Lecture 26

Hypothesis Testing

Manju M. Johny

STAT 330 - Iowa State University

Hypothesis Testing

Definition:
A statistical *hypothesis* is a statement about a parameter θ

There are 2 competing hypotheses in a testing problem:

- *Null Hypothesis (H_0):* the default/pre-data view about the parameter.
- *Alternative Hypothesis (H_A):* usually what you want your data/study to show.

Note: H_0 and H_A have to be disjoint. There can not be any outcomes in common between the null and alternative hypotheses.

Motivating Example

Example 1: I have a coin and I'm interested in the probability of flipping a "head". I flip a coin 100 times and record the number of heads obtained.

$$X = \# \text{ of heads}$$
$$X \sim Bin(n = 100, p)$$

where $p = P(\text{"}heads\text{"})$ is unknown

By default, we assume coin is fair $p = 0.5$ (null hypothesis).

Alternative hypothesis should contradict the null hypothesis.

Hypotheses:

- $H_0 : p = 0.5$ (coin is fair)
- $H_A : p \neq 0.5$ (coin is unfair)

Motivating Example Continued

Data: Out of 100 flips, I get 71 heads. $\hat{p} = 0.71$

Idea of Hypothesis Testing:

- Assume H_0 (our default belief) is true until our *data* tells us otherwise.
- Ask ourselves "what is the probability of getting 71 heads if the null hypothesis is true (coin is fair)?"

  $\rightarrow$ probability $= 0.000032$ (called the "$p - value$")
- There is a 0.000032 probability that we observed our data if the null hypothesis that the coin is fair is true.

  $\rightarrow$ Now we have evidence against the null hypothesis (that coin is fair), and in favor of the alternative hypothesis (that coin is unfair).

General Hypothesis Testing Procedure

Hypothesis Tests

We will look at 4 different hypothesis testing scenarios.
Their null hypotheses are given below:

- $H_0 : \mu = \#$
- $H_0 : p = \#$
- $H_0 : \mu_1 - \mu_2 = \#$
- $H_0 : p_1 - p_2 = \#$

The above all follow the same general hypothesis testing procedure.

Testing Procedure

General Hypothesis Testing Procedure

1. Determine the Null and Alternative Hypotheses:

   $H_0 : \theta = \#$

   $H_A : \theta \underset{\neq}{\overset{<}{>}} \#$

2. Gather data and calculate a *test statistic* under the assumption that H_0 is true. Test statistic has general form:

$$Z = \frac{\hat{\theta} - \#}{SE(\hat{\theta})}$$

3. Calculate the *p-value*. Use $p-$value to determine whether you have enough evidence to reject the null hypothesis.
   - small $p-$value $\rightarrow$ H_0 unlikely $\rightarrow$ Reject H_0
   - large $p-$value $\rightarrow$ No evidence against H_0

Calculating p-values

Calculating $p-$value

Definition: $p-$value

The *$p-value$* is the probability of observing your test statistic or *more extreme* if the null hypothesis (H_0) is true.

"*more extreme*" can be bigger, smaller or both depending on the the sign in the alternative hypothesis (H_A)

- Small $p - value$ indicates a small probability of seeing your data if H_0 is true. The data is evidence against H_0 (Reject H_0)
- Large $p - value$ indicates a large probability of seeing your data if H_0 is true. No evidence against H_0 (Do Not Reject H_0)
- $P - value$ is often *wrongly* interpreted as the probability of the null hypothesis. (Don't make this mistake)

Calculating the $p - value$

- By central limit theorem, the estimator follows a normal distribution. Standardizing the estimator gives us the test statistic Z, which follows $N(0,1)$ distribution
- Obtain $p - value$ from the $z-$table as left-hand area, right-hand area or both (depending on sign in H_A)

Left-sided Hypothesis Test

$H_0 : \theta = \#$

$H_A : \theta < \#$

$Z = \frac{\hat{\theta} - \#}{SE(\hat{\theta})}$

Calculating $p-$value Cont.

Right-sided Hypothesis Test

$H_0 : \theta = \#$

$H_A : \theta > \#$

$Z = \frac{\hat{\theta} - \#}{SE(\hat{\theta})}$

2-sided Hypothesis Test

$H_0 : \theta = \#$

$H_A : \theta \neq \#$

$Z = \frac{\hat{\theta} - \#}{SE(\hat{\theta})}$

Types of Errors

In the testing framework, it is possible to make errors that are inherent to the testing procedure (not calculation mistakes).

Types of errors

- Type I Error (wrongly reject H_0)
  $\rightarrow$ P(Type I error) $= \alpha$
- Type II Error (wrongly fail to reject H_0)
  $\rightarrow$ P(Type II error) $= \beta$

Note:

- α (significance level) can be viewed as a cut-off for how small the $p-$value needs to be to reject H_0. Reject H_0 if $p - value < \alpha$. (α set before conducting the test).
- In this class, we use a strength of evidence argument without a "cut-off" for $p - value$.

Hypothesis Testing Summary

Null Hypothesis	Test-Statistic	Reference Dist.
$H_0 : \mu = \#$	$Z = \frac{\bar{X} - \#}{s/\sqrt{n}}$	$Z \sim N(0,1)$
$H_0 : p = \#$	$Z = \frac{\hat{p} - \#}{\sqrt{\frac{\#(1-\#)}{n}}}$	$Z \sim N(0,1)$
$H_0 : \mu_1 - \mu_2 = \#$	$Z = \frac{(\bar{X}_1 - \bar{X}_2) - \#}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$	$Z \sim N(0,1)$
$H_0 : p_1 - p_2 = \#$	$Z = \frac{(\hat{p}_1 - \hat{p}_2) - \#}{\sqrt{\hat{p}_{pool}(1-\hat{p}_{pool})}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ where $\hat{p}_{pool} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}$	$Z \sim N(0,1)$