

1. The site also contains an R script fitting a widely-used Bayesian model, in Stan or JAGS, using data from a blood pressure GWAS meta-analysis, combining linear regression estimates across 6 studies. It assumes the sampling model for the observed $\hat{\beta}_i$ is $N(\beta_i, \sigma_i^2)$ with known σ_i^2 . The prior for the β_i is specified hierarchically, and (as the data have been suitably scaled) can be written as

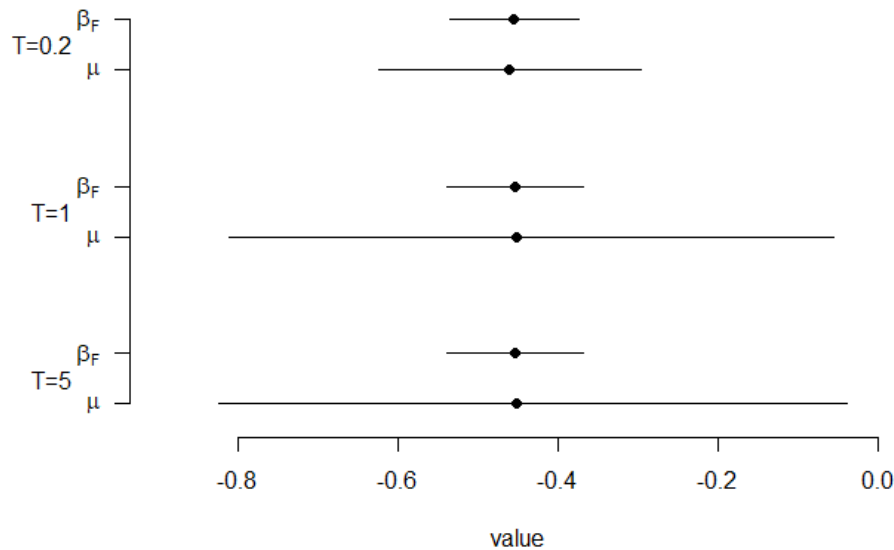
$$\beta_i \sim N(\mu, \tau^2) \text{ for each } i,$$

$$\mu \sim N(0,1)$$

$$\tau \sim U(0, T), \text{ i.e. a flat distribution between 0 and } T$$

For $T=1/5, 1$ and 5 , what is the posterior for $\beta_F = \sum_i \beta_i / \sigma_i^2 / (\sum_i 1/\sigma_i^2)$? For the same values of T , what is the posterior for μ , the mean of the random effects distribution?

Adapting the code in the script by changing T , the posterior medians and 95% credible intervals for these two parameters are as follows:



