

HW6

```
library(plyr)
library(ggplot2)
```

Section 1: Exploring Generalized Linear Mixed Effects Model (GLMM) Analysis

We will use data from a clinical trial comparing auranofin therapy (3 mg of oral gold, twice daily) and placebo for the treatment of rheumatoid arthritis (Bombardier et al., 1986). In this six-month, randomized, double-blind trial, 303 patients with classic or definite rheumatoid arthritis were randomized to one of the two treatment groups and followed over time. The outcome variable is a global impression scale (Arthritis Categorical Scale, ACS) measured at baseline (month 0), month 2, month 4, and month 6. This is a self-assessment of a patient's current arthritis, measured on a 5-level ordinal scale: (1) very good, (2) good, (3) fair, (4) poor, and (5) very poor. Baseline data on this outcome variable are available for 303 of the patients who participated in this trial; follow-up data at 6 months are available for 294 patients.

Data Preprocessing

```
df = read.csv("arthritis-data.csv")

df = reshape(df, timevar = 'Month',
             idvar = c('ID', 'Treatment', 'Age'),
             direction = "long",
             varying = c('Baseline', 'Month2', 'Month4', 'Month6'),
             time = c(0, 2, 4, 6),
             v.names = 'ACS')

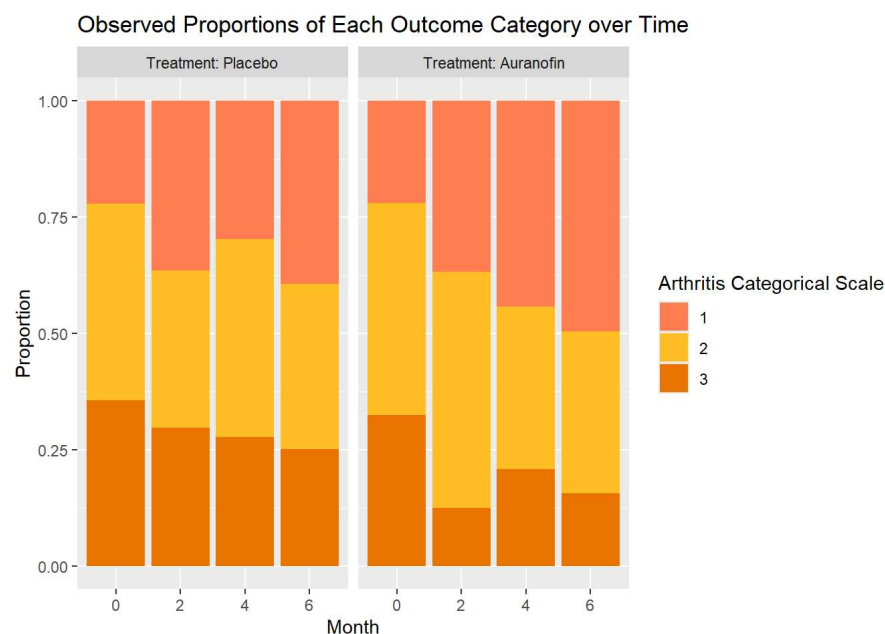
df$sacs = ifelse(df$ACS <= 2, 1, ifelse(df$ACS == 3, 2, 3))
df = df[is.na(df$sacs) == F, ]
df = df[order(df$ID),]
```

Preliminary Analysis: Observed Proportions over Time

```
#Calculate observed proportions of each category over time by treatment group

arth_obs <- count(df, c("Treatment", "Month", "acs"))
arth_obs$Treatment <- factor(arth_obs$Treatment,
                             levels=c(0,1),
                             labels=c("Placebo", "Auranofin"))

#Plot a stacked bar graph of observed proportions of each category over time
p1 <- ggplot(data=arth_obs, aes(x=factor(Month), y=freq, fill=factor(acs)))
p1 + facet_wrap(~Treatment, labeller=label_both) +
  geom_bar(position="fill", stat="identity") +
  labs(title="Observed Proportions of Each Outcome Category over Time",
       x="Month",
       y="Proportion",
       fill="Arthritis Categorical Scale") +
  scale_fill_manual(values = c("coral", "goldenrod1", "darkorange2"))
```



It appears that in the placebo group, ACS proportion for category 3 steadily declines over time. In the treatment group, the converse is true as proportion for category 3 increases following month 2. Also, the proportion of category 1 is higher in the treatment group compared to the control group across different time points (except at baseline). At month 2, the category 2 proportion is substantially higher in treatment compared to control.

Equations for GLMM:

Let's write subject-specific equations for the GLMM, starting with thresholds, integrating covariates, beta coefficients, and random effects, and eventually converting equations to combined intercepts, factoring in each cumulative logit.

