Worksheet 2

Foundations of Bayesian Methodology

Wenjie Tu

Spring Semester 2022

Exercise 1 (Individual project (Part 2A))

Table 1: ASAS20 responders at week 6

Group	n	Responders x (%)	Posterior response rate
Secukinumab	23	14 (60.9%)	59.5%
Placebo	6	1 (16.7%)	24.5%

1(a) Compute 95% CIs for the true probability of response in the Secukinumab and in the placebo groups:

```
library(DescTools)

## Secukinumab group
BinomCI(x=14, n=23, conf.level=0.95, method="wilson")
```

```
##            est    lwr.ci    upr.ci
## [1,] 0.6086957 0.4078552 0.7784238
```

```
## Placebo group
BinomCI(x=1, n=6, conf.level=0.95, method="wilson")
```

```
##            est     lwr.ci    upr.ci
## [1,] 0.1666667 0.03005337 0.5635028
```

Interpretation of CIs:

- In *Secukinumab* group, for repeated random samples from a binomial distribution with unknown but fixed parameters (i.e. the number of subjects and the response rate), the 95% confidence interval $(0.4079, 0.7784)$ will cover the response rate in 95% of all cases.
- In *Placebo* group, for repeated random samples from a binomial distribution with unknown but fixed parameters (i.e. the number of subjects and the response rate), the 95% confidence interval $(0.0301, 0.5635)$ will cover the response rate in 95% of all cases.

Wilson method is used to compute the 95% CIs for the true probability of response in Secukinumab and in the placebo groups. Wilson confidence interval is recommended in this setting since it does not suffer from problems of overshoot and zero-width intervals that afflict the normal interval and it can also be safely used with small samples and skewed observations. To show the difference among CIs computed from other methods, the proportion test and the exact binomial test are conducted for comparison.

```r
## Comparison

## Proportion test
prop.test(x=14, n=23, conf.level=0.95, correct=TRUE)
```

```
##
##  1-sample proportions test with continuity correction
##
## data:  14 out of 23, null probability 0.5
## X-squared = 0.69565, df = 1, p-value = 0.4042
## alternative hypothesis: true p is not equal to 0.5
## 95 percent confidence interval:
##  0.3878251 0.7953232
## sample estimates:
##         p
## 0.6086957
```

```r
## Exact binomial test
binom.test(x=14, n=23)
```

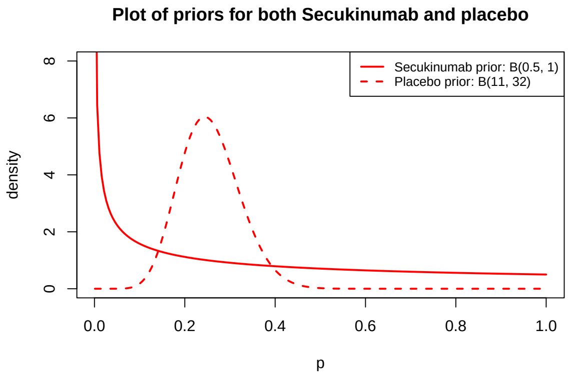
```
##
##  Exact binomial test
##
## data:  14 and 23
## number of successes = 14, number of trials = 23, p-value = 0.4049
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
##  0.3854190 0.8029236
## sample estimates:
## probability of success
##              0.6086957
```

1(b) Plot prior $\mathrm{Beta}(0.5, 1)$ for Secukinumab group and prior $\mathrm{Beta}(11, 32)$ for placebo group:

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

## Plot several Beta distributions
plot(p, dbeta(p, 0.5, 1), ylab="density", type="l", col="red", lwd=2,
     ylim=c(0, 8), main="Plot of priors for both Secukinumab and placebo")
lines(p, dbeta(p, 11, 32), col="red", lwd=2, lty=2)

## Add legend
legend("topright", col="red", lwd=2, lty=c(1, 2), cex=.8,
       legend=c("Secukinumab prior: B(0.5, 1)", "Placebo prior: B(11, 32)"))
```




```r
## 2.5%, 50%, 97.5% quantiles of Beta(0.5, 1)
qbeta(c(0.025, 0.5, 0.975), 0.5, 1)
```

```
## [1] 0.000625 0.250000 0.950625
```

```r
## 2.5%, 50%, 97.5% quantiles of Beta(11, 32)
qbeta(c(0.025, 0.5, 0.975), 11, 32)
```

