

# Analysemodeller

## 1 Lineær regression (kap. 12)

- Ofte scatter plot
- Ofte indgår ordet prædikere under delopgaverne!
- Middelværdi og standard afvigelse er ofte angivet.
- Når der er sammenhæng mellem 2 variabler, f.eks: alder og kolesterol, så er alder den uafhængige (X) og kolesterol den afhængige. Eller sammenhæng mellem forældres højde (uafhængig) og børnenes højde (afhængig).
  - Altid  $H_0: \beta_1 = 0$  og  $H_a: \beta_1 \neq 0$
  - $df = n - 2$  (n er antal individer)
  - tabel 4!
- FORUDSÆTNINGER:
  - SKAL VÆRE TILFÆLDIGE og uafhængige observationer.
  - Sammenhængen skal være lineær, som kan ses ud fra scatter plot.
  - Y skal være normalt fordelt udfra hver x, som kan ses ud fra QQ-plot.
  - Varianserne skal være ens, som kan ses ud fra plottet med residualer.

## 2 Anova / F-test (s. 426)

- Analyse
- Middelværdi skal være angivet.
  - når der er mere end 2 middelværdier, som skal sammenlignes med hinanden.
  - Ofte boxplot i starten.
    - Altid!  $H_0: \mu_1 = \mu_2 = \mu_3$  (alt efter hvor mange grupper)
    - Altid!  $H_a: \mu_i \neq \mu_j$  for mindst 2.
    - Df between:  $i - 1$  (i er antal grupper og n antal individer)
    - Df within:  $n - i$
    - Tabel 10!
  - FORUDSÆTNINGER:
    - Skal være uafhængige observationer (som kan antages ud fra indledning.)
    - Der skal være normal fordeling, og det kan ses ud fra normalfordelingsplot.
    - Varianserne skal være ens for alle grupper (ud fra heavens test, som plejer at være opgivet, så kigger man på værdien som skal være større end 0.1)

## 3 T-test for 2 uafhængige stikprøver (kap 7)

- Middelværdi skal være angivet.
- Når der kun er 2 middelværdier, som skal sammenlignes med hinanden.
  - $H_0: \mu_1 = \mu_2$
  - $H_a: \mu_1 \neq \mu_2$
  - $df = n_1 + n_2 - 2$
  - tabel 4!

- FORUDSÆTNINGER:
  - SKAL VÆRE TILFÆLDIGE og uafhængige observationer.
  - Varianserne skal være ens. Man regner dem ud, og sammenligner med hinanden, og der skal være mindre end faktor 4 til forskel.
  - D-variablen skal være normalfordelt. Må antages når der ikke er QQ-plot.

#### 4 Goodness of fit, chi square ( kap 9, formel på s. 351)

- Tal med f.eks procenter, farver : f.eks: eksemplet opgaven med "fisk"(en række) .
- Ingen middelværdi.
  - $H_0$ : altid det omvendte af indledningen (se opg. 3 i juni 2010 sættet)
  - $H_a$ : altid det omvendte af  $H_0$ .
  - $df = \text{ANTAL KATEGORIER eller grupper} - 1$
  - Tabel 9!
- FORUDSÆTNINGER:
  - SKAL VÆRE TILFÆLDIGE og uafhængige observationer.
  - Forventet antal skal ALTID være større end 5.

#### 5 Chi square uafhængighed – krydstabel ( formel findes på hhv. s. 367 og s. 351)

- Ofte er der en tabel med antal kolonner og rækker.
- Ingen middelværdi
- $H_0$ : altid det omvendte af indledningen.
- $H_a$ : altid det omvendte af  $H_0$ .
- $Df$  altid 1 når der er tale om 2 x 2 tabel, mens når det er andet (f.eks 2 x 3 tabel) så gælder følgende:  $df = (\text{rækker} - 1) * (\text{kolonner} - 1)$
- Tabel 9!
- Forudsætninger:
  - SKAL VÆRE TILFÆLDIGE og uafhængige observationer.
  - Forventet antal skal ALTID være større end 5

#### Parret t-test

- Ofte står der i indledningen eller opgaveformuleringen at man skal finde en differens (forskell) mellem 2 middelværdier, så 2 middelværdier skal være angivet.
  - F.eks: En før of efter måling (differens)
  - Eller en måling mellem 2 tvillinger (2 personer), hvor differensen af middelværdierne skal findes.
  - Differencen i middelværdi = 0 ( $\mu_{\text{diff}} = 0$ )
  - Differencen i middelværdi  $\neq 0$  ( $\mu_{\text{diff}} \neq 0$ )
  - $Df = n_d - 1$
  - Tabel 4!
- Forudsætninger:
  - SKAL VÆRE TILFÆLDIGE og uafhængige par.
  - Differencen skal være normalt fordelt, og det kan ses udefra QQ-plot for begge personer.

Middelværdi i Population er  $\mu$ , og standard afvigelse er  $\sigma$ . Hælden kaldes  $\beta$ .

Middelværdi i stikprøver er  $\bar{y}$ , og standard afvigelse i stikprøve er  $s$ . Hælden her kaldes  $b_1$ .

## august 2010:

- Opg 1: Lineær regression
- OPG 2: ANOVA
- OPG3: Goodness of fit Chi square og krydstabel.

## juni 2010

- Opg 1: Lineær regression
- Opg 2: Parret t-test og t-test for 2 uafhængige stikprøver.
- OPG 3: Goodness of fit Chi square og krydstabel.

## august 2009

- OPG 1: Parret t-test.
- OPG 2: Lineær regression
- OPG 3: Krydstabel og goodness of fit Chi square.

## juni 2009

- OPG 1: Anova
- OPG 2: Lineær regression
- OPG 3: Goodness of fit Chi square og krydstabel.

## januar 2009

- OPG 1: Parret t-test
- OPG 2: ANOVA
- OPG 3: Chi krydstabel og Goodness of fitt chi square.

## 2008

- OPG 1: ANOVA
- OPG 2: Lineær regression
- OPG 3: Goodness of fit Chi square og krydstabel.

## 2007

- OPG 1: Goodness of fit Chi square og krydstabel.
- OPG 2: Lineær regression
- OPG 3: ANOVA

## 2006

- OPG 1: ANOVA
- OPG 2: Chi krydstabel og Goodness of fit chi square.
- OPG 3: Lineær regression