ST517-HW6

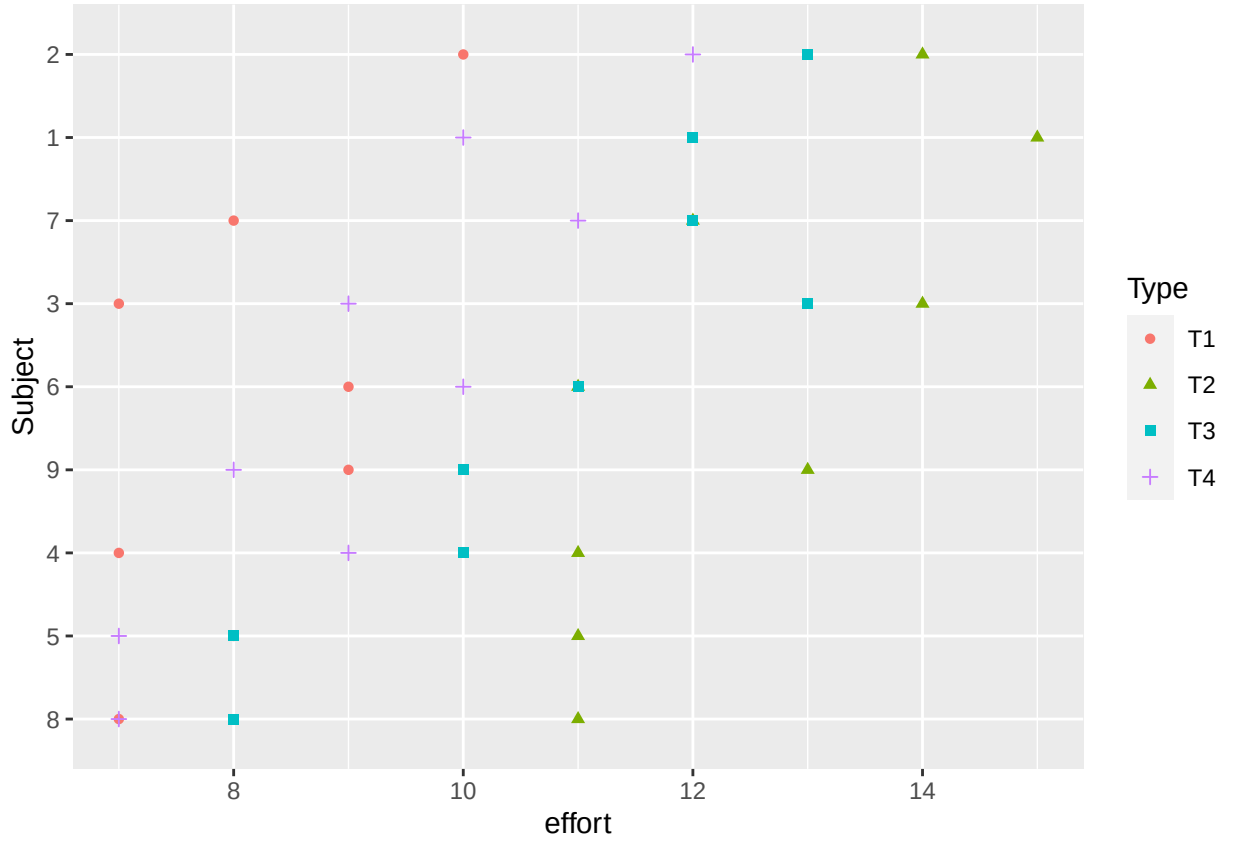
Nora Quick

1. The ergoStool dataset in the nlme package contains information on an experiment comparing the effort required by each of 9 Subjects to get up from each of 4 Types of stools.

