STA365 Assignment #1 Ruike Xu.

Problem 1. (25 points)

Suppose that $y \sim \text{Exp}(\theta)$ with density

$$\pi(y) = \theta \exp(-\theta y).$$

Suppose that the prior for θ is $\text{Gamma}(\alpha, \beta)$ with density

$$\pi(\theta) \propto \theta^{\alpha-1} \exp(-\beta\theta).$$

(a) We observe only that $y \geq 10$ without observing the exact value of y.

(i) What is the posterior distribution $\pi(\theta|y \geq 10)$, as a function of α and β?

(ii) What is the posterior mean of θ?

(iii) What is the posterior variance of θ?

(10 points)

(b) We are now told that y is exactly 15.

(i) What is the posterior distribution $\pi(\theta|y = 15)$?

(ii) What is the posterior mean of θ?

(iii) What is the posterior variance of θ?

(10 points)

(c) We assign weakly informative priors to θ in both parts (a) and (b) by setting $\alpha = 0.001$ and $\beta = 0.001$. Which posterior variance of θ is higher, the one from part (a) or the one from part (b)?

(5 points)

(a)

(i) $\pi(\theta|y \geq 10) \propto \pi(\theta) \cdot \pi(y \geq 10 \,|\theta)$

$$\propto \theta^{\alpha-1} \exp(-\beta\theta) \cdot \theta \cdot \exp(-\theta y)$$

$$\propto \theta^{\alpha} \exp(-(\beta + y)\theta)$$

$$\propto \theta^{\alpha+1-1} \exp(-(\beta + y)\theta)$$

This is a recognizable form as a gamma distribution $\text{Gamma}(\alpha+1, \beta+y)$

(ii) $E(\theta|y \geq 10) = \frac{\cancel{\alpha}}{\beta} = \frac{\alpha+1}{\beta+y}$

(iii) $Var(\theta|y \geq 10) = \frac{\alpha}{\beta^2} = \frac{\alpha+1}{(\beta+y)^2} = \frac{\alpha+1}{\beta^2 + 2\beta y + y^2}$

(b) Since y is known as exactly $y=15$, we can directly use it in the likelihood function

(i) $\pi(\theta|y=15) \propto \pi(\theta) \cdot \pi(y=15|\theta)$

$$\propto \theta^{\alpha-1} \exp(-\beta\theta) \cdot \theta \cdot (-15\theta)$$

$$\propto \theta^{\alpha+1-1} \exp(-(\beta+15)\theta)$$

The posterior distribution of $\pi(\theta|y=15)$ is $\text{Gamma}(\alpha+1, \beta+15)$

(ii) The Posterior mean of θ is $E(\theta|Y=15) = \frac{\alpha}{\beta} = \frac{\alpha+1}{\beta+15}$

(iii) The Posterior Variance of θ is $Var(\theta|Y=15) = \frac{\alpha}{\beta^2} = \frac{\alpha+1}{(\beta+15)^2} = \frac{\alpha+1}{\beta^2+30\beta+225}$

(c) Given that $\alpha = 0.001$ and $\beta = 0.001$

In Part (a),

$$Var(\theta|Y\geq 10) = \frac{\alpha+1}{\beta^2+2\beta Y+Y^2} = \frac{1.001}{0.000001+0.002Y+Y^2}$$

In Part (b)

$$Var(\theta|Y=15) = \frac{\alpha+1}{\beta^2+30\beta+225} = \frac{1.001}{0.000001+0.03+225} = \frac{1.001}{225.030001} \approx 0.004448$$

Since $Y \geq 10$ in Part (a), if we want to have $Var(\theta|Y\geq 10) > Var(\theta|Y=15)$

$$\frac{1.001}{(0.001+Y)^2} > 0.004448$$

$$1.001 > 0.004448(0.001+Y)^2$$

$$(0.001+Y)^2 < 225.045$$

$$0.001+Y < 15.001$$

$$Y < 15.001 - 0.001$$

$$Y < 15$$

Thus, if $10 \leq Y < 15$, $Var(\theta|Y\geq 10) > Var(\theta|Y=15)$. otherwise, for $Y > 15$, $Var(\theta|Y\geq 10) > Var(\theta|Y=15)$