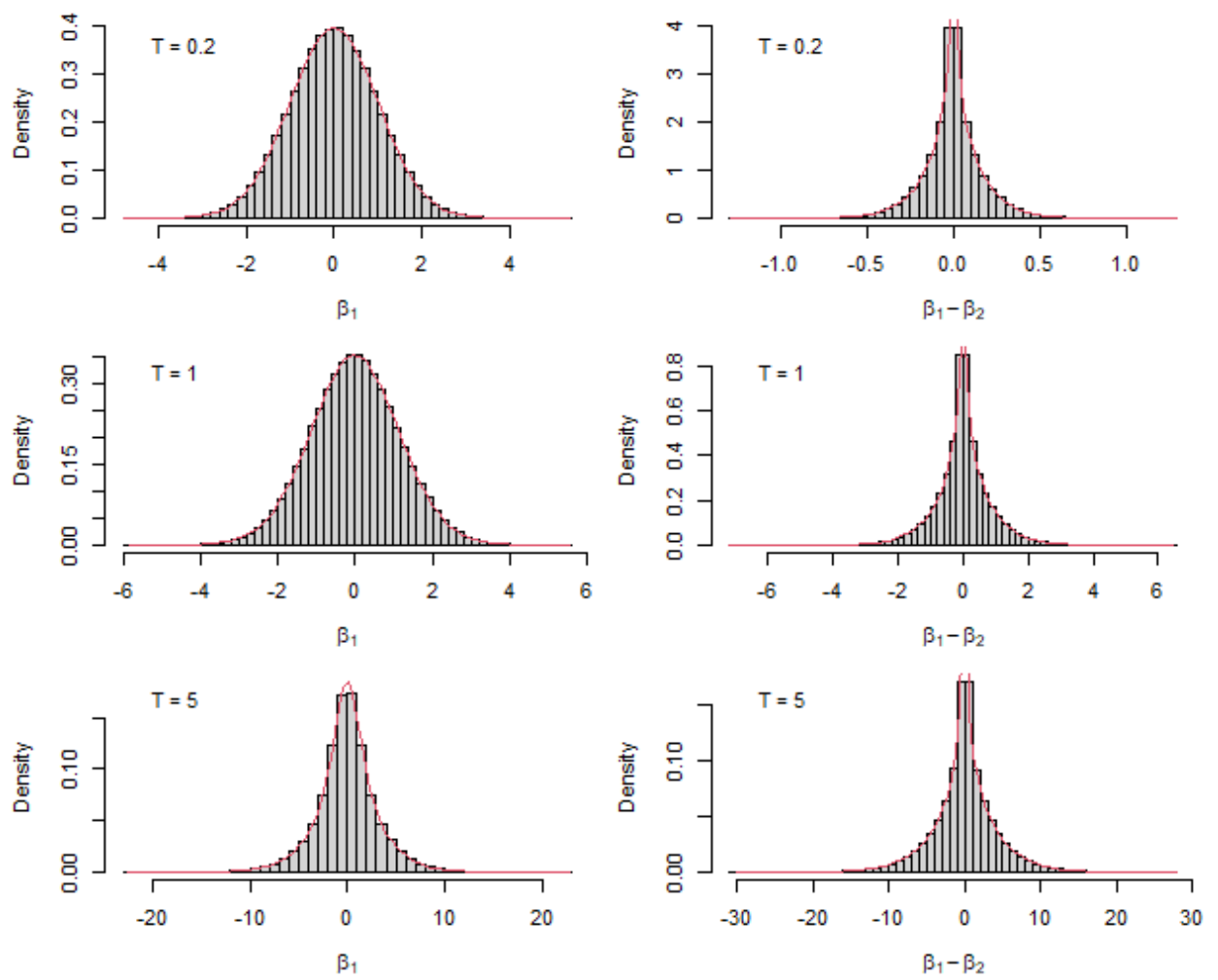
Clearly, while the point estimates are always similar, the spread of the posterior for β_F is much less sensitive to the prior than that for μ . (The posterior for β_F also closely agrees with a “fixed-effects” meta-analysis, as expected.)

2. Using the same prior as in Q1, what is the marginal prior for any individual β_i ? For the same values of T what is the marginal prior for any $\beta_i - \beta_j$, i.e. the difference between two study parameters? Do any differences surprise you?

Hint: use the Monte Carlo method in base R: take a large sample from the prior, and make a histogram with many bars.

Sampling from the priors, we get the following histograms (the exact prior is also superimposed, though this requires more complex calculation). Some users are surprised by the notable “pointiness” of the prior on the differences between parameters: the small values of T force this

difference to be close to zero, as the effects are quite homogeneous. The prior on each β_i retains a familiar bell-shaped curve, for most values of T.



