Name: ____________________________________

STAT 206B Midterm 1, February 3rd, 2022 (Thursday)

You MUST show all your work and justify each of your steps for the full credit. You may use any results from lecture without proving, but please state the results and why you use them.

Some pdfs: Unless pdfs with different parameterizations are stated, use the following pdfs for Gamma distributions and Inverse-Gamma distributions.

- Inverse-Gamma distribution with α and β

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} \exp\left(-\frac{\beta}{x}\right), \quad x \in \mathbb{R}^+$$

  where $\mathrm{E}(X) = \beta/(\alpha-1)$ for $\alpha > 1$ and $\mathrm{Var}(X) = \beta^2/\{(\alpha-1)^2(\alpha-2)\}$ for $\alpha > 2$.

1. Let $x_1, \ldots, x_n$ be an iid sample from $\mathrm{N}(\theta, \sigma^2)$ with $\sigma > 0$ known and suppose that θ has the normal prior, $\theta \sim \mathrm{N}(\mu, \tau^2)$, with fixed $\mu \in \mathbb{R}$ and $\tau^2 > 0$.

   Fact: The posterior distribution of θ, $\pi(\theta \mid \boldsymbol{x})$ is $\mathrm{N}(\mu_1, \tau_1^2)$ with

$$\tau_1^2 = \left(\frac{1}{\sigma^2/n} + \frac{1}{\tau^2}\right)^{-1}, \text{ and } \quad \mu_1 = \tau_1^2 \left(\frac{\bar{x}}{\sigma^2/n} + \frac{\mu}{\tau^2}\right),$$

   where $\boldsymbol{x} = (x_1, \ldots, x_n)$ and $\bar{x} = \sum_{i=1}^n x_i/n$. Consider a squared error loss $L(\theta, a) = (\theta - a)^2$.

   (a) (15 pts) Derive the Bayes estimator $\delta^\pi(\boldsymbol{x})$ under the loss function and compute the frequentist risk of $\delta^\pi(\boldsymbol{x})$, $R(\theta, \delta^\pi)$.

   (b) (5 pts) The maximum likelihood estimator (MLE) for the normal mean estimation problem is $\delta_1(\boldsymbol{x}) = \bar{x}$. Compute the frequentist risk of $\delta_1(\boldsymbol{x})$, $R(\theta, \delta_1)$.

   (c) (5 pts) Compare δ_1 and δ^π. Is δ^π admissible? Explain why.

   (d) (5 pts) Consider δ_1 and δ^π. Which one is a minimax estimator? Explain why.

   Note: You may plot $R(\theta, \delta^\pi)$ and $R(\theta, \delta_1)$ for parts (c) and (d).

2. Assume that observations, $y_i \mid x_i, \theta, \sigma \overset{indep}{\sim} \mathrm{N}(\theta x_i, \sigma^2)$, $i = 1, \ldots, n$, where $\theta \in \mathbb{R}$ and $\sigma^2 > 0$ are unknown, and $x_1, \ldots, x_n$ are known and fixed. Consider $\pi(\theta, \sigma^2) = \pi_1(\theta \mid \sigma^2)\pi_2(\sigma^2)$ and let $\pi_1(\theta \mid \sigma^2) = \mathrm{N}(\mu, \kappa\sigma^2)$ and $\pi_2(\sigma^2) = \mathrm{IG}(a, b)$, where $\mu \in \mathbb{R}$, $\kappa > 0$, $a > 0$ and $b > 0$ are fixed.

   (a) (8 pts) Find the joint posterior distribution $\pi(\theta, \sigma^2 \mid \boldsymbol{x}, \boldsymbol{y})$ up to a proportionality constant, where $\boldsymbol{y} = (y_1, \ldots, y_n)$ and $\boldsymbol{x} = (x_1, \ldots, x_n)$.

   (b) (15 pts) Find the posterior distributions $\pi(\theta \mid \boldsymbol{x}, \boldsymbol{y}, \sigma^2)$.

   (c) (15 pts) Find the posterior distributions $\pi(\sigma^2 \mid \boldsymbol{x}, \boldsymbol{y})$.

