STA3000 : 2020 Solutions - Exercises 2

1. (i) The likelihood function is

$$L(\mu,\sigma^2 \mid \underset{\sim}{x}) = (\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\}$$
$$= (\sigma^2)^{-n/2} \exp\{-\frac{n}{2\sigma^2}(\bar{x}-\mu)^2\}\exp\{-\frac{1}{2\sigma^2}\|\underset{\sim}{x}-\bar{x}\underset{\sim}{1}\|^2\}$$

So given $(\bar{x}, \|\underset{\sim}{x}-\bar{x}\underset{\sim}{1}\|^2)$ we know the likelihood fn and from the likelihood we can compute $(\bar{x}, \|\underset{\sim}{x}-\bar{x}\underset{\sim}{1}\|^2/n)$ the MLE of (μ,σ^2). Therefore $(\bar{x}, \|\underset{\sim}{x}-\bar{x}\underset{\sim}{1}\|^2/n)$ is a mss and thus $(\bar{x}, \|\underset{\sim}{x}-\bar{x}\underset{\sim}{1}\|^2)$ is also a mss.
By well-known results we have $\bar{x} \sim N(\mu, \sigma^2/n)$ stat. ind. of $\|\underset{\sim}{x}-\bar{x}\underset{\sim}{1}\|^2 \sim \sigma^2$ chi-squared $(n-1)$.

(ii) The posterior dist. depends on the data only through a mss so the joint posterior of $(\mu, 1/\sigma^2)$ is proportional to

$$(\sigma^2)^{-\frac{1}{2}}\exp\{-\frac{n}{2\sigma^2}(\bar{x}-\mu)^2\}(\sigma^2)^{-\frac{1}{2}}\exp\{-\frac{1}{2\tau_0^2\sigma^2}(\mu-\mu_0)^2\} \times$$
$$\left(\frac{1}{\sigma^2}\right)^{\frac{n-1}{2}}\exp\{-\frac{\|\underset{\sim}{x}-\bar{x}\underset{\sim}{1}\|^2}{2\sigma^2}\}\left(\frac{1}{\sigma^2}\right)^{\alpha_0-1}\exp\{-\frac{\beta_0}{\sigma^2}\}$$

and using $a(\mu+b)^2 + c(\mu+d)^2$
$= (a+c)(\mu+(a+c)^{-1}(ab+cd))^2 + ac(a+c)^{-1}(b-d)^2$ with

$$a = \frac{n}{\sigma^2},\ b = -\bar{x},\ c = \frac{1}{\tau_0^2\sigma^2},\ d = -\mu_0$$

Then $\frac{n}{\sigma^2}(\bar{x}-\mu)^2 + \frac{1}{\tau_0^2\sigma^2}(\mu-\mu_0)^2$ =

$$\left(\frac{n}{\sigma^2}+\frac{1}{\sigma^2\tau_0^2}\right)\left(\mu-\left(\frac{n}{\sigma^2}+\frac{1}{\sigma^2\tau_0^2}\right)^{-1}\left(\frac{n\bar{x}}{\sigma^2}+\frac{\mu_0}{\tau_0^2\sigma^2}\right)\right)^2$$

$$+\frac{n}{\sigma^4\tau_0^2}\left(\frac{n}{\sigma^2}+\frac{1}{\sigma^2\tau_0^2}\right)^{-1}(\bar{x}-\mu_0)^2$$

$$=\frac{1}{\sigma^2}\left(n+\frac{1}{\tau_0^2}\right)\left(\mu-\left(n+\frac{1}{\tau_0^2}\right)^{-1}\left(n\bar{x}+\mu_0/\tau_0^2\right)\right)^2$$

$$+\frac{n}{\sigma^2\tau_0^2}\left(n+\frac{1}{\tau_0^2}\right)^{-1}(\bar{x}-\mu_0)^2$$

From this and the expression for the joint posterior we conclude

$$\mu\,|\,\underline{x},\frac{1}{\sigma^2}\sim N\left(\left(n+\frac{1}{\tau_0^2}\right)^{-1}\left(n\bar{x}+\frac{\mu_0}{\tau_0^2}\right),\ \sigma^2\left(n+\frac{1}{\tau_0^2}\right)^{-1}\right)$$

$$\frac{1}{\sigma^2}\,\Big|\,\underline{x}\sim \text{gamma}_{\text{rate}}\left(\frac{n}{2}+\alpha_0,\ \beta_0+\frac{\|\underline{x}-\bar{x}\underline{1}\|^2}{2}+\frac{(\tau_0^2+\frac{1}{n})^{-1}(\bar{x}-\mu_0)^2}{2}\right)$$

