

# final\_project

Apoorva Srinivasan

11/30/2019

```
alcdep_df = read.table("ALCDEP.txt") %>%
  mutate(Gender = as.factor(Gender),
         Treatment = as.factor(Treatment),
         Relapse = as.factor(Relapse))

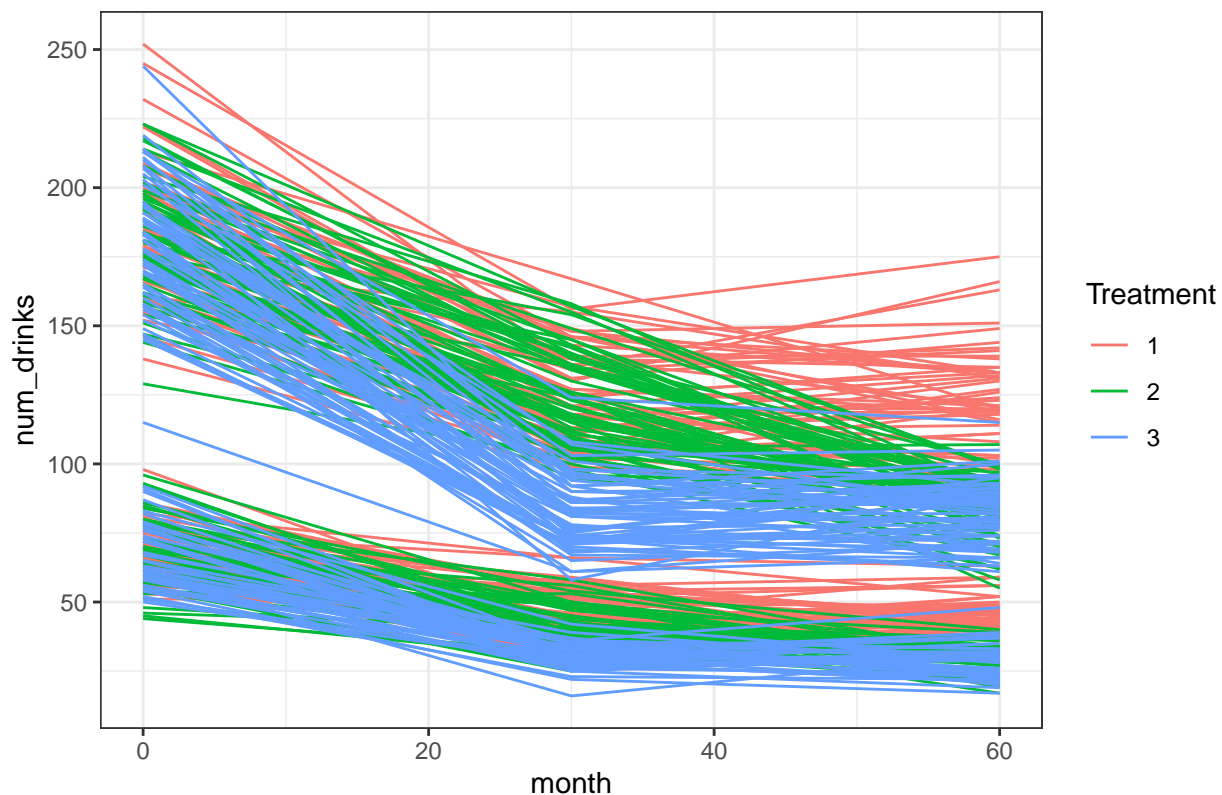
alcdep_long =
  pivot_longer(
    alcdep_df,
    'NDO':'ND60',
    names_to = "month",
    values_to = "num_drinks"
  ) %>%
  mutate(
    month = ifelse(month == "NDO", 0, ifelse(month == "ND30", 30, 60)),
    Gender = as.factor(Gender),
    Relapse = as.factor(Relapse),
    Treatment = as.factor(Treatment)
  )

##no missing values
```

## Drinks between treatment groups over time

```
ggplot(alcdep_long,
  aes(x = month,
    y = num_drinks,
    group = sid,
    color = Treatment)) +
  geom_line() +
  theme_bw() +
  ggtitle("Number of drinks over time between the treatment groups")
```

