HOMEWORK 5

[Wei] 3.17. In this problem, we write the likelihood function as

$$L(\theta|\mathbf{X}) = \prod_{i=1}^{n} \mathbf{1}_{(\theta-1/2,\theta+1/2)}(X_i) = \mathbf{1}_{(X_{(n)}-1/2,X_{(1)}+1/2)}(\theta)$$

So, for any $0 < \lambda < 1$,

$$\hat{\theta}^*(\mathbf{X}) = \lambda(X_{(n)} - \frac{1}{2}) + (1-\lambda)(X_{(1)} + \frac{1}{2})$$

is an MLE estimator of θ.

[Wei] 7.3. In this problem, suppose that the samples $X_1, \ldots, X_8$ *i.i.d.* $\sim$ *Bernoulli*(θ), where $X_i = 1$ if the product is useless, else $X_i = 0$. Then the observation $X = \sum_i X_i \sim B(8, \theta)$ denotes the number of useless products in those 8 samples.

Since $\theta = 0.1$ or 0.2 alternatively, we only need to calculate one of the posterior probability. For example,

$$\begin{aligned}\pi(\theta = 0.1|X = 2) &= \frac{P(X = 2|\theta = 0.1)\pi(0.1)}{\sum_{i=0.1,0.2} P(X = 2|\theta = i)\pi(i)} \\ &= \frac{\binom{8}{2}0.1^2 0.9^6 0.7}{\binom{8}{2}0.1^2 0.9^6 0.7 + \binom{8}{2}0.2^2 0.8^6 0.3} \\ &= 0.5418.\end{aligned}$$

Therefore, we derive that

$$\begin{aligned}\pi(\theta = 0.2|X = 2) &= 1 - \pi(\theta = 0.1|X = 2) \\ &= 0.4582.\end{aligned}$$

$\square$

[Wei] 7.4. In this problem, suppose the observation $X \sim P(\lambda)$ denotes the number of errors in a record.

Since $\lambda = 1.0$ or 1.5 alternatively, we only need to calculate one of the posterior probability. For example,

$$\begin{aligned}\pi(\lambda = 1.0|X = 3) &= \frac{P(X = 3|\lambda = 1.0)\pi(1.0)}{\sum_{i=1.0,1.5} P(X = 3|\lambda = i)\pi(i)} \\ &= \frac{e^{-1.0}\frac{1.0^3}{3!}0.4}{e^{-1.0}\frac{1.0^3}{3!}0.4 + e^{-1.5}\frac{1.5^3}{3!}0.6} \\ &= 0.2457.\end{aligned}$$

Thanks to Weiyu Li is a graduate of the School of the Gifted Young, University of Science and Technology of China, Hefei, Anhui 230026, China. Corresponding Email: liweiyu@mail.ustc.edu.cn.

Therefore, we derive that

$$\begin{aligned}\pi(\lambda = 1.5|X = 3) &= 1 - \pi(\lambda = 1.0|X = 3)\\ &= 0.7543.\end{aligned}$$

□

[Wei] 7.5. (1) We first calculate the kernel

$$\begin{aligned}\pi(\theta|x) &\propto p(x|\theta)\pi(\theta)\\ &\propto \frac{1}{\theta^2} I_{(x,1)}(\theta) I_{(0,1)}(\theta)\\ &\propto \frac{1}{\theta^2} I_{(x,1)}(\theta),\end{aligned}$$

where we use the fact that $2x$ is constant with respect to θ for the second line and $x \in (0,1)$ for the third line. Thus we get $\pi(\theta|x) = c(x)\frac{1}{\theta^2} I_{(x,1)}(\theta)$.

From

$$\int_x^1 \frac{1}{\theta^2} = \frac{1}{x} - 1,$$

and the normalization condition $\int \pi(\theta|x) = 1$, we have the posterior distribution of θ

$$\pi(\theta|x) = \frac{x}{(1-x)\theta^2} I_{(x,1)}(\theta), \quad 0 < x < 1.$$

□

(2) Similarly, first calculate the kernel

$$\begin{aligned}\pi(\theta|x) &\propto p(x|\theta)\pi(\theta)\\ &\propto \frac{1}{\theta^2} I_{(x,1)}(\theta) 3\theta^2 I_{(0,1)}(\theta)\\ &\propto 3I_{(x,1)}(\theta),\end{aligned}$$

where we use the fact that $x \in (0,1)$ for the last line. Thus we get $\pi(\theta|x) = c(x)3I_{(x,1)}(\theta)$.

