

Forskningstræning

Statistik og forskningsmetodik

2025

Program – Del 1

- Dag 1 (9:30-15:30)
 - Hvorfor statistik/forskningsmetodologi
 - Øvelse med retningslinjer forståelse
 - Statistik
 - Basal forståelse for variation og stikprøver, observeret vs. forventet
 - T-test/2x2 tabeller, chi2, binomial test
 - AI integration i læring og løsning af tekniske udfordringer
 - Overlevelses & tid-til-event statistik
 - Databaser
 - Litteratursøging
 - Opgave

- Dag 2
 - 9.30-12.45
 - Opgave
 - Repetition
 - Regression

Program – Del 2

- Dag 3
 - Repetition
 - Power-beregninger

- Introduktion til basal statistik
 - Basale koncepter omkring kausalitet, variation og stikprøver

Test 1

3 ; 5 ; 8 ; 20 ; 36 ; 64 ; 256

- Hvad er medianen og gennemsnittet?
 - A. Median 56 & gennemsnit 20
 - B. Median 8.23 & gennemsnit 55
 - C. Median 20 & gennemsnit 56
 - D. Median 63.5 & gennemsnit 16

Test 2

- En patient har en positive test.
 - Hvad er sandsynligheden for at patienten fejler sygdommen?
 - Testen har en sensitivitet på 90% og en specifitet på 90%.
 - Pre-test sandsynligheden for sygdommen er 1:100

- A.~1%**
- B.~10%**
- C.~50%**
- D.~90%**
- E.~100%**

Test 2

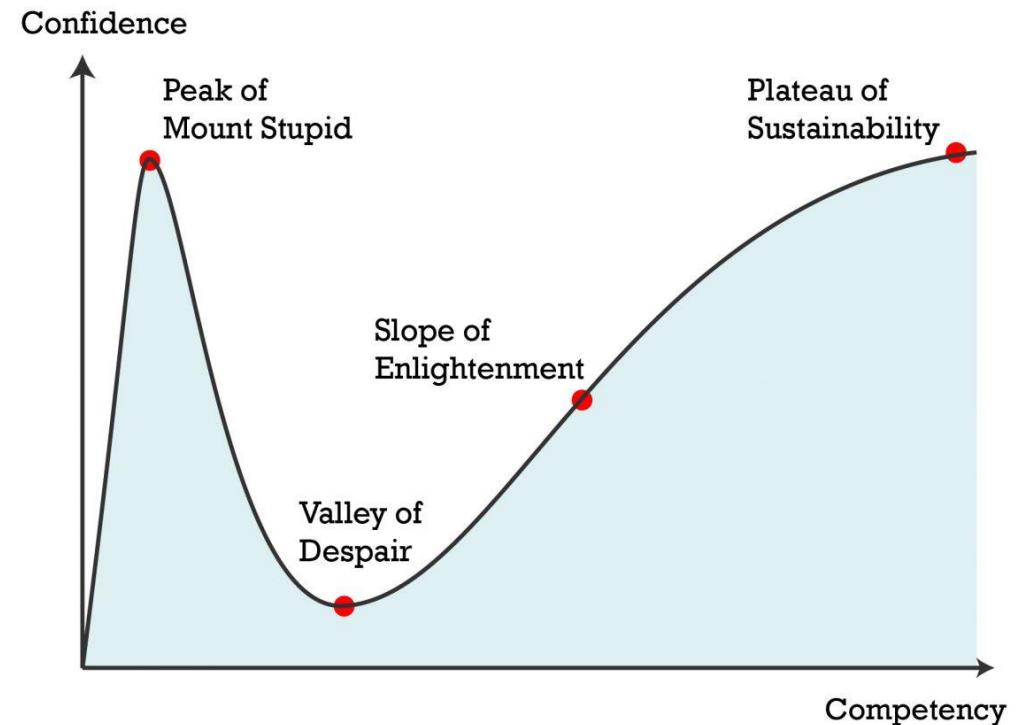
- En patient har en positive test.
 - Hvad er sandsynligheden for at patienten fejler sygdommen?
 - Testen har en sensitivitet på 90% og en specifitet på 90%.
 - Pre-test sandsynligheden for sygdommen er 1:100

- A.~1%**
- B.~10%**
- C.~50%**
- D.~90%**
- E.~100%**



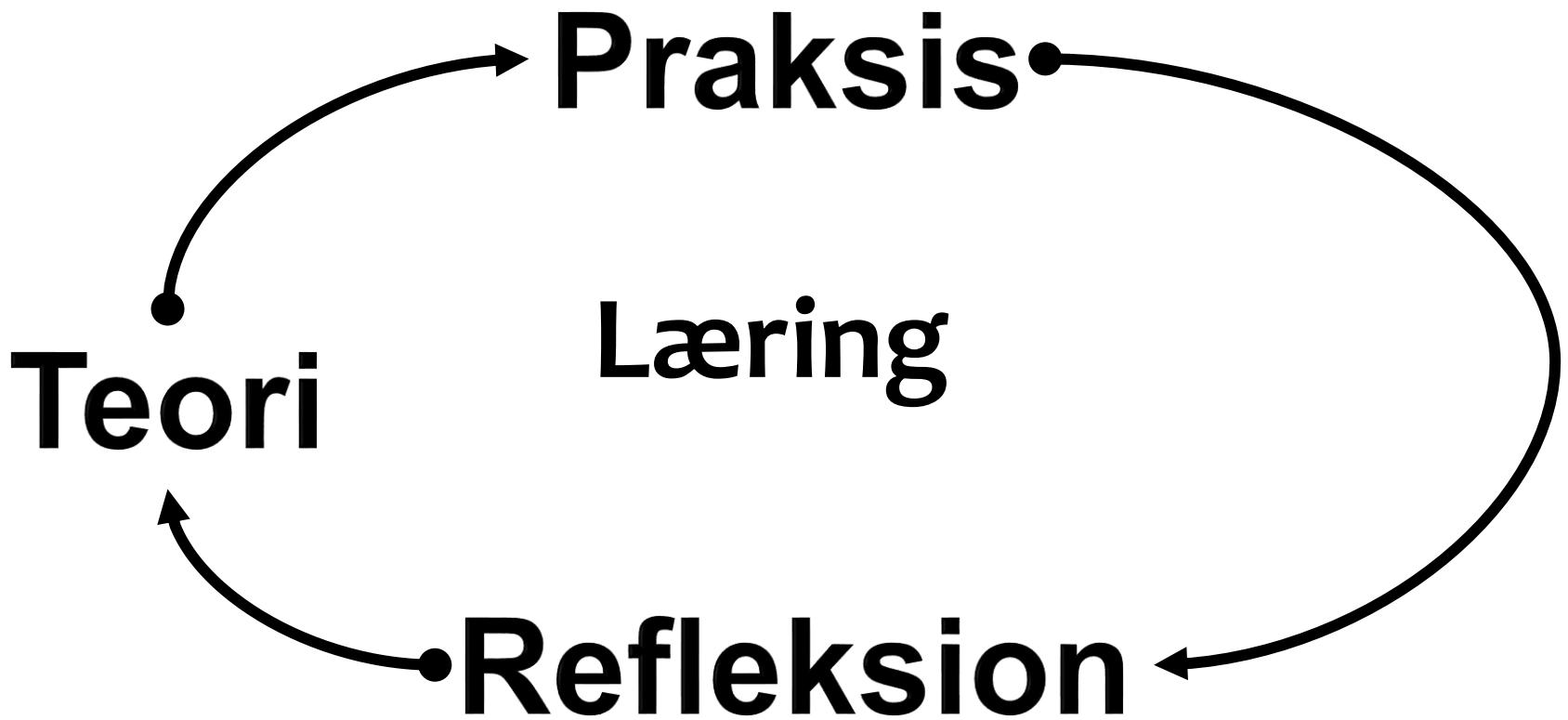
Disclaimer

- Mange genveje
 - Nogle af dem tvivlsomme
- Mange muligheder for at lærer videre
 - Youtube
 - AI
 - Hjælpe filer
 - Basis bøger



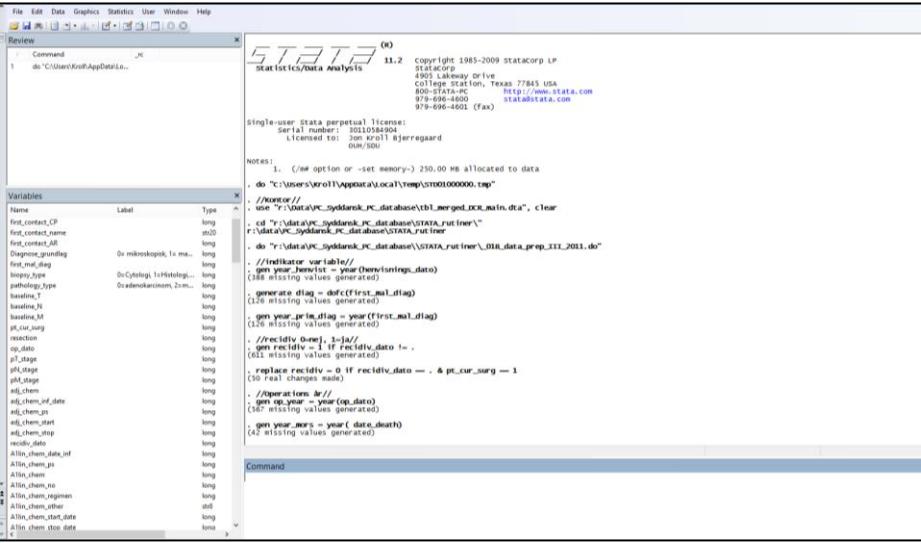
Lidt om programmer

- Kommercielle
 - SAS
 - STATA
 - SPSS
 - Graphpad
 - Excel
 -
- Open Source
 - R
 - R studio
 - CDC – epi-info
 - SOFA
 -



Udfordringer

- Datahåndtering
 - Databaser
- Statistik programmer
 - IT kundskaber
- Glemt viden
 - $f(x) = a_0 + \sum_{n=1}^{\infty} (a_n + b_n)$
- Misforståelser



```
clear
set obs 10
gen zero = -0.01
gen truemean = 178
gen tal1 = rnormal(178,15)
sum tal1, detail
scalar mean1 = r(mean)
scalar sd1 = r(sd)
gen mean1=r(mean)

ci tal1
gen cil1 = r(lb)
gen ciu1 = r(ub)

scalar cil1 = r(lb)
scalar ciu1 = r(ub)
scalar number1=r(N)

twoway (histogram tal1, width(1) color(green)) || /*
*/ function y=normalden(x,mean1,sd1), range (140 240) lstyle(foreground) lwidth(thick) || /*
*/ function y=178, range(-0.05 0) horizontal lwidth(vthick) lcolor(maroon) || /*
*/ function y=`=mean1', range(-0.05 0) horizontal lwidth(thick) lcolor(red) || /*
*/ function y=`=ciu1', range(-0.05 0) horizontal lwidth(thick) lcolor(gs0) lpattern(dash) || /*
*/ function y=`=cil1', range(-0.05 0) horizontal lwidth(thick) lcolor(gs0) lpattern(dash) || /*
*/ , name("tal001", replace) title ('=number1') legend(off)
```

Læge ≠ Statistiker

- Som kliniker er det vigtigt at kunne tolke resultater
- Som forsker er det vigtigt at udfører arbejde af høj kvalitet
- Som statistiker er det vigtigt at kende alt baggrunden

Læge ≠ Statistiker



AMERICAN STATISTICAL ASSOCIATION
Promoting the Practice and Profession of Statistics®

732 North Washington Street, Alexandria, VA 22314 • (703) 684-1221 • Toll Free: (888) 231-3473 • www.amstat.org • www.twitter.com/AmstatNews

AMERICAN STATISTICAL ASSOCIATION RELEASES STATEMENT ON STATISTICAL SIGNIFICANCE AND P-VALUES

*Provides Principles to Improve the Conduct and Interpretation of Quantitative
Science*

March 7, 2016

“Over time it appears the *p*-value has become a gatekeeper for whether work is publishable, at least in some fields,” said Jessica Utts, ASA president. “This apparent editorial bias leads to the ‘file-drawer effect,’ in which research with statistically significant outcomes are much more likely to get published, while other work that might well be just as important scientifically is never seen in print. It also leads to practices called by such names as ‘*p*-hacking’ and ‘data dredging’ that emphasize the search for small *p*-values over other statistical and scientific reasoning.”

STATISTICS DONE WRONG

**The Woefully
Complete Guide**

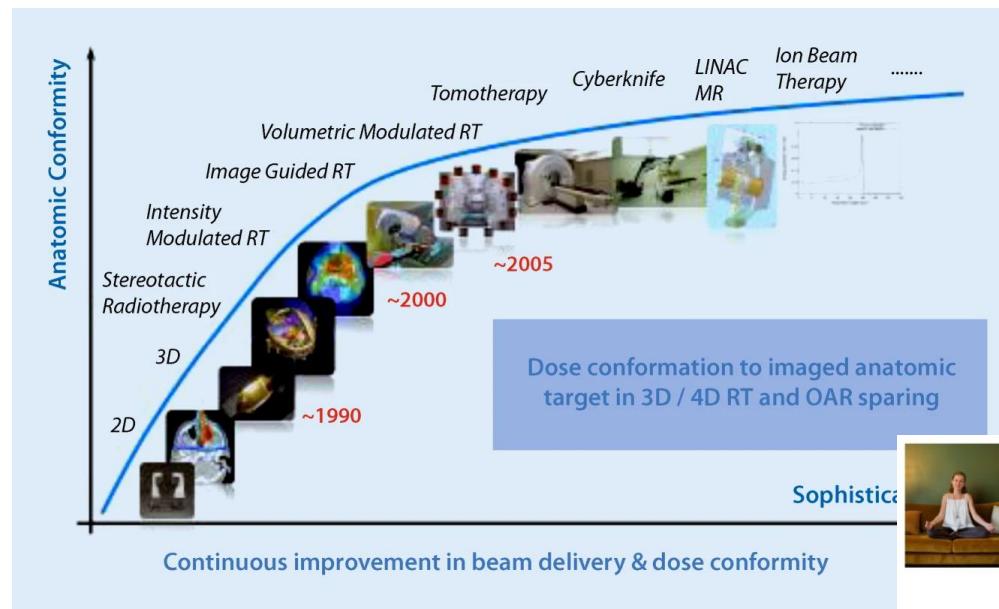
by Alex Reinhart



no starch
press

San Francisco

Hvorfor statistik/forskning?



Gongbad v/Elin Karbech sted Pakhuset i Næstved
Et gongbad er en beroligende og beklagende helt ned på collegen. Du kan sidde eller ligge. Slip det af stedne noget. Slip præstation og at være noget bestemt. Modtag. Lad dig forlade fra krystaliske, gong man. og meditationer.
Onsdag d. 26. november 2025 kl. 19:00 til 21:00
Pakhuset, Næstved
Lukket for tilmelding

Fra 75 kr.

Læs mere

ZOOM - Keto og Kræft - strategi eller stress for kroppen? v/ Christina Santini
Kan en ketogen kost trænge en strategisk støtte under kræfteforløb? Risikerer man i virkeligheden at belaste kroppen yderligere? Hvorør kan keto virke som støtte til behandlingen? Hvorør virker det som stressfaktor, for kroppen og dens immunforsvar?
Torsdag d. 27. november 2025 kl. 19:00 til 20:30
Online via ZOOM
Aflyst

Fra 25 kr.

Læs mere

ZOOM - Cancerpakken v. Milena Penkova
In livsstil er vigtigere end din genetik i kampen mod cancer? Cancerpakken af specifik kosttilskud, som forskning har vist kan bekæmpe cancer, for alle, der har cancer eller har kræfttraum på råreværd.
d. 22. januar 2026 kl. 19:00 til 21:00
via ZOOM
tilmelding

Fra 25 kr.

Læs mere

elvhjælpskursus - Lær at bruge Qigong, irk og meditation i hverdagen
visklig medicin ser på hele mennesket - krop, sind og ånd. Ubalancer i riger kan skabe helbredsmæssige udfordringer. Du lærer enkle øvelser, der trænger til system og fremmer balance og velvære.

500 kr.

Læs mere

BT
Lena helbredt for kræft i Kina - pris: 1,2 millioner
49

MEST LESTE LIGE NU:

- 1. Lena Leonard - indleder behandling efter 10-årig behandling i Kina. Foto: Scanpix
- 2. Lena Leonard - indleder behandling efter 10-årig behandling i Kina. Foto: Scanpix
- 3. Lena Leonard - indleder behandling efter 10-årig behandling i Kina. Foto: Scanpix
- 4. Lena Leonard - indleder behandling efter 10-årig behandling i Kina. Foto: Scanpix
- 5. Lena Leonard - indleder behandling efter 10-årig behandling i Kina. Foto: Scanpix
- 6. Lena Leonard - indleder behandling efter 10-årig behandling i Kina. Foto: Scanpix
- 7. Lena Leonard - indleder behandling efter 10-årig behandling i Kina. Foto: Scanpix
- 8. Lena Leonard - indleder behandling efter 10-årig behandling i Kina. Foto: Scanpix
- 9. Lena Leonard - indleder behandling efter 10-årig behandling i Kina. Foto: Scanpix
- 10. Lena Leonard - indleder behandling efter 10-årig behandling i Kina. Foto: Scanpix

LIGE NU:

- 1. Lena Leonard - indleder behandling efter 10-årig behandling i Kina. Foto: Scanpix
- 2. Lena Leonard - indleder behandling efter 10-årig behandling i Kina. Foto: Scanpix
- 3. Lena Leonard - indleder behandling efter 10-årig behandling i Kina. Foto: Scanpix
- 4. Lena Leonard - indleder behandling efter 10-årig behandling i Kina. Foto: Scanpix
- 5. Lena Leonard - indleder behandling efter 10-årig behandling i Kina. Foto: Scanpix
- 6. Lena Leonard - indleder behandling efter 10-årig behandling i Kina. Foto: Scanpix
- 7. Lena Leonard - indleder behandling efter 10-årig behandling i Kina. Foto: Scanpix
- 8. Lena Leonard - indleder behandling efter 10-årig behandling i Kina. Foto: Scanpix
- 9. Lena Leonard - indleder behandling efter 10-årig behandling i Kina. Foto: Scanpix
- 10. Lena Leonard - indleder behandling efter 10-årig behandling i Kina. Foto: Scanpix

Lena Leonard mør solgte hus og bil, så den 22-årige datter kunne blive helbredt for kræft i Kina.
Svindede 1,2 millioner kroner. Så meget har Lena Leonard betalt for at blive behandleret for kræft i Kina.

ORTHOCARE
DIN ORTHOMOLEKYLÆRE SPECIALISTGEKLINK

Velkommen til Orthocare
Vi tilbyder helhedsorienteret og professionel hjælp, hvor hver patient modtager individuel tilpasset behandling

Kontakt os

Telefonløber
Mandag-Torsdag
12:00-14:00

Adresse
Slottsherrevej 215
2610 Rødovre

ORTHCARE
DIN ORTHOMOLEKYLÆRE SPECIALISTGEKLINK

Vores Behandlinger

- Redox - Hejdosis C-Vitamin
- EDTA
- Myers cocktail

BiologischeKrebsabwehr e.V.
menschlich - ganzheitlich - unabhängig
KONGRES Mai 2017

Weltkrebstag: Ja, Ich kann

Ansværlig lege Irene Camilla Hage

Trine Hegelund
Samarbejdspartner ved BiologischeKrebsabwehr e.V.

Naturvidenskabelig metode

- 
- Observation
 - Ide
 - Hypotese
 - Statistisk hypotese
 - Observation – systematisk
 - Statistisk hypotese test

Tese



Statistik skal hjælpe os med at finde ud af, om den **forskel vi observerer**, er **kausalt forbundet** med en **intervention/proces**, eller et resultat af **tilfældig variation**.



Opgave 1

- Læs overordnet de 2 retningslinjer
 - Farmakologisk behandling af kvalme med ukendt årsag
 - hos voksne patienter med dissemineret cancer i palliativt forløb
 - Opfølgning efter kurativ intenderet behandling for esophagus-/ventrikelcancer
- Brug 10 minutter på at læse dem overordnet
 - Diskuter følgende i gruppen (10 minutter)
 - Hvad – udover emnet – adskiller de to retningslinjer?
 - Vurdering af kvalitet/klinisk betydning?
 - Plenum diskussion 5 minutter (næste slide)

Diskussion

- Statistik skal hjælpe os med at finde ud af, om den forskel vi observere, er kausalt forbundet med en intervention/proces, eller et resultat af tilfældig variation.

Observation

Ide

Hypotese

Statistisk hypotese

Observation – systematisk

Statistisk hypotese test



Basal statistik

I sidder i uro-amb

I ser en mand med ondt i ryggen gennem 4 uger

- Hvad fejler han?

I sidder i almen praksis

I ser en mand med ondt i ryggen gennem 4 uger

- Hvad fejler han?

Pre-test sandsynligheder

I sidder i uro-amb

I ser en mand med ondt i ryggen gennem 4 uger

- Hvad fejler han?

I sidder i almen praksis

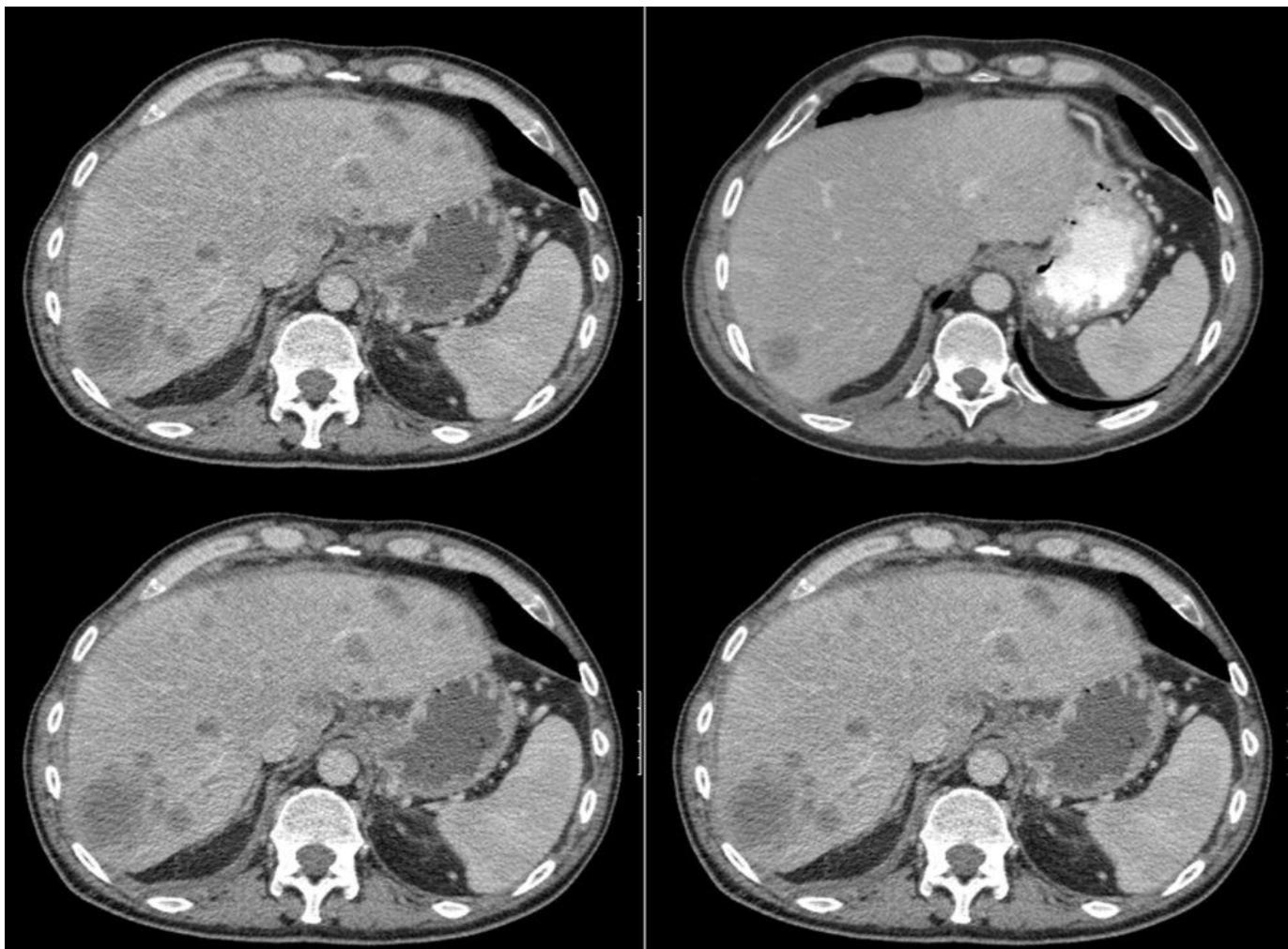
I ser en mand med ondt i ryggen gennem 4 uger

- Hvad fejler han?

Hvilken behandling har størst sandsynlighed for at påvirke overlevelsen?