Problem 2. (25 points)

Suppose $y \sim$ Galenshore(a, θ), given by

$$\pi(y) = \frac{2}{\Gamma(a)} \theta^{2a} y^{2a-1} \exp(-\theta^2 y^2)$$

for $y > 0, \theta > 0, a > 0$. Assume that a is known. It may be helpful to know that for this density,

$$\mathrm{E}[y] = \frac{\Gamma(a+1/2)}{\theta\Gamma(a)}, \text{ and } \mathrm{E}[y^2] = \frac{a}{\theta^2}.$$

You can also assume that $y_1, ..., y_n, y_{n+1} \sim$ Galenshore(a, θ) for any $n > 1$.

a) Identify the natural conjugate prior for θ.

b) Find $\pi(\theta | y_1, ... y_n)$.

c) Find $\mathrm{E}[\theta | y_1, ... y_n]$.

d) Find $\pi(\tilde{y} | y_1, ... y_n)$.

e) Identify the Jeffreys' prior for θ.

(5 points each)

(a) $\pi(y) = \frac{2}{\Gamma(\alpha)} \theta^{2\alpha} y^{2\alpha-1} \exp(-\theta^2 y^2)$

$\pi(y_1, y_2, \cdots y) \doteq \prod_{i=1}^{n} \pi(y_i | \theta)$

$= \frac{2}{\Gamma(\alpha)} \theta^{2\alpha} y_1^{2\alpha-1} \exp(-\theta^2 y_1^2) \cdots\cdots \frac{2}{\Gamma(\alpha)} \theta^{2\alpha} y_n^{2\alpha-1} \exp(-\theta^2 y_n^2)$

$= \frac{2^n}{\Gamma(\alpha)^n} \theta^{2\alpha n} \exp(-\theta^2 \cdot \sum_{i=1}^{n} y_i^2) \prod_{i=1}^{n} y_i^{2\alpha-1}$

The distribution for the given likelihood $\pi(y)$ and unknown prior distribution $\pi(\theta)$ is

$\pi(\theta | y_1, y_2, \cdots y_n, \alpha) \propto \pi(y_1, y_2, \cdots y_n | \alpha, \theta) \cdot \pi(\theta)$

$\propto \frac{2^n}{\Gamma(\alpha)^n} \theta^{2\alpha n} \exp(-\theta^2 \sum_{i=1}^{n} y_i^2) \prod_{i=1}^{n} y_i^{2\alpha-1} \cdot \pi(\theta)$

$\propto \theta^{2\alpha n} \cdot \exp(-\theta^2 \sum_{i=1}^{n} y_i^2) \cdot \pi(\theta)$

For a natural conjugate prior, we seek a prior that has the same functional form.

So $\pi(\theta) = \frac{2}{\Gamma(a)} b^{2a} \theta^{2a-1} \exp(-b^2 \theta^2)$

Thus, the natural conjugate prior for $\theta \sim$ Galenshore(a, b).

(b) $\pi(y|\theta) = \pi(y_1, y_2, \cdots y_n | \theta) = \frac{2^n}{\Gamma(\alpha)^n} \theta^{2\alpha n} \exp(-\theta^2 \sum_{i=1}^{n} y_i^2) \prod_{i=1}^{n} y_i^{2\alpha-1}$

$\propto \theta^{2\alpha n} \exp(-\theta^2 \sum_{i=1}^{n} y_i^2)$

$\pi(\theta | a, b) = \frac{2}{\Gamma(a)} b^{2a} \theta^{2a-1} \exp(-b^2 \theta^2)$

$\propto \theta^{2a-1} \exp(-b^2 \theta^2)$

$\pi(\theta | y) \propto \pi(y|\theta) \cdot \pi(\theta | \alpha, \beta)$

$\propto \theta^{2\alpha n} \exp(-\theta^2 \sum_{i=1}^{n} y_i^2) \cdot \theta^{2a-1} \exp(-b^2 \theta^2)$

