

ASA DataFest 2019

In our project, we first tried to examine whether the current overall measurement of fatigue, “Monitoring Score” can be improved or not. To do this, we noticed that if “Monitoring Score” is indeed a good overall measure of fatigue, then for groups of players who believe they were performing better than usual in the current game, we should expect their “Monitoring Score” are on average larger than that of groups of players who think they didnot perform as well as usual in the games.

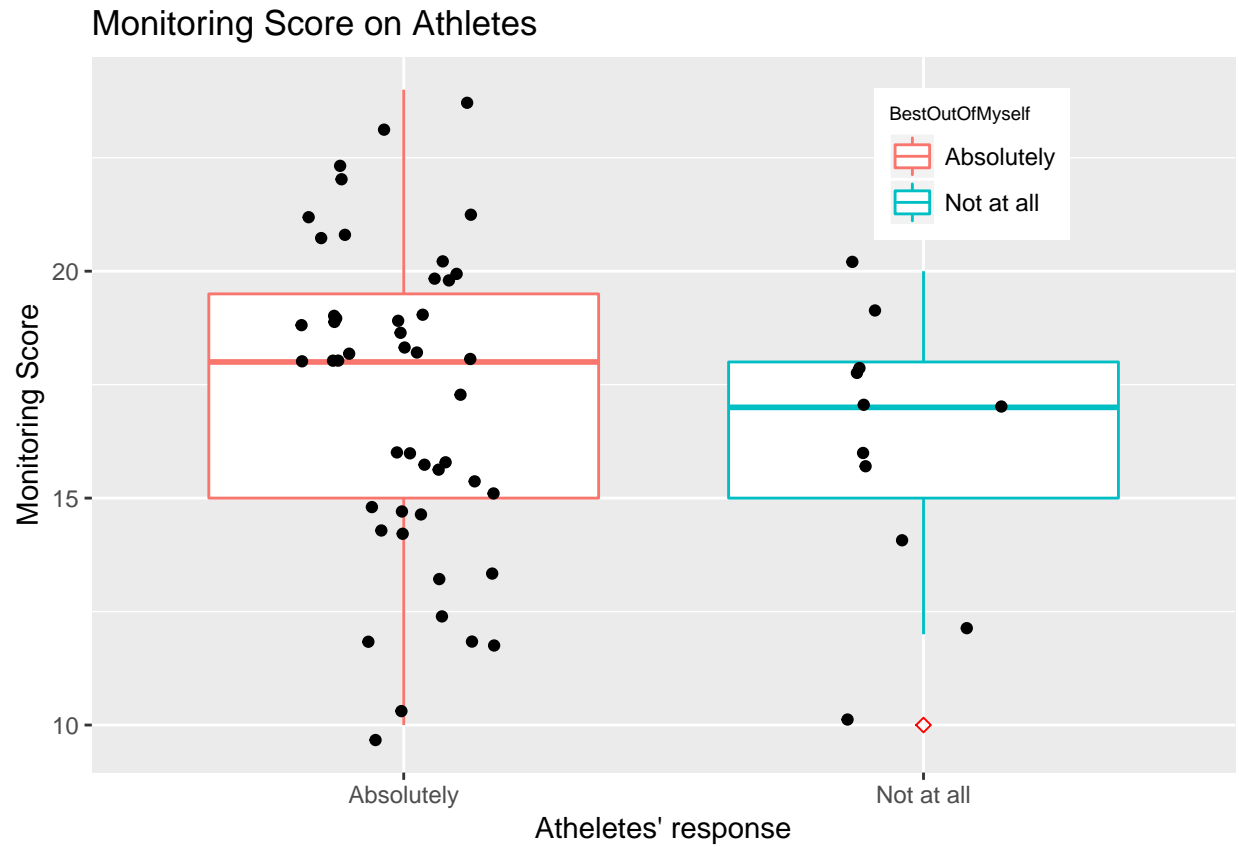
We can use two sample t-test as follows:

```
##
## F test to compare two variances
##
## data: ANOVA$MonitoringScore[ANOVA$BestOutOfMyself == "Not at all"] and ANOVA$MonitoringScore[ANOVA$
## F = 0.77258, num df = 10, denom df = 46, p-value = 0.692
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
## 0.3299648 2.4979691
## sample estimates:
## ratio of variances
## 0.7725843
```

Since the two groups have variances not significantly different from each other, we can use pooled variance two sample t-tests.

