Name: ______________________________________

STAT 206B Midterm 2, March 1st, 2022 (Tuesday)

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.

- Normal distribution $\mathrm{N}_p(\boldsymbol{\theta}, \Sigma)$ with $\boldsymbol{\theta} \in \mathbb{R}^p$ and a $(p \times p)$ symmetric positive-definite matrix Σ,
$$f(\boldsymbol{x} \mid \boldsymbol{\theta}, \Sigma) = |\Sigma|^{-1/2}(2\pi)^{-p/2} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\theta})'\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\theta})\right\},$$
where $\mathrm{E}(\boldsymbol{x}) = \boldsymbol{\theta}$ and $\mathrm{Var}(\boldsymbol{x}) = \Sigma$

1. Consider a regression of Y on X. Assume x is fixed (that is, we can treat x as a constant) and y is from $\mathrm{N}(\beta x, \sigma^2)$, where β and σ are unknown.

   (a) (20 pts) Let $\boldsymbol{\theta} = (\beta, \sigma)$. Assume that β and σ are **a priori independent**. Find the corresponding joint noninformative prior $\pi(\boldsymbol{\theta})$ of $\boldsymbol{\theta}$.

   (b) (5 pts) As in part (a), assume that β and σ are independent a priori. Find the joint noninformative prior for β and σ^2. (*Hint:* You may use your answer from part (a))

2. (25 pts) Suppose that X has density
$$f(x \mid \theta) = e^{-(x-\theta)}, \qquad \text{for } \theta \leq x.$$
Consider a prior for θ,
$$\pi(\theta) = \frac{1}{\beta}\exp\left(-\frac{\theta}{\beta}\right), \qquad \text{for } \theta > 0.$$
Assume $\beta = 3$. Suppose it is desired to test $H_0 : \theta \leq 3$ versus $H_1 : \theta > 3$. Also, suppose $x = 4$ is observed. Compute the **(i)** prior and **(ii)** posterior probabilities of the two hypotheses, and **(iii)** the Bayes factor. Briefly interpret what the Bayes factor means for this particular testing in **one sentence**.

3. Consider the usual regression model, $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, $i = 1, \ldots, n$ with $n > 1$, where $\epsilon_i \overset{iid}{\sim} \mathrm{N}(0, \sigma^2)$, a covariate x_i and regression coefficients $\beta_0 \in \mathbb{R}$ and $\beta_1 \in \mathbb{R}$.

   We may write the model in a form of vectors, $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where $\boldsymbol{X}$ is a $n \times 2$ matrix whose ith row is $\boldsymbol{x}_i = (1, x_i)'$, $\boldsymbol{y}$ the $n-$dimensional column vector of $(y_1, \ldots, y_n)$, a coefficient vector $\boldsymbol{\beta} = (\beta_0, \beta_1)'$, and $\boldsymbol{\epsilon}$ the $n-$dimensional column vector of $(\epsilon_1, \ldots, \epsilon_n)$. Since $\epsilon_i \overset{iid}{\sim} \mathrm{N}(0, \sigma^2)$, $\boldsymbol{y} \sim \mathrm{N}_n(\boldsymbol{X}\boldsymbol{\beta}, \sigma^2 \mathrm{I}_n)$, where I_n the identity matrix of size n.

Consider the following prior on $(\boldsymbol{\beta}, \sigma^2)$, $p(\boldsymbol{\beta}, \sigma^2) \propto \sigma^2$.

Fact 1: You may write the likelihood function using $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$.

Fact 2: We have the marginal distribution $m(\boldsymbol{y} \mid \boldsymbol{X}) < \infty$ under the prior $p(\boldsymbol{\beta}, \sigma^2) \propto \sigma^2$ if $n > 1$.

(a) (5 pts) Find the joint posterior distribution of $\boldsymbol{\beta}$ and σ^2, $\pi(\boldsymbol{\beta}, \sigma^2 \mid \boldsymbol{y}, \boldsymbol{X})$, up to a proportionality constant.

(b) (15 pts) Describe a sampling procedure of drawing $(\boldsymbol{\beta}, \sigma^2)$ from their joint posterior distribution. Specify the full conditional distributions of the parameters, $(\boldsymbol{\beta}, \sigma^2)$ with all necessary parameters. Do not write any code, and showing a sampling scheme with brief explanation is enough. But be specific as much as possible.

