Relative Belief

Demonstrating the feasibility of automating experimental bias evaluations within probabilistic programming frameworks

Feifan Liu and Javier Mencia
Co-Supervisors: Mike Evans and Scott Schwartz

Summer 2024

Table of Contents

1. Introduction to Relative Belief

Relative Belief

$$RB(A|B) = \frac{Pr(A|B)}{Pr(A)}$$

$$\begin{cases} > 1 & \text{B gives evidence } \textbf{in favor} \text{ of A} \\ < 1 & \text{B gives evidence } \textbf{against} \text{ A} \\ = 1 & \text{A and B are independent} \end{cases}$$

Analytic Approach:
Using Posteriors and Priors,

$$RB(\theta|x) = \frac{p(\theta|x)}{p(\theta)}$$

Simulated Approach:
Using Marginal Likelihoods,

$$RB(\theta|x) = \frac{p(x|\theta)}{p(x)}$$

where $\quad p(x) = \int p(x|\theta)p(\theta)d\theta$

2.Hypothesis Assessment vs Estimation

- **H: Hypothesis assessment** of θ_0, where the "a posteriori strength of evidence" can be given by the posterior probability

$$\begin{aligned}
\Pr(\ RB(\theta_0|x) \geq RB(\theta|x) \mid x) &= \int 1_{RB(\theta|x) \leq RB(\theta_0|x)}(\theta) p(\theta|x) d\theta \\
\text{Analytic Approach} &= \int 1_{\frac{p(\theta|x)}{p(\theta)} \leq \frac{p(\theta_0|x)}{p(\theta_0)}}(\theta) p(\theta|x) d\theta \\
\text{Simulated Approach} &= \int 1_{\frac{p(x|\theta)}{p(x)} \leq \frac{p(x|\theta_0)}{p(x)}}(\theta) p(\theta|x) d\theta
\end{aligned}$$

- **E: Estimation** where $\hat{\theta} = \sup_\theta RB(\theta|x)$ belongs to the **plausible region (**Pl**)**

$$Pl_\theta(x) = \{\theta : RB(\theta|x) > \overset{\text{or } c}{1}\} = \{\theta : \Pr(\theta|x) > \Pr(\theta)\}$$

which naturally shrinks as the posterior concentrates (with increasing data) and can be given a credible interval level.

3.1 Bias Calculations: Hypothesis assessment

Bias Against (Type I):

$$p_\alpha = \Pr(\ RB(\theta_0|x) \leq 1 \mid \theta_0) = \int 1_{RB(\theta_0|x)\leq 1}(x)p(x|\theta_0)dx$$

$$\text{Analytic Approach} = \int 1_{p(\theta_0|x)\leq p(\theta_0)}(x)p(x|\theta_0)dx$$

$$\text{Simulation Approach} = \int 1_{p(x|\theta_0)\leq p(x)}(x)p(x|\theta_0)dx$$

Bias In Favor (Type II):

$$p_{1-\beta} = \sup_{\theta:d(\theta,\theta_0)\geq\delta} \Pr(\ RB(\theta_0|x) \geq 1 \mid \theta)$$

$$= \sup_{\theta:d(\theta,\theta_0)\geq\delta} \int 1_{RB(\theta_0|x)\geq 1}(x)p(x|\theta)dx$$

where $d(\theta, \theta_0) \geq \delta$ denotes the "meaningful" difference between θ and θ_0 (as opposed to statistically significant difference).

4.1 Beta Binomial Model

Set-up

Prior: $\theta \sim Beta(\alpha_0, \beta_0)$
Likelihood: $x = (x_1, \ldots, x_n) \in \{0, 1\}^n$, $x_i \in Ber(\theta)$
Posterior: $\theta \sim Beta(\alpha_0 + \overline{x}, n - \overline{x} + \beta_0)$

Beta binomial model from Evans (2016)

- $n = 20$
- $\alpha_0 = 4; \beta_0 = 4$
- $\theta_0 = 0.5$

Since this is a priori bias, there is no need to look at observed data. Simulated data is generated using the above structure with each run of the model

Michael Evans (2016), Measuring statistical evidence using relative belief, Computational and Structural Biotechnology Journal, https://doi.org/10.1016/j.csbj.2015.12.001

4.1.2 Beta Binomial Model: Bias Against (Type I)

Evidence against θ_0, given that $\theta_0 = 0.5$ is true,
Analytic Result: p $= 0.265$