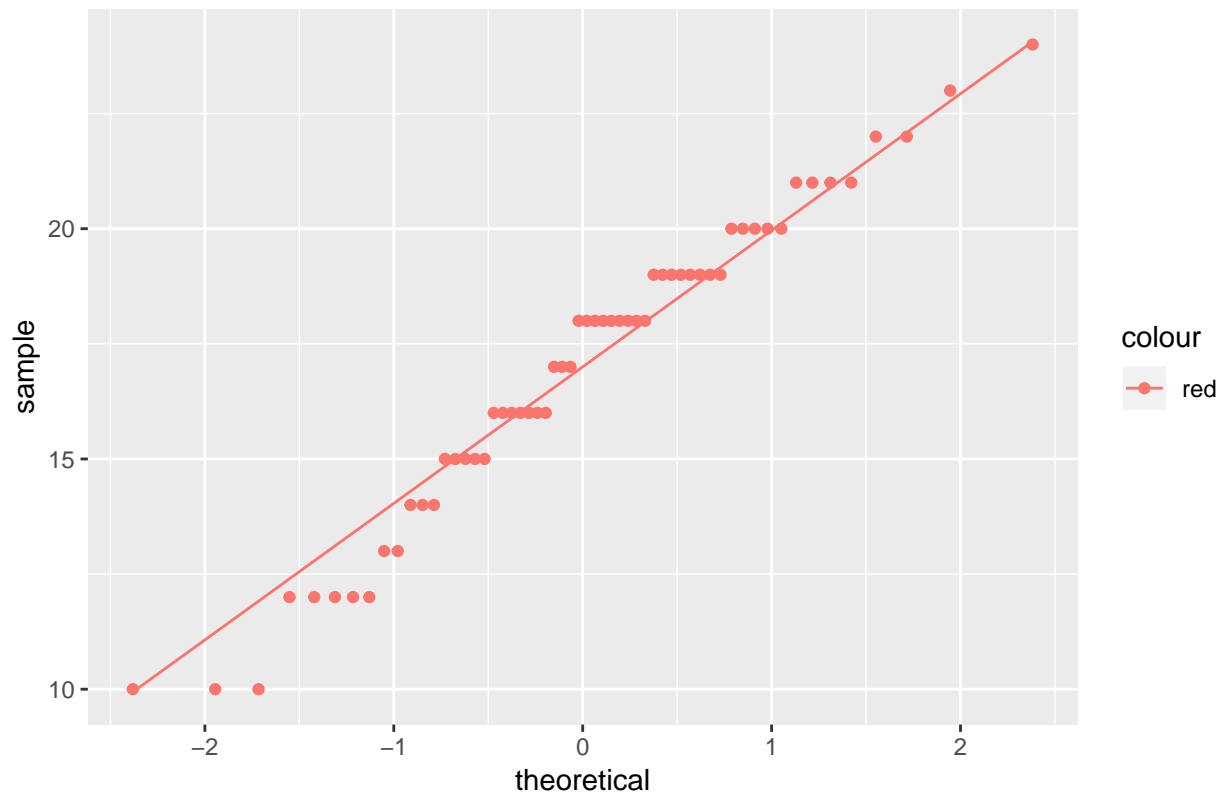
```
t.test(ANOVA$MonitoringScore[ANOVA$BestOutOfMyself=="Not at all"],
       ANOVA$MonitoringScore[ANOVA$BestOutOfMyself=="Absolutely"], var.equal = T)
```

```
##
## Two Sample t-test
##
## data: ANOVA$MonitoringScore[ANOVA$BestOutOfMyself == "Not at all"] and ANOVA$MonitoringScore[ANOVA$
## t = -0.97797, df = 56, p-value = 0.3323
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -3.354976 1.153816
## sample estimates:
## mean of x mean of y
## 16.09091 17.19149
```



We can then check the assumptions of the two sample t-tests.

