Mathematical Statistics II	**Winter 2023**
Assignment 3	
Handed Out: Dec 15 2023	*Due: Dec 22 2023*

- Feel free to talk to other students in the class when doing this assignment. You should, however, write down your solution yourself.
- Only homeworks **submitted in the tutorial of week 6** are graded.

Task 1. Consider observing data X. Assume you have a parametric model, which depends on a parameter θ. The likelihood function is defined over values of $\theta \in \{-3, -2, 1, 0, 1, 2, 3, 4\}$. Likelihood and prior distribution values are given in the table below.

θ	-3	-2	-1	0	1	2	3	4
$L(\theta\|X)$	0.5	2	1	3	1	3	2	0.5
$\pi(\theta)$	0.1	0.3	0.05	0.15	0.05	0.1	0.2	0.05

(a) Compute the posterior distribution over a valid range of θ.

(b) Compute the posterior mean.

Solution.

(a) We know that

$$\pi(\theta|X) = \frac{L(\theta|X) \cdot \pi(\theta)}{\int L(\theta|X) \cdot \pi(\theta) d\theta}.$$

Using the table above leads to

θ	-3	-2	-1	0	1	2	3	4	$\sum$
$L(\theta\|X)$	0.5	2	1	3	1	3	2	0.5	
$\pi(\theta)$	0.1	0.3	0.05	0.15	0.05	0.1	0.2	0.05	
$L(\theta\|X) \cdot \pi(\theta)$	0.05	0.6	0.05	0.45	0.05	0.3	0.4	0.025	1.85
$\pi(\theta\|X)$	0.027	0.324	0.027	0.243	0.027	0.162	0.216	0.014	

(b) The posterior mean is given by

$$\begin{aligned}\int_\Theta \theta\pi(\theta|X)d\theta &= \sum_\theta \theta\pi(\theta|X) \\ &= (-3 \cdot 0.027) + (-2 \cdot 0.324) + (-1 \cdot 0.027) + (0 \cdot 0.243) \\ &\quad + (1 \cdot 0.027) + (2 \cdot 0.162) + (3 \cdot 0.216) + (4 \cdot 0.014) \\ &\approx -0.303.\end{aligned}$$

Task 2. A researcher is trying to estimate the number of accidents per month within 100 feet of the Gervais Street/Assembly Street intersection in Columbia. She assumes a Poisson(λ) model for the number of accidents Y per month, so that the density function for Y given λ is

$$p(y|\lambda) = \frac{\lambda^y e^{-\lambda}}{y!}, \quad y = 0, 1, 2, \ldots, \ \lambda \geq 0$$

(a) She uses a standard exponential prior distribution for λ, i.e.,

$$\pi(\lambda) = \begin{cases} e^{-\lambda} & \lambda \geq 0 \\ 0 & \lambda < 0. \end{cases}$$

Derive the general form of a posterior distribution for λ given a random sample $y_1, \ldots, y_n$ for n months. Note that $X \sim \Gamma(\alpha, \beta)$ if the pdf is given by

$$f_{\alpha,\beta}(x) = \frac{x^{\alpha-1} e^{-\beta x} \beta^\alpha}{(\alpha - 1)!} = c \cdot x^{\alpha-1} \cdot e^{-\beta x}$$

for $\alpha > 0$. Then $EX = \alpha/\beta$.

(b) If she gathers the following accident counts from 15 randomly selected months

1 0 4 1 4 2 5 3 0 3 1 2 2 4 1

find the posterior mean and the 95% credible interval for λ using the standard exponential prior, along with these data.

Solution.

(a) The posterior distribution is given by Bayes' theorem:

$$\pi(\lambda|y_1, \ldots, y_n) \propto \pi(\lambda) \cdot \prod_{i=1}^{n} p(y_i|\lambda)$$

Given that the prior is an exponential distribution with parameter 1:

$$\pi(\lambda) = e^{-\lambda} \quad \text{for } \lambda \geq 0$$

And the likelihood function is Poisson:

$$p(y_i|\lambda) = \frac{\lambda^{y_i} e^{-\lambda}}{y_i!}$$

Now, multiply the prior and likelihood:

$$\pi(\lambda|y_1, \ldots, y_n) \propto e^{-\lambda} \cdot \prod_{i=1}^{n} \frac{\lambda^{y_i} e^{-\lambda}}{y_i!} \qquad = e^{-n\lambda} \cdot \frac{\lambda^{\sum_{i=1}^n y_i} e^{-\lambda n}}{\prod_{i=1}^n y_i!} \quad = \frac{\lambda^{\sum_{i=1}^n y_i} e^{-\lambda(n+1)}}{\prod_{i=1}^n y_i!}$$

