Jan 15, 2021

Estimation of μ_p, μ_t

A New Adhoc Algorithm.

Step 0. Initialize μ_p, μ_t (perhaps using the method in the report).

Step 1. Calculate the total number of judge-days and subtract the expected number of trial judge-days. Let N denote the number of remaining judge-days.

Step 2. Set $$\theta = \frac{\text{Total number of pleas in the data}}{N}.$$

We model the plea demand for a judge as $D \sim \text{Poisson}(\theta)$, whereas the number of pleas a judge can serve in a day is denoted by X, $X \sim \text{Poisson}(\mu_p)$.

Step 3. Let $S_i = \min(D_i, X_i)$ denote the number of pleas sentenced for judge-day combination $i = 1, \dots, N$. We have that

$$P(S_i = s) = P(X_i = s \mid X_i \leq D_i) \cdot P(X_i \leq D_i) + P(D_i = s \mid X_i > D_i) \cdot P(X_i > D_i).$$

Step 4 Maximize loglikelihood $\mathcal{L}(\mu_p) = \sum_{i=1}^{N} \log P(S_i = s_i)$.

Set $\mu_p = \arg\max \mathcal{L}(\mu_p)$ and calculate μ_t as in the report.

- Iterate until convergence of μ_p, μ_t.

2/

EM Algorithm

This analysis will focus on the clean days (for pleas) and consider two distributional assumptions for D, X: Poisson and Normal. First, consider the Poisson distribution.

E-Step. Compute $\mathbb{E}_{X,D}\{\log \mathbb{P}(X_i, D_i, S_i \mid \theta, \mu_p) \mid S_i = s, \theta, \mu_p\}$. Note that where $X_i \sim \text{Poisson}(\mu_p)$, $D_i \sim \text{Poisson}(\theta)$, $S_i = \min(X_i, D_i)$.

$$p(X_i=x, D_i=d, S_i=s) = \begin{cases} p(X_i=x, D_i=d) & \text{if } s=\min(x,d) \\ 0 & \text{otherwise,} \end{cases}$$

where $p(X_i=x, D_i=d) = e^{-\mu_p x} \dfrac{\mu_p^x}{x!} e^{-\theta d} \dfrac{\theta^d}{d!}$.

For $x, d, s \geq 0$ such that $s = \min(d,x)$ we have that

$$\log p(X_i=x, D_i=d, S_i=s) = x(\log \mu_p - \mu_p) - \sum_{j=1}^{x} \log j + d(\log\theta - \theta) - \sum_{j=1}^{d} \log j.$$

To compute the conditional expectation $\mathbb{E}_{X,D}[\cdot \mid S_i = s]$, consider the conditional pmf $p(X_i=x, D_i=d \mid S_i=s)$, denoted by $p(x,d \mid s)$

3/

$$p(x, d \mid s) = \frac{\mathbb{P}(X_i = x, D_i = d)}{\mathbb{P}(S_i = s)}$$

$$= \frac{\mathbb{P}(X_i = x)\, \mathbb{P}(D_i = d)}{\mathbb{P}(X_i = s)\, \mathbb{P}(D_i \geq s) + \mathbb{P}(X_i > s)\, \mathbb{P}(D_i = s)}$$

Then letting $A(s) = \{(z, d) : d \geq z\} \cup \{(x, z) : x > z\}$, we have that

$$\mathbb{E}\left[\log \mathbb{P}(X_i, D_i, S_i) \mid S_i = s\right]$$

$$= \mathbb{E}_{X,D}\left[(\log \mu_p - \mu_p) X_i - \sum_{j=1}^{X_i} \log j + (\log\theta - \theta) D_i - \sum_{j=1}^{D_i} \log j \mid S_i = s\right]$$

$$= \sum_{(x,d) \in A(s)} \left[(\log \mu_p - \mu_p) x - \sum_{j=1}^{x} \log j + (\log\theta - \theta) d - \sum_{j=1}^{d} \log j\right] p(x, d \mid s)$$

<u>M-Step:</u> Optimize over μ_p, θ, the following

$$\sum_{i=1}^{N} \mathbb{E}_{X,D}\left[\log(X_i, D_i, S_i) \mid S_i = s_i\right],$$

and iterate until convergence.

