Statistics 582, Problem Set 5 Solutions

Wellner; 2/14/2007; Corrected 2/24/07

1. Suppose that $X_n \equiv X \sim \text{Multinomial}_k(n, \underline{\theta})$.
(a) Suppose that the prior distribution on θ is given by a Dirichlet distribution, Dirichlet($\underline{\alpha}$):

$$\lambda(\underline{\theta}) = \frac{\Gamma(\alpha_1 + \cdots + \alpha_k)}{\prod_{j=1}^k \Gamma(\alpha_j)} \theta_1^{\alpha_1 - 1} \cdots \theta_k^{\alpha_k - 1} 1_{[\underline{\theta}:\ \sum \theta_i = 1]} \, .$$

Verify the computation of the Bayes estimator for squared error loss given in example 4.3.4
(b) What is the posterior distribution for θ? Find the mode of the posterior distribution (along the lines of our computations of the MLE of the multinomial) and compare it with the MLE.
(c) Find a minimax estimator d_M of $\underline{\theta}$.

Solution: (a) If $\underline{\theta} \sim$ Dirichlet($\underline{\alpha}$) then $\theta_j \sim$ Beta($\alpha_j, \sum_{j' \neq j} \alpha_{j'}$), and hence from our computations of the mean of a Beta, $E(\theta_j) = \alpha_j / \sum_{i=1}^k \alpha_i$, and as a vector $E(\underline{\theta}) = \underline{\alpha} / \sum_{i=1}^k \alpha_i$ Since the posterior distribution of $\underline{\theta}$ is Dirichlet($\underline{\alpha} + \underline{X}$), the posterior mean is

$$d_\Lambda(\underline{X}) = E(\underline{\theta} | \underline{X}) = (\underline{\alpha} + \underline{X}) / (\sum_i \alpha_i + n).$$

(b) As noted in (a), the posterior density is Dirichlet($\underline{\alpha} + \underline{X}$):

$$\lambda(\underline{\theta} | \underline{X}) = \frac{\Gamma(\alpha_1 + \cdots + \alpha_k + n)}{\prod_{j=1}^k \Gamma(\alpha_j + X_j)} \theta_1^{\alpha_1 + X_1 - 1} \cdots \theta_k^{\alpha_k + X_k - 1} 1_{[\underline{\theta}:\ \sum \theta_j = 1]}.$$

To find the mode of the posterior, we need to find the value of $\underline{\theta}$ which maximizes $\lambda(\underline{\theta} | \underline{X})$ over the set $\sum_j \theta_j = 1$, or equivalently which maximizes

$$\sum_{j=1}^k (\alpha_j + X_j - 1) \log \theta_j + c(\sum_{j=1}^k \theta_j - 1).$$

Thus we need to solve

$$\frac{\alpha_j + X_j - 1}{\theta_j} + c = 0, \qquad j = 1, \ldots, k. \tag{1}$$

and

$$\sum_{j=1}^{k} \theta_j = 1. \tag{2}$$

The first equation yields

$$\theta_j^{mode} = \frac{\alpha_j + X_j - 1}{-c}, \qquad j = 1, \ldots, k;$$

substitution of this into (2) yields

$$1 = \sum_{j=1}^{k} \theta_j^{mode} = \frac{1}{-c} \left\{ \sum_{j=1}^{k} \alpha_j + n - k \right\},$$

and hence $-c = \sum_j \alpha_j + n - k$. Thus the mode of the posterior is given by

$$\underline{\theta}^{mode} = \frac{\underline{\alpha} + \underline{X} - \underline{1}}{\sum \alpha_j + n - k}.$$

When $\underline{\alpha} = \underline{1}$ (the vector of all 1's), then the mode of the posterior equals the MLE $\hat{\theta} = \underline{X}/n$. Note that $\underline{\alpha} = \underline{1}$ yields a uniform prior over θ.

