200B HW#3 solution

7.2 Prior and Posterior Distribution

2. The likelihood function is

$$\tilde{L}(\theta) = f_n(\mathbf{x}|\theta) = \theta^{\sum x_i}(1-\theta)^{n-\sum x_i} = \theta^2(1-\theta)^6$$

Therefore,

$$\begin{aligned}
\xi(0.1|x) = P(\theta = 0.1|x) &= \frac{\xi(0.1)f_n(\mathbf{x}|0.1)}{\xi(0.1)f_n(\mathbf{x}|0.1) + \xi(0.2)f_n(\mathbf{x}|0.2)} \\
&= \frac{(0.7)(0.1)^2(0.9)^6}{(0.7)(0.1)^2(0.9)^6 + (0.3)(0.2)^2(0.8)^6} \\
&= 0.5418.
\end{aligned}$$

And $\xi(0.2|x) = 1 - \xi(0.1|x) = 0.4582$.

10. The p.d.f. of X is

$$f_n(x|\theta) = \begin{cases} 1 & \text{for } \theta - 1/2 \le x \le \theta + 1/2, \\ 0 & \text{otherwise.} \end{cases}$$

and the prior p.d.f. of θ is

$$\xi(\theta) = \begin{cases} \frac{1}{10} & \text{for } 10 \le \theta \le 20, \\ 0 & \text{otherwise.} \end{cases}$$

The condition that $\theta - 1/2 \le x \le \theta + 1/2$ is the same as the conditions that $x - 1/2 \le \theta \le x + 1/2$. Therefore, $f(x|\theta)\xi(\theta)$ is a positive constant only for values of θ which satisfy both conditions that $x - 1/2 \le \theta \le x + 1/2$ and $10 \le \theta \le 20$. Since $X = 12$, $\xi(\theta|x) \propto f(x|\theta)\xi(\theta)$ is a positive constant only for $11.5 \le \theta \le 12.5$. In other words, the posterior distribution of θ is a uniform distribution on the interval $[11.5, 12.5]$.

7.3 Conjugate Prior Distributions

18. Suppose that the prior distribution of θ is the Pareto distribution with parameters x_0 and α ($x_0 > 0$ and $\alpha > 0$). Then the prior p.d.f. $\xi(\theta)$ has the form

$$\xi(\theta) \propto 1/\theta^{\alpha+1} \text{ for } \theta \geq x_0.$$

If $X_1, X_2, \ldots, X_n$ form a random sample from uniform distribution on $[0, \theta]$, the likelihood function is

$$f_n(\mathbf{x}|\theta) \propto 1/\theta^n \text{ for } \theta \geq max\{x_1, \ldots, x_n\}.$$

Hence, the posterior p.d.f. of θ has the form

$$\xi(\theta|\mathbf{x}) \propto \xi(\theta) f_n(\mathbf{x}|\theta) \propto 1/\theta^{\alpha+n+1} \text{ for } \theta \geq max\{x_0, x_1, \ldots, x_n\}$$

We recognize the posterior distribution of θ as Pareto distribution with parameters $max\{x_0, x_1, \ldots, x_n\}$ and $\alpha + n$.

19. Since the prior distribution of θ is a gamma distribution with parameters α and β,

$$\xi(\theta) \propto \theta^{\alpha-1} e^{-\beta\theta} \text{ for } \theta > 0,$$

the likelihood function is

$$f_n(\mathbf{x}|\theta) \propto \theta^n (\prod_{i=1}^{n} x_i)^\theta,$$

so that the posterior distribution is

$$\xi(\theta|\mathbf{x}) \propto \theta^{\alpha+n-1} e^{-(\beta - \sum_{i=1}^n \log x_i)\theta} \text{ for } \theta > 0$$

and it is recognized as a gamma distribution with parameters $\alpha + n$, $\beta - \sum_{i=1}^n \log x_i$. Then mean and variances of the posterior distribution of θ are given by
$E(\theta|\mathbf{x}) = \frac{\alpha+n}{\beta - \sum_{i=1}^n \log x_i}$ and $\text{Var}(\theta|\mathbf{x}) = \frac{\alpha+n}{(\beta - \sum_{i=1}^n \log x_i)^2}$ respectively.

23.
The likelihood function is

$$f_n(\mathbf{x}|\theta) \propto a(\theta)^n \prod_{i=1}^{n} b(x_i) \exp[c(\theta)(\sum_{i=1}^{n} d(x_i))]$$

and the prior distribution of θ is

$$\xi(\theta) \propto a(\theta)^{\alpha} \exp[c(\theta)\beta].$$

So the posterior is

$$\xi(\theta|\mathbf{x}) \propto a(\theta)^{\alpha+n} \prod_{i=1}^{n} b(x_i) \exp[c(\theta)(\beta + \sum_{i=1}^{n} d(x_i))] \propto a(\theta)^{\alpha+n} \exp[c(\theta)(\beta + \sum_{i=1}^{n} d(x_i))].$$

$\prod_{i=1}^{n} b(x_i)$ can be omitted because we only need to look at terms including θ and $b(x_i)$ can be treated as constant.
So it is a conjugate family and posterior hyper parameters are $\alpha_1 = \alpha_0 + n$, $\beta_1 = \beta_0 + \sum_{i=1}^{n} d(x_i)$, where $(\alpha_1, \beta_1) \in H$ is guaranteed because the joint distribution calculated from marginal and conditional distribution has integral 1.

