Homework1

Chen Xupeng

2

2.5

a

$$\begin{aligned}
\Pr(y=k) &= \int_0^1 \Pr(y=k|\theta)d\theta \\
&= \int_0^1 \binom{n}{k} \theta^k(1-\theta)^{n-k}d\theta \\
&= \binom{n}{k} \frac{\Gamma(k+1)\Gamma(n-k+1)}{\Gamma(n+2)} \\
&= \frac{1}{n+1}
\end{aligned}$$

b

from the geometriy arribute, we know a point between can be written as $\frac{\alpha+y}{\alpha+\beta+n} = \lambda\frac{\alpha}{\alpha+\beta} + (1-\lambda)\frac{y}{n}$ and $\lambda \in [0,1]$

We can also write $ $ as $ + (-)$

c

$\text{Uniform prior distribution: } \alpha = \beta = 1. \text{ Prior variance is } \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} = \frac{1}{12}$

$$\begin{aligned}
\text{Posterior variance } &= \frac{(1+y)(1+n-y)}{(2+n)^2(3+n)} \\
&= \left(\frac{1+y}{2+n}\right)\left(\frac{1+n-y}{2+n}\right)\left(\frac{1}{3+n}\right)
\end{aligned}$$

The first two factors in are two numbers that sum to 1, so their product is at most $\frac{1}{4}$. And, since $n \geq 1$, the third factor is less than $\frac{1}{4}$. So the product of all three factors is less than $\frac{1}{16}$.

d

n = n and y = 1. α = 1, β = 3, then prior variance is $\frac{3}{80}$, and posterior variance is $\frac{1}{20}$.

2.7

a

$q(\theta) = p\left(\frac{e^\phi}{1+e^\phi}\right)\left|\frac{d}{d\theta}\log\left(\frac{\theta}{1-\theta}\right)\right| \propto \theta^{-1}(1-\theta)^{-1}$

b.

If $y=0$, $p(\theta|y) \propto \theta^{-1}(1-\theta)^{n-1}$ has infinite integral over any interval near $\theta = 0$. When $y = n$ similar result happens at $\theta = 1$.

2.8

a

$$\theta \Big| y \sim \mathrm{N}\left(\frac{\frac{1}{40^2}180 + \frac{n}{20^2}150}{\frac{1}{40^2} + \frac{n}{20^2}}, \frac{1}{\frac{1}{40^2} + \frac{n}{20^2}}\right)$$

b

$$\tilde{y} \Big| y \sim \mathrm{N}\left(\frac{\frac{1}{40^2}180 + \frac{n}{20^2}150}{\frac{1}{40^2} + \frac{n}{20^2}}, \frac{1}{\frac{1}{40^2} + \frac{n}{20^2}} + 20^2\right)$$

c

95% posterior interval for $\theta|\overline{y} = 150, n = 10 :\quad 150.7 \pm 1.96(6.25) = [138, 163]$

95% posterior interval for $\tilde{y}|\overline{y} = 150, n = 10 :\quad 150.7 \pm 1.96(20.95) = [110, 192]$

d

95% posterior interval for $\theta|\overline{y} = 150, n = 100 :\quad [146, 154]$

95% posterior interval for $\tilde{y}|\overline{y} = 150, n = 100 :\quad [111, 189]$

2.19

a

We have $p(y|\theta) = \theta \cdot e^{-\theta y} I_{(0,\infty)}(y)$, let the prior be Gamma distribution: $P(\theta) = \frac{\beta^\alpha}{P(\alpha)}\theta^{\alpha-1}e^{-\beta\theta}I_{(0,\infty)} \sim Gamma(\alpha, \beta)$ So, $P(\theta|y) \propto p(y/\theta)p(\theta) \propto \theta^\alpha e^{-(\beta+y)\theta}I_{(0,\infty)}(y, \theta)$ So, $p(y|\theta) \sim Gamma(\alpha + 1, \beta + y)$ is conjugate prior distribution

b

$$P(\phi) = p(\theta)\left|\frac{d\theta}{d\phi}\right|, \quad \theta = \frac{1}{\phi}$$

$$P(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta}I_{[0,\infty)}(\theta), \frac{d\theta}{d\phi} = -\frac{1}{\phi^2}$$