(c) As shown in class, if $\underline{X} \sim \text{Mult}_k(n; \underline{\theta})$ and $\underline{\theta} \sim \text{Dirichlet}(\underline{\alpha})$, then the Bayes estimator of $\underline{\theta}$ for squared error loss is $d_\Lambda(\underline{X}) = (\underline{\alpha} + \underline{X})/(\sum \alpha_i + n)$ For $\alpha_1 = \cdots = \alpha_k = \alpha$, this yields the Bayes estimator

$$d_\Lambda(\underline{X}) \;=\; \frac{\alpha \underline{1} + \underline{X}}{k\alpha + n} = \frac{k\alpha}{k\alpha + n} \frac{\underline{1}}{k} + \frac{n}{k\alpha + n} \frac{\underline{X}}{n}.$$

Note that $d_{\Lambda,i}(\underline{X}) = (\alpha + X_i)/(k\alpha + n)$ has

$$Var_{\underline{\theta}}(d_{\Lambda,i}(X) = \frac{n\theta_i(1-\theta_i)}{(k\alpha + n)^2},$$

$$E_{\underline{\theta}}(d_{\Lambda,i}(X)) = \frac{\alpha + n\theta_i}{k\alpha + n},$$

$$\text{bias}_{\underline{\theta}}(d_{\Lambda,i}(X)) = \frac{\alpha - k\alpha\theta_i}{k\alpha + n}.$$

Thus the risk is

$$\begin{aligned}
R(\underline{\theta}, \underline{d}_\Lambda) &= E_{\underline{\theta}}|\underline{\theta} - \underline{d}_\Lambda(\underline{X})|^2 \\
&= \sum_{i=1}^{k} \left\{Var_{\underline{\theta}}(d_{\Lambda,i}(\underline{X})) + \text{bias}^2_{\underline{\theta}}(d_{\Lambda,i})\right\} \\
&= \frac{1}{(k\alpha+n)^2} \sum_{i=1}^{k} \left\{n\theta_i(1-\theta_i) + (\alpha - k\alpha\theta_i)^2\right\} \\
&= \frac{1}{(k\alpha+n)^2} \left\{n - k\alpha^2 + (\alpha^2 k^2 - n)\sum_{i=1}^{k} \theta_i^2\right\} \qquad \text{since} \quad \sum \theta_i = 1 \\
&= \frac{(1-1/k)}{(1+\sqrt{n})^2} \qquad \text{if} \;\; \alpha = \frac{\sqrt{n}}{k}.
\end{aligned}$$

which is constant in $\underline{\theta}$. Hence by corollary 5.6.3

$$\begin{aligned}
d_\Lambda(\underline{X}) &= \frac{\sqrt{n}}{\sqrt{n}+n}\frac{1}{k} + \frac{n}{\sqrt{n}+n}\frac{\underline{X}}{n} \\
&= (1-\lambda_n)\frac{1}{k} + \lambda_n \underline{\hat{p}}_n
\end{aligned}$$

is minimax for estimation of $\underline{\theta}$.

2. Find the limit distribution of the minimax estimator d_M in problem 1 (i.e. $\sqrt{n}(d_M(X_n) - p) \to_d$ "something" and find "something"). Is d_M a regular estimator of p?

   Solution: Note that $\sqrt{n}(1-\lambda_n) = \lambda_n \to 1$. Hence

$$\begin{aligned}
\sqrt{n}(d_M(\underline{X}_n) - \underline{\theta}) &= \sqrt{n}\{\lambda_n \underline{\hat{p}}_n + (1-\lambda_n)\frac{1}{k} - (\lambda_n + 1 - \lambda_n)\underline{\theta}\} \\
&= \lambda_n \sqrt{n}(\underline{\hat{\theta}}_n - \underline{\theta}) + \sqrt{n}(1-\lambda_n)(\frac{1}{k} - \underline{\theta}) \\
&\to_d N_k(0, \Sigma) + \frac{1}{k} - \underline{\theta} \\
&= N_k(\frac{1}{k} - \underline{\theta}, \Sigma)
\end{aligned}$$

   where $\Sigma = \text{diag}(\underline{\theta}) - \underline{\theta}\underline{\theta}^T$. To see that $d_M(\underline{X}_n)$ is a regular estimator of θ, let $\theta_n = \theta_0 + tn^{-1/2}$ where $1't = 0$ (so that $1'\theta_n = 1$). Then since $\hat{p}_n$ is a regular estimator of θ with

$$\sqrt{n}(\hat{p}_n - \theta_n) \to_d Z \sim N_k(0, \text{diag}(\theta_0) - \theta_0\theta_0')$$

under P_{θ_n} (which follows from the Liapunov CLT together with the Cram'er-Wold device, or from contiguity theory), it follows that

$$\begin{aligned}
\sqrt{n}(d_M(\underline{X}_n) - \theta_n) &= \sqrt{n}((1-\lambda_n)(1/k) + \lambda_n \hat{p}_n - \theta_n) \\
&= \lambda_n \sqrt{n}(\hat{p}_n - \theta_n) + \sqrt{n}(1-\lambda_n)(1/k - \theta_n) \\
&\to_d 1 \cdot Z + 1 \cdot (1/k - \theta_0) \\
&\sim N_k((1/k - \theta_0), \mathrm{diag}(\theta_0) - \theta_0\theta_0'),
\end{aligned}$$

where we used $\sqrt{n}(1-\lambda_n) = \lambda_n \to 1$ and $\theta_n \to \theta_0$. Since this limiting distribution does not depend on t, $d_M(\underline{X}_n)$ is regular.

