Worksheet 2

Foundations of Bayesian Methodology

Group: Goliath

Spring Semester 2022

Exercise 3 (Conjugate Bayes: analytical derivation)

$$\begin{aligned} y_1, \cdots, y_n \mid m &\overset{i.i.d}{\sim} \mathcal{N}(m, \kappa^{-1}) \\ m &\sim \mathcal{N}(\mu, \lambda^{-1}) \end{aligned}$$

$$\begin{aligned} f(y_1, \cdots, y_n \mid m) &= \prod_{i=1}^{n} f(y_i \mid m) \\ &= \prod_{i=1}^{n} \left(\sqrt{\frac{\kappa}{2\pi}} \exp\left(-\frac{\kappa}{2}(y_i - m)^2\right) \right) \\ &= \left(\frac{\kappa}{2\pi}\right)^{\frac{n}{2}} \exp\left(-\frac{\kappa}{2} \sum_{i=1}^{n} (y_i - m)^2 \right) \end{aligned}$$

$$f(m) = \sqrt{\frac{\lambda}{2\pi}} \exp\left(-\frac{\lambda}{2}(m-\mu)^2\right)$$

$$\begin{aligned} & f(y_1, \cdots, y_n \mid m) f(m) \\ &= \left(\frac{\kappa}{2\pi}\right)^{\frac{n}{2}} \exp\left(-\frac{\kappa}{2} \sum_{i=1}^{n} (y_i - m)^2 \right) \cdot \sqrt{\frac{\lambda}{2\pi}} \exp\left(-\frac{\lambda}{2}(m-\mu)^2\right) \\ &= \left(\frac{\kappa}{2\pi}\right)^{\frac{n}{2}} \left(\frac{\lambda}{2\pi}\right)^{\frac{1}{2}} \exp\left(-\frac{\kappa}{2} \sum_{i=1}^{n} (y_i - m)^2 - \frac{\lambda}{2}(m-\mu)^2 \right) \\ &= \left(\frac{\kappa}{2\pi}\right)^{\frac{n}{2}} \left(\frac{\lambda}{2\pi}\right)^{\frac{1}{2}} \exp\left(-\frac{\kappa}{2} \sum_{i=1}^{n} (y_i^2 - 2y_i m + m^2) - \frac{\lambda}{2}(m^2 - 2m\mu + \mu^2) \right) \\ &= \left(\frac{\kappa}{2\pi}\right)^{\frac{n}{2}} \left(\frac{\lambda}{2\pi}\right)^{\frac{1}{2}} \exp\left(-\frac{\kappa}{2} \left(\sum_{i=1}^{n} y_i^2 - 2n\bar{y}m + nm^2 \right) - \frac{\lambda}{2}(m^2 - 2m\mu + \mu^2) \right) \\ &= \left(\frac{\kappa}{2\pi}\right)^{\frac{n}{2}} \left(\frac{\lambda}{2\pi}\right)^{\frac{1}{2}} \exp\left(-\frac{1}{2} \left((n\kappa + \lambda)m^2 - 2m(\kappa n\bar{y} + \lambda\mu) + \kappa \sum_{i=1}^{n} y_i^2 + \lambda\mu^2 \right) \right) \\ &= \left(\frac{\kappa}{2\pi}\right)^{\frac{n}{2}} \left(\frac{\lambda}{2\pi}\right)^{\frac{1}{2}} \exp\left(-\frac{n\kappa + \lambda}{2} \left(m^2 - 2m\frac{\kappa n\bar{y} + \lambda\mu}{n\kappa + \lambda} + \frac{\kappa \sum_{i=1}^{n} y_i^2 + \lambda\mu^2}{n\kappa + \lambda} \right) \right) \\ &= \left(\frac{\kappa}{2\pi}\right)^{\frac{n}{2}} \left(\frac{\lambda}{2\pi}\right)^{\frac{1}{2}} \exp\left(-\frac{n\kappa + \lambda}{2} \left(\left(m - \frac{\kappa n\bar{y} + \lambda\mu}{n\kappa + \lambda} \right)^2 - \left(\frac{\kappa n\bar{y} + \lambda\mu}{n\kappa + \lambda} \right)^2 + \frac{\kappa \sum_{i=1}^{n} y_i^2 + \lambda\mu^2}{n\kappa + \lambda} \right) \right) \\ &= \underbrace{\left(\frac{\kappa}{2\pi}\right)^{\frac{n}{2}} \left(\frac{\lambda}{2\pi}\right)^{\frac{1}{2}} \exp\left(\frac{\kappa \sum_{i=1}^{n} y_i^2 + \lambda\mu^2}{n\kappa + \lambda} - \left(\frac{\kappa n\bar{y} + \lambda\mu}{n\kappa + \lambda} \right)^2 \right)}_{\text{constant}} \exp\left(-\frac{n\kappa + \lambda}{2} \left(m - \frac{\kappa n\bar{y} + \lambda\mu}{n\kappa + \lambda} \right)^2 \right) \end{aligned}$$