so we have

$$P(\phi) = \frac{\beta^\alpha}{\Gamma(\alpha)}\phi^{-(\alpha-1)}e^{-\beta/\phi}\frac{1}{\phi^2}I_{[0,\infty)}(\phi)$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)}\phi^{-(\alpha+1)} \cdot e^{-\frac{\beta}{\phi}}I_{[\cos j}(\phi) \sim Inv - Gamma(\alpha, \beta)$$

c

$CV = \alpha^{-\frac{1}{2}} = 0.5 \Rightarrow \alpha = 4, p(\theta|y) \sim G_{ama}(\alpha + n, \beta + n\overline{y})$, from $CV = (\alpha + n)^{-\frac{1}{2}} = 0.1 \Rightarrow n = 96$

d

$P(\phi) \sim Inv - Gamma(\alpha, \beta), CV = (\alpha - 2)^{-\frac{1}{2}} = 0.5 \Rightarrow \alpha = 6$ And,

$$P(\phi|y) \propto p(\phi)p(y|\phi)$$
$$\propto \phi^{-n}e^{-\frac{n\bar{y}}{\phi}}I_{(0,\infty)}(y)\phi^{-(\alpha+1)}e^{-\frac{\beta}{\phi}}I_{(0,\infty)}(\phi)$$
$$\sim Inv - Gamma(\alpha + n, \beta + n\overline{y})$$

Now, $CV = (\alpha + n - 2)^(- 0.5) = 0.1 \Rightarrow n = 96$

3

3.1

a

Label the prior distribution $p(\theta)$ as Dirichlet $(a_1, \ldots, a_n)$. The posterior distribution is $p(\theta|y) = \text{Dirichlet}(y_1 + a_1, \ldots, y_n + a_n)$., the marginal posterior distribution of $(\theta_1, \theta_2, 1 - \theta_1 - \theta_2)$ is also Dirichlet:

$$p\left(\theta_1, \theta_2 | y\right) \propto \theta_1^{y_1 + a_1 - 1} \theta_2^{y_2 + a_2 - 1} (1 - \theta_1 - \theta_2)^{y_{\text{rest}} + a_{\text{rest}} - 1}$$

where $y_{\text{rest}} = y_3 + \ldots + y_J, a_{\text{rest}} = a_3 + \ldots + a_J$

Change variables to $(\alpha, \beta) = \left(\frac{\theta_1}{\theta_1 + \theta_2}, \theta_1 + \theta_2\right)$. The Jacobian of this transformation is $|1/\beta|$, so the transformed density is

$$\begin{aligned} p(\alpha, \beta | y) &\propto \beta (\alpha\beta)^{y_1 + a_1 - 1} ((1 - \alpha)\beta)^{y_2 + a_2 - 1} (1 - \beta)^{y_{\text{rest}} + a_{\text{rest}} - 1} \\ &= \alpha^{y_1 + a_1 - 1} (1 - \alpha)^{y_2 + a_2 - 1} \beta^{y_1 + y_2 + a_1 + a_2 - 1} (1 - \beta)^{y_{\text{rest}} + a_{\text{rest}} - 1} \\ &\propto \text{Beta}(\alpha | y_1 + a_1, y_2 + a_2) \, \text{Beta}(\beta | y_1 + y_2 + a_1 + a_2, y_{\text{rest}} + a_{\text{rest}}) \end{aligned}$$

since the posterior density divides into separate factors for α and β, they are independent, and, and shown above, $\alpha \, | y \sim \text{Beta}(y_1 + a_1, y_2 + a_2)$

b

The Beta $(y_1 + a_1, y_2 + a_2)$ posterior distribution can also be derived from a Beta (a_1, a_2) prior distribution and a binomial observation y_1 with sample size $y_1 + y_2$.

3.9

$$\begin{aligned} p\left(\mu, \sigma^2 | y\right) &\propto p(y | \mu, \sigma^2) p\left(\mu, \sigma^2\right) \\ &\propto \left(\sigma^2\right)^{-n/2} \exp\left(-\frac{(n-1)s^2 + n(\mu - \bar{y})^2}{2\sigma^2}\right) \sigma^{-1} \left(\sigma^2\right)^{-(\nu_0/2+1)} \exp\left(-\frac{\nu_0\sigma_0^2 + \kappa_0(\mu - \mu_0)^2}{2\sigma^2}\right) \\ &\propto \sigma^{-1} \left(\sigma^2\right)^{-((\nu_0+n)/2+1)} \exp\left(-\frac{\nu_0\sigma_0^2 + (n-1)s^2 + \frac{n\kappa_0(\bar{y}-\mu_0)^2}{n+\kappa_0} + (n+\kappa_0)\left(\mu - \frac{\mu_0\kappa_0 + n\bar{y}}{n+\kappa_0}\right)^2}{2\sigma^2}\right) \end{aligned}$$