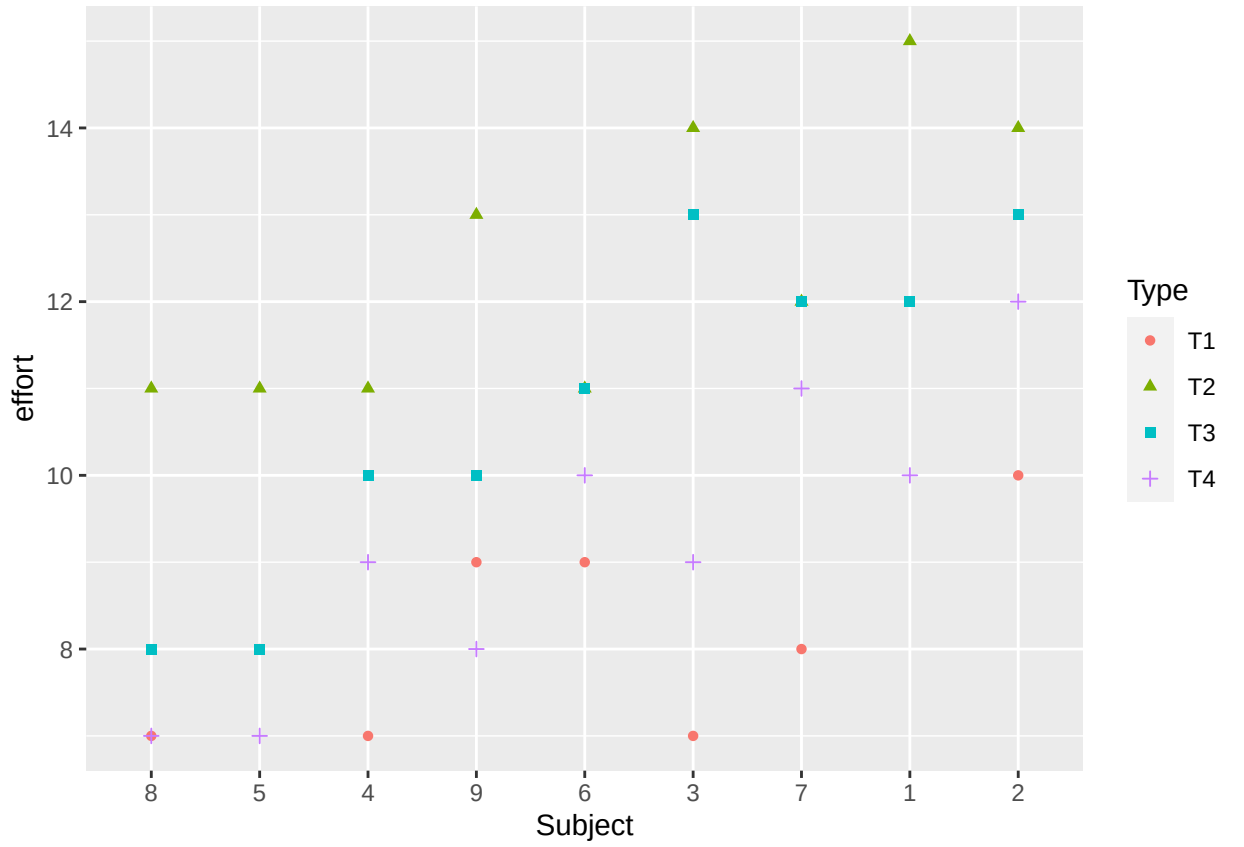
```
library(nlme)
?ergoStool
#ergoStool

effort <- ergoStool$effort
Type <- ergoStool$Type
Subject <- ergoStool$Subject
```

(a) Create at least one plot to explore the effects of Type and Subject on effort. Does there seem to be evidence of differences between stools? What about differences between subjects?

```
qplot(effort, Subject, data = ergoStool, colour = Type, shape = Type)
```



```
qplot(Subject, effort, data = ergoStool, colour = Type, shape = Type)
```



As we can see from both graphs above, there seems to be a difference between stools. We can see that T2 is always the highest on the effort scale and T1 is almost always the lowest.