(END)

1. (a) Bayes estimator $\Leftrightarrow$ minimize the posterior expected loss

$$\rho(\pi, d) = E^{\pi}\left((\theta - d)^2 \mid x \right)$$

$$= E^{\pi}\left((\theta - E(\theta \mid x) + E(\theta \mid x) - d)^2 \mid x \right)$$

$$= E^{\pi}\left((\theta - E(\theta \mid x))^2 \mid x \right) + E\left((E(\theta \mid x) - d)^2 \mid x \right)$$
$$+ 2 E^{\pi}\left(\underbrace{\cancel{\theta - E(\theta \mid x)} \mid x}_{0} \right) \cdot \left(E(\theta \mid x) - d \right)$$

$$= \underbrace{\mathrm{Var}^{\pi}(\theta \mid x)}_{\geq 0} + \underbrace{\left(E(\theta \mid x) - d \right)^2}_{\geq 0}$$

$$(E(\theta \mid x) - d)^2 = 0 \quad \text{iff} \quad d = E^{\pi}(\theta \mid x)$$

$$\Rightarrow \quad \rho(\pi, d) \geq \mathrm{Var}^{\pi}(\theta \mid x) \ \forall d \quad \& \quad \rho(\pi, d) = \mathrm{Var}^{\pi}(\theta \mid x) \quad \text{iff} \quad d = E^{\pi}(\theta \mid x)$$

$$\Rightarrow \quad \delta^{\pi} = E^{\pi}(\theta \mid x) = \left(\frac{1}{\sigma^2/n} + \frac{1}{\tau^2} \right)^{-1} \left(\frac{\bar{x}}{\sigma^2/n} + \frac{\mu}{\tau^2} \right)$$

$$= \underbrace{\frac{\tau^2}{\sigma^2/n + \tau^2}}_{=a} \bar{x} + \underbrace{\frac{\sigma^2/n}{\sigma^2/n + \tau^2}}_{=b \,= (1-a)} \mu \qquad a + b = 1 \qquad 0 < a, b < 1$$

$$E_{\theta}(\delta^{\pi}) = E_{\theta}\left[\left(\frac{1}{\sigma^2/n} + \frac{1}{\tau^2} \right)^{-1} \left(\frac{\bar{x}}{\sigma^2/n} + \frac{\mu}{\tau^2} \right) \right] = \left(\frac{1}{\sigma^2/n} + \frac{1}{\tau^2} \right)^{-1} \left(\frac{\theta}{\sigma^2/n} + \frac{\mu}{\tau^2} \right)$$

$$= \frac{\tau^2}{\sigma^2/n + \tau^2} \theta + \frac{\sigma^2/n}{\sigma^2/n + \tau^2} \mu$$

$$= a\theta + b\mu$$

$$R(\theta, \delta^\pi) = E_\theta\left((\theta - \delta^\pi)^2 \right)$$

$$= E_\theta\left((\theta - E_\theta(\delta^\pi) + E_\theta(\delta^\pi) - \delta^\pi)^2 \right)$$

$$= (\theta - a\theta - b\mu)^2 + \mathrm{Var}_\theta(\delta^\pi)$$

$$= b^2(\theta-\mu)^2 + \mathrm{Var}_\theta(a\bar{x} + b\mu)$$

$$= b^2(\theta-\mu)^2 + a^2\frac{\sigma^2}{n}$$

$$= \left(\frac{\sigma^2/n}{\sigma^2/n+\tau^2}\right)^2(\theta-\mu)^2 + \left(\frac{\tau^2}{\sigma^2/n+\tau^2}\right)^2\frac{\sigma^2}{n}$$