$$
\begin{aligned}
f(m \mid y_1, \cdots, y_n) &= \frac{f(y_1, \cdots, y_n \mid m) f(m)}{f(y_1, \cdots, y_n)} \\
&= \frac{f(y_1, \cdots, y_n \mid m) f(m)}{\int_{-\infty}^{\infty} f(y_1, \cdots, y_n \mid m) f(m) \mathrm{d}m} \\
&= \frac{\exp\left(-\frac{n\kappa+\lambda}{2}\left(m - \frac{\kappa n \bar{y} + \lambda\mu}{n\kappa+\lambda}\right)^2\right)}{\int_{-\infty}^{\infty} \exp\left(-\frac{n\kappa+\lambda}{2}\left(m - \frac{\kappa n \bar{y} + \lambda\mu}{n\kappa+\lambda}\right)^2\right)} \\
&= \sqrt{\frac{n\kappa+\lambda}{2\pi}} \exp\left(-\frac{n\kappa+\lambda}{2}\left(m - \frac{\kappa n \bar{y} + \lambda\mu}{n\kappa+\lambda}\right)^2\right)
\end{aligned}
$$

$$
m \mid y_1, \cdots, y_n \sim \mathcal{N}\left(\frac{\kappa n \bar{y} + \lambda\mu}{n\kappa+\lambda}, \frac{1}{n\kappa+\lambda}\right)
$$

Note:

$$
\int_{-\infty}^{\infty} e^{-a(x+b)^2} \mathrm{d}x = \sqrt{\frac{\pi}{\alpha}}
$$

Source: Gaussian Integral

Exercise 4 (Conjugate Bayesian analysis in practice)

$$
\begin{aligned}
y_1, \cdots, y_n \mid m &\overset{i.i.d}{\sim} \mathcal{N}(m, \kappa^{-1}) \\
m &\sim \mathcal{N}(\mu, \lambda^{-1})
\end{aligned}
$$

$$
\kappa = \frac{1}{900}, \quad \mu = 161, \quad \lambda = \frac{1}{70}
$$

```r
height <- c(166, 168, 168, 177, 160, 170, 172, 159, 175, 164, 175, 167, 164)
table(height)
```

4(a)

```
## height
## 159 160 164 166 167 168 170 172 175 177
##   1   1   2   1   1   2   1   1   2   1
```

```r
hist(height, breaks=length(height), freq=FALSE)
lines(density(height), col="red")
```

```r
## Sample size
sample.n <- length(height); sample.n
```

```
## [1] 13
```

```r
## Sample meam
sample.mean <- mean(height); sample.mean
```

```
## [1] 168.0769
```

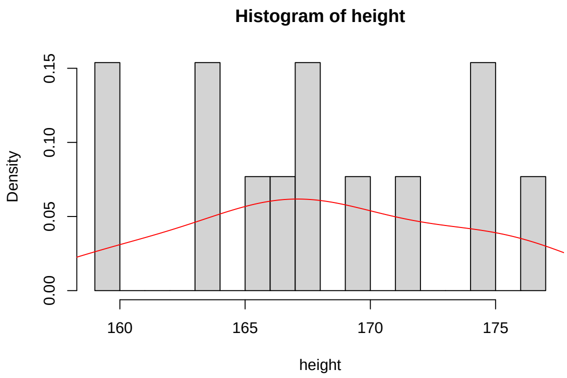


Figure 1: Histogram of the height measurements

```
## Sample standard deviation
sample.sd <- sd(height); sample.sd
```

```
## [1] 5.634145
```

```
## Sample standard error
sample.se <- sample.sd / sqrt(sample.n); sample.se
```

```
## [1] 1.562631
```

```
## Significance level
alpha <- 0.05

## Degrees of freedom
df <- sample.n - 1

## t-score
t.score <- qt(p=alpha/2, df=df, lower.tail=FALSE); t.score
```

```
## [1] 2.178813
```

```
## Contract the 95% confidence interval
lower.bound <- sample.mean - t.score * sample.se
upper.bound <- sample.mean + t.score * sample.se
print(c(lower.bound, upper.bound))
```

```
## [1] 164.6722 171.4816
```

Interpretation: there is a probability of 95% that the true mean will fall into the interval $(164.6722, 171.4816)$.

