05: MAXIMUM LIKELIHOOD ESTIMATION

Stat250-S25

---

1 Warm up

Example: You are testing seeds from a new plant variety. You plant 10 seeds (n=10) and observe that 3 of them successfully germinate (k=3). Consider two hypotheses about the true germination probability (p) for this variety:

Hypothesis A:

Hypothesis B:

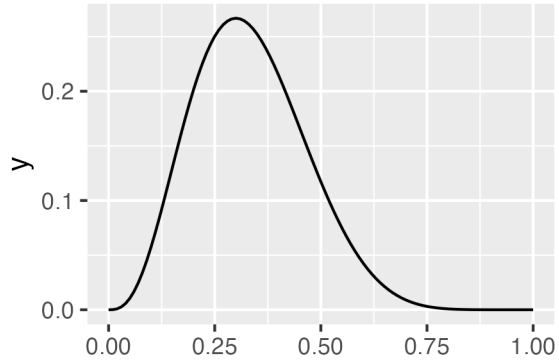
Given you observed 3 germinations out of 10 seeds, which hypothesis (A or B) is more supported by the data?

In general, the *likelihood* of seeing $X = 3$ given p is:

3 ways of finding the maximum:

1. Approximate solution graphically

```
germ_function <- function(p) {choose(10,3) * p^3 * (1-p)^7} # define function
ggplot() +
  geom_function(fun = germ_function) +
  xlim(c(0,1))
```



2. Find a numerical approximation

```
optimize(germ_function, interval = c(0,1), maximum = TRUE)
```

```
$maximum
[1] 0.3000157

$objective
[1] 0.2668279
```

3. Use calculus to find an exact maximum

Can we generalize to *any* data?

2 Definitions

Likelihood function

Let $f(x;\theta)$ denote the probability mass function for a discrete distribution with associated parameter θ. Suppose $X_1, X_2, ..., X_n$ are a random sample from this distribution and $x_1, x_2, ..., x_n$ are the actual observed values. Then, the *likelihood function* of θ is:

Maximum likelihood estimate

A maximum likelihood estimate (MLE), $\hat{\theta}_{MLE}$ is the value of θ that maximizes the likelihood function, or equivalently, that maximizes the log-likelihood $l(\theta) = \ln L(\theta)$

Bernoulli Data Example

Estimator

Estimate

3 Exercise

Let $X_1, ..., X_n$ be an iid random sample from a distribution with PDF

$$f(x|\theta) = (\theta + 1)x^{\theta}, 0 \leq x \leq 1$$

1. Find the maximum likelihood estimator for a random sample of size n using calculus.
2. Suppose we observe a sample of size 5: {.83, .49, .72, .57, .66}. Find the maximum likelihood estimate and verify with a graph or numerical approximation