Monotoring Score Normality



It seems that normality assumption is satisfied.

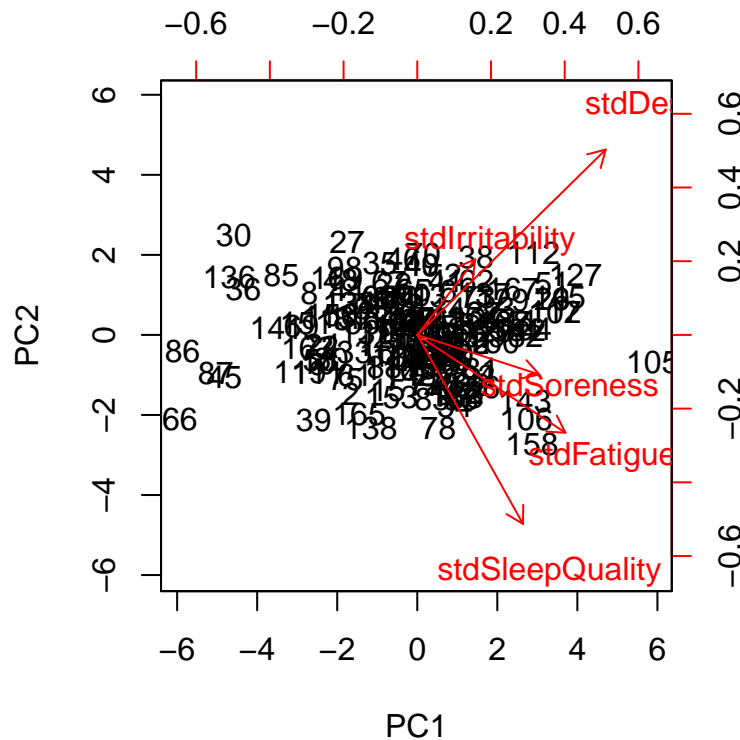
However, based on the boxplots, these two groups seem to have similar “Monitoring Scores” instead. Does that mean these five subjective measures of fatigues are totally useless? Maybe not! Since we know that even though those measures are subjective and may even be biased in some ways, they are still very good proxy to the real physical fatigue levels of players. Therefore, we just need to find a more efficient way to use this information.

To do that, we decided to apply PCA on these five standardized subjective features, and the reason that we standardized these features is to take in to account the individual’s variations. Using R, we found all of the principal component vectors, and we noticed that the first principal component vector alone explains the most of variability among these features.

```
X <- select(PlayerinGame,"stdFatigue","stdSoreness","stdDesire","stdIrritability","stdSleepQuality")
PCA <- prcomp(X)
summary(PCA)
```

```
## Importance of components:
##              PC1    PC2    PC3    PC4    PC5
## Standard deviation  1.9446 1.0402 0.8810 0.72307 0.63539
## Proportion of Variance 0.5759 0.1648 0.1182 0.07962 0.06148
## Cumulative Proportion 0.5759 0.7407 0.8589 0.93852 1.00000
```

```
biplot(PCA$x,PCA$rotation[,1:2])
```



```
X1 <- mutate(X, Subjective_Score = as.matrix(X) %*% PCA$rotation[,1])
X1$SleepHours <- PlayerinGame$SleepHours
PlayerinGame$Sscore <- X1$Subjective_Score
Weights_for_Subjective <- PCA$rotation[,1]
```

Therefore, we will assign the components of first loading vector as weights to these five features and compute a weighted sum of them, and that weighted sum will be our new measure of fatigue called “Sscore”. Now, we claim that this “Sscore” computed using PCA, will be much better than “Monitoring Scores”, since first principal component vector by definition is the best direction in terms of efficiently utilizing all the information in the feature space. To confirm that, we can do a pooled t-test between Scores of players with higher self-rating and Sscores of players with lower self-rating, since all the diagnostics such as normal qqplots and F test of variance, showed that our pooled t-test is valid. Consequently, we find that now we have strong evidence to say that these two groups are on average with different Sscores which we failed to conclude when we used “Monitoring Scores”. This result strongly suggests that we should think about using our “Sscore” to measure the overall fatigue instead of the original Monitoring Scores.

