1 Derivation of posterior mean of Beta distribution

- The posterior mode is $Beta(\bar{\alpha}, \bar{\beta})$
- The posterior mean is

$$
\begin{aligned}
\bar{\theta} &= \int \theta p(\theta|y) d\theta \\
&= \int_0^1 \theta \frac{1}{B(\bar{\alpha}, \bar{\beta})} \theta^{\bar{\alpha}-1} (1-\theta)^{\bar{\beta}-1} d\theta \\
&= \frac{1}{B(\bar{\alpha}, \bar{\beta})} \int_0^1 \theta^{\bar{\alpha}} (1-\theta)^{\bar{\beta}-1} d\theta \\
&= \frac{B(\bar{\alpha}+1, \bar{\beta})}{B(\bar{\alpha}, \bar{\beta})}
\end{aligned}
$$

Note that $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$ and $\Gamma(x) = (x-1)!$

Thus, we find that:

$$
\begin{aligned}
\bar{\theta} &= \frac{B(\bar{\alpha}+1, \bar{\beta})}{B(\bar{\alpha}, \bar{\beta})} \\
&= \frac{\Gamma(\bar{\alpha}+1)\Gamma(\bar{\beta})}{\Gamma(\bar{\alpha}+1+\bar{\beta})} \frac{\Gamma(\bar{\alpha}+\bar{\beta})}{\Gamma(\bar{\alpha})\Gamma(\bar{\beta})} \\
&= \frac{\Gamma(\bar{\alpha}+1)}{\Gamma(\bar{\alpha})} \frac{\Gamma(\bar{\alpha}+\bar{\beta})}{\Gamma(\bar{\alpha}+1+\bar{\beta})} \\
&= \frac{\bar{\alpha}!}{(\bar{\alpha}-1)!} \frac{(\bar{\alpha}+\bar{\beta}-1)!}{(\bar{\alpha}+\bar{\beta})!} \\
&= \frac{\bar{\alpha}}{\bar{\alpha}+\bar{\beta}}
\end{aligned}
$$

2 Posterior median in R

The posterior median can be obtained in R for different distributions, e.g.

- $qbeta(p = 0.5, shape1 = \alpha, shape2 = \beta)$
- $qgamma(p = 0.5, shape = \alpha, rate = \beta)$
- $dnorm(p = 0.5, mean = \mu, sd = \sigma)$

3 Calculate Credibility Interval: R Code

In this section, I will show how to calculate the equal tail interval and HPD interval for the setting of a beta posterior $Beta(\alpha = 19, \beta = 133)$

3.1 Equal Tail Interval

```
# Set the posterior parameters of the beta distribution
alpha<-19
beta<-133
# Calculate the 2.5\% and 97.5\% quantiles of the beta distribution
cimin<-qbeta(p=0.025,shape1=alpha,shape2=beta)
cimax<-qbeta(p=0.975,shape1=alpha,shape2=beta)
# The equal tail credibility interval is:
cimin;cimax
```

3.2 HPD Interval

```
# Set the posterior parameters of the beta distribution
alpha<-19
beta<-133
# Define function f(a) that calculates (f(a)-f(b))^2,
# where b is such that F(a)-F(b)=0.95
f <- function(a,p,q){
   b<-qbeta(pbeta(a,p,q)+0.95,p,q)
   (dbeta(a,p,q)-dbeta(b,p,q))^2
   }
# Minimise the function f() with respect to a (indeed, we want f(a)=f(b))
hpdmin <- optimize(f,lower=0,upper=qbeta(p=0.05,shape1=alpha,shape2=beta),
```

```
       p=alpha,q=beta)$minimum
#  Define b such that F(a)-F(b)=0.95
hpdmax <- qbeta(p=pbeta(hpdmin,alpha,beta)+0.95,shape1=alpha,shape2=beta)
#  The HPD interval is:
hpdmin;hpdmax
```

4 Reject/Accept Method: R program

Suppose that a study based on $N = 200$ individuals resulted in $x = 25$ successes and $N-x = 175$ failures. Assume now that the prior information is a normal prior for the success rate θ is with mean 0.3 and standard deviation 0.10. Sample from the posterior distribution using the rejection/acceptance method.

In this exercise, the exact posterior distribution cannot be derived analytically anymore.

