

Biostatistics in animal experiments

PD Dr. Kathrin Büttner

Unit of Biomathematics and Data Processing

Faculty of Veterinary Medicine

Justus Liebig University, Gießen

Kathrin.Buettner@vetmed.uni-giessen.de



Agenda

I. Introduction

Planning and conducting an animal experiment from a statistical point of view

II. Statistical methodology

Functioning of statistical test procedures

III. Sample size calculation or estimation

Various examples depending on the hypothesis and the study design

IV. Sample size calculation in SAS

Coding examples

V. Conclusion



I. Introduction

Biostatistics in animal experiments



I. Introduction: *Biostatistics in animal experiments*

**Steps in planning and conducting animal experiments
from a statistical perspective**

General question:

→ Where do we need or use biostatistics in animal experiments?

Answer:

→ (Nearly) at each phase!!!



I. Introduction: *Biostatistics in animal experiments*

Different phases of an animal experiment

- Phase 1: Planning
- Phase 2: Conducting the animal experiment
- Phase 3: Statistical analysis
- Phase 4: Publishing the results (Interpretation and presentation of the findings)



I. Introduction: *Biostatistics in animal experiments*

Different phases of an animal experiment

- Phase 1: Planning

Design of the study

Consideration of
confounding factors

Sample size
calculation



I. Introduction: *Biostatistics in animal experiments*

Different phases of an animal experiment

- Phase 1: Planning

Design of the study

Consideration of
confounding factors

Sample size
calculation

- Formulation of research questions and choice of appropriate study design
- Definition of dependent and independent variables (Which is the primary objective criterion and which are the influencing factors?)
- Independent samples or dependent samples (repeated measurements)
- Randomisation: Random classification of the study units to the different treatments (classification by chance, avoidance of systematic faults)
- ...



I. Introduction: *Biostatistics in animal experiments*

Different phases of an animal experiment

- Phase 1: Planning

Design of the study

Consideration of
confounding factors

Sample size
calculation

- Which factors may have an influence on the objective criterion? (e.g. sex or age of the animals)
- All possible confounding factors should be recorded (if applicable) or should be avoided by an appropriate study design
- If there are confounding factors, they have to be considered in the statistical analysis (Reduction of the random error)



I. Introduction: *Biostatistics in animal experiments*

Different phases of an animal experiment

- Phase 1: Planning

Design of the study

Consideration of
confounding factors

Sample size
calculation

- Not more animals than necessary!
- But: Not less animals than necessary!

→ **Upper and lower boundary for the number of animals which should be included in the animal experiment**



I. Introduction: *Biostatistics in animal experiments*

Different phases of an animal experiment

- Phase 1: Planning

Design of the study

Consideration of
confounding factors

Sample size
calculation

→ **All these aspects have to be clearly specified
in the study protocol**



I. Introduction: *Biostatistics in animal experiments*

Different phases of an animal experiment

- Phase 2: Conducting the animal experiment
(with special focus on data recording)
 - Appropriate data storage
 - Regular transfer of data in digital format
 - Regular backups
 - Data checking (Typing errors und plausibility check)
 - Validation of the used measurement methodology (e.g. recalibration)

Calculate simple statistical parameters (e.g. mean, standard deviation, minimum and maximum) to detect transcription errors



I. Introduction: *Biostatistics in animal experiments*

Different phases of an animal experiment

- Phase 3: Statistical analysis

→ The best data are worthless without the appropriate data analysis

Descriptive

Confirmatory

Explorative



I. Introduction: *Biostatistics in animal experiments*

Different phases of an animal experiment

- Phase 3: Statistical analysis

→ The best data are worthless without the appropriate data analysis

Descriptive

Confirmatory

Explorative

- Arrange data
- Condensate data
- Create graphics of the date

Can and should be performed already in
Phase 2: data recording (plausibility check)

Is part of every statistical analysis!!!!



I. Introduction: *Biostatistics in animal experiments*

Different phases of an animal experiment

- Phase 3: Statistical analysis

→ The best data are worthless without the appropriate data analysis

Descriptive

Confirmatory

Explorative

- Research questions are translated in hypotheses
which are tested with the recorded data



I. Introduction: *Biostatistics in animal experiments*

Different phases of an animal experiment

- Phase 3: Statistical analysis

→ The best data are worthless without the appropriate data analysis

Descriptive

Confirmatory

Explorative

- Extra (new) hypotheses are created based on the recorded data

- Interesting patterns arise during data analysis which were not considered yet

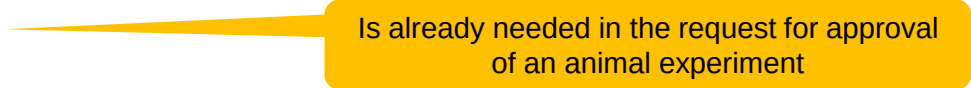
- New research questions for the next study

- Attention: New hypotheses have to be tested on a new data set



I. Introduction: *Biostatistics in animal experiments*

Different phases of an animal experiment

- Phase 4: Publishing the results (Interpretation and presentation of the findings)
 - Correct description of the performed statistical analysis (comprehensive materials & methods section)
 Is already needed in the request for approval of an animal experiment
 - Correct interpretation of the results (e.g. interpretation of interactions between different effects, disclosure of all confounding factors, ...)
 - Correct interpretation of statistical significances (p values)
 - Attention: Multiple comparisons
 - Suitable illustration of the results (e.g. tables and figures)
 - What happens with insignificant results?
 - Insignificant results are also results, try to avoid publication bias



II. Statistical methodology

Functioning of statistical test procedures?



II. Statistical methodology

Basic scheme of statistical tests

Research question → Hypotheses → Test statistics
→ Test decision → Results / Interpretation

- **Intend:** With an animal experiment, it should be tested, whether specific hypotheses can be accepted or should be rejected
 - In order that the statistical test is able to answer this question, the amount of data should be large enough
 - **Attention:** Sample size too large
 - Duration of the experiment ↑, Costs ↑, Ethical aspects?
- Sample size calculation prior to study begin (→ request for approval)



II. Statistical methodology

Basic scheme of statistical tests

Research question → Hypotheses → Test statistics
→ Test decision → Results / Interpretation

- **Research question:**
 - Exact and detailed definition of the scientific research question
 - Depending on research question you have to choose the appropriate study design
 - Depending of research question (and data basis) the appropriate statistical test has to be chosen



II. Statistical methodology

Basic scheme of statistical tests

Research question → **Hypotheses** → Test statistics
→ Test decision → Results / Interpretation

- **Hypotheses** (Translation of the practical research question)
 - Null hypothesis H_0 (e.g. assumption that on average there are no differences between two treatments) vs. Alternative hypothesis H_A (Opposite of H_0)
 - The statistical test gives an objective decision whether to choose H_0 or H_A which is based on the underlying sample and the probability theory
- Unavoidable disadvantage: Erroneous decisions for H_0 or H_A are possible



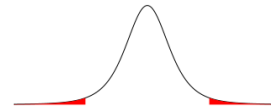
II. Statistical methodology

Basic scheme of statistical tests

Research question → **Hypotheses** → Test statistics
→ Test decision → Results / Interpretation

- **Hypotheses:** One-sided or two-sided?

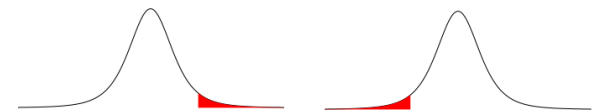
- **Two-sided:** Test of differences



- **Example:** Differ two breeds in their milk yield?

- $H_0 : \mu_{Breed\ 1} = \mu_{Breed\ 2}; H_A : \mu_{Breed\ 1} \neq \mu_{Breed\ 2}$

- **One-sided:** Test of increase or decrease



- **Example:** Has Breed 1 a higher milk yield compared to Breed 2?

- $H_0 : \mu_{Breed\ 1} \leq \mu_{Breed\ 2}; H_A : \mu_{Breed\ 1} > \mu_{Breed\ 2}$



II. Statistical methodology

Probabilities of error in statistical procedures

- Decision for H_0 or H_A is always related with statistical uncertainty
 - Wrong decisions are possible in every statistical test

		Test decision	
		Do not reject H_0	Reject H_0
Real situation	H_0 right	$1 - \alpha$	α
	H_0 wrong	β	$1 - \beta$

Type I error / α error:
Probability of erroneous rejection of H_0
Probability of finding a difference between the treatments when the difference does not exist

Type II error / β error:
Probability of erroneously not reject H_0
Probability of failing to find a difference between the treatments when a difference exist

Power of the test:
Probability of the correct rejection of H_0
Probability that H_0 can be rejected if the treatments differ

- α and β depend on each other $\rightarrow$ Minimising of both errors is problematic
- Constant value for α (e.g. $\alpha = 0.05$), Minimising β ($\rightarrow$ sample size calculation)



II. Statistical methodology

Basic scheme of statistical tests

Research question → Hypotheses → Test statistics
→ Test decision → Results / Interpretation

Test statistics

- Depends on the test problem → Has to be chosen according to the research question and the data basis
- Value of the test statistics (which is calculated from the data) determines the test decision



II. Statistical methodology

Basic scheme of statistical tests

Research question → Hypotheses → Test statistics
→ **Test decision** → Results / Interpretation

Test decision

- Three different alternatives with the same result
 1. Comparison of the test statistic with the critical value (classic → lectures in biostatistics)
 2. Comparison of p values with the significance level α (most common → statistical software packages)
 3. Calculation of confidence limits (rarely used)



II. Statistical methodology

Basic scheme of statistical tests

Research question → Hypotheses → Test statistics
→ Test decision → Results / Interpretation

Results / Interpretation

- (Back-) Translation of the Test results to the original research question
→ What does a significant result mean?



III. Sample size calculation



III. Sample size calculation: *General aspects*

Phase 1: Planning

Design of the study

Consideration of
confounding factors

Sample size
calculation

- Not more animals than necessary!
- But: Not less animals than necessary!



