

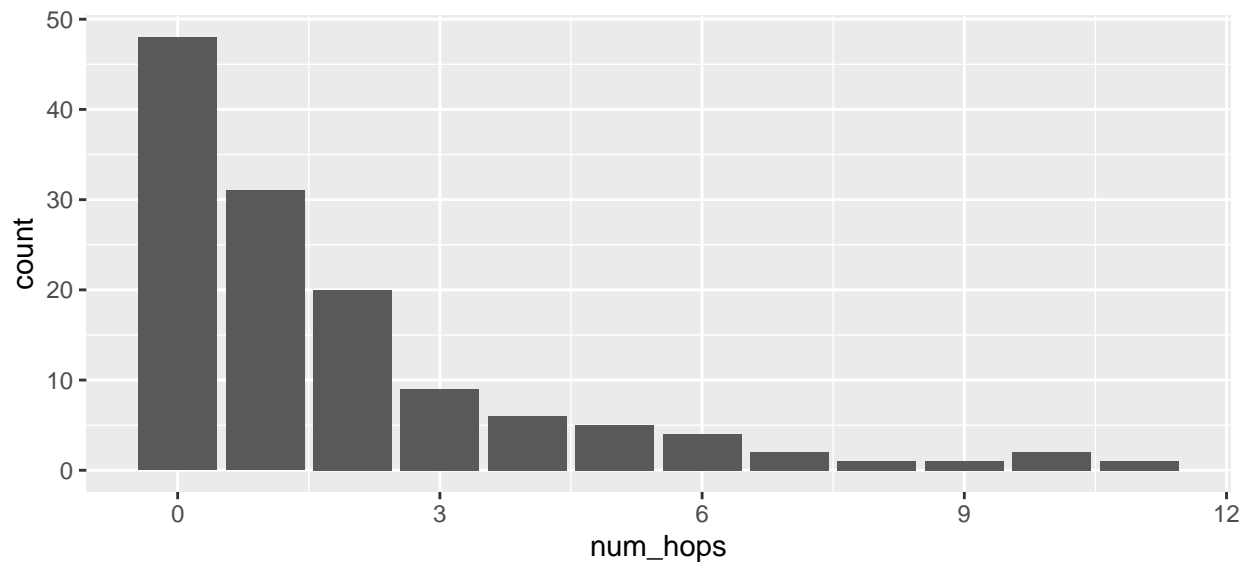
Stat 343 MLE Practice: Saplings

Bird Hops (Adapted from Rice problem 8.8)

In an ecological study of the feeding behavior of birds, researchers counted the number of hops each of several birds took between flights. The R code below reads in the data and makes an initial plot:

```
bird_hops <- read_csv("http://www.evanlray.com/data/rice/Chapter%208/hops.csv") %>%
  mutate(
    num_hops = num_hops - 1
  )

ggplot(data = bird_hops, mapping = aes(x = num_hops)) +
  geom_bar()
```



```
bird_hops %>%
  count(num_hops)
```

```
## # A tibble: 12 x 2
##   num_hops     n
##   <dbl> <int>
## 1      0     48
## 2      1     31
## 3      2     20
## 4      3      9
## 5      4      6
## 6      5      5
## 7      6      4
## 8      7      2
## 9      8      1
## 10     9      1
## 11    10      2
## 12    11      1
```

After subtracting 1 from the originally reported counts (as is done in the code above), the number of hops is 0 if the bird took off directly (0 hops before flight), 1 if the bird took one hop before taking off, and so on. The researchers wanted to learn about the distribution of the number of hops taken before flight by this species.

Let X_i be the number of hops taken before flight by bird number i .

(1) What distribution might you use to model X_i ?

$X_i \sim \text{Geometric}(p)$, because we have a count of the number of “failures” (not taking off) before the first “success” (taking off). With the definition of the Geometric distribution we are using in this class, if $X \sim \text{Geometric}(p)$, then X represents the number of failures before the first success in a sequence of independent identically distributed Bernoulli trials each with probability of success p . This distribution is supported on the set $\{0, 1, 2, 3, \dots\}$.

(2) Find the likelihood function

The density function for each X_i is $f(x_i|p) = (1-p)^{x_i}p$. The likelihood function is given by:

$$\begin{aligned}\mathcal{L}(p|x_1, \dots, x_n) &= f(x_1, \dots, x_n|p) \\ &= \prod_{i=1}^n f(x_i|p) \\ &= \prod_{i=1}^n (1-p)^{x_i} p \\ &= (1-p)^{\sum_{i=1}^n x_i} p^n\end{aligned}$$

(3) Find the log-likelihood function

$$\begin{aligned}\ell(p|x_1, \dots, x_n) &= \log\{\mathcal{L}(p|x_1, \dots, x_n)\} \\ &= \log\{(1-p)^{\sum_{i=1}^n x_i} p^n\} \\ &= \left(\sum_{i=1}^n x_i\right) \log(1-p) + n \log(p)\end{aligned}$$

(4) Find a critical point of the log-likelihood function

The first derivative of the log-likelihood function is:

$$\begin{aligned}\frac{d}{dp} \ell(p|x_1, \dots, x_n) &= \frac{d}{dp} \left\{ \left(\sum_{i=1}^n x_i\right) \log(1-p) + n \log(p) \right\} \\ &= -\frac{\sum_{i=1}^n x_i}{1-p} + \frac{n}{p}\end{aligned}$$

Setting equal to 0, we obtain:

$$\begin{aligned}\frac{\sum_{i=1}^n x_i}{1-p} &= \frac{n}{p} \\ \Rightarrow p \sum_{i=1}^n x_i &= n - np \\ \Rightarrow p \left(\sum_{i=1}^n x_i + n\right) &= n \\ \Rightarrow p &= \frac{n}{\sum_{i=1}^n x_i + n} = \frac{1}{\bar{x} + 1}\end{aligned}$$

(5) Verify that the critical point you found above gives a maximum of the log-likelihood.

The second derivative of the log-likelihood function is:

$$\begin{aligned}\frac{d^2}{dp^2}L(p|x_1, \dots, x_n) &= \frac{d}{dp} \left\{ -\frac{\sum_{i=1}^n x_i}{1-p} + \frac{n}{p} \right\} \\ &= -\frac{\sum_{i=1}^n x_i}{(1-p)^2}(-1)(-1) - \frac{n}{p^2} \\ &= -\frac{\sum_{i=1}^n x_i}{(1-p)^2} - \frac{n}{p^2} < 0\end{aligned}$$

Since the second derivative is negative, the critical value identified above is a maximum of the likelihood function for a particular sample with observed values x_1, \dots, x_n .

(6) Write down your final answer for what the maximum likelihood estimator is.

The maximum likelihood estimator is a random variable, which we express by replacing the lower case x_1, \dots, x_n in the critical point from part (4) with capital X_1, \dots, X_n .

$$\hat{p}_{MLE} = \frac{n}{\sum_{i=1}^n X_i + n} = \frac{1}{\bar{X} + 1}$$