Econ 1042 PS1

Owen Asnis

2023-02-01

Data Wrangling

```
kickers <- read_csv("kickers_v2.csv",
                    show_col_types = FALSE) %>%
  rename(., "ID" = "...1") %>%
  mutate(Grass = if_else(Grass == TRUE, 1, 0))
```

Question 1

a.)

```
min <- min(kickers$Distance)
max <- max(kickers$Distance)
mean <- round(mean(kickers$Distance),
              digits = 3)
median <- median(kickers$Distance)

sprintf("Minimum distance: %s yards.", min)
```

```
## [1] "Minimum distance: 18 yards."
```

```
sprintf("Maximum distance: %s yards.", max)
```

```
## [1] "Maximum distance: 76 yards."
```

```
sprintf("Mean distance: %s yards", mean)
```

```
## [1] "Mean distance: 36.897 yards"
```

```
sprintf("Median distance: %s yards", median)
```

```
## [1] "Median distance: 37 yards"
```

b.) The minimum distance (18 yards) isn't lower for two reasons: First, the kicker must kick the ball through the end zone (accounting for 10 yards) in addition to the distance from the ball to the goal line, and second the ball is snapped a distance of 7 yards to the holder, who holds the ball for the kicker. Therefore, an 18 yard field goal actually begins at the 1 yard line, but the total distance of the kick is 10 yards (the end zone) + 7 yards (the snap) + 1 yard (the distance to the goal line) = 18 yards.

c.)

```
max_distance_kick <- kickers %>%
  filter(Distance == max(Distance))

print(max_distance_kick)
```

```
## # A tibble: 1 x 9
##      ID Team   Year GameMinute Kicker     Distance ScoreDiff Grass Success
##   <dbl> <chr> <dbl>      <dbl> <chr>         <dbl>     <dbl> <dbl>   <dbl>
## 1  3558 OAK    2008         30 Janikowski       76        15     1       0
```

The special circumstance that explains the maximum is that it occured in the 30th minute of the game, the last minute of the 2nd quarter before halftime. Therefore, in this instance, the Oakland Raiders were forced to try to score, because the half was ending and possession is assigned based on the result of the coin toss after the 2nd quarter, and decided they had a better chance of scoring on a very long field goal rather than a hail mary.

Question 2

```
forty_to_fortyfive <- kickers %>%
  filter(Distance %in% (40:45))

forty_to_fortyfive_success <- round((mean(forty_to_fortyfive$Success) * 100),
                                    digits = 3)

over_fortyfive <- kickers %>%
  filter(Distance > 45)

over_fortyfive_success <- round((mean(over_fortyfive$Success) * 100),
                                digits = 3)

sprintf("Kicks from 40 to 45 yards made: %s percent", forty_to_fortyfive_success)
```

```
## [1] "Kicks from 40 to 45 yards made: 79.215 percent"
```

```
sprintf("Kicks over 45 yards made: %s percent", over_fortyfive_success)
```

```
## [1] "Kicks over 45 yards made: 64.448 percent"
```

Question 3

```
grass_only <- kickers %>%
  filter(Grass == 1)

make_rate_grass <- round((mean(grass_only$Success) * 100),
                         digits = 3)

turf_only <- kickers %>%
  filter(Grass == 0)

make_rate_turf <- round((mean(turf_only$Success) * 100),
                        digits = 3)

sprintf("Make rate on grass: %s percent", make_rate_grass)
```

```
## [1] "Make rate on grass: 82.393 percent"
```

```
sprintf("Make rate on turf: %s percent", make_rate_turf)
```

```
## [1] "Make rate on turf: 84.326 percent"
```

The make rate was slightly higher on turf.

```
# Source: "Dealing with heteroskedasticity; regression with robust standard
# errors using R" by Bruno Rodrigues

m1 <- lm(Success ~ Grass, data = kickers)
coeftest(m1, vcov = vcovHC(m1))
```

```
##
## t test of coefficients:
##
##               Estimate Std. Error  t value  Pr(>|t|)
## (Intercept)  0.8432614  0.0051154 164.8473 < 2.2e-16 ***
## Grass       -0.0193292  0.0070587  -2.7384  0.006184 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The difference is statistically significant at the 0.001 level. It could be the true effect of surface, because turf is consistantly a smoother and more solid surface, whereas grass is more imperfect with divits and inconsistencies. Therefore, in some ways, it seems like turf provides a better condition for kickers. However, there are clear counterarguments, including that grass might be better in rainy and snowy conditions, because it would be less slippery and easier to get good footing than turf. Also, this model is incredibly simple, and doesn't account for many other variables.

Question 4

