Bayesian Variable Selection Method for Modeling Dose-Response Microarray Data Under Simple Order Restrictions

Bayes2013, Rotterdam

Martin Otava

I-Biostat, Universiteit Hasselt

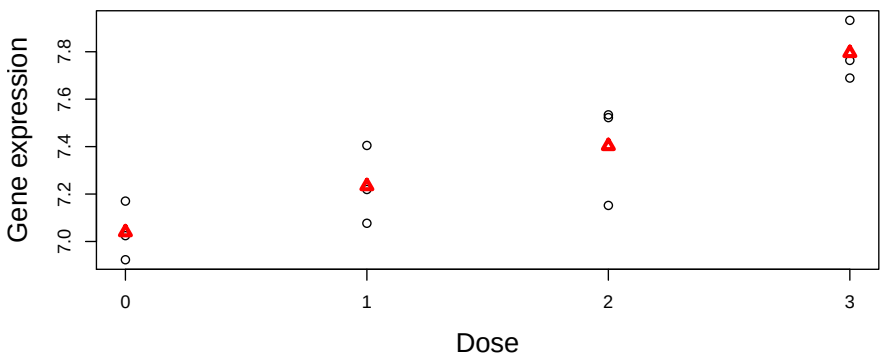
22.05.2013

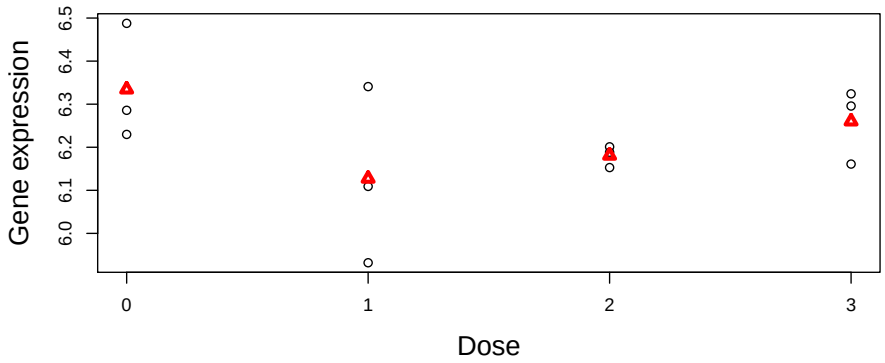
Research team

- Hasselt University (Belgium):
  - Martin Otava
  - Ziv Shkedy
- Durham University (UK):
  - Adetayo Kasim
- Imperial College London (UK):
  - Bernet Kato
- Zoetis (Belgium):
  - Dan Lin
- Janssen Pharmaceutica (Belgium):
  - Luc Bijnens
  - Hinrich W.H. Göhlmann
  - Willem Talloen

Dose-response modeling

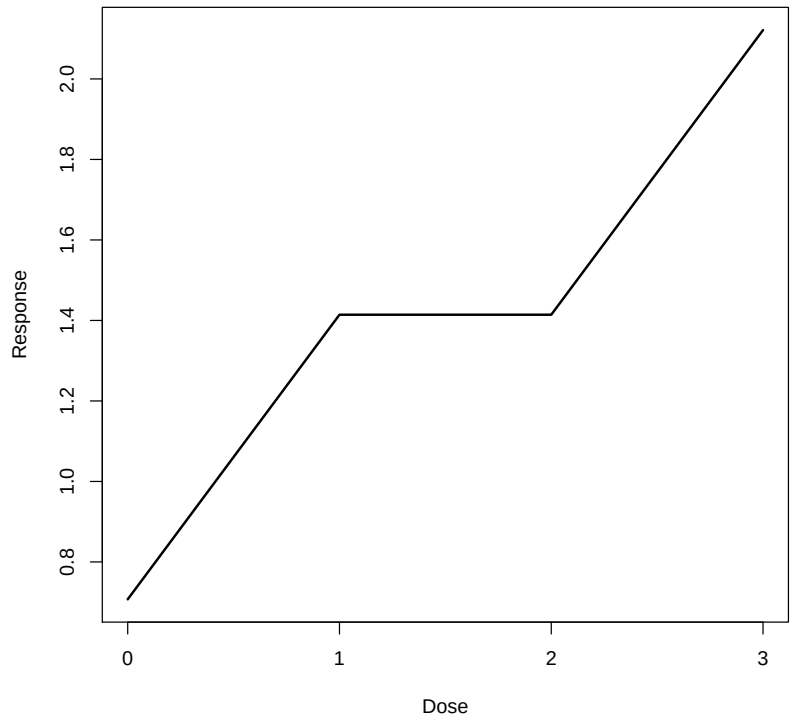
- Increasing dose of therapeutical compound.
- Variety of possible responses:
  - Toxicity.
  - Inhibition or stimulation.
  - Gene expression level.
- Goal:
  - Determine if there is any relationship.
  - If so, what is the shape of the profile.
  - Select threshold doses (e.g. MED).





