**%------------%**

**% Preamble %**

**%------------%**

**\documentclass[final,11pt]{article}**

**\usepackage[paperwidth=9.0in, top=1.2in, bottom=1.2in, left=1.2in, right=1.2in]{geometry}**

**\usepackage{amsmath}**

**\usepackage{color}**

**\usepackage{multirow}**

**\usepackage{setspace}**

**\usepackage{fancyhdr}**

**\usepackage{longtable}**

**\usepackage{array}**

**\usepackage{booktabs}**

**\usepackage{mathpazo}**

**\usepackage{threeparttable}**

**\usepackage{eurosym}**

**\usepackage[colorlinks, linkcolor=blue, anchorcolor=blue, citecolor=blue]{hyperref}**

**\renewcommand{\headrulewidth}{0pt}**

**\setlength{\arraycolsep}{10pt}**

**\setlength\headheight{0.5cm}**

**\setlength\headsep{0.8cm}**

**\setlength\footskip{1.0cm}**

**\setlength{\parindent}{0em}**

**\pagestyle{fancy}**

**\chead{\textcolor[rgb]{0.5,0.5,0.5}{\sc Spring 2025: ECON 3120}}**

**%------------%**

**% Document %**

**%------------%**

**\begin{document}**

**\thispagestyle{empty}**

**\begin{spacing}{1.25}**

**\textbf{Your Name: Dennis Kwadzode \hfill Problem Set 1 Due: Feb. 18, 2025}\\**

**(1) The Poisson distribution has probability mass function**

**\begin{gather}**

**p(y\_i|\theta)=\frac{\theta^{y\_i}e^{-\theta}}{y\_i!},\qquad \theta>0,\qquad y\_i=0,1,\ldots**

**\end{gather}**

**and let $y\_1,\ldots,y\_n$ be random sample from this distribution.**

**\begin{enumerate}**

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

**\item Show that $\bar{y}$ is the MLE for $\theta$.**

**\item Write the mean of the posterior distribution as a weighted average of the mean of the prior distribution and the MLE.**

**\item What happens to the weight on the prior mean as $n$ becomes large?**

**\end{enumerate}**

**\begin{}**

**\textbf{SOLUTION}**

**\subsection\*{1. Conjugate Prior}**

**The likelihood based on the sample is**

**\[**

**L(\theta) = \prod\_{i=1}^n \frac{\theta^{y\_i}e^{-\theta}}{y\_i!} \propto \theta^{\sum\_{i=1}^n y\_i}e^{-n\theta}.**

**\]**

**Assume a Gamma prior for \( \theta \) with parameters \(\alpha\) and \(\beta\):**

**\[**

**p(\theta) \propto \theta^{\alpha-1}e^{-\beta\theta}, \quad \theta>0.**

**\]**

**Multiplying the likelihood and the prior gives the posterior (ignoring normalizing constants):**

**\[**

**p(\theta \mid y) \propto \theta^{\alpha-1+\sum\_{i=1}^n y\_i}e^{-(\beta+n)\theta}.**

**\]**

**Since this is the kernel of a Gamma distribution, we have**

**\[**

**\theta \mid y \sim \text{Gamma}\Bigl(\alpha+\sum\_{i=1}^n y\_i,\,\beta+n\Bigr).**

**\]**

**Thus, the Gamma prior is conjugate to the Poisson likelihood.**

**\subsection\*{2. Maximum Likelihood Estimate (MLE)}**

**The log-likelihood is**

**\[**

**\ell(\theta) = \sum\_{i=1}^n \left[ y\_i\ln\theta - \theta - \ln(y\_i!) \right]**

**= \left(\sum\_{i=1}^n y\_i\right)\ln\theta - n\theta + \text{const.}**

**\]**

**Differentiate with respect to \(\theta\):**

**\[**

**\frac{d\ell}{d\theta} = \frac{\sum\_{i=1}^n y\_i}{\theta} - n.**

**\]**

**Setting the derivative equal to zero:**

**\[**

**\frac{\sum\_{i=1}^n y\_i}{\theta} - n = 0 \quad\Longrightarrow\quad \hat{\theta} = \frac{1}{n}\sum\_{i=1}^n y\_i = \bar{y}.**

**\]**

**Thus, the MLE for \(\theta\) is \(\bar{y}\).**

**\subsection\*{3. Posterior Mean as a Weighted Average}**

**The posterior distribution is**

**\[**

**\theta \mid y \sim \text{Gamma}\Bigl(\alpha+\sum\_{i=1}^n y\_i,\,\beta+n\Bigr),**

**\]**

**so its mean is**

**\[**

**E[\theta \mid y] = \frac{\alpha+\sum\_{i=1}^n y\_i}{\beta+n}.**

**\]**

**Since \(\sum\_{i=1}^n y\_i = n\bar{y}\) and the prior mean is \(\mu\_0 = \alpha/\beta\), we can write:**

**\[**

**E[\theta \mid y] = \frac{\alpha}{\beta+n} + \frac{n\bar{y}}{\beta+n}**

**= \frac{\beta}{\beta+n}\left(\frac{\alpha}{\beta}\right) + \frac{n}{\beta+n}\bar{y}.**

**\]**

**Thus, the posterior mean is a weighted average of the prior mean \(\alpha/\beta\) and the MLE \(\bar{y}\), with weights \(\beta/(\beta+n)\) and \(n/(\beta+n)\), respectively.**