GLMM for the log-odds of a more favorable outcome with random slope and random intercept:

$$\begin{aligned}
 \text{logit}\{E(Y_{ij} \leq c|b_i)\} &= \log \left\{ \frac{\Pr(Y_{ij} \leq c|b_i)}{\Pr(Y_{ij} > c|b_i)} \right\} \\
 &= a_c - (\beta_1 + \beta_2 \text{Treatment}_i + \beta_3 \sqrt{\text{Month}_{ij}} \\
 &\quad + \beta_4 \text{Treatment} * \sqrt{\text{Month}_{ij}} + b_{1i} + b_{3i} \sqrt{\text{Month}_{ij}}) \\
 &= (a_c - \beta_1) + (\beta_2 \text{Treatment}_i + \beta_3 \sqrt{\text{Month}_{ij}} \\
 &\quad + \beta_4 \text{Treatment} * \sqrt{\text{Month}_{ij}} + b_{1i} + b_{3i} \sqrt{\text{Month}_{ij}}) \\
 &= a_c^\circ - (\beta_2 \text{Treatment}_i + \beta_3 \sqrt{\text{Month}_{ij}} \\
 &\quad + \beta_4 \text{Treatment} * \sqrt{\text{Month}_{ij}} + b_{1i} + b_{3i} \sqrt{\text{Month}_{ij}})
 \end{aligned}$$

$$b_i \sim \text{MVN}(0, G), \text{ where } G = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} = \begin{pmatrix} \sigma_{b_{1i}}^2 & \text{Cov}(b_{1i}, b_{3i}) \\ \text{Cov}(b_{1i}, b_{3i}) & \sigma_{b_{3i}}^2 \end{pmatrix}$$

Incorporating Estimated Parameters:

We can transpose the equations using the estimated parameters extracted from the GLMM output, providing a nuanced understanding of the model's dynamics.

```
library(ordinal)

df$SMONTH <- sqrt(df$Month)
df$TXSMTH <- df$SMONTH*df$Treatment

M1.1 <- clmm(as.factor(acs) ~ Treatment + SMONTH
              + TXSMTH + (SMONTH|ID),
              data=df)

summary(M1.1)
```

Cumulative Link Mixed Model fitted with the Laplace approximation

formula: as.factor(acs) ~ Treatment + SMONTH + TXSMTH + (SMONTH | ID)
data: df

```
link threshold nobs logLik AIC niter max.grad cond.H
logit flexible 1194 -1115.48 2246.97 522(2625) 2.82e-04 8.3e+02
```

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
ID	(Intercept)	2.55362	1.5980	
	SMONTH	0.09292	0.3048	0.978

Number of groups: ID 303

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
Treatment	-0.1607	0.2919	-0.551	0.581969
SMONTH	-0.3916	0.1012	-3.870	0.000109 ***
TXSMTH	-0.3228	0.1416	-2.280	0.022632 *

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

Threshold coefficients:

	Estimate	Std. Error	z value
1 2	-1.9178	0.2331	-8.227
2 3	1.0128	0.2172	4.662

From R output: $\hat{\sigma}_{b_{1i}}^2 = 2.55362$; $\hat{\sigma}_{b_{3i}}^2 = 0.09292$; $Corr(\hat{b}_{1i}, \hat{b}_{3i}) = 0.978$

$\alpha_c^{\odot} = \alpha_c^{\odot} - \beta_1 : \alpha_c^{\odot}$ is a combined new intercept, α_c^{\odot} is a threshold, β_1 is Intercept

$\alpha_1^{\odot} = \alpha_1 - \beta_1, \alpha_1 = 0, \hat{\alpha}_1^{\odot} = -1.9178$, hence $\hat{\beta}_1$ is 1.9178

$\alpha_2^{\odot} = \alpha_2 - \beta_1, \hat{\alpha}_2^{\odot} = 1.0128$, hence $\hat{\alpha}_2 = 1.0128 + 1.9178 = 2.9306$

Average of Subject-specific logit for a 'typical' subject, $b_{1i} = 0, b_{3i} = 0$

$logit\{E(Y_{ij} \leq c | b_i)\} = a_c^{\odot} - (\beta_2 * Treatment_i + \beta_3 \sqrt{(Month)_{ij}} + \beta_4 * Treatment * \sqrt{(Month)_{ij}})$

$logit\{E(Y_{ij} \leq 1 | b_i = 0)\} = -1.9178 - (-0.1607 * Treatment_i - 0.3916 \sqrt{(Month)_{ij}} - 0.3228 * Treatment * \sqrt{(Month)_{ij}})$

$logit\{E(Y_{ij} \leq 2 | b_i = 0)\} = 1.0128 - (-0.1607 * Treatment_i - 0.3916 \sqrt{(Month)_{ij}} - 0.3228 * Treatment * \sqrt{(Month)_{ij}})$

Random Effects Significance Test

A pivotal step involved testing the relevance of both random intercept and slope components within the model, we will examine the significance of each and decipher their implications on the overall model structure. We will compare the two models using likelihood ratio test.