7.4 Bayes Estimators

2. Based on Theorem 7.3.1, the posterior distribution of θ is the beta distribution with parameters $5 + 1 = 6$ and $10 + 19 = 29$. Therefore, the Bayes estimate of θ is the mean of this posterior distribution, which is $6/(1 + 29) = 6/35$.
10. Let α and β denote the parameters of the prior gamma distribution of θ. Then $\alpha/\beta = 0.2$ and $\alpha/\beta^2 = 1$. Therefore, $\beta = 0.2$ and $\alpha = 0.04$. Furthermore, the total time required to serve the sample of 20 customers is $y = 20(3.8) = 76$. Therefore, by Theorem 7.3.4, the posterior distribution of θ is the gamma distribution with parameters $0.04 + 20 = 20.04$ and $0.2 + 76 = 76.2$. The Bayes estimate under the square error loss is the mean of this posterior distribution and is equal to $20.04/76.2 = 0.263$.
12.

(a) A's Prior distribution of θ is the beta distribution with parameters $\alpha = 2$ and $\beta = 1$. Based on Theorem 7.3.1, A's posterior distribution for θ is the beta distribution with parameters $2+710=712$ and $1+290=291$. B's prior distribution for θ is a beta distribution with parameters $\alpha=4$ and $\beta=1$. Similarly, B's posterior distribution for θ is the beta distribution with parameters $4+710=714$ and $1+290=291$.

(b) Under squared error loss, A's Bayes estimate of θ is $712/(712+291)=712/1003$. B's Bayes estimate of θ is $714/(714+291)=714/1005$.

(c) If y denotes the number in the sample who were in favor of the proposition, then A's posterior distribution for θ will be the beta distribution with parameters $2+y$ and $1+1000-y=1001-y$, and B's posterior distribution for θ will be a beta distribution with parameters $4+y$ and $1+1000-y=1001-y$. Therefore, A's Bayes estimate of θ will be $(2+y)/1003$ and B's Bayes estimate of θ will be $(4+y)/1005$. Then

$$|\frac{4+y}{1005}-\frac{2+y}{1003}|=\frac{2(1001-y)}{(1005)(1003)}.$$

This difference is a maximum when $y=0$, but even then its value is only

$$\frac{2(1001)}{(1005)(1003)}<\frac{2}{1000}.$$

7.10 Supplementary Exercises

9. It is easy to know the posterior distribution is

$$\xi(\theta|x)\propto e^{-\theta},\ \text{for } \theta>x.$$

By $\int_x^{+\infty}e^{-\theta}d\theta=e^{-x}$ we know $\xi(\theta|x)=e^{x}e^{-\theta}$, for $\theta>x$.
The Bayesian estimator with respect to mean squared error loss is the mean of posterior distribution

$$\hat{\theta}_{mse}=E_{\theta\sim\xi(\theta|x)}(\theta)=\int_x^{+\infty}\theta e^{x}e^{-\theta}d\theta=-\int_x^{+\infty}\theta e^{x}de^{-\theta}=x+\int_x^{+\infty}e^{x}e^{-\theta}d\theta=x+1$$

The Bayesian estimator with respect to mean absolute error loss is the median of posterior distribution so we have

$$\int_{x}^{\hat{\theta}_{mae}} e^{x}e^{-\theta}d\theta = \frac{1}{2},$$

from which we can solve $\hat{\theta}_{mae} = x + \log 2$.

19. The likelihood function is given by,

$$f(x|n) = \binom{n}{x} p^{x}(1-p)^{n-x}$$

Now, following the hint, we have,

$$\frac{f(x|n+1)}{f(x|n)} = \frac{n+1}{n+1-x}(1-p).$$

The ratio is less than 1 when $n > \frac{x}{p} - 1$, so choosing $n+1$ in such a case would not give you a maximum likelihood estimate. Thus, given x and p, the maximum likelihood estimate of n is $\lfloor \frac{x}{p} \rfloor$ (rounding down) unless $\frac{x}{p}$ is an integer, in which case the two likelihoods, $f(x|n+1)$ and $f(x|n)$ are equal, giving (non-unique) MLEs $\frac{x}{p} - 1$ and $\frac{x}{p}$.

(Here $\lfloor a \rfloor$ is the largest integer less or equal to a.)