Statistics 582, Problem Set 5 Solutions

Wellner; 2/8/2006

1. Let X be a random variable with finite first moment: $E|X| < \infty$. Show that $f(b) \equiv E|X - b|$ is minimized by $b =$ any median of the distribution F of X. [A median m of F is any value satisfying $F(m) = P(X \le m) \ge 1/2$ and $1 - F(m-) = P(X \ge m) \ge 1/2$; See Lehmann and Casella, TPE, page 62, problems 1.7 and 1.8.]

   Solution: Suppose that m is a median of F. From Lehmann and Casella problem 1.7, it follows that $m_0 \le m \le m_1$ so that the set of medians is a closed interval. This is easily proved as follows: suppose that $\mathcal{M}$ is the set of medians of F. Note that $\mathcal{M}$ is always non-empty since $m_0 \equiv \inf\{x : \ F(x) \ge 1/2\} \in \mathcal{M}$. If $\mathcal{M} = \{m_0\}$, then $[m_0, m_0] = \{m_0\}$ is closed. If If $a, b \in \mathcal{M}$ with $a < b$, then if $c \in (a, b)$ we have $P(X \le c) \ge P(X \ge a) \ge 1/2$ (since $a \in \mathcal{M}$), and $P(X \ge c) \ge P(X \ge b) \ge 1/2$ (since $b \in \mathcal{M}$). Thus $c \in \mathcal{M}$ and hence $(a, b) \subset \mathcal{M}$. Let $(m_0, m_1) = \cup_{a,b \in \mathcal{M}}(a, b)$ be the union of all the open intervals contained in $\mathcal{M}$. Then if $m \in (m_0, m_1)$

$$\begin{aligned} &1/2 \le P(X \le m) = E1\{X \le m\} \to E1\{X < m_1\} \le E1\{X \le m_1\} = P(X \le m_1), \quad \text{and} \\ &1/2 \le P(X \ge m) = E1\{X \ge m\} \to E1\{X \ge m_1\} = P(X \ge m_1) \end{aligned}$$

   as $m \nearrow m_1$ by the dominated convergence theorem. Thus $m_1 \in \mathcal{M}$. Similarly,

$$\begin{aligned} &1/2 \le P(X \le m) = E1\{X \le m\} \to E1\{X \le m_0\} \le P(X \le m_0), \quad \text{and} \\ &1/2 \le P(X \ge m) = E1\{X \ge m\} \to E1\{X > m_0\} \le P(X \ge m_0) \end{aligned}$$

   as $m \searrow m_0$ by the dominated convergence theorem. Thus $m_0 \in \mathcal{M}$ and we conclude that $[m_0, m_1] \subset \mathcal{M}$. On the other hand $\mathcal{M} \subset [m_0, m_1]$ with $m_0 \equiv \inf\{x : F(x) \ge 1/2\}$ and $m_1 \equiv \inf\{x : F(x) > 1/2\}$.
   Suppose that $c > m_1$. Then by examining the graphs of $|x - c|$ and $|x - m|$ we see that

$$\begin{aligned} |x-c| - |x-m| &= (m-c)1_{[x \ge c]} + (c-m)1_{[x \le m]} + \{(c-x) - (x-m)\}1_{[m<x<c]} \\ &= (c-m)\left\{1_{[x \le m]} - 1_{[x \le c]}\right\} + (c+m-2x)1_{[m<x<c]} \\ &= (c-m)\left\{1_{[x \le m]} - 1_{[x \le c]}\right\} + 2(c-x)1_{[m<x<c]} - (c-m)1_{[m<x<c]} \\ &= (c-m)\left\{1_{[x \le m]} - 1_{[x > m]}\right\} + 2(c-x)1_{[m<x<c]}. \end{aligned}$$

