Kapitel 9
Test for normalitet

Peter Tibert Stoltze
stat@peterstoltze.dk

Elementær statistik
F2011

Indledning

- Vi antager ofte, at data kan beskrives ved en normalfordeling — nu skal vi kontrollere om denne antagelse er rimelig
- To muligheder for at sammenligne en empirisk med en teoretiske fordeling:
  - χ^2-test for *goodness-of-fit*
  - Kolmogorov-Smirnov testet
- Begge tests kan også benyttes for andre fordelingstyper

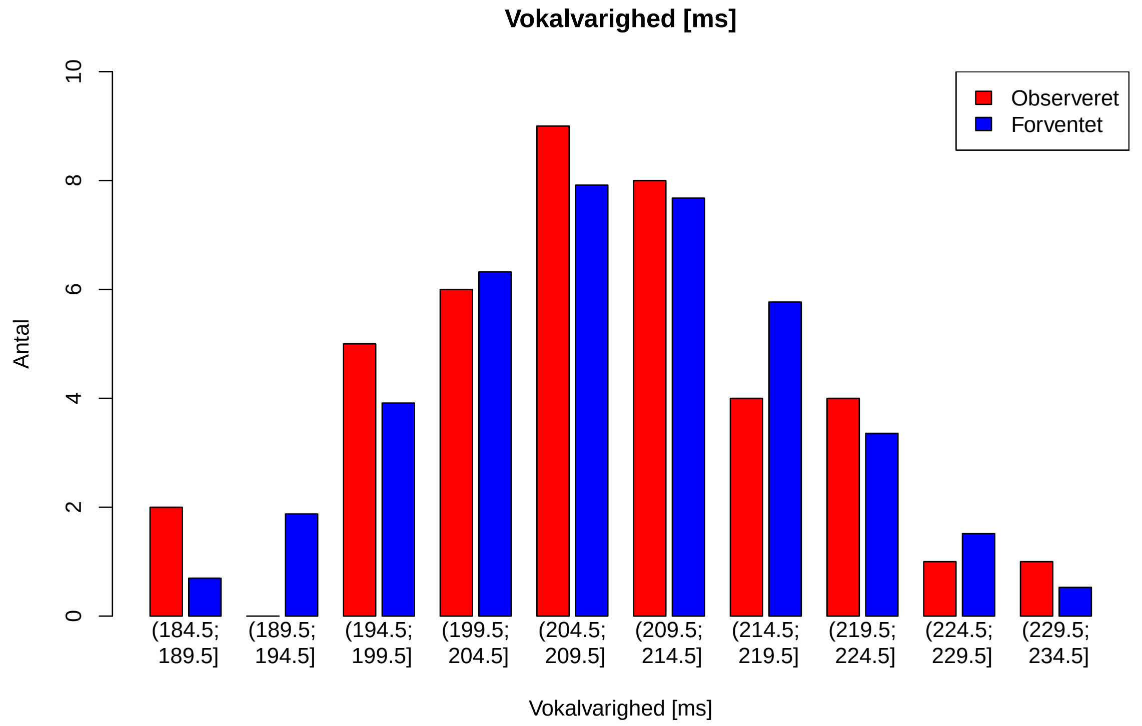
χ^2-test for *goodness-of-fit*

- Vi har tidligere lavet grafisk modelkontrol ved at beregne et forventet antal observationer under antagelse af normalitet (altså at stikprøven stammer fra en normalfordelt population)
- Vi har også tidligere set på χ^2 testet for *goodness-of-fit*, hvor et forventet antal observationer blev sammenlignet med et forventet antal observationer
- Nu kombinerer vi disse to ting...