(END)

1 - (a) $\quad f(y \mid \beta, \sigma) = \frac{1}{\sqrt{2\pi}\,\sigma} \cdot \left(\exp\left(- \frac{(y - x\beta)^2}{2\sigma^2} \right) \right)$

β and σ are a priori independent $\Rightarrow \pi(\beta, \sigma) = \pi_1(\beta)\,\pi_2(\sigma)$

① Jeffreys prior for β

$$\pi(\beta) \propto \left(I(\beta) \right)^{1/2}$$

$$\log f(y \mid \beta, \sigma) = -\frac{1}{2}\log(2\pi) - \log(\sigma) - \frac{(y - \beta x)^2}{2\sigma^2}$$

$$\frac{\partial \log f(y \mid \beta, \sigma)}{\partial \beta} = \frac{x(y - \beta x)}{\sigma^2}$$

$$\frac{\partial^2 \log f(y \mid \beta, \sigma)}{\partial \beta^2} = -\frac{x^2}{\sigma^2}$$

$$I(\beta) = E_\beta\left(- \frac{\partial^2 \log f(y \mid \beta, \sigma)}{\partial \beta^2} \right) = \frac{x^2}{\sigma^2}$$

$$\Rightarrow \pi_1(\beta) \propto \sqrt{\frac{x^2}{\sigma^2}} \propto 1$$

② Jeffreys prior for σ.

$$\frac{\partial \log f(y \mid \beta, \sigma)}{\partial \sigma} = -\frac{1}{\sigma} + \frac{(y - x\beta)^2}{\sigma^3}$$

$$\frac{\partial^2 \log f(y \mid \beta, \sigma)}{\partial \sigma^2} = +\frac{1}{\sigma^2} - \frac{3(y - x\beta)^2}{\sigma^4}$$

$$I(\sigma) = E_\sigma\left(-\frac{1}{\sigma^2} + \frac{3(y - x\beta)^2}{\sigma^4} \right) = -\frac{1}{\sigma^2} + \frac{3\cancel{\sigma^2}}{\sigma^{\cancel{4}2}}$$

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

$$\Rightarrow \boxed{\pi_2(\sigma) \propto \sqrt{\frac{2}{\sigma^2}} \propto \frac{1}{\sigma}}$$

$$\Rightarrow \boxed{\pi(\beta, \sigma) = \pi_1(\beta)\,\pi_2(\sigma) \propto \frac{1}{\sigma}}, \quad \beta \in \mathbb{R} \text{ and } \sigma \in \mathbb{R}^+$$

(b) Jeffreys priors are invariant under reparameterization.

$$z = \sigma^2 \Rightarrow \sigma = \sqrt{z} \Rightarrow d\sigma = \frac{1}{2\sqrt{z}}\,dz$$

$$\tilde{\pi}(\beta, z) \propto \frac{1}{\sqrt{z}} \cdot \frac{1}{2\sqrt{z}} \propto \frac{1}{z}$$

$$\Rightarrow \boxed{\tilde{\pi}(\beta, \sigma^2) \propto \frac{1}{\sigma^2}}, \quad \beta \in \mathbb{R} \;\&\; \sigma^2 \in \mathbb{R}^+$$

2. $$P(H_0) = P(\theta \leq 3) = \int_0^3 \frac{1}{\beta} \exp\left(-\frac{\theta}{\beta}\right) d\theta$$

$$= -\exp\left(-\frac{\theta}{\beta}\right)\Big|_0^3$$

$$= 1 - \exp\left(-\frac{3}{\beta}\right) \quad \beta=3$$

$$= 1 - e^{-1}$$

$$= \boxed{0.632}$$

$$P(H_1) = P(\theta > 3) = 1 - P(H_0) = \exp\left(-\frac{3}{\beta}\right) = \boxed{0.368}$$

$$m(x) = \int_0^x e^{-(x-\theta)} \frac{1}{\beta} e^{-\theta/\beta}\, d\theta$$

$$= \frac{1}{\beta} e^{-x} \int^x e^{-(1/\beta - 1)\theta}\, d\theta$$

$$= \frac{-(1/\beta - 1)^{-1}}{\beta} e^{-x} \quad e^{-(1/\beta - 1)\theta} \Big|_0^x$$

$$= \frac{1}{\beta(1/\beta - 1)} e^{-x} \left(1 - e^{-(1/\beta - 1)x} \right)$$

$$\pi(\theta \mid x) = \frac{e^{-(x-\theta)} \cdot \cancel{1/\beta} \cdot \exp\left(-\frac{\theta}{\beta}\right)}{\frac{1}{\cancel{\beta}(1/\beta - 1)} e^{-x} \left(e^{-(1/\beta - 1)x} - 1 \right)}$$

$$= \frac{(1/\beta - 1)\, e^{-(\frac{1}{\beta} - 1)\theta}}{\left(1 - e^{-(1/\beta - 1)x}\right)} , \qquad 0 < \theta \le x$$

$$P(H_0 \mid x) = P_r(\theta \le 3 \mid x)$$

$$= \frac{1 - e^{-(\frac{1}{\beta} - 1)3}}{1 - e^{-(\frac{1}{\beta} - 1)x}} \qquad = \frac{6}{3} = 2$$

$$\beta = 3 \quad \& \quad x = 4 \qquad = \frac{1 - e^{-\overbrace{(\frac{1}{3} - 1)\,3}}}{1 - e^{-\underbrace{(\frac{1}{3} - 1) \cdot 4}_{= \frac{8}{3}}}} = \underline{0.477}$$

$$P(H_1 \mid x) = 1 - P(H_0 \mid x) = 1 - 0.477 = 0.523 .$$

$$B_{01}^{\pi}(x) = \text{posterior odds} / \text{prior odds} = \frac{0.477 / 0.523}{0.632 / 0.368} = \underline{0.531}$$

$\Rightarrow$ The odds of H_0 against H_1 decreased by a factor of 0.531 after observing data, i.e., data does not support H_0.

