Short Assignment 2

This is an individual assignment.

For the analytical problems, you can write your answers in markdown cells with $\LaTeX$ or `push` a single pdf to your repository with all handwritten answers.

Objectives

- Define observed data likelihoods.
- Derive the parameter estimation with MLE and MAP approaches.
- Derive online prior update.

Suppose you have a training set with N data points $\{x_i\}_{i=1}^N$, where $x_i \in \mathbb{R}^+$ (set of positive real numbers). Assume the samples are independent and identically distributed (i.i.d.), and each sample is drawn from a Gamma random variable with probability density function:

$$p(x|\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$$

where $\alpha, \beta > 0$.

Moreover, consider another Gamma density as the prior probability on the hyperparameter β,

$$p(\beta|a,b) = \frac{b^a}{\Gamma(a)} \beta^{a-1} e^{-b\beta}$$

where $a, b > 0$.

Answer the following questions:

1. (3 points) **Derive the maximum likelihood estimate (MLE) for the parameter β. Show your work.**

In [1]:
```
from IPython.display import Image
```

In [4]:
```
Image('4.png', width=900)
```

Out[4]:

$$① \quad \mathcal{L}^0 = \prod_{i=1}^{N} \frac{\beta^{\alpha}}{\Gamma(\alpha)} \cdot x_i^{a-1} \cdot e^{-\beta x_i}$$

$$\mathcal{L} = \ln \mathcal{L}^0 = \sum_{i=1}^{N} \left((\alpha \cdot \ln\beta - \ln\Gamma(\alpha)) - (a-1) \cdot \ln x_i - \beta x_i \right)$$

$$\frac{\partial \mathcal{L}}{\partial \beta} = \sum_{i=1}^{N} \left(\alpha \cdot \frac{1}{\beta} - x_i \right)$$

$$\frac{\partial \mathcal{L}}{\partial \beta} = 0 \iff \beta = \frac{\alpha}{\sum x_i}$$

In [ ]:

1. (3 points) **Derive the maximum a posteriori (MAP) estimate for the parameter β. show your work.**

In [5]:

```
Image('5.png', width=900)
```

Out[5]:

$$② \quad \mathcal{L}^0 = p(x \mid \alpha, \beta) \cdot p(\beta \mid a, b)$$

$$= \prod_{i=1}^{N} \left(\frac{\beta^{\alpha}}{\Gamma(\alpha)} \cdot x_i^{a-1} \cdot e^{-\beta x} \right) \cdot \frac{b^a}{\Gamma(a)} \beta^{a-1} \cdot e^{-b\beta}$$

$$\mathcal{L} = \ln \mathcal{L}^0 = \prod_{i=1}^{N} \left(\alpha \cdot \ln\beta - \ln\Gamma(a) + (a-1) \ln x - \beta x_i \right) + a \cdot \ln b - \ln\Gamma(a) + (a-1) \cdot \ln\beta - b\beta$$

$$\frac{\partial \mathcal{L}}{\partial \beta} = \prod_{i=1}^{N} \left(\frac{\alpha}{\beta} - x_i \right) + \frac{a-1}{\beta} - b$$

$$\frac{\partial \mathcal{L}}{\partial \beta} = 0 \iff \frac{\alpha}{\beta} - \prod_{i=1}^{N} x_i + \frac{a-1}{\beta} - b = 0$$

$$\frac{\alpha + a - 1}{\beta} = \prod_{i=1}^{N} x_i + b$$

$$\beta = \frac{\alpha + a - 1}{\prod_{i=1}^{N} x_i + b}$$

In [ ]:

In [ ]:

```

```

1. (3 points) **Is the Gamma distribution a conjugate prior for the parameter β, of the Gamma data likelihood distribution? Why or why not? If yes, write down the pseudo-code for the online update of the prior parameters (including the equations for the new parameters of the prior).**

In [6]:

```
Image('6.png', width=900)
```

Out[6]:

③ Yes, because we can find a conjngate relationship.

$$\mathcal{L}_0 = p(x|\alpha,\beta) \cdot p(\beta|a,b)$$

$$= \prod_{i=1}^{N} \left(\frac{\beta^{\alpha}}{\Gamma(\alpha)} \cdot x_i^{\alpha-1} \cdot e^{-\beta x_i}\right) \cdot \frac{b^a}{\Gamma(a)} \beta^{a-1} \cdot e^{-b\beta}$$

$$= \frac{1}{\Gamma(\alpha)\Gamma(a)} \cdot \beta^{\alpha+a-1} \cdot \prod_{i=1}^{N} x_i^{a-1} \cdot e^{-\beta\left(\prod_{i=1}^{N} x_i + b\right)}$$

$$p(\beta|a,b) = \frac{b^a}{\Gamma(a)} \beta^{a-1} \cdot e^{-b\beta}.$$

from comparison we know:

$$a^{(t+1)} \rightarrow a^{(t)} + \alpha$$
$$b^{(t+1)} \rightarrow b^{(t)} + \prod_{i=1}^{N} x_i$$

$$t \leftarrow t+1$$

In [ ]:

In [ ]:

On-Time (1 point)

Submit your assignment before the deadline.

Submit Your Solution

Confirm that you've successfully completed the assignment.

Along with the Notebook, include a PDF of the notebook with your solutions.

`add` and `commit` the final version of your work, and `push` your code to your GitHub repository.

Submit the URL of your GitHub Repository as your assignment submission on Canvas.