Now, we claim that this “Sscore” computed using PCA, will be much better than “Monitoring Scores”, since first principal component vector by definition is the best direction in terms of efficiently utilizing all the information in the feature space. To confirm that, we can do a pooled t-test between Sscores of players with higher self-rating and Sscores of players with lower self-rating.

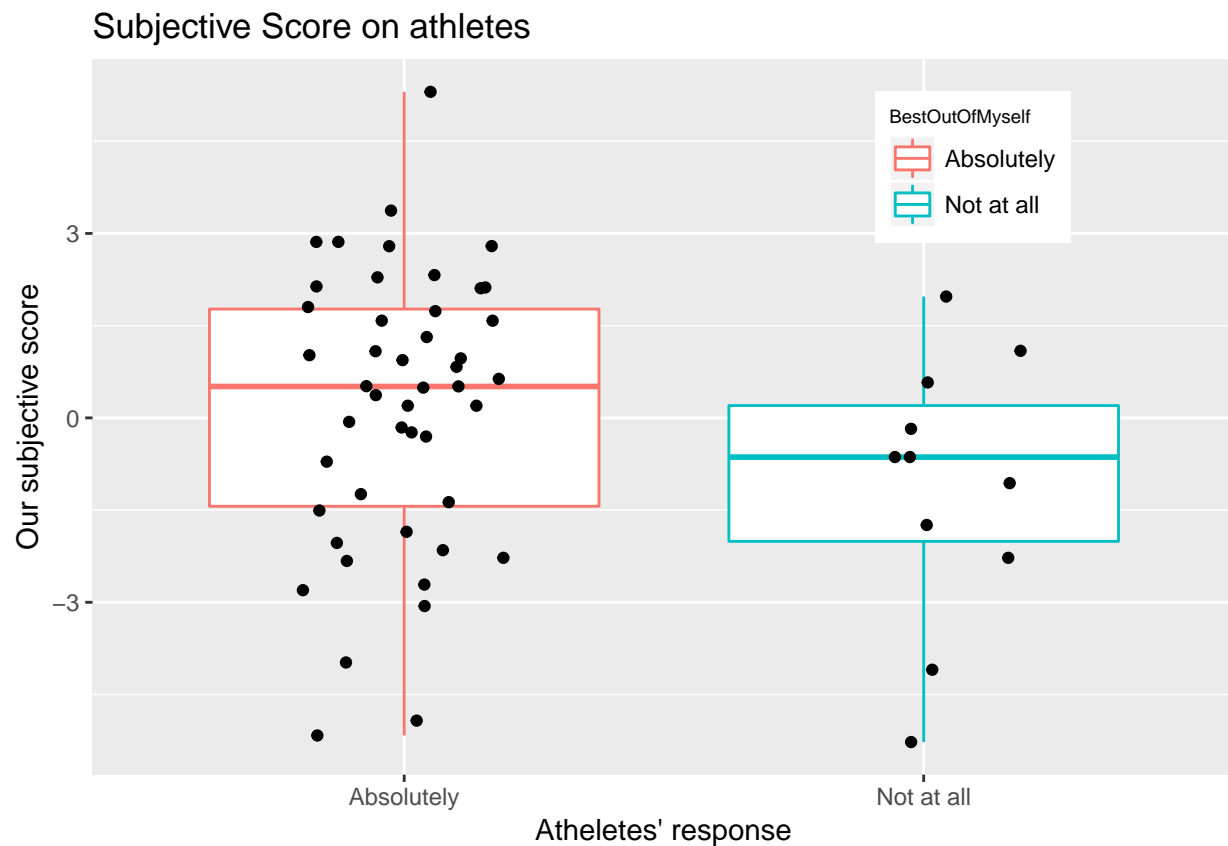
```
##
## F test to compare two variances
##
## data: ANOVA$Sscore[ANOVA$BestOutOfMyself == "Not at all"] and ANOVA$Sscore[ANOVA$BestOutOfMyself ==
## F = 0.91529, num df = 10, denom df = 46, p-value = 0.9449
```

```
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.3909133 2.9593745
## sample estimates:
## ratio of variances
##          0.9152901
```

