

Exercise 6B: Comparing Multivariable Linear Models

**Problem 1.**

Model One: `lm(formula = T2DM ~ log_ubpa + ucr + age + factor(gender), data = bpa, na.action = na.omit)`

Model Two: `lm(formula = T2DM ~ log_ubpa + ucr + age + factor(gender) + factor(raceeth) + factor(educat) + factor(inccat) + factor(SMK) + bmi + waist, data = bpa, na.action = na.omit)`

Odds ratios and 95% confidence intervals for a one standard deviation unit (1-SD) increase in BPA in each model:

Model One:

OR 1.018 (1.008, 1.029) with a p-value of 8.168405e-04

**Interpretation:** The odds of having type 2 diabetes increase by 1.8% for each 1-SD unit increase in BPA. Because the p-value for this OR is lower than our alpha threshold of statistical significance (0.05), we can conclude that this OR is statistically significant.

Model Two:

OR 1.015 (1.004, 1.026) with a p-value of 5.609570e-03

**Interpretation:** The odds of having type 2 diabetes increase by 1.5% for each 1-SD unit increase in BPA. Because the p-value for this OR is lower than our alpha threshold of statistical significance (0.05), we can conclude that this OR is statistically significant. This OR is very slightly lower than the OR from model 1 because we are taking more factors into account and therefore we decrease the amount of odds attributed specifically to urinary bpa in this model.

## CODE.

```
>#EPID 674 | HW 6B | Stephanie Mecham
>
>setwd("\\Users\\smecham\\Desktop\\Hw2") #setting working directory
>
>#Installing and Calling Packages
>
>library(sas7bdat
> #Installing Epicalc
> install.packages("C:/Users/smecham/Desktop/epicalc_2.15.1.0.tar.gz", repos = NULL, type = "source")
>install.packages("C:/Users/smecham/Desktop/epicalc_2.15.1.0.tar.gz", repos = NULL, type = "source")
* installing *source* package 'epicalc' ...
** package 'epicalc' successfully unpacked and MD5 sums checked
** R
** data
** demo
** preparing package for lazy loading
** help
*** installing help indices
** building package indices
** testing if installed package can be loaded
* DONE (epicalc)
>library(epicalc)
>
>#Importing file & attaching
>
>options(max.print=999999)
>read.sas7bdat("bpa.sas7bdat")
> bpa <- read.sas7bdat("bpa.sas7bdat")
> save(bpa, file="bpa.rda") #saving for future use
> attach(bpa)
>
>#Creating binary T2DM variable:
> diabetic <- ifelse(a1c >= 6.5, 1, 0)
> diabetic_meds <- ifelse (dmmed == 1,1,0)
> T2DM <- ifelse (!is.na(diabetic) & (diabetic==1) | (!is.na(diabetic_meds) & (diabetic_meds ==1)) ,1, 0)
>
>#Creating log UBPA variable:
>log_ubpa <- log(ubpa)
>
>#Creating Model 1
>lm(formula = T2DM ~ log_ubpa + ucr + age + factor(gender), data = bpa, na.action = na.omit)
```

Call:

```
lm(formula = T2DM ~ log_ubpa + ucr + age + factor(gender), data = bpa,
```

```
na.action = na.omit)
```

Coefficients:

```
(Intercept)    log_ubpa      ucr
-0.095296     0.016824   -0.000130
    age factor(gender)2
  0.004585   -0.006572
```

```
>
```

```
>
```

```
>#Creating Model 2
```

```
> lm(formula = T2DM ~ log_ubpa + ucr + age + factor(gender) + factor(raceeth) + factor(educat) + factor(inccat) + factor(SMK) + bmi + waist, data = bpa, na.action = na.omit)
```

Call:

```
lm(formula = T2DM ~ log_ubpa + ucr + age + factor(gender) + factor(raceeth) + factor(educat) + factor(inccat) + factor(SMK) + bmi + waist, data = bpa, na.action = na.omit)
```

