# Solutions to selected exercises from BDA3

Geir Storvik

Spring 2020

## Exam STK4021 2013, Exercise 2

($a$) We have

$$\ell(\theta) = \sum_{i=1}^{n}[-\frac{1}{2}\theta y_i^2 + \log(\theta) + \log(y_i)]$$
$$\frac{\partial}{\partial\theta}\ell(\theta) = \sum_{i=1}^{n}[-\frac{1}{2}y_i^2 + \frac{1}{\theta}] = -\frac{1}{2}\sum_{i=1}^{n} y_i^2 + \frac{n}{\theta}$$
$$\frac{\partial^2}{\partial\theta^2}\ell(\theta) = -\frac{n}{\theta^2}$$

so the function is unimodal and we obtain

$$\hat{\theta}_{ML} = \frac{2n}{\sum_{i=1}^{n} y_i^2} = 2w_n^{-1}$$

($b$) We have

$$p(\theta|\boldsymbol{y}) \propto \theta^{a-1}e^{-b\theta}\prod_{i=1}^{n} e^{-0.5\theta y_i^2}\theta y_i \propto \theta^{a+n-1}e^{[b+0.5\sum_{i=1}^{n} y_i^2]\theta}$$
$$\propto \text{Gamma}(a_n, b_n)$$

with

$$a_n = a + n \quad , b_n = b + 0.5\sum_{i=1}^{n} y_i^2$$

($c$) Under square root loss, the Bayes rule is the posterior expectation, that is

$$\hat{\theta}_B = \frac{a+n}{b + 0.5\sum_{i=1}^{n} y_i^2} = \frac{2a+2n}{2b + \sum_{i=1}^{n} y_i^2}$$

90% creidibility intervals can be found be taking the 0.05 and 0.95 quantiles in the Gamma distribution about.

(*d*) We have

$$\begin{aligned}p(y_{n+1}|\boldsymbol{y}) =& \int_\theta p(y_{n+1}|\theta)p(\theta|\boldsymbol{y})d\theta\\ =& \int_\theta e^{-0.5\theta y_{n+1}^2}\theta y_{n+1}\frac{b_n^{a_n}}{\Gamma(a_n)}\theta^{a_n-1}e^{-b_n\theta}d\theta\\ =& a_n y_{n+1}\frac{b_n^{a_n}}{[b_n+0.5y_{n+1}^2]^{a_n+1}}\times\\ &\int_\theta \frac{[b_n+0.5y_{n+1}^2]^{a_n+1}}{\Gamma(a_n+1)}\theta^{a_n+1-1}e^{-[b_n+0.5y_{n+1}^2]\theta}d\theta\\ =& a_n y_{n+1}\frac{b_n^{a_n}}{[b_n+0.5y_{n+1}^2]^{a_n+1}}\end{aligned}$$

## Exam STK4021 2014, Exercise 2

(*a*) Cauchy density

$$p(y) = \frac{1}{\pi[1+(y-\theta)^2]}$$

We have

$$p(\theta|\boldsymbol{y}) \propto \prod_{i=1}^2 \frac{1}{\pi[1+(y_i-\theta)^2]}I(0\le\theta\le 1)$$

The log-posterior is for $\theta\in[0,1]$

$$\begin{aligned}\log p(\theta|\boldsymbol{y}) =& -2\log(\pi)-\sum_{i=1}^2\log[1+(y_i-\theta)^2]\\ \frac{\partial}{\partial\theta}\log p(\theta|\boldsymbol{y}) =& -2\sum_{i=1}^2\frac{(y_i-\theta)}{1+(y_i-\theta)^2}\\ \frac{\partial^2}{\partial\theta^2}\log p(\theta|\boldsymbol{y}) =& -2\sum_{i=1}^2\frac{-\theta[1+(y_i-\theta)^2]+2(y_i-\theta)^2}{[1+(y_i-\theta)^2]^4}\end{aligned}$$

We also see that since $y_2 = 1-y_1$ that $\log p(\theta|\boldsymbol{y}) = -\log p(1-\theta|\boldsymbol{y})$, so antisymmetric around $\theta = 0.5$. Further, the derivative is zero at $\theta = 0.5$ and

$$\frac{\partial^2}{\partial\theta^2}\log p(\theta|\boldsymbol{y}) = -4\left[\frac{-0.5(1+0.5^2)+2*0.5^2}{[1+0.5^2]^4}\right] = -2.97$$

Since the second derivative is zero, we get a maximumpoint.
The Gaussian approximation becomes $N(0.5, 1/\sqrt{2.97})$.

