OCN 750 HW 12 - Revisiting Heat Shock

Maia Kapur

Thursday, November 12, 2015

#### Fit a GLMM that incorporates random effects for isofemale and population, and also includes fixed effects and interactions that test the 6 questions. We also need to account for potential overdispersion in the binomial response. The best way to do this in lme4 is with an 'individual level random effect'. The thing to do is make a new factor where each row of the dataset gets a level. This is done in the replicateID column which is just a single number as a factor that gets popped into the dataframe. Other ways of doing this include the seq() function.

require(lme4)  
require(car)  
require(effects)  
require(lattice)  
  
heat = read.csv("heat\_shock\_all.csv")  
heat$replicateID = factor(1:nrow(heat))  
#heat$proportion[heat$proportion != 0] <- 1 ## change any positive proportion of survivorship to 1  
  
mod1 = glmer(proportion ~ (1|isofemale) + (1|population) + sex + region + treatment + treatment\*region + treatment\*sex + sex\*region + (1|replicateID), data = heat, family = "binomial")

#### Use likelihood ratio tests to see which of the interactions is significant (LRTs tend to be anti-conservative for fixed effects, but this model has a lot of data, and doing a parametric bootstrap would take a long time with this model). Drop the non-significant interactions, which will allow you to test and interpret the main effects more clearly. Now use LRTs to test the significance of the remaining fixed effects in the model. these are taking a weirdly long time to run.

summary(mod1)

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: proportion ~ (1 | isofemale) + (1 | population) + sex + region +   
## treatment + treatment \* region + treatment \* sex + sex \*   
## region + (1 | replicateID)  
## Data: heat  
##   
## AIC BIC logLik deviance df.resid   
## 479.0 532.8 -229.5 459.0 1600   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -0.2872 -0.1111 0.0966 0.8578 9.1713   
##   
## Random effects:  
## Groups Name Variance Std.Dev.   
## replicateID (Intercept) 7.124e-08 0.0002669  
## isofemale (Intercept) 3.662e-08 0.0001914  
## population (Intercept) 0.000e+00 0.0000000  
## Number of obs: 1610, groups:   
## replicateID, 1610; isofemale, 56; population, 9  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)  
## (Intercept) -4.4986 0.7311 -6.154 7.58e-10  
## sexMales -1.3257 1.0087 -1.314 0.1887  
## regionSouthwest 0.4030 0.8371 0.481 0.6302  
## treatmenthardening 1.3809 0.8214 1.681 0.0927  
## regionSouthwest:treatmenthardening 0.2198 0.9241 0.238 0.8120  
## sexMales:treatmenthardening 0.2754 0.7351 0.375 0.7080  
## sexMales:regionSouthwest 1.0273 0.8635 1.190 0.2341  
##   
## (Intercept) \*\*\*  
## sexMales   
## regionSouthwest   
## treatmenthardening .   
## regionSouthwest:treatmenthardening   
## sexMales:treatmenthardening   
## sexMales:regionSouthwest   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) sexMls rgnSth trtmnt rgnSt: sxMls:  
## sexMales -0.235   
## reginSthwst -0.827 0.007   
## trtmnthrdnn -0.850 0.041 0.688   
## rgnSthwst:t 0.704 0.186 -0.851 -0.827   
## sxMls:trtmn 0.140 -0.594 0.212 -0.164 -0.227   
## sxMls:rgnSt 0.176 -0.749 -0.213 0.068 -0.064 -0.011  
## convergence code: 0  
## Model failed to converge with max|grad| = 0.00636713 (tol = 0.001, component 1)

## non significant FIXED interactions are sex + region; sex alone and region alone are also insignificant   
  
mod2 = glmer(proportion ~ (1|isofemale) + (1|population) + treatment + sex + region + region\*treatment + sex\*treatment + (1|replicateID), data = heat, family = "binomial")  
summary(mod2)

## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: proportion ~ (1 | isofemale) + (1 | population) + treatment +   
## sex + region + region \* treatment + sex \* treatment + (1 |   
## replicateID)  
## Data: heat  
##   
## AIC BIC logLik deviance df.resid   
## 479.8 528.3 -230.9 461.8 1601   
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -0.2847 -0.1106 0.1561 0.9826 6.5716   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## replicateID (Intercept) 2.051e-03 0.04529   
## isofemale (Intercept) 4.435e-04 0.02106   
## population (Intercept) 2.251e-06 0.00150   
## Number of obs: 1610, groups:   
## replicateID, 1610; isofemale, 56; population, 9  
##   
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)  
## (Intercept) -4.4033 0.6562 -6.710 1.94e-11  
## treatmenthardening 1.0212 0.7698 1.327 0.185  
## sexMales -0.7058 0.6215 -1.136 0.256  
## regionSouthwest 0.6165 0.7175 0.859 0.390  
## treatmenthardening:regionSouthwest 0.2502 0.8302 0.301 0.763  
## treatmenthardening:sexMales 0.5085 0.6952 0.731 0.464  
##   
## (Intercept) \*\*\*  
## treatmenthardening   
## sexMales   
## regionSouthwest   
## treatmenthardening:regionSouthwest   
## treatmenthardening:sexMales   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) trtmnt sexMls rgnSth trtm:S  
## trtmnthrdnn -0.852   
## sexMales -0.261 0.223   
## reginSthwst -0.832 0.709 -0.076   
## trtmnthrd:S 0.719 -0.836 0.066 -0.864   
## trtmnthrd:M 0.233 -0.275 -0.894 0.068 -0.070  
## convergence code: 0  
## Model failed to converge with max|grad| = 4.35126 (tol = 0.001, component 1)  
## failure to converge in 10000 evaluations

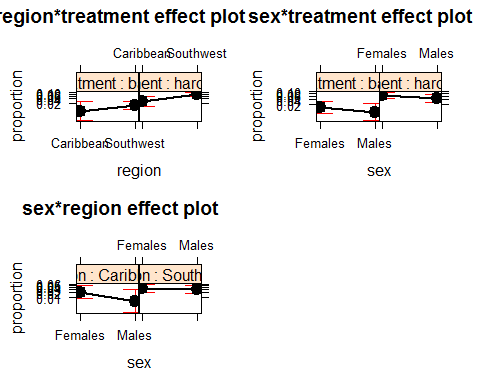
anova(mod1, mod2, test = "Chisq")

## Data: heat  
## Models:  
## mod2: proportion ~ (1 | isofemale) + (1 | population) + treatment +   
## mod2: sex + region + region \* treatment + sex \* treatment + (1 |   
## mod2: replicateID)  
## mod1: proportion ~ (1 | isofemale) + (1 | population) + sex + region +   
## mod1: treatment + treatment \* region + treatment \* sex + sex \*   
## mod1: region + (1 | replicateID)  
## Df AIC BIC logLik deviance Chisq Chi Df Pr(>Chisq)   
## mod2 9 479.80 528.25 -230.90 461.80   
## mod1 10 478.97 532.81 -229.48 458.97 2.8284 1 0.09261 .  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

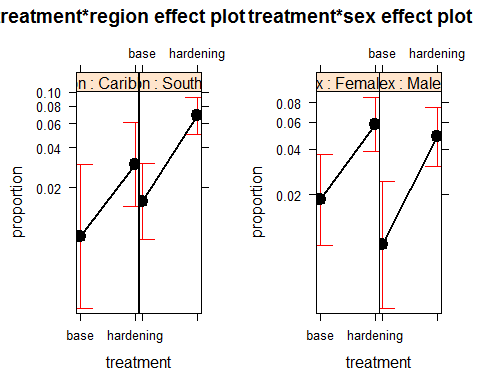
## The second model has a slighly lower AIC value.

#### Plot the fitted fixed effects using the effects package. How do you interpret these results, in light of the 6 questions listed above?

plot(allEffects(mod1))



plot(allEffects(mod2))

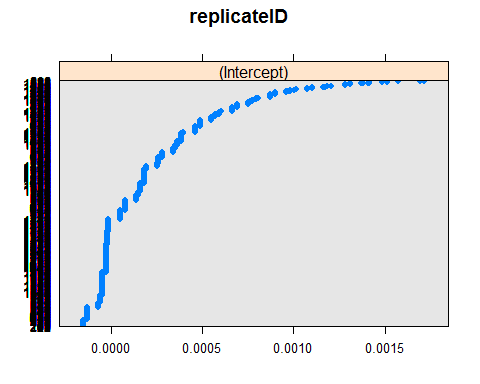


## It appears from the effects plot of the full model that the only significant differences are between SW Hardening and Caribbean Base (eg treatment \* region) and differences between treatments alone but not between sexes and treatments.

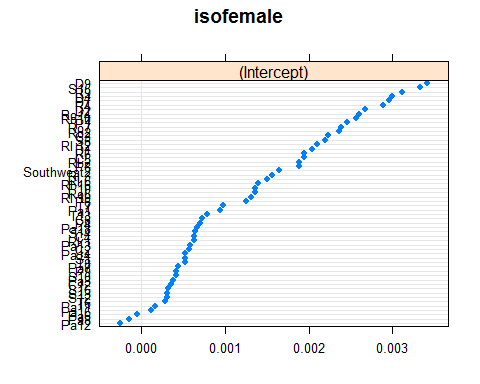
#### Use the random effects estimates, and dotplot() of those estimates, to discuss how much genetic variation there is among populations, and among isofemales within populations, and how much extra (overdispersed) variation there is among replicates. Plotting the individual-level random effect for replicate is not advised; there are too many levels for the function. How does the spread for the random effects compare to the effect sizes of the fixed effects?

