Solutions for the exam in Statistical Reasoning 2020/2021

SOLUTION EXERCISE 1(a):
Compute the joint density:

$$p(y_1,\dots,y_n|\theta) = \prod_{i=1}^{n} p(y_i|\theta) = \prod_{i=1}^{n} \theta\cdot(1-\theta)^{y_i} = \theta^n\cdot(1-\theta)^{\sum_{i=1}^n y_i}$$

For the posterior density we have:

$$\begin{aligned} p(\theta|y_1,\dots,y_n) &\propto p(y_1,\dots,y_n|\theta)\cdot p(\theta) \\ &\propto \theta^{a+n}\cdot(1-\theta)^{b+\sum_{i=1}^n y_i} \end{aligned}$$

The posterior density is proportional to a Beta density with parameters: $\tilde{a} = a+n$ and $\tilde{b} = b+\sum_{i=1}^n y_i$. Thus,

$$\theta|(Y_1=y_1,\dots,Y_n=y_n) \sim \text{BETA}(a+n, b+\sum_{i=1}^n y_i)$$

Interpretation of hyperparaneters in terms of 'pseudo-counts':
We have a more observations, whose sum is equal to b.

SOLUTION EXERCISE 1(b):
Compute the density of the marginal probability

$$\begin{aligned} p(y_1,\dots,y_n) &= \int_0^1 p(y_1,\dots,y_n|\theta)\cdot p(\theta)\, d\theta \\ &= \int_0^1 \theta^n\cdot(1-\theta)^{\sum_{i=1}^n y_i}\cdot\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\cdot\theta^{a-1}\cdot(1-\theta)^{b-1}\, d\theta \\ &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\cdot A^{-1}\cdot\int_0^1 A\cdot\theta^{a+n-1}\cdot(1-\theta)^{b+\sum_{i=1}^n y_i}\, d\theta \end{aligned}$$

where

$$A := \frac{\Gamma(a+n+b+\sum_{i=1}^n y_i)}{\Gamma(a+n)\cdot\Gamma(b+\sum_{i=1}^n y_i)}$$

The integral is over a Beta density with parameters $a+n$ and $b+\sum_{i=1}^n y_i$. it follows:

$$p(y_1,\dots,y_n) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\cdot\frac{\Gamma(a+n)\cdot\Gamma(b+\sum_{i=1}^n y_i)}{\Gamma(a+n+b+\sum_{i=1}^n y_i)}$$

SOLUTION EXERCISE 2:

(a) We have the relationship: $\text{EXP}(\lambda) = \text{GAM}(1,\lambda)$.

(b) For the likelihood we have:

$$p(y_1,\dots,y_n|\theta) = \prod_{i=1}^{n} p(y_i|\theta) = \prod_{i=1}^{n}\frac{\theta^{y_i}\cdot e^{-\theta}}{y_i!} = \frac{\theta^{\sum_{i=1}^n y_i}\cdot e^{-n\theta}}{\prod_{i=1}^n y_i!}$$

For the power posterior density we have:

$$\begin{aligned} p_\tau(\theta|y_1,\ldots,y_n) &\propto p(y_1,\ldots,y_n|\theta)^\tau \cdot p(\theta) \\ &\propto \left(\frac{\theta^{\sum_{i=1}^n y_i} \cdot e^{-n\theta}}{\prod_{i=1}^n y_i!}\right)^\tau \cdot \lambda \cdot e^{-\lambda\theta} \\ &\propto \theta^{\tau\sum_{i=1}^n y_i} \cdot e^{-n\tau\theta} \cdot e^{-\lambda\theta} \\ &\propto \theta^{1+\tau\sum_{i=1}^n y_i - 1} \cdot e^{-(\lambda+n\tau)\theta} \end{aligned}$$

in which we recognize the shape of the density of a Gamma distribution. We get:

$$\theta|(\tau, Y_1 = y_1, \ldots, Y_n = y_n) \sim \text{GAM}(1 + \tau\sum_{i=1}^n y_i, \lambda + n\tau)$$

