Methods 3: Multilevel Statistical Modeling and Machine Learning

Week 1: *Introduction and Why are we here?*

September 14, 2021

by: Lau Møller Andersen

Outline

1) The overall idea of the course
2) Presenting the instructor and myself
3) Getting to know your names
4) Academic regulations (*studieordning*) and exam
5) How much do you remember?

The foundations we build on

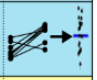
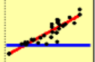
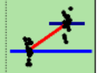
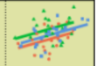
- You have now learnt to code (*Methods 1*), and you have learnt about the General Linear Model (*Methods 2*)
- You can thus evaluate any given model

From Jonas Lindeløv

Common statistical tests are linear models

Last updated: 28 June, 2019.Also check out the Python version!

See worked examples and more details at the accompanying notebook: https://lindeloev.github.io/tests-as-linear

	Common name	Built-in function in R	Equivalent linear model in R	Exact?	The linear model in words	Icon
Simple regression: lm(y ~ 1 + x)	**y is independent of x** P: One-sample t-test N: Wilcoxon signed-rank	t.test(y) wilcox.test(y)	lm(y ~ 1) lm(signed_rank(y) ~ 1)	✓ for N >14	One number (intercept, i.e., the mean) predicts **y**. - (Same, but it predicts the *signed rank* of **y**.)	
	P: Paired-sample t-test N: Wilcoxon matched pairs	t.test(y_1, y_2, paired=TRUE) wilcox.test(y_1, y_2, paired=TRUE)	lm(y_2 - y_1 ~ 1) lm(signed_rank(y_2 - y_1) ~ 1)	✓ for N >14	One intercept predicts the pairwise **y_2-y_1** differences. - (Same, but it predicts the *signed rank* of **y_2-y_1**.)	
	y ~ continuous x P: Pearson correlation N: Spearman correlation	cor.test(x, y, method='Pearson') cor.test(x, y, method='Spearman')	lm(y ~ 1 + x) lm(rank(y) ~ 1 + rank(x))	✓ for N >10	One intercept plus **x** multiplied by a number (slope) predicts **y**. - (Same, but with *ranked* **x** and **y**)	
	y ~ discrete x P: Two-sample t-test P: Welch's t-test N: Mann-Whitney U	t.test(y_1, y_2, var.equal=TRUE) t.test(y_1, y_2, var.equal=FALSE) wilcox.test(y_1, y_2)	lm(y ~ 1 + G_2)[A] gls(y ~ 1 + G_2, weights=…[B])[A] lm(signed_rank(y) ~ 1 + G_2)[A]	✓ ✓ for N >11	An intercept for **group 1** (plus a difference if **group 2**) predicts **y**. - (Same, but with one variance *per group* instead of one common.) - (Same, but it predicts the *signed rank* of **y**.)	
Multiple regression: lm(y ~ 1 + x_1 + x_2 +…)	P: One-way ANOVA N: Kruskal-Wallis	aov(y ~ group) kruskal.test(y ~ group)	lm(y ~ 1 + G_2 + G_3 +…+ G_N)[A] lm(rank(y) ~ 1 + G_2 + G_3 +…+ G_N)[A]	✓ for N >11	An intercept for **group 1** (plus a difference if group ≠ 1) predicts **y**. - (Same, but it predicts the *rank* of **y**.)	
	P: One-way ANCOVA	aov(y ~ group + x)	lm(y ~ 1 + G_2 + G_3 +…+ G_N + x)[A]	✓	- (Same, but plus a slope on **x**.) *Note: this is discrete AND continuous. ANCOVAs are ANOVAs with a continuous x.*	
	P: Two-way ANOVA	aov(y ~ group * sex)	lm(y ~ 1 + G_2 + G_3 +…+ G_N + S_2 + S_3 +…+ S_K + G_2*S_2 + G_3*S_3 + … + G_N*S_K)	✓	Interaction term: changing **sex** changes the **y ~ group** parameters. *Note: $G_{2 \text{ to } N}$ is an indicator (0 or 1) for each non-intercept levels of the **group** variable. Similarly for $S_{2 \text{ to } K}$ for sex. The first line (with G_i) is main effect of group, the second (with S_i) for sex and the third is the **group × sex** interaction. For two levels (e.g. male/female), line 2 would just be "S_2" and line 3 would be S_2 multiplied with each G_i.*	*[Coming]*
	Counts ~ discrete x N: Chi-square test	chisq.test(groupXsex_table)	**Equivalent log-linear model** glm(y ~ 1 + G_2 + G_3 + … + G_N + S_2 + S_3+ … + S_K + G_2*S_2 + G_3*S_3 +…+ G_N*S_K, family=…)[A]	✓	Interaction term: (Same as Two-way ANOVA.) *Note: Run glm using the following arguments: `glm(model, family=poisson())` As linear-model, the Chi-square test is $log(y_i) = log(N) + log(\alpha_i) + log(\beta_j) + log(\alpha_i\beta_j)$ where α_i and β_j are proportions. See more info in the accompanying notebook.*	*Same as Two-way ANOVA*
	N: Goodness of fit	chisq.test(y)	glm(y ~ 1 + G_2 + G_3 +…+ G_N, family=…)[A]	✓	(Same as One-way ANOVA and see Chi-Square note.)	*1W-ANOVA*

https://lindeloev.github.io/tests-as-linear/linear_tests_cheat_sheet.pdf

... so what is left to learn?

Overall idea of the course

- **Master complex data:** build models that work despite missing data entries, and which model the nested and hierarchical structures of the data
- **Create generalisable models:** do model comparisons and learn the difference between *explanation* and *prediction*
- **Reduce the dimensionality of data:** extract the meaningful dimensions of data and separate it from the noise

Overall idea of the course

- The aim is that you learn the practical skills (not just the theory), such that you yourself can build and evaluate models in *R* and *Python.*
- In a wider sense, the idea is that you should be able to critically assess models and understand their advantages and limitations.

My expectations towards you – the do’s!

- Do show up for the lectures and classes!
- Do the exercises and the assignments and hand them in on time!
- Do take responsibility for your learning!
- Do ask questions if things are unclear!
- Do challenge me!
- Do make sure that I provide a foundation of learning for you all!

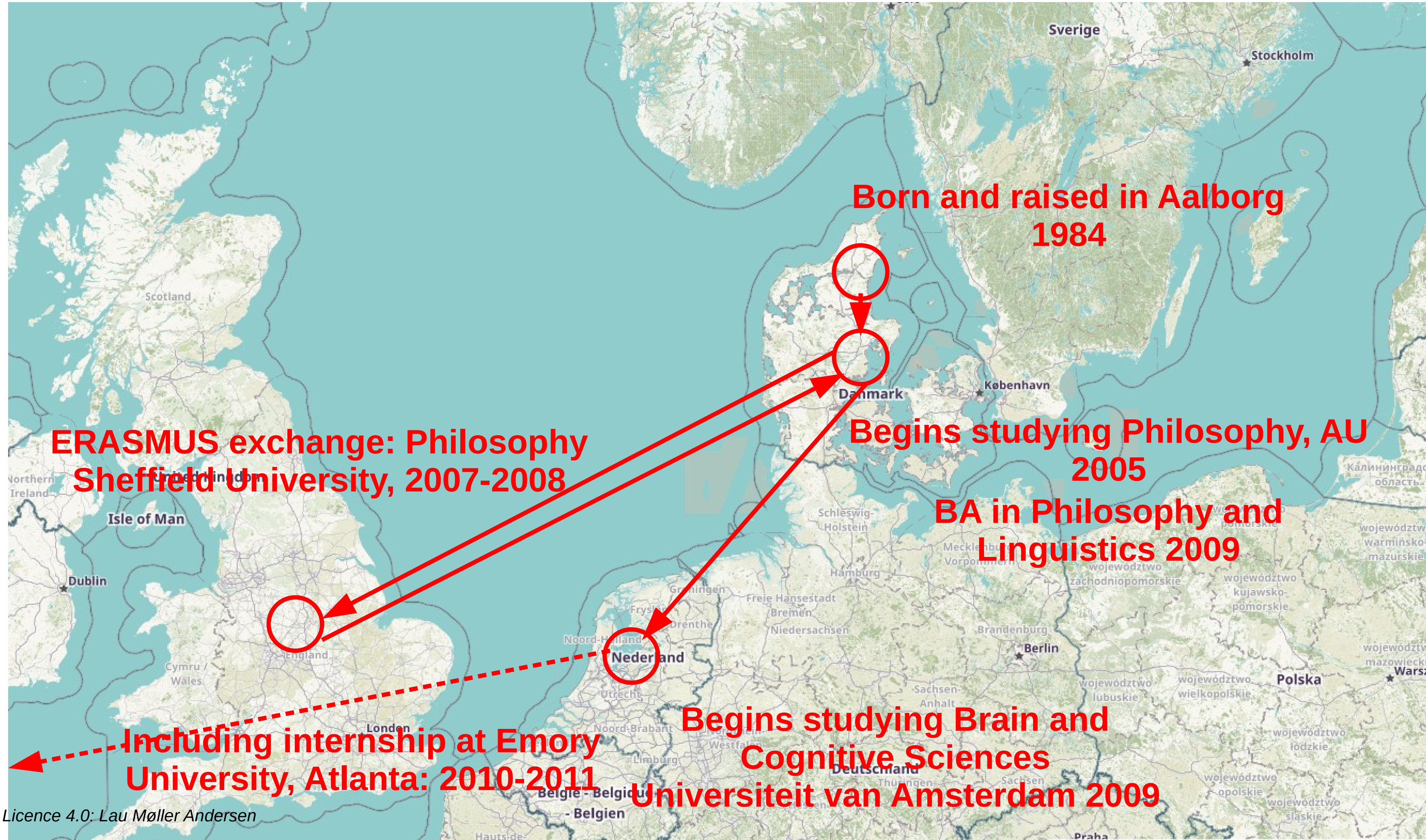
Who are we?