$= \theta^{2\alpha n + 2a - 1} \exp(-\theta^2 (\sum_{i=1}^{n} y_i^2 + b^2))$

$= \theta^{2(\alpha n + a) - 1} \exp(-\theta^2 (\sum_{i=1}^{n} y_i^2 + b^2))$

Therefore, the Posterior distribution follows as the same form of function as Galenshore, which is Galenshore $\left(\alpha n + a, \left(\sum_{i=1}^{n} y_i^2 + b^2\right)\right)$

The Posterior distribution also provide that Galenshore (a, b) is the natural Conjugate prior.

(c) Since $E(Y) = \frac{\Gamma(\alpha+\frac{1}{2})}{\theta\Gamma(\alpha)}$, given that $\pi(\theta | y_1, y_2, \cdots y_n)$ follows Galenshore $\left(\alpha n + a, \sqrt{\sum_{i=1}^{n} y_i^2 + b^2}\right.$

$$E[\theta | y_1, y_2, \cdots y_n] = \frac{\Gamma(\alpha n + a + \frac{1}{2})}{\sqrt{\sum_{i=1}^{n} y_i^2 + b^2} \cdot \Gamma(\alpha n + a)}$$

(d)
$$\pi(\tilde{y} | y_1, y_2, \cdots y_n) = \int \pi(\tilde{y}|\theta) \cdot \pi(\theta | y_1, y_2, \cdots y_n)\, d\theta$$

$$= \int \frac{2}{\Gamma(\alpha)} \theta^{2\alpha} \tilde{y}^{2\alpha-1} \exp(-\theta^2 \tilde{y}^2) \cdot \frac{2}{\Gamma(\alpha n + a)} \theta^{2(\alpha n + a)-1} \cdot \exp\left(-\theta^2\left(\sum_{i=1}^{n} y_i^2 + b^2\right)\right) d\theta$$

$$\propto \int \theta^{2\alpha} \exp(-\theta^2 \tilde{y}^2) \cdot \theta^{2(\alpha n + a)-1} \exp\left(-\theta^2\left(\sum_{i=1}^{n} y_i^2 + b^2\right)\right) d\theta$$

$$\propto \int \theta^{2(\alpha n + \alpha + a)-1} \cdot \exp\left(-\theta^2\left(\sum_{i=1}^{n} y_i^2 + \tilde{y}^2 + b^2\right)\right) d\theta$$

$$\propto -\frac{1}{2} \cdot \theta^{2(\alpha n + \alpha + a)} \left[\left(\sum_{i=1}^{n} y_i^2 + \tilde{y} + b^2\right)\theta\right]^{-\alpha(n+1)-a} \cdot \Gamma\left(\alpha n + \alpha + a, \theta^2\left(\sum_{i=1}^{n} y_i^2 + \tilde{y}^2 + b^2\right)\right)$$

$$\propto -\frac{1}{2} \theta^{2(\alpha n + \alpha + a)} \left[\left(\sum_{i=1}^{n} y_i^2 + \tilde{y} + b^2\right)\theta^2\right]^{-\alpha(n+1)-a} \frac{2}{\Gamma(\alpha n + \alpha + a)} \cdot \theta^2\left(\sum_{i=1}^{n} y_i^2 + \tilde{y}^2 + b^2\right)^{2(\alpha n + \alpha + a)}$$

Thus, we can conclude that the $\pi(\tilde{y} | y_1, y_2, \cdots y_n)$ is Galenshore $\left(\alpha n + \alpha + a + 1, \sqrt{\sum_{i=1}^{n} y_i^2 + \tilde{y}^2 + b^2}\right)$

