Solutions to Practical 3
Using WinBUGS

Section 1: Estimating the sex ratio

A combined version of BUGS code for (a) and (b):

```
model
{
      y ~ dbin(theta, n)
      theta ~ dbeta(alpha, beta)
      ratio <- (1-theta)/theta
}
```

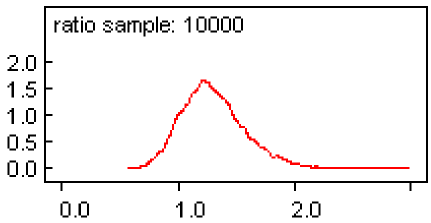
Data for (a):
```
list(n=98, y=43, alpha=1, beta=1)
```

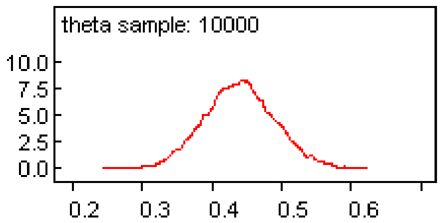
Data for (b):
```
list(n=98, y=43, alpha=48.5, beta=51.5)
```

Inits:
```
list(theta=0.5)
```

Results for (a):

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
ratio	1.302	0.2695	0.002608	0.8465	1.274	1.907	1	10000
theta	0.4401	0.04999	4.875E-4	0.344	0.4397	0.5416	1	10000

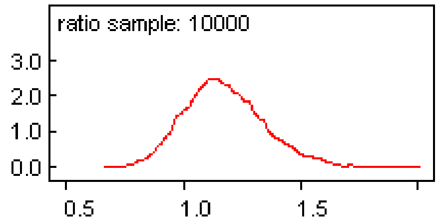


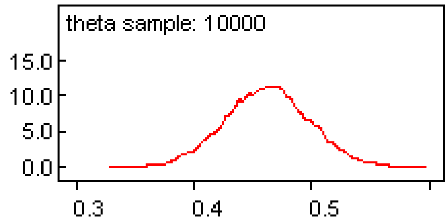


Parameter	*posterior mean*	*posterior SD*	*95% credible interval*
theta	0.44	0.05	0.34 – 0.54
ratio	1.30	0.27	0.84 – 1.91

Results for (b):

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
ratio	1.175	0.1705	0.001913	0.8772	1.162	1.55	1	10000
theta	0.4625	0.03569	3.985E-4	0.3922	0.4625	0.5328	1	10000





Parameter	*posterior mean*	*posterior SD*	*95% credible interval*
theta	0.46	0.04	0.39 – 0.53
ratio	1.18	0.17	0.88 – 1.55

(c) *non-conjugate prior:*

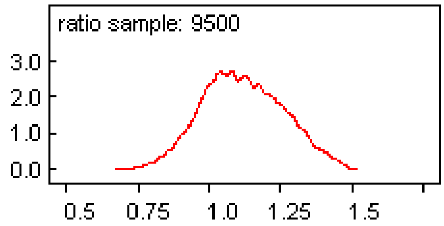
```
model
{
        y ~ dbin(theta, n)
        theta1 ~ dunif(0.2, 0.3)
        theta2 ~ dunif(0.2, 0.3)
        theta <- theta1+theta2
        ratio <- (1-theta)/theta
}
```

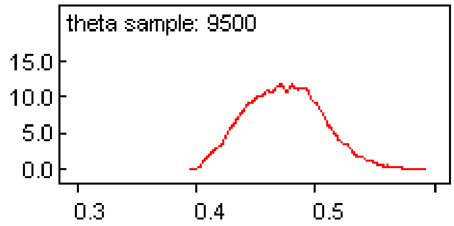
Data for (c):
```
list(n=98, y=43)
```

Inits:
```
list(theta1=0.21, theta2=0.29)
```

Results:

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
ratio	1.118	0.1417	0.001493	0.8537	1.112	1.399	501	9500
theta	0.4742	0.03187	3.427E-4	0.4168	0.4734	0.5395	501	9500





