# Estimation vs hypothesis testing

**In the case of estimation**, we ask a question about the value of a particular parameter.

**In hypothesis testing** the question is preceded by a statement concerning the population; the question then is whether this statement is true or false.

# Hypotheses

In statistical theory a hypothesis is a supposition about the population.

A statement that a certain population parameter is equal to a given value. This hypothesis is called **the null hypothesis**: $H_{0.}$

Since the null hypothesis is a testable proposition, there must exist a counterproposition to it.

The counterproposition is called **the alternative hypothesis**: $H_1$.

# Procedure for testing hypotheses

1. State the null hypothesis and the alternative hypothesis,

e.g.

$$H_0\colon \mu = \mu_0$$

$$H_1\colon \mu \neq \mu_0$$

In our example we want to test if mean is still 200 ml, so

$$H_0\colon \mu = 200$$
$$H_1\colon \mu \neq 200$$

# Procedure for testing hypotheses

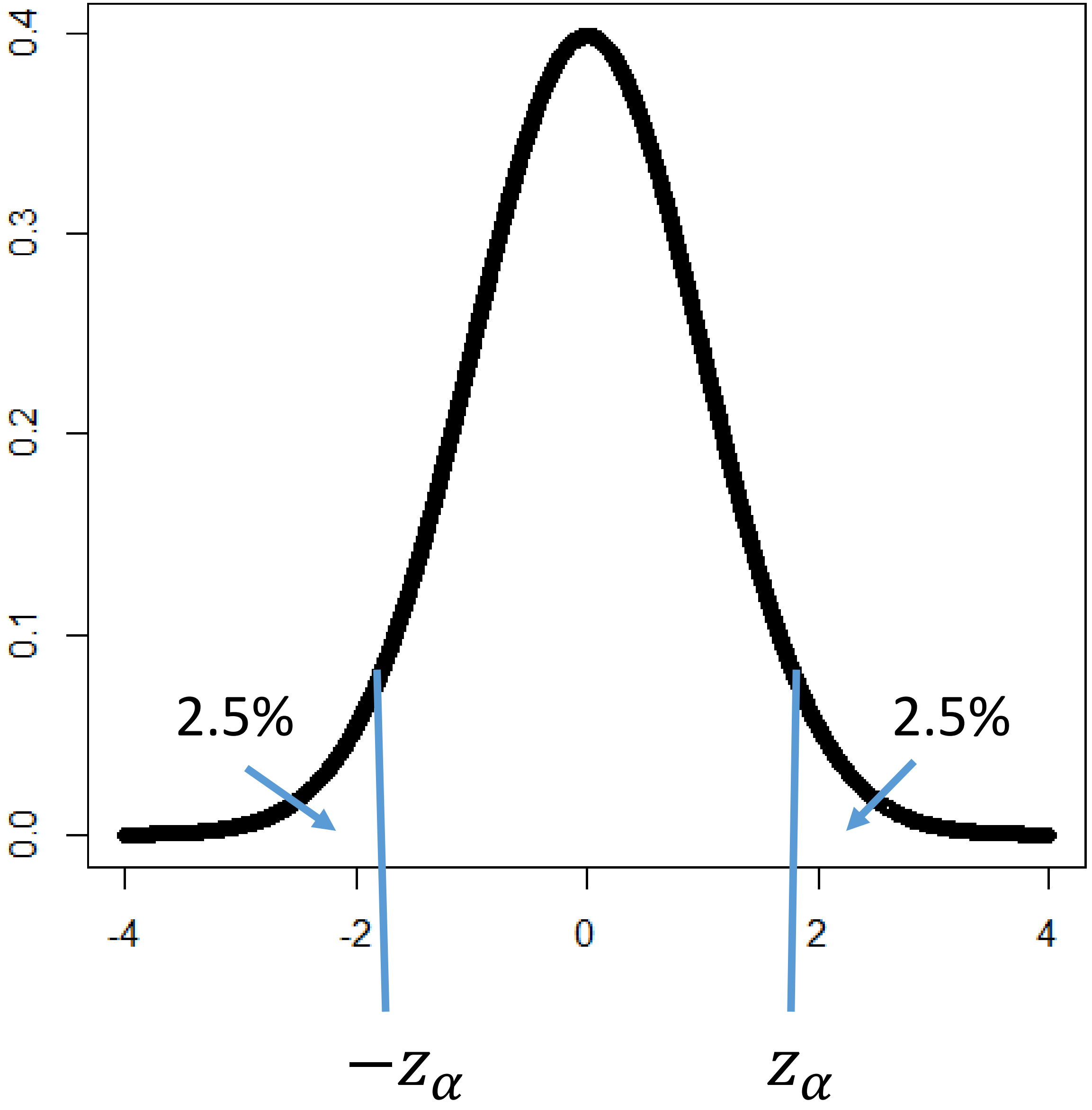
2. Choose the level of significance and determine **the rejection region**.