(e)
$$-E\left[\frac{d^2 \log \pi(y_1, \cdots y_n | \theta)}{d\theta^2}\right] = -E\left[\frac{d^2}{d\theta^2} \log\left(\frac{2^n}{\Gamma(\alpha)^n} \theta^{2\alpha n} \exp\left(-\theta^2 \sum_{i=1}^{n} y_i^2\right) \prod_{i=1}^{n} y_i^{2\alpha-1}\right)\right]$$

$$= -E\left[\frac{d^2}{d\theta^2}\, n\log(2) - n\log(\Gamma(\alpha)) + 2\alpha n \log(\theta) - \theta^2 \sum_{i=1}^{n} y_i^2 + \log\left(\prod_{i=1}^{n} y_i^{2\alpha-1}\right)\right]$$

$$= -E\left[\frac{d}{d\theta} \frac{2\alpha n}{\theta} - 2\theta \sum_{i=1}^{n} y_i^2\right]$$

$$= -E\left[-2\sum_{i=1}^{n} y_i^2 - \frac{2\alpha n}{\theta^2}\right]$$

$$= E\left[\frac{2\alpha n}{\theta^2} + 2\sum_{i=1}^{n} y_i^2\right]$$

$$= \frac{2\alpha n}{\theta^2} + 2E\left[\sum_{i=1}^{n} y_i^2\right]$$

$$= \frac{2\alpha n}{\theta^2} + 2E\left[y_1^2 + y_2^2 + \cdots + y_n^2\right]$$

$$= \frac{2\alpha n}{\theta^2} + 2 \cdot \left(\frac{\alpha}{\theta^2}\right) \cdot n$$

$$= \frac{2\alpha n}{\theta^2} + \frac{2\alpha n}{\theta^2}$$

$$= \frac{4\alpha n}{\theta^2}$$

$$\pi(\theta) \propto \sqrt{I(\theta)} \propto \sqrt{\frac{4\alpha n}{\theta^2}} \propto \frac{1}{\theta^2}$$

Thus the Jeffreys' prior for θ is $\frac{1}{\theta^2}$.

Problem 3. (25 points) normal distribution with 2 unknowns.

(a) Suppose that we are interested in flight times between Toronto and Washington DC. We will assume that the times are normally distributed and wish to make inferences on μ. We know from previous observations that the average flight time is about 1.5 hours and so we set μ_0 to 1.5. Suppose we set τ_0 to 1. Hence, the prior for μ is Normal(μ_0, τ_0^2). $\pi(\theta) \sim N(\mu_0, \tau_0^2)$

Now suppose we observe $y_1 ... y_n$ which can be summarized by $\bar{y} = 1.6$ and $s^2 = 0.01$. Find $\pi(\mu | y_1, ... y_n, \sigma^2 = s^2)$. (5 points)

(b) It is a common problem for measurements to be observed in rounded form. For a simple example, suppose we weigh an object five times and measure weights, rounded to the nearest pound, of 10,10,12,11,9. Assume that unrounded measurements are normally distributed with a noninformative prior distribution on the mean μ and variance σ^2, given by

$$\pi(\mu, \sigma^2) \propto (\sigma^2)^{-1}.$$

(i) Give the posterior distribution for (μ, σ^2) obtained by pretending that the observations are exact, unrounded measurements. (10 points)

(ii) Give the posterior distribution for (μ, σ^2) treating the measurements as rounded. (10 points)

(a) Prior $\mu \sim N(\mu_0 = 1.5,\ \tau_0^2 = 1)$

$Y_1, Y_2, \cdots Y_n \sim N(\mu, \sigma^2 = s^2 = 0.01)$, $\bar{y} = 1.6$

Want to find $\pi(\mu | Y_1, \cdots Y_n, \sigma^2 = s^2)$.

