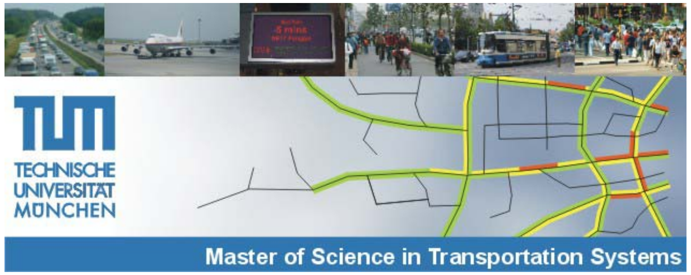
M.Sc. in ‚Transportation Systems'



Applied Statistics in Transport
Hypothesis testing

Regine Gerike
Technische Universität München, mobil.TUM
regine.gerike@tum.de

Munich, 10/01/2012

Plan for Today’s Lecture: Hypothesis Testing

- Introduction
- Types of statistical hypotheses
- One-sided and two-sided hypotheses
- α- and β-error
- General procedure for Hypothesis testing

Hypothesis Testing - Introduction

- So far we illustrated how to construct a confidence interval estimate of a parameter from sample data.
- However, many problems in engineering require that we decide whether to accept or reject a statement about some parameter.
- The statement is called a **hypothesis**.
- The decision-making procedure about the hypothesis is called **hypothesis testing**.
- This is one of the most useful aspects of statistical inference, since many types of decision-making problems, tests, or experiments in the engineering world can be formulated as hypothesis-testing problems.
- There is a very close connection between hypothesis testing and confidence intervals.

Hypothesis Testing - Introduction

- Different types of hypotheses:
- Hypothesis which analyse the difference (today)
- Hypotheses that analyse the relation (correlation, regression)

Examples:

- Drivers of cars with airbag drive faster than drivers of cars without air bag
- Households with higher income own more cars
- Cyclists with helmet drive faster and riskier than cyclists without helmet

Tests of Statistical Hypotheses

- A statistical hypothesis is a statement about the parameters of one or more populations.
- Statistical hypotheses are always formulated in pairs:
- If we reject one of the two, the other is automatically accepted.
- Since it is easier to reject a hypothesis, all test procedures are designed to reject a hypothesis.
- The hypothesis which we try to reject is called the null hypothesis H_0.
- Its counterpart is the alternative hypothesis H_1.
- H_1 is regarded as proven if H_0 is rejected.
- The null hypothesis H_0 says that there are no differences in the population and that the computed difference is only randomly other than zero (with the given statistical certainty).
- The following pairs of hypotheses can be formulated:

Type of Statistical Test	Null Hypothesis	Alternative Hypothesis
Two-sided	$\mu = \mu_0$	$\mu \neq \mu_0$
One-sided	$\mu \leq \mu_0$ $\mu \geq \mu_0$	$\mu > \mu_0$ $\mu < \mu_0$

Decisions in Hypothesis Testing: type I and type II error, α- and β-error

- Rejecting the null hypothesis H_0 when H_0 is true is defined as a type I error.
- Failing to reject the null hypothesis H_0 when it is false is defined as type II error.

Decision	H_0 is true	H_0 is false
Fail to reject H_0	No error / Right decision	Type II error
Reject H_0	Type I error	No error / Right decision

- α = P(type I error) = P(reject H_0 when H_0 is true)
- β = P(type II error) = P(fail to reject H_0 when H_0 is false)