```
cor(kickers$Distance, kickers$Grass)
```

```
## [1] -0.002551996
```

```
cor(kickers$Distance, kickers$Success)
```

```
## [1] -0.336934
```

Question 5

```
# short-form(coefficient) = long-form(coefficient) + (omitted variable coefficient)(correlation)

lm2 <- lm_robust(Success ~ Grass + Distance, data = kickers)
summary(lm2)
```

```
##
## Call:
## lm_robust(formula = Success ~ Grass + Distance, data = kickers)
##
## Standard error type:  HC2
##
## Coefficients:
##             Estimate Std. Error t value   Pr(>|t|) CI Lower CI Upper    DF
## (Intercept)  1.29988  0.0112628 115.413  0.000e+00  1.27780  1.32195 11184
## Grass       -0.01997  0.0066494  -3.004  2.671e-03 -0.03301 -0.00694 11184
## Distance    -0.01237  0.0003322 -37.218 6.044e-286 -0.01302 -0.01171 11184
##
## Multiple R-squared:  0.1142 ,    Adjusted R-squared:  0.1141
## F-statistic: 695.7 on 2 and 11184 DF,  p-value: < 2.2e-16
```

```
# becomes more negative (adding distance is omitted variable bias)
```

Question 6

```
lm3 <- lm_robust(Success ~ Distance + Grass + ScoreDiff + GameMinute,
                 data = kickers)
summary(lm3)
```

```
##
## Call:
## lm_robust(formula = Success ~ Distance + Grass + ScoreDiff +
##     GameMinute, data = kickers)
```

```
##
## Standard error type:  HC2
##
## Coefficients:
##               Estimate Std. Error  t value   Pr(>|t|)   CI Lower   CI Upper
## (Intercept)  1.299e+00  0.0128000 101.4621  0.000e+00  1.2736278  1.3238085
## Distance    -1.237e-02  0.0003330 -37.1581 4.484e-285 -0.0130245 -0.0117192
## Grass       -1.997e-02  0.0066489  -3.0039  2.672e-03 -0.0330054 -0.0069394
## ScoreDiff   -1.022e-04  0.0003593  -0.2846  7.760e-01 -0.0008064  0.0006020
## GameMinute   4.408e-05  0.0001991   0.2213  8.248e-01 -0.0003463  0.0004344
##                 DF
## (Intercept) 11182
## Distance    11182
## Grass       11182
## ScoreDiff   11182
## GameMinute  11182
##
## Multiple R-squared:  0.1142 ,    Adjusted R-squared:  0.1139
## F-statistic: 347.8 on 4 and 11182 DF,  p-value: < 2.2e-16
```

```
lm4 <- lm_robust(Success ~ Distance + Grass + ScoreDiff + GameMinute + Kicker,
                 data = kickers)
summary(lm4)
```

```
##
## Call:
## lm_robust(formula = Success ~ Distance + Grass + ScoreDiff +
##     GameMinute + Kicker, data = kickers)
##
## Standard error type:  HC2
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|) CI Lower CI Upper    DF
## (Intercept)        1.266e+00        NaN     NaN      NaN      NaN      NaN 11100
## Distance          -1.246e-02        NaN     NaN      NaN      NaN      NaN 11100
## Grass             -2.456e-02        NaN     NaN      NaN      NaN      NaN 11100
## ScoreDiff         -3.034e-05        NaN     NaN      NaN      NaN      NaN 11100
## GameMinute         4.117e-05        NaN     NaN      NaN      NaN      NaN 11100
## KickerAndersen     4.382e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerAndrus      -3.664e-01        NaN     NaN      NaN      NaN      NaN 11100
## KickerBailey       1.116e-01        NaN     NaN      NaN      NaN      NaN 11100
## KickerBarth        7.746e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerBironas      7.605e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerBoswell      1.248e-01        NaN     NaN      NaN      NaN      NaN 11100
## KickerBrien       -4.597e-01        NaN     NaN      NaN      NaN      NaN 11100
## KickerBrindza     -2.372e-01        NaN     NaN      NaN      NaN      NaN 11100
## KickerBrown        3.113e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerBryant       5.718e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerBuehler     -3.196e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerBullock      3.496e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerCarney       3.711e-03        NaN     NaN      NaN      NaN      NaN 11100
## KickerCarpenter    6.895e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerCatanzaro    9.254e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerCoons        6.276e-02        NaN     NaN      NaN      NaN      NaN 11100
```

