Final STOR 590

Ted Henson

4/28/2020

I will neither give nor receive unauthorized aid on this exam Ted Henson

Question 1

a)

```
## # A tibble: 6 x 5
##   Treat Judge Replicate   Day Score
##   <chr> <dbl>     <dbl> <dbl> <dbl>
## 1 0         1         1     1     2
## 2 0         1         1     2     4
## 3 dcs       1         1     1     4
## 4 dcs       1         1     2     4
## 5 d         2         1     1     5
## 6 d         2         1     2     4
```

```
## # A tibble: 8 x 2
##   Treat mean.score
##   <chr>      <dbl>
## 1 dcs         3.5
## 2 dc          3.25
## 3 s           3.12
## 4 0           3
## 5 ds          3
## 6 d           2.75
## 7 c           2.5
## 8 cs          2.5
```

Treatment dcs had the highest mean score by the judges.

b)

This method would not work as there would be too many factor variables and not enough observations. There would be one observation per nested factored level so the regression would break as the number of features would exceed the number of observations.

c)

```
##               Df Sum Sq Mean Sq F value  Pr(>F)
## factor(Judge) 15  31.61   2.107   2.698 0.00599 **
## factor(Treat)  7  11.23   1.604   2.054 0.07102 .
## Residuals     41  32.02   0.781
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##  factor(Judge)10   factor(Judge)6   factor(Judge)4  factor(Judge)12
##      -2.82291667      -1.77083333      -1.31250000      -1.11458333
##   factor(Judge)3  factor(Judge)13  factor(Judge)14  factor(Treat)cs
##      -0.81250000      -0.59375000      -0.53125000      -0.45833333
##  factor(Judge)11   factor(Judge)8   factor(Judge)2   factor(Treat)c
##      -0.42708333      -0.39583333      -0.37500000      -0.14583333
##   factor(Judge)9   factor(Judge)7  factor(Treat)ds   factor(Judge)5
##      -0.13541667       0.02083333       0.10416667       0.39583333
##  factor(Judge)15  factor(Judge)16 factor(Treat)dcs   factor(Treat)s
##       0.53125000       0.59375000       0.83333333       0.93750000
##  factor(Treat)dc   factor(Treat)d      (Intercept)
##       1.02083333       1.04166667       3.08333333
```

Treatment cs shapes up the best in terms of minimizing the mean score compared to other treatments. This gives a different result than a) because of the effects posed by the judges.

d)

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Score ~ Treat + (1 | Judge)
##    Data: soap
##
## REML criterion at convergence: 178.2
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.0502 -0.5994  0.1411  0.5490  1.8070
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Judge    (Intercept) 0.5074   0.7123
##  Residual             0.7809   0.8837
## Number of obs: 64, groups:  Judge, 16
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  2.78468    0.41533   6.705
## Treatc      -0.39711    0.56464  -0.703
## Treatcs     -0.52037    0.52145  -0.998
## Treatd       0.43058    0.56464   0.763
## Treatdc      0.69957    0.52145   1.342
## Treatdcs     0.59511    0.53401   1.114
## Treatds      0.06286    0.52145   0.121
## Treats       0.47691    0.56464   0.845
##
## Correlation of Fixed Effects:
##          (Intr) Treatc Tretcs Treatd Tretdc Trtdcs Tretds
## Treatc   -0.680
## Treatcs  -0.628  0.462
## Treatd   -0.680  0.574  0.519
## Treatdc  -0.628  0.462  0.500  0.462
## Treatdcs -0.643  0.551  0.428  0.551  0.428
## Treatds  -0.628  0.519  0.500  0.462  0.500  0.428
```

```
## Treats   -0.680  0.574  0.462  0.574  0.519  0.551  0.462
```

In treating the judge as a random effect, treatment cs again would expect to yield the lowest score compared to other treatments. This would not be the case if two treatments were subject to different judges.

e)

```
##
## Call:
## glm(formula = success ~ factor(Treat) + factor(Judge), family = "binomial",
##     data = soap)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -1.17741  -0.58633  -0.00005   0.00006   2.22238
##
## Coefficients:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)       6.564e-15  1.414e+00   0.000    1.000
## factor(Treat)c   -1.020e+00  2.799e+00  -0.364    0.716
## factor(Treat)cs   7.938e-02  1.872e+00   0.042    0.966
## factor(Treat)d   -4.002e+01  8.296e+03  -0.005    0.996
## factor(Treat)dc  -2.039e+00  1.839e+00  -1.109    0.267
## factor(Treat)dcs -7.878e+01  1.284e+04  -0.006    0.995
## factor(Treat)ds  -1.020e+00  1.959e+00  -0.521    0.603
## factor(Treat)s    1.744e+01  7.286e+03   0.002    0.998
## factor(Judge)2    1.997e+01  8.013e+03   0.002    0.998
## factor(Judge)3    1.020e+00  2.799e+00   0.364    0.716
## factor(Judge)4   -1.744e+01  7.286e+03  -0.002    0.998
## factor(Judge)5   -1.139e+00  2.060e+00  -0.553    0.580
## factor(Judge)6    5.940e+01  1.081e+04   0.005    0.996
## factor(Judge)7   -3.772e+01  1.307e+04  -0.003    0.998
## factor(Judge)8    3.655e-01  2.484e+00   0.147    0.883
## factor(Judge)9   -6.541e-01  1.981e+00  -0.330    0.741
## factor(Judge)10   6.042e+01  1.418e+04   0.004    0.997
## factor(Judge)11  -1.926e+01  1.089e+04  -0.002    0.999
## factor(Judge)12   9.799e-01  2.426e+00   0.404    0.686
## factor(Judge)13  -3.418e-01  1.939e+00  -0.176    0.860
## factor(Judge)14   2.052e+01  6.659e+03   0.003    0.998
## factor(Judge)15  -3.772e+01  1.309e+04  -0.003    0.998
## factor(Judge)16  -2.114e+01  1.387e+04  -0.002    0.999
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 79.499  on 63  degrees of freedom
## Residual deviance: 32.654  on 41  degrees of freedom
## AIC: 78.654
##
## Number of Fisher Scoring iterations: 20
```

```
##  factor(Judge)10   factor(Judge)6  factor(Judge)14   factor(Judge)2
##     6.041912e+01     5.940458e+01     2.052226e+01     1.997276e+01
##   factor(Treat)s   factor(Judge)3  factor(Judge)12   factor(Judge)8
##     1.744123e+01     1.019623e+00     9.799312e-01     3.655377e-01
##  factor(Treat)cs      (Intercept)  factor(Judge)13   factor(Judge)9
```

