Smoking_Data

Jason Ma

October 16, 2018

```
# read in the data

puff_probability = read.csv("puff-probability.csv", header=TRUE)
puff_episode = read.csv("puff-episode.csv", header=TRUE)
random_ema = read.csv("random-ema.csv", header=TRUE)
eod_ema = read.csv("eod-ema.csv", header=TRUE)
event_ema = read.csv("eventcontingent-ema.csv", header=TRUE)

puff_probability_alt = read.csv("puff-probability-alternative.csv", header=TRUE)
puff_episode_alt = read.csv("puff-episode-alternative.csv", header=TRUE)
random_ema_alt = read.csv("random-ema-alternative.csv", header=TRUE)
eod_ema_alt = read.csv("eod-ema-alternative.csv", header=TRUE)
event_ema_alt = read.csv("eventcontingent-ema-alternative.csv", header=TRUE)
```

```
# fix some bug in the data before running summaries
puff_probability  <- puff_probability[puff_probability$event != 224, ]
puff_episode <- puff_episode[puff_episode$event != 224, ]
```

puff_probability

```
summary(puff_probability$event)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.3069  0.4747  0.6936  0.6790  0.8906  1.0000
```

```
summary(puff_probability_alt$event)
```

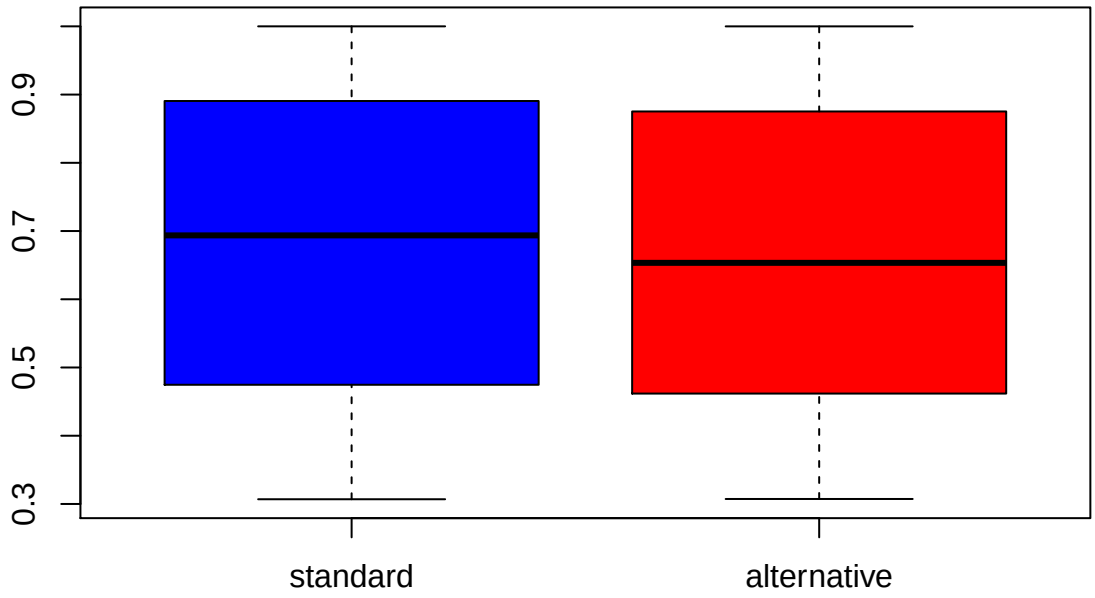
```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.3073  0.4618  0.6534  0.6635  0.8753  1.0000
```

```
t.test(puff_probability$event-puff_probability_alt$event)
```

```
## Warning in puff_probability$event - puff_probability_alt$event: longer
## object length is not a multiple of shorter object length
```

```
##
##  One Sample t-test
##
## data:  puff_probability$event - puff_probability_alt$event
## t = 5.0717, df = 10616, p-value = 4.009e-07
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  0.009456533 0.021371279
## sample estimates:
##  mean of x
## 0.01541391
```

```
boxplot(puff_probability$event, puff_probability_alt$event, names=c('standard', 'alternative'), col=c("
```



The puff_probability data seem ok and consistent across the two files.

puff_episode

random_ema

How many random EMAs on average?

```
summary(random_ema$status)
```

```
## ABANDONED_BY_TIMEOUT    ABANDONED_BY_USER            COMPLETED
##                    1                    3                  529
##               MISSED
##                  158
```

```
summary(random_ema_alt$status)
```

```
## ABANDONED_BY_USER         COMPLETED            MISSED
##                 1               154                45
```

```
at_most_three <- function(ema) {
    print("The following participants have random EMAs of more than 3 times on some days")
    count = 1
    max_count = 1
    for (i in 2:nrow(ema)) {
        current = ema[i,]
        prev = ema[i-1, ]
        if (current$participant_id == prev$participant_id) {
            if ((current$day_of_week == prev$day_of_week)) {
```

```
                count = count + 1
                max_count = max(max_count, count)
            }
            else {
                count = 1
            }
        }
        else {
            if (max_count > 3) {
                print(prev$participant_id)
            }
            count = 1
            max_count = 1
        }
    }
}

at_most_three(random_ema)
```

```
## [1] "The following participants have random EMAs of more than 3 times on some days"
## [1] 202
## [1] 211
## [1] 212
## [1] 218
## [1] 231
## [1] 233
```

```
at_most_three(random_ema_alt)
```

```
## [1] "The following participants have random EMAs of more than 3 times on some days"
## [1] 223
## [1] 224
## [1] 225
## [1] 226
```

eod_ema