Data Processing for Low Carb High Fat Diet

Repeated measures on subjects

Mark Connolly

2016-08-28

# Introduction

Analysis of stuff.

# Data Preparation

Set up for combining several spreadsheets into a single data frame. The sheets must all have the same structure and the same column names.

excelfilter <- matrix(c("Excel", "\*.xls\*"), ncol=2)  
ingestvo2 <- function(afilename) {  
 parts <- unlist(strsplit(basename(afilename)," "))  
 columns <-   
 data.frame("subj" = as.numeric(parts[1]),  
 "testnumber" = as.numeric(parts[2]),  
 "testdate" = as.Date(parts[3],format="%m%d%y")  
 )   
 obsvo2 <-   
 read\_excel(afilename, sheet=1, skip=4, col\_names=FALSE)  
 names(obsvo2) <-   
 trim(names(read\_excel(afilename, sheet=1, skip=2, col\_names=TRUE)))  
 obsvo2$interval <- 0:(nrow(obsvo2) - 1)  
 return(cbind(columns,obsvo2))  
}

Create a list that contains user-selected files, with each file processed by the ingestvo2 function. The data frames in the list are then bound together by row.

# user selects source files (all with the same structure)  
list\_of\_files <-   
 choose.files(filters = excelfilter,  
 caption = "Select VO2 files to combine")  
# each file is ingested to create a data frame  
# each data frame is added to a list   
list\_of\_data\_frames <-   
 lapply(list\_of\_files,   
 FUN=ingestvo2)  
# the data frames in the list are bound together by rows  
vo2data <- do.call('rbind', list\_of\_data\_frames)

Make the subject a factor.

vo2data$subj = as.factor(vo2data$subj)

## Crossover Plot

The subject heartrate when percent kcal / min from fat equals the percent kcal / min from carbohydrates.

Define the subject and the test number that will be plotted.

testnumber = 1  
subject = "123456"

Subset the data, selecting for test number and subject and for only the needed columns. The data from the subset are persisted as the dataframe fuel.

fuel <- subset(vo2data, subj==subject & testnumber == testnumber)[,c("%Fat","%Carbs","HR", "Rel.VO2")]

Create a "long" form of the data to be used for the plot. Functions head and tail are used to convey the structure changes to the data frame. The long form is used for the plot.

fuel\_long <- gather(fuel, # the source data frame  
 percentage, # the column name of the key  
 `%KCal / min`, # the column name of the value  
 c(`%Fat`,`%Carbs`)) # the columns that are the source of the key name

## Warning: failed to assign NativeSymbolInfo for env since env is already  
## defined in the 'lazyeval' namespace

# and the source of the value.  
head(fuel\_long) # shows the top part of the data frame

## HR Rel.VO2 percentage %KCal / min  
## 1 76 0.78 %Fat 72.4  
## 2 75 8.95 %Fat 74.5  
## 3 74 8.52 %Fat 75.1  
## 4 72 13.76 %Fat 87.7  
## 5 70 9.74 %Fat 74.6  
## 6 71 12.55 %Fat 75.8

tail(fuel\_long) # shows the bottom part of the data frame

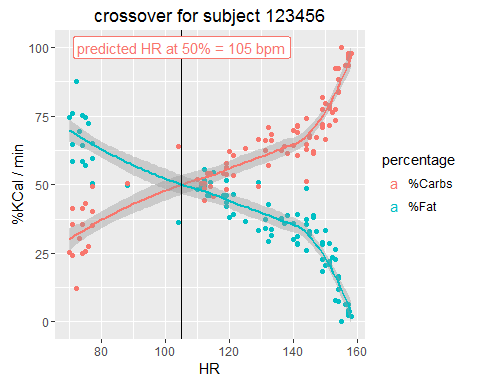
## HR Rel.VO2 percentage %KCal / min  
## 169 157 57.77 %Carbs 96.0  
## 170 157 59.64 %Carbs 96.6  
## 171 157 60.05 %Carbs 96.0  
## 172 157 60.26 %Carbs 97.8  
## 173 155 57.15 %Carbs 100.0  
## 174 158 59.80 %Carbs 97.9

Find the heartrate where the lines cross. The lines cross when the %fat and %carbohydrates are each 50. A loess(R Core Team 2016) function fits a model to predict %Fat for the heartrate. Then the predicted %Fat value that is closest to 50 is used to index the heartrate value. The heartrate at 50% is used to mark the crossover point on the graph.

# fuel data frame is ordered by HR and the result  
# is used to construct polynomial regression fitting of %Fat ~ HR  
fatfromhrloess <-  
 loess(data=fuel[order(fuel$HR),], `%Fat` ~ HR)  
# A data frame is constructed for predicting %Fat for sequenc of heartrate values.  
# The sequence is from the lowest observed value to the hightest observed value,   
# incremented by 1.  
predictedfat <-   
 data.frame(HR=c(min(fuel$HR):max(fuel$HR)))  
# Predicted values of %Fat (fat) are added to the data frame. The values are predicted  
# using the loess fit model and the sequence of heartrates.  
predictedfat$fat <-   
 predict(loess(data=fuel,`%Fat` ~ HR),   
 newdata = predictedfat)  
# The predicted value of fat is used to find the closest value to 50%. This value is used to  
# index the heartrate at that value. This is the imputed heartrate at 50%.  
crossover <-   
 predictedfat$HR[which.min(abs(predictedfat$fat - 50))]

Now plot the crossover with a label.

ggplot(data=fuel\_long,aes(x=HR,y=`%KCal / min`,colour=percentage)) +   
 geom\_point() +   
 geom\_smooth() +  
 geom\_vline(xintercept = crossover) +  
 ggtitle(paste("crossover for subject",subject)) +   
 geom\_vline(xintercept = crossover) +  
 geom\_label(x=crossover, y=100, label=paste("predicted HR at 50% =",crossover,"bpm"))



Rel.VO2 ~ Heartrate

# References

R Core Team. 2016. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.