

I06a - Hypothesis tests

with binomial example

STAT 5870 (Engineering)
Iowa State University

August 28, 2024

Statistical hypothesis testing

A **hypothesis test** consists of two hypotheses,

- null hypothesis (H_0) and
- an alternative hypothesis (H_A),

which make claims about parameter(s) in a model, and a decision to either

- reject the null hypothesis or
- fail to reject the null hypothesis.

Binomial model

If $Y \sim \text{Bin}(n, \theta)$, then some hypothesis tests are

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_A : \theta \neq \theta_0$$

or

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_A : \theta > \theta_0$$

or

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_A : \theta < \theta_0$$

Small data

Let $Y \sim \text{Bin}(n, \theta)$ with

$$H_0 : \theta = 0.5 \quad \text{versus} \quad H_A : \theta \neq 0.5.$$

You collect data and observe $y = 6$ out of $n = 13$ attempts. Should you reject H_0 ? Probably not since $6 \approx E[Y] = 6.5$ if H_0 is true.

What if you observed $y = 2$? Well, $P(Y = 2) \approx 0.01$.

Large data

Let $Y \sim \text{Bin}(n, \theta)$ with

$$H_0 : \theta = 0.5 \quad \text{versus} \quad H_A : \theta \neq 0.5.$$

You collect data and observe $y = 6500$ out of $n = 13000$ attempts. Should you reject H_0 ?
Probably not since $6500 = E[Y]$ if H_0 is true. But $P(Y = 6500) \approx 0.007$.

p-values

p-value: the probability of observing a **test** statistic as or more extreme than observed if the **null hypothesis** is true

The **as or more extreme** region is determined by the alternative hypothesis.

For example, if $Y \sim \text{Bin}(n, \theta)$ and $H_0 : \theta = \theta_0$ then

$$H_A : \theta < \theta_0 \implies Y \leq y$$

or

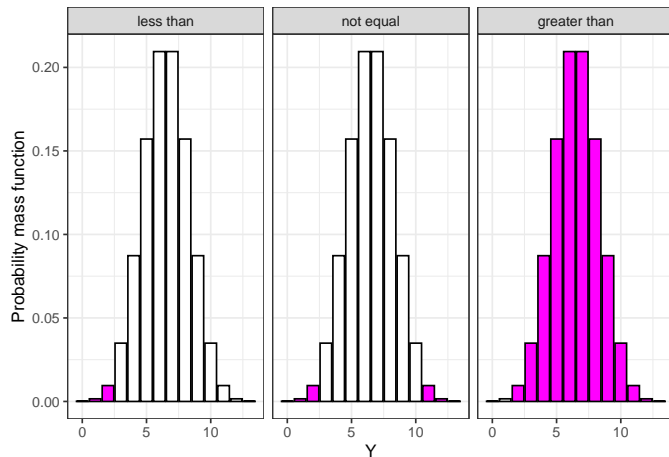
$$H_A : \theta > \theta_0 \implies Y \geq y$$

or

$$H_A : \theta \neq \theta_0 \implies |Y - n\theta_0| \geq |y - n\theta_0|.$$

as or more extreme regions

As or more extreme regions for $Y \sim \text{Bin}(13, 0.5)$ with $y = 2$



R “hand” calculation

$$H_A : \theta < 0.5 \implies p\text{-value} = P(Y \leq y)$$

```
pbinom(y, size = n, prob = theta0)
```

```
[1] 0.01123047
```

$$H_A : \theta > 0.5 \implies p\text{-value} = P(Y \geq y) = 1 - P(Y \leq y - 1)$$

```
1-pbinom(y-1, size = n, prob = theta0)
```

```
[1] 0.998291
```

$$H_A : \theta \neq 0.5 \implies p\text{-value} = P(|Y - n\theta_0| \leq |y - n\theta_0|)$$

```
2*pbinom(y, size = n, prob = theta0)
```

```
[1] 0.02246094
```


R Calculation

$$H_A : \theta < 0.5$$

```
binom.test(y, n, p = theta0, alternative = "less")$p.value
```

```
[1] 0.01123047
```

$$H_A : \theta > 0.5$$

```
binom.test(y, n, p = theta0, alternative = "greater")$p.value
```

```
[1] 0.998291
```

$$H_A : \theta \neq 0.5$$

```
binom.test(y, n, p = theta0, alternative = "two.sided")$p.value
```

```
[1] 0.02246094
```

Significance level

Make a decision to either

- reject the null hypothesis or
- fail to reject the null hypothesis.

Select a **significance level** α and

- reject if $p\text{-value} < \alpha$ otherwise
- fail to reject.

Decisions

Decision	Truth	
	H_0 true	H_0 not true
reject H_0	type I error	correct
fail to reject H_0	correct	type II error

Then

significance level α is $P(\text{reject } H_0 | H_0 \text{ true})$

and

power is $P(\text{reject } H_0 | H_0 \text{ not true})$.

Interpretation

The null hypothesis is a model. For example,

$$H_0 : Y \sim \text{Bin}(n, \theta_0)$$

if we **reject H_0** , then we are saying the **data are incompatible with this model**.

Recall that $Y = \sum_{i=1}^n X_i$ for $X_i \overset{\text{ind}}{\sim} \text{Ber}(\theta)$.

So, possibly

- the X_i are not independent or
- they don't have a common θ or
- $\theta \neq \theta_0$ or
- you just got unlucky.

If we **fail to reject H_0** , insufficient evidence to say that the data are incompatible with this model.

Die tossing example

You are playing a game of Dragonwood and a friend rolled a four 3 times in 6 attempts. Did your friend (somehow) increase the probability of rolling a 4?

Let Y be the number of fours rolled and assume $Y \sim \text{Bin}(6, \theta)$. You observed $y = 3$ and are testing

$$H_0 : \theta = \frac{1}{6} \quad \text{versus} \quad H_A : \theta > \frac{1}{6}.$$

```
binom.test(3, 6, p = 1/6, alternative = "greater")$p.value
```

```
[1] 0.06228567
```

With a significance level of $\alpha = 0.05$, you fail to reject the null hypothesis.

Summary

- Hypothesis tests:

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_A : \theta \neq \theta_0$$

- Use p -values to determine whether to
 - reject the null hypothesis or
 - fail to reject the null hypothesis.
- More assessment is required to determine if other model assumptions hold.