

Faculty of Health Sciences

Basal Statistik - SPSS

Begreber. Parrede sammenligninger.

Lene Theil Skovgaard

4. februar 2019



APPENDIX

med instruktioner til **SPSS**-analyse svarende til nogle af slides

- ▶ Indlæsning af vitamin D datasæt, s. 3-4
- ▶ Tegninger vedrørende vitamin D, s. 5-7
- ▶ Udregning af summary statistics og fraktildiagram, s. 8-9
- ▶ Indlæsning af MF-SV data, s. 10
- ▶ Tegninger vedrørende MF-SV, s. 11-13
- ▶ Parret T-test, s. 14-16



Indlæsning af Vitamin D datasæt

Slide 6

Vi indlæser filen fra nettet ved at benytte

File/Open/Internet Data, hvorefter man skriver stien
`http://publicifsv.sund.ku.dk/~lts/basal/data/VitaminD.txt`
i Web location... samt det ønskede navn på datasættet i
Dataset Name to Assign.

Derefter går man til File/Open/Data og sætter Files of Type
til All Files og følger derefter instruktionerne.

I SPSS benyttes variabelnavnet land til de numeriske 1,2,4 og 6,
medens country defineres som DK, SF, EI, PL eller med det
fulde navn på landet.



Indlæsning af Vitamin D datasæt, 2

Slide 6, fortsat

Vi skal kun se på Irske kvinderne (`land=4` og `category=2`), så vi benytter

Data/Select Cases, hvor der afkrydses i If og skrives `land=4 & category=2`.

Herefter transformeres ved hjælp af

Transform/Compute Variable, idet man sætter `log10vitd` i Target Variable og skriver `Lg10(vitd)` i definitionsfeltet.

Tilsvarende med $\log_2$, hvor man sætter `log2vitd` i Target Variable og skriver `Lg10(vitd)(Lg10(2))` i definitionsfeltet.



Histogram

Slide 9

Benyt Graphs/Chart Builder og vælg Histogram (det enkleste, dobbeltklik det op i det store felt), sæt vidt over på X-aksen, og tryk OK, så figuren fremkommer.

For at lægge en normalfordelingskurve oveni, dobbeltklikker man på figuren, klikker på show distribution curve-ikonet og afkrydser Normal og trykker Apply/Close.

Box-plot

Slide 11

Boxplottet skal vise alle landene, så vi må ind i Data/Select Cases, hvor der afkrydses i If og rettes til category=2.

Herefter benytter vi

Analyze/Descriptive Statistics/Explore, hvor vi sætter vitd i Dependent List, country i Factor List samt sætter hak i Plots

Scatter plot med linier

Slide 12

Benyt Graphs/Chart Builder og vælg Scatter (det, der er nummer to fra venstre), og dobbeltklik det op i det store felt. Sæt bmi over på X-aksen, vitd over på Y-aksen, og country over i Set Color.

For at lægge regressionslinier oveni, dobbeltklikker man på figuren, klikker på Add Fit line at Subgroups-ikonet og afkrydser Linear. Jeg plejer også at fjerne fluebenet i Attach Label to Line, før jeg trykker Apply.



Udregning af summary statistics

Slide 17

For at få opdelt udregningerne i de enkelte lande, går vi først ind i Data/Split File, vælger Compare groups og sætter country over i Groups Based on.

Herefter benyttes

Analyze/Descriptive Statistics/Frequencies, hvor vi sætter vitd over i Variable(s), fjerner fluebenet i Display Frequency Tables, og under Statistics vælges Mean, Median, Quartiles og Std.deviation.



Specielle fraktiler og fraktildiagram

Slide 21 og 24

Nu skal vi igen kun se på Irske kvinder, så vi må tilbage i Data/Select Cases, hvor der afkrydses i If og skrives `land=4 & category=2`.

Herefter benyttes

Analyze/Descriptive Statistics/Frequencies, hvor vi sætter vitd over i Variable(s), fjerner fluebenet i Display Frequency Tables, og under Statistics vælges Percentiles, hvor vi skriver 2,5, klikker Add og skriver 97,5 og klikker Add/Continue

Til fraktildiagrammet benyttes

Analyze/Descriptive Statistics/Q-Q Plots og sætter vitd over i Variables og der trykkes OK



Datafilen vedr. MF og SV

Slide 37

Indlæsning: se tilsvarende her i appendix, s. 3-4

Definition af to ny variable:

Brug Transform/Compute Variable, idet man sætter dif i Target Variable og skriver `mf-sv` i definitionsfeltet.

Tilsvarende med gennemsnittet, hvor man sætter average i Target Variable og skriver $(mf+sv)/2$ i definitionsfeltet.



Scatter plot med identitetslinie

Slides 38

Benyt Graphs/Chart Builder og vælg Scatter (det simple længst til venstre), og dobbeltklik det op i det store felt. Sæt sv over på X-aksen og mf over på Y-aksen.

For at lægge identitetslinien oveni, dobbeltklikker man på figuren, klikker på Properties-ikonet og vælger

Add a reference line from Equation, hvorefter man i Custom Equation skriver $y=1*x+0$ (måske står det der allerede) og herefter Apply.

Evt kan man vælge liniens farve under Linear.



Box-plot og Spaghetti-plot

Slide 39

Dette er vanskeligt at beskrive i SPSS....

Man skal benytte Data/Restructure og følge anvisningerne, herunder vælge

- ▶ person som betegnelse for Case group identification
- ▶ flow som betegnelse for Target Variable
- ▶ metode som betegnelse for Index Variable

Box plot:

Herefter benyttes Graphs/Chart Builder/Boxplot (det længst til venstre), hvor flow sættes på Y-aksen og metode på X-aksen.

Spaghettiplot:

Herefter benyttes Graphs/Chart Builder/Line (det opdelte, nr. 2 fra venstre), hvor flow sættes på Y-aksen, metode på X-aksen og person som Set color.



Bland-Altman plot

Slides 41

Benyt Graphs/Chart Builder og vælg Scatter (det simple længst til venstre), og dobbeltklik det op i det store felt. Sæt average over på X-aksen og dif over på Y-aksen.

For at lægge en vandret linie i 0 oveni, dobbeltklikker man på figuren, klikker på Properties-ikonet og vælger Add a reference line from Equation, hvorefter man i Custom Equation skriver $y=0$ og herefter Apply.

Evt kan man vælge liniens farve under Linear.



Estimat med konfidensgrænser

Slide 51

Benyt Analyze/Descriptive Statistics/Explore og under Statistics afkrydses Confidence Interval for Mean 95%



Parret T-test

af MV vs. SV, med differenser `dif`

Slide 59 og 60

Benyt Analyze/Compare Means/Paired Samples T-test,
marker både `mf` og `sv` samtidig og før dem over til `Pair1`.
Klik OK



Parret non-parametrisk test

af MV vs. SV, med differenser dif

Slide 76

Brug Analyze/Nonparametric Tests/
Legacy Dialogs/2 Related Samples, sæt mf over i
Variable1, sv over i Variable2 og gå dernæst ind i Test Type
og afkryds Wilcoxon samt i Exact og afkryds Exact

Jeg har også forsøgt at benytte
Analyze/Nonparametric Tests/Related Samples,
men **den giver noget helt uforståeligt** (P-værdi på 747....??)

