STOR 665 HW 7

Brian N. White

4/23/2021

Problem 7

(a)

Explain the data briefly.

```
#load data
ergoStool <- ergoStool

#examine the data structure
str(ergoStool)
```

```
## 'data.frame':    36 obs. of  3 variables:
##  $ effort : num  12 15 12 10 10 14 13 12 7 14 ...
##  $ Type   : Factor w/ 4 levels "T1","T2","T3",..: 1 2 3 4 1 2 3 4 1 2 ...
##  $ Subject: Factor w/ 9 levels "A","B","C","D",..: 1 1 1 1 2 2 2 2 3 3 ...
```

```
#data summary
summary(ergoStool)
```

```
##      effort      Type      Subject
##  Min.   : 7.00   T1:9   A      : 4
##  1st Qu.: 8.00   T2:9   B      : 4
##  Median :10.00   T3:9   C      : 4
##  Mean   :10.25   T4:9   D      : 4
##  3rd Qu.:12.00          E      : 4
##  Max.   :15.00          F      : 4
##                         (Other):12
```

Inspection of the contingency table below reveals that the factors 'Type' and 'Subject' are completely crossed (i.e. there is at least one observation for each combination of factor levels). Further, this is an unreplicated design.

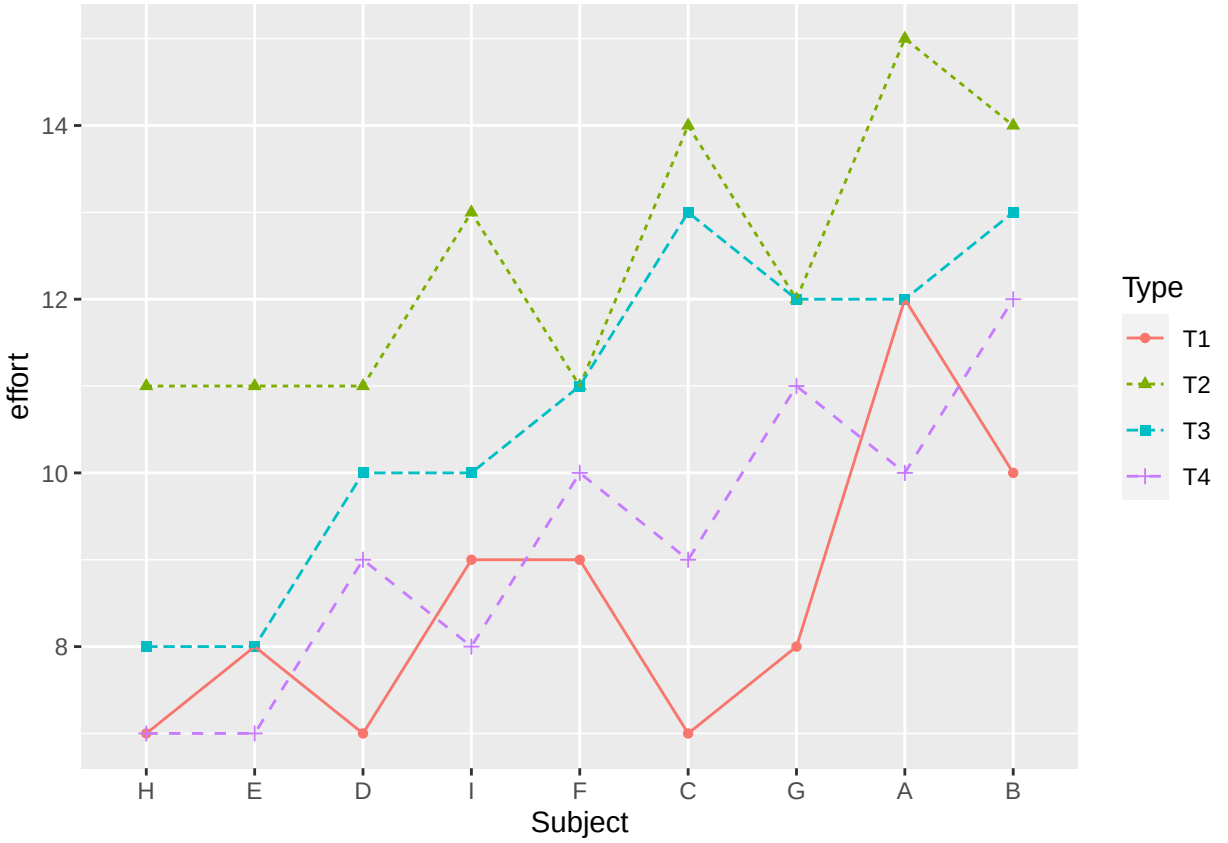
```
xtabs(~Type + Subject, ergoStool)
```

```
##     Subject
## Type A B C D E F G H I
##   T1 1 1 1 1 1 1 1 1 1
##   T2 1 1 1 1 1 1 1 1 1
##   T3 1 1 1 1 1 1 1 1 1
##   T4 1 1 1 1 1 1 1 1 1
```

(b)

```
#determine the average effort for each level of the factor Subject
ergoStool %>%
  group_by(Subject) %>%
  summarize(avg_effort=mean(effort)) -> avg_effort_df

ergoStool %>%
  mutate(avg_effort=rep(avg_effort_df$avg_effort, each=4, times=1)) %>%
  mutate(Subject=fct_reorder(Subject, avg_effort)) %>%
  ggplot(aes(x=Subject, y=effort, shape=Type, color=Type)) +
  geom_point() +
  geom_line(aes(x=Subject, y=effort, group=Type, color=Type, linetype=Type))
```



