Bayes Methodology: Mixed Methods in Drag

Isadore Newman

Florida International University

Anthony J. Onwuegbuzie

Sam Houston State University

John H. Hitchcock

Indiana University

Paper presented at the First Annual meeting of the Mixed Methods International Research Association, June 2014, Boston, MA

Purpose

- To describe how Bayesian Statistical Inference can be logically connected to principles of Mixed Methods Research.
  - Discuss connection via philosophical aspects of Bayesian work and qualitative inquiry.
  - Describe how Bayesian analysts can benefit from a good command of qualitative techniques.
  - We plan to work out more specific guidance on integration over time.
- Disclaimer: right now we feel more comfortable of philosophical than procedural grounds.
  - We’re working on the *how to* part.
  - Later on we discuss reflexivity and part of this entails self-disclosure; in that spirit, we’re not Bayesian methodologists (yet)…or are we?

Let’s begin by reviewing classical statistical inference

- The classical approach conceptualizes theoretical distributions (normal, uniform, binomial, Poisson, etc.) that characterize the results of, say, an experiment many times (e.g., 1,000).
- We of course don’t run these experiments, but instead rely on what the distribution of their results would be **if** the null hypothesis is true.
- We then set up an actual experiment, and see how likely it is that our results are compared to what we would expect if the null is true.
- If our observed result would be unusual (say chances are less than 1 in 20 if the null is true), then we reject the idea that the null is true.
- Recall that we never know for sure if we’re making a correct decision (there are Type 1 and 2 errors, we go all the way up to Type 6 – and you can ask Isadore about this idea).
- This approach assumes no prior information about what we expect to find, which is generally not true.

Bayesian Inference

- For now it is important to understand that the Bayesian work entails using *informed priors* (or Bayesian Priors), rather than acting as if we have no information.
- Such information requires making judgments (often by making subjective assumptions) about context.
- The idea is to use available information to help make better predictions.
- There are many classes or types of Bayesian Priors, but they generally involve making informed judgment about related context and have the potential to improve prediction.
  - They do no yet account for any collected data.
- We'll work through an example later on.

Broad elaborations on Bayesian compared to Classical Inference

- Bayesian and classical/frequentist approaches (i.e., hypothesis testing and confidence intervals) both entail making inferences about a population based on collected, or observed data.
- Both approaches deal with the degree of uncertainty (think of error) analysts have when making inferences.
- Both rely on probability theory.
- But how probability is thought about differs across the two approaches.

With that background in mind, allow a general overview of some qualitative ideas

- We are making a case for connections between Bayesian and Qualitative Inquiry…
- Just to warm up, recall that general qualitative inquiry is often described by using words like:
- Emergent;
- Naturalistic;
- Interpretive.

  (e.g., Brantlinger, Jiminez, Klingner, Pugach, & Richardson, 2005; Creswell, 2009; Johnson & Christensen, 2012; Onwuegbuzie & Leech, 2007; Patton, 2002)

Some elaborations about qualitative work…

- Naturalistic inquiry means that qualitative research designs generally attempt to study phenomenon in typical settings, without contriving or manipulating circumstances (although one can do so).
- Emergent: refers to the idea that qualitative researchers adapt questions and methods as their understanding of constructs, context and so on improves.
- Interpretive: among other things, interest in how people construct realities and construe events.

- We can go on but our point is to discuss Bayesian work and mixed methods research, and not offer a primer on qualitative approaches.

Can Bayesian work benefit from Qualitative Inquiry?

- We see a fit.
- Consider some quotes drawn from Iversen (1984):

  "The two types of probabilities relevant here can be called empirical [classical] and *subjective* [emphasis added] probabilities" (p.8).

  "The subjective view takes probabilities as personal measures of uncertainty, based on available evidence" (p.9).

  "Subjective probabilities are more general than empirical probabilities since they can also be used to measure uncertainty we have about single, unique events" (p.9).

  "Bayesian analysis is cumulative" (p. 16).

On Bayesian priors…

- “It can be claimed that research is never done in a vacuum, and if nothing were known about a parameter we would not have thought of doing the research in the first place” (p. 62).
- “A prior distribution expresses the analyst's opinion of a population parameter before any new data are available” (p. 66).
- “Bayesian Inference…reflects the very human fact that if two people have different opinions about something, then even the same limited empirical facts will modify these opinions differently” (p. 67).
- “But the prior distributions bring the subjective aspects of the analysis out in the open for everyone to see” (p.67).
- “…there is no reason why we should shy away from our personal choice of a more informative distribution if we have information that warrants such a choice” (p. 68).

Some Clarifications

- O'Hagan & Luce (2003) point out that subjective probabilities are not arbitrary, but orderly and informed by prior data and need to be defensible.
- They also note that over the long run, Bayesian and Frequentist approaches yield similar results although they are interpreted differently.
- But they still argue that science cannot be truly objective. We add that qualitative reasoning, while allowing for subjectivity, should in general be equally defensible and transparent (Hitchcock & Newman, 2012; Newman & Hitchcock, 2011; Newman, Hitchcock & Nastasi, 2013).