```
##     7.938332e-02     6.563925e-15    -3.418057e-01    -6.540851e-01
##   factor(Treat)c  factor(Treat)ds   factor(Judge)5  factor(Treat)dc
##    -1.019623e+00    -1.019623e+00    -1.138698e+00    -2.039246e+00
##   factor(Judge)4  factor(Judge)11  factor(Judge)16   factor(Judge)7
##    -1.744123e+01    -1.926408e+01    -2.113523e+01    -3.771734e+01
##  factor(Judge)15   factor(Treat)d factor(Treat)dcs
##    -3.772233e+01    -4.002490e+01    -7.878427e+01
```

Treatment s would yield the largest probability of a successful result. The problem with this model is that there are too many factored variables and not enough observations, resulting in large standard errors and p values so it is difficult to conclude which treatment is best from this model.

f)

```
## F-test with Kenward-Roger approximation; computing time: 0.10 sec.
## large : Score ~ Treat + (1 | Judge)
## small : Score ~ 1 + (1 | Judge)
##          stat     ndf     ddf F.scaling p.value
## Ftest  1.4388  7.0000 53.5164   0.99783  0.2094
```

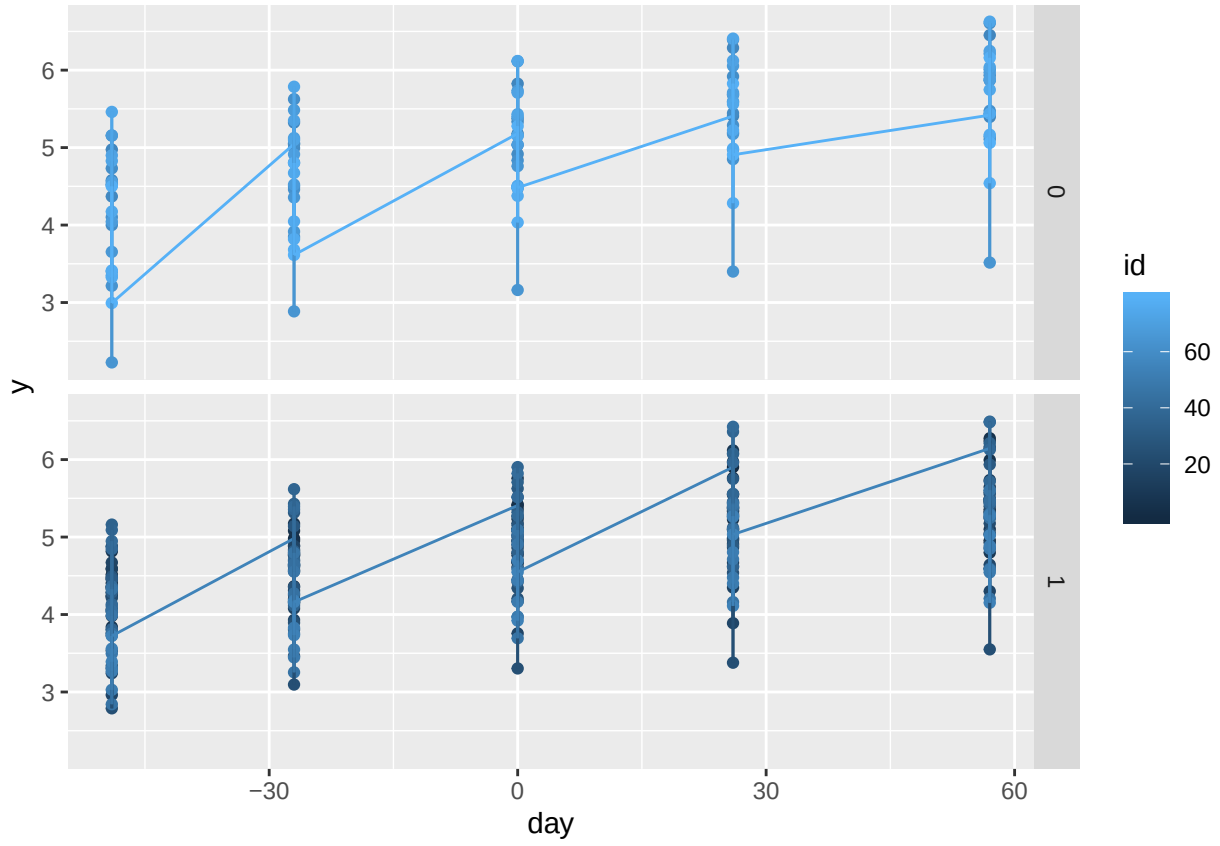
```
## Parametric bootstrap test; time: 19.92 sec; samples: 1000 extremes: 291;
## large : Score ~ Treat + (1 | Judge)
## small : Score ~ 1 + (1 | Judge)
##            stat     df    ddf p.value
## PBtest   9.7814                0.2917
## Gamma    9.7814                0.2861
## Bartlett 8.5511 7.0000         0.2865
## F        1.3973 7.0000 2.2854  0.4611
## LRT      9.7814 7.0000         0.2013
```

The Kenward-Rodger test and parametric bootstrap test both do not show statistical significance for the effects of the treatments. Treatment cs may be the treatment yielding the lowest score, but there is not statistical evidence to conclude that considering the large standard error in the mixed effects model and the lack of statistical evidence for effects of the treatments.

Question 2

a)

