

Faculty of Health Sciences

Basal Statistik - SPSS

Regressionsanalyse.

Lene Theil Skovgaard

25. februar 2019



APPENDIX

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

- ▶ Indlæsning og scatter plot, s. 3
- ▶ Regressionsanalyse, s. 4-6
- ▶ Korrelation, s. 7
- ▶ Konfidens- og prediktionsgrænser, s. 8
- ▶ Modelkontrol, s. 9-11
- ▶ Diagnostics, s. 12



Indlæsning og scatter plot i SPSS

Slide 8

Indlæsning fra nettet

File/Open/Internet Data, hvorefter man skriver stien
`http://publicifsv.sund.ku.dk/~lts/basal/data/vcf.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.

Scatter plot

Man går ind i Graph/Chart Builder/Scatter og i den
fremkomne boks trækker man vcf over på Y-aksen, og
blodsukker over på X-aksen.



Regressionsanalyse i SPSS

Slide 13 og 14

Man går ind i menuen `Analyze/Regression/Linear`, og i boksen sættes `vcf` som `Dependent` og `blodsukker` som `Independent(s)` (et uheldigt navn til forklarende variable...)

Man skal efterfølgende huske at gå ind i `Statistics` og afkrydse `Parameter Estimates` og `Confidence intervals`

Man kan også benytte den mere generelle

`Analyze/Genereal Linear Model/Univariate`
(mere om denne senere)



Regressionslinie i SPSS

Slide 16

Først laves plottet ved at gå ind i Graph/Chart Builder/Scatter og i den fremkomne boks trækker man vcf over på Y-aksen, og blodsukker over på X-aksen.

For at tegne linien dobbeltklikker man efterfølgende på grafen og klikker på ikonet Add Fit Line at Total og derefter i Properties-boksen afkrydse Linear og klikke Apply.

Man kan endvidere vælge, om man vil have liniens ligning skrevet på linien eller ej (flueben ved Attach label to line kan fjernes) Farven på linien kan vælges under Lines: klik på farven og Apply/Close



Forventet værdi for specifikke værdier af kovariaten

Slide 18

Her for en blodsukkerværdi på 10: I SPSS kan man finde et sådant estimat ved at snyde den til at tro, at det er interceptet.

Dette gør vi ved at flytte nulpunktet hen i 10 mmol/l, dvs. ved at benytte en ny X-variabel, der er blodsukker-10. Denne udregnes i Transform/Compute, og herefter gentages regressionen med denne nye X-variabel.

Alternativt kan man tilføje en observation med blodsukker=10 og manglende outcome (`vcf=.`). Man kan nu vælge at gemme predikterede værdier, med tilhørende konfidensintervaller (se s. 9), heriblandt for den fiktive observation, der ikke ændrer på noget, da outcome er missing.



Korrelationer i SPSS

Slide 38-39

Pearson: betegner den *sædvanlige* korrelation
baseret på en todimensional normalfordeling

Spearman: betegner den nonparametriske korrelation

Korrelationer udregnes i SPSS ved at benytte
Analyze/Correlate/Bivariate. Man markerer de variable, der
ønskes korrelationer imellem og fører dem over i Variables,
hvorefter man afkrydser den type korrelationer, man ønsker at
udregne, her Pearson og Spearman.



Konfidens- og prediktionsgrænser i SPSS

Slide 45

Den pæne figur s. 45 er **yderst vanskelig** at få frem i SPSS, men man kan tegne de to sæt grænser hver for sig:

Konfidensgrænser:

fås ved at dobbeltklikke på scatterplottet fra tidligere (se s. 3), og benytte Add Fit Line at Total og derefter i Properties-boksen afkrydse Linear og Confidence Intervals/Mean.

Prediktionsgrænser:

fås som ovenfor, blot skal man i stedet vælge Confidence Intervals/Individual

Man kan i begge tilfælde overveje at fjerne fluebenet i Attach label to line



Modelkontrol i SPSS

Slide 49 ff

Nogle enkelte modelkontroltegninger kan fås frem direkte fra regressionsanalysen, ved at klikke på Plots
Betegnelserne er her:

- DEPENDNT: the dependent variable
- *ZPRED Standardized predicted values
- *ADJPRED Adjusted predicted values
- *ZRESID Standardized residuals
- *DRESID Deleted residuals
- *SRESID Studentized residuals
- *SDRESID Studentized deleted residuals

Vælg nu X og Y blandt de ovenstående variable, eller afkryds
Histogram og Normal probability plot.



Modelkontrol, fortsat

Slide 51

For at få fuld kontrol over modelkontrollen, anvender vi Save-knappen, og krydser det af, vi gerne vil bruge, hvorefter de ønskede figurer laves i Graph/ChartBuilder.

Man kan vælge at indlægge en vandret linie i 0 ved at dobbeltklikke på en graf og klikke

Add a reference line to the Y axis, hvorefter der fremkommer en

Properties-boks, hvor man skriver 0 i Position.



Check af varianshomogenitet i SPSS

Slide 54

De normerede residualer hedder SRE_1, og for at definere kvadratroden af de numeriske residualer går vi ind i Transform/Compute, skriver `sqrtres` i Target Variable og `sqrt(abs(SRE_1))` i definitionsboksen.

Herefter fremstilles figurerne som sædvanligt i Graph/Chart Builder, og ved at dobbeltklikke på grafen efterfølgende, kan man lægge udglattede kurver på ved at klikke på Add Fit Line at Total og derefter i Properties-boksen afkrydse Loess/Apply.

Regression diagnostics i SPSS

Slide 66

Vi anvender her igen Save-knappen i regressionsopsætningen, og krydser det af, vi gerne vil bruge, nemlig Cook's og under Influence Statistics: Standardized DfBetas (og evt. også DfBetas(s))

Herefter benyttes Graph/Chart builder/Scatter som tidligere.