M358K: Hypothesis Testing.

October 14^{th}, 2020.

Proof by Contradiction

K... claim that I'm trying to prove as true

Q: What if K is not true?

Assume not K

$\Downarrow \quad \Downarrow \quad \Downarrow$

fact Q ⋮ fact not Q

These cannot coexist.

We say that we have reached a contradiction!

$\Rightarrow\Leftarrow$

Our assumption of not K was wrong!

Special Case:

Assume not K

$\Downarrow$

K

Hypothesis Testing.

Claim which you're trying to substantiate: $\mu < \mu_0$

alternative hypothesis

μ: the population parameter

e.g., the population mean cholesterol level after tretment

μ_0: a #

e.g., the average cholesterol level before treatment

Assume: $\mu = \mu_0$ — null hypothesis

⋮

collect data

statistical analysis

⋮

Figure out the probability of seeing the data that you saw if $\mu = \mu_0$.

If this probability is "small", you have evidence to the contrary.

Population Model.

$X \sim \text{Normal}(\text{mean} = \mu, \text{variance} = \sigma^2)$

unknown

assumed to be known/given

Hypothesis Testing Procedure.

Set our hypotheses:

Null hypothesis: $H_0 : \mu = \mu_0$

Alternative hypothesis: $H_a : \begin{cases} \mu < \mu_0 & \text{(lower or left-sided)} \\ \mu \neq \mu_0 & \text{(two-sided)} \\ \mu > \mu_0 & \text{(upper or right-sided)} \end{cases}$

Let the sample size be n.

The sample will be $X_1, X_2, \dots, X_n$.

Since our parameter of interest is the population mean μ, it's useful to focus on the sample mean $\bar{X}$: $\bar{X} = \frac{1}{n}(X_1 + \dots + X_n)$

We know: $\bar{X} \sim \text{Normal}\left(\text{mean} = \mu, \text{variance} = \frac{\sigma^2}{n}\right)$

Under the null hypothesis, i.e., if $\mu = \mu_0$:

$$\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1) \text{ ... std normal.}$$

Say that $\bar{x}$ is the observed sample average.

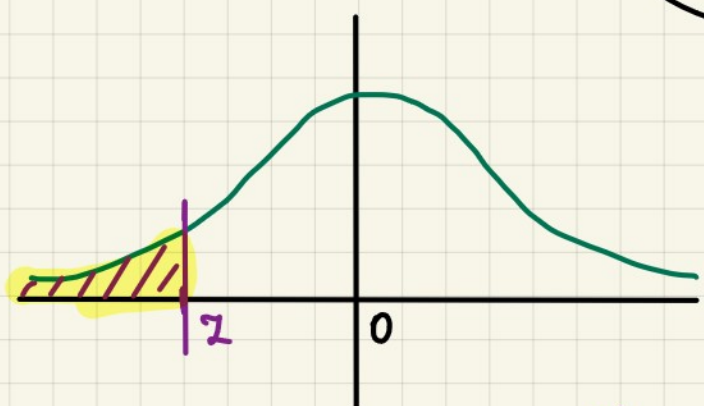
Q: What is the probability of observing that value or something more extreme under the null?

The interpretation for this depends on the structure of your alternative hypothesis.

We always calculate the z-score of the observed value of our test statistic $\bar{x}$.

$$\boxed{z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}}$$

- Lower-sided alternative: $H_a: \mu < \mu_0$



$\mathbb{P}[Z \le z] = p\text{-value}$

z-score (observed)

the smaller the p-value, the stronger your evidence for rejecting the null hypothesis.