Observing Gap Year Students from the POV of a Prospective Student

Ariana Chae and Louis Yi

Background

Each year, Olin Admissions allows admitted students to take a gap year before enrolling at Olin, and gives waitlisted students the opportunity to take a gap year for guaranteed admittance the following year. The number of students who take a gap year fluctuates each year. In the four classes currently enrolled at Olin, there are 43 gap year students. The class of 2015 has 17; the class of 2016 has 2; the class of 2017 has 6; and the class of 2018 has 18.

Scenario

Julie is a prospective student from Virginia and is visiting Olin for a few days. Her flight arrives in the morning and her first stop at Olin is the dining hall for breakfast. As she eats meals with other students and attends classes, she meets students from all different graduating classes. As she meets more students, she begins to learn who is a gap year student and who is not, and she develops a hypothesis as to the percentage and number of gap year students at Olin. How correct are her hypotheses?

Implementation and Results

First, we look at Julie's perception of the percentage of gap year students there are at Olin. Every time Julie meets a new Olin student, she learns whether they are a gap year student or not, and her knowledge of the ratio between gap year students and non-gap year students at Olin shifts. Assuming that there is an equal probability of meeting any current student in the school, and that it is not possible to meet someone more than once, we update the suite given how many gap year and non-gap year students Julie has met. In this case, Julie met 4 gap year students and 20 non-gap year students.

```
def RunUpdate(suite, gapYear=4, nonGapYear=20):
    """Updates the Suite with the given number of gap year and non-gap year
students.

    suite: Suite object
    gapYear: int
    nonGapYear: int
    """
    dataset = 'G' * gapYear + 'N' * nonGapYear

    for data in dataset:
        suite.Update(data)
```

In the likelihood function for this suite, the probability of the student being a gap year student is x/100 and the probability of them not being a gap year student is 1 − x/100.

```
def Likelihood(self, data, hypo):
    """Computes the likelihood of the data under the hypothesis.
    """
    x = hypo / 100.0
    if data == 'G':
        return x
    else:
        return 1-x
```

If Julie has already met 20 non-gap year students and 4 gap year students, she would estimate the probability of gap year students at Olin to be 19.23 percent, based on the peak of the curve of the Pdf in Figure 1.