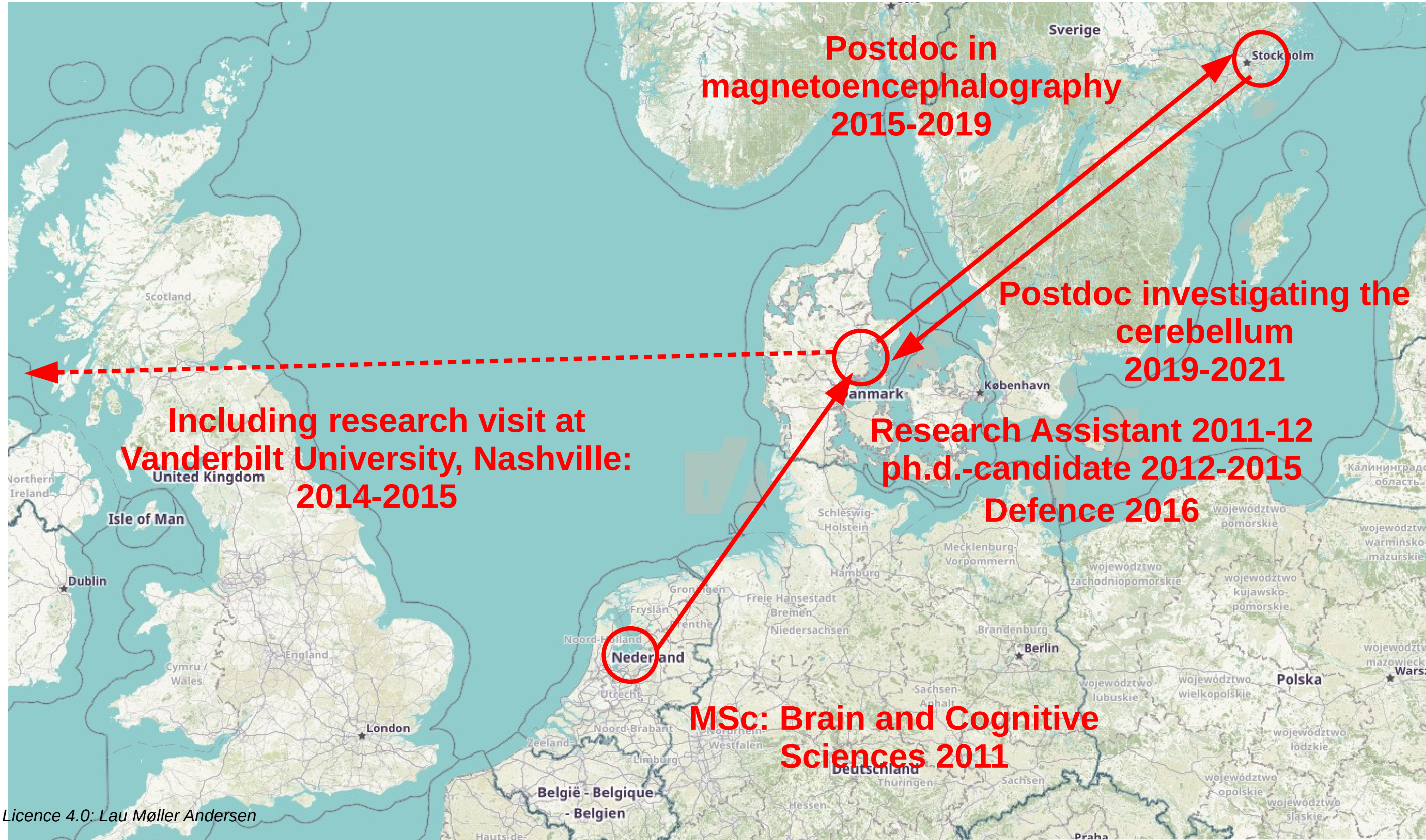
Instructor

- Emil Trenckner Jessen
  - Master's degree student (1st semester)

Lecturer


Lau Møller Andersen - PhD
Front Page
Publications & CV
Workshops
Open Science
Teaching Material
Contact details
Email: lmandersen[at]cfin.au.dk
Phone: +45 4042 7564
Visiting address:
CFIN cfin.au.dk
Nørrebrogade 44, Building 1A
8000 Aarhus C
Denmark
or
AIAS aias.au.dk
Høegh-Guldbergsgade 6B
8000 Aarhus C
Denmark
ABOUT ME
I am a post-doc working at CFIN (Center of Functionally Integrative Neuroscience) at Aarhus University, Denmark and I am a fellow at the Aarhus Institute of Advanced Studies (AIAS). I'm interested in anything related to magnetoencephalography (MEG), but sensory expectations, consciousness, cerebellar MEG and on-scalp MEG must be highlighted among my interests. I have also worked and published on on-scalp MEG and machine learning. Finally, I take a keen interest in coding using language such as MATLAB, Python and R and have published tutorial papers on how to do MEG analysis in FieldTrip and MNE-Python.
CFIN
www.cfin.au.dk
LAU MØLLER ANDERSEN
Postdoc: Karolinska Institutet 2015-2019: working on on-scalp MEG and somatosensation
PhD: Aarhus University 2016
Thesis: Spatio-temporal localization and task specificity in the search for neural correlates of perceptual consciousness
MSc (Cum Laude): University of Amsterdam 2011
Programme: Brain and Cognitive Sciences
BA: Aarhus University 2009
Programmes: Philosophy and Lingustics
AARHUS UNIVERSITET
©Villy Fink Isaksen CC-BY-SA 4.0, file link


Born and raised in Aalborg
1984
Begins studying Philosophy, AU
2005
BA in Philosophy and
Linguistics 2009
ERASMUS exchange: Philosophy
Sheffield University, 2007-2008
Including internship at Emory
University, Atlanta: 2010-2011
Begins studying Brain and
Cognitive Sciences
Universiteit van Amsterdam 2009
Sverige
Stockholm
Danmark
København
Scotland
Isle of Man
Dublin
England
London
Nederland
Berlin
Polska
Hamburg
Deutschland


Postdoc in
magnetoencephalography
2015-2019
Postdoc investigating the
cerebellum
2019-2021
Including research visit at
Vanderbilt University, Nashville:
2014-2015
Research Assistant 2011-12
ph.d.-candidate 2012-2015
Defence 2016
MSc: Brain and Cognitive
Sciences 2011
Sverige
Stockholm
Scotland
Danmark
København
United Kingdom
Northern
Ireland
Isle of Man
Dublin
Cymru /
Wales
England
London
Nederland
Noord-Holland
Fryslân
Groningen
Drenthe
Utrecht
Noord-Brabant
Zeeland
Limburg
België - Belgique
- Belgien
Nordrhein-
Westfalen
Deutschland
Thüringen
Hessen
Sachsen
Sachsen-
Anhalt
Niedersachsen
Freie Hansestadt
Bremen
Hamburg
Schleswig-
Holstein
Mecklenburg-
Vorpommern
Brandenburg
Berlin
Polska
województwo
pomorskie
województwo
zachodniopomorskie
województwo
kujawsko-
pomorskie
województwo
lubuskie
województwo
wielkopolskie
województwo
łódzkie
województwo
opolskie
województwo
śląskie
województwo
warmińsko-
mazurskie
Калининградская
область
Praha

2021 – onwards
Department of Cognitive Science

Examples of my interests

Classifying when visual stimuli become consciously perceptible using *machine learning*

Occipital MEG Activity in the Early Time Range (<300 ms) Predicts Graded Changes in Perceptual Consciousness FREE

Lau M. Andersen ✉, Michael N. Pedersen, Kristian Sandberg, Morten Overgaard

Cerebral Cortex, Volume 26, Issue 6, June 2016, Pages 2677–2688, https://doi.org/10.1093/cercor/bhv108

Published: 24 May 2015

Examples of my interests

Finding the somatosensory activation to absent stimulation using *magnetoencephalography*

ELSEVIER

NeuroImage
Volume 184, 1 January 2019, Pages 78-89

Somatosensory responses to nothing: An MEG study of expectations during omission of tactile stimulations

Lau M. Andersen, Daniel Lundqvist

Show more

https://doi.org/10.1016/j.neuroimage.2018.09.014

Get rights and content

Examples of my interests

Providing pipelines for group analysis in *Python* and *MATLAB*

PROTOCOLS ARTICLE
Front. Neurosci., 22 January 2018 | https://doi.org/10.3389/fnins.2018.00006

Group Analysis in MNE-Python of Evoked Responses from a Tactile Stimulation Paradigm: A Pipeline for Reproducibility at Every Step of Processing, Going from Individual Sensor Space Representations to an across-Group Source Space Representation

Lau M. Andersen*
NatMEG, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

PROTOCOLS ARTICLE
Front. Neurosci., 01 May 2018 | https://doi.org/10.3389/fnins.2018.00261

Group Analysis in FieldTrip of Time-Frequency Responses: A Pipeline for Reproducibility at Every Step of Processing, Going From Individual Sensor Space Representations to an Across-Group Source Space Representation

Lau M. Andersen*
NatMEG, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

Examples of my interests

Benchmarking new technologies in *magnetoencephalography*

ELSEVIER
NeuroImage
Volume 221, 1 November 2020, 117157

On-scalp MEG SQUIDs are sensitive to early somatosensory activity unseen by conventional MEG

Lau M. Andersen [a, b], Christoph Pfeiffer [c, d], Silvia Ruffieux [c], Bushra Riaz [e], Dag Winkler [c], Justin F. Schneiderman [e], Daniel Lundqvist [a]

Show more

+ Add to Mendeley Share Cite

https://doi.org/10.1016/j.neuroimage.2020.117157 Get rights and content

PLOS | ONE PUBLISH ABOUT BROWSE

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

Similarities and differences between on-scalp and conventional in-helmet magnetoencephalography recordings

Lau M. Andersen, Robert Oostenveld, Christoph Pfeiffer, Silvia Ruffieux, Veikko Jousmäki, Matti Hämäläinen, Justin F. Schneiderman, Daniel Lundqvist

Published: July 24, 2017 • https://doi.org/10.1371/journal.pone.0178602

Examples of my interests

Using *mixed effects modelling*

ELSEVIER

Consciousness and Cognition

Volume 71, May 2019, Pages 59-69

Visual expectations change subjective experience without changing performance

Lau Møller Andersen [a, b], Morten Overgaard [b], Frank Tong [c]

Show more

\+ Add to Mendeley Share Cite

https://doi.org/10.1016/j.concog.2019.03.007 Get rights and content

LibreOffice – (I work in Linux)

- Slides will be in odp- or pdf-format.

https://www.libreoffice.org/download/download/

... and I avoid Google as much as I can

Search engine

CryptPad
Online document collaboration

File management and version control

Getting to know your names

I am Lau and I brought a Lama with me

Languages

https://cran.r-project.org/mirrors.html

My recommendation:
https://docs.conda.io/en/latest/miniconda.html

Python environment

Create environment (yml-file can be downloaded from GitHub)

```
(base) lau@lau:~$ conda env create -f methods3_environment.yml
```