   Replacing x by X and taking expectations across the identity with respect to X yields

$$\begin{aligned} E|X-c| - E|X-m| &= (c-m)\{P(X \le m) - P(X > m)\} + 2E\{(c-X)1_{[m<X<c]}\} \\ &> 0 + 0 = 0 \end{aligned}$$

since m is a median of F implies that $P(X \le m) - P(X > m) \ge 0$ and $c > m_1 \ge m_0$ implies that $E\{(c-X)1_{[m<X<c]}\} = E\{(c-X)1_{[m_1<X<c]}\} > 0$. Similarly, if $c < m_0$,

$$|x-c| - |x-m| \; = \; (m-c)(1_{[x\ge m]} - 1_{[x<m]}) + 2(x-c)1_{[c<x<m]},$$

and taking expectations yields

$$\begin{aligned} E|X-c| - E|X-m| &= (m-c)\{P(X\ge m) - P(X<m)\} + 2E\{(X-c)1_{[c<X<m]}\} \\ &> 0. \end{aligned}$$

Thus $E|X-b|$ is minimized by any median of the distribution F of X.

2. Suppose that $X_1,\ldots,X_n$ are i.i.d. Exponential(θ) (so the X's have density $p_\theta(x) = \theta e^{-\theta x}1_{(0,\infty)}(x)$. with respect to Lebesgue measure on R, and that $\theta \sim \Gamma(\alpha,\beta)$:

$$\lambda(\theta) = \beta\frac{(\beta\theta)^{\alpha-1}}{\Gamma(\alpha)}\exp(-\beta\theta)1_{[0,\infty)}(\theta).$$

A. Find the Bayes rule $d_B(\underline{X})$ for estimation of θ with squared error loss $L(\theta,a) = |\theta-a|^2$. Find the Bayes rule $d_{Bw}(\underline{X})$ for estimation of θ with weighted squared error loss $L(\theta,a) = (\theta-a)^2/\theta$. Is the maximum likelihood estimator among either of these families of Bayes estimators?
B. Are the Bayes estimators d_B and d_{Bw} consistent? What are the limit distributions of d_B and d_{Bw}? Compare them with the maximum likelihood estimator.
C. Suppose that instead of the Gamma prior distribution, θ has the Pareto(θ_0,α) distribution with density λ given by

$$\lambda(\theta) = (\frac{\alpha}{\theta_0})(\frac{\theta_0}{\theta})^{\alpha+1}1_{(\theta_0,\infty)}(\theta);$$

here $E(\theta) = \frac{\alpha}{\alpha-1}\theta_0$ where $\alpha > 1$ and $\theta_0 > 0$ are known. What can you say about the Bayes estimator for squared error loss with this prior? For what values of θ_0 is the Bayes rule consistent?

Solution: A. The posterior distribution is Gamma($\alpha+n, \beta+\sum X_i$). Thus the Bayes rule for $L(\theta,a) = (\theta-a)^2$ is

$$d_B(\underline{X}) = \frac{\alpha+n}{\beta+\sum X_i}.$$

For $L(\theta,a) = (\theta-a)^2/\theta$, the Bayes rule is

$$d_{Bw}(\underline{X}) = \frac{E(\theta K(\theta)|\underline{X})}{E(K(\theta)|\underline{X})} = \frac{1}{E(1/\theta|\underline{X})} = \frac{\alpha+n-1}{\beta+\sum X_i}$$

since, for $\theta \sim$ Gamma(α, β) we have

$$E(1/\theta) = \frac{\beta}{\alpha - 1}$$

if $\alpha > 1$. Thus the MLE $1/\overline{X}_n$ is *not* among either of these families of estimators.
B. Both d_B and d_{Bw} are consistent and asymptotically equivalent to the MLE $1/\overline{X}_n$:

$$\begin{aligned} \sqrt{n}\left\{d_B(\underline{X}) - 1/\overline{X}_n\right\} &= \sqrt{n}\left\{\frac{1+n^{-1}\alpha}{\overline{X}_n + n^{-1}\beta} - \frac{1}{\overline{X}_n}\right\} \\ &= n^{-1/2}\frac{\alpha\overline{X}_n - \beta}{\overline{X}_n(\overline{X}_n + n^{-1}\beta)} = O(n^{-1/2})O_p(1) = o_p(1)\,, \end{aligned}$$

and similarly for d_{Bw}. Thus, for $d = d_B$ or $d = d_{Bw}$ we have, since $I(\theta) = \theta^{-2}$,

$$\sqrt{n}(d(\underline{X}) - \theta) = \sqrt{n}(\frac{1}{\overline{X}_n} - \theta) + o_p(1) \to_d N(0, 1/I(\theta)) = N(0, \theta^2)\,.$$

