Conditional probability

There are two ways that we can calculate transition probabilities in the seed bank with data from seed burial experiments and viability tests. Both take into account the fact that outcomes in these experiments are conditional on previous outcomes. For example, seeds that germinate in the first January must first have remained intact and viable.

One approach to this problem is to change the denominator for all of our calculations. In this case, the rate is calculated based on the subset of possible seeds. For example, the number of seeds that survive to January is the sum of seeds that germinate and seeds that are estimated to be viable. This value is the numerator in calculating s_1, and then becomes the denominator in calculating g_1. This method states that the number of seeds that are viable is known with the same amount of certainty as the number of seeds that germinate. In fact, our estimate of viability in January probably has a greater amount of uncertainty.

A second approach to this problem is to include uncertainty about estimates in all of the calculations. We can do this by calculating the transition probabilities as conditional probabilities.

Below, I calculate each transition probability using both methods. The figure below shows the results of a simulation study indicating that the two approaches are directly comparable.
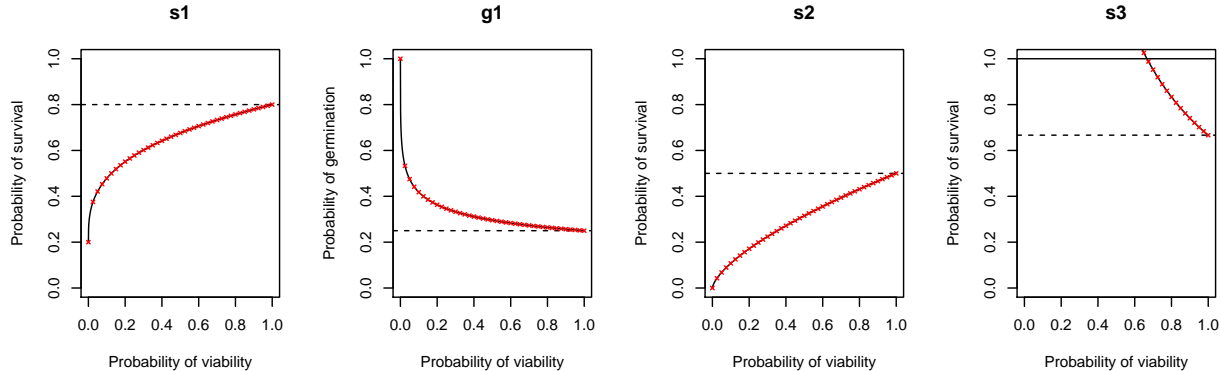

Figure 1: (A) This plot compares estimated probabilities for the following values (total = 100, yjan = 60, ygerm = 20, yoct = 50) for a range of possible viability. The line is the changing denominator approach, the red crosses are the conditional probability approach. The estimate for s3 is for $\nu_2 = 1$.

s1

We define s_1 as the probability of a seed being intact and viable from October in year t to January in year $t+1$. We want this estimate to include seeds that germinated and seeds that were intact and viable. To obtain this estimate, we have data on the following quantities:

- n^{total} = the number of seeds that started the seed bag experiments

- y^{germ} = the number of seedlings in January

- y^{jan1} = the number of intact seeds in January of year $t+1$

- ν = the probability that a seed that is intact in October of year $t+1$ is viable (estimated elsewhere)

We can use this information to calculate the s_1 in two ways. First, we can estimate the fraction of seeds that are intact and viable in January as $y^{jan1} \times \nu^{1/3}$. We then calculate

$$s_1 = \frac{y^{jan1} \times \nu^{1/3} + y^{germ}}{n^{total}} \tag{1}$$

Alternatively, we can compose an estimate for s_1 using two conditional probabilities. First, we estimate the probability of a seed being intact to January as $p^{intact} = (y^{germ} + y^{jan1})/n^{total}$. Second, we estimate the probability of a plant emerging, conditional on being intact as $p^{germ} = y^{germ}/(y^{germ} + y^{jan1})$. We then calculate

$$s_1 = p^{intact} \times (p^{germ} + (1 - p^{germ}) \times \nu^{1/3}) \tag{2}$$

g1

We define g_1 as the probability of germination for a seed that is intact and viable. To obtain this estimate, we have data on the following quantities:

- n^{total} = the number of seeds that started the seed bag experiments

- y^{germ} = the number of seedlings in January

- y^{jan1} = the number of intact seeds in January of year $t+1$

- ν = the probability that a seed that is intact in October of year $t+1$ is viable (estimated elsewhere)

