Bios 767 Homework 6

Problem 14.1.1

Using GEE and assuming exchangeable log odds ratios, we fit a marginal model of the following form:

$$\text{logit}\{Pr(Y_{ij}=1)\} = \beta_1 + \beta_2 Month_{ij} + \beta_3 Treatment_i Month_{ij},$$
$$\text{where } i = 1, ..., 294 \text{ and } j = 1, ..., n_i.$$

Parameter estimates and empirical standard errors are provided in the table below. The reference group is $Treatment = 0$ and $Month = 0$. α is the parameter corresponding to the log odds ratio for the association among the repeated binary responses.

Table 1: Parameter estimates, standard errors for marginal model.

Parameter	Estimate	Std. Error
β_1	-0.5209	0.1215
β_2	-0.1712	0.0275
β_3	-0.0757	0.0456
α	3.229	0.2901

In this data, we note that 1 month corresponds to 4 weeks, or 28 days.

Problem 14.1.2

$\beta_2 = \text{logit}\{Pr(Y_{ij}=1|Month = m+1, Treatment = 0)\} - logit\{Pr(Y_{ij}=1|Month = m, Treatment = 0)\} = (\beta_1 + \beta_2) - (\beta_1)$

β_2 is the difference in the log odds of moderate or severe degree of onycholysis per 4 week increase for subjects in the Itraconazole treatment group.

Problem 14.1.3

$\beta_3 = \text{logit}\{Pr(Y_{ij}=1|Month = 1, Treatment = 1)\} - \text{logit}\{Pr(Y_{ij}=1|Month = 1, Treatment = 0)\} = (\beta_1 + \beta_2 + \beta_3) - (\beta_1 + \beta_2)$

β_3 is the difference in the log odds of moderate or severe degree of onycholysis at week 4 between subjects receiving Terbinafine treatment and subjects receiving Itraconazole treatment.

Problem 14.1.4

To assess whether treatment has a significant effect on changes in the log odds of moderate or severe onycholysis over time, we perform a Wald test of the hypothesis $H_0 : \beta_3 = 0$. We obtain a test statistic of $2.75 \sim \chi^2_2$ with p-value 0.0972. At a 5% level of significance, we do not have sufficient evidence to reject the null hypothesis. The Terbinafine treatment group does not appear to have significantly different changes in log odds of moderate or severe onycholysis over time from the Itraconazole treatment group. In other words, both treatments appear to have the same effect on changes in the log odds of moderate or severe onycholysis over time.

Problem 14.1.5

Using maximum-likelihood estimation based on 50-point adaptive Gaussian quadrature, we fit a mixed effects logistic regression model with randomly varying intercepts of the following form:

$$\text{logit}\{Pr(Y_{ij} = 1|b_i)\} = (\beta_1 + b_i) + \beta_2 Month_{ij} + \beta_3 Treatment_i Month_{ij},$$
$$\text{where } i = 1, ..., 294, j = 1, ..., n_i,\ b_i \sim N(0, \sigma_b^2)$$
$$\text{and } Y_{ij}|b_i \text{ is assumed to have a Bernoulli distribution.}$$

Parameter estimates and standard errors are provided in the table below. The reference group is $Treatment = 0$ and $Month = 0$.

Table 2: Parameter estimates, standard errors for generalized linear mixed model.

Parameter	Estimate	Std. Error
β_1	-1.697	0.3298
β_2	-0.389	0.0433
β_3	-0.142	0.0649
σ_b^2	16.03	3.0395

Again, we note that 1 month corresponds to 4 weeks, or 28 days.

Problem 14.1.6

From Table 2, we see that the estimated variance of the random intercepts is relatively large: $\hat{\sigma}_b^2 = 16.03$. This implies that there is substantial variability in the underlying propensity for a moderate or severe onycholysis at baseline. At baseline, approximately 95% of individuals have a log odds of moderate or severe onycholysis that varies from -9.546 to 6.151 (or $-1.697 \pm 1.96\sqrt{16.03}$). Translating from the log odds scale to the probability scale, approximately 95% of individuals have a baseline probability of moderate or severe onycholysis that varies from

$$\frac{\exp(-1.697 - 1.96\sqrt{16.03})}{1 + \exp(-1.697 - 1.96\sqrt{16.03})}$$

to

$$\frac{\exp(-1.697 + 1.96\sqrt{16.03})}{1 + \exp(-1.697 + 1.96\sqrt{16.03})},$$

or 0.007% to 99.79%

Problem 14.1.7

$\hat{\beta}_2 = -0.3885$ is an estimate of the change in log odds of moderate or severe onycholysis for each 4 week increase in time within an individual in the Itraconazole treatment group.

Problem 14.1.8

Note that $\beta_2 + \beta_3$ is the change in log odds of moderate or severe onycholysis at 4 weeks within an individual in the Terbinafine treatment group.

$\hat{\beta}_3 = -0.1424$ is an estimate of the difference in the change in log odds of moderate or severe onycholysis at 4 weeks between an individual in the Terbinafine treatment group and an individual in the Itraconazole treatment group, given that they both have the same underlying propensity for a moderate or severe onycholysis at baseline.

