Lecture 10:
Poisson regression with random coefficients

Did the German health-care reform reduce the number of doctor visits?

- In order to reduce medical expenditures, a major health-care reform took place in 1997
- The reform raised the co-payments for prescription drugs by up to 200% and imposed upper limits for reimbursement of physicians by the state insurance
- The goal is to investigate whether the number of doctor visits decreased after the reform

Study design

- To address this research question, Wilkelmann (2004) analyzed data from the German socio-economic panel
- Data includes women working full time in the 1996 panel wave or occasion preceding the reform and the 1998 panel wave succeeding the reform

Data

- Outcome: self-reported number of visits to a doctor during the three month prior the interview (y)
- Reform: dummy variable for interview being during the year after the reform versus the year prior the reform (x2)
- Age: age in years (x3)
- Educ: education in years (x4)
- Married: dummy variable for being married (x5)
- Badh: dummy variable for self-reported current health being classified as poor versus good or fair (x6)
- Loginc: log of household income (x7)

Note: fewer than half of the subjects provide data for both occasions

Poisson regression ignoring clustering and overdispersion

- Occasion *(i), subject (j)*

The number of doctor visits y_{ij} is assumed to have a Poisson distribution with mean μ_{ij}

$$\log \mu_{ij} = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3ij} + \beta_4 x_{4ij}$$

$$+\beta_5 x_{5ij} + \beta_6 x_{6ij} + \beta_7 x_{7ij}$$

Poisson regression with level-1 random intercept to account for overdispersion

- Occasion *(i), subject (j)*

$$\log \mu_{ij} = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3ij} + \beta_4 x_{4ij}$$

$$+\beta_5 x_{5ij} + \beta_6 x_{6ij} + \beta_7 x_{7ij} + \varsigma_{ij}^{(1)}$$

$$\varsigma_{ij}^{(1)} \sim N(0, \tau_{1}^{2})$$

Connection between model with random intercept and marginal model for count data

$$\mu_{ij}^{M} = E[y_{ij}^{\cancel{M}} \mid x_{ij}] = \exp(\beta_1 + \tau^2/2 + \sum_{p=1}^{7} \beta_p x_{pij})$$

$$\beta_1^{M} = \beta_1 + \tau^2/2$$

Under a marginal model and model with random intercept the interpretation of the parameters is the same expect for the intercept

$$\mu_{ij}^{M} = E[y_{ij} \mid x_{ij}] = \int E[y_{ij} \mid \varsigma_{ij}^{(1)}, x_{ij}] \times N(0, \tau^2) d\varsigma_{ij}^{(1)}$$

Connection between model with random intercept and marginal model for count data

Additive overdispersion

$$\mathrm{var}(y_{ij} \mid x_{ij}) = \mu_{ij}^{M} + (\mu_{ij}^{M})^2 \left\{ \exp(\tau^2) - 1 \right\}$$

$$\mathrm{var}(y_{ij} \mid x_{ij}) > \mathrm{var}^{M}(y_{ij} \mid x_{ij}) = \mu_{ij}^{M} \Leftrightarrow \tau^2 > 0$$

Note: including a level-1 random intercept is equivalent to fitting a Poisson regression model with an "additive" overdispersion

Poisson regression with a level 2 random intercept

(i) Is the occasion, (j) is the person

$$\log \mu_{ij} = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3ij} + \beta_4 x_{4ij}$$
$$+\beta_5 x_{5ij} + \beta_6 x_{6ij} + \beta_7 x_{7ij} + \varsigma_{1j}^{(2)}$$

The standard Poisson regression model makes the unrealistic assumption that the number of doctor visits before the reform (y1j) is independent on the number of doctor visits after the reform (y2j) for the same person given the covariates

Quasi-likelihood (Marginal Model)