```
## [1] 0.1386101 0.2520003 0.3945024
```

We know that if $X \sim \text{Beta}(\alpha, \beta)$, then the expectation is given by:

$$\mathbb{E}X = \frac{\alpha}{\alpha + \beta}$$

We can apply this formula to obtain the prior means:

- Prior mean for $\text{Beta}(0.5, 1)$: 0.3334
- Prior mean for $\text{Beta}(11, 32)$: 0.2558

Table 2: Summary statistics of the prior distributions

Group	Prior	Mean	Median	Equi-tailed 95% interval
Secukinumab	$B(0.5, 1)$	0.3334	0.2500	$(0.0006, 0.9506)$
Placebo	$B(11, 32)$	0.2558	0.2520	$(0.1386, 0.3945)$

1(c) & 1(d) Prior distribution:

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

Posterior distribution:

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

Since y_i is a binary variable and only takes 0 or 1, $\bar{y}$ corresponds to the proportion of the responders (i.e. response rate).

- $n\bar{y}$ corresponds to the number of responders (i.e. $\sum_{i=1}^{n} \mathbf{1}[y_i = 1]$)
- $n - n\bar{y}$ corresponds to the number of non-responders (i.e. $n - \sum_{i=1}^{n} \mathbf{1}[y_i = 1]$)

Table 3: Prior vs. Posterior

Group	Prior Beta(α, β)	# of Subjects n	# of Responders x	Posterior Beta($\alpha + x, \beta + n - x$)
Secukinumab	Beta(0.5, 1)	23	14	Beta(14.5, 10)
Placebo	Beta(11, 32)	6	1	Beta(12, 37)

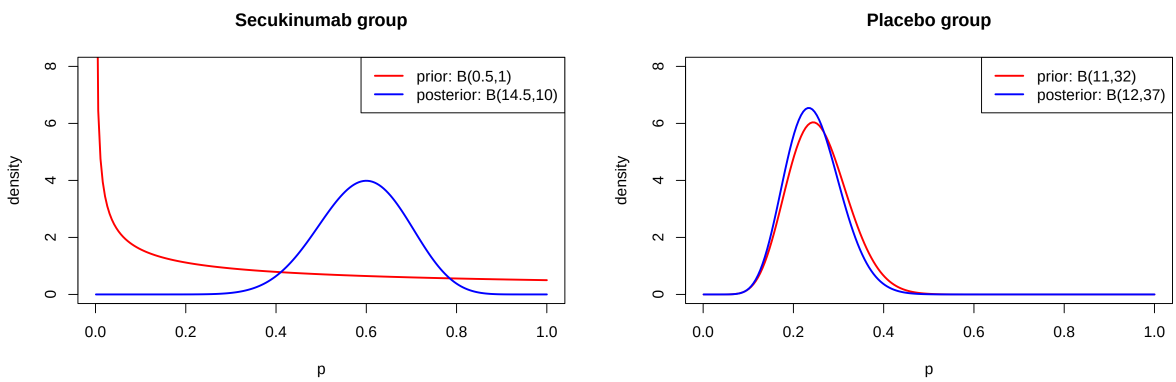
```r
p <- seq(1e-3,1, length=200)

plot(p, dbeta(p, 0.5, 1), type="l", ylab="density", col="red",
     lwd=2, ylim=c(0,8), main="Secukinumab group")
lines(p, dbeta(p, 14.5,10), ylab="density", col="blue", lwd=2)

legend("topright", legend=c("prior: B(0.5,1)", "posterior: B(14.5,10)"),
       col=c("red", "blue"), lwd=2)

plot(p, dbeta(p, 11, 32), type="l", ylab="density", col="red",
     lwd=2, ylim=c(0, 8), main="Placebo group")
lines(p, dbeta(p, 12, 37), ylab="density", col="blue", lwd=2)

legend("topright", legend=c("prior: B(11,32)", "posterior: B(12,37)"),
       col=c("red", "blue"), lwd=2)
```




```r
# Secukinumab group
## 2.5%, 50%, 97.5% quantiles of Beta(14.5, 10)
qbeta(c(0.025, 0.5, 0.975), 14.5, 10)
```

```
## [1] 0.3958401 0.5943750 0.7736253
```

```r
# Placebo group
## 2.5%, 50%, 97.5% quantiles of Beta(12, 37)
qbeta(c(0.025, 0.5, 0.975), 12, 37)
```

