EXERCISE 1 - SOLUTION:

(a) Joint PDF (likelihood):

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

For the posterior PDF we have:

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

The posterior PDF is proportional to the density of a $Beta(a+n\cdot r, b+\sum_{i=1}^n y_i)$ distribution; i.e. the posterior distribution of θ is a Beta distribution with parameters:

$$\begin{aligned} \tilde{a} &= a+n\cdot r \\ \tilde{b} &= b+\sum_{i=1}^n y_i \end{aligned}$$

(b) For $r=1$ we have:

$$\begin{aligned} \tilde{a} &= a+n \\ \tilde{b} &= b+\sum_{i=1}^n y_i \end{aligned}$$

In terms of pseudo counts: There are a pseudo observations, whose sum is b.

(c) Compute the marginal likelihood (using $r=1$, $a=1$ and $b=1$):

$$\begin{aligned} p(y_1,\dots,y_n) &= \int p(y_1,\dots,y_n|\theta)\cdot p(\theta)\ d\theta \\ &= \int \theta^{n\cdot r} \cdot (1-\theta)^{\sum_{i=1}^n y_i} \cdot \prod_{i=1}^{n} \binom{r+y_i-1}{y_i} \cdot \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \theta^{a-1} \cdot (1-\theta)^{b-1}\ d\theta \\ &= \int \theta^{n\cdot 1} \cdot (1-\theta)^{\sum_{i=1}^n y_i} \cdot \prod_{i=1}^{n} \binom{1+y_i-1}{y_i} \cdot \frac{\Gamma(2)}{\Gamma(1)\Gamma(1)} \cdot \theta^{1-1} \cdot (1-\theta)^{1-1}\ d\theta \\ &= \int \theta^{n} \cdot (1-\theta)^{\sum_{i=1}^n y_i}\ d\theta \\ &= \frac{\Gamma(n+1)\cdot\Gamma(\sum_{i=1}^n y_i+1)}{\Gamma(n+\sum_{i=1}^n y_i+2)} \cdot \int \frac{\Gamma(n+\sum_{i=1}^n y_i+2)}{\Gamma(n+1)\cdot\Gamma(\sum_{i=1}^n y_i+1)} \cdot \theta^n \cdot (1-\theta)^{\sum_{i=1}^n y_i}\ d\theta \\ &= \frac{\Gamma(n+1)\cdot\Gamma(\sum_{i=1}^n y_i+1)}{\Gamma(n+\sum_{i=1}^n y_i+2)} \end{aligned}$$

EXERCISE 2 - SOLUTION:

(a) For the full conditional PDF:

$$
\begin{aligned}
p(b|y_1,\dots,y_n,a) &\propto p(y_1,\dots,y_n|a,b)\cdot p(a)\cdot p(b) \\
&\propto p(y_1,\dots,y_n|a,b)\cdot p(b) \\
&\propto \left(\prod_{i=1}^{n} p(y_i|a,b)\right)\cdot p(b) \\
&\propto \left(\prod_{i=1}^{n} \frac{b^a}{\Gamma(a)}\cdot y_i^{a-1}\cdot e^{-b\cdot y_i}\right)\cdot \delta\cdot e^{-\delta\cdot b} \\
&\propto \frac{b^{n\cdot a}}{\Gamma(a)^n}\cdot\left(\prod_{i=1}^{n} y_i\right)^{a-1}\cdot e^{-b\cdot\sum_{i=1}^{n} y_i}\cdot\delta\cdot e^{-\delta\cdot b} \\
&\propto b^{n\cdot a}\cdot e^{-b\cdot(\delta+\sum_{i=1}^{n} y_i)}
\end{aligned}
$$

The full conditional PDF is proportional to the density of a $GAM(a\cdot n+1,\delta+\sum_{i=1}^{n} y_i)$ distribution; i.e. the posterior distribution of b is a Gamma distribution with parameters:

$$
\begin{aligned}
\tilde{a} &= a\cdot n+1 \\
\tilde{b} &= \delta+\sum_{i=1}^{n} y_i
\end{aligned}
$$

(b) Marginal likelihood for $a=1$, $\delta=2$, $n=5$ and $y_1=3$, $y_2=2$, $y_3=0.5$, $y_4=2$, $y_5=0.5$. Note that we here have: $\sum y_i=8$ and $\prod y_i=3$.

$$
\begin{aligned}
p(y_1,\dots,y_n) &= \int p(y_1,\dots,y_n|b)\cdot p(b)\;db \\
&= \int \frac{b^{n\cdot a}}{\Gamma(a)^n}\cdot\left(\prod_{i=1}^{n} y_i\right)^{a-1}\cdot e^{-b\cdot\sum_{i=1}^{n} y_i}\cdot\delta\cdot e^{-\delta\cdot b}\;db \\
&= \int \frac{b^{5\cdot 1}}{\Gamma(1)^5}\cdot(3)^{1-1}\cdot e^{-b\cdot 8}\cdot 2\cdot e^{-2\cdot b}\;db \\
&= 2\cdot\int b^5\cdot e^{-b\cdot 10}\;db \\
&= 2\cdot\frac{\Gamma(6)}{10^6}\cdot\int\frac{10^6}{\Gamma(6)}\cdot b^5\cdot e^{-b\cdot 10}\;db \\
&= 2\cdot\frac{\Gamma(6)}{10^6} \\
&= \frac{240}{10^6} = 0.00024
\end{aligned}
$$

(b) Predictive probability (for $\tilde{y} = 2$). For the posterior parameters we now have: $\tilde{a} = 6$ and $\tilde{b} = 10$.

