Proportion Inference

Modeling a count of successes

Download the section 11.Rmd handout to `STAT240/lecture/sect11-single-proportion`.

Download the file `chimpanzee.csv` to `STAT240/data`.

Material in this section is covered by Chapter 12 on the notes website.

A chimpanzee can choose to act selfishly, or socially.

Modeling a binary population:

- Based on the binomial + normal
- p is the unknown parameter of interest

The chimpanzee data is based on an Emory University experiment.

- Chimpanzees choose from colored tokens
- One color is **selfish**, the other is **prosocial**
- Different combinations were studied

p is the probability of a prosocial choice.

- Is $p_{partner}$ greater than 0.5?
- Is $p_{partner}$ the same as $p_{nopartner}$?

Let X be the count of prosocial choices and $\hat{p}$ be the proportion of prosocial choices out of n.

$$X \sim Binom(n, p), \qquad \hat{p} = \frac{X}{n}$$

which relies on the BINS assumptions. n is known and p is the parameter of interest.

Chimpanzee A had 90 trials with a partner. Let's suppose $p = 0.6$.

What would we observe for $\hat{p}$? Let's try a simulation.

Recall that the normal is a good approximation to the binomial when $np(1-p)$ is large.

$$X \dot{\sim} N(np,\ \sqrt{np(1-p)})$$

$$\hat{p} \dot{\sim} N\left(p,\ \sqrt{\frac{p(1-p)}{n}}\right)$$

$$\hat{p} \dot{\sim} N\left(p,\ \sqrt{\frac{p(1-p)}{n}}\right)$$

- $E(\hat{p}) = p$
- Error decreases with n
- Error is a function of p

Inference for p is based on the binomial and normal.

- CI critical values
- Null distribution for a hypothesis test

To use the normal, we like to have $np(1-p) \geq 10$.

Back to chimpanzee A. We want to study $p_{A,partner}$.

We pretended $p_{A,partner} = 0.6$ and simulated different values of $\hat{p}$.

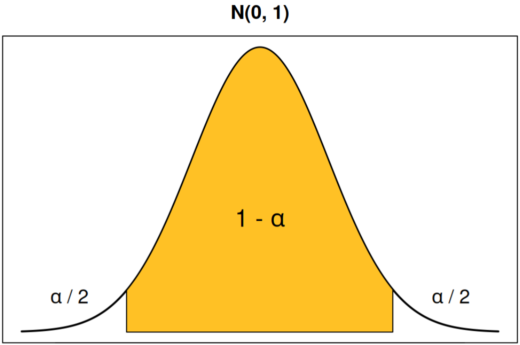
In practice, we don't have $p_{A,partner}$, just a single $\hat{p} = 60/90$.

CI formula:

$$\text{point estimate} \;\pm\; \text{critical value} \;\times\; \text{standard error}$$

- For p, the point estimate is $\hat{p}$
- The standard error is $\sqrt{\frac{p(1-p)}{n}}$
- The critical value is from $N(0,1)$

Find a critical value with `qnorm()`.



For a 90% CI, use 1.645.

Problem: don't have $\sqrt{\frac{p(1-p)}{n}}$. Use $\hat{p}$ instead:

$$\hat{se}(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

This is the **Wald** adjustment.

Wald CI for p:

$$\hat{p} \ \pm \ z_{\alpha/2} \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

We are 90% confident that $p_{A,partner}$ is within (0.585, 0.748).

Another adjustment: add 2 successes and failures to "stabilize" data.

$$n_{AC} = n + 4, \qquad \hat{p}_{AC} = \frac{X + 2}{n + 4}$$

$$\hat{se}(\hat{p}_{AC}) = \sqrt{\frac{\hat{p}_{AC}(1 - \hat{p}_{AC})}{n_{AC}}}$$

Agresti-Coull adjusted CI for p:

$$\hat{p}_{AC} \pm z_{\alpha/2} \times \sqrt{\frac{\hat{p}_{AC}(1-\hat{p}_{AC})}{n_{AC}}}$$

We are 90% confident that $p_{A,partner}$ is within (0.579, 0.74).