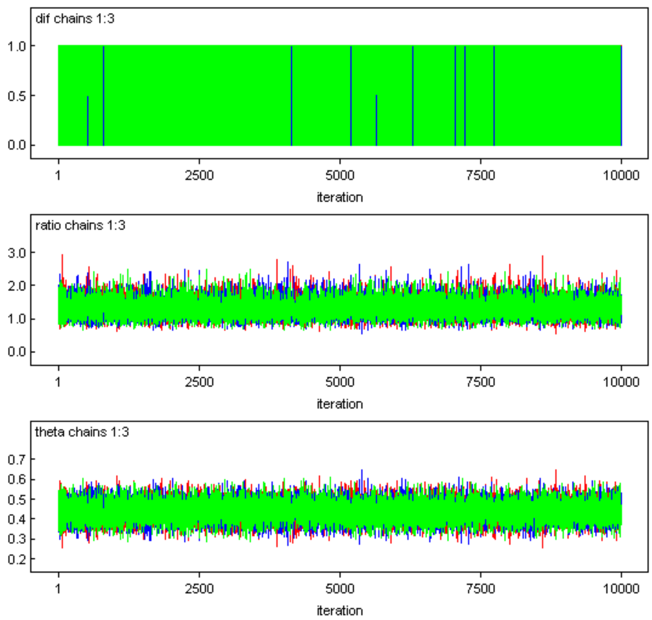
Parameter	*posterior mean*	*posterior SD*	*95% credible interval*
theta	0.47	0.03	0.42 – 0.54
ratio	1.12	0.14	0.85 – 1.40

(d) $\Pr\left(\theta < 0.485 \mid y\right)$

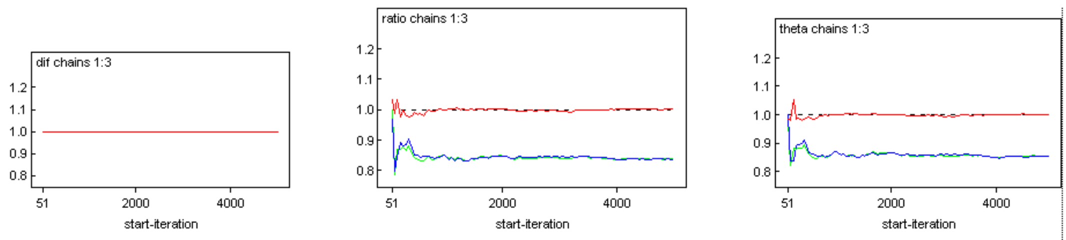
Prior	*posterior prob. that* $\theta < 0.485$
uniform	0.81
Beta(48,5, 51.5)	0.74
triangular	0.63

Section 2: MCMC Diagnostics for estimating the sex ratio

History



Bgr plots with y-axis adjusted



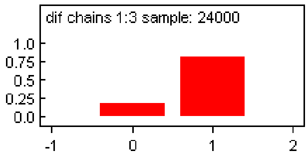
Chain appears to converge relatively rapidly. In the bgr plot, the pooled and within values only appear to be really steady after about 2000 runs so could select a burn-in of 2000.

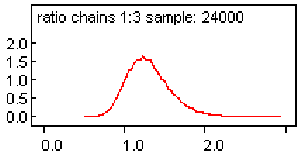
Using the values 2,001 to 10,000 to produce summary statistics we get:

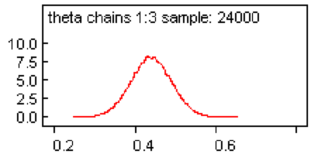
node	mean	sd	MC error	2.5%	median	97.5%	start	sample
dif	0.8166	0.387	0.002484	0.0	1.0	1.0	2001	24000
ratio	1.303	0.2686	0.001755	0.8608	1.275	1.914	2001	24000
theta	0.44	0.04958	3.216E-4	0.3432	0.4396	0.5374	2001	24000

Note that this uses 8,000 values from each chain giving 24,000 values in total.

MC error also less than 1% of the SD in each case.







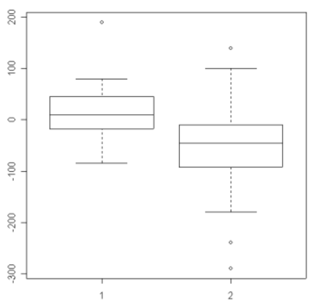
Section 3:

Infant weight gain

Means for the two groups: 18 -52

Sds for the two groups: 60 88

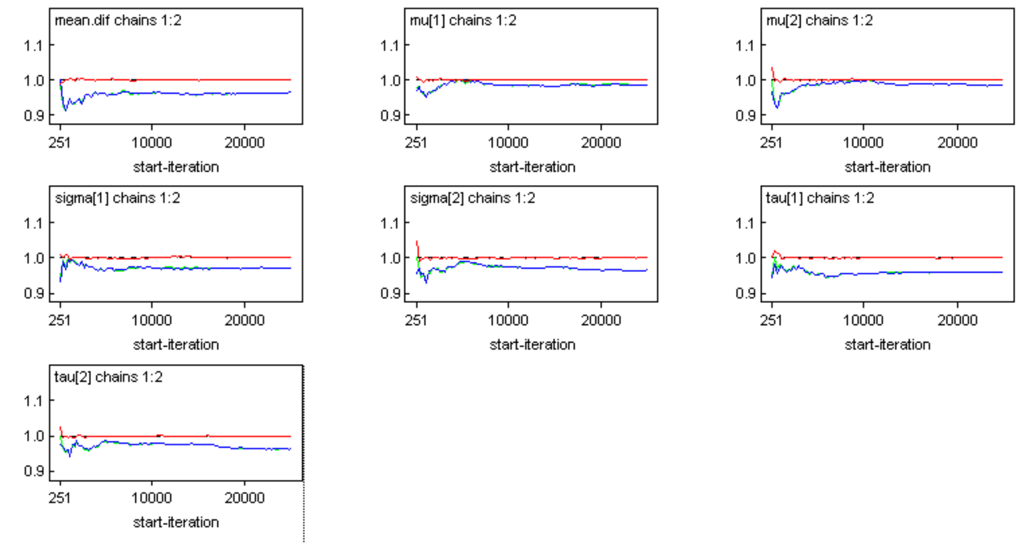
Boxplot of the weight change for the two groups