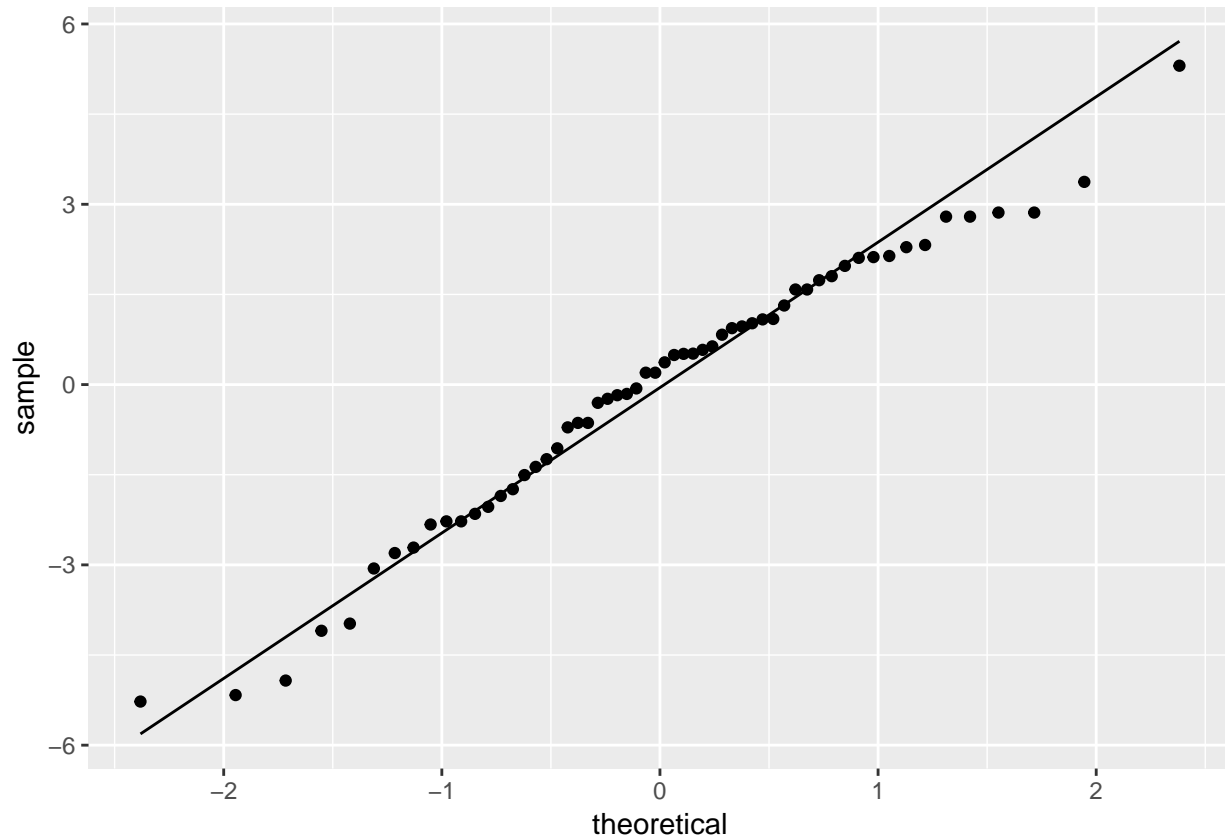
This shows that we can use pooled two sample t-tests,

```
t.test(ANOVA$Sscore[ANOVA$BestOutOfMyself=="Not at all"],
       ANOVA$Sscore[ANOVA$BestOutOfMyself=="Absolutely"],var.equal = T)
```

```
##
## Two Sample t-test
##
## data: ANOVA$Sscore[ANOVA$BestOutOfMyself == "Not at all"] and ANOVA$Sscore[ANOVA$BestOutOfMyself ==
## t = -1.7097, df = 56, p-value = 0.09286
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -2.7834244  0.2200434
## sample estimates:
## mean of x mean of y
## -1.1142097  0.1674807
```