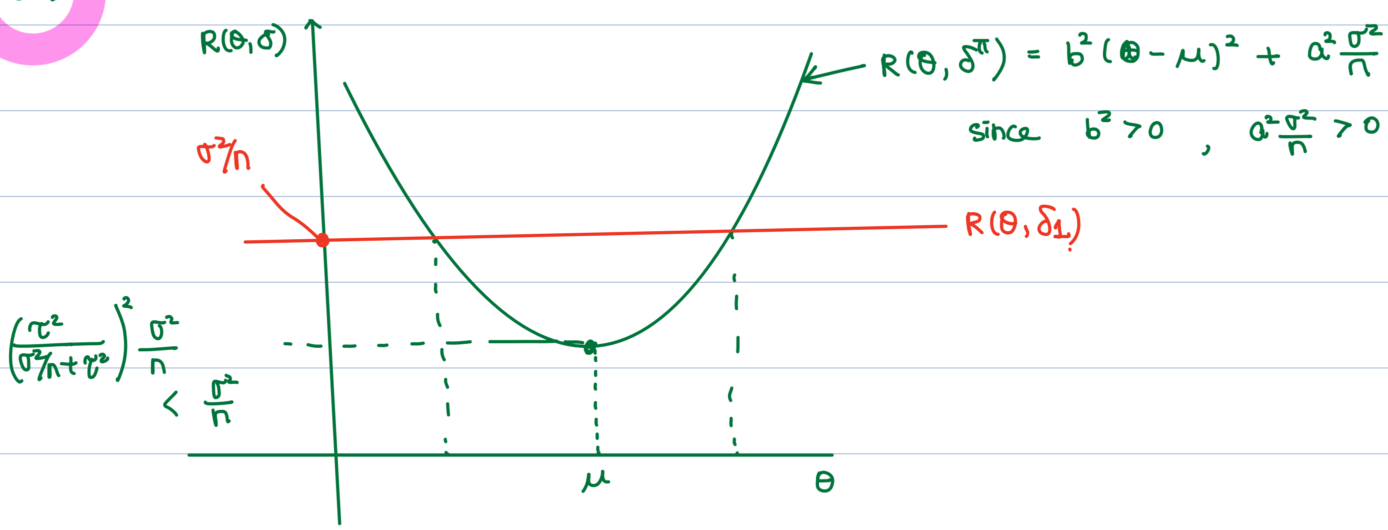
(b) $\delta_1 = \bar{x} \Rightarrow E_\theta(\delta_1) = \theta$

$$R(\theta, \delta_1) = E_\theta\left((\theta - \delta_1)^2 \right)$$

$$= E_\theta\left((\theta - \bar{x})^2 \right)$$

$$= \frac{\sigma^2}{n}$$

(c)



$$b^2(\theta-\mu)^2 + a^2\frac{\sigma^2}{n} = \frac{\sigma^2}{n} \qquad \left(\text{i.e, } R(\theta,\delta^\pi) = R(\theta,\delta_1)\right)$$

$$b^2(\theta-\mu)^2 = (1-a^2)\frac{\sigma^2}{n} \qquad (1-a)(1+a)$$

$$(\theta-\mu)^2 = \frac{(1+a)}{b}\frac{\sigma^2}{n}$$

$$\theta = \mu \pm \sqrt{\frac{(1+a)}{b}\frac{\sigma^2}{n}}$$

$\Rightarrow$ if $\theta \in \left(\mu \pm \sqrt{\frac{(1+a)}{b}\frac{\sigma^2}{n}}\right)$, $R(\theta,\delta^\pi) < R(\theta,\delta_1)$

Otherwise, $R(\theta,\delta^\pi) \geq R(\theta,\delta_1)$

$\Rightarrow$ δ^π is admissible.

(d) $\max R(\theta,\delta^\pi) > \max R(\theta,\delta_1)$ $\Rightarrow$ δ_1 is the minimax

2-(a)

$$\pi(\theta, \sigma^2 \mid x, y) \propto \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_i-\theta x_i)^2}{2\sigma^2}\right) \cdot \frac{1}{\sqrt{2\pi K\sigma^2}} \exp\left(-\frac{(\theta-\mu)^2}{2K\sigma^2}\right)$$

$$\times (\sigma^2)^{-a-1} \exp\left(-\frac{b}{\sigma^2}\right)$$

$$\propto (\sigma^2)^{-n/2-a-\frac{1}{2}-1} \exp\left(-\sum_{i=1}^{n}\frac{(y_i-\theta x_i)^2}{2\sigma^2} - \frac{(\theta-\mu)^2}{2K\sigma^2} - \frac{b}{\sigma^2}\right)$$

