

Modeling Ordinal Data: Model Building Hands On

DJNickens

29 Sept 2018

Preparation for Analysis

```
rm(list=ls())
```

Libraries

```
library(MASS)      #' polr() function , stepAIC() function
library(ordinal)   #' clm() function
library(dplyr)     #' data transformation functions
library(boot)      #' logit(), inv.logit() functions
library(ggplot2)   #' plotting functions
library(GGally)    #' scatterplot matrices
```

Functions

```
stuniq <- function(x) { sort(unique(x)) }
indicator <- function(condition) { ifelse(condition,1,0) }
mean_2 <- function(x) { round(mean(x,na.rm=T),2) }
median_2 <- function(x) { round(median(x,na.rm=T),2) }
min_2 <- function(x) { round(min(x,na.rm=T),2) }
max_2 <- function(x) { round(max(x,na.rm=T),2) }
sum_2 <- function(x) { sum(x,na.rm=T) }
nobs_2 <- function(x) { sum(!is.na(x)) }

AICr <- function(minModel,model) {
  # function to compute relative likelihood
  AICr <- exp((as.numeric(as.character(minModel$info$AIC))-
    as.numeric(as.character(model$info$AIC)))/2)
  return(AICr)
}
```

Data

```
wd <- "C:/..."; setwd(wd)
```

```
getwd()
```

```
## [1] "C:/Program Files (x86)/ePharm/ePharmRunWorkspace/AnalysisStep726570"
```

```

dsn <- "AEs-ordinal-data-ver3.csv"
d <- read.csv(dsn, as.is=T)
d$AEGrade <- as.factor(d$DV)
d$Sex <- factor(d$SEX,levels=c(1,2), labels=c("Male","Female"))
d$Smoker <- factor(d$SMOK,levels=c(1,2,3), labels=c("Current","Former","Never"))
d$Dose <- as.factor(d$DOSE)

```

```
head(d)
```

```

##   ID DOSE      BIO SEX AGE   WT SMOK   CMAXSS   CAVGSS   BIO2 DV
## 1  1  50 20.09565  1  64 76.66   3 13.215195 13.378387 55.16573  2
## 2  2  30 25.37029  2  69 73.50   1 15.867221 12.778449 52.19035  1
## 3  3  10 15.98545  2  54 76.00   1  6.620023  6.490101 58.65953  1
## 4  4 100 13.36886  2  50 65.90   1 34.478728 16.450041 37.92031  1
## 5  5  50 10.21455  1  63 67.00   3 18.182994 11.171032 38.19847  2
## 6  6 100 28.84815  2  73 74.00   1 26.420695 18.208066 53.85277  0
##   AEGrade   Sex   Smoker Dose
## 1         2   Male   Never   50
## 2         1 Female Current   30
## 3         1 Female Current   10
## 4         1 Female Current  100
## 5         2   Male   Never   50
## 6         0 Female Current  100

```

```
summary(d)
```

```

##           ID           DOSE           BIO           SEX
##  Min.   : 1.00   Min.   : 0.0   Min.   : 0.9467   Min.   :1.000
## 1st Qu.: 50.75   1st Qu.: 10.0   1st Qu.:12.0209   1st Qu.:2.000
## Median :100.50   Median : 30.0   Median :16.1634   Median :2.000
## Mean   :100.50   Mean   : 40.9   Mean   :16.4975   Mean   :1.785
## 3rd Qu.:150.25   3rd Qu.: 50.0   3rd Qu.:20.8966   3rd Qu.:2.000
## Max.   :200.00   Max.   :100.0   Max.   :33.1072   Max.   :2.000
##           AGE           WT           SMOK           CMAXSS
##  Min.   :19.00   Min.   : 41.50   Min.   :1.000   Min.   : 0.000
## 1st Qu.:48.00   1st Qu.: 62.00   1st Qu.:1.000   1st Qu.: 7.124
## Median :56.00   Median : 71.35   Median :1.000   Median :14.991
## Mean   :53.65   Mean   : 73.26   Mean   :1.375   Mean   :15.096
## 3rd Qu.:61.00   3rd Qu.: 82.12   3rd Qu.:2.000   3rd Qu.:20.696
## Max.   :73.00   Max.   :128.40   Max.   :3.000   Max.   :63.131
##           CAVGSS           BIO2           DV           AEGrade           Sex
##  Min.   : 0.000   Min.   : 9.199   Min.   :0.00   0:29   Male : 43
## 1st Qu.: 4.748   1st Qu.:30.721   1st Qu.:1.00   1:57   Female:157
## Median :10.334   Median :39.101   Median :2.00   2:78
## Mean   : 9.347   Mean   :40.725   Mean   :1.71   3:15
## 3rd Qu.:12.790   3rd Qu.:50.189   3rd Qu.:2.00   4:21
## Max.   :21.403   Max.   :76.248   Max.   :4.00
##           Smoker           Dose
## Current:148   0 :39
## Former : 29   10 :31
## Never  : 23   30 :39
##                               50 :48
##                               100:43

```

```
##
```

Question 1:

Create at least 1 graph (a) to explore relationships between covariates and
(b) between covariates and the response.

Observed Covariate-to-Covariate Relationships

Continuous Variables

```
gg <- ggpairs(d[,c(2,3,5,6,8,9,10)],lower=list(continuous='smooth_loess'),  
             diag=list(continuous='density'),  
             axisLabels = 'show')
```

```
## Warning in check_and_set_ggpairs_defaults("diag", diag, continuous =  
## "densityDiag", : Changing diag$continuous from 'density' to 'densityDiag'
```

```
ggsave("Cov_Cov_02.png",gg,height=9,width=12)
```

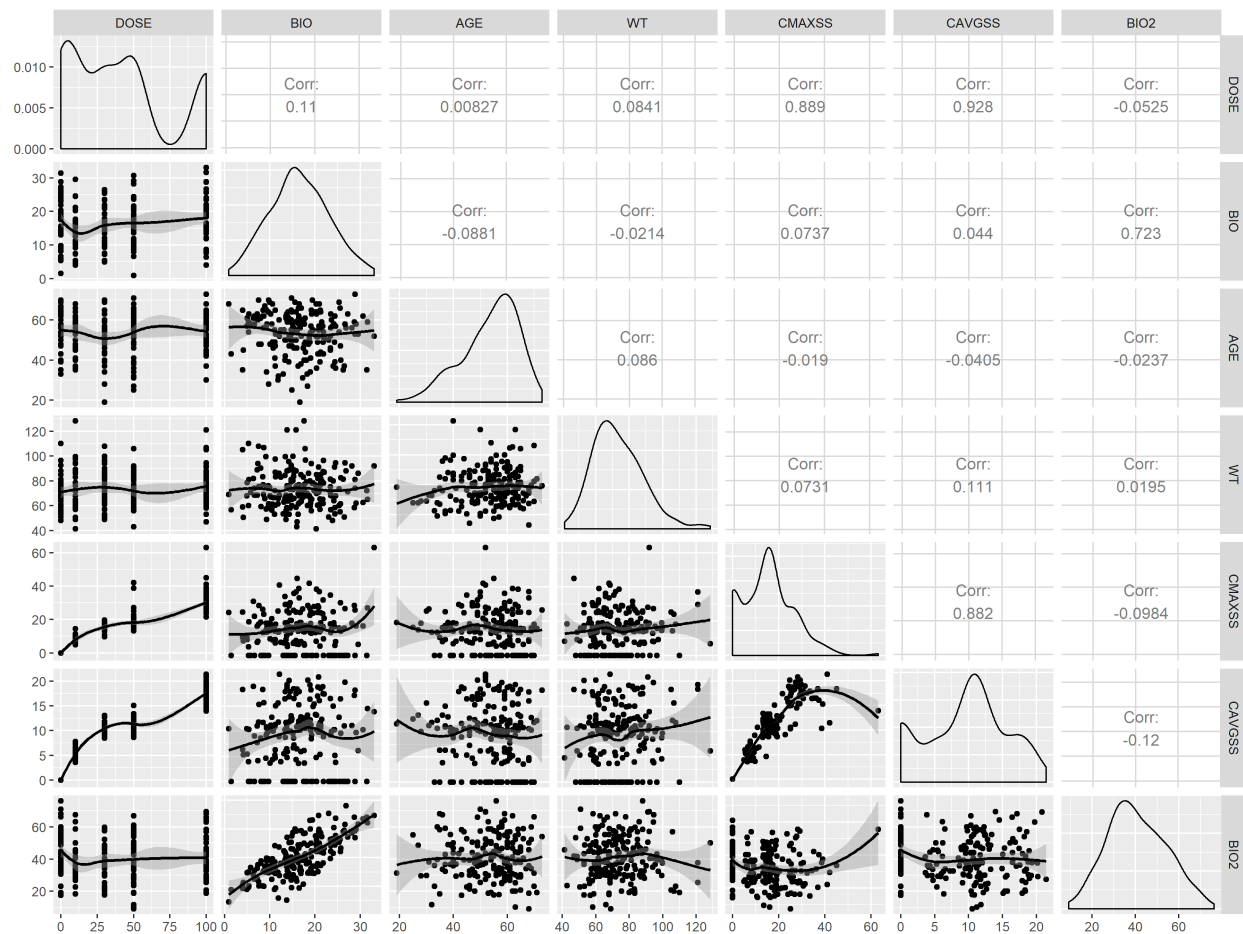


Figure 1. Covariate Relationships - Continuous.

