

# BIOS 6312: Modern Regression Analysis

**Andrew J. Spieker, Ph.D.**

Assistant Professor of Biostatistics  
Vanderbilt University Medical Center

Set 7 supplementary slides for R enthusiasts

Version: 03/14/2021

# EXAMPLES FOR SET 7

## **Examples for R enthusiasts:**

- ▶ GEE in REACH (Slide 710)
- ▶ Mixed model in DFMO (Slide 729)

# EXAMPLES FOR SET 7

## Examples for R enthusiasts:

- ▶ **GEE in REACH (Slide 710)**
- ▶ Mixed model in DFMO (Slide 729)

## Reading in the REACH data:

- ▶ Read in data:

```
reach.data <- read.csv("reach.csv",  
                      header = TRUE,  
                      stringsAsFactors = FALSE)
```

## Convert REACH data to long form:

- ▶ Don't forget to reorder the data and rename...

```
reach.long <- reshape(reach.data,  
                      idvar = "id",  
                      direction = "long",  
                      varying = list(c(12:14), c(15:17)),  
                      timevar = "t")
```

```
reach.long <- reach.long[order(reach.long$id),]
```

```
K <- dim(reach.long)[2]  
names(reach.long) <- c(names(reach.long)[1:(K - 2)],  
                      "alc", "sdsca")
```

## **GEE (working independence):**

- ▶ geese function in R (honk, honk!)

```
library("geepack")
```

```
zz.indep <- geese(a1c ~ factor(t)*factor(reach),  
                 corstr = "independence",  
                 data = reach.long)
```

# GEE IN REACH

## GEE (working independence):

- ▶ Output should match Stata results.

```
> summary(zz.indep)
```

```
Call:
```

```
geese(formula = a1c ~ factor(t) * factor(reach), data = reach.long,  
      corstr = "independence")
```

```
Mean Model:
```

```
Mean Link:          identity  
Variance to Mean Relation: gaussian
```

```
Coefficients:
```

	estimate	san.se	wald	p
(Intercept)	8.539357	0.11842	5.1997e+03	0.0000e+00
factor(t)2	0.178742	0.13653	1.7140e+00	1.9047e-01
factor(t)3	0.044121	0.14029	9.8908e-02	7.5314e-01
factor(reach)1	0.169586	0.16987	9.9668e-01	3.1811e-01
factor(t)2:factor(reach)1	-0.802617	0.18568	1.8684e+01	1.5424e-05
factor(t)3:factor(reach)1	-0.192031	0.20088	9.1384e-01	3.3910e-01

## GEE (working exchangeable):

- ▶ Update the correlation structure!

```
zz.exch <- geese(a1c ~ factor(t)*factor(reach),  
                corstr = "exchangeable",  
                data = reach.long)
```



# GEE IN REACH

## GEE (working exchangeable):

- ▶ Output should match Stata results.

```
> summary(zz.exch)
```

```
Call:
```

```
geese(formula = a1c ~ factor(t) * factor(reach), data = reach.long,  
      corstr = "exchangeable")
```

```
Mean Model:
```

```
Mean Link:          identity  
Variance to Mean Relation: gaussian
```

```
Coefficients:
```

	estimate	san.se	wald	p
(Intercept)	8.534138	0.11806	5224.99181	0.0000e+00
factor(t)2	0.178341	0.13523	1.73913	1.8725e-01
factor(t)3	0.044573	0.13986	0.10157	7.4996e-01
factor(reach)1	0.178013	0.16942	1.10405	2.9338e-01
factor(t)2:factor(reach)1	-0.819732	0.18268	20.13617	7.2120e-06
factor(t)3:factor(reach)1	-0.194139	0.19812	0.96025	3.2713e-01

# EXAMPLES FOR SET 7

## Examples for R enthusiasts:

- ▶ *GEE in REACH (Slide 710)*
- ▶ **Mixed model in DFMO (Slide 729)**

# MIXED MODEL IN DFMO

## Reading in the DFMO data:

- ▶ Read in data (subsetting to twelve months):

```
dfmo.data <- read.csv("dfmo.csv",  
                      header = TRUE,  
                      stringsAsFactors = FALSE)
```

```
dfmo.data <- subset(dfmo.data, dfmo.data$time != 15)
```

# MIXED MODEL IN DFMO

## Random intercepts model:

- ▶ Including cluster-robust standard errors:

```
library("lme4")
library("clubSandwich")

zz.1 <- lmer(spd ~ dose*time + (1|ptid),
             REML = FALSE,
             data = dfmo.data)

summary(zz.1)
sqrt(diag(vcovCR(zz.1, type = "CR0")))
```

# MIXED MODEL IN DFMO

## Random intercepts model:

► Output:

Fixed effects:

	Estimate	Std. Error	t	value
(Intercept)	3.26550	0.17790	18.36	
dose	0.52476	0.79889	0.66	
time	-0.00714	0.02074	-0.34	
dose:time	-0.31906	0.09611	-3.32	

```
> sqrt(diag(vcovCR(zz.1, type = "CR0")))
(Intercept)      dose      time  dose:time
  0.184968    0.911460    0.017868    0.102780
```

- The standard errors do not match Stata's exactly. It's probably off by some degrees-of-freedom correction that I was unable to figure out this time.

# MIXED MODEL IN DFMO

## Mixed-effects model:

- ▶ Including cluster-robust standard errors:

```
library("lme4")
library("clubSandwich")

zz.2 <- lmer(spd ~ dose*time + (time|ptid),
             REML = FALSE,
             data = dfmo.data)

summary(zz.2)
sqrt(diag(vcovCR(zz.2, type = "CR0")))
```

# MIXED MODEL IN DFMO

## Mixed-effects model:

### ► Output:

Fixed effects:

	Estimate	Std. Error	t	value
(Intercept)	3.26737	0.19527	16.73	
dose	0.51463	0.87632	0.59	
time	-0.00809	0.02103	-0.38	
dose:time	-0.31524	0.09689	-3.25	

```
> sqrt(diag(vcovCR(zz.2, type = "CR0")))
(Intercept)      dose      time  dose:time
  0.185329    0.911482    0.017949    0.102898
```

- The standard errors do not match Stata's exactly. It's probably off by some degrees-of-freedom correction that I was unable to figure out this time.