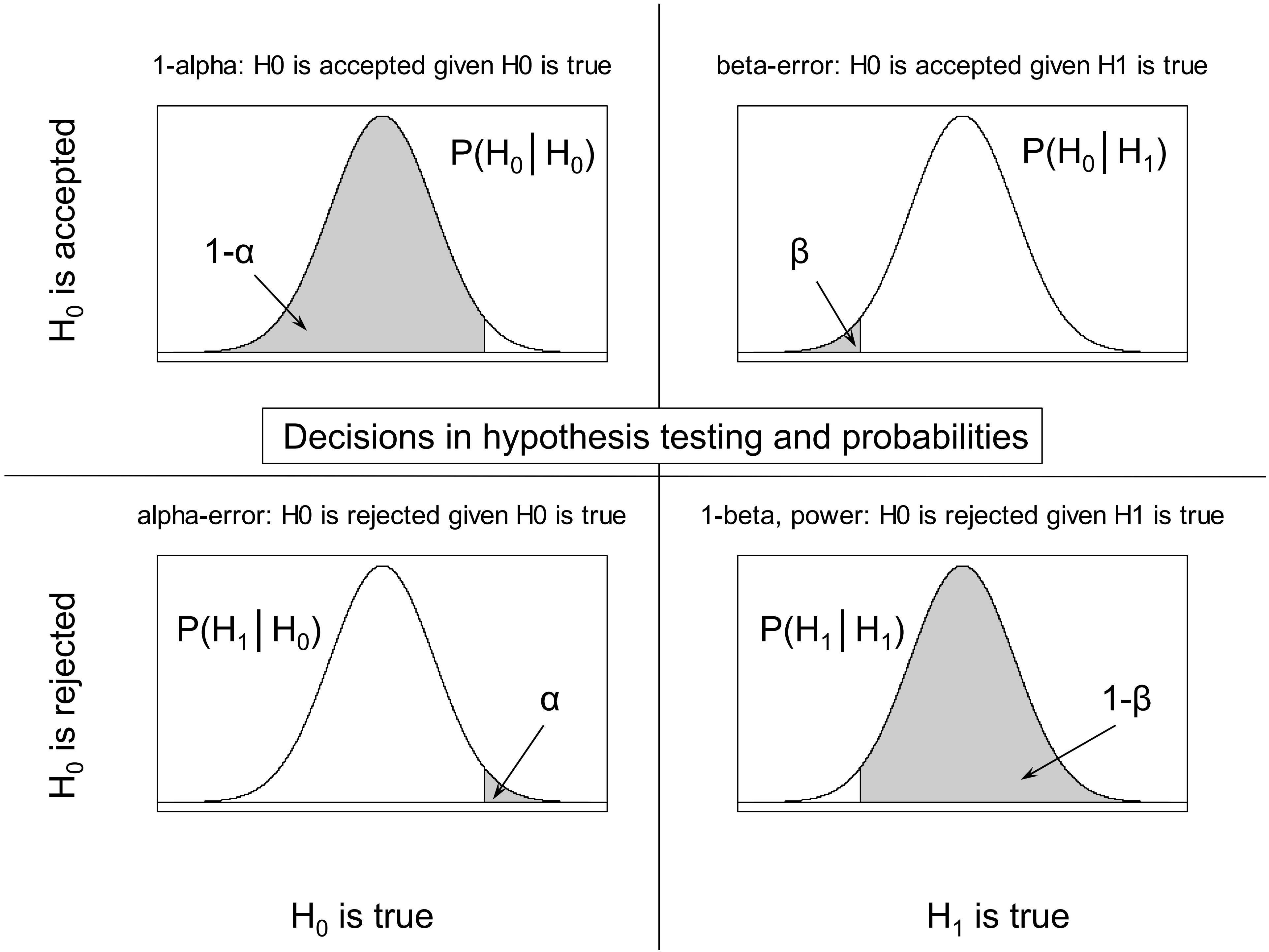
Example: α- and β-error

H_1: $\mu_0 < \mu_1$: The new method of teaching is better than the old one (e.g. the students get better marks when solving the same exercises)

H_0: $\mu_0 \geq \mu_1$: The is no difference between the methods or the new one is worse than the old one.

Type 1 error: H_0 is right (no difference) but we reject it and postulate that the new method is better. Probably there will be re-education/retraining of the teachers, new material will be purchased – all those investments are not justifiable, since the new method is not better.

Type 2 error: H_0 is accepted and consequently H_1 is rejected even though it is true and the new method is better. Then there will be no investment but we miss the opportunity to improve teaching.


