

Research Methods D
Statistics
PSM2RD

Dietmar Heinke

Contact

Room: 2.09 (Hills building)

Email: d.g.heinke@bham.ac.uk

Phone: 44920

Appointments (via email): any time

This term's contents

- **Basics for the first half of the term (today)**
- **One-way ANOVA (next week)**
- **Two-way ANOVA within-participants (3. week)**
- **Two-way factorial ANOVA (4. week)**
- **Split-plot ANOVA (5. week)**
- **Multiple Comparisons (6. + 7. week)**
- **Non-parametric test (8.+9. week)**
- **Preparation for class test (10. week)**
- **11.week: class test in lecture slot**

This term's labs

- Starting **next** week (20. January)

- **9 labs, matching contents of lecture**

Assessment of course work

- **Every second** lab is marked (5 labs)
- Answers of 6 questions in each lab are marked (bold font)
- Each correct answer is a point
- The sum of all points is halved and the result is directly translated into grades (15 = A+, 14 = A,... 1 = F-, 0 = X).

Assessment of course work

- The answer sheets are collected by the TA and return in the next lab (only for the marked lab)
- **But remember: all questions are relevant for the class test**

Attendance of labs

- **If you have missed a marked lab, you can ask the TA to mark an originally unmarked lab.**
- **You must ask at the beginning of your own lab slot and at the earliest possible time.**
- **The next, normally unmarked, lab will be marked (possibly on the same day).**
- **You have to make sure yourself that you hand in your answers.**
- **Normal mitigation procedures apply as well**

Books

- **Christine P. Dancey & John Reidy**
Statistics Without Maths for Psychology – Using SPSS for Windows

- **David C. Howell**
Statistical Methods for Psychology

Handouts

- For today's and next week's lecture on WebCT and paper
- After that only on WebCT

Basics for Analysis of Variance (ANOVA)

Learning objectives

- **Short revision of concepts which are necessary to understand the first part of this course**

- **Independent variable ?**
- **Dependent variable ?**
- **Mean value ?**
- **Variation ?**
- **Degrees of Freedom ?**
- **Distribution ?**
- **Probability ?**
- **Normal Distribution ?**
- **Hypothesis testing ?**

Variable

- **Variable:**
a property of an object or event that can take on different values
- **Independent Variable:**
Variable under the control of the experimenter
- **Dependent variable:**
Variable **not** under the control of the experimenter
(data)

Mean

- **= average**
- **Indication of typical score in data set**
- **Calculation:**
 - **Sum all scores in a sample**
 - **Divide by the number of scores in the sample**

Variation

- **Information of how a sample is spread out**
- **One way of measuring spread is Variance:**
 - **Compute squared differences between mean and score**
 - **Compute of sum of squared differences**
 - **Divide by number of scores minus 1**
- **Alternatively, Standard deviation:**
 - **Square root of variance**

Degrees of Freedom (df)

- **Number of values that can vary for a given measure**
- **Examples:**
 - **Mean: $df = \text{number of scores}$**
 - **Variance: $df = \text{number of scores} - 1$**

