Automated Forecasting

Junshi

March 9, 2022

Abstract

1 Introduction

Most of the important questions answered by forecasters are still done manually. This 1) is time-consuming and inefficient 2) may lead to subjective decisions since personal judgements are made frequently on when and how to adjust the probability. Therefore I want to discover to possibility of following a quantative/statistical approach while at the same adopt the conventional reference class method. There are certain constraints due to the scalability and accessibility to useful datasets. And most importantly, whether this model can outperform human forecasters(as stated in Superforecasting, his superforecasters already beat experts in Math, Stat and other fields).

2 Testing Simple Models

This section is dedicated to testing this statistical model method in forecasting. Two resolved questions are selected here, **Election Fraud** and **Cultured Meat**.

2.1 Simple Model for Election Fraud

Here is the Metaculus link for this question. I now try to use a very simple and inaccurate model, and do a forecast with data accessible on Google Search, keyword set as "election fraud".

2.1.1 Model Setting

Our problem of interest is "Probability of greater than 20% people believe election rigged". Let's define b as a random variable corresponding to the individual's belief in a rigged election. $b = 1$ if one believes this firmly, and $b = 0$ otherwise. Then $p = \mathbf{E}(b)$ is the proportion of people firmly believe in the fraud election. So $\mathbf{P}(p > 0.2)$ is what we cares about. This p also follows a normal distribution(I adopt the Bayesian approach).

As for Google Search, denoted as I, has a relationship with b and p. Intuitively, all those with $b = 1$ with search this topic definitely while the others will pay attention to this with some probability(random variable s). I assume s to follow a normal distribution in the month when everyone is talking about this and it becomes a constant 0.

To summarize, Statistical Model is

$$
\begin{aligned}
b &\overset{iid}{\sim} \mathbf{Bino}(1, p) \\
Prior : p &\overset{iid}{\sim} \mathbf{N}(1/2, 1/9) \\
I &\sim \mathbf{N}(p + s(1-p), 1/9) \\
s &\sim \mathbf{N}(1/2, 1/9)\mathbf{1}\{T \leq 1\}
\end{aligned}
$$

Why 1/9? Only to simulate a truncated normal.

2.1.2 Result

In the month when election was really trendy(Nov 2021), the posterior distribution of p is $p \stackrel{iid}{\sim} \mathbf{N}(\frac{2}{5}\hat{I}+\frac{9}{20}, 4/45)$, $\hat{I} = 0.28$. To estimate the believers of a rigged election in the following year(except Nov 2021),$\hat{I} = 0.05$, $p \stackrel{iid}{\sim} \mathbf{N}(\frac{1}{2}\hat{I}+\frac{1}{4}, 1/18)$.

The final estimation of $\mathbf{P}(p > 0.2)$ is then 0.76 and 0.59 respectively. According to Reuters, in Nov 2020, 68% of Republicans said they were concerned that the election was "rigged," while only 16% of Democrats and one-third of independents were similarly worried. However, in 2021 Nov, report shows that only 31% of the interviewees are still suspicious.

For reference,the Metaculus Community gave a almost consistent 80% confidence for this question and the question was resolved positively.

2.1.3 Comments

The model is extremely simplified. Better work can be done to have some hyper-parameters for p and more data should be used in prediction since the final outcome in this model is largely relied on our prior believe.

2.2 Simple Model for Cultured Meat

Here is the Metaculus link for this question. I now try to use a very simple and inaccurate model, and do a forecast based on the information provided by Wikipedia, keyword set as "cultured meat".

2.2.1 Methodology

In this question, we are asked to predict the probability that cultured meat will be served in a restaurant before 2021/1/1. We think of this question abstractly as "what is the possibility that a scientific invention can be brought to the public within certain time period", so as to find our reference class, "scientific inventions", biological scientific interventions or inventions regarded as a novel consumption good.

Further unveiling the requirements of a successful invention, we may break it down into small stages of breakthroughs and assume that, on average an invention will be successful after n_{break} times of small breakthroughs have occurred. For example, in the case of cultured meat, the first patent in 1991 and the first public trial on TV in 2013 should be counted. To get a rough list of all breakthroughs, most events included in History section on Wikipedia are counted(This of course, is just a rough list, since we does not even have a rigorous definition of breakthroughs. However, there are some potential ways to do this, i.e. list their search times on google, ask an expert).

2.2.2 Model Setting

For this breakthrough model, we can readily adopt a regular poisson process model, where the probability of a breakthrough happening in any time unit is a constant λ. A bayesian's approach is again used here to estimate and update our estimation of λ over time. $N(t)$ is defined as the total number of breakthroughs at time.

To summarize, Statistical Model is

$$\begin{aligned} Prior : \lambda &\sim \mathbf{Gamma}(\alpha, \beta) \\ N(t) &\sim \mathbf{Poi}(\lambda t) \\ \lambda | N(t) &\sim \mathbf{Gamma}(\lambda + N(t), \beta + t) \end{aligned}$$

Hyperparameter α and β are chosen such that the prior belief imitate a uniform distribution.

2.2.3 Result

If we acknowledge the time of breakthroughs are recognized to be 1991,2001,2003,2008,2013,2017,2018, 2019,2019,2020,2020(Without the one when cultured meat had its debut at Singapore), the model predictions the following probability at different combinations of n_break and t_p(the end of each year) which means the time of forecasting.

n_{break}	2016	2017	2018	2019	2020
7	0.82	1.00	1.00	1.00	1.00
8	0.57	0.76	1.00	1.00	1.00
9	0.35	0.49	0.69	1.00	1.00
10	0.20	0.29	0.43	1.00	1.00
11	0.12	0.18	0.28	0.64	1.00
12	0.09	0.14	0.21	0.42	1.00
13	0.07	0.12	0.19	0.33	0.66
14	0.07	0.11	0.18	0.30	0.55
15	0.07	0.11	0.18	0.30	0.52

Table 1: Predictions based on Poisson Model

Type	2016	2017	2018	2019	2020
Community	0.38	0.80	0.72	?	?
Metaculus	0.22	0.48	0.50	?	?

