

Cluster Models

Contents

1	Getting Set Up	1
1.1	Setting chunk options and generating R script	1
1.2	Installing Packages	1
1.3	Reading the Data	2
2	Exploring the Data	2
2.1	Table describing the data	2
3	Statistical Analyses	5
3.1	Fitting a homoscedastic model	5
3.2	Display the random effects (EBLUPs) from the model.	7
3.3	Fitting a heteroscedastic model	7
3.4	Heteroscedastic versus homoscedastic model	8
3.5	High-low dose: Equal residual variance	9
3.6	Fitting the final model using ML	11

1 Getting Set Up

1.1 Setting chunk options and generating R script

```
knitr::opts_chunk$set(warning=FALSE, message=FALSE)
knitr::purl("longitudinal-models.Rmd")
```

1.2 Installing Packages

```
install.packages("Hmisc")
install.packages("lattice")
install.packages('grid')
install.packages('nlme')
```

1.3 Reading the Data

```
ratpup = read.table("rat_pup.dat", h = T)
ratpup$sex1[ratpup$sex == "Female"] = 1
ratpup$sex1[ratpup$sex == "Male"] = 0
attach(ratpup)
```

2 Exploring the Data

2.1 Table describing the data

```
library(Hmisc)
g = function(x){
  c(N=length(x), Mean=mean(x,na.rm=TRUE), SD=sd(x,na.rm=TRUE), Min=min(x,na.rm=TRUE), Max=max(x,na.rm=TRUE))
}
summarize(weight,by=llist(treatment,sex),g)
```

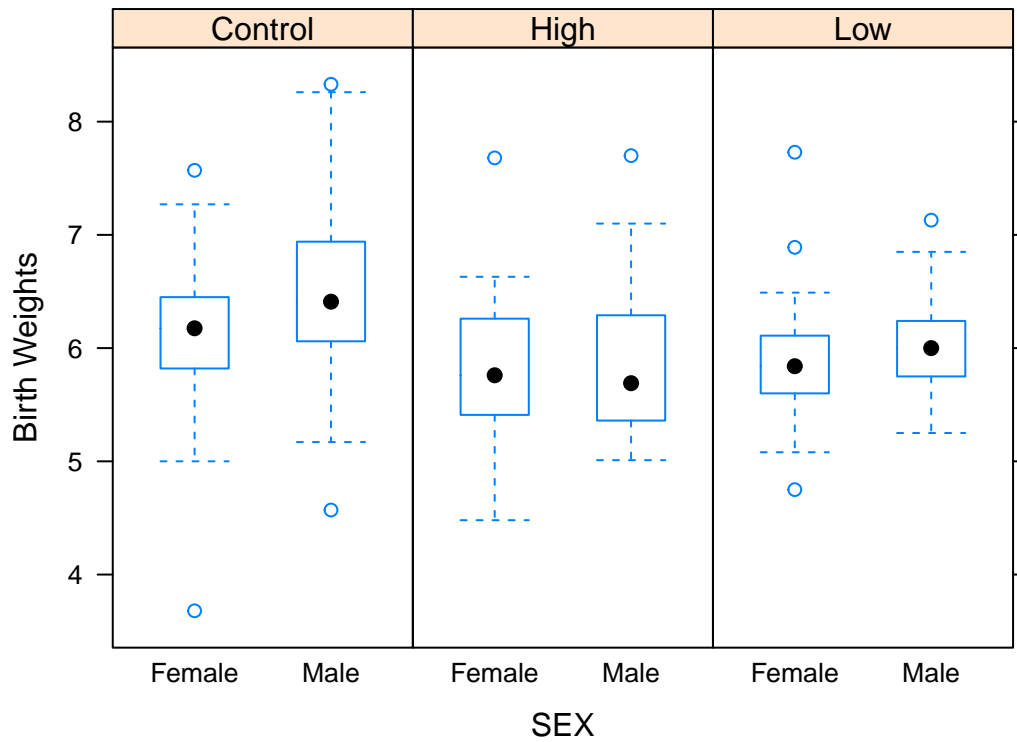
##	treatment	sex	weight	Mean	SD	Min	Max
## 1	Control	Female	54	6.116111	0.6851179	3.68	7.57
## 2	Control	Male	77	6.471039	0.7537880	4.57	8.33
## 3	High	Female	32	5.851562	0.6001887	4.48	7.68
## 4	High	Male	33	5.918485	0.6909058	5.01	7.70
## 5	Low	Female	65	5.837538	0.4504964	4.75	7.73
## 6	Low	Male	61	6.025082	0.3803403	5.25	7.13

