Your Name: Robby Kevlishvili **Problem Set 1 Due: Feb. 18, 2025**

(1) The Poisson distribution has probability mass function

$$p(y_i|\theta) = \frac{\theta^{y_i} e^{-\theta}}{y_i!}, \qquad \theta > 0, \qquad y_i = 0, 1, \ldots \tag{1}$$

and let $y_1, \ldots, y_n$ be random sample from this distribution.

1. Show that the gamma distribution $\mathcal{G}(\alpha, \beta)$ is a conjugate prior distribution for the Poisson distribution.

   Solution:

   (a) The Gamma prior is given by

   $$\pi(\theta|\alpha, \beta) \propto \theta^{\alpha-1} \cdot e^{-\theta/\beta} \tag{2}$$

   The likelihood function is given by

   $$f(y|\theta) \propto \theta^{\sum y_i} \cdot e^{-n\theta} \tag{3}$$

   Multiplying these two equations gives

   $$\pi(\theta|y) \propto \theta^{\alpha-1+\sum y_i} \cdot e^{-\theta(n/\beta)} \tag{4}$$

   We see that

   $$\theta|y \sim \mathcal{G}(\alpha', \beta') \qquad \text{where } \alpha' = \alpha + \sum y_i, \qquad \beta' = n + (1/\beta) \tag{5}$$

2. Show that $\bar{y}$ is the MLE for θ.

   Solution:

   (a) The likelihood function is given by

   $$f(y|\theta) \propto \theta^{\sum y_i} \cdot e^{-n\theta} \tag{6}$$

   Taking the log gives us

   $$log f(y|\theta) = \sum y_i log(\theta) - n\theta - log(\sum y_i) \tag{7}$$

Taking the derivative with respect to θ gives

$$log f(y|\theta)' = \frac{\sum y_i}{\theta} - n \tag{8}$$

Setting this equal to 0 and solving for θ gives

$$\theta = \frac{\sum y_i}{n} = \bar{y} \tag{9}$$
$$Q.E.D \tag{10}$$

3. Write the mean of the posterior distribution as a weighted average of the mean of the prior distribution and the MLE.

   Solution:

   (a) We want to find weights (w and (1-w)) so that

   $$E(\theta|y) = w \cdot E[0] + (1-w) \cdot \bar{y} \tag{11}$$

   Substituting answers from parts (1) and (2) gives

   $$\frac{\alpha + \sum y_i}{\beta + n} = w \cdot (\alpha/\beta) + (1-w) \cdot \frac{\sum y_i}{n} \tag{12}$$

   Using WolframAlpha to solve for w gives

   $$w = \frac{\beta}{\beta + n} \tag{13}$$

   This means the posterior distribution can be expressed as

   $$E(\theta|y) = \frac{\beta}{\beta + n} \cdot E[0] + (1 - \frac{\beta}{\beta + n}) \cdot \bar{y} \tag{14}$$

4. What happens to the weight on the prior mean as n becomes large?

   As n becomes larger, the weight on the prior mean decreases. If n becomes sufficiently large, the weight will tend to 0.

(2) Consider the following two sets of data obtained after tossing a die 100 and 1000 times, respectively:

n	1	2	3	4	5	6
100	19	12	17	18	20	14
1000	190	120	170	180	200	140

Suppose you are interested in θ_1, the probability of obtaining a one spot. Assume your prior for all the probabilities is a Dirichlet distribution, where each $\alpha_i = 2$. Compute the posterior distribution for θ_1 for each of the sample sizes in the table. Plot the resulting distribution and compare the results. Comment on the effect of having a larger sample.

Solution:

1. The prior distribution is expressed as

$$Dirichlet(2,2,2,2,2,2) \tag{15}$$

The posterior distribution for $n = 100$ is expressed as

$$Dirichlet(21,14,19,20,22,16) \tag{16}$$

The posterior distribution for $n = 1000$ is expressed as

$$Dirichlet(192,122,172,182,202,142) \tag{17}$$

The marginal distribution can be expressed as $\theta_1 \sim Beta(\alpha_1, \sum \alpha)$, or specifically

$$for\ n = 100 : \theta_1 \sim Beta(21,91) \tag{18}$$

$$for\ n = 1000 : \theta_1 \sim Beta(192,820) \tag{19}$$

Plotting the graphs:

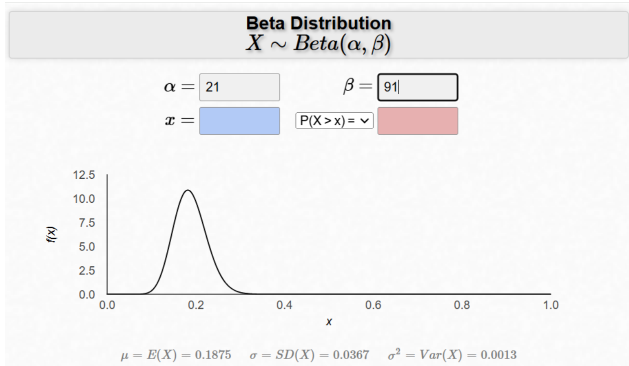

Figure 1: Beta(21,91)

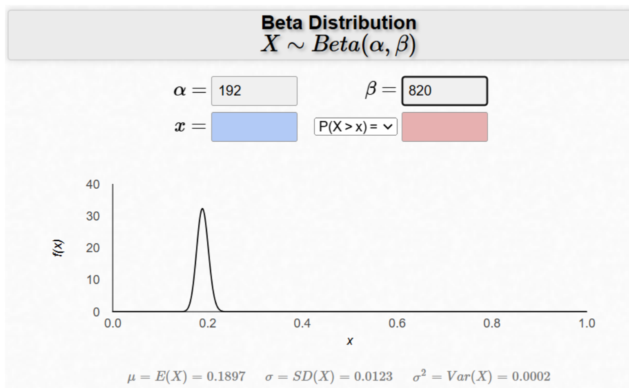

Figure 2: Beta(192,820)

Comparing these two figures, we clearly see the larger sample size gives a more accurate and robust estimation range. The estimation seems to converge closer to the actual parameter value. This also makes intuitive sense- more data results in more accurate estimations.