STATISTICAL TESTS

Overview

Hypothesis-testing
Null hypothesis (H0)
Alternative hypothesis (H1)
Test statistic
Types of error
level of significance
Power

STATISTICAL HYPOTHESIS

A statistical hypothesis is an assertion or conjecture concerning one or more populations.

To prove that a hypothesis is true or false, we would need absolute knowledge with absolute certainty. That is, we would have to examine the entire population.

Instead, hypothesis testing concerns how to use a random sample to judge if it is evidence that supports or not the hypothesis

HYPOTHESIS TESTING

Hypothesis testing is formulated in terms of two hypotheses:

- H0: the null hypothesis;
- H0 : $p = 0.05$;
- H1: the alternate hypothesis;
- H1 : $p > 0.05$.

HYPOTHESIS TESTING

The hypothesis we want to test is if H1 is “likely” true. So, there are two possible outcomes:

- Reject H0 and accept H1 because of sufficient evidence in the sample in favor of H1;

- Do not reject H0 because of insufficient evidence to support H1

Very important!!

Failure to reject H0 does not mean the null hypothesis is true. There is no formal outcome that says “accept H0.” It only means that we do not have sufficient evidence to support H1.

HYPOTHESIS TESTING

Example
In a jury trial the hypotheses are:

- H0: the defendant is innocent;

- H1: the defendant is guilty.

H0 (innocent) is rejected if H1 (guilty) is supported by evidence beyond “reasonable doubt.” Failure to reject H0 (prove guilty) does not imply innocence, only that the evidence is insufficient to reject it

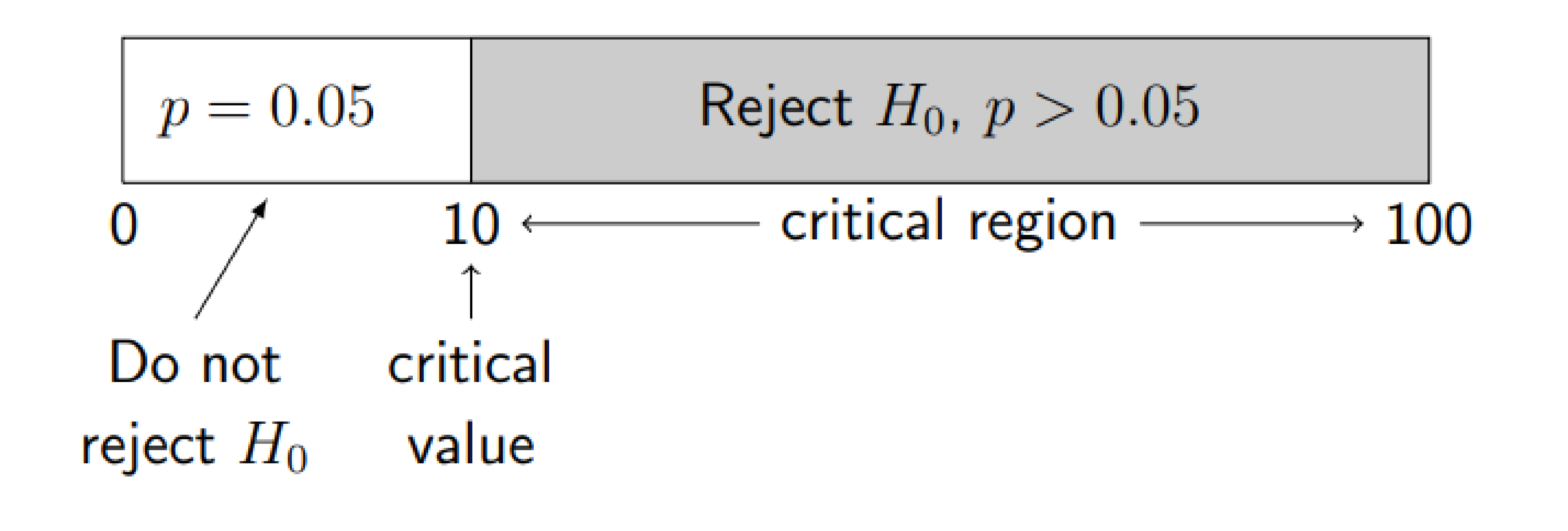
A company manufacturing RAM chips claims the defective rate of the population is 5%. Let p denote the true defective probability. We want to test if:

- $H_0 : p = 0.05$
- $H_1 : p > 0.05$

We are going to use a sample of 100 chips from the production to test.