C. When the prior is Pareto(θ_0, α), the posterior density is of the form

$$\begin{aligned} \lambda(\theta|\underline{X}) &= \frac{\theta^n \exp(-\theta\sum X_i)(\alpha\theta_0^{-1})(\theta_0/\theta)^{\alpha+1}1_{(\theta_0,\infty)}(\theta)}{\int_{\theta_0}^{\infty} s^n \exp(-s\sum X_i)(\alpha\theta_0^{-1})(\theta_0/s)^{\alpha+1}ds} \\ &= \frac{\theta^{n-\alpha-1}\exp(-\theta\sum X_i)1_{(\theta_0,\infty)}(\theta)}{\int_{\theta_0}^{\infty} s^{n-\alpha-1}\exp(-s\sum X_i)ds}\,, \end{aligned}$$

which is concentrated on (θ_0, ∞). Thus the Bayes rule $d_B(\underline{X}) = E(\theta|\underline{X})$ takes values in (θ_0, ∞) a.s.. Similar to the argument in class in the Bernoulli(θ) example, $Z_n = d_B(\underline{X}) = E(\theta|X_1, \ldots, X_n)$ is a martingale and hence $Z_n = d_B(\underline{X}) \to E(\theta|X_1, X_2, \ldots)$. But $\widehat{\theta} = \overline{X}_n^{-1} \to_{a.s.} \theta$ for each fixed $\theta \in (0, \infty)$, and hence

$$P_\Lambda(\widehat{\theta}_n \to \theta) = \int P_\theta(\widehat{\theta}_n \to \theta)d\Lambda(\theta) = 1\,.$$

Hence $\widehat{\theta}_n \to \theta$ a.s. P_Λ, and this implies that θ is $\mathcal{F}_\infty \equiv \sigma(X_1, X_2, \ldots)$ measurable. Therefore $E(\theta|X_1, X_2, \ldots) = \theta$ a.s. and $d_B(\underline{X}) \to \theta$ a.s. P_Λ. This in turn implies that $d_B(\underline{X}) \to_{a.s.} \theta$ for Λ-a.e. θ. this suggests that d_B might be inconsistent for $\theta \in (0, \theta_0)$, and this is in fact the case since $d_B(\underline{X}) < \theta_0$. When the true $\theta < \theta_0$, it is possible to show that $d_B(\underline{X}) \to_{a.s.} \theta_0 > \theta$ and that the posterior distributions convergen to point mass at θ_0.

3. Let $\Theta = (0, \infty)$, $\mathbf{A} = [0, \infty)$, let X have the discrete distribution

$$p(x, \theta) = \binom{r+x-1}{x} \theta^x (\theta+1)^{-(r+x)}, \qquad x = 0, 1, 2, \ldots$$

where r is some known positive integer; this is the negative binomial distribution reparametrized so that $E_\theta X = r\theta$. Suppose that

$$L(\theta, a) = \frac{(\theta - a)^2}{\theta(\theta+1)}.$$

(a) Show that the usual estimator, $d_0(X) = X/r$ is an equalizer rule.
(b) Show that the usual estimator d_0 is generalized Bayes with respect to Lebesgue measure on $(0, \infty)$ provided $r > 1$. (What happens if $r = 1$?)
(c) Find Bayes decision rules with respect to the prior distributions $\Lambda_{\alpha,\beta}$ with densities

$$\lambda_{\alpha,\beta}(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1}(\theta+1)^{-(\alpha+\beta)} 1_{(0,\infty)}(\theta),$$

the distribution of $\theta = Z/(1-Z)$ where $Z \sim$ Beta(α, β).
(d) Show that $d(X) = X/(r+1)$ is minimax. [Note that d_0 is not minimax, hence not admissible.]