III. Sample size calculation: *General aspects*

**Guidelines for the request of approval of an animal experiment
„Leitfaden zur Erstellung eines Antrages auf Genehmigung eines
Tierversuches“ (Regierungspräsidium Gießen, Stand März 2016)**

- The number of animals which should be included in the animal experiment has to be chosen as low as possible, but as high as necessary
- A detailed biometrical planning is needed with specification of the planned statistical analyses and the primary objective criterion



III. Sample size calculation: *General aspects*

**Guidelines for the request of approval of an animal experiment
„Leitfaden zur Erstellung eines Antrages auf Genehmigung eines
Tierversuches“ (Regierungspräsidium Gießen, Stand März 2016)**

- The number of animals which should be included in the animal experiment has to be chosen as low as possible, but as high as necessary
 - A detailed biometrical planning is needed with specification of the planned statistical analyses and the primary objective criterion
-
- **Not more animals than necessary,
but also not less animals than necessary!**
 - **How can these requirements be fulfilled?**



III. Sample size calculation: *General aspects*

An appropriate sample size ensures that

... an existing effect can be detected

... with a specific certainty one can assume that there is no relevant effect, if the statistical analysis does not reveal significant differences

		Test decision	
		Do not reject H_0	Reject H_0
Real situation	H_0 right	$1 - \alpha$	α
	H_0 wrong	β	$1 - \beta$

Type I error / α error:
Probability of erroneous rejection of H_0
Probability of finding a difference between the treatments when the difference does not exist

Type II error / β error:
Probability of erroneously not reject H_0
Probability of failing to find a difference between the treatments when a difference exist

Power of the test:
Probability of the correct rejection of H_0
Probability that H_0 can be rejected if the treatments differ



III. Sample size calculation: *General aspects*

First step planning!!!

- **Prospective studies**

- Sample size calculation **prior to** data recording
- Sample size calculation is part of the study protocol
 - A priori power analysis
 - Mandatory for clinical studies and animal experiments

- **Retrospective studies**

- Data are already recorded (e.g. patient data from clinical units)
- Sample size calculation is also meaningful in retrospective studies to achieve the optimal data basis for analyses



III. Sample size calculation: *General aspects*

Requirements for sample size calculation

- **Main objective criterion** and **research question** (one-sided or two-sided)
- Choice of the **appropriate statistical test** according to the research question and the data basis
- **Smallest detectable difference:** Determination of the smallest clinically or biologically relevant difference which should be detected
- **Standard deviation** of the main objective criterion
- **Significance level α :** Probability to achieve a significant test result, although no difference exists (usually $\alpha = 0.05$)
- **Power of the study ($1 - \beta$):** Probability to detect a real difference with the statistical test



III. Sample size calculation: *General aspects*

Requirements for sample size calculation

- For every sample size calculation a close collaboration with the statistician is needed and mandatory
- Sample size calculation is highly interdisciplinary
- Precise knowledge about the data is needed in order to achieve a meaningful sample size calculation
- Assumptions about the expected effects, the distribution of the data can be found in literature review, pilot studies or expert opinions

→ **Every sample size calculation is customized for every single study!!!**



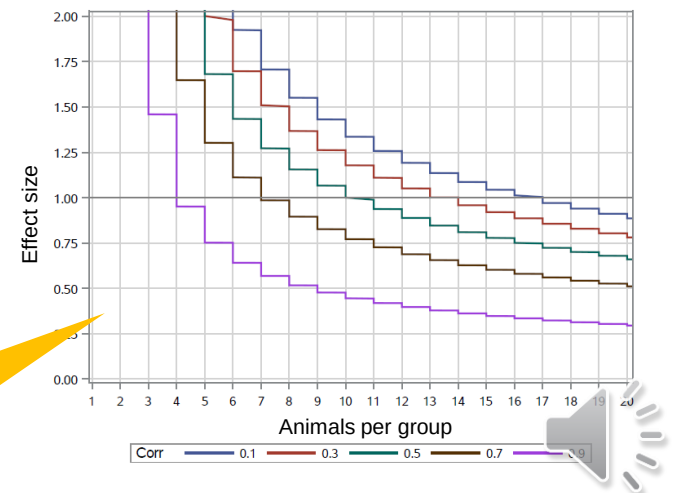
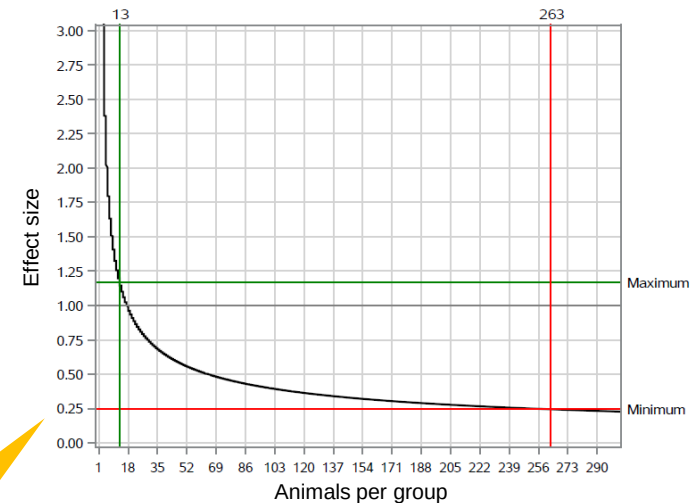
III. Sample size calculation: *General aspects*

It is still an estimation

- Also with exact planning and high expertise the sample size calculation remains frequently a rough estimate
- Sensitivity analysis can determine the effects of changing input parameters in order to get a better impression of the needed sample size

Impact of effect size (smallest detectable difference / standard deviation) on the number of animals needed

Effects of different correlations between two parameters in combination with the effect size (smallest detectable difference / standard deviation) on the number of animals needed



III. Sample size calculation: *General aspects*

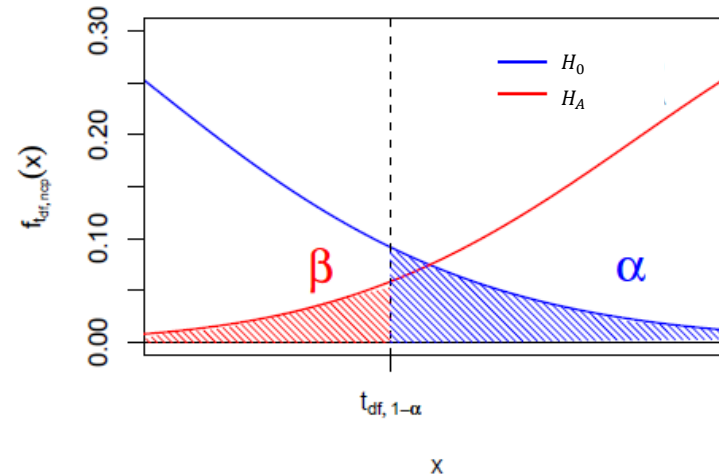
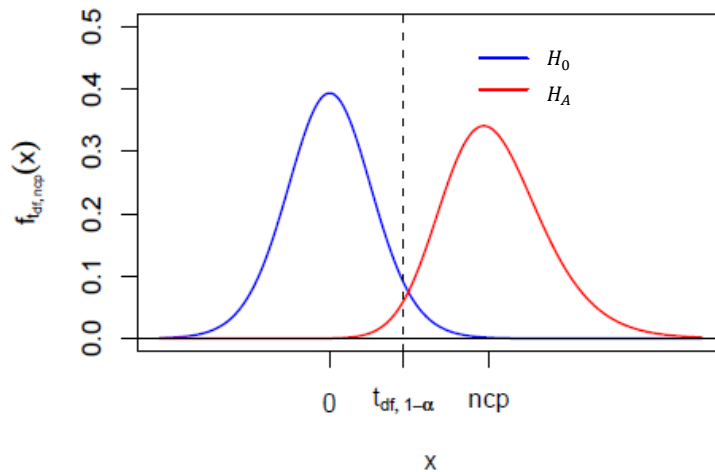
Two types of sample size calculation

- **Classical type:** With assumptions about the expected difference or the expected dependencies the needed sample size is estimated
 - **Question: Which sample size is needed in order to significantly detect this difference/dependency?**
- **Detectable difference or dependency with a fixed sample size:** Based on the possible sample size, it is calculated which difference or dependency can be significantly detected
 - **Which difference/dependency can be detected with the available sample size?**



