Homework #8 Solutions STAT 506 Spring 2015

The nlme package in R contains datasets MathAchieve and MathAchSchool. Our goal is to model the individual and school level effects.

1. We seem to agree that a good model includes:
   Gender, Minority Status (yes/no), , SES, and SES by Minority interaction as main effects. (SES is socio-economic status of the individual student.)

	Estimate	Std. Error	t value
(Intercept)	14.08	0.18	76.53
Minority	-3.21	0.26	-12.28
SES	2.40	0.13	19.18
Sex	-1.23	0.16	-7.58
Minority:SES	-1.13	0.22	-5.14

2. First look at school effects. Available predictors are:

   Size a numeric vector giving the number of students in the school

   Sector a factor with levels Public Catholic

   PRACAD a numeric vector giving the percentage of students on the academic track

   DISCLIM a numeric vector measuring the discrimination climate

   HIMINTY where 1 means high proportion of minorities, 0 otherwise

   MEANSES a numeric vector giving the mean SES score.

	Estimate	Std. Error	t value
(Intercept)	-2.25	0.33	-6.88
size	0.57	0.15	3.81
sectorCatholic	0.91	0.24	3.76
pracad	2.38	0.55	4.34
meanses	0.60	0.27	2.21

Table 1: n = 160 rank = 5 resid sd = 1.1 R-Squared = 0.5

	Estimate	Std. Error	t value
(Intercept)	-0.26	0.10	-2.68
sectorCatholic	0.57	0.17	3.32
disclim	-0.33	0.09	-3.80
meanses	0.67	0.16	4.29

Table 2: n = 160 rank = 4 resid sd = 0.7 R-Squared = 0.5

I used AIC's and conditional t-tests to reduce a model using all school – level predictors. to these terms:
Intercept modeled by size, sector, pracad, and meanses
Minority slope modeled by sector, disclim, and meanses.

In addition, I fit the 2 stage model as one, using the above terms as fixed effects. This does change estimates quite a bit.

	Estimate	Std. Error	t value
(Intercept)	11.09	0.49	22.54
minority	-3.87	0.38	-10.05
ses	2.23	0.13	17.03
sex	-1.23	0.16	-7.81
size	0.87	0.22	3.98
sectorCatholic	0.83	0.35	2.34
pracad	3.23	0.81	4.00
meanses	0.92	0.47	1.96
minority:ses	-1.20	0.23	-5.15
minority:sectorCatholic	1.28	0.61	2.10
minority:disclim	-0.82	0.29	-2.85
minority:meanses	0.31	0.58	0.53

	Std.Dev	Corr
Intercept	1.15	
Minority	0.95	-0.04
sigma	5.96	

3. Pick a good model from your work above and fit it using jags. Provide enough information to show that it has converged, and compare output to that from the fit above.

   I started with just fixed effects.

   The jags "fixed effects" fit agrees very closely with a fit in ‘lm‘. Next I added random effects for intercept and minority effect across schools. Fitting a correlation really slowed down convergence, so I left it out.

   Intercept estimate (and its SE) and slope on Minority (and its SE) are half what lmer gives.

	Mean	SD	Rhat	Estimate	Std. Error	t value	effects
beta0	5.133	0.361	1.026	11.086	0.492	22.538	(Intercept)
beta1	-1.514	0.171	1.033	-3.866	0.385	-10.054	minority
beta2	1.536	0.178	1.009	2.233	0.131	17.034	ses
beta3	-1.212	0.165	1.003	-1.231	0.158	-7.806	sex
beta4	0.079	0.307	1.026	0.873	0.219	3.978	size
beta5	0.610	0.626	1.014	0.828	0.355	2.335	sectorCatholic
beta6	-1.688	1.193	1.002	3.232	0.809	3.996	pracad
beta7	0.583	0.698	1.006	0.919	0.468	1.961	meanses
beta8	-0.794	0.211	1.004	-1.205	0.234	-5.148	minority:ses
beta9	-0.127	0.562	1.022	1.281	0.61	2.1	minority:sectorCatholic
beta10	-0.044	0.234	1.001	-0.818	0.287	-2.85	minority:disclim
beta11	0.268	0.567	1.004	0.305	0.58	0.527	minority:meanses
sigma.b1	1.964	0.147	1.003	1.151			Intercept SD
sigma.b2	0.466	0.275	1.203	0.945	-0.041		Minority SD / Corr
sigma.y	5.985	0.049	1.000	5.958			Resid SD

Either way, convergence took quite a while. The estimates for the two fits seem to agree on the gender effect and the underlying residual standard deviation.

4. Explanation:

   (a) Describe the data.