Order constraints

- Compound effect becomes stronger when dose is increased.
- Monotone restriction (non-decreasing or non-increasing).
- Zero effect is meaningful.
- No parametrical assumptions about dose-response curve shape.



Basic Model

- One-way ANOVA model formulation:

$$Y_{ij} = \mu_i + \varepsilon_{ij} \qquad i = 0, \ldots, K-1$$
$$\varepsilon_{ij} \sim N(0, \sigma^2) \qquad j = 0, 1, 2, \ldots, n_i$$

  $\Longrightarrow$ necessary to incorporate order constraints.

- Testing the hypothesis

$$H_0 : \mu_0 = \mu_1 = \mu_2 = \ldots = \mu_{K-1}$$

  against ordered alternative (one inequality strict)

$$H^{up} : \mu_0 \leq \mu_1 \leq \mu_2 \leq \ldots \leq \mu_{K-1}$$
$$H^{dn} : \mu_0 \geq \mu_1 \geq \mu_2 \geq \ldots \geq \mu_{K-1}$$

Reformulation of model

- New notation (non-decreasing trend):

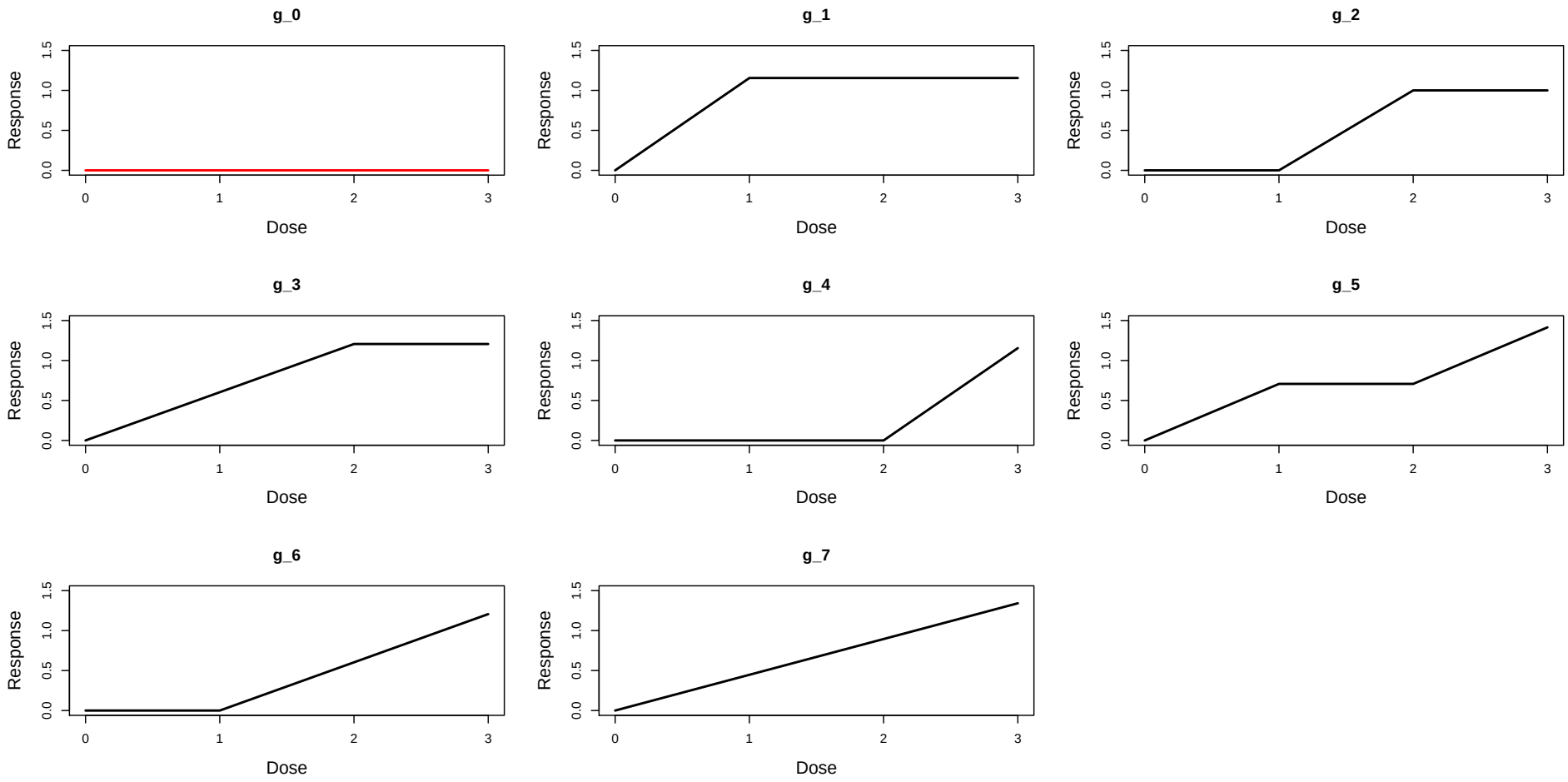
$$E(Y_{ij}) = \mu_i = \begin{cases} \mu_0, & i = 0, \\ \mu_0 + \sum_{\ell=1}^{i} \delta_\ell, & i = 1, \ldots, K-1 \end{cases}$$