Significance level is denoted $\alpha$. Usually it is set equal to 5% (or 1%, sometimes 10%).

For our example let's choose $\alpha = 5\%$.

# Rejection region

$\alpha = 5\%$




$z_\alpha = 1.96$

# Procedure for testing hypotheses

3. Draw a sample and calculate the value of the test statistic.

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

$$z = \frac{202.5 - 200}{5}\sqrt{16} = 2$$

# Procedure for testing hypotheses

4. Reach a conclusion:

- if the sample statistic falls into the rejection region

$$|z| \geq z_\alpha$$

reject the null hypothesis

or

- if the sample statistic falls outside the rejection region

$$|z| < z_\alpha$$

state that the sample does not provide evidence against the null hypothesis so there is no reason to reject it.

# Procedure for testing hypotheses

For our example:

$$z_\alpha = 1.96$$

$$z = 2$$

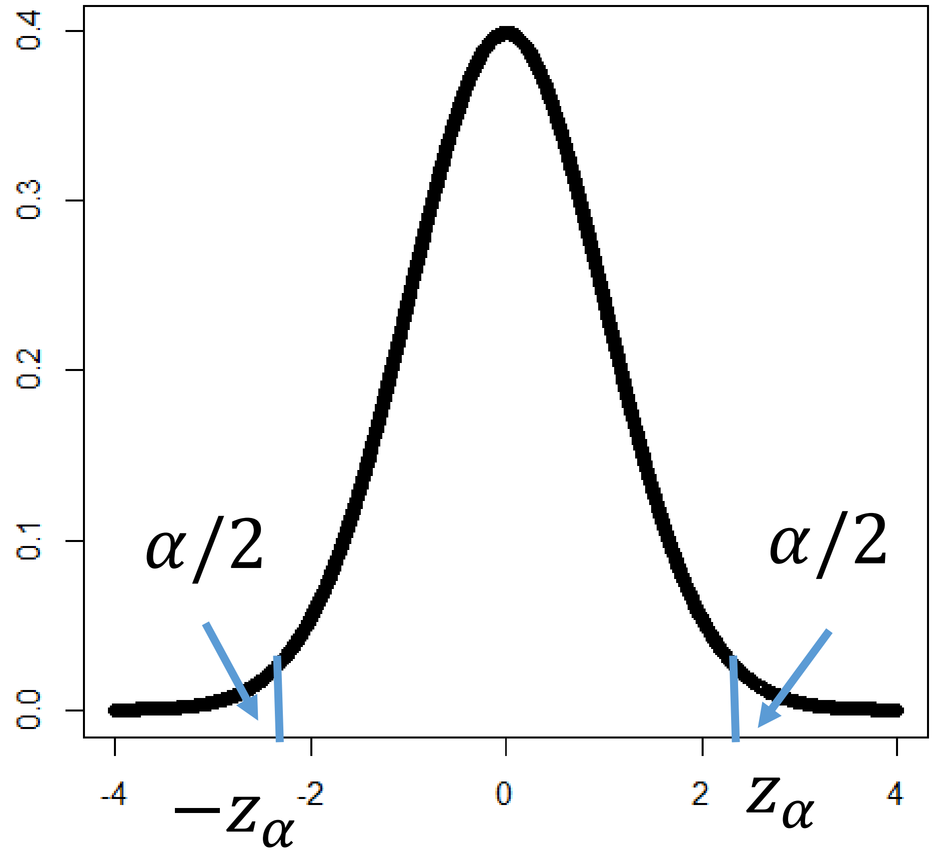
$$|z| \geq z_\alpha$$

so we reject null hypothesis at 5% significance level, and conclude the mean is not equal to 200 ml.