(b) $\pi(\theta \mid \sigma^2, x, y) \propto \pi(\theta, \sigma^2 \mid x, y)$

$$\propto \exp\left(-\sum_{i=1}^{n}\frac{(y_i-\theta x_i)^2}{2\sigma^2} - \frac{(\theta-\mu)^2}{2K\sigma^2}\right)$$

$$\propto \exp\left[-\frac{1}{2}\left\{\left(\frac{\Sigma x_i^2}{\sigma^2}+\frac{1}{K\sigma^2}\right)\theta^2 - 2\left(\frac{\Sigma x_i y_i}{\sigma^2}+\frac{\mu}{K\sigma^2}\right)\theta\right\}\right]$$

$\Rightarrow$ observe a kernel for $N\left(\left(\frac{\Sigma x_i^2}{\sigma^2}+\frac{1}{K\sigma^2}\right)^{-1}\left(\frac{\Sigma x_i y_i}{\sigma^2}+\frac{\mu}{K\sigma^2}\right), \left(\frac{\Sigma x_i^2}{\sigma^2}+\frac{1}{K\sigma^2}\right)^{-1}\right)$

$$\Rightarrow \pi(\theta \mid \sigma^2, x, y) \text{ is } N\Bigg(\underbrace{\left(\frac{\Sigma x_i^2}{\sigma^2}+\frac{1}{K\sigma^2}\right)^{-1}\left(\frac{\Sigma x_i y_i}{\sigma^2}+\frac{\mu}{K\sigma^2}\right)}_{=\left(\Sigma x_i^2+\frac{1}{K}\right)^{-1}\left(\Sigma x_i y_i+\frac{\mu}{K}\right)}, \underbrace{\left(\frac{\Sigma x_i^2}{\sigma^2}+\frac{1}{K\sigma^2}\right)^{-1}}_{=\sigma^2\left(\Sigma x_i^2+\frac{1}{K}\right)^{-1}}\Bigg)$$

(c) $\pi(\sigma^2 \mid x, y) = \int_{-\infty}^{\infty} \pi(\theta, \sigma^2 \mid x, y)\, d\theta$

$$\propto \int_{-\infty}^{\infty} (\sigma^2)^{-n/2 - a - 1 - \frac{1}{2}} \exp\left(- \sum_{i=1}^{n} \frac{(y_i - \theta x_i)^2}{2\sigma^2} - \frac{(\theta - \mu)^2}{2K\sigma^2} - \frac{b}{\sigma^2} \right) d\theta$$

$$= (\sigma^2)^{-n/2 - a - 1 - \frac{1}{2}} \exp\left(- \frac{b}{\sigma^2} - \frac{\Sigma y_i^2}{2\sigma^2} - \frac{\mu^2}{2K\sigma^2} \right)$$

$$\cdot \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2} \left\{ \left(\frac{\Sigma x_i^2}{\sigma^2} + \frac{1}{K\sigma^2} \right) \theta^2 - 2 \left(\frac{\Sigma x_i y_i}{\sigma^2} + \frac{\mu}{K\sigma^2} \right) \theta \right\} \right] d\theta$$

$$\propto (\sigma^2)^{-n/2 - a - 1 - \frac{1}{2}} \exp\left(- \frac{b}{\sigma^2} - \frac{\Sigma y_i^2}{2\sigma^2} - \frac{\mu^2}{2K\sigma^2} \right)$$

$$\times \quad (\sigma^2)^{+1/2} \cdot \exp\left(+ \frac{1}{2\sigma^2} \left(\Sigma x_i^2 + \frac{1}{K} \right)^{-1} \left(\Sigma x_i y_i + \frac{\mu}{K} \right)^2 \right)$$

$$\Rightarrow \quad IG\left(\frac{n}{2} + a, \quad b + \frac{\Sigma y_i^2}{2} + \frac{\mu^2}{2K} - \frac{1}{2} \left(\Sigma x_i^2 + \frac{1}{K} \right)^{-1} \left(\Sigma x_i y_i + \frac{\mu}{K} \right)^2 \right)$$