Activate environment

```
(base) lau@lau:~$ conda activate methods3
```

This course

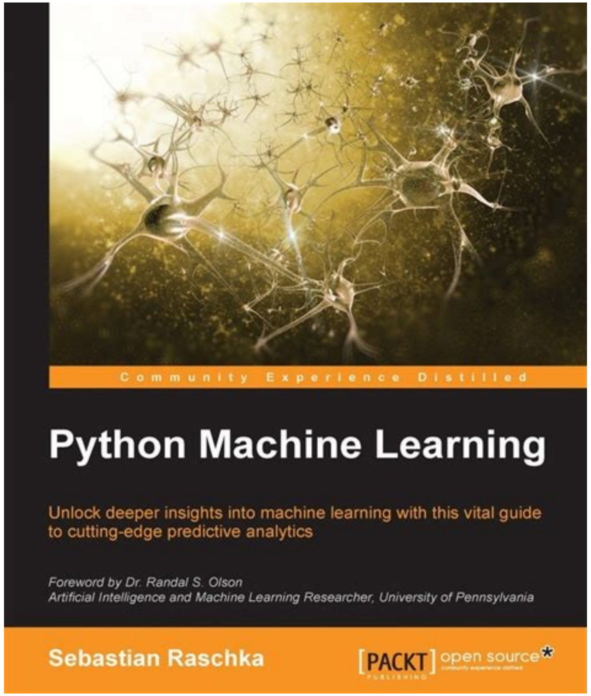
R

Linear mixed models

Python

Machine learning

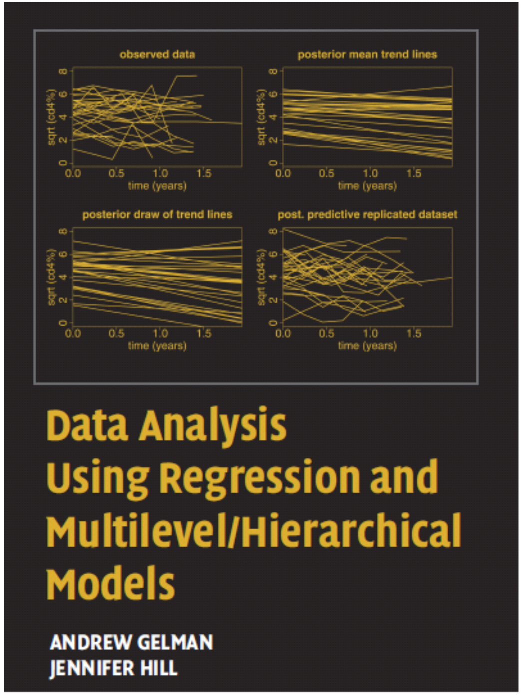
Book for latter half



- Has been ordered from Stakbogladen
  - Support the local book stores and not space cowboys

Not necessary – but really good if you are interested

- Can be ordered from Stakbogladen
  - Let me know if you are interested!
    - Support the local book stores and not space cowboys



Programming languages interpret *human readable code*, such that complex operations can be performed in a systematic way

Example – estimating π

$$\pi = 4 \cdot \sum_{n=1}^{\infty} (-1)^{n+1} \cdot \frac{1}{2n-1} = 4 \cdot \left(\frac{1}{1} - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{11} \cdots \right)$$

https://da.wikipedia.org/wiki/Pi_(tal)

Scripts and output

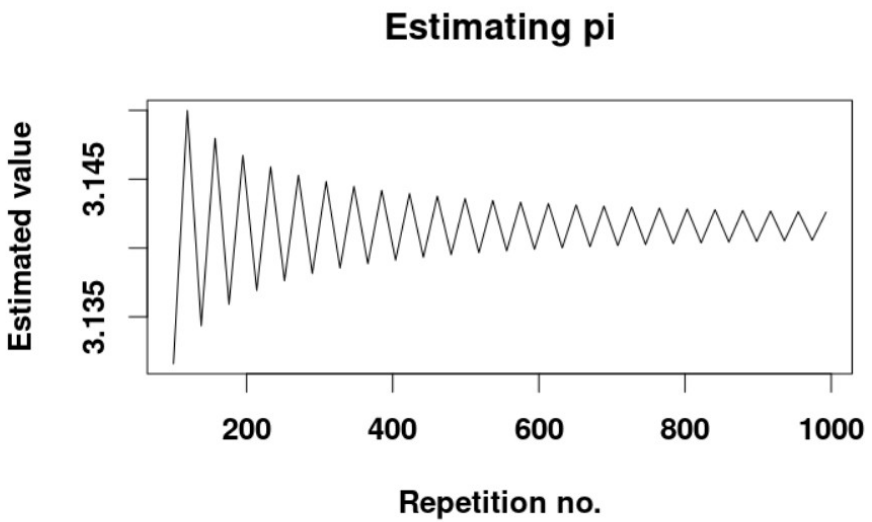
estimate_pi.R

Source on Save

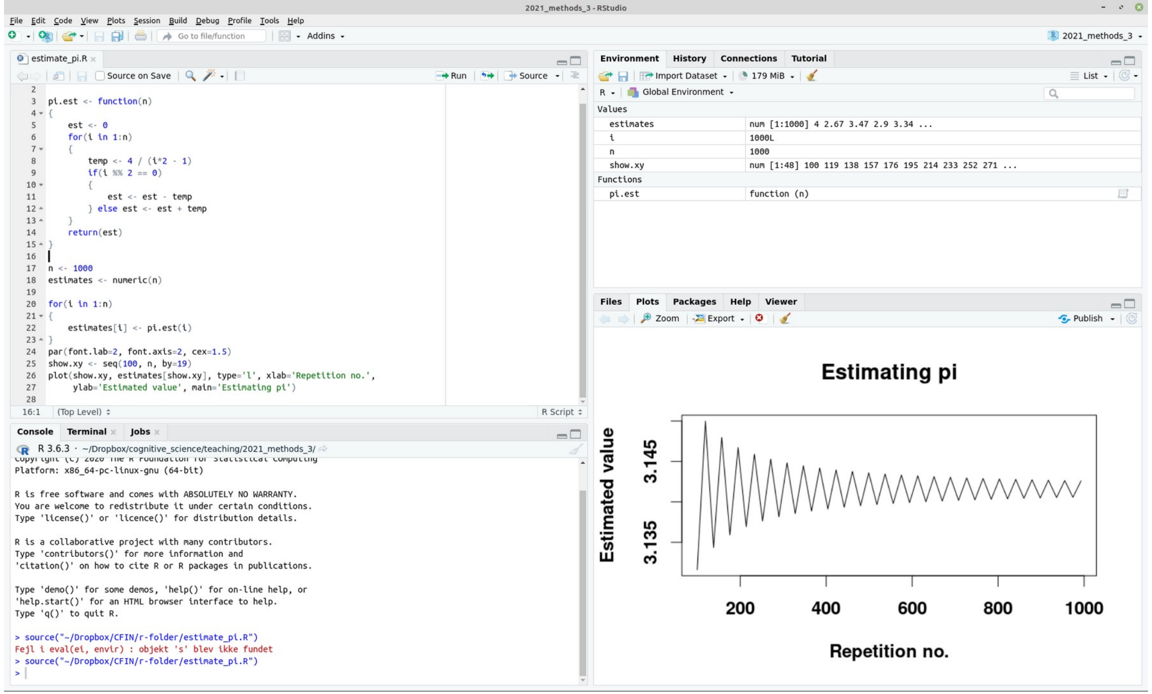
```r

pi.est <- function(n)
{
    est <- 0
    for(i in 1:n)
    {
        temp <- 4 / (i*2 - 1)
        if(i %% 2 == 0)
        {
            est <- est - temp
        } else est <- est + temp
    }
    return(est)
}
s
n <- 1000
estimates <- numeric(n)

for(i in 1:n)
{
    estimates[i] <- pi.est(i)
}
par(font.lab=2, font.axis=2, cex=1.5)
show.xy <- seq(100, n, by=19)
plot(show.xy, estimates[show.xy], type='l', xlab='Repetition no.',
     ylab='Estimated value', main='Estimating pi')
```