# Two tailed test

$H_0$: $\mu = \mu_0$

$H_1$: $\mu \neq \mu_0$

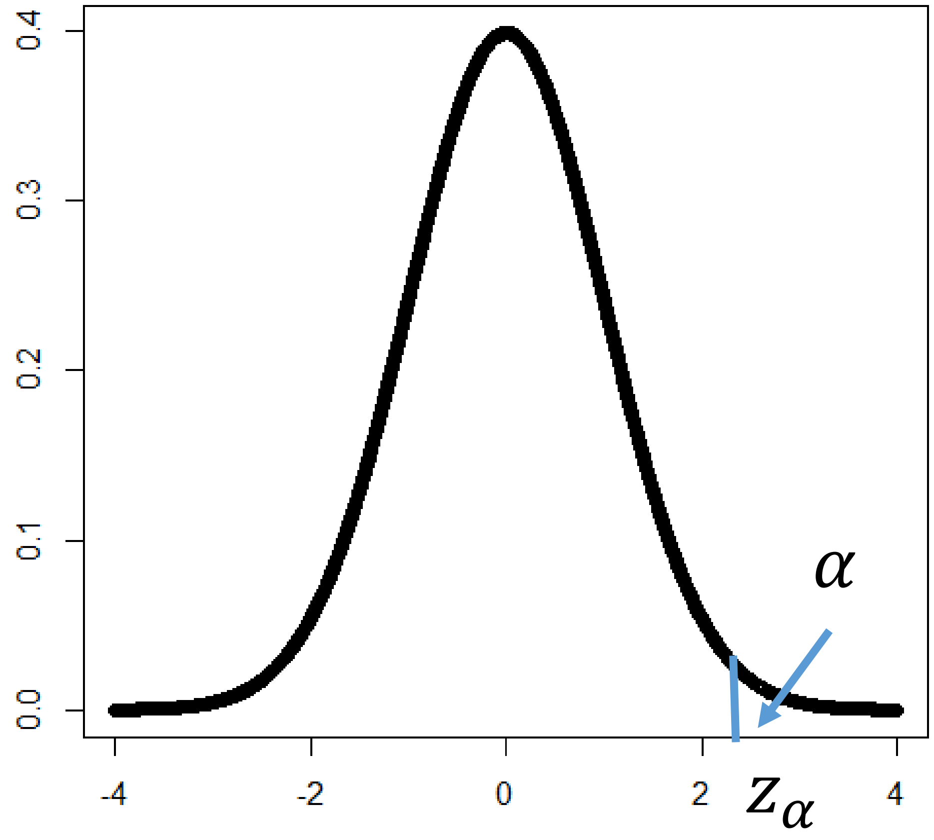



If $|z| \geq z_\alpha$
reject the null hypothesis

# One tailed test

$H_0\colon \mu = \mu_0$

$H_1\colon \mu > \mu_0$

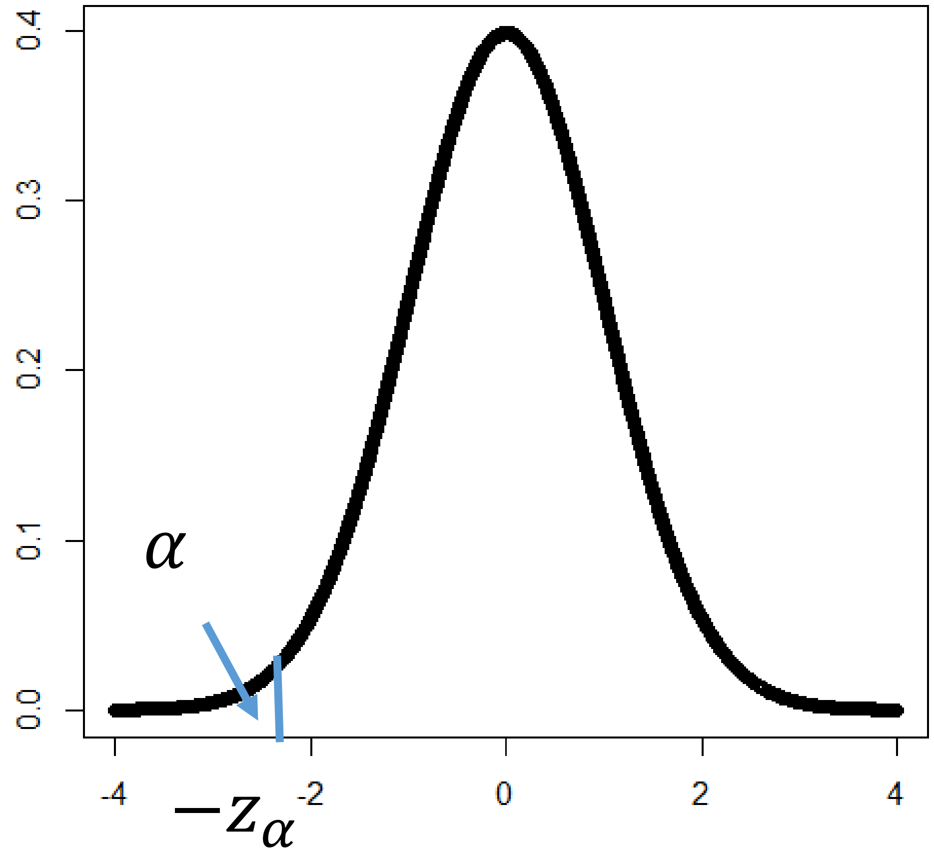



If $z \geq z_\alpha$
reject the null hypothesis

# One tailed test

$H_0$: $\mu = \mu_0$

$H_1$: $\mu < \mu_0$




If $z \leq -z_\alpha$

reject the null hypothesis

# Two tailed vs one tailed tests

- Two tailed test

$H_0$: $\mu = \mu_0$

$H_1$: $\mu \neq \mu_0$

divide $\alpha$ by 2

Reject $H_0$ if $|z| \geq z_\alpha$

- One tailed test

do not divide $\alpha$ by 2

Reject $H_0$ if

$H_0$: $\mu = \mu_0$

$H_1$: $\mu > \mu_0$

$z \geq z_\alpha$

or

$H_0$: $\mu = \mu_0$

$H_1$: $\mu < \mu_0$

$z \leq -z_\alpha$

# Parametric tests

Tests for one parameter

- Mean
- Proportion
- Variance

Tests for two parameters

- Two means
- Two proportions
- Two variances

# Test for population mean

Model 1

Normally distributed population with known variance $\sigma^2$

$$H_0\colon \ \mu = \mu_0$$

$$H_1\colon \ \mu \neq \mu_0 \text{ or } H_1\colon \ \mu > \mu_0 \text{ or } H_1\colon \ \mu < \mu_0$$

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

$z_\alpha$ from Normal distribution

# Test for population mean

Model 2

Population with unknown variance $\sigma^2$, big sample

$$H_0:\ \mu = \mu_0$$

$$H_1:\ \mu \neq \mu_0 \text{ or } H_1:\ \mu > \mu_0 \text{ or } H_1:\ \mu < \mu_0$$