4(b)

$$P[m > 200] = 1 - P[m \leq 200]$$

```
## 2.5%, 50%(i.e. median), 97.5% quantiles of N(161,70)
qnorm(c(0.025, 0.5, 0.975), mean=161, sd=sqrt(70))
```

```
## [1] 144.6018 161.0000 177.3982
```

```r
## P[m>200]
pnorm(200, mean=161, sd=sqrt(70), lower.tail=FALSE)
```

```
## [1] 1.570393e-06
```

- Plot: see Figure 2
- Summary statistics: see Table 1

4(c)

$$
\begin{aligned}
y_1, \cdots, y_n \mid m &\overset{i.i.d}{\sim} \mathcal{N}(m, \kappa^{-1}) \\
m &\sim \mathcal{N}(\mu, \lambda^{-1}) \\
m \mid y_1, \cdots, y_n &\sim \mathcal{N}\left(\frac{\kappa n \bar{y} + \lambda\mu}{n\kappa + \lambda}, (n\kappa + \lambda)^{-1}\right)
\end{aligned}
$$

```r
## Posterior mean
((1/900)*2185+(1/70)*161 ) / (13*(1/900)+(1/70) )
```

```
## [1] 164.558
```

```r
## Posterior variance
(13*(1/900)+(1/70))^(-1)
```

```
## [1] 34.80663
```

$$
\begin{aligned}
y_1, \cdots, y_n &\sim \mathcal{N}(m, 900) \\
m &\sim \mathcal{N}(161, 70) \\
m \mid y_1, \cdots, y_n &\sim \mathcal{N}(164.558, 34.80663)
\end{aligned}
$$

```r
m <- seq(101, 221)
plot(m, dnorm(m, mean=161, sd=sqrt(70)), type="l", ylab="density", col="red",
     ylim=c(0,0.08), main="Prior vs Posterior")
lines(m, dnorm(m, mean=164.558, sd=sqrt(34.80663)), col="blue")
legend("topright", legend=c("prior: N(161,70)", "posterior: N(164.56, 34.81)"),
       lty=1, col=c("red", "blue"), cex=.8)
```

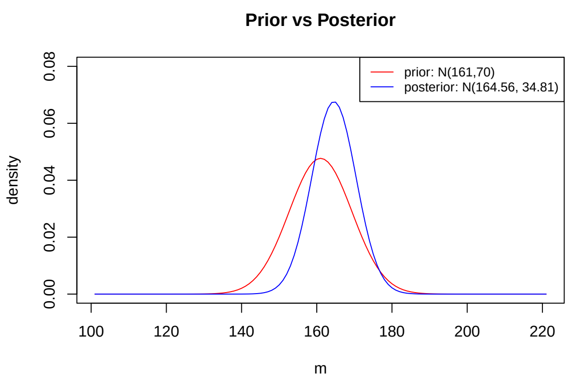


Figure 2: Prior vs. Posterior

```r
qnorm(c(0.025, 0.5, 0.975), mean=164.558, sd=sqrt(34.80663))
```

```
## [1] 152.9948 164.5580 176.1212
```

Table 1: Summary statistics of the prior and posterior distributions

	Mean	SD	Median	Equi-tailed 95% CI/CrI
m	161.000	8.3666	161.000	$(144.6018, 177.3982)$
$m \mid y_i, \cdots, y_n$	164.558	5.8997	164.558	$(152.9948, 176.1212)$

Interpretation: there is a posterior probability of 95% that the true mean falls into the interval between 152.9948 and 176.1212, given a $\mathcal{N}(m, 900)$ prior is assumed.

```r
## P[m>200|y1,...,yn]
pnorm(200, mean=164.558, sd=sqrt(34.80663), lower.tail=FALSE)
```

4(d)

```
## [1] 9.425552e-10
```

4(e)

$$\text{Prior} \rightarrow \text{Posterior}$$
$$\mathcal{N}(161, 70) \rightarrow \mathcal{N}(164.558, 34.80663)$$
$$P[m > 200] \rightarrow P[m > 200 \mid y_1, \cdots, y_n]$$
$$1.570393 \times 10^{-6} \rightarrow 9.425552 \times 10^{-10}$$