III. Sample size calculation: *General aspects*

Influencing factors on the power of a two sample *t*-Tests



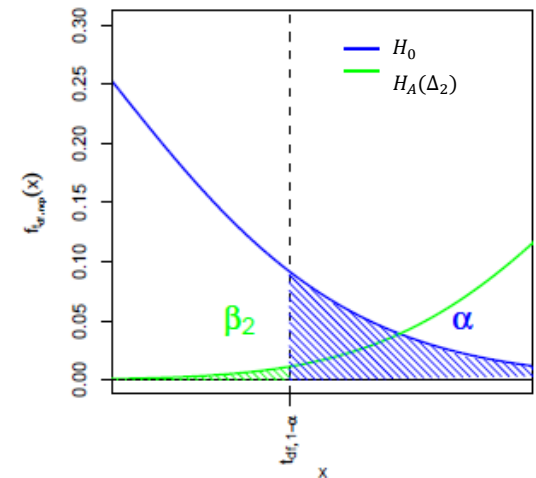
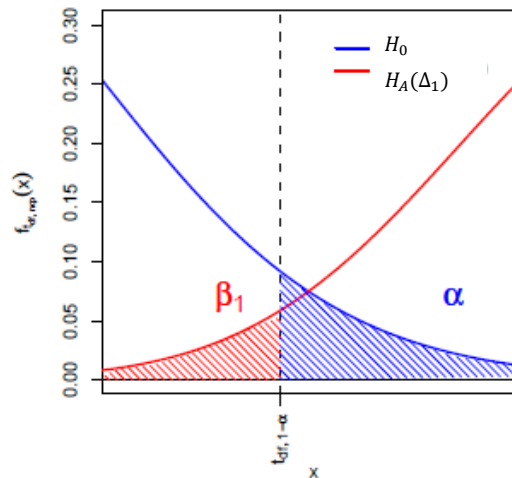
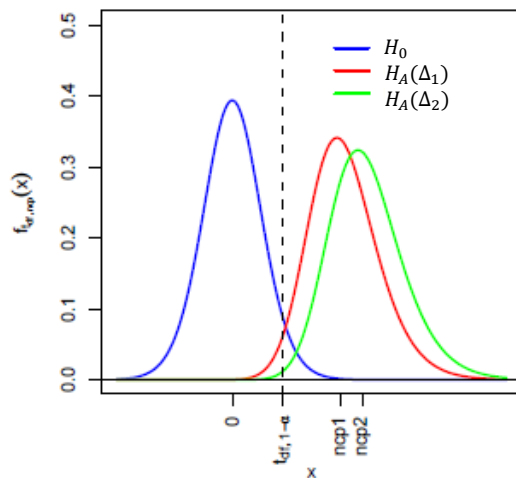
- α and β error depend on each other (α increases, β decreases and vice versa)
 - Simultaneous minimising of both errors is not possible
 - Way of proceeding: chose constant value for α & minimising β
- Non-central distribution & $ncp \rightarrow$ Non-centrality parameter
 - Distribution of the test statistics under the alternative hypothesis



III. Sample size calculation: *General aspects*

Influencing factors on the power of a two sample *t*-Tests

Impact of the detectable difference $\Delta = \mu_X - \mu_Y$ $(\Delta_1 < \Delta_2)$



- Larger $\Delta \rightarrow$ larger $ncp \rightarrow$ smaller $\beta \rightarrow$ higher power
- The higher Δ becomes, the higher is the power of the test

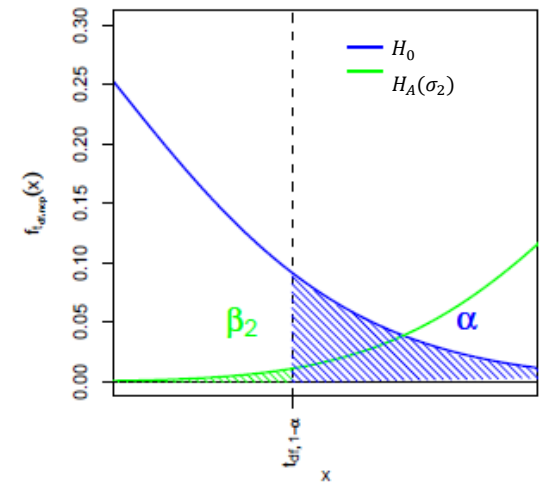
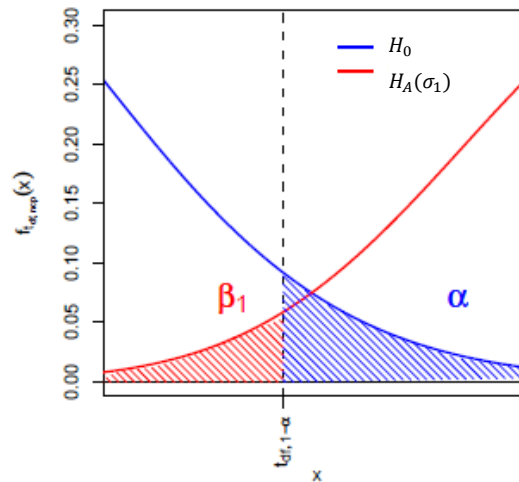
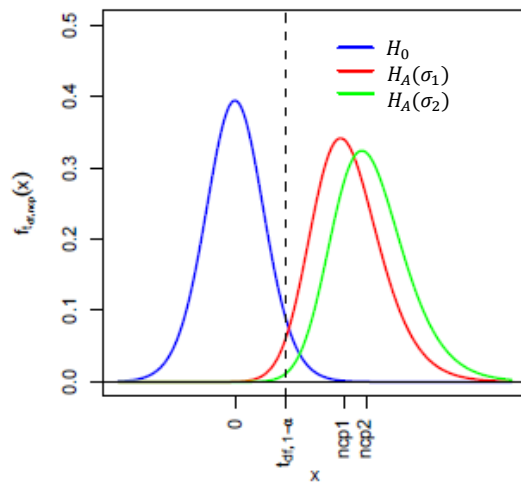


III. Sample size calculation: *General aspects*

Influencing factors on the power of a two sample t -Tests

Impact of the standard deviation σ

$(\sigma_1 > \sigma_2)$



- Smaller $\sigma \rightarrow$ larger $ncp \rightarrow$ smaller $\beta \rightarrow$ higher power
- The smaller σ , the higher the power of the test

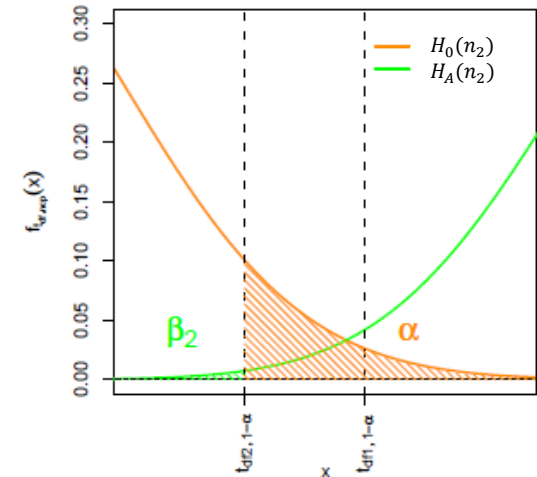
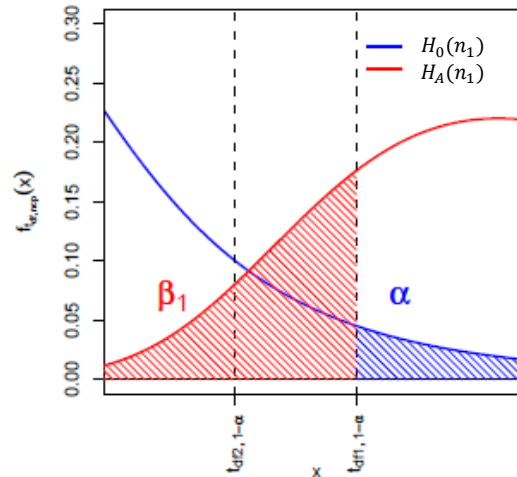
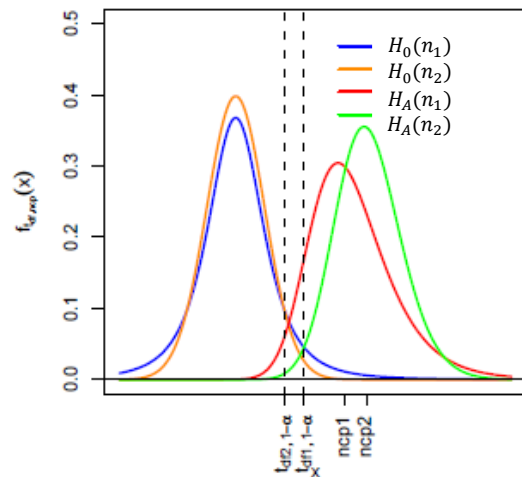


III. Sample size calculation: *General aspects*

Influencing factors on the power of a two sample *t*-Tests

Impact of the sample size n

$(n_1 < n_2)$



- Larger $n \rightarrow$ larger ncp & larger $df \rightarrow$ smaller $\beta \rightarrow$ higher power
- The larger n , the higher the power of the test



III. Sample size calculation: *General aspects*

**Connection between the influencing factors
and the power of the test or sample size**

- **Power increases (β error decreases) with ...**
 - ... increasing α error
 - ... increasing detectable difference
 - ... increasing sample size n
 - ... decreasing standard deviation σ
- **Number of needed samples increases with ...**
 - ... decreasing α error
 - ... decreasing detectable difference
 - ... increasing power $1 - \beta$
 - ... increasing standard deviation σ



III. Sample size calculation

Practical implementation



III. Sample size calculation: *Practical implementation*

Sample size calculation: 3 questions which have to be answered

- Which risk are you willing to take, to say that two treatments are significantly different, although they are not?
→ α error
- Which risk are you willing to take, to not detect a difference between two treatments when there is a significant difference?
→ β error
- Which is the smallest detectable difference between the two treatments which should be detected?
→ Δ = smallest detectable difference



III. Sample size calculation: *Practical implementation*

Sample size calculation: 3 questions which have to be answered

- Which risk are you willing to take, to say that two treatments are significantly different, although they are not?

→ α error

No risk

- Which risk are you willing to take, to not detect a difference between two treatments when there is a significant difference?

→ β error

No risk

- Which is the smallest detectable difference between the two treatments which should be detected?

→ Δ = smallest detectable difference

All differences are important,
even the smallest ones!



III. Sample size calculation: *Practical implementation*

Sample size calculation: 3 questions which have to be answered

- Which risk are you willing to take, to say that two treatments are significantly different, although they are not?

→ α error

No risk

- Which risk are you willing to take, to not say that two treatments are significantly different, although there is a significant difference?

→ β error

No risk

**With these assumptions
an infinite number of observations
in each treatment group is needed!**

- Which is the smallest detectable difference between the two treatments which should be detected?