$$=\text{gamma}_{\text{rate}}(\alpha_0(\underline{x}),\beta_0(\underline{x}))$$

(iii) Since we are using quadratic loss the Bayes rule is given by the posterior mean of μ which is

$$E(\mu\,|\,\underline{x})\overset{\text{TTE}}{=}E\left(E\left(\mu\,|\,\underline{x},\frac{1}{\sigma^2}\right)\right)$$

$$=E\left(\left(n+\frac{1}{\tau_0^2}\right)^{-1}\left(n\bar{x}+\frac{\mu_0}{\tau_0^2}\right)\right)=\left(n+\frac{1}{\tau_0^2}\right)^{-1}\left(n\bar{x}+\frac{\mu_0}{\tau_0^2}\right)=b(\underline{x})$$

The risk of this estimate is

$$R((\mu,\sigma^2),b)=E_{(\mu,\sigma^2)}((b(\underline{x})-\mu)^2)=$$

(using bias^2 + variance formula)

$$\left(\left(n+\frac{1}{\tau_0^2}\right)^{-1}\left(n\mu+\frac{\mu_0}{\tau_0^2}\right)-\mu\right)^2+\left(n+\frac{1}{\tau_0^2}\right)^{-2}n\sigma^2$$

$$=\frac{1}{\tau_0^4}\left(n+\frac{1}{\tau_0^2}\right)^{-2}(\mu-\mu_0)^2+\left(n+\frac{1}{\tau_0^2}\right)^{-2}n\sigma^2$$

$$=\left(n+\frac{1}{\tau_0^2}\right)^{-2}\left[\frac{(\mu-\mu_0)^2}{\tau_0^4}+n\sigma^2\right]$$ and

then the Bayes risk is the expectation of this wrt

the prior which equals

$$E\left[E\left[\cdot \mid \tfrac{1}{\sigma^2}\right]\right]$$

$$=\left(n+\frac{1}{\tau_0^2}\right)^{-2}\left(\frac{1}{\tau_0^2}+n\right)E[\sigma^2]=\left(n+\frac{1}{\tau_0^2}\right)^{-1}E(\sigma^2)$$

and $$E(\sigma^2)=\int_0^\infty \sigma^2\,\frac{\beta_0^{\alpha_0}}{\Gamma(\alpha_0)}\left(\frac{1}{\sigma^2}\right)^{\alpha_0-1}e^{-\beta_0/\sigma^2}\,d\!\left(\frac{1}{\sigma^2}\right)$$

$$=\frac{\beta_0^{\alpha_0}}{\Gamma(\alpha_0)}\int_0^\infty v^{-1}v^{\alpha_0-1}e^{-\beta_0 v}\,dv$$

$$=\frac{\beta_0^{\alpha_0}}{\Gamma(\alpha_0)}\int_0^\infty v^{(\alpha_0-1)-1}e^{-\beta_0 v}\,dv=\frac{\beta_0^{\alpha_0}}{\Gamma(\alpha_0)}\,\frac{\Gamma(\alpha_0-1)}{\beta_0^{\alpha_0-1}}$$

$=\beta_0/(\alpha_0-1)$ Therefore the Bayes risk

is $\left(n+\frac{1}{\tau_0^2}\right)^{-1}\frac{\beta_0}{(\alpha_0-1)}$ (provided $\alpha_0>1$ and is ∞ otherwise).

(iv) Again the Bayes rule is given by

$$E(\mu^3 \mid \underline{x}) = E\left(E\left(\mu^3 \mid \underline{x}, \tfrac{1}{\sigma^2}\right)\right)$$

If $X \sim N(\mu_y, \sigma_y^2)$ then $Z = (X-\mu_y)/\sigma_y \sim N(0,1)$

so $E(X^3) = E\left((\mu_y + \sigma_y Z)^3\right)$

$$= E\left(\mu_y^3 + 3\mu_y\sigma_y^2 Z^2 + 3\mu_y^2\sigma_y Z + \sigma_y^3 Z^3\right)$$

$$= \mu_y^3 + 3\mu_y\sigma_y^2 \quad \text{since } E(Z) = E(Z^3) = 0.$$

Therefore $E\left(\mu^3 \mid \underline{x}, \frac{1}{\sigma^2}\right)$

$$= \left(n + \tfrac{1}{\tau_0^2}\right)^{-3}\left(n\bar{x} + \tfrac{\mu_0}{\tau_0^2}\right)^3 + 3\left(n + \tfrac{1}{\tau_0^2}\right)^{-2}\left(n\bar{x} + \tfrac{\mu_0}{\tau_0^2}\right)\sigma^2$$

