# The Power of a Hypothesis Test

## POWER

The probability that a fixed level $\alpha$ significance test will reject $H_0$ when a particular alternative value of the parameter is true is called the **power** of the test to detect that alternative.

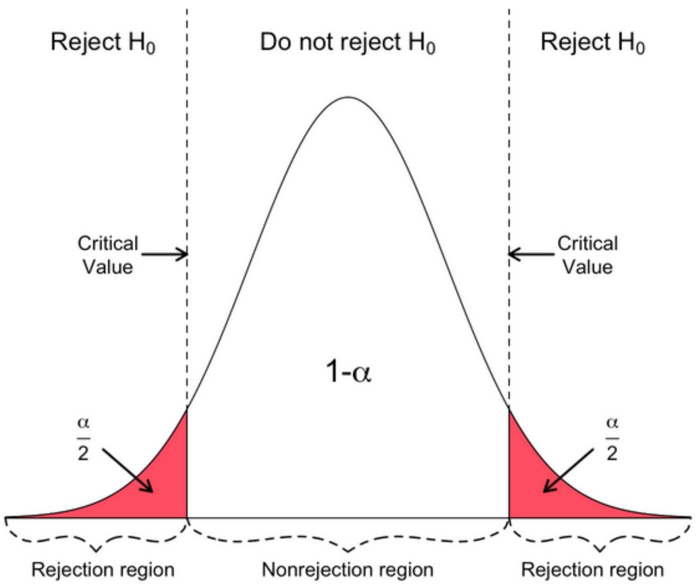
NOTE: Power = 1 – P(Type II error)

# Power of a Test

Consider the hypotheses

$H_0$: μ = 10
$H_a$: μ ≠ 10

Also given: *n*=25 (sample size) σ=1 (std.dev) α=0.05




Power(μ) = P(Reject $H_0$ ) = P($\bar{X} < 9.608$) + P($\bar{X} > 10.392$)

# Confidence Intervals Overview

- Provides a range of likely values of a given parameter.
- When we have a specific question in mind, a hypothesis test is usually more appropriate, particularly if the alternative hypothesis is one-sided.
- For questions involving a two-sided hypothesis, the conclusion obtained using a 1-$\alpha$ confidence interval is the same as the conclusion obtained using a hypothesis test with significance level $\alpha$.
- If we are interested in how strong a demonstrated effect is, or much a parameter of interest is different from a given value, a confidence interval is more informative than a hypothesis test.

# Hypothesis Testing Overview

- Appropriate for answering a specific question, such as "Is there evidence that the mean is greater that 5?".
- When a one-sided alternative can be used, a hypothesis test is more likely to demonstrate an effect than the corresponding confidence interval.
- The *p*-value says something about our confidence in the decision made. The smaller the *p*-value is the more confident we are that $H_0$ is false.
- A hypothesis test does not give any information about how strong a demonstrated effect is, or much the parameter of interest is different from the value assumed under the null hypothesis.

# Hypothesis Testing Overview (cont.)

- Not rejecting a null hypothesis does not provide evidence that the null hypothesis is true.
- A null hypothesis can be false even when the *p*-value is large if the test has small power.
- The best way to investigate the degree to which the "null-hypothesis it true" is by calculating and interpreting the confidence interval for the parameter being tested.
- The significance level $\alpha$ is the probability of incorrectly rejecting a null hypothesis that is true (Type I error). It does not accurately reflect the probability of a Type I error when a test is repeated several times. A Type I error is very likely to occur when a test is repeated many times (be aware of "searching for significance".)

# DOs and DON'Ts for
## Hypothesis Tests

Each case below involves a statistical inference.

Determine whether the interpretation is correct. If not explain what is wrong.

- **Given:** In testing the hypotheses

  $H_0: \mu = 20$

  $H_A: \mu \neq 20$

  a test statistic z = 2.9 is observed.

- **Interpretation:** *Since $z > 2.58$ ($z_{\alpha/2}$ with $\alpha$=0.01) there is sufficient evidence that $\mu > 20$.*

CORRECT!

# DOs and DON'Ts for

## Hypothesis Tests

Each case below involves a statistical inference.

Determine whether the interpretation is correct. If not explain what is wrong.

- In the previous example the *p*-value is very small (0.0038).
- **Interpretation:** *Therefore there is a small probability that the null hypothesis is true.*

WRONG! The null hypothesis is either true or it is false. It is not a random event that sometimes happen and sometimes don't. We can say that there is a small probability that we would observe this large value of the test statistic if the null hypothesis is true.

# DOs and DON'Ts for

## Hypothesis Tests

Each case below involves a statistical inference.
Determine whether the interpretation is correct. If not explain what is wrong.

- **Given:** A research wants to demonstrate that a mean is greater than 20 using a significance level $\alpha = 0.05$ and the hypotheses

  $H_0: \mu = 20$
  $H_A: \mu > 20$

  The first test does not come out significant (p-value = 0.07), so the researcher decides to take another sample and in this case the test is significant (p-value =0.04)

- **Interpretation:** *The researcher concludes that $\mu > 20$ based on extensive research.*

  WRONG! When a test is repeated multiple times then the probability of making a wrong decision is increased.