A Wisdom of the Crowd Approach to Forecasting

Funded by the Intelligence Advanced Research Projects Activity (IARPA) via Department of Interior National Business Center contract number D11PC20059

Brandon Turner
and
Mark Steyvers

UC, Irvine

December 17th, 2011

UNIVERSITY of CALIFORNIA IRVINE

The Research

UCI is one member of Team ARA, along with six other universities:

The Research

We work together to

- Investigate good elicitation methods
- Build models that use this information to predict the future

The Research

- Everyday people log on to a website
- They make predictions about items (IFPs) they are interested in
- We record lots of data and analyze it

The Research

The goal is to beat MITRE, a data collection company, at making predictions

- MITRE uses the unweighted linear average on their own data
- Team ARA competes against four other teams to beat MITRE's ULinOP

The Research

The data comes in a variety of forms

- Binary IFPs
- Multi-Choice IFPs
- Continuous IFPs

UNIVERSITY of CALIFORNIA IRVINE

The Research

- We currently have over 50 models
- To evaluate them, we compare them to our own ULinOp
- We are now past the burn-in period

Outline

UNIVERSITY of CALIFORNIA IRVINE

Outline

UNIVERSITY of CALIFORNIA IRVINE

Motivation

The Wisdom of the Crowd Effect

- Groups of people make an estimate about a quantity
- The “correctness” of these participants will vary
- The mean of the estimates is better than the majority of the group

Motivation

WoC effects have been found in a variety of interesting problems

- Static judgments
- Rank-ordering tasks
- Event recall
- Scene reconstruction
- Combinatorial problems

Motivation

Can the WoC effect be harnessed to predict the future?

- Build on previous “shared truth” models
- Build on classic JDM confidence literature

Outline

Data

- 817 participants (general public)
- Provided estimates of the probability of the occurrence of future events
- 51 (binary) questions
- Judgments made over a one-month period

Complications

- At first, there are no known answers
- Questions are designed to eventually resolve
  - "Who will win the January 2012 Taiwan Presidential election?"
  - "By 1 January 2012 will the Iraqi government sign a security agreement that allows US troops to remain in Iraq?"
- 18 questions resolved during the one-month period
- Focused on binary items only

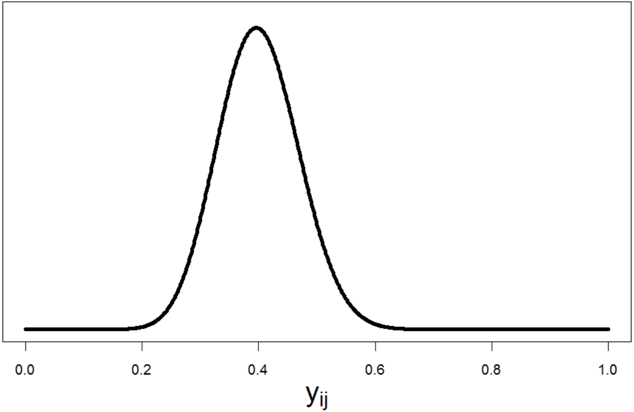
Outline

UNIVERSITY of CALIFORNIA IRVINE

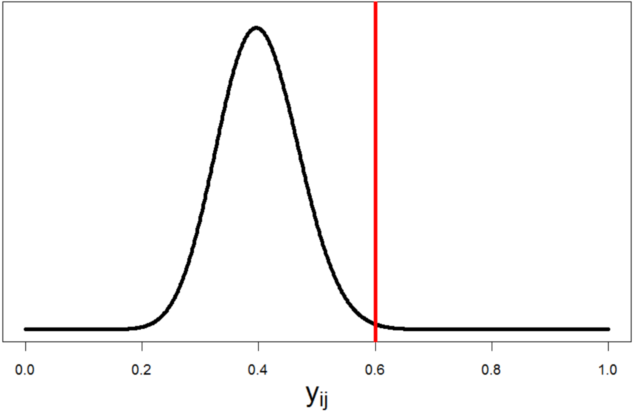
Modeling Approach

- Assume some latent shared truth (CCT)
- Model the aggregate of the judgments (WoC)
- Assume the shared truth is systematically inaccurate
- Assume a distortion occurs, prohibiting accurate forecasting
  - By Question
  - By Subject

