Horsing Around:
## Race Betting for the Casual Bettor
● ● ●

Zi Yang
Oct 2020

Question

Can I make money using tipsters in horse race betting?

Does it matter if the tipster is active?

What is a Tipster?

- A tipster is someone who charges to regularly provides information (tips) on the likely outcomes of sporting events to deliver profit over the long term to their clients

- Some tipsters use statistical based estimations about the outcome of a game, and compare this estimation with the bookmaker's odds.

- Tipster "Active" vs "Inactive" : Aggregation sites online rank tipsters by their profit performance value. At a certain point of negative profit, a tipster becomes "Inactive."

    - For this data, the performance index used to determine Active/Inactive status is unknown.
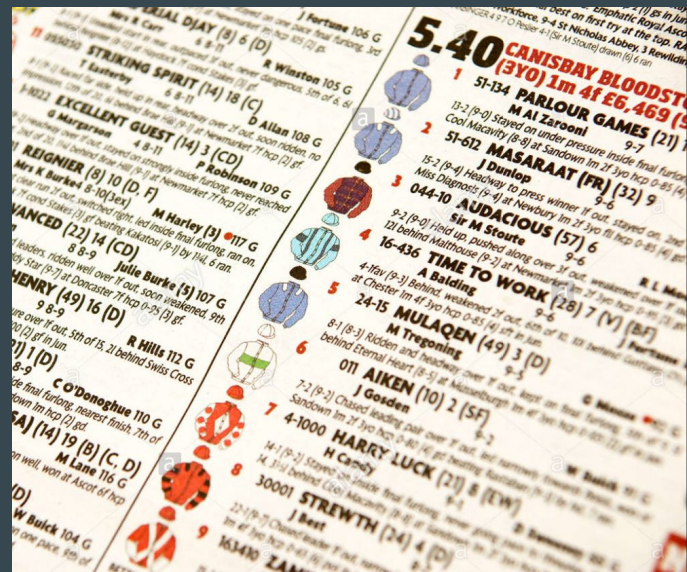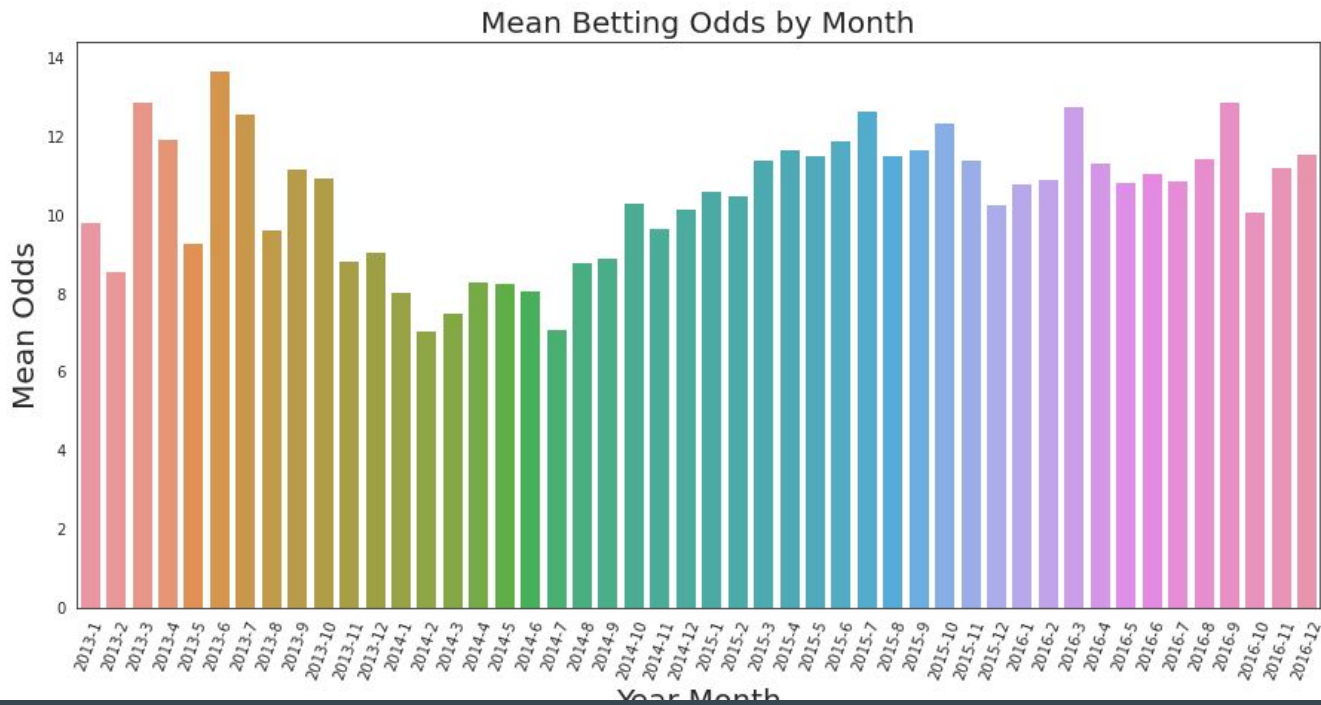
Data