For normal model with unknown mean and known variance.

$$\pi(\mu | Y_1, \cdots Y_n, \sigma^2 = s) \propto \pi(\mu | \sigma^2) \cdot f(y_1, \cdots y_n | \mu, \sigma^2)$$

$$\propto \exp\left[-\frac{1}{2\tau_0^2}(\mu - \mu_0)^2\right] \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^n (Y_i - \mu)^2\right]$$

$$\propto \exp\left[-\frac{1}{2}\left[\frac{1}{\tau_0^2}(\mu^2 - 2\mu\mu_0 + \mu_0^2) + \frac{1}{\sigma^2}\left(\sum_{i=1}^n Y_i^2 - 2\mu \sum_{i=1}^n Y_i + n\mu^2\right)\right]\right]$$

Let's assume that $a = \frac{1}{\tau_0^2} + \frac{n}{\sigma^2}$, $b = \frac{\mu_0}{\tau_0^2} + \frac{\sum_{i=1}^n Y_i}{\sigma^2}$

$$\propto \exp\left[-\frac{1}{2}(a\mu^2 - 2b\mu)\right]$$

$$\propto \exp\left[-\frac{1}{2} a\left(\mu^2 - \frac{2b\mu}{a} + \frac{b^2}{a^2}\right) + \frac{1}{2} b^2 / a\right]$$

$$= \exp\left[-\frac{1}{2}\left(\frac{\mu - b/a}{1/a^{\frac{1}{2}}}\right)^2\right]$$

As the above function has exactly the same shape as a normal distribution, $\pi(\mu | Y_1, \cdots Y_n, \sigma^2)$ is a normal density function.

$$\mu_1 = \frac{\frac{1}{\tau_0^2}\mu_0 + \frac{n}{\sigma^2}\bar{y}}{\frac{1}{\tau_0^2} + \frac{n}{\sigma^2}} = \frac{1.5 + \frac{n}{0.01} \cdot 1.6}{1 + \frac{n}{0.01}} = \frac{160n + 1.5}{1 + 100n}$$

$$\tau_1^2 = \frac{1}{\frac{1}{\tau_0^2} + \frac{n}{\sigma^2}} = \frac{1}{1 + \frac{n}{0.01}} = \frac{1}{1 + 100n}$$

Therefore, We Can Conclude that $\pi(\mu|y_1, \cdots y_n, \sigma^2 = s^2)$ follows Normal $\left(\frac{160n+1.5}{1+100n}, \frac{1}{1+100n}\right)$

$$\pi(\mu|y_1, \cdots\cdots y_n, \sigma^2 = s^2) = \left(2\pi\left(\frac{160n+1.5}{1+100n}\right)\right)^{-\frac{1}{2}} \exp\left[-\frac{1}{2\frac{1}{(1+100n)}}\left(\mu - \frac{160n+1.5}{1+100n}\right)^2\right]$$

(b)(i) Uninformative Prior : $\pi(\mu, \sigma^2) \propto (\sigma^2)^{-1}$

likelihood information: $y_1, y_2, \cdots\cdots y_5 = 10, 10, 12, 11, 9.$

$$\pi(\mu, \sigma^2|y_1, \cdots y_5) \propto \pi(\mu, \sigma^2) \cdot \pi(y_1, \cdots y_5|\mu, \sigma^2)$$

$$\propto (\sigma^2)^{-1} \cdot \left[(2\pi\sigma^2)^{-\frac{1}{2}} \exp\left[-\frac{1}{2\sigma^2}(y_1-\mu)^2\right] \cdots\cdots (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left[-\frac{1}{2\sigma^2}(y_5-\mu)^2\right]\right]$$

$$\propto (\sigma^2)^{-1} (2\pi)^{-\frac{5}{2}} \cdot (\sigma^2)^{-\frac{5}{2}} \cdot \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{5}(y_i-\mu)^2\right]$$

$$\propto \sigma^{-2}\sigma^{-5} \cdot \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{5}(y_i-\mu)^2\right]$$

$$\propto \sigma^{-7} \exp\left[-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{5}(y_i-\bar{y}+\bar{y}-\mu)^2\right]\right]$$

$$\propto \sigma^{-7} \exp\left[-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{5}(y_i-\bar{y})^2 + \sum_{i=1}^{5}(\bar{y}-\mu)^2 + 2\sum_{i=1}^{5}(y_i-\bar{y})(\bar{y}-\mu)\right]\right]$$

$$\propto \sigma^{-7} \exp\left[-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{5}(y_i-\bar{y})^2 + 5(\bar{y}-\mu)^2 + 2(\bar{y}-\mu)\cancel{\sum_{i=1}^{5}(y_i-\bar{y})}\right]\right]$$

$$\propto \sigma^{-7} \exp\left[-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{5}(y_i-\bar{y})^2 + 5(\bar{y}-\mu)^2\right]\right]$$

