{(0.6)(1)(0.8)}{0.72} \\
&= \frac{2}{3} \doteq 0.6667
\end{aligned}$$

(c) The last part is to find $\mathbb{P}(O_2 = T \mid O_1 = T)$. If these events were independent, which would be the case of the two offspring had different randomly chosen parents, the answer would just be 0.72 as in (a). However, they share the same randomly chosen parent and the given information that $O_1 = T$ affects the probability distribution of the genotype of $P_1$. In fact, the simplest solution is to use the new posterior probabilities for the genotype of $P_1$ from (b) and simply to compute the probability that the next offspring has the trait as follows.

$$\begin{aligned}
\mathbb{P}(O_2 = T \mid O_1 = T) &= \mathbb{P}(P_1 = a_1a_1 \text{ AND } O_2 = a_1a_1 \text{ AND } O_2 = T \mid O_1 = T) \\
&\quad + \mathbb{P}(P_1 = a_1a_2 \text{ AND } O_2 = a_1a_1 \text{ AND } O_2 = T \mid O_1 = T) \\
&\quad + \mathbb{P}(P_1 = a_1a_2 \text{ AND } O_2 = a_1a_2 \text{ AND } O_2 = T \mid O_2 = T) \\
&= \frac{2}{3}(1)(0.8) + \frac{1}{3}(0.5)(0.8) + \frac{1}{3}(0.5)(0.4) \\
&= \frac{1.6 + 0.40 + 0.20}{3} \\
&\doteq 0.7333
\end{aligned}$$

If you did not see this short cut, we can again use the rule for conditional probability as follows.

$$\mathbb{P}(O_2 = T \mid O_1 = T) = \frac{\mathbb{P}(O_1 = T \text{ AND } O_2 = T)}{\mathbb{P}(O_1 = T)}$$

The denominator is 0.72 from (a), so we need only compute the numerator. To do so, we need to condition on the genotype of $P_1$ as then the offspring are independent. If $P_1 = a_1 a_1$, then we have the following.

$$\mathbb{P}\left(O_1 = T \text{ AND } O_2 = T \mid P_1 = a_1 a_1\right) = (0.8)^2$$

When $P_1 = a_1 a_2$, the case is trickier as each offspring can have either genotype. In this case,

$$\mathbb{P}\left(O_1 = T \mid P_1 = a_1 a_2\right) = (0.5)(0.8) + (0.5)(0.4) = 0.6 \ ,$$

so

$$\mathbb{P}\left(O_1 = T \text{ AND } O_2 = T \mid P_1 = a_1 a_2\right) = (0.6)^2 \ .$$

When we put this together,

$$\begin{aligned}
\mathbb{P}\left(O_2 = T \mid O_1 = T\right) &= \frac{\mathbb{P}\left(O_1 = T \text{ AND } O_2 = T\right)}{\mathbb{P}\left(O_1 = T\right)} \\
&= \frac{(0.6)(0.8)^2 + (0.4)(0.6)^2}{0.72} \\
&\doteq 0.7333 \ .
\end{aligned}$$