$$\mu, \sigma^2 \Big| y \sim \text{N} - \text{Inv\_}\chi^2 \left(\frac{\mu_0\kappa_0 + n\bar{y}}{n+\kappa_0}, \frac{\sigma_n^2}{n+\kappa_0}; n + \nu_0, \sigma_n^2\right)$$

$$\sigma_n^2 = \frac{\nu_0\sigma_0^2 + (n-1)s^2 + \frac{n\kappa_0(\bar{y}-\mu_0)^2}{n+\kappa_0}}{n+\nu_0}$$

and prove

$\mu | \sigma^2, y \sim N(\bar{y}, \sigma^2/n)$

$P(u | \sigma^2, y) = P\left(u, \sigma^2 | y\right) / P\left(\sigma^2 | y\right)$, for non-informative, $p\left(x, \sigma^2 | y\right) \propto p(y | y, \sigma^2) p\left(u, \sigma^2\right)$

$$\begin{aligned} p\left(x, \sigma^2 | y\right) &\propto \sigma^{-(n+2)} \exp\left\{-\frac{1}{\sigma^2} \sum_{i=1}^{n} (y_i - u)^2\right\} \\ &= \sigma^{-(n+2)} \exp\left\{-\frac{1}{2\sigma^2}\left[(n-1)s^2 + n(\bar{y} - \mu)^2\right]\right\} \end{aligned}$$

and we have $p(\sigma^2|y) = \int p(\mu, \sigma^2|y)d\mu \propto \sigma^{-(n+2)} \exp -\frac{1}{2\sigma^2}(n-1)s^2\sqrt{2\pi\sigma^2/n}$

Addition

Prove that Jeffreys' prior for $\mu|\sigma^2, y \sim N(\bar{y}, \sigma^2/n)$ is \$p (,) (^ { 2 }) ^ { - 1 } \$ without any (independence) assumptions. This result was presented on the blackboard today. Please derive it.

$$\theta = [\mu, \sigma]$$
$$P(\theta) \propto [J(\theta)]^{\frac{1}{2}}$$

and we have

$$\log p(y/\theta) = \log\left[\frac{1}{\left(2\pi\sigma^2\right)^{\frac{n}{2}}}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-u)^2\right)\right]$$
$$= C - n\log\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-u)^2$$
$$\Rightarrow \frac{\partial \log P(y|\theta)}{\partial \mu} = \frac{\sum_{i=1}^{n}(y_i-u)}{\sigma^2} \Rightarrow \frac{\partial^2 \log p(y|\theta)}{\partial u^2} = -\frac{n}{\sigma^2}$$

$\Rightarrow \frac{\partial \log P(y|\theta)}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(y_i-u)^2 \ \frac{\partial^2 \log p(y|\theta)}{\partial \sigma^{2.}} = \frac{\theta n}{\sigma^2} - \frac{3}{\sigma^4}\sum_{i=1}^{n}(y_i-u)^2 \Rightarrow \frac{\partial^2 \log p(y|\theta)}{\partial u \partial \sigma} = -\frac{2\sum_{i=1}^{n}(y_i-u)}{\sigma^3}$

So $J(\theta)$

$$= \begin{bmatrix} -E\left(\frac{\partial^2 \log p(y|\theta)}{\partial \mu^2}\right) & -E\left(\frac{\partial^2 \log P(y|\theta)}{\partial u d\theta}\right) \\ -E\left(\frac{\partial^2 \log P(y|\theta)}{\partial u d\theta}\right) & -E\left(\frac{\partial^2 \log p(y|\theta)}{\partial \theta^2}\right) \end{bmatrix}$$
$$= \begin{bmatrix} \frac{n}{\sigma^2} & 0 \\ 0 & \frac{2n}{\sigma^2} \end{bmatrix}$$

So

$$\det[J(\theta)^{\frac{1}{2}}] = \begin{bmatrix} \frac{\sqrt{n}}{\sigma} & 0 \\ 0 & \frac{\sqrt{2}n}{\sigma} \end{bmatrix}$$
$$= \sqrt{2}\eta\sigma^{-2}\sigma \cdot \sigma^{-2}$$

$\Rightarrow p(\mu, \sigma) \propto \left(\sigma^2\right)^{-1}$