	index	UID	ID	Tipster	Date	Track	Horse	Bet Type	Odds
0	20268	20269	43	Tipster T	2013-01-06	Epsom	Galileo Rock	Win	23.0
1	20267	20268	42	Tipster T	2013-01-06	Epsom	Captaine Dunne	Win	9.0
2	20266	20267	41	Tipster T	2013-01-06	Epsom	St Paul De Vence	Win	11.0
3	23412	23413	72	Tipster X	2013-01-12	Fairyhouse	Minella Foru	Win	6.5

Odds are read as a ratio to the dollar.

- For a 'Win' result, Profit = Odds - 1.00
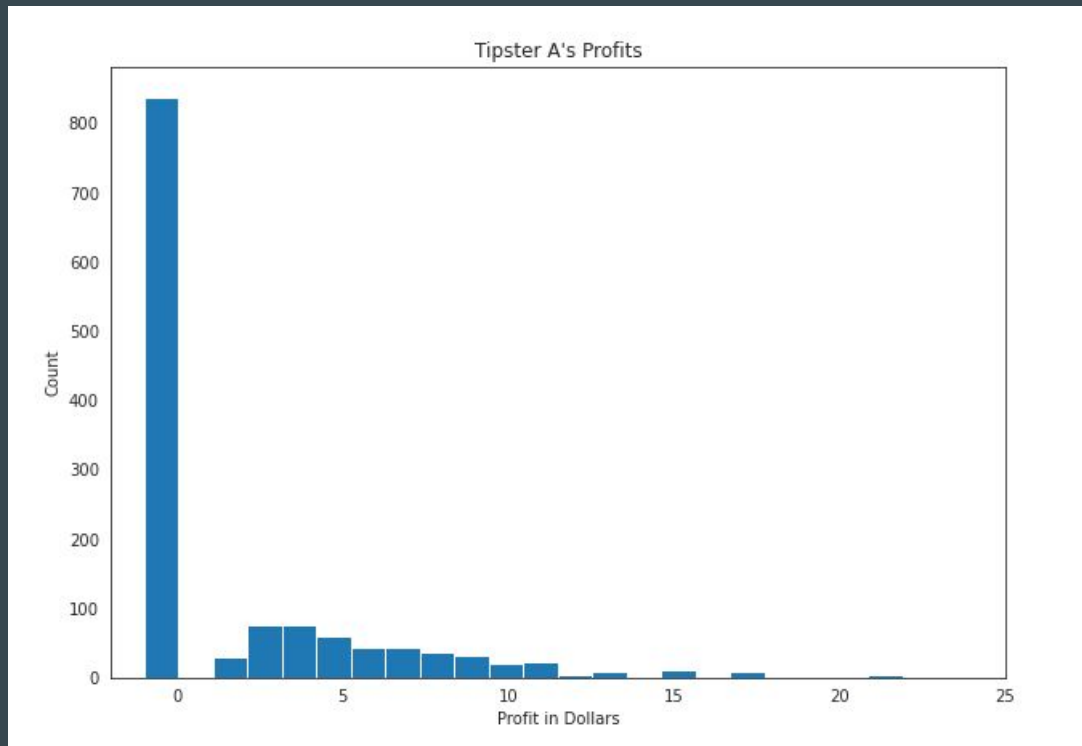- For a 'Lose' result, Profit = -$1.00

Data

- Betting Odds center around 10.00

Data

- Tipsters earn money over time. Most of the time their bets lose.