(*b*) We have

$$\int_{-\infty}^{\infty}\prod_{i=1}^2\frac{1}{\pi[1+(y_i-\theta)^2]}d\theta \le \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{1}{\pi[1+(y_i-\theta)^2]}d\theta = \frac{1}{\pi}$$

so a finite integral that we can normalize.
With the same argument as before, the derivative will be zero for $\hat\theta = \frac{y_1+y_2}{2} = \bar y$. Define $d = |y_1 - y_2$. Then we need $(y_i-\hat\theta)^2 = d^2/4$. We then need

$$-\bar y(1+d^2/4)+2d^2/4 > 0$$

so $\bar y < 8d^2/(4+d^2)$.

## Exam STK4021 2014, Exercise 3

($a$) We have

$$
\begin{aligned}
p(y|\alpha,\beta) &= \int_0^1 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}\frac{n!}{y!(n-y)!}\theta^y(1-\theta)^{n-y}d\theta\\
&=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{n!}{y!(n-y)!}\frac{\Gamma(\alpha+y)\Gamma(\beta+n-y)}{\Gamma(\alpha+\beta+n)}\\
&\quad\int_0^1 \frac{\Gamma(\alpha+\beta+n)}{\Gamma(\alpha+y)\Gamma(\beta+n-y)}\theta^{\alpha+y-1}(1-\theta)^{\beta+n-y-1}d\theta\\
&=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{n!}{y!(n-y)!}\frac{\Gamma(\alpha+y)\Gamma(\beta+n-y)}{\Gamma(\alpha+\beta+n)}\\
&=\frac{\Gamma(\alpha+\beta)n!}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\alpha+\beta+n)}\frac{\Gamma(\alpha+y)\Gamma(\beta+n-y)}{\Gamma(y+1)\Gamma(n-y+1)}
\end{aligned}
$$

where the last term can only be constant if $\alpha=\beta=1$.

($b$) The ML estimate is $y/n$. The posterior is

$$
p(\theta|y) \propto\theta^{\alpha-1}(1-\theta)^{\beta-1}\theta^y(1-\theta)^{n-y} \propto \text{Beta}(\alpha+y,\beta+n-y)
$$

Under square error loss, the Bayes rule is the expectation, so

$$
\hat{\theta}_B = \frac{\alpha+y}{\alpha+\beta+n}
$$

($c$) We have

$$
\begin{aligned}
R(\hat{\theta}_{ML}) =&E[(\hat{\theta}_{ML}-\theta)^2] = E[(\frac{y}{n}-\theta)^2] = \text{var}[y]/n^2 = \theta(1-\theta)/n\\
R(\hat{\theta}_B) =&E[(\hat{\theta}_B-\theta)^2] = E[(\frac{\alpha+y}{\alpha+\beta+n}-\theta)^2]\\
=&E[(\frac{\alpha+y}{\alpha+\beta+n}-\frac{\alpha+n\theta}{\alpha+\beta+n}+\frac{\alpha+n\theta}{\alpha+\beta+n}-\theta)^2]\\
=&\text{Var}[\frac{\alpha+y}{\alpha+\beta+n}]+[(\frac{\alpha+n\theta}{\alpha+\beta+n}-\theta)^2]\\
=&\frac{n\theta(1-\theta)}{(\alpha+\beta+n)^2}+\frac{(\alpha(1-\theta)+\beta\theta)^2}{(\alpha+\beta+n)^2}.
\end{aligned}
$$

Note that the first term will be smaller than $R(\hat{\theta}_{ML})$ while the last term is $O(n^{-2})$ so for large enough $n$ it will be smaller.

For $\alpha=\beta$, we get

$$
R(\hat{\theta}_B) =\frac{n\theta(1-\theta)}{(2\alpha+n)^2}+\frac{\alpha^2}{(2\alpha+n)^2}
$$

We then get that

$$
\theta(1-\theta) > \frac{n\alpha^2}{(2\alpha+n)^2-n^2}
$$

which will occur if $\theta$ is close to 0.5, that is close to the prior expectation.