SOLUTION EXERCISE 3:

(a) For the density of the $\mathcal{N}(\mu,\sigma^2)$, we have (see Lecture Notes 1):

$$p(x) \propto e^{-\frac{1}{2}\cdot\frac{1}{\sigma^2}\cdot x^2 + \frac{\mu}{\sigma^2}\cdot x}$$

From

$$p(x) \propto \exp\{-2x^2 - 4x\} = \exp\{-\frac{1}{2}\cdot 4\cdot x^2 - 4x\}$$

we read that $\sigma^2 = \frac{1}{4}$ and $\mu = \frac{1}{4}\cdot -4 = -1$ so that $X \sim \mathcal{N}(-1, \frac{1}{4})$.

(b) From

$$p(x) \propto 2\cdot\exp\{-\frac{1}{2}x - 2\} \propto x^{1-1}\cdot\exp\{-\frac{1}{2}\cdot x\}$$

it follows that X must have the following Gamma distribution: $X \sim \text{GAM}(1, \frac{1}{2})$.

(c) We have:

$$\begin{aligned} p(\lambda^{-1}) &\propto (2\pi)^{-n/2}\cdot\det(\lambda\cdot\mathbf{I})^{-1/2}\cdot\exp\{-\frac{1}{2}\cdot(\mathbf{x}-\boldsymbol{\mu})^T(\lambda\cdot\mathbf{I})^{-1}(\mathbf{x}-\boldsymbol{\mu})\} \\ &\propto (2\pi)^{-n/2}\cdot\lambda^{-n/2}\cdot\det(\mathbf{I})^{-1/2}\cdot\exp\{-\frac{1}{2}\cdot(\mathbf{x}-\boldsymbol{\mu})^T\cdot\lambda^{-1}\cdot\mathbf{I}^{-1}(\mathbf{x}-\boldsymbol{\mu})\} \\ &\propto \lambda^{-n/2}\cdot\exp\{-\lambda^{-1}\left(\frac{1}{2}\cdot(\mathbf{x}-\boldsymbol{\mu})^T(\mathbf{x}-\boldsymbol{\mu})\right)\} \\ &\propto (\lambda^{-1})^{n/2}\cdot\exp\{-\lambda^{-1}\left(\frac{1}{2}\cdot(\mathbf{x}-\boldsymbol{\mu})^T(\mathbf{x}-\boldsymbol{\mu})\right)\} \end{aligned}$$

This is the shape of the density of a Gamma distribution with parameters: $a = \frac{n}{2}+1$ and $b = \frac{1}{2}\cdot(\mathbf{x}-\boldsymbol{\mu})^T(\mathbf{x}-\boldsymbol{\mu})$. Therefore we have

$$\lambda^{-1} \sim \text{GAM}\left(\frac{n}{2}+1, \frac{1}{2}\cdot(\mathbf{x}-\boldsymbol{\mu})^T(\mathbf{x}-\boldsymbol{\mu})\right) \Leftrightarrow \lambda \sim \text{INV-GAM}\left(\frac{n}{2}+1, \frac{1}{2}\cdot(\mathbf{x}-\boldsymbol{\mu})^T(\mathbf{x}-\boldsymbol{\mu})\right)$$

(d) We have:

$$p(\mathbf{x}) \propto \exp\{-2x_1^2 - 4x_2^2 - 4x_1x_2\} = \exp\{-\frac{1}{2}(4x_1^2 + 8x_2^2 + 8x_1x_2)\}$$

And we can write:

$$4x_1^2 + 8x_2^2 + 8x_1x_2 = (x_1, x_2)\begin{pmatrix} 4 & 4 \\ 4 & 8 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \mathbf{x}^T\mathbf{A}\mathbf{x}$$

So it follows: $\mathbf{X} \sim \mathcal{N}_2(\mathbf{0}, \mathbf{A}^{-1})$ where $\mathbf{A}^{-1} = \begin{pmatrix} 0.5 & -0.25 \\ -0.25 & 0.25 \end{pmatrix}$