→ Δ = smallest detectable difference

All differences are important,
even the smallest ones!



t -Test for independent samples and equal variances

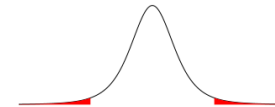
**Approximate solution
for the sample size calculation**



III. Sample size calculation: *t*-Test (for independent samples)

Approximate solution for sample size calculation

→ **Two-sided test direction: Test of differences**



$$N_0 = \frac{[u_{1-\alpha/2} + u_{1-\beta}]^2}{[\Delta/(2 \times \sigma)]^2}$$

with

$$\Delta = \mu_1 - \mu_2$$

σ = Standard deviation

$u_{1-\alpha/2}$ = $(1 - \alpha/2)$ -quantile of the standard normal distribution

$u_{1-\beta}$ = $(1 - \beta)$ -quantile of the standard normal distribution

q -quantiles u_q of the standard normal distribution

q	0.750	0.800	0.900	0.950	0.975	0.990	0.995
u_q	0.6745	0.8416	1.2812	1.6449	1.9600	2.3264	2.5758

Example: $\alpha = 0.05$
 $u_{1-\alpha/2} = u_{1-0.05/2} = u_{0.975} = 1.96$

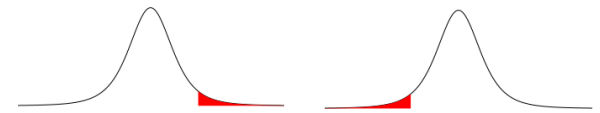


III. Sample size calculation: *t*-Test (for independent samples)

Approximate solution for sample size calculation

→ **One-sided test direction: Test of decrease or increase**

$$N_0 = \frac{[u_{1-\alpha} + u_{1-\beta}]^2}{[\Delta/(2 \times \sigma)]^2}$$



mit

$$\Delta = \mu_1 - \mu_2$$

σ = Standard deviation

$u_{1-\alpha}$ = $(1 - \alpha)$ -quantile of the standard normal distribution

$u_{1-\beta}$ = $(1 - \beta)$ -quantile of the standard normal distribution

q -quantiles u_q of the standard normal distribution

q	0.750	0.800	0.900	0.950	0.975	0.990	0.995
u_q	0.6745	0.8416	1.2812	1.6449	1.9600	2.3264	2.5758



III. Sample size calculation: *t*-Test (for independent samples)

Approximate solution for sample size calculation

→ Total sample size N for the one- and two-sided *t*-Test with an $\alpha = 0.05$ and specified Power ($1 - \beta$). The total sample size has to be separated into two equal sized groups.

Effect size

Two sample *t*-Test

Δ/σ	Power = 0.80		Power = 0.90		Power = 0.95	
	one-sided	two-sided	one-sided	two-sided	one-sided	two-sided
0.20	620	788	858	1054	1084	1302
0.25	398	506	550	676	694	834
0.30	278	352	382	470	484	580
0.35	204	260	282	346	356	428
0.40	156	200	216	266	272	328
0.45	124	158	172	210	216	260
0.50	102	128	140	172	176	210
0.55	84	106	116	142	146	174
0.60	72	90	98	120	122	148
0.65	60	78	84	102	104	126
0.70	52	68	72	88	90	110
0.75	46	58	64	78	80	96
0.80	42	52	56	68	70	84
0.85	36	46	50	62	62	74
0.90	32	42	44	54	56	68
0.95	30	38	40	50	50	60
1.00	28	34	36	46	46	54
1.05	24	32	34	42	42	50
1.10	22	30	30	38	38	46

Δ/σ	Power = 0.80		Power = 0.90		Power = 0.95	
	one-sided	two-sided	one-sided	two-sided	one-sided	two-sided
1.15	22	26	28	34	36	42
1.20	20	24	26	32	32	40
1.25	18	24	24	30	30	36
1.30	18	22	22	28	28	34
1.35	16	20	22	26	26	32
1.40	16	20	20	24	24	30
1.45	14	18	18	24	24	28
1.50	14	18	18	22	22	26
1.55	12	16	16	20	20	24
1.60	12	16	16	20	20	24
1.65	12	14	16	18	18	22
1.70	12	14	14	18	18	22
1.75	10	14	14	16	16	20
1.80	10	12	14	16	16	20
1.85	10	12	12	16	16	18
1.90	10	12	12	14	14	18
1.95	10	12	12	14	14	16
2.00	8	12	12	14	14	16
2.05	8	10	10	14	12	16

III. Sample size calculation: *t*-Test (for independent samples)

Approximate solution for sample size calculation

→ Total sample size N for the one- and two-sided *t*-Test with an $\alpha = 0.05$ and specified Power ($1 - \beta$). The total sample size has to be separated into two equal sized groups.

Effect size

Two sample *t*-Test

Δ/σ	Power = 0.80		Power = 0.90		Power = 0.95	
	one-sided	two-sided	one-sided	two-sided	one-sided	two-sided
0.20	620	788	858	1054	1084	1302
0.25	398	506	550	676	694	834
0.30	278	352	382	470	484	580
0.35	204	260	282	346	356	428
0.40	156	200	216	266	272	328
0.45	124	158	172	210	216	260
0.50	102	128	140	172	176	210
0.55	84	106	116	142	146	174
0.60	72	90	98	120	122	148
0.65	60	78	84	102	104	126
0.70	52	68	72	88	90	110
0.75	46	58	64	78	80	96
0.80	42	52	56	68	70	84
0.85	36	46	50	62	62	74
0.90	32	42	44	54	56	68
0.95	30	38	40	50	50	60
1.00	28	34	36	46	46	54
1.05	24	32	34	42	42	50
1.10	22	30	30	38	38	46

Δ/σ	Power = 0.80		Power = 0.90		Power = 0.95	
	one-sided	two-sided	one-sided	two-sided	one-sided	two-sided
1.15	22	24	28	34	36	42
1.20	20	24	28	32	32	40
1.25	18	24	28	30	30	36
1.30	18	22	28	28	28	34
1.35	16	22	28	28	28	34
1.40	16	22	28	28	28	34
1.45	14	22	28	28	28	34
1.50	14	22	28	28	28	34
1.55	12	22	28	28	28	34
1.60	12	22	28	28	28	34
1.65	12	22	28	28	28	34
1.70	12	22	28	28	28	34
1.75	10	22	28	28	28	34
1.80	10	22	28	28	28	34
1.85	10	22	28	28	28	34
1.90	10	22	28	28	28	34
1.95	10	12	12	14	14	16
2.00	8	12	12	14	14	16
2.05	8	10	10	14	12	16

- Both group do not have to be of equal size
- However, two equal sized groups provide the highest power
- Unequal sized groups need with the same settings a larger sample size

t -Test for independent samples and equal variances

Example of application:

Feeding experiment in fattening pigs



III. Sample size calculation: *Feeding experiment*

Example of application: Feeding experiment in pigs –
Effect of two different types of feed on the daily weight gain
of fattening pigs



- Two independent groups of pigs fed with one of the two feed types
- Potential confounding factors (e.g. breed, husbandry, ...) are kept constant

→ **Question: How many pigs have to be included in the study, if two groups of the same size are planned?**



III. Sample size calculation: *Feeding experiment*

Example of application: Feeding experiment in pigs –
Effect of two different types of feed on the daily weight gain
of fattening pigs



→ **Which information is needed for the sample size calculation?**

- **Objective criterion:** Daily weight gain → Assumption: Normal distribution
→ Two-sample t -Test for independent samples
- **Smallest relevant difference:** $\Delta = 30 \text{ g}$
→ A difference of 30 g in the daily weight gain
between the two groups represents a relevant benefit
- **Standard deviation:** $\sigma = 60 \text{ g}$

Effect size: $\Delta/\sigma = 30/60 = 0.5$



III. Sample size calculation: *Feeding experiment*

Example of application: Feeding experiment in pigs –
Effect of two different types of feed on the daily weight gain
of fattening pigs



→ **Which information is needed for the sample size calculation?**

- **Type I error:** $\alpha = 0.05$
- **Type II error:** $\beta = 0.20$
→ $\text{Power} = 1 - \beta = 0.80$
- **Test direction:** two-sided

		Test decision	
		Do not reject H_0	Reject H_0
Real situation	H_0 right	$1 - \alpha$	α
	H_0 wrong	β	$1 - \beta$

Type I error / α error:
Probability of erroneous rejection of H_0
Probability of finding a difference between the treatments when the difference does not exist

Type II error / β error:
Probability of erroneously not reject H_0
Probability of failing to find a difference between the treatments when a difference exist

Power of the test:
Probability of the correct rejection of H_0
Probability that H_0 can be rejected if the treatments differ



III. Sample size calculation: *Feeding experiment*

Example of application: Feeding experiment in pigs –
Effect of two different types of feed on the daily weight gain
of fattening pigs



Approximate solution for sample size calculation

→ **Two-sided**

1. Determination of the starting point

$$N_0 = \frac{[u_{1-\alpha/2} + u_{1-\beta}]^2}{[\Delta/(2 \times \sigma)]^2} = \frac{[u_{0.975} + u_{0.80}]^2}{[30/(2 \times 60)]^2} = \frac{[1.96 + 0.8416]^2}{[30/(2 \times 60)]^2}$$

q -quantiles u_q of the standard normal distribution

q	0.750	0.800	0.900	0.950	0.975	0.990	0.995
u_q	0.6745	0.8416	1.2812	1.6449	1.9600	2.3264	2.5758

Summary

- **Objective criterion:**
daily weight gain
- **Statistical test:**
t-Test for independent samples
- **Relevant difference:**
 $\Delta = 30 \text{ g}$
- **Standard deviation:**
 $\sigma = 60 \text{ g}$
- **Type I error:** $\alpha = 0.05$
- **Type II error:** $\beta = 0.20$
- **Test direction:**
two-sided



III. Sample size calculation: *Feeding experiment*

Example of application: Feeding experiment in pigs –
Effect of two different types of feed on the daily weight gain
of fattening pigs



