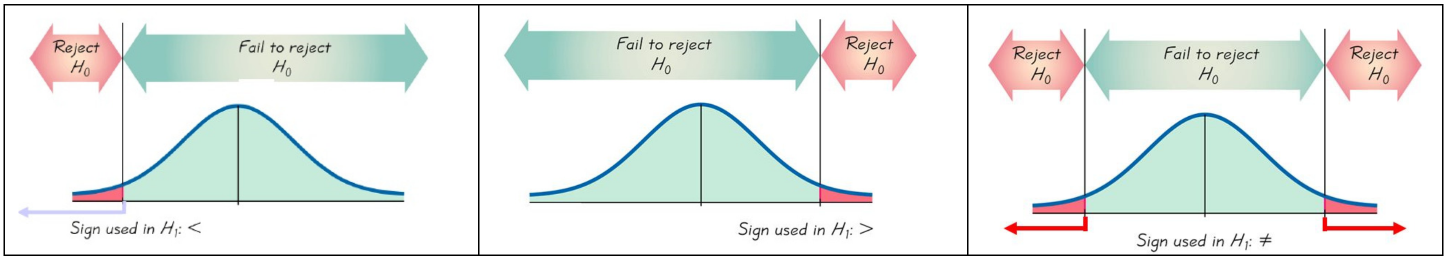
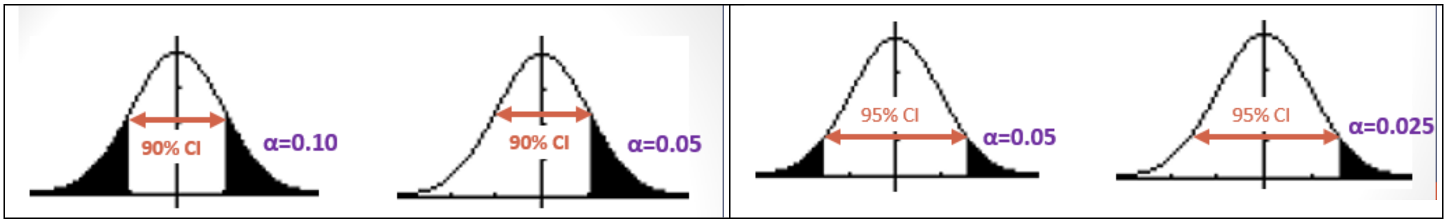
CI and HT

One tail vs two-tail





Using Confidence Intervals for Testing:

- If Ho is in the interval, fail to reject it
- If Ho is not in the interval, reject it

Example 1

p.535 #18

Is the Euro fair? Soon after the Euro was introduced as currency in Europe, it was widely reported that someone had spun a Euro coin 250 times and gotten heads 140 times. We wish to test a hypothesis about the fairness of spinning the coin.

a) Estimate the true proportion of heads. Use a 95% confidence interval. Don't forget to check the conditions.

$\hat{p} = \frac{140}{250} = 0.56$ – proportion of heads.	• the sample is random • spins are independent • 250 spins are less than 10% of all coin spins • np = 140 and nq = 110 are both greater than 10 The conditions have been satisfied
$\hat{p} = 0.56 \quad \hat{q} = 0.44 \quad n = 250$ $SE = \sqrt{\frac{0.56 * 0.44}{250}} = 0.0314$ $0.56 \pm 1.96 * 0.0314$ $(0.498, 0.621)$	*I am 95% confident that the true proportion of heads is between 49.8% and 62.1%.*

b) Does your confidence interval provide evidence that the coin is unfair when spun? Explain.