SOLUTION EXERCISE 4:

(a) Apply the marginalization rule:

$$Y_i \sim \mathcal{N}(0.5 \cdot 2, 0.5 + 0.5 \cdot 2 \cdot 0.5) = \mathcal{N}(1, 1) \qquad (i = 1, 2)$$

(b) Given μ, we have two iid random variables:

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} |\mu \sim \mathcal{N}_2\left(\begin{pmatrix} 0.5\mu \\ 0.5\mu \end{pmatrix}, \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}\right) = \mathcal{N}_2\left(\begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}\mu, \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}\right)$$

For $\mu = 2$ we get:

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} |(\mu = 2) \sim \mathcal{N}_2\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}\right)$$

(c) Apply the marginalization rule, using the result from (b):

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \sim \mathcal{N}_2\left(\begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} \cdot 2, \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix} + \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} 2\ (0.5\ \ 0.5)\right) = \mathcal{N}_2\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}\right)$$

SOLUTION EXERCISE 5:

(a) We re-write:

$$p(\boldsymbol{\beta}) \propto \exp\{-\boldsymbol{\beta}^T\boldsymbol{\beta}\} = \exp\{-\frac{1}{2} \cdot \boldsymbol{\beta}^T(2\mathbf{I})\boldsymbol{\beta}\}$$

and recognize the shape of the density of the Gaussian distribution:

$$\boldsymbol{\beta} \sim \mathcal{N}_{k+1}\left(\mathbf{0}, (2\mathbf{I})^{-1}\right) = \mathcal{N}_{k+1}\left(\mathbf{0}, 0.5\mathbf{I}\right)$$

(b) As a function of $\boldsymbol{\beta}$ we have for the likelihood:

$$\begin{aligned} p(\mathbf{y}|\boldsymbol{\beta}) &\propto \exp\{-\frac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T\boldsymbol{\Sigma}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})\} \\ &\propto \exp\{-\frac{1}{2}(\mathbf{y}^T\boldsymbol{\Sigma}^{-1}\mathbf{y} - \mathbf{y}^T\boldsymbol{\Sigma}^{-1}\mathbf{X}\boldsymbol{\beta} - \boldsymbol{\beta}^T\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{y} + \boldsymbol{\beta}^T\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X}\boldsymbol{\beta})\} \\ &\propto \exp\{-\frac{1}{2}(-2\boldsymbol{\beta}^T\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{y} + \boldsymbol{\beta}^T\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X}\boldsymbol{\beta})\} \\ &\propto \exp\{\boldsymbol{\beta}^T\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{y} - \frac{1}{2}\boldsymbol{\beta}^T\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X}\boldsymbol{\beta}\} \end{aligned}$$

For the posterior density of $\boldsymbol{\beta}$ we have:

$$\begin{aligned}
p(\boldsymbol{\beta}|\mathbf{y}) &\propto p(\mathbf{y}|\boldsymbol{\beta}) \cdot p(\boldsymbol{\beta}) \\
&\propto \exp\{\boldsymbol{\beta}^T \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{y} - \frac{1}{2} \boldsymbol{\beta}^T \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X} \boldsymbol{\beta}\} \cdot \exp\{-\boldsymbol{\beta}^T \boldsymbol{\beta}\} \\
&\propto \exp\{\boldsymbol{\beta}^T \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{y} - \frac{1}{2} \boldsymbol{\beta}^T \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X} \boldsymbol{\beta} - \frac{1}{2} \boldsymbol{\beta}^T (2\mathbf{I}) \boldsymbol{\beta}\} \\
&\propto \exp\{\boldsymbol{\beta}^T \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{y} - \frac{1}{2} \boldsymbol{\beta}^T \left(\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X} + 2\mathbf{I} \right) \boldsymbol{\beta}\}
\end{aligned}$$

From the shape of the posterior density it follows that the posterior distribution of $\boldsymbol{\beta}$ is the following **multivariate Gaussian distribution**:

$$\boldsymbol{\beta}|\mathbf{y} \sim \mathcal{N}_{k+1}(\boldsymbol{\mu}^\dagger, \boldsymbol{\Sigma}^\dagger)$$

where $\boldsymbol{\Sigma}^\dagger = \left(\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X} + 2\mathbf{I} \right)^{-1}$ and $\mu^\dagger = \left(\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X} + 2\mathbf{I} \right)^{-1} \cdot \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{y}$.

