S620 - Introduction To Statistical Theory - Homework 4

Enrique Areyan

February 13, 2014

(3.1) Let θ be a random variable in $(0, \infty)$ with density

$$\pi(\theta) \propto \theta^{\gamma-1} e^{-\beta\theta},$$

where $\beta, \gamma \in (1, \infty)$

(i) Calculate the mean and mode of θ.

Solution: Since θ is proportional to $\theta^{\gamma-1}e^{-\beta\theta}$, we know that $\theta \sim Gamma(\gamma, \beta)$, where $\beta > 1, \gamma > 1$. Therefore,

$$E[\theta] = \frac{\gamma}{\beta}, \text{ and Mode} = \frac{\gamma - 1}{\beta}$$

(ii) Suppose that $X_1, \ldots, X_n$ are random variables, which, conditional on θ are independent and each have the Poisson distribution with parameter θ. Find the form of the posterior density of θ given observed values $X_1 = x_1, \ldots X_n = x_n$. What is the posterior mean?

Solution: We know that $x|\theta \sim Poisson(\theta)$. Let us compute the posterior distribution:

$$\begin{aligned}
\theta|x &\sim \frac{\pi(\theta)f(x;\theta)}{\int_\Theta \pi(\theta')f(x;\theta')d\theta'} && \text{by definition of } \textit{posterior density} \\
&= \frac{\dfrac{\theta^{\gamma-1}e^{-\beta\theta}}{constant}\dfrac{e^{-\theta}\theta^x}{x!}}{constant} && \text{replacing prior and Poisson distribution} \\
&= constant \times \theta^{\gamma-1+x}e^{-\beta\theta-\theta} && \text{grouping like terms} \\
&= constant \times \theta^{(\gamma+x)-1}e^{-(\beta+1)\theta} && \text{rearranging exponents} \\
&\sim Gamma(\gamma + x, \beta + 1), && \text{by definition of Gamma distribution}
\end{aligned}$$

Therefore, the posterior mean is the mean of the gamma distribution with parameters $\gamma + x$ and $\beta + 1$, i.e.,

$$\boxed{E[\theta|x] = \frac{\gamma + x}{\beta + 1}}$$

Note that $\gamma > 1$ and $x = 0, 1, 2, \cdots$, so $\gamma + x > 0$ and $\beta > 1$ thus $\beta + 1 > 0$.

(ii) Suppose now that $T_1, \ldots, T_n$ are random variables, which, conditional on θ are independent, each exponentially distributed with parameter θ. What is the mode of the posterior distribution of θ, given $T_1 = t_1, \ldots, T_n = t_n$?

Solution: In this case we know that $x|\theta \sim Exp(\theta)$. The posterior distribution is:

$$\begin{aligned}
\theta|x &\sim \frac{\pi(\theta)f(x;\theta)}{\int_\Theta \pi(\theta')f(x;\theta')d\theta'} && \text{by definition of } \textit{posterior density} \\
&= \frac{\dfrac{\theta^{\gamma-1}e^{-\beta\theta}}{constant}\theta e^{-\theta t}}{constant} && \text{replacing prior and Exp. distribution} \\
&= constant \times \theta^{\gamma-1+1}e^{-\beta\theta-\theta t} && \text{grouping like terms} \\
&= constant \times \theta^{\gamma}e^{-(\beta+t)\theta} && \text{rearranging exponents} \\
&\sim Gamma(\gamma + 1, \beta + t), && \text{by definition of Gamma distribution}
\end{aligned}$$

Since $\gamma > 1$ then $\gamma + 1 > 0$, and we have a close form for the mode:

$$\boxed{\text{Mode} \;=\; \frac{(\gamma+1)-1}{\beta+t} = \frac{\gamma}{\beta+t}}$$

(3.3) Find the form of the Bayes rule in an estimation problem with loss function

$$L(\theta, d) = \begin{cases} a(\theta - d) & \text{if } d \le \theta, \\ b(d-\theta) & \text{if } d > \theta, \end{cases}$$

where a and b are given positive constants.

