

# Evolution of information in prediction markets

APS March Meeting 2018

Avaneesh Narla

Dept. of Physics, UC San Diego

Benjamin Machta

Dept. of Physics and Systems Biology, Yale University

6th March, 2018

# PredictIt

- 1 PredictIt is a political prediction marketplace based in Australia. It is a research project of Victoria University of Wellington and lets its users make and trade predictions on political events.



 **2020 Democratic nominee?**

Top Predictions ■ 8883 Comments

<b>Sanders</b> <span>■</span> 46	18¢	NC
<b>Biden</b> <span>■</span> 24	15¢ <span>↓</span> 1¢	

 **Will Trump be president at year-end 2018?**


Current Market ■ 5274 Comments

<b>Yes</b>	80¢ <span>↑</span> 1¢
<b>No</b>	20¢ <span>↓</span> 1¢

 **2020 presidential winner?**

Top Predictions ■ 1595 Comments

<b>Trump</b> <span>■</span> 96	33¢ <span>↑</span> 1¢
<b>Biden</b> <span>■</span> 1	12¢ NC

 **GOP Senate seats after midterms?**

Top Predictions ■ 1709 Comments

<b>49-3</b>	33¢ <span>↑</span> 2¢
<b>50-5</b>	16¢ NC

 **Government shutdown on 2/9?**

Current Market ■ 14554 Comments

<b>Yes</b>	16¢ <span>↑</span> 2¢
<b>No</b>	84¢ <span>↓</span> 2¢

 **2020 Republican nominee?**


Top Predictions ■ 351 Comments

<b>Trump</b> <span>■</span> 21	55¢ <span>↑</span> 1¢
<b>Pence</b> <span>■</span> 1	13¢ <span>↓</span> 1¢

 **Who will be the next justice to leave the Supreme Court?**

Top Predictions ■ 1077 Comments

<b>Kennedy</b> <span>■</span> 31	47¢	NC
<b>Ginsburg</b> <span>■</span> 9	28¢	NC

 **GOP House seats after midterms?**

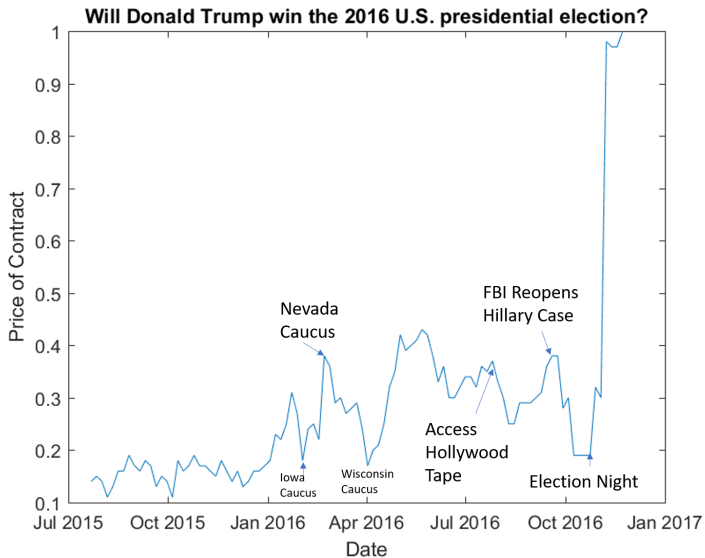
Top Predictions ■ 433 Comments

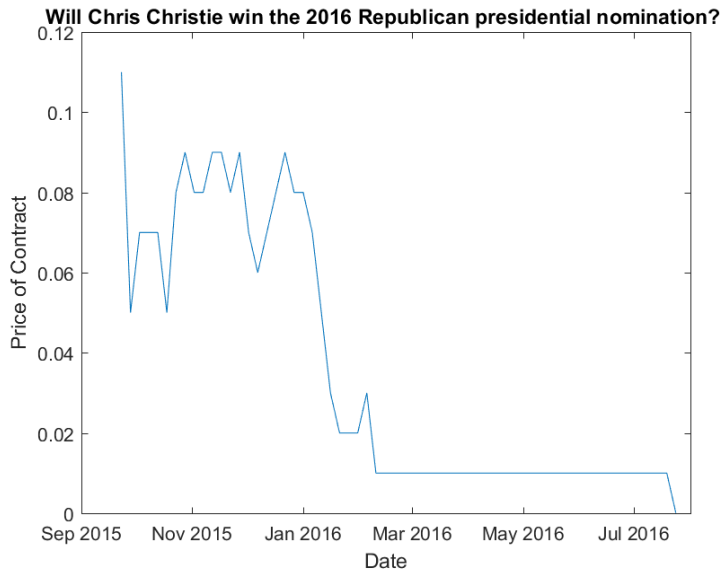
<b>217-11</b>	58¢	NC
<b>218-225</b>	19¢	NC

