Socially-inclusive foundations of statistics

Dan J. Spitzner

September 9, 2020

human and machine intelligence group

meetings: wednesdays 12:15 – 1:30 in wilson 142

FALL 2017 SCHEDULE

Wed 18-Oct	Dan Spitzner(Statistics)	Why analyze data? How variety in the objectives of data analysis points to complementary roles for statistics and data science. [**slides**]

- Alignment (and realignment?) between practice and worldview
- Criticism of the **scientific worldview:** emphasis on control and for exclusionary effects
- **Socially-inclusive research:**

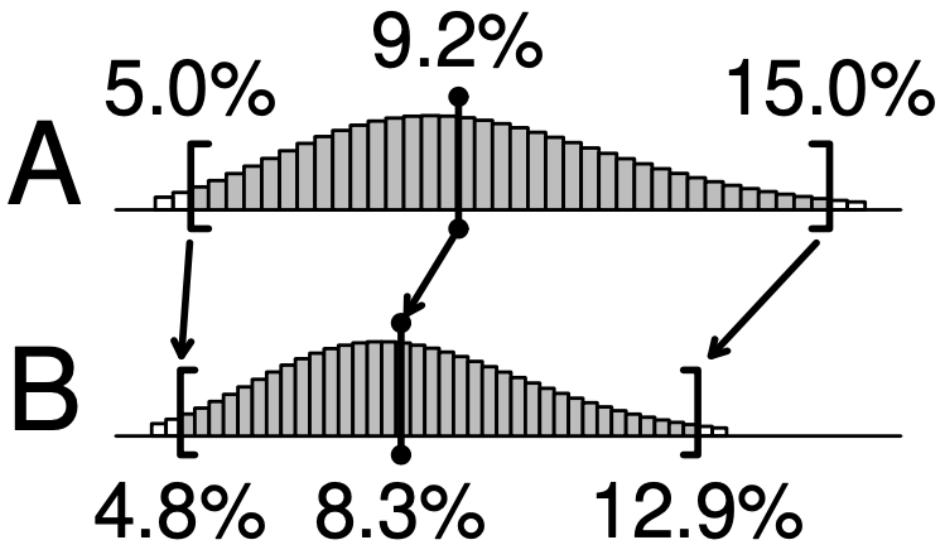
  > "unity, cohesion, civic engagement, togetherness, or bridging the gap between 'us' and 'the other'."
  >
  > — Koikkalainen (2011)

- **Goals:** Ignite creativity, diversify my discipline, open new collaborations

- **Community elicitation:** knowledge is made meaningful through a social process
- **Bayesian analysis:** A story of transformation

  STEP 1: State probabilistically what is known
  STEP 2: Look to newly-measured data
  STEP 3: Update what is known

- Prior knowledge is transformed into posterior knowledge



- **Elicitation:** bringing forth knowledge, especially prior knowledge.
- **Traditional Bayesian elicitation:** individualistic, knowledge is in the minds of experts.
- **Community elicitation:** knowledge resides in community

Traditional Bayesian elicitation:

- Tversky and Kahneman: Human judgement of uncertainty
- **Heurstics:** approximate strategies that use only limited information
- **Biases:** predictable violations of probability theory
- Goal: reduce the elicitation “bias”

Community elicitation:

- Mercier and Sperber: “argumentative” theory of reasoning

  “Reasoning has evolved and persisted mainly because it makes human communication more effective and advantageous.”

- Main function of reasoning is not to “enhance individual cognition” or correct the mistakes of intuition
- Even isolated individuals are woven into community.

Traditional Bayesian elicitation:

- Individualistic process of eliciting knowledge from an expert's mind

Community elicitation:

- Procession of **arguments** and **counterarguments**, accompanied by proposed descriptions of knowledge
- Goal of an individual researcher is to formulate an argument, which could rest on sources other than experts

Statistical inference:

- **Bayesianism:** regarded as a normative theory of statistical inference
- **Frequentism:** Logically flawed; dominates statistical practice
- **"Objective" Bayes:** Enormously popular; negates normativity

Observation:

- Tenets of positivist-type scientific worldview dictate what is allowed to flourish

First steps of a new project:

- Identify communities that implement statistical practice under a socially-inclusive worldview
- **Mixed methods research:** intersection between qualitative and quantitative methodologies

> "...actively invites [us] to participate in dialogue...multiple ways of seeing and hearing, multiple ways of making sense of the social world, and multiple standpoints on what is important and to be valued and cherished."
>
> — Greene (2007)

Integrating qualitative and quantitative methodologies:

POSITION 1: Incompatibility: qualitative and quantitative methods are inherently incompatible and cannot be integrated

POSITION 2: Dialectical: qualitative and quantitative approaches can be put into dialogue while respecting their interpretive distinctions

POSITION 3: Pragmatism: “what works”

Pragmatism (**of a certain flavor**) is most widely adopted

- Emphasizes expediency
- Falls short of informing substantial decision-making
- Hinders integration

Motivations and effects:

- Adopted to “secure funding for their research interests and publish their findings.” — Bryman (2007)
- Marginalization of qualitative research
- Lack of integration having become accepted $\Rightarrow$ a majority of mixed methods studies “use the analytic and prescriptive style of positivism.” — Giddings (2006)

Integrating qualitative and quantitative methodologies:

POSITION 1: Incompatibility...

POSITION 2: Dialectical...

POSITION 3: Pragmatism...

POSITION 4: Realign statistical practice to a socially-inclusive worldview

Blueprint: Bayesian analysis using community elicitation

- Prior knowledge bridges to positionality
- Shifts perspective to a deeper level of meaning
- Gathering prior knowledge and uncovering positionality hold potential to inform each other, set up a seamless integration of methodology.

Statistical reporting:

- **Pool reduction** formalizes an idea expressed in the law literature
- Substitute for (Bayesian versions of) things like p-values

Example: Thompson (2018)

- "One-in-a-million" DNA profile
- In the U.S. perhaps 300 people share the profile
- To further reduce the pool, look beyond forensic analysis

Pool reduction emphasizes. . .

- Responsibilities of decision-making
- Actual stakes of inquiry

Tasks:

- Improve ease-of-access to Bayesian methodology
  - Lower-level methods that would be used routinely
  - Advanced techniques applied with consulting statistician
- Clarify a qualitative understanding of statistical methodology
- Work toward holistic collaborations