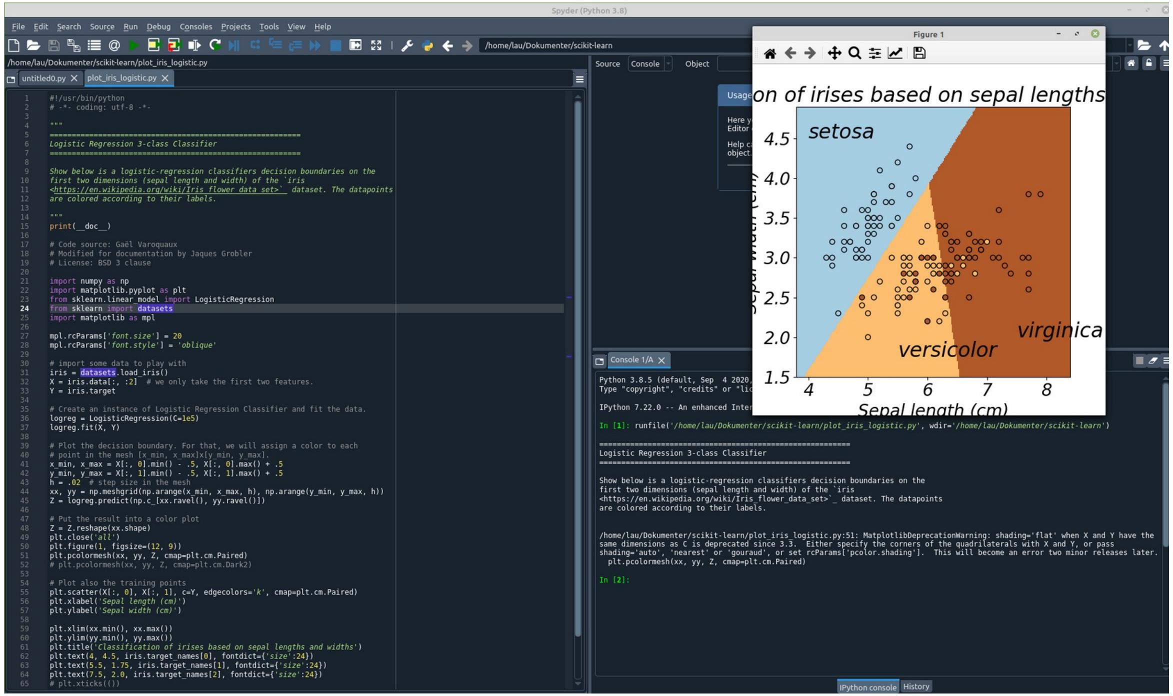


RStudio – Integrated Development Environment



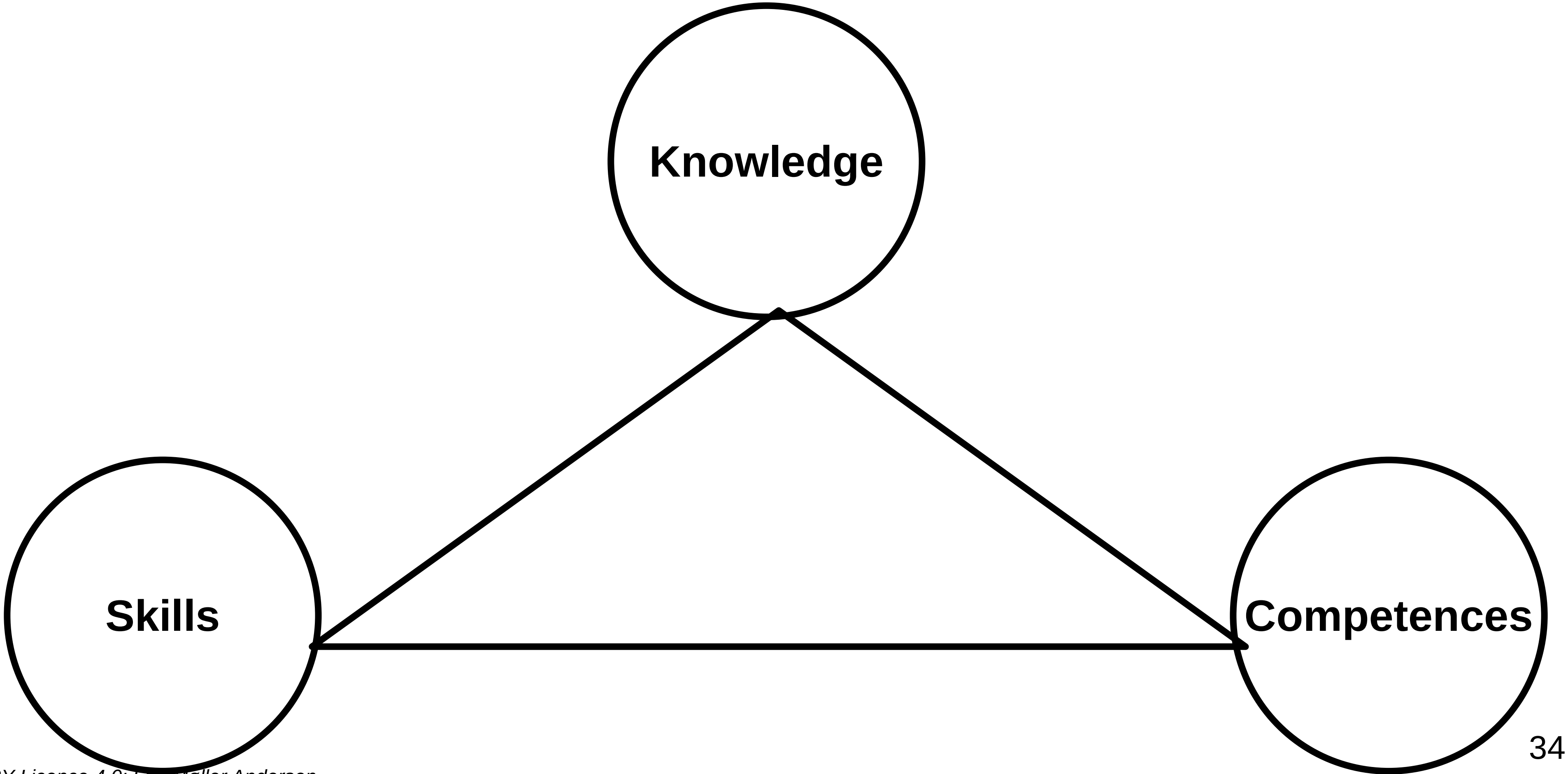
https://www.rstudio.com/products/rstudio/download/

Spyder – Integrated Development Environment



Part of installed environment

Academic regulations – objectives



Academic regulations – objectives

Knowledge:

- demonstrate understanding of statistical techniques relying on the Generalised Linear Model

- demonstrate understanding of hierarchical modeling methods

- demonstrate understanding of basic machine learning concepts.

Academic regulations – KNOWLEDGE

You should thus be able to answer questions like:

What is logistic regression?

What are some naturally occurring hierarchies within research data?

What is cross-validation?

What is a mixed-effects model?

Academic regulations – objectives

Skills:

- build and evaluate models of hierarchically structured data

- integrate machine learning procedures in data analysis

- communicate analysis processes, results and interpretation.

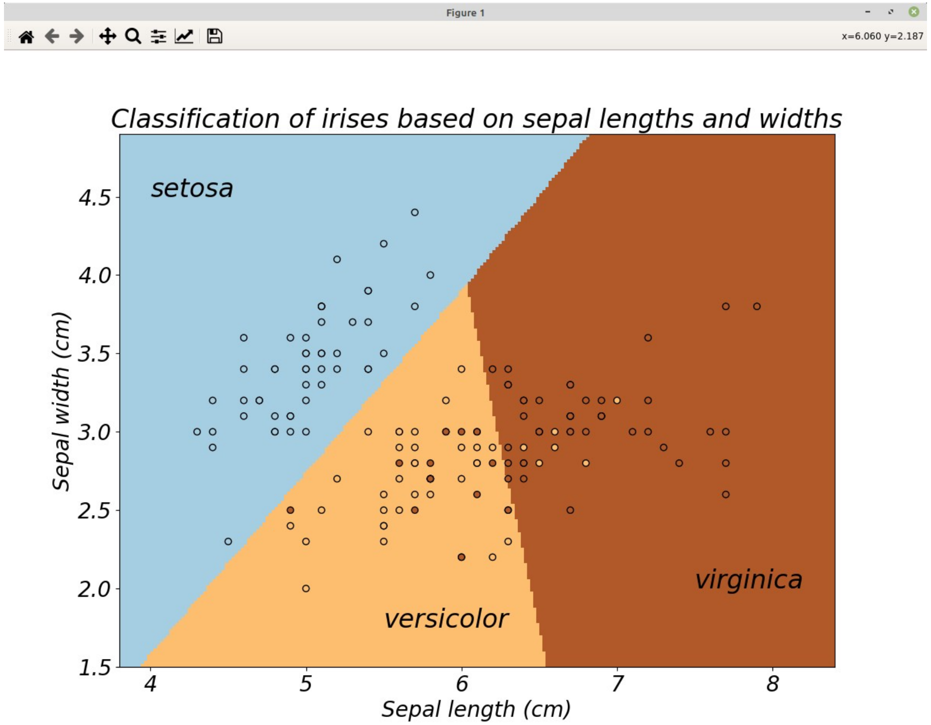
Academic regulations – SKILLS

build and evaluate models of hierarchically structured data

```
Linear mixed model fit by REML ['lmerMod']
Formula: Height ~ Gender + (1 | Family)
   Data: height
```

Academic regulations – SKILLS

integrate machine learning procedures in data analysis



Academic regulations – SKILLS

communicate analysis processes, results and interpretation

(Andersen et al., 2019)

Mixed model analyses (McCulloch & Neuhaus, 2005) were applied to investigate how top-down expectations (No. of Possible Targets) affected subjective experience and objective performance. We performed model comparisons between models that did or did not include the relevant fixed effects and interactions to find the best compromise between an explanatory and a parsimonious model. This was done using the log-likelihood ratio between two models because this ratio approximates a chi-square distribution. A chi-square test can thus be used to assess whether two models differ significantly, where the test statistic is the log-likelihood-ratio and the degrees of freedom is the difference in free parameters of the two models.

Academic regulations – objectives

Competences:

- independently decide on data analysis methods, given a data set and a research question

- justify decision making when pre-processing messy data for data analysis.

Difference between *skills* and *competences:*
Skills are specific
Competences are more generic

Academic regulations – COMPETENCES

independently decide on data analysis methods, given a data set and a research question

justify decision making when pre-processing messy data for data analysis.

EXAMPLE QUESTION: Can native Danish speakers tell apart soft d’s (ð) and l’s if they are embedded in English speech, e.g. maðfunction/malfunction

EXAMPLE DATA: Dataset with response times and discrimination responses

Exam; portfolio

- Ongoing assignments to be solved in small groups
  - Will be done in *R Markdown* (.Rmd)
- Final portfolio:
  - Revised assignments, handed in as short reports with reproducible code on GitHub
    - Assignment 1: Using mixed effects modelling to model hierarchical data (Winter & Grawunder, 2012)
    - Assignment 2: Mixed effects modelling of response times, response counts, and accuracy (Andersen et al., 2019)
    - Assignment 3: Using logistic regression to classify subjective experience from brain data
    - Assignment 4: Dimensionality reduction, finding the signal among the noise

Re-examination
Will be done during the summer (tip: stay on schedule, such that you can enjoy your summer)

The structure of lectures

- Tuesdays 11.00-13.00 (NB: Academic quarter)
- A mix between the general theoretical framework and a few formulas
- Code will be incorporated
- Group discussions
- Please don’t hold back on your questions
- Slides will be shared *just before* the lecture

The structure of a practical exercise

- Wednesdays 10.00-12.00 and 16.00-18.00
- The practical exercises are led by either Emil or me.
  - An introduction to the task
  - Group work (3-4 people)
  - Data and an instructions file are shared
  - Intermittent discussion of pressing issues
- Feedback is given per assignment and collectively

The structure of a practical exercise

- You will be separated into two groups
  - Emil and I will take turns leading the instructor sessions
    - Your time slot is fixed, either 10-12, (Class 2) or 16-18 (Class 1)
    - We, Emil and Lau, show up alternately.
  - Lau: 10-12 – Weeks 3, 5, 7 and 9
  - Emil: 10-12 – Weeks 2, 4, 8, 10
  - Together: Weeks 1 (intro), 6 (mid-way evaluation), 11 (final evaluation)

Homework

- Focus is on developing data analysis and reporting skill
- Readings:
  - Mainly covers background and are sometimes more in-depth than the lectures
  - One have-to-read
  - ... and some recommended (that are not strictly necessary)
- Practical exercises and assignments
  - Hand them in on GitHub on the following Monday (23:59)

Questions about homework and the subject in general

- Lau: Office hours between 8 and 9 on Mondays
- We, Emil and I, collect questions during the week and try to answer them in the classes
- https://cryptpad.fr/pad/#/2/pad/edit/U21qNTbLgfkRiGZU1bnmDE2o/

Student counsellors

You can contact any of us for help - but each of us has an academic specialisation, so we may refer you to another student counsellor for certain questions. If we can't answer your questions, we will normally be able to direct you to someone who can.

Sofia Madsen

- Scandinavian Languages and Literature

Josephine Brunsgaard

- Cognitive Science

Kristina Duun

- Linguistics and Cognitive Semiotics

Emilie Vinther

- Experience Economy and Cultur Of Events

The course plan

Week 1: *Introduction and Why are we here?*, September 14 & 15

Instructor sessions: *Setting up* R *and* Python *and recollection of the general linear model*

Week 2: *Linear Mixed Effects Models*, September 21 & 22

Instructor sessions: *Modelling random effects – and how do they differ from fixed effects?*

Week 3: *Generalized Linear Mixed Effects Models*, September 28 & 29

Instructor sessions: *What to do when the response variable is not continuous?*

Week 4: *Explanation and Prediction*, October 5 & 6

Instructor sessions: *Why are good explanations sometimes bad?*

Week 5: *Evaluating and comparing models,* October 12 & 13

Instructor sessions: *How do we assess how models compare to one another?*

Week 6: *Mid-way evaluation and Machine Learning Intro,* November 2 & 3

Instructor sessions: *Moving the goal away from explanation towards prediction and getting Python running*

Week 7: *Linear regression revisited (machine learning)*, November 9 & 10

Instructor sessions: *How to constrain our models to make them more predictive*

Week 8: *Logistic regression (machine learning)*, November 16 & 17

Instructor sessions: *Categorizing responses based on informed guesses*

Week 9: *Dimensionality Reduction, Principled Component Analysis (PCA)*, November 23 & 24