Comparing the distributions of birth weights for each treatment by sex combination.

```
library(lattice) # trellis graphics
library(grid)

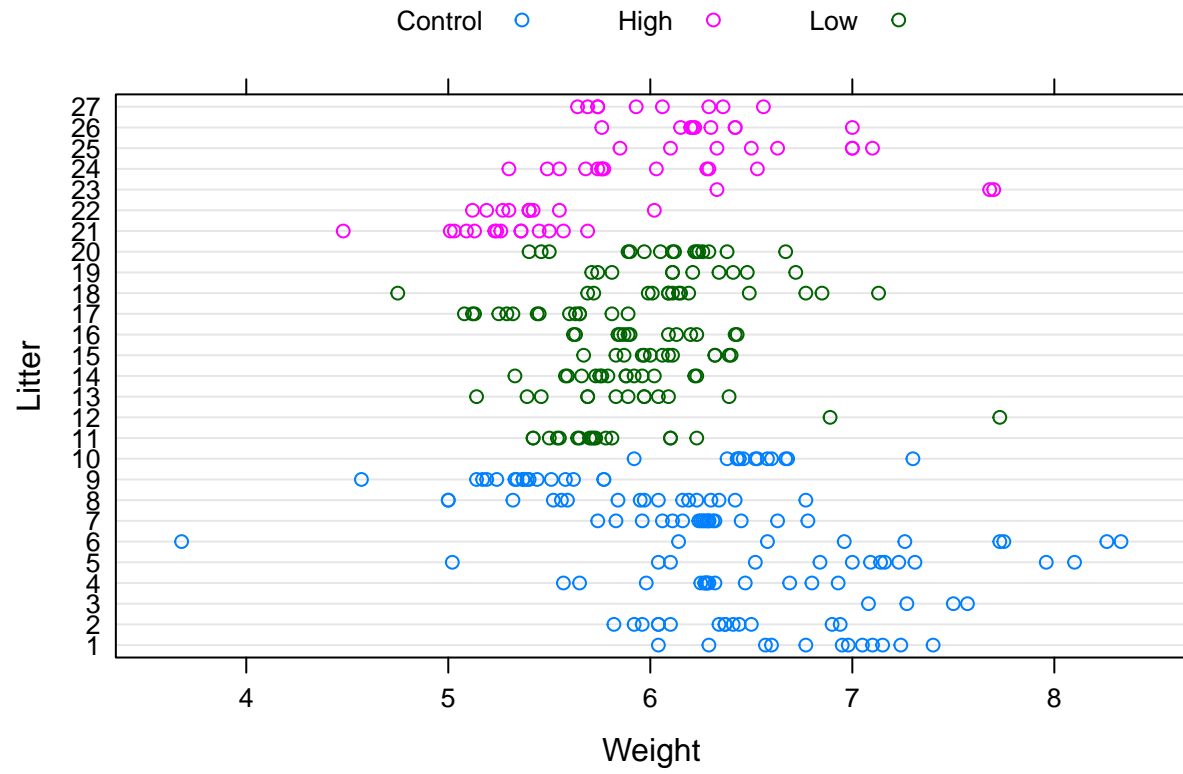
bwplot(weight ~ sex|treatment, data=ratpup,aspect = 2, ylab="Birth Weights", xlab="SEX", main = "Boxplot of Birth Weights by Treatment and Sex")
```