Catergorcal

```
gg <- ggpairs(data=d,columns=c(3,5:6,9,10,13), title = "By Sex",
  upper = list(continuous = "cor"),
  diag = list(continuous= "density"),
  lower = list(continuous = "smooth_loess", mapping=aes(colour = Sex)),
  axisLabels="show")
```

```
## Warning in check_and_set_ggpairs_defaults("diag", diag, continuous =
## "densityDiag", : Changing diag$continuous from 'density' to 'densityDiag'
```

```
ggsave("Cov_Cov_10.png",gg,height=9,width=12)
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Figure 2. Covariate Relationships - By Sex.

```
gg <- ggpairs(data=d,columns=c(3,5:6,9,10,14), title = "By Smoker",
  upper = list(continuous = "cor"),
  diag = list(continuous= "density"),
  lower = list(continuous = "smooth_loess", mapping=aes(colour = Smoker)),
  axisLabels="show")
```

```
## Warning in check_and_set_ggpairs_defaults("diag", diag, continuous =
## "densityDiag", : Changing diag$continuous from 'density' to 'densityDiag'
```

```
ggsave("Cov_Cov_12.png",gg,height=9,width=12)
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

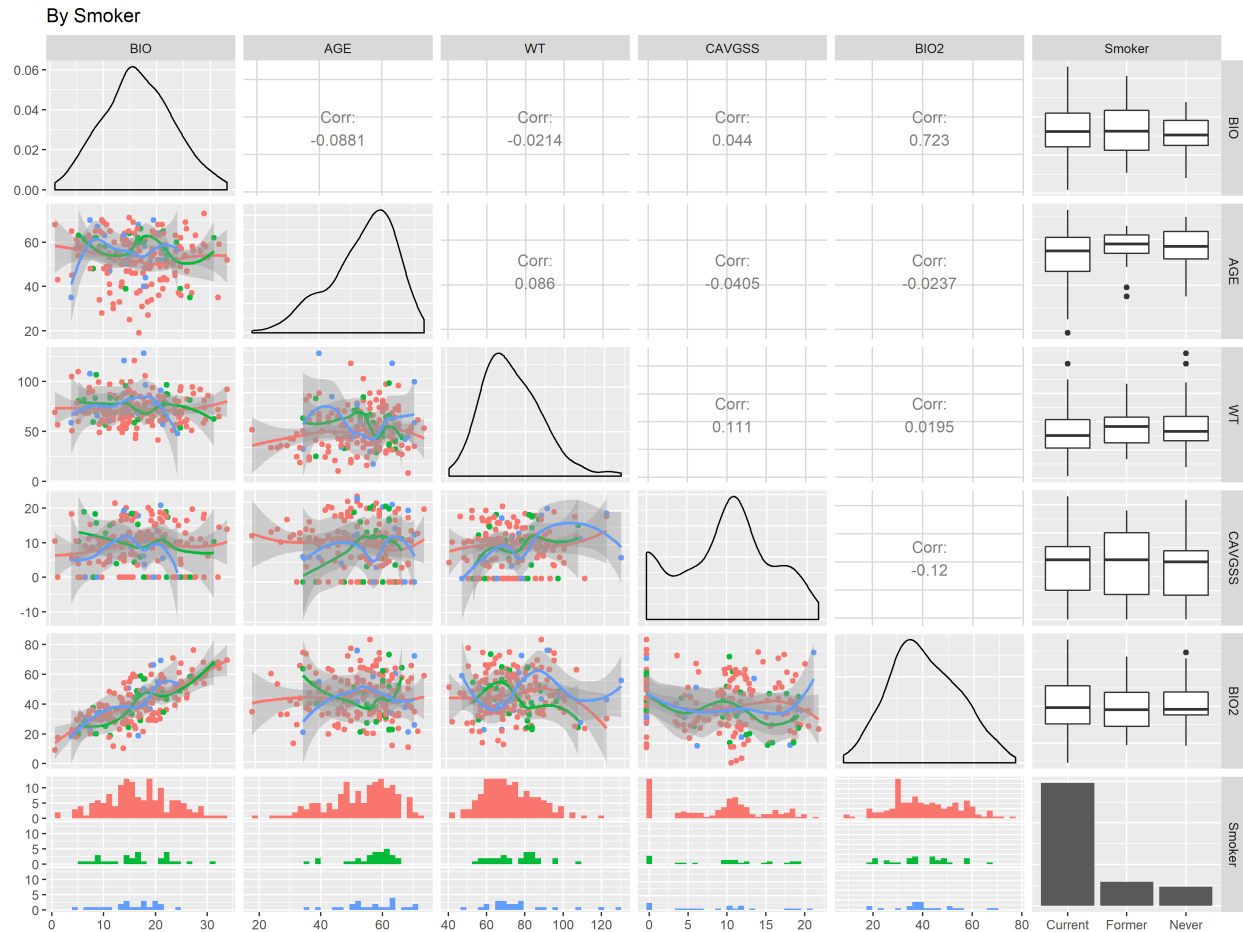


Figure 3. Covariate Relationships - By Smoker.

Observed Covariate-to-Response Relationships

Continuous

Frequency & cumulative sum by DOSE

```
d1 <- d %>% group_by(DOSE,DV) %>% summarize(N.AEGrade=nobs_2(DV)) %>%  
  mutate(cFreq = cumsum(N.AEGrade)) %>%  
  group_by(DOSE) %>% mutate(N.Tot = sum_2(N.AEGrade))  
  
d1$AEGrade <- as.factor(d1$DV)  
d1$p <- d1$N.AEGrade/d1$N.Tot  
d1$cp <- d1$cFreq/d1$N.Tot  
d1$Dose <- as.factor(d1$DOSE)  
  
gg <- ggplot(data=d1, aes(x=AEGrade, y=cp, colour=Dose, group=Dose)) +  
  geom_point() +  
  geom_line() +  
  theme(axis.text.x = element_text(angle=0, colour="black", vjust=0, size=20),  
        axis.text.y = element_text(angle=0, colour="black", vjust=0, size=20),  
        axis.title.x = element_text(size=20),  
        axis.title.y = element_text(size=20),  
        legend.position = "bottom",  
        legend.title=element_text(colour="black", size=18, face="bold"),  
        legend.text = element_text(colour="black", size=18, face="bold")) +  
  xlab("AE Grade") +  
  ylab("Cumulative Probability")  
ggsave("Cov_Resp_1_02.png",gg,height=9,width=12)
```

