Bayesian Variable Selection for Order Restricted Dose-Response Models: Model Selection and Model Complexity

Belay Birlie, Martin Otava, Ziv Shkedy, Adetayo Kasim

BSS 2016, Namur

October 2016

Outline of the presentation

Dose-response modeling

Bayesian Variable selection

Order Restricted information Criteria

Application

Further Topics

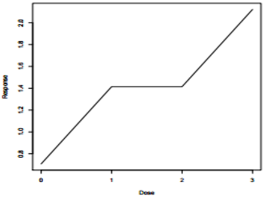
Dose-response modeling

- Increasing dose of therapeutical compound.
- Variety of possible responses:
  - Toxicity.
  - Inhibition or stimulation.
- Goal:
  - Determine if there is any relationship.
  - If so, what is the shape of the profile.

Dose-response modeling

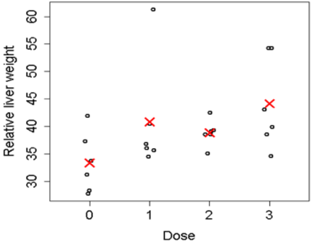
- Increasing dose of therapeutical compound.
- Variety of possible responses:
  - Toxicity.
  - Inhibition or stimulation.
- Goal:
  - Determine if there is any relationship.
  - If so, what is the shape of the profile.

- Order Constraint
  - Compound effect becomes stronger when dose is increased.
  - Monotone restriction (non-decreasing or non-increasing).
  - Zero effect is meaningful.



Motivating Example

- Toxicity study (Yanagawa and Kikuchi 2001)
  - N = 24 dogs randomized in to 4 groups
  - Each group of six receive one treatment regime
    - Placebo and three active doses of Mosapride citrate (12.5, 50, and 200 mg/kg)
  - Response: Liver weight relative body weight
    - High value of the response suggest higher toxicity of Mosapride citrate

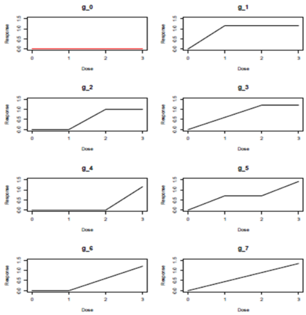


Model Formulation

- One-way ANOVA model formulation:

$$Y_{ij} = \mu_i + \epsilon_{ij}\epsilon_{ij} \sim N(0, \sigma^2)$$

- Order restriction:
  $H^{up} = \mu_0 \leq \mu_1 \leq ... \leq \mu_{K-1}$ or $H^{dn} = \mu_0 \geq \mu_1 \geq ... \geq \mu_{K-1}$



Model	Up: Mean structure
g_0	$\mu_0 = \mu_1 = \mu_2 = \mu_3$
g_1	$\mu_0 < \mu_1 = \mu_2 = \mu_3$
g_2	$\mu_0 = \mu_1 < \mu_2 = \mu_3$
g_3	$\mu_0 < \mu_1 < \mu_2 = \mu_3$
g_4	$\mu_0 = \mu_1 = \mu_2 < \mu_3$
g_5	$\mu_0 < \mu_1 = \mu_2 < \mu_3$
g_6	$\mu_0 = \mu_1 < \mu_2 < \mu_3$
g_7	$\mu_0 < \mu_1 < \mu_2 < \mu_3$

Bayesian Variable selection (BVS)

- Basic Model

$$Y_{ij} \sim N(\mu_i, \tau^{-1})$$

- Modeling of the mean

$$E(Y_{ij}) = \mu_0 + \sum_{h=1}^{i} z_h \delta_h, h = 1, ..., K - 1$$

Bayesian Variable selection (BVS)

- Basic Model

$$Y_{ij} \sim N(\mu_i, \tau^{-1})$$

- Modeling of the mean

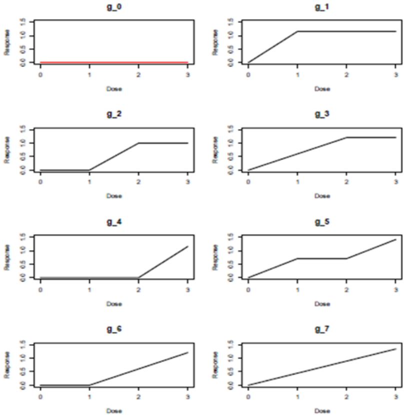
$$E(Y_{ij}) = \mu_0 + \sum_{h=1}^{i} z_h \delta_h, h = 1, ..., K-1$$