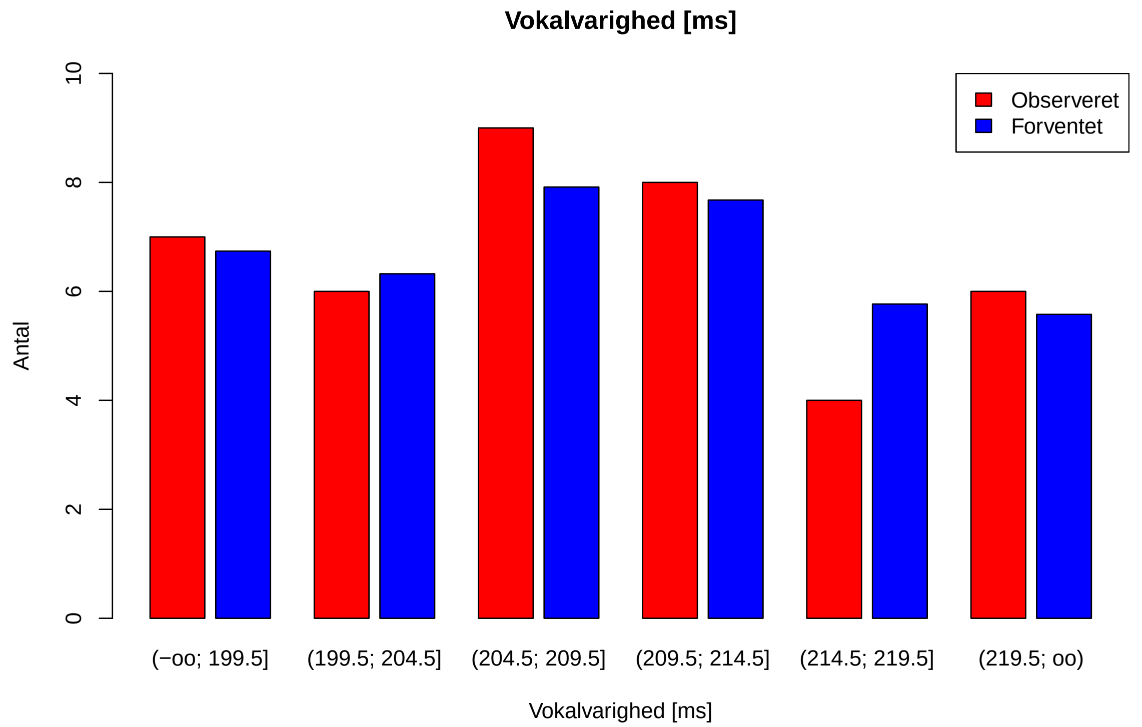
χ^2-test for *goodness-of-fit*

- Beregningerne er præcist som tidligere — udfordringen er nu at finde en passende klassebredde
- Er der for mange klasser bliver de for tynde, og er der for få mister vi for megen information
- Vi starter med at lave klasserne tilpas små, hvorefter vi kan lave sammenlægning efter behov
- Opstilling i regneark giver gode muligheder for at forsøge med forskellige inddelinger uden alt for meget arbejde

Grafisk sammenligning



Grafisk sammenligning — efter sammenslåning



Beregning af teststørrelse og beregning af p

- Nulhypotesen er, at stikprøven stammer fra en population med fordeling (type og parametre) svarende til dem, vi anvendte for at beregne forventede antal
- Alternativhypotesen er, at stikprøven stammer fra en population med en anden fordeling
- Teststørrelsen er den sædvanlige $\sum(O-E)^2/E$
- Der sammenlignes med en χ^2 fordeling med frihedsgrader svarende til antal klasser minus én

Eksempel med vokalvarighed

Klasse	(-∞; 199,5]	(199,5; 204,5]	(204,5; 209,5]	(209,5; 214,5]	(214,5; 219,5]	(219,5; ∞)
O	7	6	9	8	4	6
E	6,74	6,32	7,92	7,68	5,77	5,58
$(O-E)^2/E$	0,01	0,02	0,15	0,01	0,55	0,03

- I eksemplet beregnes teststørrelsen til 0,763
- Kritiske værdier (med $df = 6-1$) er 11,07 ($p = 0,05$) og 15,09 ($p = 0,01$)
- Vi er slet ikke i nærheden af at kunne afvise nulhypotesen — meget tyder altså på, at stikprøven er trukket fra en normalfordelt population