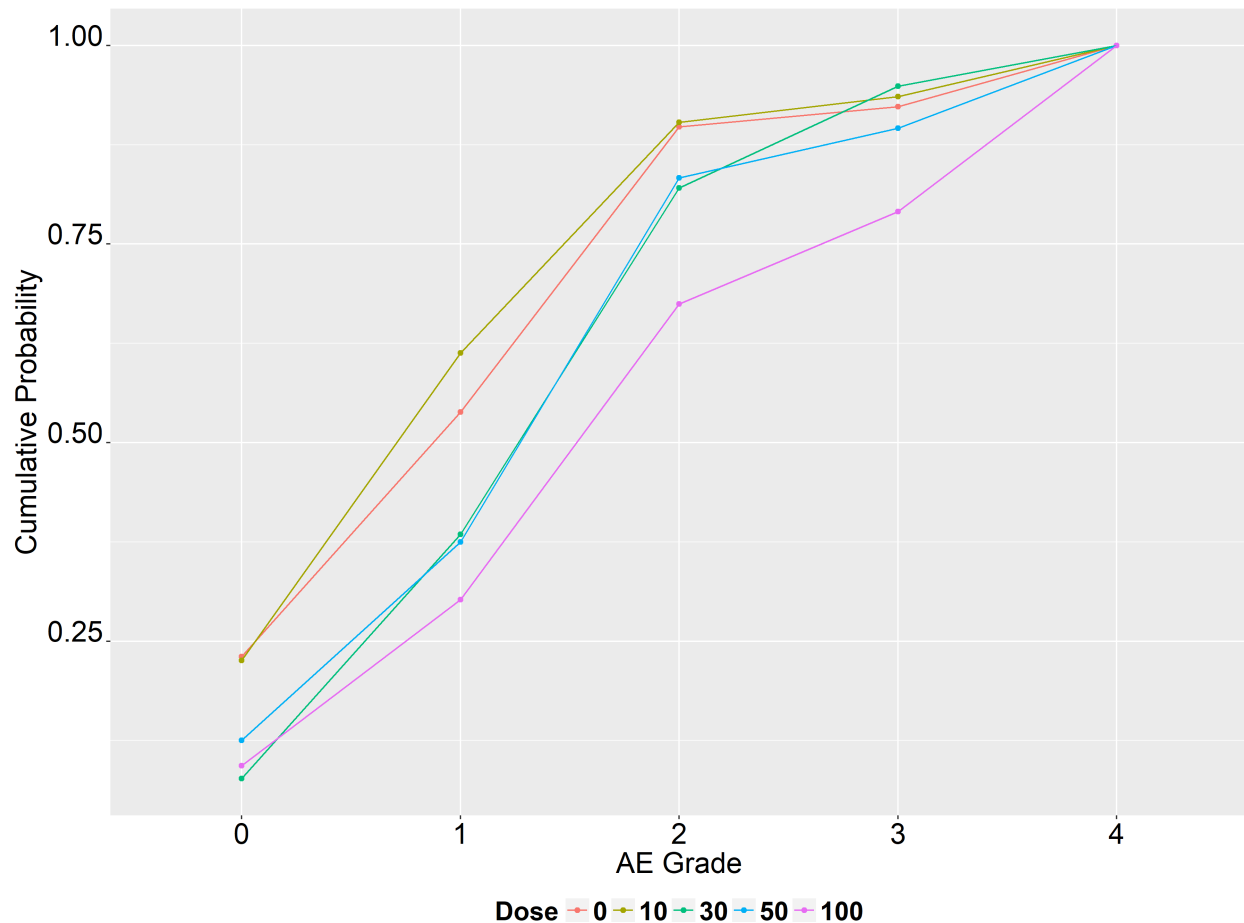


Figure 4. Cumulative Probability Plot - AE Grade vs Dose.

```
gg <- ggplot(data=d, aes(x=AEGrade, y=DOSE, colour=AEGrade)) +
  geom_boxplot(fill="white", alpha=0.5) +
  # geom_point(pch=0, size=4) +
  geom_jitter(pch=1, size=5, width = 0.05) +
  theme(axis.text.x = element_text(angle=0, colour="black", vjust=0, size=20),
        axis.text.y = element_text(angle=0, colour="black", vjust=0, size=20),
        axis.title.x = element_text(size=20),
        axis.title.y = element_text(size=20),
        legend.position = "none") +
  xlab("AEGrade") +
  ylab("Dose (mg)")
ggsave("Cov_Resp_1_04.png", gg, height=9, width=12)
```

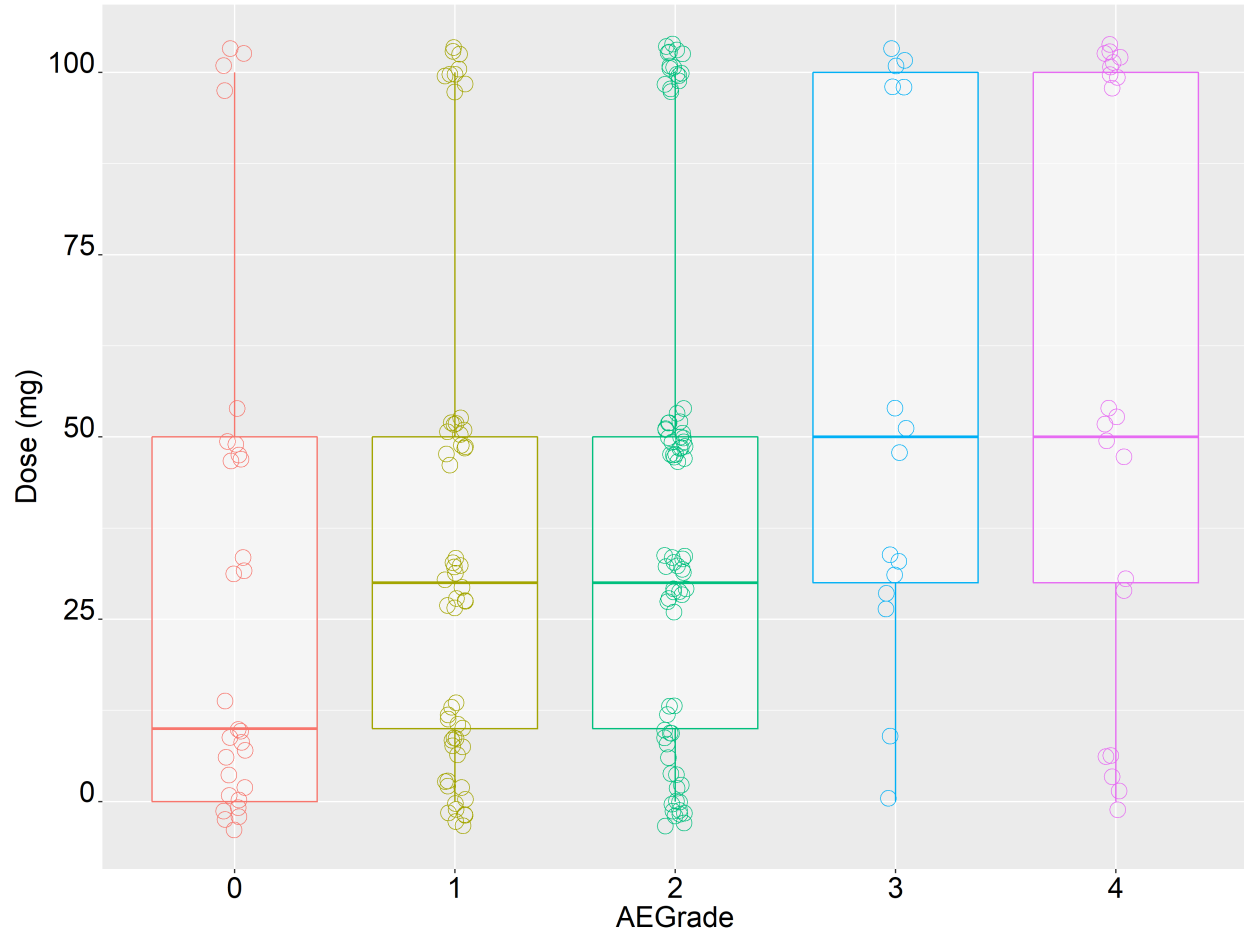


Figure 5. *AE Grade vs Dose.*

Frequency & cumulative sum by AGE

```
d2 <- d %>% mutate(q20 = quantile(AGE, prob=0.2),
  q40 = quantile(AGE, prob=0.4),
  q60 = quantile(AGE, prob=0.6),
  q80 = quantile(AGE, prob=0.8))

d2$AGEgn <- 1
d2$AGEgn[d2$AGE>d2$q20 & d2$AGE<=d2$q40] <- 2
d2$AGEgn[d2$AGE>d2$q40 & d2$AGE<=d2$q60] <- 3
d2$AGEgn[d2$AGE>d2$q60 & d2$AGE<=d2$q80] <- 4
d2$AGEgn[d2$AGE>d2$q80] <- 5

d2 <- d2 %>% group_by(AGEgn) %>% mutate(AGEgmin = min_2(AGE), AGEgm = median_2(AGE),
  AGEgmax = max_2(AGE))
d2$AGE_Group <- paste(d2$AGEgmin,d2$AGEgmax,sep=" - ")

d3 <- d2 %>% group_by(AGE_Group,DV) %>% summarize(N.AEGrade=nobs_2(DV)) %>%
  mutate(cFreq = cumsum(N.AEGrade)) %>%
  group_by(AGE_Group) %>% mutate(N.Tot = sum_2(N.AEGrade))

d3$AEGrade <- as.factor(d3$DV)
d3$p <- d3$N.AEGrade/d3$N.Tot
```



```

d3$cp <- d3$cFreq/d3$N.Tot
d3$Age <- as.factor(d3$AGE_Group)

gg <- ggplot(data=d3, aes(x=AEGrade, y=cp, colour=Age, group=Age)) +
  geom_point() +
  geom_line() +
  theme(axis.text.x = element_text(angle=0, colour="black", vjust=0, size=20),
        axis.text.y = element_text(angle=0, colour="black", vjust=0, size=20),
        axis.title.x = element_text(size=20),
        axis.title.y = element_text(size=20),
        legend.position = "bottom",
        legend.title=element_text(colour="black", size=18, face="bold"),
        legend.text = element_text(colour="black", size=18, face="bold")) +
  xlab("AE Grade") +
  ylab("Cumulative Probability")
ggsave("Cov_Respl_1_08.png", gg, height=9, width=12)