AC is usually preferable to Wald.

- Both methods are approximate
- AC tends to lead to intervals with coverage probability greater than $1 - \alpha$.

Now, let's look at a hypothesis test approach for p.
Two ways:

- Binomial null
- Normal null (Z test)

Both use BINS, and Z test also depends on n and p.

Model the count of prosocial choices made by chimpanzee A (with a partner) as

$$Binom(90, p)$$

Is chimpanzee A more prosocial than selfish?

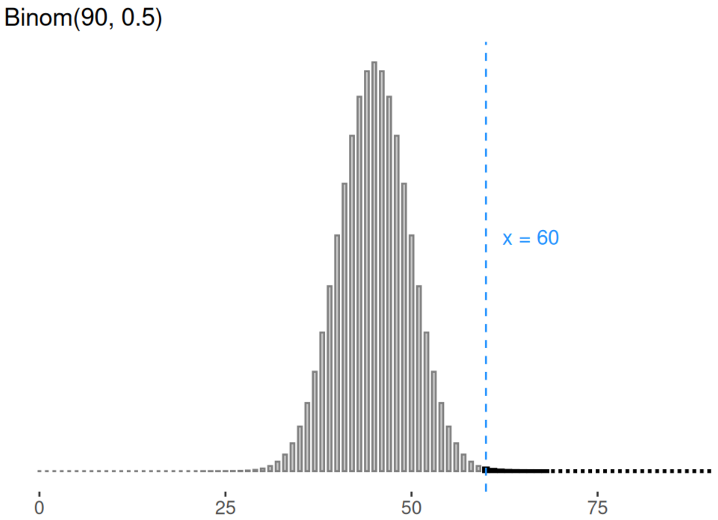
$$H_0 : p = 0.5 \quad \text{versus} \quad H_A : p > 0.5$$

Let's use $\alpha = 0.1$.

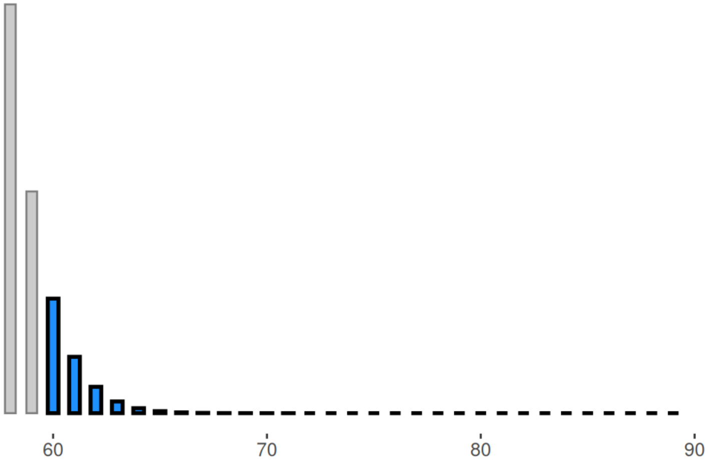
Our test statistic is X has null distribution

$$Binom(90, 0.5)$$

and an observed test statistic $x_{obs} = 60$. Values of X higher than 60 give more evidence for H_A.


Binom(90, 0.5)
x = 60
0
25
50
75

zoom... and enhance!



The p-value is the probability above (and including) $x_{obs} = 60$ on our null distribution.

This is 0.001, which is strong evidence against the null. Chimpanzee A appears to be prosocial more than half the time.

The average prosocial rate with a partner is 0.59.
Chimpanzee F has a rate of $47/90$.

Is this significantly *less* than 0.59 with $\alpha = 0.05$?

- Set up hypotheses
- Identify null distribution
- Use test statistic $x_{obs} = 47$
- Calculate p-value in the *lower* direction

The direction of p-value calculation depends on H_A.

What if H_A is symmetric?