Some parallels and connections between Bayesian Theory and Qualitative Inquiry

- Research is always situated in context; it is not done in a vacuum.
  - This connects to the idea of naturalistic inquiry.
- If one has a strong understanding of context, prediction (and thus probability) are better understood.
  - Pick your qualitative method for understanding context.
- Subjectivity is fine, even encouraged in a way, so long as it is made plain.
  - Connect back to notion of reflexivity, analyzing implicit assumptions, etc.
- Emergent design is not *winging it*; the approach owns up to a rigorous account of what one does and does not know.
  - See the parallel to Bayesian inference based on cumulative evidence.

Or more visually…

Qualitative Inquiry	Parallels	**Bayesian Inference**
Naturalistic	<- ->	Context
Interpretive	<- ->	Subjectivity
Emergent	<- ->	Cumulative
Member Checking	<- ->	Intermittent Checks
Reflexivity	<- ->	Make assumptions plain

How does Bayesian inference essentially work?

- The basic idea, expressed as formula, relies on algebra where there are in essence three known (or at least estimated) values and one unknown one.
- From Silver (2012):
  - X = prior probability
  - Y = probability of a conditional event
  - Z = probability of the event not occurring
- This yields the posterior probability:
- XY/XY+Z(1-X)

Alternatively

Probability of (P|D)*Prob (P)/Prob (D)

Where:

P=Population

D=Data

Iversen (1984, p. 12)

Suppose we have three communities and we want to know from which one sample data were drawn

One person is selected. She happens to be Catholic.

So which community is she from?

Without knowing anything, we might guess about 0.33 probability from community one (P1), two (P2) or three (P3).

But what if we had some more information?

P1 = .30 Catholic, P2 = .50 Catholic and P3 = .70 Catholic

Three Posterior Probabilities using the above formula:

Initial Prior is 0.33

Prob (P1|D) = 0.3(0.33)/0.3(0.33)+0.5(0.33)+0.7(0.33) = 0.099/0.4995 = <u>0.20</u>
Prob (P2|D) = 0.5(0.33)/0.4995 = <u>0.33</u>
Prob (P3|D) = 0.7(0.33)/0.4995 = <u>0.47</u>

We're more likely to say the person is from P3

Cumulative Inference: Now suppose we get more data…

- N=10; 80% are Catholics
- Use the new observation + the old posteriors become the new priors.
- Yields:

- Prob (P1|Data) = 0.002
- Prob (P2|Data) = 0.115
- Prob (P3|Data) = 0.883

- Highly unlikely the sample is drawn from community 1; probably from community 3

From Iversen (1984, pp. 12-16)
See also Silver (2012)

A key connection

- <u>How</u> can one have some confidence that community 1 has fewer Catholics, than Community 2, and Community 3 has the most number of Catholics?

- That is, how can we develop informed priors?

Drumroll…

Qualitative Inquiry can help

- An operating idea is this requires contextual knowledge, which can be won with qualitative methods.
- The credibility of such knowledge can be promoted with techniques like member checks, triangulation, prolonged engagement, reflexivity, etc.
- Both Bayesian and Qualitative Inquiry (really any research) can be improved via intermittent checks on the veracity of findings and assumptions.

A larger picture

- We see several parallels between common ideas in qualitative inquiry and Bayesian conceptualizations.
- This discussion is emblematic of the idea that good research is based on solid and transparent logic, defensible conclusions, and checks on intermittent and overall findings.
- Our opening joke is that Bayesian methodology is “mixed methods in drag.” We think this because we view mixed methods as an exercise in using solid and transparent logic, and verification of findings (again Hitchcock & Newman, 2013; Newman & Hitchcock, 2012).

References

Brantlinger, E., Jiminez, R., Klingner, J., Pugach, M., & Richardson, V. (2005). Qualitative studies in special education. *Exceptional Children, 71*, 195-207. doi:10.1177/001440290507100205

Creswell, J. W. (2009). *Research Design: Qualitative, quantitative, and mixed methods approaches*, (3rd Edition). Thousand Oaks, CA: Sage.

Hitchcock, J.H., & Newman, I. (2013). Applying an interactive quantitative-qualitative framework: How identifying common intent can enhance inquiry. *Human Resources Development Review, 12(1),* 36-52. dOI: 10.1177/1534484312462127

Iversen, G.R. (1984). *Bayesian statistical inference*. Sage.

Johnson, R. B., & Christenson, L. (2012). *Educational research: Quantitative, qualitative, and mixed approaches* (4th ed.). Thousand Oaks, CA: Sage.

Newman, I. & Hitchcock, J.H. (2011). Underlying agreements between quantitative and qualitative research: The short and tall of it all. *Human Resources Development Review, 10*(4), 381-398. doi: 10.1177/1534484311413867

Newman, I. Hitchcock, J.H., & Nastasi, B.K. (2013, April). Using phenomenological principles to assess stakeholder perceptions of probability: Applying mixed methods to translational research. 2013 annual meeting of the American Education Research Association, San Francisco, CA.

O'Hagan, A., & Luce, B.R. (2003). *A primer on Bayesian Statistics in health economics and outcomes research*. Bethesda, MD: MEDTAP International.

Onwuegbuzie, A. J., & Leech, N. L. (2007). Validity and qualitative research: An oxymoron? *Quality & Quantity, 41*, 233-249. doi:10.1007/s11135-006-9000-3

Patton, M. Q. (2002). *Qualitative research and evaluation methods*. Thousand Oaks, CA: Sage.

Silver, N. (2012). *The Signal and the noise: Why so many predictions fail-but some don't*. New York, NY: The Penguin Press.