Solution: (a) First note that $E_\theta(X) = r\theta$ and $Var_\theta(X) = r\theta(\theta+1)$; this follows from the facts that if X has a negative binomial distribution with mass function

$$p(x; p) = \binom{x+r-1}{x} p^r q^x, \quad x \in \{0, 1, \ldots\},$$

then $EX = rq/p$ and $Var(X) = rq/p^2$ with $q \equiv 1 - p$. Thus for the weighted squared error loss $L(\theta, a) = (\theta - a)^2/(\theta(\theta+1))$ the rule $d_0(X) = X/r$ has risk

$$R(\theta, d_0) = \frac{1}{\theta(\theta+1)} Var_\theta(X/r) = \frac{1}{r^2\theta(\theta+1)} r\theta(\theta+1) = \frac{1}{r};$$

since the risk function of the rule d_0 is constant in θ, it is "an equalizer rule".
(b) For $\lambda(\theta) = 1_{(0,\infty)}(\theta)$ (corresponding to Λ Lebesgue measure on $(0, \infty)$, the (generalized) Bayes rule is

$$d_\Lambda(X) = \frac{E\{K(\theta)\theta|X\}}{E\{K(\theta)|X\}} = \frac{E\{(\theta+1)^{-1}|X\}}{E\{\theta^{-1}(\theta+1)^{-1}|X\}}$$

where the posterior density is

$$\lambda(\theta|X) = \frac{\Gamma(X+r)}{\Gamma(X+1)\Gamma(r-1)} \theta^{X+1-1}(\theta+1)^{-(r+X)}.$$

Thus we compute the numerator as

$$\begin{aligned} & E\{(\theta+1)^{-1}|X\} \\ = & \int_0^\infty \theta^{X+1-1}(\theta+1)^{-(r+X+1)}\frac{\Gamma(X+r+1)}{\Gamma(X+1)\Gamma(r)}d\theta\cdot\frac{\Gamma(X+r)}{\Gamma(X+r+1)}\cdot\frac{\Gamma(r)}{\Gamma(r-1)} \\ = & \frac{r-1}{X+r}, \end{aligned}$$

and the denominator is

$$\begin{aligned} & E\{\theta^{-1}(\theta+1)^{-1}|X\} \\ = & \int_0^\infty \theta^{X-1}(\theta+1)^{-(r+X+1)}\frac{\Gamma(X+r+1)}{\Gamma(X)\Gamma(r+1)}d\theta\cdot\frac{\Gamma(X+r)}{\Gamma(X+r+1)}\cdot\frac{\Gamma(X)}{\Gamma(X+1)}\cdot\frac{\Gamma(r+1)}{\Gamma(r-1)} \\ = & \frac{1}{X+r}\cdot\frac{1}{X}\cdot r(r-1). \end{aligned}$$

Putting these together yields $d_\Lambda(X) = X/r = d_0(X)$. Thus d_0 is a "generalized Bayes rule" with respect to the (improper) prior given by Lebesgue measure on $(0,\infty)$. This argument works when $r > 1$ (because of the factor $\Gamma(r-1)$ in the denominator). When $r = 1$ the corresponding posterior is

$$\lambda(\theta|X) = \frac{\Gamma(X+1)}{\Gamma(X+1)\Gamma(0)}\theta^{X+1-1}(\theta+1)^{-(1+X)} = 0$$

since $\Gamma(0) = \int_0^\infty x^{-1}e^{-x}dx = \infty$.

(c) By straightforward calculation the posterior density of θ for the given prior is

$$\lambda(\theta|X) = \frac{\Gamma(X+\alpha+r+\beta)}{\Gamma(X+\alpha)\Gamma(r+\beta)}\theta^{X+\alpha-1}(\theta+1)^{-(r+X+\alpha+\beta)}1_{(0,\infty)}(\theta).$$

The Bayes rule with respect to the loss function $L(\theta,a) = (\theta-a)^2/[\theta(\theta+1)] \equiv K(\theta)(\theta-a)^2$ is given by

$$d_\Lambda(X) = \frac{E\{K(\theta)\theta|X\}}{E\{K(\theta)|X\}} = \frac{E\{(\theta+1)^{-1}|X\}}{E\{\theta^{-1}(\theta+1)^{-1}|X\}}$$