As we can see from both graphs above, there seems to be a difference between subjects as well. 1 and 2 both have overall higher effort than the others while 8, 5, and 4 are on the lower end of the effort scale.

(b) Fit a fixed effects model with effects for Type and Subject. Use this model to answer the question: are there any significant differences in effort between stool Type T1 and the other stool types? (Remember to quantify your answer with the relevant confidence intervals).

```
fix_fit <- lm(effort ~ Type + Subject, data = ergoStool)
summary(fix_fit)
```

```
##
## Call:
## lm(formula = effort ~ Type + Subject, data = ergoStool)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.0556 -0.7361  0.2222  0.7361  1.7222
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)  8.55556    0.36676  23.327  < 2e-16 ***
## TypeT2       3.88889    0.51868   7.498 9.75e-08 ***
## TypeT3       2.22222    0.51868   4.284 0.000256 ***
## TypeT4       0.66667    0.51868   1.285 0.210951
## Subject.L    4.00208    0.55015   7.275 1.63e-07 ***
## Subject.Q    0.02849    0.55015   0.052 0.959128
## Subject.C    0.11124    0.55015   0.202 0.841468
## Subject^4    0.05587    0.55015   0.102 0.919949
## Subject^5   -0.57781    0.55015  -1.050 0.304047
## Subject^6   -0.33710    0.55015  -0.613 0.545808
## Subject^7   -0.27312    0.55015  -0.496 0.624100
## Subject^8   -0.26444    0.55015  -0.481 0.635099
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.1 on 24 degrees of freedom
## Multiple R-squared:  0.8356, Adjusted R-squared:  0.7603
## F-statistic: 11.09 on 11 and 24 DF,  p-value: 6.352e-07
```

There is strong evidence that there is a difference between Type 1 and Type 2 as well as Type 1 and Type 3 due to their significant p-values.

(c) Why would it be more appropriate to treat Subject as a random effect? How will the inferences about the stool types change if Subject is a random effect?

If we randomly assign subjects then we would have a better understanding of a population and stool types. If subject is the random effect then we can make inferences about stool types for entire populations.

(d) Fit a mixed effects model where Type is a fixed effect, and there are random intercepts for each Subject. Interpret the estimates of the random effect standard deviations in the context of the study.

```r
rand_fit <- lme(effort ~ Type, data = ergoStool, random = ~ 1 | Subject)
summary(rand_fit)
```

```
## Linear mixed-effects model fit by REML
##   Data: ergoStool
##        AIC      BIC    logLik
##   133.1308 141.9252 -60.56539
##
## Random effects:
##  Formula: ~1 | Subject
##         (Intercept) Residual
## StdDev:    1.332465 1.100295
##
## Fixed effects:  effort ~ Type
##                Value Std.Error DF   t-value p-value
## (Intercept) 8.555556 0.5760123 24 14.853079  0.0000
## TypeT2      3.888889 0.5186838 24  7.497610  0.0000
## TypeT3      2.222222 0.5186838 24  4.284348  0.0003
## TypeT4      0.666667 0.5186838 24  1.285304  0.2110
##  Correlation:
##        (Intr) TypeT2 TypeT3
```

```
## TypeT2 -0.45
## TypeT3 -0.45   0.50
## TypeT4 -0.45   0.50   0.50
##
## Standardized Within-Group Residuals:
##         Min          Q1         Med          Q3         Max
## -1.80200345 -0.64316591  0.05783115  0.70099706  1.63142054
##
## Number of Observations: 36
## Number of Groups: 9
```

The standard deviation of the random effects is estimated to be about 1.33 and the standard deviation of the residuals is estimated to be about 1.10. This shows that the variation can be attributed to the differences in Subject instead of variation in Subject.

(e) Fit a mixed effects model with just a single mean, and a random intercept for each subject. Compare this model with the model in part (d) using anova().