and so $E(\mu^3 \mid \underline{x}) = b(\underline{x})\left(b^2(\underline{x}) + 3\left(n + \tfrac{1}{\tau_0^2}\right)^{-1} E(\sigma^2 \mid \underline{x})\right)$

and using the computation on the previous page

$$E(\sigma^2 \mid \underline{x}) = \frac{\beta_0(\underline{x})}{\alpha_0(\underline{x})-1} = \frac{\beta_0 + \|\underline{x} - \bar{x}\underline{1}\|^2/2 + \left(\tau_0^2 + \frac{1}{n}\right)^{-1}(\bar{x}-\mu_0)^2/2}{\frac{n}{2} + \alpha_0 - 1}$$

So the Bayes rule is $b^3(\underline{x}) + 3\, b(\underline{x})\left(n + \tfrac{1}{\tau_0^2}\right)^{-1}\dfrac{\beta_0(\underline{x})}{\alpha_0(\underline{x})-1} \neq b^3(\underline{x})$,

so Bayes rules are not invariant.

(v) The posterior risk is

$$\int_0^\infty \frac{(\sigma^2-a)^2}{\sigma^2} \frac{\beta_0(\underline{x})^{\alpha_0(\underline{x})}}{\Gamma(\alpha_0(\underline{x}))} \left(\frac{1}{\sigma^2}\right)^{\alpha_0(\underline{x})-1} e^{-\beta_0(\underline{x})/\sigma^2} d\left(\frac{1}{\sigma^2}\right)$$

$$= \frac{\beta_0^{-1}(\underline{x})}{\Gamma(\alpha_0(\underline{x}))} \Gamma(\alpha_0(\underline{x})+1) \int_0^\infty (\sigma^2-a)^2 \frac{\beta_0(\underline{x})^{\alpha_0(\underline{x})+1}}{\Gamma(\alpha_0(\underline{x})+1)} \left(\frac{1}{\sigma^2}\right)^{\alpha_0(\underline{x})+1-1} e^{-\beta_0(\underline{x})/\sigma^2} d\left(\frac{1}{\sigma^2}\right)$$

which is minimized by the mean of σ^2 when $\frac{1}{\sigma^2} \sim$ gamma$(\alpha_0(\underline{x})+1, \beta_0(\underline{x}))$ distribution so the Bayes rule is $\sigma^2(\underline{x}) = \frac{\beta_0(\underline{x})}{\alpha_0(\underline{x})}$ and the posterior risk is

$$\left(\frac{\alpha_0(\underline{x})}{\beta_0(\underline{x})}\right)\left(\frac{\beta_0^2(\underline{x})}{\alpha_0(\underline{x})(\alpha_0(\underline{x})-1)}\right) = \frac{\beta_0(\underline{x})}{\alpha_0(\underline{x})(\alpha_0(\underline{x})-1)}$$

Then the Bayes risk is $E\left(\frac{\beta_0(\underline{x})}{\alpha_0(\underline{x})\,(\alpha_0(\underline{x})-1)}\right)$

$$= \frac{1}{(\alpha_0+\frac{n}{2})(\alpha_0+\frac{n}{2}-1)}\left(\beta_0 + \tfrac{1}{2}E(\|\underline{x}-\bar{x}\|^2) + \frac{(\tau_0^2+\frac{1}{n})^{-1}}{2} E((\bar{x}-\mu_0)^2)\right)$$

where E is the expectation operator wrt the prior predictive distribution of $(\bar{x}, \|\underline{x}-\bar{x}\|^2)$

and $E((\bar{x}-\mu_0)^2) = E_\pi(E((\bar{x}-\mu_0)^2 \mid \mu, \sigma^2))$

$$= E_\pi\left((\mu-\mu_0)^2 + \frac{\sigma^2}{n}\right) = (\tau_0^2+\tfrac{1}{n}) E_\pi(\sigma^2) = (\tau_0^2+\tfrac{1}{n})\frac{\beta_0}{\alpha_0-1}$$

and $E(\|x - \bar{x}1\|^2) = E_\pi\left(E(\|x - \bar{x}1\|^2 \mid \mu, \sigma^2)\right)$

