

Lambda-koefficienten med -lambda-

[2. udgave, marts 2013]

Kim Mannemar Sønderskov

Institut for Statskundskab, Aarhus Universitet
ks@ps.au.dk

Goodman and Kruskals lambda kan bruges som sammenhængsmål i bivariante situationer, hvor mindst én af variablene er på nominalskala og har mere end to kategorier¹ (Babbie, 1995: 417-18). I Stata har man ikke umiddelbart mulighed for at beregne lambda, men ved at installere et såkaldt *user written program*, kan lambda nemt beregnes. User written programs beskrives ikke nærmere her, men se evt. afsnit 8.3 i Sønderskov (2011).

For at installere lambda-programpakken i Stata skal du skrive følgende i kommandolinjen (kræver internetadgang):

```
. *installerer programpakken lambda  
. net install lambda, from(http://fmwww.bc.edu/RePEc/bocode/l)
```

Dette vil installere programmet (med mindre det allerede er installeret), og man vil kunne bruge det som enhver anden Stata-kommando.

Lambda-koefficienten for sammenhængen mellem to variable, *rækkevar* og *kolonnevar* (hvor *rækkevar* indikerer at variabelen behandles som rækkevariabel i en krydstabulering), beregnes med følgende kommando:

```
. *lambda-koefficient for sammenhængen mellem rækkevar og kolonnevar  
. lambda rækkevar kolonnevar
```

Kommandoen producerer en krydstabel for de to variable og tre lambda-mål: **lambda_a**, **lambda_b** samt **lambda**. **lambda_a** er lambda-koefficienten, hvis *rækkevar* antages at være den afhængige variabel, **lambda_b** er lambda-koefficienten, hvis *kolonnevar* antages at være den afhængige variabel, og **lambda** er den symmetriske lambda-koefficient.

Et eksempel: Boligtype og stemmeadfærd

I følgende eksempel analyseres sammenhængen mellem boligtype og vælgeradfærd med lambda og surveydata (Den danske valgundersøgelse fra 2005). Konkret undersøges det, om der er forskel i tilbøjeligheden på at støtte partier fra 'blå' blok ved det dengang netop overståede folketingsvalg for respondenter, der bor i hhv. lejet bolig, andelsbolig eller ejerbolig (hvor sidstnævnte er opdelt i hhv.

¹ Bemærk at diverse former for regressionsanalyse vil være alternativer til lambda i den nævnte situation – jf. Tabel 6.1 i Sønderskov (2011). Bemærk også, at lambda kan bruges i andre end de nævnte tilfælde, men informationsværdien i målet er ikke videre stor, hvorfor man vil bruge andre mål når det er muligt.

hus/gård eller lejlighed). Boligtype er tydeligvis på nominalskala og behandles som uafhængig variabel.

1a: Installer lambda-programpakken.

1b: Åbn datasættet 'Valg05Kap5.dta', der kan hentes fra <http://ps.au.dk/soenderskov/stata/>.

1c: Inspicer variablene `blå` og `v335` med kommandoen `codebook` (se evt. afsnit 2.2.2 i Sønderskov 2011 om codebook).

`blå` er en dikotom variabel, der antager værdien 1, hvis respondenterne stemte på blå blok (`blå` er altså en dummyvariabel) mens `v6` er en nominalskalavariabel med information om respondentens boligforhold.

2: Brug `tabulate` til at lave en krydstabulering mellem de to variable. Procentuér på den uafhængige variabel (se evt. afsnit 6.1 i Sønderskov, 2011).

Der synes at være en sammenhæng mellem de to variable, idet der for næsten alle boligtyper er forskel i andelen, der stemte på blå blok. Sammenhængen synes endvidere intuitiv – hvis man eksempelvis kigger på fordelingen indenfor lejerboliger, hvor hovedparten af respondenterne stemte på rød blok.²

3: Brug `lambda` til at udregne sammenhængens lambda-værdi.

Løsninger

*Opgave 1

```
net install lambda, from(http://fmwww.bc.edu/RePEc/bocode/l) //a
use Valg05Kap5.dta //b
codebook blå v335 //c
```

*Opgave 2

```
tabulate blå v335, column
```

*Opgave 3

```
lambda blå v335
```

Referencer

Babbie, E. (1995). *The Practice of Social Research*, 7. ed. Belmont: Wadsworth.

Sønderskov, K.M. (2011). *Stata – En praktisk introduktion*. København: Hans Reitzels.

² Bemærk, at man blandt andet med en *chi2-test* kan undersøge om der er statistisk uafhængighed mellem variablene – jf. afsnit 6.5 i Sønderskov (2011).