Tipster A's Profits

Assumptions/Methods - Bootstrap Hypothesis Testing

- **I want to be 95% confident that I can trust a tipster's tips to win money at a Horse Race**
- 31 tipsters, that is a max 5% Type I Error rate across 31 tests
  - use the *Bonferroni Correction* method to get confidence level :
  - 0.05/31 = 99.92
- Ran simulations for each tipster with
  - Bootstrap with resample of sample size (n) 20,000 from each tipster
  - resample 1,000 times for each tipster.
  - look at the sums and means of the 1,000 resamples
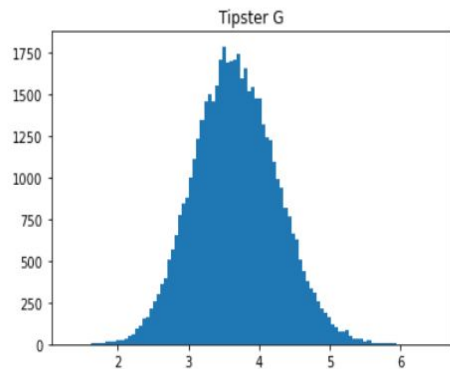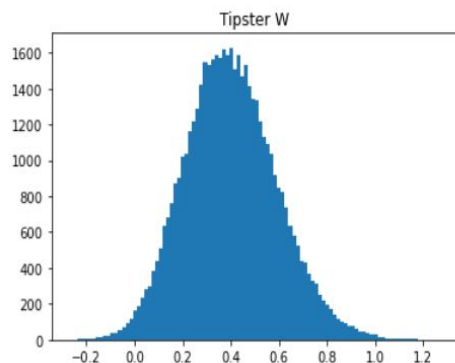 - Calculate confidence level
  - draw conclusion

Results
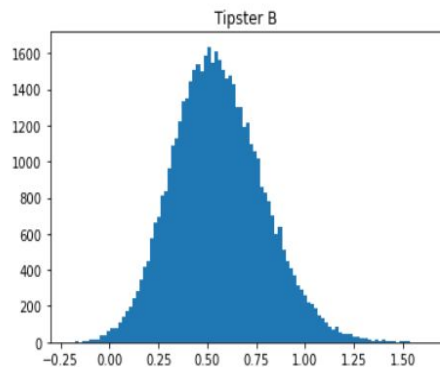
Tipster B bootstrap confidence intervals: -0.103, 1.414
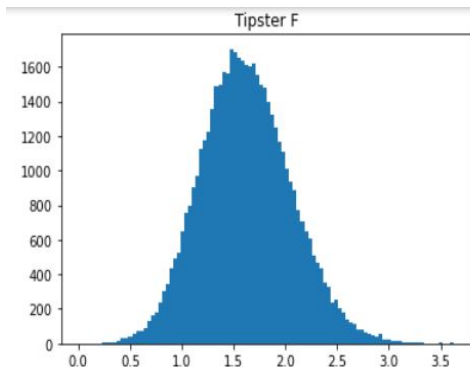
Tipster F bootstrap confidence intervals: 0.363, 3.265

Tipster G bootstrap confidence intervals: 1.841, 5.774

Tipster W bootstrap confidence intervals: -0.161, 1.116

	lower	higher
0	1.095063	1.925420
1	-0.103142	1.414358
2	1.671899	2.877297
3	0.319838	1.860999
4	-0.021306	0.203062
5	0.363010	3.265411
6	1.841378	5.774143
7	-0.226448	0.718583
8	0.244568	1.893202
9	0.456892	1.211131
10	-0.139230	1.028051
11	-0.023958	0.548330
12	0.859685	1.955298
13	0.737397	2.222971
14	-0.171391	0.856947
15	-0.115750	0.559656

	lower	higher
16	-0.006115	0.629537
17	0.305328	0.729364
18	0.165404	1.621726
19	-0.206149	0.553108
20	0.308612	1.993812
21	-0.511207	0.673399
22	-0.161218	1.116329
23	0.196968	0.567456
24	0.234574	0.859572
25	0.407882	1.211540
26	0.062771	0.906423
27	-0.063384	0.223533
28	0.008122	1.153773
29	1.560051	2.781638
30	0.503395	1.679936