Distribution

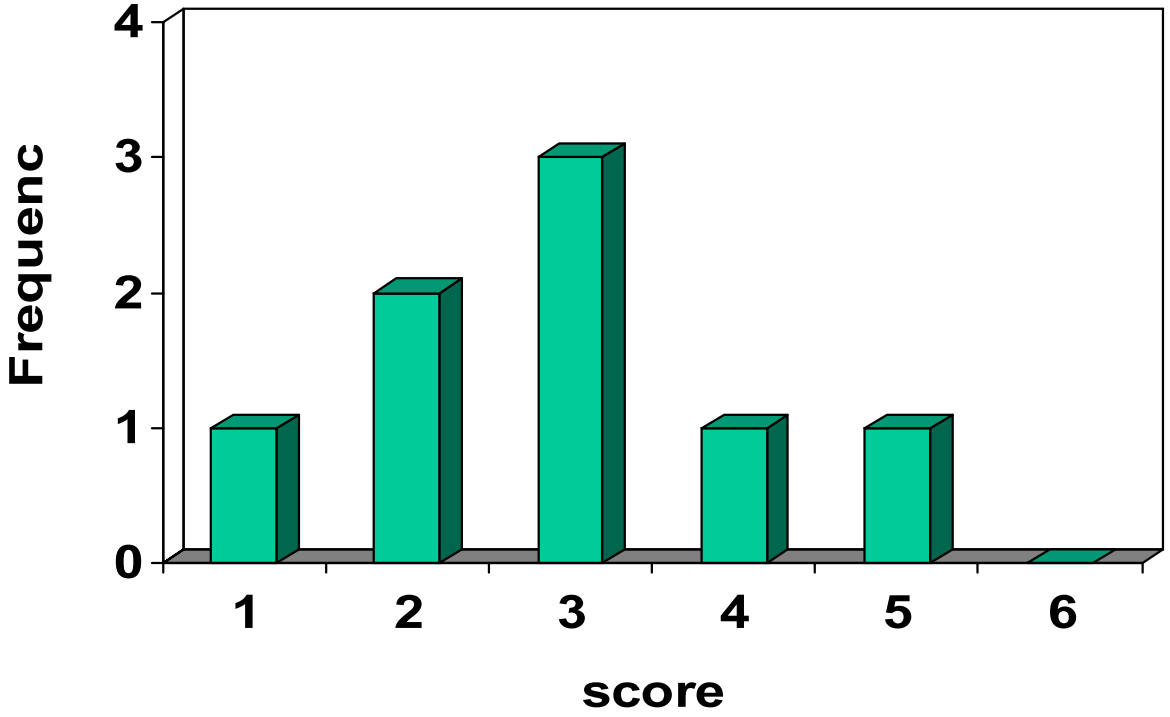
- = “Histogram”
- Frequency of scores
- Important distribution:
normal distribution

Example

- **Sample:** 4, 5, 2, 1, 3, 3, 2, 3
- **Mean:** 2.86
- **Variance:** 1.55
- **Distribution:**