Approximate solution for sample size calculation

→ **Two-sided**

1. Determination of the starting point

$$N_0 = \frac{[u_{1-\alpha/2} + u_{1-\beta}]^2}{[\Delta/(2 \times \sigma)]^2} = \frac{[u_{0.975} + u_{0.80}]^2}{[30/(2 \times 60)]^2} = \frac{[1.96 + 0.8416]^2}{[30/(2 \times 60)]^2}$$
$$= \frac{7.8490}{0.0625} = 125.584 \approx 126 \text{ animals}$$

Summary

- **Objective criterion:**
daily weight gain
- **Statistical test:**
t-Test for independent samples
- **Relevant difference:**
 $\Delta = 30 \text{ g}$
- **Standard deviation:**
 $\sigma = 60 \text{ g}$
- **Type I error:** $\alpha = 0.05$
- **Type II error:** $\beta = 0.20$
- **Test direction:**
two-sided

Iterative procedure → Formula provides only a
start value for the total sample size



III. Sample size calculation: *Feeding experiment*

Example of application: Feeding experiment in pigs –
Effect of two different types of feed on the daily weight gain
of fattening pigs



Approximate solution for sample size calculation

→ **Two-sided**

2. Iterative procedure continued

– with quantiles of the t distribution

$$126 \geq \frac{[t_{1-\alpha/2, df} + t_{1-\beta, df}]^2}{[\Delta/(2 \times \sigma)]^2} = \frac{[t_{0.975, 124} + t_{0.80, 124}]^2}{[30/(2 \times 60)]^2}$$
$$= \frac{[1.979 + 0.845]^2}{[30/(2 \times 60)]^2} = \frac{7.975}{0.0625} = 127.6$$

The inequality for $N = 126$ is not fulfilled
→ another iteration step is needed

Summary

- **Objective criterion:**
daily weight gain
- **Statistical test:**
t-Test for independent samples
- **Relevant difference:**
 $\Delta = 30 \text{ g}$
- **Standard deviation:**
 $\sigma = 60 \text{ g}$
- **Type I error:** $\alpha = 0.05$
- **Type II error:** $\beta = 0.20$
- **Test direction:**
two-sided



III. Sample size calculation: *Feeding experiment*

Example of application: Feeding experiment in pigs –
Effect of two different types of feed on the daily weight gain
of fattening pigs



Approximate solution for sample size calculation

→ **Two-sided**

3. Iterative procedure continued

– with $N = 128$

$$128 \geq \frac{[t_{1-\alpha/2, df} + t_{1-\beta, df}]^2}{[\Delta/(2 \times \sigma)]^2} = \frac{[t_{0.975, 126} + t_{0.80, 126}]^2}{[30/(2 \times 60)]^2}$$
$$= \frac{[1.979 + 0.844]^2}{[30/(2 \times 60)]^2} = \frac{7.969}{0.0625} = 127.5$$

For a total sample size $N = 128$
the inequality is fulfilled

Summary

- **Objective criterion:**
daily weight gain
- **Statistical test:**
t-Test for independent samples
- **Relevant difference:**
 $\Delta = 30 \text{ g}$
- **Standard deviation:**
 $\sigma = 60 \text{ g}$
- **Type I error:** $\alpha = 0.05$
- **Type II error:** $\beta = 0.20$
- **Test direction:**
two-sided



III. Sample size calculation: *Feeding experiment*

Example of application: Feeding experiment in pigs –
Effect of two different types of feed on the daily weight gain
of fattening pigs



Approximate solution for sample size calculation

→ **Two-sided**

Two sample *t*-Test

Δ/σ	Power = 0.80		Power = 0.90		Power = 0.95	
	one-sided	two-sided	one-sided	two-sided	one-sided	two-sided
0.20	620	788	858	1054	1084	1302
0.25	398	506	550	676	694	834
0.30	278	352	382	470	484	580
0.35	204	260	282	346	356	428
0.40	156	200	216	266	272	328
0.45	124	158	172	210	216	260
0.50	102	128	140	172	176	210
0.55	84	106	116	142	146	174
0.60	72	90	98	120	122	148
0.65	60	78	84	102	104	126

The result fits with the table entries.
In each group, 64 animals are needed with the given settings.

Summary

- **Objective criterion:**
daily weight gain
- **Statistical test:**
t-Test for independent samples
- **Relevant difference:**
 $\Delta = 30\text{ g}$
- **Standard deviation:**
 $\sigma = 60\text{ g}$
- **Type I error:** $\alpha = 0.05$
- **Type II error:** $\beta = 0.20$
- **Test direction:**
two-sided



t -Test for independent samples and equal variances

Example of application:

Feeding experiment in fattening pigs

(changed assumptions!)



III. Sample size calculation: *Feeding experiment* (Part 2)

Example of application: Feeding experiment in pigs –
Effect of two different types of feed on the daily weight gain
of fattening pigs



Two sample *t*-Test

Δ/σ	Power = 0.80		Power = 0.90		Power = 0.95	
	one-sided	two-sided	one-sided	two-sided	one-sided	two-sided
0.20	620	788	858	1054	1084	1302
0.25	398	506	550	676	694	854
0.30	278	352	382	470	484	580
0.35	204	260	282	346	356	428
0.40	156	200	216	266	272	328
0.45	124	158	172	210	216	260
0.50	102	128	140	172	176	210
0.55	84	106	116	142	146	174
0.60	72	90	98	120	122	148
0.65	60	78	84	102	104	126
0.65	60	78	84	102	104	126
0.70	52	68	72	88	90	110
0.75	46	58	64	78	80	96
0.80	42	52	56	68	70	84
0.85	36	46	50	62	62	74
0.90	32	42	44	54	56	68
0.95	30	38	40	50	50	60
1.00	28	34	36	46	46	54
1.05	24	32	34	42	42	50
1.10	22	30	30	38	38	46

Increasing the power to 90 %
($\rightarrow \beta = 0.10$):
 $N = 172$

Summary

- **Objective criterion:**
daily weight gain
- **Statistical test:**
t-Test for independent samples
- **Relevant difference:**
 $\Delta = 30 \text{ g}$
- **Standard deviation:**
 $\sigma = 60 \text{ g}$
- **Type I error:** $\alpha = 0.05$
- **Type II error:** $\beta = 0.20$
- **Test direction:**
two-sided



III. Sample size calculation: *Feeding experiment* (Part 2)

Example of application: Feeding experiment in pigs –
Effect of two different types of feed on the daily weight gain
of fattening pigs



Two sample *t*-Test

Δ/σ	Power = 0.80		Power = 0.90		Power = 0.95	
	one-sided	two-sided	one-sided	two-sided	one-sided	two-sided
0.20	620	788	858	1054	1084	1302
0.25	398	506	550	676	694	854
0.30	278	352	382	470	484	580
0.35	204	260	282	346	356	428
0.40	156	200	216	266	272	328
0.45	124	158	172	210	216	260
0.50	102	128	140	172	176	210
0.55	84	106	116	142	146	174
0.60	72	90	98	120	122	148
0.65	60	78	84	102	104	126
0.65	60	78	84	102	104	126
0.70	52	68	72	88	90	110
0.75	46	58	64	78	80	96
0.80	42	52	56	68	70	84
0.85	36	46	50	62	62	74
0.90	32	42	44	54	56	66
0.95	30	38	40	50	52	60
1.00	28	34	36	46	46	54
1.05	24	32	34	42	42	50
1.10	22	30	30	38	38	46

Increasing the power to 90 %
($\rightarrow \beta = 0.10$):
 $N = 172$

A difference of 60 g is seen as
relevant, e.g. $\Delta/\sigma = 60/60 = 1$:
 $N = 34$

Summary

- **Objective criterion:**
daily weight gain
- **Statistical test:**
t-Test for independent samples
- **Relevant difference:**
 $\Delta = 30 \text{ g}$
- **Standard deviation:**
 $\sigma = 60 \text{ g}$
- **Type I error:** $\alpha = 0.05$
- **Type II error:** $\beta = 0.20$
- **Test direction:**
two-sided



t -Test for independent samples and equal variances

Example of application:

Ketosis in dairy cattle



III. Sample size calculation: *Ketosis*

Example of application: Treatment for ketosis in dairy cattle –
Testing the effect of two different medications
(new medication S_1 vs. standard treatment S_2)



- **Objective criterion:** Glucose concentration in serum (in mmol/l)
- **Smallest relevant difference:** The new medication is assumed to be more effective, if the level of glucose rises about 2 mmol/l in a predefined time period
- **Information from previous studies:** The standard deviation in glucose concentration in dairy cattle with ketosis is 1.6 mmol/l

→ **Question: How many animals have to be included in the study?**



III. Sample size calculation: *Ketosis*

Example of application: Treatment for ketosis in dairy cattle –
Testing the effect of two different medications
(new medication S_1 vs. standard treatment S_2)



- **Objective criterion:** Glucose concentration in serum (in mmol/l)
- **Smallest relevant difference:** The new medication is assumed to be more effective, if the level of glucose rises about 2 mmol/l in a predefined time period
- **Information from previous studies:** The standard deviation in glucose concentration in dairy cattle with ketosis is 1.6 mmol/l

Test direction: one-sided

$\Delta = 2 \text{ mmol/l}$

$\sigma \approx 1.6 \text{ mmol/l}$

→ **Question: How many animals have to be included in the study?**



III. Sample size calculation: *Ketosis*

Example of application: Treatment for ketosis in dairy cattle –
Testing the effect of two different medications
(new medication S_1 vs. standard treatment S_2)



- **Objective criterion:** Glucose concentration in serum (in mmol/l)
- **Smallest relevant difference:** $\Delta = 2 \text{ mmol/l}$
- **Standard deviation:** $\sigma \approx 1.6 \text{ mmol/l}$
- **Test direction:** one-sided
- **Further assumptions:**

→ **Effect size:**
 $\Delta/\sigma = 2/1.6 = 1.25$

- Is the glucose concentration normally distributed? → Yes
- **Type I error:** $\alpha = 0.05$
- **Type II error:** $\beta = 0.20$

t-Test for independent samples

III. Sample size calculation: *Ketosis*

Example of application: Treatment for ketosis in dairy cattle –
Testing the effect of two different medications
(new medication S_1 vs. standard treatment S_2)