Boxplots of birth weights for levels of treatment by sex

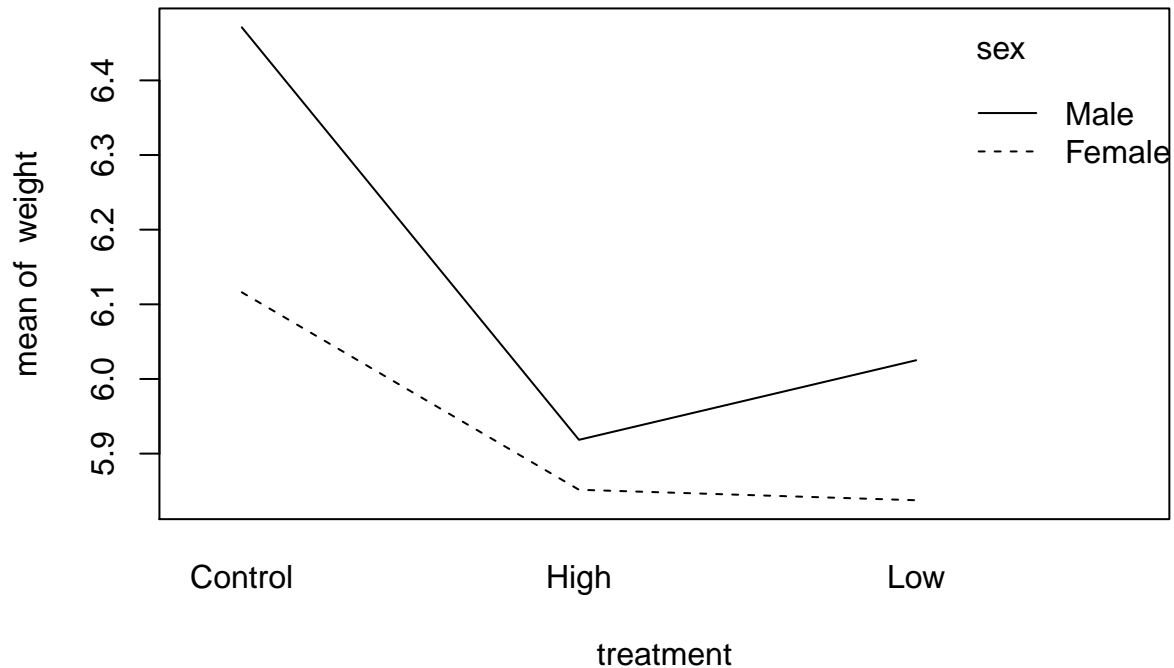


Comparing the distributions of birth weights for each treatment

```
dotplot(litterid ~ weight, group=treatment, data=ratpup, xlab="Weight", ylab="Litter", auto.key=list(spa
```



```
with(ratpup, interaction.plot(treatment,sex,weight))
```



3 Statistical Analyses

3.1 Fitting a homoscedastic model

At this moment the `lmer()` function does not allow users to fit models with heterogeneous error variance structures. Therefore, we will work with the function `lme()` from the package `nlme`. `lme()` and `lmer()` are similar but there are some differences in syntax and output that will be explained in what follows.

```
library(nlme)

meanfull.hom <- lme(weight ~ treatment + sex1 + litsize + treatment:sex1, random=~1 | litterid, ratpup,
```

The `factor()` function is not necessary for `treatment`, because the original `treatment` variable has string values `High`, `Low`, and `Control`, and will therefore be considered as a factor automatically. We also do not need to declare `sex1` as a factor, because it is an indicator variable having only values of 0 and 1.

`lme()` treats the lowest level (alphabetically or numerically) of a factor as the reference category. This means that “Control” will be the reference category of `treatment`. The reference level can be changed using “`treatment=relevel(treatment,ref=“High”)`”.

