Assignment_06 - STAT 689

Philip Anderson; panders2@tamu.edu

2/14/2018

```
# import third-party modules
library("HRW")
library("tidyverse")
library("mgcv")
library("nlme")
```

Read in our data

```
fram <- read.csv("/Users/panders2/Documents/schools/tamu/stat_689/homework/semiparametric-regression/hw
names(fram) <- tolower(names(fram))
dim(fram)
```

```
## [1] 1615   10
```

Create new variables for the systolic blood pressure readings and the two cholesterol measurements.

```
# systolic blood pressure first
fram$lsbp <- log(((fram$sbp21 + fram$sbp22 + fram$sbp31 + fram$sbp32) / 4) - 50)
# cholesterol measurements second
fram$lcholest <- log((fram$cholest2 + fram$cholest3) / 2)
```

Keep only the variables that we will be working with directly and make sure everything seems reasonable.

```
fram2 <- fram %>%
  dplyr::select(chd, age, smoker, lsbp, lcholest)
head(fram2)
```

```
##   chd age smoker     lsbp lcholest
## 1   0  56      0 4.317488 5.654242
## 2   0  38      1 4.241327 5.451038
## 3   0  54      1 4.347047 5.654242
## 4   0  42      1 4.185860 5.541264
## 5   0  47      1 4.454347 5.583496
## 6   0  43      1 4.269697 5.298317
```

Question 1

Fit a multiple linear regression of LSBP on lcholest and smoker via "lm" function. Produce a estimates, standard errors, and p-values.

```
# fit model
lin_mod <- lm(lsbp ~ smoker + lcholest
              , data=fram2
              )
# produce summary
summary(lin_mod)
```

```
##
## Call:
```

```
## lm(formula = lsbp ~ smoker + lcholest, data = fram2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.79148 -0.14009 -0.02043  0.10915  0.93289
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.55569    0.17029  20.880  < 2e-16 ***
## smoker      -0.03796    0.01251  -3.034  0.00246 **
## lcholest     0.15540    0.03140   4.949 8.22e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2107 on 1612 degrees of freedom
## Multiple R-squared:  0.02036,    Adjusted R-squared:  0.01915
## F-statistic: 16.75 on 2 and 1612 DF,  p-value: 6.299e-08
```

Question 2

Conduct web research on whether smokers have higher or lower blood pressure on average compared to non-smokers.

WebMD indicates that individuals who smoke tend to have higher blood pressure than those who do not. This is not consistent with my findings from Question 1, which indicate that participants who smoke have lower blood pressure than those who do not, conditional on our transformed cholesterol variable. There is nothing in the documentation to indicate that smoker is not encoded with '1' as the positive class. This is suspicious, and suggests that we need to check our data or revisit our model specification.

Question 3

The model produces the expectation of LSBP given smoking status *conditional on* cholesterol.

Question 4

Recreate the same model as in Question 1, but add in an interaction amongst the independent variables.

```
lin_mod2 <- lm(lsbp ~ smoker + lcholest + smoker:lcholest
               , data=fram2
               )

summary(lin_mod2)
```

```
##
## Call:
## lm(formula = lsbp ~ smoker + lcholest + smoker:lcholest, data = fram2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.79151 -0.14039 -0.02046  0.10916  0.93280
```

```
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)      3.55075    0.33525  10.591   <2e-16 ***
## smoker          -0.03130    0.38907  -0.080   0.9359
## lcholest         0.15632    0.06191   2.525   0.0117 *
## smoker:lcholest -0.00123    0.07184  -0.017   0.9863
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2108 on 1611 degrees of freedom
## Multiple R-squared:  0.02036,    Adjusted R-squared:  0.01854
## F-statistic: 11.16 on 3 and 1611 DF,  p-value: 2.993e-07
```

The smoking indicator is still negatively associated with our blood pressure variable, but is no longer significant in the presence of the interaction term.

Question 5

Run a semiparametric ANCOVA with mgcv, the semiparametric version of an ANCOVA without an interaction.

```
semi_mod <- mgcv::gam(lsbp ~ factor(smoker) +
                        s(lcholest, k=23, bs="cr")
                      , data=fram2
                      , method="REML"
                     )
summary(semi_mod)
```