$H_{\{0\}}$: Variance of the random slope (b_{3i}) is zero $H_{\{a\}}$: Variance of the random slope (b_{3i}) is not zero

```
M1.2 <- clmm(as.factor(acs) ~ Treatment + SMONTH
              + TXSMTH + (1|ID),
              data = df)
```

```
test.statistics <- -2*(M1.2$logLik-M1.1$logLik)
test.statistics
```

[1] 8.804148

```
p.value <- 0.5*(1-pchisq(test.statistics,1)) + 0.5*(1-pchisq(test.statistics,2))
p.value
```

[1] 0.007628685

With a p-value < than 0.05 and a test statistic of 8.804, we reject the null hypothesis. Thus, the model with random intercept and slopes is a better model based on the LRT.

The random effects can be interpreted as: - slope: heterogeneity among patients at baseline for change in log-odds of favorable response over time - intercept: heterogeneity among patients at baseline for log-odds of favorable response

Intervention Group Significance Test

With the optimal model, let's further examine the impact of intervention group on the subject-specific odds of a more favorable response over the study duration. We do this by formulating hypotheses, executing tests, and dissecting the results for insightful interpretation.

$H_0 : \beta_4 = 0$ $H_1 : \beta_4 \neq 0$

```
M1.3 <- clmm(as.factor(acs) ~ Treatment + SMONTH
              + (SMONTH|ID),
              data = df)
```

```
test.statistics <- -2*(M1.3$logLik-M1.1$logLik)
test.statistics
```

```
[1] 5.22242
```

```
p.value <- 0.5*(1-pchisq(test.statistics,1)) + 0.5*(1-pchisq(test.statistics,2))
p.value
```

```
[1] 0.04787157
```

With a p-value < 0.05 and test statistic of 5.2242 for a likelihood ratio test, we reject the null hypothesis. We conclude that there is significant effect of intervention group on the change in subject-specific odds of a more favorable response over the duration of the study.

Interpreting Combined Intercepts

The first cumulative $\text{logit}\{Pr(Y_{ij} \leq 1)\}$ of a typical patient at baseline in the placebo group = $\hat{a}_1^{\odot} = -1.9178$. The second second cumulative $\text{logit}\{Pr(Y_{ij} \leq 2)\}$ of a typical patient at baseline in the placebo group = $\hat{a}_2^{\odot} = 1.0128$

Visualizing the Proportionality Assumption

```
#Output common intercepts for each logit:
alpha1_C<-M1.1$Theta[1]
alpha2_C<-M1.1$Theta[2]

Beta2<-M1.1$beta[1]
Beta3<-M1.1$beta[2]
Beta4<-M1.1$beta[3]

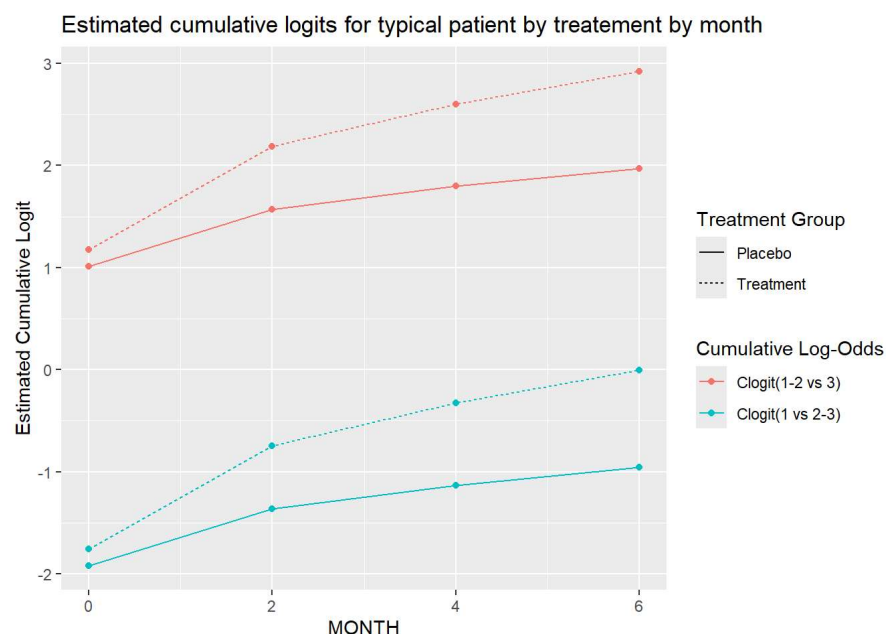
time<-c(0,2,4,6)

est.cum.logit<-data.frame(MONTH=rep(time,times=2),
                          SMONTH=rep(sqrt(time),times=2),
                          treatment=c(rep(0,times=4),rep(1,times=4)),
                          alpha12_C=c(rep(alpha1_C,times=8),rep(alpha2_C,times=8)),
                          Clogit_type=c(rep("Clogit(1 vs 2-3)", times=8),
                                         rep("Clogit(1-2 vs 3)",times=8)))

est.cum.logit$TSMTH<-est.cum.logit$SMONTH*est.cum.logit$treatment
est.cum.logit$Clogit<-est.cum.logit$alpha12_C-(Beta2*est.cum.logit$treatment
                                             +Beta3*est.cum.logit$SMONTH
                                             +Beta4*est.cum.logit$TSMTH)

est.cum.logit$ctx <- factor(est.cum.logit$treatment,
                           levels=c(0,1),
                           labels=c("Placebo", "Treatment"))

p2 <- ggplot(est.cum.logit, aes(x = MONTH,
                              y = Clogit,
                              group = interaction(ctx, Clogit_type)))
p2 + geom_line(aes(color = Clogit_type,linetype =ctx )) +
  geom_point((aes(color = Clogit_type))) +
  labs(title = "Estimated cumulative logits for typical patient by treatment by month",
       y = "Estimated Cumulative Logit",
       color="Cumulative Log-Odds",
       linetype ="Treatment Group")
```