```
## [1] 0.1363723 0.2414054 0.3731202
```

Interpretation of CrIs:

- In Secukinumab group, the posterior probability of the response rate lies between 0.3958 and 0.7736 with probability 95%, when a Beta$(0.5, 1)$ prior is assumed.
- In Placebo group, the posterior probability of the response rate lies between 0.1364 and 0.3731 with probability 95%, when a Beta$(11, 32)$ prior is assumed.

Table 4: Summary statistics of the posterior distributions

Group	Posterior	Mean	Median	95% CrI
Secukinumab	$B(14.5, 10)$	0.5918	0.5944	$(0.3958, 0.7736)$
Placebo	$B(12, 37)$	0.2449	0.2414	$(0.1364, 0.3731)$

Exercise 2 (Individual project (Part 2B))

2(a) Beta distribution:

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

Density function of Beta distribution:

$$f(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1}$$

Expectation of X:

$$\begin{aligned}
\mathbb{E}X &= \int_0^1 x f(x) \mathrm{d}x \\
&= \int_0^1 x \frac{1}{B(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1} \mathrm{d}x \\
&= \frac{1}{B(\alpha, \beta)} \int_0^1 x^{\alpha}(1-x)^{\beta-1} \mathrm{d}x \\
&= \frac{B(\alpha+1, \beta)}{B(\alpha, \beta)} \\
&= \frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+\beta+1)} \cdot \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \\
&= \frac{\alpha!(\beta-1)!}{(\alpha+\beta)!} \cdot \frac{(\alpha+\beta-1)!}{(\alpha-1)!(\beta-1)!} \\
&= \frac{\alpha}{\alpha+\beta}
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}[X^2] &= \int_0^1 x^2 f(x) \mathrm{d}x \\
&= \int_0^1 x^2 \frac{1}{B(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1} \mathrm{d}x \\
&= \frac{1}{B(\alpha, \beta)} \int_0^1 x^{\alpha+1}(1-x)^{\beta-1} \mathrm{d}x \\
&= \frac{B(\alpha+2, \beta)}{B(\alpha, \beta)} \\
&= \frac{\Gamma(\alpha+2)\Gamma(\beta)}{\Gamma(\alpha+\beta+2)} \cdot \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \\
&= \frac{(\alpha+1)!(\beta-1)!}{(\alpha+\beta+1)!} \cdot \frac{(\alpha+\beta-1)!}{(\alpha-1)!(\beta-1)!} \\
&= \frac{(\alpha+1)\alpha}{(\alpha+\beta+1)(\alpha+\beta)}
\end{aligned}$$

$$
\begin{aligned}
\mathrm{Var}X &= \mathbb{E}\left[[X - \mathbb{E}X]^2\right] \\
&= \mathbb{E}[X^2] - [\mathbb{E}X]^2 \\
&= \frac{(\alpha+1)\alpha}{(\alpha+\beta+1)(\alpha+\beta)} - \left(\frac{\alpha}{\alpha+\beta}\right)^2 \\
&= \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}
\end{aligned}
$$

We know that:

$$
\begin{cases}
\mu = \frac{\alpha}{\alpha+\beta} \\
\sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}
\end{cases}
$$

$$
\mu = \frac{\alpha}{\alpha+\beta} \implies \beta = \frac{\alpha}{\mu} - \alpha
$$

Plug $\beta = \frac{\alpha}{\mu} - \alpha$ into $\sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$:

$$
\begin{aligned}
\sigma^2 &= \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \\
&= \frac{\alpha\left(\frac{\alpha}{\mu} - \alpha\right)}{\left(\frac{\alpha}{\mu}\right)^2\left(\frac{\alpha}{\mu}+1\right)} \\
&= \frac{\mu^2(1-\mu)}{\alpha+\mu}
\end{aligned}
$$

Rearranging above equation yields:

$$
\begin{cases}
\alpha = \left(\frac{1-\mu}{\sigma^2} - \frac{1}{\mu}\right)\mu^2 \\
\beta = \alpha\left(\frac{1}{\mu} - 1\right)
\end{cases}
$$

2(b) Implementation with a function in `R`:

- Input: sample mean and sample variance of a Beta distribution
- Output: α and β shape parameters of the Beta distribution

```
## Define a function
## Args: mean, var
estBetaParams <- function(mean, var) {
  alpha <- ((1 - mean) / var - 1 /mean) * mean ^ 2
  beta <- alpha * (1 / mean - 1)
  return(params = c(alpha=alpha, beta=beta))
}
```

2(c) Input:

- mean = 0.255814
- variance = 0.004326663

```
params <- estBetaParams(0.255814, 0.004326663); params
```

```
## alpha  beta
##    11    32
```

The resulting values of α and β parameters of the Beta distribution are 11 and 32 respectively.