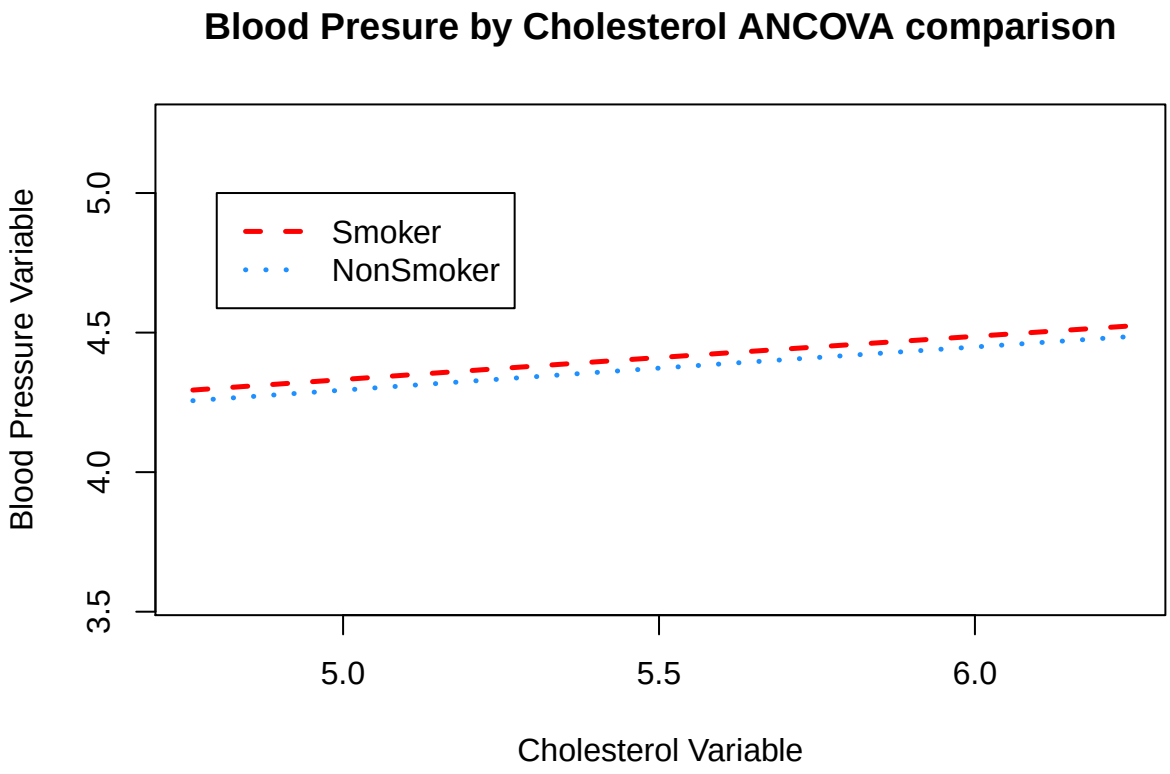
```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## lsbp ~ factor(smoker) + s(lcholest, k = 23, bs = "cr")
##
## Parametric coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)      4.39713    0.01100 399.735  < 2e-16 ***
## factor(smoker)1 -0.03799    0.01251  -3.036  0.00244 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                edf Ref.df     F  p-value
## s(lcholest) 1.064  1.126 22.27 1.73e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.0192   Deviance explained = 2.04%
## -REML = -214.85  Scale est. = 0.044402  n = 1615
```

Question 6

For the data in part 5, display a plot of the two lines, but without the data

```
# first, generate the data required for plotting

x_var <- seq(from=min(fram2$lcholest)
             , to=max(fram2$lcholest)
             , len=1000
             )
f_hat_smoker <- predict(semi_mod
                        , newdata=data.frame(
                         smoker=rep('0', 1000)
                         , lcholest=x_var
                        )
                        )

f_hat_nosmoke <- predict(semi_mod
                         , newdata=data.frame(
                            smoker=rep('1', 1000)
                            , lcholest=x_var
                          )
                         )
```

```
lineColors <- c("red", "dodgerblue")

plot(fram2$lcholest, fram2$lsbp, type="n"
     , xlab="Cholesterol Variable"
     , ylab = "Blood Pressure Variable"
     , main="Blood Presure by Cholesterol ANCOVA comparison"
       )

lines(x_var, f_hat_smoker, lty=2, lwd=2.5, col=lineColors[1])
lines(x_var, f_hat_nosmoke, lty=3, lwd=2.5, col=lineColors[2])

legend(4.8, 5
       , c("Smoker", "NonSmoker")
       , lty=c(2,3)
       , lwd=rep(2.5, 2)
       , col=c(lineColors[1], lineColors[2])
       )
```