- with priors:

$$\begin{aligned} \mu_0 &\sim N(\eta_\mu, \sigma^2_\mu), \\ \delta_i &\sim N(\eta_{\delta_i}, \sigma^2_{\delta_i}) I(0, A), \quad i = 1, \ldots, K-1. \end{aligned}$$

$$\Rightarrow \quad \delta_i \geq 0.$$

Set of all models



Sub-hypotheses

$$H^{up} : \mu_0 \leq \mu_1 \leq \mu_2 \leq \ldots \leq \mu_{K-1}$$

Model	Up: Mean Structure	**z**
g_0	$\mu_0 = \mu_1 = \mu_2 = \mu_3$	(0,0,0)
g_1	$\mu_0 < \mu_1 = \mu_2 = \mu_3$	(1,0,0)
g_2	$\mu_0 = \mu_1 < \mu_2 = \mu_3$	(0,1,0)
g_3	$\mu_0 < \mu_1 < \mu_2 = \mu_3$	(1,1,0)
g_4	$\mu_0 = \mu_1 = \mu_2 < \mu_3$	(0,0,1)
g_5	$\mu_0 < \mu_1 = \mu_2 < \mu_3$	(1,0,1)
g_6	$\mu_0 = \mu_1 < \mu_2 < \mu_3$	(0,1,1)
g_7	$\mu_0 < \mu_1 < \mu_2 < \mu_3$	(1,1,1)

Modification to BVS

- The distribution of δ is continuous.

  $\Longrightarrow$ probability of all models except one equals zero!

- Instead of only sampling δ_i we need to select which δ_i occurs in model.

- Let be z_i indicator of δ_i occurring in the model.

$$z_i = \begin{cases} 1, & \delta_i \quad \text{is included in the model,} \\ 0, & \delta_i \quad \text{is not included in the model.} \end{cases}$$

$$\Rightarrow \quad E(Y_{ij}) = \mu_0 + \sum_{\ell=1}^{i} z_\ell \delta_\ell.$$

BVS model formulation

- Basic model:

$$Y_{ij} \sim N(\mu_i, \sigma^2)$$

- Modeling of mean:

$$E(Y_{ij}) = \mu_i = \mu_0 + \sum_{\ell=1}^{i} z_\ell \delta_\ell.$$

- Priors:

$$\begin{aligned} \mu_0 &\sim N(\eta_\mu, \sigma_\mu^2), \\ \delta_i &\sim N(\eta_{\delta_i}, \sigma_{\delta_i}^2) I(0, A), \\ z_i &\sim \text{Bernoulli}(\pi_i), \end{aligned}$$

