

OPGAVE 2

I bogen “En benchmarking af arbejdsløshedskasserne”, som blev udgivet af Direktoratet for Arbejdsløshedsforsikringen i efteråret 1999, undersøgtes blandt meget andet antallet af marginaliserede medlemmer i de enkelte A-kasser. Et marginaliseret medlem defineres som et medlem, der inden for de sidste tre år højst har haft ordinær beskæftigelse 20% af tiden.

I opgaven analyseres for de i alt $n = 36$ A-kasser i Danmark antallet af marginaliserede medlemmer som funktion af antal ledighedsberørte medlemmer, dvs. medlemmer der inden for de sidste tre år har haft mindst én ledighedsperiode. Som en model for antal marginaliserede, y_i , ud af x_i ledighedsberørte medlemmer i den i 'te A-kasse, anvendes Poissonfordelingen.

Imidlertid kan man forestille sig, at andelen af marginaliserede medlemmer stiger med antallet af ledighedsberørte medlemmer, fordi en A-kasse med mange ledighedsberørte har færre ressourcer til rådighed for aktivering etc. En model for dette kunne være, at intensiteten er λx_i^{γ} , således at punktsandsynligheden for alle $n = 36$ A-kasser er

MODEL 1

$$\prod_{i=1}^n \frac{(\lambda x_i^{\gamma})^{y_i}}{y_i!} \exp(-\lambda x_i^{\gamma}) .$$

Nu er der imidlertid forskel på A-kassernes organisering. I denne opgave betragtes to muligheder for A-kassernes administrationsform, idet der skelnes mellem de 22 A-kasser med en central administration, dvs. hvor hovedkontoret træffer alle betydende afgørelser og de 14 A-kasser med decentral administration, hvor også de enkelte afdelinger har beslutningskompetence.

Antag at de $n_1 = 22$ A-kasser med central form følger MODEL 1 med parametre λ_1 og γ_1 mens de $n_2 = 14$ A-kasser med decentral form følger MODEL 1 med parametre λ_2 og γ_2 . Derved fås en model med i alt fire parametre med punktsandsynligheden

MODEL 2

$$\prod_{j=1}^{n_1} \frac{(\lambda_1 x_j^{\gamma_1})^{y_j}}{y_j!} \exp(-\lambda_1 x_j^{\gamma_1}) \prod_{k=1}^{n_2} \frac{(\lambda_2 x_k^{\gamma_2})^{y_k}}{y_k!} \exp(-\lambda_2 x_k^{\gamma_2}) .$$

I MODEL 2 estimeres de fire parametre ved et kald af PROC NLP og de tilhørende standardafvigelser til

Form	Parameter	Estimat	Standardafvigelse
Central	λ_1	0.0243	0.0018
	γ_1	1.1681	0.0080
Decentral	λ_2	0.0196	0.0014
	γ_2	1.1572	0.0063

Hvis det antages at $\gamma_1 = \gamma_2 = \gamma$ fås en model med tre parametre med punktsandsynligheden