 **Which party will win the White House in 2020?**

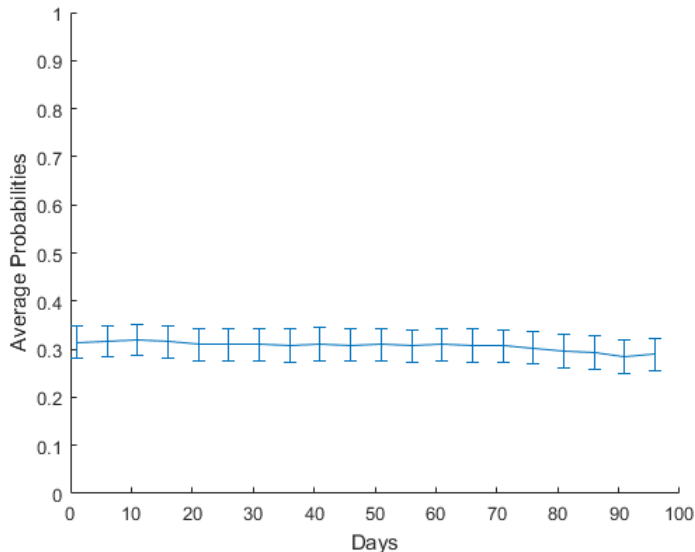
Top Predictions ■ 2373 Comments

<b>Democratic</b> <span>■</span> 47	58¢ <span>↑</span> 1¢	
<b>Republican</b> <span>■</span> 30	44¢	NC

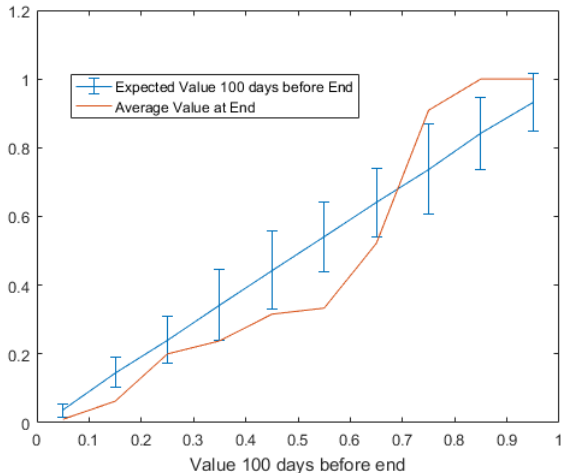




# Plot of average probabilities over last 100 days



# Average value at end of contract for binned predicted values 100 days before end



# A few terms

- 1 Entropy ( $S$ ) is a measure of unpredictability of the average information content of a state ( $X$ ).

$$S(X) = - \sum_{i=1}^n P(x_i) \log P(x_i) = -\pi \log \pi - (1 - \pi) \log(1 - \pi)$$

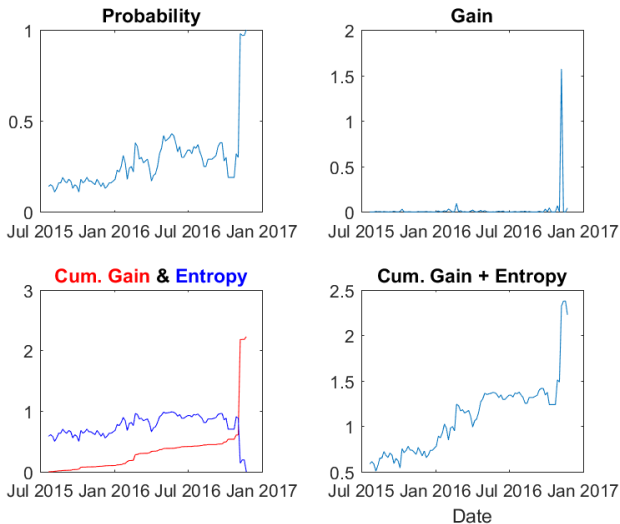
The Kullback-Leibler Divergence,  $D_{KL}(P||Q)$  is a measure of the information gained when one revises one's beliefs from the prior probability distribution  $Q$  to the posterior distribution  $P$

$$D_{KL}(P||Q) = \sum_{x_i} P(x_i) \log \frac{P(x_i)}{Q(x_i)}$$

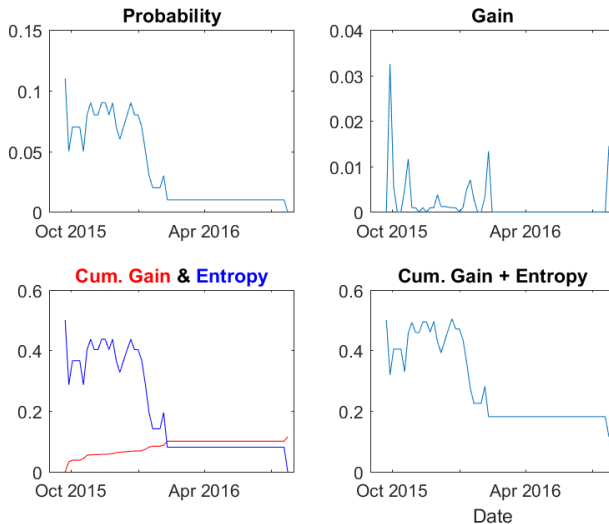
- 2 We introduce *Gain*  $G$  over time increment  $t \rightarrow t + \Delta t$