Question 7

Skipping for now

```
# form confidence intervals for the regression lines
conf_int_smoke <- predict(semi_mod, newdata=data.frame(
                        smoker=rep('1', 1000)
                        , lcholest=x_var
                          )
                        , interval="confidence"
                        , level=0.95
                      )
```

```
# the key is the polygon command in hw02
#https://stackoverflow.com/questions/14069629/plotting-confidence-intervals
glimpse(conf_int_smoke)
```

```
##  num [1:1000(1d)] 4.26 4.26 4.26 4.26 4.26 ...
##  - attr(*, "dimnames")=List of 1
##   ..$ : chr [1:1000] "1" "2" "3" "4" ...
```

```
conf_int_smoke
```

```
##        1        2        3        4        5        6        7        8
## 4.256062 4.256298 4.256534 4.256769 4.257005 4.257241 4.257476 4.257712
##        9       10       11       12       13       14       15       16
## 4.257948 4.258183 4.258419 4.258654 4.258890 4.259126 4.259361 4.259597
##       17       18       19       20       21       22       23       24
## 4.259833 4.260068 4.260304 4.260540 4.260775 4.261011 4.261247 4.261482
```

```
##       25       26       27       28       29       30       31       32
## 4.261718 4.261953 4.262189 4.262425 4.262660 4.262896 4.263132 4.263367
##       33       34       35       36       37       38       39       40
## 4.263603 4.263839 4.264074 4.264310 4.264545 4.264781 4.265017 4.265252
##       41       42       43       44       45       46       47       48
## 4.265488 4.265724 4.265959 4.266195 4.266431 4.266666 4.266902 4.267137
##       49       50       51       52       53       54       55       56
## 4.267373 4.267609 4.267844 4.268080 4.268316 4.268551 4.268787 4.269022
##       57       58       59       60       61       62       63       64
## 4.269258 4.269494 4.269729 4.269965 4.270201 4.270436 4.270672 4.270908
##       65       66       67       68       69       70       71       72
## 4.271143 4.271379 4.271614 4.271850 4.272086 4.272321 4.272557 4.272792
##       73       74       75       76       77       78       79       80
## 4.273028 4.273264 4.273499 4.273735 4.273971 4.274206 4.274442 4.274677
##       81       82       83       84       85       86       87       88
## 4.274913 4.275149 4.275384 4.275620 4.275855 4.276091 4.276327 4.276562
##       89       90       91       92       93       94       95       96
## 4.276798 4.277033 4.277269 4.277505 4.277740 4.277976 4.278212 4.278447
##       97       98       99      100      101      102      103      104
## 4.278683 4.278918 4.279154 4.279390 4.279625 4.279861 4.280096 4.280332
##      105      106      107      108      109      110      111      112
## 4.280567 4.280803 4.281039 4.281274 4.281510 4.281745 4.281981 4.282217
##      113      114      115      116      117      118      119      120
## 4.282452 4.282688 4.282923 4.283159 4.283395 4.283630 4.283866 4.284101
##      121      122      123      124      125      126      127      128
## 4.284337 4.284572 4.284808 4.285044 4.285279 4.285515 4.285750 4.285986
##      129      130      131      132      133      134      135      136
## 4.286221 4.286457 4.286693 4.286928 4.287164 4.287399 4.287635 4.287870
##      137      138      139      140      141      142      143      144
## 4.288106 4.288341 4.288577 4.288813 4.289048 4.289284 4.289519 4.289755
##      145      146      147      148      149      150      151      152
## 4.289990 4.290226 4.290461 4.290697 4.290932 4.291168 4.291404 4.291639
##      153      154      155      156      157      158      159      160
## 4.291875 4.292110 4.292346 4.292581 4.292817 4.293052 4.293288 4.293523
##      161      162      163      164      165      166      167      168
## 4.293759 4.293994 4.294230 4.294465 4.294701 4.294936 4.295172 4.295407
##      169      170      171      172      173      174      175      176
## 4.295643 4.295878 4.296114 4.296349 4.296585 4.296821 4.297056 4.297292
##      177      178      179      180      181      182      183      184
## 4.297527 4.297763 4.297998 4.298233 4.298469 4.298704 4.298940 4.299175
##      185      186      187      188      189      190      191      192
## 4.299411 4.299646 4.299882 4.300117 4.300353 4.300588 4.300824 4.301059
##      193      194      195      196      197      198      199      200
## 4.301295 4.301530 4.301766 4.302001 4.302237 4.302472 4.302708 4.302943
##      201      202      203      204      205      206      207      208
## 4.303178 4.303414 4.303649 4.303885 4.304120 4.304356 4.304591 4.304827
##      209      210      211      212      213      214      215      216
## 4.305062 4.305297 4.305533 4.305768 4.306004 4.306239 4.306474 4.306710
##      217      218      219      220      221      222      223      224
## 4.306945 4.307180 4.307416 4.307651 4.307886 4.308122 4.308357 4.308592
##      225      226      227      228      229      230      231      232
## 4.308828 4.309063 4.309298 4.309533 4.309769 4.310004 4.310239 4.310474
##      233      234      235      236      237      238      239      240
## 4.310710 4.310945 4.311180 4.311415 4.311650 4.311885 4.312121 4.312356
```

```
##      241      242      243      244      245      246      247      248
## 4.312591 4.312826 4.313061 4.313296 4.313531 4.313766 4.314001 4.314236
##      249      250      251      252      253      254      255      256
## 4.314471 4.314706 4.314941 4.315176 4.315411 4.315646 4.315881 4.316115
##      257      258      259      260      261      262      263      264
## 4.316350 4.316585 4.316820 4.317055 4.317290 4.317525 4.317759 4.317994
##      265      266      267      268      269      270      271      272
## 4.318229 4.318464 4.318698 4.318933 4.319168 4.319403 4.319637 4.319872
##      273      274      275      276      277      278      279      280
## 4.320107 4.320341 4.320576 4.320811 4.321045 4.321280 4.321514 4.321749
##      281      282      283      284      285      286      287      288
## 4.321983 4.322218 4.322453 4.322687 4.322922 4.323156 4.323391 4.323625
##      289      290      291      292      293      294      295      296
## 4.323860 4.324094 4.324329 4.324563 4.324797 4.325032 4.325266 4.325501
##      297      298      299      300      301      302      303      304
## 4.325735 4.325970 4.326204 4.326438 4.326673 4.326907 4.327141 4.327376
##      305      306      307      308      309      310      311      312
## 4.327610 