# Biostat 803 Homework 2

David (Daiwei) Zhang

February 12, 2018

## 1 Exponential distribution with a change point

### 1.1 Posterior for $\xi$

For a variable $\theta$, let $d\theta = \mu(d\theta)$ where $\mu$ is the Lebesgue measure. We have

$$\begin{aligned}\pi(\boldsymbol{x}|\xi,\eta) =& \prod_{i}^{\xi} \eta \exp\{-\eta x_i\} \prod_{\xi+1}^{n} c\eta \exp\{-c\eta x_i\} \\ =& \eta^n c^{n-\xi} \exp\Big\{ -\eta\Big[\sum_{i}^{\xi} x_i + c\sum_{\xi+1}^{n} x_i\Big]\Big\},\end{aligned}$$

so

$$\begin{aligned}\pi(\xi,\eta|\boldsymbol{x}) \propto& \pi(\boldsymbol{x}|\xi,\eta)\pi(\xi,\eta) \\ =& \pi(\boldsymbol{x}|\xi,\eta)\pi(\xi)\pi(\eta) \\ =& \eta^n c^{n-\xi} \exp\Big\{ -\eta\Big[\sum_{i}^{\xi} x_i + c\sum_{\xi+1}^{n} x_i\Big]\Big\}\pi(\xi)\pi(\eta),\end{aligned}$$

and

$$\begin{aligned}\pi(\xi|\boldsymbol{x}) =& \int \pi(\xi,\eta|\boldsymbol{x})d\eta \\ \propto& \int \eta^n c^{n-\xi} \exp\Big\{ -\eta\Big[\sum_{i}^{\xi} x_i + c\sum_{\xi+1}^{n} x_i\Big]\Big\}\pi(\xi)\pi(\eta)d\eta \\ =& \frac{c^{n-\xi} n! \pi(\xi)}{\Big[\sum_1^{\xi} x_i + c\sum_{\xi+1}^{n} x_i\Big]^{n+1}} \\ \propto& \frac{c^{n-\xi} \pi(\xi)}{\Big[\sum_1^{\xi} x_i + c\sum_{\xi+1}^{n} x_i\Big]^{n+1}}\end{aligned}$$

$$
\begin{aligned}
\pi(\boldsymbol{x}|\xi) =& \int \pi(\boldsymbol{x}|\xi,\eta)\pi(\eta)d\eta \\
\propto& \int \pi(\boldsymbol{x}|\xi,\eta)d\eta \\
=& \frac{c^{n-\xi}n!}{[\sum_1^\xi x_i + c\sum_{\xi+1}^n x_i]^{n+1}} \\
=& \frac{n!c^{n-\xi}}{x_1^{n+1}}[\sum_1^\xi z_i + c\sum_{\xi+1}^n z_i]^{-n-1} \\
\propto& \frac{c^{n-\xi}}{x_1^{n+1}}[\sum_1^\xi z_i + c\sum_{\xi+1}^n z_i]^{-n-1}
\end{aligned}
$$

and thus

$$
\begin{aligned}
\pi(\xi|\boldsymbol{x}) \propto& \pi(\boldsymbol{x}|\xi)\pi(\xi) \\
\propto& \frac{\pi(\xi)}{x_1^{n+1}}c^{n-\xi}[\sum_1^\xi z_i + c\sum_{\xi+1}^n z_i]^{-n-1} \\
\propto& \pi(\xi)c^{n-\xi}[\sum_1^\xi z_i + c\sum_{\xi+1}^n z_i]^{-n-1}
\end{aligned}
$$

### 1.2 $\boldsymbol{Z}$ is ancillary with respect to $\eta$

We know that the exponential distribution family is a location family with respect to $\eta$, so for $1 \leq i \leq \xi$, the distribution of $W_i = \eta X_i$ does not depend on $\eta$. Then $Z_i = W_i/W_1$ does not depend on $\eta$ for $2 \leq i \leq \xi$. Similarly, for $\xi + 1 \leq i \leq n$, the distribution of $W_i = c\eta X_i$ does not depend on $\eta$. Then $Z_i = W_i/(cW_1)$ does not depend on $\eta$ for $\xi+1 \leq i \leq n$. Thus $\boldsymbol{Z} = (Z_2, \ldots, Z_n)$ is ancillary wrt $\eta$.