- Hyper Priors:

$$\begin{aligned} \sigma^{-2} &\sim \Gamma(10^{-3}, 10^{-3}), \\ \eta_\mu &\sim N(0, 10^6), \\ \sigma_\mu^{-2} &\sim \Gamma(10^{-3}, 10^{-3}), \\ \eta_{\delta_i} &\sim N(0, 10^6), \\ \sigma_{\delta_i}^{-2} &\sim \Gamma(10^{-3}, 10^{-3}). \\ \pi_i &\sim \text{U}(0, 1). \end{aligned}$$

Posterior mean of μ_i

- Posterior distribution for all dose-specific means.
- Use posterior mean of such distribution as our estimation.
- Connection of Bayesian model averaging.

  $\Longrightarrow$ posterior model probabilities are weights.

$$\hat{\boldsymbol{\mu}}_{BVS} = \sum_{r=0}^{R} w_r \hat{\boldsymbol{\mu}}_r$$

Posterior probability of model

- Vector $\mathbf{z} = (z_1, \ldots, z_{K-1})$ uniquely defines the model.
- Transformation $G(\mathbf{z}) = 1 + \sum_{i=1}^{K-1} z_i \, 2^{i-1} \Longrightarrow$ unique value for each model.
- In each MCMC iteration we sample one vector $\mathbf{z} = (z_1, \ldots, z_{K-1})$.
- Posterior mean of indicator $G(\mathbf{z}) = r + 1$ translates into posterior probability of the model g_r.

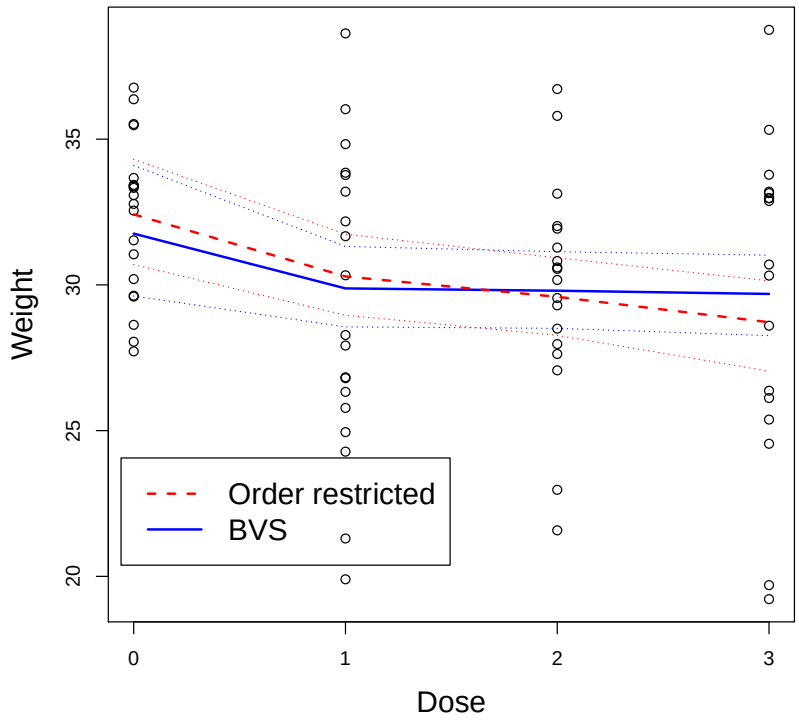
  $\Longrightarrow$ For posterior probabilities holds:

$$P[G(\mathbf{z}) = r + 1|\text{data}] = P(g_r|\text{data}).$$

Example: BVS model

- Incorporating models with equal means results into less decreasing profile.
- Posterior means are averages of means of particular models at each MCMC iteration.