Solution:

We wish to minimize the following expression:

$$\int_{\Theta} L(\theta, d)\pi(\theta|x)d\theta = \int_{d}^{\infty} a(\theta - d)\pi(\theta|x)d\theta + \int_{-\infty}^{d} b(d-\theta)\pi(\theta|x)d\theta$$

As a function of d, i.e., let $f(d)$ be defined as follow:

$$\begin{aligned} f(d) &= \int_{d}^{\infty} a(\theta - d)\pi(\theta|x)d\theta + \int_{-\infty}^{d} b(d-\theta)\pi(\theta|x)d\theta \\ &= a\left[\int_{d}^{\infty} \theta\pi(\theta|x)d\theta - d\int_{d}^{\infty} \pi(\theta|x)d\theta\right] + b\left[d\int_{-\infty}^{d} \pi(\theta|x)d\theta - \int_{-\infty}^{d} \theta\pi(\theta|x)d\theta\right] \\ &= d\left[b\int_{-\infty}^{d} \pi(\theta|x)d\theta - a\int_{d}^{\infty} \pi(\theta|x)d\theta\right] + a\int_{d}^{\infty} \theta\pi(\theta|x)d\theta - b\int_{-\infty}^{d} \theta\pi(\theta|x)d\theta \end{aligned}$$

Find $\frac{df}{d(d)}$ (derivative of f with respect to d) using product rule and fundamental theorem of calculus:

$$\begin{aligned} \frac{df}{d(d)} &= \left(b\int_{-\infty}^{d} \pi(\theta|x)d\theta - a\int_{d}^{\infty} \pi(\theta|x)d\theta\right) + d\left[b\pi(d|x) + a\pi(d|x)\right] - ad\pi(d|x) - bd\pi(d|x) \\ &= b\int_{-\infty}^{d} \pi(\theta|x)d\theta - a\int_{d}^{\infty} \pi(\theta|x)d\theta + d(a+b)\pi(d|x) - d(a+b)\pi(d|x) \\ &= b\int_{-\infty}^{d} \pi(\theta|x)d\theta - a\int_{d}^{\infty} \pi(\theta|x)d\theta \end{aligned}$$

Set $\frac{df}{d(d)} = 0 \iff b\int_{-\infty}^{d} \pi(\theta|x)d\theta - a\int_{d}^{\infty} \pi(\theta|x)d\theta = 0 \iff b\int_{-\infty}^{d} \pi(\theta|x)d\theta = a\int_{d}^{\infty} \pi(\theta|x)d\theta \iff \frac{b}{a}\int_{-\infty}^{d} \pi(\theta|x)d\theta = \int_{d}^{\infty} \pi(\theta|x)d\theta$

This expression represents a quantile of posterior $\pi(\theta|x)$ in terms of a and b. Concretely:

$$\int_{-\infty}^{d} \pi(\theta|x)d\theta + \int_{d}^{\infty} \pi(\theta|x)d\theta = 1 \qquad \text{since prob. adds to 1} \qquad \Longrightarrow$$

$$\int_{-\infty}^{d} \pi(\theta|x)d\theta + \frac{b}{a}\int_{-\infty}^{d} \pi(\theta|x)d\theta = 1 \qquad \text{replacing for our minimun} \qquad \Longrightarrow$$

$$\int_{-\infty}^{d} \pi(\theta|x)d\theta = \frac{a}{a+b} \qquad \text{Solving for the posterior quantile}$$

Hence, the Bayes rule d is the quantile posterior $\frac{a}{a+b}$ of $\pi(\theta|x)$.

Note that we should check whether the minimality of the value obtain. Using the second derivative test:

$$\frac{d^2 f}{d(d)^2} = \frac{d}{d(d)}\left(b\int_{-\infty}^{d}\pi(\theta|x)d\theta - a\int_{d}^{\infty}\pi(\theta|x)d\theta\right) = b\pi(d|x) + a\pi(d|x) = (a+b)\pi(d|x) = (a+b)$$

The last equality is holds since the probability of taking action $d(x)$ given x is 1. By hypothesis a and b are given positive constants, and so is $a+b$, so we have a local minimun. Moreover, this expression is always positive and hence, the minimum found is the global minimum.

(3.4) Suppose that X is distributed as a binomial random variable with index n and parameter θ. Calculate the Bayes rule (based on the single observation X) for estimating θ when the prior distribution is the uniform distribution on $[0,1]$ and the loss function is

$$L(\theta, d) = (\theta - d)^2/\{\theta(1-\theta)\}$$

Is the rule you obtain minimax?