Decisions in hypothesis testing and probabilities
1-alpha: H0 is accepted given H0 is true
$P(H_0 | H_0)$
1-α
beta-error: H0 is accepted given H1 is true
$P(H_0 | H_1)$
β
alpha-error: H0 is rejected given H0 is true
$P(H_1 | H_0)$
α
1-beta, power: H0 is rejected given H1 is true
$P(H_1 | H_1)$
1-β
H_0 is accepted
H_0 is rejected
H_0 is true
H_1 is true

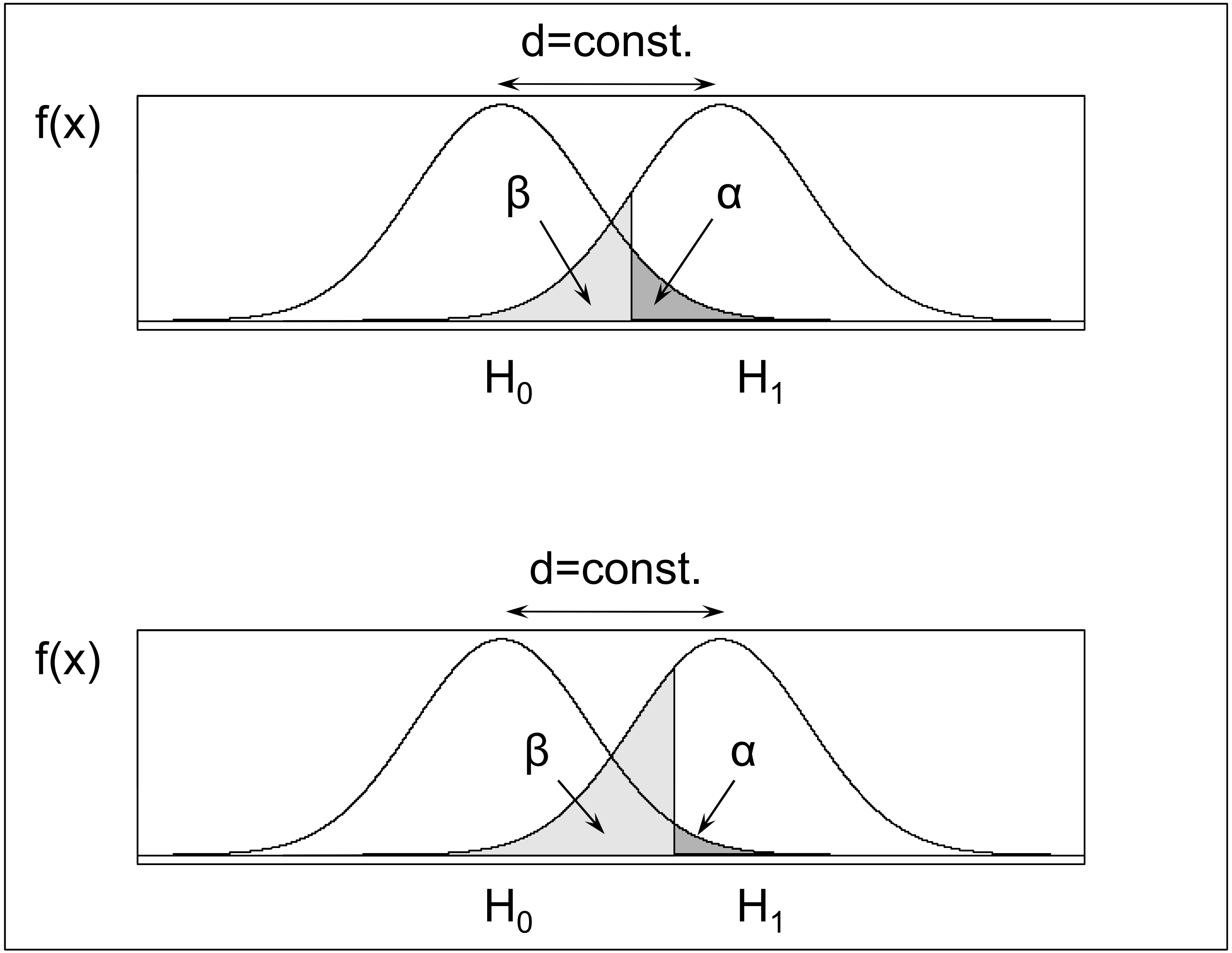
Decisions in Hypothesis Testing, α- and β-error

α- and β-error are directly related:
the lower the α-error the higher the β-error

The size of the β-error depends on:

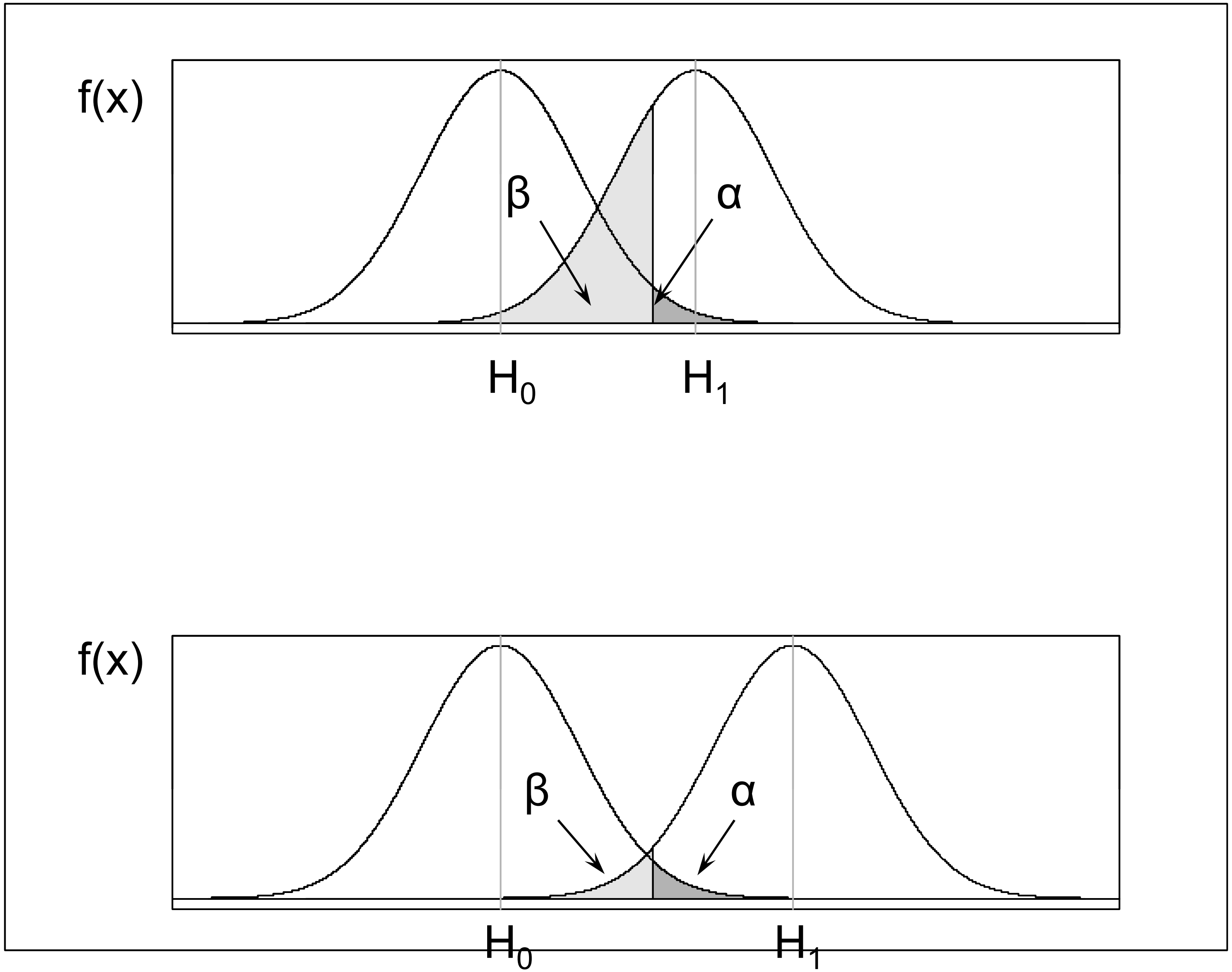
- The level of significance α
- The sample size n
- The size of the effect (e.g. $d = \frac{\mu_1 - \mu_2}{\sigma_X}$)