```

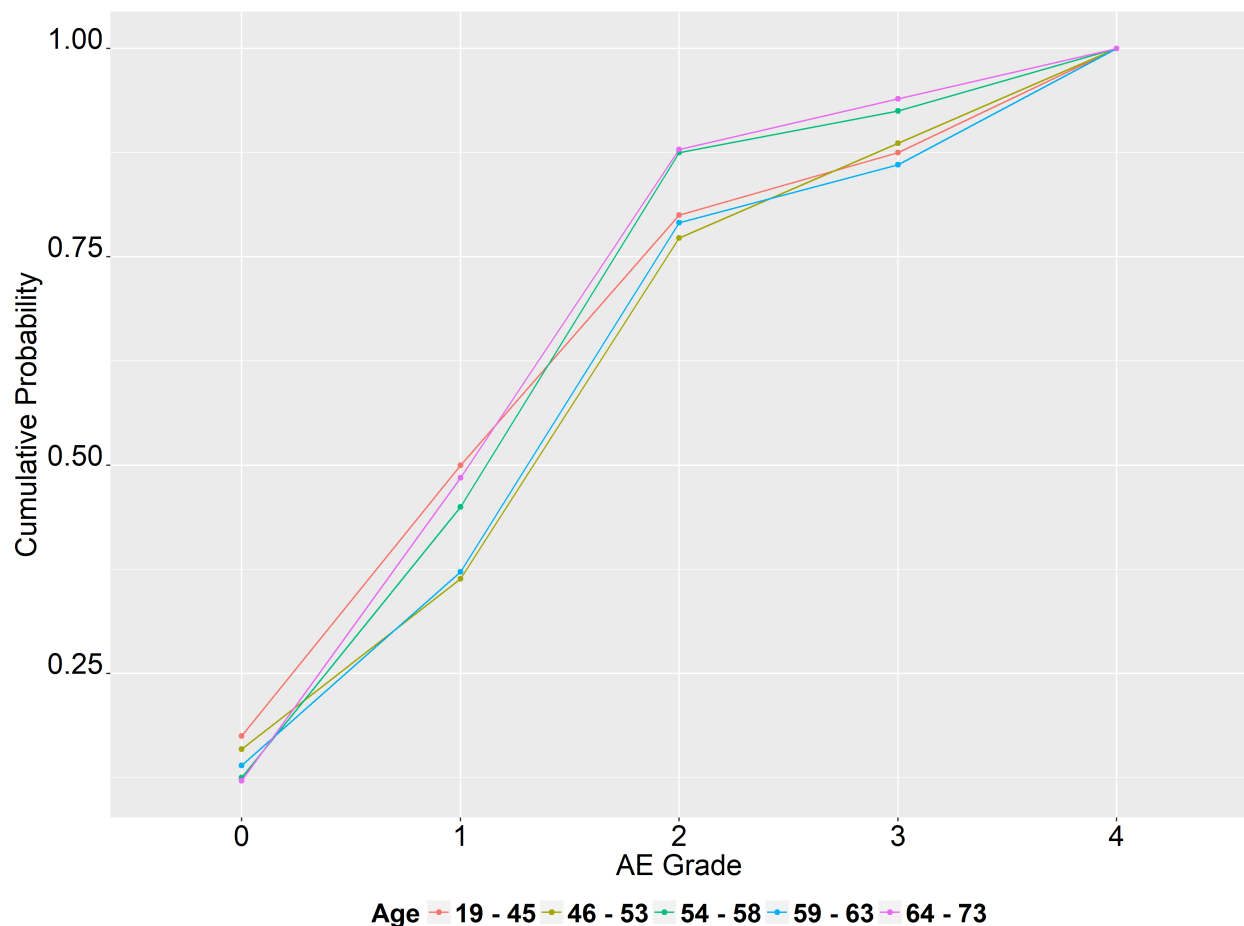


Figure 6. Cumulative Probability Plot - AE Grade vs Age.

```

gg <- ggplot(data=d, aes(x=AEGrade, y=AGE, colour=AEGrade)) +
  geom_boxplot(fill="white", alpha=0.5) +
  # geom_point(pch=0, size=4) +
  geom_jitter(pch=1, size=5, width = 0.05) +
  theme(axis.text.x = element_text(angle=0, colour="black", vjust=0, size=20),

```

```

axis.text.y = element_text(angle=0, colour="black", vjust=0, size=20),
axis.title.x = element_text(size=20),
axis.title.y = element_text(size=20),
legend.position = "none") +
xlab("AEGrade") +
ylab("Age (yrs)")
ggsave("Cov_Resp_1_10.png",gg,height=9,width=12)

```

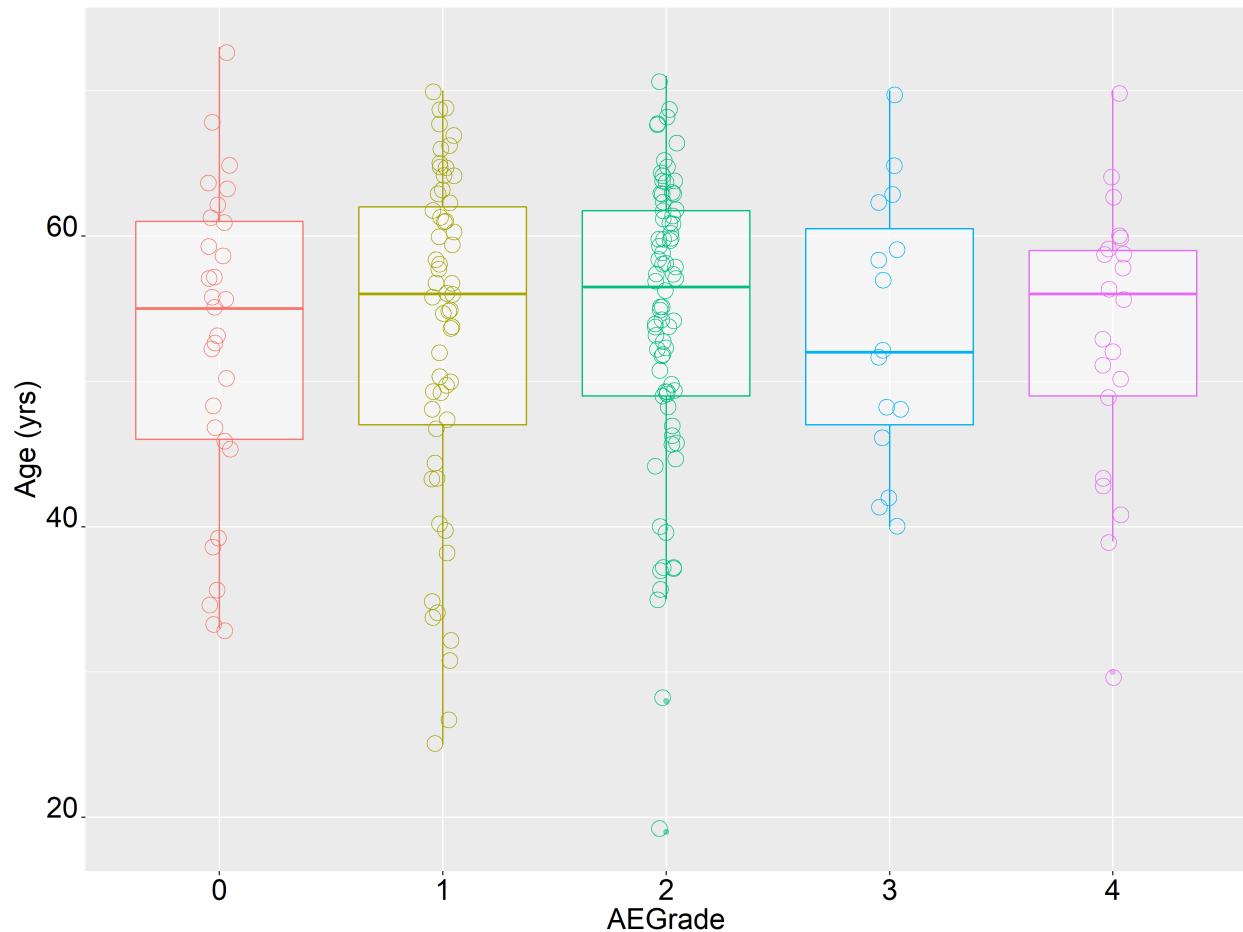


Figure 7. *AE Grade vs Age.*

Categorical

Frequency & cumulative sum by SEX

```

d2 <- d %>% group_by(SEX,DV) %>% summarize(N.AEGrade=length(DV)) %>%
  mutate(cFreq = cumsum(N.AEGrade)) %>%
  group_by(SEX) %>% mutate(N.Tot = sum(N.AEGrade))

d2$AEGrade <- as.factor(d2$DV)
d2$p <- d2$N.AEGrade/d2$N.Tot
d2$cp <- d2$cFreq/d2$N.Tot
d2$Sex <- factor(d2$SEX,levels=c(1,2), labels=c("Male","Female"))

