Question 2

Joseph Froelicher

4/30/2021

Part A

$$Y_{counts} = \beta_0 + \beta_{baseline} X_1 + \beta_{treatment} X_2 + \beta_{age} X_3 + \beta_{time,\, j} X_{4,\, j}$$

Part B

```
##                        estimates standard errors   P-values
## intercept             0.74488289     0.312888576 0.01728142
## baseline              0.02270315     0.001264422 0.00000000
## treatment group = 1 -0.18079611     0.170875448 0.29002907
## age                   0.01914748     0.010158101 0.05943700
## timepoint = 2        -0.06858711     0.109119070 0.52964179
## timepoint = 3        -0.05849621     0.155852383 0.70741443
## timepoint = 4        -0.19835856     0.098702169 0.04446638
```

The estimate correlation coefficient for the model that controls for baseline seizures, with treatment, age and time as predictors is 0.411.

Part C

working_cor	qic	term	estimate	std.error	p.value
unst	-5897.169	(Intercept)	0.7565568	0.2676969	0.0047108
exch	-5899.351	(Intercept)	0.7448829	0.3128886	0.0172814
unst	-5897.169	age	0.0187315	0.0090981	0.0395096
exch	-5899.351	age	0.0191475	0.0101581	0.0594370
unst	-5897.169	base	0.0227330	0.0011797	$<$ 2.22e-16
exch	-5899.351	base	0.0227031	0.0012644	$<$ 2.22e-16
unst	-5897.169	timey2	-0.0685871	0.1091191	0.5296418
exch	-5899.351	timey2	-0.0685871	0.1091191	0.5296418
unst	-5897.169	timey3	-0.0584962	0.1558524	0.7074144
exch	-5899.351	timey3	-0.0584962	0.1558524	0.7074144
unst	-5897.169	timey4	-0.1983586	0.0987022	0.0444664
exch	-5899.351	timey4	-0.1983586	0.0987022	0.0444664
unst	-5897.169	trt1	-0.1842484	0.1303238	0.1574284
exch	-5899.351	trt1	-0.1807961	0.1708754	0.2900291

While the correlations are roughly the same, they are not identical, and it may be of interest to examine some other correlation structure. Here we are examining unstructured correlation, allowing for the correlations to vary. The estimates are very similar between unstructured and exchangeable correlation, so it really won't matter which model we choose. However, seeing as the correlation to seem to be exchageable, and the QIC value of the unstructured model is not significantly better (< 6) than the model with exchangeable correlation structure, and therefore we should stick with our model that best fits the correlation structure.

Part D

- $e^{\beta_0} = 2.1061948$: The expected rate ratio of seizures at the first time point for patients with no treatment, baseline number of seizures equal to zero, and age equal to zero (note we should not be interpreting this coeefficent because age zero and baseline zero are extrapolating outside of the scope of our data).
- $e^{\beta_{baseline}} = 1.0229628$: The expected increase in rate ratio of seizures at baseline when compared to the first timepoint, controlling for timepoint, age, and treatment, for a unit increase in the number of seizures at baseline.
- $e^{\beta_{treatment=1}} = 0.8346055$: The expected increase in rate ratio of seizures for treatment group 1 over treatment group 0 when controling for baseline, age, and time point.
- $e^{\beta_{age}} = 1.019332$: The expected increase in rate ratio of seizures for one year increase in age when controlling for baseline, treatment group, and timepoint.
- $e^{\beta_{time=2}} = 0.9337121$: The expected increase in rate ratio of seizures when compared to the first timepoint, while controlling for baseline, age and treatment group.
- $e^{\beta_{time=3}} = 0.9431818$: The expected increase in rate ratio of seizures when compared to the first timepoint, while controlling for baseline, age and treatment group.
- $e^{\beta_{time=4}} = 0.8200758$: The expected increase in rate ratio of seizures when compared to the first timepoint, while controlling for baseline, age and treatment group.

Apendix

```
knitr::opts_chunk$set(include = TRUE, echo = TRUE, warning = FALSE)

library(geepack)
library(tidyr)
library(tidyverse)

data(seizure)
seizure$id = as.factor(1:dim(seizure)[1])
data = pivot_longer(data = seizure, cols = c("y1", "y2", "y3", "y4"), names_to = "time", values_to = "co
data$trt = factor(data$trt)
data = as.data.frame(data)

# part b
model0 = geeglm(
  counts ~ base + trt + age + time,
  id = id,
  data = data,
  family = poisson,
  corstr = "exchangeable"
)

results = summary(model0)$coefficients
results = results[,-3]
colnames(results) = c("estimates", "standard errors", "P-values")
rownames(results) = c("intercept","baseline", "treatment group = 1", "age", "timepoint = 2", "timepoint

results

summary = summary(model0)
corr = summary$corr

# part c
model1 = geeglm(
  counts ~ base + trt + age + time,
  id = id,
  data = data,
  family = poisson,
  corstr = "unstructured"
)

corr_comparison = bind_rows(
  mutate(tidy(model1), working_cor =  "unst", qic = QIC(model1)[1]),
  mutate(tidy(model0), working_cor =  "exch", qic = QIC(model0)[1])
) %>%
  select(working_cor, qic, term, estimate, std.error, p.value) %>%
  arrange(term) %>%
  mutate(p.value = format.pval(p.value)) %>%
  knitr::kable()

corr_comparison
```

```r
# part d
results2 = summary(model0)$coefficients
results2 = results2[,-3]
colnames(results2) = c("estimates", "standard errors", "P-values")
rownames(results2) = c("intercept","baseline", "treatment group = 1", "age", "timepoint = 2", "timepoin

b0 = exp(results[1, 1])
b1 = exp(results[2, 1])
b2 = exp(results[3, 1])
b3 = exp(results[4, 1])
b4 = exp(results[5, 1])
b5 = exp(results[6, 1])
b6 = exp(results[7, 1])
```