Instructor sessions: *What to do with very rich data?*

Week 10: *Organising and preprocessing messy data,* November 30 and December 1

Instructor sessions: *How to clean up?*

Week 11: *Final evaluation and wrap-up of course*, December 7 & 8

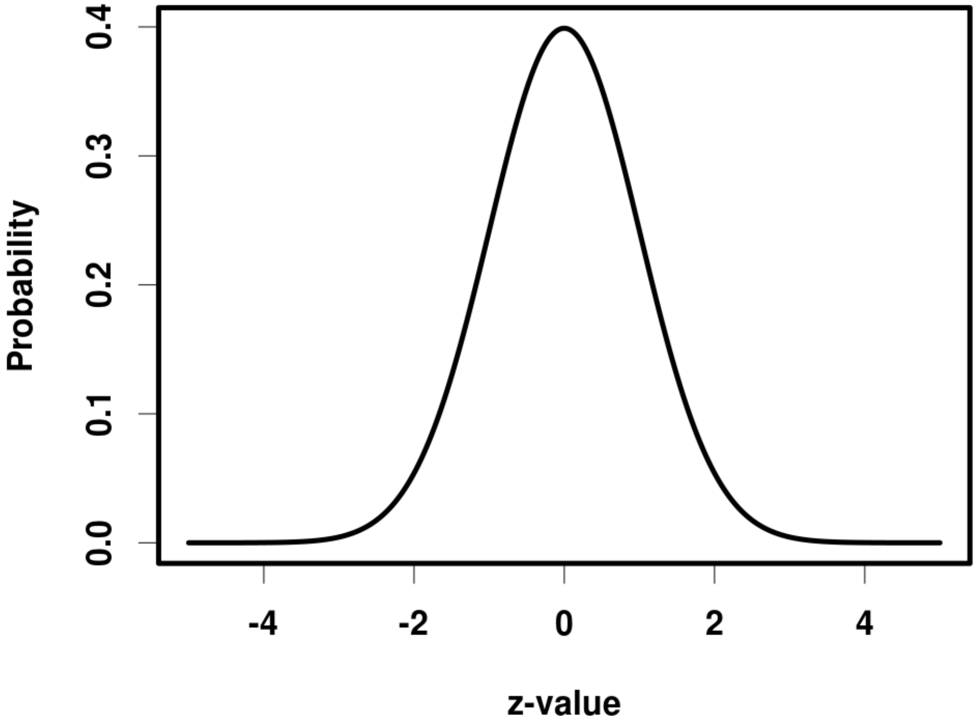
Instructor sessions: *Ask anything!*

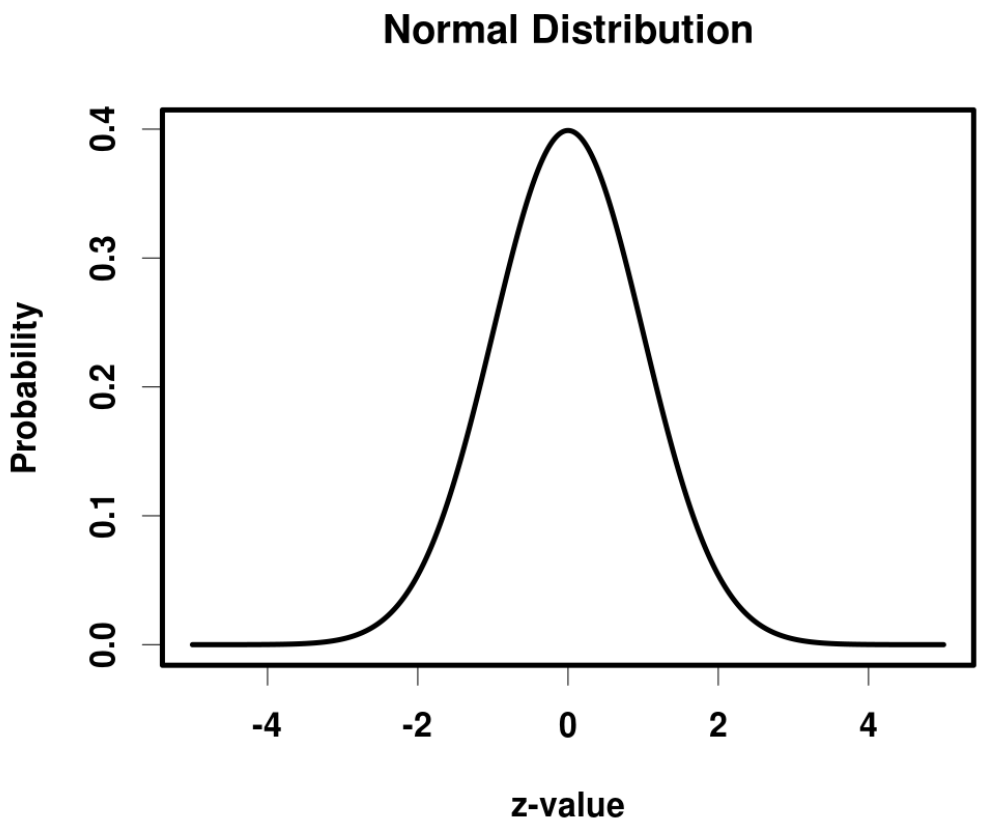
What do you remember?

Group discussions

- What does a *z*-value express?
- What does a *p*-value express?
- What is the relation between a *t*-value and a *z*-value?
- What do the following symbols signify in these equations?
  - $Y = X\beta + \epsilon$
  - $\hat{\beta} = \left(X^T X\right)^{-1} X^T Y$

What does a z-value express?

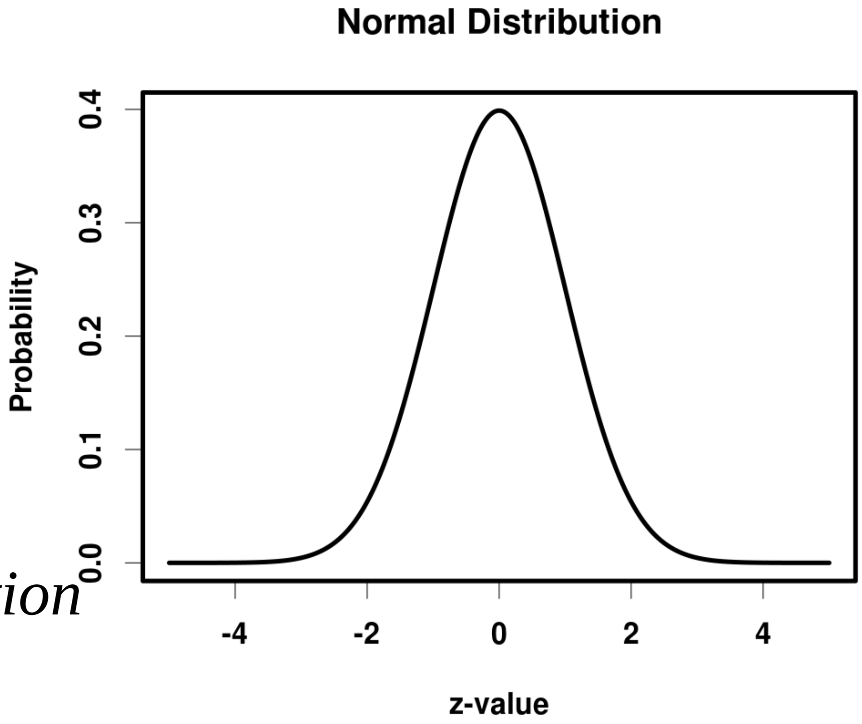

Probability
0.0
0.1
0.2
0.3
0.4
-4
-2
0
2
4
z-value


Normal Distribution
Probability
0.0
0.1
0.2
0.3
0.4
-4
-2
0
2
4
z-value

$$z = \frac{\bar{X} - P}{SE}$$

$$SE = \frac{\bar{\sigma}}{\sqrt{n}}$$

$\bar{X}$: *sample mean*
P : *population mean*
SE : *standard error*
$\bar{\sigma}$: *estimated standard deviation*
n : *number of observations*



$$z = \frac{\bar{X} - P}{SE}$$

$$SE = \frac{\bar{\sigma}}{\sqrt{n}}$$

$\bar{X}: sample\,mean$
$P: population\,mean$
$SE: standard\,error$
$\bar{\sigma}: estimated\,standard\,deviation$
$n: number\ of\ observations$

If we are testing the null hypothesis, then what is *z*?

$$P = 0$$

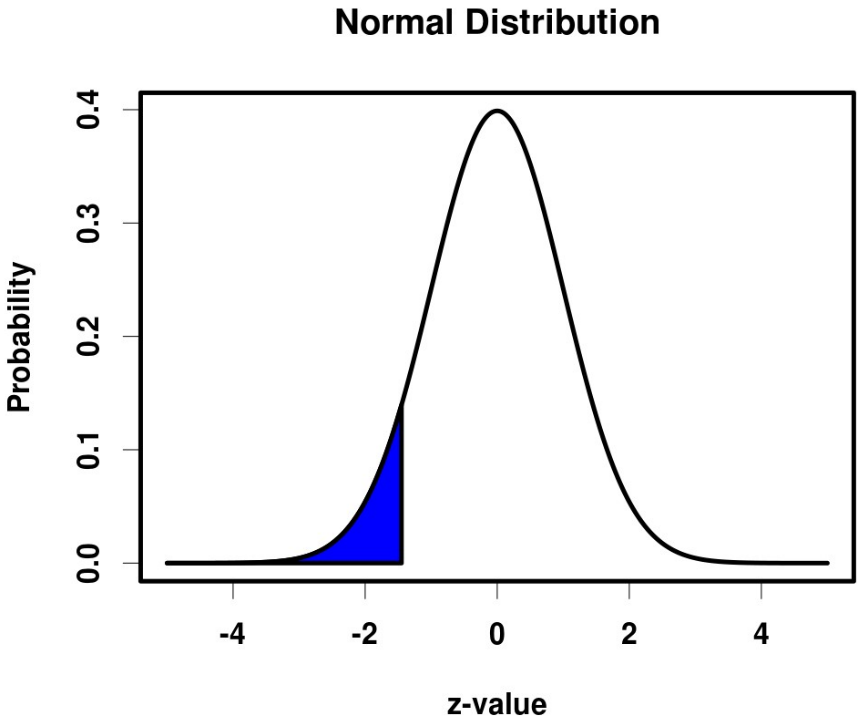
$$z = \frac{\bar{X}}{SE}$$

What does a *p*-value express?

When assuming the truth of the null hypothesis, the *p*-value is the probability of observing a given *z*-value or one that is more extreme.

If $z = -1.45$, what is *p* then?

The area of the blue curve, $p = 0.0735$



Bonus questions:
1) what is the total area under the curve equal to?
2) why does the *p*-value *not* say about the truth of the null hypothesis?

What is the relation between a *t*-value and a *z*-value?