$$\begin{aligned}
p(\tilde{y}|y_1, \ldots, y_n) &= \int p(\tilde{y}|b) \cdot p(b|y_1, \ldots, y_n)\, db \\
&= \int \frac{b^1}{\Gamma(1)} \cdot 2^{1-1} \cdot e^{-b \cdot 2} \cdot \frac{10^6}{\Gamma(6)} \cdot b^{6-1} \cdot e^{-10 \cdot b}\, db \\
&= \frac{10^6}{\Gamma(6)} \cdot \int b^6 \cdot e^{-b \cdot 12}\, db \\
&= \frac{10^6}{\Gamma(6)} \cdot \frac{\Gamma(7)}{12^7} \cdot \int \frac{12^7}{\Gamma(7)} \cdot b^6 \cdot e^{-b \cdot 12}\, db \\
&= \frac{6}{12} \cdot \left(\frac{10}{12}\right)^6 \approx 0.17
\end{aligned}$$

EXERCISE 3 - SOLUTION:

Initialisation: Set $\theta^{(0)} = \theta$, where $\theta \in \mathbb{R}$.
Iterations For $t = 1, \ldots, T$

- Sample random number u from Uniform distribution over $[-\epsilon, \epsilon]$, $u \sim \text{UNI}([-\epsilon, \epsilon])$.
- Propose new candidate: $\theta^\star = \theta^{(t-1)} + u$
- Compute the acceptance probability $A(\theta^{(t-1)}, \theta^\star) = \min\{1, R\}$, where $R = \frac{\prod_{i=1}^n p(y_i|\theta^\star)}{\prod_{i=1}^n p(y_i|\theta^{(t-1)})} \cdot \frac{p(\theta^\star)}{p(\theta^{(t-1)})} \cdot 1$
- Sample random number v from Uniform distribution over $[0, 1]$, $v \sim \text{UNI}([0, 1])$.
- IF $v < A(\theta^{(t-1)}, \theta^\star)$, set $\theta^{(t)} = \theta^\star$. ELSE Set $\theta^{(t)} = \theta^{(t-1)}$. ...

Output: $\theta^{(0)}, \ldots, \theta^{(T)}$

EXERCISE 4 - SOLUTION:
(a)

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

(b)

$$
\begin{aligned}
p(\mathbf{x}) &= (2\pi)^{-n/2} \cdot \det(\sigma^2\mathbf{I})^{-1/2} \cdot \exp\{-\frac{1}{2} \cdot (\mathbf{x}-\mu\mathbf{1})^T(\sigma^2\mathbf{I})^{-1}(\mathbf{x}-\mu\mathbf{1})\} \\
&= (2\pi)^{-n/2} \cdot \sigma^{-n} \det(\mathbf{I})^{-1/2} \cdot \exp\{-\frac{1}{2} \cdot (\mathbf{x}-\mu\mathbf{1})^T\sigma^{-2}\mathbf{I}(\mathbf{x}-\mu\mathbf{1})\} \\
&= (2\pi)^{-n/2} \cdot \sigma^{-n} \cdot \exp\{-\frac{1}{2} \cdot \frac{1}{\sigma^2} \cdot (\mathbf{x}-\mu\mathbf{1})^T(\mathbf{x}-\mu\mathbf{1})\} \\
&= (2\pi)^{-n/2} \cdot \sigma^{-n} \cdot \exp\{-\frac{1}{2} \cdot \frac{1}{\sigma^2} \cdot \sum_{i=1}^{n}(x_i-\mu)^2\}
\end{aligned}
$$

On the other hand, we have for the joint PDF of an i.i.d sample from the $N(\mu, \sigma^2)$:

$$
\begin{aligned}
p(x_1, \ldots, x_n) &= \prod_{i=1}^{n} p(x_i) \\
&= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sigma} \cdot \exp\{-\frac{1}{2} \cdot \frac{(x_i-\mu)^2}{\sigma^2}\} \\
&= \left(\frac{1}{\sqrt{2\pi}}\right)^n \cdot \frac{1}{\sigma^n} \cdot \exp\{-\frac{1}{2}\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2}\} \\
&= (2\pi)^{-n/2} \cdot \sigma^{-n} \cdot \exp\{-\frac{1}{2} \cdot \frac{1}{\sigma^2} \cdot \sum_{i=1}^{n}(x_i-\mu)^2\}
\end{aligned}
$$

(c) First note that:

$$p(x) \propto e^{-2x^2-4x+7} \propto e^{-2x^2-4x}$$

Part (a) implies for the one-dimensional Gaussian:

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

Therefore X must have a Gaussian distribution. For the parameters we solve:

$$
\begin{aligned}
-\frac{1}{2}\sigma^{-2} &= -2 \\
\sigma^{-2} \cdot \mu &= -4
\end{aligned}
$$

Solution is $\sigma^2 = \frac{1}{4}$ and $\mu = -1$, and it follows: $X \sim N(-1, \frac{1}{4})$.

(c) First note that:

$$p(x) \propto e^{-4x+7} \propto e^{-4x}$$

The density has thus the shape of the density of a Gamma distribution with parameters $a = 1$ and $b = 4$. So: $X \sim \text{GAM}(1, 4)$. This, by the way, is the exponential distribution with parameter b.

(d) We have:

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

This is the shape of the density of a Gamma distribution with parameters:

$$\begin{aligned}
a &= \frac{n}{2} + 1 \\
b &= \frac{1}{2} \cdot \mathbf{x}^T \cdot \mathbf{W}^{-1} \mathbf{x}
\end{aligned}$$

Therefore $\lambda \sim GAM(\frac{n}{2} + 1, \frac{1}{2} \cdot \mathbf{x}^T \cdot \mathbf{W}^{-1}\mathbf{x})$.