```

```
gg <- ggplot(data=d2, aes(x=AEGrade, y=cp, colour=Sex, group=Sex)) +
  geom_point() +
  geom_line() +
  theme(axis.text.x = element_text(angle=0, colour="black", vjust=0, size=20),
        axis.text.y = element_text(angle=0, colour="black", vjust=0, size=20),
        axis.title.x = element_text(size=20),
        axis.title.y = element_text(size=20),
        legend.position = "bottom",
        legend.title=element_text(colour="black", size=18, face="bold"),
        legend.text = element_text(colour="black", size=18, face="bold")) +
  xlab("AE Grade") +
  ylab("Cumulative Probability")
ggsave("Cov_Resp_2_01.png", gg, height=9, width=12)
```

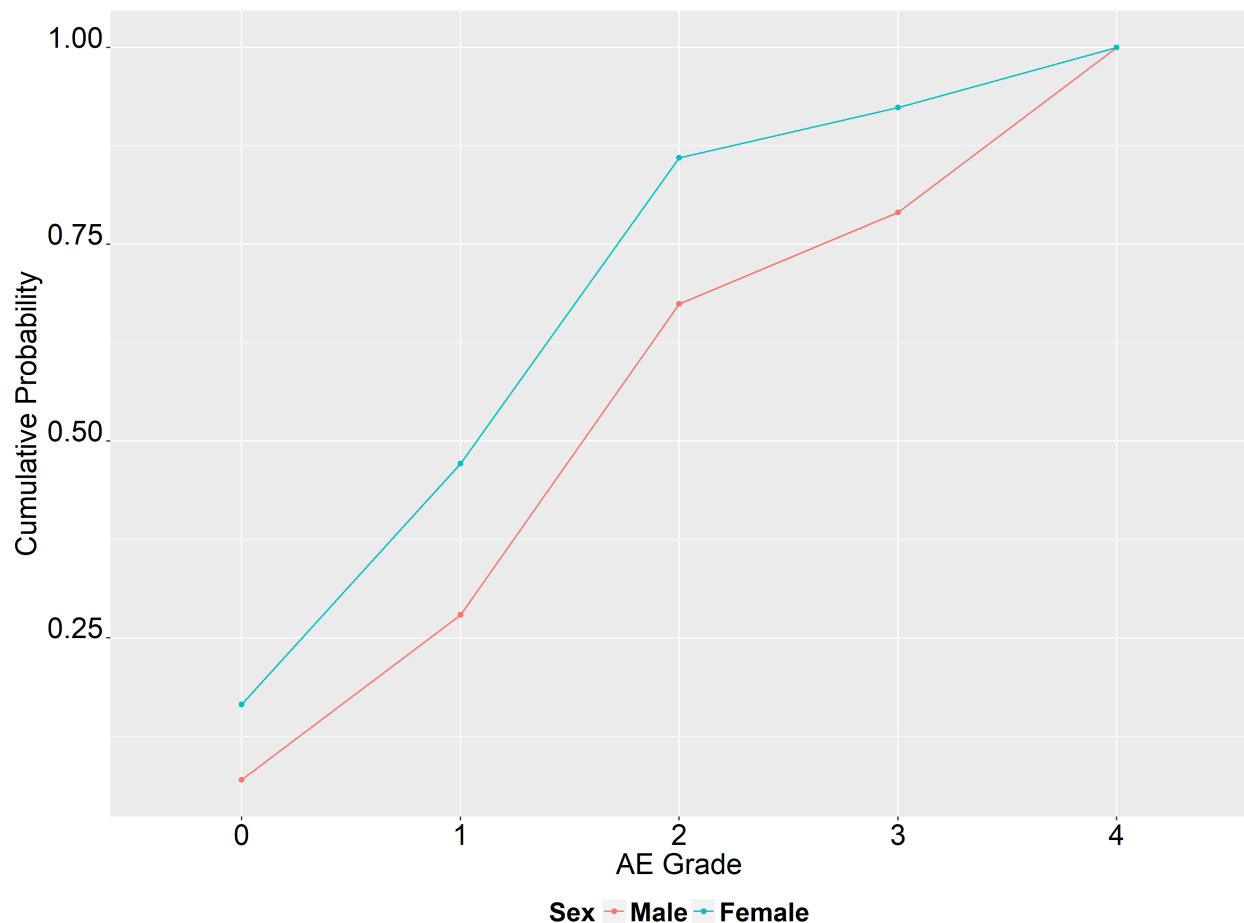


Figure 8. Cumulative Probability Plot - AE Grade vs Sex.

```
gg <- ggplot(data=d, aes(x=Sex, y=DV, colour=AEGrade)) +
  geom_point() +
  # geom_point(pch=0, size=4) +
  geom_jitter(pch=1, size=5, width = 0.05) +
  theme(axis.text.x = element_text(angle=0, colour="black", vjust=0, size=20),
        axis.text.y = element_text(angle=0, colour="black", vjust=0, size=20),
        axis.title.x = element_text(size=20),
        axis.title.y = element_text(size=20),
```

```

    legend.position = "none") +
  xlab("Sex") +
  ylab("AE Grade")
ggsave("Cov_Resp_2_03.png", gg, height=9, width=12)