4.327844 4.328079 4.328313 4.328547 4.328782 4.329016 4.329250
##      313      314      315      316      317      318      319      320
## 4.329484 4.329719 4.329953 4.330187 4.330421 4.330656 4.330890 4.331124
##      321      322      323      324      325      326      327      328
## 4.331358 4.331592 4.331826 4.332061 4.332295 4.332529 4.332763 4.332997
##      329      330      331      332      333      334      335      336
## 4.333231 4.333465 4.333700 4.333934 4.334168 4.334402 4.334636 4.334870
##      337      338      339      340      341      342      343      344
## 4.335104 4.335338 4.335572 4.335806 4.336040 4.336274 4.336508 4.336742
##      345      346      347      348      349      350      351      352
## 4.336976 4.337210 4.337443 4.337677 4.337911 4.338145 4.338379 4.338613
##      353      354      355      356      357      358      359      360
## 4.338846 4.339080 4.339314 4.339548 4.339781 4.340015 4.340249 4.340482
##      361      362      363      364      365      366      367      368
## 4.340716 4.340950 4.341183 4.341417 4.341650 4.341884 4.342117 4.342351
##      369      370      371      372      373      374      375      376
## 4.342584 4.342818 4.343051 4.343284 4.343518 4.343751 4.343984 4.344218
##      377      378      379      380      381      382      383      384
## 4.344451 4.344684 4.344917 4.345150 4.345383 4.345616 4.345849 4.346082
##      385      386      387      388      389      390      391      392
## 4.346315 4.346548 4.346781 4.347014 4.347247 4.347480 4.347713 4.347946
##      393      394      395      396      397      398      399      400
## 4.348178 4.348411 4.348644 4.348876 4.349109 4.349342 4.349574 4.349807
##      401      402      403      404      405      406      407      408
## 4.350039 4.350272 4.350504 4.350736 4.350969 4.351201 4.351433 4.351666
##      409      410      411      412      413      414      415      416
## 4.351898 4.352130 4.352362 4.352594 4.352826 4.353058 4.353290 4.353522
##      417      418      419      420      421      422      423      424
## 4.353754 4.353986 4.354218 4.354450 4.354682 4.354913 4.355145 4.355377
##      425      426      427      428      429      430      431      432
## 4.355608 4.355840 4.356071 4.356303 4.356534 4.356766 4.356997 4.357228
##      433      434      435      436      437      438      439      440
## 4.357460 4.357691 4.357922 4.358153 4.358384 4.358615 4.358846 4.359077
##      441      442      443      444      445      446      447      448
## 4.359308 4.359539 4.359770 4.360001 4.360232 4.360462 4.360693 4.360924
##      449      450      451      452      453      454      455      456
## 4.361154 4.361385 4.361615 4.361845 4.362076 4.362306 4.362536 4.362767
```

```
##      457      458      459      460      461      462      463      464
## 4.362997 4.363227 4.363457 4.363687 4.363917 4.364147 4.364377 4.364607
##      465      466      467      468      469      470      471      472
## 4.364836 4.365066 4.365296 4.365525 4.365755 4.365984 4.366214 4.366443
##      473      474      475      476      477      478      479      480
## 4.366673 4.366902 4.367131 4.367360 4.367589 4.367818 4.368047 4.368276
##      481      482      483      484      485      486      487      488
## 4.368505 4.368734 4.368963 4.369192 4.369420 4.369649 4.369878 4.370106
##      489      490      491      492      493      494      495      496
## 4.370335 4.370563 4.370791 4.371020 4.371248 4.371476 4.371704 4.371932
##      497      498      499      500      501      502      503      504
## 4.372160 4.372388 4.372616 4.372844 4.373072 4.373300 4.373527 4.373755
##      505      506      507      508      509      510      511      512
## 4.373983 4.374210 4.374438 4.374665 4.374893 4.375120 4.375347 4.375575
##      513      514      515      516      517      518      519      520
## 4.375802 4.376029 4.376256 4.376483 4.376711 4.376938 4.377165 4.377392
##      521      522      523      524      525      526      527      528
## 4.377619 4.377845 4.378072 4.378299 4.378526 4.378753 4.378979 4.379206
##      529      530      531      532      533      534      535      536
## 4.379433 4.379659 4.379886 4.380112 4.380339 4.380565 4.380792 4.381018
##      537      538      539      540      541      542      543      544
## 4.381244 4.381471 4.381697 4.381923 4.382150 4.382376 4.382602 4.382828
##      545      546      547      548      549      550      551      552
## 4.383054 4.383280 4.383507 4.383733 4.383959 4.384185 4.384411 4.384637
##      553      554      555      556      557      558      559      560
## 4.384862 4.385088 4.385314 4.385540 4.385766 4.385992 4.386217 4.386443
##      561      562      563      564      565      566      567      568
## 4.386669 4.386895 4.387120 4.387346 4.387572 4.387797 4.388023 4.388248
##      569      570      571      572      573      574      575      576
## 4.388474 4.388700 4.388925 4.389151 4.389376 4.389602 4.389827 4.390053
##      577      578      579      580      581      582      583      584
## 4.390278 4.390503 4.390729 4.390954 4.391180 4.391405 4.391630 4.391856
##      585      586      587      588      589      590      591      592
## 4.392081 4.392306 4.392531 4.392757 4.392982 4.393207 4.393432 4.393658
##      593      594      595      596      597      598      599      600
## 4.393883 4.394108 4.394333 4.394559 4.394784 4.395009 4.395234 4.395459
##      601      602      603      604      605      606      607      608
## 4.395684 4.395909 4.396135 4.396360 4.396585 4.396810 4.397035 4.397260
##      609      610      611      612      613      614      615      616
## 4.397485 4.397710 4.397935 4.398160 4.398385 4.398610 4.398836 4.399061
##      617      618      619      620      621      622      623      624
## 4.399286 4.399511 4.399736 4.399961 4.400186 4.400411 4.400636 4.400861
##      625      626      627      628      629      630      631      632
## 4.401086 4.401311 4.401536 4.401761 4.401985 4.402210 4.402435 4.402660
##      633      634      635      636      637      638      639      640
## 4.402885 4.403110 4.403335 4.403560 4.403785 4.404010 4.404235 4.404460
##      641      642      643      644      645      646      647      648
## 4.404685 4.404910 4.405134 4.405359 4.405584 4.405809 4.406034 4.406259
##      649      650      651      652      653      654      655      656
## 4.406484 4.406708 4.406933 4.407158 4.407383 4.407608 4.407833 4.408058
##      657      658      659      660      661      662      663      664
## 4.408282 4.408507 4.408732 4.408957 4.409182 4.409407 4.409631 4.409856
##      665      666      667      668      669      670      671      672
## 4.410081 4.410306 4.410531 4.410755 4.410980 4.411205 4.411430 4.411655
```