par(mfrow = c(2,2))  
lattice::dotplot(ranef(mod2))

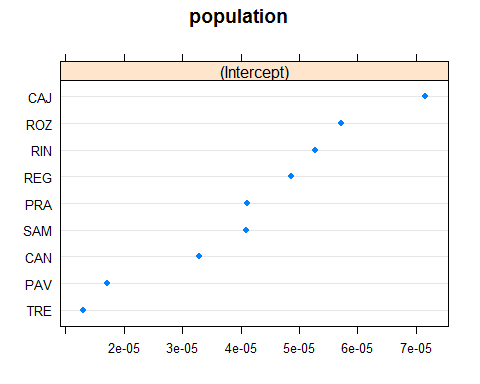
## $replicateID



##   
## $isofemale



##   
## $population



fixef(mod2)

## (Intercept) treatmenthardening   
## -4.4032935 1.0212341   
## sexMales regionSouthwest   
## -0.7057720 0.6165114   
## treatmenthardening:regionSouthwest treatmenthardening:sexMales   
## 0.2501921 0.5085087

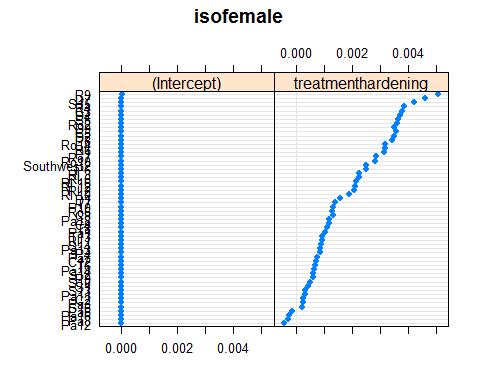
## fixed effects estimates are generally lower but all generally in keeping with the random effects of isofemale and region.

#### The authors were interested in another question as well: Does acclimation to high temperature (i.e., the treatment effect) exhibit genetic variation among isofemale lines? To test this, you need to allow treatment to vary by isofemale. In other words, you need to include a fixed\*random interaction. Put this in the model, and use a likelihood ratio test to see if it is important. Use dotplot() to visualize how the treatment effect varies by isofemale. Does the genetic variation seem substantial? What is the correlation between the random Intercept and random treatment effect among isofemales? Can you explain what this correlation implies? Does including the treatment-by-isofemale interaction change the results for the fixed effects in the model? If so, how?

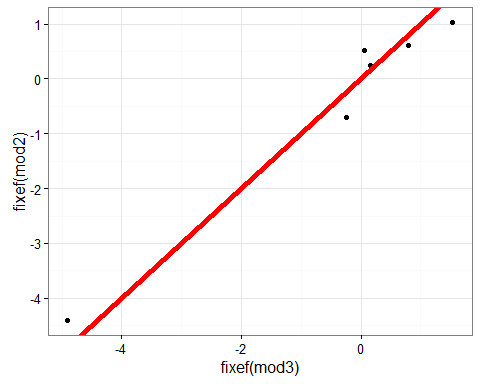
mod3 = glmer(proportion ~ (1|isofemale) + (1|population) + (treatment|isofemale) + treatment + sex + region + region\*treatment + sex\*treatment + (1|replicateID), data = heat, family = "binomial")  
  
anova(mod3, mod2)

## Data: heat  
## Models:  
## mod2: proportion ~ (1 | isofemale) + (1 | population) + treatment +   
## mod2: sex + region + region \* treatment + sex \* treatment + (1 |   
## mod2: replicateID)  
## mod3: proportion ~ (1 | isofemale) + (1 | population) + (treatment |   
## mod3: isofemale) + treatment + sex + region + region \* treatment +   
## mod3: sex \* treatment + (1 | replicateID)  
## Df AIC BIC logLik deviance Chisq Chi Df Pr(>Chisq)  
## mod2 9 479.80 528.25 -230.90 461.80   
## mod3 12 485.86 550.47 -230.93 461.86 0 3 1

dotplot(ranef(mod3))$isofemale



# plot(fixef(mod3), fixef(mod2))  
# abline(0,1)  
  
require(ggplot2)  
ggplot(data = NULL, aes(x = fixef(mod3), y = fixef(mod2)))+  
 geom\_point()+  
 geom\_abline(slope = 1, intercept = 0, lwd = 2, col = 'red') +  
 theme\_bw()



## there's almost a perfect 1:1 relationship between the fixed effects for the model with and without the fixed\*random interaction. So it doesn't add a very significant difference in the value.