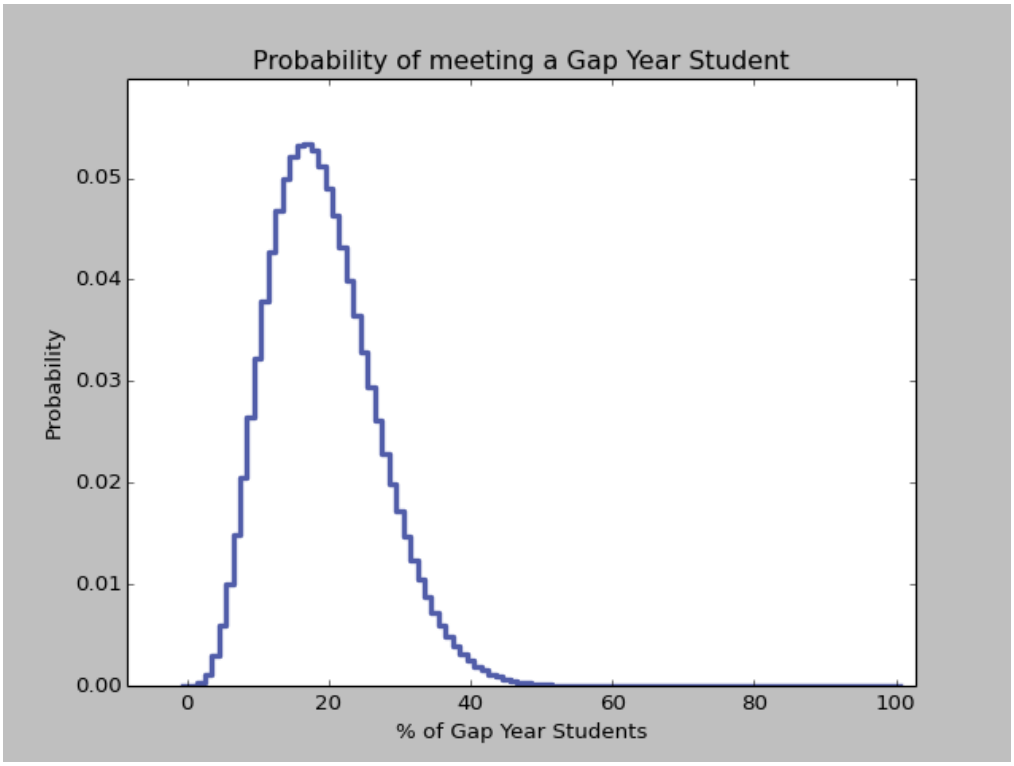

Figure 1: This figure shows the probability of there being a certain percentage of gap year students at Olin.

At the end of her visit, Julie thinks back on everyone she met at Olin. Based on the Olin students she met during her three day visit, she thinks she has an idea of how many gap year students there might be at Olin.

The likelihood function to determine the number of gap year students at Olin uses the EvalBinomialPmf method where *p* is estimated as the number of gap year students Julie met divided by the hypothesis of how many gap year students there are at Olin.

```
def Likelihood(self, data, hypo):

    g, n, o = data
    if hypo < g + n - o:
        return 0

    p = g / hypo
    like = thinkbayes2.EvalBinomialPmf(o, n, p)
    return like
```

As shown in Figure 2, Julie infers that there are a different number of gap year students at the end of each day at Olin. As the length of her visit increases, she becomes more confident of her estimation of the number of gap year students, and the lower bound for her estimation increases as she meets more gap year students. After two days, Julie believes there are 39 gap year students at Olin, based on the Pdf in Figure 2.

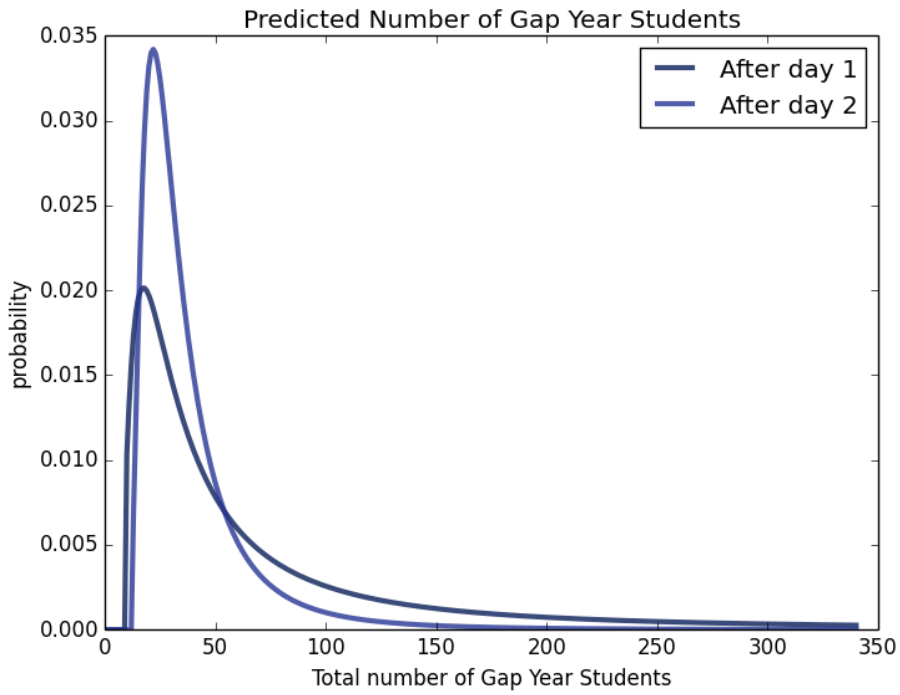

Figure 2: This figure shows the predicted number of gap year students based on a prospective student's observations over the course of a 2 day visit

Summary
After meeting 20 non-gap year students and 4 gap year students, Julie thinks that there is a 19.23 percent chance of meeting another gap year. The actual percentage of gap year students at Olin is around 13 percent. At the end of her visit, Julie thinks there are a total of 39 gap year students at Olin, whereas the actual number of gap year students is 43. From these predictions we can see that the estimated and actual values do not always match up. In order for Julie to get a better assessment of the percentage and number of gap year students, she would need a larger sample size from Olin.