```
##      673      674      675      676      677      678      679      680
## 4.411879 4.412104 4.412329 4.412554 4.412779 4.413003 4.413228 4.413453
##      681      682      683      684      685      686      687      688
## 4.413678 4.413902 4.414127 4.414352 4.414577 4.414801 4.415026 4.415251
##      689      690      691      692      693      694      695      696
## 4.415476 4.415700 4.415925 4.416150 4.416375 4.416599 4.416824 4.417049
##      697      698      699      700      701      702      703      704
## 4.417274 4.417498 4.417723 4.417948 4.418173 4.418397 4.418622 4.418847
##      705      706      707      708      709      710      711      712
## 4.419072 4.419296 4.419521 4.419746 4.419970 4.420195 4.420420 4.420644
##      713      714      715      716      717      718      719      720
## 4.420869 4.421094 4.421319 4.421543 4.421768 4.421993 4.422217 4.422442
##      721      722      723      724      725      726      727      728
## 4.422667 4.422891 4.423116 4.423341 4.423565 4.423790 4.424015 4.424239
##      729      730      731      732      733      734      735      736
## 4.424464 4.424689 4.424913 4.425138 4.425363 4.425587 4.425812 4.426036
##      737      738      739      740      741      742      743      744
## 4.426261 4.426486 4.426710 4.426935 4.427160 4.427384 4.427609 4.427833
##      745      746      747      748      749      750      751      752
## 4.428058 4.428283 4.428507 4.428732 4.428957 4.429181 4.429406 4.429630
##      753      754      755      756      757      758      759      760
## 4.429855 4.430080 4.430304 4.430529 4.430753 4.430978 4.431203 4.431427
##      761      762      763      764      765      766      767      768
## 4.431652 4.431876 4.432101 4.432325 4.432550 4.432775 4.432999 4.433224
##      769      770      771      772      773      774      775      776
## 4.433448 4.433673 4.433897 4.434122 4.434346 4.434571 4.434796 4.435020
##      777      778      779      780      781      782      783      784
## 4.435245 4.435469 4.435694 4.435918 4.436143 4.436367 4.436592 4.436816
##      785      786      787      788      789      790      791      792
## 4.437041 4.437266 4.437490 4.437715 4.437939 4.438164 4.438388 4.438613
##      793      794      795      796      797      798      799      800
## 4.438837 4.439062 4.439286 4.439511 4.439735 4.439960 4.440184 4.440409
##      801      802      803      804      805      806      807      808
## 4.440633 4.440858 4.441082 4.441307 4.441531 4.441756 4.441980 4.442205
##      809      810      811      812      813      814      815      816
## 4.442429 4.442654 4.442878 4.443103 4.443327 4.443552 4.443776 4.444001
##      817      818      819      820      821      822      823      824
## 4.444225 4.444449 4.444674 4.444898 4.445123 4.445347 4.445572 4.445796
##      825      826      827      828      829      830      831      832
## 4.446021 4.446245 4.446470 4.446694 4.446919 4.447143 4.447367 4.447592
##      833      834      835      836      837      838      839      840
## 4.447816 4.448041 4.448265 4.448490 4.448714 4.448939 4.449163 4.449387
##      841      842      843      844      845      846      847      848
## 4.449612 4.449836 4.450061 4.450285 4.450510 4.450734 4.450958 4.451183
##      849      850      851      852      853      854      855      856
## 4.451407 4.451632 4.451856 4.452081 4.452305 4.452529 4.452754 4.452978
##      857      858      859      860      861      862      863      864
## 4.453203 4.453427 4.453651 4.453876 4.454100 4.454325 4.454549 4.454774
##      865      866      867      868      869      870      871      872
## 4.454998 4.455222 4.455447 4.455671 4.455896 4.456120 4.456344 4.456569
##      873      874      875      876      877      878      879      880
## 4.456793 4.457018 4.457242 4.457466 4.457691 4.457915 4.458139 4.458364
##      881      882      883      884      885      886      887      888
## 4.458588 4.458813 4.459037 4.459261 4.459486 4.459710 4.459935 4.460159
```

```
##      889      890      891      892      893      894      895      896
## 4.460383 4.460608 4.460832 4.461056 4.461281 4.461505 4.461730 4.461954
##      897      898      899      900      901      902      903      904
## 4.462178 4.462403 4.462627 4.462851 4.463076 4.463300 4.463524 4.463749
##      905      906      907      908      909      910      911      912
## 4.463973 4.464198 4.464422 4.464646 4.464871 4.465095 4.465319 4.465544
##      913      914      915      916      917      918      919      920
## 4.465768 4.465992 4.466217 4.466441 4.466665 4.466890 4.467114 4.467338
##      921      922      923      924      925      926      927      928
## 4.467563 4.467787 4.468012 4.468236 4.468460 4.468685 4.468909 4.469133
##      929      930      931      932      933      934      935      936
## 4.469358 4.469582 4.469806 4.470031 4.470255 4.470479 4.470704 4.470928
##      937      938      939      940      941      942      943      944
## 4.471152 4.471377 4.471601 4.471825 4.472050 4.472274 4.472498 4.472723
##      945      946      947      948      949      950      951      952
## 4.472947 4.473171 4.473396 4.473620 4.473844 4.474069 4.474293 4.474517
##      953      954      955      956      957      958      959      960
## 4.474742 4.474966 4.475190 4.475415 4.475639 4.475863 4.476088 4.476312
##      961      962      963      964      965      966      967      968
## 4.476536 4.476761 4.476985 4.477209 4.477434 4.477658 4.477882 4.478107
##      969      970      971      972      973      974      975      976
## 4.478331 4.478555 4.478780 4.479004 4.479228 4.479452 4.479677 4.479901
##      977      978      979      980      981      982      983      984
## 4.480125 4.480350 4.480574 4.480798 4.481023 4.481247 4.481471 4.481696
##      985      986      987      988      989      990      991      992
## 4.481920 4.482144 4.482369 4.482593 4.482817 4.483042 4.483266 4.483490
##      993      994      995      996      997      998      999     1000
## 4.483715 4.483939 4.484163 4.484388 4.484612 4.484836 4.485060 4.485285
```

