May 27, 2021 Meeting Agenda

May 26, 2021

1 Joint estimation of μ_p and μ_t

I was mostly able to implement step 4 of the ad-hoc algorithm. This is the step where for a given θ and number of plea days, N, we find the μ_p that minimizes the negative log likelihood (NLL). I'm using a gradient descent algorithm, so the immediate next step is to figure out a good stopping criterion for the gradient descent algorithm. I was planning on stopping once there has been no improvement after 5 iterations. For context, this is what we minimize in step 4:

$$\min L(\mu_p) = -\sum_{i=1}^{N} \log P(S_i = s)$$

$$P(S_i = s) = \frac{\theta^s e^{-\theta}}{s!}[1 - \sum_{k=1}^{s-1} \frac{\mu_p^k e^{-\mu_p}}{k!}] + \frac{\mu_p^s e^{-\mu_p}}{s!}[1 - \sum_{k=1}^{s} \frac{\theta^k e^{-\theta}}{k!}]$$

1.1 Problems

- Our current best guess of μ_p is 10.7. However, as part of our probability calculation, we have to compute μ_p^s, where s is the number of pleas a judge heard in a day. Trying to compute μ_p^s for values of s greater than 20 raises an error. I think this is because the largest integer value allowed by the deep learning package I'm using for the gradient calculation is a little less than 10^{18}. I'm currently only using values of $s \leq 15$.
- Step 4 is currently a bit slow, it takes 1 or 2 minutes to find the value of μ_p that minimizes the NLL for a given value of θ.

1.2 Next Steps

- Figure out a stopping criteria for gradient descent algorithm.
- Figure out what to do about days with more than 20 pleas. For context, 14, 16, and 17 are the 96th, 97th, and 98th percentile for pleas sentenced in a day in tha dataset.
- Finish implementing full algorithm (i.e. include updating of μ_t, θ, and N)
- Implement exclusion criteria for clean days.