

Kursus 02403

Introduktion til Statistik

Forelæsning 13: Summary

Jan Kloppenborg Møller

DTU Informatik

Bygning 305 - rum 219

Danmarks Tekniske Universitet

2800 Lyngby – Danmark

e-mail: jkm@imm.dtu.dk

Oversigt

- 1 Kapitel 2: Simpel grafik og nøgletal
- 2 Kapitel 4: Diskrete fordelinger
- 3 Kapitel 5: Kontinuerte fordelinger
- 4 Kapitel 6: Stikprøvefordelinger
- 5 Kapitel 7 og 8: Een og to stikprøver
- 6 Kapitel 9, Inferens for varianser
- 7 Kapitel 10, Inferens for andele
- 8 R-note , Statistik ved simulering
- 9 Kapitel 11, Regressionsanalyse
- 10 Kapitel 12, Variansanalyse

Oversigt

- 1 Kapitel 2: Simpel grafik og nøgletal
- 2 Kapitel 4: Diskrete fordelinger
- 3 Kapitel 5: Kontinuerte fordelinger
- 4 Kapitel 6: Stikprøvefordelinger
- 5 Kapitel 7 og 8: Een og to stikprøver
- 6 Kapitel 9, Inferens for varianser
- 7 Kapitel 10, Inferens for andele
- 8 R-note , Statistik ved simulering
- 9 Kapitel 11, Regressionsanalyse
- 10 Kapitel 12, Variansanalyse

Kapitel 2: Simpel grafik og nøgletal

- Kigge på data som de er! (Eksplorativt)
 - Nøgletal
 - Gennemsnit $\bar{x}$
 - spredning s , varians s^2
 - Median, øvre- og nedre kvartil.
- Simpel grafik
 - Scatter plot (xy plot)
 - Histogram, Kumulativ fordeling
 - Boxplots, Bar charts, Pie charts

Oversigt

- 1 Kapitel 2: Simpel grafik og nøgletal
- 2 **Kapitel 4: Diskrete fordelinger**
- 3 Kapitel 5: Kontinuerte fordelinger
- 4 Kapitel 6: Stikprøvefordelinger
- 5 Kapitel 7 og 8: Een og to stikprøver
- 6 Kapitel 9, Inferens for varianser
- 7 Kapitel 10, Inferens for andele
- 8 R-note , Statistik ved simulering
- 9 Kapitel 11, Regressionsanalyse
- 10 Kapitel 12, Variansanalyse

Kapitel 4: Diskrete fordelinger

- Generelle begreber
 - Definition af en stokastisk variabel
 - Tæthedsfunktion
 - Fordelingsfunktion
- Specifikke fordelinger
 - Binomial fordeling
 - Hypergeometrisk fordeling
 - Poisson fordeling

Oversigt

- 1 Kapitel 2: Simpel grafik og nøgletal
- 2 Kapitel 4: Diskrete fordelinger
- 3 Kapitel 5: Kontinuerte fordelinger**
- 4 Kapitel 6: Stikprøvefordelinger
- 5 Kapitel 7 og 8: Een og to stikprøver
- 6 Kapitel 9, Inferens for varianser
- 7 Kapitel 10, Inferens for andele
- 8 R-note , Statistik ved simulering
- 9 Kapitel 11, Regressionsanalyse
- 10 Kapitel 12, Variansanalyse

Kapitel 5: Kontinuerte fordelinger

- Generelle begreber:
 - Tæthedsfunktion, Fordelingsfunktion
 - Middelværdi, Varians
 - Regneregler for stokastiske variable
- Specifikke fordelinger:
 - Normal
 - Log-Normal, Uniform, Eksponential
- "Datametode:"
 - Normalplot

Oversigt

- 1 Kapitel 2: Simpel grafik og nøgletal
- 2 Kapitel 4: Diskrete fordelinger
- 3 Kapitel 5: Kontinuerte fordelinger
- 4 Kapitel 6: Stikprøvefordelinger**
- 5 Kapitel 7 og 8: Een og to stikprøver
- 6 Kapitel 9, Inferens for varianser
- 7 Kapitel 10, Inferens for andele
- 8 R-note , Statistik ved simulering
- 9 Kapitel 11, Regressionsanalyse
- 10 Kapitel 12, Variansanalyse

Kapitel 6: Stikprøvefordelinger

- Generelle begreber:
 - Population og tilfældig stikprøve
 - Statistisk inferens
 - Central Grænseværdisætning
 - Link mellem matematiske modeller og observerede data
- Specifikke stikprøvefordelinger: (Een og to stikprøver)
 - "z-fordelingen" (=Normalfordelingen)
 - t-fordelingen
 - χ^2 -fordelingen, F-fordelingen

Oversigt

- 1 Kapitel 2: Simpel grafik og nøgletal
- 2 Kapitel 4: Diskrete fordelinger
- 3 Kapitel 5: Kontinuerte fordelinger
- 4 Kapitel 6: Stikprøvefordelinger
- 5 Kapitel 7 og 8: Een og to stikprøver**
- 6 Kapitel 9, Inferens for varianser
- 7 Kapitel 10, Inferens for andele
- 8 R-note , Statistik ved simulering
- 9 Kapitel 11, Regressionsanalyse
- 10 Kapitel 12, Variansanalyse

Kapitel 7 og 8: Een og to stikprøver

- Generelle begreber:
 - Estimation, Konfidensinterval
 - Hypotesetest, P-værdi, Type I og Type II fejl
- Specifikke analysemetoder, Én stikprøve:
 - Test og konfidensinterval for gennemsnit (t -test)
 - Planlægning: beregning af stikprøvestørrelse
- Specifikke analysemetoder, To UAFHÆNGIGE stikprøver:
 - Test og konfidensinterval for gennemsnitsforskel
- Specifikke analysemetoder, To PARREDE stikprøver:
 - "Tag differens" $\Rightarrow$ "Én stikprøve"