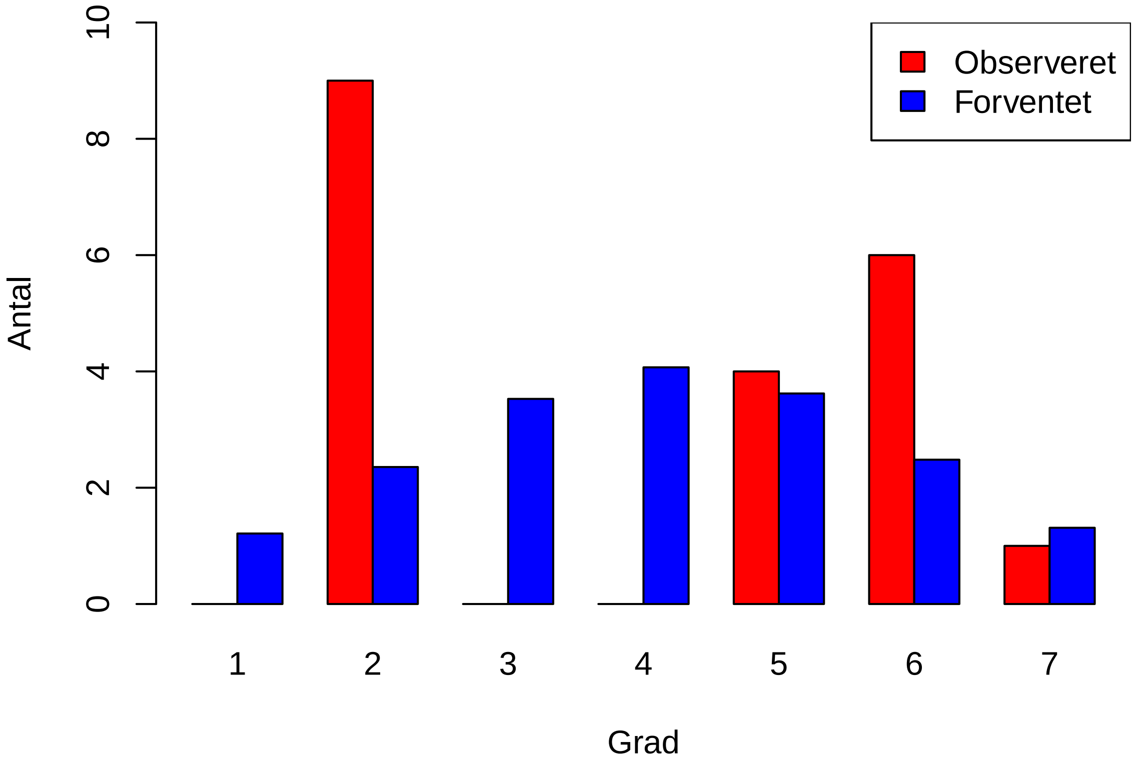
Kolmogorov-Smirnov testet

- Vi danner den kumulative frekvensfordeling for vores stikprøve, og den skaleres herefter til en empirisk fordelingsfunktion
- Vi beregner også fordelingsfunktionen for en normalfordeling med parametre svarende til det, som vi har estimeret for stikprøven
- Vi beregner den største absolutte forskel mellem de to funktioner og sammenligner med en tabel...