`random = 1 | litterid`, includes a random effect (intercept) for each level of litter in the model. `method = “REML”`, specifies that the default REML estimation method is to be used.

```
summary(meanfull.hom)
```

```
## Linear mixed-effects model fit by REML
## Data: ratpup
##      AIC      BIC    logLik
## 419.1043 452.8775 -200.5522
##
## Random effects:
## Formula: ~1 | litterid
##      (Intercept) Residual
## StdDev:   0.3106722 0.404337
##
## Fixed effects: weight ~ treatment + sex1 + litsize + treatment:sex1
##              Value Std.Error DF   t-value p-value
## (Intercept)   8.323340 0.27333009 292 30.451605 0.0000
## treatmentHigh -0.906057 0.19154238 23 -4.730320 0.0001
## treatmentLow  -0.467040 0.15818328 23 -2.952521 0.0071
## sex1          -0.411688 0.07315410 292 -5.627679 0.0000
## litsize       -0.128382 0.01875336 23 -6.845819 0.0000
## treatmentHigh:sex1 0.107023 0.13176318 292 0.812239 0.4173
## treatmentLow:sex1 0.083866 0.10568189 292 0.793568 0.4281
## Correlation:
##              (Intr) trtmnH trtmnL sex1  litsiz trtH:1
## treatmentHigh -0.562
## treatmentLow  -0.297 0.404
## sex1          -0.111 0.158 0.191
## litsize       -0.916 0.363 0.022 0.001
## treatmentHigh:sex1 0.043 -0.323 -0.106 -0.555 0.020
## treatmentLow:sex1 0.044 -0.096 -0.320 -0.692 0.036 0.385
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -7.47250744 -0.50014749 0.02911668 0.57348178 3.00962055
##
## Number of Observations: 322
## Number of Groups: 27
```

```
anova(meanfull.hom)
```

```
##              numDF denDF F-value p-value
## (Intercept)      1   292 9093.772 <.0001
## treatment        2    23   5.082 0.0149
## sex1             1   292  52.602 <.0001
## litsize          1    23  47.374 <.0001
## treatment:sex1    2   292   0.466 0.6282
```

The `anova()` function performs a series of Type I (or sequential) F-tests for the fixed effects in the model, each of which are conditional on the preceding terms in the model specification. For example, the F-test for `sex1` is conditional on the treatment effects, but the F-test for treatment is not conditional on the `sex1` effect.

The model was fitted using REML and, therefore, different mean structures cannot be compared.

3.2 Display the random effects (EBLUPs) from the model.

```
random.effects(meanfull.hom)
```

```
##      (Intercept)
## 1    0.17480024
## 2   -0.07362296
## 3   -0.17490203
## 4   -0.05376249
## 5    0.34446954
## 6   -0.05480208
## 7    0.39153638
## 8   -0.02616704
## 9   -0.61772106
## 10   0.09017150
## 11   0.04136696
## 12   0.02072931
## 13  -0.35981737
## 14   0.01847368
## 15   0.03549783
## 16   0.06416884
## 17  -0.39636862
## 18   0.42802095
## 19  -0.18110865
## 20   0.32903707
## 21  -0.27813901
## 22  -0.49096620
## 23   0.26053476
## 24   0.17537803
## 25   0.23748827
## 26   0.22966911
## 27  -0.13396497
```

3.3 Fitting a heteroscedastic model

```
meanfull.het <- lme(weight ~ treatment + sex1 + litsize + treatment:sex1, random=~1 | litterid, ratpup,
```