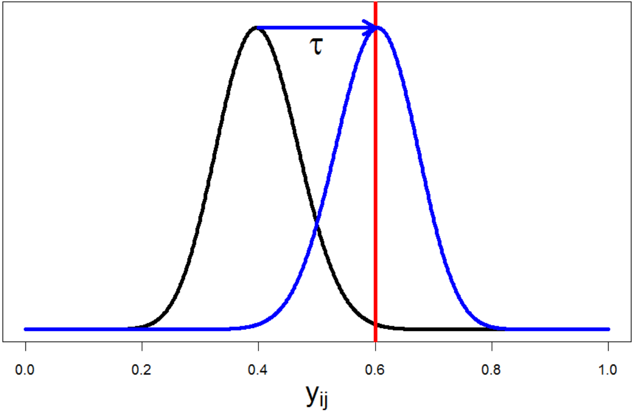
Modeling Approach



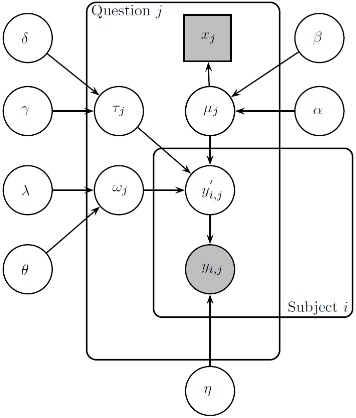
Modeling Approach

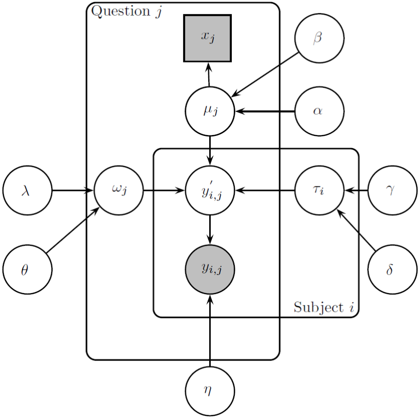


Modeling Approach



New Modeling Attempts



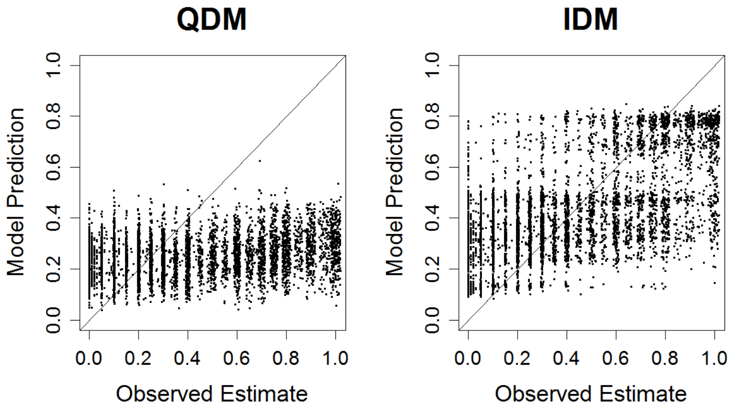


Outline

Results

- Distortion by Question
  - Performed 4.7% better than unweighted average
  - Mean predictive error was 0.337
- Distortion by Subject
  - Performed 9.6% better than unweighted average
  - Mean predictive error was 0.320

Posterior Predictive Distributions



Outline

Conclusions

- An accurate shared truth does not perform well
- A distorted version of the shared truth does well
- Distortion by subject is better than by question

Future (Current) Directions

- Exploit non-stationarity
  - Judgments might change over time and recent judgment might be more accurate; track opinions over time
- Recalibrate judgments
  - Recalibrate individual judgments before aggregating
  - Recalibrate the aggregate
- Exploit individual differences
  - Estimate expertise from resolved IFPs and user profiles
  - Match between user profile and IFP profile

Future (Current) Directions

- Model missingness
  - Incorporate information about the specific IFPs a user chooses to forecast, along with information about the number of IFPs that a user forecasts
- Supervised learning algorithms
  - Enter a large number of features in various supervised learning algorithms, determine which are related to individuals Brier scores
- Bayesian nonparametrics
  - Isolate subgroups of users with different forecasts/opinions, aggregate based on these subgroups