Solution: We wish to minimize the following expression:

$$\int_{\Theta} L(\theta, d)\pi(\theta|x)d\theta = \int_0^1 \frac{(\theta-d)^2}{\theta(1-\theta)}\pi(\theta|x)d\theta$$

Let us first find the posterior distribution $\pi(\theta|x)$.

$$\begin{aligned}
\theta|x &\sim \frac{\pi(\theta)f(x;\theta)}{\int_\Theta \pi(\theta')f(x;\theta')d\theta'} && \text{by definition of } \textit{posterior density} \\
&= \frac{1\cdot\binom{n}{x}\theta^x(1-\theta)^{n-x}}{\int_\Theta \pi(\theta')f(x;\theta')d\theta'} && \text{replacing prior and Binomial distribution} \\
&= \textit{constant} \times \theta^x(1-\theta)^{n-x} && \text{grouping like terms} \\
&\sim \textit{Beta}(x+1, n-x+1), && \text{by definition of Beta distribution}
\end{aligned}$$

Now we can minimize the function:

$$\begin{aligned}
f(d) &= \int_0^1 \frac{(\theta-d)^2}{\theta(1-\theta)}\pi(\theta|x)d\theta && \text{define } f(d) \text{ this way} \\
&= \int_0^1 \frac{(\theta-d)^2}{\theta(1-\theta)}\frac{\theta^x(1-\theta)^{n-x}}{B(x+1,n-x+1)}d\theta && \text{replace posterior} \\
&= \frac{1}{B(x+1,n-x+1)}\int_0^1(\theta-d)^2\theta^{x-1}(1-\theta)^{n-x-1}d\theta && \text{simplifying}
\end{aligned}$$

Note that in this context $\frac{1}{B(x+1,n-x+1)}$ is a constant, and so we can ignore this term for optimization purposes. Hence, optimize $f(d)$ is equivalent to optimize $g(d)$ defined as:

$$\begin{aligned}
g(d) &= \int_0^1(\theta-d)^2\theta^{x-1}(1-\theta)^{n-x-1}d\theta \\
&= \int_0^1(\theta^2 - 2\theta d + d^2)\theta^{x-1}(1-\theta)^{n-x-1}d\theta \\
&= \int_0^1\theta^{x+1}(1-\theta)^{n-x-1}d\theta - 2d\int_0^1\theta^x(1-\theta)^{n-x-1}d\theta + d^2\int_0^1\theta^{x-1}(1-\theta)^{n-x-1}d\theta
\end{aligned}$$

Find $\frac{dg}{d(d)}$ (derivative of g with respect to d):

$$\begin{aligned}
\frac{dg}{d(d)} &= \frac{d}{d(d)}\left[\int_0^1 \theta^{x+1}(1-\theta)^{n-x-1}d\theta - 2d\int_0^1 \theta^{x}(1-\theta)^{n-x-1}d\theta + d^2\int_0^1 \theta^{x-1}(1-\theta)^{n-x-1}d\theta\right] \\
&= -2\int_0^1 \theta^{x}(1-\theta)^{n-x-1}d\theta + 2d\int_0^1 \theta^{x-1}(1-\theta)^{n-x-1}d\theta
\end{aligned}$$

Set $\frac{dg}{d(d)} = 0 \iff -2\int_0^1 \theta^{x}(1-\theta)^{n-x-1}d\theta + 2d\int_0^1 \theta^{x-1}(1-\theta)^{n-x-1}d\theta = 0 \iff d = \frac{\int_0^1 \theta^{x}(1-\theta)^{n-x-1}d\theta}{\int_0^1 \theta^{x-1}(1-\theta)^{n-x-1}d\theta}$

We can write this condition in terms of the Beta function, i.e., $B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1}dt$, as follows:

$$d = \frac{B(x+1, n-x)}{B(x, n-x)}$$

Finally, using identities relating the Beta and Gamma function and the fact that n and x are positive integers:

$$d = \frac{B(x+1, n-x)}{B(x, n-x)} = \frac{\frac{\Gamma(x+1)\Gamma(n-x)}{\Gamma(n+1)}}{\frac{\Gamma(x)\Gamma(n-x)}{\Gamma(n)}} = \frac{\Gamma(x+1)\Gamma(n)}{\Gamma(n+1)\Gamma(x)} = \frac{x!(n-1)!}{n!(x-1)!} = \frac{x}{n}$$

And hence, the Bayes rule is just the mean $d_\pi = \frac{X}{n}$. A similar argument to that in (3.3), using second derivative test, shows that this is in fact the minimum.