3 måneder

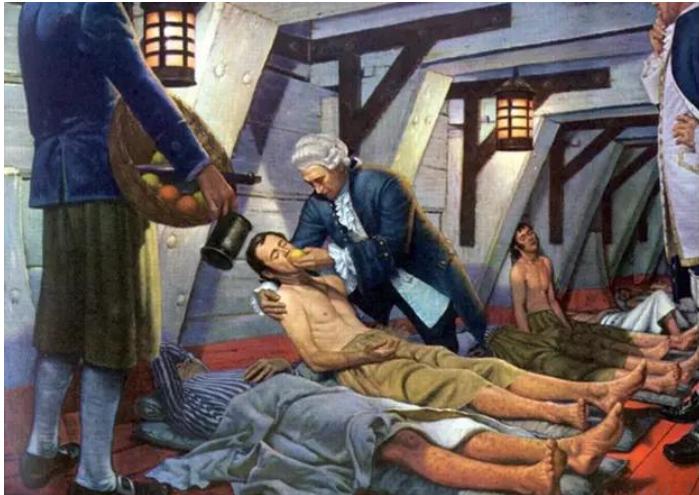
Behandling A



Behandling B

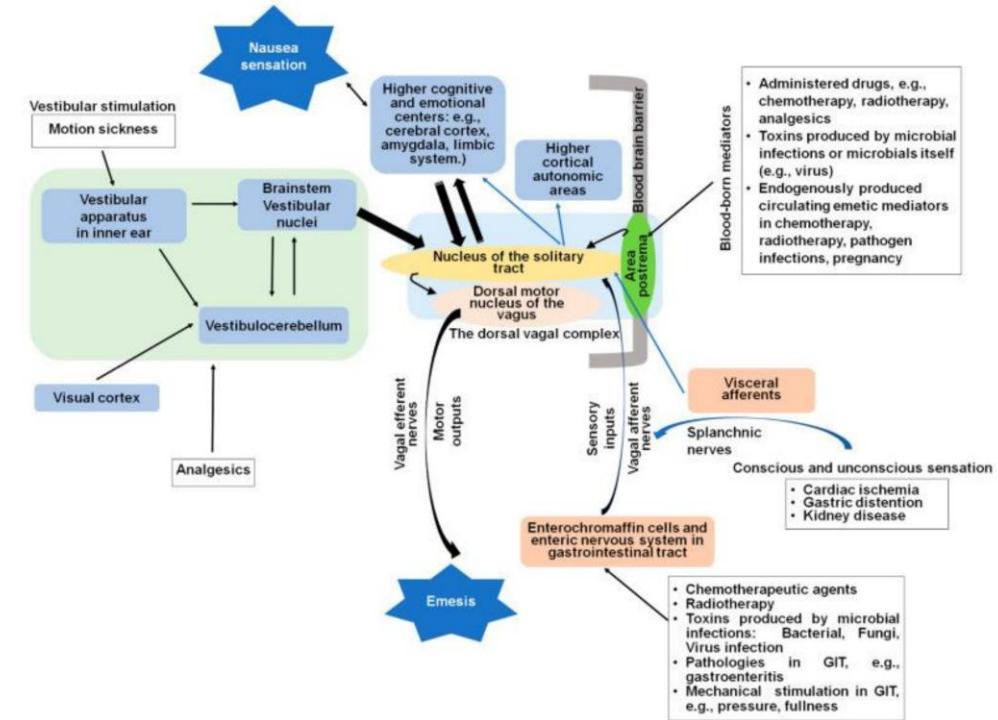
Pre-test sandsynligheder

- Et veletableret rationale for at en interventions kan have en effekt er centralt for statistiske tests værdi.



James Lind

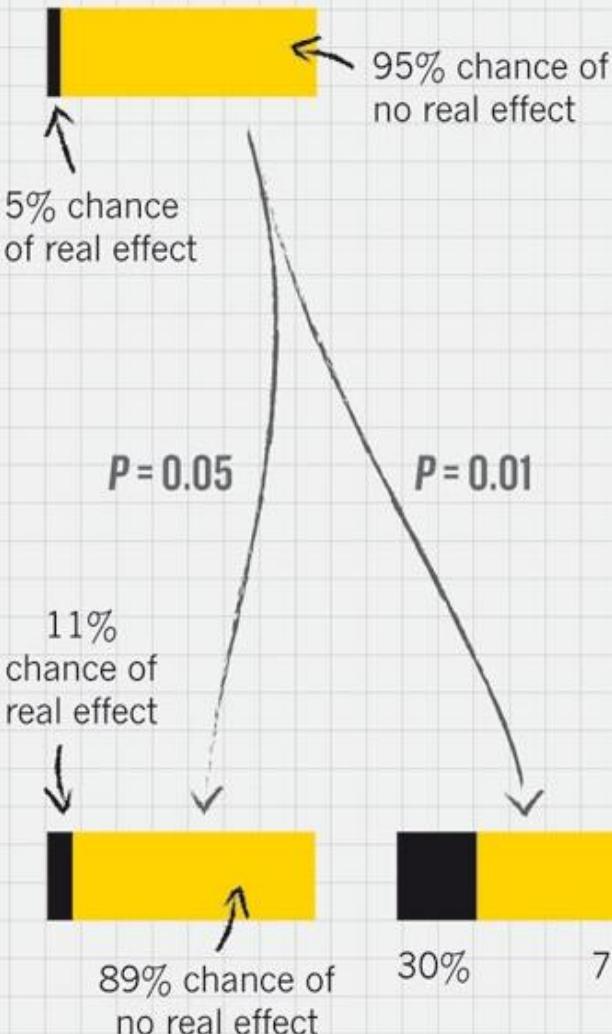
- a quart of cider a day
- 25 drops of elixir of vitriol, three times a day
- half a pint of sea-water a day
- a nutmeg-sized paste of garlic, mustard seed, horse-radish, balsam of Peru, and gum myrrh three times a day
- two spoonfuls of vinegar, three times a day
- two oranges and one lemon a day



Figur 1: Mechanisms of Nausea and Vomiting (7)

THE LONG SHOT

19-to-1 odds against



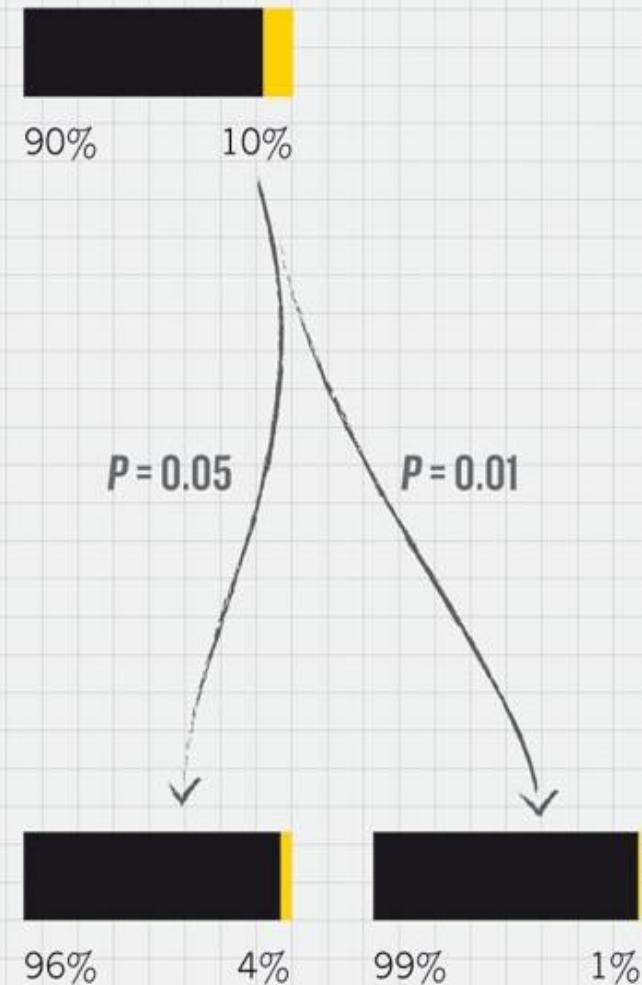
THE TOSS-UP

1-to-1 odds



THE GOOD BET

9-to-1 odds in favour



God nyhed for klinikeren! Forstå rationalet - ikke p-værdien

AMERICAN STATISTICAL ASSOCIATION RELEASES STATEMENT ON STATISTICAL SIGNIFICANCE AND P-VALUES

Provides Principles to Improve the Conduct and Interpretation of Quantitative Science

1. *P-values can indicate how incompatible the data are with a specified statistical model.*
2. *P-values do not measure the probability that the studied hypothesis is true, or the probability that the data were produced by random chance alone.*
3. *Scientific conclusions and business or policy decisions should not be based only on whether a p-value passes a specific threshold.*
4. *Proper inference requires full reporting and transparency.*
5. *A p-value, or statistical significance, does not measure the size of an effect or the importance of a result.*
6. *By itself, a p-value does not provide a good measure of evidence regarding a model or hypothesis.*

Take-home Forsknings metodologi 1

Ethvert forsøg skal baserer sig på et videnskabeligt solidt grundlag.

- Statistik skal hjælpe os med at finde ud af, om den forskel vi observerer, er kausalt forbundet med en intervention/proces, eller et resultat af tilfældig variation.



Projekt

Ide

Hypoteze

Beskrivelse af projekt

Dataindsamling

Dataanalyse

Hypotesetest

Præsentation

Statistik

- Median
- Gennemsnit
- Spredning
- SE
- SEM/SD
- Korrelation/sammenhæng
- Tid til hændelse modeller
- Generaliserede liniære modeller
 - Multivariat analyser

Centrale prinsipper

- Normal fordeling
- Stikprøve
- Variabel typer

Opgave 1 - Normalfordeling

- Normal fordeling er et centralt princip i biologisk forskning
 - Hvad er normal fordeling og hvorfor er det så smart?
 - Nogen der kan huske noget om det?

– FORKLAR DIN SIDEMAKKER HVAD NORMAL FORDELING ER OG HVORDAN DU VIL BESKRIVE DET/EN?

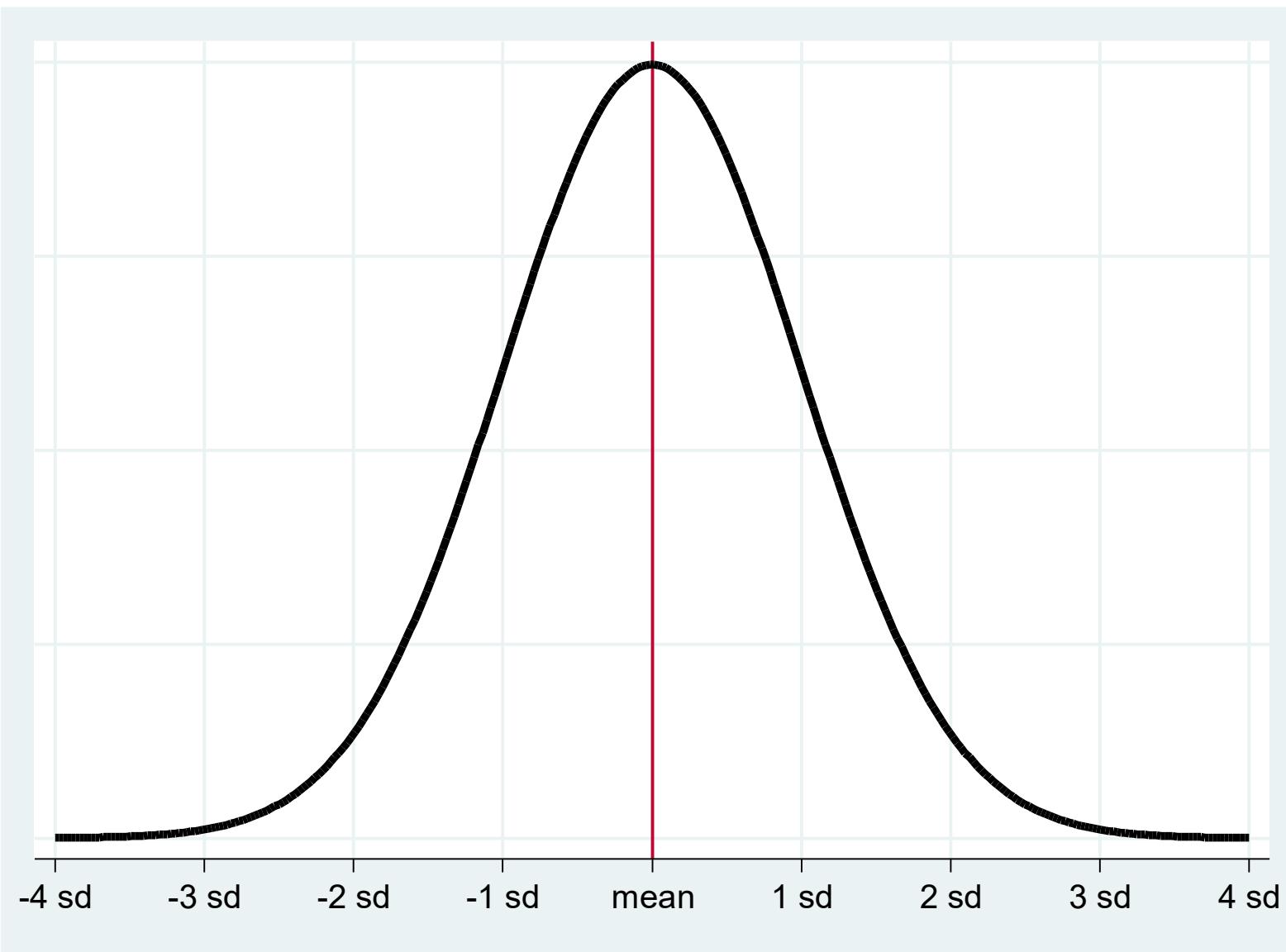
A photograph of a dense tropical forest. Sunlight filters down from the canopy of tall trees, creating bright highlights on the green leaves and ferns in the foreground. The ground is covered in a mix of dark soil and fallen brown leaves. The overall atmosphere is lush and serene.

Dyb indånding

Ingen Panik

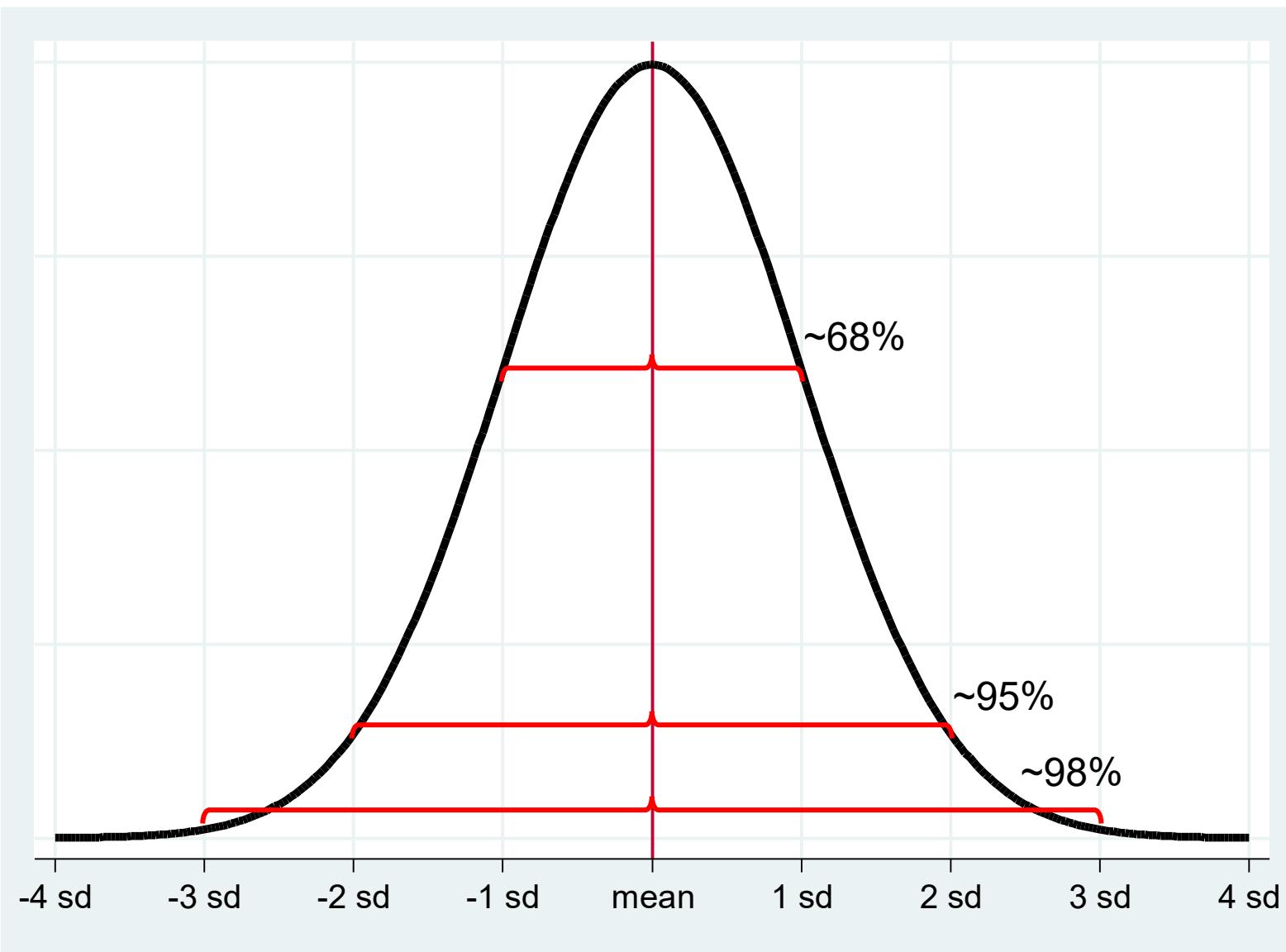
Normalfordeling

Kan beskrives med 2 værdier



Normalfordeling

Kan beskrives med 2 værdier

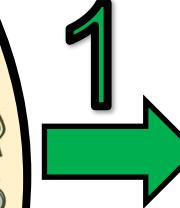
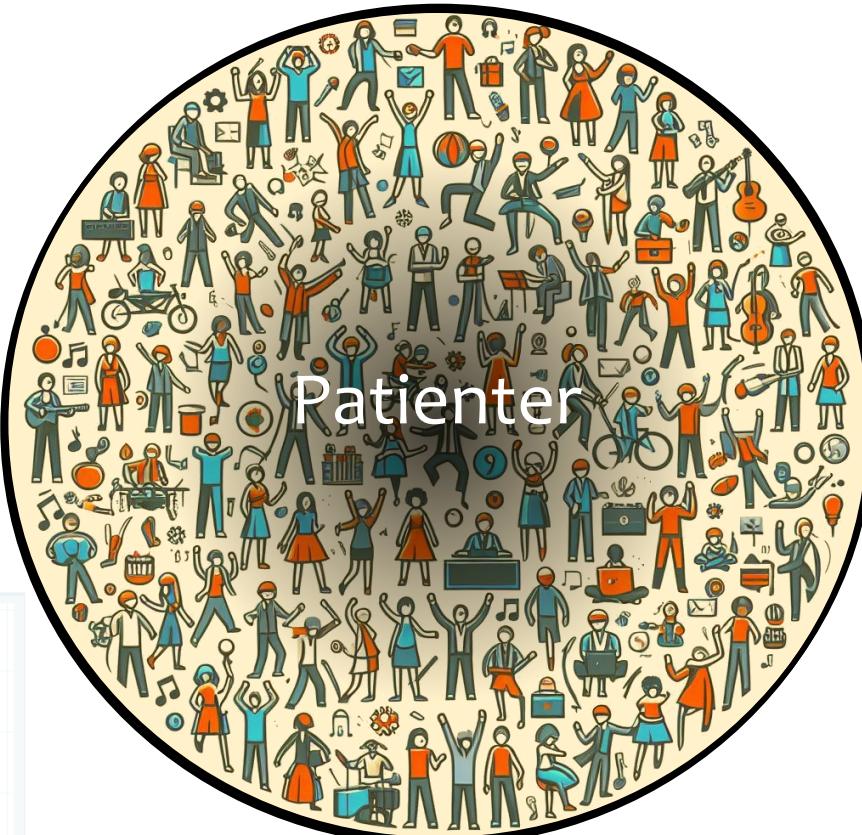
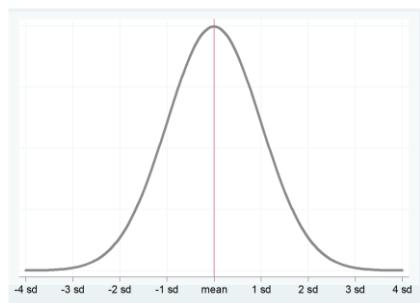


Centrale principper

- Normal fordeling ✓
- Stikprøve
- Variabel typer

Stikprøve

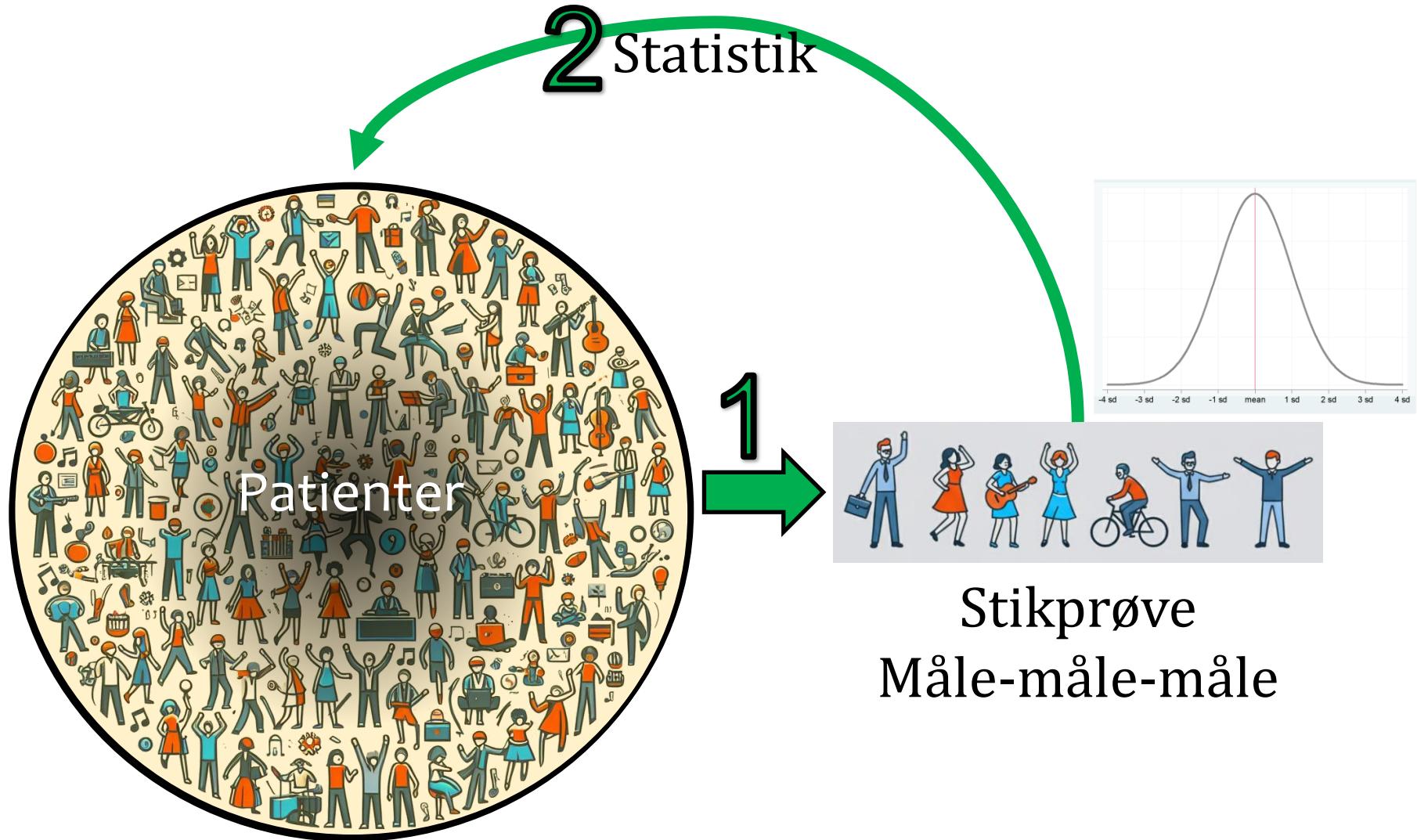
Anvendes til at generalisere til alle



Måle-måle-måle

Stikprøve

Anvendes til at generalisere til alle



Centrale principper

- Normal fordeling ✓
- Stikprøve ✓
- Variabel typer

Datatyper

- Kontinuerlige variabler
 - $1, 2, 3, \dots, \infty$
- Kategoriske variabler
 - Let, Moderat, Svær

Opgave

- Visuel Analog Skala (VAS)
 - Anvendes hyppigt til beskrivelse af smerte intensitet.
 - Går fra 1-10
 - 1 = ingen smerter
 - 10 = Værst tænkelige smerter

- Fortæl/Forklar din sidemakker hvilken variabel type VAS er?
 - 2 minutter

Centrale prinsipper

- Normal fordeling ✓
- Stikprøve ✓
- Variabel typer ✓

Stikprøve teori

Øvelse

- Gennemsnit højden for en dansker er:
 - 175 cm (*opfundet tal*)
 - Spørgsmål:
 - Er lokalet højere end gennemsnit danskere?
- 1
 - Beregn gennemsnit højden i dette lokale
 - Datasæt skal hentes
- 2
 - Er lokalet højere eller lavere end gennemsnittet i DK?
 - Se senere

AI/LLM baseret hjælp/læring

- Hvad er AI/LLM
 - <https://youtu.be/LPZh9BOjkQs>
The logo for 3Blue1Brown, featuring a stylized eye-like icon composed of blue and brown concentric circles, set against a dark background with a grid pattern.