Approximate solution for sample size calculation

→ **One-sided**

1. Determination of the starting point

$$N_0 = \frac{[u_{1-\alpha} + u_{1-\beta}]^2}{[\Delta/(2 \times \sigma)]^2} = \frac{[u_{0.95} + u_{0.80}]^2}{[2/(2 \times 1.6)]^2} = \frac{[1.6449 + 0.8416]^2}{[2/(2 \times 1.6)]^2}$$

q -quantiles u_q of the standard normal distribution

q	0.750	0.800	0.900	0.950	0.975	0.990	0.995
u_q	0.6745	0.8416	1.2812	1.6449	1.9600	2.3264	2.5758

Summary

- **Objective criterion:**
Glucose concentration
- **Statistical Test:**
t-Test for independent samples
- **Relevant difference:**
 $\Delta = 2$ mmol/l
- **Standard deviation:**
 $\sigma = 1.6$ mmol/l
- **Type I error:** $\alpha = 0.05$
- **Type II error:** $\beta = 0.20$
- **Test direction:**
one-sided



III. Sample size calculation: *Ketosis*

Example of application: Treatment for ketosis in dairy cattle –
Testing the effect of two different medications
(new medication S_1 vs. standard treatment S_2)



Approximate solution for sample size calculation

→ **One-sided**

1. Determination of the starting point

$$N_0 = \frac{[u_{1-\alpha} + u_{1-\beta}]^2}{[\Delta/(2 \times \sigma)]^2} = \frac{[u_{0.95} + u_{0.80}]^2}{[2/(2 \times 1.6)]^2} = \frac{[1.6449 + 0.8416]^2}{[2/(2 \times 1.6)]^2}$$
$$= \frac{6.1827}{0.3906} = 15.83 \approx 16 \text{ animals}$$

Summary

- **Objective criterion:**
Glucose concentration
- **Statistical Test:**
t-Test for independent samples
- **Relevant difference:**
 $\Delta = 2 \text{ mmol/l}$
- **Standard deviation:**
 $\sigma = 1.6 \text{ mmol/l}$
- **Type I error:** $\alpha = 0.05$
- **Type II error:** $\beta = 0.20$
- **Test direction:**
one-sided

Iterative procedure → Formula provides only a start value for the total sample size



III. Sample size calculation: *Ketosis*

Example of application: Treatment for ketosis in dairy cattle –
Testing the effect of two different medications
(new medication S_1 vs. standard treatment S_2)



Approximate solution for sample size calculation

→ **One-sided**

2. Iterative procedure continued

– with quantiles of the t distribution

$$16 \geq \frac{[t_{1-\alpha,df} + t_{1-\beta,df}]^2}{[\Delta/(2 \times \sigma)]^2} = \frac{[t_{0.95,14} + t_{0.80,14}]^2}{[2/(2 \times 1.6)]^2}$$
$$= \frac{[1.761 + 0.868]^2}{[2/(2 \times 1.6)]^2} = \frac{6.9116}{0.3906} = 17.69$$

The inequality for $N = 16$ is not fulfilled
→ another iteration step is needed

Summary

- **Objective criterion:**
Glucose concentration
- **Statistical Test:**
t-Test for independent samples
- **Relevant difference:**
 $\Delta = 2$ mmol/l
- **Standard deviation:**
 $\sigma = 1.6$ mmol/l
- **Type I error:** $\alpha = 0.05$
- **Type II error:** $\beta = 0.20$
- **Test direction:**
one-sided



III. Sample size calculation: *Ketosis*

Example of application: Treatment for ketosis in dairy cattle –
Testing the effect of two different medications
(new medication S_1 vs. standard treatment S_2)



Approximate solution for sample size calculation

→ **One-sided**

3. Iterative procedure continued

– with $N = 18$

$$18 \geq \frac{[t_{1-\alpha,df} + t_{1-\beta,df}]^2}{[\Delta/(2 \times \sigma)]^2} = \frac{[t_{0.95,16} + t_{0.80,16}]^2}{[2/(2 \times 1.6)]^2}$$
$$= \frac{[1.746 + 0.865]^2}{[2/(2 \times 1.6)]^2} = \frac{6.8173}{0.3906} = 17.45$$

For a total sample size $N = 18$
the inequality is fulfilled

Summary

- **Objective criterion:**
Glucose concentration
- **Statistical Test:**
t-Test for independent samples
- **Relevant difference:**
 $\Delta = 2$ mmol/l
- **Standard deviation:**
 $\sigma = 1.6$ mmol/l
- **Type I error:** $\alpha = 0.05$
- **Type II error:** $\beta = 0.20$
- **Test direction:**
one-sided



III. Sample size calculation: *Ketosis*

Example of application: Treatment for ketosis in dairy cattle –
Testing the effect of two different medications
(new medication S_1 vs. standard treatment S_2)



Approximate solution for sample size calculation

→ **One-sided**

Two sample *t*-Test

Δ/σ	Power = 0.80		Power = 0.90		Power = 0.95	
	one-sided	two-sided	one-sided	two-sided	one-sided	two-sided
1.15	22	26	28	34	36	42
1.20	20	24	26	32	32	40
1.25	18	24	24	30	30	36
1.30	18	22	22	28	28	34
1.35	16	20	22	26	26	32
1.40	16	20	20	24	24	30
1.45	14	18	18	24	24	28
1.50	14	18	18	22	22	26
1.55	12	16	16	20	20	24

The result fits with the table.

Per treatment group 9
animals are needed with the
predefined settings

Summary

- **Objective criterion:**
Glucose concentration
- **Statistical Test:**
t-Test for independent samples
- **Relevant difference:**
 $\Delta = 2 \text{ mmol/l}$
- **Standard deviation:**
 $\sigma = 1.6 \text{ mmol/l}$
- **Type I error:** $\alpha = 0.05$
- **Type II error:** $\beta = 0.20$
- **Test direction:**
one-sided



***t*-Test for dependent samples**

Example of application:

Ketosis in dairy cattle - β -hydroxybutyric acid



III. Sample size calculation: *Ketosis II*

Example of application: Treatment for ketosis in dairy cattle –
Evaluation of the effectiveness of a new treatment for the
normalisation of the fat and energy metabolism



- **Objective criterion:** Serum level of β -hydroxybutyric acid (in mmol/l), a product of the fatty acid catabolism
 - 2 repeated measurements:
 - 1 measurement prior to the treatment (baseline value)
 - 1 measurement three days after administration of the drug
 - **Dependent sample**
 - **Repeated measurements on the same animals**
 - ***t*-Test for dependent samples**



III. Sample size calculation: *Ketosis II*

Example of application: Treatment for ketosis in dairy cattle – Evaluation of the effectiveness of a new treatment for the normalisation of the fat and energy metabolism



- **Objective criterion:** Serum level of β -hydroxybutyric acid (in mmol/l)
- **Relevant difference:** Decrease of β -hydroxybutyric acid of $\Delta = 1$ mmol/l
- **Test direction:** one-sided
- **Standard deviation:** $\sigma \approx 1.8$ mmol/l
- **Type I error:** $\alpha = 0.05$
- **Type II error:** $\beta = 0.05$



III. Sample size calculation: *Ketosis II*

Example of application: Treatment for ketosis in dairy cattle –
Evaluation of the effectiveness of a new treatment for the
normalisation of the fat and energy metabolism



Approximate solution for sample size calculation

→ **One-sided**

1. Determination of the starting point

$$N_0 = \frac{[u_{1-\alpha} + u_{1-\beta}]^2}{[\Delta/\sigma]^2} = \frac{[u_{0.95} + u_{0.95}]^2}{[1/1.8]^2} = \frac{[1.6449 + 1.6449]^2}{[1/1.8]^2}$$

q -quantiles u_q of the standard normal distribution

q	0.750	0.800	0.900	0.950	0.975	0.990	0.995
u_q	0.6745	0.8416	1.2812	1.6449	1.9600	2.3264	2.5758

Summary

- **Objective criterion:**
Serum level of β -hydroxybutyric acid
- **Statistical test:**
t-Test for dependent samples
- **Relevant difference:**
 $\Delta = 1$ mmol/l
- **Standard deviation:**
 $\sigma = 1.8$ mmol/l
- **Type I error:** $\alpha = 0.05$
- **Type II error:** $\beta = 0.05$
- **Test direction:**
one-sided



III. Sample size calculation: *Ketosis II*

Example of application: Treatment for ketosis in dairy cattle – Evaluation of the effectiveness of a new treatment for the normalisation of the fat and energy metabolism



Approximate solution for sample size calculation

→ **One-sided**

1. Determination of the starting point

$$N_0 = \frac{[u_{1-\alpha} + u_{1-\beta}]^2}{[\Delta/\sigma]^2} = \frac{[u_{0.95} + u_{0.95}]^2}{[1/1.8]^2} = \frac{[1.6449 + 1.6449]^2}{[1/1.8]^2}$$
$$= \frac{10.8228}{0.3086} = 35.07 \approx 36 \text{ animals}$$

Iterative procedure → Formula provides only a start value for the total sample size

Summary

- **Objective criterion:**
Serum level of β -hydroxybutyric acid
- **Statistical test:**
t-Test for dependent samples
- **Relevant difference:**
 $\Delta = 1 \text{ mmol/l}$
- **Standard deviation:**
 $\sigma = 1.8 \text{ mmol/l}$
- **Type I error:** $\alpha = 0.05$
- **Type II error:** $\beta = 0.05$
- **Test direction:**
one-sided



III. Sample size calculation: *Ketosis II*

Example of application: Treatment for ketosis in dairy cattle –
Evaluation of the effectiveness of a new treatment for the
normalisation of the fat and energy metabolism