Number of drinks over time between the treatment groups



Is there evidence to suggest that the treatments differ in their effects on alcohol dependence, as reflected by the number of drinks consumed in a given 30 day period?

```
gee.model = geeglm(num_drinks ~ Treatment*month, family = poisson(link = "log"), corstr = "exchangeable")
summary(gee.model)
```

```
##
## Call:
## geeglm(formula = num_drinks ~ Treatment * month, family = poisson(link = "log"),
## data = alcdep_long, id = sid, corstr = "exchangeable")
##
## Coefficients:
##              Estimate      Std. err      Wald Pr(>|W|)
## (Intercept)  4.7831442  0.0478863 9977.114 <2e-16 ***
## Treatment2    0.1133927  0.0641791   3.122  0.0773 .
## Treatment3   -0.0086921  0.0654931   0.018  0.8944
## month        -0.0073235  0.0002202 1106.097 <2e-16 ***
## Treatment2:month -0.0060401  0.0003146  368.681 <2e-16 ***
## Treatment3:month -0.0078966  0.0003602  480.681 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Estimated Scale Parameters:
##              Estimate Std. err
## (Intercept)  22.04  0.6832
##
## Correlation: Structure = exchangeable Link = identity
```

```

##
## Estimated Correlation Parameters:
##      Estimate Std.err
## alpha    0.8548 0.02739
## Number of clusters: 314 Maximum cluster size: 3
glm.model = glmer(num_drinks ~ Treatment*month + (1 | sid),
family = 'poisson',
data = alcdep_long)

## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl =
## control$checkConv, : Model failed to converge with max|grad| = 0.00604473
## (tol = 0.001, component 1)

## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is nearly unidentifiable:
##  - Rescale variables?

summary(glm.model)

## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: poisson ( log )
## Formula: num_drinks ~ Treatment * month + (1 | sid)
## Data: alcdep_long
##
##      AIC      BIC    logLik deviance df.resid
##    8988    9022    -4487    8974     935
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -4.678 -0.749  0.111  0.833  3.529
##
## Random effects:
## Groups Name      Variance Std.Dev.
## sid      (Intercept) 0.264    0.514
## Number of obs: 942, groups: sid, 314
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    4.657192   0.050658   91.93   <2e-16 ***
## Treatment2      0.119250   0.071436    1.67    0.095 .
## Treatment3      0.026585   0.072538    0.37    0.714
## month          -0.007308   0.000234  -31.23   <2e-16 ***
## Treatment2:month -0.006052   0.000338  -17.92   <2e-16 ***
## Treatment3:month -0.007641   0.000357  -21.39   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr) Trtmn2 Trtmn3 month  Trtm2:
## Treatment2  -0.709
## Treatment3  -0.698  0.495
## month       -0.118  0.084  0.083
## Trtmnt2:mnt  0.082 -0.113 -0.057 -0.693
## Trtmnt3:mnt  0.078 -0.055 -0.114 -0.655  0.454
## convergence code: 0

```

```
## Model failed to converge with max|grad| = 0.00604473 (tol = 0.001, component 1)
## Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?
```

doesn't converge— explain why

Is there a difference in the pattern of change in the number of drinks consumed between the various treatment groups over the duration of the study?

Not much

Alcohol-use disorders are among the most disabling disease categories for the global burden of disease especially for men. Is there evidence to suggest that males tend to have a higher alcohol dependence than females?

### Drinks between genders over time

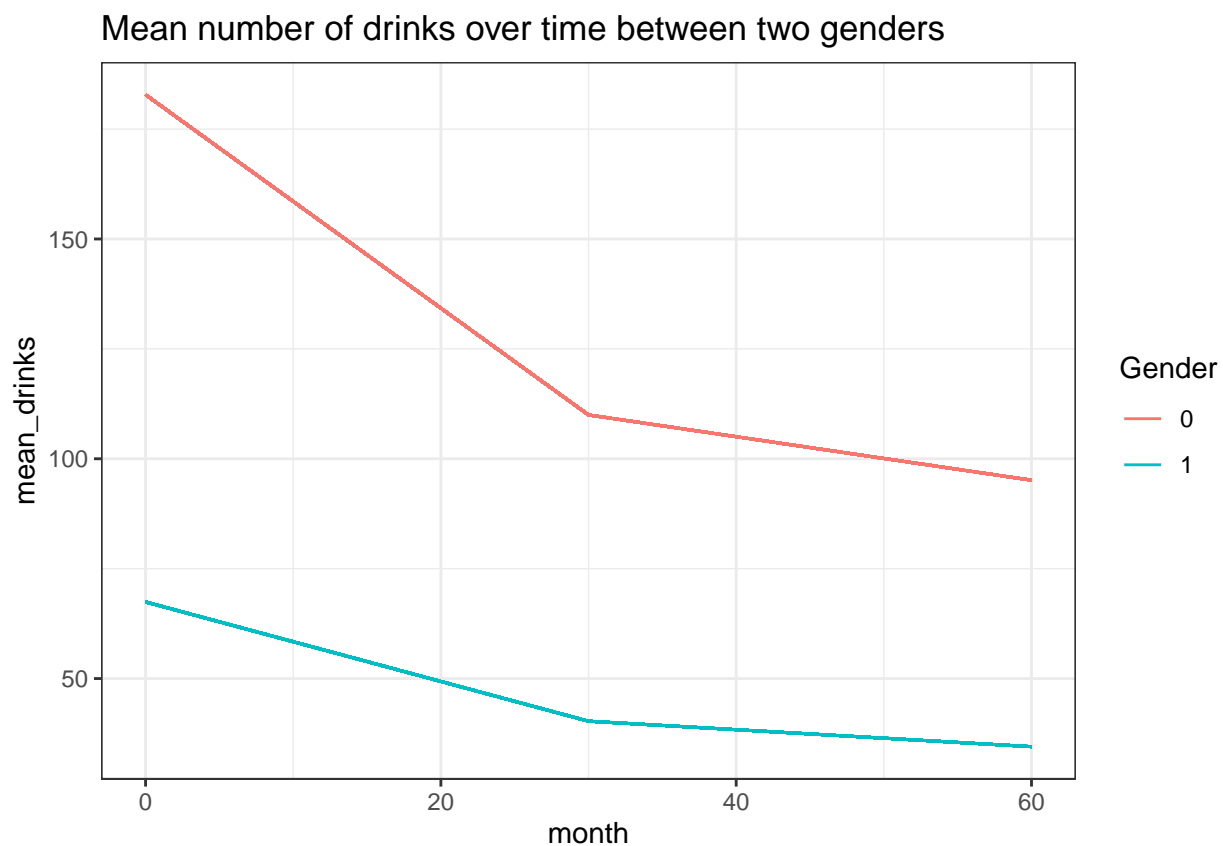
```
ggplot(alcdep_long,
  aes(x = month,
    y = num_drinks,
    group = sid,
    color = Gender)) +
  geom_line() +
  theme_bw() +
  ggtitle("Number of drinks over time between two genders")
```