In the quasi-likelihood approach, we do not specify a statistical model, but instead we merely specify the expectation and the variance of the counts

$$\log \mu_{ij} = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3ij} + \beta_4 x_{4ij}$$
$$+\beta_5 x_{5ij} + \beta_6 x_{6ij} + \beta_7 x_{7ij}$$

$$\text{var}(y_{ij} \mid x_{ij}) = \phi^* \mu_{ij}$$

Multiplicative overdispersion

Table 6.1: Estimates for generalized estimating equations Poisson regression, random-intercept Poisson regression, and random-coefficient Poisson regression

Estimated incidence rate ratio: there is a 12% reduction in the number of doctor visits between 1996 and 1998 for a given value of the covariates

	Marginal effects			Conditional effects			
	Poisson	GEE-Poisson		RI-Poisson		RC-Poisson	
	Est	Est	(95% CI)†	Est	(95% CI)†	Est	(95% CI)†
Fixed part: rate ratios							
$\exp(\beta_2)$ [reform]	0.88	0.88	(0.80, 0.98)	0.95	(0.86, 1.06)	0.90	(0.81, 1.00)
$\exp(\beta_3)$ [age]	1.00	1.01	(1.00, 1.01)	1.01	(1.00, 1.01)	1.00	(1.00, 1.01)
$\exp(\beta_4)$ [educ]	0.99	0.99	(0.97, 1.01)	1.01	(0.98, 1.03)	1.01	(0.98, 1.03)
$\exp(\beta_5)$ [married]	1.04	1.04	(0.90, 1.19)	1.08	(0.95, 1.23)	1.09	(0.97, 1.22)
$\exp(\beta_6)$ [badh]	3.11	3.02	(2.54, 3.58)	2.47	(1.96, 3.10)	3.03	(2.58, 3.55)
$\exp(\beta_7)$ [loginc]	1.16	1.15	(0.98, 1.34)	1.10	(0.93, 1.30)	1.14	(0.98, 1.31)
$\exp(\beta_8)$ [summer]	1.01	0.97	(0.82, 1.16)	0.87	(0.74, 1.02)	0.91	(0.78, 1.07)
Random part							
$\psi_{11}^{(2)}$				0.82		0.91	
$\psi_{22}^{(2)}$						0.86	
$\psi_{21}^{(2)}$						−0.43	
Log likelihood	−5942.69			−4643.34		−4513.80	

†based on sandwich estimator

The RC model is preferred

Random coefficient Poisson regression

With this model we allow the effect of the reform to vary across individuals

$$\log \mu_{ij} = \beta_1 + \beta_2 x_{2ij} + \beta_3 x_{3ij} + \beta_4 x_{4ij}$$

$$+\beta_5 x_{5ij} + \beta_6 x_{6ij} + \beta_7 x_{7ij} + \varsigma_{1j}^{(2)} + \varsigma_{2j}^{(2)} x_{2ij}$$

Interpretation

- Instead of thinking of this model as a random coefficients model, we could view it as a model with two different random intercepts in 1996 and in 1998.

$$1996: \varsigma_{1j}^{(2)}$$

$$1998: \varsigma_{1j}^{(2)} + \varsigma_{2j}^{(2)}$$

$$1996: \text{var}(\varsigma_{1j}^{(2)}) = 0.909$$

$$1998: \text{var}(\varsigma_{1j}^{(2)} + \varsigma_{2j}^{(2)}) = 0.904$$

$$\text{cov}(\varsigma_{1j}^{(2)}, \varsigma_{1j}^{(2)} + \varsigma_{2j}^{(2)}) = 0.475$$

Interpretation

- The model with random intercept only has the same random intercept in 1996 and 1998 and had a single parameter representing the random intercept variance at both occasions which is equal to 0.817
- The random coefficient model can be viewed as a model allowing separate accommodation for overdispersion and dependence

Interpretation

- Not surprisingly the estimate (0.817) from the random intercept model was intermediate between the estimates of the two variances (0.909,0.904) and the covariance (0.475) from the random coefficient model.