By straightforward calculation the numerator and denominator are given by

$$\begin{aligned} E\{K(\theta)\theta|X\} &= \frac{r+\beta}{X+\alpha+r+\beta}, \\ E\{K(\theta|X\} &= \frac{(r+\beta+1)(r+\beta)}{(X+\alpha+r+\beta)(X+\alpha-1)}. \end{aligned}$$

Thus the Bayes rule with respect to this weighted loss function and prior Λ is

$$d_\Lambda(X) = \frac{X+\alpha-1}{r+\beta+1}.$$

Since $E_\theta d_\Lambda(X) = (r\theta + \alpha - 1)/(r + \beta + 1)$ and

$$Var_\theta(d_\Lambda(X)) = \frac{r\theta(\theta+1)}{(r+\beta+1)^2},$$

The (ordinary) risk of the rule d_Λ is

$$\begin{aligned}
R(\theta, d_\Lambda) &= \frac{\frac{r\theta(\theta+1)}{(r+\beta+1)^2} + \left(\frac{r\theta+\alpha-1}{r+\beta+1} - \theta\right)^2}{\theta(\theta+1)} \\
&= \frac{1}{(r+\beta+1)^2}\left\{r + \frac{[\alpha - 1 - \theta(\beta+1)]^2}{\theta(\theta+1)}\right\} \\
&= \frac{1}{(r+\beta+1)^2}\left\{r + \frac{(\alpha-1)^2}{\theta(\theta+1)} - \frac{2(\alpha-1)(\beta+1)}{\theta+1} + \frac{\theta(\beta+1)^2}{\theta+1}\right\}.
\end{aligned}$$

Thus after calculation of

$$\begin{aligned}
\int_0^\infty \frac{1}{\theta(\theta+1)}\lambda(\theta)d\theta &= \frac{\beta(\beta+1)}{\alpha(\alpha+\beta+1)}, \\
\int_0^\infty \frac{1}{\theta+1}\lambda(\theta)d\theta &= \frac{\beta}{\alpha+\beta}, \quad \text{and} \\
\int_0^\infty \frac{\theta}{\theta+1}\lambda(\theta)d\theta &= \frac{\alpha}{\alpha+\beta},
\end{aligned}$$

we find the Bayes risk of the Bayes rule d_Λ to be

$$\begin{aligned}
\mathcal{R}(\Lambda, d_\Lambda) &= \frac{1}{(r+\beta+1)^2}\left\{r + (\alpha-1)^2\frac{\beta(\beta+1)}{\alpha(\alpha+\beta+1)}\right. \\
&\qquad \left. - 2(\alpha-1)(\beta+1)\frac{\beta}{\alpha+\beta} + (\beta+1)^2\frac{\alpha}{\alpha+\beta}\right\} \\
&\to \frac{1}{(r+1)^2}\{r+1\} = \frac{1}{r+1} \qquad \text{as } \alpha \to 1,\ \beta \to 0. \qquad (1)
\end{aligned}$$

(d) The rule $d(X) = X/(r+1)$ corresponding to the limiting Bayes risk in (1) has risk

$$R(\theta, d) = \frac{1}{(r+1)^2}\left\{r + \frac{\theta}{\theta+1}\right\}$$

with supremum risk

$$\sup_{\theta>0} R(\theta, d) = \frac{1}{r+1}.$$

Thus by theorem 6.2 the rule d is minimax.

4. A. Let $(X|\sigma^2) \sim N(0, \sigma^2)$. Show that the conjugate prior for σ^2 is the distribution of $1/Y$ where Y has a gamma distribution.
B. Suppose that $(X|\theta, \kappa) \sim N(\theta, 1/\kappa)$, $(\theta|\kappa) \sim N(\mu, \tau/\kappa)$, and $\kappa \sim \text{Gamma}(\alpha, \beta)$. Show that the posterior distribution of (θ, κ) has the same form as the prior.
C. Find the marginal posterior distribution for θ in B.
D. If $X_1, \ldots, X_n$ are i.i.d. as X in A, find the limiting distribution of the Bayes estimator of θ.