Approximate solution for sample size calculation

→ **One-sided**

2. Iterative procedure continued

– with quantiles of the t distribution

$$\begin{aligned} 36 &\geq \frac{[t_{1-\alpha,df} + t_{1-\beta,df}]^2}{[\Delta/\sigma]^2} = \frac{[t_{0.95,35} + t_{0.95,35}]^2}{[1/1.8]^2} \\ &= \frac{[1.690 + 1.690]^2}{[1/1.8]^2} = \frac{11.4244}{0.3086} = 37.02 \end{aligned}$$

The inequality for
 $N = 36$ is not fulfilled
→ another iteration
step is needed

Summary

- **Objective criterion:**
Serum level of
 β -hydroxybutyric acid
- **Statistical test:**
t-Test for dependent
samples
- **Relevant difference:**
 $\Delta = 1 \text{ mmol/l}$
- **Standard deviation:**
 $\sigma = 1.8 \text{ mmol/l}$
- **Type I error:** $\alpha = 0.05$
- **Type II error:** $\beta = 0.05$
- **Test direction:**
one-sided



III. Sample size calculation: *Ketosis II*

Example of application: Treatment for ketosis in dairy cattle –
Evaluation of the effectiveness of a new treatment for the
normalisation of the fat and energy metabolism



Approximate solution for sample size calculation

→ **One-sided**

3. Iterative procedure continued

– with $N = 37$

$$\begin{aligned} 37 &\geq \frac{[t_{1-\alpha,df} + t_{1-\beta,df}]^2}{[\Delta/\sigma]^2} = \frac{[t_{0.95,36} + t_{0.95,36}]^2}{[1/1.8]^2} \\ &= \frac{[1.688 + 1.688]^2}{[1/1.8]^2} = \frac{11.3974}{0.3086} = 36.93 \end{aligned}$$

For a total sample size $N = 37$
the inequality is fulfilled

Summary

- **Objective criterion:**
Serum level of β -hydroxybutyric acid
- **Statistical test:**
t-Test for dependent samples
- **Relevant difference:**
 $\Delta = 1$ mmol/l
- **Standard deviation:**
 $\sigma = 1.8$ mmol/l
- **Type I error:** $\alpha = 0.05$
- **Type II error:** $\beta = 0.05$
- **Test direction:**
one-sided



III. Sample size calculation: *Ketosis II*

Example of application: Treatment for ketosis in dairy cattle –
Evaluation of the effectiveness of a new treatment for the
normalisation of the fat and energy metabolism



Approximate solution for sample size calculation

→ **One-sided**

***t*-Test dependent samples**

Δ/σ	Power = 0.80		Power = 0.90		Power = 0.95	
	one-sided	two-sided	one-sided	two-sided	one-sided	two-sided
0.10	620	787	858	1053	1084	1302
0.15	277	351	382	469	483	580
0.20	156	199	216	265	272	327
0.25	101	128	139	171	175	210
0.30	71	90	97	119	122	147
0.35	52	67	72	88	90	109
0.40	41	52	55	68	70	84
0.45	32	41	44	54	55	67
0.50	27	34	36	44	45	54
0.55	22	28	30	37	38	45
0.60	19	24	26	32	32	39

$$\Delta/\sigma = 1/1.8 \approx 0.56$$

Summary

- **Objective criterion:**
Serum level of β -hydroxybutyric acid
- **Statistical test:**
t-Test for dependent samples
- **Relevant difference:**
 $\Delta = 1 \text{ mmol/l}$
- **Standard deviation:**
 $\sigma = 1.8 \text{ mmol/l}$
- **Type I error:** $\alpha = 0.05$
- **Type II error:** $\beta = 0.05$
- **Test direction:**
one-sided



***t*-Test for dependent samples**

Example of application:

Ketosis in dairy cattle - β -hydroxybutyric acid
(changed assumptions!)



III. Sample size calculation: *Ketosis II* (Part 2)

Example of application: Treatment for ketosis in dairy cattle –
Evaluation of the effectiveness of a new treatment for the
normalisation of the fat and energy metabolism



t-Test for dependent samples

Δ/σ	Power = 0.80		Power = 0.90		Power = 0.95	
	one-sided	two-sided	one-sided	two-sided	one-sided	two-sided
0.10	620	787	858	1053	1084	1302
0.15	277	351	382	469	483	580
0.20	156	199	216	265	272	327
0.25	101	128	139	171	175	210
0.30	71	90	97	119	122	147
0.35	52	67	72	88	90	109
0.40	41	52	55	68	70	84
0.45	32	41	44	54	55	67
0.50	27	34	36	44	45	54
0.55	22	28	30	37	38	45
0.60	19	24	26	32	32	39

New assumption for the
standard deviation:
 $\sigma = 2 \text{ mmol/l}$

$$\rightarrow \Delta/\sigma = 1/2 \approx 0.50$$
$$\rightarrow N = 45$$

- Even small changes in the input parameters can have an enormous impact of the needed sample size
- Most exact choice of input parameters!

Summary

- **Objective criterion:**
Serum level of β -hydroxybutyric acid
- **Statistical test:**
t-Test for dependent samples
- **Relevant difference:**
 $\Delta = 1 \text{ mmol/l}$
- **Standard deviation:**
 $\sigma = 1.8 \text{ mmol/l}$
- **Type I error:** $\alpha = 0.05$
- **Type II error:** $\beta = 0.05$
- **Test direction:**
one-sided



χ^2 -Test – Comparison of two prevalences for independent samples



III. Sample size calculation: *Mammary tumours*

Example of application: Examination of the relationship between the exposure to electromagnetic fields and increased risk for cancer (mammary tumours in rats)



- **2 groups:** Both groups are treated with cancer inducing substances, one of the groups are additionally exposed to an electromagnetic field
- **Objective criterion:** Mammary tumours (yes/no)

Dichotomous variable
Follows a binomial distribution

→ **Question:** Are there differences in the frequency of mammary tumours between the two groups?

(Comparison of probabilities or proportions)

→ **Statistical test:** χ^2 -Test or Fisher's exact test in a fourfold table



III. Sample size calculation: *Mammary tumours*

Example of application: Examination of the relationship between the exposure to electromagnetic fields and increased risk for cancer (mammary tumours in rats)



- **2 groups:** Both groups are treated with cancer inducing substances, one of the groups are additionally exposed to an electromagnetic field
- **Objective criterion:** Mammary tumours (yes/no)
- **Relevant difference: Which differences in the tumour frequency between the two groups is seen to be biologically relevant?**
 - Strongly depends on the research question
 - For specific diseases a very small difference can be clinically relevant, whereas for other research questions doubling or halving the proportions can be of interest



III. Sample size calculation: *Mammary tumours*

Example of application: Examination of the relationship between the exposure to electromagnetic fields and increased risk for cancer (mammary tumours in rats)



- **2 groups:** Both groups are treated with cancer inducing substances, one of the groups are additionally exposed to an electromagnetic field
- **Objective criterion:** Mammary tumours (yes/no)
- **Relevant difference:** Difference in the tumour frequency
 - Control group: $P_1 = 0.35$
 - Exposed group: $P_2 = 0.55$
- **Type I error:** $\alpha = 0.05$
- **Type II error:** $\beta = 0.20$

$$\Delta = P_2 - P_1 = 0.55 - 0.35 = 0.2$$



III. Sample size calculation: *Mammary tumours*

Example of application: Examination of the relationship between the exposure to electromagnetic fields and increased risk for cancer (mammary tumours in rats)



$$\bar{P} = (P_1 + P_2)/2$$

$$N = \frac{2 \left[u_{1-\alpha/2} \sqrt{2 \times \bar{P}(1 - \bar{P})} + u_{1-\beta} \sqrt{P_1(1 - P_1) + P_2(1 - P_2)} \right]^2}{\Delta^2}$$

$$\Delta = P_2 - P_1 = 0.55 - 0.35 = 0.2$$

$$= \frac{2 \left[u_{0.975} \sqrt{2 \times 0.45 (1 - 0.45)} + u_{0.80} \sqrt{0.35(1 - 0.35) + 0.55 (1 - 0.55)} \right]^2}{0.2^2}$$

$$= \frac{2 \left[1.9600 \sqrt{2 \times 0.45 (1 - 0.45)} + 0.8416 \sqrt{0.35(1 - 0.35) + 0.55 (1 - 0.55)} \right]^2}{0.2^2}$$

q -quantiles u_q of the standard normal distribution

q	0.750	0.800	0.900	0.950	0.975	0.990	0.995
u_q	0.6745	0.8416	1.2812	1.6449	1.9600	2.3264	2.5758



III. Sample size calculation: *Mammary tumours*

Example of application: Examination of the relationship between the exposure to electromagnetic fields and increased risk for cancer (mammary tumours in rats)



$$\bar{P} = (P_1 + P_2)/2$$

$$N = \frac{2 \left[u_{1-\alpha/2} \sqrt{2 \times \bar{P}(1 - \bar{P})} + u_{1-\beta} \sqrt{P_1(1 - P_1) + P_2(1 - P_2)} \right]^2}{\Delta^2}$$

$$\Delta = P_2 - P_1 = 0.55 - 0.35 = 0.2$$

$$= \frac{2 \left[u_{0.975} \sqrt{2 \times 0.45 (1 - 0.45)} + u_{0.80} \sqrt{0.35(1 - 0.35) + 0.55 (1 - 0.55)} \right]^2}{0.2^2}$$

$$= \frac{2 \left[1.9600 \sqrt{2 \times 0.45 (1 - 0.45)} + 0.8416 \sqrt{0.35(1 - 0.35) + 0.55 (1 - 0.55)} \right]^2}{0.2^2}$$

$$= \frac{7.6755}{0.04} = 191.89 \approx 192 \text{ animals}$$

Preliminary sample size



III. Sample size calculation: *Mammary tumours*

Example of application: Examination of the relationship between the exposure to electromagnetic fields and increased risk for cancer (mammary tumours in rats)



- **Preliminary total sample size:** $N = 192$

Formula according to Casagrande et al. (1978)

$$n_1^* = \frac{n_1}{4} \left[1 + \sqrt{1 + \frac{4}{n_1 \Delta}} \right]^2 \quad \text{mit } n_1 = \frac{N}{2}$$

$$n_1^* = \frac{192/2}{4} \left[1 + \sqrt{1 + \frac{4}{(192/2) 0.2}} \right]^2 = 105.76 \approx 106 \text{ animals per group}$$

212 animals in total



IV. Sample size calculation in SAS



IV. Sample size calculation in SAS

- **Statistical programs which enable a sample size calculation:**
SAS, SPSS, BMDP, StatExact, Stata, nQuery, ...
- **SAS:**
 - Free of charge for scientific purposes:
SAS for Academics on Demand
 - PROC POWER
Procedure for sample size calculation for different statistical tests



IV. Sample size calculation in SAS

ONESAMPLEMEANS	One-sample t-Test Confidence intervals
TWOSAMPLEMEANS	Two-sample t-Test Confidence intervals
PAIRED MEANS	t-Test for dependent samples Confidence intervals
ONEWAYANOVA	One way analysis of variance
TWOSAMPLESURVIVAL	Survival analysis
ONESAMPLEFREQ	One-sample binomial test
PAIREFREQ	McNemar Test
TWOSAMPLEFREQ	Fisher's exact test χ^2 -Test
ONECORR	Pearson's Correlation
...	