Question 8

Run the semiparametric version of ANCOVA but with an interaction.

```
semi_mod2 <- mgcv::gam(lsbp ~ factor(smoker)*lcholest +
                         s(lcholest, k=23, bs="cr")
                       , data=fram2
#                        , method="REML"
                      )
summary(semi_mod2)
```

```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## lsbp ~ factor(smoker) * lcholest + s(lcholest, k = 23, bs = "cr")
##
## Parametric coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.3664545  0.0218837  16.746   <2e-16 ***
## factor(smoker)1            -0.0330055  0.3891394  -0.085    0.932
## lcholest                    0.7444406  0.0044560 167.063   <2e-16 ***
## factor(smoker)1:lcholest -0.0009226  0.0718517  -0.013    0.990
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##               edf Ref.df     F p-value
## s(lcholest) 1.015  1.091 94.01  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Rank: 25/26
## R-sq.(adj) =  0.0186   Deviance explained = 2.05%
## GCV = 0.044542  Scale est. = 0.044429  n = 1615
```

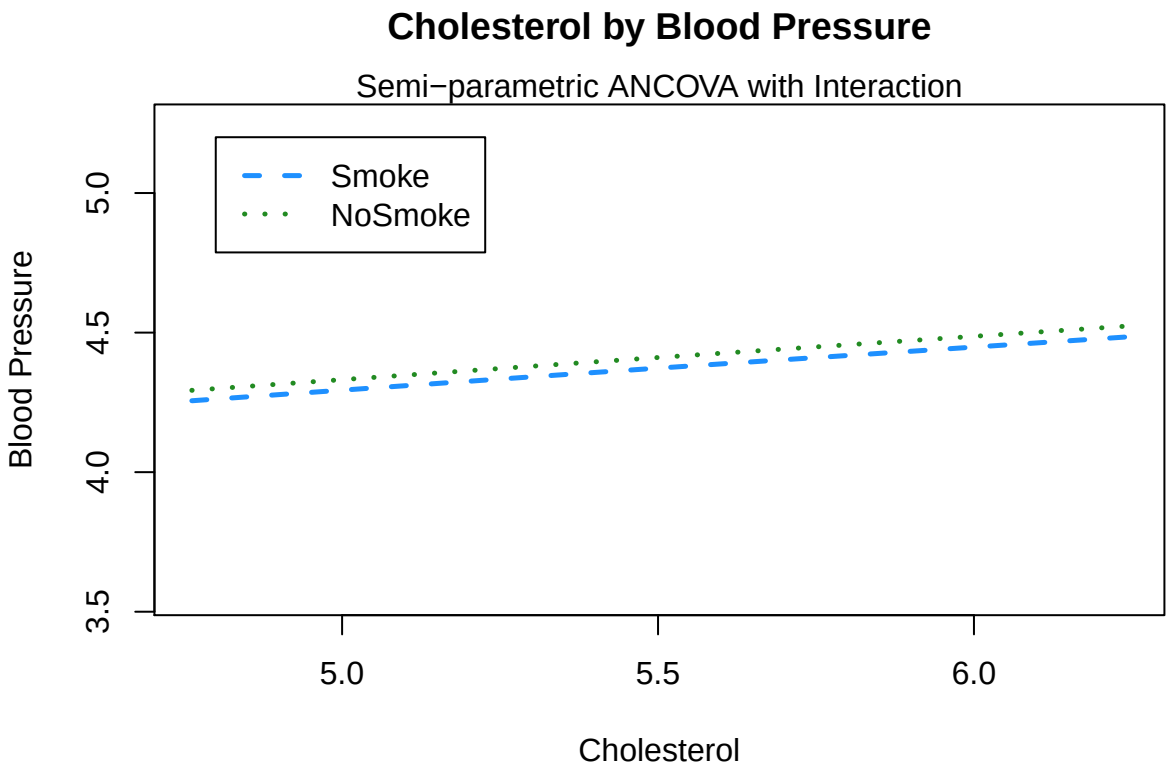
From the above results, it does not appear that there is any sort of interaction present, as the p-value on the interaction term between *smoker* and *lcholest* is not significant.