Yeah, from the above graph, men have higher occurrences than females at all timepoints including baseline

```
alc_mean = alcdep_long %>%
  group_by(Gender, month) %>%
  mutate(mean_drinks = mean(num_drinks))

ggplot(alc_mean,
  aes(x = month,
    y = mean_drinks,
    group = sid,
    color = Gender)) +
  geom_line() +
  theme_bw() +
  ggtitle("Mean number of drinks over time between two genders")
```



Do men and women respond differently to treatment ?

```
gee.gender.model = geeglm(num_drinks ~ Gender*month, family = 'poisson', corstr = "exchangeable", id = sid, data = alcdep_long)
summary(gee.gender.model)
```

```
##
## Call:
## geeglm(formula = num_drinks ~ Gender * month, family = "poisson",
##   data = alcdep_long, id = sid, corstr = "exchangeable")
##
## Coefficients:
##              Estimate      Std.err      Wald Pr(>|W|)
```

```

## (Intercept)    5.167992  0.009743  2.81e+05   <2e-16 ***
## Gender1       -0.996734  0.016842  3.50e+03   <2e-16 ***
## month         -0.011532  0.000293  1.55e+03   <2e-16 ***
## Gender1:month -0.000287  0.000538  2.80e-01     0.59
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Estimated Scale Parameters:
##           Estimate Std.err
## (Intercept)    4.25   0.253
##
## Correlation: Structure = exchangeable Link = identity
##
## Estimated Correlation Parameters:
##           Estimate Std.err
## alpha      0.277   0.0353
## Number of clusters: 314 Maximum cluster size: 3
mem.gender.model = glmer(num_drinks ~ Gender*month + (1 | sid),
family = 'poisson',
data = alcdep_long)

## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl =
## control$checkConv, : Model failed to converge with max|grad| = 0.00440803
## (tol = 0.001, component 1)

## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is nearly unidentifiable:
##  - Rescale variables?

summary(mem.gender.model)

## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: poisson ( log )
## Formula: num_drinks ~ Gender * month + (1 | sid)
## Data: alcdep_long
##
##           AIC      BIC   logLik deviance df.resid
##           8748     8772    -4369     8738      937
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -5.056 -0.893 -0.073  0.895  5.193
##
## Random effects:
## Groups Name          Variance Std.Dev.
## sid      (Intercept) 0.0182   0.135
## Number of obs: 942, groups: sid, 314
##
## Fixed effects:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept)  5.159087   0.011659  442.50   <2e-16 ***
## Gender1     -0.995871   0.018942  -52.57   <2e-16 ***
## month       -0.011528   0.000163  -70.65   <2e-16 ***
## Gender1:month -0.000286   0.000337   -0.85     0.4
## ---

```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##           (Intr) Gendr1 month
## Gender1    -0.615
## month      -0.325  0.200
## Gendr1:mnth  0.157 -0.411 -0.484
## convergence code: 0
## Model failed to converge with max|grad| = 0.00440803 (tol = 0.001, component 1)
## Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?
```

Is there any evidence to suggest that the treatments differ in their effects on subjects with regard to relapsing into alcohol dependence ?

###Mixed model