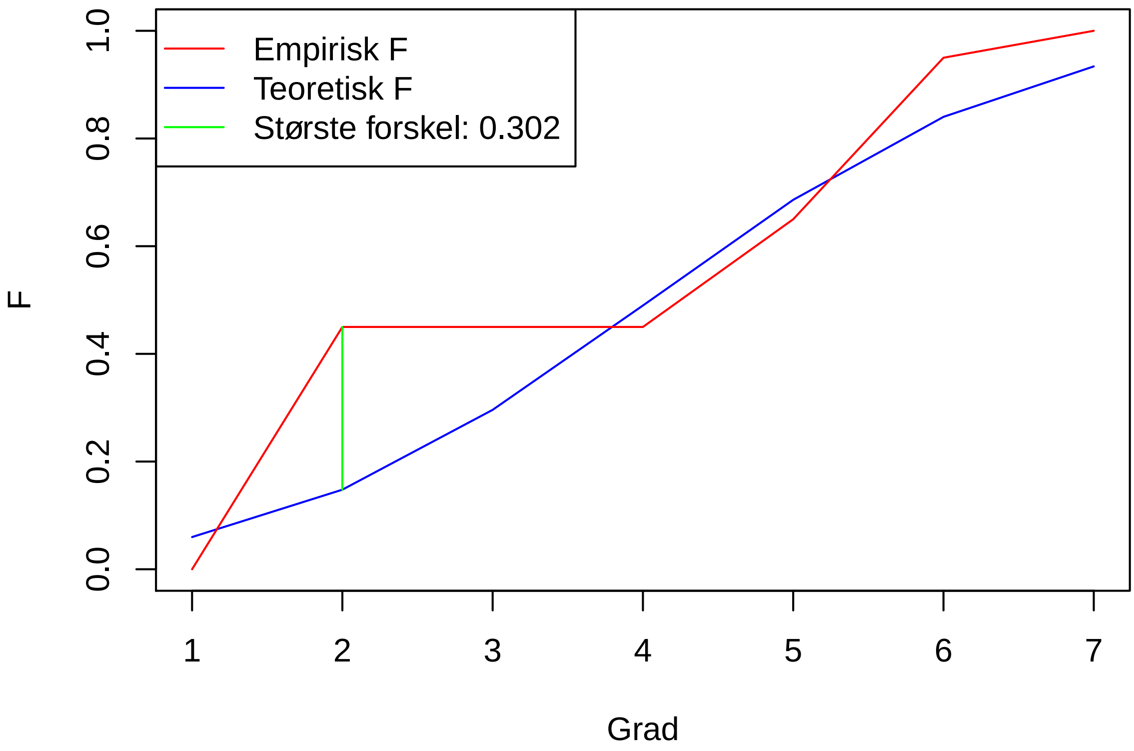
Tabel 9.3

	Observeret			Forventet	
Grad	f_e	F_e	F_e/n	z	F_{teo}
1	0	0	0,000	-1,557	0,060
2	9	9	0,450	-1,046	0,148
3	0	9	0,450	-0,536	0,296
4	0	9	0,450	-0,026	0,490
5	4	13	0,650	0,485	0,686
6	6	19	0,950	0,995	0,840
7	1	20	1,000	1,506	0,934