- Take a sample $\tilde{\theta}$ from the prior distribution. This is the envelope density q.

```
qtheta<-rnorm(n=10000,mean=0.3,sd=0.10)
```

- Calculate the constant A such that $p(\theta|x) \leq Aq(\theta)$. We can take $A = L(\widehat{\theta}|x)$, with $\widehat{\theta} = 25/200$ the MLE of θ. This is the envelope constant A.

```
A<-dbinom(x=25,size=200,prob=25/200)
```

- Draw independently a sample u from the uniform distribution $U[0, 1]$

```
u<-runif(n=10000,min=0,max=1)
```

- Accept when $u \leq \frac{p(\theta|x)}{Aq(\theta)} = \frac{L(\tilde{\theta}|x)}{A}$

```
L<-dbinom(x=25,size=200,prob=qtheta)
accept<-(u<=L/A)
```

- We then make a histogram of the envelope distribution

```
hist(qtheta,class=50,col="blue",main="",freq=FALSE)
```

- And a histogram of the posterior distribution obtained from the rejection/acceptance sampling, this is by using only the accepted values of the envelope distribution:

```
hist(qtheta[accept],nclass=50,col="blue",main="",freq=FALSE)
```

In summary, this is the code:

```
# Rejection/acceptance method
qtheta<-rnorm(n=10000,mean=0.3,sd=0.10)
u<-runif(n=10000,min=0,max=1)
A<-dbinom(x=25,size=200,prob=25/200)
L<-dbinom(x=25,size=200,prob=qtheta)
accept<-(u<=L/A)
hist(qtheta,nclass=50,col="blue",main="",freq=FALSE,xlim=c(0,0.40))
theta<-seq(0,0.5,0.001)
lines(theta,dnorm(theta,mean=0.3,sd=0.10))
hist(qtheta[accept],nclass=50,col="blue",main="",freq=FALSE,xlim=c(0,0.40))
lines(theta,dnorm(theta,mean=0.3,sd=0.10))
```

The result is shown in Figure 1. The posterior mean based on this sample is 0.142.

Note: in some of the samples, the likelihood L cannot be calculated. The reason for this is that some of the sampled values for qtheta are smaller than 0. We just ignore these samples (only very few).

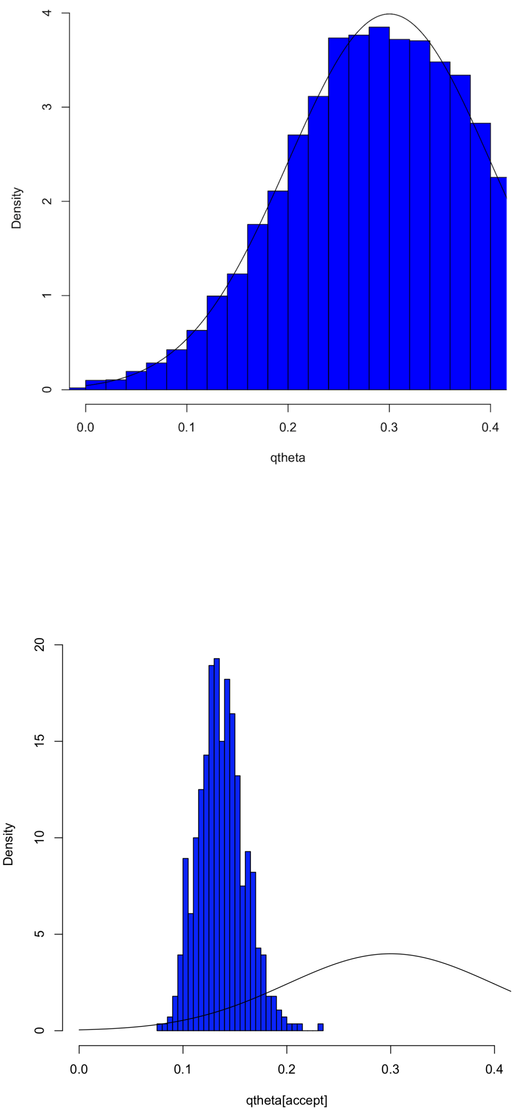

Figure 1: *Illustration of the rejection/acceptance sampling. Left: sample from the prior distribution (used as envelope density). Right: histogram of the posterior distribution, overlayed with the prior distribution (line)*

5 Weighted Resampling Method: R program

Suppose that a study based on $N = 200$ individuals resulted in $x = 25$ successes and $N - x = 175$ failures. Assume now that the prior information is a normal prior for the success rate θ is with mean 0.3 and standard deviation 0.10. Sample from the posterior distribution using the weighted resampling method.