Coefficients:

```
(Intercept)    log_ubpa      ucr
-0.4497411     0.0137553   -0.0003299
    age factor(gender)2 factor(raceeth)2
  0.0040073   -0.0007105    0.0093321
factor(raceeth)3 factor(raceeth)4 factor(raceeth)5
-0.0438284     0.0542704   -0.0001397
factor(educat)2 factor(educat)3 factor(educat)NaN
-0.0283878     -0.0320485   -0.1420890
factor(inccat)2 factor(inccat)3 factor(inccat)4
-0.0008769     0.0009475    0.0051074
factor(inccat)5 factor(SMK)1    factor(SMK)2
  0.0055327    0.0047454    0.0120021
factor(SMK)NaN      bmi      waist
  0.0655480    0.0025855    0.0036373
```

```
>#Standardizing log_ubpa so one unit increase = 1-SD unit increae
```

```
> zlogubpa <- (log_ubpa - mean(log_ubpa)) / sd(log_ubpa)
```

```
>
```

```
>#Setting models with standardized log_ubpa's to variable names
```

```
> zmodel1 <- lm(formula = T2DM ~ zlogubpa + ucr + age + factor(gender), data = bpa, na.action = na.omit)
```

```
>
```

```
> zmodel2 <- lm(formula = T2DM ~ zlogubpa + ucr + age + factor(gender) + factor(raceeth) + factor(educat) + factor(inccat) + factor(SMK) + bmi + waist, data = bpa, na.action = na.omit)
```

```
>#Obtaining ORs and 95% CIs for 1-SD unit increases in BPA on T2DM
```

```
> logistic.display(zmodel1)
```

	OR	lower95ci	upper95ci	Pr(> Z )
zlogubpa	1.0184428	1.0076090	1.029393	8.168405e-04
ucr	0.9998700	0.9997328	1.000007	6.353298e-02
age	1.0045952	1.0040907	1.005100	2.516464e-69
factor(gender)2	0.9934496	0.9750765	1.012169	4.902156e-01

```
> logistic.display(zmodel2)
```

	OR	lower95ci	upper95ci	Pr(> Z )
zlogubpa	1.0150533	1.0043830	1.0258370	5.609570e-03
ucr	0.9996702	0.9995318	0.9998085	3.066555e-06
age	1.0040154	1.0034490	1.0045821	3.641495e-43
factor(gender)2	0.9992898	0.9789006	1.0201036	9.461466e-01
factor(raceeth)2	1.0093757	0.9685457	1.0519270	6.578168e-01
factor(raceeth)3	0.9571182	0.9324851	0.9824019	9.931447e-04
factor(raceeth)4	1.0557701	1.0252334	1.0872163	2.931177e-04
factor(raceeth)5	0.9998603	0.9522773	1.0498210	9.955199e-01
factor(educat)2	0.9720113	0.9475208	0.9971348	2.928319e-02
factor(educat)3	0.9684596	0.9455001	0.9919767	8.873013e-03
factor(educat)NaN	0.8675440	0.6126229	1.2285415	4.234927e-01
factor(inccat)2	0.9991235	0.9720708	1.0269292	9.500807e-01
factor(inccat)3	1.0009480	0.9743044	1.0283201	9.451245e-01
factor(inccat)4	1.0051205	0.9775888	1.0334275	7.185435e-01
factor(inccat)5	1.0055480	0.9659147	1.0468075	7.874316e-01
factor(SMK)1	1.0047566	0.9822781	1.0277495	6.810470e-01
factor(SMK)2	1.0120744	0.9886841	1.0360181	3.144505e-01
factor(SMK)NaN	1.0677440	0.8159352	1.3972644	6.329250e-01
bmi	1.0025889	0.9990837	1.0061063	1.479821e-01
waist	1.0036439	1.0021694	1.0051206	1.283336e-06

```
> #End of code
```