### 1.3 Likelihood for $\xi$ under $Z$

We transform $(x_1, \ldots, x_n)$ to $(x_1, z_2, \ldots, z_n)$. We have

$$
\begin{aligned}
\pi(x_1, z_2, \ldots, z_n|\xi,\eta) =& \pi(x_1, z_2, \ldots, z_n|\xi,\eta)|\boldsymbol{J}| \\
=& \eta^n c^{n-\xi} \exp\Big\{-\eta\Big[\sum_i^\xi x_i + c\sum_{\xi+1}^n x_i\Big]\Big\}x^{n-1} \\
=& \eta^n c^{n-\xi} \exp\Big\{-\eta x_1\Big[\sum_i^\xi z_i + c\sum_{\xi+1}^n z_i\Big]\Big\}x^{n-1},
\end{aligned}
$$

since

$$\boldsymbol{J} = \begin{bmatrix} \frac{\partial x_1}{\partial x_1} & \frac{\partial x_1}{\partial z_2} & \cdots & \frac{\partial x_1}{\partial z_n} \\ \frac{\partial x_2}{\partial x_1} & \frac{\partial x_2}{\partial z_2} & \cdots & \frac{\partial x_2}{\partial z_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial x_1} & \frac{\partial x_n}{\partial z_2} & \cdots & \frac{\partial x_n}{\partial z_n} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ z_2 & x_1 & 0 & \cdots & 0 \\ z_3 & 0 & x_1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ z_n & 0 & 0 & \cdots & 0 \end{bmatrix},$$

so $|\boldsymbol{J}| = x_1^{n-1}$. Then

$$\begin{aligned} \pi(z_2, \ldots, z_n|\xi, \eta) = & \int \pi(x_1, z_2, \ldots, z_n|\xi, \eta) dx_1 \\ = & \int \eta^n c^{n-\xi} \exp \Big\{ -\eta x_1 \Big[ \sum_i^{\xi} z_i + c \sum_{\xi+1}^{n} z_i \Big] \Big\} x^{n-1} dx_1 \\ = & \eta^n c^{n-\xi} n! \eta^{-n} \Big[ \sum_i^{\xi} z_i + c \sum_{\xi+1}^{n} z_i \Big]^{-n} \\ \propto & c^{n-\xi} \Big[ \sum_i^{\xi} z_i + c \sum_{\xi+1}^{n} z_i \Big]^{-n}, \end{aligned}$$

so $\pi(\boldsymbol{z}|\eta) = \pi(z_2, \ldots, z_n|\eta) = \pi(z_2, \ldots, z_n|\xi, \eta)$. Notice that $\pi(z_2, \ldots, z_n|\xi, \eta)$ does not depend on $\eta$, so this is another way to show that $(z_2, \ldots, z_n)$ is an ancillary statistic with respect to $\eta$. Furthermore,

$$\begin{aligned} \pi(\xi|\boldsymbol{z}) \propto & \pi(\boldsymbol{z}|\eta)\pi(\xi) \\ = & c^{n-\xi} \Big[ \sum_i^{\xi} z_i + c \sum_{\xi+1}^{n} z_i \Big]^{-n} \pi(\xi) \end{aligned}$$

### 1.4 Reconciliation of the two likelihoods

Under $\pi(\eta) \propto 1$, we have $\pi(\xi|\boldsymbol{z})/\pi(\xi|\boldsymbol{x}) = \sum_i^\xi z_i + c\sum_{\xi+1}^n z_i \neq 1$ in general, so the two posteriors cannot reconcile. However, if we use $\pi(\eta) \propto \eta^{-1}$, then

$$\begin{aligned}
\pi(\xi|\boldsymbol{x}) \propto &\int \eta^n c^{n-\xi} \exp\Big\{-\eta\Big[\sum_i^\xi x_i + c\sum_{\xi+1}^n x_i\Big]\Big\}\pi(\xi)\pi(\eta)d\eta \\
=&\int \eta^n c^{n-\xi} \exp\Big\{-\eta\Big[\sum_i^\xi x_i + c\sum_{\xi+1}^n x_i\Big]\Big\}\eta^{-1}\pi(\xi)d\eta \\
=&\int \eta^{n-1} c^{n-\xi} \exp\Big\{-\eta\Big[\sum_i^\xi x_i + c\sum_{\xi+1}^n x_i\Big]\Big\}\pi(\xi)d\eta \\
=&\frac{c^{n-\xi}n!\pi(\xi)}{\Big[\sum_1^\xi x_i + c\sum_{\xi+1}^n x_i\Big]^n} \\
=&\frac{c^{n-\xi}n!\pi(\xi)}{x_1^n\Big[\sum_1^\xi z_i + c\sum_{\xi+1}^n z_i\Big]^n} \\
\propto&\frac{c^{n-\xi}\pi(\xi)}{\Big[\sum_1^\xi z_i + c\sum_{\xi+1}^n z_i\Big]^n} \\
=&\pi(\xi|\boldsymbol{z}).
\end{aligned}$$