3Blue1Brown
Animated math
 - Ekstremt komplekse sandsynlighedsmodeller
 - Hvis I har tid til overs så se videoen og de relaterede andre videoer



Praktiske eksempler -

Product	Model Family	Pricing & Free Tier	Privacy Warning ▲	Open/Proprietary
ChatGPT	GPT-4o / o1	Free: Unlimited (lower intelligence). Plus: \$20/mo (Top reasoning models).	Cloud: Trains on data by default. Enterprise/Team has privacy.	Proprietary (OpenAI)
DeepSeek	DeepSeek-V3 / R1	Free: 100% Free (Web). Very generous API. Paid: API is ~10x cheaper than US rivals.	Cloud: Chinese jurisdiction. Do not upload sensitive/patient data. High censorship on political topics.	Open Weights (Model is open; Service is closed)
Claude	Claude 3.5 / 3.7	Free: Low daily limit. Pro: \$20/mo (High reliability).	Cloud: High privacy focus. "Projects" feature isolates data.	Proprietary (Anthropic)
Gemini	Gemini 1.5 / 2.0	Free: Unlimited (Flash model). Adv: \$20/mo (Integrates w/ Google Docs).	Cloud: Google uses free-tier data for training.	Proprietary (Google)
Perplexity	GPT-4 / Claude / Sonar	Free: Basic search. Pro: \$20/mo (Select your preferred model).	Cloud: Best for research. Citations are real. Privacy varies by setting.	Proprietary Wrapper
Mistral (Le Chat)	Mistral Large / Pixtral	Free: Very generous free beta access. Paid: Enterprise tiers.	Cloud: EU-based (GDPR compliant). Good alternative to US tech.	Open Weights (mostly)
Grok	Grok 2 / 3	Free: Limited (X/Twitter users). Prem: ~\$16/mo (via X).	Cloud: Real-time access to X data. "Fun mode" is uncensored.	Proprietary (xAI)

Tool Name	Purpose	Pricing	Tech Skills Required
Claude Code	Agentic coding tool that lives in your terminal. Can edit files, run git commands, and refactor code autonomously.	Requires Anthropic API credits (Pay-as-you-go).	Medium: Command Line Interface (CLI) usage, understanding API keys.
Gemini CLI	Google's command-line interface for accessing Gemini models directly in the terminal (often via Google Cloud SDK).	Free tier available via Google AI Studio; then Pay-as-you-go.	Medium: Google Cloud Platform (GCP) basics, CLI navigation.

Name	Type	Pricing	Privacy	Self-Host?
Hugging Face	The Repository. Not a model itself, but the "GitHub of AI." You download models (Llama, Mistral) here.	Free to download models. Paid for hosted GPU computing.	High: If you download and run locally, no data leaves your PC.	Yes (Source of models)
Oobabooga	The Interface. A web UI tool to run the models you download from Hugging Face.	Free (Open Source software).	Maximum: Runs 100% offline. Zero data leakage.	Yes (This is the host)
Ollama	The Interface. A web UI tool to run the models you download from Hugging Face	Free	Runs locally	Self-host

Tips og tricks

- **Fortæl hvad du laver og kontekst**

- Du er nu min personlige statistik-tutor. Jeg er læge og deltager i et introduktionskursus i statistik (bruger [INDSÆT PROGRAM: R/SPSS/Stata]).
- **Regler:**
 - Giv mig ikke bare den færdige løsning med det samme. Hjælp mig med at forstå logikken.
 - Når du forklarer begreber (f.eks. p-værdi eller konfidensintervaller), så brug analogier fra den medicinske verden (f.eks. kliniske studier eller diagnostik).
 - Hvis jeg viser dig kode med fejl, så forklar fejlen pædagogisk, før du retter den.
 - Hold svarene korte og præcise."

Tips og tricks

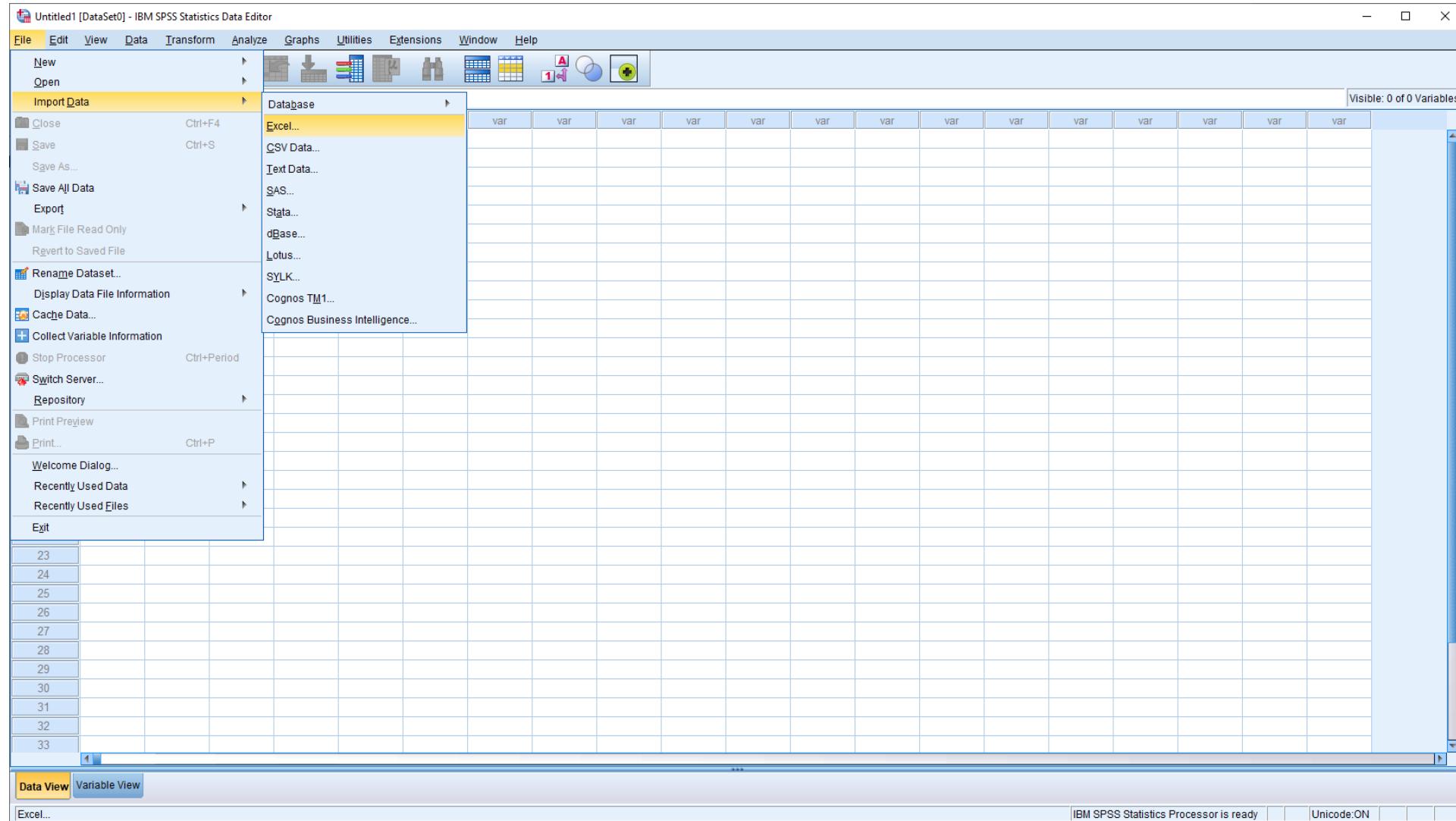
- **Fejlfinding**
- **Scenarie A: Du har en fejlmeldelse**
 - Jeg får denne fejlmeldelse: **[INDSÆT FEJLMEDDEELSE]**. Her er den kode/kommando jeg brugte: **[INDSÆT KODE]**. Forklar på simpelt dansk, hvad jeg har gjort galt, og hvordan jeg retter det.”
 - Du kan også copy/paste koden ind – alternativt lave screen-shot.
- **Scenarie B: Data ser forkert ud (Format-hjælp)**
 - Min datafil driller. Jeg prøver at importere en Excel-fil, men mine variable bliver indlæst som tekst (characters) i stedet for tal (numeric). Hvilken kommando skal jeg bruge for at fikse det? Jeg bruger **[R/SPSS/STATA]**.

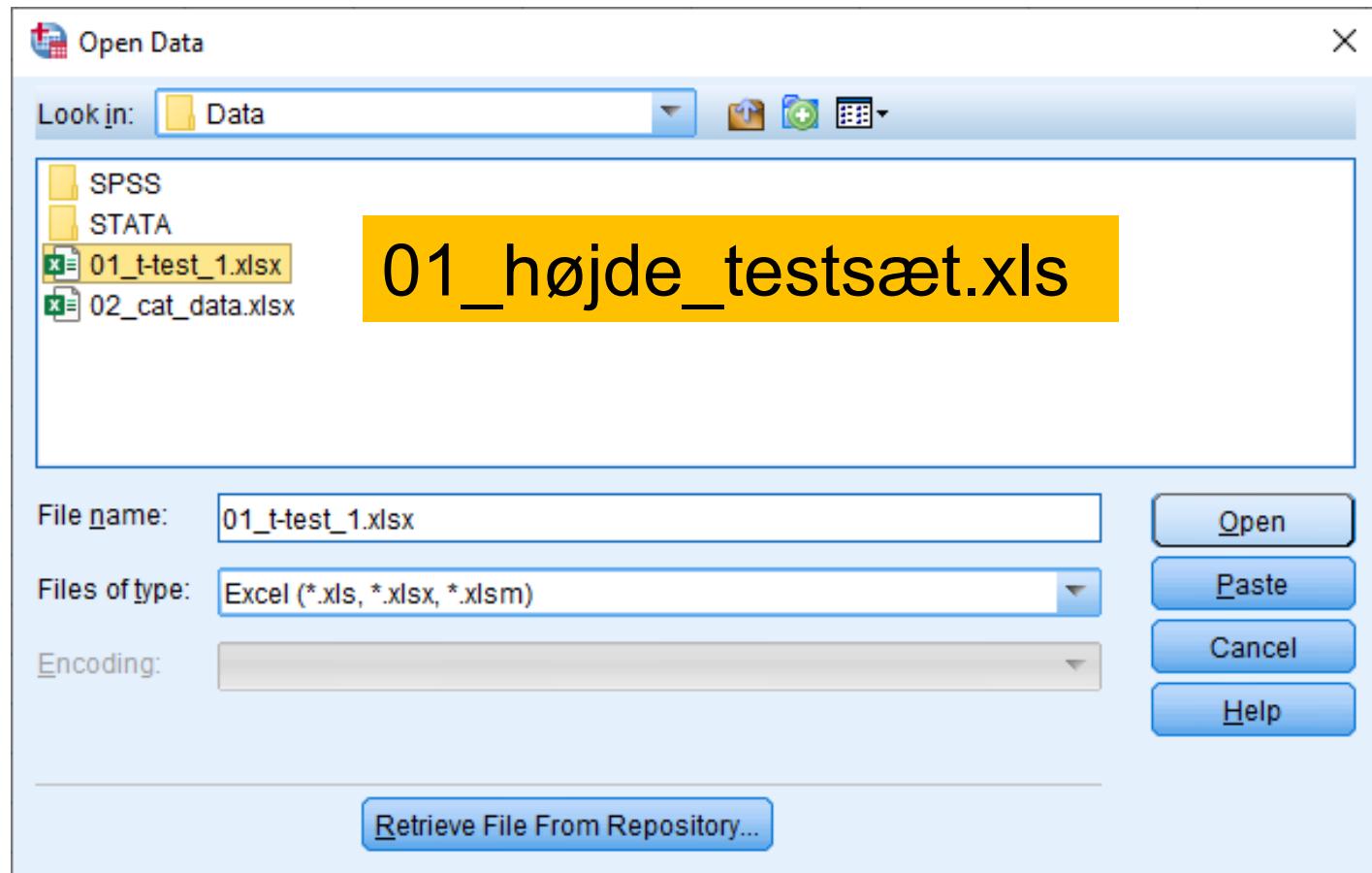
Tips og tricks

- **Forståelse af Output:**
 - Jeg har lavet en **[TEST-TYPE, f.eks. T-test]** og fået dette output (se nedenfor). Kan du hjælpe mig med at tolke resultatet? Jeg er især i tvivl om, hvordan jeg skal læse 'Estimate' og 'Pr(>|t|)' i en klinisk sammenhæng. Er resultatet signifikant?"
[Indsæt kopi af resultattabel her]
- **Begrebsafklaring**
 - Jeg har glemt forskellen på Standard Deviation (SD) og Standard Error (SE). Kan du forklare forskellen meget kort, som om jeg skal forklare det til en kollega? Giv mig et eksempel med blodtryksmålinger.
- **HUSK AT TÆNKE DIG OM!!!!!! OG SPECIELT** om svaret er meningsfyldt. Hvis du ikke forstår det..så lad være med at bruge det.

Øvelse

Statistik-program





 Read Excel File X

C:\Users\JBJE0062\OneDrive - Region Hovedstaden\Public\A_ &_ P\Forsk_træning\Forskningstræning\Forskningstræning 2021\Forskningstræning_2021_Kursist-mappe\Statistik&Forskningsmetodologi\Data\01_t-test_1.xlsx

Worksheet:

Range:

Read variable names from first row of data

Percentage of values that determine data type:

Ignore hidden rows and columns

Remove leading spaces from string values

Remove trailing spaces from string values

Preview

	ID	Height	Gender	Weight	Blond
1	1	169	0	58	0
2	2	171	0	62	1
3	3	167	0	71	0
4	4	170	0	67	1
5	5	177	0	78	0
6	6	164	0	69	1
7	7	165	0	66	0

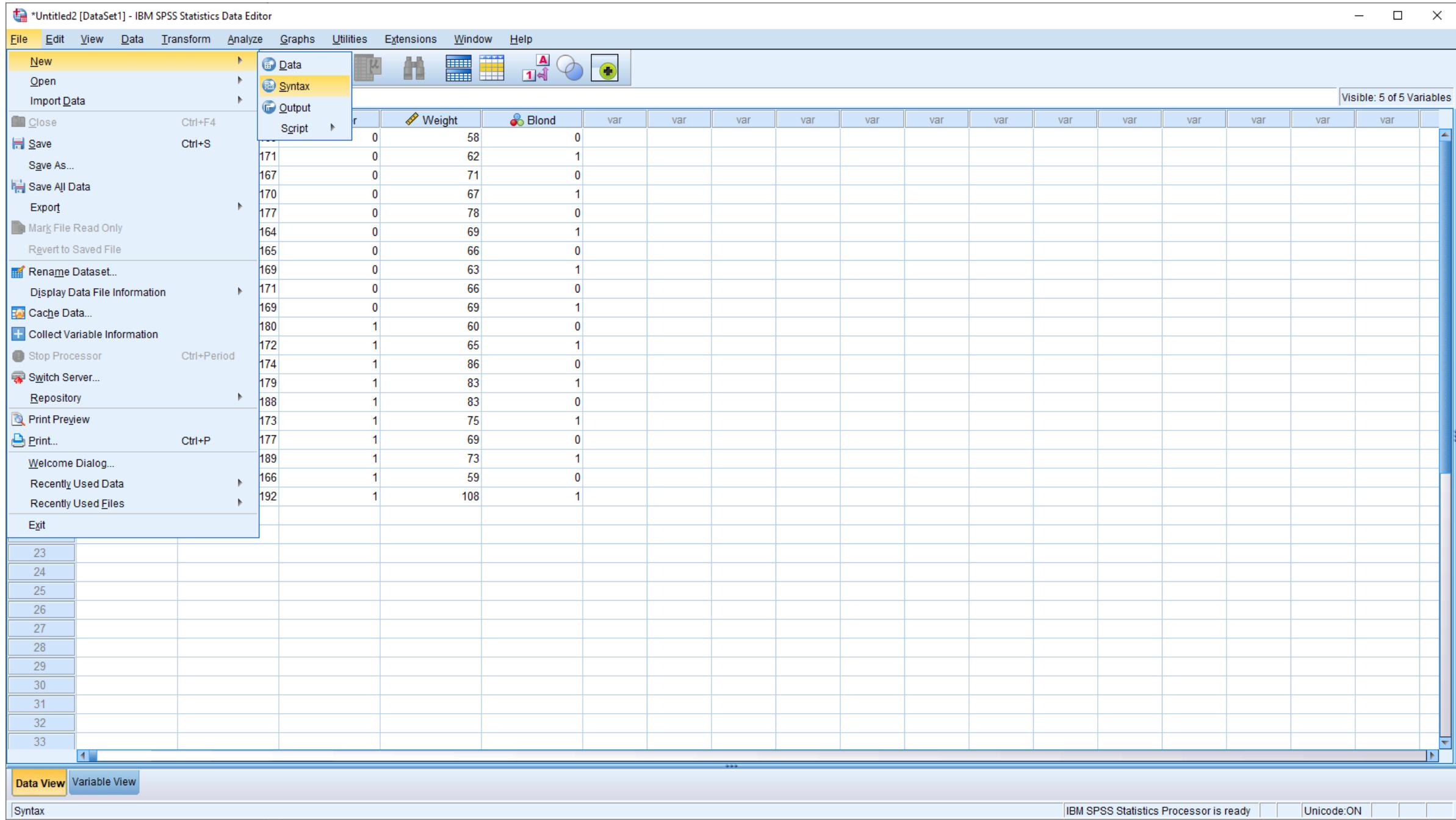
 Final data type is based on all data and can be different from the preview, which is based on the first 200 data rows. The preview displays only the first 500 columns.

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

Visible: 5 of 5 Variables

	ID	Height	Gender	Weight	Blond	var															
1	1	169	0	58	0																
2	2	171	0	62	1																
3	3	167	0	71	0																
4	4	170	0	67	1																
5	5	177	0	78	0																
6	6	164	0	69	1																
7	7	165	0	66	0																
8	8	169	0	63	1																
9	9	171	0	66	0																
10	10	169	0	69	1																
11	11	180	1	60	0																
12	12	172	1	65	1																
13	13	174	1	86	0																
14	14	179	1	83	1																
15	15	188	1	83	0																
16	16	173	1	75	1																
17	17	177	1	69	0																
18	18	189	1	73	1																
19	19	166	1	59	0																
20	20	192	1	108	1																
21																					
22																					
23																					
24																					
25																					
26																					
27																					
28																					
29																					
30																					
31																					
32																					
33																					

Data View Variable View



*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

Data View Variable View

IBM SPSS Statistics Processor is ready | Unicode:ON

Syntax1 - IBM SPSS Statistics Syntax Editor

File Edit View Data Transform Analyze Graphs Utilities Run Tools Extensions Window Help

Active: DataSet1

*Output1 [Document1] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

Output Log

```
GET DATA  
  /TYPE=XLSX  
  /FILE='C:\Users\JBJE0062\OneDrive - Region Hovedstaden\Public\A & P\Forsk_træning\Forskningstræning 2021\Forskningstræning_2021_Kursist-mappe\Statistik  
&Forskningsmetodologi\Data\01_t-test_1.xlsx'  
  /SHEET=name 'Sheet1'  
  /CELLRANGE=FULL  
  /READNAMES=ON  
  /DATATYPEMIN PERCENTAGE=95.0  
  /HIDDEN IGNORE=YES.  
EXECUTE.  
DATASET NAME DataSet1 WINDOW=FRONT.
```

IBM SPSS Statistics Processor is ready | Unicode:ON

Indbakke - jon.kroll.. Screenshots in Win... F_&_U Data Program.docx - W... 2021-12-02-FSK-Sta... *Output1 [Docume... *Untitled2 [DataSet... Syntax1 - IBM SPSS... 17:32 29-11-2021

File Edit View Data Transform Insert Format

Analyze Graphs Utilities Extensions Window Help

Reports

Descriptive Statistics

Bayesian Statistics

Tables

Compare Means

General Linear Model

Generalized Linear Models

Mixed Models

Correlate

Regression

Loglinear

Neural Networks

Classify

Dimension Reduction

Scale

Nonparametric Tests

Forecasting

Survival

Multiple Response

Missing Value Analysis...

Multiple Imputation

Complex Samples

Simulation...

Quality Control

ROC Curve...

Spatial and Temporal Modeling...

→ Frequencies

[DataSet1]

Statistics

Height

	N	Valid	Missing
Valid	164	1	5,0

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	164	1	5,0	5,0
	165	1	5,0	10,0
	166	1	5,0	15,0
	167	1	5,0	20,0
	169	3	15,0	35,0
	170	1	5,0	40,0
	171	2	10,0	50,0
	172	1	5,0	55,0
	173	1	5,0	60,0
	174	1	5,0	65,0
	177	2	10,0	75,0

Frequencies...

Descriptives...

Explore...

Crosstabs...

TURF Analysis

Ratio...

P-P Plots...

Q-Q Plots...



```
/HIDDEN IGNORE=YES.  
EXECUTE.  
DATASET NAME DataSet1 WINDOW=FRONT.  
FREQUENCIES VARIABLES=Height  
/ORDER=ANALYSIS.
```

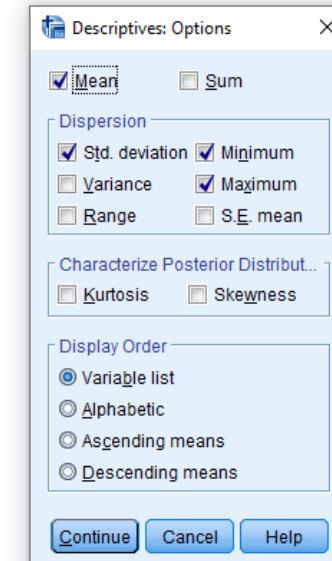
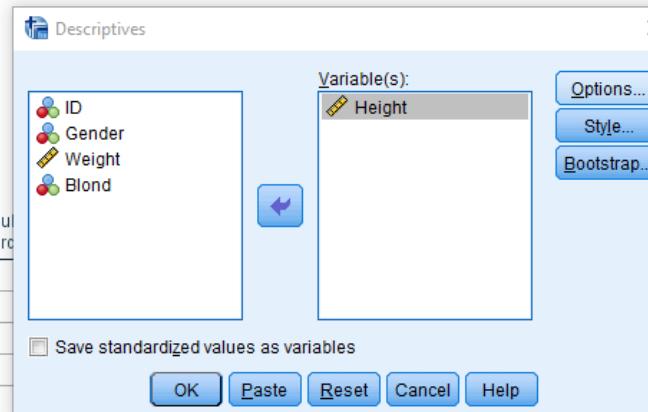
→ Frequencies

[DataSet1]

Statistics

Height	N	Valid	20
	Missing		0

Height				
	Frequency	Percent	Valid Percent	Cumul. Perc.
Valid	164	1	5,0	5,0
	165	1	5,0	5,0
	166	1	5,0	5,0
	167	1	5,0	5,0
	169	3	15,0	15,0
	170	1	5,0	5,0
	171	2	10,0	10,0
	172	1	5,0	5,0
	173	1	5,0	5,0
	174	1	5,0	5,0
	177	2	10,0	10,0
	179	1	5,0	5,0
	180	1	5,0	5,0
	188	1	5,0	5,0
	189	1	5,0	5,0
	192	1	5,0	5,0
Total		20	100,0	100,0





Output

- Log
- Frequencies
 - Title
 - Notes
- Active Dataset
- Log
- Descriptives
 - Title
 - Notes
 - Descriptive Statist

```
GET DATA
/TYPE=XLSX
/FILE='C:\Users\JBJE0062\OneDrive - Region Hovedstaden\Public\A_&_P\Forsk_træning\Forskningstræning\Forskningstræning 2021\Forskningstræning_2021_Kursist-mappe\Statistik &Forskningsmetodologi\Data\01_t-test_1.xlsx'
/SHEET=name 'Sheet1'
/CELLRANGE=FULL
/READNAMES=ON
/DATATYPEMIN PERCENTAGE=95.0
/HIDDEN IGNORE=YES.
EXECUTE.
DATASET NAME DataSet1 WINDOW=FRONT.
FREQUENCIES VARIABLES=Height
/ORDER=ANALYSIS.
```

Frequencies

→ [DataSet1]

```
DESCRIPTIVES VARIABLES=Height
/STATISTICS=MEAN STDDEV MIN MAX.
```

Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Height	20	164	192	174,10	8,052
Valid N (listwise)	20				

A photograph of a dense forest. Sunlight filters through the tall, thin trunks of the trees, creating bright rays of light that illuminate patches of mossy rocks and low-lying plants on the forest floor. The overall atmosphere is serene and natural.