Oversigt

- 1 Kapitel 2: Simpel grafik og nøgletal
- 2 Kapitel 4: Diskrete fordelinger
- 3 Kapitel 5: Kontinuerte fordelinger
- 4 Kapitel 6: Stikprøvefordelinger
- 5 Kapitel 7 og 8: Een og to stikprøver
- 6 Kapitel 9, Inferens for varianser**
- 7 Kapitel 10, Inferens for andele
- 8 R-note , Statistik ved simulering
- 9 Kapitel 11, Regressionsanalyse
- 10 Kapitel 12, Variansanalyse

Kapitel 9, Inferens for varianser

- Specifikke analysemetoder, Én og to stikprøver
- Konfidensinterval for én varians
- Hypotesetest for én varians
- Hypotesetest for to varianser

Oversigt

- 1 Kapitel 2: Simpel grafik og nøgletal
- 2 Kapitel 4: Diskrete fordelinger
- 3 Kapitel 5: Kontinuerte fordelinger
- 4 Kapitel 6: Stikprøvefordelinger
- 5 Kapitel 7 og 8: Een og to stikprøver
- 6 Kapitel 9, Inferens for varianser
- 7 Kapitel 10, Inferens for andele**
- 8 R-note , Statistik ved simulering
- 9 Kapitel 11, Regressionsanalyse
- 10 Kapitel 12, Variansanalyse

Kapitel 10, Inferens for andele

- Specifikke analysemetoder, Én, to og $k > 2$ stikprøver
 - Binær/kategorisk respons
- Estimation af andele
- Hypotesetest for én andel
- Hypotesetest for to andele
- Analyse af antalstabeller

Oversigt

- 1 Kapitel 2: Simpel grafik og nøgletal
- 2 Kapitel 4: Diskrete fordelinger
- 3 Kapitel 5: Kontinuerte fordelinger
- 4 Kapitel 6: Stikprøvefordelinger
- 5 Kapitel 7 og 8: Een og to stikprøver
- 6 Kapitel 9, Inferens for varianser
- 7 Kapitel 10, Inferens for andele
- 8 R-note , Statistik ved simulering**
- 9 Kapitel 11, Regressionsanalyse
- 10 Kapitel 12, Variansanalyse

R-note , Statistik ved simulering

- Introduktion til simulering
- Fejlophobningslove
- Konfidensintervaller ved hjælp af simulering:
bootstrapping
- Hypotesetest ved hjælp af simulering
 - bootstrap konfidensintervaller
 - ved hjælp af permutationstest
- Konkrete situationer:
 - One-sample data
 - Two-sample data

Oversigt

- 1 Kapitel 2: Simpel grafik og nøgletal
- 2 Kapitel 4: Diskrete fordelinger
- 3 Kapitel 5: Kontinuerte fordelinger
- 4 Kapitel 6: Stikprøvefordelinger
- 5 Kapitel 7 og 8: Een og to stikprøver
- 6 Kapitel 9, Inferens for varianser
- 7 Kapitel 10, Inferens for andele
- 8 R-note , Statistik ved simulering
- 9 Kapitel 11, Regressionsanalyse**
- 10 Kapitel 12, Variansanalyse

Kapitel 11, Regressionsanalyse

- To kvantitative variable, x og y .
- Beregning af bedste rette linie (least squares)
- Inferens i en simpel lineær regressionsmodel
 - Statistisk model: $y_i = \alpha + \beta x_i + \varepsilon_i$
 - Intervalestimation og test for α og β .
 - Konfidensinterval for den forventede linie.
 - Prædiktions-interval.
- Korrelation r
 - r udtrykker graden af lineær sammenhæng.
 - r^2 udtrykker "forklaringsgraden".

Oversigt

- 1 Kapitel 2: Simpel grafik og nøgletal
- 2 Kapitel 4: Diskrete fordelinger
- 3 Kapitel 5: Kontinuerte fordelinger
- 4 Kapitel 6: Stikprøvefordelinger
- 5 Kapitel 7 og 8: Een og to stikprøver
- 6 Kapitel 9, Inferens for varianser
- 7 Kapitel 10, Inferens for andele
- 8 R-note , Statistik ved simulering
- 9 Kapitel 11, Regressionsanalyse
- 10 Kapitel 12, Variansanalyse**

Kapitel 12, Variansanalyse

- Specifikke analysemetoder, K UAFHÆNGIGE stikprøver
- Envejs variansanalyse
 - Tester forskel på gruppernes gennemsnit
 - ANOVA-tabel: $SST = SS(Tr) + SSE$
 - F -test.
 - Post hoc test: parvise t -test
- Specifikke analysemetoder, K AFHÆNGIGE stikprøver
- Tovejs variansanalyse
 - SST , $SS(Tr)$ og $SS(Bl)$ beregnes som for énsidet
 - $SSE = SST - SS(Tr) - SS(Bl)$

Oversigt

- 1 Kapitel 2: Simpel grafik og nøgletal
- 2 Kapitel 4: Diskrete fordelinger
- 3 Kapitel 5: Kontinuerte fordelinger
- 4 Kapitel 6: Stikprøvefordelinger
- 5 Kapitel 7 og 8: Een og to stikprøver
- 6 Kapitel 9, Inferens for varianser
- 7 Kapitel 10, Inferens for andele
- 8 R-note , Statistik ved simulering
- 9 Kapitel 11, Regressionsanalyse
- 10 Kapitel 12, Variansanalyse