Since $s^2 = \frac{1}{5-1}\sum_{i=1}^{5}(y_i-\bar{y})^2$ is an unbiased estimator,

$$\pi(\mu, \sigma^2|y_1, \cdots y_5) \propto \sigma^{-7} \exp\left[-\frac{1}{2\sigma^2}\ 4s^2 + 5(\bar{y}-\mu)^2\right]$$

By using the given sample information, we would be able to derive s^2 and $\bar{y}$

$$\bar{y} = \frac{10+10+12+11+9}{5} = 10.4, \quad s^2 = \frac{1}{5-1}\sum_{i=1}^{5}(y_i-\bar{y})^2$$

$$= \frac{1}{4}(0.4^2 + 0.4^2 + 1.6^2 + 0.6^2 + 1.4^2)$$

$$= \frac{1}{4} \times \frac{26}{5}$$

$$= 1.3$$

$$\pi(\mu, \sigma^2|y_1, \cdots y_5) \propto \sigma^{-7} \exp\left[-\frac{1}{2\sigma^2}\left[4 \times 1.3 + 5(10.4-\mu)^2\right]\right]$$

$$\propto \sigma^{-7} \exp\left[-\frac{1}{2\sigma^2}\left[5.2 + 5(10.4-\mu)^2\right]\right]$$

$$\propto (\sigma^2)^{-\frac{7}{2}} \exp\left[-\frac{1}{2\sigma^2}\left[5.2 + 5(10.4-\mu)^2\right]\right]$$

This Part is cited from Chapter 5. bayesian Statistic (II).

Therefore, we can conclude that the Posterior distribution for (μ, σ^2) by Pretending the observations are exact, unrounded measurements is $\pi(\mu, \sigma^2|y_1, \cdots y_5) \propto (\sigma^2)^{-\frac{7}{2}} \exp\left[-\frac{1}{2\sigma^2}\left[5.2 + 5(10.4-\mu)^2\right]\right]$

(C)(vi) Let's assume that W_i is unrounded, exact measurement of weight to Y_i for $i = 1, 2, \cdots\cdots 5$.

Given that $W_i \sim N(\mu, \sigma^2)$ and $W_i = (Y_i - 0.5,\ Y_i + 0.5)$ for $i = 1, 2, \cdots\cdots 5$.

Given the uninformative Prior $\pi(\mu, \sigma^2) \propto (\sigma^2)^{-1}$

$$\pi(\mu, \sigma^2 \mid w_1, w_2, \cdots\cdots w_5) \propto \pi(\mu, \sigma^2) \cdot \pi(w_1, w_2, \cdots\cdots w_5 \mid \mu, \sigma^2)$$
$$\propto \pi(\mu, \sigma^2) \cdot \prod_{i=1}^{5} \pi(Y_i - 0.5 < y_i < Y_i + 0.5 \mid \mu, \sigma^2)$$
$$\propto \pi(\mu, \sigma^2) \cdot \prod_{i=1}^{5} \int_{Y_i - 0.5}^{Y_i + 0.5} N(\mu, \sigma^2) dy_i$$

Given that $w_i \sim N(\mu, \sigma^2)$, here we define $Z = \frac{X - \mu}{\sigma}$ is a standard random variable such that $Z \sim N(0, 1)$.