Table 2: Belief updating: Comparision between prior and posterior

	Prior	Posterior
Mean	161	164.558
Variance	70	34.80663
$\mathcal{N}(\mu, \sigma^2)$	$\mathcal{N}(161, 70)$	$\mathcal{N}(164.558, 34.80663)$
$P[m > 200]$	1.570393×10^{-6}	9.425552×10^{-10}

Exercise 5 (Bayesian learning)

Prior distribution:

$$p \sim \text{Beta}(\alpha, \beta)$$

Posterior distribution:

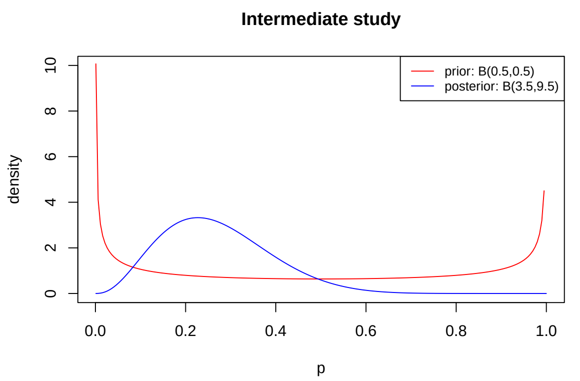
$$p \mid y_1, \cdots, y_n \sim \text{Beta}(\alpha + n\bar{y}, \beta + n - n\bar{y})$$

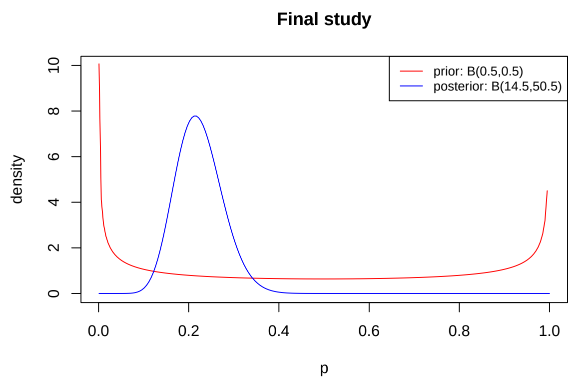
5(a) $\alpha = \beta = 0.5, \quad \text{Beta}(0.5, 0.5)$

Table 3: Response rate for each stage under $B(0.5, 0.5)$ prior

Stage	n	Responders x (%)	Prior $\mathrm{B}(\alpha, \beta)$	Posterior $\mathrm{B}(\alpha+x, \beta+n-x)$
Interim	12	3 (25%)	$\mathrm{B}(0.5, 0.5)$	$\mathrm{B}(3.5, 9.5)$
Final	64	14 (21.875%)	$\mathrm{B}(0.5, 0.5)$	$\mathrm{B}(14.5, 50.5)$

```r
p <- seq(1e-3,1, length=200)

plot(p, dbeta(p, 0.5, 0.5), type="l", ylab="density", col="red",
     ylim=c(0, 10), main="Intermediate study")
lines(p, dbeta(p, 3.5, 9.5), ylab="density", col="blue")

legend("topright", legend=c("prior: B(0.5,0.5)", "posterior: B(3.5,9.5)"),
       col=c("red", "blue"), lty=1, cex=.8)

plot(p, dbeta(p, 0.5, 0.5), type="l", ylab="density", col="red",
     ylim=c(0, 10), main="Final study")
lines(p, dbeta(p, 14.5, 50.5), ylab="density", col="blue")

legend("topright", legend=c("prior: B(0.5,0.5)", "posterior: B(14.5,50.5)"),
       col=c("red", "blue"), lty=1, cex=.8)
```





$$P[\text{response rate} > 0.4] = P[p > 0.4] = 1 - P[p \le 0.4]$$

```r
# Intermediate study
pbeta(0.4, 3.5, 9.5, lower.tail=FALSE)
```

```
## [1] 0.1437649
```

```r
# Final study
pbeta(0.4, 14.5, 50.5, lower.tail=FALSE)
```

```
## [1] 0.001075757
```

```r
# Intermediate study
## 2.5%, 50%, 97.5% quantiles of Beta(3.5, 9.5)
qbeta(c(0.025, 0.5, 0.975), 3.5, 9.5)
```

```
## [1] 0.07594233 0.25711895 0.52919108
```

```r
# Final study
## 2.5%, 50%, 97.5% quantiles of Beta(14.5, 50.5)
qbeta(c(0.025, 0.5, 0.975), 14.5, 50.5)
```

```
## [1] 0.1312669 0.2202242 0.3310055
```