The arguments of the `lme()` function are the same as those used to fit Model 1, with the addition of the `weights` argument. The argument “`weights = varIdent(form = ~ 1 | treatment)`” sets up a heterogeneous residual variance structure, with observations at different levels of treatment having different residual variance parameters.

```
summary(meanfull.het)
```

```
## Linear mixed-effects model fit by REML
## Data: ratpup
##      AIC      BIC    logLik
## 381.8847 423.163 -179.9423
##
## Random effects:
```

```

## Formula: ~1 | litterid
##      (Intercept) Residual
## StdDev:  0.3134846 0.5147948
##
## Variance function:
## Structure: Different standard deviations per stratum
## Formula: ~1 | treatment
## Parameter estimates:
##   Control      Low      High
## 1.0000000 0.5649830 0.6394383
## Fixed effects: weight ~ treatment + sex1 + litsize + treatment:sex1
##              Value Std.Error DF   t-value p-value
## (Intercept)    8.345294 0.27464753 292 30.385468  0.0000
## treatmentHigh  -0.903277 0.19215903  23 -4.700672  0.0001
## treatmentLow   -0.466292 0.15908908  23 -2.931013  0.0075
## sex1           -0.408131 0.09303486 292 -4.386865  0.0000
## litsize        -0.130007 0.01848708  23 -7.032332  0.0000
## treatmentHigh:sex1 0.094666 0.12919527 292  0.732737  0.4643
## treatmentLow:sex1  0.076013 0.10811858 292  0.703053  0.4826
## Correlation:
##              (Intr) trtmnH trtmnL sex1  litsiz trtH:1
## treatmentHigh  -0.598
## treatmentLow   -0.362  0.466
## sex1           -0.140  0.200  0.241
## litsize        -0.905  0.376  0.056  0.000
## treatmentHigh:sex1 0.087 -0.304 -0.173 -0.720  0.015
## treatmentLow:sex1  0.100 -0.164 -0.303 -0.860  0.022  0.620
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -5.88670114 -0.52493419  0.02123518  0.57307286  2.56409983
##
## Number of Observations: 322
## Number of Groups: 27

```

Random effects portion of the output: Estimated residual standard deviation equal to 0.5147948. Parameter estimates: Values by which the residual standard deviation should be multiplied to obtain the estimated standard deviation of the residuals in each treatment group. This multiplier is 1.0 for the control group (the reference). Multipliers for the low and high treatment groups are very similar.

3.4 Heterocedastic versus homocedastic model

```
anova(meanfull.hom, meanfull.het)
```

```

##           Model df      AIC      BIC    logLik   Test  L.Ratio p-value
## meanfull.hom    1  9 419.1043 452.8775 -200.5522
## meanfull.het    2 11 381.8847 423.1630 -179.9423 1 vs 2 41.21964 <.0001

```


3.5 High-low dose: Equal residual variance

```
ratpup$trtgrp[treatment=="Control"] <- 1
ratpup$trtgrp[treatment == "Low" | treatment == "High"] <- 2

meanfull.hilo <- lme(weight ~ treatment + sex1 + litsize + treatment:sex1, random=~1 | litterid, ratpup)

summary(meanfull.hilo)
```

```
## Linear mixed-effects model fit by REML
## Data: ratpup
##      AIC      BIC    logLik
## 381.0807 418.6065 -180.5404
##
## Random effects:
## Formula: ~1 | litterid
##      (Intercept) Residual
## StdDev:  0.3145678 0.5147878
##
## Variance function:
## Structure: Different standard deviations per stratum
## Formula: ~1 | trtgrp
## Parameter estimates:
##      1      2
## 1.0000000 0.5905488
## Fixed effects: weight ~ treatment + sex1 + litsize + treatment:sex1
##      Value Std.Error DF  t-value p-value
## (Intercept)      8.350351 0.27567830 292 30.290200  0.0000
## treatmentHigh    -0.901844 0.19140143  23 -4.711794  0.0001
## treatmentLow     -0.466596 0.15999335  23 -2.916348  0.0078
## sex1             -0.408195 0.09303538 292 -4.387529  0.0000
## litsize          -0.130383 0.01856366  23 -7.023575  0.0000
## treatmentHigh:sex1  0.092026 0.12461723 292  0.738473  0.4608
## treatmentLow:sex1  0.076397 0.10939796 292  0.698337  0.4855
## Correlation:
##      (Intr) trtmnH trtmnL sex1  litsiz trtH:1
## treatmentHigh    -0.605
## treatmentLow     -0.358  0.466
## sex1             -0.140  0.201  0.240
## litsize          -0.906  0.382  0.054  0.000
## treatmentHigh:sex1  0.092 -0.292 -0.178 -0.747  0.013
## treatmentLow:sex1  0.097 -0.162 -0.306 -0.850  0.023  0.635
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -5.88725948 -0.52398488  0.01731336  0.56412336  2.55014722
##
## Number of Observations: 322
## Number of Groups: 27
```

```
anova(meanfull.hilo)
```

```
##      numDF denDF  F-value p-value
```

```
## (Intercept)      1    292 9027.743 <.0001
## treatment        2     23   4.241 0.0271
## sex1             1    292  61.568 <.0001
## litsize          1     23  49.577 <.0001
## treatment:sex1    2    292   0.317 0.7288
```

```
anova(meanfull.het, meanfull.hilo)
```

```
##           Model df      AIC      BIC    logLik  Test  L.Ratio p-value
## meanfull.het    1  11 381.8847 423.1630 -179.9423
## meanfull.hilo    2  10 381.0807 418.6065 -180.5404 1 vs 2 1.196053 0.2741
```