Table 2: Predictions based on Poisson Model

We can see that when n_{break} is set to be 9, 10 or 11, the prediction agrees best with the community.

2.2.4 Comments

Better work can be done to estimate hyperparameters(the ones we set manually, such as n_{break}). It would be possible if n_{break} itself is random(follow a normal distribution maybe, as suggested by Will in the last call on Feb10th). Mostly important, a stricter definition or model for breakthroughs should be discussed. The full timeline is not accessible on Metaculus, so how well the model truly performs is hard to say at this point.

3 Important Questions

In the previous section, some preliminary results are shown and the simplest statistical models can lead to overall satisfactory predictions.

In this section, I will continue explore the limit of this approach, starting by generalizing our model so that it is equipped with great flexibility to cover most of the forecasting questions. To achieve this, case studies on typical and important questions need to be done, in order that a few common traits will be revealed. I start with those shortlisted by this Lesswrong post. Notice that the problems are classified according to its **cause area** and **question type**.

3.1 Leader of China in 2030 (Great Power; Binary)

Here is the Metaculus link for this question. My table of data(Figure1) looks like includes key information such as length of reign and lifetime.

3.1.1 Methodology

In this question, we are asked to predict whether or not Xi Jinping will continue to be the leader of China in 2030. To tackle this problem, I have chosen a rather simple and straight forward idea of reference class. First retrieving data on political party and countries of similarities(Russian, Soviet Union apart from China). We can roughly think of each leaders' ruling time and their lifetime as **i.i.d** normal random variables. Then our task is to estimate the mean and variance. Then plugging in year 2030(or other time points as well) will give us the probability. Secondly, considering the different nature of Chinese and Russian political system, I only made use of the data about Chinese leader in the end, but it's worth trying to use a full data set(and the results will be reported). Thirdly, one adjustments was made on the final estimated Xi's mean year of reign: I add **5** to the simple average,

	Leader	Country	Start	End	Inpower	Born	Death
6	Andropov	Soviet	1982	1984	2	NA	NA
7	Chernenko	Soviet	1984	1985	1	NA	NA
8	Gorbachev	Soviet	1985	1991	6	NA	NA
9	Yeltsin	Soviet	1991	2000	9	NA	NA
10	Putin	Soviet	2000	2022	22	NA	NA
11	Maozedong	China	1949	1976	27	1893	1976
12	Huaguofeng	China	1976	1980	4	1921	2008
13	Dengxiaoping	China	1980	1991	11	1904	1997
14	Jiangzemin	China	1991	2004	13	1926	2022
15	Hu Jintao	China	2004	2012	8	1942	2022
16	Xi Jinping	China	2012	2022	10	1953	2022

Figure 1: Basic Information of Leaders

considering the huge impact made by reforming constitution and Xi's obvious distinction. **5** can be justified by his almost sure winning in the upcoming term. But I understand this can be a controversial point so some experiments are done without this modification.

3.1.2 Model Setting

To estimate $\mathbf{P}(Xistillin2030)$, I propose breaking it down into $\mathbf{P}(Inpower_{Xi} \geq 2030 - 2012) * \mathbf{P}(Life_{Xi} \geq 2030 - 1953)$(age and reign are presumably independent). These two can be further revised into a conditional probability, given he has already been the president since 2013 and is still alive.

To summarize, Statistical Model is

$$\begin{aligned} Prior : Inpower &\sim \mathbf{N}(\mu_{C\_Inpower}, \sigma^2_{C\_Inpower}) \\ Posterior : Inpower &\sim \mathbf{N}(\mu_{Xi\_Inpower}, \sigma^2_{Xi\_Inpower}) \\ Prior : Life &\sim \mathbf{N}(\mu_{C\_Life}, \sigma^2_{C\_Life}) \\ Posterior : Life &\sim \mathbf{N}(\mu_{Xi\_Life}, \sigma^2_{Xi\_Life}) \end{aligned}$$

and in specific, $\hat{\mu}_{Xi\_Inpower} = \hat{\mu}_{C\_Inpower} + C$, where C can be a constant(say 5), subject to our belief, $\sigma^2_{Xi\_Inpower} = \hat{\sigma^2}_{C\_Inpower}$.

3.1.3 Result

Agreeing on the ruling period of leaders, the model predicts the following probability. In C1 and C2, I used $C = 5$ versus $C = 0$ in C3 and C4. While in C1 and C2, the set of all Chinese leaders are considered as the reference class, only Mao, Deng and Jiang are included in C3 and C4's. Personally, I want to choose somewhere between C1 and C2.

1. C1 and C2 reflects the distinct present situation

2. C1 includes some moderate and unpopular leaders who are considered to be very different from Xi, Mao, Deng and Jiang

3. C2 has too small a dataset

Scenario	Prob	Scenario	Prob
C1	0.56	C3	0.41
C2	0.72	C4	0.56
9	0.35	0.49	0.69
10	0.20	0.29	0.43

Table 3: Predictions based on Different Assumptions

In conclusion, a simple average of C1 and C2 generates a probability of 0.64, which is very close to the community prediction on Metaculus of 0.60.

3.1.4 Comments

Better work can be done to blabla.

References

All the code used are available here(https://github.com/wjshku/Forecasting).