Conclusion

- Out of 31 tipsters, 12 did not fall in my 95% confidence level
  - I would not trust 12 out of 31, or 38. 7% of tipsters to win me any money had I followed their advice
  - **not enough evidence to support that ~40% of tipsters have skill**
- Not factored: the cost of employing the tipster

Conclusion: bet at your own risk, and pick tipsters at your own risk!

Additional Analysis

- Find out how many tipsters were active in a certain time
  - *"at this time there were <x> active tipsters, only <y> of them would yield a profit if you followed every one of their tips."*
- Plot distribution of odds for losing tips vs winning tips within each tipster
  - if the distributions look the same, then it builds the story that it is really hard to distinguish winning tips from losing tips
  - if you want to make money then you have to follow every tip
- Look up and include cost of employing a tipster, factor that into profit calculation

Bootstrap

Bootstrap one tipster with the biggest win

- Sample with replacement from the sample of one tipster's tips
- Sample size = total number of tips
- Repeat 1000 times
- Sum each 1000 resamples
- Plot histogram of 1000 sums
- Use percentile function (0.025 to 0.975 percentile)

Build confidence intervals for mean profit for each tipster (using just the data from the csv)

- Mean profit +/- 1.96*std(ddof=1)/sqrt(n)
- If 0 is in the interval between the high and low -> not confident that I will make money following their tips (as a casual gambler)
-

What is a Tipster?

A **tipster** is someone who regularly provides information (**tips**) on the likely outcomes of sporting events on internet sites or special betting places.

In the past tips were bartered for and traded but nowadays, thanks largely to the Internet and premium rate telephone lines, they are usually exchanged for money, and many tipsters operate websites. Some of them are free and some require subscription.

A tip in gambling is a bet suggested by a third party who is perceived to be more knowledgeable about that subject than the bookmaker who sets the initial odds.

The Tipster must overcome the profit margin integrated into sports betting odds by bookmakers trading teams and then also obtain an additional edge to deliver profit over the long term.

Some tipsters use statistical based estimations about the outcome of a game, and compare this estimation with the bookmaker's odds. If there is a gap between the estimate odds and the bookmakers odds, the tipster is said to identify "value", and a person who bets on such odds when they perceive not a certainty but a "gap in the book" is said to be a "value bettor". When value is found, the tipster is recommending the bettor to place a bet.

A tip that is considered to be a **racing certainty**, that is, almost completely certain to be true, is also called a **nap** and tipsters in newspapers will tend to indicate the "nap".

Tipster "Active" vs "Inactive" : Aggregation sites online rank tipsters by their profit performance value. At a certain point of negative profit, a tipster becomes "Inactive." For this data, the performance index used to determine Active/Inactive status is unknown.

Forming a Hypothesis Statement

1. State a scientific Y/N question
   a. "Can I make money using tipsters in horse race betting?"
   b. Does using a tipster increase your chance at making money in horse races?
   c. Does it matter if they're in the Active or Inactive status?
2. State skeptical Null hypothesis
   a. H0: No, using a tipster does not give net positive profit from betting in horse races.
3. State alternate hypothesis
   a. Ha: Yes
4. Create probabilistic model of situation when the null is true
   a. Monte Carlo - simulation model
5. Determine alpha - how surprised you'd have to be to reject the null
   a. Alpha = 5%
6. Collect Data (Simulate Data in this case)
7. Calculate p-value: conditional probability of finding a result equal or more extreme than what you actually observed, assuming the null is true

Grading Rubric

- Describe the practical application of the project.

- Document all assumptions about the data.

- Lay out steps for future work.

- Identify and describe data sources.

- Provide instructions to obtain the data and replicate the data preparation.

- If appropriate, provide scripts to obtain the data.

- Include a small representative sample of data in the repository.