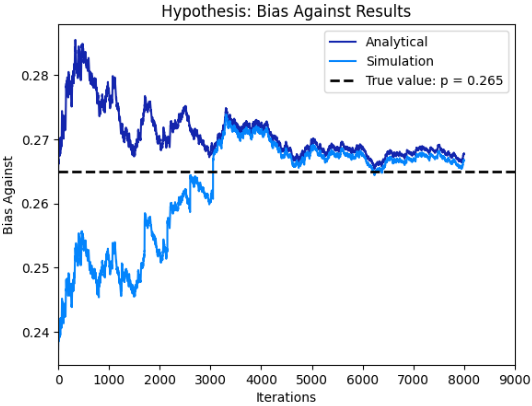


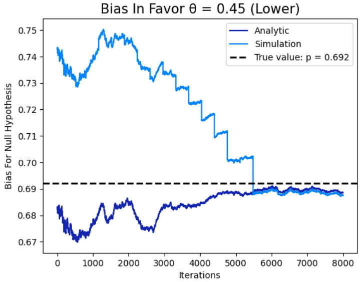
Figure: Results for Bias Against (Type I) for Simulated and Analytic Approaches

4.1.3 Beta Binomial Model: Bias In Favor (Type II)

Let $\sigma = 0.05$ be meaningful difference
Find maximum probability of finding evidence in favor of θ_0, when θ_0 is meaningfully wrong

Analytic Result: $\text{p} = 0.692$

This is a one-dimensional problem and the bias in favor (Type II) supremum occurs at points δ away from the parameter, so only consider:
$\theta_0 - \delta$ (Lower Value) and $\theta_0 + \delta$ (Upper Value)



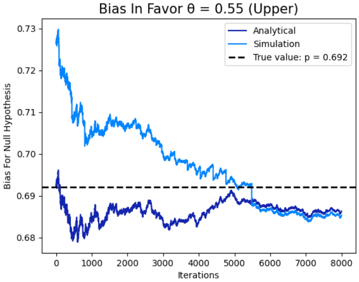


Figure: Results for Bias In Favor (Type II) for Simulated and Analytic Approaches

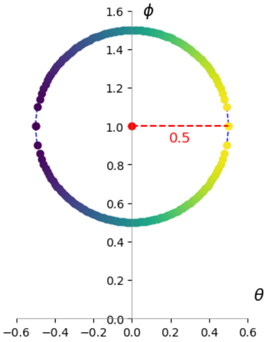
4.2 Two-Parameter Model

Set-up

Data:: $x = (x_1, \ldots, x_n)$, $x_i \in N(\theta, \phi)$
Prior: $\theta \sim N(\theta_0, \tau)$; $\phi \sim Gamma(\alpha_0, \beta_0)$
Likelihood: $x_i \sim N(\theta, \phi)$
Posterior:
$p(\theta|x, \theta_0, \tau, \phi) = \mathrm{N}(\frac{(\tau\theta_0 + \phi\sum_{i=1}^{n} x_i)}{(\tau + n\phi)}, \sigma^{-2} = \tau + n\phi)$
$p(\phi|x, \alpha, \beta, \theta) = \mathrm{Gamma}(\alpha + \frac{n}{2}, \beta + \frac{1}{2}\sum_{i=1}^{n}(x_i - \theta)^2)$

- $n = 20$
- hyper-parameter:$\alpha_0 = 1; \beta_0 = 1; \theta_0 = 0; \tau = 1$
- $\theta = 0; \phi = 1$

4.2.2 Two-Parameter Model - Bias in favor (Type II)



Formula for bias in favor:

$$p_{1-\beta} = \sup_{\theta: d(\theta,\theta_0)\geq\delta} \Pr(\ RB(\theta_0|x) \geq 1 \mid \theta)$$

$$= \sup_{\theta: d(\theta,\theta_0)\geq\delta} \int 1_{RB(\theta_0|x)\geq 1}(x) p(x|\theta) dx$$

How to Pick our circle

- **L2 norm**
  $(\phi_{upper_i} - \phi_0)^2 + (\mu_i - \mu_0)^2 = 0.5^2$
- **code:**
  for i in range(-50, 51):
  $\mu_i = \mu_0 + \frac{i}{100}$

$\phi_{upper_i} = \phi_0 + (0.25 - \frac{i^2}{10000})^{0.5}$

$\phi_{lower_i} = \phi_0 - (0.25 - \frac{i^2}{10000})^{0.5}$

4.3.1 Linear Regression Model

Set-up

Prior: $p(\boldsymbol{\beta}) = \mathcal{MVN}(E[\boldsymbol{\beta}] = \beta_0, \text{Cov}[\beta] = \Sigma_\beta)$
$\quad\text{p}(\ \sigma^2) = \text{InvGamma}(\alpha = \alpha_0, \beta = \beta_0)$