Can the random effects (b0i) associated with the litter-specific intercepts be omitted from Model 3? One does not directly test the significance of the random litter-specific intercepts, but rather tests a hypothesis related to the variance of the random litter effects. Although hypothesis tests are often phrased in terms of parameter restrictions, they basically compare the quality of the fit obtained from two nested models. Likelihood ratio tests (LRTs) are a valuable tool to compare nested models.

```
meanfull.hilo.nolitter <- gls(weight ~ treatment + sex1 + litsize + treatment:sex1, data = ratpup, weight = weight)
summary(meanfull.hilo.nolitter)
```

```
## Generalized least squares fit by REML
##   Model: weight ~ treatment + sex1 + litsize + treatment:sex1
##   Data: ratpup
##           AIC      BIC    logLik
##   489.6521 523.4252 -235.826
##
## Variance function:
##   Structure: Different standard deviations per stratum
##   Formula: ~1 | trtgrp
##   Parameter estimates:
##           1      2
##   1.0000000 0.7060188
##
## Coefficients:
##           Value Std.Error   t-value p-value
## (Intercept)   8.201712 0.15902776  51.57409 0.0000
## treatmentHigh -0.976414 0.10624042  -9.19060 0.0000
## treatmentLow  -0.456018 0.08700180  -5.24147 0.0000
## sex1          -0.339911 0.10616682  -3.20167 0.0015
## litsize       -0.121478 0.01008518 -12.04524 0.0000
## treatmentHigh:sex1 0.180960 0.14941228   1.21114 0.2267
## treatmentLow:sex1 0.076386 0.13035758   0.58597 0.5583
##
## Correlation:
##           (Intr) trtmnH trtmnL sex1  litsiz trtH:1
## treatmentHigh -0.574
## treatmentLow  -0.344 0.506
## sex1          -0.265 0.408 0.503
## litsize       -0.903 0.331 0.010 -0.012
## treatmentHigh:sex1 0.142 -0.613 -0.357 -0.711 0.059
```

```
## treatmentLow:sex1    0.172 -0.316 -0.667 -0.815  0.058  0.582
##
## Standardized residuals:
##      Min      Q1      Med      Q3      Max
## -5.16394307 -0.58298255 -0.05202491  0.62941251  2.85718985
##
## Residual standard error: 0.5980885
## Degrees of freedom: 322 total; 315 residual
```

```
anova(meanfull.hilo.nolitter,meanfull.hilo)
```

```
##              Model df      AIC      BIC    logLik    Test  L.Ratio
## meanfull.hilo.nolitter      1  9 489.6521 523.4252 -235.8260
## meanfull.hilo              2 10 381.0807 418.6065 -180.5404 1 vs 2 110.5713
##              p-value
## meanfull.hilo.nolitter
## meanfull.hilo              <.0001
```

