

3.36pt

Hvorfor specielle metoder for varighedsdata ?

- ▶ varigheder/tider er positive og oftest højreskæve
- ▶ data næsten altid *censurerede* (vi kan ikke observere for evigt)
- ▶ - dvs. *ufuldstændig/manglende* observation

Nye begreber !

- ▶ Overlevelsesfunktion
- ▶ Hazard rate/funktion samt kumuleret hazard
- ▶ middel resterende levetid/mean residual life time
- ▶ forskellige typer af censureringer samt antagelser vedr. disse

Specielt hazard rate og mean residual life time giver kun praktisk mening i relation til varighedsdata (betingelser med overlevelse op til givet tidspunkt).

Overlevelsesfunktionen blot 1 minus fordelingsfunktionen.

Nye statistiske metoder: ikke-parametrisk

- ▶ Aktuar estimatet af overlevelsesfunktionen (grupperede data)
- ▶ Kaplan-Meier estimat af overlevelsesfunktionen (grænseværdi af aktuar-estimatet)
- ▶ Nelson-Aalen estimat af kumuleret hazard
- ▶ Greenwoods estimat af usikkerhed på Kaplan-Meier/Nelson-Aalen og konfidensintervaller
- ▶ log rank testet (for ens hazard i to grupper)

Nye metoder: Cox's proportional hazards model

- ▶ semi-parametrisk: kombinerer fordele ved ikke-parametriske og parametriske metoder
- ▶ fleksibel pga. helt uspecificeret reference/baseline hazard funktion h_0
- ▶ likelihood-baseret estimation ikke mulig pga. uspecificeret reference hazard
- ▶ *partial* likelihood: eliminerer behovet for at kende h_0
- ▶ forskellige udledninger af partial likelihood: Cox's faktoriseringsargument, likelihood baseret på range (Kalbfleisch og Prentice), profil-likelihood (S. Johansen)
- ▶ håndtering af sammenfaldende tider (ties)

Modelcheck for Cox's proportional hazard model

- ▶ Andersen plots (cumulative hazards, stratified Cox ph)
- ▶ Martingale residuals (scoring af variable)
- ▶ Cox-Snell
- ▶ Deviance
- ▶ Score process residuals (giver dfbetas - influential observations)
- ▶ Schoenfeld residuals (time-varying effects ?)

Udvidelser af Cox's proportional hazards

- ▶ Tidsvarierende forklarende variable
- ▶ Random effects - frailty modeller

Parametrisk estimation

- ▶ minder meget om sædvanlig statistik baseret på lineære/generaliserede lineære modeller
- ▶ dog særligt udtryk for likelihood-funktionen pga. censurerede observationer

Tælleprocesser

- ▶ teoretiske grundlag for asymptotiske resultater for Kaplan-Meier, Nelson-Aalen, partial likelihood.
- ▶ data formuleres som tælleproces, derudfra dannes martingal og endelige benyttes martingal convergence theorem
- ▶ nyttig ramme for inkludering af tidsvarierende variable
- ▶ nyttig ramme for håndtering af gentagne hændelser (ikke-dødelige hændelser for den samme person)
- ▶ nyttig ramme for komplekse observations-mønstre (andet end: i risiko \rightarrow ikke i risiko)

Hvad har vi (bla.) ikke dækket ?

- ▶ venstre- og interval censurede data
- ▶ additive hazard modeller
- ▶ gentagne hændelser

Gik endvidere ret hurtigt hen over

- ▶ frailty-modeller/modeller for korrelerede data
- ▶ tælleprocesser: kunne bruge et helt kursus til at dække teorien bag tælleprocesser
- ▶ competing risks

Overordnet set har I fået indblik i de vigtigste aspekter/modeller/metoder vedr. varigheds/overlevelsesdata.