$$G_{t \rightarrow t + \Delta t} \equiv D_{KL}(P_{t + \Delta t} || P_t)$$

# Will Donald Trump win the 2016 U.S. Presidential Election?



# Will Chris Christie win the 2016 Republican presidential nomination?



Expected gain in information is equal to the expected decrease in entropy for a time-evolving probability distribution

## Theorem

*The expected gain from time  $t$  to  $(t + \Delta t)$ ,  $\langle G_{t \rightarrow t + \Delta t} \rangle_t$ , given the probability distribution at time  $t$  is the difference between the entropy at time  $t$  and the expected entropy at time  $(t + \Delta t)$*

$$\langle G_{t \rightarrow t + \Delta t} \rangle_t = H_t - \langle H_{t + \Delta t} \rangle_t$$

## Lemma

$$\left\langle \sum_{x_i} P_{t + \Delta t}(x_i) \log P_t(x_i) \right\rangle_t = -H_t$$

## Lemma

$$\left\langle \sum_{x_i} P_{t+\Delta t}(x_i) \log P_t(x_i) \right\rangle_t = -H_t$$

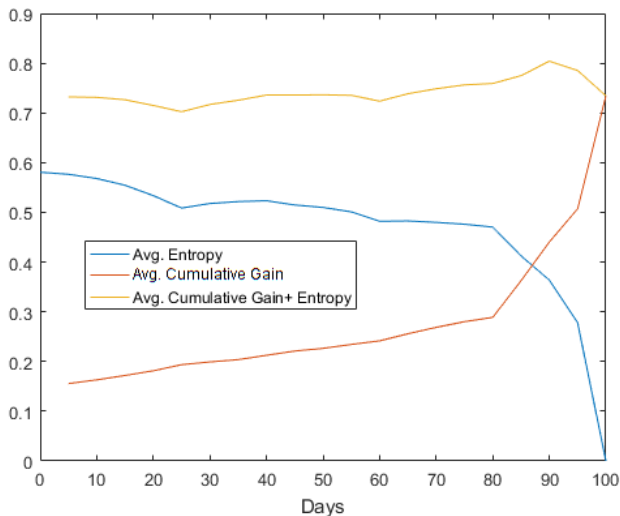
## Sketch of Argument

- 1 *The probability distribution at time  $t$  is independent of future events*
- 2 *The expected probability at time  $t + \Delta t$ , averaged over all events that could take place between  $t$  and  $t + \Delta t$ , is the probability at time  $t$*

# Hypothesis

- ① On average, if the markets do reflect the actual probabilities of an event taking place and the updates reflect new collective information, the sum of entropy and the cumulative gain will be constant.
- ② Though they will fail for certain events, if the markets are able to assess probabilities, on average the information gained should reflect a reduction in entropy.

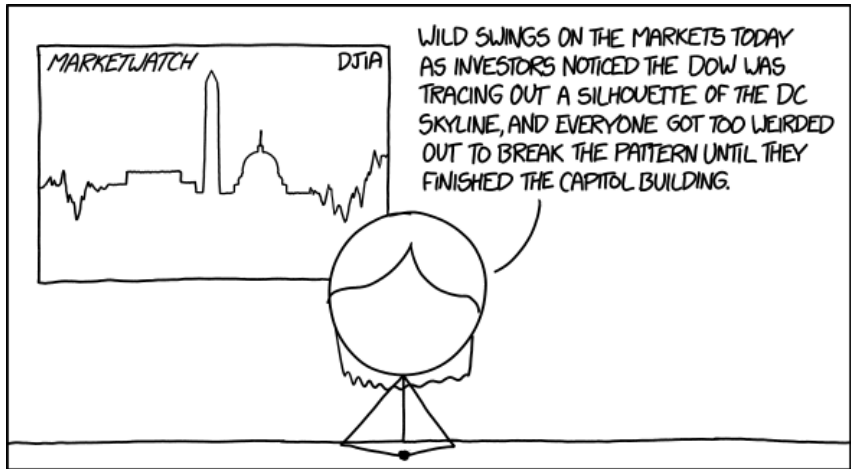
# Plot of average values over last 100 days for chosen markets



# Conclusion

- 1 Expected information gain is the expected decrease in entropy for a time-evolving probability distribution.
- 2 Thus, sum of expected gain and entropy is constant.
- 3 We tested this property for political prediction data from PredictIt
- 4 PredictIt data seems to reflect probabilities with Bayesian updates

# Thanks!



Source: [xkcd.com](http://xkcd.com)