Therefore, the posterior distribution is a Gamma distribution:

$$\pi(\lambda|y_1, \ldots, y_n) \sim \text{Gamma}\left(\sum_{i=1}^{n} y_i + 1, n + 1\right)$$

(b) For the given data:

$$y = 1, 0, 4, 1, 4, 2, 5, 3, 0, 3, 1, 2, 2, 4, 1.$$

The sum of the data is:

$$\sum_{i=1}^{15} y_i = 38.$$

So, the posterior distribution is a Gamma distribution with parameters:

$$\alpha = \sum_{i=1}^{15} y_i + 1 = 39$$
$$\beta = n + 1 = 15 + 1 = 16$$

Now, the posterior mean of a Gamma distribution with shape α and scale β is given by:

$$\text{Posterior mean} = \int \lambda \pi(\lambda | y_1, \ldots, y_n) d\lambda = \frac{\alpha}{\beta} = \frac{39}{16} \approx 2.44$$

Using

```
alpha_posterior <- 39
beta_posterior <- 16

credible_interval <- qgamma(c(0.025, 0.975),
shape=alpha_posterior, rate=beta_posterior)
```

leads to a 95% credible interval of the form

$$[1.733, 3.260]$$

Task 3. Consider the coin flipping problem, in which we treat the probability of heads θ as a random variable sampled from some prior distribution $\pi(\theta)$. We represent the ith coin flip by a random variable $X_i \in \{0, 1\}$, where $X_i = 1$ if the ith flip is a head. We assume that the X_i's are conditionally independent given θ.

(a) Suppose our prior distribution on θ is Beta(h, t) for some $h, t > 0$ and that our sequence of flips $\mathcal{D}$ has n_h heads and n_t tails. Show that the posterior distribution of θ is Beta$(h + n_h, t + n_t)$.

(b) Give expressions of the MAP and the posterior mean estimate of θ. You may use the fact that a Beta(a, b) distribution has mean $a/(a + \beta)$ and mode (i.e. the most likely value of the distribution) $(\alpha - 1)/(\alpha + \beta - 2)$.

(c) What happens to the MAP and the posterior mean estimate as the number of coin flips $n = n_h + n_t$ approaches infinity?

Solution.

(a) We know that

$$\pi(\theta|\mathcal{D}) \propto L(\theta|\mathcal{D})\pi(\theta).$$

As $\theta \sim \text{Beta}(h, t)$, its pdf is given by

$$\pi(\theta) = c \cdot \theta^{h-1}(1-\theta)^{t-1}.$$

Further we know that X_i describes the ith coin flip, therefore $X_i \sim Ber(\theta)$. Thus $\mathcal{D} \sim B(n_h + n_t, \theta)$ with corresponding likelihood function

$$L(\theta|\mathcal{D}) = \binom{n_h + n_t}{n_h} \theta^{n_h}(1-\theta)^{n_t}.$$

Then we get

$$\begin{aligned}\pi(\theta|\mathcal{D}) \propto L(\theta|\mathcal{D})\pi(\theta) &\propto \theta^{n_h}(1-\theta)^{n_t} \cdot \theta^{h-1}(1-\theta)^{t-1} \\ &= \theta^{h+n_h-1}(1-\theta)^{n_t+t-1} \\ &\sim \text{Beta}(h+n_h, t+n_t).\end{aligned}$$

(b) By definition the posterior mean is

$$\hat{\theta} = \int_\Theta \theta\pi(\theta|X)d\theta = E(\theta|X).$$

From (a) we know that $\theta|X \sim \text{Beta}(h+n_h, t+n_t)$. Using the hint we get

$$E(\theta|X) = \frac{h+n_h}{h+n_h+t+n_t}.$$

By definition the posterior mean is given by

$$\hat{\theta}_{MAP} \in \arg\max_{\theta\in\Theta} \pi(\theta|X).$$

Using the hint about the mode of the Beta distribution we get

$$\hat{\theta} = \frac{h+n_h-1}{h+n_h+t+n_t-2}.$$

(c) The posterior mean is given by

$$\hat{\theta} = \frac{h+n_h}{h+n_h+t+n_t}.$$

As $n = n_h + n_t$ becomes large, the term n dominates the expression. The impact of the prior parameters h and t diminishes and the posterior mean converges to n_h/n. The same happens to the MAP estimator.

Grading:

	1		2		3			Total
	(a)	(b)	(a)	(b)	(a)	(b)	(c)	
Points	1	1	2	1	1	0.5	0.5	7