```
relapse.mm = glm(Relapse ~ Treatment,
                  family = "binomial",
                  data = alcdep_long)
summary(relapse.mm)
```

```
##
## Call:
## glm(formula = Relapse ~ Treatment, family = "binomial", data = alcdep_long)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.854  -0.546  -0.546   0.628   1.988
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.521     0.146   10.41  <2e-16 ***
## Treatment2     -1.690     0.184   -9.18  <2e-16 ***
## Treatment3     -3.348     0.221  -15.12  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 1304.02  on 941  degrees of freedom
## Residual deviance:  985.68  on 939  degrees of freedom
## AIC: 991.7
##
## Number of Fisher Scoring iterations: 4
```

Even in the case that the treatments might differ in their pattern of change or on how subjects relapse into alcohol dependence, is there any evidence to suggest that any of the treatments might be beneficial once the treatment has stopped.

noo,since the difference above is not significant. there isn't enough evidence that the treatment is beneficial.

###Main effects model

#### ##GEE MODEL

```
gee.model = geeglm(num_drinks ~ Treatment*month + Gender*month, family = poisson(link = "log"), corstr = "exchangeable", data = alcdep_long)
summary(gee.model)
```

```
##
## Call:
## geeglm(formula = num_drinks ~ Treatment * month + Gender * month,
##        family = poisson(link = "log"), data = alcdep_long, id = sid,
##        corstr = "exchangeable")
##
## Coefficients:
##              Estimate      Std.err      Wald Pr(>|W|)
## (Intercept)    5.174435    0.015314 1.14e+05  <2e-16 ***
## Treatment2      0.030770    0.018798 2.68e+00   0.102
## Treatment3     -0.050472    0.020111 6.30e+00   0.012 *
## month          -0.007090    0.000238 8.89e+02  <2e-16 ***
## Gender1        -0.996159    0.016632 3.59e+03  <2e-16 ***
## Treatment2:month -0.006104    0.000314 3.79e+02  <2e-16 ***
## Treatment3:month -0.007692    0.000352 4.78e+02  <2e-16 ***
## month:Gender1   -0.000801    0.000347 5.35e+00   0.021 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Estimated Scale Parameters:
##              Estimate Std.err
## (Intercept)    2.85    0.153
##
## Correlation: Structure = exchangeable Link = identity
##
## Estimated Correlation Parameters:
##              Estimate Std.err
## alpha         0.143  0.0359
## Number of clusters: 314 Maximum cluster size: 3
```

#### ##MIXED EFFCT MODEL

```
glm.model = glmer(num_drinks ~ Treatment*month + Gender*month + (1 | sid),
family = 'poisson',
data = alcdep_long)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl =
## control$checkConv, : Model failed to converge with max|grad| = 0.0085953
## (tol = 0.001, component 1)
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is nearly unidentifiable:
##  - Rescale variables?
```

```
summary(glm.model)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: poisson ( log )
## Formula: num_drinks ~ Treatment * month + Gender * month + (1 | sid)
## Data: alcdep_long
##
##              AIC      BIC    logLik deviance df.resid
##           8062     8106     -4022     8044      933
```



```

##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -4.739 -0.725  0.049  0.878  3.793
##
## Random effects:
##   Groups Name      Variance Std.Dev.
##   sid      (Intercept) 0.00978  0.0989
## Number of obs: 942, groups:  sid, 314
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    5.169101   0.014411  358.69  <2e-16 ***
## Treatment2      0.030457   0.018109    1.68   0.093 .
## Treatment3     -0.048500   0.018561   -2.61   0.009 **
## month          -0.007090   0.000252  -28.19  <2e-16 ***
## Gender1        -0.995757   0.015878  -62.71  <2e-16 ***
## Treatment2:month -0.006104   0.000339  -18.03  <2e-16 ***
## Treatment3:month -0.007684   0.000358  -21.47  <2e-16 ***
## month:Gender1   -0.000799   0.000338   -2.36   0.018 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr) Trtmn2 Trtmn3 month  Gendr1 Trtm2: Trtm3:
## Treatment2   -0.671
## Treatment3   -0.650  0.500
## month        -0.439  0.314  0.304
## Gender1      -0.436  0.070  0.058  0.181
## Trtmnt2:mnt  0.296 -0.444 -0.224 -0.667 -0.032
## Trtmnt3:mnt  0.279 -0.217 -0.445 -0.627 -0.025  0.456
## month:Gndr1  0.139 -0.018 -0.011 -0.366 -0.489  0.065  0.051
## convergence code: 0
## Model failed to converge with max|grad| = 0.0085953 (tol = 0.001, component 1)
## Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?

```