```
## # A tibble: 6 x 5
##       y   day    tx chamber    id
##   <dbl> <dbl> <dbl>   <dbl> <dbl>
## 1  4.51   -49     1       1     1
## 2  4.98   -27     1       1     1
## 3  5.41     0     1       1     1
## 4  5.90    26     1       1     1
## 5  6.14    57     1       1     1
## 6  4.24   -49     1       1     2
```



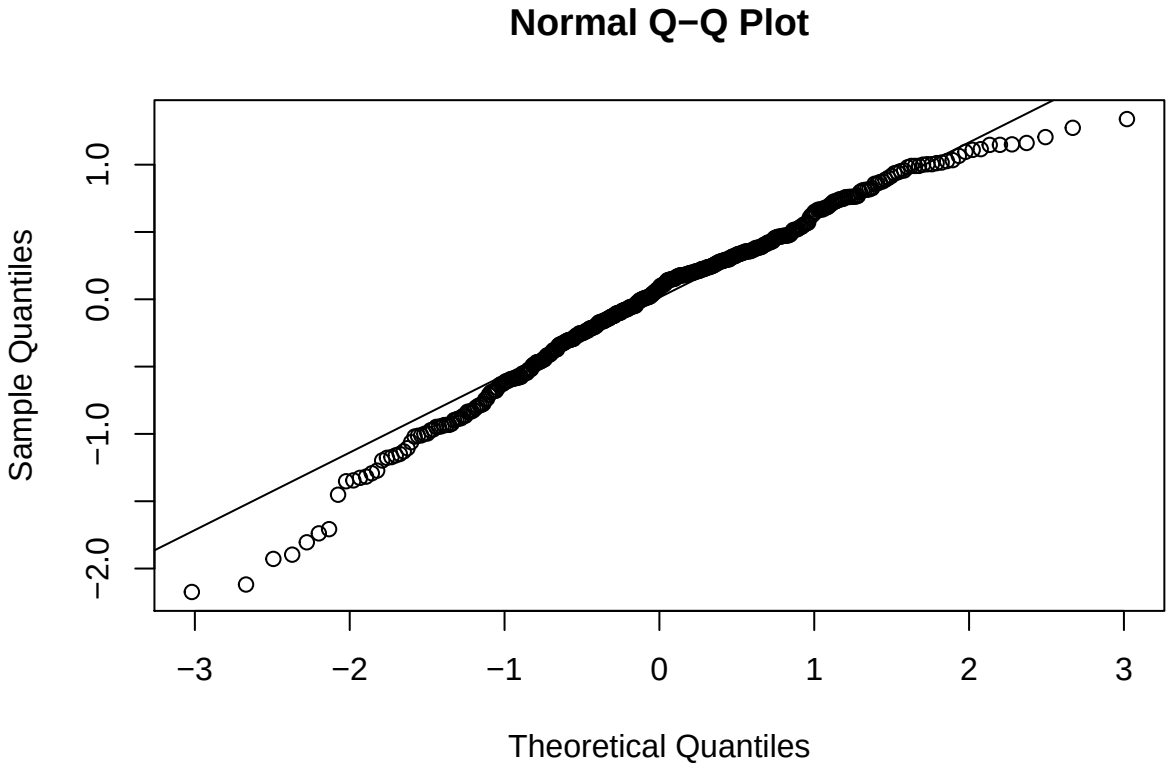
The size of the tree does appear to be larger for the lower ozone environment as well as the growth rate. The growth rate of the high ozone environment gradually levels off and almost becomes flat. The total size of most trees is less at each day for the higher ozone trees compared to the lower ozone trees.

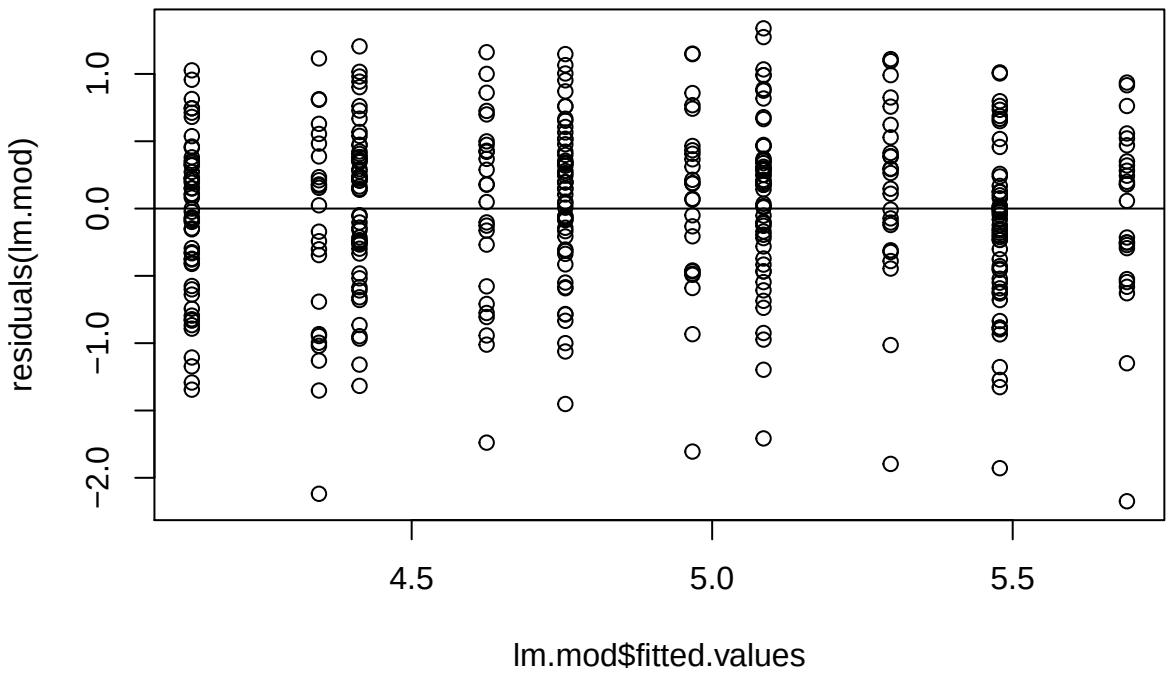
b)

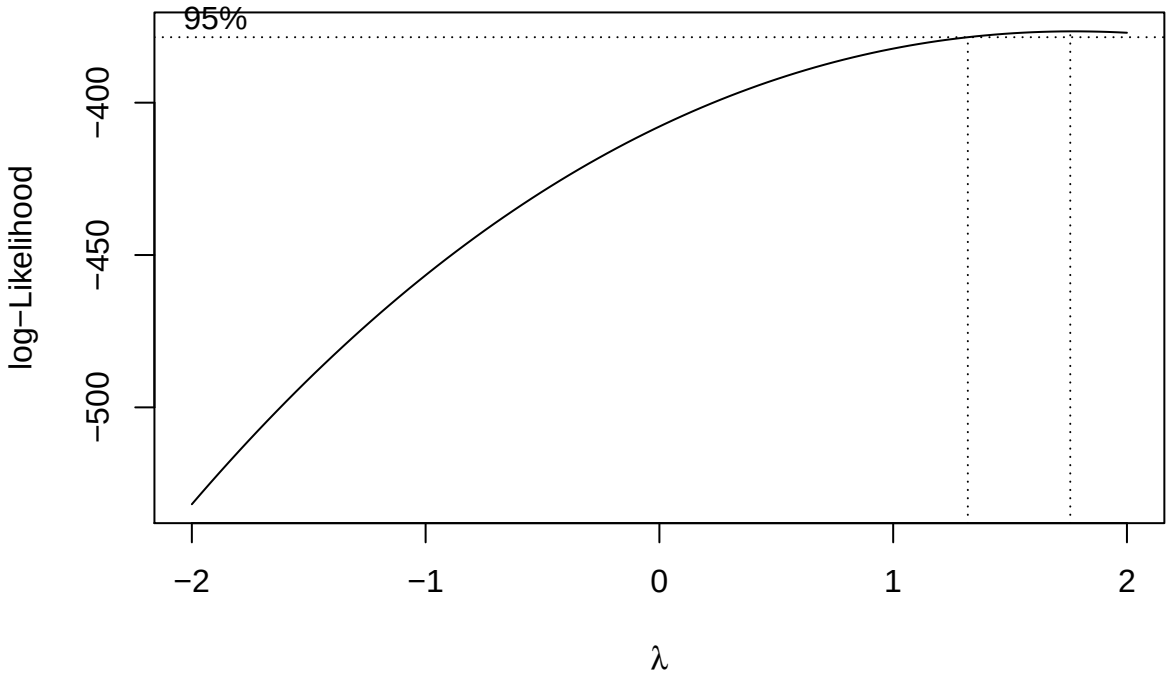
```
##
## Call:
## lm(formula = y ~ day + factor(chamber), data = spruce)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.08321 -0.37517  0.08635  0.40459  1.29889
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)      4.7160298  0.0544872  86.553  < 2e-16 ***
## day              0.0126824  0.0008484  14.949  < 2e-16 ***
## factor(chamber)2 0.0793333  0.0770382   1.030 0.303746
## factor(chamber)3 0.1602815  0.0982049   1.632 0.103463
## factor(chamber)4 0.3350610  0.0955542   3.507 0.000507 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6329 on 390 degrees of freedom
## Multiple R-squared:  0.3774, Adjusted R-squared:  0.3711
```