   (b) What have we learned about effects of different variables on kids scores (individual level)?
   There are strong effects for minority, gender, SES, and minority by SES interaction (less improvement for minorities). Explain those.

   (c) How much do school variables impact kids scores?
   Effects are not large. School to school SD is only 1/5th to 1/3rd that of the residual, and all of the school-level effects have small t-values in the JAGs output. If you are discussing the lmer output, then size, percent Academic, Minority by sector interaction have t-ratios over 2 in magnitude.

   (d) Scope of inference?
   Non-causal inference to just this sample.

R Code

```
data(MathAchieve,  package="nlme")
data(MathAchSchool,package="nlme")
options(show.signif.stars = FALSE)
require(lme4,quietly=TRUE)
require(R2jags, quietly=TRUE)
MathAchieve$Minority <- as.numeric(MathAchieve$Minority ) -1
MathAchieve$Sex <- as.numeric(MathAchieve$Sex ) -1
math.fit1 <- lmer(MathAch ~ Minority * SES + Sex + (1|School), data = MathAchieve)
math.fit2 <- lmer(MathAch ~ Minority * SES + Sex + (1 + SES|School), data = MathAchieve)
##anova(math.fit1,math.fit2)
math.fit3 <- lmer(MathAch ~ Minority * SES + Sex + (1 + Sex|School), data = MathAchieve)
##anova(math.fit1,math.fit3)
math.fit4 <- lmer(MathAch ~ Minority * SES + Sex + (1 + Minority|School), data = MathAchieve)
##anova(math.fit1,math.fit4)
math.fit5 <- lmer(MathAch ~ Minority * SES + Sex + (1 + Minority + Sex|School), data = MathAchieve)
##anova(math.fit1,math.fit4, math.fit5)
xtable(summary(math.fit4)$coefficients)
```

```
## use math.fit4
RE4s <- ranef(math.fit4)$School
colnames(RE4s) <- c("intcpt","minorty")
RE4s$School <- rownames(RE4s)
schoolData<- merge(MathAchSchool, RE4s)
names(schoolData) <- tolower(names(schoolData))
schoolData$size <- schoolData$size/1000

intModel2 <- update(intModel1 <- lm(intcpt ~ ., data = schoolData[,2:8]), .~.-himinty)
intModel3 <- update(intModel2, .~. - disclim)
##anova(intModel1,intModel2,intModel3)
display.xtable(intModel3)

minModel2 <- update(minModel1 <- lm(minorty ~ ., data = schoolData[,c(2:7,9)]), .~.-himinty)
minModel3 <- update(minModel2, .~. - size)
minModel4 <- update(minModel3, .~. - pracad)
##anova(minModel1,minModel4)
display.xtable(minModel4)
```

```r
## one step fit
bigAchieve <- merge(MathAchieve, MathAchSchool)
bigAchieve$Size <- bigAchieve$Size/1000
names(bigAchieve) <- tolower(names(bigAchieve))
bigmath.fit <- lmer(mathach ~ minority + ses + sex + size +sector + pracad + meanses  + minority:ses + min

xtable(summary(bigmath.fit)$coefficients)
vc <-  VarCorr(bigmath.fit)$school

 varCorrs <-  matrix( c(sqrt(diag(vc[1:2,1:2])),NA, attr(vc,"correlation")[1,2]) ,2,2,
              dimnames=list(c("Intercept","Minority"),c("Std.Dev", "Corr")))
sigma <- c(summary( bigmath.fit)$sigma, NA)
 varCorrs <- rbind(varCorrs, sigma)
 xtable(varCorrs)
```

```r
require(R2jags, quietly=TRUE)
cat(" model{
  for(ndx in 1:n){
     y[ndx] ~ dnorm(y.hat[ndx], tau.y)
     y.hat[ndx] <- beta0 + beta1 * Minority[ndx] +beta2 * SES[ndx] + beta3 * Sex[ndx] + beta4 * SES[ndx] *
  }
  tau.y <- pow(sigma.y, -2)
  sigma.y ~ dunif(0,100)
  beta0 ~ dnorm(0, .00001)
  beta1 ~ dnorm(0, .00001)
  beta2 ~ dnorm(0, .00001)
  beta3 ~ dnorm(0, .00001)
  beta4 ~ dnorm(0, .00001)