Dyb indånding
Ingen Panik

Variation

- Variation i stikprøven

$$- sd = \sqrt{\frac{(obs\ 1-gennemsnit)^2 + (obs\ 2-gennemsnit).....}{antal\ observationer - 1}}$$

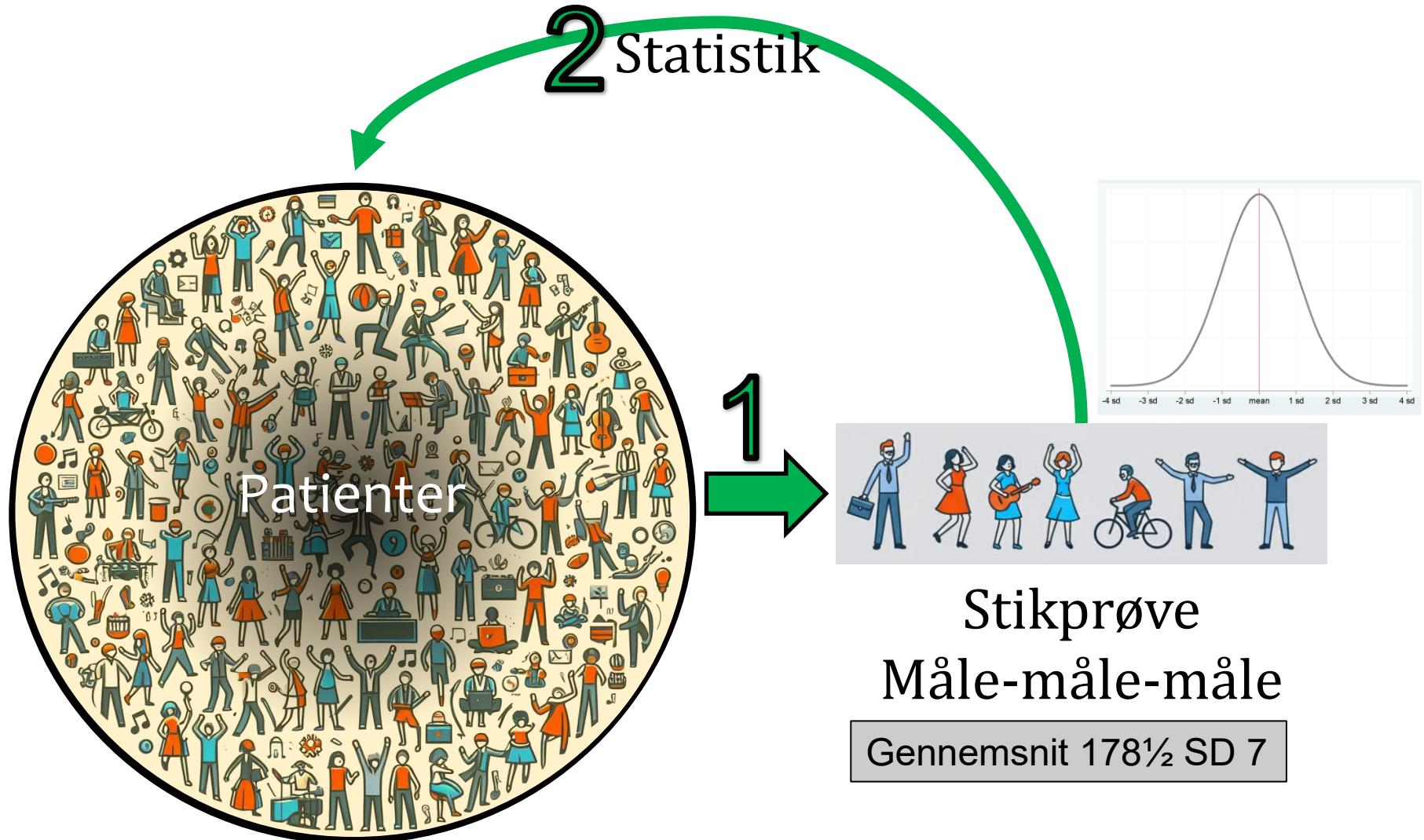


Øvelse

- Gennemsnit højden for en dansker er:
 - 175 cm

Stikprøve

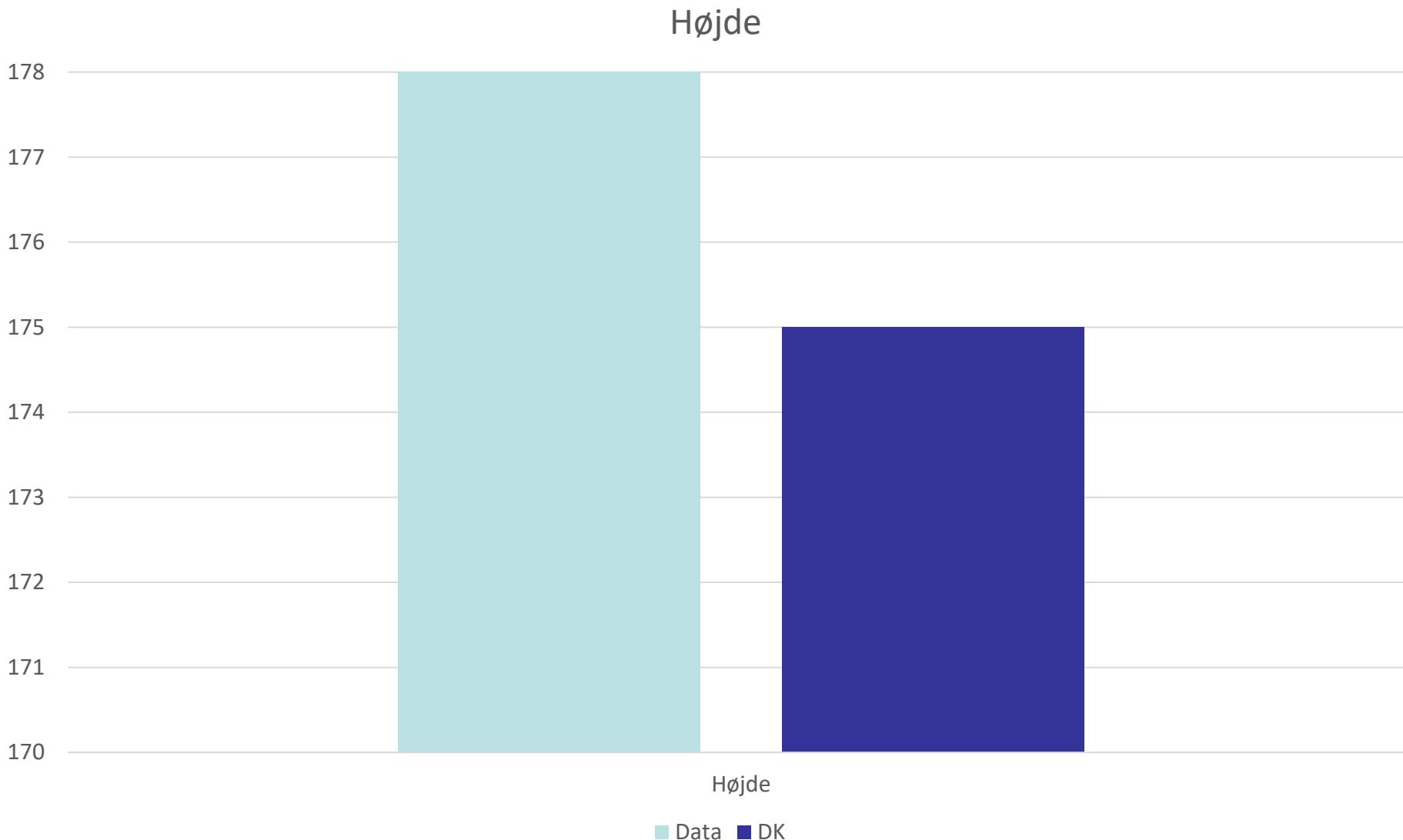
Anvendes til at generalisere til alle



Normalfordeling og stikprøver

- Hele gruppen af patienter har et "ægte"
 - gennemsnit
 - variation
- Vores stikprøve anvendes som bedste bud på dette
 - Jo større stikprøve – des bedre bud

Er data højere/lavere end gennemsnit danskere? Hvad er svaret? - 2 minutter



Klinisk sammenhæng

- Vi ønsker ofte at få en ide om vores måling er tæt på den "ægte"



?



Klinisk problem

Forklar din sidemakker hvad et konfidens interval er og hvordan du bruger det klinisk?



Konfidens interval

- Et interval hvor der er sandsynligt at den "ægte" værdi befinder sig.
- Spørgsmål
 - Hvor præcist har vi målt højden?

Standard error of the mean

- Et mål for præcisionen

$$SEM = \frac{SD}{\sqrt{n}}$$



Konsekvenser

- Variation i måling
 - Stor
 - SEM højere
 - Lille
 - SEM mindre
- Antal patienter
 - Stort
 - SEM falder
 - Lille
 - SEM stiger

Konfidens interval

$$SEM = \frac{SD}{\sqrt{n}}$$

$$95\% CI = gennemsnit \pm 1,96 \times SEM$$

$$95\% CI = gennemsnit \pm 1,96 \times \frac{SD}{\sqrt{n}}$$

Konfidens interval

- 194,176,182,168,169,175,183; Gennemsnit =178,3 ; SD=9,04 ; n=7

$$95\% CI = \text{gennemsnit} \pm 1,96 \times \frac{SD}{\sqrt{n}}$$

$$95\% CI = 178,3 \pm 1,96 \times \frac{9,04}{\sqrt{7}}$$

$$95\% CI = 178,3 \pm 6,7 (172 - 185)$$



Output

- Log
- Frequencies
 - Title
 - Notes
- Active Dataset
- Log
- Descriptives
 - Title
 - Notes
 - Descriptive Statist...

```
GET DATA
  /TYPE=XLSX
  /FILE='C:\Users\JBJE0062\OneDrive - Region Hovedstaden\Public\A & P\Forsk_træning\Forskningstræning 2021\Forskningstræning_2021_Kursist-mappe\Statistik
&Forskningsmetodologi\Dataset\01_t-test_1.xlsx'
  /SHEET=name 'Sheet1'
  /CELLRANGE=FULL
  /READNAMES=ON
  /DATATYPEMIN PERCENTAGE=95.0
  /HIDDEN IGNORE=YES.

EXECUTE.

DATASET NAME DataSet1 WINDOW=FRONT.
FREQUENCIES VARIABLES=Height
  /ORDER=ANALYSIS.
```

Frequencies

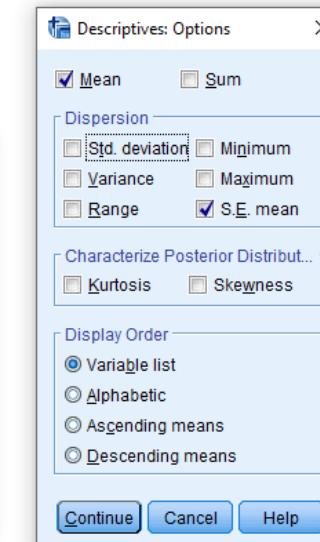
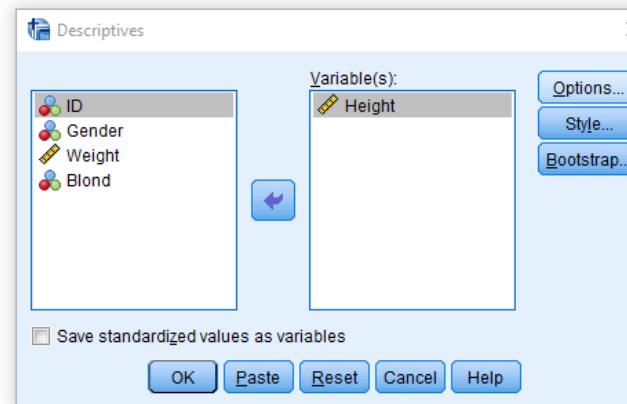
→ [DataSet1]

```
DESCRIPTIVES VARIABLES=Height
  /STATISTICS=MEAN STDDEV MIN MAX.
```

Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Height	20	164	192	174,10	8,052
Valid N (listwise)	20				



Øvelse

*Output1 [Document1] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

Reports Descriptive Statistics Bayesian Statistics Tables Compare Means General Linear Model Generalized Linear Models Mixed Models Correlate Regression Loglinear Neural Networks Classify Dimension Reduction Scale Nonparametric Tests Forecasting Survival Multiple Response Missing Value Analysis... Multiple Imputation Complex Samples Simulation... Quality Control ROC Curve... Spatial and Temporal Modeling...

One-Sample T Test...

edstaden\Public\A_&_P\Forsk_træning\Forskningsstræning\Forskningsstræning 2021\Forskningsstræning_2021_Kursist-mappe\Statistik

EXECUTE.
DATASET NAME Dat
FREQUENCIES VARI
/ORDER=ANALYSIS.

Frequencies

[DataSet1]

DESCRIPTIVES VAR
/STATISTICS=MEAN SEMEAN.

Descriptives

Height
Valid N (listwise)

Std. Deviation
8,052

DESCRIPTIVES VARIABLES=Height
/STATISTICS=MEAN SEMEAN.

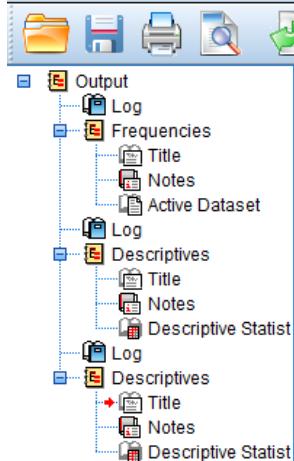
→ **Descriptives**

Descriptive Statistics

	N	Mean	
	Statistic	Statistic	Std. Error
Height	20	174,10	1,800
Valid N (listwise)	20		

One-Sample T Test...

IBM SPSS Statistics Processor is ready | Unicode:ON | H: 105, W: 190 pt.



```

(FILE='C:\Users\JBJE0062\OneDrive - Region Hovedstaden\Public\A_&_P\Forsk_træning\Forskningsstræning\Forskningsstræning 2021\Forskningsstræning_2021_Kursist-mappe\Statistik
&Forskningsmetodologi\Data\01_t-test_1.xlsx'
/SHEET=name 'Sheet1'
/CELLRANGE=FULL
/READNAMES=ON
/DATATYPEMIN PERCENTAGE=95.0
/HIDDEN IGNORE=YES.

EXECUTE.

DATASET NAME DataSet1 WINDOW=FRONT.
FREQUENCIES VARIABLES=Height
/ORDER=ANALYSIS.

```

Frequencies

[DataSet1]

DESCRIPTIVES VARIABLES=Height
/STATISTICS=MEAN STDDEV MIN MAX.

Descriptives

Descriptive Statistics

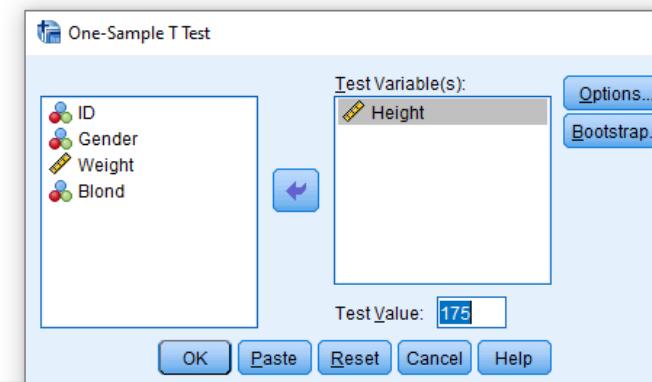
	N	Minimum	Maximum	Mean	
Height	20	164	192	174,10	8,052
Valid N (listwise)	20				

DESCRIPTIVES VARIABLES=Height
/STATISTICS=MEAN SEMEAN.

Descriptives

Descriptive Statistics

	N	Mean	
	Statistic	Statistic	Std. Error
Height	20	174,10	1,800
Valid N (listwise)	20		



Forklar din sidemakker hvad du ser

*Output1 [Document1] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

Output

Log

Frequencies

Descriptives

T-Test

One-Sample Test

Height

Valid N (listwise)

N Minimum Maximum Mean Std. Deviation

Height	20	164	192	174,10	8,052
Valid N (listwise)	20				

DESCRIPTIVES VARIABLES=Height
/STATISTICS=MEAN SEMEAN.

Descriptives

Descriptive Statistics

	N	Mean	
	Statistic	Statistic	Std. Error
Height	20	174,10	1,800
Valid N (listwise)	20		

T-TEST
/TESTVAL=175
/MISSING=ANALYSIS
/VARIABLES=Height
/CRITERIA=CI (.95).

T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Height	20	174,10	8,052	1,800

One-Sample Test

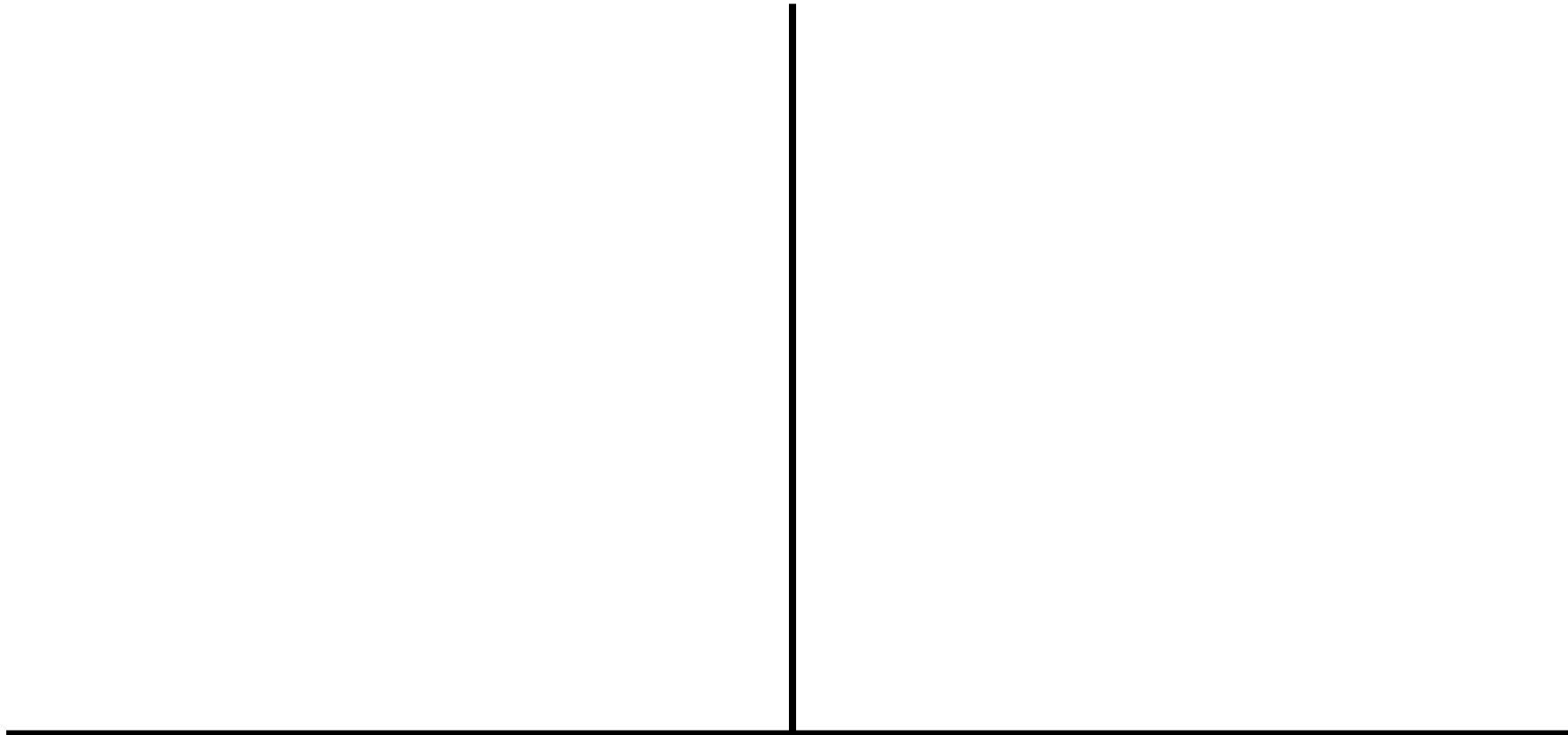
Test Value = 175

	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
				Lower	Upper	
Height	-,500	19	,623	-,900	-4,67	2,87

IBM SPSS Statistics Processor is ready Unicode:ON H: 104, W: 1421 pt.