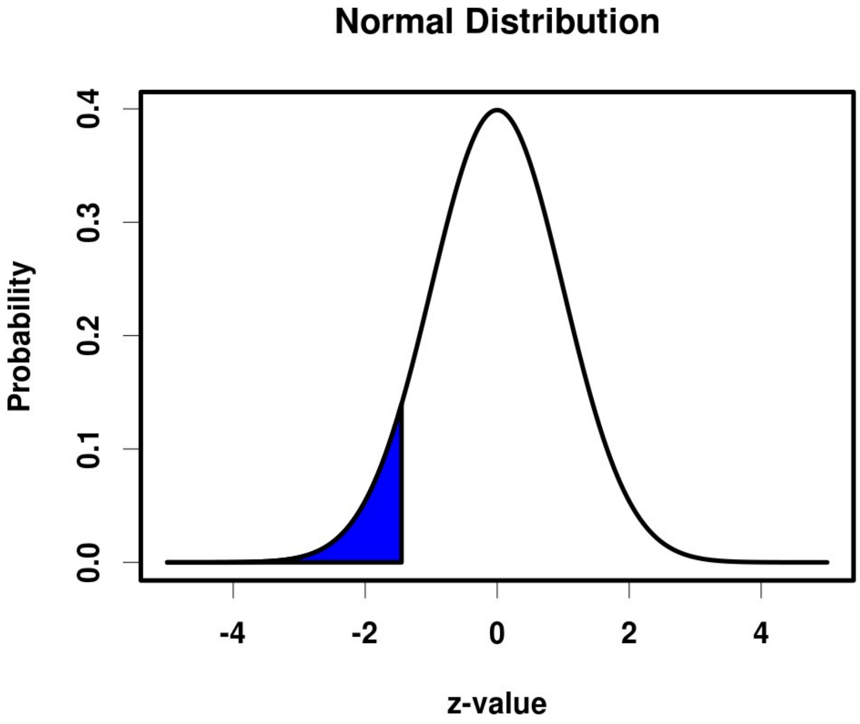
It is the identity relation

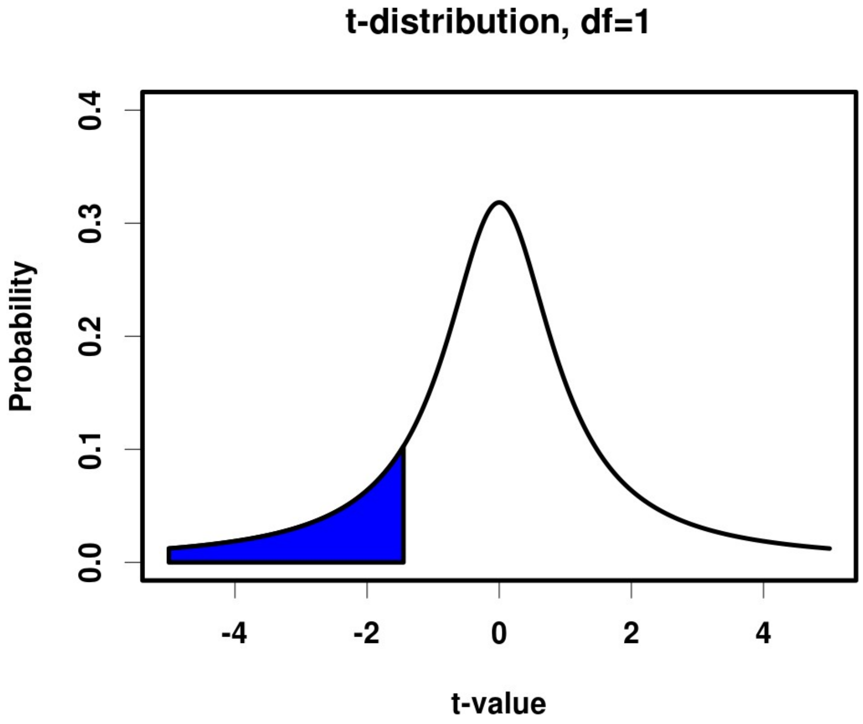
$$t = \frac{\bar{X} - P}{SE} = z$$

... But the *p*-values associated with *t*- and *z*-values differ. For *t*-values, the *p*-values depend on the degrees of freedom; for the *z*-values, they do not

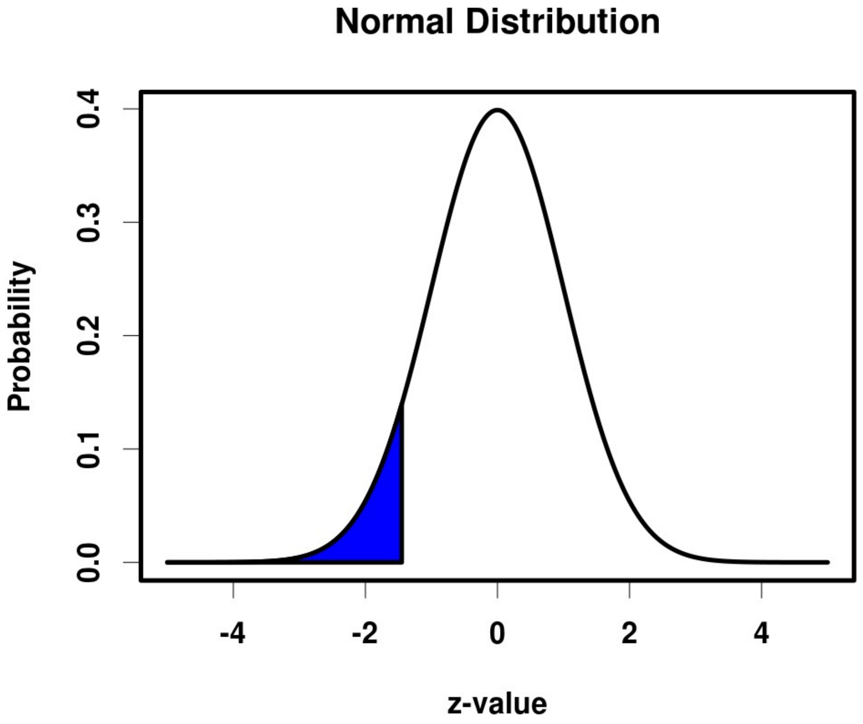
Crucial assumption when you are performing a *z*-test is that the population variance is known. When this is not known a *t*-test is more appropriate, as it takes the uncertainty of the estimated variance into the equation