Now we can check the normality assumption,



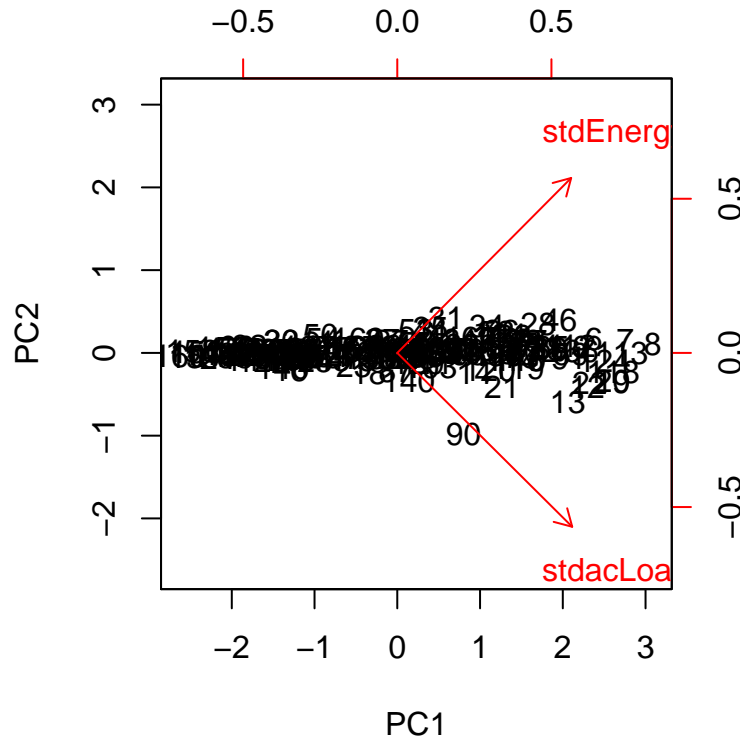
This shows that under $\alpha = 0.10$, athletes who feel absolutely best out of themselves and athletes who feel not at all best out of themselves have significantly different Sscore. This result strongly suggests that we should think about using our “Score” to measure the overall fatigue instead of the original Monitoring Scores.

Furthermore, with our new overall measurement of player’s fatigue, some more intriguing relationship between fatigues and performance can be found and utilized to form useful auxiliary tools for the coaches to refer to during the games. We constructed a new variable to measure the use of overall energy a player used during every half game called “Oscore”. Again, it is calculated using weighted sum of objective measurements from gps data, and the weights are given by the first principal component.

```
X2 <- select(PlayerinGame, "stdEnergy", "stdacLoad")
PCA2 <- prcomp(X2)
summary(PCA2)
```

```
## Importance of components:
##              PC1      PC2
## Standard deviation    1.4014 0.16228
## Proportion of Variance 0.9868 0.01323
## Cumulative Proportion 0.9868 1.00000
```

```
biplot(PCA2$x, PCA2$rotation[,1:2])
```



```
X3 <- mutate(X2, Objective_Score = as.matrix(X2) %*% PCA2$rotation[,1])
X1$SleepHours <- PlayerinGame$SleepHours
PlayerinGame$Oscore <- X3$Objective_Score
Weights_for_Objective <- PCA2$rotation[,1]
```

It is important to understand the intuitions of the two measures that we gave: “Sscore” represents self-perceived mental status at the beginning of a day, the higher it is, the more positive the player perceives herself to be, and “Oscore” represents physical energy consumed at the end of a half game, the higher it is, the more fatigue the player should feel. How are those quantitative measures useful in an actual game?

Finally, we can apply the two measures in a real-world scenario; We use SVM classifier to classify that will there be a significant drop of performance for a player in the 2nd half of the game given the two scores, this prediction could be a very useful reference when a coach is trying to decide whether to substitute a player during half time. The measure of performance is important here, and we use average speed of a player in a half game for that, the intuition is that, if a player is performing well, she shouldn’t remain idle or a low speed for too long. Now You can see our predictions in the plot; the classifier predicts a significant drop of performance if the player has low “Sscore” from the very beginning or starts with a high “Sscore” but also has a high “Oscore” at the first half, this is not perfect, but falls in line with our understanding of those two measures.