The slope of change in the cumulative log-odds for the favorable response is the same across (C-1) cumulative log-odds (logits). The log-odds are increasing proportionally across different cumulative logits, so that the ratios remain constant. The plot shows the predicted cumulative log-odds for the favorable response across time for two subject-specific ($b_{1i} = b_{3i} = 0$) log-odds.

Probability Estimation Plot

```
#Create a wide dataframe 'eta' from a long dataframe 'est.cum.logit'
eta1<-est.cum.logit[1:8,]
eta2<-est.cum.logit[9:16,]

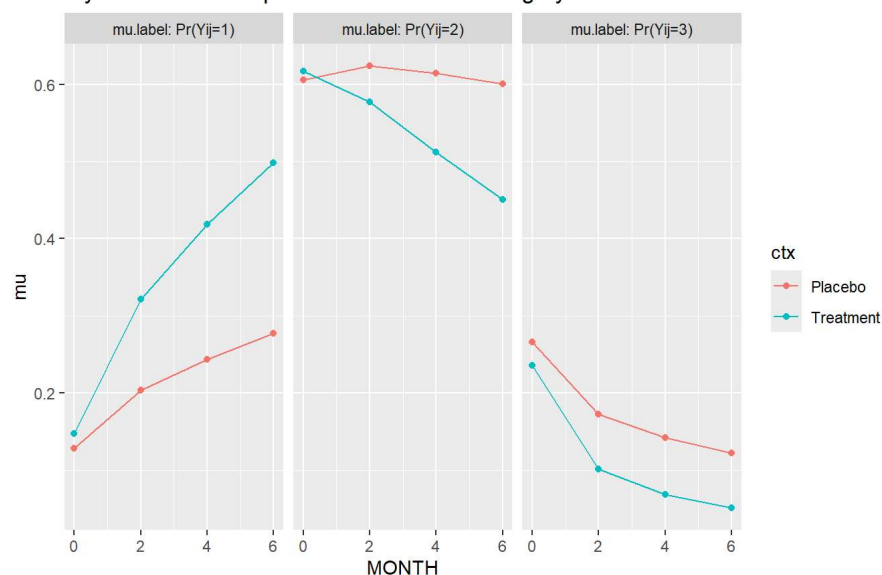
eta<-eta1[,c("MONTH", "SMONTH", "treatment", "alpha12_C", "TSMTH", "Clogit", "ctx")]
names(eta)<-c("MONTH", "SMONTH", "treatment", "alpha12_C", "TSMTH", "eta1", "ctx")
eta$eta2 <- eta2$Clogit

eta$mu_1<-1/(1+exp(-eta$eta1))
eta$mu_2<-1/(1+exp(-eta$eta2))-1/(1+exp(-eta$eta1))
eta$mu_3<-1-1/(1+exp(-eta$eta2))

est.mu.data<-data.frame(MONTH=rep(eta$MONTH, times=3),
                        ctx=rep(eta$ctx, times=3),
                        mu=c(eta$mu_1, eta$mu_2, eta$mu_3),
                        mu.label=c(rep("Pr(Yij=1)", times=8),
                                   rep("Pr(Yij=2)", times=8), rep("Pr(Yij=3)", times=8)))
for (i in 1:nrow(est.mu.data)) {
  est.mu.data$label[i]<-paste(as.character(est.mu.data$ctx[i]),
                             as.character(est.mu.data$mu.label[i]))
}

p5 <- ggplot(est.mu.data, aes(x=MONTH, y=mu, group=ctx))
p5 + facet_wrap(~mu.label, labeller=label_both) +
  geom_line(aes(color=ctx)) + geom_point(aes(color=ctx)) +
  labs(title="Predicted Probabilities for a Typical Subject over Time \n by Treatment Group for Each Outcome Category")
```

