Discussion of: "Intrinsic tests for the equality of two correlated proportions" by Guido Consonni and Luca La Rocca

Mario Peruggia, The Ohio State University

Supported by
NSF SES-0437251 and DMS-0605052

The Data

- Paired binary data, e.g. presidential support before and after the State of the Union address

		After		
		No	Yes	Total
Before	No	n_{00}	n_{01}	n_{0+}
	Yes	n_{10}	n_{11}	n_{1+}
	Total	n_{+0}	n_{+1}	n_{++}

The Question

- Did the proportion of voters in favor of the president change?

The Question

- Did the proportion of voters in favor of the president change?

- Testing problem
  Null: Before prop. $=$ After prop.
  Alternative: Before prop. $\neq$ After prop.

The Models

- Null is equivalent to $\pi_{01} = \pi_{10}$
  (marg. hmg. $\iff$ symm. in a 2×2 table)

The Models

- Null is equivalent to $\pi_{01} = \pi_{10}$
  (marg. hmg. $\iff$ symm. in a 2×2 table)

- Collapse $(0, 0)$ and $(1, 1)$ cells because π_{00} and π_{11} do not appear in the hypothesis.

The Models

- Null is equivalent to $\pi_{01} = \pi_{10}$
  (marg. hmg. $\iff$ symm. in a 2×2 table)

- Collapse $(0,0)$ and $(1,1)$ cells because π_{00} and π_{11} do not appear in the hypothesis.

- Define $n_\Delta = n_{01} + n_{10}$.
  Model reduces to product of two binomial prob's:
  $P(n_{01}|n_\Delta, \theta) \sim Bin(n_\Delta, \theta), \quad \theta = \pi_{01}/(\pi_{01} + \pi_{10})$
  times
  $P(n_\Delta|n_{++}, \eta) \sim Bin(n_{++}, \eta), \quad \eta = \pi_{01} + \pi_{10}$

Two Analyses (plus one)

- Null becomes $H_0 : \theta = 1/2$

Two Analyses (plus one)

- Null becomes $H_0 : \theta = 1/2$
- Conditional Intrinsic (CI) analysis:
  Based on $P(n_{01}|n_\Delta, \theta)$

Two Analyses (plus one)

- Null becomes $H_0 : \theta = 1/2$
- Conditional Intrinsic (CI) analysis:
  Based on $P(n_{01}|n_\Delta, \theta)$
- Intrinsic (I) analysis:
  Based on $P(n_{01}|n_\Delta, \theta) \times P(n_\Delta|n_{++}, \eta)$

Two Analyses (plus one)

- Null becomes $H_0 : \theta = 1/2$
- Conditional Intrinsic (CI) analysis:
  Based on $P(n_{01}|n_\Delta, \theta)$
- Intrinsic (I) analysis:
  Based on $P(n_{01}|n_\Delta, \theta) \times P(n_\Delta|n_{++}, \eta)$
- Naive Intrinsic (NI) analysis:
  Based on $P(n_{01}|n_\Delta, \theta) \times P(n_\Delta|n_{++}, \eta)$
  with imaginary data perfectly consistent with H_0

Items for Discussion

- η is called a *nuisance* parameter

Items for Discussion

- η is called a *nuisance* parameter
- Yet, this parameter is what will eventually determine the size, and hence the relevance, of the swing

Items for Discussion

- η is called a *nuisance* parameter
- Yet, this parameter is what will eventually determine the size, and hence the relevance, of the swing
- Compare with the difference between *absolute* risk and *relative* risk in epidemiological studies

Items for Discussion

- Is collapsing of the diagonal cells justified?

Items for Discussion

- Is collapsing of the diagonal cells justified?
- Question: Do n_{00} and n_{11} separately carry any information about θ?

Items for Discussion

- Is collapsing of the diagonal cells justified?
- Question: Do n_{00} and n_{11} separately carry any information about θ?
- Arguably they might. E.g., if n_{11} is much larger than n_{00} then, given that a swing occurs, the swing will more likely be against the president than in in favor of the president

Items for Discussion

- The I approach produces an intrinsic prior where η and θ are dependent

Items for Discussion

- The I approach produces an intrinsic prior where η and θ are dependent
- Dependence is such that the "effective support" of θ, always centered around $1/2$, becomes more concentrated as η gets closer to 1

Items for Discussion

- The I approach produces an intrinsic prior where η and θ are dependent
- Dependence is such that the "effective support" of θ, always centered around $1/2$, becomes more concentrated as η gets closer to 1
- Is this a desirable property?
  Perhaps, if η is large, more extreme values of θ, in either direction, should be weighted more heavily

Items for Discussion

- A big swing will typically be associated with a major occurrence in the address that will sway most viewers in one specific direction

Items for Discussion

- A big swing will typically be associated with a major occurrence in the address that will sway most viewers in one specific direction
- E.g., President unveils plan to ...

Items for Discussion

- A big swing will typically be associated with a major occurrence in the address that will sway most viewers in one specific direction
- E.g., President unveils plan to ...
- Smaller swings, on the other hand, will tend to go in either direction

Items for Discussion

- How to assess the performance of the intrinsic methodology?

Items for Discussion

- How to assess the performance of the intrinsic methodology?
- Examine the features of the intrinsic prior and see if the resulting model is adequate

Items for Discussion

- How to assess the performance of the intrinsic methodology?
- Examine the features of the intrinsic prior and see if the resulting model is adequate
- Possibly modify some modeling decisions

Items for Discussion

- How to assess the performance of the intrinsic methodology?
- Examine the features of the intrinsic prior and see if the resulting model is adequate
- Possibly modify some modeling decisions
- In this application one could, for example, avoid collapsing the diagonal cells and use a prior other than the "default" Dirichlet distribution.