(i) Fit a mixed effects model with just a single mean, and a random intercept for each subject. Compa

```
mixed_fit <- lme(effort ~ 1, random = ~ 1 | Subject, data = ergoStool)

anova(rand_fit, mixed_fit)
```

```
## Warning in anova.lme(rand_fit, mixed_fit): fitted objects with different fixed
## effects. REML comparisons are not meaningful.
```

```
##           Model df      AIC      BIC    logLik   Test  L.Ratio p-value
## rand_fit      1  6 133.1308 141.9252 -60.56539
## mixed_fit     2  3 163.8380 168.5040 -78.91898 1 vs 2 36.70718  <.0001
```

(ii) What hypothesis is being tested? What do you conclude?

We are checking that the mean type of each subject is the same. In other words, we're testing between the two models is if the stools are different between subjects by checking a null hypothesis that they are the same.

(f) Repeat the inferences from part (b) using the mixed model from part (d).

```
rep_b <- lme(effort ~ Type, data = ergoStool, random = ~ 1 | Subject)
summary(rep_b)
```

```
## Linear mixed-effects model fit by REML
##   Data: ergoStool
##        AIC      BIC    logLik
##   133.1308 141.9252 -60.56539
##
## Random effects:
##  Formula: ~1 | Subject
##         (Intercept) Residual
```

```
## StdDev:    1.332465 1.100295
##
## Fixed effects:  effort ~ Type
##                Value Std.Error DF   t-value p-value
## (Intercept) 8.555556 0.5760123 24 14.853079  0.0000
## TypeT2      3.888889 0.5186838 24  7.497610  0.0000
## TypeT3      2.222222 0.5186838 24  4.284348  0.0003
## TypeT4      0.666667 0.5186838 24  1.285304  0.2110
##  Correlation:
##        (Intr) TypeT2 TypeT3
## TypeT2 -0.45
## TypeT3 -0.45   0.50
## TypeT4 -0.45   0.50   0.50
##
## Standardized Within-Group Residuals:
##         Min          Q1         Med          Q3         Max
## -1.80200345 -0.64316591  0.05783115  0.70099706  1.63142054
##
## Number of Observations: 36
## Number of Groups: 9
```

```
intervals(rep_b)
```

```
## Approximate 95% confidence intervals
##
##  Fixed effects:
##                  lower      est.    upper
## (Intercept)  7.3667247 8.5555556 9.744386
## TypeT2       2.8183781 3.8888889 4.959400
## TypeT3       1.1517114 2.2222222 3.292733
## TypeT4      -0.4038442 0.6666667 1.737177
## attr(,"label")
## [1] "Fixed effects:"
##
##  Random Effects:
##   Level: Subject
##                     lower     est.   upper
## sd((Intercept)) 0.7495009 1.332465 2.36886
##
##  Within-group standard error:
##     lower      est.     upper
## 0.8294505 1.1002946 1.4595785
```

Similar to (b) there is strong evidence that there is a difference between Type 1 and Type 2 as well as Type 1 and Type 3 due to their significant p-values.

With 95% confidence the mean effort of Type 1 is between 2.82 and 4.96 lower than Type 2. With 95% confidence the mean effort of Type 1 is between 1.15 and 3.29 lower than Type 3.

(g) Imagine you have a 10th subject.

(i) What would you predict for the mean effort of getting out of stool T1 for this subject? What would

Based on the lme() output above the mean effort of getting out of stool T1 for the 10th subject would be 8.56 with an interval between 7.37 and 9.74. This seems like a reasonable interval because it we are treating subjects at random and therefore we are making a conclusion for a whole populatoin above.

(ii) What would you predict for the difference in mean effort between getting out of stool T1 and T2 :

Based on the lme() output above the mean effort of getting out of stool T1 and T2 would be 12.5 with an interval of 10.2 and 14.7.