Solution: A. Now $p(x|\sigma^2) = (2\pi)^{-1/2}\sigma^{-1}\exp(-x^2/2\sigma^2)$, so a conjugate prior is of the form

$$\lambda(\sigma^2) = (\sigma^2)^{-a}\exp(-b/2\sigma^2).$$

If $Y \sim \Gamma(\alpha, \beta)$, then $Z \equiv 1/Y$ has density

$$p_Z(z; \alpha, \beta) = \frac{1}{z^2} p_Y(1/z; \alpha, \beta) = \frac{z^{-\alpha-1}}{\Gamma(\alpha)}\beta^\alpha \exp(-\beta/z).$$

Identifying a with $\alpha + 1$ and b with $\beta/2$, the claim follows. Equivalently, if we reparametrize the normal density by $\kappa \equiv 1/\sigma^2$ so that

$$p(x|\kappa) = (\kappa/2\pi)^{1/2}\exp(-(\kappa/2)x^2)$$

and suppose that $\kappa \sim \Gamma(\alpha, \beta)$. then

$$\begin{aligned} p(x|\kappa)\lambda(\kappa) &= \left(\frac{\kappa}{2\pi}\right)^{1/2}\exp(-\kappa x^2/2)\frac{\kappa^{\alpha-1}}{\Gamma(\alpha)}\beta^\alpha\exp(-\beta\kappa)1_{(0,\infty)}(\kappa) \\ &= \frac{\kappa^{\alpha-1/2}\beta^\alpha}{\sqrt{2\pi}\Gamma(\alpha)}\exp(-(\beta + \frac{x^2}{2})\kappa)1_{(0,\infty)}(\kappa), \end{aligned}$$

and hence $(\kappa|X) \sim \Gamma(\alpha + 1/2, \beta + X^2/2)$.
B. Since it is not more difficult and is needed in part D, we will take X to have the distribution of $\overline{X}_n$ with $X_1, \ldots, X_n$ i.i.d $N(\theta, 1/\kappa)$, namely $N(\theta, 1/\kappa)$. Then the result for this part follows by taking $n = 1$. The joint density of $\overline{X}_n, \theta, \kappa$ is given by

$$\begin{aligned} &p(x|\theta,\kappa)\lambda(\theta|\kappa)\lambda(\kappa) \\ &= \sqrt{\frac{n\kappa}{2\pi}}\exp(-\frac{n\kappa}{2}(x-\theta)^2)\sqrt{\frac{\kappa}{2\pi\tau}}\exp\left(-\frac{\kappa}{2\tau}(\theta-\mu)^2\right)\frac{\kappa^{\alpha-1}\beta^\alpha}{\Gamma(\alpha)}\exp(-\beta\kappa) \\ &= \frac{\kappa^\alpha\beta^\alpha}{2\pi\Gamma(\alpha)}\sqrt{\frac{n}{\tau}}\exp(-\kappa(\beta + \frac{n}{2}(x-\theta)^2 + \frac{1}{2\tau}(\theta-\mu)^2)) \\ &= \frac{\kappa^\alpha\beta^\alpha}{2\pi\Gamma(\alpha)}\sqrt{\frac{n}{\tau}}\exp\left(-\frac{\kappa}{2}(n + \frac{1}{\tau})\left(\theta - \frac{nx + \frac{\mu}{\tau}}{n + \frac{1}{\tau}}\right)^2\right) \\ &\qquad \cdot \exp\left(-\kappa\left(\beta + \frac{1}{2}\frac{1/\tau}{(n+1/\tau)}(x-\mu)^2\right)\right), \end{aligned}$$

after some (careful!) algebra, and it follows that

$$(\theta, \kappa|\overline{X}) \sim N(\mu(\overline{X};\tau), \kappa^{-1}(n+1/\tau)^{-1}) \cdot \text{Gamma}\left(\alpha+\frac{1}{2},\ \frac{1/\tau}{(n+1/\tau)}(\overline{X}-\mu)^2\right)$$

where

$$\mu_n(x;\tau) = \frac{nx+\frac{\mu}{\tau}}{n+\frac{1}{\tau}} = \frac{x+\mu/(\tau n)}{1+1/(\tau n)} \ .$$

