

Basal Statistik - SPSS

Multipel regression.

Lene Theil Skovgaard

11. marts 2019

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

- ▶ Figurer: s. 3, 6, 7, 10, 11
- ▶ Regressionsanalyse: s. 4, 8, 12
- ▶ Modelkontrol: s. 5, 6, 7
- ▶ Prediktion: s. 9, 10
- ▶ Diagnostics: s. 11, 12

1 / 12



2 / 12



Tredimensionalt plot

Slide 7

Benyt Graphs/Graphboard Template Choser, marker de 3 involverede variable (vaegt, bpd og ad) og klik på Scatterplot.

For at bytte om på akserne, klikkes på Detailed, hvor man så kan sætte vaegt på Y-aksen, bpd på X-aksen og ad på Z-aksen.

3 / 12



Regressionsanalyse

Slide 9

Den *minimale* opsætning:

- ▶ Benyt Analyze/Regression/Linear
- ▶ Sæt vaegt over i Dependent
- ▶ Sæt både bpd og ad over i Independent(s)

Som minimum bør man dog også klikke Statistics og afkrydse Confidence Intervals.

Adskillige ekstra options kan tilføjes, som det vil fremgå af nogle af de efterfølgende sider (s. 5, 11, 12)

4 / 12



Modelkontrol

Slide 12

Den automatiske modelkontrol i SPSS er ikke særlig god.....

Brug i stedet nedenstående:

I regressionsopsætningen fra s. 4 klikkes på Save, hvorefter man typisk vil vælge

- ▶ Under Predicted Values: vælg Unstandardized, som så kommer til at ligge i datasættet under navnet PRE_1
- ▶ Under Residuals er der 5 muligheder, f.eks.
 - ▶ Unstandardized: De sædvanlige, altså observeret minus forventet, kaldet RES_1 i datasættet
 - ▶ Studentized deleted: Normerede Press-residualer (altså hvor den pågældende observation er udeladt ved fit af modellen), kaldet SDR_1 i datasættet

5 / 12



Modelkontrol, II

Slide 15-17

Ud fra det udbyggede datasæt dannet s. 5 kan vi nu lave diverse figurer:

- ▶ Benyt Graph/Chart Builder, vælg Scatter, klik Unstandardized Predicted Values over på X-aksen og Unstandardized Residuals over på Y-aksen.
- ▶ Vælg Analyze/Descriptive Statistics/Q-Q Plots og sæt Unstandardized Residuals over i Variables.
- ▶ Benyt Graph/Chart Builder, vælg Histogram og klik det ønskede residual (f.eks. RES_1) ned på X-aksen. Klik så på Distribution Curve og afkryds Normal, klik Apply og Close

6 / 12



Check af varianshomogenitet

Slide 18

De standardiserede residualer hedder ZRE_1, og for at definere kvadratroden af de numeriske residualer går vi ind i Transform/Compute, skriver sqrtres i Target Variable og $\sqrt{\text{abs}(\text{ZRE}_1)}$ 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.

7 / 12



Regressionsanalyse på logaritmer

Slide 20

Her skal vi definere logaritmetransformerede variable ved at benytte Transform/Compute, hvor vi først i Target Variable skriver logvaegt og i Numeric Expression skriver $\text{Ln}(\text{vaegt})$

Derefter gør vi det igen, men denne gang sætter vi logad som Target Variable, og i Numeric Expression skriver vi $\text{Ln}(\text{ad}/100)/\text{Ln}(1.1)$

Og endelig en sidste gang, hvor vi sætter logbpd som Target Variable, og i Numeric Expression skriver $\text{Ln}(\text{bpd}/90)/\text{Ln}(1.1)$

8 / 12



Prediktioner til brug for illustrationer

Slide 29

- ▶ Tilføj fiktive personer til datasættet, med manglende værdier af logvaegt, men med de værdier af logbpd og logad (dvs. af bpd og ad), som man gerne vil lave prediktioner for (her 3 værdier af bpd og adskillige af ad)
- ▶ Lav en kolonne, kaldet f.eks. ny, som skal have værdien 0 for alle de oprindelige observationer, men 1 for de nye fiktive
- ▶ Estimer nu i samme model som før i det udbyggede datasæt og gem de predikterede værdier (nu hedder de sikkert PRE_2, fordi vi har gemt nogen tidligere)
- ▶ Transformer de predikterede værdier tilbage til oprindelig skala ved at benytte Transform/Compute Variable, skriv predikteret i Target Variable og exp(PRE_2) i Numeric Expression

9 / 12



Prediktioner til brug for illustrationer, II

Slide 30

Selve figurerne dannes således:

- ▶ Begræns data til de fiktive personer ved at gå ind i Data/Select Cases, klik i If og skriv ny=1
- ▶ **Venstre figur:**
Gå ind i Graph/Chart Builder, vælg Lines (det til højre, der kan opdeles efter gruppe), og i den fremkomne boks trækkes PRE_2 over på Y-aksen, logad over på X-aksen og bpd over i Set Color
- ▶ **Højre figur:**
Som venstre, blot med predikteret på Y-aksen i stedet for PRE_2, og ad på X-aksen i stedet for logad

10 / 12



Regression diagnostics

Slide 32

Vi anvender her igen Save-knappen i regressionsopsætningen (se s. 4 og evt. 5), 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.

11 / 12



Variabelselektion og flere options

Slide 52

I regressionsopsætningen (se s. 4) klikkes på Method, der nu kan skiftes fra Enter ("*den kontrollerede og anbefalede metode*") til Backward eller Forward

Slide 66

I regressionsopsætningen (se s. 4) klikkes på Statistics, hvor man kan vælge Collinearity Diagnostics

12 / 12