```
## F-statistic: 59.11 on 4 and 390 DF,  p-value: < 2.2e-16
```

```
##
## Call:
## lm(formula = y ~ day + factor(tx), data = spruce)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.17410 -0.37493  0.08074  0.40202  1.33856
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.96720    0.05673  87.562  < 2e-16 ***
## day          0.01268    0.00085  14.921  < 2e-16 ***
## factor(tx)1 -0.21150    0.06860  -3.083  0.00219 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6341 on 392 degrees of freedom
## Multiple R-squared:  0.3719, Adjusted R-squared:  0.3687
## F-statistic: 116.1 on 2 and 392 DF,  p-value: < 2.2e-16
```




residuals(lm.mod)
1.0
0.0
−1.0
−2.0
4.5
5.0
5.5
lm.mod$fitted.values




```
##
##  Shapiro-Wilk normality test
##
## data:  spruce$y
## W = 0.99088, p-value = 0.01526
```

```
##
##  Shapiro-Wilk normality test
##
## data:  1/spruce$y
## W = 0.90054, p-value = 2.177e-15
```

Since there were only low ozone observations for chamber 3 and 4, only separate models could be built to compare the effects of tx and the chamber. Based on the p values it appears that the tx and day is significant, but the chambers are not so the model with days and tx will be used for analysis. The residuals looked skewed at the higher and lower ends of the response variable. A box cox transformation would yield a lambda of about 1.5. The Shapiro test for normality of the transformed response yielded a p value closer to zero than the standard response.

c)

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: y ~ day + factor(tx) + (1 | chamber) + (1 | id) + (1 | chamber:id)
##    Data: spruce
##
## REML criterion at convergence: 164.8
##
```

```
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.0028 -0.5212  0.1663  0.5545  4.3846
##
## Random effects:
##  Groups     Name        Variance Std.Dev.
##  chamber:id (Intercept) 0.33358  0.5776
##  id         (Intercept) 0.03587  0.1894
##  chamber    (Intercept) 0.00000  0.0000
##  Residual               0.03922  0.1980
## Number of obs: 395, groups:  chamber:id, 79; id, 79; chamber, 4
##
## Fixed effects:
##               Estimate Std. Error t value
## (Intercept)  4.9671966  0.1228498  40.433
## day          0.0126824  0.0002655  47.776
## factor(tx)1 -0.2115001  0.1485899  -1.423
##
## Correlation of Fixed Effects:
##             (Intr) day
## day         -0.003
## factor(tx)1 -0.827  0.000
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

The fixed effect term of the intercept would be 4.9671966 with a standard error of 0.1228498, the days term would be 0.0126824 with a standard error of 0.0002655 and the fixed effects of tx would be -0.2115001 with a standard error of 0.1485899. Theses coefficients are more or less the same compared to the previous model, although there are higher standard errors for the intercept and tx in this model.

d)

```
## Parametric bootstrap test; time: 30.76 sec; samples: 1000 extremes: 197;
## Requested samples: 1000 Used samples: 197 Extremes: 197
## large : y ~ day + factor(tx) + (1 | chamber) + (1 | id) + (1 | chamber:id)
## small : y ~ day + factor(tx) + (1 | id)
##          stat          df    ddf p.value
## PBtest      0                           1
## Gamma       0                           1
## Bartlett    0  2.0000e+00             1
## F           0  2.0000e+00 -1.333
## LRT         0  2.0000e+00             1
##
##  simulated finite sample distribution of RLRT.
##
##  (p-value based on 10000 simulated values)
##
## data:
## RLRT = 613.91, p-value < 2.2e-16
```

```
## Parametric bootstrap test; time: 30.63 sec; samples: 1000 extremes: 0;
## Requested samples: 1000 Used samples: 448 Extremes: 0
## large : y ~ day + factor(tx) + (1 | chamber) + (1 | id) + (1 | chamber:id)
## small : y ~ day + factor(tx) + (1 | chamber)
##             stat      df     ddf   p.value
```

```
## PBtest    614.09                  0.002227 **
## Gamma     614.09                 < 2.2e-16 ***
## Bartlett 1508.35    2.00         < 2.2e-16 ***
## F         307.05    2.00 -8.7674
## LRT       614.09    2.00         < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
##  simulated finite sample distribution of RLRT.
##
##  (p-value based on 10000 simulated values)
##
## data:
## RLRT = 5.4001e-13, p-value = 0.3487
```

The RLRT test show that the random effect posed by the id of the tree is not significant, but the parametric boostrap does. For the chamber, the RLRT shows it is significant, but the parametric boot strap test does not.

e)

