Project Part 3

Were the top 20 golfers from the 2020-21 PGA Tour season better than the average PGA Tour pro at putting?

Background and Data Description

Golf is an incredibly difficult sport both mentally and physically. With the new analytics boom in all sports, golf has followed suit with consistently new technology in golf balls, clubs, and swing analysis. I want to research, as so much new emphasis is being placed on hitting the golf ball further on every shot, if putting still has an impact in forming the difference between the top golfers consistently competing for tournament titles and the rest of the golfers just trying to stay on the PGA Tour. I collected the PGA Tour data by copying the statistics from both the ESPN (1) and PGA Tour (2) websites into an excel spreadsheet. They were converted into csv's and then merged by the name of each golfer in the dataset. The ESPN data included basic statistics for the golfers - statistics like final ranking, number of rounds played, number of wins, average driving distance, and many more. The PGA Tour data contained more advanced and modern analytics like club head speed, ball speed, and strokes gained. Strokes gained is a modern way of comparing golfers directly with each other (3). For every round in each tournament, the net number of strokes gained by all players that day will be zero. If a golfer played better than the rest of the field, their strokes gained will be positive and if they didn't play well, it will be negative. I will be focusing on its putting-specific component to analyze if the top 20 PGA Tour golfers have an above average strokes gained.

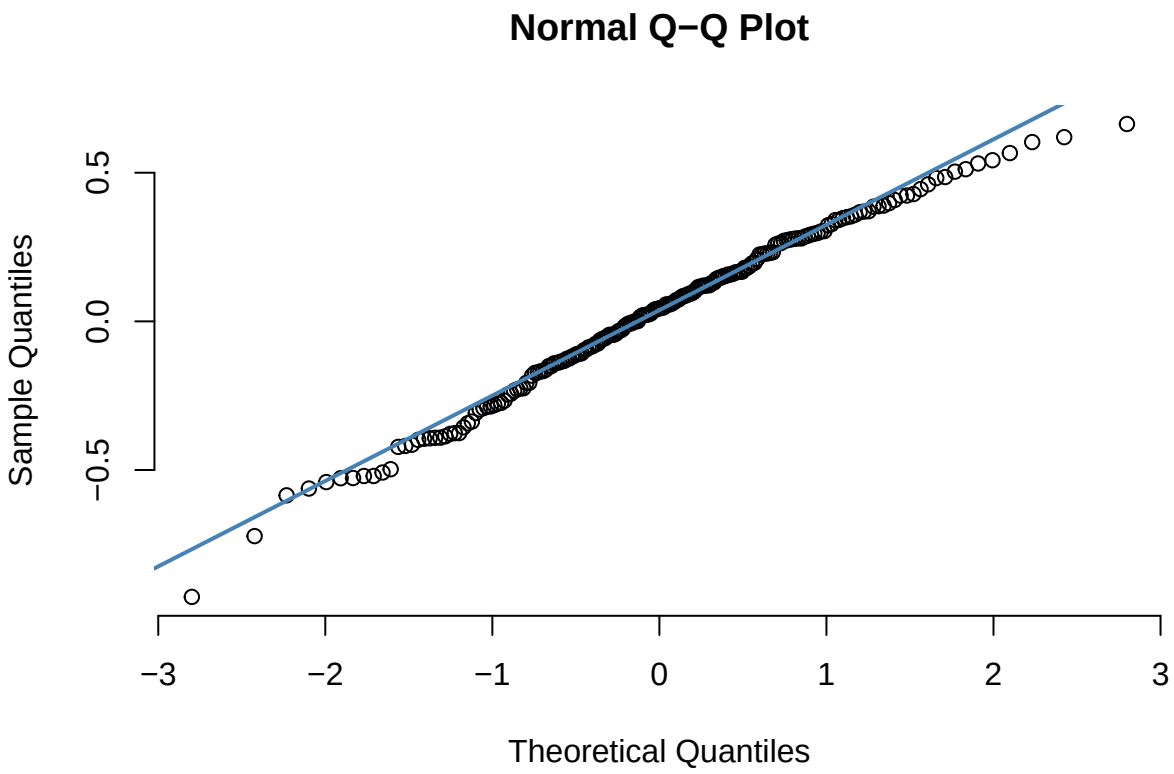
Statistical Test

One-Sample t-test

I will be testing the if mean of the average strokes gained putting of the top 20 golfers is equal to or greater than the mean of the average strokes gained of all golfers in the 2020-21 PGA

Tour season. As stated earlier, the net strokes gained is always zero, so that will be the value of the mean for my null hypothesis. Because my dataset does not contain every single golfer that participated in the season of interest, I do not know the population standard deviation and thus will perform a one sample t-test. For the assumptions of this test, the data is continuous, there are no significant outliers in the sample, and the data is approximately normally distributed (shown in qqplot below). The strokes gained statistic is somewhat dependent one other golfers (must sum to zero), but I feel comfortable continuing with this test and trusting the results because the physical putting of each golfer is independent and that's what causes the differences within this statistic.

```
qqnorm(big_pga$AVERAGE.SG.P, pch = 1, frame = FALSE)
qqline(big_pga$AVERAGE.SG.P, col = "steelblue", lwd = 2)
```



Hypotheses

H0: mu = 0 vs. HA: mu > 0

Test code

```
topPutting <- big_pga %>% filter(RK <= 20) %>% select(AVERAGE.SG.P)
t.test(topPutting, alternative = 'greater', mu = 0, conf.level = 0.95)
```

```
##
##  One Sample t-test
##
## data:  topPutting
## t = 3.8056, df = 19, p-value = 0.0005972
## alternative hypothesis: true mean is greater than 0
## 95 percent confidence interval:
##  0.1293706       Inf
## sample estimates:
## mean of x
## 0.2371007
```

Result and Conclusion

Test stat: t = 3.8056; p-value: p = 0.0005972

Because the p-value is less than 0.05, there is enough evidence to reject the null hypothesis and suggest that the true mean strokes gained of top 20 golfers on the PGA Tour in the 2020-21 season is greater than zero. This means that the top golfers are better putters on average than all PGA Tour golfers as a whole. The results from the test even suggests that the true mean for the top 20 golfers is actually 0.237. This means that the top golfers gain about .237 strokes per round on the rest of the field from putting alone (almost a stroke per tournament, which are usually four rounds). Putting is clearly important for success on the PGA Tour.

Further application of the data

This test can be used to explain the putting of top golfers during the 2020-21 PGA Tour season, but I believe it can be generalized past just the population of golfers from this season. Because they play most tournaments on the same courses every year and most top golfers have little year-to-year variation, I feel comfortable claiming that the top golfers on the PGA Tour are always above average putters for any season from the past and in the future.

References

1. https://www.espn.com/golf/stats/player/_/season/2021
2. https://www.pgatour.com/stats.html
3. https://www.pgatour.com/news/2016/05/31/strokes-gained-defined.html