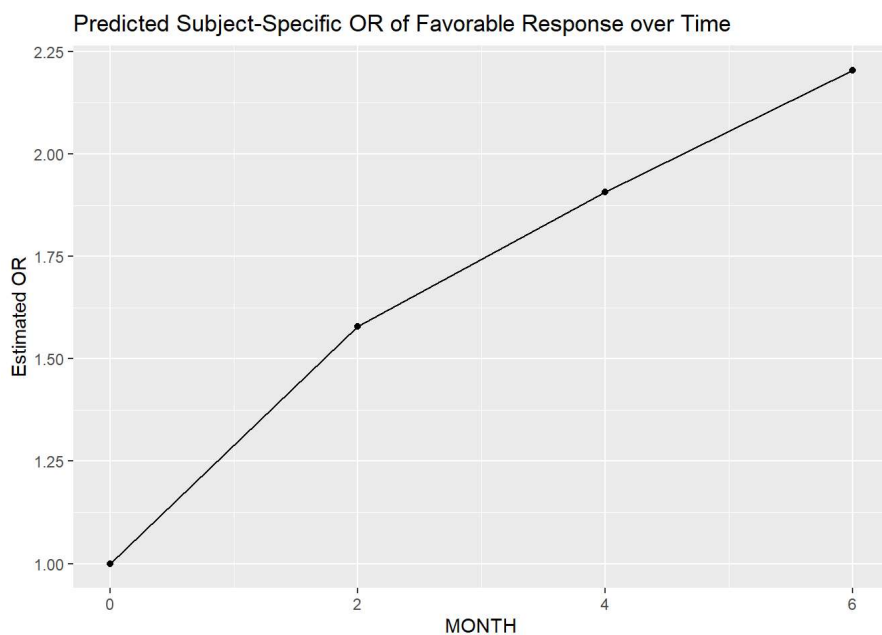
Predicted Probabilities for a Typical Subject over Time
by Treatment Group for Each Outcome Category



For the Very Good category, probability of patients increased over time and the increase was sharper among the treatment group compared to the placebo. For the Good and Fair category, probability of patients decreased over time with a larger decrease in the treatment group compared to the placebo group. Compared to 1b, the proportion of patients in Very Good and Fair is higher across time. The direction of change (increase or decrease) is maintained, however it is much smoother in the predicted probabilities vs 1b.

Analyzing Subject-specific Odds Ratios

```
OR <- exp(sqrt(c(0,2,4,6))*-Beta4)
OR_DATA <- data.frame(MONTH=c(0,2,4,6), OR)
p4 <- ggplot(OR_DATA, aes(x=MONTH, y=OR))
p4 + geom_line() + geom_point()+
  labs(title="Predicted Subject-Specific OR of Favorable Response over Time",
        y="Estimated OR")
```



At 4 months, a patient assigned to treatment is approximately 2 times more likely to have a favorable response when compared to a similar patient treated with placebo.

Comparing with Reference Analysis

Compared to the “Applied Longitudinal Analysis, 2nd Edition” by Fitzmaurice, Laird, and Ware estimates, the estimated β 's for the random effects are about half in magnitude and the variance in double in magnitude. The dissimilarity likely arises from the fact that the book's model had 5 ordinal levels of ACS while ours had 3 ordinal levels. Even with a smaller level of granularity, the drug has significant effect in the change in subject-specific odds of a more favorable response over time. However, the effects are more pronounced and less variable when there are more ACS levels.

Section 2: Exploring Generalized Linear Mixed Effects Model for Nominal Response

Subject-specific Equations for GLMM:

Let's write out the subject-specific equations for the GLMM using covariates, beta coefficients, and random effects, considering “Poor or Very Poor” as the reference category.

Three-part MGLM specification for Nominal Outcome, $Y_{i,j}$

1. A distributional assumption of $Y_{ij}|b_i \sim \text{Multinomial}(p_{ij}^{(1)}, p_{ij}^{(2)}, p_{ij}^{(3)})$, where $p_{ij}^{(c)} = \mu_{ij}^{*(c)} = E(Y_{ij} = c|b_i^{(c)}) = \text{Pr}(Y_{ij} = c|b_i^{(c)})$;
2. $\eta_{ij}^{*(c)} = X_{ij}\beta^{(c)} + Z_{ij}b_i^{(c)}$;
3. $g(\mu_{ij}^{*(c)}) = \text{logit}\{\text{Pr}(Y_{ij} = c|b_i^{(c)})\} = \log\left\{\frac{\text{Pr}(Y_{ij}=c|b_i^{(c)})}{\text{Pr}(Y_{ij}=C|b_i^{(c)})}\right\}$