```
## F-test with Kenward-Roger approximation; computing time: 0.06 sec.
## large : y ~ day + factor(tx) + (1 | id)
## small : y ~ day + (1 | id)
##         stat    ndf    ddf F.scaling p.value
## Ftest  2.026  1.000 77.000         1  0.1587
```

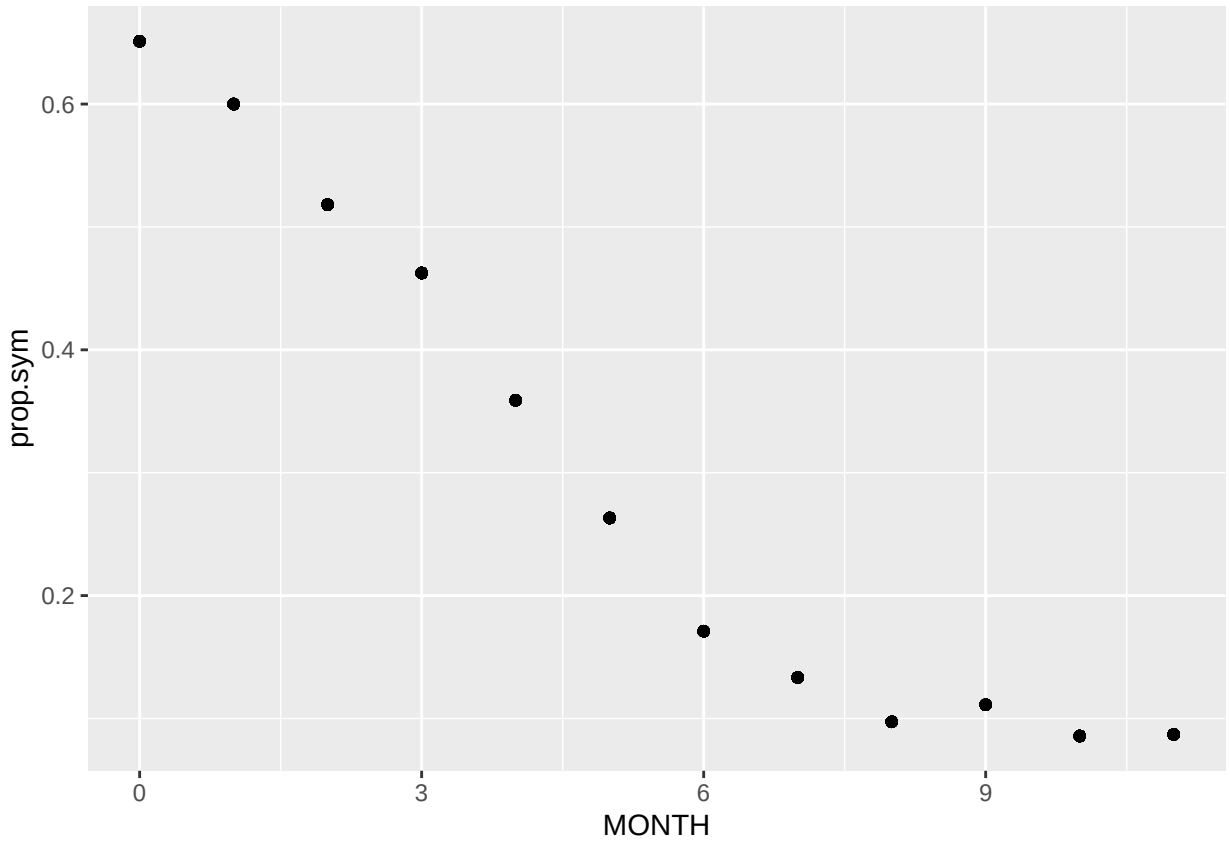
The Kenward-Rodgers test does not show statistical significance for the effect of the tx variable, but it was fairly close so it should not be thrown out of discussion completely.

f)

One disadvantage to a random effects model is that it decreases interpretability, particularly for non statistician oriented audiences; however, without doing the random effects analysis one may conclude for certain that the tx variable is significant and causes differences in the growth of a tree. As shown by the Kenward-Rodgers test, we do not currently have statistical evidence to conclude this. So a random effects model can disprove relationships one may think are significant based on a fixed effects model, uncover significant effects a fixed effects model may not find, but at the cost of interpretability.

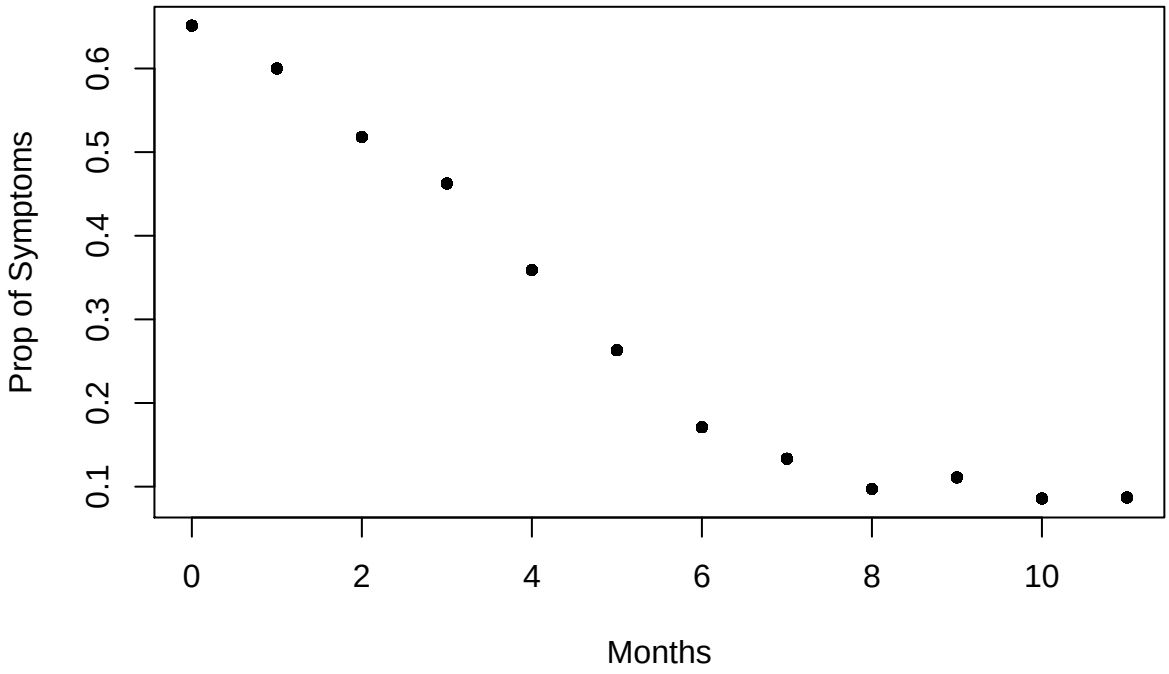
Question 3

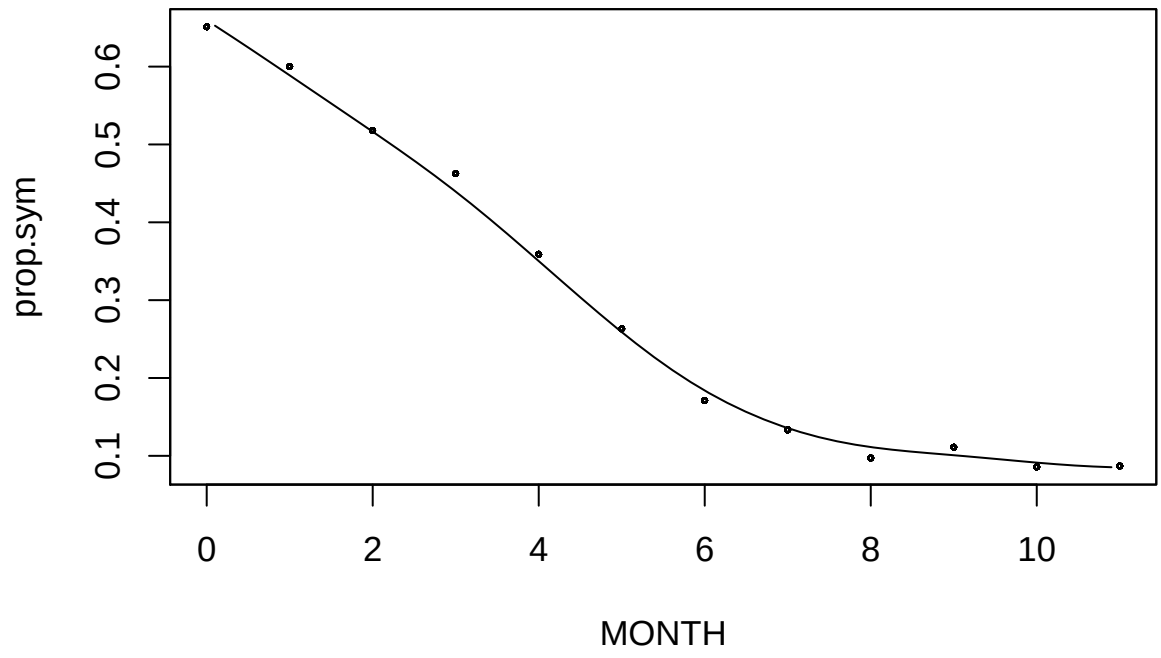
a)

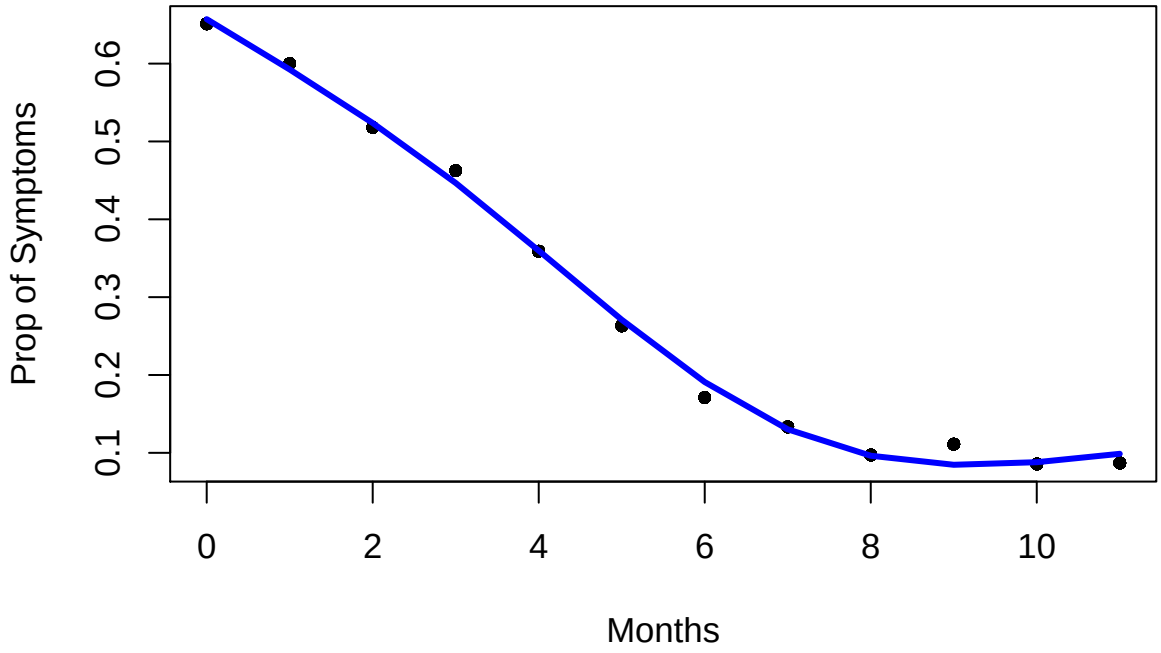


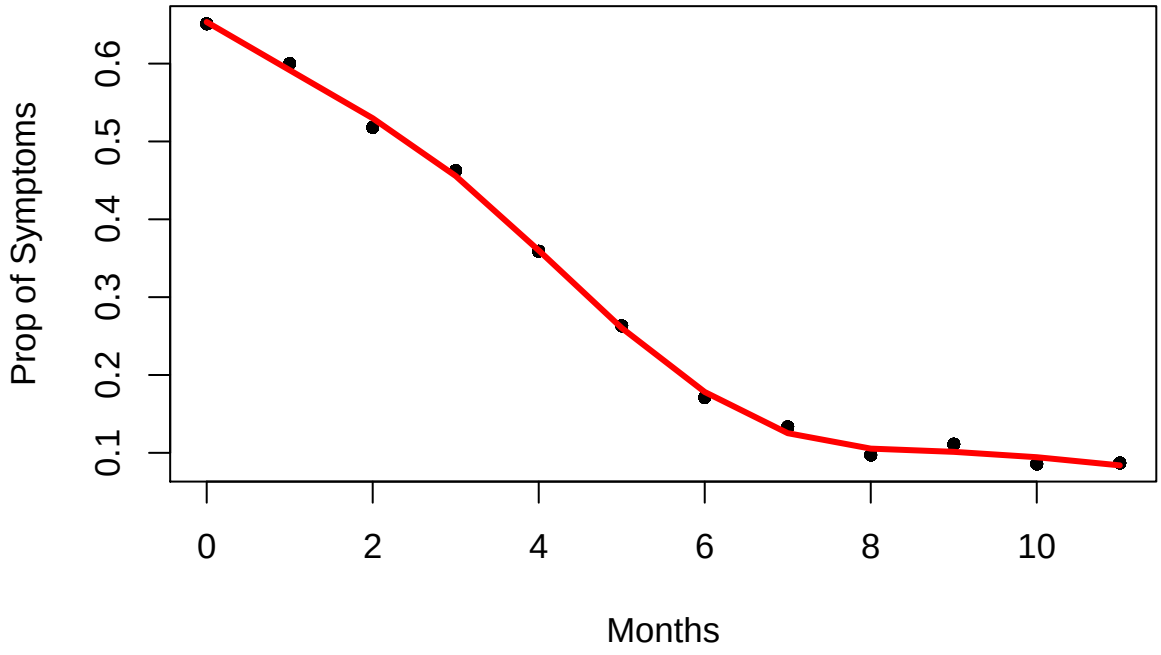
The proportion of patients exhibiting symptoms decreases over the months in a linear fashion, but begins to level off by month 8.

b)


Prop of Symptoms
0.1
0.2
0.3
0.4
0.5
0.6
0
2
4
6
8
10
Months


prop.sym
0.1
0.2
0.3
0.4
0.5
0.6
0
2
4
6
8
10
MONTH


Prop of Symptoms
0.1
0.2
0.3
0.4
0.5
0.6
0
2
4
6
8
10
Months



6 degrees of freedom makes the regression spline curve approximately the same as the fitted smoothed curve, but other larger numbers do fairly well as well.

c)