The following steps are being taken:

- We first take a large sample $\tilde{\theta}$ from the prior distribution:

```
qtheta<-rnorm(n=10000,mean=0.3,sd=0.10)
```

- The weights $L(\tilde{\theta}_i|x)/\sum_i \left(L(\tilde{\theta}_i|x)\right)$ are being calculated:

```
L<-dbinom(x=25,size=200,prob=qtheta)
w<-L/sum(L)
```

- We re-sample each of the sampled values $\tilde{\theta}$ using the weights calculated from a multinomial distribution:

```
freq<-rmultinom(n=1,size=10000,prob=w)
rtheta<-rep(qtheta,freq)
```

- Both the prior distribution and posterior distribution are presented in a histogram:

```
hist(qtheta,nclass=50,col="blue",main="",freq=FALSE)
hist(rtheta,nclass=50,col="blue",main="",freq=FALSE)
lines(theta,y,type="l",lwd=2,ylab="Posterior density")
```

In summary, the code is as follows:

```
# Weighted resampling
qtheta<-rnorm(n=10000,mean=0.3,sd=0.10)
qtheta<-qtheta[qtheta>=0]

L<-dbinom(x=25,size=200,prob=qtheta)
w<-L/sum(L)
freq<-rmultinom(n=1,size=10000,prob=w)
rtheta<-rep(qtheta,freq)

hist(qtheta,nclass=50,col="blue",main="",freq=FALSE,xlim=c(0,0.40))
lines(theta,dnorm(theta,mean=0.3,sd=0.10))
hist(rtheta,nclass=50,col="blue",main="",freq=FALSE,xlim=c(0,0.40))
lines(theta,dnorm(theta,mean=0.3,sd=0.10))
```

The result is shown in Figure 2. The posterior mean based on this sample is 0.141.

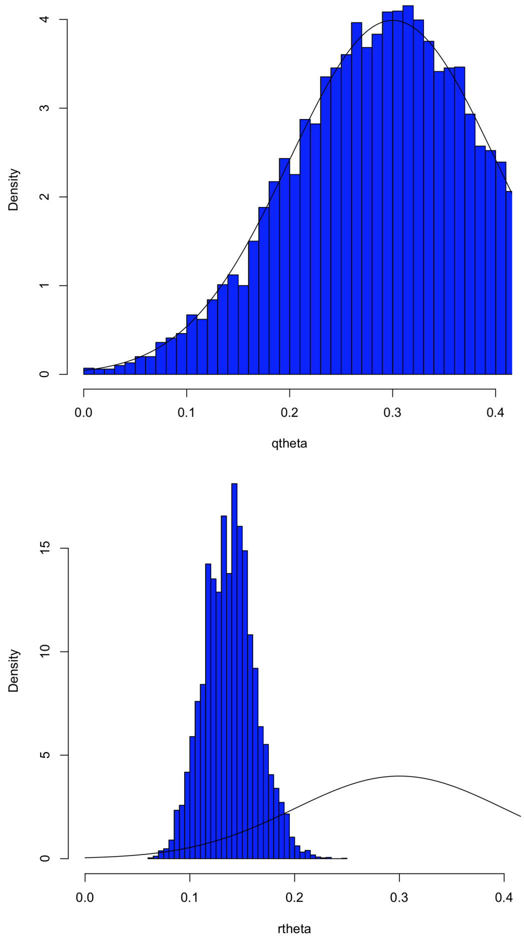

Figure 2: *Illustration of the resampling method. Left: sample from the prior distribution. Right: histogram of the posterior distribution, overlayed with the prior distribution (line)*

6 Sample from transformed parameter

Suppose that a study based on $N = 200$ individuals resulted in $x = 25$ successes and $N - x = 175$ failures. Assume now that the prior information is a normal prior for the success rate θ is with mean 0.3 and standard deviation 0.10. Sample from the posterior distribution of the odds $\theta/(1-\theta)$.

Based on the accept/reject or weighted resampling method, we have a sample `rtheta` of the θ parameter. A sample of the odds $\theta/(1-\theta)$ can be obtained by calculating the odds for every value in `rtheta`. This is obtained in R in the following way:

```
# Derive sample for odds
odds<-rtheta/(1-rtheta)
hist(odds,nclass=50,col="blue",main="")
title("Sampling from odds")
```

This is presented in Figure 3. This can be done based on any of the previous samples from the posterior distribution.

