M358K: October 13th, 2023.

Hypothesis Testing.

Proof by Contradiction.

K... the claim we are trying to PROVE to be true

Q: What if K were not true?

Assume not K

fact Q ...... fact (not Q)

The cannot coexist!

We say that we reached contradiction!

$\Rightarrow\Leftarrow$

Our assumption of not K was wrong!

Hypothesis Testing.

Claim we're trying to SUBSTANTIATE.

μ... the population mean parameter
(say the mean cholesterol level after treatment)

μ_0... the null population mean (a number)
(say the mean cholesterol level w/ old treatment)

$\boxed{\mu < \mu_0}$ ← Alternative Hypothesis

Assume $\boxed{\mu = \mu_0}$ ← Null Hypothesis

collect data
statistical analysis

Figure out the probability of seeing the data that you saw (or something more extreme) if $\mu = \mu_0$

If this probability is "small", we have evidence against $\mu = \mu_0$,

The smaller this probability, the stronger the evidence.

p-value

The Normal Case.

Population model: $X \sim \text{Normal}(\boxed{\text{mean} = \mu}, \text{sd} = \sigma)$

known

unknown and of interest

Hypothesis Testing Procedure.

First: Set the hypothesis.

2^{nd} Null Hypothesis: $\boxed{H_0 : \mu = \mu_0}$

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

Second: Figure out the appropriate TEST STATISTIC (TS).

Natural choice:

$$\bar{X} \sim \text{Normal}\left(\text{mean} = \mu, \text{sd} = \frac{\sigma}{\sqrt{n}}\right)$$

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

$$\frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

Third: Consider the observed value of the test statistic.

In this case, it's $\bar{x}$, the observed sample average.

Q: What's the probability of observing $\bar{x}$ or something $\boxed{\text{more extreme}}$ under the null?

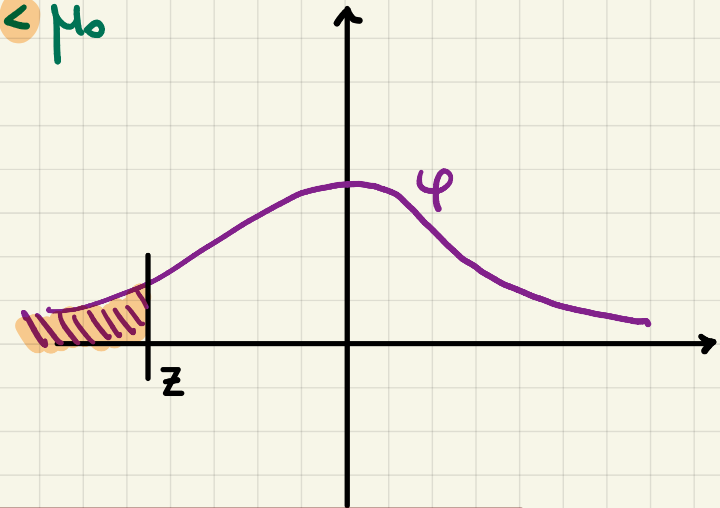
The exact interpretation depends on the structure of the alternative hypothesis!

Regardless:

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

Left-Sided Alternative.

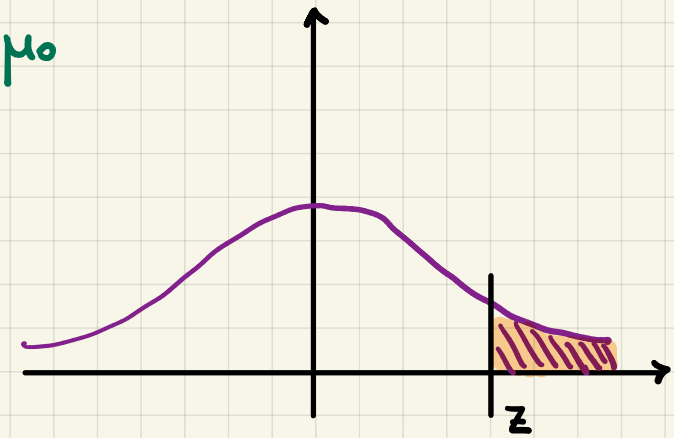
$H_a: \mu < \mu_0$



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

Right-Sided Alternative.

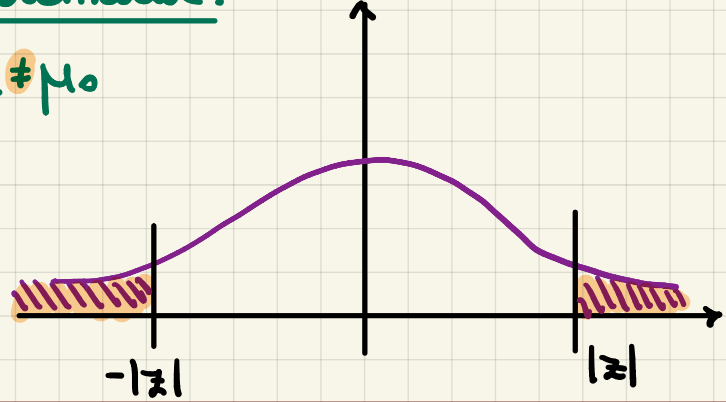
$H_a: \mu > \mu_0$



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

Two-Sided Alternative.

$H_a: \mu \neq \mu_0$



$$\boxed{\mathbb{P}[Z \geq |z|] + \mathbb{P}[Z \leq -|z|] = 2 \cdot \mathbb{P}[Z \leq -|z|] = p\text{-value}}$$

UNIVERSITY OF TEXAS AT AUSTIN

Problem Set # 10

Hypothesis testing. p–value.

Problem 10.1. The null hypothesis is a statement about the population parameter. *True or false?*

Problem 10.2. The null and alternative hypotheses are stated in terms of the statistics obtained from the random sample. *True or false?*

Complete the following statements:

Problem 10.3. When we state the alternative hypothesis to look for a difference in a parameter in any direction, we are doing a ____two____-sided test.

Problem 10.4. When choosing between a one-sided alternative hypothesis and a two-sided alternative hypothesis, you should base the decision on ____the research question you're trying to answer.

Problem 10.5. The ____smaller____ the p–value, the **stronger** the evidence against the null hypothesis provided by the data.

Provide your **complete solution** for the following problems.

Problem 10.6. The square footage of several thousand apartments in a new development is advertised to be 1250 square feet, on average. A tenant group thinks that the apartments are smaller than advertised. They hire an engineer to measure a sample of apartments to test their suspicions. Let μ represent the "true" mean area (in square feet) of these apartments. What are the appropriate null and alternative hypotheses?

$H_0: \mu = 1250$ vs. $H_a: \mu < 1250$

Problem 10.7. Is the mean height for all adult American males between the ages of 18 and 21 now over 6 feet? Let μ denote the population mean height of all adult American males between the ages of 18 and 21. What are the appropriate null and alternative hypotheses?

$H_0: \mu = 6$ vs. $H_a: \mu > 6$

Problem 10.8. The hypotheses are $H_0 : \mu = 10$ versus $H_a : \mu > 10$. The value of the test statistic for the population mean is $z = -2.12$. What is the corresponding p–value?

Right sided : $\mathbb{P}[Z > -2.12] = 1 - \Phi(-2.12) = 1 - 0.017 = 0.983$

Problem 10.9. The value of the test statistic for a **two-sided** test for a population mean is $z = -2.12$. What is the corresponding p–value?

$\Downarrow$

$2 \cdot \mathbb{P}[Z < -2.12] = 2 \cdot (0.017) = 0.034$ $\square$