Likelihood: $\text{y}_{n\times 1} \sim \mathcal{MVN}(\mathcal{X}_{n\times p}\beta_{p\times 1}, \Sigma_{n\times n} = \sigma^2 I_{n\times n})$

$$\textbf{Posterior:}\ p(\beta \mid \Sigma, \mathcal{X}, y) = \mathcal{MVN}\left(\overbrace{\text{Cov}[\beta \mid \Sigma, \mathcal{X}, y]\left(\mathcal{X}^\top \Sigma^{-1}\mathbf{y} + \Sigma_\beta^{-1}\beta\right)}^{\text{E}[\beta|\Sigma,\mathcal{X},y]=}, \right.$$

$$\left. \text{Cov}[\beta|\Sigma,\mathcal{X},y]= \left(\mathcal{X}^\top \Sigma^{-1}\mathcal{X} + \Sigma_\beta^{-1}\right)^{-1}\right)$$

- $n = 10,\ p = 2,\ y_{n\times 1} = X\beta + \epsilon$, where X is fixed
- $\beta_{true} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\Sigma_\beta = I_{p\times p}$
- $\delta_{true} = 1$ with $\alpha_0 = 3, \beta_0 = 2$ and $\Sigma = I_{n\times n}$

4.3.2 Linear Regression Model: Bias In Favor (Type II) Approaches

Points to consider:

- Should all parameters use the same δ?
- What should each δ be?
- What is the relationship between both parameters? (Choice of Metric)

<u>Some examples for various δs:</u> (0.05, 0.15, 0.5 and 1)

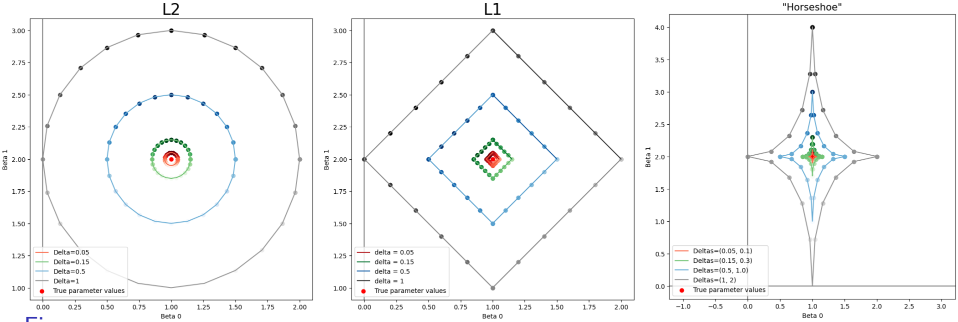
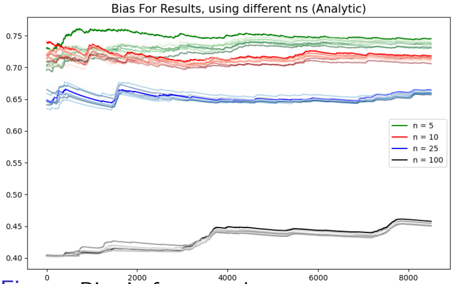


Figure: The first two graphs use the same δ for both parameters with $\delta_{\beta_0} = \delta_{\beta_1} = \delta$. In the third graph, $\delta_{\beta_0} = \delta_0$ and $\delta_{\beta_1} = 2 \times \delta_0$.

4.3.4 Linear Regression Model: Bias In Favor (Type II) Results

Both biases are decreasing functions with larger n's when δ is fixed

Does the same happen for larger δ's when n is fixed?



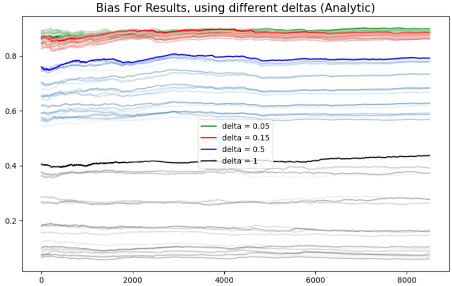


Figure: Bias in favor results.

Graph on the left shows bias in favor for different n's (10 lines for each n) using $\delta = 0.05$.
Graph on the right shows results for every point of each circle (25 lines for each δ) in first figure in previous slide using n = 10

Thank You!

We would like to extend our heartfelt thanks to:

- Professor Scott Schwartz
- Professor Michael Evans
- Qiaoyu Liang (PhD Candidate)
- Siqi Zheng (Graduate Student)
- Matthew Pronyshyn (Undergraduate Student)

Feifan and Javier