Decisions in Hypothesis Testing, α- and β-error



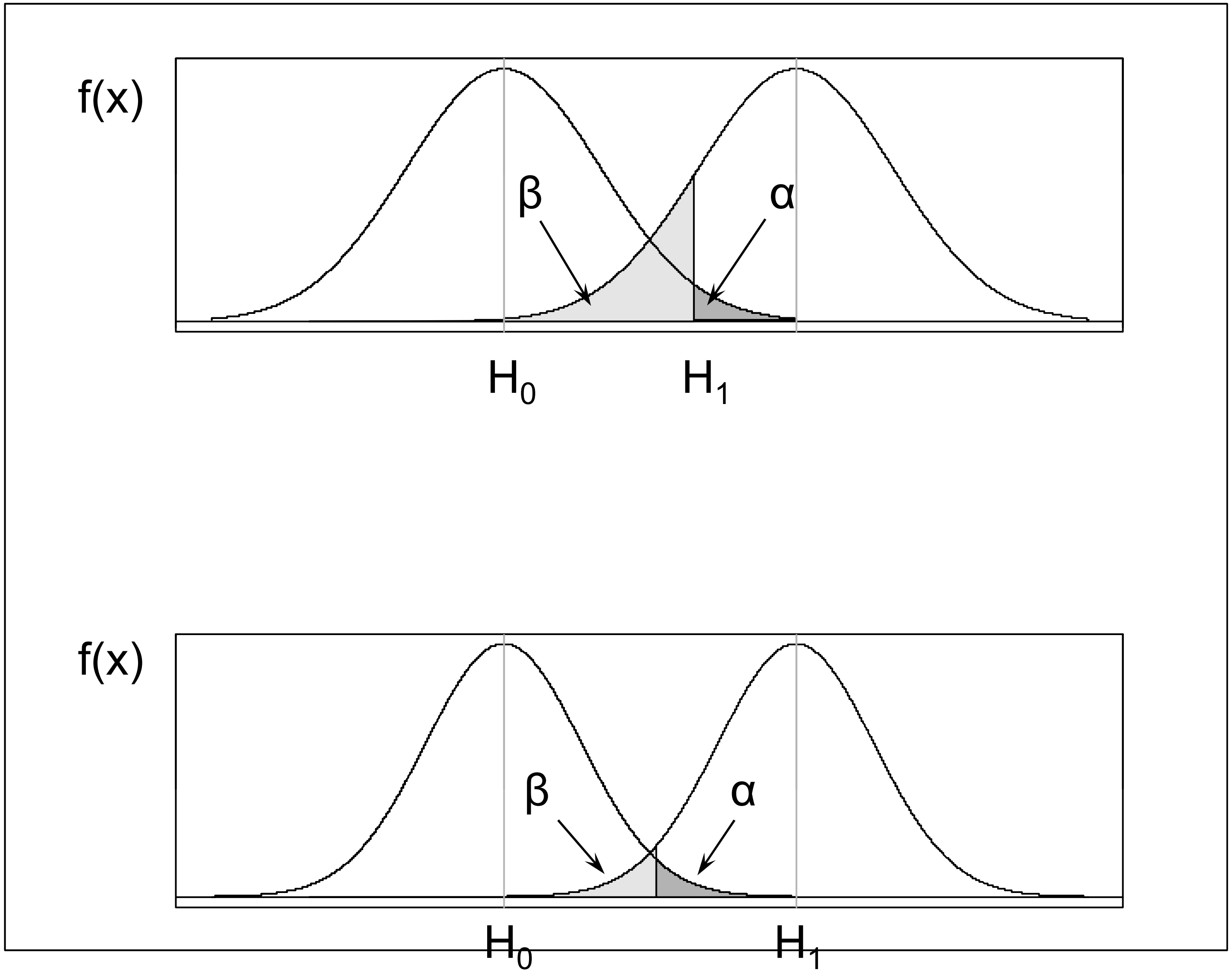
The lower the level of significance the higher the β-error.

Decisions in Hypothesis Testing, α- and β-error



The higher the effect the lower the β-error.

Decisions in Hypothesis Testing, α- and β-error



The higher sample size the lower variance, the lower the β-error.

Decisions in Hypothesis Testing, α- and β-error

- How to choose the α- and β-error?

- Example:
- H_0: Device indicates that a person has no serious disease.
- H_1: Device indicates that a person has serious disease.