    }", file = "schoolFE.jags")

schoolDataJags1 <- with(MathAchieve, list("y" = MathAch, "n" = nrow(MathAchieve), "Minority" =  Minority,
 parms1 <- c("beta0", "beta1", "beta2", "beta3", "beta4", "sigma.y")
 jags1 <- jags(data=schoolDataJags1, model.file="schoolFE.jags", parameters.to.save = parms1, n.chains=4,
 jags1
summary(lm(MathAch ~ Minority * SES + Sex, MathAchieve))$coef
```

```r
cat(" model{
  for(ndx in 1:n){
     y[ndx] ~ dnorm(y.hat[ndx], tau.y)
     y.hat[ndx] <- beta0 + beta1 * Minority[ndx] + beta2 * SES[ndx] + beta3 * Sex[ndx] +
                     beta4 * SES[ndx] * Minority[ndx] + b0[school[ndx]] + b1[school[ndx]] * Minority[ndx]
  }
  tau.y <- pow(sigma.y, -2)
  sigma.y ~ dunif(0,100)
  beta0 ~ dnorm(0, .00001)
  beta1 ~ dnorm(0, .00001)
  beta2 ~ dnorm(0, .00001)
  beta3 ~ dnorm(0, .00001)
  beta4 ~ dnorm(0, .00001)
  for(j in 1:J){
    b0[j] ~ dnorm(b0.hat[j], tau.b0)
    b0.hat[j] <- beta0
    b1[j] ~ dnorm(b1.hat[j], tau.b1)
    b1.hat[j] <- beta1
```

```
  }
  tau.b0 <- pow(sigma.b0, -2)
  tau.b1 <- pow(sigma.b1, -2)
  sigma.b0 ~ dunif(0,10)
  sigma.b1 ~ dunif(0,10)
    }", file = "schoolME2.jags")

schoolDataJags2 <- with(MathAchieve,
   list("y" = MathAch, "n" = nrow(MathAchieve), "Minority" = Minority,
        "SES" = SES, "Sex" = Sex, "school" = as.numeric(School),
        "J" = length(levels(School))))
 parms2 <- c("beta0", "beta1", "beta2", "beta3", "beta4",
             "sigma.y", "sigma.b0","sigma.b1")
 jags2 <- jags(data=schoolDataJags2, model.file="schoolME2.jags", parameters.to.save = parms2,  n.chains=4
 #  jags2 <- autojags(jags2)
```

```
 jags2
 summary(math.fit4)
```

```
cat(" model{
 for(ndx in 1:n){
     y[ndx] ~ dnorm(y.hat[ndx], tau.y)
     y.hat[ndx] <-  beta2 * SES[ndx] + beta3 * Sex[ndx] +  beta8 * SES[ndx] * Minority[ndx] +
       b1[school[ndx]] + b2[school[ndx]] * Minority[ndx]
  }
  tau.y <- pow(sigma.y, -2)
  sigma.y ~ dunif(0,100)
  beta0 ~ dnorm(0, .00001)  ## Intercept
  beta1 ~ dnorm(0, .00001)  ## Minority
  beta2 ~ dnorm(0, .00001)  ## SES
  beta3 ~ dnorm(0, .00001)  ## Sex
  beta4 ~ dnorm(0, .00001)  ## size
  beta5 ~ dnorm(0, .00001)  ## sector
  beta6 ~ dnorm(0, .00001)  ## pracad
  beta7 ~ dnorm(0, .00001)  ## meanSES
  beta8 ~ dnorm(0, .00001)  ## min*SES
  beta9 ~ dnorm(0, .00001)  ## min*sector
  beta10 ~ dnorm(0, .00001)  ## min*disclim
  beta11 ~ dnorm(0, .00001)  ## min*meanSES
  for(j in 1:J){
    b1[j] ~ dnorm(b1.hat[j], tau.b1)
    b1.hat[j] <- beta0 + beta4 * size[j] + beta5 * sector[j] + beta6 * pracad[j] + beta7 * meanSES[j]
    b2[j] ~ dnorm(b2.hat[j], tau.b2)
    b2.hat[j] <- beta1 +  beta9*sector[j] + beta10 * disclim[j] + beta11 * meanSES[j]
  }
  tau.b1 <- pow(sigma.b1,-2)
  tau.b2 <- pow(sigma.b2,-2)
  sigma.b1 ~ dunif(0,10)
  sigma.b2 ~ dunif(0,10)
    }", file = "schoolME3.jags")
 schoolDataJags3 <- with(MathAchieve,
   list("y" = MathAch, "n" = nrow(MathAchieve), "Minority" =  as.numeric(Minority)-1,
        "SES" = SES, "Sex" = as.numeric(Sex)-1, "school" = as.numeric(School),
        "J" = length(levels(School))))
 schoolDataJags3 <- c(schoolDataJags3,  with(MathAchSchool,
        list("size" = Size/1000, "sector" = as.numeric(Sector) -1, "pracad" = PRACAD,
         "disclim" = DISCLIM + .01512, meanSES = MEANSES)))

 parms3 <- c("beta0", "beta1", "beta2", "beta3", "beta4",