We normally do not know the population variance

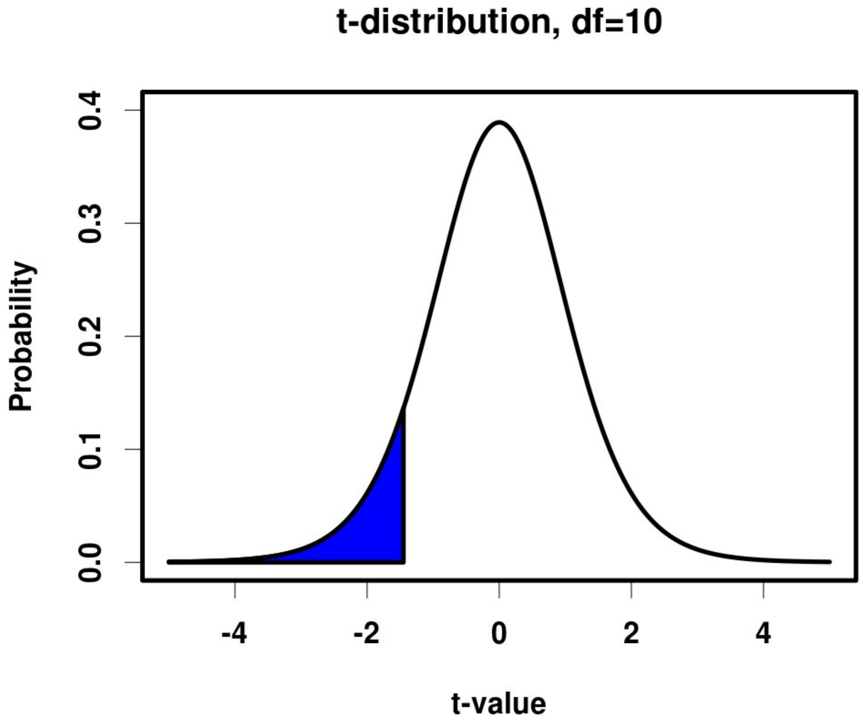




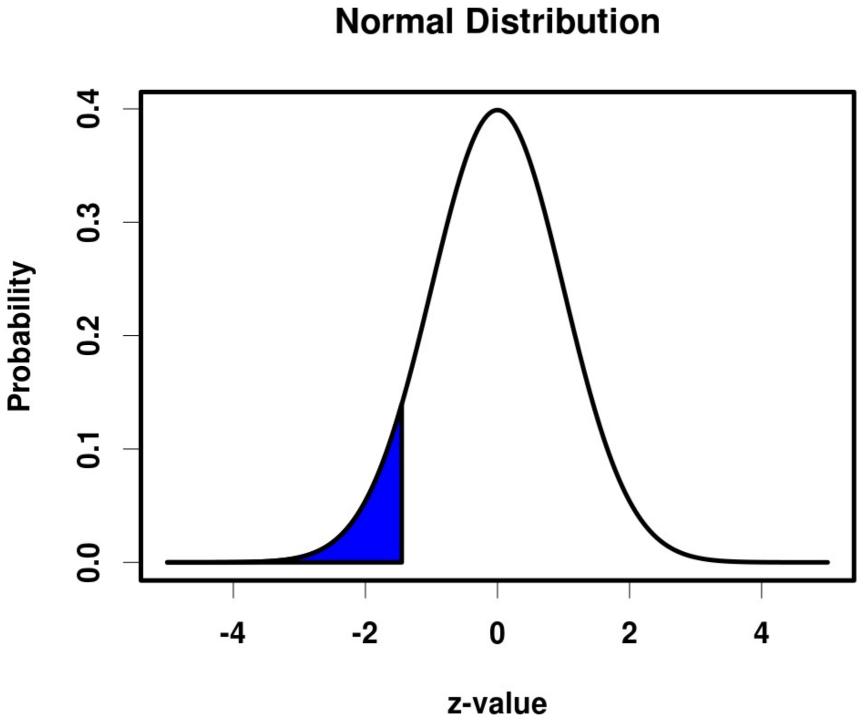
$p = 0.192$



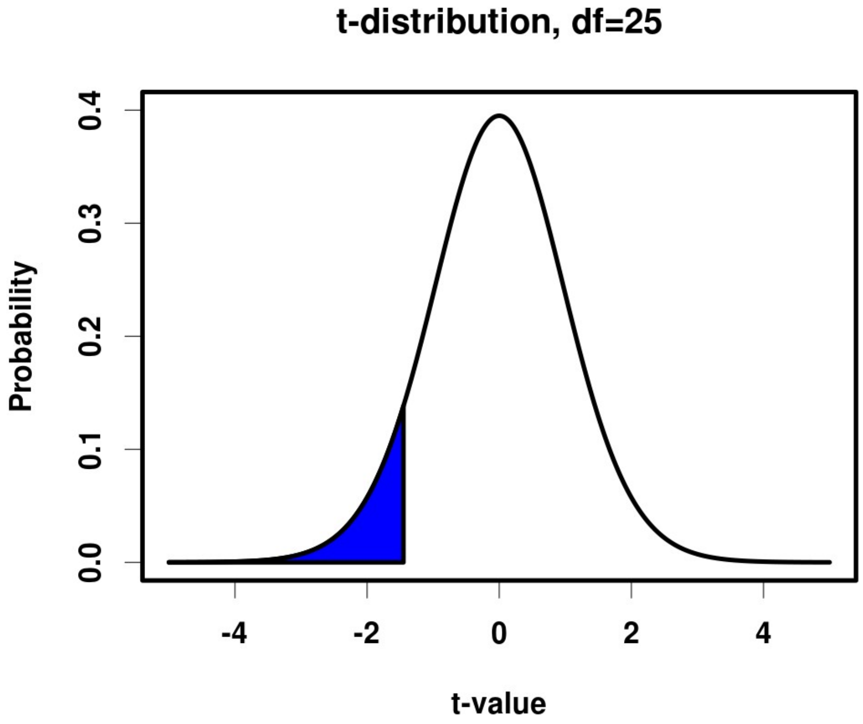
$p = 0.0735$



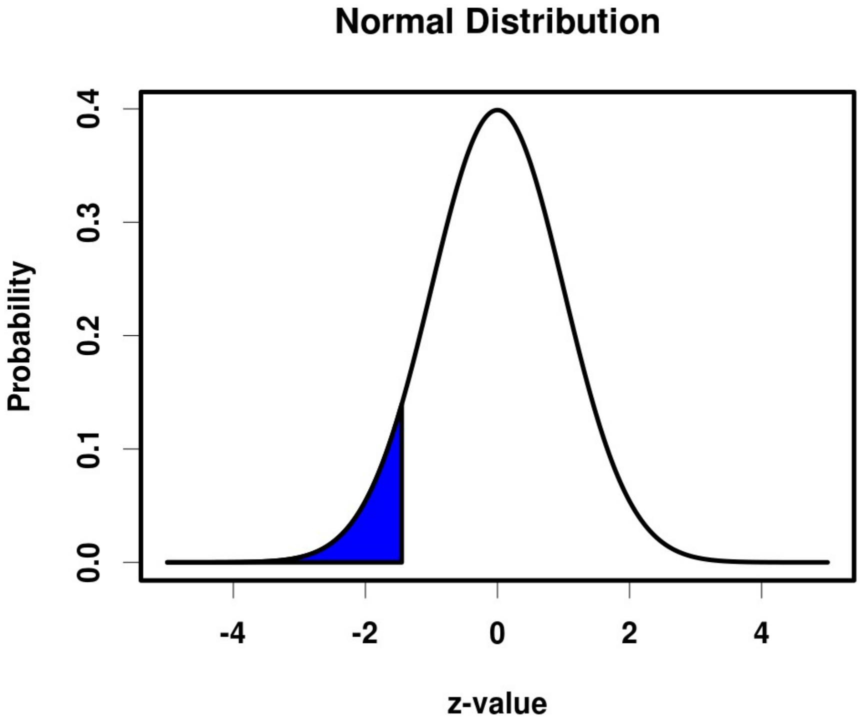
$p = 0.0888$



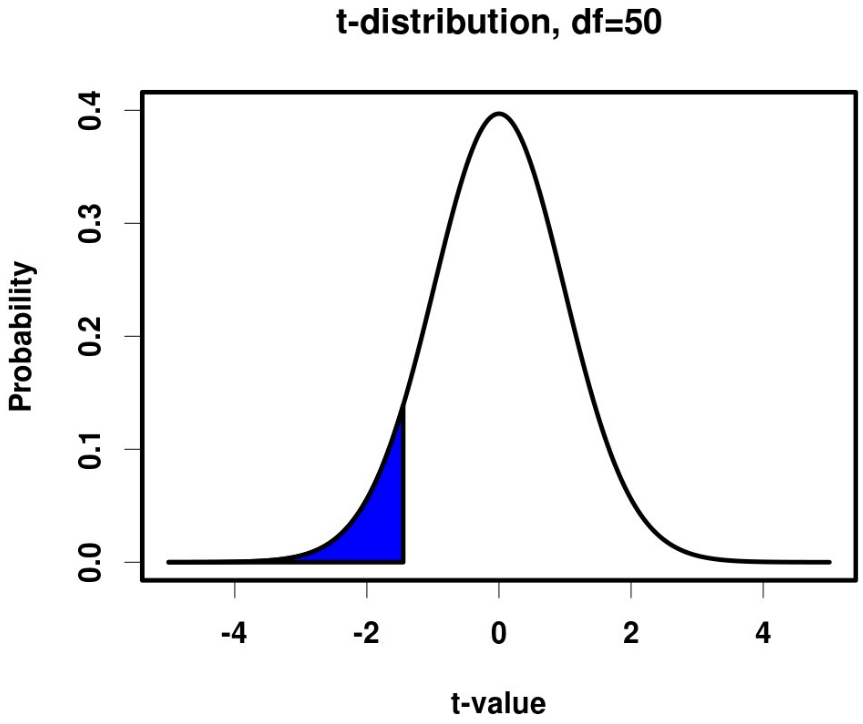
$p = 0.0735$

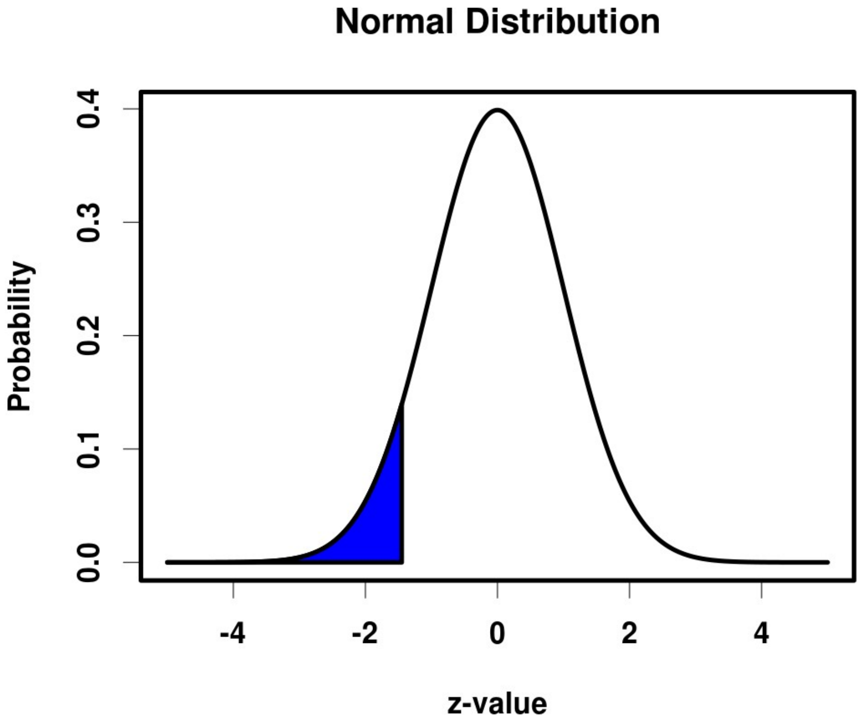


$p = 0.0797$



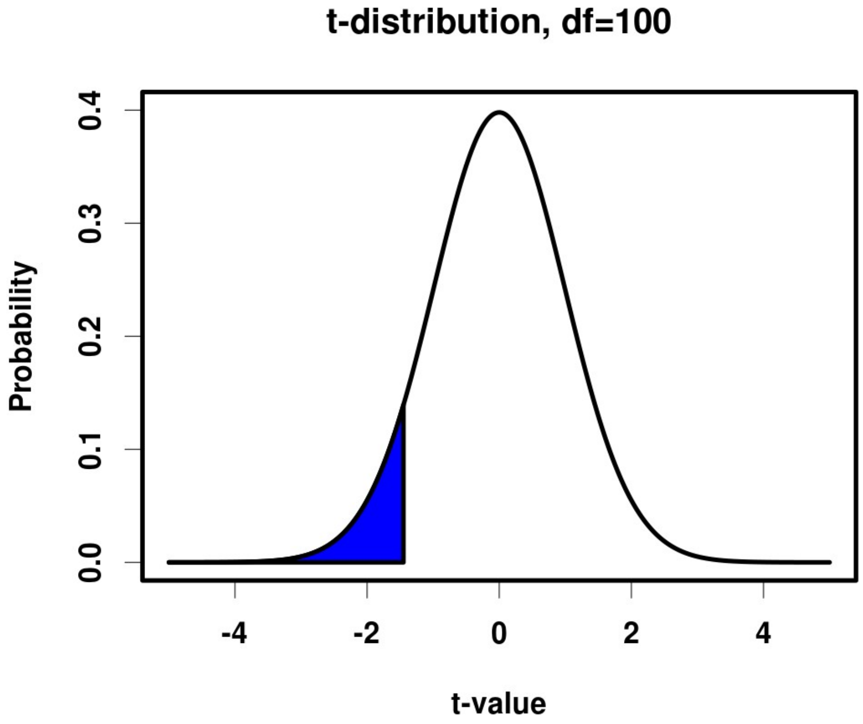
$p = 0.0735$



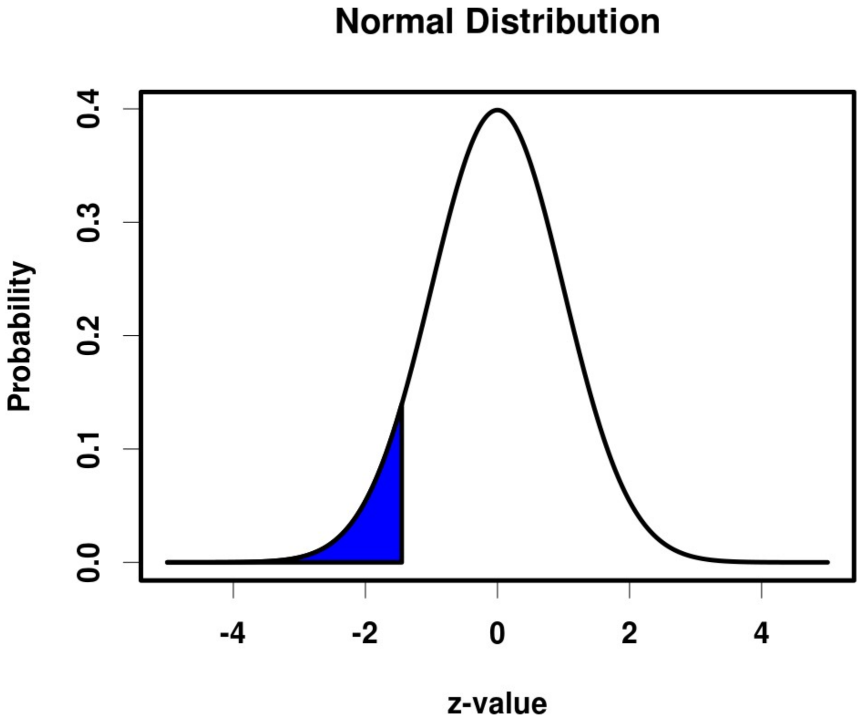


$p = 0.0767$

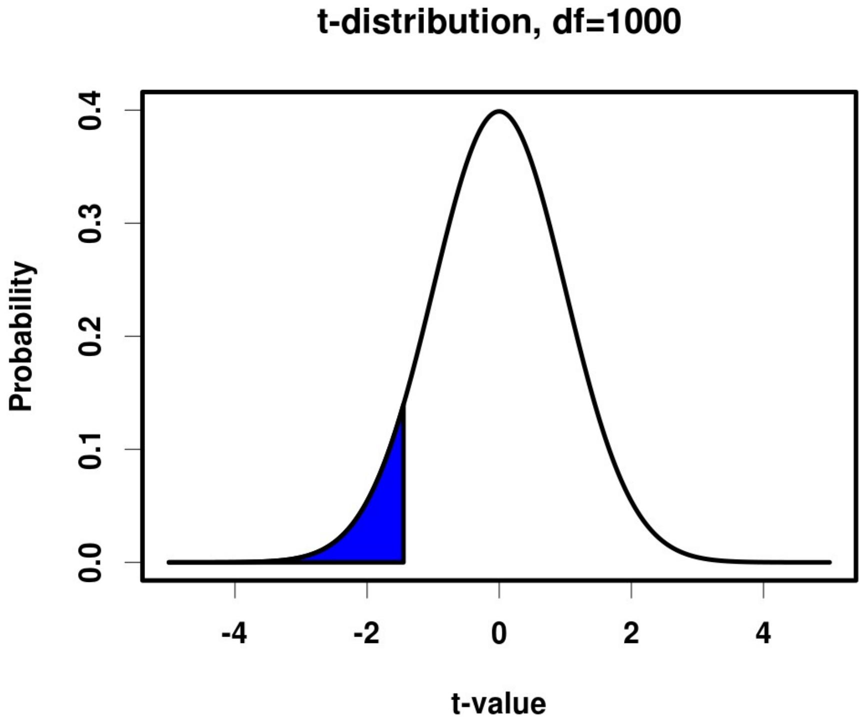
$p = 0.0735$



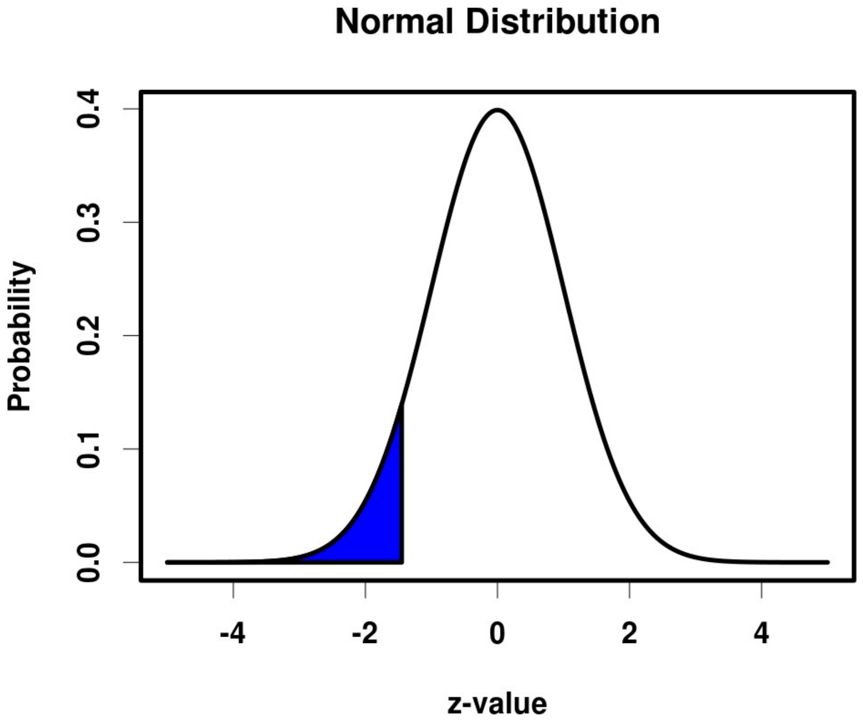
$p = 0.0751$



$p = 0.0735$



$p = 0.0737$



$p = 0.0735$

Bonus question:
Which distribution does the *t*-distribution converge on when the degrees of freedom goes towards infinity?

What do the following symbols signify in these equations?

$$Y = X\beta + \epsilon$$

$$\hat{\beta} = \left(X^T X\right)^{-1} X^T Y$$

$\boldsymbol{Y = X\beta + \epsilon}$ **(The General Linear Model)**

Y: a column vector with J observations (known)

X: the design matrix (known), size: $J \times L$

β: a column vector with L (unknown) model parameters

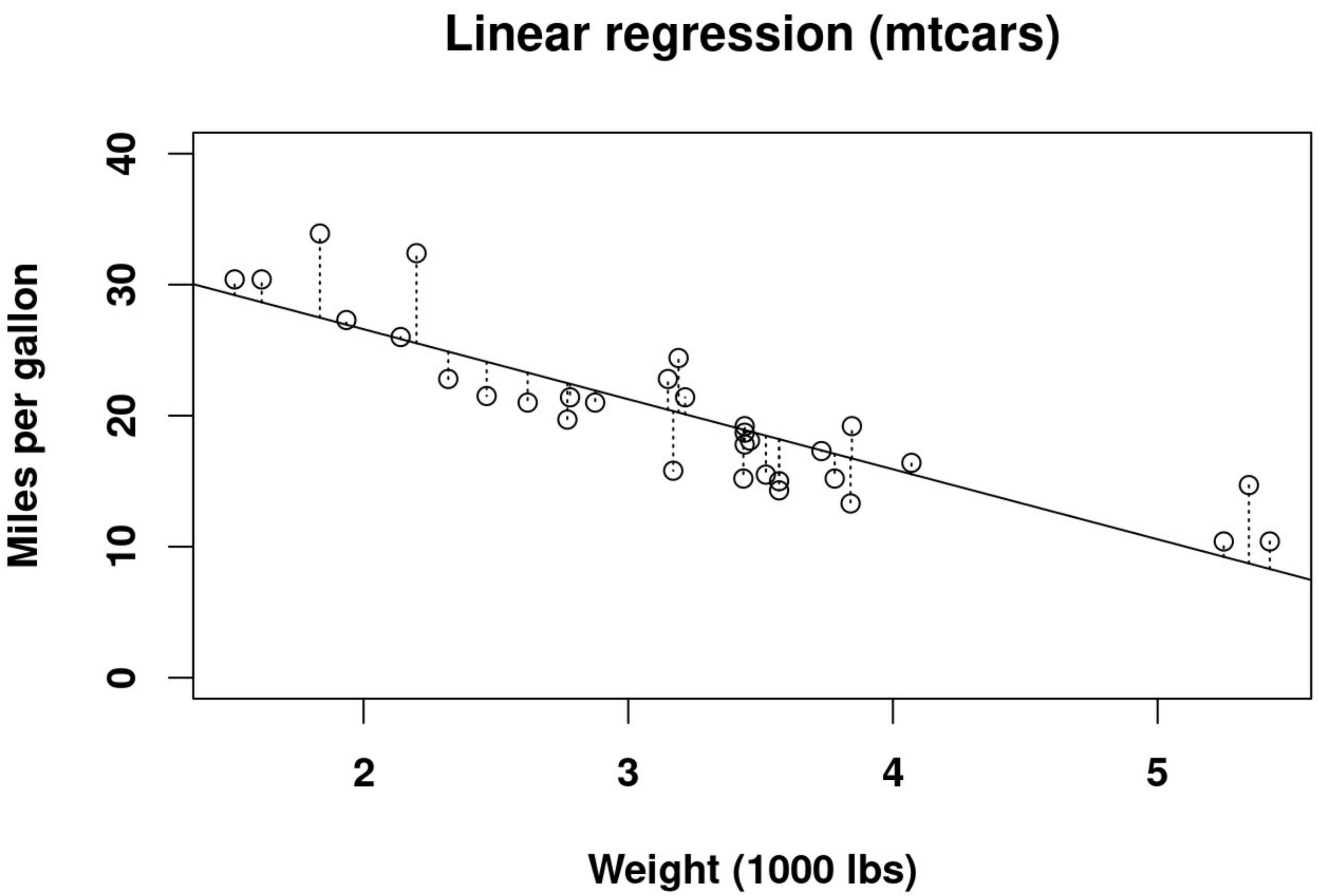
ϵ: a column vector with J residuals

$\boldsymbol{\hat{\beta} = (X^T X)^{-1} X^T Y}$ **(least squares solution)**

$\hat{\beta}$: the estimated L model parameters

X^T: the transpose of the design matrix

$(X^T X)^{-1}$: the inversion of the $L \times L$ matrix


Linear regression (mtcars)
Miles per gallon
Weight (1000 lbs)
0
10
20
30
40
2
3
4
5

Next time

- Introducing mixed effects modelling
  - Fixed effects
  - Random effects
  - Nested effects

References

- Andersen, L.M., 2018a. Group Analysis in FieldTrip of Time-Frequency Responses: A Pipeline for Reproducibility at Every Step of Processing, Going From Individual Sensor Space Representations to an Across-Group Source Space Representation. Front. Neurosci. 12. https://doi.org/10.3389/fnins.2018.00261