- Priors and Hyper Priors

$$\mu_0 \sim N(\eta_{\mu_0}, \tau_{\mu_0}^{-1})$$
$$\delta_h \sim N(\eta_{\delta_h}, \tau_{\delta_h}^{-1}) I(0, A)$$
$$\eta_{\mu_0} \sim N(0, 10^6)$$
$$\eta_{\delta_h} \sim N(0, 10^6)$$

$$\tau \sim \Gamma(10^{-3}, 10^{-3})$$
$$z_h \sim Bernoulli(\pi_h)$$
$$\tau_{\mu_0} \sim \Gamma(1, 1)$$
$$\tau_{\delta_h} \sim \Gamma(1, 1)$$
$$\pi_h \sim U(0, 1)$$

Bayesian Variable selection (BVS)



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)

$$H_{up}: \mu_0 < \mu_1, \dots < \mu_{k-1}$$

Bayesian Variable selection (BVS)

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

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

Bayesian Variable selection (BVS)

- Model Selection
  - Vector $z = (z_1, ..., z_(K1))$ uniquely defines the model.
  - Transformation $G(z) = 1 + \sum_{i-1}^{K1} z_i 2^{i-1}$ unique value for each model.
  - In each MCMC iteration we sample one vector $z = (z_1, ..., z_{K1})$.
  - Posterior mean of indicator $G(z) = r + 1$ translates into posterior probability of the model g_r
    - For posterior probabilities holds:

$$P(G(z) = r + 1|data) = P(g_r|data)$$

Order Restricted information Criteria

- In the case of order restricted hypothesis, the common information criteria's such as AIC, BIC, etc. is not suitable for model selection
- ORIC (Anraku ,1999) take in to account monotonicity induced in the mean structure

$$ORIC = -2logL_{OR}(\theta|data) + \sum_{l=1}^{K} lP(l, k, w)$$

- Level Probability - $P(l, k, w)$
  - Express the probability of obtaining number l of distinct means
  - For example, for $K = 4$, $P(1, k, w) = 0.25$, $P(2, k, w) = 0.46$, $P(3, k, w) = 0.25$, $P(4, k, w) = 0.04$

Order Restricted information Criteria

- Model selection
  - Approximation of Posterior Model probability

$$w_r = P_{IC}(g_r|data) = \frac{exp(-\frac{1}{2}\Delta ORIC_r)}{\sum_{s=1}^{R} exp(-\frac{1}{2}\Delta ORIC_s)}$$

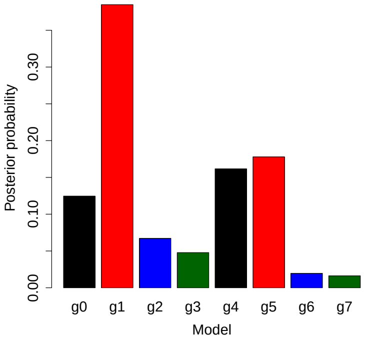
  where $\Delta ORIC_r = ORIC_r - ORIC_{min}$

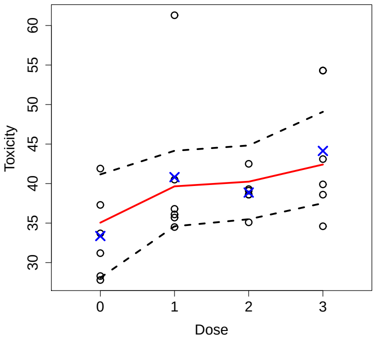
- Model complexity: weighted sum of number of levels

$$EC = \sum_{l=1}^{K} lP(l, k, w)$$

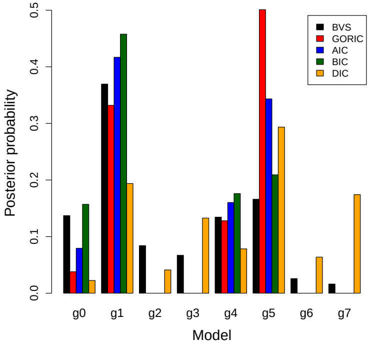
  - For example, for $K = 4$, $EC = 2.083$
  - It is the expected number of levels when isotonic regression is used to estimate the means and the data are generated under the null hypothesis

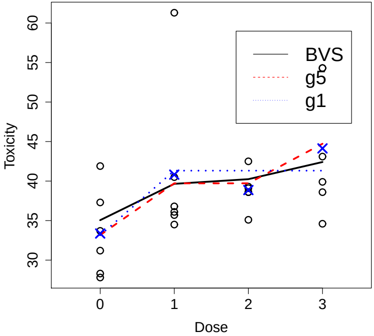
Application





Application





Further Topics

- Model Complexity
- Hypothesis Testing