$$\pi(\mu, \sigma^2 \mid w_1, w_2, \cdots\cdots w_5) \propto (\sigma^2)^{-1} \cdot \prod_{i=1}^{5} F_Y(Y_i)$$
$$\propto (\sigma^2)^{-1} \cdot \prod_{i=1}^{5} [F_Y(Y_i + 0.5) - F_Y(Y_i - 0.5)]$$
$$\propto (\sigma^2)^{-1} \cdot \prod_{i=1}^{5} \left[\Phi\left(\frac{Y_i + 0.5 - \mu}{\sigma}\right) - \Phi\left(\frac{Y_i - 0.5 - \mu}{\sigma}\right)\right]$$
$$\propto (\sigma^2)^{-1} \left[\Phi\left(\frac{10.5 - \mu}{\sigma}\right) - \Phi\left(\frac{9.5 - \mu}{\sigma}\right)\right] \left[\Phi\left(\frac{10.5 - \mu}{\sigma}\right) - \Phi\left(\frac{9.5 - \mu}{\sigma}\right)\right]$$
$$\left[\Phi\left(\frac{12.5 - \mu}{\sigma}\right) - \Phi\left(\frac{11.5 - \mu}{\sigma}\right)\right] \left[\Phi\left(\frac{11.5 - \mu}{\sigma}\right) - \Phi\left(\frac{10.5 - \mu}{\sigma}\right)\right]$$
$$\left[\Phi\left(\frac{9.5 - \mu}{\sigma}\right) - \Phi\left(\frac{8.5 - \mu}{\sigma}\right)\right]$$

This Part is cited from Normal distribution: Gaussian: Normal random Variables. PDF

Problem 4. (25 points)

(a) Find Jeffreys' priors for the unknown parameters in the following models:

(i) $Y \sim \mathrm{N}(\mu, \sigma^2)$ with μ known and σ^{-2} (precision) as the parameter that we need to specify the prior for. (5 points)

(ii) $Y \sim \mathrm{N}(\mu, \sigma^2)$ and we need a prior for (μ, σ^{-2}). (10 points)

(b) Let

$$L(\theta, a) = \begin{cases} k_0(\theta - a), & \text{if } \theta \geq a, \\ k_1(a - \theta), & \text{if } \theta < a. \end{cases}$$

Find the Bayes estimator under this loss function. Recall that the Bayes estimator a is the quantity that minimizes the posterior expected loss given by

$$\mathrm{E}[L(\theta, a)|y] = \int_\Omega L(\theta, a)\pi(\theta, y)d\theta.$$

(10 points)

(a)(i) $Y \sim N(\mu, \sigma^2)$

$$-E\left[\frac{d^2 \log \pi(Y|\sigma^{-2})}{d(\sigma^{-2})^2}\right] = -E\left[\frac{d^2}{d(\sigma^{-2})^2}\left(-\frac{\log(\frac{2\pi}{\sigma^{-2}})}{2} - \frac{\sigma^{-2}}{2}(y-\mu)^2\right)\right]$$

$$= -E\left[\frac{d}{d\sigma^{-2}}\left(\frac{1}{2\sigma^{-2}} - \frac{1}{2}(y-\mu)^2\right)\right]$$

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

$$= \frac{1}{2}(\sigma^{-2})^{-2}$$

$$\pi(\sigma^{-2}) \propto \sqrt{I(\sigma^{-2})} \propto \sqrt{\frac{1}{2}(\sigma^{-2})^{-2}}$$

$$\propto (\sigma^{-2})^{-1}$$

(ii) Let $Y \sim N(\mu, \sigma^2)$,

Since the Jeffreys' Prior only takes the part that is proportional to the Prior Parameters,

we would first take the constant out of the log likelihood function.

$$\psi(Y, \mu, \sigma^{-2}) = -\log\left(\frac{1}{\sigma^{-2}}\right) - \frac{\sigma^{-2}}{2}(Y-\mu)^2.$$

This idea is partially cited from Part II bayesian Statistics.

