EM Algorithm and ZIP models

Recap: EM algorithm for ZIP models

EM algorithm

E-step: Given $\gamma^{(k)}$ and $\beta^{(k)}$,

$$Z_i^{(k)} = \mathbb{E}[Z_i \mid Y_i, \gamma^{(k)}, \beta^{(k)}]$$

M-step: Given $Z_i^{(k)}$,

$$\gamma^{(k+1)} = \underset{\gamma}{\operatorname{argmax}}\ \ell(\gamma; Y, Z^{(k)})$$

$$\beta^{(k+1)} = \underset{\beta}{\operatorname{argmax}}\ \ell(\beta; Y, Z^{(k)})$$

M-step

$$\beta^{(k+1)} = \underset{\beta}{\arg\max} \sum_{i=1}^{n} (1-Z_i^{(k)}) \left[Y_i \beta^T X_i - e^{\beta^T X_i} \right]$$

weighted Poisson regression w/ weights $w_i = 1 - Z_i^{(k)}$

$$\gamma^{(k+1)} = \underset{\gamma}{\arg\max} \sum_{i=1}^{n} \mathbb{1}\{Y_i = 0\} Z_i^{(k)} \gamma^T X_i + \sum_{i=1}^{n} 0 \cdot (1-Z_i^{(k)}) \gamma^T X_i$$
$$- \sum_{i=1}^{n} Z_i^{(k)} \log(1+e^{\gamma^T X_i}) - \sum_{i=1}^{n} (1-Z_i^{(k)}) \log(1+e^{\gamma^T X_i})$$

weighted logistic regression of $Y^* = (\mathbb{1}\{Y_1 = 0\}, \dots, \mathbb{1}\{Y_n = 0\}, 0, 0, \dots, 0)^T$

weights $w = (Z_1^{(k)}, \dots, Z_n^{(k)}, 1-Z_1^{(k)}, \dots, 1-Z_n^{(k)})^T$

EM algorithm in general

Let θ be an unknown parameter we want to estimate. Let Y be a set of observed data, Z a set of unobserved latent/missing data.

$$L(\theta) = P(Y \mid \theta) = \int_z P(Y, Z=z \mid \theta)\, dz$$

want to maximize $L(\theta)$, but this is challenging when Z is unobserved

$$= \int_z P(Y \mid Z=z, \theta)\, P(Z=z \mid \theta)\, dz$$

EM algorithm

E step: Let $\theta^{(k)}$ be current estimate of θ

$$Q(\theta \mid \theta^{(k)}) = \mathbb{E}_{Z \mid Y, \theta^{(k)}}[\log L(\theta; Z, Y)]$$

M step: $\theta^{(k+1)} = \underset{\theta}{\operatorname{argmax}}\ Q(\theta \mid \theta^{(k)})$

<u>Motivation</u>:

$$\log P(Y|\theta) = \log\left(\int_z P(Y, Z=z|\theta)\, dz\right)$$

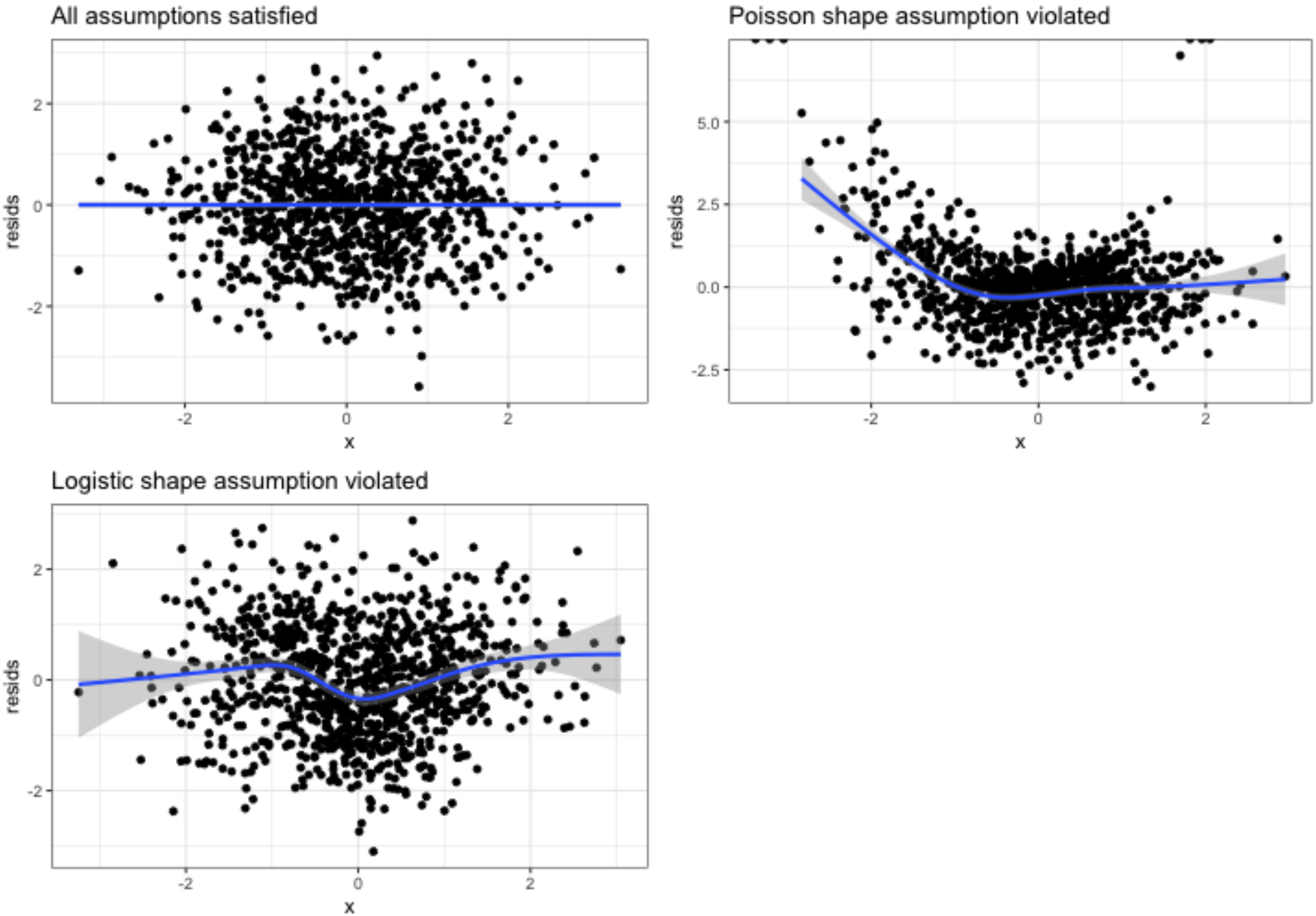
$$= \log\left(\int_z \frac{P(Y, Z=z|\theta)}{P(Z=z|Y, \theta_{old})} P(Z=z|Y, \theta_{old})\, dz\right)$$

$$= \log \mathbb{E}_{Z|Y, \theta_{old}}\left[\frac{P(Y, Z|\theta)}{P(Z|Y, \theta)}\right]$$

Class activity

https://sta712-f22.github.io/class_activities/ca_lecture_31.html

Assessing the shape assumption



Logistic component vs. Poisson component