```
##
## Call:
## glm(formula = prop.sym ~ factor(AGE) + factor(GENDER) + factor(MONTH),
##     family = "binomial", data = schiz)
##
## Deviance Residuals:
##        Min          1Q      Median          3Q         Max
## -1.702e-08   0.000e+00   0.000e+00   8.013e-09   1.333e-08
##
## Coefficients:
##                   Estimate Std. Error z value Pr(>|z|)
## (Intercept)      6.242e-01  2.437e-01   2.561 0.010446 *
## factor(AGE)1    -7.859e-17  1.684e-01   0.000 1.000000
## factor(GENDER)1  1.901e-16  1.618e-01   0.000 1.000000
## factor(MONTH)1  -2.187e-01  3.166e-01  -0.691 0.489675
## factor(MONTH)2  -5.518e-01  3.154e-01  -1.750 0.080137 .
## factor(MONTH)3  -7.744e-01  3.186e-01  -2.431 0.015054 *
## factor(MONTH)4  -1.204e+00  3.270e-01  -3.682 0.000231 ***
## factor(MONTH)5  -1.654e+00  3.451e-01  -4.793 1.65e-06 ***
## factor(MONTH)6  -2.202e+00  3.795e-01  -5.804 6.49e-09 ***
## factor(MONTH)7  -2.496e+00  4.082e-01  -6.115 9.66e-10 ***
```