            "beta5", "beta6", "beta7", "beta8","beta9","beta10","beta11",
            "sigma.y", "sigma.b1","sigma.b2")

 jags3 <- jags(data=schoolDataJags3, model.file="schoolME3.jags", parameters.to.save = parms3,  n.chains=4

intModel3 <- update(intModel3, . ~ size + I(as.numeric(sector)-1) + pracad + disclim + meanses)
minModel4 <- update(minModel4, . ~ I(as.numeric(sector)-1) +  I(disclim+.01512) + meanses)

jags3  <- autojags(jags3)
##rbind(summary(intModel3)$coef, rep(NA,4), summary(minModel4)$coef)
```

```
cat(" model{
 for(ndx in 1:n){
     y[ndx] ~ dnorm(y.hat[ndx], tau.y)
     y.hat[ndx] <- beta0 + beta1 * Minority[ndx] + beta2 * SES[ndx] + beta3 * Sex[ndx] +
         beta4 * size[school[ndx]] + beta5 * sector[school[ndx]] + beta6 * pracad[school[ndx]] +
         beta7 * meanSES[school[ndx]] +   beta8 * SES[ndx] * Minority[ndx] +
         beta9 * Minority[ndx] * sector[school[ndx]] + beta10 * Minority[ndx]  * disclim[school[ndx]] +
        beta11 * Minority[ndx] * meanSES[school[ndx]] + b1[school[ndx]] + b2[school[ndx]] * Minority[ndx]
  }
  tau.y <- pow(sigma.y, -2)
  sigma.y ~ dunif(0,100)
  beta0 ~ dnorm(0, .00001)  ## Intercept
  beta1 ~ dnorm(0, .00001)  ## Minority
  beta2 ~ dnorm(0, .00001)  ## SES
  beta3 ~ dnorm(0, .00001)  ## Sex
  beta4 ~ dnorm(0, .00001)  ## size
  beta5 ~ dnorm(0, .00001)  ## sector
  beta6 ~ dnorm(0, .00001)  ## pracad
  beta7 ~ dnorm(0, .00001)  ## meanSES
  beta8 ~ dnorm(0, .00001)  ## min*SES
  beta9 ~ dnorm(0, .00001)  ## min*sector
  beta10 ~ dnorm(0, .00001)  ## min*disclim
  beta11 ~ dnorm(0, .00001)  ## min*meanSES
  for(j in 1:J){
    b1[j] ~ dnorm(b1.hat[j], tau.b1)
    b1.hat[j] <- beta0
    b2[j] ~ dnorm(b2.hat[j], tau.b2)
    b2.hat[j] <- beta1
  }
  tau.b1 <- pow(sigma.b1,-2)
  tau.b2 <- pow(sigma.b2,-2)
  sigma.b1 ~ dunif(0,10)
  sigma.b2 ~ dunif(0,10)
    }", file = "schoolME4.jags")
 schoolDataJags3 <- with(MathAchieve,
   list("y" = MathAch, "n" = nrow(MathAchieve), "Minority" =  as.numeric(Minority)-1,
        "SES" = SES, "Sex" = as.numeric(Sex)-1, "school" = as.numeric(School),
        "J" = length(levels(School))))
 schoolDataJags3 <- c(schoolDataJags3,  with(MathAchSchool,
        list("size" = Size/1000, "sector" = as.numeric(Sector) -1, "pracad" = PRACAD,
         "disclim" = DISCLIM + .01512, "meanSES" = MEANSES)))

 parms4 <- c("beta0", "beta1", "beta2", "beta3", "beta4", "beta5","beta6", "beta7", "beta8", "beta9",
             "beta10", "beta11",  "sigma.y", "sigma.b1","sigma.b2")
 jags4 <- jags(data=schoolDataJags3, model.file="schoolME4.jags", parameters.to.save = parms4,  n.chains=4
 ##  jags4
 jags4  <- autojags(jags4, n.iter = 10000, n.update = 5, Rhat = 1.05, n.thin = 20)
 traceplot(jags4, varname = "sigma")
```

```
 out4 <- summary(as.mcmc(jags4))
 Rhat  <- gelman.diag(as.mcmc(jags4))[[1]][,1]
 effects <- c(names(fixef(bigmath.fit)),"Intercept SD",
                                       "Minority SD / Corr", "Resid SD")
 results <- cbind(out4[[1]][,1:2], Rhat)[c(1:2,5:12,3:4,14:16),]
 results <- cbind( data.frame(results),
```

```
                  cbind(
                      rbind(round(summary(bigmath.fit)$coefficients,3),
                        cbind(round(varCorrs,3), c(NA,NA,NA))),effects))

xtable(results,  digits = 3, align = "lrrr|rrrl")
```