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; i.e. show that it has a risk function $R(\theta, d_0)$ which is constant in θ.
(b) Show that the usual estimator d_0 is generalized Bayes with respect to Lebesgue measure on $(0, \infty)$ provided $r > 1$. (A generalized Bayes rule is a rule that minimizes the posterior Bayes risk even when starting with an improper prior; see e.g. Ferguson, MS, page 50.) (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

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

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

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

(corrected on 2/24/07, thanks to Krisztian Sebestyen), we find the Bayes risk of the Bayes rule d_Λ to be

$$
\begin{aligned}
\mathcal{R}(\Lambda, d_\Lambda) &= \frac{1}{(r+\beta+1)^2}\left\{r + (\alpha-1)^2\frac{\beta(\beta+1)}{(\alpha-1)(\alpha+\beta)}\right. \\
&\qquad \left. - 2(\alpha-1)(\beta+1)\frac{\beta}{\alpha+\beta} + (\beta+1)^2\frac{\alpha}{\alpha+\beta}\right\} \\
&= \frac{1}{r+\beta+1} \qquad (3) \\
&\to \frac{1}{r+1} \qquad \text{as} \;\; \beta \to 0. \qquad (4)
\end{aligned}
$$

(d) The rule $d(X) = X/(r+1)$ corresponding to the limiting Bayes risk in (4) 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

$$\beta_n(x,\tau) \equiv \beta+\frac{1}{2}\frac{1/\tau}{(n+1/\tau)}(x-\mu)^2. \tag{5}$$

Hence

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

(c) Thus the marginal posterior distribution of θ is

$$
\begin{aligned}
\lambda(\theta|\overline{X}) &= \int_0^\infty \lambda(\theta,\kappa|\overline{X}) d\kappa \\
&= \int_0^\infty \kappa^\alpha \exp\left\{-\kappa\left(\beta_n(\overline{X},\tau)+\frac{n+1/\tau}{2}(\theta-\mu_n(\overline{X},\tau))^2\right)\right\} d\kappa \\
&\quad \sqrt{\frac{n+1/\tau}{2\pi}}\frac{\beta_n(\overline{X},\tau)^{\alpha+1/2}}{\Gamma(\alpha+1/2)} \\
&= \int_0^\infty \kappa^{\alpha+1-1}\frac{\tilde{\beta}^{\alpha+1}}{\Gamma(\alpha+1)} \exp(-\tilde{\beta}\kappa) d\kappa \cdot \frac{\Gamma(\alpha+1)}{\tilde{\beta}^{\alpha+1}}\sqrt{\frac{n+1/\tau}{2\pi}}\frac{\beta_n(\overline{X},\tau)^{\alpha+1/2}}{\Gamma(\alpha+1/2)} \\
&= \frac{\Gamma(\alpha+1)}{\tilde{\beta}^{\alpha+1}}\sqrt{\frac{n+1/\tau}{2\pi}}\frac{\beta_n(\overline{X},\tau)^{\alpha+1/2}}{\Gamma(\alpha+1/2)} \\
&= \frac{\Gamma(\alpha+1)}{\Gamma(\alpha+1/2)}\sqrt{\frac{n+1/\tau}{2\pi}}\frac{\beta_n(\overline{X},\tau)^{\alpha+1/2}}{\tilde{\beta}^{\alpha+1}} \\
&= \frac{\Gamma(\alpha+1)}{\Gamma(\alpha+1/2)}\sqrt{\frac{n+1/\tau}{2\pi}}\frac{\beta_n(\overline{X},\tau)^{\alpha+1/2}}{\left(\beta_n(\overline{X},\tau)+\frac{n+1/\tau}{2}(\theta-\mu_n(\overline{X},\tau))^2\right)^{\alpha+1}} \qquad (6)
\end{aligned}
$$

with $\beta_n(x,\tau)$ as defined in (5)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 (6),

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

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 $\overline{X}$ 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

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

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