The obtained rule d_π is minimax since it has constant risk, as the following computation shows:

$$\begin{aligned}
R(\theta, d_\pi) &= E_\theta L(\theta, d_\pi) && \text{definition of risk} \\
&= E_\theta L\left(\theta, \frac{X}{n}\right) && \text{replacing for the rule } d_\pi \\
&= E_\theta \left[\frac{\left(\theta - \frac{X}{n}\right)^2}{\theta(1-\theta)}\right] && \text{by loss function} \\
&= E_\theta \left[\frac{\theta^2 - 2\theta\frac{X}{n} + \left(\frac{X}{n}\right)^2}{\theta(1-\theta)}\right] && \text{squaring} \\
&= \frac{\theta^2 - 2\theta(\frac{\theta n}{n}) + \frac{n\theta(1-\theta)+n^2\theta^2}{n^2}}{\theta(1-\theta)} && \text{expectation of a Binomial and Binomial squared} \\
&= \frac{-\theta^2 + \frac{n\theta(1-\theta)+n^2\theta^2}{n^2}}{\theta(1-\theta)} \\
&= \frac{\frac{n\theta - n\theta^2 + n^2\theta^2 - n^2\theta^2}{n^2}}{\theta(1-\theta)}
\end{aligned}$$

$$= \frac{n\theta - n\theta^2}{n^2\theta(1-\theta)}$$

$$= \frac{n(1-\theta)}{n^2(1-\theta)}$$

$$= \frac{1}{n} \qquad \text{Since } n \text{ is fixed, this is a constant.}$$

(3.5) At a critical stage in the development of a new airplane, a decision must be taken to continue or to abandon the project. The financial viability of the project can be measured by a parameter θ, $0 < \theta < 1$, the project being profitable if $\theta > \frac{1}{2}$. Data x provide information about θ.

(i) If $\theta < \frac{1}{2}$, the cost to the taxpayer of continuing the project is $(\frac{1}{2} - \theta)$ (in units of \$bilion), whereas if $\theta > \frac{1}{2}$ it is zero (since the project will be privatized if profitable). If $\theta > \frac{1}{2}$ the cost of abandoning the project is $(\theta - \frac{1}{2})$ (due to contractual arrangements for purchasing the airplane from the French), whereas if $\theta < \frac{1}{2}$ it is zero. Derive the Bayes decision rule in terms of the posterior mean of θ given x.

Solution: Let $\mathcal{A} = \{0, 1\}$, where 0 denotes abandoning the project and 1 continuing the project. Let us write the loss function as follow:

$$L(\theta, 0) = \begin{cases} 0 & \text{if } 0 \le \theta \le \frac{1}{2} \\ (\theta - \frac{1}{2}) & \text{if } \frac{1}{2} < \theta \le 1 \end{cases} \qquad L(\theta, 1) = \begin{cases} (\frac{1}{2} - \theta) & \text{if } 0 \le \theta \le \frac{1}{2} \\ 0 & \text{if } \frac{1}{2} < \theta \le 1 \end{cases}$$

We want to minimize $\int\limits_{\Theta} L(\theta, d)\pi(\theta|x)d\theta$, i.e.,

$$\int_{\Theta} L(\theta, d)\pi(\theta|x)d\theta = \begin{cases} \int\limits_{0}^{\frac{1}{2}} (\frac{1}{2} - \theta)\pi(\theta|x)d\theta & \text{if } d = 1 \\ \\ \int\limits_{\frac{1}{2}}^{1} (\theta - \frac{1}{2})\pi(\theta|x)d\theta & \text{if } d = 0 \end{cases}$$

The Bayes rule is that which minimizes the above expression. That is, the Bayes rule will choose to go on with the protect if $\int\limits_{0}^{\frac{1}{2}} (\frac{1}{2} - \theta)\pi(\theta|x)d\theta < \int\limits_{\frac{1}{2}}^{1} (\theta - \frac{1}{2})\pi(\theta|x)d\theta$ or, equivalently, the Bayes rule will choose to abandon the protect if $\int\limits_{\frac{1}{2}}^{1} (\theta - \frac{1}{2})\pi(\theta|x)d\theta < \int\limits_{0}^{\frac{1}{2}} (\frac{1}{2} - \theta)\pi(\theta|x)d\theta$.

(ii) The minister of aviation has prior density $6\theta(1 - \theta)$ for θ. The Prime Minister has prior density $4\theta^3$. The prototype aeroplane is subjected to trials, each independently having probability θ of success, and the data x consist of the total number of trials required for the first successful result to be obtained. For what values of x will there be serious ministerial disagreement?

Solution: Clearly, $X \sim Geometric(\theta)$, where $X = 1, 2, \ldots$. Hence $\pi(x) = (1 - \theta)^{x-1}\theta$.