Question 9

Display the fits of the above regressions, but without the data points.

```
ele_num <- 1000
x_vec <- seq(from=min(fram2$lcholest), to=max(fram2$lcholest), len=ele_num)
fHat_smoke <- predict(semi_mod2, newdata=data.frame(
                        lcholest=x_vec
                        , smoker=rep('1', ele_num)
                         )
                       )

fHat_nosmoke <- predict(semi_mod2, newdata=data.frame(
                        lcholest=x_vec
                        , smoker=rep('0', ele_num)
                         )
                       )
```

```
plot(fram2$lcholest, fram2$lsbp, type='n'
     , xlab="Cholesterol"
     , ylab="Blood Pressure"
     , main="Cholesterol by Blood Pressure"
     )
mtext("Semi-parametric ANCOVA with Interaction")

col_vec <- c("dodgerblue", "forestgreen")
lines(x_vec, fHat_smoke, col=col_vec[1], lwd=2.5, lty=2)
lines(x_vec, fHat_nosmoke, col=col_vec[2], lwd=2.5, lty=3)

legend(4.8, 5.2, c("Smoke", "NoSmoke")
       , col=col_vec
       , lwd=rep(2.5, 2)
       , lty=c(2,3)
      )
```



Question 10

What does the interaction mean in the case when the factors are binary?

When we have a binary factor for our ANCOVA model, this indicates that the interaction term's coefficient is reflecting what happens to our outcome variable for that factor's non-reference class only.

Question 11

Run an analysis of whether our two regression lines are significantly different.

```
# First, fit the null model
contrast_mod1 <- mgcv::gam(lsbp ~ s(lcholest), data=fram2)

# indicator of the smoke varible taking positive class
smoke_ind <- as.numeric(fram2$smoker==1)
# now, fit the alternative model
contrast_mod2 <- mgcv::gam(lsbp ~ s(lcholest, smoke_ind), data=fram2)

anova(contrast_mod1, contrast_mod2, test="F")
```

```
## Analysis of Deviance Table
##
## Model 1: lsbp ~ s(lcholest)
## Model 2: lsbp ~ s(lcholest, smoke_ind)
##   Resid. Df Resid. Dev Df Deviance      F   Pr(>F)
## 1      1613     71.988
```

```
## 2      1612     71.579  1  0.40861 9.2021 0.002456 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

It appears that there is a significant difference between the two fits, for smoker and non-smoker.

Question 12

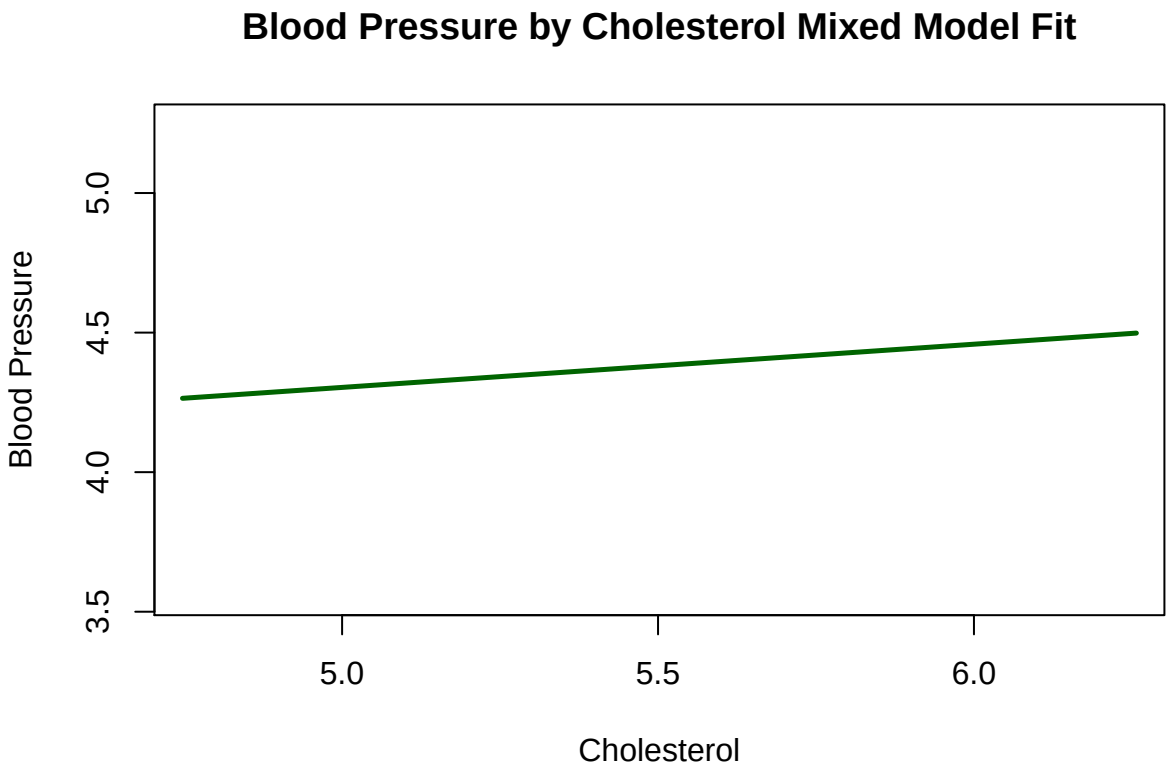
Run native code using lme to fit LSBP vs. age.