Scores	1	2	3	4	5	6
Frequency of scores	1	2	3	1	1	0

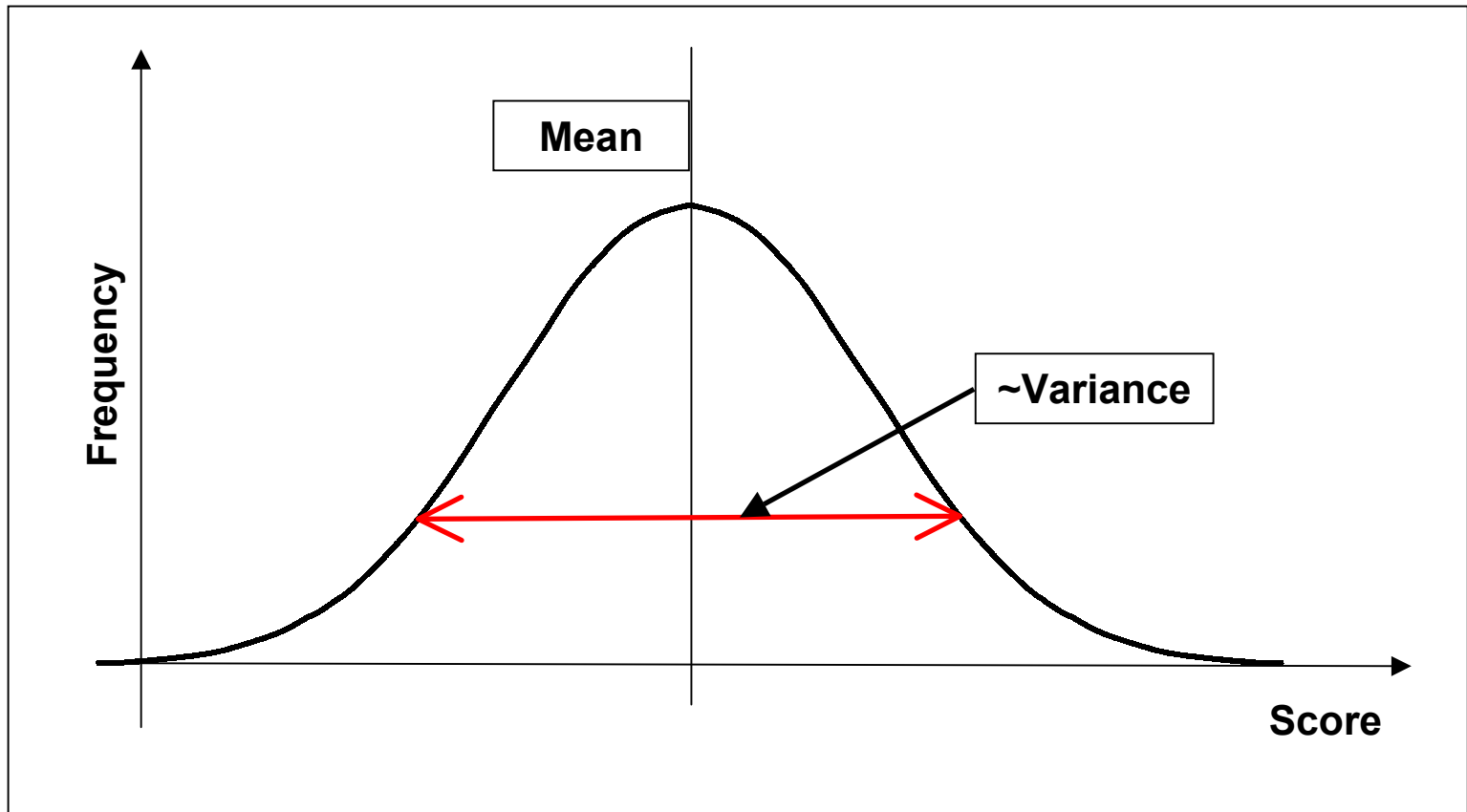
Distribution



Normal distribution

- **Many “things” are normally distributed**
- **Bell-shaped distribution**
- **Can be described by mean and variance**

Normal distribution



Probability

- **Likelihood of an event expressed by a decimal number between 0 and 1**
- **0 = event definitely will not happen**
- **1 = event definitely will happen**
- **Also expressed in percentage**

Conditional probability

- **Probability of one event given the occurrence of some other event**
- **Examples:**
 - **The probability of Arsenal winning might be 80%, if they were playing Birmingham City, but might be only 50%, if they were playing Manchester United**
 - **The probability of getting lung cancer, if you smoke**

Hypothesis-Testing

Hypothesis testing

- **Inferential statistic**
- **Draw conclusions from experiments**
- **Competition between two hypothesis**

Hypothesis

- **Null Hypothesis (H_0):**
 - Assumption of no experimental effect
 - “All being normal”
- **Alternative Hypothesis (H_1):**
 - Being adopted when H_0 is rejected
 - Also: research or experimental hypothesis

Significance level (p-value)

- **p-value is a conditional probability:**
The probability that an event (experimental result) occurs given that H_0 is true.
- **Significant vs. non-significant decided by α -level:**
 - If p is smaller than α the result is “significant”
 - If p is larger than α the result is “non-significant”

Errors

- **Type I error: The error of rejecting H_0 when it is true (p-value)**
- **Type II error: The error of not rejecting H_0 when it is false**