We can compute the posterior probability density $\pi(\theta|x)$ for the minister of aviation (MA) and the Prime Minister (PM):

MA: $\pi(\theta|x) = constant \times \theta(1-\theta)(1-\theta)^{x-1}\theta = \theta^2(1-\theta)^x = Beta(3, x+1)$

PM $\pi(\theta|x) = constant \times \theta^3(1-\theta)^{x-1}\theta = \theta^4(1-\theta)^{x-1} = Beta(5, x)$

We want to find the values of x for which the Bayes rule for MA and MP will disagree. Let $\pi_{MA}(\theta|x)$ and $\pi_{MP}(\theta|x)$ be the prior of the minister of aviation and Prime Minister respectively. Then, the values of x we are interested in are:

$$x \text{ for which: } \begin{cases} \int\limits_0^{\frac{1}{2}} (\frac{1}{2}-\theta)\pi_{MA}(\theta|x)d\theta < \int\limits_{\frac{1}{2}}^{1} (\theta-\frac{1}{2})\pi_{MA}(\theta|x)d\theta \\ \\ \int\limits_0^{\frac{1}{2}} (\frac{1}{2}-\theta)\pi_{MP}(\theta|x)d\theta > \int\limits_{\frac{1}{2}}^{1} (\theta-\frac{1}{2})\pi_{MP}(\theta|x)d\theta \end{cases}$$

Solving for x using the fact that $\pi(\theta|x)$ is a probability distribution and computing its mean:

$$\begin{cases} \frac{1}{2}\int\limits_0^{\frac{1}{2}} \pi_{MA}(\theta|x)d\theta - \int\limits_0^{\frac{1}{2}} \theta\pi_{MA}(\theta|x)d\theta < \int\limits_{\frac{1}{2}}^{1} \theta\pi_{MA}(\theta|x)d\theta - \frac{1}{2}\int\limits_{\frac{1}{2}}^{1} \pi_{MA}(\theta|x)d\theta \\ \\ \frac{1}{2}\int\limits_0^{\frac{1}{2}} \pi_{MP}(\theta|x)d\theta - \int\limits_0^{\frac{1}{2}} \theta\pi_{MP}(\theta|x)d\theta > \int\limits_{\frac{1}{2}}^{1} \theta\pi_{MP}(\theta|x)d\theta - \frac{1}{2}\int\limits_{\frac{1}{2}}^{1} \pi_{MP}(\theta|x)d\theta \end{cases} \implies \begin{cases} \frac{1}{2}\int\limits_0^{1} \pi_{MA}(\theta|x)d\theta - \int\limits_0^{1} \theta\pi_{MA}(\theta|x)d\theta < 0 \\ \\ \frac{1}{2}\int\limits_0^{1} \pi_{MP}(\theta|x)d\theta - \int\limits_0^{1} \theta\pi_{MP}(\theta|x)d\theta > 0 \end{cases}$$

Note that $\frac{1}{2}\int\limits_0^{1} \pi_{MA}(\theta|x)d\theta = \frac{1}{2}$ and $\int\limits_0^{1} \theta\pi_{MA}(\theta|x)d\theta = E[\pi_{MA}(\theta|x)]$. Same for MP. We know each of these distributions are Beta and so we know its mean:

$$E[\pi_{MA}(\theta|x)] = E[Beta(3, x+1)] = \frac{3}{x+4} \text{ and } E[\pi_{MP}(\theta|x)] = E[Beta(5, x)] = \frac{5}{5+x}$$

Replacing into our conditions:

$$\begin{cases} \frac{1}{2} - E[\pi_{MA}(\theta|x)] < 0 \\ \\ \frac{1}{2} - E[\pi_{MP}(\theta|x)] > 0 \end{cases} \implies \begin{cases} \dfrac{1}{2} - \dfrac{3}{x+4} < 0 \\ \\ \dfrac{1}{2} - \dfrac{5}{5+x} > 0 \end{cases} \implies \begin{cases} \dfrac{1}{2} < \dfrac{3}{x+4} \\ \\ \dfrac{1}{2} > \dfrac{5}{5+x} \end{cases} \implies \begin{cases} x+4 < 6 \\ \\ x+5 > 10 \end{cases} \implies \begin{cases} x < 2 \\ \\ x > 5 \end{cases}$$

Hence, the minister of aviation will go on with the project if $x < 2$ while the Prime Minister will choose to abandon the project if $x > 5$. This means that MA will choose to abandon if $x > 2$. So, for values $x = 3, 4$; there will be a disagreement: MA would want to abandon but PM will not. A decision will have to be reached through some other argument.

Note that this result makes intuitive sense: the prime minister prior is a cubic function that places lower emphasis on smaller values of θ and thus, is more tolerant to risk while the opposite is try for the minister of aviation whose prior is a quadratic function.