From

$$\int_x^1 3 = 3(1-x),$$

and the normalization condition $\int \pi(\theta|x) = 1$, we obtain the posterior distribution

$$\pi(\theta|x) = \frac{1}{(1-x)} I_{(x,1)}(\theta), \quad 0 < x < 1.$$

□

[Wei] 7.12. Let $X \sim B(100, \theta)$ denotes the number of useless products in 100 samples, then the kernel of the posterior distribution is

$$\begin{aligned}\pi(\theta|X = 3) &\propto f_{X|\theta}(3|\theta)\pi(\theta)\\ &\propto \theta^3(1-\theta)^{100-3}\theta^{2-1}(1-\theta)^{200-1} I_{(0,1)}(\theta)\\ &\propto \theta^4(1-\theta)^{296} I_{(0,1)}(\theta).\end{aligned}$$

Notice that the kernel is the same as that of $Y \sim Be(5, 297)$. We conclude that
$$\pi(\theta|X = 3) \sim Be(5, 297).$$
$\square$

[Wei] 7.13. (1) Suppose $\pi(\theta) \sim N(\mu, \sigma^2)$, $\sigma^2 > 0$. Then the kernel of the posterior distribution is
$$\begin{aligned}
\pi(\theta|\mathbf{x}) &\propto f(\mathbf{x}|\theta)\pi(\theta) \\
&\propto \exp\left\{-\frac{\sum_i (x_i-\theta)^2}{2\times 2^2}\right\}\exp\left\{-\frac{(\theta-\mu)^2}{2\sigma^2}\right\} \\
&\propto \exp\left\{-\theta^2\left(\frac{n}{2\times 2^2}+\frac{1}{2\sigma^2}\right)+b(\mathbf{x},\mu,\sigma^2)\theta\right\}.
\end{aligned}$$
Notice that the kernel is the same as that of a normal distribution with variance $\left(\frac{n}{4}+\frac{1}{2\sigma^2}\right)^{-1} < \frac{4}{n} = \frac{1}{25}$. Therefore we conclude that the standard deviation must be less than 1/5. $\square$

(2) With $\sigma^2 = 1$ and the conclusion in (1), we require the sample size n satisfying
$$\left(\frac{n}{4}+\frac{1}{2\sigma^2}\right)^{-1} \le 0.1.$$
The solution of the equation is $n \geq 36$. Thus the sample size should be at least 36. $\square$

[Wei] 7.8. If the prior distribution is $\lambda \sim \Gamma(\alpha, \beta)$, then the kernel of the posterior distribution is
$$\begin{aligned}
\pi(\lambda|\mathbf{x}) &\propto f_{\mathbf{X}|\lambda}(\mathbf{x})\pi(\lambda) \\
&\propto \prod_{i=1}^{n}\left(\lambda e^{-\lambda x_i}\right)\times \lambda^{\alpha-1}e^{-\beta\lambda} \\
&\propto \lambda^{n+\alpha-1}e^{-\lambda(\beta+\sum_i x_i)}.
\end{aligned}$$
Notice that the kernel is the same as that of $Y \sim \Gamma(\alpha+n, \beta+\sum_i x_i)$. We conclude that the conjugate prior distribution family of λ is gamma distribution family. $\square$

[Wei] 7.11. We say $\theta \sim Pareto(\theta_0, \alpha)$ if the density function of θ is in the form in the problem. If the prior distribution is $\theta \sim Pareto(\theta_0, \alpha)$, then the kernel of the posterior distribution is
$$\begin{aligned}
\pi(\theta|\mathbf{x}) &\propto f_{\mathbf{X}|\theta}(\mathbf{x})\pi(\theta) \\
&\propto \frac{1}{\theta^n}I(0 < x_{(1)} \le x_{(n)} < \theta)\times\frac{1}{\theta^{\alpha+1}}I_{(\theta_0,+\infty)}(\theta) \\
&\propto \frac{1}{\theta^{n+\alpha+1}}I_{(\tilde{\theta}_0,+\infty)}(\theta),
\end{aligned}$$
where $\tilde{\theta}_0 = \max\{\theta_0, x_{(n)}\}$. Notice that the kernel is the same as that of $Pareto(\tilde{\theta}_0, \alpha+n)$. We conclude that the conjugate prior distribution family of θ is Pareto distribution. $\square$