M378K: April 16th, 2025.

Hypothesis Testing.

Proof by Contradiction.

$\boxed{K}$ ... the claim we're trying to PROVE to be true

Q: What if K were not true?

Assume $\boxed{\text{not } K}$

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

These cannot coexist!

We say that we reached a 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 w/ pills)

μ_0 ... the null population mean (A number) (say, a healthy benchmark)

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

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

collect data
statistical analysis

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

p.value

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

The smaller the probability the STRONGER THE EVIDENCE.

The Normal Case w/ σ known.

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

unknown
and of
interest

Hypothesis Testing Procedure.

First: Set the hypotheses.

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

1st 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:

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

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

$$Z = \boxed{\frac{\bar{Y} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)}$$

Third: Consider the observed value of the TS.
In this case, it's $\bar{y}$, i.e., the observed sample average.

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

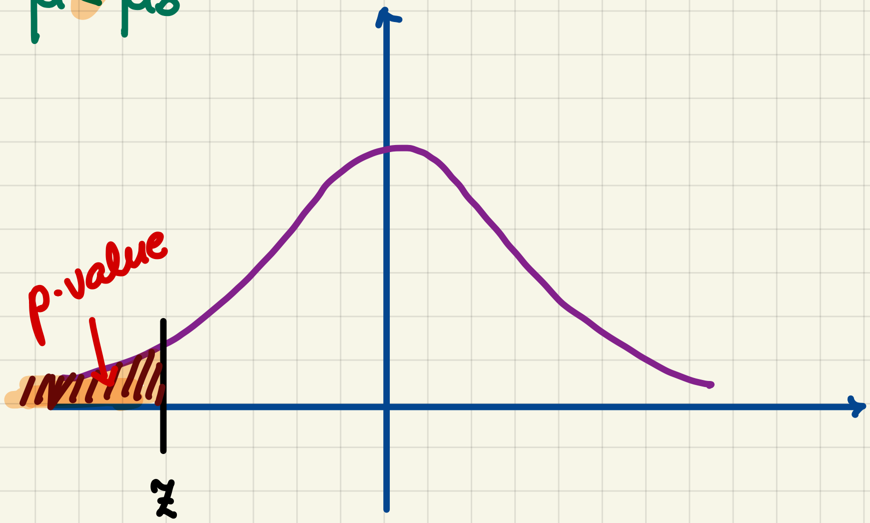
Exact interpretation depends on the structure of the alternative hypothesis.

Regardless:

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

Left-Sided Alternative:

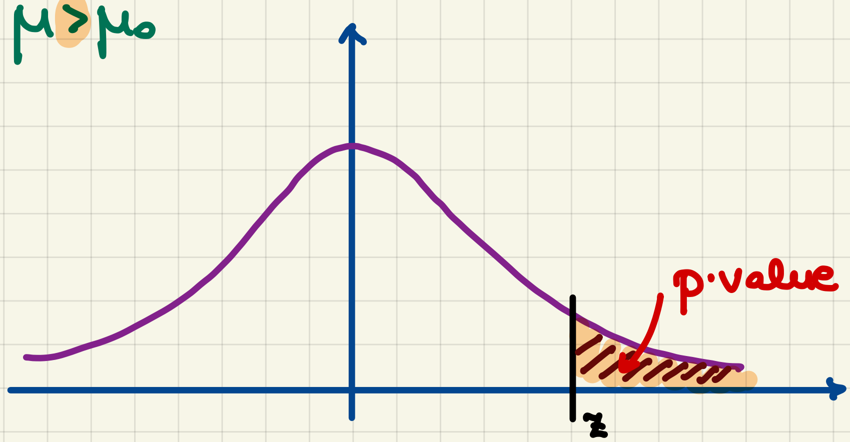
$H_a: \mu < \mu_0$



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

Right-Sided Alternative:

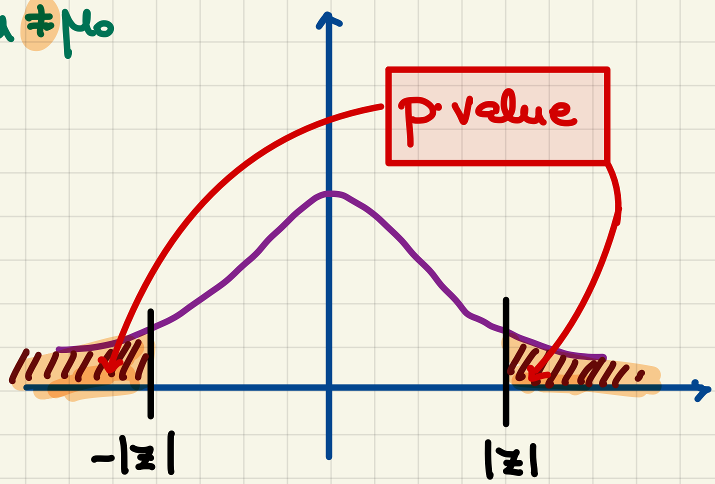
$H_a: \mu > \mu_0$



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

Two-Sided Alternative.

$H_a: \mu \neq \mu_0$



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

Test of Significance.

Set α ... significance level

Typically: $\alpha = \boxed{0.05}, 0.01, 0.10$

Decision Process.

If p-value $\leq \alpha$, we REJECT the null hypothesis.

If p-value $> \alpha$, we FAIL TO REJECT the null hypothesis.

Note: The p-value corresponding to an observed value of the test statistic is the LOWEST significance level @ which the null hypothesis would still be REJECTED.