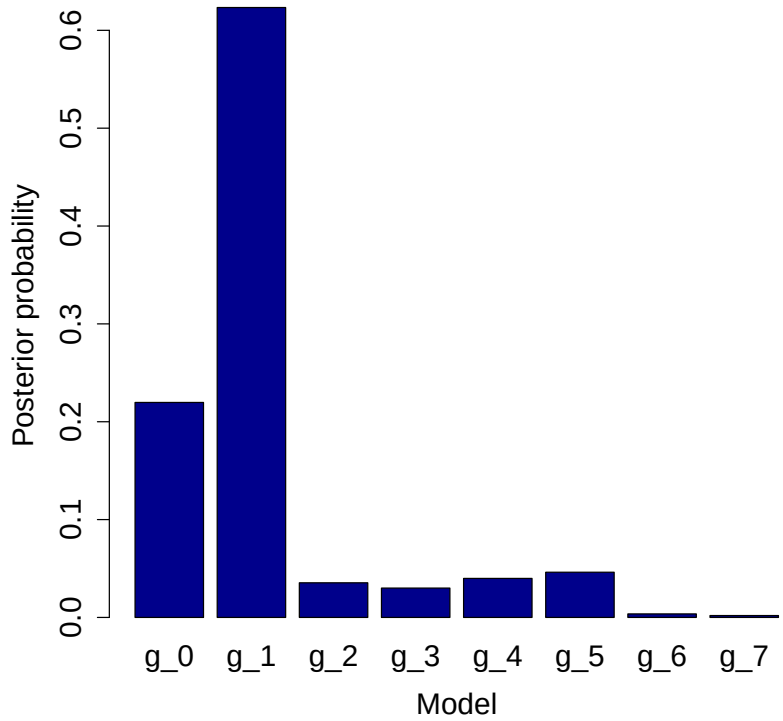
$$\hat{\boldsymbol{\mu}}_{BVS} = \sum_{r=0}^{R} \bar{P}(g_r|\text{data})\hat{\boldsymbol{\mu}}_r$$



- Connection to model averaging.

Example: Posterior probabilities

- Posterior probabilities of particular models.
- Model g_0 represents H_0.
- Model g_1 is strongly supported by the data.
- Connection to model selection.