We also define

$$\beta_n(x,\tau) \equiv \beta + \frac{1}{2}\frac{1/\tau}{(n+1/\tau)}(x-\mu)^2. \tag{2}$$

Hence

$$\begin{aligned}\lambda(\theta,\kappa|\overline{X}_n) \quad &= \quad \sqrt{\frac{\kappa(n+1/\tau)}{2\pi}} \exp\left(-\frac{\kappa(n+1/\tau)}{2}(\theta-\mu_n(\overline{X},\tau))^2\right) \\ &\quad \cdot \frac{\kappa^{\alpha-1/2}\beta(\overline{X},\tau)^{\alpha+1/2}}{\Gamma(\alpha+1/2)} \exp(-\beta_n(\overline{X},\tau)\kappa).\end{aligned}$$

C. Thus the marginal posterior distribution of θ is

$$\begin{aligned}\lambda(\theta|\overline{X}) \quad &= \quad \int_0^\infty \lambda(\theta,\kappa|\overline{X})d\kappa \\ &= \quad \int_0^\infty \kappa^\alpha \exp\left\{-\kappa\left(\beta_n(\overline{X},\tau) + \frac{n+1/\tau}{2}(\theta-\mu_n(\overline{X},\tau))^2\right)\right\} d\kappa \\ &\qquad \sqrt{\frac{n+1/\tau}{2\pi}}\frac{\beta_n(\overline{X},\tau)^{\alpha+1/2}}{\Gamma(\alpha+1/2)} \\ &= \quad \int_0^\infty \kappa^{\alpha+1-1}\frac{\tilde{\beta}^{\alpha+1}}{\Gamma(\alpha+1)}\exp(-\tilde{\beta}\kappa)d\kappa \cdot \frac{\Gamma(\alpha+1)}{\tilde{\beta}^{\alpha+1}}\sqrt{\frac{n+1/\tau}{2\pi}}\frac{\beta_n(\overline{X},\tau)^{\alpha+1/2}}{\Gamma(\alpha+1/2)} \\ &= \quad \frac{\Gamma(\alpha+1)}{\tilde{\beta}^{\alpha+1}}\sqrt{\frac{n+1/\tau}{2\pi}}\frac{\beta_n(\overline{X},\tau)^{\alpha+1/2}}{\Gamma(\alpha+1/2)} \\ &= \quad \frac{\Gamma(\alpha+1)}{\Gamma(\alpha+1/2)}\sqrt{\frac{n+1/\tau}{2\pi}}\frac{\beta_n(\overline{X},\tau)^{\alpha+1/2}}{\tilde{\beta}^{\alpha+1}} \\ &= \quad \frac{\Gamma(\alpha+1)}{\Gamma(\alpha+1/2)}\sqrt{\frac{n+1/\tau}{2\pi}}\frac{\beta_n(\overline{X},\tau)^{\alpha+1/2}}{\left(\beta_n(\overline{X},\tau)+\frac{n+1/\tau}{2}(\theta-\mu_n(\overline{X},\tau))^2\right)^{\alpha+1}} \qquad (3)\end{aligned}$$

with $\beta_n(x,\tau)$ as defined in (2)and

$$\tilde{\beta} \equiv \tilde{\beta}_n(x,\theta,\tau,\beta) = \left(\beta_n(x,\tau) + \frac{n+1/\tau}{2}(\theta-\mu_n(x,\tau))^2\right).$$