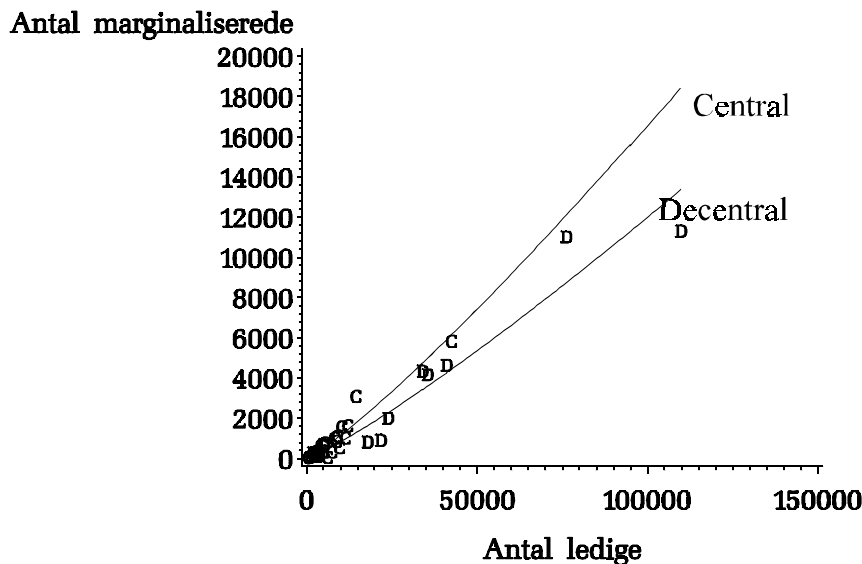
MODEL 3

$$\prod_{j=1}^{n_1} \frac{(\lambda_1 x_j^\gamma)^{y_j}}{y_j!} \exp(-\lambda_1 x_j^\gamma) \prod_{k=1}^{n_2} \frac{(\lambda_2 x_k^\gamma)^{y_k}}{y_k!} \exp(-\lambda_2 x_k^\gamma) .$$

I MODEL 3 estimeres parametrene og de tilsvarende standardafvigelse, atter ved et kald af PROC NLP, til

Form	Parameter	Estimat	Standardafvigelse
Central	λ_1	0.0259	0.00123
Decentral	λ_2	0.0187	0.00102
Fælles	γ	1.1614	0.0050

Figuren viser de 36 observationer med de to kurver estimeret i MODEL 3



a)

Test i MODEL 2 de to hypoteser $\gamma_1 = 1$ og $\gamma_2 = 1$ hver for sig.

b)

Idet estimatorerne i MODEL 2 for de centralt administrerede A-kasser er uafhængige af estimatorerne for de decentralt administrerede, skal det tillige testes om $\gamma_1 = \gamma_2$, altså om MODEL 2 kan reduceres til MODEL 3.

c)

Vis at scorevektoren i MODEL 3 er givet ved

$$\frac{\partial}{\partial \lambda_1} \log L = - \sum_{j=1}^{n_1} x_j^\gamma + \frac{\sum_{j=1}^{n_1} y_j}{\lambda_1}$$

$$\frac{\partial}{\partial \lambda_2} \log L = - \sum_{k=1}^{n_2} x_k^\gamma + \frac{\sum_{k=1}^{n_2} y_k}{\lambda_2}$$

og

$$\frac{\partial}{\partial \gamma} \log L = \sum_{j=1}^{n_1} y_j \log(x_j) - \sum_{j=1}^{n_1} \lambda_1 x_j^\gamma \log(x_j) + \sum_{k=1}^{n_2} y_k \log(x_k) - \sum_{k=1}^{n_2} \lambda_2 x_k^\gamma \log(x_k)$$

Husk at $\frac{d}{d\gamma} x^\gamma = \log(x) x^\gamma$.

I MODEL 3 beregnes korrelationen imellem estimatorerne ud fra matricen $\mathbf{Q}^{-1}$ for λ_1 og λ_2 til - 0.9895.

d)

Beregn Wald teststørrelsen for hypotesen $\lambda_1 = \lambda_2$, altså hypotesen om, at MODEL 3 kan reduceres til MODEL 1.

I MODEL 3 er maksimum for loglikelihoodfunktionen lig -3269.051, mens den i MODEL 1 er givet ved -3702.704.

e)

Beregn Likelihood Ratio teststørrelsen for hypotesen $\lambda_1 = \lambda_2$.

Indsættes estimatorerne for parametrene i MODEL 1 $\lambda = \lambda_1 = \lambda_2 = 0.0528$ og $\gamma = 1.0717$ i scorevektoren for MODEL 3, fås numerisk

$$\mathbf{q} = (50553 \ -50553 \ 0) .$$

Desuden fås med $\lambda = \lambda_1 = \lambda_2 = 0.0528$ og $\gamma = 1.0717$, atter for scorevektoren i MODEL 3:

$$\left(\sum_{i=1}^n \mathbf{q}_i^T \mathbf{q}_i \right)^{-1} = \begin{pmatrix} 2.9858178\text{E}-8 & 3.3137915\text{E}-8 & -5.756343\text{E}-8 \\ 3.3137915\text{E}-8 & 3.7727791\text{E}-8 & -6.514311\text{E}-8 \\ -5.756343\text{E}-8 & -6.514311\text{E}-8 & 1.1315923\text{E}-7 \end{pmatrix}$$

(Notationen E-8 betyder, at tallet før skal ganges med 10^{-8})

f)

Hvorfor er der et 0 (nul) i denne scorevektor $\mathbf{q}$?

g)

Beregn ud fra disse tal Lagrange Multiplikator teststørrelsen for hypotesen $\lambda_1 = \lambda_2$.