Bayesian Machine Learning Assignment 2

1. Congratulations! You just assumed the position of Dean of Clinical Affairs at University of Chicago Mitchell Hospital. Your task is to determine the number of ER personnel to be present at the hospital during the 8-hour shifts.

   a. (10p) Since you have no idea how many patients will show up on 8-hour shifts, you check the patient logs to have a "prior" belief of the λ. You learned, many years ago from a very fruitful Bayesian Machine Learning class, that the prior distribution for λ of a Poisson process is Gamma. You calculated the parameters of this prior distribution from your logs to be $\alpha = 20$ and $\beta = 20$. Plot the prior distribution $p(\lambda)$ for these parameters between $0 < \lambda < 4$. (HINT: you can use `scipy.stats.gamma.pdf` for this)

   b. First, you will use the $E(\lambda)$ as your estimate for λ^* to determine the number of patients you expect.

      i. (10p) Plot the predictive probability distribution of patients $p(x|\lambda^*)$ for this initial estimate between (0<x<10). Remember, the distribution is Poisson. You can use `scipy.stats.poisson.pmf`

      ii. (10p) Calculate E(X) – this is your expected number of patients on an 8-hour shift. (HINT: You can do the expectation calculation numerically, too – like below)

```
y = [i*j for i,j in zip(probabilities, x)]
E_x = np.sum(y)/np.sum(probabilities)
```

   c. (10p) After your first day at work, you get some negative feedback from the crew that they felt understaffed. You ask for the first day's records, and your staff report that the numbers of patients showed on the three 8-hour shifts that day were [3, 4, 1]. Now, you have your first data! (yay) $D = x_1, x_2, x_3$. What is the likelihood of this first day, based on your λ^*?

   d. (20p) Calculate the Maximum Likelihood estimate of your new λ, based on your first day.

   e. (10p) Now calculate and plot your posterior distribution of $p(\lambda|D)$. What are the new α, β?

   f. (20p) Calculate the Maximum a Posteriori estimate of your new λ, based on your first day and your prior information.

   g. (10p) Plot the posterior predictive distribution of the expected number of patients $p(x|D)$. This becomes negative binomial with parameters $n = \alpha'$ and $p = \frac{\beta'}{\beta'+1}$, where α' and β' are the posterior parameters of $p(\lambda|D)$ you calculated at part e. So, you will use `scipy.stats.nbinom.pmf`