(1) Information aggregation—a statistical perspective

(2) Distinctions between different scenarios of group decision-making

Andrew Gelman, Roger Clemen, Roger Cooke, Jim Hammitt, David Krantz, and Francis Tuerlinckx

(1) A statistician's perspective on information aggregation

- Bayesian data analysis
- Public opinion, elections, policy
- How this differs from what you do
- A paradox of Bayesianity
- Challenges of working with expert judgments

Examples

- Election forecasts given national opinion polls and political conditions
  - Regression model calibrated using election data with national, regional, statewide errors
- Decision making for home radon
  - National and state survey data
  - Geological predictors
  - Lac Qui Parle County (very high estimate, but based on only two measurements)
- Combining several sources of information
- Hierarchical regression models

Bayesian data analysis differs from what's been said at this workshop

- The truth is rarely binary
  - We don't think about "false positives" etc.
  - Death sentences
- Instead of "experts," we have numerical data
- The data model is typically measurable and clear

A Bayesian wants everybody else to be non-Bayesian

- It's easy to combine several unbiased estimates
- Linear regression
- Also, more complicated models
  - (e.g., y = x*b + e, with x and y both measured with error)
- For example, national polls, economic and political conditions, and state polls

A Bayesian doesn’t want to deal with other Bayesians

- Need to “subtract out” the prior distributions
- Simplest to deal with separate unbiased estimates
- Consider Laq Qui Parle County
- “You give me the data, I’ll do the adjusting”

Motivating sincerity

- How to get experts to give honest, unbiased estimates?
- Proper scoring rules, moral hazards, . . .
- Can we motivate a norm of zero bias?
- What about a norm of moderate, known bias?
- Downweighting extremes

One of the challenges of aggregating expert judgments

- We don’t know what we *should* be doing!
- Need a model for the experts’ statements, given the true “parameter values”

Model of unbiased judgments

- True parameter value is T_j
- Expert i's estimate is y_ij
- Examples:
  - % of people in the U.S. who are "black"
  - Age guessing
- Unbiased: E(y_ij | T_j) = T_j

  or E(y_ij | T_j) = T_j + b_j (bias)
- y_ij = T_j + b_j + e_ij (bias + idiosyncratic error)

Model of calibrated judgments

- Recall unbiasedness: $E(y \mid T) = T$
  - For all questions where T=0.20, the average guess will be 0.20
- Calibration: $E(T \mid y) = y$
  - For all guesses of y=0.20, the average true value will be 0.20
- Unbiasedness and calibration are not the same thing (Wallsten etc)
- People are typically underconfident
  - If a true value T is 0.1, people typically guess values like 0.2

and overconfident
  - If someone gives you an estimate of 0.2, the true value is probably closer to 0.2

Incentives for motivating calibration and unbiasedness

- If your goal is to minimize Sum $((y_{ij} - T_j)^2)$, you should be calibrated
- Can be done using feedback:
  - For each question j, your information is I_{ij}
  - With enough data, you can figure out the empirical average value of T_j as a function of I_{ij}
  - These are calibrated estimates and minimize your loss
- Incentive system to motivate unbiased estimates?

Incentives for unbiasedness?

- Mechanisms for unbiased estimation:
  - Direct measurement ($y = T + b + error$)
  - Inverting the Bayes procedure:
    - Create calibrated estimates (using feedback)
    - Use past info to get prior distribution, $p(T)$
    - Inflate the calibrated estimate to get unbiased est
- Is there a direct incentive?

A new norm and payoff structure?

- Is there a reasonable and realistic norm between unbiasedness and calibration?
- If so,
  - can we enforce this norm with a payoff structure?
  - how can we combine these judgments in statistical inference?

How does this apply to aggregation?

- The pre-calibration debate:
  - Clemen: recalibrate each expert, then combine (using some “combination rule”)
  - Cooke: take experts at their word, then use a weighted average
- What are reasonable models for experts’ statements?

Summary of part 1

- Use statistical data combination as a template for combining personal judgments
- Bayesian inference is not about “false positives” etc.
- A Bayesian wants everybody else to be non-Bayesian
- How can we get people to give us data we can easily and reliably combine?

(2) Distinctions between different scenarios of group decision-making

- Most of this conference has been about comparing different *rules* for making a group decision
- Rules include statistical modeling, numerical averaging, majority rule, pick the best decision maker, repeated voting, …
- But not much talk about the different *scenarios* of group decision making
- Our claim: the scenario matters when evaluating aggregation rules

Disturbed by blurring

- We are disturbed by the blurring of distinctions among the following:
  1. combining *information* (as in perception and estimation tasks)
  2. combining *attitudes* (as in national elections)
  3. combining *interests* (as in competitive games and distributive politics)

Three different group decision scenarios

- In each scenario, several people are getting together to make a single decision
- They can have similar or divergent *goals*
- They can have different *information*, *attitudes*, or *goals*
- 3 scenarios:
  1. Inference
  2. Difference of opinion
  3. Conflict of interest

Scenario 1: inference

- The different people have a common *goal* but dispersed *information*
- Examples:
  - Memory
  - Perception
  - Military simulation tasks

Scenario 2: difference of opinion

- The different people have different *attitudes*
- Examples:
  - National elections

Scenario 3: conflict of interest

- The different people have different *goals*
- Examples:
  - Competitive games
  - Business negotiation

Demarcation points

- Real situations often have features of more than one of the pure scenarios
- For example, a jury trial fits into Scenario 1 (inference), but different jurors have different attitudes (e.g., liberal or conservative) which puts the problem somewhat into Scenario 2
- Nonetheless, we will clarify the distinctions between the scenarios by defining demarcation rules

Scenario 1 (inference) or 2 (difference of opinion)?

- Demarcation between scenarios 1 and the others: Can you easily persuade me with data?
- If Yes, then it's an inference problem (and it would be optimal for *you* as well as me to share all your information with me)
- If No, then it's a difference-of-opinion or conflict-of-interest problem (and there is little reason for you to share info with me)

Scenario 2 (difference of opinion) or Scenario 3 (conflict of interest)?

- Demarcation between scenario 3 and the others: Would a side payment solve this problem?
- If Yes, then it’s a conflict of interest problem
- If No, then it’s an inference or difference-of-opinion problem (imagine the absurdity of saying, “if you give me $50, then I’ll change my estimate of theta”)

Example of the scenarios

Consider the strategy of *hiding information*

- In scenario 1 (inference), this is a *cognitive error* (it actually hurts *you* if you hide information from me). We want to train people not to do it.
- In scenario 2 or 3 (difference of opinion or conflict of interest), hiding info is *strategic behavior* that can help you but have negative consequences for the group. We want to construct decision rules that acknowledge that players will be strategic, or adjust the incentive structure so that they will be sincere.

Summary of part 2

- Decision aggregation rules have different implications in different scenarios
- Voting is a combination rule, but inferences are not elections
- Demarcation points:
  - Can you easily persuade me with data?
  - Can a side payment solve the problem?