## 2 Uniform distribution with Pareto prior

### 2.1 Posterior

Let $X_{(n)} = \max(X_i)$. For the likelihood, we have

$$\pi(\boldsymbol{x}|\theta) = \theta^{-n}I[X_{(n)} \leq \theta],$$

so the posterior is

$$\begin{aligned}
\pi(\theta|\boldsymbol{x}) \propto&\pi((x)|\theta)\pi(\theta) \\
\propto&\theta^{-n}I[X_{(n)} \leq \theta]\alpha\beta^\alpha\theta^{-\alpha-1}I[\beta \leq \theta] \\
=&\alpha\beta^\alpha\theta^{-a-n-1}I[\theta \geq \tilde{\beta}]
\end{aligned}$$

where $\tilde{\beta} = \max(X_{(n)}, \beta)$. To find the normalizing constant, we have

$$\begin{aligned}
\pi(\boldsymbol{x}) &= \int \pi(\boldsymbol{x}|\theta)\pi(\theta)d\theta \\
&= \int \alpha\beta^{\alpha}\theta^{-a-n-1}I[\theta \geq \tilde{\beta}]d\theta \\
&= \alpha\beta^{\alpha}\int_{\tilde{\beta}}^{\infty}\theta^{-a-n-1}d\theta \\
&= \alpha\beta^{\alpha}(\alpha+n)^{-1}\tilde{\beta}^{-\alpha-n}.
\end{aligned}$$

Then

$$\begin{aligned}
\pi(\theta|\boldsymbol{x}) &= \frac{\pi(\boldsymbol{x}|\theta)\pi(\theta)}{\pi(\boldsymbol{x})} \\
&= (a+n)\tilde{\beta}^{\alpha+n}\theta^{-a-n-1}I[\theta \geq \tilde{\beta}] \\
&\sim PA(\alpha+n, \max(X_{(n)}, \beta)).
\end{aligned}$$

## 2.2 Bayes estimator

For the Bayes estimator under the square loss function, we have the posterior mean

$$\begin{aligned}
\hat{\theta}_{Bayes} &= E[\theta|\boldsymbol{x}] \\
&= \int \theta(a+n)\tilde{\beta}^{\alpha+n}\theta^{-a-n-1}I[\theta \geq \tilde{\beta}]d\theta \\
&= (\alpha+n)\tilde{\beta}^{\alpha+n}\int_{\tilde{\beta}}^{\infty}\theta^{-a-n}d\theta \\
&= (\alpha+n)\tilde{\beta}^{\alpha+n}(\alpha+n-1)^{-1}\tilde{\beta}^{-\alpha-n+1} \\
&= \frac{\alpha+n}{\alpha+n-1}\max(X_{(n)}, \beta)
\end{aligned}$$

## 2.3 Compare Bayes with MLE

Since $\beta < X_{(n)}$, we have

$$\hat{\theta}_{Bayes} = \frac{\alpha+n}{\alpha+n-1}X_{(n)} = (1+\epsilon)X_{(n)}.$$

where $\epsilon = (\alpha+n-1)^{-1}$. On the other hand, by looking at $\pi(\boldsymbol{x}|\theta)$, we know

$$\hat{\theta}_{MLE} = X_{(n)}.$$

We need to know the distribution of $X_{(n)}$. For the CDF,

$$P(X_{(n)} \leq x) = \frac{x^n}{\theta^n},$$

so the PDF is

$$\pi(X_{(n)}) = \theta^{-n} n x^{n-1}.$$

Then

$$\begin{aligned}
E[X_{(n)}] =& \int_0^\theta x\theta^{-n} n x^{n-1} dx \\
=& \theta^{-n} n \int_0^\theta x^n dx \\
=& \theta^{-n} n (n+1)^{-1} \theta^{n+1} \\
=& \frac{\theta n}{n+1}, \\
E[X_{(n)}^2] =& \frac{\theta^2 n}{n+2}, \\
Var[X_{(n)}] =& E[X_{(n)}^2] - E[X_{(n)}]^2 \\
=& \frac{\theta^2 n}{(n+1)^2(n+2)}
\end{aligned}$$

Hence

$$\begin{aligned}
MSE(\hat{\theta}_{MLE}) =& Bias[\hat{\theta}_{MLE}]^2 + Var[\hat{\theta}_{MLE}] \\
=& \frac{\theta^2}{(n+1)^2} + \frac{\theta^2 n}{(n+1)^2(n+2)} \\
=& \frac{2\theta^2}{(n+1)(n+2)}.
\end{aligned}$$