3.6 Fitting the final model using ML

```
meanfull.hilo.ml <- lme(weight ~ treatment + sex1 + litsize + treatment:sex1, random = ~1 | litterid, reml = FALSE)
summary(meanfull.hilo.ml)
```

```
## Linear mixed-effects model fit by maximum likelihood
## Data: ratpup
##      AIC      BIC    logLik
##  357.1317 394.8773 -168.5659
##
## Random effects:
## Formula: ~1 | litterid
##      (Intercept) Residual
## StdDev:   0.2882595 0.5123784
##
## Variance function:
## Structure: Different standard deviations per stratum
## Formula: ~1 | trtgrp
## Parameter estimates:
##      1      2
## 1.0000000 0.5897706
## Fixed effects: weight ~ treatment + sex1 + litsize + treatment:sex1
##              Value Std.Error DF t-value p-value
## (Intercept)   8.350608 0.26150064 292 31.93341  0.0000
## treatmentHigh -0.904757 0.18092616  23 -5.00070  0.0000
## treatmentLow  -0.466869 0.15105108  23 -3.09080  0.0052
## sex1          -0.406590 0.09357754 292 -4.34495  0.0000
## litsize       -0.130402 0.01755814  23 -7.42689  0.0000
## treatmentHigh:sex1  0.093026 0.12521954 292  0.74290  0.4581
## treatmentLow:sex1  0.075602 0.10998665 292  0.68737  0.4924
## Correlation:
##      (Intr) trtmnH trtmnL sex1  litsiz trtH:1
```

```
## treatmentHigh      -0.606
## treatmentLow       -0.359  0.470
## sex1               -0.148  0.213  0.255
## litsize            -0.906  0.383  0.053  0.000
## treatmentHigh:sex1  0.097 -0.310 -0.190 -0.747  0.015
## treatmentLow:sex1   0.103 -0.172 -0.326 -0.851  0.025  0.636
##
## Standardized Within-Group Residuals:
##           Min           Q1           Med           Q3           Max
## -5.921828512 -0.521040722  0.004717487  0.567700472  2.571416659
##
## Number of Observations: 322
## Number of Groups: 27
```

```
summary(meanfull.hilo.ml)
```

```
## Linear mixed-effects model fit by maximum likelihood
## Data: ratpup
##      AIC      BIC    logLik
##  357.1317 394.8773 -168.5659
##
## Random effects:
## Formula: ~1 | litterid
##      (Intercept) Residual
## StdDev:   0.2882595 0.5123784
##
## Variance function:
## Structure: Different standard deviations per stratum
## Formula: ~1 | trtgrp
## Parameter estimates:
##           1           2
## 1.00000000 0.5897706
## Fixed effects: weight ~ treatment + sex1 + litsize + treatment:sex1
##           Value Std.Error DF t-value p-value
## (Intercept)   8.350608 0.26150064 292 31.93341 0.0000
## treatmentHigh -0.904757 0.18092616 23 -5.00070 0.0000
## treatmentLow  -0.466869 0.15105108 23 -3.09080 0.0052
## sex1          -0.406590 0.09357754 292 -4.34495 0.0000
## litsize       -0.130402 0.01755814 23 -7.42689 0.0000
## treatmentHigh:sex1 0.093026 0.12521954 292 0.74290 0.4581
## treatmentLow:sex1  0.075602 0.10998665 292 0.68737 0.4924
## Correlation:
##           (Intr) trtmnH trtmnL sex1  litsiz trtH:1
## treatmentHigh -0.606
## treatmentLow  -0.359  0.470
## sex1          -0.148  0.213  0.255
## litsize       -0.906  0.383  0.053  0.000
## treatmentHigh:sex1 0.097 -0.310 -0.190 -0.747  0.015
## treatmentLow:sex1  0.103 -0.172 -0.326 -0.851  0.025  0.636
##
## Standardized Within-Group Residuals:
##           Min           Q1           Med           Q3           Max
## -5.921828512 -0.521040722  0.004717487  0.567700472  2.571416659
##
```

```
## Number of Observations: 322
## Number of Groups: 27
```

```
anova(meanfull.hilo.ml)
```

##	numDF	denDF	F-value	p-value
## (Intercept)	1	292	10274.678	<.0001
## treatment	2	23	4.810	0.0180
## sex1	1	292	59.906	<.0001
## litsize	1	23	55.438	<.0001
## treatment:sex1	2	292	0.315	0.7303