(1) The Poisson distribution has probability mass function:

$$p(y_i|\theta) = \frac{\theta^{y_i} e^{-\theta}}{y_i!}, \quad \theta > 0, \quad y_i = 0, 1, 2, \ldots \tag{1}$$

Let $y_1, \ldots, y_n$ be a random sample from this distribution.

1. **Conjugate Prior:** The gamma distribution $\mathcal{G}(\alpha, \beta)$ has density:

$$p(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta}, \quad \theta > 0. \tag{2}$$

The likelihood function for n observations is:

$$L(\theta) = \prod_{i=1}^{n} \frac{\theta^{y_i} e^{-\theta}}{y_i!} = \theta^{\sum y_i} e^{-n\theta} \prod_{i=1}^{n} \frac{1}{y_i!}. \tag{3}$$

The posterior is proportional to $\theta^{(\alpha+\sum y_i)-1} e^{-(\beta+n)\theta}$, which follows a gamma distribution $\mathcal{G}(\alpha + \sum y_i, \beta + n)$. Thus, the gamma distribution is a conjugate prior.

2. **MLE for θ:** The log-likelihood function is:

$$\ell(\theta) = \sum_{i=1}^{n} \left(y_i \log\theta - \theta - \log y_i! \right). \tag{4}$$

Differentiating and solving for θ gives:

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^{n} y_i = \bar{y}. \tag{5}$$

3. **Posterior Mean as Weighted Average:** The mean of a gamma distribution $\mathcal{G}(a, b)$ is $\frac{a}{b}$. The posterior mean is:

$$\mathbb{E}[\theta|y] = \frac{\alpha + \sum y_i}{\beta + n}. \tag{6}$$

Rewriting:

$$\mathbb{E}[\theta|y] = \left(\frac{\beta}{\beta+n}\right) \frac{\alpha}{\beta} + \left(\frac{n}{\beta+n}\right) \bar{y}. \tag{7}$$

This shows the posterior mean is a weighted average of the prior mean and the MLE.

4. **Effect of Large n:** As $n \to \infty$, the weight on the prior mean $\frac{\beta}{\beta+n}$ approaches zero, while the weight on the MLE $\frac{n}{\beta+n}$ approaches one. Thus, for large n, the posterior mean converges to the MLE $\bar{y}$, reducing the influence of the prior.

article amsmath amssymb pgfplots width=10cm, compat=1.18

Problem 2: Posterior Analysis for Die Toss Probability

Given Data

Observed frequencies for die tosses:

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

Prior Specification

We assume a symmetric Dirichlet prior:

$$\boldsymbol{\alpha} = (2, 2, 2, 2, 2, 2)$$

Posterior Distribution

For a Dirichlet-Multinomial conjugate pair:

$$\boldsymbol{\alpha}^{\text{post}} = \boldsymbol{\alpha} + \text{observed counts}$$

- For $n = 100$:

$$\alpha_1^{\text{post}} = 2 + 19 = 21$$

- For $n = 1000$:

$$\alpha_1^{\text{post}} = 2 + 190 = 192$$

Posterior Mean

The marginal posterior mean for θ_1 is given by:

$$E[\theta_1] = \frac{\alpha_1^{\text{post}}}{\sum \alpha_i^{\text{post}}}$$

- $n = 100$:

$$E[\theta_1] = \frac{21}{2 \cdot 6 + 100} = \frac{21}{112} \approx 0.1875$$

- $n = 1000$:

$$E[\theta_1] = \frac{192}{2 \cdot 6 + 1000} = \frac{192}{1012} \approx 0.1897$$

Posterior Density Plot

[domain=0:0.4, samples=200, xlabel=θ_1, ylabel=Density, legend pos=north east, title=Posterior Distribution of θ_1, grid=major, ymajorgrids=true, xmin=0.1, xmax=0.3, ymin=0, ymax=25, restrict y to domain=0:25] [blue, thick] x$^{(}21 - 1) * (1 - x)^{(}91 - 1)/beta(21, 91); 100tosses$

[red, thick, dashed] x$^{(}192 - 1) * (1 - x)^{(}820 - 1)/beta(192, 820); 1000tosses$

Effect of Sample Size

Key observations:

- The posterior for $n = 1000$ (red dashed) is much narrower than for $n = 100$ (blue), demonstrating reduced uncertainty with larger samples
- Both posteriors are centered near 0.19, consistent with the observed frequencies
- The prior ($\alpha = 2$) has negligible influence for $n = 1000$, as expected in large-sample Bayesian analysis

Plotting the Posterior Distributions