Table 4: Summary statistics of posterior distribution ($B(0.5, 0.5)$ prior)

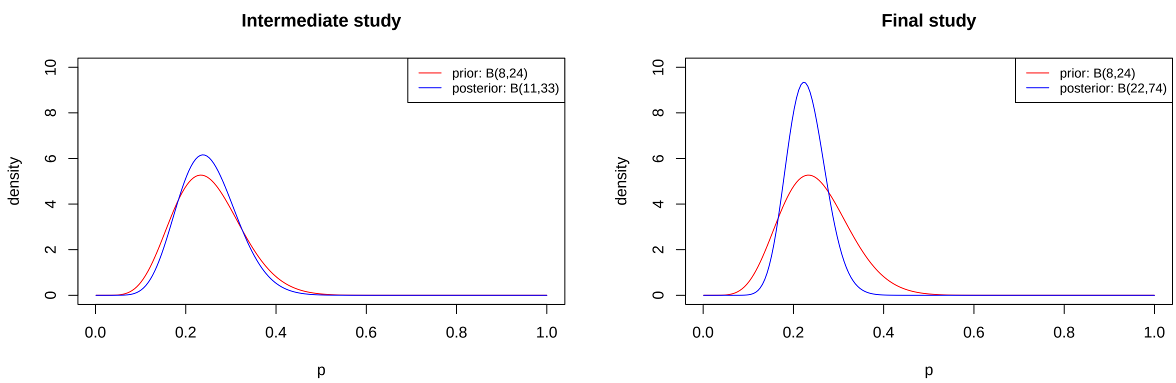
Stage	$B(\alpha, \beta)$	Mean	Median	95% CrI
Interim	$B(3.5, 9.5)$	0.2692	0.2571	$(0.0759, 0.5292)$
Final	$B(22, 74)$	0.2292	0.2202	$(0.1313, 0.3310)$

5(b) $\alpha = 8, \beta = 24$, Beta$(8, 24)$

Table 5: Response rate for each stage under $B(8, 24)$ prior

Stage	n	Responders x (%)	Prior B(α, β)	Posterior B$(\alpha + x, \beta + n - x)$
Interim	12	3 (25%)	B$(8, 24)$	B$(11, 33)$
Final	64	14 (21.875%)	B$(8, 24)$	B$(22, 74)$

```r
p <- seq(1e-3,1, length=200)

plot(p, dbeta(p, 8, 24), type="l", ylab="density", col="red",
     ylim=c(0, 10), main="Intermediate study")
lines(p, dbeta(p, 11, 33), ylab="density", col="blue")

legend("topright", legend=c("prior: B(8,24)", "posterior: B(11,33)"),
       col=c("red", "blue"), lty=1, cex=.8)

plot(p, dbeta(p, 8, 24), type="l", ylab="density", col="red",
     ylim=c(0, 10), main="Final study")
lines(p, dbeta(p, 22, 74), ylab="density", col="blue")

legend("topright", legend=c("prior: B(8,24)", "posterior: B(22,74)"),
       col=c("red", "blue"), lty=1, cex=.8)
```



$$P[\text{response rate} > 0.4] = P[p > 0.4] = 1 - P[p \leq 0.4]$$

```r
# Intermediate study
pbeta(0.4, 11, 33, lower.tail=FALSE)
```

```
## [1] 0.01621346
```

```r
# Final study
pbeta(0.4, 22, 74, lower.tail=FALSE)
```

```
## [1] 0.0001727695
```

```r
# Intermediate study
## 2.5%, 50%, 97.5% quantiles of Beta(11, 33)
qbeta(c(0.025, 0.5, 0.975), 11, 33)
```

```
## [1] 0.1351860 0.2461854 0.3863082
```

```r
# Final study
## 2.5%, 50%, 97.5% quantiles of Beta(22, 74)
qbeta(c(0.025, 0.5, 0.975), 22, 74)
```

```
## [1] 0.1511774 0.2272801 0.3178360
```

Table 6: Summary statistics of posterior distribution ($B(8, 24)$ prior)

Stage	Posterior	Mean	Median	95% CrI
Interim	$B(11, 33)$	0.2500	0.2461	$(0.1352, 0.3863)$
Final	$B(22, 74)$	0.2292	0.2273	$(0.1512, 0.3178)$