```

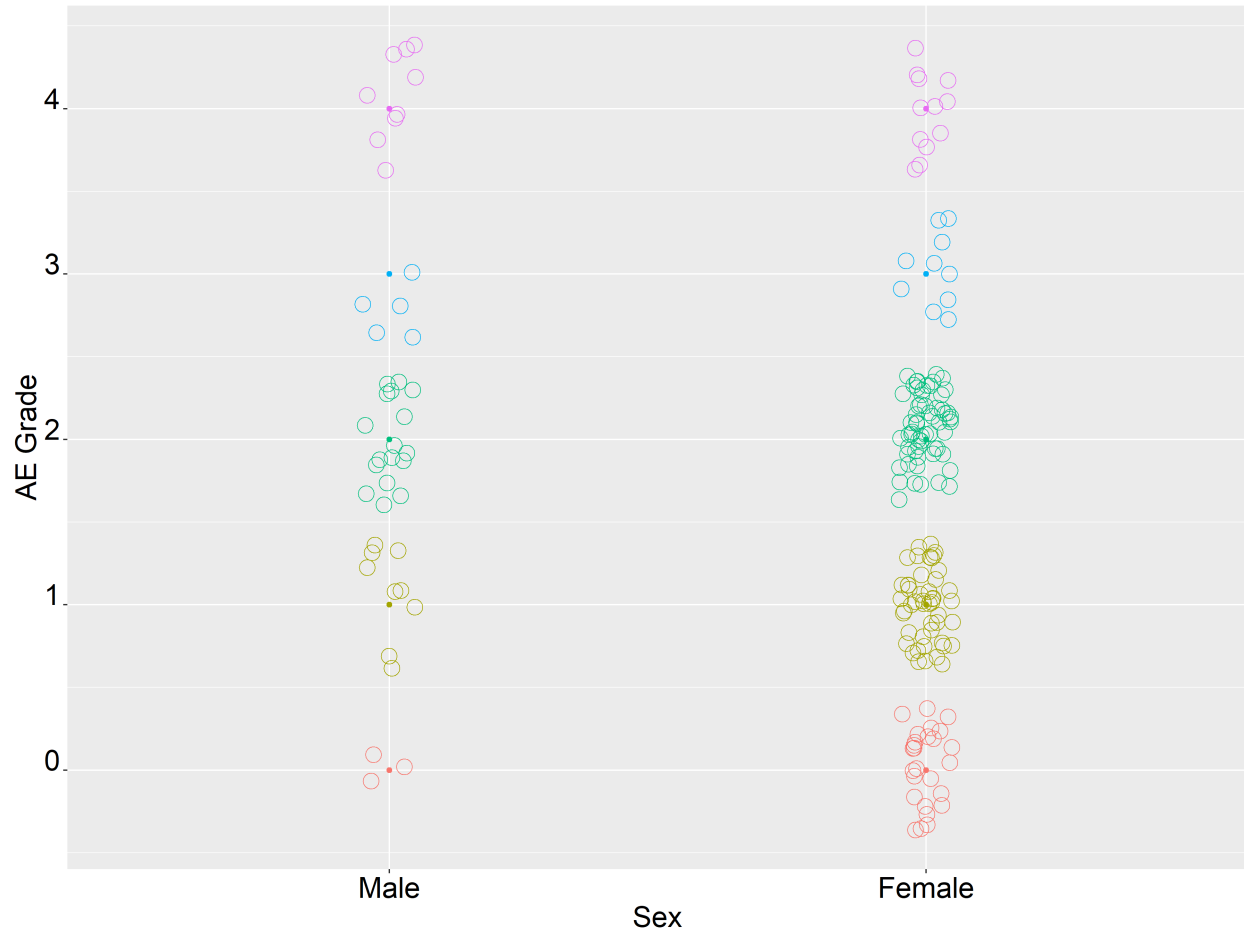


Figure 9. *AE Grade vs Sex.*

Frequency & cumulative sum by SMOKER

```

d3 <- d %>% group_by(SMOK,DV) %>% summarize(N.AEGrade=length(DV)) %>%
  mutate(cFreq = cumsum(N.AEGrade)) %>%
  group_by(SMOK) %>% mutate(N.Tot = sum(N.AEGrade))

d3$AEGrade <- as.factor(d3$DV)
d3$p <- d3$N.AEGrade/d3$N.Tot
d3$cp <- d3$cFreq/d3$N.Tot
d3$Smoker <- factor(d3$SMOK,levels=c(1,2,3), labels=c("Current","Former","Never"))

gg <- ggplot(data=d3, aes(x=AEGrade, y=cp, colour=Smoker, group=Smoker)) +
  geom_point() +
  geom_line() +
  theme(axis.text.x = element_text(angle=0, colour="black", vjust=0, size=20),
        axis.text.y = element_text(angle=0, colour="black", vjust=0, size=20),

```

```

axis.title.x = element_text(size=20),
axis.title.y = element_text(size=20),
legend.position = "bottom",
legend.title=element_text(colour="black", size=18, face="bold"),
legend.text = element_text(colour="black", size=18, face="bold")) +
xlab("AE Grade") +
ylab("Cumulative Probability")
ggsave("Cov_Resp_2_04.png",gg,height=9,width=12)

```

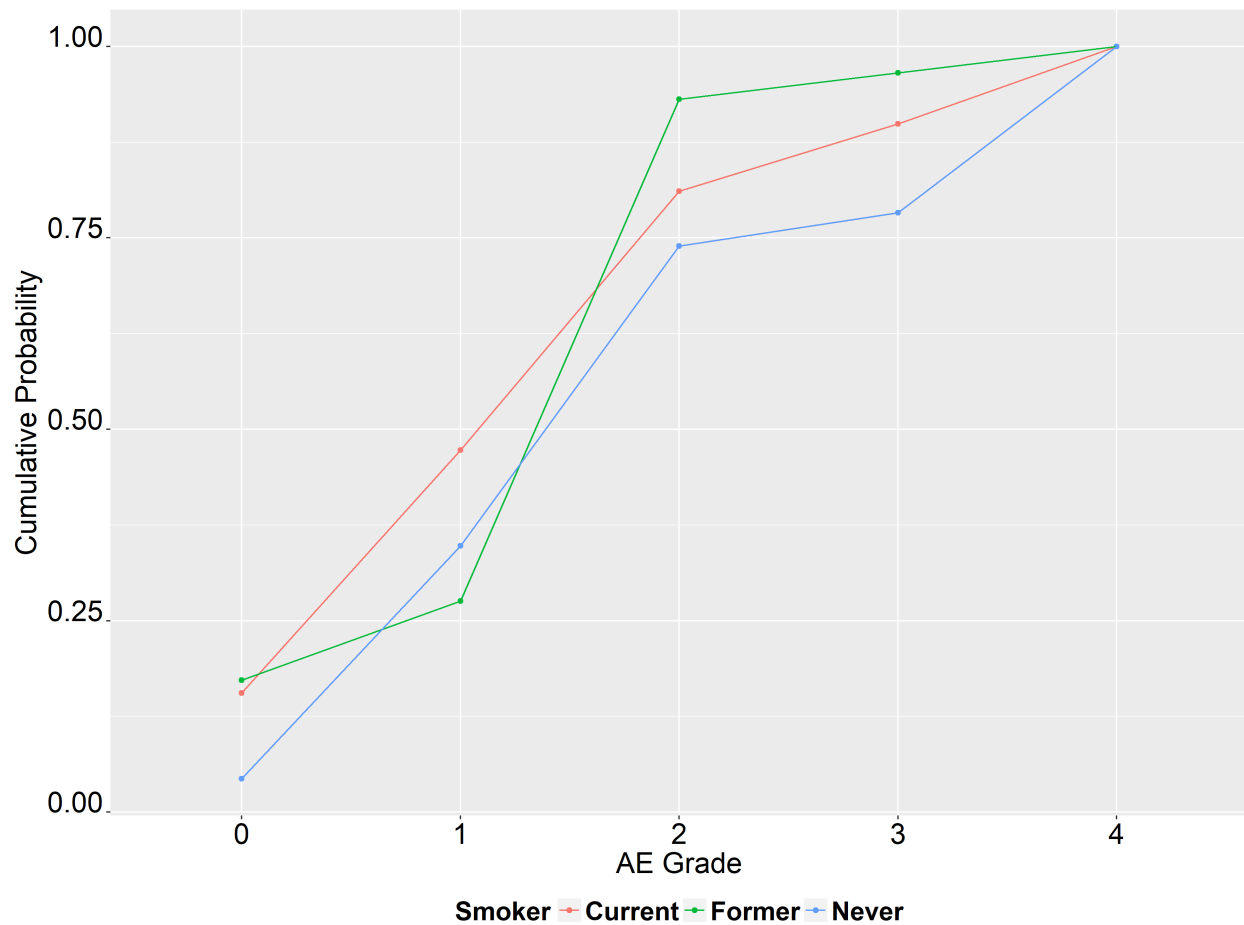


Figure 10. Cumulative Probability Plot - AE Grade vs Smoker.

```

gg <- ggplot(data=d, aes(x=Smoker, y=DV, colour=AEGrade)) +
  geom_point() +
  # geom_point(pch=0, size=4) +
  geom_jitter(pch=1, size=5, width = 0.05) +
  theme(axis.text.x = element_text(angle=0, colour="black", vjust=0, size=20),
        axis.text.y = element_text(angle=0, colour="black", vjust=0, size=20),
        axis.title.x = element_text(size=20),
        axis.title.y = element_text(size=20),
        legend.position = "none") +
  xlab("Smoker") +
  ylab("AE Grade")
ggsave("Cov_Resp_2_06.png",gg,height=9,width=12)