Baseline category logit with “Poor or Very Poor” (3) as the reference, we have 2 generalized logits:

$$\log\left\{\frac{\text{Pr}(Y_{ij}=c|b_i^{(c)})}{\text{Pr}(Y_{ij}=3|b_i^{(c)})}\right\} = \beta_1^{(c)} + \beta_2^{(c)} * \text{Treatment}_i + \beta_3^{(c)} * \sqrt{(\text{Month})}_{ij} + \beta_4^{(c)} * \text{Treatment} * \sqrt{(\text{Month})}_{ij} + b_{1i}^{(c)}, c = 1, 2$$

$$b_{1i}^{(c)} \sim N(0, \sigma_{b_{1i}^{(c)}}^2)$$

Interpretation of Random Effects

```
library(mclgfit)
library(memisc)

df$acs2[df$acs==1]<-3
df$acs2[df$acs==2]<-2
df$acs2[df$acs==3]<-1
df$acsn <- factor(df$acs)
df$acsn2 <- factor(df$acs2)
df$id <- factor(df$ID)

M2.1 <- mblogit(formula = acsn2 ~ Treatment + SMONTH + TXSMTH,
  random=~1|id, catCov = "free", data=df,
  control = mmclogit.control(epsilon=1e-08, maxit = 25,
    trace=TRUE,trace.inner=FALSE,
    avoid.increase=FALSE,
    break.on.increase=FALSE,
    break.on.infinite=FALSE,
    break.on.negative=FALSE
  )
)
```

```
Iteration 1 - deviance = 2354.677 - criterion = 0.794223
Iteration 2 - deviance = 2344.297 - criterion = 0.03072274
Iteration 3 - deviance = 2331.546 - criterion = 6.921906e-05
Iteration 4 - deviance = 2331.038 - criterion = 1.273286e-09
converged
```

```
summary(M2.1)
```

```
Call:
mblogit(formula = acsn2 ~ Treatment + SMONTH + TXSMTH, data = df,
  random = ~1 | id, catCov = "free", control = mmclogit.control(epsilon = 1e-08,
    maxit = 25, trace = TRUE, trace.inner = FALSE, avoid.increase = FALSE,
    break.on.increase = FALSE, break.on.infinite = FALSE,
    break.on.negative = FALSE))
```

Equation for 2 vs 1:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	0.1547	0.2078	0.744	0.457
Treatment	0.3334	0.2938	1.135	0.256
SMONTH	0.1023	0.1115	0.917	0.359
TXSMTH	0.0941	0.1631	0.577	0.564

Equation for 3 vs 1:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0.65045	0.25823	-2.519	0.01177 *
Treatment	0.08273	0.36795	0.225	0.82211
SMONTH	0.41872	0.12976	3.227	0.00125 **
TXSMTH	0.35545	0.18914	1.879	0.06020 .

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

(Co-)Variances:

Grouping level: id	Estimate	Std.Err.
2~1	1.2206	0.04749
3~1	0.3608 2.0948	0.12543 0.21400

Approximate residual deviance: 2331
Number of Fisher scoring iterations: 4
Number of observations
Groups by id: 303
Individual observations: 1194

mtable(M2.1)

Calls:

```
M2.1: mblogit(formula = acsn2 ~ Treatment + SMONTH + TXSMTH, data = df,
  random = ~1 | id, catCov = "free", control = mmclogit.control(epsilon = 1e-08,
    maxit = 25, trace = TRUE, trace.inner = FALSE, avoid.increase = FALSE,
    break.on.increase = FALSE, break.on.infinite = FALSE,
    break.on.negative = FALSE))
```

	2/1	3/1
(Intercept)	0.155	-0.650*
	(0.208)	(0.258)
Treatment	0.333	0.083
	(0.294)	(0.368)
SMONTH	0.102	0.419**
	(0.111)	(0.130)
TXSMTH	0.094	0.355
	(0.163)	(0.189)
2/1 x VCov(~1,~1)	1.221	0.361
	(0.047)	(0.125)
3/1 x VCov(~1,~1)	0.361	2.095
	(0.125)	(0.214)
Groups by id	303	
Deviance	2331.0	
N	1194	

Significance: *** = p < 0.001;
** = p < 0.01;
* = p < 0.05

- Equation for Fair vs Poor: $\log\left\{\frac{Pr(Y_{ij}=2|b_{1i}^{(c)})}{Pr(Y_{ij}=3|b_{1i}^{(c)})}\right\} = -0.65045 + 0.08273 * \text{Treatment}_i + 0.41872\sqrt{(\text{Month})_{ij}} + 0.35545 * \text{Treatment} * \sqrt{(\text{Month})_{ij}} + b_{1i}^{(c)}$
where $b_{1i}^{(c)} \sim N(0, 2.0948)$
- Equation for Good or Very Good vs Poor:
 $\log\left\{\frac{Pr(Y_{ij}=1|b_{1i}^{(c)})}{Pr(Y_{ij}=3|b_{1i}^{(c)})}\right\} = 0.1547 + 0.3334 * \text{Treatment}_i + 0.1023\sqrt{(\text{Month})_{ij}} + 0.0941 * \text{Treatment} * \sqrt{(\text{Month})_{ij}} + b_{1i}^{(c)}$ where $b_{1i}^{(c)} \sim N(0, 1.2206)$

Intervention Group Effect Test

For each relevant equation, let's test the significant effect of the intervention group on the change in subject-specific odds of a more favorable response.