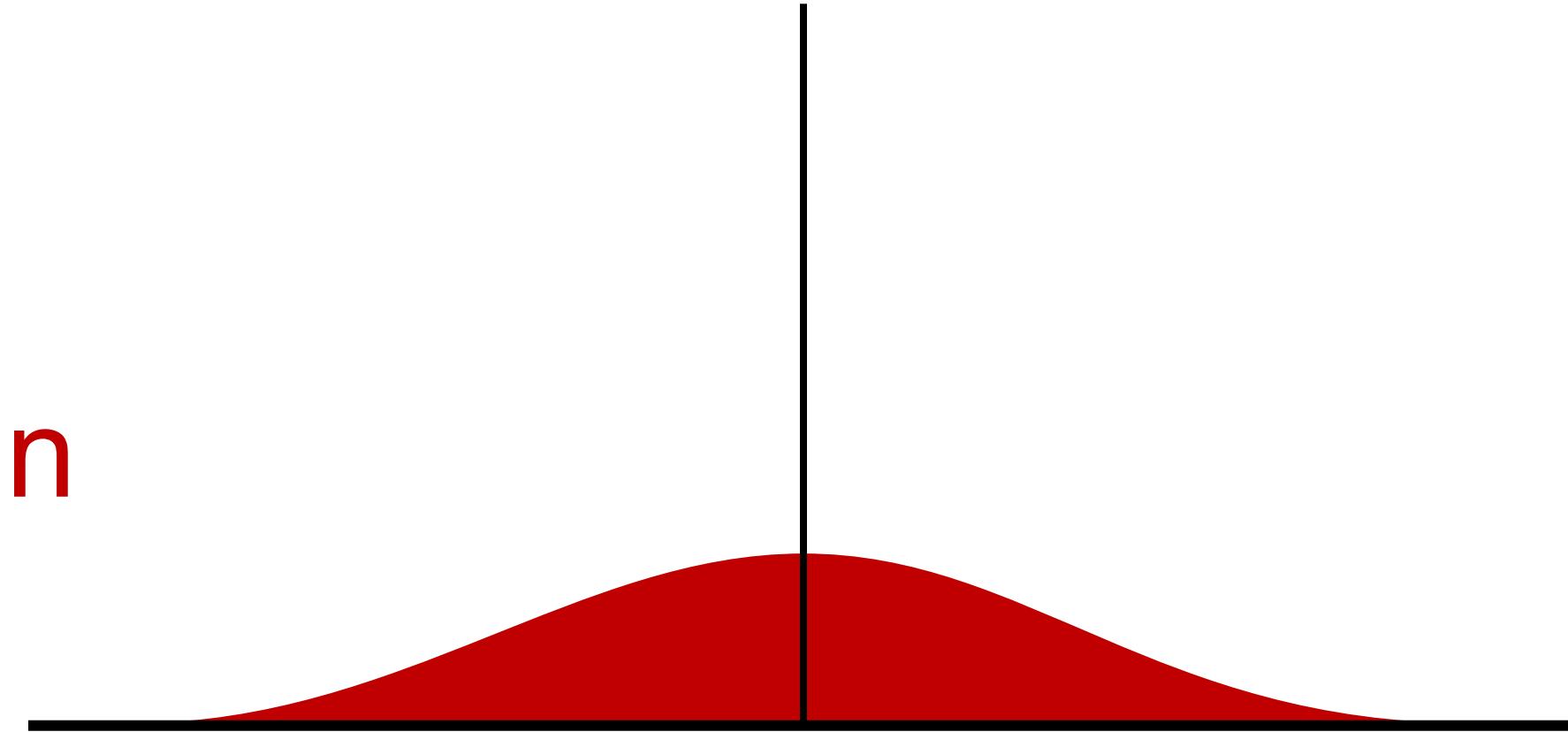
Stikprøve teori

- Jo større stikprøve
 - Jo større sandsynlighed for at den passer



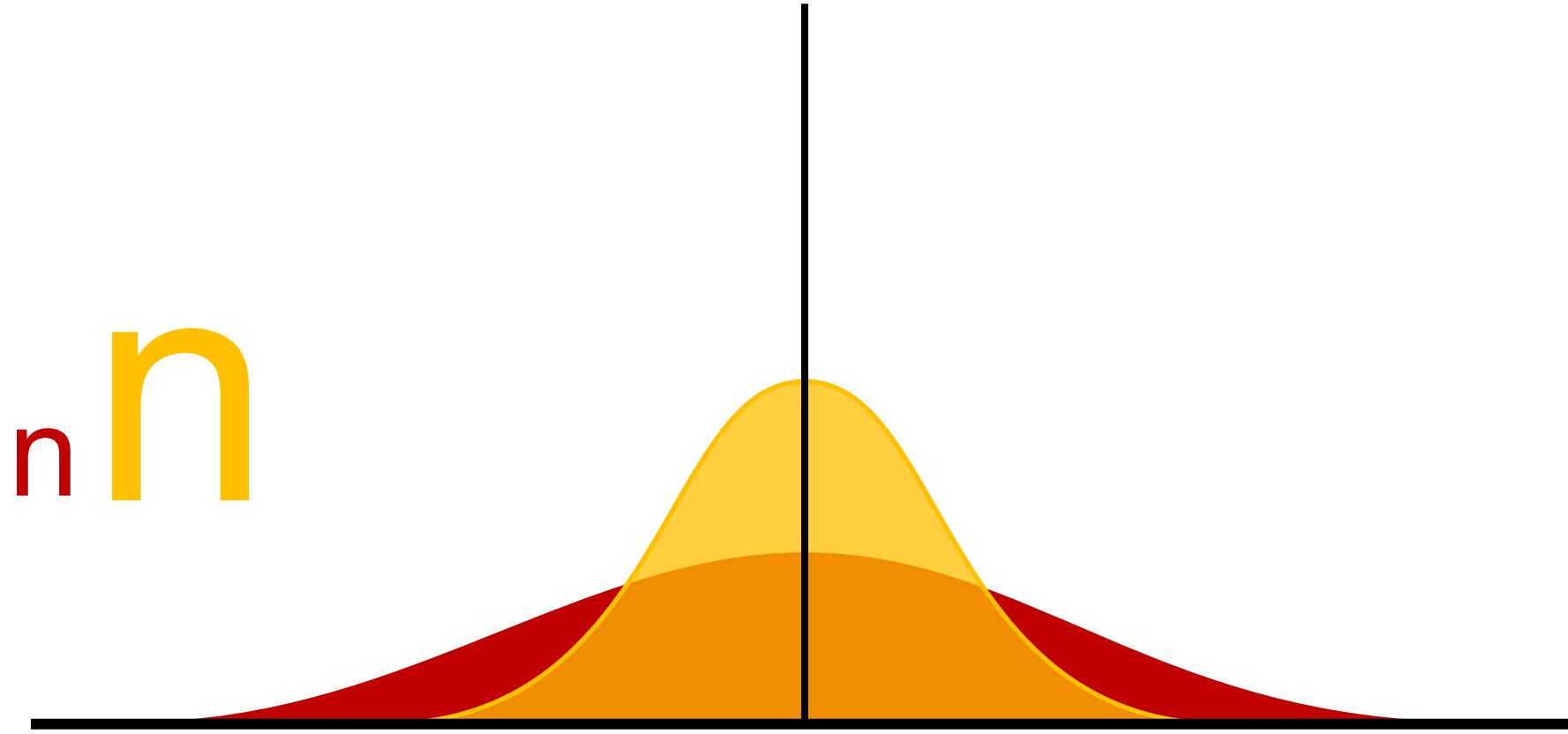
Stikprøve teori

- Jo større stikprøve
 - Jo større sandsynlighed for at den passer



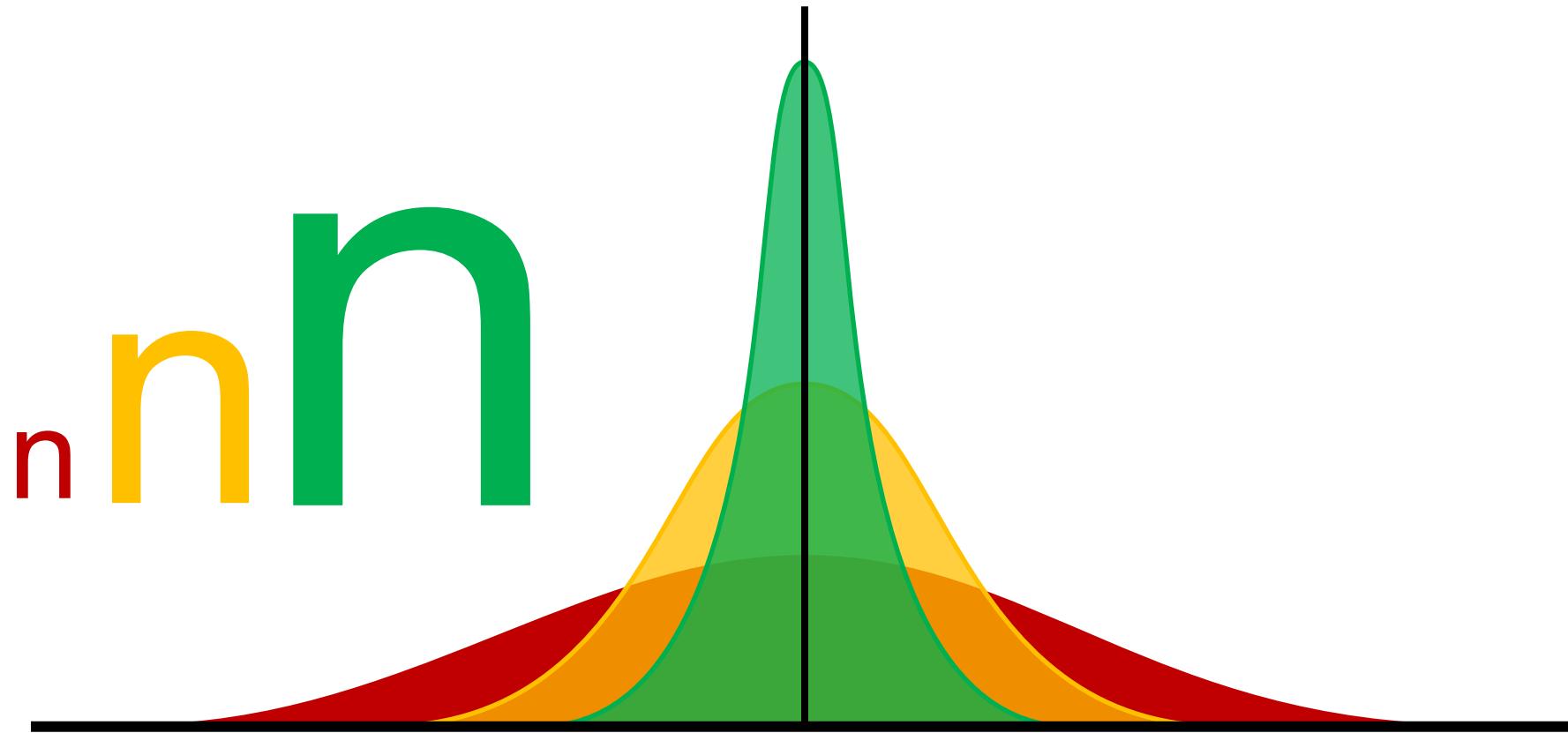
Stikprøve teori

- Jo større stikprøve
 - Jo større sandsynlighed for at den passer



Stikprøve teori

- Jo større stikprøve
 - Jo større sandsynlighed for at den passer



Øvelse

- 95% CI
 - Hvorfor ikke 50%?
- Hvilke problemer kan der være i tolkningen af konfidens intervaller?

Variation - hvor går det galt

- Stikprøve
 - Sted
 - Børnehave vs. plejehjem vs. lokalet
 - Andre
 - Køn
 - Race etc.
- Man kan tænke om variationen i stikprøven kan forklares fra andre faktorer.
 - Alternativt kan man kalde det "støj" i målingen
 - Mere om dette senere

Take-home

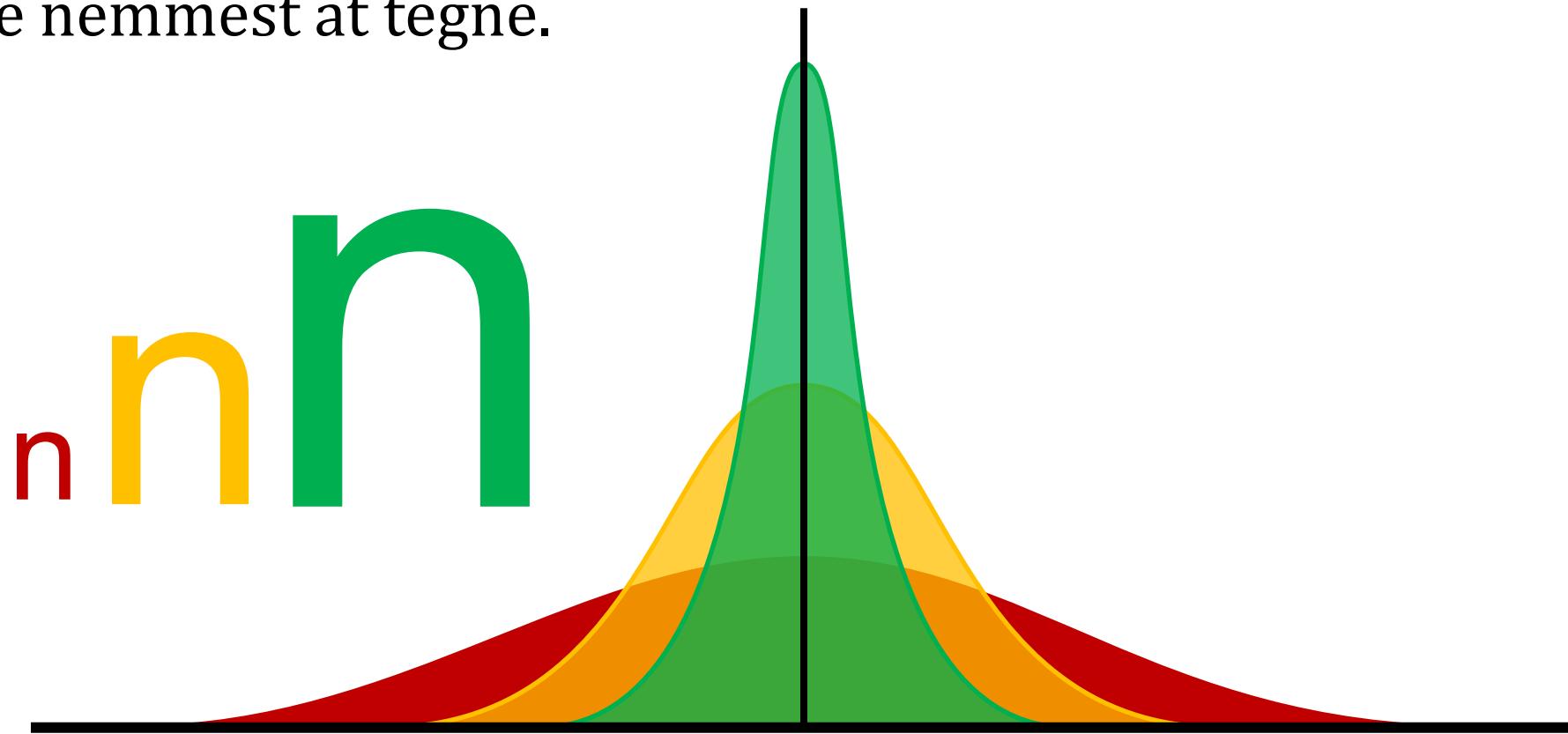
- Normafordeling er et centralt princip
- Fordelingen kan beskrives udfra
 - Gennemsnit og spredning(SD)
- En stikprøve bruges til at estimere den "ægte" værdi
 - Stikprøvens præcision øges ved at lave den større

Øvelse

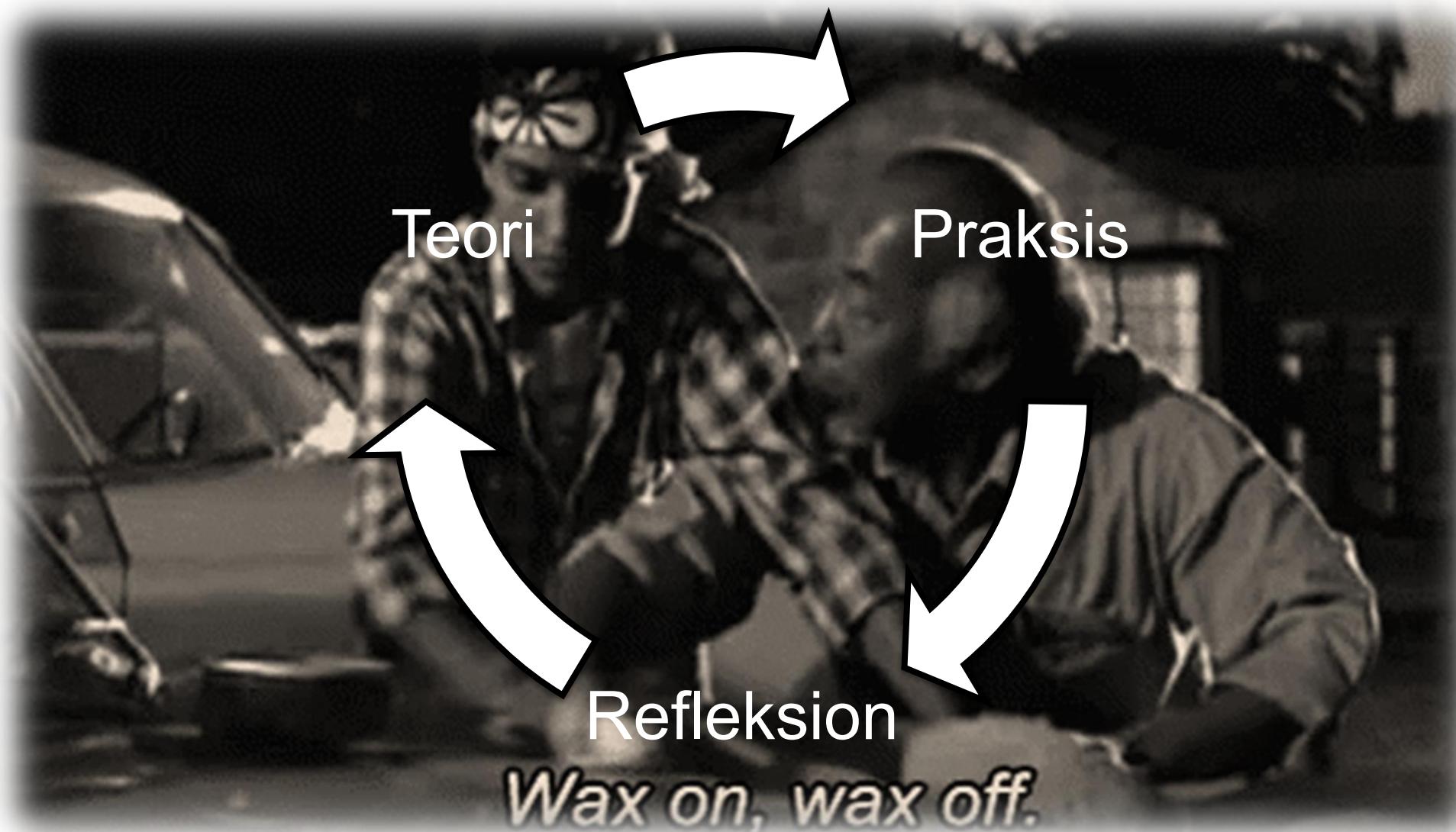
- Indlæs filen (igen)
 - 01_højde_testsæt.xls
- Bestem
 1. Konfidens interval for vægt
 2. Fordelt på køn
- 1. Brug menu
 1. STATA
 1. Statistics -> Summary &....
 2. SPSS
 1. Analyze – compare means

Øvelse

- Udfra jeres/din viden om fordeling af gennemsnit i stikprøver – forklar din sidemakker hvad du tror en t-test laver?
 - Måske nemmest at tegne.



Pause (måske)



Hvilke variable har du i dit datasæt?

Kategoriske variabler

- Anvendes ved ikke kontinuerlige data.
 - Eksempler
 - Performance status
 - TNM
 - Køn
 - Beskrives ofte som en fraktion/procent

Kategoriske variable - øvelse

- Indlæs datasæt
 - 02_cat_data.xlsx
- Se på køn og T stadie
 - Gennemsnit giver ikke mening
 - Find ud af hvor mange procent er
 - mænd/kvinder
 - T1-2 vs T3-4 (kodet hhv 1/2)

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

9 : OROPHARYNX 2 Visible: 10 of 10 Variables

	AGE	KOEN	TC B. A.	PACKYEARS	PERFORM	OROPHARYNX	T_STAGE	N_STAGE	STAGE	COMO	var										
1	51	11	2	0	2	1	2	2	2	0											
2	59	11	2	0	1	2	1	2	2	1											
3	48	12	2	1	0	2	2	2	2	1											
4	51	11	2	0	1	1	1	2	2	1											
5	56	12	2	0	2	2	2	2	2	2											
6	64	10	1	0	1	1	1	2	2	2											
7	65	11	2	0	2	2	2	2	2	2											
8	63	11	2	0	1	1	1	2	2	2											
9	63	11	2	1	2	2	2	2	2	0											
10	58	10	1	1	1	1	2	2	2	2											
11	58	11	.	0	2	2	2	1	2	2											
12	44	12	2	0	2	2	1	2	2	2											
13	65	10	1	0	1	1	2	2	2	2											
14	62	12	2	0	1	1	1	2	2	2											
15	42	22	2	0	1	1	1	2	2	2											
16	65	11	2	1	2	2	2	2	2	2											
17	57	11	.	0	1	2	2	2	2	2											
18	68	11	2	0	2	1	2	2	2	2											
19	48	22	1	0	2	2	1	2	2	1											
20	48	10	1	0	1	1	1	2	2	2											
21	59	21	2	0	2	1	1	2	2	2											
22	60	11	2	0	1	1	1	2	2	2											
23	57	12	2	1	1	1	2	2	2	2											
24	60	12	2	0	1	1	1	2	2	2											
25	65	10	1	0	1	1	1	2	2	2											
26	64	10	1	0	1	1	1	2	2	2											
27	59	12	2	0	2	2	2	2	2	2											
28	63	10	1	0	1	1	1	2	2	2											
29	61	10	1	0	1	1	1	2	2	2											
30	51	11	2	0	1	1	2	2	2	2											
31	55	20	1	0	1	1	2	2	2	2											
32	59	11	2	0	1	1	1	2	2	2											
33	62	21	2	0	1	1	1	2	2	2											
34	55	11	2	0	1	1	2	2	2	2											
35	57	11	2	0	2	2	2	2	2	2											
36	61	10	1	0	1	1	1	2	2	2											
37	46	12	.	1	2	2	1	2	1	2											
38	55	22	2	0	1	2	1	2	1	2											
39	54	11	2	0	1	2	2	2	2	2											

Data View

Variable View

IBM SPSS Statistics Processor is ready

Unicode:ON

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor

Analyze menu open, showing the **Frequencies...** option selected.

The Data View window shows a dataset with 39 rows and 10 variables. The variables are: OROPHARYNX, AGE, T_STAGE, N_STAGE, STAGE, COMO, and four unnamed variables labeled 'var'.

The Frequencies dialog box is displayed, with the following settings:

- Variable(s):** KOEN
- Statistics...**
- Charts...**
- Format...**
- Style...**
- Bootstrap...**
- Display frequency tables** checkbox is checked.

Buttons at the bottom of the dialog box: OK, Paste, Reset, Cancel, Help.

Bottom status bar: Data View, Variable View, Frequencies..., IBM SPSS Statistics Processor is ready, Unicode:ON.

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

Output Log Frequencies Title Notes Active Dataset Statistics KOEN

```
/TYPE=XLSX
/FILE='H:\Forskningstræning_2021_Kursist-mappe\Datagenerering\Statistik&Forskningsmetodologi\Data\02_cat_data.xlsx'
/SHEET=name 'testsæt'
/CELLRANGE=FULL
/READNAMES=ON
/DATATYPEMIN PERCENTAGE=95.0
/HIDDEN IGNORE=YES.

EXECUTE.

DATASET NAME DataSet1 WINDOW=FRONT.
FREQUENCIES VARIABLES=KOEN
/ORDER=ANALYSIS.
```

→ **Frequencies**

[DataSet1]

Statistics

KOEN

N	Valid	273
	Missing	0

KOEN

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	227	83,2	83,2
	2	46	16,8	100,0
Total	273	100,0	100,0	

IBM SPSS Statistics Processor is ready Unicode:ON

T-stadie

*Output1 [Document1] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

Missing 0

KOEN

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	227	83,2	83,2
	2	46	16,8	100,0
Total	273	100,0	100,0	

FREQUENCIES VARIABLES=T_STAGE
/ORDER=ANALYSIS.

→ **Frequencies**

Statistics

T_STAGE

N	Valid	273
	Missing	0

T_STAGE

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	166	60,8	60,8
	2	107	39,2	100,0
Total	273	100,0	100,0	

IBM SPSS Statistics Processor is ready Unicode:ON

Test

- Vi vil gerne vide om mænd har højere stadie end kvinder ved diagnosen.

	T1-T2	T3-T4	
Mand	134	93	
Kvinde	32	14	

Test

- Vi vil gerne vide om mænd har højere stadie end kvinder ved diagnosen.

	T1-T2	T3-T4	
Mand	134	93	227
Kvinde	32	14	46
	166	107	273

Test

- Vi vil gerne vide om mænd har højere stadie end kvinder ved diagnosen.
 - 2x2 tabel
 - Er det en forskel?
 - Chi² test

	T1-T2	T3-T4	
Mand	59%	41%	227
Kvinde	69%	31%	46
	166	107	273

Baggrund for test

- Observeret vs. Forventet
- Hvis to grupper er ens vil fordelingen af variablen også være ens.
 - Hvis fordelingen er anderledes end forventet
 - Efter kompenstation for usikkerhed
 - Bestemt af størrelsen på stikprøven
 - Er der en signifikant forskel

Test

- Vi vil gerne vide om mænd har højere stadie end kvinder ved diagnosen.
 - 2x2 tabel
 - Er det en forskel?
 - Chi² test

	T1-T2	T3-T4	
Mand	134	93	227
Kvinde	32	14	46
	166	107	273

Test

- Vi vil gerne vide om mænd har højere stadie end kvinder ved diagnosen.
 - 2x2 tabel
 - Er det en forskel?
 - Chi² test

	T1-T2	T3-T4	
Mand	134	93	227
Kvinde	32	14	46
	61%	39%	273

Test

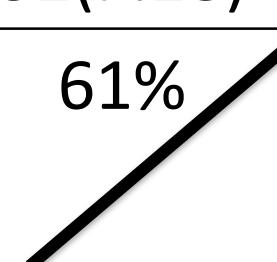
- Vi vil gerne vide om mænd har højere stadie end kvinder ved diagnosen.
 - 2x2 tabel
 - Er det en forskel?
 - Chi² test

	T1-T2	T3-T4	
Mand	134(F:138)	93(F:88)	227
Kvinde	32(F:28)	14(F:18)	46
	61%	39%	273

Test

- Vi vil gerne vide om mænd har højere stadie end kvinder ved diagnosen.
 - 2x2 tabel
 - Er det en forskel?
 - Chi² test

	T1-T2	T3-T4	
Mand	134(F:138)	93(F:88)	227
Kvinde	32(F:28)	14(F:18)	46
	61%	39%	273



Under 10 så Fischer's eksakte test

Øvelse - gerne med SPSS/STATA

- Rygning og Køn
- T og N
- Performance status og Køn
- Tolk fund med sidemand/kvinde

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

Reports Descriptive Statistics Bayesian Statistics Tables Compare Means General Linear Model Generalized Linear Models Mixed Models Correlate Regression Loglinear Neural Networks Classify Dimension Reduction Scale Nonparametric Tests Forecasting Survival Multiple Response Missing Value Analysis... Multiple Imputation Complex Samples Simulation... Quality Control ROC Curve... Spatial and Temporal Modeling...

Frequencies... Descriptives... Explore... Crosstabs... TURF Analysis Ratio... P-P Plots... Q-Q Plots...

9 : OROPHARYNX 2 Visible: 10 of 10 Variables

	AGE	OROPHARYNX	T_STAGE	N_STAGE	STAGE	COMO	var
1	51		2	1	2	2	0
2	59		1	1	2	2	1
3	48		2	2	2	2	1
4	51		1	1	2	2	1
5	56		2	2	2	2	1
6	64		0	1	1	2	1
7	65		0	2	2	2	1
8	63		0	1	2	2	1
9	63		1	1	2	2	1
10	58		1	1	2	2	1
11	58		1	1	2	2	1
12	44		1	1	2	2	1
13	65		1	1	2	2	1
14	62		1	1	2	2	1
15	42		1	1	2	2	1
16	65		1	1	2	2	1
17	57		1	1	2	2	1
18	68		1	1	2	2	1
19	48		1	1	2	2	1
20	48		1	1	2	2	1
21	59		1	1	2	2	1
22	60		1	1	2	2	1
23	57	12	2	1			
24	60	12	2	0			
25	65	10	1	0			
26	64	10	1	0			
27	59	12	2	0			
28	63	10	1	0			
29	61	10	1	0			
30	51	11	2	0			
31	55	20	1	0			
32	59	11	2	0			
33	62	21	2	0			
34	55	11	2	0			
35	57	11	2	0			
36	61	10	1	0			
37	46	12	-	1			
38	55	22	2	0			
39	54	11	2	0			

Crosstabs... Reports Descriptive Statistics Bayesian Statistics Tables Compare Means General Linear Model Generalized Linear Models Mixed Models Correlate Regression Loglinear Neural Networks Classify Dimension Reduction Scale Nonparametric Tests Forecasting Survival Multiple Response Missing Value Analysis... Multiple Imputation Complex Samples Simulation... Quality Control ROC Curve... Spatial and Temporal Modeling...

Exact... Statistics... Cells... Format... Style... Bootstrap...

Row(s): KOEN

Column(s): TOBAK

Layer 1 of 1 Previous Next

Display layer variables in table layers

Display clustered bar charts

Suppress tables

OK Paste Reset Cancel Help

Crosstabs Statistics

Chi-square Correlations

Contingency coefficient Gamma

Phi and Cramer's V Somers' d

Lambda Kendall's tau-b

Uncertainty coefficient Kendall's tau-c

Nominal by Interval Kappa

Eta Risk

Cochran's and Mantel-Haenszel statistics McNemar

Test common odds ratio equals: 1

Continue Cancel Help

Data View Variable View

Crosstabs... IBM SPSS Statistics Processor is ready Unicode:ON



- Output
 - Log
 - Frequencies
 - Title
 - Notes
 - Active Dataset
 - Statistics
 - KOEN
 - Log
 - Frequencies
 - Title
 - Notes
 - Statistics
 - T_STAGE
 - Log
 - Crosstabs
 - Title
 - Notes
 - Case Processing
 - KOEN * TOBAK CI
 - Chi-Square Tests

→ Crosstabs

Case Processing Summary

	Cases					
	Valid	Percent	Missing	Percent	N	Percent
KOEN * TOBAK	273	100,0%	0	0,0%	273	100,0%

KOEN * TOBAK Crosstabulation

Count		TOBAK			Total
		0	1	2	
KOEN	1	37	120	70	227
	2	8	17	21	46
Total		45	137	91	273

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	4,475 ^a	2	,107
Likelihood Ratio	4,424	2	,109
N of Valid Cases	273		

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 7,58.

Bi-nomial test

- Grundlag
 - Baserer sig på observeret vs. forventet
 - Kræver data er på bi-nært format (0/1)
 - Der testes på om outcome (0/1) er significant forskellig fra hypotetisk værdi

Konfidens interval på kategorisk variabel

- Fordeling af T-stadiet

The screenshot shows the IBM SPSS Statistics Data Editor interface. The window title is "*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor". The menu bar is visible with options: File, Edit, View, Data, Transform, Analyze, Graphs, Utilities, Extensions, Window, and Help. The "Analyze" menu is currently open, displaying various statistical procedures. The "Nonparametric Tests" option is highlighted, and its sub-options are shown: One Sample..., Independent Samples..., Related Samples..., Legacy Dialogs, Simulation..., Quality Control, and Spatial and Temporal Modeling.... A tooltip for "One Sample..." is displayed. The main data area shows a dataset with 249 rows and 5 columns: AGE, KOEN, T, O, and BA. The "Data View" tab is selected at the bottom. The status bar at the bottom right indicates "IBM SPSS Statistics Processor is ready" and shows the date and time as "02-00 28-11-2023".

 One-Sample Nonparametric Tests

Objective Fields Settings

Identifies differences in single fields using one or more nonparametric tests. Nonparametric tests do not assume your data follow the normal distribution.

What is your objective?

Each objective corresponds to a distinct default configuration on the Settings Tab that you can further customize, if desired.

- Automatically compare observed data to hypothesized
- Test sequence for randomness
- Customize analysis

Description

'Customize analysis' allows you fine-grained control over the tests performed and their options. The Wilcoxon Signed-Rank test is also available on the Settings tab.