```

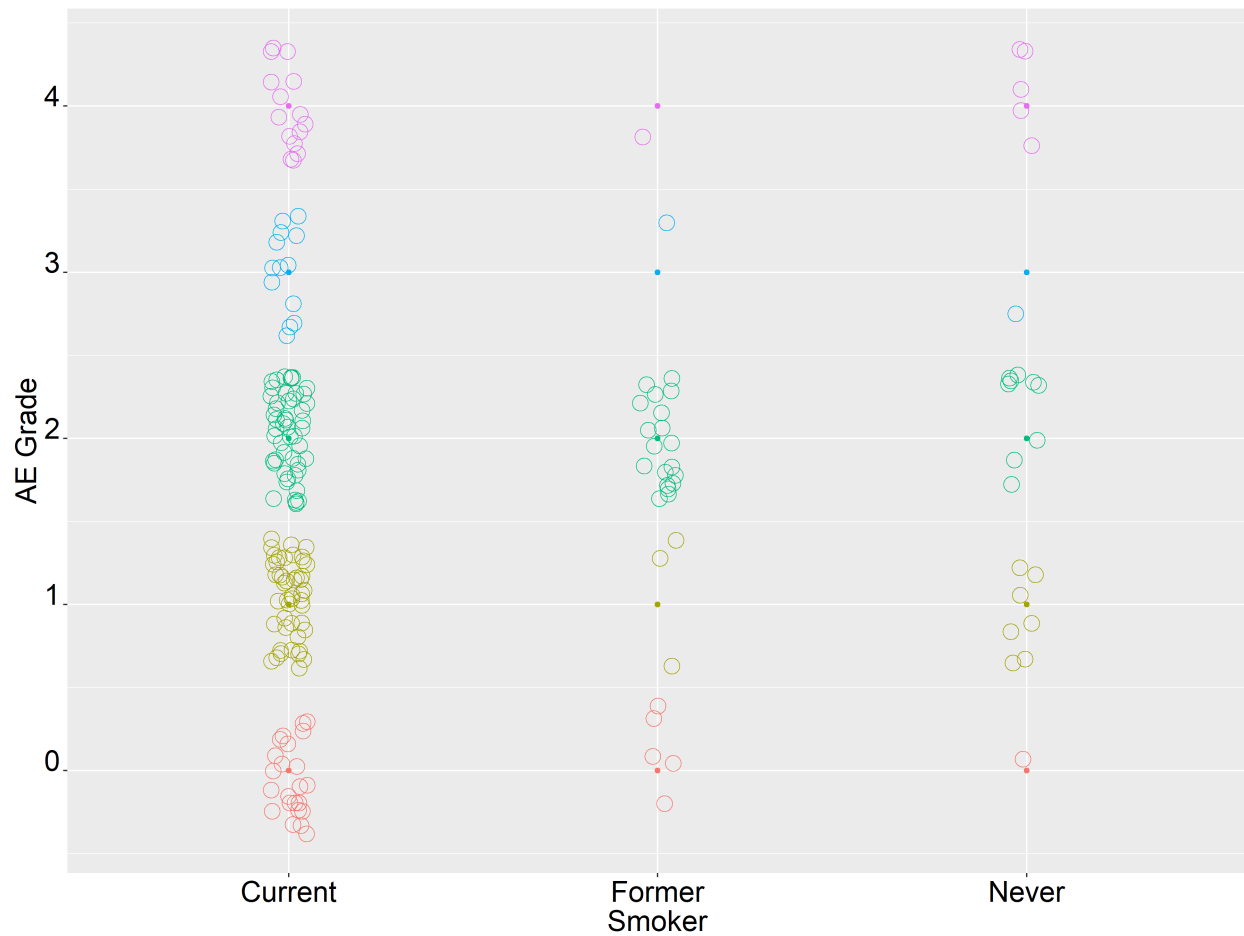


Figure 11. *AE Grade vs Smoker.*

Question 2:

Do a scatter plot for the 3 exposure measures.

Which measure would you choose for further covariate model building?

```
gg <- ggpairs(d[,c(2,8,9)],lower=list(continuous='smooth_loess'),
             diag=list(continuous='density'),
             axisLabels = 'show')
```

```
## Warning in check_and_set_ggpairs_defaults("diag", diag, continuous =
## "densityDiag", : Changing diag$continuous from 'density' to 'densityDiag'
```

```
ggsave("Cov_Cov_04.png",gg,height=9,width=12)
```

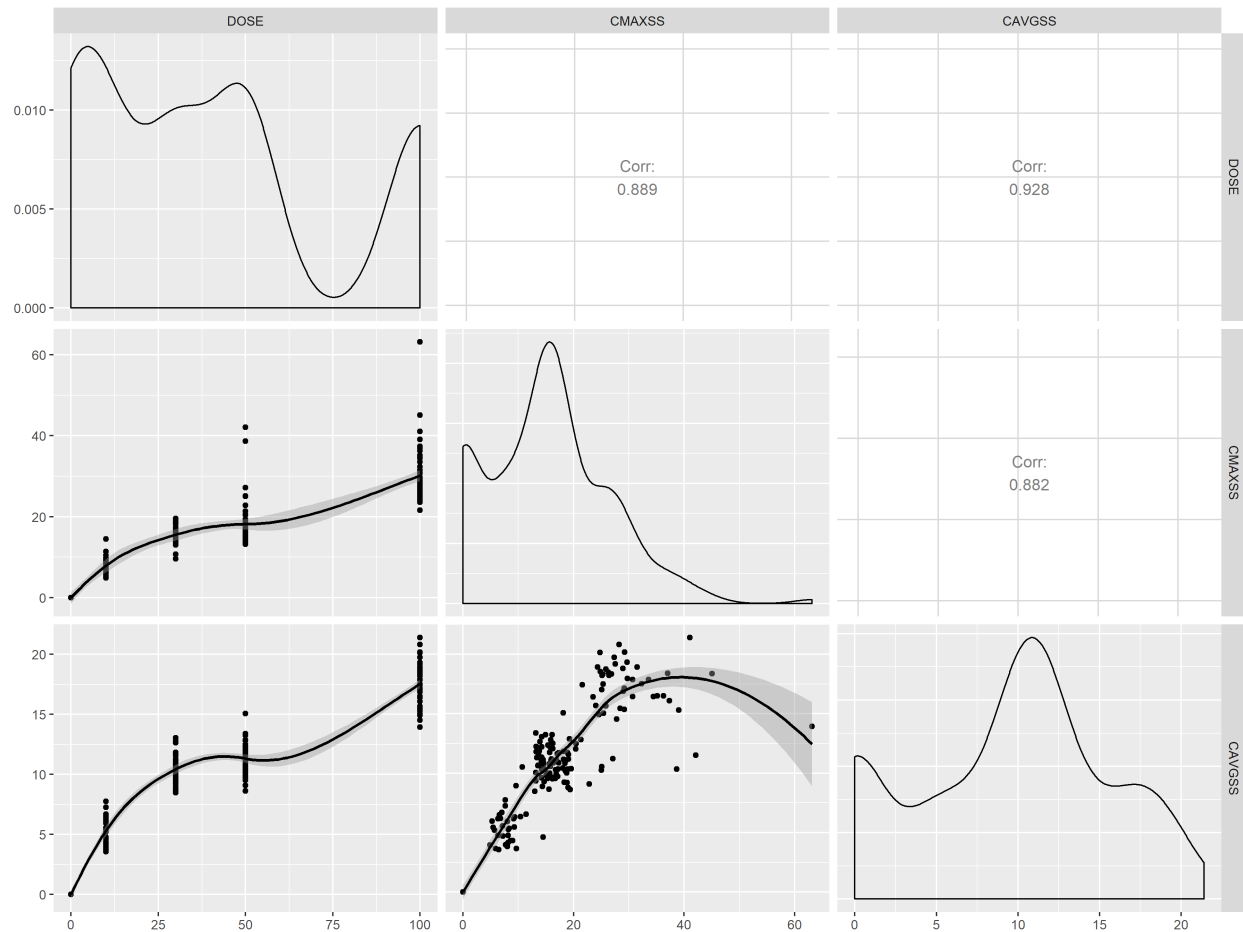


Figure 12. Covariate Relationships - Exposure.

Question 3

With CMAXSS as the exposure measure in the base model, use `stepAIC()` to do a complete forward selection followed by a complete backward selection to select a final model.

Models

Exposure = CMAXSS

```
d$EXP <- d$CMAXSS
```

Null Model

```
mod0 <- clm(AEGrade ~ 1, data = d, link = "logit", threshold="flex", Hess=T)
summary(mod0, digits = 3)
```

```
## formula: AEGrade ~ 1
## data:    d
##
## link threshold nobs logLik AIC      niter max.grad cond.H
## logit flexible  200 -287.18 582.36 5(0)  2.88e-08 8.7e+00
##
## Threshold coefficients:
##      Estimate Std. Error z value
## 0|1  -1.7744     0.2008  -8.835
## 1|2  -0.2819     0.1428  -1.973
## 2|3   1.5163     0.1841   8.239
## 3|4   2.1429     0.2307   9.290
```

```
inv.logit(mod0$coef[1:4])
```

```
##    0|1    1|2    2|3    3|4
## 0.145 0.430 0.820 0.895
```

Base Model

```
mod1 <- clm(AEGrade~ EXP , data = d, link = "logit", threshold="flex", Hess=T)
summary(mod1, digits = 3)
```

```
## formula: AEGrade ~ EXP
## data:    d
##
## link threshold nobs logLik AIC      niter max.grad cond.H
## logit flexible  200 -277.92 565.85 5(0)  4.73e-07 4.9e+03
##
## Coefficients:
##      Estimate Std. Error z value Pr(>|z|)
## EXP  0.05240     0.01239   4.229 2.34e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Threshold coefficients:
##      Estimate Std. Error z value
## 0|1  -1.0934     0.2537  -4.311
## 1|2   0.4735     0.2289   2.069
## 2|3   2.3994     0.2894   8.291
## 3|4   3.0612     0.3293   9.295
```

```
exp(mod1$beta)
```

```
##      EXP
## 1.053793
```

```
inv.logit(mod1$coef[1:4])
```

```
##      0|1      1|2      2|3      3|4
## 0.2509704 0.6162227 0.9167782 0.9552623
```