$$z = \frac{\bar{x} - \mu_0}{\hat{s}}\sqrt{n}$$

$z_\alpha$ from Normal distribution

# Test for population mean

Model 3

Normally distributed population with unknown variance $\sigma^2$, small sample

$$H_0\colon\ \mu = \mu_0$$

$$H_1\colon\ \mu \neq \mu_0 \text{ or } H_1\colon\ \mu > \mu_0 \text{ or } H_1\colon\ \mu < \mu_0$$

$$t = \frac{\bar{x} - \mu_0}{\hat{s}}\sqrt{n}$$

$t_\alpha$ from student t distribution, $v = n - 1$

# Test for population proportion

Big sample

$$H_0 : p = p_0$$

$H_1 : p \neq p_0$ or $H_1 : p > p_0$ or $H_1 : p < p_0$

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$$

$z_\alpha$ from Normal distribution

# Test for population variance

Normally distributed population with unknown variance $\sigma^2$

$$H_0\colon\ \sigma^2 = \sigma_0^2$$
$$H_1\colon \sigma^2 \neq \sigma_0^2 \text{ or } \sigma^2 > \sigma_0^2 \text{ or } \sigma^2 < \sigma_0^2$$

$$\chi^2 = \frac{(n-1)\hat{s}^2}{\sigma_0^2}$$

$\chi_\alpha^2$ from $\chi^2$ distribution, $v = n - 1$

# Errors

| | No reason to reject $H_0$ | Reject $H_0$ |
|---|---|---|
| $H_0$ is true | OK | Error type I |
| $H_0$ is false | Error type II | OK |

Error type I: we reject null hypothesis while it is true.
Probability of this error is $\alpha$, the significance level.

Error type II: we don't reject null hypothesis while it is false.
Probability of this error is $\beta$

Power of a test: probability of rejecting null hypothesis while it is false
Power of a test = $1 - \beta$