Using a Wald Test: $H_0 : \beta_4^{(1)} = 0$ $H_a : \beta_4^{(1)} \neq 0$

1. Fair vs Poor: With a z-value of 0.577 and p-value > 0.05 (0.564), we fail to reject the null hypothesis. This means that there is not a significant effect of intervention group on the change in subject-specific odds of a more favorable response over the duration of the study.
2. Good or Very Good vs Poor: With a z-value of 1.879 and p-value > 0.05 (0.06), we fail to reject the null hypothesis. This means that there is not a significant effect of intervention group on the change in subject-specific odds of a more favorable response over the duration of the study.

Interpretation of Intercepts

1. The log-odds of the Fair category relative to the Poor category is estimated to change by $\widehat{\beta_1^{(2)}} = -0.65045$ for an individual with propensity to respond higher relative to the baseline category.
2. The log-odds of the Good or Very Good category relative to the Fair category is estimated to change by $\widehat{\beta_1^{(1)}} = 0.1547$ for an individual with propensity to respond higher relative to the baseline category.

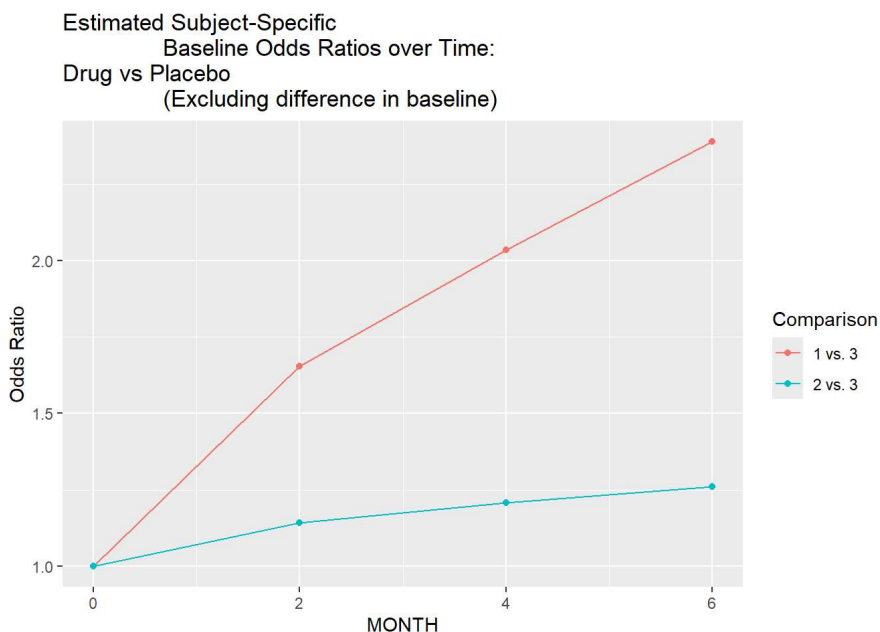
The interpretations are different from Section 1 because Section 1 looks at the cumulative logits and is not specific to which $C - 1$ logit is being looked at while 3iii looks at logit specific to each $C - 1$.

Odds Ratios Plot

```
MONTH<-rep(c(0,2,4,6),2)
SM=sqrt(MONTH)
Numerator=c(rep(1,4),rep(2,4))
Beta4_hat<-c(rep(0.35545,4),rep(0.094,4))
OR_data<-cbind(Numerator, MONTH,SM,Beta4_hat)

rownames(OR_data) <- 1:8
OR_data <- as.data.frame(apply(OR_data[, 1:4], 2, as.numeric))
OR_data$OR <- exp(OR_data$SM*OR_data$Beta4_hat)
#Create a factor variable for Comparison
OR_data$Comparison <- factor(OR_data$Numerator,
                             levels=c(1,2),
                             labels=c("1 vs. 3", "2 vs. 3"))

p2 <- ggplot(OR_data, aes(x=MONTH, y=OR, group=Comparison))
p2 + geom_line(aes(color=Comparison)) + geom_point(aes(color=Comparison)) +
  labs(title="Estimated Subject-Specific
           Baseline Odds Ratios over Time: \nDrug vs Placebo
           (Excluding difference in baseline)", y="Odds Ratio")
```



At 4 months, a patient assigned to the treatment is approximately 1.2 times more likely to feel Fair vs. Poor compared to a similar patient treated with placebo. At 4 months, a patient assigned to the treatment is approximately 2 times more likely to feel Very Good or Good vs. Poor compared to a similar patient treated with placebo.