- Andersen, L.M., 2018b. Group Analysis in MNE-Python of Evoked Responses from a Tactile Stimulation Paradigm: A Pipeline for Reproducibility at Every Step of Processing, Going from Individual Sensor Space Representations to an across-Group Source Space Representation. Front. Neurosci. 12. https://doi.org/10.3389/fnins.2018.00006
- Andersen, L.M., Overgaard, M., Tong, F., 2019. Visual expectations change subjective experience without changing performance. Consciousness and Cognition 71, 59–69. https://doi.org/10.1016/j.concog.2019.03.007
- Andersen, L.M., Lundqvist, D., 2019. Somatosensory responses to nothing: An MEG study of expectations during omission of tactile stimulations. NeuroImage 184, 78–89. https://doi.org/10.1016/j.neuroimage.2018.09.014
- Andersen, L.M., Overgaard, M., Tong, F., 2019. Visual expectations change subjective experience without changing performance. Consciousness and Cognition 71, 59–69. https://doi.org/10.1016/j.concog.2019.03.007
- Andersen, L.M., Oostenveld, R., Pfeiffer, C., Ruffieux, S., Jousmäki, V., Hämäläinen, M., Schneiderman, J.F., Lundqvist, D., 2017. Similarities and differences between on-scalp and conventional in-helmet magnetoencephalography recordings. PLOS ONE 12, e0178602. https://doi.org/10.1371/journal.pone.0178602
- Andersen, L.M., Pedersen, M.N., Sandberg, K., Overgaard, M., 2016. Occipital MEG Activity in the Early Time Range (<300 ms) Predicts Graded Changes in Perceptual Consciousness. Cereb. Cortex 26, 2677–2688. https://doi.org/10.1093/cercor/bhv108
- Andersen, L.M., Pfeiffer, C., Ruffieux, S., Riaz, B., Winkler, D., Schneiderman, J.F., Lundqvist, D., 2020. On-scalp MEG SQUIDs are sensitive to early somatosensory activity unseen by conventional MEG. NeuroImage 221, 117157. https://doi.org/10.1016/j.neuroimage.2020.11715
- Winter, B., Grawunder, S., 2012. The phonetic profile of Korean formal and informal speech registers. Journal of Phonetics 40, 808–815. https://doi.org/10.1016/j.wocn.2012.08.006