Causally Sound Priors for Binary Experiments

Nicholas J. Irons and Carlos Cinelli (University of Washington)

CENTER for STATISTICS and the SOCIAL SCIENCES

CENTER FOR STUDIES IN DEMOGRAPHY & ECOLOGY

NIH *Eunice Kennedy Shriver* National Institute of Child Health and Human Development

Physicians' Health Study (1989)

Does low-dose aspirin reduce risk of fatal heart attack?

- Large-scale randomized placebo-controlled trial
- 22,000 healthy male physicians randomly assigned to receive 325 mg aspirin or placebo every other day
- 26 of 11,034 prescribed placebo experienced fatal myocardial infarction; 10 of 11,037 prescribed aspirin did.
- Estimated risk ratio (θ_1/θ_0): 0.39 [0.19, 0.80], ***rejects the null*** with P-value 0.008.
- **Frequentist: *yes!***

Main Bayesian approaches for comparing proportions:

- Independent beta (IB) approach using beta(1,1) priors:
  - Bayes factor BF_{01} = 20.27
  - ***Strong*** evidence in favor of the null (H_0: $\theta_0 = \theta_1$)
- Logit transform (LT) approach using N(0,1) priors:
  - Bayes factor BF_{10} = 5.24
  - ***Moderate*** evidence in favor of alternative (H_1: $\theta_0 \neq \theta_1$)
- **Bayesian: **shrug****

Binary experiments

Statistical setup:

- $N = N_0 + N_1$ denotes the number of subjects in the study
- Z_i is a binary treatment (Z=1, treat.),
- Y_i is a binary outcome (Y=1, death);
- $Y_i(z)$ is the potential outcome of subject *i* under experimental condition *z*
- **Assumption 1 (*Consistency*):** $Y_i = Y_i(Z_i)$.
- **Assumption 2 (*Super-population*):** Subjects drawn IID from a common population.
- **Assumption 3 (*Ignorability*):** $\{Y_i(0), Y_i(1)\} \perp\!\!\!\perp Z_i$
- Under A1-3, observed counts are independent binomials:

$$y_0 = \sum_{i:Z_i=0} Y_i \sim \text{Binomial}(N_0, \theta_0) \quad \perp\!\!\!\perp \quad y_1 = \sum_{i:Z_i=1} Y_i \sim \text{Binomial}(N_1, \theta_1)$$

where we define the ***baseline risk*** and ***risk of treatment***,

$$\theta_0 = \mathbb{P}(Y_i(0) = 1), \quad \theta_1 = \mathbb{P}(Y_i(1) = 1)$$

Independent Beta (IB) approach: $\theta_0 \sim \text{Beta}(a_0, b_0) \quad \perp\!\!\!\perp \quad \theta_1 \sim \text{Beta}(a_1, b_1)$

- Pro: trivial estimation, posterior simulation, Bayes factors
- Con: knowledge about risks in each study arm assumed independent *a priori* (and therefore *a posteriori*)

Logit transform (LT) approach: $\log\left(\frac{\theta_0}{1-\theta_0}\right) = \beta - \frac{\psi}{2}, \quad \log\left(\frac{\theta_1}{1-\theta_1}\right) = \beta + \frac{\psi}{2}$

- Pro: encodes dependence across treatment arms
- Con: not intuitive and not analytically tractable

Baseline Risk, Efficacy, and Adverse Side Effects (BREASE)

BREASE parametrization:

- We define the *efficacy* of treatment as the probability that treatment is sufficient to save or cure a patient:
$$\eta_e = \mathbb{P}(Y_i(1) = 0 | Y_i(0) = 1)$$
- Similarly, the *risk of adverse side effects* is the probability that treatment causes death of an otherwise healthy patient:
$$\eta_s = \mathbb{P}(Y_i(1) = 1 | Y_i(0) = 0)$$
- Applying the law of total probability, we can decompose the risk of treatment in terms of the BREASE parameters as:
$$\theta_1 = \theta_0(1 - \eta_e) + (1 - \theta_0)\eta_s$$

Partial identification:

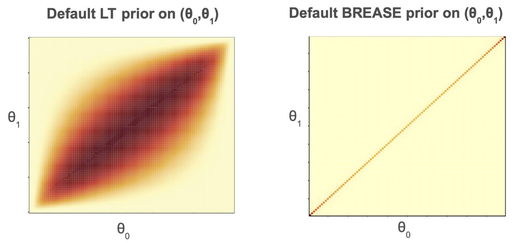
- The counterfactual parameters η_e and η_s are only partially identified by the observed data,

$$\max\left\{0, 1 - \frac{\theta_1}{\theta_0}\right\} \leq \eta_e \leq \min\left\{\frac{1-\theta_1}{\theta_0}, 1\right\}, \quad \max\left\{0, 1 - \frac{\theta_1 - \theta_0}{1-\theta_0}\right\} \leq \eta_s \leq \min\left\{\frac{\theta_1}{1-\theta_0}, 1\right\}$$

- However, (θ_0, θ_1) are still identified in this model.
- Placing priors on (η_e, η_s) induces sensible priors on (θ_0, θ_1) respecting scientific knowledge of the problem.