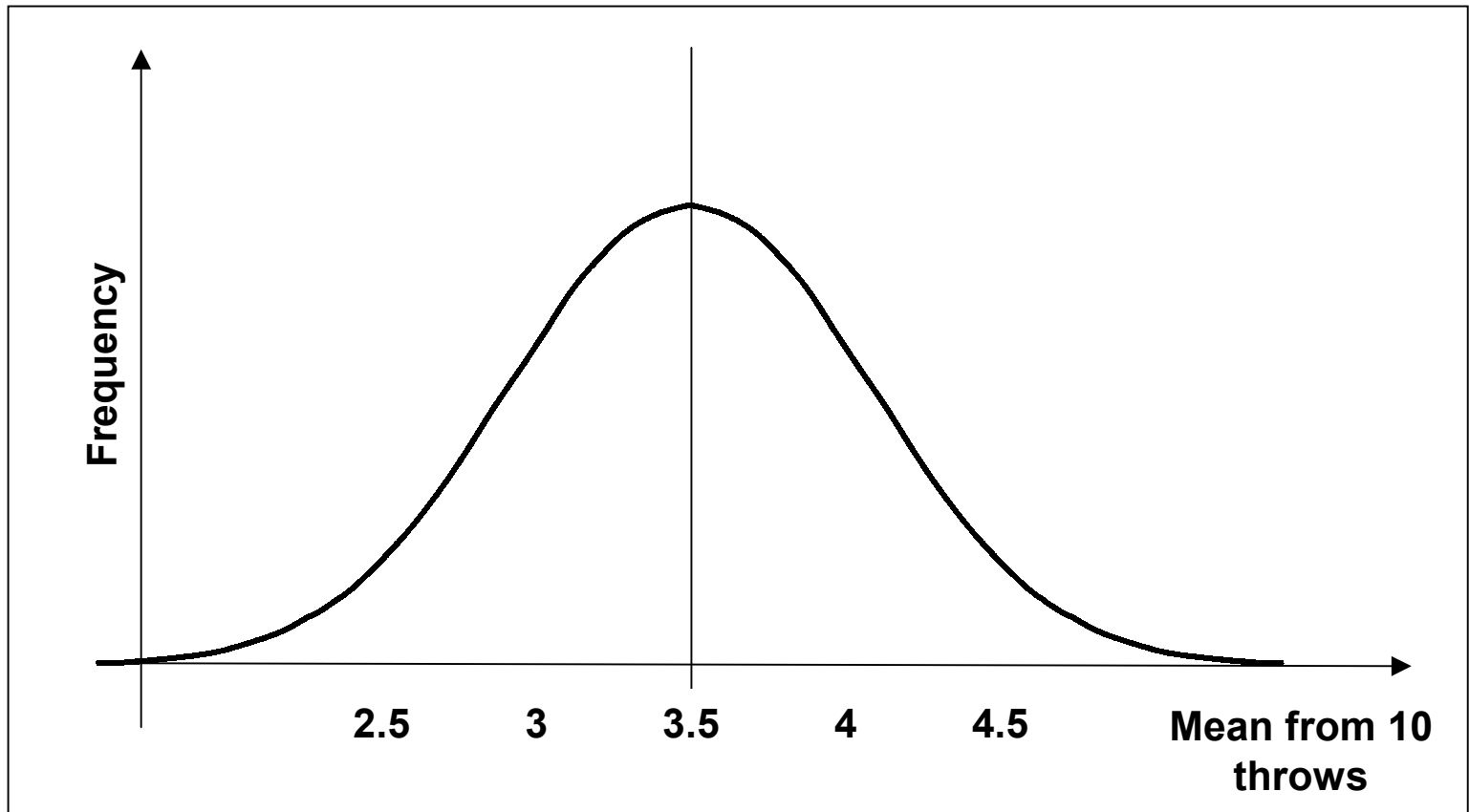
Errors

Decision	State of the world: H_0 true	State of the world: H_0 false
Reject H_0	Type I error	Correct decision
Fail to reject H_0	Correct decision	Type II error

