

Note On Updated FUI Code

Edward Gunning

2023-07-23

Note on FUI code used in comparison

In going through the FUI code supplied as supplementary material to the FUI paper here, we noticed an irregularity with how the function `lfosr3s()` was labelling subjects during the bootstrap. As illustrated in Figure 1, each time a given subject appears in a bootstrap sample, their original subject ID can be used for model fitting (Option 1) or they can be assigned a new pseudo-ID for model fitting (Option 2). Currently, `lfosr3s()` uses Option 1, but we noted that Option 2 is more appropriate because it preserves the number of clusters and distribution of cluster sizes. This is also the convention we use in our implementation of the bootstrap of subjects. Therefore, we have provided an updated version, `lfosr3s_updated()`, which implements option 2 for the comparison in our manuscript.

Gunning/RISC1-fda-manuscript-01-code/code/functions/FUI-functions/bootstrap-of-subjects.pdf

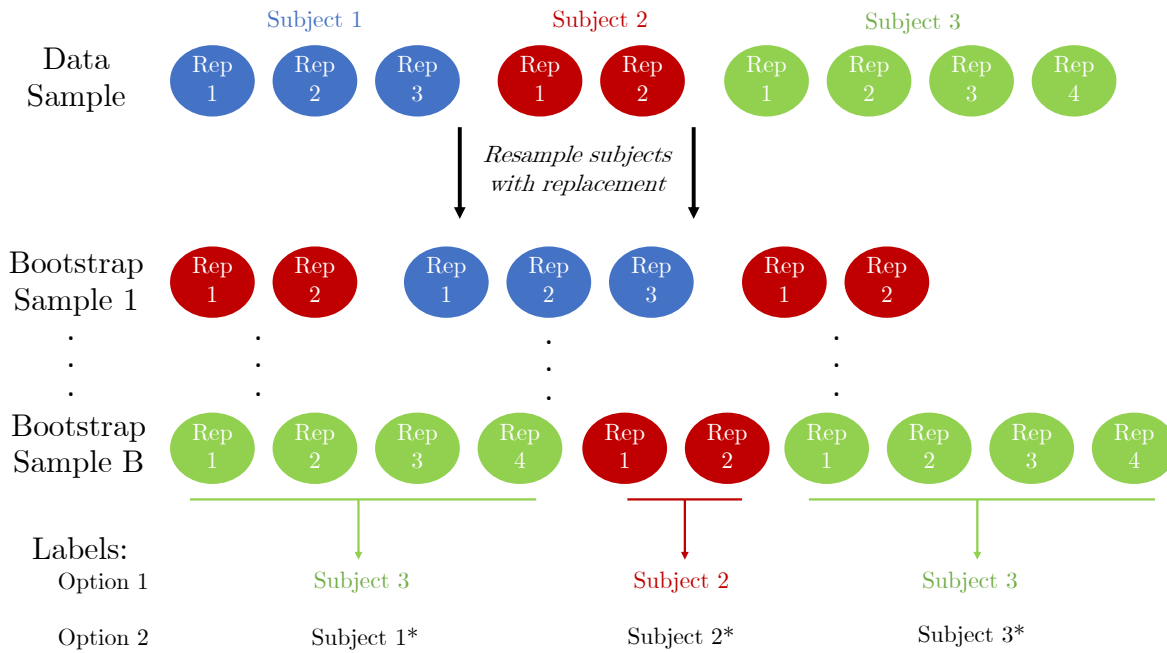


Figure 1: Schematic of bootstrap of subjects resampling.

This shows the snippet of code for the bootstrap in `lfosr3s()`:

```
for(boots in 1:B){  
  pb$tick()  
  sample.ind <- sample(1:length(ID.number), size = length(ID.number), replace = TRUE)  
  dat.ind <- c()
```

```

for(i in 1:length(ID.number)){
  dat.ind <- c(dat.ind, which(data[,group] == ID.number[sample.ind[i]])) ## subject-level bootstrap
}
fit_boot <- lfossr3s(formula = formula, data = data[dat.ind,], family = family, var = FALSE,
                    parallel = parallel, silent = TRUE)
betaHat_boot[, ,boots] <- fit_boot$betaHat
}

```

Here is the corresponding code in `lfossr3s_updated()`:

```

for(boots in 1:B){
  pb$tick()
  sample.ind <- sample(1:length(ID.number), size = length(ID.number), replace = TRUE)
  # previous way of resampling:
  # dat.ind <- c()
  # for(i in 1:length(ID.number)){
  #   dat.ind <- c(dat.ind, which(data[,group] == ID.number[sample.ind[i]])) ## subject-level bootstrap
  # }
  df2 <- data.frame()
  for(i in 1:length(sample.ind)){
    df2 <- rbind(df2, cbind(data[data[,group] == ID.number[sample.ind[i]], ], subj_b = i)) # subj_b is
  }
  names(df2)[names(df2) == group] <- "old_group"
  names(df2)[names(df2) == "subj_b"] <- group
  stopifnot(length(unique(df2[, group])) == length(ID.number)) # check that I've done this labelling co
  fit_boot <- lfossr3s(formula = formula, data = df2, family = family, var = FALSE,
                      parallel = parallel, silent = TRUE)
  betaHat_boot[, ,boots] <- fit_boot$betaHat
}

```

We have contacted the authors directly regarding this update which they have said will be implemented in the next version of the software. For the comparisons in this work we used `lfossr3s_updated()`.