## Exam STK4021 2017, Exercise 1

(*a*) We have

$$\begin{aligned}
E[y_i|\theta] =& \theta \\
\text{Var}[y_i|\theta] =& \theta(1-\theta) \\
E[z|\theta] =& n\theta \\
\text{Var}[z|\theta] =& n\theta(1-\theta)
\end{aligned}$$

(*b*) (*i*)

$$\begin{aligned}
E[\theta] =& \frac{2}{2+2} = \frac{1}{2} \\
\text{Var}[\theta] =& \frac{2\cdot 2}{(2+2)^2(2+2+1)} = \frac{1}{20}
\end{aligned}$$

(*ii*)

$$\begin{aligned}
E[y_i] =& E[E[y_i|\theta]] = E[\theta] = \frac{1}{2} \\
\text{Var}[y_i] =& E[\text{Var}[y_i|\theta]] + \text{Var}[E[y_i|\theta]] \\
=& E[\theta(1-\theta)] + \text{Var}[\theta] = E[\theta] - E[\theta^2] + E[\theta^2] - (E[\theta])^2 = \frac{1}{2} - (\frac{1}{2})^2 = \frac{1}{4}
\end{aligned}$$

Further, $y_1$ is a discrete variable with only two outcomes, 0 and 1 and thereby has to be a Bernoulli variable. Further, since the expectation is 0.5, the sucess probabilyt has to be 0.5 as well, thereby we get a Bernoulli distribution with success probability 0.5.

(*iii*)

$$\begin{aligned}
\text{Cov}[y_i, y_j] =& E[\text{Cov}[y_i, y_j|\theta]] + \text{Cov}[E[y_i|\theta], E[y_j|\theta]] \\
=& 0 + \text{Cov}[\theta, \theta] = \frac{1}{20} \\
\text{Cor}[y_i, y_j] =& \frac{1}{5}
\end{aligned}$$

(*iv*)

$$\begin{aligned}
E[z] =& \sum_{i=1}^{n} E[y_i] = n\theta \\
\text{Var}[z] =& \sum_{i=1}^{n} \text{Var}[y_i] + \sum_{i\neq j} \text{Cov}[y_i, y_j] = \frac{n}{4} + \frac{n(n-1)}{20} = \frac{n}{5} + \frac{n^2}{20}
\end{aligned}$$

Alternatively,

$$\begin{aligned}
\text{Var}[z] =& E[\text{Var}[z|\theta]] + \text{Var}[E[z|\theta]] \\
=& E[n\theta(1-\theta)] + \text{Var}[n\theta] \\
=& n[E[\theta] - E[\theta^2]] + n^2\text{Var}[n\theta] \\
=& n[E[\theta] - \text{Var}[n\theta] - (E[\theta])^2] + n^2\text{Var}[\theta] \\
=& n[\frac{1}{2} - \frac{1}{20} - \frac{1}{4}] + n^2\frac{1}{20} \\
=& \frac{n}{5} + \frac{n^2}{20}
\end{aligned}$$

(c)

$$\begin{aligned}
p(z) =& \int_0^1 p(z|\theta)p(\theta)d\theta \\
=& \int_0^1 \binom{n}{z}\theta^z(1-\theta)^{n-z}\frac{\Gamma(4)}{\Gamma(2)\Gamma(2)}\theta(1-\theta)d\theta \\
=& \binom{n}{z}\frac{\Gamma(4)}{\Gamma(2)\Gamma(2)}\frac{\Gamma(z+2)\Gamma(n-z+2)}{\Gamma(n+4)}\int_0^1\frac{\Gamma(n+4)}{\Gamma(z+2)\Gamma(n-z+2)}\theta^{z+2-1}(1-\theta)^{n-z+2-1} \\
=& \frac{n!3!(z+1)!(n-z+1)!}{z!(n-z)!(n+3)!} = \frac{6(z+1)(n-z+1)}{(n+3)(n+2)(n+1)}
\end{aligned}$$

(d)

$$\begin{aligned}
p(\theta|z) \propto& p(\theta)p(z|\theta) \propto \qquad \theta^z(1-\theta)^{n-z}\theta(1-\theta) \propto \text{Beta}(z+2, n-z+2) \\
E[\theta|z] =& \frac{z+2}{n+4}
\end{aligned}$$

(e) The risk function is defined as the expected loss when expectation is with respect to the random variable conditional on the parameter. This gives

$$R(\theta) = E[(\tilde{\theta}-\theta)^2|\theta] = E[[(\tfrac{z}{n}-\theta)^2]|\theta] = \text{Var}[\tfrac{z}{n}] = \frac{1}{n}\theta(1-\theta)$$