```
x <- fram2$lcholest
y <- fram2$lsbp
numIntKnots <- 23
intKnots <- quantile(unique(x), seq(0, 1, length=numIntKnots+2))[-c(1, numIntKnots+2)]
a <- 1.01 * min(x) - 0.01*max(x)
b <- 1.01 * max(x) - 0.01*min(x)
Z <- HRW::ZOSull(x, range.x=c(a, b), intKnots=intKnots)
dummyID <- factor(rep(1, length(x)))
mm_fit <- nlme::lme(y ~ x, random=list(dummyID=pdIdent(~ -1+Z)))
```

Question 13

Display the fit without data points. Also provide a confidence band.

```
# extract our model coefficient
betaHat <- mm_fit$coefficients$fixed
uHat <- unlist(mm_fit$coefficients$random)
ng <- 1001
xg <- seq(a, b, length.out=ng)
Xg <- cbind(rep(1, ng), xg)
Zg <- HRW::ZOSull(xg, range.x=c(a,b), intKnots=intKnots)
fHatg <- as.vector(Xg %*% betaHat + Zg %*% uHat)
plot(x, y, type='n'
     , xlab="Cholesterol"
     , ylab="Blood Pressure"
     , main='Blood Pressure by Cholesterol Mixed Model Fit'
     )

lines(xg, fHatg, col="darkgreen", lwd=2.5)
```



Question 14

Is the fit of our model statistically significant?

```
anova(mm_fit)
```

```
##             numDF denDF  F-value p-value
## (Intercept)     1  1613 690343.1  <.0001
## x               1  1613     24.2  <.0001
```

The F-test indicates that the fit of our model is significant.

Question 15

Test whether the fit is linear or quadratic vs. the need to do a semiparametric fit.

```
mm_fit2 <- nlme::lme(y ~ x + x**2, random=list(dummyID=pdIdent(~ -1+Z)))
summary(mm_fit2)
```

```
## Linear mixed-effects model fit by REML
##  Data: NULL
##         AIC       BIC   logLik
##   -418.6971 -397.1536 213.3485
##
## Random effects:
##  Formula: ~-1 + Z | dummyID
##  Structure: Multiple of an Identity
##                   Z1           Z2           Z3           Z4           Z5
```

```
## StdDev: 0.0002778356 0.0002778356 0.0002778356 0.0002778356 0.0002778356
##                   Z6           Z7           Z8           Z9          Z10
## StdDev: 0.0002778356 0.0002778356 0.0002778356 0.0002778356 0.0002778356
##                  Z11          Z12          Z13          Z14          Z15
## StdDev: 0.0002778356 0.0002778356 0.0002778356 0.0002778356 0.0002778356
##                  Z16          Z17          Z18          Z19          Z20
## StdDev: 0.0002778356 0.0002778356 0.0002778356 0.0002778356 0.0002778356
##                  Z21          Z22          Z23          Z24          Z25
## StdDev: 0.0002778356 0.0002778356 0.0002778356 0.0002778356 0.0002778356
##          Residual
## StdDev: 0.2112576
##
## Fixed effects: y ~ x + x^2
##                Value  Std.Error   DF   t-value p-value
## (Intercept) 3.529742 0.17050870 1613 20.701240       0
## x           0.154779 0.03147687 1613  4.917213       0
##  Correlation:
##   (Intr)
## x -1
##
## Standardized Within-Group Residuals:
##        Min         Q1        Med         Q3        Max
## -3.7875892 -0.6517280 -0.1098031  0.5482766  4.3742402
##
## Number of Observations: 1615
## Number of Groups: 1
```