To understand this marginal posterior distribution, we first calculate the marginal *prior* distribution of θ:

$$\begin{aligned}
\lambda(\theta) &= \int_0^\infty \lambda(\theta|\kappa)\lambda(\kappa)d\kappa \\
&= \frac{\beta^\alpha}{\sqrt{2\pi\tau}\Gamma(\alpha)}\int_0^\infty \frac{\kappa^{\alpha+1/2-1}\tilde{\beta}^{\alpha+1/2}}{\Gamma(\alpha+1/2)}\exp(-\tilde{\beta}\kappa)d\kappa \cdot \frac{\Gamma(\alpha+1/2)}{\tilde{\beta}^{\alpha+1/2}} \\
&= \frac{\Gamma(\alpha+1/2)}{\Gamma(\alpha)}\frac{1}{\sqrt{2\pi\tau\beta}}\frac{\beta^{\alpha+1/2}}{\tilde{\beta}^{\alpha+1/2}} \\
&= \frac{\Gamma(\alpha+1/2)}{\Gamma(\alpha)}\frac{1}{\sqrt{2\pi\tau\beta}}\frac{\beta^{\alpha+1/2}}{[\beta+\frac{1}{2\tau}(\theta-\mu)^2]^{\alpha+1/2}} \\
&= \frac{\Gamma((2\alpha+1)/2)}{\Gamma((2\alpha)/2)}\frac{1}{\sqrt{\pi 2\alpha}}\sqrt{\frac{\alpha}{\tau\beta}}\frac{1}{\left\{1+\frac{\left(\sqrt{\frac{\alpha}{\tau\beta}}(\theta-\mu)\right)^2}{2\alpha}\right\}^{(2\alpha+1)/2}} \\
&= t_{2\alpha}\left(\sqrt{\frac{\alpha}{\tau\beta}}(\theta-\mu)\right)
\end{aligned}$$

where $t_{2\alpha}(x)$ is the t-density with 2α degrees of freedom. Similarly, for the marginal posterior density derived in (3),

$$\begin{aligned}
\lambda(\theta|X) &= \frac{\Gamma\left(\frac{(2\alpha+1)+1}{2}\right)}{\Gamma\left(\frac{2\alpha+1}{2}\right)}\frac{1}{\sqrt{\pi(2\alpha+1)}}\sqrt{\frac{(2\alpha+1)(n+1/\tau)}{\beta_n(\overline{X},\tau)}}\frac{1}{\left\{1+\frac{\left(\sqrt{\frac{(n+1/\tau)(2\alpha+1)}{\beta_n(\overline{X},\tau)}}(\theta-\mu(X,\tau))\right)^2}{2\alpha+1}\right\}^{\frac{(2\alpha+1)+1}{2}}} \\
&= t_{2\alpha+1}\left(\sqrt{\frac{(n+1/\tau)(2\alpha+1)}{\beta_n(\overline{X},\tau)}}(\theta-\mu_n(\overline{X},\tau))\right)
\end{aligned}$$

where $t_{2\alpha+1}(x)$ is the t density with $2\alpha+1$ degrees of freedom.
D. Since the t distribution is symmetric about zero, the posterior distribution of θ given $\overline{\overline{X}}$ is symmetric about

$$\mu_n(\overline{X}_n, n\tau) = \frac{\overline{X}_n + \frac{\mu}{n\tau}}{1+\frac{1}{n\tau}} = \frac{1}{1+1/(n\tau)}\overline{X}_n + \frac{1/(n\tau)}{1+1/(n\tau)}\mu,$$

and hence for any $\alpha > 0$ (since the mean of a t_r distribution is finite if $r > 1$) the resulting Bayes estimator $d_\Lambda(\underline{X}) = E\{\theta|\underline{X}\}$ of θ is $\mu_n(\overline{X}_n, \tau)$ But

$$\begin{aligned}
\sqrt{n}\{E(\theta|\underline{X}) - \theta\} &= \frac{1}{1+1/(n\tau)}\sqrt{n}(\overline{X}_n - \theta) + \sqrt{n}\left(\frac{1}{1+1/(n\tau)}\theta - \theta + \frac{1/(n\tau)}{1+1/(n\tau)}\mu\right) \\
&= \frac{1}{1+1/(n\tau)}\sqrt{n}(\overline{X}_n - \theta) + o(1) \\
&\to_d 1 \cdot N(0, 1/\kappa),
\end{aligned}$$

so the Bayes estimator is again asymptotically equivalent to the usual estimator, $\overline{X}_n$.