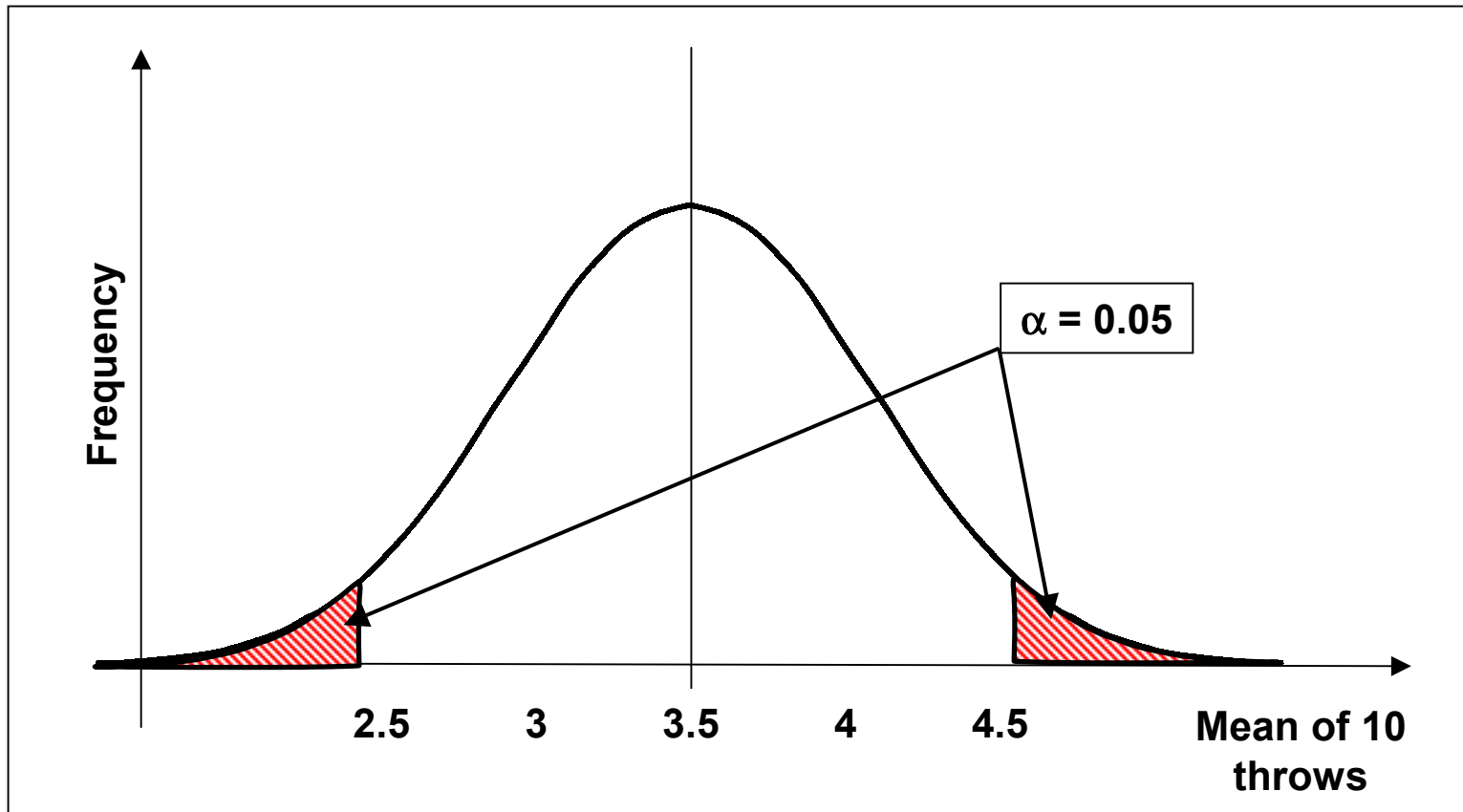
Experiment

- **Is the die loaded?**
- **Approach: Calculate the mean from throws**
- **Null Hypothesis (H_0):**
The die is completely normal
- **Alternative Hypothesis (H_1):**
The die produces a smaller or larger mean than a normal die. Hence, die is loaded

Distribution of a normal die



Distribution of a normal die



Hypothesis-testing

Example	General concept
Assumption: Small or large mean indicates a loaded die	Research hypothesis; experimental or alternative hypothesis
Assumption of the die being normal; Usage of mean and normal distribution	Null hypothesis
p-value is the probability the mean from 10 throws is generated from a normal die.	p-value is the probability that a experimental result is produced when the null hypothesis is

true

Next time

- **One-way ANOVA**