**\subsection\*{4. Behavior as \(n\) Increases}**

**As \(n\) becomes large,**

**\[**

**\frac{\beta}{\beta+n}\to 0 \quad\text{and}\quad \frac{n}{\beta+n}\to 1.**

**\]**

**Thus, for large \(n\) the posterior mean is dominated by the MLE \(\bar{y}\) and the influence of the prior diminishes.**

**%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%**

**\end{SOLUTION}**

**\newpage**

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

**\begin{center}**

**\begin{tabular}{ r r r r r r r }**

**\hline**

**$n$ & 1 & 2 & 3 & 4 & 5 & 6 \\**

**\hline**

**100 & 19 & 12 & 17 & 18 & 20 & 14 \\**

**1000 & 190 & 120 & 170 & 180 & 200 & 140 \\**

**\hline**

**\end{tabular}**

**\end{center}**

**Suppose you are interested in $\theta\_1$, 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 $\theta\_1$ 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.**

**\end{spacing}**

**\begin{}**

**\textbf{SOLUTION}**

**\section\*{Dirichlet Posterior for a Die Toss}**

**%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%**

**We are given the following data from tossing a die:**

**\[**

**\begin{array}{rcccccc}**

**\text{Outcome} & 1 & 2 & 3 & 4 & 5 & 6 \\[0.5em]**

**\text{Counts (n=100)} & 19 & 12 & 17 & 18 & 20 & 14 \\[0.5em]**

**\text{Counts (n=1000)} & 190 & 120 & 170 & 180 & 200 & 140 \\**

**\end{array}**

**\]**

**We are interested in \(\theta\_1\), the probability of obtaining a one spot. The prior for the probability vector \(\boldsymbol{\theta}=(\theta\_1,\ldots,\theta\_6)\) is a Dirichlet distribution with all parameters equal to 2:**

**\[**

**\boldsymbol{\theta} \sim \text{Dirichlet}(2,2,2,2,2,2).**

**\]**

**\subsection\*{Posterior Distribution}**

**Because the Dirichlet prior is conjugate to the multinomial likelihood, the posterior distribution is**

**\[**

**\boldsymbol{\theta} \mid y \sim \text{Dirichlet}(2+n\_1,\,2+n\_2,\,\dots,\,2+n\_6).**

**\]**

**\subsection\*{Marginal Posterior for \(\theta\_1\)}**

**The marginal posterior for \(\theta\_1\) is a Beta distribution:**

**\[**

**\theta\_1 \mid y \sim \text{Beta}\Bigl(2+n\_1,\,\sum\_{i\neq 1}(2+n\_i)\Bigr).**

**\]**

**\subsubsection\*{For \(n=100\):}**

**\[**

**n\_1 = 19 \quad\Longrightarrow\quad 2+n\_1 = 21.**

**\]**

**For the other outcomes:**

**\[**

**2+12=14,\quad 2+17=19,\quad 2+18=20,\quad 2+20=22,\quad 2+14=16.**

**\]**

**The second parameter is:**

**\[**

**14+19+20+22+16 = 91.**

**\]**

**Thus, for \(n=100\):**

**\[**

**\theta\_1 \mid y \sim \text{Beta}(21,91).**

**\]**

**\subsubsection\*{For \(n=1000\):}**

**\[**

**n\_1 = 190 \quad\Longrightarrow\quad 2+n\_1 = 192.**

**\]**

**For the other outcomes:**

**\[**

**2+120=122,\quad 2+170=172,\quad 2+180=182,\quad 2+200=202,\quad 2+140=142.**

**\]**

**The sum for the remaining outcomes is:**

**\[**

**122+172+182+202+142 = 820.**

**\]**

**Thus, for \(n=1000\):**

**\[**

**\theta\_1 \mid y \sim \text{Beta}(192,820).**

**\]**

**\subsection\*{Plotting the Posterior Densities}**

**Below is a plot comparing the posterior densities of \(\theta\_1\) for the two sample sizes. The blue curve corresponds to the \(n=100\) case (Beta(21,91)) and the red curve to the \(n=1000\) case (Beta(192,820)).**

**\begin{figure}[h]**

**\centering**

**\begin{tikzpicture}**

**\begin{axis}[**

**domain=0,0.4,**

**samples=100,**

**xlabel={\(\theta\_1\)},**

**ylabel={Density},**

**legend pos=north east,**

**title={Posterior Density for \(\theta\_1\)}**

**]**

**\addplot [blue, thick] { (x^(21-1)\*(1-x)^(91-1)) / (exp(lngamma(21)+lngamma(91)-lngamma(112)) ) };**

**\addlegendentry{\(\text{Beta}(21,91)\)};**

**\addplot [red, thick] { (x^(192-1)\*(1-x)^(820-1)) / (exp(lngamma(192)+lngamma(820)-lngamma(1012)) ) };**

**\addlegendentry{\(\text{Beta}(192,820)\)};**

**\end{axis}**

**\end{tikzpicture}**

**\caption{Posterior densities for \(\theta\_1\) for \(n=100\) (blue) and \(n=1000\) (red).}**

**\end{figure}**

**\subsection\*{Discussion}**

**Both posteriors are centered near the observed proportion of ones:**

**\[**

**\hat{\theta}\_1 \approx \frac{19}{100}=0.19 \quad \text{and} \quad \frac{190}{1000}=0.19.**

**\]**

**However, the posterior for \(n=1000\) is much more concentrated around 0.19, indicating a smaller variance due to the larger sample size. This illustrates that as the sample size increases, the data increasingly dominate the prior, resulting in a more precise (less variable) estimate of \(\theta\_1\).**

**\end{document}**