- α-error?
- β-error?

Decisions in Hypothesis Testing, α- and β-error

- How to choose the α- and β-error?

- Example:
- H_0: Device indicates that a person has no serious disease.
- H_1: Device indicates that a person has serious disease.

- α-error = P(Device indicates disease when there is no disease.)
- β-error = P(Device indicates no disease when it is present.)

- How to choose the α- and β-error?

- Choose a large α-error (0.10) since β-error has more serious consequences.

Example 2: How to choose the α- and β-error?

H$_1$: $\mu_0 < \mu_1$: The new method of teaching is better than the old one (e.g. the students get better marks when solving the same exercises)

H$_0$: $\mu_0 \geq \mu_1$: The is no difference between the methods or the new one is worse than the old one.

Done: Describe the α- and β-error (and the consequences) for this example.

Type 1 error: H_0 is right (no difference) but we reject it and postulate that the new method is better. Probably there will be re-education/retraining of the teachers, new material will be purchased – all those investments are not justifiable, since the new method is not better.

Type 2 error: H_0 is accepted and consequently H_1 is rejected even though it is true and the new method is better. Then there will be no investment but we miss the opportunity to improve teaching.

New: What level of significance would you choose in this example?

My suggestion: Choose a small α since teachers in the administration in this field are very conservative (at least in Germany) and you should only introduce new methods when you are really sure that they have an effect.

When there is no difference between α-error and β-error we often choose α=0.05.

Hypothesis testing, α- and β-error, summary

- The size of α-error can be controlled by choosing the appropriate level of significance.
- The α-error and β-error are related. A decrease of one type of error always results in an increase of the probability of the other (provided the sample size n does not change).
- An increase in the sample size will reduce both α-error and β-error (provided the level of significance does not change).
- When the null hypothesis is false, the β-error increases as the true value of the parameter approaches the value hypothesized in the null hypothesis.

General Procedure for Hypotheses Tests: Tests for the Mean of a Normal Distribution, Variance known

- Step 1:
  Formulate the null hypothesis H_0, and the alternative hypothesis H_1:
  H_0: $\mu_0=\mu_1$; H_1: $\mu_0\neq\mu_1$
- Step 2:
  Choose a significance level α.
- Step 3:
  Determine the appropriate test statistic. $\hat{z} = \dfrac{\overline{x} - \mu}{\sigma/\sqrt{n}}$
- Step 4:
  State the rejection region for the test statistic.
- Step 5:
  Compute the necessary sample quantities, substitute these into the equation for the test statistic, and compute that value.
- Step 6:
  Decide whether or not H_0 should be rejected. Formulate your conclusion in terms of original problem.

General Procedure for Hypotheses Tests: Tests for the Mean of a Normal Distribution, Variance unknown

- Step 1:
  Formulate the null hypothesis H_0, and the alternative hypothesis H_1.
  H_0: $\mu_0 = \mu_1$; H_1: $\mu_0 \neq \mu_1$
- Step 2:
  Choose a significance level α.
- Step 3:
  Determine the appropriate test statistic. $\hat{t} = \frac{\bar{x} - \mu_0}{\hat{\sigma}_{\bar{X}}} = \frac{\bar{x} - \mu_0}{\hat{\sigma}/\sqrt{n}}$ (with d.f.=n-1)
- Step 4:
  State the rejection region for the test statistic.
- Step 5:
  Compute the necessary sample quantities, substitute these into the equation for the test statistic, and compute that value.
- Step 6:
  Decide whether or not H_0 should be rejected. Formulate your conclusion in terms of original problem.

General Procedure for Hypotheses Tests: Example

For a sample of 50 passenger cars we got $\overline{x} = 76\frac{km}{h}; s^2 = 120 km^2/h^2$.
Does this result differ significantly from $\mu = 80\ km/h$?

a) Assume that we know the population variance from comprehensive pre-studies with $\sigma^2 = 100 km^2/h^2$.

b) Check the assumption that the speed at this section is lower than 80 km/h. Assume that we know the population variance from comprehensive pre-studies with $\sigma^2 = 100 km^2/h^2$.

c) Assume that the population variance is unknown.

d) What would be the result in c) for a reduced sample size of n=30?