TEST STATISTICS

Let X denote the number of defective in the sample of 100. Reject H0 if X ≥ 10 (chosen “arbitrarily” in this case). X is called the test statistic



Why did we choose a critical value of 10 for this example?

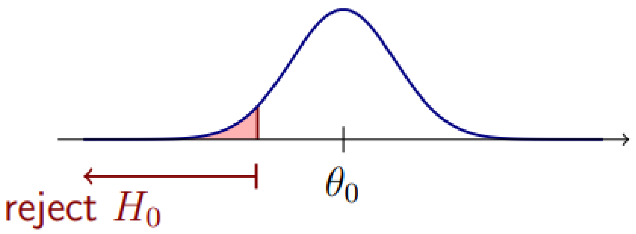
Because this is a Bernoulli process, the expected number of defectives in a sample is np. So, if $p = 0.05$ we should expect $100 \times 0.05 = 5$ defectives in a sample of 100 chips. Therefore, 10 defectives would be strong evidence that $p > 0.05$. The problem of how to find a critical value for the desired level of significance of the hypothesis test

Test Statics

Another one-tailed test could have the form,

- H_0: $\theta = \theta_0$
- H_1: $\theta < \theta_0$,

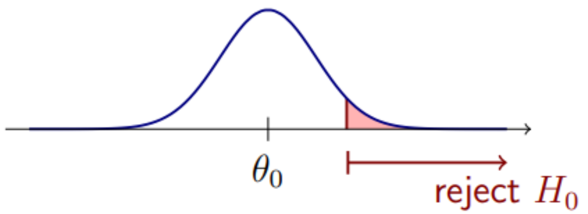
in which the critical region is in the left-tail.



In our examples so far we have considered:

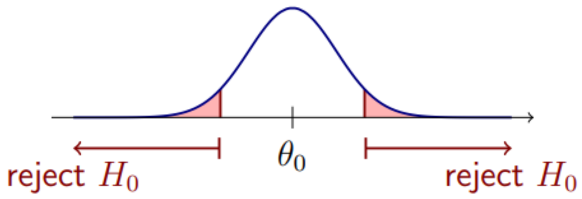
- H_0: $\theta = \theta_0$
- H_1: $\theta > \theta_0$.

This is a one-tailed test with the critical region in the right-tail of the test statistic X.



In a two-tailed test check for differences:

- H_0: $\theta = \theta_0$
- H_1: $\theta \neq \theta_0$,



TYPES OF ERROR

The possible outcomes are:

	H0 is true	H1 is true
Do not reject H0	Correct decision	Type II error
Reject H0	Type I error	Correct decision

Conclusion. If H_0 is rejected, we conclude that H_A is true. If H_0 is not rejected, we conclude that H_0 may be true.

TYPES OF ERROR

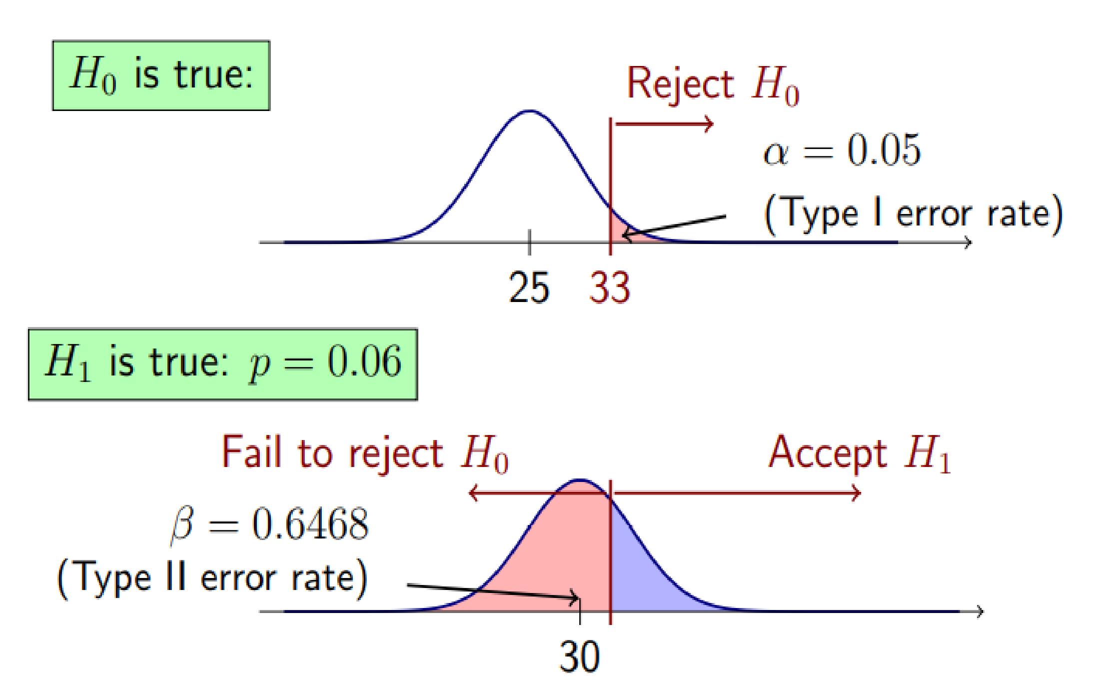
The acceptance of H1 when H0 is true is called a Type I error. The probability of committing a type I error is called the level of significance and is denoted by α.