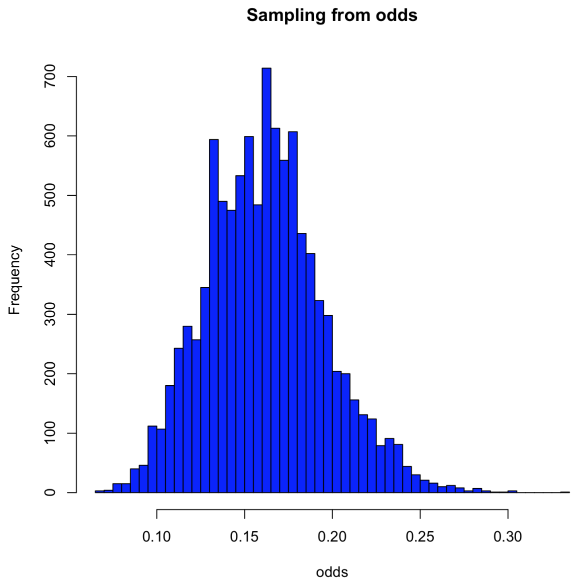

Figure 3: *Sampling from the odds*

7 Bayesian hypothesis testing

Assume that 30 patients with systolic hypertension are better of with treatment 1 than with treatment B. We assume that the likelihood is described by a binomial likelihood with parameter θ.

Calculate Bayes factor for testing the $H_0 : \theta \leq 0.5$ versus $H_a : \theta > 0.5$, for a continuous parameter θ. Assume that a prior $p(H_0) = p(H_a) = 0.5$

The Bayes factor is defined as $\frac{p(y|H_0)}{p(y|H_a)}$, so we need to calculate $p(y|H_0)$ and $p(y|H_0)$. Is is assumed that, under the null hypothesis, any value smaller than 0.5 is equally likely, thus $p(\theta|H_0) = U(0, 0.5)$, and thus $p(\theta|H_0) = 2$ if $\theta \leq 0.5$, and 0 otherwise.

$$\begin{aligned} p(y|H_0) &= \int_0^1 L(\theta|y)p(\theta|H_0)d\theta \\ &= \int_0^1 \binom{30}{21}\theta^{21}(1-\theta)^9 p(\theta|H_0)d\theta \end{aligned}$$

$$
\begin{aligned}
&= 2\int_0^{0.5} \binom{30}{21}\theta^{21}(1-\theta)^9 d\theta \\
&= 2\binom{30}{21}\int_0^{0.5} \theta^{21}(1-\theta)^9 d\theta \\
&= 2\binom{30}{21}B(22,10)\frac{1}{B(22,10)}\int_0^{0.5} \theta^{21}(1-\theta)^9 d\theta
\end{aligned}
$$

The blue part can be recognised as the distribution function of the beta-distribution, and is calculated in R with the function `pbeta(0.5,22,10)`, leading to:

$$
p(y|H_0) = 2\binom{30}{21}B(22,10)0.01472
$$

Similarly, $p(y|H_a)$ is calculated as

$$
\begin{aligned}
p(y|H_a) &= \int_0^1 L(\theta|y)p(\theta|H_0)d\theta \\
&= 2\int_{0.5}^{1} \binom{30}{21}\theta^{21}(1-\theta)^9 d\theta \\
&= 2\binom{30}{21}B(22,10)\frac{1}{B(22,10)}\int_{0.5}^{1} \theta^{21}(1-\theta)^9 d\theta \\
&= 2\binom{30}{21}B(22,10)\left(1-\frac{1}{B(22,10)}\int_0^{0.5} \theta^{21}(1-\theta)^9 d\theta\right) \\
&= 2\binom{30}{21}B(22,10)\left(1-0.01472\right)
\end{aligned}
$$

As a result, the BF is:

$$
\begin{aligned}
BF &= \frac{p(y|H_0)}{p(y|H_a)} \\
&= \frac{2\binom{30}{21}B(22,10)0.01472}{2\binom{30}{21}B(22,10)\left(1-0.01472\right)} \\
&= \frac{0.01472}{(1-0.01472)} \\
&= 0.0149
\end{aligned}
$$

As a result, there is substantial evidence to favor the alternative hypothesis over the null hypothesis!