```
## KickerCortez     -1.175e-01        NaN     NaN      NaN      NaN      NaN 11100
## KickerCoutu      -7.064e-01        NaN     NaN      NaN      NaN      NaN 11100
## KickerCrosby      1.127e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerCundiff    -3.313e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerDawson      6.070e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerEdinger    -7.195e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerElam        3.501e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerElling     -5.950e-01        NaN     NaN      NaN      NaN      NaN 11100
## KickerFeely       4.763e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerFolk        1.124e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerForbath     5.427e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerFrance     -4.723e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerFranks      3.613e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerFreese     -3.562e-01        NaN     NaN      NaN      NaN      NaN 11100
## KickerGano        1.730e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerGostkowski  6.010e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerGould       6.641e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerGraham      3.745e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerGramatica   1.663e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerHall        2.865e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerHanson      6.731e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerHartley     1.052e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerHauschka    7.434e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerHenery      1.834e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerHocker     -9.164e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerHopkins     1.042e-01        NaN     NaN      NaN      NaN      NaN 11100
## KickerJanikowski  5.138e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerKaeding     4.317e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerKasay       6.534e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerKoenen     -4.115e-01        NaN     NaN      NaN      NaN      NaN 11100
## KickerLambo       7.430e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerLindell     3.252e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerLongwell    4.079e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerMare       -9.964e-03        NaN     NaN      NaN      NaN      NaN 11100
## KickerMcManus     6.306e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerMedlock    -1.423e-01        NaN     NaN      NaN      NaN      NaN 11100
## KickerMehlhaff   -1.027e-01        NaN     NaN      NaN      NaN      NaN 11100
## KickerMurray      1.003e-01        NaN     NaN      NaN      NaN      NaN 11100
## KickerMyers       6.213e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerNedney      7.429e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerNovak       3.248e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerNugent      8.303e-03        NaN     NaN      NaN      NaN      NaN 11100
## KickerParkey      6.656e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerPeterson    5.377e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerPettrey    -4.034e-01        NaN     NaN      NaN      NaN      NaN 11100
## KickerPotter     -6.652e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerPrater      5.309e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerRackers     5.314e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerRayner     -6.415e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerReed        2.355e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerSantos      5.739e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerSchmitt    -1.605e-01        NaN     NaN      NaN      NaN      NaN 11100
## KickerScifres     2.547e-01        NaN     NaN      NaN      NaN      NaN 11100
## KickerScobee      3.290e-02        NaN     NaN      NaN      NaN      NaN 11100
```

```
## KickerStitser    -9.251e-03        NaN     NaN      NaN      NaN      NaN 11100
## KickerStover      4.535e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerSturgis     7.998e-03        NaN     NaN      NaN      NaN      NaN 11100
## KickerSuccop      4.197e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerSuisham     2.981e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerTucker      1.047e-01        NaN     NaN      NaN      NaN      NaN 11100
## KickerTynes      -1.301e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerVanderjagt -2.028e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerVinatieri   6.188e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerWalsh       7.112e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerWilkins     4.380e-02        NaN     NaN      NaN      NaN      NaN 11100
## KickerZuerlein    2.767e-02        NaN     NaN      NaN      NaN      NaN 11100
##
## Multiple R-squared:  0.1273 ,    Adjusted R-squared:  0.1205
## F-statistic:    NA on 86 and 11100 DF,  p-value: NA
```

```
# Accounts for the skill of the kicker
```

Question 7

```
new_data <- kickers %>%
  filter(Kicker == "Tucker",
         ScoreDiff == -11,
         Year == 2015,
         GameMinute == 30)

predict(lm4, newdata = new_data) * 100
```

```
##        1
## 99.85399
```

```
# estimate - Tucker coefficient
```