SOLUTION EXERCISE 6:

We have the likelihood:

$$\begin{aligned}
p(\mathbf{y}_1, \dots, \mathbf{y}_n | \boldsymbol{\mu}) &= \prod_{i=1}^{n} p(\mathbf{y}_i | \boldsymbol{\mu}) \\
&= \prod_{i=1}^{n} (2\pi)^{-N/2} \cdot \det(\mathbf{I})^{-1/2} \cdot \exp\{-\frac{1}{2} \cdot (\mathbf{y}_i - \boldsymbol{\mu})^T \mathbf{I}^{-1} (\mathbf{y}_i - \boldsymbol{\mu})\} \\
&= (2\pi)^{-n \cdot N/2} \cdot \exp\{-\frac{1}{2} \cdot \sum_{i=1}^{n} (\mathbf{y}_i - \boldsymbol{\mu})^T (\mathbf{y}_i - \boldsymbol{\mu})\} \\
&= (2\pi)^{-n \cdot N/2} \cdot \exp\{-\frac{1}{2} \cdot \sum_{i=1}^{n} (\mathbf{y}_i^T \mathbf{y}_i - 2\boldsymbol{\mu}^T \mathbf{y}_i + \boldsymbol{\mu}^T \boldsymbol{\mu})\} \\
&= (2\pi)^{-n \cdot N/2} \cdot \exp\{-\frac{1}{2} \left(\sum_{i=1}^{n} \mathbf{y}_i^T \mathbf{y}_i \right) + \boldsymbol{\mu}^T \left(\sum_{i=1}^{n} \mathbf{y}_i \right) - \frac{n}{2} \boldsymbol{\mu}^T \boldsymbol{\mu})\}
\end{aligned}$$

As a function of $\boldsymbol{\mu}$ we have:

$$p(\mathbf{y}_1, \dots, \mathbf{y}_n | \boldsymbol{\mu}) \propto \exp\{\boldsymbol{\mu}^T \left(\sum_{i=1}^{n} \mathbf{y}_i \right) - \frac{n}{2} \boldsymbol{\mu}^T \boldsymbol{\mu})\}$$

And for the posterior density we have:

$$\begin{aligned}
p(\boldsymbol{\mu} | \mathbf{y}_1, \dots, \mathbf{y}_n) &\propto p(\mathbf{y}_1, \dots, \mathbf{y}_n | \boldsymbol{\mu}) \cdot p(\boldsymbol{\mu}) \\
&\propto \exp\{\boldsymbol{\mu}^T \left(\sum_{i=1}^{n} \mathbf{y}_i \right) - \frac{n}{2} \boldsymbol{\mu}^T \boldsymbol{\mu})\} \\
&\propto \exp\{\boldsymbol{\mu}^T \left(\sum_{i=1}^{n} \mathbf{y}_i \right) - \frac{1}{2} \boldsymbol{\mu}^T (n\mathbf{I}) \boldsymbol{\mu})\}
\end{aligned}$$

in which we recognize the shape of the density of the following Gaussian distribution:

$$\boldsymbol{\mu}|\mathbf{y} \sim \mathcal{N}_N(\boldsymbol{\mu}^\dagger, \boldsymbol{\Sigma}^\dagger)$$

where $\boldsymbol{\Sigma}^\dagger = (n\mathbf{I})^{-1} = \frac{1}{n}\mathbf{I}$ and $\mu^\dagger = \frac{1}{n}\mathbf{I} \cdot \sum_{i=1}^{n} \mathbf{y}_i =: \bar{\mathbf{y}}$.