<u>Remark:</u> I'm not sure how manageable this computation is.

EM-Algorithm for the Normal Distribution Case:

Under the normal distribution assumption, we write

$$f_X(x) = \frac{1}{\sigma_1\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu_p}{\sigma_1}\right)^2\right\} = \phi\left(\frac{x-\mu_p}{\sigma_1}\right)$$

$$f_D(d) = \frac{1}{\sigma_2\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{d-\theta}{\sigma_2}\right)^2\right\} = \phi\left(\frac{d-\theta}{\sigma_2}\right).$$

Because $S = \min(X, D)$, we write

$$\mathbb{P}(S \le s) = \mathbb{P}(X \le s) \cdot P(D \le s) = \Phi\left(\frac{s-\mu_p}{\sigma_1}\right) \Phi\left(\frac{s-\theta}{\sigma_2}\right).$$

Thus, $f_S(s) = \frac{d}{ds}\mathbb{P}(S \le s)$ is given as follows:

$$f_S(s) = \frac{1}{\sigma_1}\phi\left(\frac{s-\mu_p}{\sigma_1}\right)\Phi\left(\frac{s-\theta}{\sigma_1}\right) + \frac{1}{\sigma_2}\phi\left(\frac{s-\theta}{\sigma_2}\right)\Phi\left(\frac{s-\mu_p}{\sigma_1}\right).$$

For the E-step, we need to calculate

$$\mathbb{E}_{X,D}\left[\log f(X, D, S \mid \theta, \mu) \mid S, \theta, \mu_p\right].$$

Note that

$$\log f(X, D, S) = -\log\sigma_1 - \frac{1}{2}\left(\frac{X-\mu_p}{\sigma_1}\right)^2 - \log\sigma_2 - \frac{1}{2}\left(\frac{D-\theta}{\sigma_2}\right)^2.$$

5/

Thus, we are interested in computing

$$-\log \sigma_1 - \log \sigma_2 - \frac{1}{2} \mathbb{E}\left[\left(\frac{X-\mu_p}{\sigma_1} \right)^2 + \left(\frac{D-\theta}{\sigma_2} \right)^2 \,\middle|\, S = s \right].$$

To this end, lets first derive the conditional density $f(x, D \mid s)$.

$$f(x, D \mid s) \propto \begin{cases} \frac{f(x, D)}{f(s)} & \text{if } s = \min(x, D) \\ 0 & \text{otherwise} \end{cases}$$

In the first case, we have that

$$f(x, D \mid s) \propto \frac{\phi\left(\frac{x-\mu_p}{\sigma_1} \right) \cdot \phi\left(\frac{D-\theta}{\sigma_2} \right)}{\frac{1}{\sigma_1} \phi\left(\frac{s-\mu_p}{\sigma_1} \right) \Phi\left(\frac{s-\theta}{\sigma_2} \right) + \frac{1}{\sigma_2} \phi\left(\frac{s-\theta}{\sigma_2} \right) \Phi\left(\frac{s-\mu_p}{\sigma_2} \right)}.$$

Then, we have that

$$\mathbb{E}[\log(X, D, S) \mid S] = -\log \sigma_1 - \log \sigma_2 \qquad (*)$$

$$-\frac{1}{2}\left[\int_s^\infty f(x, s \mid s) \left[\left(\frac{x-\mu_p}{\sigma_1} \right)^2 + \left(\frac{s-\theta}{\sigma_2} \right)^2 \right] dx \right.$$

$$\left. + \int_s^\infty \left[\left(\frac{s-\mu_p}{\sigma_1} \right)^2 + \left(\frac{d-\theta}{\sigma_2} \right)^2 \right] f(s, d \mid s)\, dd \right]$$

6/

M-Step. Minimize

$$\sum_{i=1}^{N} \mathbb{E}\left[\log f(X_i, D_i, S_i) \mid S_i\right]$$ (see (*)) over

$\mu_p, \theta, \sigma_1, \sigma_2$. To do so, we can use Gauss-Hermite integration (see Judd's book) for Normal densities.

Then we iterate until the parameters converge.