```
## factor(MONTH)8  -2.853e+00  4.577e-01  -6.233 4.58e-10 ***
## factor(MONTH)9  -2.704e+00  4.380e-01  -6.173 6.72e-10 ***
## factor(MONTH)10 -2.991e+00  4.832e-01  -6.190 6.01e-10 ***
## factor(MONTH)11 -2.976e+00  4.835e-01  -6.154 7.55e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1.9348e+02  on 921  degrees of freedom
## Residual deviance: 3.8241e-14  on 908  degrees of freedom
## AIC: 619.69
##
## Number of Fisher Scoring iterations: 5
```

```
##
## Call:
## glm(formula = prop.sym ~ factor(AGE) + factor(GENDER) + factor(MONTH),
##     family = "quasibinomial", data = schiz)
##
## Deviance Residuals:
##        Min          1Q      Median          3Q         Max
## -1.702e-08   0.000e+00   0.000e+00   8.013e-09   1.333e-08
##
## Coefficients:
##                   Estimate Std. Error    t value Pr(>|t|)
## (Intercept)      6.242e-01  7.463e-14  8.363e+12   <2e-16 ***
## factor(AGE)1    -7.859e-17  5.156e-14 -2.000e-03    0.999
## factor(GENDER)1  1.901e-16  4.956e-14  4.000e-03    0.997
## factor(MONTH)1  -2.187e-01  9.693e-14 -2.256e+12   <2e-16 ***
## factor(MONTH)2  -5.518e-01  9.656e-14 -5.715e+12   <2e-16 ***
## factor(MONTH)3  -7.744e-01  9.754e-14 -7.939e+12   <2e-16 ***
## factor(MONTH)4  -1.204e+00  1.001e-13 -1.202e+13   <2e-16 ***
## factor(MONTH)5  -1.654e+00  1.057e-13 -1.565e+13   <2e-16 ***
## factor(MONTH)6  -2.202e+00  1.162e-13 -1.895e+13   <2e-16 ***
## factor(MONTH)7  -2.496e+00  1.250e-13 -1.997e+13   <2e-16 ***
## factor(MONTH)8  -2.853e+00  1.401e-13 -2.036e+13   <2e-16 ***
## factor(MONTH)9  -2.704e+00  1.341e-13 -2.016e+13   <2e-16 ***
## factor(MONTH)10 -2.991e+00  1.480e-13 -2.022e+13   <2e-16 ***
## factor(MONTH)11 -2.976e+00  1.481e-13 -2.010e+13   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for quasibinomial family taken to be 9.375881e-26)
##
##     Null deviance: 1.9348e+02  on 921  degrees of freedom
## Residual deviance: 3.8241e-14  on 908  degrees of freedom
## AIC: NA
##
## Number of Fisher Scoring iterations: 5
```

Age and Gender do not appear to have an effect on the probability of developing symptoms. The quasi binomial model reached the same conclusion to an even greater degree. All month variables appear to be significant in the quasi binomial model. All month variables aside from months 1 and 2 are significant in the standard logistic regression model.

d)