3.-(a) $$\pi(\beta, \sigma^2 \mid y, x) \propto \frac{1}{\sigma^2} \cdot \left(\frac{1}{\sigma^2}\right)^{n/2} \cdot \exp\left\{-\frac{1}{2\sigma^2}(y-x\beta)'(y-x\beta)\right\}$$

$$\propto (\sigma^2)^{-n/2-1} \cdot \exp\left\{-\frac{1}{2\sigma^2}S^2 - \frac{1}{2\sigma^2}(\beta-\hat{\beta})'x'x(\beta-\hat{\beta})\right\}$$

(b) We use the Gibbs sampler.

① $$\pi(\beta \mid \sigma^2, x, y) \propto \exp\left\{-\frac{1}{2\sigma^2}(\beta-\hat{\beta})'(x'x)(\beta-\hat{\beta})\right\}$$

$$\Rightarrow \pi(\beta \mid \sigma^2, x, y) = N_2\left(\hat{\beta}, \sigma^2(x'x)^{-1}\right)$$

② $$\pi(\sigma^2 \mid \beta, x, y) \propto (\sigma^2)^{-n/2-1} \exp\left(-\frac{1}{2\sigma^2}\left(S^2 + (\beta-\hat{\beta})'(x'x)(\beta-\beta')\right)\right)$$

$$\Rightarrow \pi(\sigma^2 \mid \beta, X, y) = IG\left(n/2, \ 1/2\left(s^2 + (\beta - \hat{\beta})'(X'X)(\beta - \beta')\right)\right)$$

Step 1 initialize $\beta^{(0)}, (\sigma^2)^{(0)}$

Step2 iterate the following two steps until $t = 1, \dots, T$.

① draw $\beta^{(t)} \sim N_2\left(\hat{\beta}, (\sigma^2)^{(t-1)}(X'X)^{-1}\right)$

② draw $(\sigma^2)^{(t)} \sim IG\left(n/2, \ 1/2\left(s^2 + (\beta^{(t)} - \hat{\beta})'(X'X)(\beta^{(t)} - \beta')\right)\right)$

Step 3 do burn in and thinning

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \qquad \varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$$

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \qquad X = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \qquad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$$

$$\underset{n\times 1}{y} = \underbrace{\underset{n\times 2}{X} \ \underset{2\times 1}{\beta}}_{n\times 1} + \varepsilon \qquad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix} \sim N_n(\mathbf{0}, \sigma^2 I_n) \qquad \mathbf{0} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$p(y \mid \beta, \sigma) = \frac{1}{(\sqrt{2\pi})^n \, |\sigma^2 I_n|^{1/2}} \cdot \exp\left(-\frac{1}{2}(y - X\beta)'(\sigma^2 I_n)^{-1}(y - X\beta)\right)$$

$$p(y \mid \beta, \sigma^2) = \prod_{i=1}^{n} p(y_i \mid \beta, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{1}{2\sigma^2}(y_i - \beta_0 - \beta_1 x_i)^2\right)$$

$$p(\beta_0, \beta_1, \sigma^2 \mid -) \propto \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{1}{2\sigma^2}(y_i - \beta_0 - \beta_1 x_i)^2\right) \times \frac{1}{\sigma^2}$$

$$\propto (\sigma^2)^{-\frac{n}{2}-1} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2\right)$$

① update β_0

$$p(\beta_0 \mid -) \propto \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n} (\beta_0 - (y_i - \beta_1 x_i))^2 \right)$$

$$\propto \exp\left(-\frac{1}{2\sigma^2}\left(n\beta_0^2 - 2\cdot\sum_{i=1}^{n}(y_i - \beta_1 x_i)\cdot \beta_0 \right)\right)$$

$$\beta_0 \mid - \sim N\left(\frac{\sum_{i=1}^{n}(y_i - \beta_1 x_i)}{n}, \frac{\sigma^2}{n} \right)$$

② update β_1

$$p(\beta_1 \mid -) \propto \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n} (\beta_1 x_i - (y_i - \beta_0))^2 \right)$$

$$\propto \exp\left(-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n} x_i^2 \cdot \beta_1^2 - 2\cdot\sum x_i(y_i - \beta_0)\beta_1 \right)\right)$$

$$\Rightarrow \beta_1 \mid - \sim N\left(\frac{\sum_{i=1}^{n} x_i(y_i - \beta_0)}{\sum_{i=1}^{n} x_i^2}, \frac{\sigma^2}{\sum_{i=1}^{n} x_i^2} \right)$$

③ update σ^2

$$\sigma^2 \mid - \sim IG\left(\frac{n}{2}, \frac{1}{2}\sum(y_i - \beta_0 - \beta_1 x_i)^2 \right)$$