Run Paste Reset Cancel Help



Search application

Visible: 10 of 10 Variables

	AGE	KOEN	T O BA S	PACKYEAR S	PERFORM	OROPHARY NX	T_STAGE	N_STAGE	STAGE	COMO	var							
211	65		1 1		2	0		2	1	2								
212	74		1 1		2	0		1	1	2								
213	69		1 1		2	1												
214	55		2 2		2	0												
215	60		1 1		2	0												
216	73		1 1		2	1												
217	78		1 1		2	0												
218	65		1 2		2	0												
219	69		1 1		.	1												
220	49		1 0		1	0												
221	54		2 1		2	1												
222	76		2 2		2	1												
223	63		1 1		2	2												
224	67		2 2		2	0												
225	53		1 2		2	0												
226	59		1 1		2	1												
227	59		1 1		1	0												
228	57		1 2		2	0												
229	63		1 1		2	2												
230	68		1 1		2	1												
231	59		2 1		2	1												
232	66		1 1		2	1												
233	55		1 1		2	1												
234	63		1 1		2	1												
235	52		1 2		2	1												
236	55		2 2		2	1												
237	71		1 2		2	1												
238	63		1 1		2	0												
239	54		1 2		2	2												
240	64		1 2		2	0												
241	74		2 2		2	1												
242	58		1 1		1	0												
243	54		2 2		2	0												
244	59		1 1		2	0												
245	55		1 2		2	1												
246	67		2 1		2	0												
247	73		1 1		1	0		2	2	1	2	1						
248	55		1 1		2	0		1	1	2	2	1						
249	66		2 2		2	1		1	2	2	2	1						

One-Sample Nonparametric Tests

 Use predefined roles Use custom field assignments

Fields:

Sort: None

- AGE
- KOEN
- TOBAK
- PACKYEARS
- PERFORM
- OROPHARYNX
- N_STAGE
- STAGE
- COMO

Test Fields:

 T_STAGE



Search application

	AGE	KOEN	T O BA	PACKYEAR S	PERFORM	OROPHARY NX	T_STAGE	N_STAGE	STAGE	COMO
211	65	1 1		2	0		2	1	2	1
212	74	1 1		2	0		1	1	2	1
213	69	1 1		2	1		2	1	1	1
214	55	2 2		2	0		1	1	2	1
215	60	1 1		2	0		1	1	2	1
216	73	1 1		2	1	2	1	1	1	1
217	78	1 1		2	0	1				
218	65	1 2		2	0	2				
219	69	1 1	.		1	1				
220	49	1 0		1	0		1			
221	54	2 1		2	1		1			
222	76	2 2		2	1		1			
223	63	1 1		2	2		2			
224	67	2 2		2	0		2			
225	53	1 2		2	0		2			
226	59	1 1		2	1		1			
227	59	1 1	1	0		1				
228	57	1 2	2	0		2				
229	63	1 1	2	2		1				
230	68	1 1	2	1		2				
231	59	2 1	2	1		2				
232	66	1 1	2	1		2				
233	55	1 1	2	1		2				
234	63	1 1	2	1		1				
235	52	1 2	2	1		2				
236	55	2 2	2	1		1				
237	71	1 2	2	1		1				
238	63	1 1	2	0		1				
239	54	1 2	2	2		2				
240	64	1 2	2	0		2				
241	74	2 2	2	1		2				
242	58	1 1	1	0	1		1	1	1	
243	54	2 2	2	0	2		1	1	1	
244	59	1 1	2	0	1		1	1	1	
245	55	1 2	2	1	2		2	2	1	
246	67	2 1	2	0	1		1	1	1	
247	73	1 1	1	0	2		2	1	1	
248	55	1 1	2	0	1		1	2	2	
249	66	2 2	2	1	1		2	2	1	

One-Sample Nonparametric Tests

Select an item:

Choose Tests

Test Options

User-Missing Values

 Automatically choose the test Customize tests Compare observed binary Compare observed proba Test observed distributio Compare median to hypo Test sequence for random

Binomial Options

Hypothesized proportion:

Confidence Interval

 Clopper-Pearson (exact) Jeffreys Likelihood ratio

Define Success for Categorical Fields

 Use first category found in data Specify success values

Success Values:



OK

Cancel

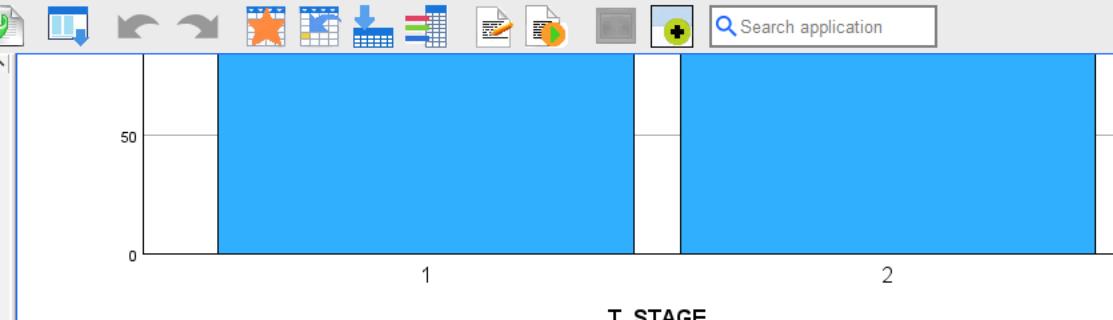
Help

Define Success for Continuous Fields

Success is equal to or less than

 Sample midpoint Custom cut point

- Title
- One-Sample Binomial
- One-Sample Binomial
- One-Sample Chi-Square Test
- Title
- TOBAK
- Title
- One-Sample Chi-Square
- One-Sample Chi-Square
- PERFORM
- Title
- One-Sample Chi-Square
- One-Sample Chi-Square
- One-Sample Kolmogorov-Smirnov
- Title
- AGE
- Title
- One-Sample Kolmogorov-Smirnov
- One-Sample Kolmogorov-Smirnov
- Continuous Field Information
- Categorical Field Information
- Nonparametric Tests
- Title
- Notes
- Hypothesis Test Summary
- Confidence Interval Summary
- One-Sample Binomial Test
- Title
- T_STAGE
- Title
- One-Sample Binomial
- One-Sample Binomial
- Categorical Field Information
- Nonparametric Tests
- Title
- Notes
- Hypothesis Test Summary
- Confidence Interval Summary
- One-Sample Binomial Test
- Title
- T_STAGE
- Title
- One-Sample Binomial
- One-Sample Binomial
- Categorical Field Information



Nonparametric Tests

Hypothesis Test Summary

Null Hypothesis	Test	Sig. ^{a,b}	Decision
1 The categories defined by T_STAGE = 1 and 2 occur with probabilities ,500 and ,500.	One-Sample Binomial Test	<,001	Reject the null hypothesis.

a. The significance level is ,050.

b. Asymptotic significance is displayed.

Confidence Interval Summary

Confidence Interval Type	Parameter	Estimate	95,0% Confidence Interval	
			Lower	Upper
One-Sample Binomial Success Rate (Clopper-Pearson)	Probability(T_STAGE=1.)	,608	,547	,666

One-Sample Binomial Test

T_STAGE

One-Sample Binomial Test Summary

Total N	273
Test Statistic	166,000
Standard Error	8,261
Standardized Test Statistic	3,510
Asymptotic Sig.(2-sided test)	<,001

One-Sample Binomial Test

Overlevelses analyse

- Vi se på en gruppe af mennesker. Vi starter med at tælle når de bliver 70 år og slutter når de dør
- Vi får følgende antal livs år

1	5
15	9
5	20
7	9
12	8
11	25
3	18

Opgave

- Lav en overlevelses graf

1	
15	
5	
7	
12	
11	
3	
5	
9	
20	
9	
8	
25	
18	

Opgave

- Lav en overlevelses graf

1	
3	
5	
5	
7	
8	
9	
9	
11	
12	
15	
18	
20	
25	

Hvad nu?

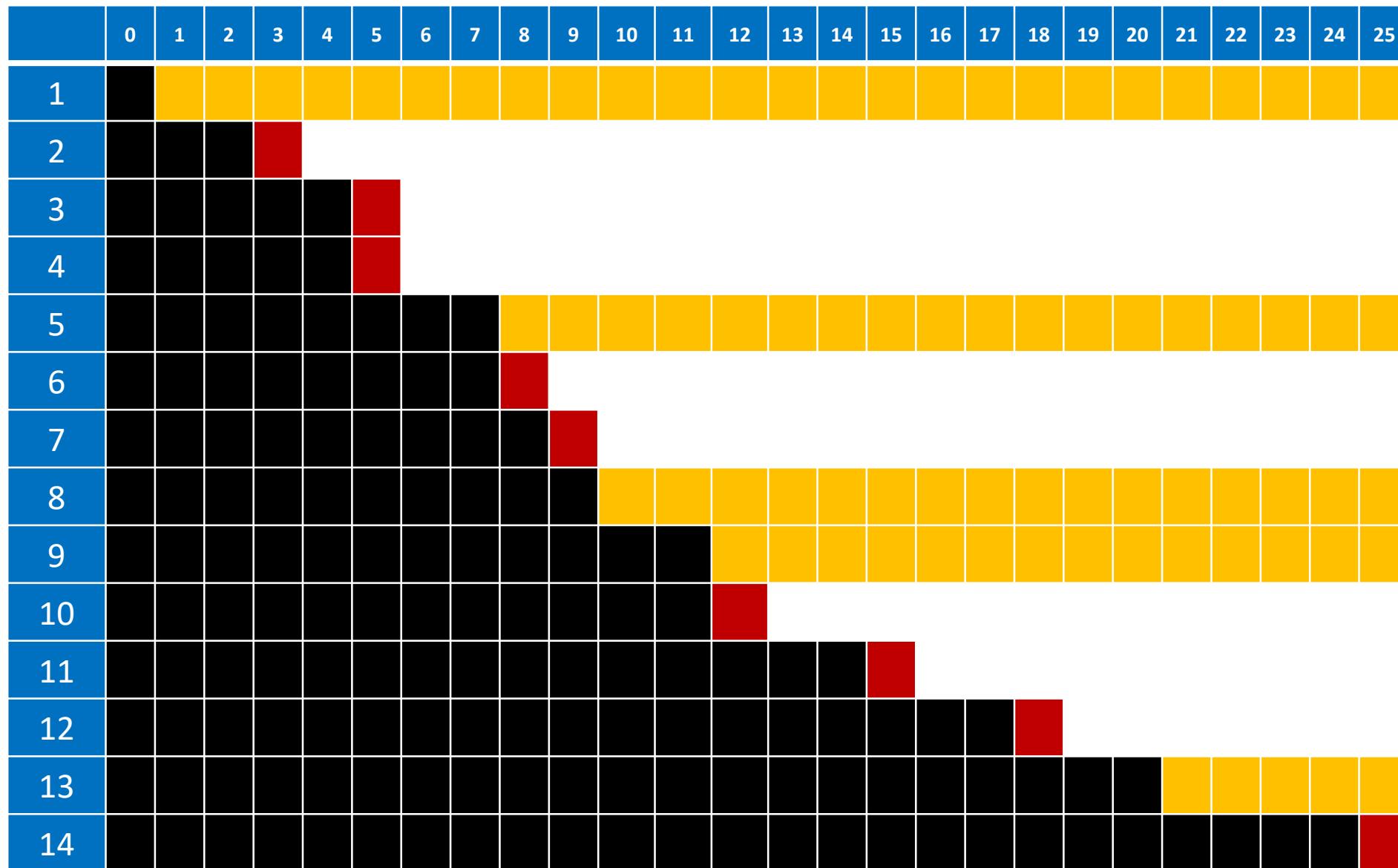
- Diskuter med sidemakker

Livstid	I live
1	Ja
3	Nej
5	Nej
5	Nej
7	Ja
8	Nej
9	Nej
9	Ja
11	Ja
12	Nej
15	Nej
18	Nej
20	Ja
25	Nej

Hvad nu?

- Vi definere at død er en hændelse – event
- I den periode patienten er observeret er der en risiko for et event
- Det kan så enten ske eller ikke

Livstid	I live
1	Ja
3	Nej
5	Nej
5	Nej
7	Ja
8	Nej
9	Nej
9	Ja
11	Ja
12	Nej
15	Nej
18	Nej
20	Ja
25	Nej