This shows clear differences in the weight gain between the two groups…

Here we just show the bgr plots after 50,000 iterations, as only after 20,000 do the pooled and within group variance for mu[1] and mu[2] appear to be stable although other variables appeared to have reached convergence earlier. So, summary statistics are calculated using runs 20,001 to 50,000 for the two chains.

95% credible interval for the mean difference is (-110.8, -28.9) suggesting that there is an increase in weight gain for the group exposed to the recorded sound of the mother’s heartbeat.



node	mean	sd	MC error	2.5%	median	97.5%	start	sample
mean.dif	-69.84	20.79	0.09702	-110.8	-69.85	-28.9	30001	40000
mu[1]	17.87	14.14	0.0682	-10.0	17.91	45.9	30001	40000
mu[2]	-51.97	15.17	0.0754	-82.12	-51.92	-22.23	30001	40000
sigma[1]	62.56	10.72	0.0587	45.71	61.2	87.51	30001	40000
sigma[2]	90.39	11.12	0.06128	71.76	89.27	115.2	30001	40000
tau[1]	2.771E-4	8.969E-5	4.812E-7	1.306E-4	2.67E-4	4.788E-4	30001	40000
tau[2]	1.278E-4	3.048E-5	1.694E-7	7.538E-5	1.255E-4	1.942E-4	30001	40000

Log-normal survival times

```
model
{
            # mean for prior of mu  -  enter value
            mu0 <- log(30)
            # parameters of prior for tau  -  enter value
            beta <- 1.5
            alpha <- 2*beta
            # prior for tau  -  enter distribution
            tau ~ dgamma(alpha, beta)
            # prior for mu, given tau  -  enter distribution
            mu ~ dnorm(mu0,tau)
           for(i in 1:N) {
                  y[i] <- log(survtime[i])
                  y[i] ~ dnorm(mu, tau)
                  }
            sigma <- 1/sqrt(tau)
            # predicted value  -  enter distribution
            y.new ~ dnorm(mu, tau)
            # greater than log(150)?
            y.dif <- step(y.new - log(150))
}
```

Data:

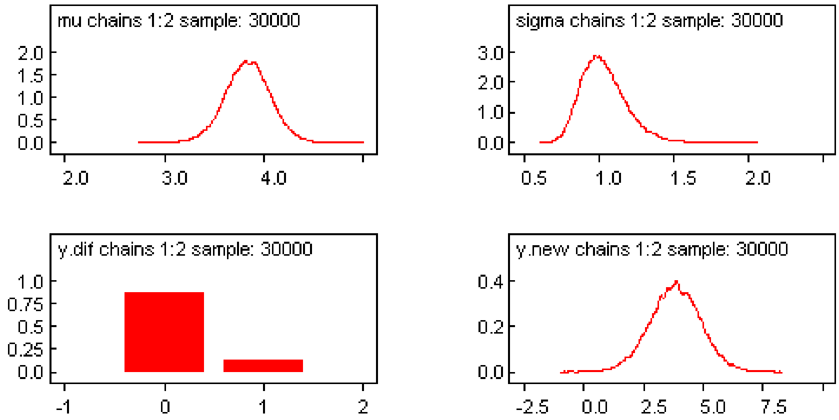
```
list(survtime=c(25, 45, 238, 194, 16, 23, 30,
                         16, 22, 123, 51, 412, 45, 162,
                         14, 72, 5, 43, 45, 91),
                          N=20)
```

Inits:

```
list(mu=0, tau=1, y.new=0)
list(mu=1, tau=2, y.new=1)
```

Results:

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
mu	3.84	0.2278	0.001261	3.387	3.841	4.288	15001	30000
sigma	1.034	0.151	9.638E-4	0.788	1.017	1.378	15001	30000
y.dif	0.1302	0.3366	0.00193	0.0	0.0	1.0	15001	30000
y.new	3.837	1.071	0.006668	1.713	3.841	5.949	15001	30000



30,000 iterations for 2 chains– converged after 15,000