IV. Sample size calculation in SAS

t-Test for independent samples with equal variances

- Example of application:** Feeding experiment in pigs – Effect of two different types of feed on the daily weight gain of fattening pigs



```
proc power;  
  twosamplemeans test=diff  
    alpha = 0.05  
    power = 0.8  
    meandiff = 30  
    stddev = 60  
    npergroup = .  
    sides = 2  
;  
run;
```

Two-sided

The POWER Procedure
Two-Sample t Test for Mean Difference

Feste Szenarioelemente	
Distribution	Normal
Method	Exact
Alpha	0.05
Mean Difference	30
Standard Deviation	60
Nominal Power	0.8
Number of Sides	2
Null Difference	0

Berechnet N per Group	
Actual Power	N per Group
0.801	64

Summary

- Objective criterion:**
daily weight gain
- Statistical test:**
t-Test for independent samples
- Relevant difference:**
 $\Delta = 30 \text{ g}$
- Standard deviation:**
 $\sigma = 60 \text{ g}$
- Type I error:** $\alpha = 0.05$
- Type II error:** $\beta = 0.20$
- Test direction:**
two-sided



IV. Sample size calculation in SAS

t-Test for independent samples with equal variances

- Example of application:** Treatment for ketosis in dairy cattle – Testing the effect of two different medications (new medication S_1 vs. standard treatment S_2)



```
proc power;  
  twosamplemeans test=diff  
    alpha = 0.05  
    power = 0.80  
    meandiff = 2  
    stddev = 1.6  
    npergroup = .  
    sides = 1  
;  
run;
```

One-sided

The POWER Procedure
Two-Sample t Test for Mean Difference

Feste Szenarioelemente	
Distribution	Normal
Method	Exact
Number of Sides	1
Alpha	0.05
Mean Difference	2
Standard Deviation	1.6
Nominal Power	0.8
Null Difference	0

Berechnet N per Group	
Actual Power	N per Group
0.814	9

Summary

- Objective criterion:** Glucose concentration
- Statistical Test:** t-Test for independent samples
- Relevant difference:** $\Delta = 2$ mmol/l (**increase**)
- Standard deviation:** $\sigma = 1.6$ mmol/l
- Type I error:** $\alpha = 0.05$
- Type II error:** $\beta = 0.20$
- Test direction:** one-sided



IV. Sample size calculation in SAS

t-Test for dependent samples

- Example of application:** Treatment for ketosis in dairy cattle – Evaluation of the effectiveness of a new treatment for the normalisation of the fat and energy metabolism



```
proc power;  
  pairedmeans test=diff  
    alpha = 0.05  
    power = 0.95  
    corr = 0.5  
    meandiff = 1  
    stddev = 1.8  
    npairs = .  
    sides = 1  
;  
run;
```

One-sided

The POWER Procedure
Paired t Test for Mean Difference

Feste Szenarioelemente	
Distribution	Normal
Method	Exact
Number of Sides	1
Alpha	0.05
Mean Difference	1
Standard Deviation	1.8
Correlation	0.5
Nominal Power	0.95
Null Difference	0

Berechnet N Pairs	
Actual Power	N Pairs
0.953	37

Summary

- Objective criterion:**
Serum level of β -hydroxybutyric acid
- Statistical test:**
t-Test for dependent samples
- Relevant difference:**
 $\Delta = 1$ mmol/l (**decrease**)
- Standard deviation:**
 $\sigma = 1.8$ mmol/l
- Type I error:** $\alpha = 0.05$
- Type II error:** $\beta = 0.05$
- Test direction:**
one-sided



IV. Sample size calculation in SAS

χ^2 -Test – Comparison of two prevalences for independent samples

- Example of application:** Examination of the relationship between the exposure to electromagnetic fields and increased risk for cancer (mammary tumours in rats)



```
proc power;  
  twosamplefreq test = pchi  
    groupproportions  
      = (0.35 0.55)  
  alpha = 0.05  
  power = 0.80  
  ntotal = .  
  sides = 2  
;  
run;
```

The POWER Procedure
Pearson Chi-square Test for Proportion Difference

Feste Szenarioelemente	
Distribution	Asymptotic normal
Method	Normal approximation
Number of Sides	2
Alpha	0.05
Group 1 Proportion	0.35
Group 2 Proportion	0.55
Nominal Power	0.8
Null Proportion Difference	0
Group 1 Weight	1
Group 2 Weight	1

Berechnet N Total	
Actual Power	N Total
0.800	192

After applying the formula according to Casagrande et al. (1978):
 $N = 212$ animals

Summary

- Objective criterion:** Tumour frequency
- Statistical test:** χ^2 -Test for independent samples
- Control group:** $P_1 = 0.35$
- Exposed group:** $P_2 = 0.55$
- Relevant difference:** $\Delta = 0.20$
- Type I error:** $\alpha = 0.05$
- Type II error:** $\beta = 0.20$
- Test direction:** two-sided



V. Conclusion



V. Conclusion

Requirements for sample size calculation

- **Main objective criterion** and **research question** (one-sided or two-sided)
- Choice of the **appropriate statistical test** according to the research question and the data basis
- **Smallest detectable difference Δ** : Determination of the smallest clinically or biologically relevant difference which should be detected
- **Standard deviation σ** of the main objective criterion
- **Significance level α** : Probability to achieve a significant test result, although no difference exists (usually $\alpha = 0.05$)
- **Power of the study ($1 - \beta$)**: Probability to detect a real difference with the statistical test



V. Conclusion

Connection between the influencing factors and the sample size

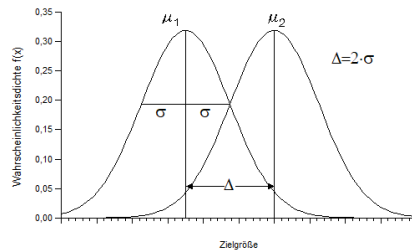
- **Number of needed samples increases with ...**
 - ... decreasing α error
 - ... decreasing detectable difference
 - ... increasing power $1 - \beta$
 - ... increasing standard deviation σ



V. Conclusion

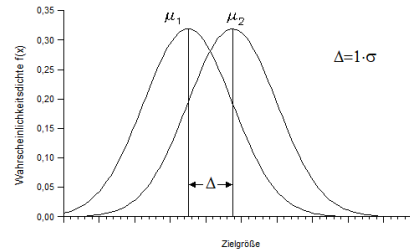
Two-sample *t*-Test (two-sided), $\alpha = 0.05$, $\beta = 0.10$

Effect size = 2



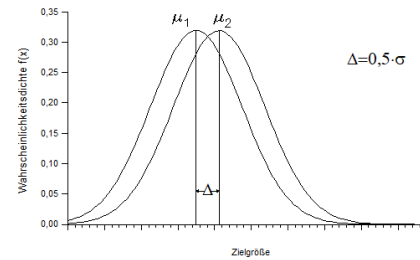
→ n = 7 animals per group

Effect size = 1



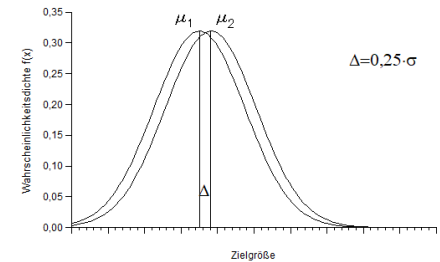
→ n = 23 animals per group

Effect size = 0.5



→ n = 86 animals per group

Effect size = 0.25



→ n = 338 animals per group

Effect size Δ/σ

Sample size



V. Conclusion

- Sample size calculation in animal experiments:
Not more animals than necessary, but also not less animals than needed!
- An appropriate sample size ensures that ...
 - ... an existing effect can be detected
 - ... with a specific certainty one can assume that there is no relevant effect, if the statistical analysis does not reveal significant differences
- Every sample size calculation is customized for every singly study!!!
 - Precise knowledge about the data is needed in order to achieve a meaningful sample size calculation
 - Assumptions about the expected effects, the distribution of the data can be found in literature review, pilot studies or expert opinions



V. Conclusion

- Sample size calculation depends on the statistical test, the type I error α , the power $(1 - \beta)$ and the specified clinical or biological relevant minimal difference
 - Complex study designs need generally higher sample sizes
- Sample size calculation has to be carried out prior to the study
 - Mandatory for animal experiments
 - Research question, study design, planned statistical analysis and resulting from these facts the sample size calculation has to be stated in the request for approval



Thank you for you attention!