We can use this information to calculate the g_1 in two ways. First, we can estimate the fraction of seeds that are intact and viable in January as $y^{jan1} \times \nu^{1/3}$. We then calculate

$$g_1 = \frac{y^{germ}}{y^{jan1} \times \nu^{1/3} + y^{germ}} \tag{3}$$

Alternatively, we can compose an estimate for g_1 using conditional probabilities. First, we estimate the probability of a plant emerging, conditional on being intact as $p^{germ} = y^{germ}/(y^{germ} + y^{jan1})$. We then calculate

$$g_1 = \frac{p^{germ}}{1 - (1 - \nu^{1/3}) \times (1 - p^{germ})} \tag{4}$$

s2

We define s_2 as the probability of survival from January to October in year t for a seed that was intact and viable in January. To obtain this estimate, we have data on the following quantities:

- n^{total} = the number of seeds that started the seed bag experiments

- y^{germ} = the number of seedlings in January

- y^{jan1} = the number of intact seeds in January of year $t + 1$

- y^{oct1} = the number of intact seeds in January of year $t + 1$

- ν = the probability that a seed that is intact in October of year $t+1$ is viable (estimated elsewhere)

We can use this information to calculate the s_2 in two ways. First, we can estimate the fraction of seeds that are intact and viable in January as $y^{jan1} \times \nu^{1/3}$. Second, we can estimate the fraction of seeds that are intact and viable in October as $y^{oct1} \times \nu$ We then calculate

$$s_2 = \frac{y^{oct1} \times \nu}{y^{jan1} \times \nu^{1/3}} \tag{5}$$

Alternatively, we can compose an estimate for s_2 using conditional probabilities. First, we estimate the probability of a seed being intact in October, conditional on being intact in January $p = y^{oct1}/y^{jan1}$. We then calculate

$$s_2 = p \times \nu^{2/3} \tag{6}$$

s3

We define s_3 as the probability of survival from October in year t to January in year $t + 1$ for a seed that was intact and viable in October. To obtain this estimate, we have data on the following quantities:

- n^{total} = the number of seeds that started the seed bag experiments

- y^{germ} = the number of seedlings in January

- y^{jan1} = the number of intact seeds in January of year $t + 1$

- y^{oct1} = the number of intact seeds in January of year $t + 1$

- y^{germ2} = the number of seedlings in January year 2

- y^{jan2} = the number of intact seeds in January of year 2

- ν = the probability that a seed that is intact in October of year $t+1$ is viable (estimated elsewhere)

- ν_2 = the probability that a seed that is intact in October of year $t+2$ is viable (estimated elsewhere)

We can use this information to calculate the s_3 in two ways. First, we can estimate the fraction of seeds that are intact and viable in October as $y^{oct1} \times \nu$. Second, we can estimate the fraction of seeds that are intact and viable in January as $y^{jan2} \times \nu_2^{1/3}$. We then calculate

$$s_3 = \frac{y^{jan2} \times \nu_2^{1/3} + y^{germ2}}{y^{oct1} \times \nu} \tag{7}$$

Alternatively, we can compose an estimate for s_3 using conditional probabilities. First, we make use the probabilities for s_1, g_1 and s_2 (described above) to normalize the event space. Second, we estimate the probability of a seed being intact to January as $p^{intact} = (y^{germ2} + y^{jan2})/n^{total}$. Finally, we estimate the probability of a plant emerging, conditional on being intact as $p^{germ2} = y^{germ2}/(y^{germ2} + y^{jan2})$. We then calculate

$$s_3 = \frac{p^{intact} \times (p^{germ2} + (1 - p^{germ2}) \times \nu_2^{1/3})}{s_1 \times (1 - g_1) \times s_2} \tag{8}$$

Adjusting viability in year 2

For the 2011 paper, the following was applied. If viability estimated in the second year was less than viability estimated in the first year, the probability of viability in January of year 2 was interpolated as $\nu^{2/3} \times \nu_2^{1/3}$. If viability estimated in the second year was equal to or greater than viability estimated in the first year, the probability of viability in January of year 2 was estimated as $\nu_2^{1/3}$.

By calculating viability using the approach described in this appendix, I did not encounter a situation where the estimate was greater than one. I thus always calculated the probability of viability in January of year 2 was interpolated as $\nu^{2/3} \times \nu_2^{1/3}$.