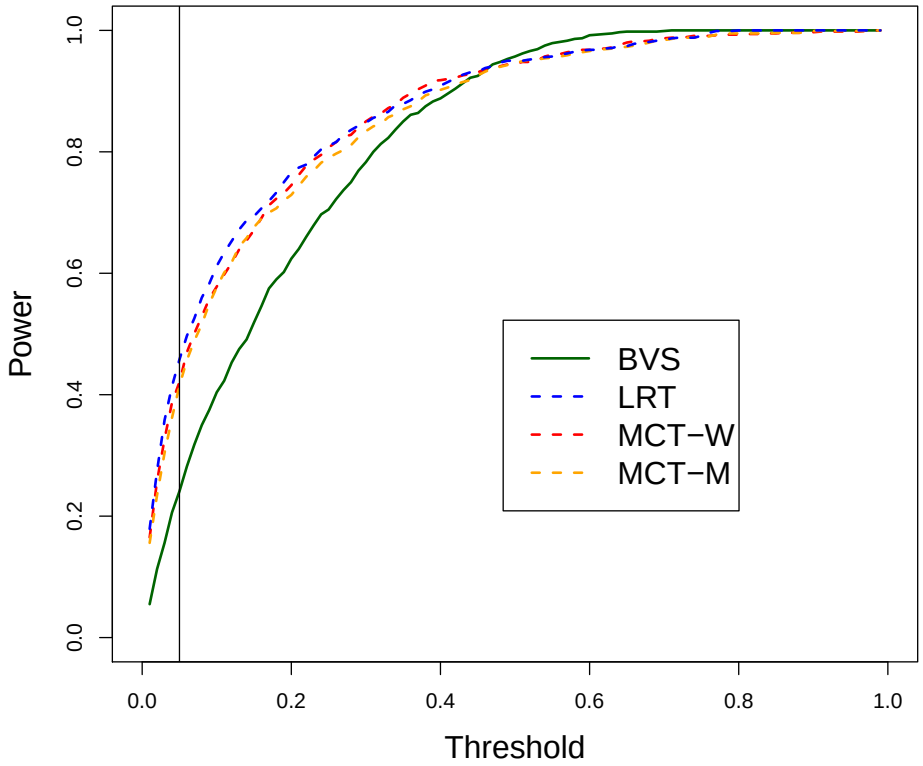
Hypothesis testing

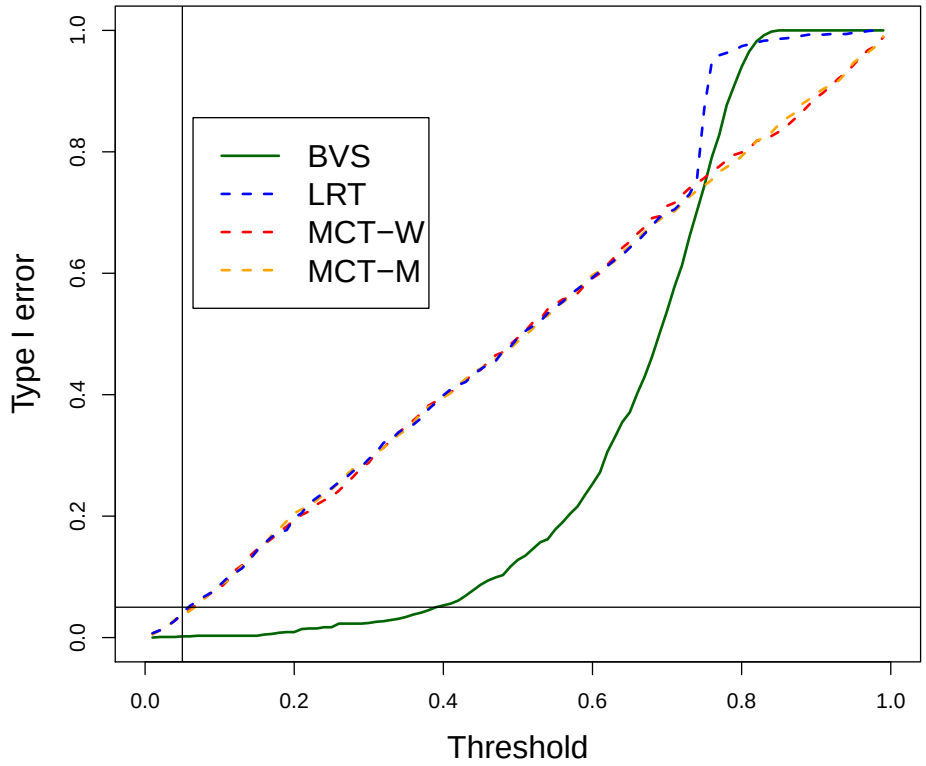
- Depends on: data on hand, prior distributions, set of alternative hypotheses.
- We use objective priors and consider the set of all possible alternative hypotheses.
- Use $\bar{P}(g_0|\text{data})$, estimation of $P(H_0|\text{data})$, to reject H_0.
- Questions:
  - How to select threshold for deciding if H_0 is rejected?
  - There is no straightforward control mechanism like Type I error.
- Simulation study can give us insight in the properties of BVS.

Simulation study

- Under the H_0 and under model g_7.
- $P(H_0|\text{data}) < \tau$ used as criterion for rejecting H_0 by BVS.
- $P_{H_0}(\text{data}^*) < \tau$ used as criterion for rejecting H_0 by LRT and MCTs.
- What happens to false rejections and false non-rejections while varying threshold τ?
- When maintaining approximately same empirical Type I error as MCTs or LRT, BVS seems to achieve similar power.
- How to select threshold for BVS in practice? $\Longrightarrow$ future research.

Simulation study - Results





Conclusion

- Model uncertainty taken into account!
- Model selection: $\bar{P}(g_r|\text{data})$.
- Estimation of means: $\hat{\boldsymbol{\mu}} = \sum_{r=0}^{R} \bar{P}(g_r|\text{data})\hat{\boldsymbol{\mu}}_r$.
- Inference: $\bar{P}(g_0|\text{data})$.
- **BVS framework address all perspectives simultaneously.**
- According to simulations seems to perform comparably with LRT and MCTs.

Future research

- How to select threshold for rejecting H_0 using $P(H_0|\text{data})$?
- How to fit BVS models with different types of restrictions (e.g. umbrella profiles)?
- How do BVS models behave when used for multiplicity adjustment?

Thank you for your attention!