```
## Generalized linear mixed model fit by maximum likelihood (Adaptive
##   Gauss-Hermite Quadrature, nAGQ = 25) [glmerMod]
##  Family: binomial  ( logit )
## Formula: Y ~ factor(MONTH) + factor(AGE) + factor(GENDER) + (1 | ID)
##    Data: schiz
##
##      AIC      BIC   logLik deviance df.resid
##    769.8    842.2   -369.9    739.8      907
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -6.9825 -0.3560 -0.1366  0.2979  6.6892
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  ID     (Intercept) 5.26     2.293
## Number of obs: 922, groups:  ID, 86
##
## Fixed effects:
##                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)       1.4033     0.5095   2.754  0.00588 **
## factor(MONTH)1   -0.3281     0.4328  -0.758  0.44836
## factor(MONTH)2   -0.8660     0.4340  -1.996  0.04599 *
## factor(MONTH)3   -1.3107     0.4427  -2.961  0.00307 **
## factor(MONTH)4   -2.0963     0.4586  -4.571 4.84e-06 ***
## factor(MONTH)5   -2.9096     0.4867  -5.979 2.25e-09 ***
## factor(MONTH)6   -3.7543     0.5300  -7.084 1.40e-12 ***
## factor(MONTH)7   -4.2476     0.5692  -7.462 8.51e-14 ***
## factor(MONTH)8   -4.7702     0.6269  -7.609 2.76e-14 ***
## factor(MONTH)9   -4.5183     0.6001  -7.530 5.09e-14 ***
## factor(MONTH)10  -4.9260     0.6545  -7.527 5.20e-14 ***
## factor(MONTH)11  -4.9097     0.6552  -7.494 6.70e-14 ***
## factor(AGE)1      0.6033     0.5904   1.022  0.30687
## factor(GENDER)1  -1.3001     0.5709  -2.277  0.02278 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## convergence code: 0
## Model failed to converge with max|grad| = 0.380234 (tol = 0.001, component 1)
```

Same As with the previous model, almost all of the months are significant in this model. Unlike the previous model, this model found that the gender was significant, but that the age was not.

e)

```
##
## Call:
## geeglm(formula = prop.sym ~ factor(AGE) + factor(GENDER) + factor(MONTH),
##     family = binomial, data = schiz, id = ID, corstr = "ar1",
##     scale.fix = T)
##
##  Coefficients:
##                   Estimate    Std.err      Wald Pr(>|W|)
## (Intercept)      6.242e-01  6.302e-17 9.810e+31   <2e-16 ***
```

```
## factor(AGE)1    -2.509e-18  6.406e-17 2.000e-03    0.969
## factor(GENDER)1  4.879e-17  6.157e-17 6.280e-01    0.428
## factor(MONTH)1  -2.187e-01  6.108e-18 1.282e+33   <2e-16 ***
## factor(MONTH)2  -5.518e-01  1.028e-17 2.879e+33   <2e-16 ***
## factor(MONTH)3  -7.744e-01  5.271e-17 2.158e+32   <2e-16 ***
## factor(MONTH)4  -1.204e+00  2.796e-17 1.854e+33   <2e-16 ***
## factor(MONTH)5  -1.654e+00  2.453e-17 4.546e+33   <2e-16 ***
## factor(MONTH)6  -2.202e+00  5.273e-17 1.744e+33   <2e-16 ***
## factor(MONTH)7  -2.496e+00  5.275e-17 2.239e+33   <2e-16 ***
## factor(MONTH)8  -2.853e+00  3.370e-17 7.165e+33   <2e-16 ***
## factor(MONTH)9  -2.704e+00  3.816e-17 5.019e+33   <2e-16 ***
## factor(MONTH)10 -2.991e+00  3.484e-17 7.373e+33   <2e-16 ***
## factor(MONTH)11 -2.976e+00  3.513e-17 7.173e+33   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation structure = ar1
## Scale is fixed.
##
##   Link = identity
##
## Estimated Correlation Parameters:
##        Estimate   Std.err
## alpha 1.385e-07 4.517e-08
## Number of clusters:   86  Maximum cluster size: 12
```

The GEE model also concludes that the age and gender are not significant and that all of the month variables are significant.

f)

```
##                     mean     sd 0.025quant 0.5quant 0.975quant    mode       kld
## (Intercept)       1.5153 0.4957     0.5545   1.5103     2.5042  1.5002 8.501e-07
## factor(AGE)1      0.5421 0.5772    -0.5881   0.5390     1.6880  0.5332 1.216e-06
## factor(GENDER)1 -1.2889 0.5576    -2.4081  -1.2820    -0.2086 -1.2685 2.805e-06
## factor(MONTH)1  -0.3999 0.4124    -1.2145  -0.3983     0.4051 -0.3951 1.056e-06
## factor(MONTH)2  -0.9799 0.4150    -1.8051  -0.9764    -0.1750 -0.9694 6.432e-07
## factor(MONTH)3  -1.4153 0.4248    -2.2632  -1.4106    -0.5947 -1.4010 6.566e-07
## factor(MONTH)4  -2.2349 0.4413    -3.1205  -2.2282    -1.3872 -2.2149 1.778e-06
## factor(MONTH)5  -3.0069 0.4673    -3.9487  -2.9983    -2.1134 -2.9813 3.553e-06
## factor(MONTH)6  -3.8814 0.5104    -4.9168  -3.8696    -2.9128 -3.8461 6.145e-06
## factor(MONTH)7  -4.3400 0.5457    -5.4525  -4.3252    -3.3108 -4.2957 7.610e-06
## factor(MONTH)8  -4.8991 0.6047    -6.1452  -4.8783    -3.7720 -4.8369 9.319e-06
## factor(MONTH)9  -4.6739 0.5814    -5.8666  -4.6556    -3.5849 -4.6192 8.573e-06
## factor(MONTH)10 -5.0914 0.6359    -6.4091  -5.0670    -3.9123 -5.0185 9.665e-06
## factor(MONTH)11 -5.0800 0.6371    -6.4000  -5.0557    -3.8984 -5.0072 9.605e-06
```

```
##                     mean      sd 0.025quant 0.5quant 0.975quant   mode
## Precision for ID 0.2052 0.05177     0.1213   0.1993     0.3239 0.1877
```

The INLA model found that the gender was significant: negative across all quantiles. It also agreed with other models that the age was not significant, but that most of the month variables were significant.

g)

Most of the conclusions were similar across all models: most or all of the month variables are highly significant while age was not. The INLA model and the Gauss-Hermite model found that the gender was significant. Although it is only one variable, these two models found it highly significant. Considering the impacts that this conclusion could have on treatments, it is a significant result and highlights the importance of using multiple methods.