(f) Under quadratic loss, the Bayes estimator is the posterior expectation. Using that

$$\begin{aligned}
E[\hat{\theta}_B|\theta] =& \frac{n\theta+2}{n+4} \\
\text{Var}[\hat{\theta}_B|\theta] =& \left(\frac{n}{n+4}\right)^2\left[\frac{n}{4}+\frac{n(n-1)}{5}\right]
\end{aligned}$$

we get

$$\begin{aligned}
R_B(\theta) =& E[(\hat{\theta}_B-\theta)^2|\theta] \\
=& E\left[\left(\frac{z+2}{n+4}-\frac{n\theta+2}{n+4}-\frac{2-4\theta}{n+4}\right)^2|\theta\right] \\
=& \text{Var}\left[\frac{z+2}{n+4}|\theta\right] + E\left(\frac{2-4\theta}{n+4}\right)^2 \\
=& \frac{n\theta(1-\theta)}{(n+4)^2} + \left(\frac{2-4\theta}{n+4}\right)^2
\end{aligned}$$

The Bayes estimator is better when its risk is smaller, i.e.

$$\begin{aligned}
\frac{n\theta(1-\theta)}{(n+4)^2} + \left(\frac{2-4\theta}{n+4}\right)^2 <& \frac{1}{n}\theta(1-\theta) \\
&\Downarrow \\
\left(\frac{2-4\theta}{n+4}\right)^2 <& \frac{8n+16}{n(n+4)^2}\theta(1-\theta) \\
&\Downarrow \\
\frac{(\theta-0.5)^2}{\theta(1-\theta)} <& \frac{n+2}{2n}
\end{aligned}$$

## Exam STK4021 2017, Exercise 4

(a) We have

$$L(\theta) = \prod_{i=1}^{n} f(y_i|\theta) = \prod_{i=1}^{n}[3\theta y_i^2 e^{-\theta y_i^3} = 3^n\theta^n e^{-\theta\sum_{i=1}^{n} y_i^3}\prod_{i=1}^{n} y_i^2$$

$$l(\theta) = n\log(3) + n\log(\theta) - \theta\sum_{i=1}^{n} y_i^3 + 2\sum_{i=1}^{n}\log(y_i)$$

showing that $\sum_{i=1}^{n} y_i^3$ is a sufficient statistic.
We further have

$$\frac{\partial}{\partial\theta}l(\theta) = \frac{n}{\theta} - \sum_{i=1}^{n} y_i^3$$

giving

$$\hat{\theta}_{ML} = \frac{n}{\sum_{i=1}^{n} y_i^3}$$

(b) We have

$$\begin{aligned} p(\theta|y) &\propto \theta^{a-1}e^{-b\theta}\theta^n e^{-\theta\sum_{i=1}^{n} y_i^3} \\ &= \theta^{a+n-1}e^{-\theta[b+\sum_{i=1}^{n} y_i^3]} \\ &\propto \text{Gamma}(a+n, b+\sum_{i=1}^{n} y_i^3) \end{aligned}$$

(c) We have with $Z = \hat{\theta}_{ML}^{-1}$

$$Z = \frac{\sum_{i=1}^{n} y_i^3}{n} \approx N(z;\mu_3,\sigma_3^2/n)$$

where

$$\mu_3 = E[y_i^3] \qquad \sigma_3^2 = \text{Var}[y_i^3]$$

Now we have in general that

$$\int_0^\infty x^n e^{-ax^b} = \frac{1}{b}a^{-(n+1)/b}\Gamma((n+1)/b)$$

giving

$$\begin{aligned} \mu_3 &= 3\theta\int y^5 e^{-\theta y^3} = 3\theta\frac{1}{3}\theta^{-2}\Gamma(2) = \theta^{-1} \\ \sigma_3^2 &= E[y_i^6] - \theta^{-2} \\ &= 3\theta\int y^8 e^{-\theta y^3} = 3\theta\frac{1}{3}\theta^{-3}\Gamma(3) = \theta^{-2} \end{aligned}$$

Using transformation rules, we then get

$$\hat{\theta}_{ML} \approx N(\hat{\theta}_{ML}^{-1}; \mu_3, \theta^{-2}/n)\hat{\theta}_{ML}^{-2}$$

Using instead the *delta* method, we get

$$\hat{\theta}_{ML} \approx N(\theta, \theta^4\theta^{-2}/n) = N(\theta, \theta^2/n)$$

Assuming now $\theta_0$ is the true $\theta$ value, we have that

$$p(\theta|y) \approx N(\theta_0, \sigma_0^2)$$

where

$$\sigma_0^2 = (nJ(\theta_0))^{-1}$$
$$J(\theta) = -\frac{\partial^2}{\partial\theta^2} l(\theta) = \frac{n}{\theta^2}$$

so

$$p(\theta|y) \approx N(\theta_0, \theta_0^2/n)$$