(c)

A linear mixed model with random effects for Type and Subject is fit below. The standard deviations for the estimates corresponding to Type, Subject and residual variability are 1.332, 1.695, and 1.100, as discerned from the output below.

```
#fit a model with random effects for Type and Subject via REML
summary(lmm_ergo <- lmer(effort ~ 1 + (1|Type) + (1|Subject), ergoStool))
```

```
## Linear mixed model fit by REML ['lmerMod']
```

```
## Formula: effort ~ 1 + (1 | Type) + (1 | Subject)
##    Data: ergoStool
##
## REML criterion at convergence: 134.3
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.87089 -0.70269  0.08069  0.68483  1.65288
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept) 1.775    1.332
##  Type     (Intercept) 2.873    1.695
##  Residual             1.211    1.100
## Number of obs: 36, groups:  Subject, 9; Type, 4
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  10.2500     0.9742   10.52
```

(d)

The model from part (c) is refit using maximum likelihood estimation. The standard deviations for the estimates corresponding to Type, Subject and residual variability are now 1.305, 1.505, and 1.101.

```
summary(lmm_ergo2 <- update(lmm_ergo, REML=FALSE))
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: effort ~ 1 + (1 | Type) + (1 | Subject)
##    Data: ergoStool
##
##      AIC      BIC   logLik deviance df.resid
##    144.0    150.4    -68.0    136.0       32
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.88436 -0.68813  0.06718  0.67820  1.65934
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept) 1.704    1.305
##  Type     (Intercept) 2.265    1.505
##  Residual             1.213    1.101
## Number of obs: 36, groups:  Subject, 9; Type, 4
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  10.2500     0.8883   11.54
```

(e)

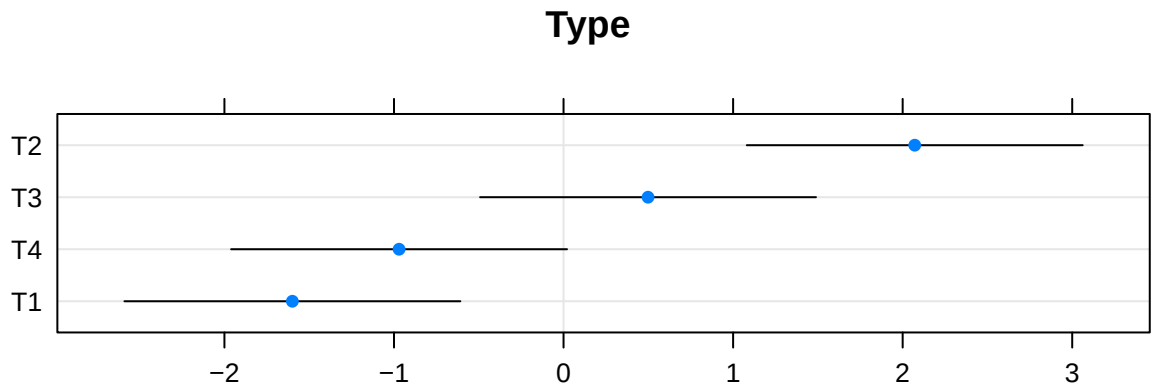
In the code-chunk below, the 95% prediction intervals corresponding to the random effects of the model from part (d) (i.e. the model fit via MLE) are generated. Based upon this plot, it is clear that stool type 1

(i.e. T1) outperforms the other types w.r.t. effort minimization.

```
## Substitute the name of your fitted model for fm in the call to ranef)
dotplot(ranef(lmm_ergo2, which = "Type", postVar = TRUE), aspect = 0.2, strip = FALSE)
```

```
## Warning in ranef.merMod(lmm_ergo2, which = "Type", postVar = TRUE): 'postVar' is
## deprecated: please use 'condVar' instead
```

```
## $Type
```



(f)

The significance of the random effect Type is assessed by comparing the model fit in Part (d) with a reduced model with only Subject as the random effect. Inspection of the anova output indicates a p-value of ~0. Thus, we fail to reject the null hypothesis $H_0 : \sigma_2 = 0$.

```
summary(lmm_ergo3 <- lmer(effort ~ 1 + (1|Subject), ergoStool, REML=FALSE))
```

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: effort ~ 1 + (1 | Subject)
##    Data: ergoStool
##
##      AIC      BIC   logLik deviance df.resid
```

```
##    164.2    168.9    -79.1    158.2       33
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.71903 -0.72603  0.01465  0.75533  1.90786
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept) 0.8264   0.9091
##  Residual             4.0833   2.0207
## Number of obs: 36, groups:  Subject, 9
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)   10.250      0.453   22.62
```

```
anova(lmm_ergo3, lmm_ergo2)
```

```
## Data: ergoStool
## Models:
## lmm_ergo3: effort ~ 1 + (1 | Subject)
## lmm_ergo2: effort ~ 1 + (1 | Type) + (1 | Subject)
##           npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## lmm_ergo3    3 164.15 168.90 -79.075   158.15
## lmm_ergo2    4 144.02 150.36 -68.011   136.02 22.128  1  2.551e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(g)

Since we are interested in making inference about four stool types, a more reasonable model is to treat Type as fix effects and Subject as the random effect. Fit this mixed effect model using maximum likelihood. Calculate the fixed-effects predictors for the four different stool types. Compare the results with that in Part (e).