Probability Estimation Equations

Let's write equations for estimating probabilities of each level of the nominal response for a typical subject. To understand the numerical workings, we can calculate the probability of a "Fair" response at 6 months for a treatment group patient with a specific random intercept.

$$\log\left\{\frac{Pr(Y_{ij}=c|b_i^{(c)})}{Pr(Y_{ij}=3|b_i^{(c)})}\right\} = \beta_1^{(c)} + \beta_2^{(c)} * \text{Treatment}_i + \beta_3^{(c)} * \sqrt{(\text{Month})_{ij}} + \beta_4^{(c)} * \text{Treatment} * \sqrt{(\text{Month})_{ij}} + b_{1i}^{(c)}, c = 1, 2$$

$$\text{For Fair vs Poor: } \log\left\{\frac{Pr(Y_{ij}=c|b_i^{(c)})}{Pr(Y_{ij}=3|b_i^{(c)})}\right\} = \beta_1^{(c)} + \beta_2^{(c)} * \text{Treatment}_i + \beta_3^{(c)} * \sqrt{(\text{Month})_{ij}} + \beta_4^{(c)} * \text{Treatment} * \sqrt{(\text{Month})_{ij}} + b_{1i}^{(c)}$$

$$\begin{aligned} Pr(Y_{ij} = c|b_i) &= \exp(\beta_1^{(c)} + \beta_2^{(c)} * \text{Treatment}_i + \beta_3^{(c)} * \sqrt{(\text{Month})_{ij}} \\ &+ \beta_4^{(c)} * \text{Treatment} * \sqrt{(\text{Month})_{ij}} + b_{1i}^{(c)}) * Pr(Y_{ij} = 3|b_i) \\ &= \exp(0.1547 + 0.3334\text{Treatment}_i + 0.1023 * \sqrt{(\text{Month})_{ij}} + 0.0941\text{Treatment} * \sqrt{(\text{Month})_{ij}} + i \\ &= \exp(\eta_{ij}^{(1)*} + b_{1i}) * Pr(Y_{ij} = 3|b_i) \end{aligned}$$

For Good or Very Good vs Poor:

$$\begin{aligned} \log\left\{\frac{Pr(Y_{ij}=c|b_i^{(c)})}{Pr(Y_{ij}=3|b_i^{(c)})}\right\} &= \beta_1^{(c)} + \beta_2^{(c)} * \text{Treatment}_i + \beta_3^{(c)} * \sqrt{(\text{Month})_{ij}} + \beta_4^{(c)} * \text{Treatment} * \sqrt{(\text{Month})_{ij}} + b_{1i}^{(c)} \\ Pr(Y_{ij} = c|b_i) &= \exp(\beta_1^{(c)} + \beta_2^{(c)} * \text{Treatment}_i + \beta_3^{(c)} * \sqrt{(\text{Month})_{ij}} \\ &+ \beta_4^{(c)} * \text{Treatment} * \sqrt{(\text{Month})_{ij}} + b_{1i}^{(c)}) * Pr(Y_{ij} = 3|b_i) \\ &= \exp(-0.65045 + 0.08273\text{Treatment}_i + 0.41872 * \sqrt{(\text{Month})_{ij}} + 0.35545\text{Treatment} * \sqrt{(\text{Month})_{ij}} \\ &= \exp(\eta_{ij}^{(1)*} + b_{1i}) * Pr(Y_{ij} = 3|b_i) \end{aligned}$$

Given $Pr(Y_{ij} = 3|b_i) = 1 - Pr(Y_{ij} = 1|b_i) - Pr(Y_{ij} = 2|b_i)$, we can derive $Pr(Y_{ij} = 3|b_i) = \frac{1}{1 + \exp(\eta_{ij}^{(1)*} + b_{1i}) + \exp(\eta_{ij}^{(2)*} + b_{1i})}$. Thus,

$$Pr(Y_{ij} = 1|b_i) = \frac{\exp(\eta_{ij}^{(1)*} + b_{1i})}{1 + \exp(\eta_{ij}^{(1)*} + b_{1i}) + \exp(\eta_{ij}^{(2)*} + b_{1i})} \text{ and } Pr(Y_{ij} = 2|b_i) = \frac{\exp(\eta_{ij}^{(2)*} + b_{1i})}{1 + \exp(\eta_{ij}^{(1)*} + b_{1i}) + \exp(\eta_{ij}^{(2)*} + b_{1i})}.$$

Using this, we can get the probability of probability of having "Fair" response at 6 months of follow up for a patient randomized to the treatment group with a random intercept $b_i = -0.17$:

$$\begin{aligned} Pr(Y_{ij} = 2|b_i = -0.17) &= \frac{\exp(0.1547 + 0.3334 + 0.1023(\sqrt{6}) + 0.0941(\sqrt{6}) - 0.17)}{1 + \exp(-0.65045 + 0.08273 + 0.41872(\sqrt{6}) + 0.35545(\sqrt{6}) - 0.17) + \exp(0.1547 + 0.3334 + 0.1023(\sqrt{6}) + 0.0941(\sqrt{6}) - 0.17)} \\ &= \frac{2.223716}{1 + 3.185475 + 2.223716} \\ &= 0.3469574 \end{aligned}$$