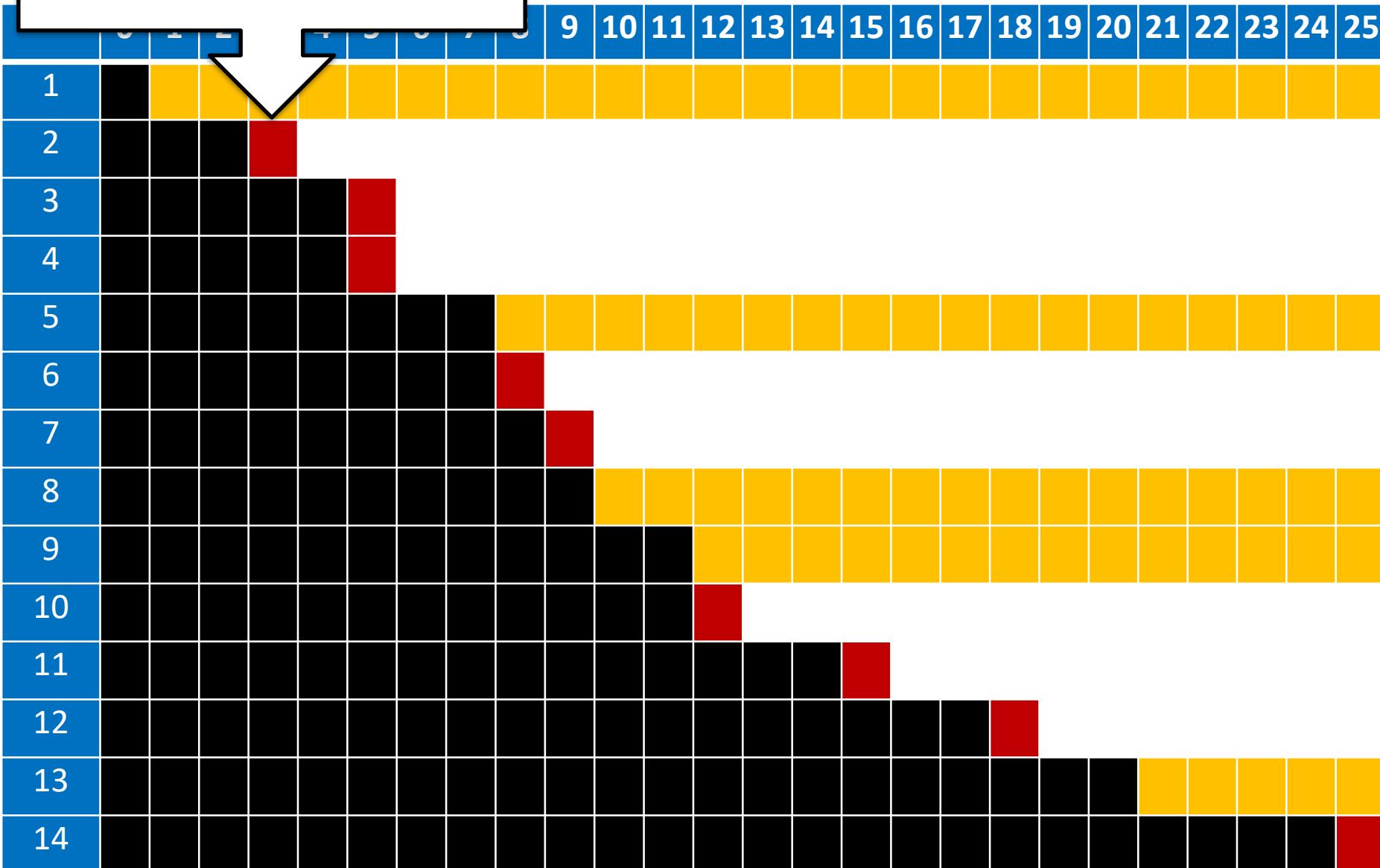


Kurven udregnes for hver rød plet

13 i risiko

1 dør

Kurven falder med 1/13 del

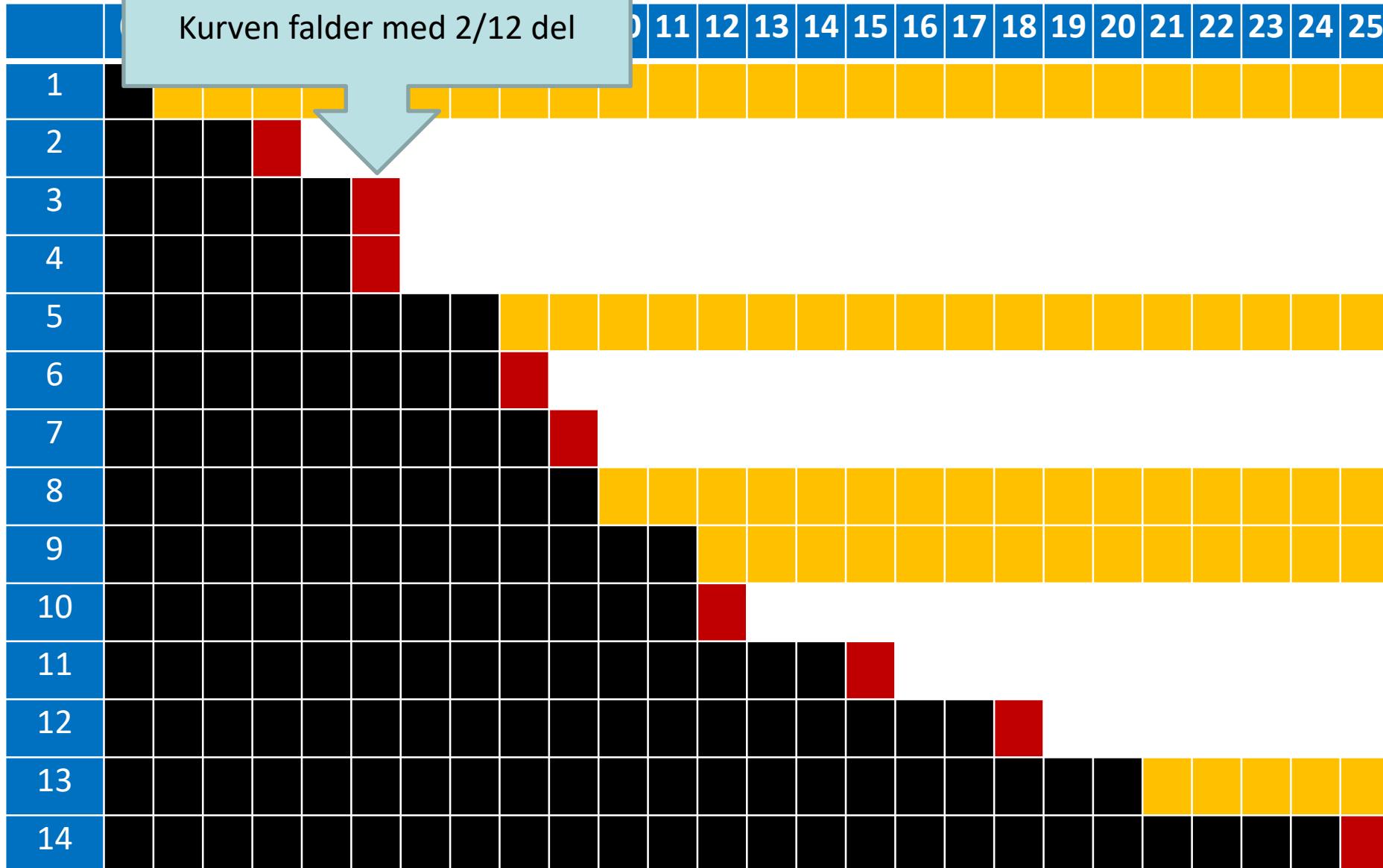


Kurven udregnes for hver rød plet

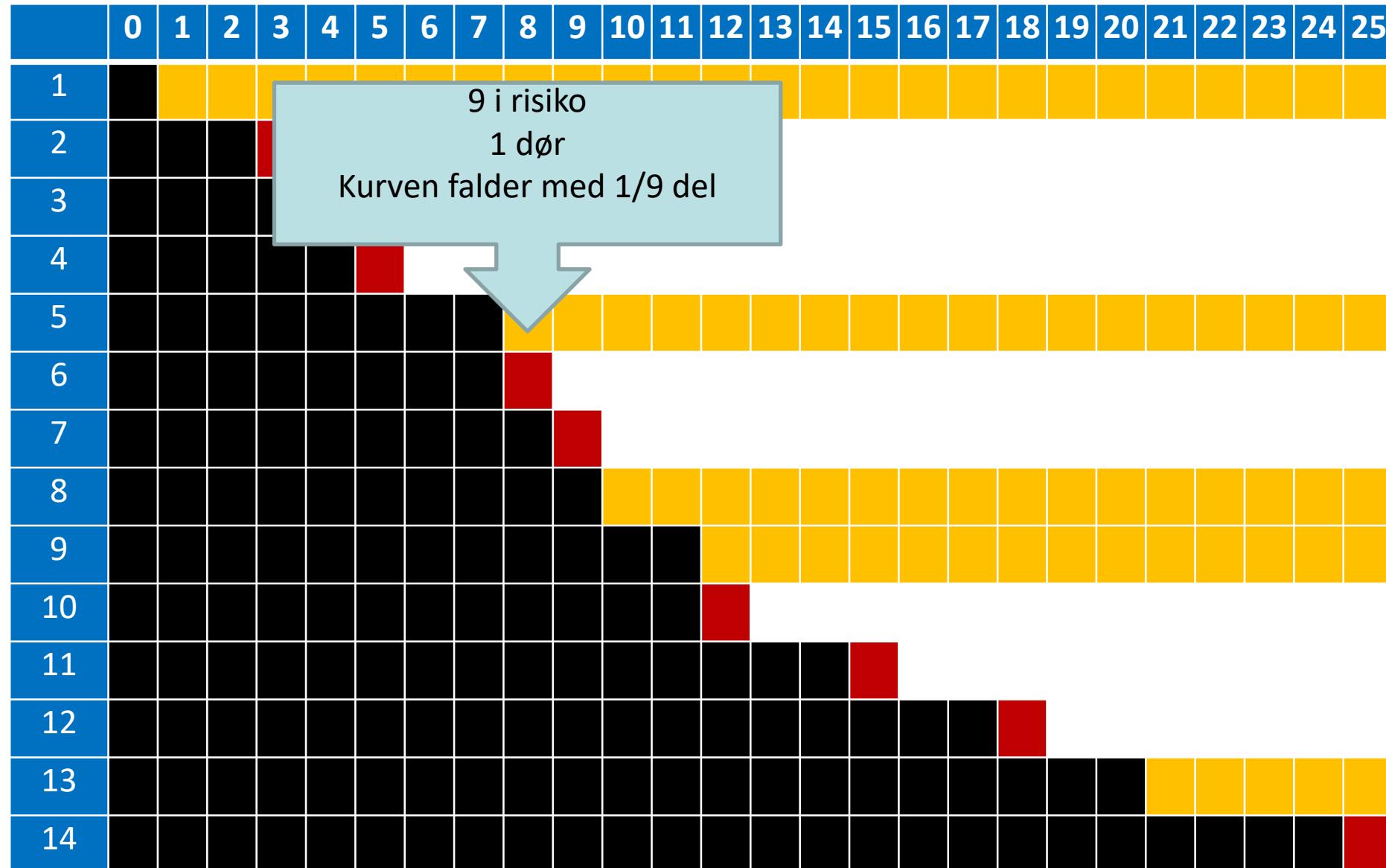
12 i risiko

2 dør

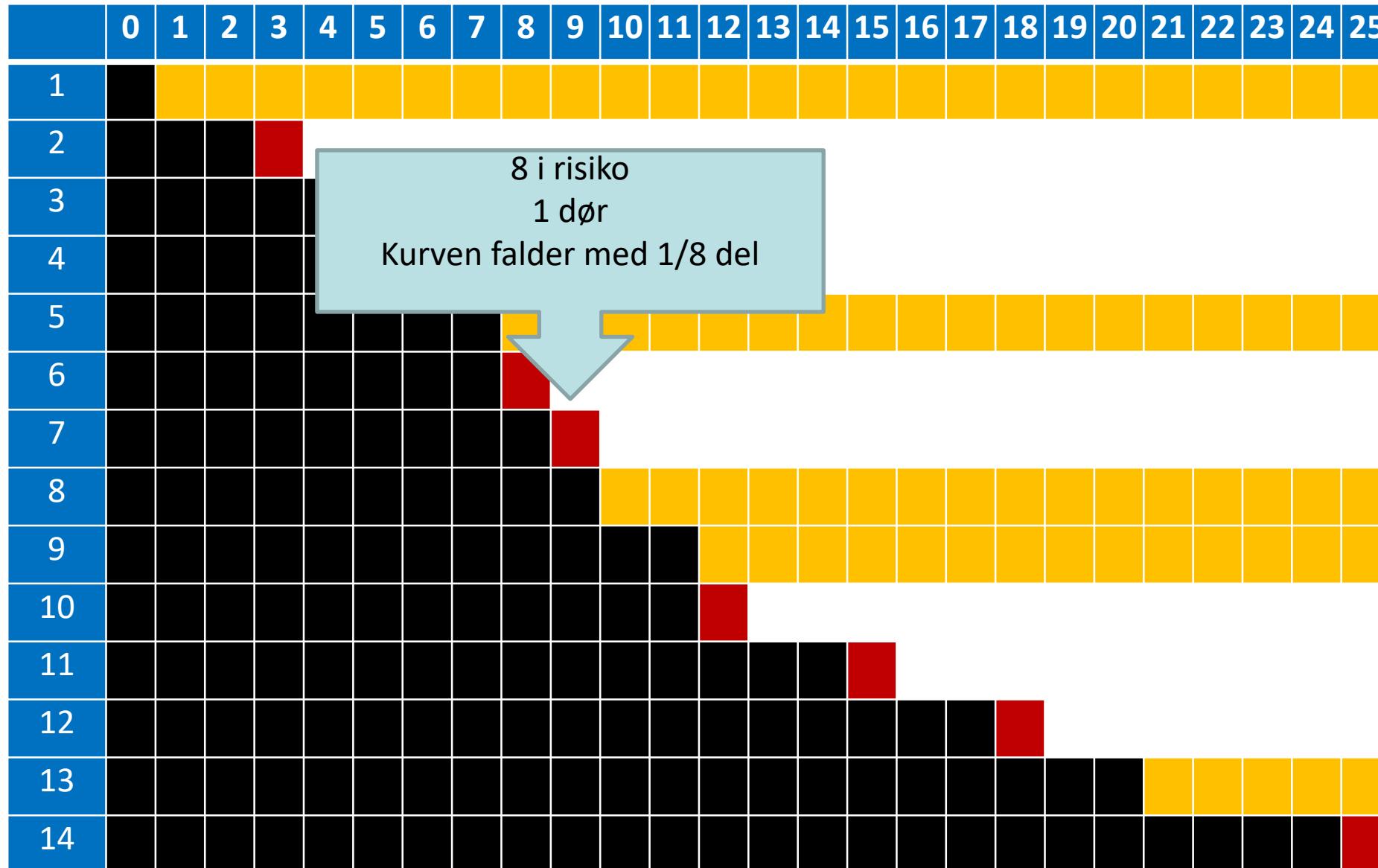
Kurven falder med 2/12 del



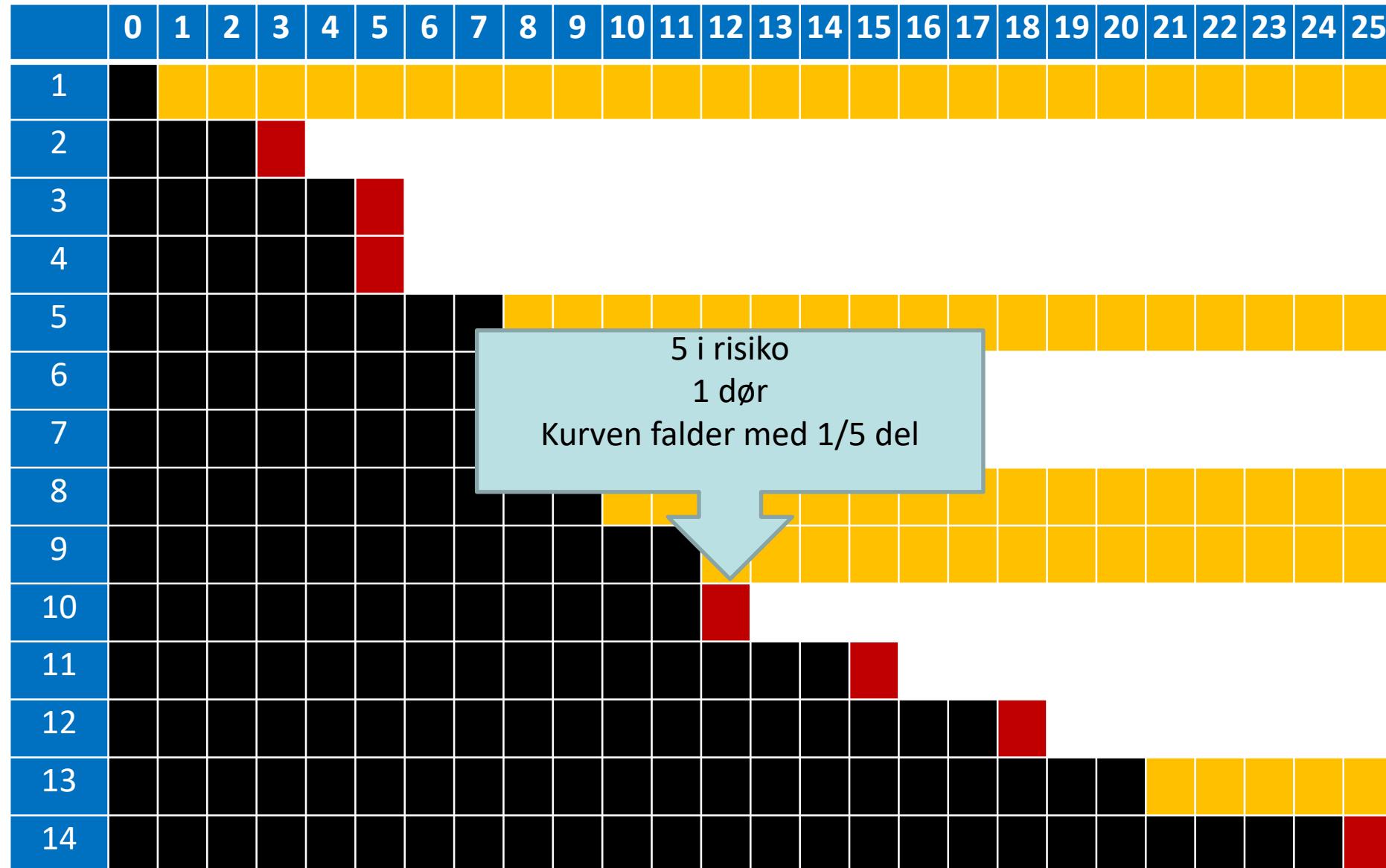
Kurven udregnes for hver rød plet



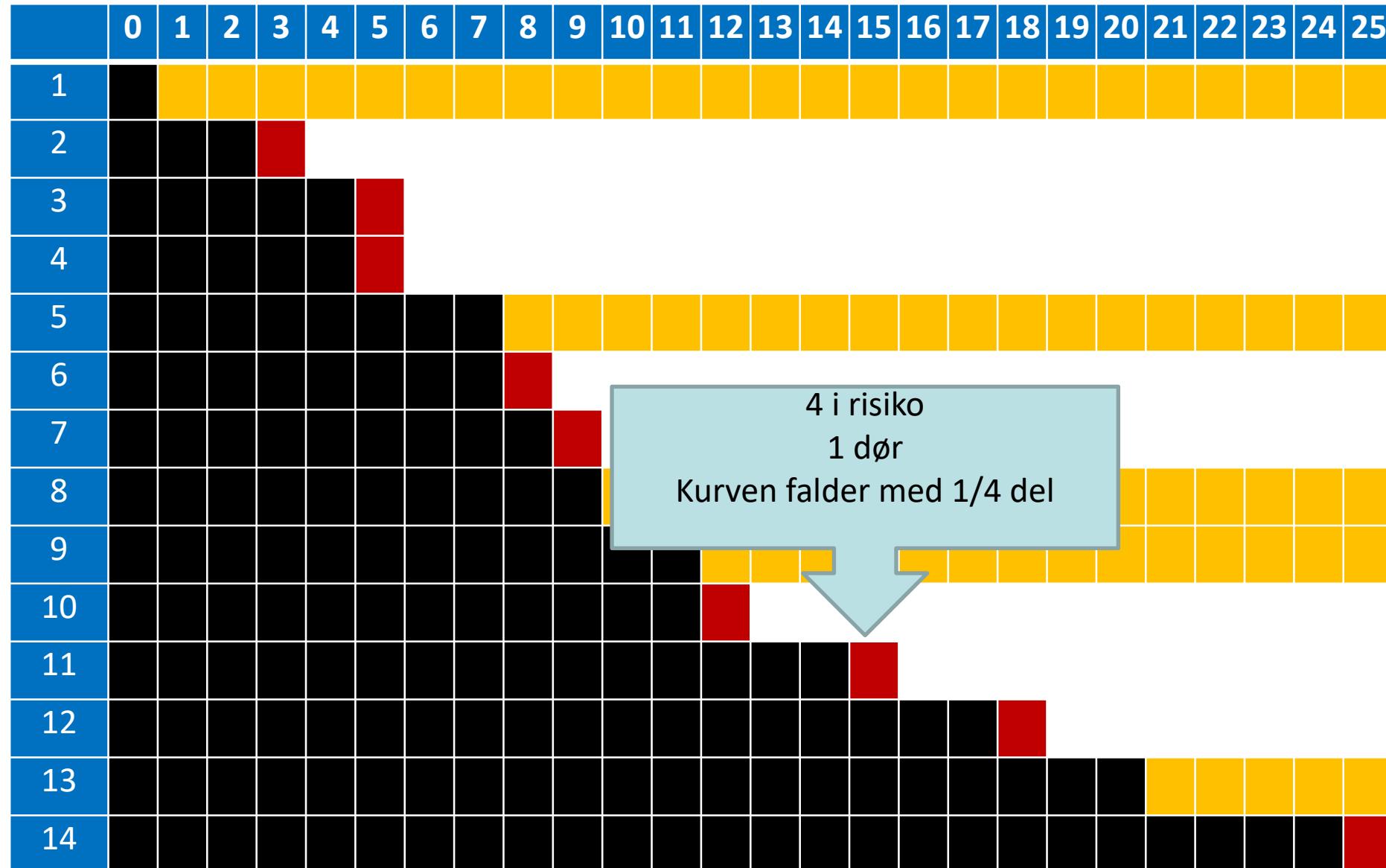
Kurven udregnes for hver rød plet



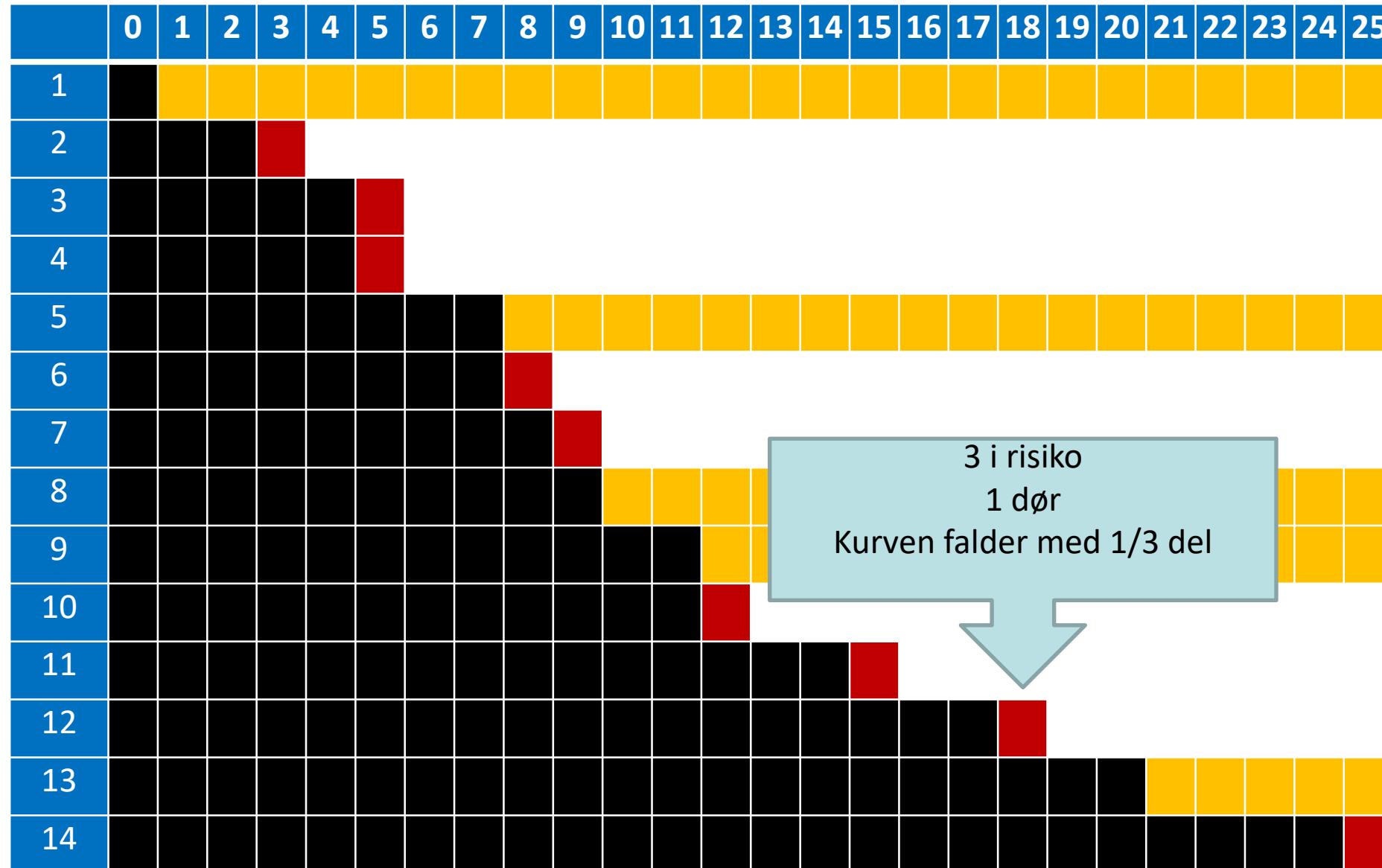
Kurven udregnes for hver rød plet



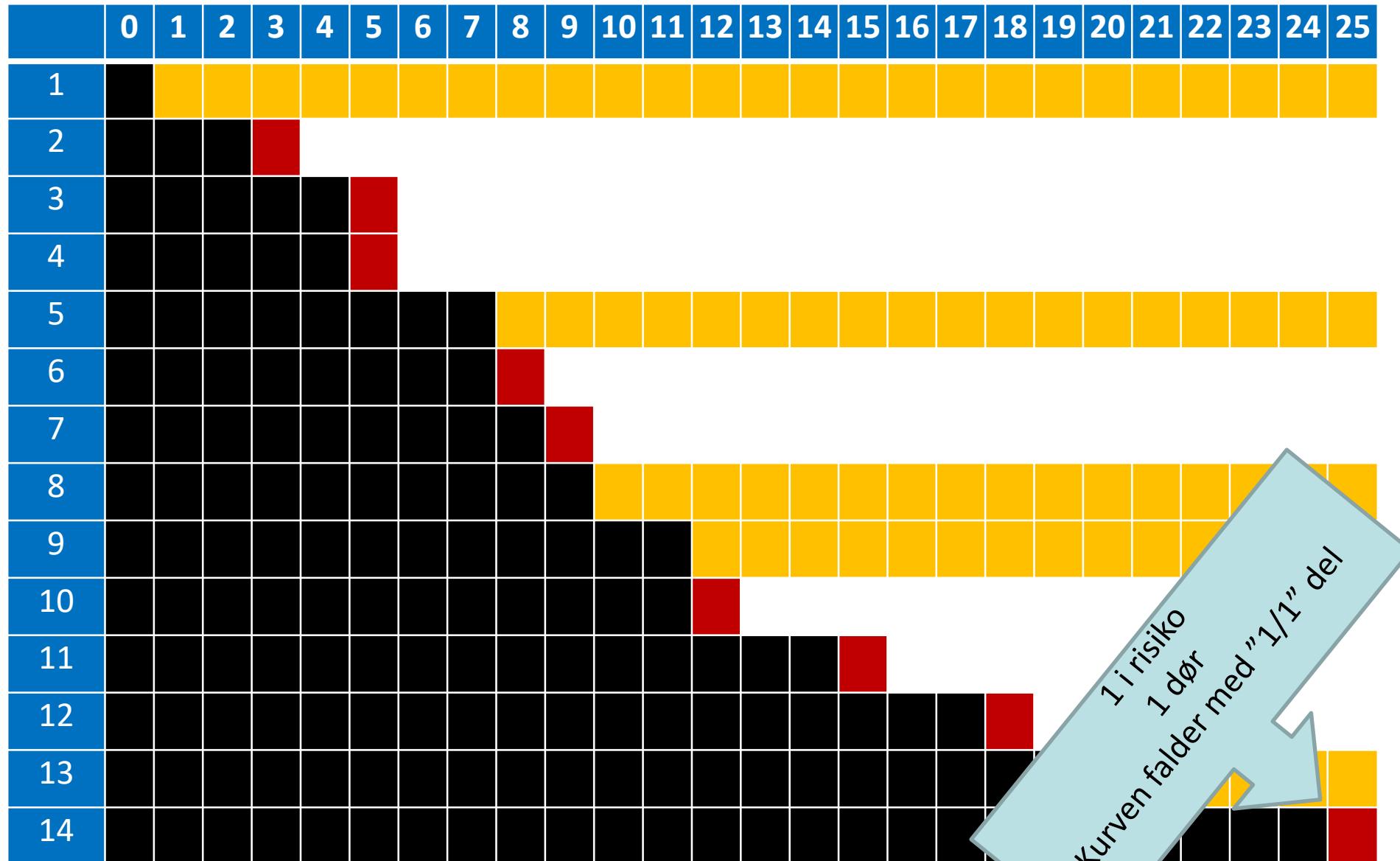
Kurven udregnes for hver rød plet



Kurven udregnes for hver rød plet



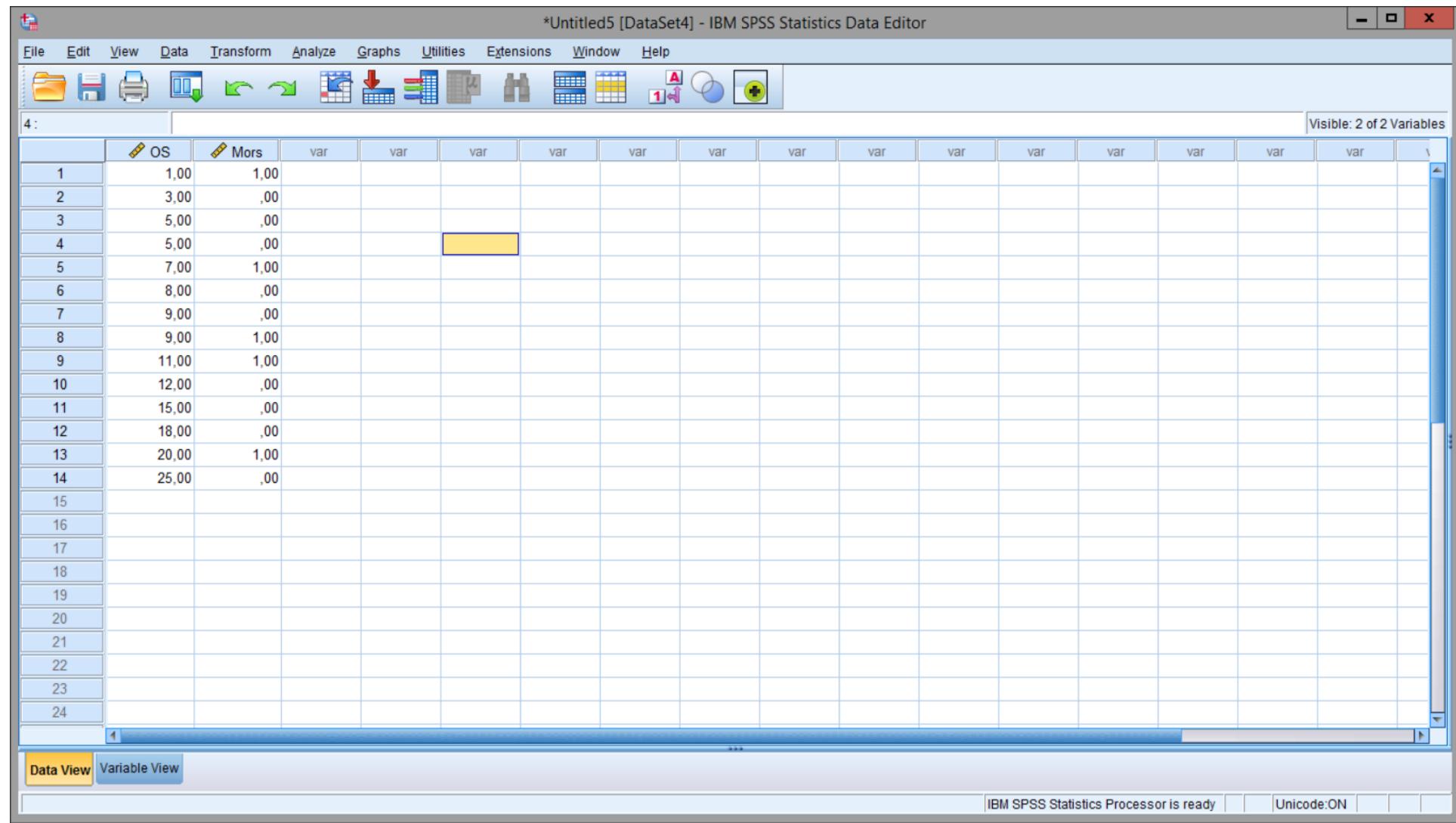
Kurven udregnes for hver rød plet



Øvelse

- Tast data ind i ...
- Lav en estimeret overlevelse kurve

risiko-tid	event
1	Ja
3	Nej
5	Nej
5	Nej
7	Ja
8	Nej
9	Nej
9	Ja
11	Ja
12	Nej
15	Nej
18	Nej
20	Ja
25	Nej



*Untitled5 [DataSet4] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

Reports Descriptive Statistics Bayesian Statistics Tables Compare Means General Linear Model Generalized Linear Models Mixed Models Correlate Regression Loglinear Neural Networks Classify Dimension Reduction Scale Nonparametric Tests Forecasting Survival Multiple Response Missing Value Analysis... Multiple Imputation Complex Samples Simulation... Quality Control ROC Curve... Spatial and Temporal Modeling...

OS Mors

	OS	Mors
1	1,00	1,00
2	3,00	,00
3	5,00	,00
4	5,00	,00
5	7,00	1,00
6	8,00	,00
7	9,00	,00
8	9,00	1,00
9	11,00	1,00
10	12,00	,00
11	15,00	,00
12	18,00	,00
13	20,00	1,00
14	25,00	,00
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		

Visible: 2 of 2 Variables

Kaplan-Meier

Time: OS
Status: Mors(?)
Factor:
Strata:
Label Cases by:

OK Paste Reset Cancel Help

Kaplan-Meier: Define Event For Stat...

Value(s) indicating event has occurred
 Single value: 1
 Range of values: through List of values:
Add Change Remove

Continue Cancel Help

Kaplan-Meier: Options

Statistics
 Survival table(s)
 Mean and median survival
 Quartiles

Plots
 Survival
 One minus survival
 Hazard
 Log Survival

Continue Cancel Help

Data View Variable View

Kaplan-Meier...

IBM SPSS Statistics Processor is ready Unicode:ON

Øvelse

- Indlæs
 - 03_Survival.xlsx
- Opgave
 - lave en KM kurve for alle patienter
 - Hint: Hvilke data skal i bruge til at tegne kurven?(transform)
 - Opdel kurven
 - P16 status
 - Performance status
 - T og N stadie
 - Tolk resultater med en ven

*Untitled6 [DataSet5] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

Visible: 18 of 18 Variables

	RANDATE	AGE	KOEN	TCA	PACKYEARS	PERFORM	OROPHARYNX	T_STAGE	N_STAGE	ST
1	11-Jan-08	51	11	2	0	2	1	2		
2	27-Feb-08	59	11	2	0	1	1	2		
3	27-Feb-08	48	12	2	1	2	2	2		
4	07-Mar-08	51	11	2	0	1	1	2		
5	28-Apr-08	56	12	2	0	2	2	2		
6	17-Apr-08	64	10	1	0	1	1	2		
7	25-Jul-08	65	11	2	0	2	2	2		
8	31-Jul-08	63	11	2	0	1	1	2		
9	27-Aug-08	63	11	2	1	2	2	2		
10	17-Sep-08	58	10	1	1	1	2	2		
11	26-Sep-08	58	11	.	0	2	2	1		
12	23-Oct-08	44	12	2	0	2	1	2		
13	14-Oct-08	65	10	1	0	1	2	2		
14	10-Oct-08	62	12	2	0	1	1	2		
15	07-Oct-08	28-May-13	0	42	22	2	0	1	1	2
16	10-Nov-08	31-May-13	0	65	11	2	1	2	2	2
17	06-Nov-08	27-May-13	0	57	11	.	0	1	2	2
18	03-Dec-08	23-Jun-13	0	68	11	2	0	2	1	2
19	28-Nov-08	18-Jun-13	0	48	22	1	0	2	2	1
20	07-Jan-09	29-May-13	0	48	10	1	0	1	1	2
21	29-Jan-09	20-Jun-13	0	59	21	2	0	2	1	2
22	05-Feb-09	27-Jun-13	0	60	11	2	0	1	1	2
23	26-Feb-09	17-Jun-13	0	57	12	2	1	1	2	2

Data View Variable View

Compute Variable... IBM SPSS Statistics Processor is ready Unicode:ON



Compute Variable



Target Variable:

OS

Type & Label...

RANDATE

Last_FU

MORS

AGE

KOEN

TOBAK

PACKYEARS

PERFORM

OROPHARYNX

T_STAGE

N_STAGE

STAGE

COMO

P16POS

TREATMENT

V16

V17

V18

Numeric Expression:

DATEDIFF(Last_FU,RANDATE,"months")

+	<	>	7	8	9
-	<=	>=	4	5	6
*	=	~=	1	2	3
/	&		0	.	.
**	~	()	Delete		

and returns an integer (with any fraction component truncated) in the specified date/time units, where datetime2 and datetime1 are both date or time format variables (or numeric values that represent valid date/time values), and "unit" is one of the following string literal values, enclosed in quotes: years, quarters, months, weeks, days, hours, minutes, seconds.