Code

```
# session 8 exercise, part 1

study.names <- c("AGES", "ARIC", "CHS", "FHS", "RS", "RES")
betahat <- c(-0.37, -0.23, -0.31, -0.79, -0.13, -0.96)
stderr <- c(0.26, 0.16, 0.33, 0.16, 0.25, 0.39)
n <- c(3219, 8047, 3277, 8096, 4737, 1760)

#install.packages("metafor")
library("metafor")
rma1 <- rma.uni(yi=betahat, vi=stderr^2, method="FE")
rma2 <- rma.uni(yi=betahat, vi=stderr^2, method="DL")

# Bayes meta in Stan

cat(file="metaanalysis.stan", "
data {
  int p; //the number of studies
  vector[p] betahat; // study-specific estimates
  vector[p] sigma; // study-specific standard errors
  real T; // upper bound of uniform prior on tau
}
parameters {
  vector[p] beta; // study-specific parameters
  real mu; // mean of the mixing distribution for beta[i]'s
  real tau; // SD of the mixing distribution
}
transformed parameters {
  vector[p] invsigma2;
  vector[p] betafpart;
  real betaf;
  for(i in 1:p){ invsigma2[i] = pow(sigma[i],-2);}
  for(i in 1:p){ betafpart[i] = beta[i]*invsigma2[i];}
  betaf = sum(betafpart[1:p])/sum(invsigma2[1:p]); // precision-
weighted mean of study-specific effects
}
model {
  for(i in 1:p){
    betahat[i] ~ normal(beta[i], sigma[i]);
    beta[i] ~ normal(mu, tau);
  }
  mu ~ normal(0,1);
  tau ~ uniform(0,T);
}
")

# do the MCMC, store the results
```

```

library("rstan")
stan.meta1 <- stan(file = "metaanalysis.stan",
data = list(betahat=betahat, sigma=stderr, p=6, T=1),
iter = 100000, chains = 1, pars=c("betaf", "mu"))

print(stan.meta1)
stan.meta0.2 <- stan(file = "metaanalysis.stan",
data = list(betahat=betahat, sigma=stderr, p=6, T=0.2),
iter = 100000, chains = 1, pars=c("betaf", "mu"))
stan.meta5 <- stan(file = "metaanalysis.stan",
data = list(betahat=betahat, sigma=stderr, p=6, T=5),
iter = 100000, chains = 1, pars=c("betaf", "mu"))

print(stan.meta1)
print(stan.meta0.2)
print(stan.meta5)
#plot(stan.meta)

do.one <- function(mmm){
  mypos <- 6:1 + c(-1,1,-1,1,-1,1)*0.2
  plot(0,0,xlim=c(-0.9,0), ylim=c(1, 6), axes=FALSE, xlab="value",
ylab="")
  segments(mmm[,2], mypos, mmm[,3], mypos)
  points(y=mypos, x=mmm[,1], pch=19)
  axis(side=1)
  axis(side=2, las=1, at=mypos, c(
    expression(beta[F]), expression(mu),
    expression(beta[F]), expression(mu),
    expression(beta[F]), expression(mu)
  ))
  mtext(side=2, las=1, line=2, at=c(5.5, 3.5, 1.5),
c("T=0.2", "T=1", "T=5"))
}

par(mar=c(4,6,0,0)+0.2)
do.one( rbind(
  summary(stan.meta0.2)$summary[1:2,c(6,4,8)],
  summary(stan.meta1)$summary[1:2,c(6,4,8)],
  summary(stan.meta5)$summary[1:2,c(6,4,8)])
)

round(unlist(rma1[2:7]),2)
round(unlist(rma2[2:7]),2)
forest(rma1, slab=study.names)
forest(rma2, slab=study.names)

# monte carlo for Q2

```

```

par(mfrow=c(3,2))
par(mar=c(4,4,1,1))
B <- 1E6
  set.seed(4)
do.one <- function(T){
  mu <- rnorm(B, 0, 1)
  tau <- runif(B, 0, T)
  beta1 <- rnorm(B, mean=mu, sd=tau)
  beta2 <- rnorm(B, mean=mu, sd=tau)
  v1 <- function(x){
    integrate(function(t){dnorm(x,0,sqrt(1+t^2))*dunif(t,0,T)},0,T)$v
alue }
  v1 <- Vectorize(v1)
  hist(beta1, xlab=expression(beta[1]), main="", freq=FALSE,
ylim=c(0,v1(0)),n=60)
#   curve(dnorm(x,0,sqrt(1+T^2)), min(beta1), max(beta1), add=TRUE,
col=4)
  curve(v1, min(beta1), max(beta1), add=TRUE, col=2)
  legend("topleft", bty="n", lty=0, paste("T =",T))

  v2 <- function(x){
    integrate(function(t){dnorm(x,0,sqrt(2*t^2))*dunif(t,0,T)},0,T)$v
alue }
  v2 <- Vectorize(v2)
  hist(beta1-beta2, xlab=expression(beta[1]-beta[2]), main="",
freq=FALSE, n=60)
#   curve(dnorm(x,0,sqrt(2*T^2)), min(beta1-beta2), max(beta1-beta2),
add=TRUE, col=4)
  curve(v2, min(beta1-beta2), max(beta1-beta2), add=TRUE, col=2)
  legend("topleft", bty="n", lty=0, paste("T =",T))
}
do.one(1/5)
do.one(1)
do.one(5)

```