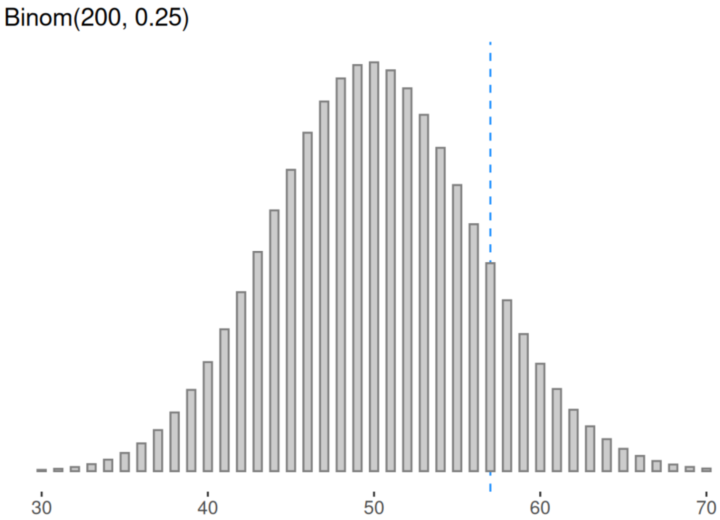
A psychic claims to be able to guess the suit of a random card without looking. 200 cards were drawn, and they guessed correctly 57 times. Model:

$$X \sim Binom(200, p)$$

The problem suggests one-sided hypotheses:

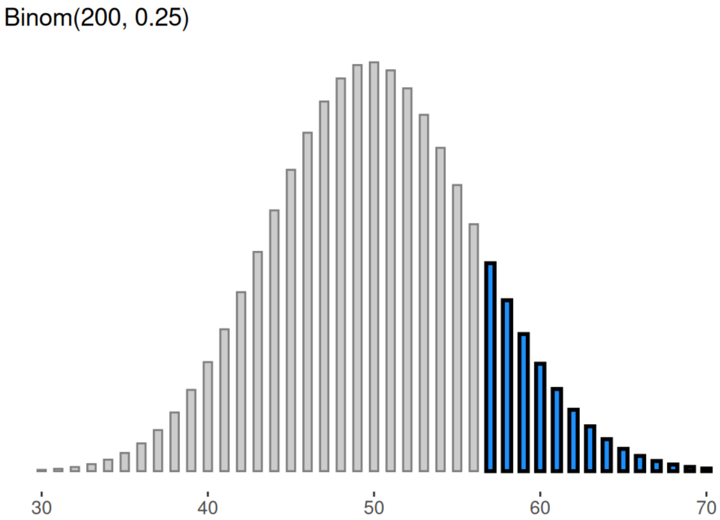
$$H_0 : p = 0.25 \quad \text{versus} \quad H_A : p > 0.25$$

Under the null distribution, which outcomes are less likely than 57?


Binom(200, 0.25)
30
40
50
60
70

The outcomes as or less likely than x_{obs} are $X \leq 42$ and $X \geq 57$.

Here, only $X \geq 57$ is relevant. We get a p-value of 0.145 and fail to reject H_0 at the 5% level.

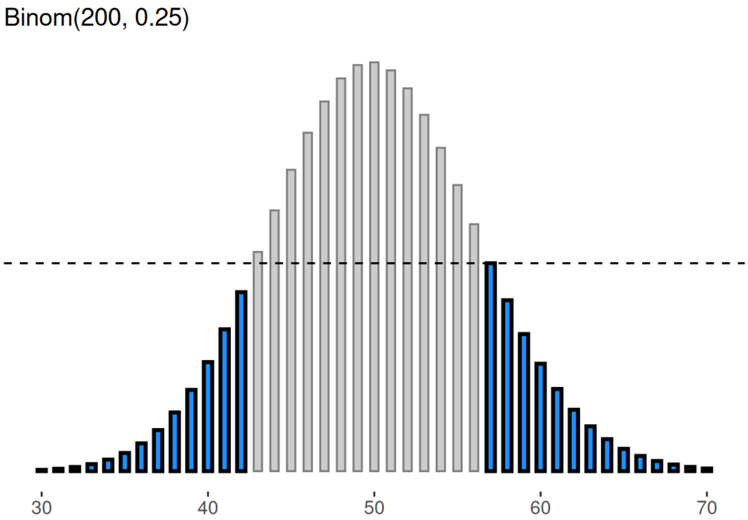

Binom(200, 0.25)
30
40
50
60
70

What if we had two-sided hypotheses?

$$H_0 : p = 0.25 \quad \text{versus} \quad H_A : p \neq 0.25$$

Now, we have to look for p being different from 0.25 in *both* directions.