Prior specification:

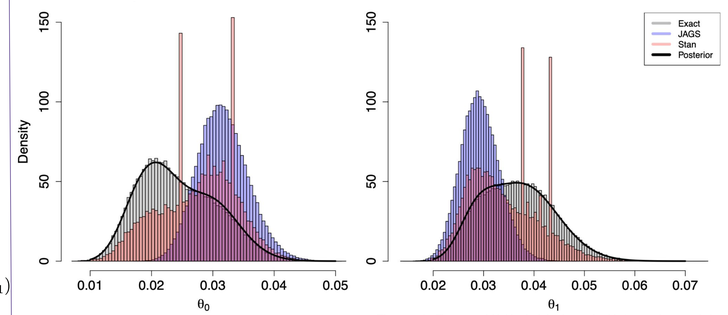
- We propose independent beta priors on the BREASE parameters:
$$\theta_0 \sim \text{Beta}^*(\mu_0, n_0) \quad \perp\!\!\!\perp \quad \eta_e \sim \text{Beta}^*(\mu_e, n_e) \quad \perp\!\!\!\perp \quad \eta_s \sim \text{Beta}^*(\mu_s, n_s)$$
- This induces a *Generalized* Dirichlet prior on the response types (Dickey, 1983).
- We propose the BREASE(½, μ, μ; 2, 1, 1) as a default. It has the following properties:
  - puts flat uniform marginals on θ_0 and θ_1 (c.f. IB)
  - induces prior correlation between θ_0 and θ_1 (c.f. LT)
  - assumes no effect on average (c.f. IB and LT)
  - depends on a single interpretable parameter μ.



Fast and exact posterior sampling, inference, sensitivity analysis

Posterior sampling: MCMC is not bulletproof

JAGS and Stan struggle to sample from the posterior in certain pathological examples, e.g., prior-data conflict.
Our exact posterior sampling algorithm and data-augmented Gibbs sampler have no issues.

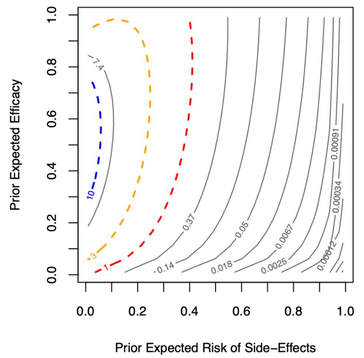


Automatic posterior inference:

We derive closed-form expressions for marginal likelihoods and Bayes factors used for Bayesian model selection and hypothesis testing

Theorem 3. *The marginal likelihood of the data under M_0 is given by a simple beta-binomial distribution. Under M_1, it is given by the following weighted sum of beta functions*[8]

$$L_1(\mathcal{D}) = \binom{N_0}{y_0}\binom{N_1}{y_1}\sum_{j=0}^{y_1}\sum_{k=0}^{N_1-y_1}\binom{y_1}{j}\binom{N_1-y_1}{k} \times \frac{B(y_0+j+k+\mu_0 n_0, N-(y_0+j+k)+(1-\mu_0)n_0)}{B(\mu_0 n_0, (1-\mu_0)n_0)} \times \frac{B(k+\mu_e n_e, j+(1-\mu_e)n_e)}{B(\mu_e n_e, (1-\mu_e)n_e)} \times \frac{B(y_1-j+\mu_s n_s, N_1-y_1-k+(1-\mu_s)n_s)}{B(\mu_s n_s, (1-\mu_s)n_s)}. \quad (3.17)$$



Prior sensitivity of the BF (Aspirin study)

Strong evidence against the null only when:

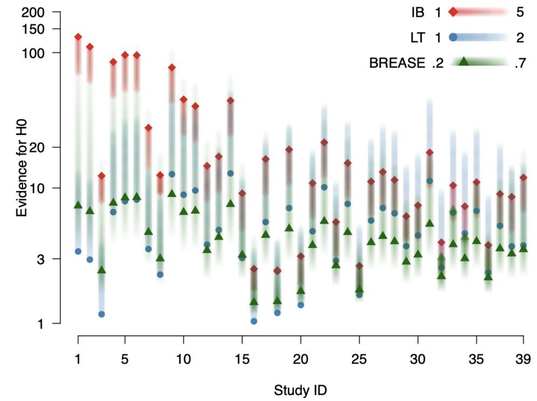
- adverse side effects are expected to be small (< 1%), and
- efficacy is expected to be relatively large (between 30% and 70%).

The results of the trial are ambiguous, and the conclusion that aspirin prevents fatal heart attack strongly depends on the prior.

Bridging the gap

A Puzzle of Proportions (Dablander et al., 2022)

BREASE reconciles disparate results reported in a Bayesian reanalysis of 39 null results in the *NEJM*. BREASE Bayes factors interpolate between the IB approach (pointing toward the null) and the LT approach (more conservative).



Conclusions

We introduce the BREASE framework for the analysis of binary experiments. Our approach has a number of desirable characteristics when compared to mainstream alternatives:

- Induces prior dependence between risks in the treatment and control groups in a causally principled way;
- Admits analytical formulae for the Bayes factor and exact posterior sampling;
- Formulated in terms of clinically meaningful parameters;
- Facilitates elicitation of prior knowledge and sensitivity analysis.

As a transparent and simple Bayesian method that elucidates the sensitivity of inferences to modeling assumptions, our proposal can

- help analysts distinguish robust from fragile findings;
- clarify what one needs to believe in order to claim that a treatment is effective;
- reconcile disparate results obtained from different methods.

Acknowledgements

Research partially supported by a Shanahan Endowment Fellowship and a Eunice Kennedy Shriver National Institute of Child Health and Human Development training grant, T32 HD101442-01, to the Center for Studies in Demography & Ecology at the University of Washington.