Steps to Run a Complete FDA with a Continuous Predictor

Carlos Gonzalez, Berkley Shands, Elena Deych, William Shannon

January 15, 2016

Below we will define the steps required to complete an FDA with our data

1. Load the **Actigraphy** package:

```
>        library(Actigraphy)
```

2. Read in the activity and covariate data and store them as the datasets covariate and activity:

```
>        data(clinic_29pt_ahi)
>        data(act_29pt)
>        covariate <- clinic_29pt_ahi
>        activity <- act_29pt
```

3. Remove any NAs in the covariate file and fix the column names in our activity file so they match the covariate dataset (We remove the first column of activity because it is an unnecessary time variable):

```
>        covariate <- na.omit(covariate)
>        activity <- as.matrix(activity[,-1])
>        colnames(activity) <- sub("X", "", colnames(activity))
```

4. Match the data in the activity and covariate datasets:

```
>        matchid <- fda.matchid(activity, covariate, "contin")
```

5. Smooth the data and plot it:

- L is the number of entries in the activity file

```
>        L <- nrow(activity)
>        FDinterest <- fda.smoothdata(matchid)
>        ts.plot(predict(FDinterest$fd$fd, c(1:L)), main="Smoothed Activity Data")
```

6. Apply FLM to the data with the **flm_cate** function:

```
>        geftFDcont <- flm_cate(FDinterest)
```

7. Set up the plot parameters for the next step:

- xlim is a vector of length 2 that contains the X-axis boundries based on the time span of the activity data

- ylim is a vector of length 2 that contains the Y-axis boundries based on the results on the `flm_cate` function

- lb is a vector of labels for the X-axis

- xat is a vector of positions for the labels (lb) on the X-axis

- legendx is the x-axis position of the left edge of the legend box

- legendy is the y-axis position of the upper edge of the legend box

```
>           predy <- as.vector(geftFDcont$freg$yhatfdobj$y)
>           xlim <- c(0, L)
>           ylim <- c(min(predy), max(predy) + 100)
>           lb <- c("Midnight", "6AM", "Noon", "6PM", "Midnight")
>           xat <- c(0, L/4, L/2, 3*L/4, L)
>           legendx <- 0
>           legendy <- max(predy) - 100
```

8. Plot results from FLM results and F-test if indicated (with the continuous flm plotting function):

```
>           cont.flm.results <- cont_flm_plot(FDinterest, matchid, geftFDcont, xlim, ylim, T
```

Notes:

- Columns in Activity file data MUST represent subjects

- Covariate file MUST only contain 2 columns; subject identifier (id) and one covariate (numeric or factor). If user is interested in 2 or more covariates, they should be put into separate datasets and analyzed separately.

- Subject identifier should be the same in both Activity (first column) and Covariate (row names) files