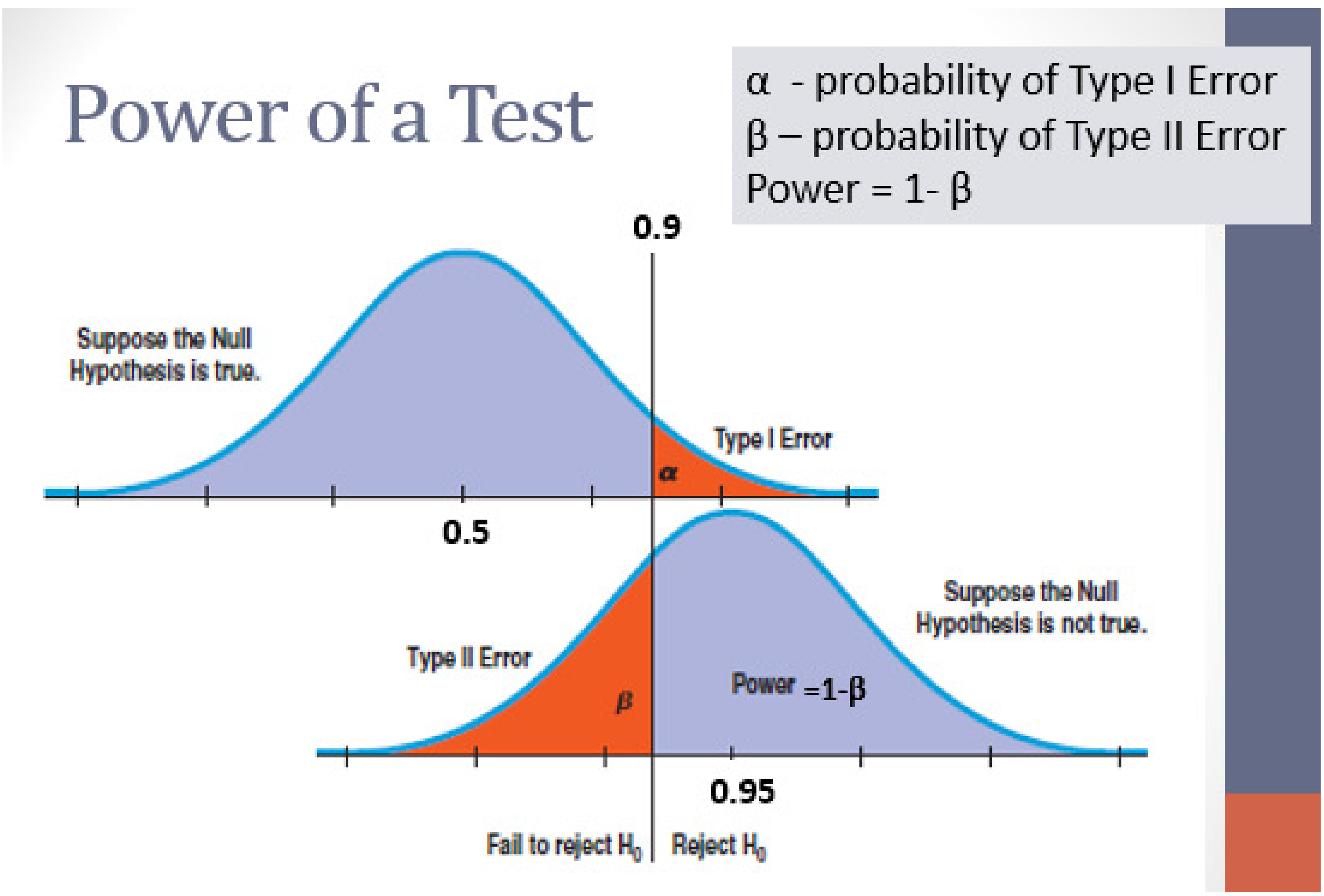
No.
Ho: p=0.5 (the coin is fair)
0.5 is within the interval, so it is a plausible value for the proportion of heads.

c) What is the rejection/significance level of this test? Explain.

α = 1-0.95 = 0.5
It is a two-tailed test based on 95% confidence interval.

Power of a Test

- The probability that the test will correctly reject a false null hypothesis



Example 2 p.553 #29,31

A clean air standard requires that vehicle exhaust emissions not exceed specified limits for various pollutants. Many states require that cars be tested annually to be sure they meet these standards. Suppose state regulators double-check a random sample of cars that a suspect repair shop has certified as okay. They will revoke the shop's license if they find significant evidence that the shop is certifying vehicles that do not meet standards.

a)In this context, what is a Type I error?

the regulators decide that the shop is not meeting standards when they actually are meeting the standards.

b) In this context, what is a Type II error?

the regulators certify the shop when they are not meeting the standards.

c) In this context, what is meant by the power of the test the regulators are conducting?

the probability of detecting that the shop is not meeting standards when they are not

d) Will the power be greater if they test 20 or 40 cars? Why?

greater when 40 cars are tested. A larger sample size increases the power of the test.

e) Will the power be greater if they use a 5% or a 10% level of significance? Why?

greater when the level of significance is 10%. There is a greater chance that the null hypothesis will be rejected.

f) Will the power be greater if the repair shop's inspectors are only a little out of compliance or a lot? Why?

greater when the shop is out of compliance "a lot". Larger problems are easier to detect.

Example 3

With a α-level of 5%, what is the power of a test against an alternative for which the probability of a Type II error is 3%?

a) 3%
b) 5%
c) 95%
d) 97%
e) The power cannot be determine from the information given.