If... (optional case selection condition)

Function group:

All

Arithmetic

CDF & Noncentral CDF

Conversion

Current Date/Time

Date Arithmetic

Date Creation

Functions and Special Variables:

Datediff

Datesum(3)

Datesum(4)

OK

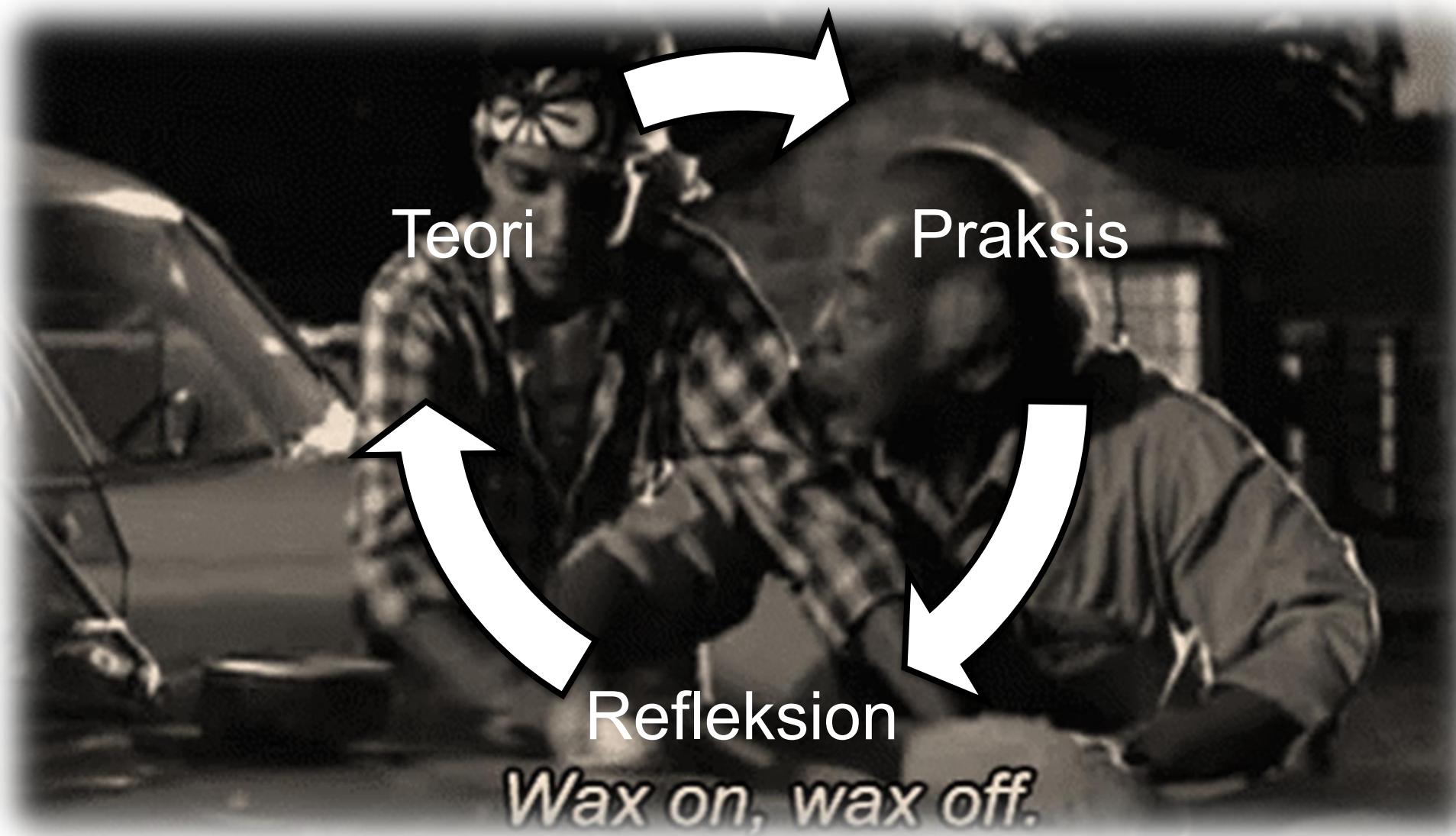
Paste

Reset

Cancel

Help

Pause/arbejd selv



Statistiske test af estimatede overlevelser

- Log-rank
- Cox regression
 - Generaliseret liniær model

Andre test

Tænk observeret vs. forventet

- Kig på data. (05-NP-test.xlsx)
 - Højde og køn
 - Lav en t-test
 - Er de ens?
- Hvis I ikke må lave en t-test
 - Hvordan kunne man så se om der var en forskel?
- 5 minutter

ID	Height	Gender	Weight	Blond
1	169	0	58	0
2	171	0	62	1
3	167	0	71	0
4	170	0	67	1
5	177	0	78	1
6	164	0	69	1
7	165	0	66	0
8	169	0	63	1
9	171	0	66	0
10	169	0	69	1
11	180	1	60	0
12	172	1	65	1
13	174	1	86	1
14	179	1	83	0
15	188	1	83	0
16	173	1	75	1
17	177	1	69	0
18	189	1	73	1
19	166	1	59	0
20	192	1	108	0

Observeret vs. forventet

- Kvinder – gul
- Mænd grøn

ID	Height	Gender	Weight	Blond
1	169	0	58	0
2	171	0	62	1
3	167	0	71	0
4	170	0	67	1
5	177	0	78	1
6	164	0	69	1
7	165	0	66	0
8	169	0	63	1
9	171	0	66	0
10	169	0	69	1
11	180	1	60	0
12	172	1	65	1
13	174	1	86	1
14	179	1	83	0
15	188	1	83	0
16	173	1	75	1
17	177	1	69	0
18	189	1	73	1
19	166	1	59	0
20	192	1	108	0

Observeret vs. forventet

- Sorter efter højde

ID	Height	Gender	Weight	Blond
6	164	0	69	1
7	165	0	66	0
19	166	1	59	0
3	167	0	71	0
1	169	0	58	0
8	169	0	63	1
10	169	0	69	1
4	170	0	67	1
2	171	0	62	1
9	171	0	66	0
12	172	1	65	1
16	173	1	75	1
13	174	1	86	0
5	177	0	78	0
17	177	1	69	0
14	179	1	83	1
11	180	1	60	0
15	188	1	83	0
18	189	1	73	1
20	192	1	108	1

Observeret vs. forventet

– Weight

ID	Height	Gender	Weight	Blond
1	169	0	58	0
2	171	0	62	1
3	167	0	71	0
4	170	0	67	1
5	177	0	78	1
6	164	0	69	1
7	165	0	66	0
8	169	0	63	1
9	171	0	66	0
10	169	0	69	1
11	180	1	60	0
12	172	1	65	1
13	174	1	86	1
14	179	1	83	0
15	188	1	83	0
16	173	1	75	1
17	177	1	69	0
18	189	1	73	1
19	166	1	59	0
20	192	1	108	0

Observeret vs. forventet

– Weight

ID	Height	Gender	Weight	Blond
1	169	0	58	0
19	166	1	59	0
11	180	1	60	0
2	171	0	62	1
8	169	0	63	1
12	172	1	65	1
7	165	0	66	0
9	171	0	66	0
4	170	0	67	1
6	164	0	69	1
10	169	0	69	1
17	177	1	69	0
3	167	0	71	0
18	189	1	73	1
16	173	1	75	1
5	177	0	78	0
14	179	1	83	1
15	188	1	83	0
13	174	1	86	0
20	192	1	108	1

Hårfarve

- Er blonde højere?

40 patienter

06_OS_LR_test.xlsx

- OS
 - Overlevelses tid
- Mors
 - Event
- Factor 0,1,2
 - Variabel vi har samlet ind
 - T stadie etc.

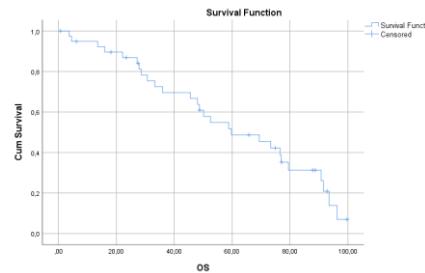
ID	OS	Mors	Factor0	Factor1	Factor2
1	3,6	1	0	1	1
2	59,7	1	1	1	0
3	65,8	0	1	0	0
4	76,9	1	1	1	0
5	48,1	1	0	1	0
6	93,5	1	1	1	1
7	13,6	1	1	0	1
8	35,9	1	1	1	0
9	16,0	1	0	1	1
10	99,7	0	1	1	1
11	30,7	1	0	1	0
12	79,5	1	0	1	1
13	92,8	0	0	0	1
14	69,4	1	1	1	0
15	52,5	1	1	1	0
16	27,5	0	0	0	0
17	91,5	1	1	1	1
18	96,2	1	1	0	1
19	50,1	1	1	0	0
20	58,8	1	1	0	0
21	48,7	1	0	0	0
22	0,7	0	0	0	1
23	22,2	1	0	1	1
24	48,8	0	1	1	0
25	74,9	0	1	1	0
26	90,8	1	1	0	1
27	73,4	1	0	0	0
28	45,6	1	0	0	0
29	87,8	0	1	0	1
30	77,1	0	1	0	1
31	33,3	1	0	0	0
32	27,2	1	1	1	1
33	4,5	1	0	0	1
34	6,2	0	0	1	1
35	18,1	0	0	0	1
36	27,8	1	0	1	0
37	23,4	0	0	0	1
38	28,6	1	0	0	0
39	88,7	0	1	1	1
40	76,6	1	1	0	0

Øvelse

- Lave en overlevelseskurve for hver af de 3 faktorer.
- Klik på log-rank under options

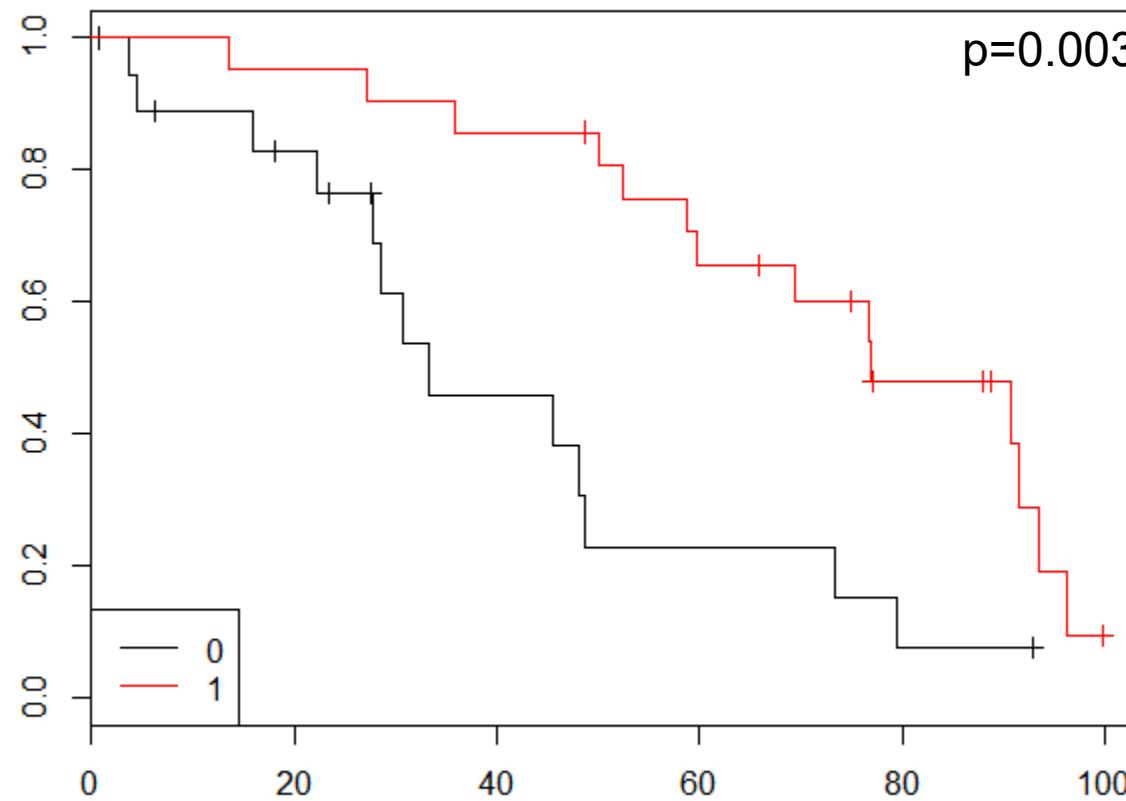
Overlevelses analyse

Forklar din sidemakker hvad du ser på



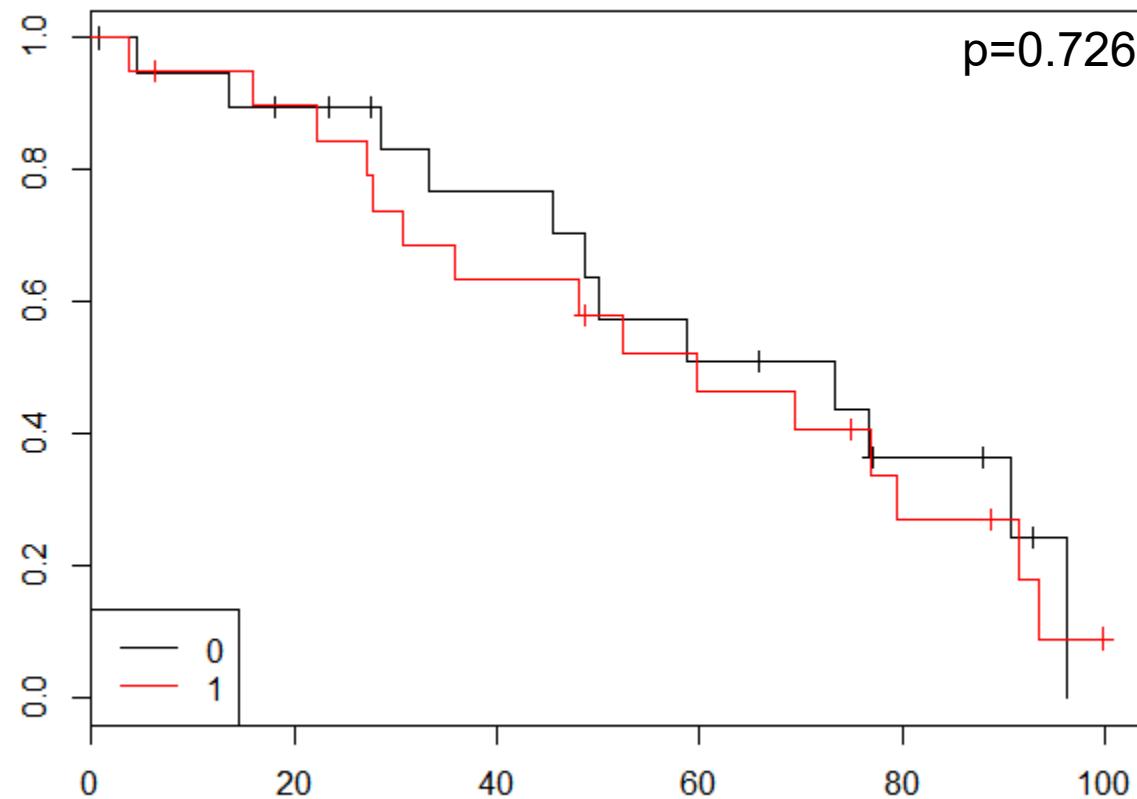
Delt på Factor0

Forklar din sidemakker hvad du ser på (skift)



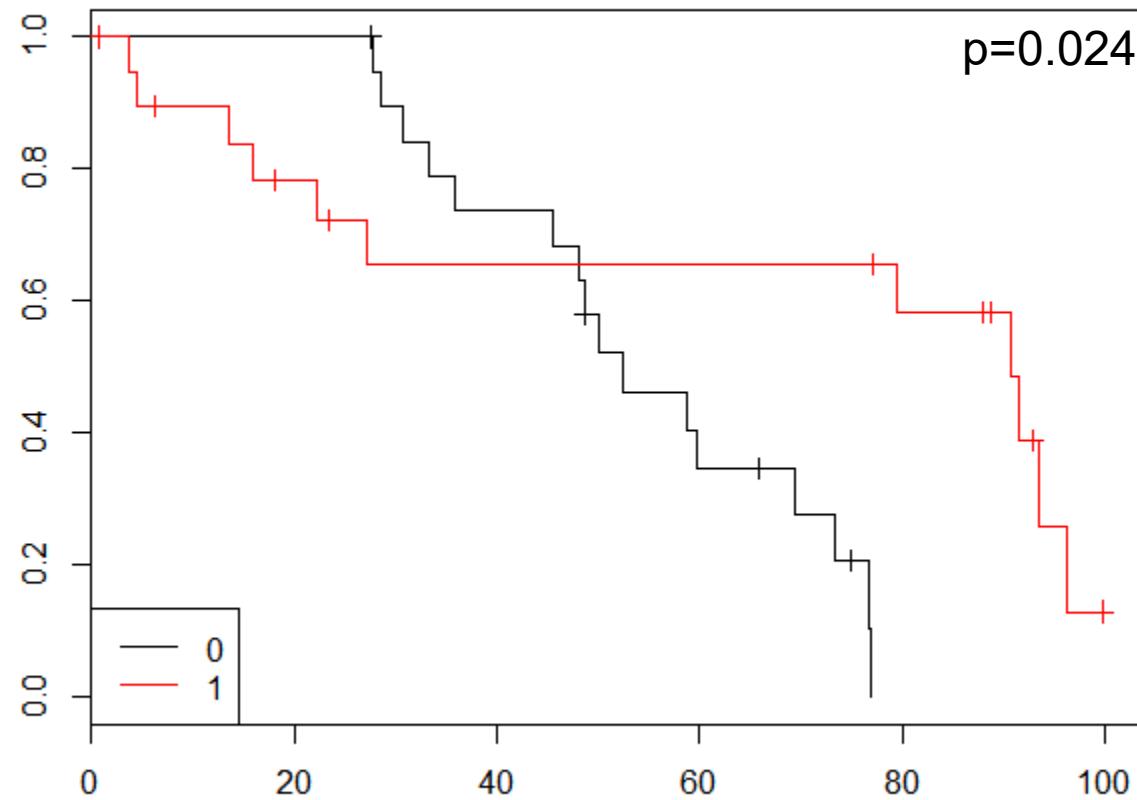
Factor1

Forklar din sidemakker hvad du ser på



Factor2

Forklar din sidemakker hvad du ser på



Hvor kom p-værdien fra?

- Diskuter omkring bordet
 - 5 minutter

40 patienter

- Sorterer efter factor
 - Sorterer
 - Sortere efter tid

	OS	Mors	Factor0	
	3,6	1	0	
	59,7	1	1	
	65,8	0	1	
	76,9	1	1	
	48,1	1	0	
	93,5	1	1	
	13,6	1	1	
	35,9	1	1	
	16,0	1	0	
	99,7	0	1	
	30,7	1	0	
	79,5	1	0	
	92,8	0	0	
	69,4	1	1	
	52,5	1	1	
	27,5	0	0	
	91,5	1	1	
	96,2	1	1	
	50,1	1	1	
	58,8	1	1	
	48,7	1	0	
	0,7	0	0	
	22,2	1	0	
	48,8	0	1	
	74,9	0	1	
	90,8	1	1	
	73,4	1	0	
	45,6	1	0	
	87,8	0	1	
	77,1	0	1	
	33,3	1	0	
	27,2	1	1	
	4,5	1	0	
	6,2	0	0	
	18,1	0	0	
	27,8	1	0	
	23,4	0	0	
	28,6	1	0	
	88,7	0	1	
	76,6	1	1	

Problem?

- Sorterer OS efter factor0
 - Sorterer
 - Sortere efter tid
- Hvad med dem uden event?

	OS	Mors	Factor0
0,7	0	0	
3,6	1	0	
4,5	1	0	
6,2	0	0	
13,6	1	1	
16	1	0	
18,1	0	0	
22,2	1	0	
23,4	0	0	
27,2	1	1	
27,5	0	0	
27,8	1	0	
28,6	1	0	
30,7	1	0	
33,3	1	0	
35,9	1	1	
45,6	1	0	
48,1	1	0	
48,7	1	0	
48,8	0	1	
50,1	1	1	
52,5	1	1	
58,8	1	1	
59,7	1	1	
65,8	0	1	
69,4	1	1	
73,4	1	0	
74,9	0	1	
76,6	1	1	
76,9	1	1	
77,1	0	1	
79,5	1	0	
87,8	0	1	
88,7	0	1	
90,8	1	1	
91,5	1	1	
92,8	0	0	
93,5	1	1	
96,2	1	1	
99,7	0	1	

Problem?

- Sorterer OS efter factor0
 - Sorterer
 - Sortere efter tid
- Hvad med dem uden event?
 - Kan ikke indgå i analysen
 - Hvad betyder det?

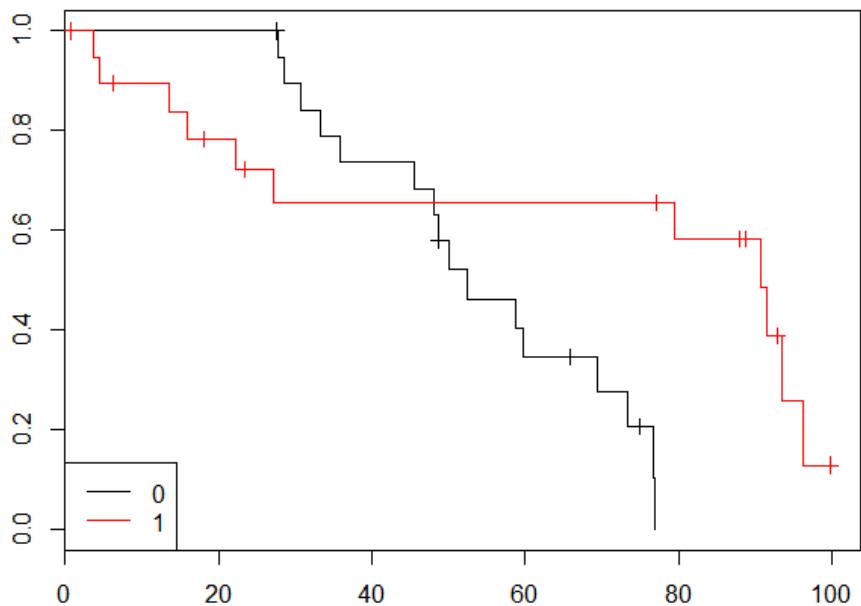
OS	Mors	Factor0
3,6	1	0
4,5	1	0
13,6	1	1
16	1	0
22,2	1	0
27,2	1	1
27,8	1	0
28,6	1	0
30,7	1	0
33,3	1	0
35,9	1	1
45,6	1	0
48,1	1	0
48,7	1	0
50,1	1	1
52,5	1	1
58,8	1	1
59,7	1	1
69,4	1	1
73,4	1	0
76,6	1	1
76,9	1	1
79,5	1	0
90,8	1	1
91,5	1	1
93,5	1	1
96,2	1	1

Lav eksemplet med Factor2

Lav eksemplet med Factor2

- Hvad nu?
 - Tolk resultater med sidemakkeren

p=0.024

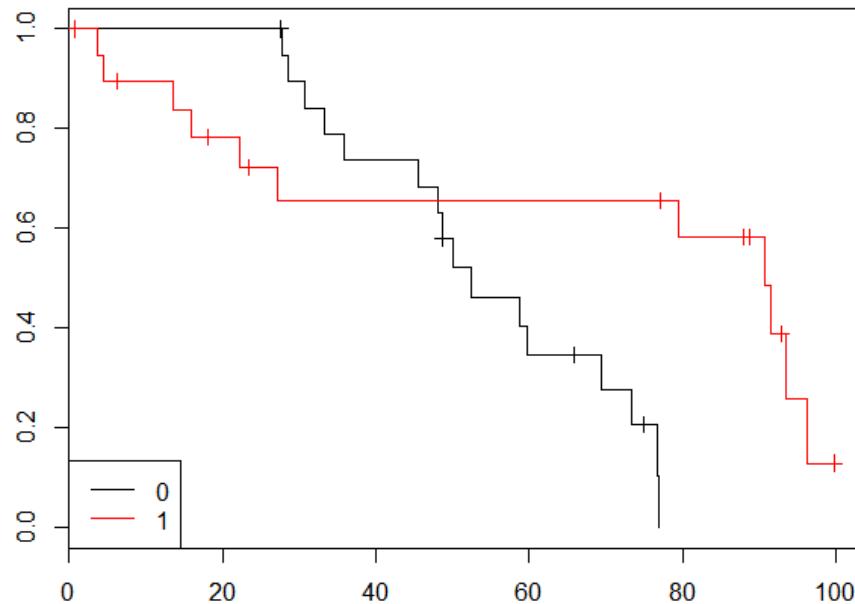


OS	Mors	Factor2
3,6	1	1
4,5	1	1
13,6	1	1
16,0	1	1
22,2	1	1
27,2	1	1
27,8	1	0
28,6	1	0
30,7	1	0
33,3	1	0
35,9	1	0
45,6	1	0
48,1	1	0
48,7	1	0
50,1	1	0
52,5	1	0
58,8	1	0
59,7	1	0
69,4	1	0
73,4	1	0
76,6	1	0
76,9	1	0
79,5	1	1
90,8	1	1
91,5	1	1
93,5	1	1
96,2	1	1

Lav eksemplet med Factor2

- Hvad nu?
 - Analysen har problemer
 - når kurverne krydser

p=0.024



OS	Mors	Factor2
3,6	1	1
4,5	1	1
13,6	1	1
16,0	1	1
22,2	1	1
27,2	1	1
27,8	1	0
28,6	1	0
30,7	1	0
33,3	1	0
35,9	1	0
45,6	1	0
48,1	1	0
48,7	1	0
50,1	1	0
52,5	1	0
58,8	1	0
59,7	1	0
69,4	1	0
73,4	1	0
76,6	1	0
76,9	1	0
79,5	1	1
90,8	1	1
91,5	1	1
93,5	1	1
96,2	1	1

Status

- Du skulle nu gerne have en basal forståelse for:
 - Stikprøver
 - Variation
 - t-test
 - Chi² test
 - Log-rank test
 - Konceptet
 - Observeret vs. forventet