Our p-value is $P(X \leq 42) + P(X \geq 57) = 0.253$.


Binom(200, 0.25)
30
40
50
60
70

In general, for different hypothesis directions:

- Test statistic does not change
- Null distribution does not change

- p-value changes
- Conclusions change

An alternative to the binomial test is a **Z test**. If the null hypothesis were true,

$$\hat{p} \mathrel{\dot\sim} N\left(p,\ \sqrt{\frac{p(1-p)}{n}}\right)$$

$$\hat{p} \mathrel{\dot\sim} N\left(0.25,\ \sqrt{\frac{0.25(0.75)}{200}}\right)$$

By standardization,

$$Z = \frac{\hat{p} - 0.25}{\sqrt{\frac{0.25(0.75)}{200}}} \;\dot{\sim}\; N(0,1)$$

Our test statistic is the observed Z. and our null is $N(0,1)$. This works because of n and p.

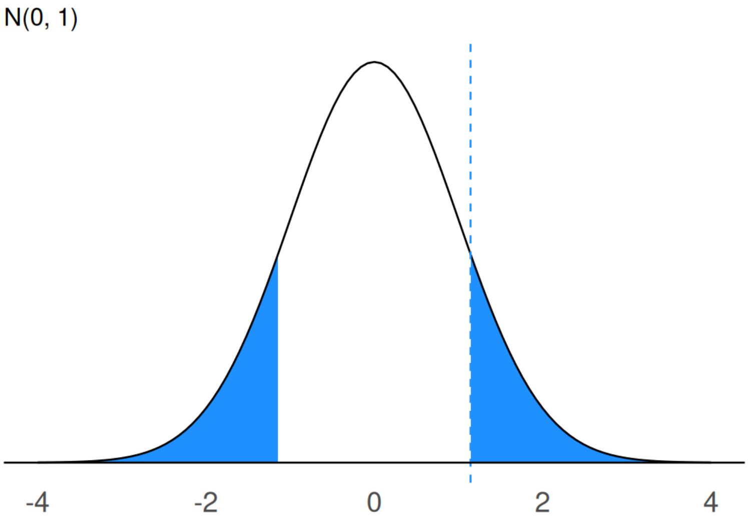
$$H_0 : p = 0.25 \quad \text{versus} \quad H_A : p \neq 0.25$$

$$z_{obs} = \frac{\frac{57}{200} - 0.25}{\sqrt{\frac{0.25(0.75)}{200}}} = 1.143$$

Is 1.143 consistent with $N(0, 1)$? For a two-sided test, we need the area outside of 1.143 in *both tails*.

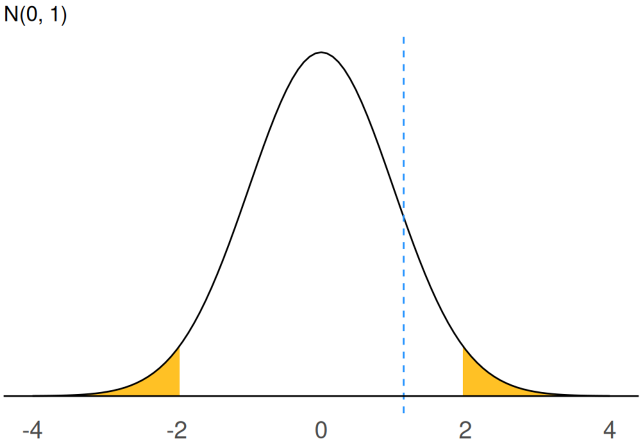
$$P(Z \leq -1.143) \; + \; P(Z \geq 1.143)$$

The p-value is 0.253, just like the exact test.


N(0, 1)
-4
-2
0
2
4

Connections to CI: Rejection region

Draw an area of size α in the tails.



Our test statistic is not in the rejection region.