Sammenligning af histogrammer



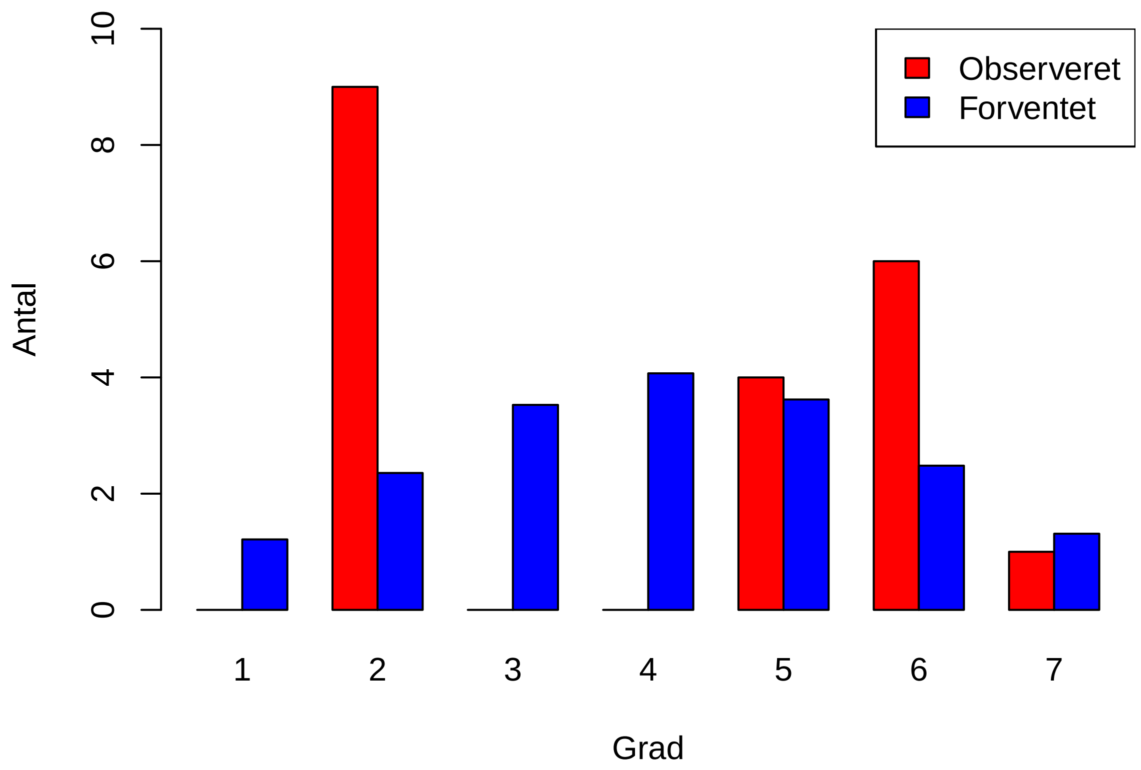
Sammenligning af fordelingsfunktioner



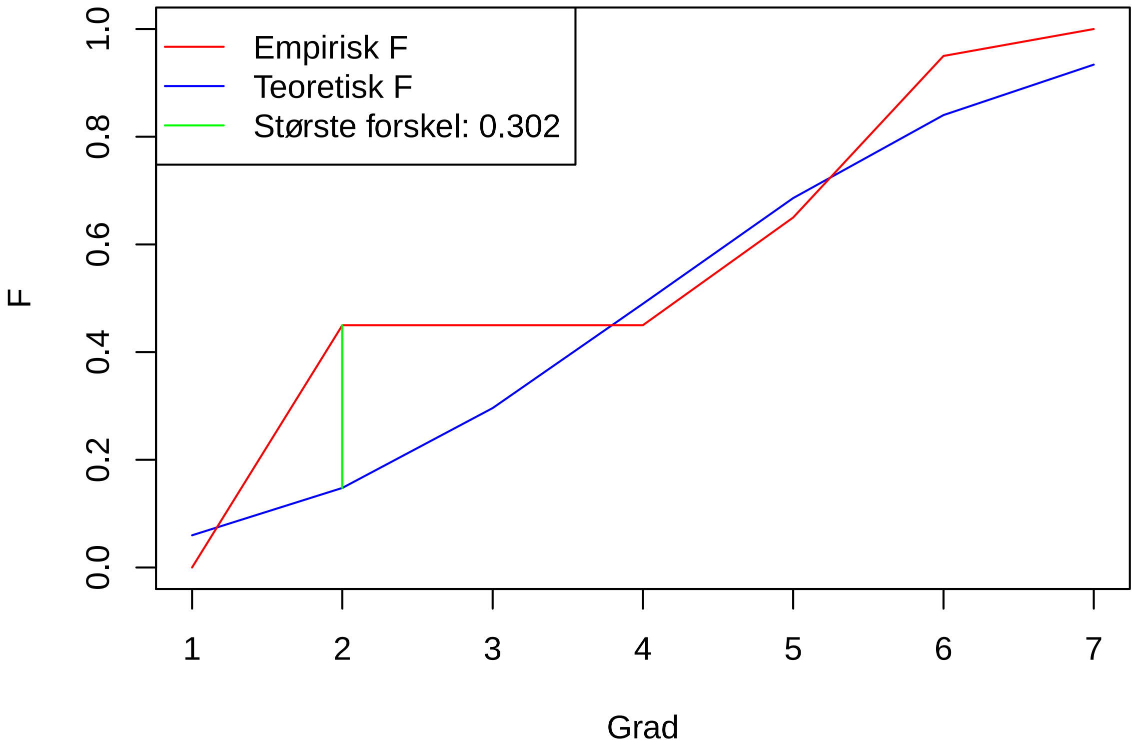
Hypoteser og teststørrelse

- Hypoteser for testet er
  - H_0: Stikprøven stammer fra en population, der er normalfordelt med middelværdi 4,05 og spredning 1,96
  - H_1: Stikprøven stammer fra en anden population
- Vi finder $D_{max} = 0,302$ og kritiske værdier til 0,294 ved 5% og 0,356 ved 1% ($n = 20$)
- Vi forkaster nulhypotesen: Det er ikke ret sandsynligt ($0,01 < p < 0,05$) at data stammer fra en population med den anførte fordeling
- Men testet er følsomt — hvad sker der hvis en enkelt af logopæderne ændrer graden fra 2 til 3?

Sammenligning af histogrammer



Sammenligning af fordelingsfunktioner



En enkelt ændring...

- Da en enkelt af logopæderne ændrer graden fra 2 til 3 ændres D_{max} til 0,252
- Testet ikke længere signifikant, og vi kan ikke forkaste nulhypotesen
- Testet er ganske følsomt, så man skal formulere sine konklusioner lidt forsigtigt når det gælder små stikprøver...