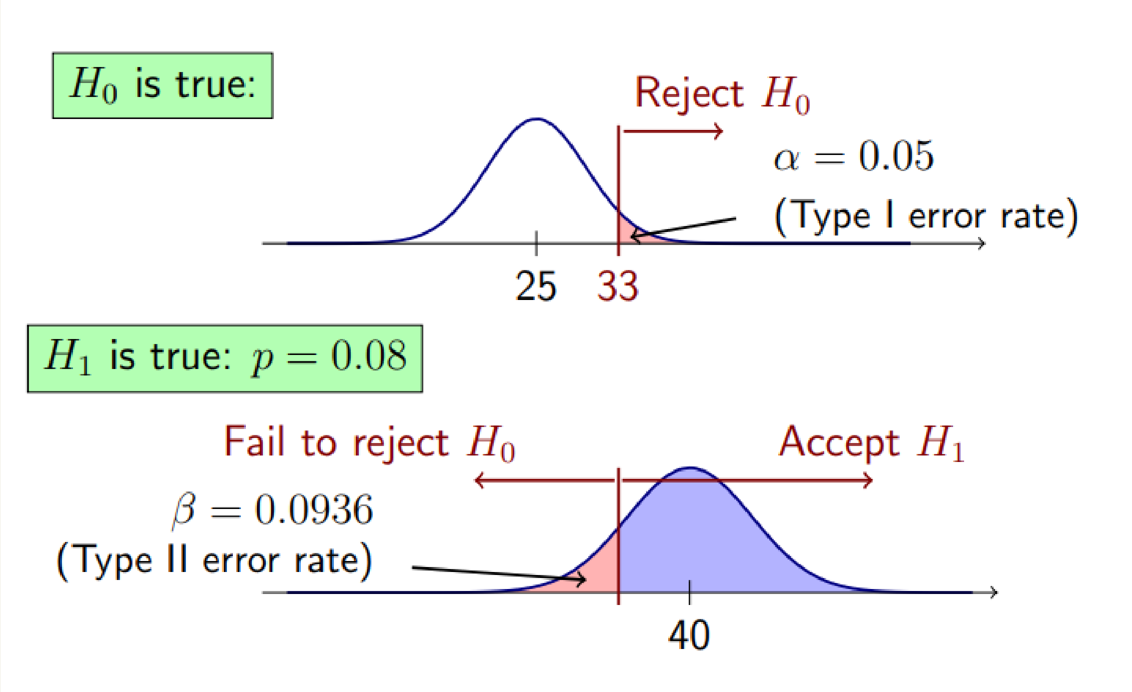
The lower significance level α, the less likely we are to commit a type I error

Failure to reject H0 when H1 is true is called a Type II error. The probability of committing a type II error is denoted by β.

It is impossible to compute β unless we have a specific alternate hypothesis.

TYPES OF ERROR





Power of a test

Definition: The power of a test is the probability of rejecting H0 given that a specific alternate hypothesis is true. That is Power = $1 - \beta$.