$= E_\pi\left((n-1)\sigma^2\right) = (n-1)\dfrac{\beta_0}{\alpha_0 - 1}$

Therefore, the Bayes risk is

$$\left(\alpha_0 + \frac{n}{2}\right)^{-1}\left(\alpha_0 + \frac{n}{2} - 1\right)^{-1}\left(\beta_0 + \frac{(n-1)}{2}\frac{\beta_0}{\alpha_0 - 1} + \frac{1}{2}\frac{\beta_0}{\alpha_0 - 1}\right).$$

$$= \left(\alpha_0 + \frac{n}{2}\right)^{-1}(\alpha_0 - 1)^{-1}\beta_0.$$

2) (i) The LF is given by $\theta^{\Sigma x_i}(1-\theta)^{n-\Sigma x_i}$
so it is determined by Σx_i. Since the MLE of θ is $\bar{x}$ this proves that knowing Σx_i is equivalent to knowing the LF so Σx_i is a mss.

(ii) $\sum_{i=1}^{n} x_i \sim$ binomial(n, θ)

(iii) $\int_0^1 (\theta - a)^2 \delta(\bar{x}, da) = \frac{n}{n+1}(\theta - \bar{x})^2 + \frac{1}{n+1}(\theta - \frac{1}{2})^2$

so $R(\theta, \delta) = \frac{n}{n+1} E_\theta((\theta - \bar{x})^2) + \frac{1}{n+1} E_\theta((\theta - \frac{1}{2})^2)$

$= \frac{n}{n+1}\frac{\theta(1-\theta)}{n} + \frac{1}{n+1}(\theta - \frac{1}{2})^2 = \frac{1}{n+1}\left[\theta(1-\theta) + (\theta - \frac{1}{2})^2\right]$

$= \frac{1}{4(n+1)}$ so $\sup_\theta R(\theta, \delta) = \frac{1}{4(n+1)}$

$< \frac{1}{4n} = \sup_\theta R(\theta, \delta)$ so $\bar{x}$ is not minimax.

(iv) The posterior is proportional to
$\theta^{n\bar{x}}(1-\theta)^{n(1-\bar{x})}\theta^{\alpha-1}(1-\theta)^{\beta-1} = \theta^{n\bar{x}+\alpha-1}(1-\theta)^{n(1-\bar{x})+\beta-1}$
and so $\theta | \underline{x} \sim$ beta$(n\bar{x}+\alpha, n(1-\bar{x})+\beta)$.

(v) Since the loss is quadratic loss the Bayes rule is the posterior mean $(n\bar{x}+\alpha)/(n+\alpha+\beta) = b(\bar{x})$

(vi)(vii) $R(\theta, b(\bar{x})) = E_\theta((b(\bar{x}) - \theta)^2)$

$=$ bias^2 + variance $= \left(\frac{n\theta + \alpha}{n+\alpha+\beta} - \theta\right)^2 + \frac{n}{(n+\alpha+\beta)^2}\theta(1-\theta)$

$= (n+\alpha+\beta)^{-2}\left[(\alpha - (\alpha+\beta)\theta)^2 + n\theta(1-\theta)\right]$

$$= (n+\alpha+\beta)^{-2}\left[\alpha^2 - 2(\alpha(\alpha+\beta)-n)\theta + ((\alpha+\beta)^2-n)\theta^2\right]$$

So $b(\bar{x})$ has constant risk when

$\alpha+\beta = \sqrt{n}$ and $\alpha(\alpha+\beta) = \alpha\sqrt{n} = n$ so $\alpha = \frac{\sqrt{n}}{2}$, $\beta = \frac{\sqrt{n}}{2}$

and the risk is $(n+\sqrt{n})^{-2}\frac{n}{4} = \frac{1}{4(\sqrt{n}+1)^2}$.

Since the Bayes rule wrt prior beta$\left(\frac{\sqrt{n}}{2}, \frac{\sqrt{n}}{2}\right)$ has constant risk this implies the Bayes rule is minimax and the prior is least favorable.

Also the Bayes rule is admissible and as it has constant risk.

(viii) Note that any prior π that has

$$\frac{\int_0^1 \theta^{x+1}(1-\theta)^{n-x}\,\pi(d\theta)}{\int_0^1 \theta^{x}(1-\theta)^{n-x}\,\pi(d\theta)} = b(x)$$

will give the same Bayes rule and this depends only on the first $n+1$ moments of π. This implies that there are many priors that have Bayes rule $b(\bar{x})$ for any choice of (α,β) and thus for $(\alpha,\beta) = \left(\frac{\sqrt{n}}{2}, \frac{\sqrt{n}}{2}\right)$. So the prior is not unique although the estimate is.