Table 1

Matthew Rhodes

Wednesday, February 19, 2020

```r
data = read.csv("C://Users//Matth//OneDrive//Desktop//electionday_tweets.csv")
true_news = data[data$is_fake_news == FALSE,]
fake_news = data[data$is_fake_news == TRUE,]

columns = c(10, 9, 16, 11, 15, 17, 5, 14)
differences = c()
p_vals = c()
for (col in columns){
  paste(col)
  sample1 = true_news[,col]
  sample2 = fake_news[,col]
  dup <- ks.test(sample2, sample1, , alternative="g")
  differences <- c(differences, dup$statistic)
  p_vals <- c(p_vals, dup$p.value)
}
```

```
## Warning in ks.test(sample2, sample1, , alternative = "g"): p-value will be
## approximate in the presence of ties

## Warning in ks.test(sample2, sample1, , alternative = "g"): p-value will be
## approximate in the presence of ties

## Warning in ks.test(sample2, sample1, , alternative = "g"): p-value will be
## approximate in the presence of ties

## Warning in ks.test(sample2, sample1, , alternative = "g"): p-value will be
## approximate in the presence of ties

## Warning in ks.test(sample2, sample1, , alternative = "g"): p-value will be
## approximate in the presence of ties

## Warning in ks.test(sample2, sample1, , alternative = "g"): p-value will be
## approximate in the presence of ties

## Warning in ks.test(sample2, sample1, , alternative = "g"): p-value will be
## approximate in the presence of ties

## Warning in ks.test(sample2, sample1, , alternative = "g"): p-value will be
## approximate in the presence of ties
```

```r
names <- c('user_followers_count', 'user_friends_count', 'num_urls', 'user_favourites_count', 'num_menti
KolmogorovSmirnov_statistic2 <- data.frame(names, differences, p_vals)
KolmogorovSmirnov_statistic2
```

```
##                   names differences       p_vals
## 1  user_followers_count 0.235713933 1.286269e-06
## 2    user_friends_count 0.020151134 9.056242e-01
## 3              num_urls 0.005241517 9.933155e-01
## 4 user_favourites_count 0.058558058 4.329595e-01
## 5          num_mentions 0.113492122 4.309171e-02
## 6             num_media 0.094754778 1.117092e-01
## 7         retweet_count 0.043494098 6.301373e-01
## 8          num_hashtags 0.035011360 7.413773e-01
```