Problem 14.1.9

The estimates of β_3 from the marginal model (given in Problem 14.1.1) and the mixed effects model (given in Problem 14.1.5) are $\hat{\beta}_3^{GEE} = -0.0757$ and $\hat{\beta}_3^{MIX} = -0.1424$, respectively. These two estimates differ because marginal models and generalized linear mixed effects models address different scientific questions of interest. Regression coefficients in generalized mixed effects models are interpreted in terms of changes in the log odds of moderate or severe onycholysis either within an individual or between two individuals who have the same values for b_i (their unobserved random effects). On the other hand, regression coefficients in marginal models have interpretations in terms of changes in the log odds of moderate or severe onycholysis in the study population.

Logistic regression coefficients in the marginal model are attenuated relative to the corresponding fixed effects in the logistic regression model with a randomly varying intercept. Due to this, we expect $\hat{\beta}_3^{GEE}$ to be closer to 0 in comparison to $\hat{\beta}_3^{MIX}$ The following approximate relationship between these two parameter estimates holds:

$$\hat{\beta}_3^{GEE} \approx \frac{\hat{\beta}_3^{MIX}}{\sqrt{1 + \frac{\hat{\sigma}_b^2}{\pi^2/3}}};$$

that is, we expect the effects of covariates are attenuated by an attenuation factor of $\frac{1}{\sqrt{1+\frac{\hat{\sigma}_b^2}{\pi^2/3}}} = \frac{1}{\sqrt{1+\frac{16.03}{\pi^2/3}}} = 0.413$ in the marginal model. The ratio of estimates, $\frac{\hat{\beta}_3^{GEE}}{\hat{\beta}_3^{MIX}} = 0.532$, is close to this attenuation factor.

Problem 14.1.10

We refit the model from Problem 14.1.5, sequentially increasing the number of quadrature points used. Parameter estimates and standard errors for each model fitting are provided in Table 3 below.

Table 3: Parameter estimates, standard errors for generalized linear mixed model for varying numbers of quadrature points.

Quadrature Pts	$\hat{\beta}_1$	$\widehat{SE}(\hat{\beta}_1)$	$\hat{\beta}_2$	$\widehat{SE}(\hat{\beta}_2)$	$\hat{\beta}_3$	$\widehat{SE}(\hat{\beta}_3)$	$\hat{\sigma}_b^2$	$\widehat{SE}(\hat{\sigma}_b^2)$
2	-1.492	0.2722	-0.3606	0.03961	-0.1303	0.05892	10.53	1.614
5	-1.521	0.2943	-0.3799	0.04224	-0.1387	0.06287	13.62	2.483
10	-1.719	0.3476	-0.3906	0.04380	-0.1432	0.06535	16.49	3.432
20	-1.697	0.3283	-0.3883	0.04325	-0.1424	0.06490	16.01	2.997
30	-1.698	0.3304	-0.3885	0.04332	-0.1424	0.06494	16.05	3.053
50	-1.697	0.3298	-0.3885	0.04330	-0.1424	0.06493	16.03	3.040

As we increase the number of quadrature points used, parameter estimates and their standard errors begin to stabilize. From 20 quadrature points onwards, parameter estimates do not differ up to the hundredths place. Any difference between parameter estimates after 20 quadrature points is very small.

Problem 14.1.11

Using the estimates from our generalized linear mixed model with 50 quadrature points, we use simulations to compute a decomposition of the total variance into "within" and "between" components, and an intra-class correlation. Additionally, we also estimate the correlation between the outcomes at weeks 0 and 48, and convert the correlations to odds ratios. These values are provided in the table below. Calculations were performed for each treatment group separately. Note that week 0 corresponds to $Month = 0$ and week 48 corresponds to $Month = 12$ in our model.

Table 4: Decomposition of variance & within-subject associations for generalized linear mixed model in both treatment groups at weeks 0 and 48.

		Variance Components				Within-Subject Association	
Treatment	Week	Total	Within	Between	ICC	Correlation	Odds Ratio
Itraconazole	0	0.2270	0.0841	0.1429	0.6296	0.3638	63.06
	48	0.0680	0.0318	0.0363	0.5330		
Terbinafine	0	0.228	0.0845	0.1431	0.6288	0.2481	125.14
	48	0.0326	0.0168	0.0157	0.4835		

From our marginal model in Problem 14.1.1, we obtained an estimated within-subject log odds ratio of $\hat{\alpha} = 3.229 \implies$ we obtained an estimated odds ratio of $\exp(\hat{\alpha}) = \exp(3.229) = 25.25$. This is lower than the within-subject odds ratios in the two treatment groups at weeks 0 and 48 from the generalized linear mixed effects model. This difference could be due to a variety of reasons. First, the within-subject odds ratio from the marginal model is exchangeable: between any two visits within a subject, regardless of treatment group assignment, the odds ratio is assumed to be the same. This structure on the odds ratio is relaxed and clearly can vary between treatment groups in the generalized linear model. Secondly, the within-subject odds ratio for the marginal model is between visits, whereas the within-subject odds ratio for the generalized linear mixed effects model is between Months. The two units are not the same.

Lastly, the problem states "The variable Month denotes the exact timing of measurements in months." Month is defined to be 4 weeks, or 28 days. We know this because every the value of Month at every occasion in the data is a multiple of $\frac{1}{28}$.