$$I(\mu, \sigma^{-2}) = -E\begin{pmatrix} \frac{d^2}{d\mu^2}\psi(Y, \mu, \sigma^{-2}) & \frac{d^2}{d\mu\, d\sigma^{-2}}\psi(Y, \mu, \sigma^{-2}) \\ \frac{d^2}{d\sigma^{-2}d\mu}\psi(Y, \mu, \sigma^{-2}) & \frac{d^2}{d(\sigma^{-2})^2}\psi(Y, \mu, \sigma^{-2}) \end{pmatrix}$$

$$= -E\begin{pmatrix} \frac{d}{d\mu}(Y-\mu)\cdot\sigma^{-2} & \frac{d}{d\mu}(Y-\mu)\sigma^{-2} \\ \frac{d}{d\mu}\frac{1}{\sigma^{-2}} - \frac{1}{2}(Y-\mu)^2 & \frac{d}{d\sigma}\frac{1}{\sigma^{-2}} - \frac{1}{2}(Y-\mu)^2 \end{pmatrix}$$

$$= -E\begin{pmatrix} -\sigma^{-2} & Y-\mu \\ Y-\mu & -\frac{1}{(\sigma^{-2})^2} \end{pmatrix}$$

$$= \begin{pmatrix} \sigma^{-2} & 0 \\ 0 & \frac{1}{(\sigma^{-2})^2} \end{pmatrix}$$

Therefore, $\pi(\mu, \sigma^{-2}) \propto \sqrt{I(\mu, \sigma^{-2})} = \sqrt{\sigma^{-2} \times \frac{1}{(\sigma^{-2})^2}} = \sqrt{\frac{1}{\sigma^{-2}}}$

$\propto (\sigma^{-2})^{-2}$

(b) $L(\theta, \alpha) = \begin{cases} k_0(\theta - \alpha), & \text{if } \theta \geq \alpha. \\ k_1(\alpha - \theta), & \text{if } \theta < \alpha \end{cases}$

$$E[L(\theta,\alpha)|Y] = \int_\Omega L(\theta,\alpha)\pi(\theta,Y)d\theta$$
$$= \int_\alpha^\infty L(\theta,\alpha)\pi(\theta,Y)\,d\theta + \int_{-\infty}^\alpha L(\theta,\alpha)\cdot\pi(\theta,Y)\,d\theta$$
$$= \int_\alpha^\infty k_0(\theta-\alpha)\cdot\pi(\theta,Y)\,d\theta + \int_{-\infty}^\alpha k_1(\alpha-\theta)\pi(\theta,Y)\,d\theta.$$
$$= \int_\alpha^\infty k_0(\theta-\alpha)\cdot\pi(\theta,Y)\,d\theta - \int_{-\infty}^\alpha k_1(\theta-\alpha)\,\pi(\theta,Y)\,d\theta.$$

To minimize the loss,

$$\int_\alpha^\infty k_0\theta\,\pi(\theta,Y)\,d\theta - k_0\alpha\,\pi(\theta,Y)d\theta = \int_{-\infty}^\alpha k_1\theta\,\pi(\theta,Y)\,d\theta - k_1\theta\pi(\theta,Y)\,d\theta$$
$$\int_\alpha^\infty k_0\theta\pi(\theta,Y)\,d\theta - \int_\alpha^\infty k_0\alpha\pi(\theta,Y)d\theta = \int_{-\infty}^\alpha k_1\theta\pi(\theta,Y)\,d\theta - \int_{-\infty}^\alpha k_1\theta\,\pi(\theta,Y)\,d\theta$$
$$k_0\int_\alpha^\infty \theta\pi(\theta,Y)d\theta - k_1\int_{-\infty}^\alpha \theta\pi(\theta,Y)d\theta = k_0\alpha\int_\alpha^\infty \pi(\theta,Y)\,d\theta - k_1\alpha\int_{-\infty}^\alpha \pi(\theta,Y)d\theta$$

Bibliography

- Chapter 5. bayesian statistics (II) - brown university. Retrieved March 5, 2022. from https://www.dam.brown.edu/people/huiwang/classes/am166/Ch5_Part2.pdf.

- Normal distribution: Gaussian: Normal random variables: PDF. Normal distribution | Gaussian | Normal random variable | PDF. Retrieved March 6, 2022. from https://www.probabilitycourse.com/

- Part II bayesian statistics - University of Oxford. Retrieved March 7, 2022. from http://www.stats.ox.ac.uk/