Full Model

```
mod2 <- clm(AEGrade ~ EXP + BIO + SEX + AGE + WT + SMOK + BIO2 , data = d,  
            link = "logit", threshold="flex", Hess=T)  
summary(mod2)
```

```
## formula: AEGrade ~ EXP + BIO + SEX + AGE + WT + SMOK + BIO2  
## data:      d  
##  
## link threshold nobs logLik AIC      niter max.grad cond.H  
## logit flexible  200 -265.75 553.49 6(0)  1.53e-12 5.6e+06  
##  
## Coefficients:  
##      Estimate Std. Error z value Pr(>|z|)  
## EXP    0.0544206  0.0128231   4.244  2.2e-05 ***  
## BIO   -0.0515692  0.0302779  -1.703   0.0885 .  
## SEX   -0.7612586  0.3661127  -2.079   0.0376 *  
## AGE   -0.0007764  0.0127251  -0.061   0.9514  
## WT     0.0133934  0.0098390   1.361   0.1734  
## SMOK   0.1082145  0.2027697   0.534   0.5936  
## BIO2  -0.0108457  0.0141631  -0.766   0.4438  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Threshold coefficients:  
##      Estimate Std. Error z value  
## 0|1    -2.767      1.494  -1.852  
## 1|2    -1.093      1.485  -0.736  
## 2|3     1.012      1.484   0.682  
## 3|4     1.719      1.492   1.152
```

```
exp(mod2$beta)
```

```
##      EXP      BIO      SEX      AGE      WT      SMOK      BIO2  
## 1.0559286 0.9497379 0.4670782 0.9992239 1.0134835 1.1142867 0.9892129
```

```
inv.logit(mod2$coef[1:4])
```

```
##      0|1      1|2      2|3      3|4  
## 0.05913922 0.25098031 0.73339518 0.84794492
```

Forward & Backward Selection Separately

Forward Selection: Start with Base model

```
stepAIC(mod1, scope=list(lower=~EXP, upper=~EXP + Sex + AGE + WT + Smoker + BIO + BIO2),  
         direction="forward", trace=1, k=2)
```

```
## Start:  AIC=565.85  
## AEGrade ~ EXP
```

```

##
##           Df      AIC
## + BIO      1 555.35
## + Sex       1 557.44
## + BIO2      1 559.39
## + WT        1 561.29
## <none>      565.85
## + Smoker    2 567.13
## + AGE       1 567.36
##
## Step: AIC=555.35
## AEGrade ~ EXP + BIO
##
##           Df      AIC
## + Sex       1 548.10
## + WT        1 551.52
## <none>      555.35
## + BIO2      1 557.12
## + AGE       1 557.22
## + Smoker    2 557.39
##
## Step: AIC=548.1
## AEGrade ~ EXP + BIO + Sex
##
##           Df      AIC
## <none>      548.10
## + WT        1 548.34
## + BIO2      1 549.60
## + AGE       1 550.10
## + Smoker    2 551.50

## formula: AEGrade ~ EXP + BIO + Sex
## data:    d
##
## link threshold nobs logLik AIC      niter max.grad cond.H
## logit flexible  200 -267.05 548.10 6(0)  2.95e-13 3.0e+04
##
## Coefficients:
##      EXP      BIO SexFemale
##  0.0570  -0.0699  -0.9756
##
## Threshold coefficients:
##      0|1      1|2      2|3      3|4
## -3.0663 -1.4059  0.6785  1.3793

```

Model from forward selection

```

mod3 <- clm(AEGrade ~ EXP + BIO + Sex, data = d, link = "logit", threshold="flex", Hess=T)
summary(mod3)

```

```

## formula: AEGrade ~ EXP + BIO + Sex
## data:    d
##

```

```
## link threshold nobs logLik AIC niter max.grad cond.H
## logit flexible 200 -267.05 548.10 6(0) 2.95e-13 3.0e+04
##
## Coefficients:
## Estimate Std. Error z value Pr(>|z|)
## EXP 0.05700 0.01252 4.554 5.28e-06 ***
## BIO -0.06990 0.02097 -3.334 0.000857 ***
## SexFemale -0.97560 0.32393 -3.012 0.002597 **
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Threshold coefficients:
## Estimate Std. Error z value
## 0|1 -3.0663 0.5081 -6.035
## 1|2 -1.4059 0.4722 -2.977
## 2|3 0.6785 0.4641 1.462
## 3|4 1.3793 0.4834 2.853
```

Backward Selection: Start with final forward model

```
summary(stepAIC(mod3, scope=list(upper=~EXP + BIO + Sex), direction="backward", trace=1, k=2))
```

```
## Start: AIC=548.1
## AEGrade ~ EXP + BIO + Sex
##
## Df AIC
## <none> 548.10
## - Sex 1 555.35
## - BIO 1 557.44
## - EXP 1 567.56

## formula: AEGrade ~ EXP + BIO + Sex
## data: d
##
## link threshold nobs logLik AIC niter max.grad cond.H
## logit flexible 200 -267.05 548.10 6(0) 2.95e-13 3.0e+04
##
## Coefficients:
## Estimate Std. Error z value Pr(>|z|)
## EXP 0.05700 0.01252 4.554 5.28e-06 ***
## BIO -0.06990 0.02097 -3.334 0.000857 ***
## SexFemale -0.97560 0.32393 -3.012 0.002597 **
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Threshold coefficients:
## Estimate Std. Error z value
## 0|1 -3.0663 0.5081 -6.035
## 1|2 -1.4059 0.4722 -2.977
## 2|3 0.6785 0.4641 1.462
## 3|4 1.3793 0.4834 2.853
```

Model 3 is the Final Model

Question 4:

Comparison of Models Using Relative Likelihood

Base model vs Final model

```
AICr(mod3,mod1)
```

```
## [1] 0.0001398416
```

Full model vs Final model

```
AICr(mod3,mod2)
```

```
## [1] 0.06754238
```

BONUS Question:

Using `clm()` with exposure, is there a way to use AIC to help decide which exposure measure to use for the AE grades data?

Compare AIC and relative likelihood for a base model for each exposure measure.
DOSE as exposure

```
mod4 <- clm(AEGrade ~ DOSE, data = d, link = "logit", threshold="flex", Hess=T)
summary(mod4)
```

```
## formula: AEGrade ~ DOSE
## data:    d
##
## link threshold nobs logLik AIC    niter max.grad cond.H
## logit flexible  200 -281.23 572.47 5(0)  5.24e-07 3.2e+04
##
## Coefficients:
##      Estimate Std. Error z value Pr(>|z|)
## DOSE  0.01291    0.00379   3.406  0.00066 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Threshold coefficients:
##      Estimate Std. Error z value
## 0|1  -1.3198    0.2380  -5.545
## 1|2   0.2203    0.2049   1.075
## 2|3   2.1010    0.2597   8.089
## 3|4   2.7494    0.3000   9.164
```

CAVGSS as exposure

```
mod5 <- clm(AEGrade ~ CAVGSS, data = d, link = "logit", threshold="flex", Hess=T)
summary(mod5)
```

```
## formula: AEGrade ~ CAVGSS
## data:    d
##
## link threshold nobs logLik AIC      niter max.grad cond.H
## logit flexible  200 -281.19 572.39 5(0)  1.67e-07 2.0e+03
##
## Coefficients:
##      Estimate Std. Error z value Pr(>|z|)
## CAVGSS  0.07571    0.02213   3.421 0.000623 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Threshold coefficients:
##      Estimate Std. Error z value
## 0|1  -1.1375    0.2696  -4.220
## 1|2   0.4078    0.2471   1.650
## 2|3   2.2845    0.2982   7.660
## 3|4   2.9304    0.3343   8.767
```

CMAXSS as exposure

```
mod6 <- clm(AEGrade ~ CMAXSS, data = d, link = "logit", threshold="flex", Hess=T)
summary(mod6)
```

```
## formula: AEGrade ~ CMAXSS
## data:    d
##
## link threshold nobs logLik AIC      niter max.grad cond.H
## logit flexible  200 -277.92 565.85 5(0)  4.73e-07 4.9e+03
##
## Coefficients:
##      Estimate Std. Error z value Pr(>|z|)
## CMAXSS  0.05240    0.01239   4.229 2.34e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Threshold coefficients:
##      Estimate Std. Error z value
## 0|1  -1.0934    0.2537  -4.311
## 1|2   0.4735    0.2289   2.069
## 2|3   2.3994    0.2894   8.291
## 3|4   3.0612    0.3293   9.295
```

Relative Likelihood, CAVGSS vs DOSE

```
AICr(mod5,mod4)
```

```
## [1] 0.9607894
```

Relative Likelihood, CMAXSS vs DOSE

```
AICr(mod6,mod4)
```

```
## [1] 0.03651617
```

Relative Likelihood, CMAXSS vs CAVGSS

```
AICr(mod6,mod5)
```

```
## [1] 0.03800643
```

fini