Combining Density Forecasts via Information Aggregation Mechanism

Discussion of Gillen, Plott, and Shum (2014)

Laura Liu

University of Pennsylvania

December 5, 2014

Background

- Forecast Combination
  - Model-based: variance-covariance, regression
  - Survey-based
  - Market-based
- More info -> better forecast: Wisdom of the crowds!

Summary

- Information Aggregation Mechanism (IAM)
  - Combine density forecasts
  - Innovative designs
- Theory
  - "A trip to Bayesland" (Nate Silver)
  - Density forecast: Dirichlet process
- Evaluation
  - Difficulty in density forecast
  - Problem: iid?

Information Aggregation Mechanism (IAM)

- Parimutuel-like betting mechanism (vs A-D security market)
  - +: Easy and intuitive to implement
  - -: Discrepency in the ideal and actual loss functions
- Fake money
  - +: Reduce the impact of risk aversion
- Non-tradable tickets
  - +: Reduce price speculation and transaction cost
- Tickets prices increase over time
  - +: Reduce information externality
  - -: Reduce information available
- Participants are chosen to be "insiders"
  - +: Reduce self-selection bias and noise trading
  - -: Maybe some "outsiders" would help too?
    - ★ More independent info set, wisdom of the crowds

Information Aggregation Mechanism (IAM)

- Between survey- and market- based forecast combinations
  - Provide money incentive
  - But with non-tradable tickets and less self-selection problem

Theory

- "A trip to Bayesland" (Nate Silver)
  - Individual:

$$\begin{aligned}\text{ind posterior} &= \text{common prior} + \text{ind private info}\\ &= \text{ind ticket placement}\end{aligned}$$

  - Aggregate:

$$\begin{aligned}\text{agg posterior} &= \text{common prior} + \text{sum of private info}\\ &= \text{agg ticket placement}\end{aligned}$$

  - More info -> better forecast
  - Bayesian updating vs invisible hand

Theory

- Density forecast: Dirichlet process
  - A set of bins
$$P(Y \in bin\ k) = \pi_k$$
  - Prior:
$$\pi \sim Dir(\alpha_1, \cdots, \alpha_K), \ E[\pi_k] = \frac{\alpha_k}{\sum_{j=1}^K \alpha_j}$$
  - Ind posterior = ind ticket placement:
$$\pi | s_n \sim Dir(\alpha_1 + m_n \hat{p}_{n,1}, \cdots, \alpha_K + m_n \hat{p}_{n,K})$$
$$E[\pi_k | s_n] = \frac{\alpha_k + m_n \hat{p}_{n,k}}{m_n + \sum_{j=1}^K \alpha_j}$$
  - Agg posterior = agg ticket placement:
$$\pi | s_{1:N} \sim Dir\left(\alpha_1 + \sum_{n=1}^N m_n \hat{p}_{n,1}, \cdots, \alpha_K + \sum_{n=1}^N m_n \hat{p}_{n,K}\right)$$
$$E[\pi_k | s_{1:N}] = \frac{\alpha_k + \sum_{n=1}^N m_n \hat{p}_{n,k}}{\sum_{n=1}^N m_n + \sum_{n=1}^N \sum_{j=1}^K \alpha_j}$$

Theory

- Some tests of the Bayesian updating in real world
  - "Belief Updating among College Students: Evidence from Experimental Variation in Information", Wiswall and Zafar (2011)
  - ...

Evaluation

- Q: whether it accurately reflects the uncertainty in sales?
  - H_0: As the forecasting horizon h decreases, density forecast should get closer to the true conditional distribution of $Y_t|\mathcal{F}_{t,t-h}$.
  - How to test?
- Difficulties in evaluating density forecast
  - the true conditional distribution of $Y_t|\mathcal{F}_{t,t-h}$ is not observable, even ex-post
  - heterogeneous info set for different period t

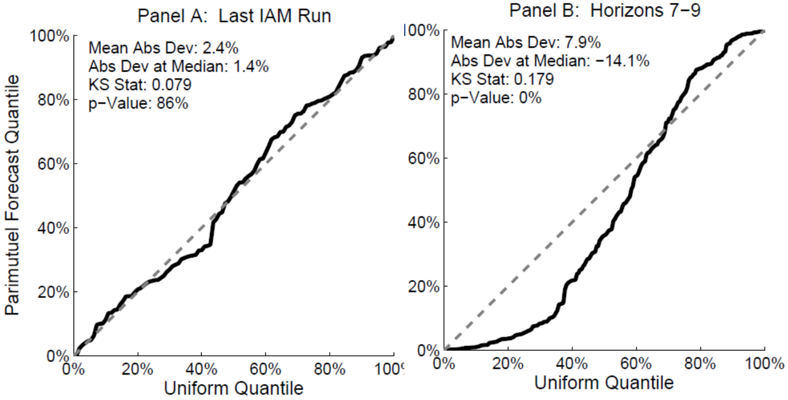
Evaluation

- Probablity integral transformation (1-step-ahead):
  - true conditional distribution of $Y_t|\mathcal{F}_{t,t-1}$: $f_t(y_t)$
  - density forecast of $Y_t|\mathcal{F}_{t,t-1}$: $p_t(y_t) = MN\left(\tilde{\eta}_{1|t-1}, \cdots, \tilde{\eta}_{K|t-1}\right)$
  - cdf of density forecast:

$$
\begin{aligned}
z_t &= \int_{-\infty}^{y_t} p_t(\tilde{y}_t)\, d\tilde{y}_t = P_t(y_t) \\
&\sim \frac{f_t\left(P_t^{-1}(z_t)\right)}{p_t\left(P_t^{-1}(z_t)\right)} \\
z_t &\sim iid\ U[0,1], \text{ if } p_t(\cdot) = f_t(\cdot)
\end{aligned}
$$

- QQ plot, KS test
  - Joint test of *iid* and $U[0,1]$
  - but for h-step-ahead forecast, *iid* would be violated...

Evaluation



Panel A: Last IAM Run
Mean Abs Dev: 2.4%
Abs Dev at Median: 1.4%
KS Stat: 0.079
p−Value: 86%
Parimutuel Forecast Quantile
100%
80%
60%
40%
20%
0%
0%
20%
40%
60%
80%
Uniform Quantile
Panel B: Horizons 7−9
Mean Abs Dev: 7.9%
Abs Dev at Median: −14.1%
KS Stat: 0.179
p−Value: 0%
100%
80%
60%
40%
20%
0%
0%
20%
40%
60%
80%
100%
Uniform Quantile

Evaluation

- h-step-ahead Forecast
  - Recall that optimal point forecast errors $MA(h-1)$
  - Similarly, if the density forecast is optimal, the z_t series would be $(h-1)$-dependent
  - And the sub-series will be *iid*: $\{z_1, z_{1+h}, z_{1+2h}, \cdots\}$, $\{z_2, z_{2+h}, z_{2+2h}, \cdots\}$, ...,$\{z_h, z_{2h}, z_{3h}, \cdots\}$

Evaluation

- Q: whether it accurately reflects the uncertainty in sales?
  - H_0: As the forecasting horizon h decreases, density forecast should get closer to the true conditional distribution of $Y_t|\mathcal{F}_{t,t-h}$.
- Cannot be directly compared via KS test due to serial correlation
- Maybe just compare the predictive likelihood

$$P_h = \prod_{t=1}^{T} p_{t|t-h}(y_t)$$

and P_h should decrease with h