Similarly, for the Bayes estimator,

$$\begin{aligned}
Bias(\hat{\theta}_{Bayes}) =& \frac{(\epsilon n - 1)\theta}{n+1} \\
Var(\hat{\theta}_{Bayes}) =& \frac{\theta^2 n (1+\epsilon)^2}{(n+1)^2(n+2)} \\
MSE(\hat{\theta}_{Bayes}) =& \frac{(\epsilon n - 1)^2(n+2) + n(1+\epsilon)^2}{(n+1)^2(n+2)} \theta^2.
\end{aligned}$$

Then comparing the two estimators,

$$\begin{aligned}
MSE(\hat{\theta}_{Bayes}) - MSE(\hat{\theta}_{MLE}) =& \frac{(\epsilon^2 n^2 - 2\epsilon n)(n+2) + n(\epsilon^2 + 2\epsilon)}{(n+1)^2(n+2)}\theta^2 \\
=& [(\epsilon n - 2)(n+2) + (\epsilon + 2)]\frac{n\epsilon\theta^2}{(n+1)^2(n+2)} \\
=& [\epsilon n^2 + 2\epsilon n + \epsilon - 2n - 2]\frac{n\epsilon\theta^2}{(n+1)^2(n+2)} \\
=& [\epsilon(n+1)^2 - 2(n+1)]\frac{n\epsilon\theta^2}{(n+1)^2(n+2)} \\
=& [\epsilon(n+1) - 2]\frac{(n+1)n\epsilon\theta^2}{(n+1)^2(n+2)} \\
=& \frac{n+1-2\alpha-2n+2}{\alpha+n+1}\frac{(n+1)n\epsilon\theta^2}{(n+1)^2(n+2)} \\
=& (3-2\alpha-n)\frac{(n+1)n\epsilon\theta^2}{(\alpha+n+1)(n+1)^2(n+2)}
\end{aligned}$$

Since the second term is always positive, we conclude that the Bayes estimator has a smaller MSE than the MLE estimator when $2\alpha + n > 3$.

# 3 Uniform distribution with a uniform prior

## 3.1 Bayes solution

For the Bayes solution, we have

$$\begin{aligned}
\pi(\theta) =& (2\alpha)^{-1} I[-\alpha \leq \theta \leq \alpha] \\
\pi(\boldsymbol{x}|\theta) =& I[X_{(1)} \geq \theta - \frac{1}{2}] I[X_{(n)} \leq \theta + \frac{1}{2}] \\
\pi(\theta|\boldsymbol{x}) \propto& (2\alpha)^{-1} I[\theta \geq \max(-\alpha, X_{(n)} - \frac{1}{2})] I[\theta \leq \min(\alpha, X_{(1)} + \frac{1}{2})] \\
\propto& I[\theta \geq \max(-\alpha, X_{(n)} - \frac{1}{2})] I[\theta \leq \min(\alpha, X_{(1)} + \frac{1}{2})],
\end{aligned}$$

so the Bayes estimator under the absolute error loss is the posterior mean, that is,

$$\delta_{\Pi_\alpha}(\boldsymbol{x}) = \frac{1}{2}[\max(-\alpha, X_{(n)} - \frac{1}{2}) + \min(\alpha, X_{(1)} + \frac{1}{2})].$$

## 3.2 Limit of the Bayes estimator

For every $\theta \in (-\infty, \infty)$, we know $X_{(n)} \geq \theta - \frac{1}{2}$ and $X_{(1)} \leq \theta + \frac{1}{2}$. Then for $\alpha$ large enough, $-\alpha < \theta - 1 \leq X_{(n)} - \frac{1}{2}$ and $\alpha > \theta + 1 \geq X_{(1)} + \frac{1}{2}$. Then

$$\delta_{\Pi_\alpha}(\boldsymbol{x}) = \frac{1}{2}[X_{(n)} - \frac{1}{2} + X_{(1)} + \frac{1}{2}] = \frac{1}{2}[X_{(1)} + X_{(n)}] = \delta(\boldsymbol{x}).$$

### 3.3 Minimax estimator

It suffices to show that the limiting Bayes estimator has constant risk with respect to $\theta$. The risk function is

$$\begin{aligned} R(\delta, \theta) =& E[|\delta(\boldsymbol{x}) - \theta|] \\ =& E\Big[\Big|\frac{1}{2}(X_{(1)} + X_{(n)}) - \theta\Big|\Big] \\ =& \frac{1}{2}E\Big[\Big|(X_{(1)} - \theta) + (X_{(n)} - \theta)\Big|\Big] \end{aligned}$$

However, $X_{(1)} - \theta$ and $X_{(n)} - \theta$ are ancillary statistics, since the distribution family is a location family with respect to $\theta$. Thus the risk function does not depend on $\theta$. This completes the proof.