Supplementary Material for:

A Bayesian nonparametric multiple testing procedure for comparing several treatments against a control

Luis Gutiérrez * Andrés F. Barrientos † Jorge González ‡
Daniel Taylor-Rodríguez §

September 7, 2018

Here, we provide the Gibbs Algorithm described in Section 3.4. We also present the image plots of the comparison between our proposal and other classical hypothesis tests of Section 4.2.

1 Appendix

Gibbs Algorithm:

We construct a Gibbs sampler algorithm with slice sampling steps as in Kalli et al. (2011) and Walker (2007) in oder to overcome the infinite-dimensionality inherent to the dependent Dirichlet process. Let the random density f_{P_x} be defined as

$$f_{P_x}(y) := \int \phi\left(y|\mu, \sigma^2\right) P_x(d\mu, d\sigma^2).$$

*Departamento de Estadística, Pontificia Universidad Católica de Chile
†Duke University, Department of Statistical Science
‡Departamento de Estadística, Pontificia Universidad Católica de Chile
§Portland State University, Department of Mathematics & Statistics

We consider an augmented model given by

$$f_{P_x}(y, u, s) = \mathbb{I}(u < \omega_s)\phi(y \mid \theta_s), \tag{1}$$

where $\theta = (\mu, \sigma^2)$, s denotes the allocation variable of y and u is a uniform random variate on $(0, \omega_s)$. Hence, the augmented likelihood can be written as

$$\mathcal{L}_{\mathsf{v},\theta}(\mathbf{y}, \mathbf{u}, \mathbf{s}) = \prod_{k \in D} \prod_{i=1}^{n_k} \mathbb{I}(u_{k,i} < \omega_{s_{k,i}})\phi(y_{k,i} \mid \theta_{s_{k,i}}), \tag{2}$$

where $D = \{c, 1, \ldots, p\}$. The main variables that need to be sampled at each step of the Gibbs algorithm are $\{\omega_j, \theta_j, j = 1, \ldots, N\}$, $s_{k,i}$ and $u_{k,i}$ for $k \in D$ and $i = 1, \ldots, n$. Here, $N := max_{k,i}\{N_{k,i}\}$ with $N_{k,i}$ being the largest integer $s_{k,i}$ for which $\{u_{k,i} < \omega_{s_{k,i}}\}$, which is equivalent to find an $N_{k,i}$ such that $\sum_{\ell=1}^{N_{k,i}} \omega_\ell > 1 - u_{k,i}$.

Updating the locations:

For the locations we have the following general posterior:

$$\pi(\theta_j \mid \ldots) \propto f_0(\theta_j) \prod_{\{k,i:s_{k,i}=j\}} \phi(y_{k,i} \mid \theta_j),$$

where $f_0(\theta_j)$ is given by the spike and slab priors of Section 3.2. Thus, given $\boldsymbol{\gamma}$ we have:

- $(\mu_{c,j}, \eta_{1,j}, \ldots, \eta_{p,j} \mid \cdots) \sim \mathrm{N}_{p+1}(\cdot \mid \tilde{\mu}_j, \Sigma_{\tilde{\mu}_j})$, where

$$\Sigma_{\tilde{\mu}_j} = \left[\frac{1}{\sigma^2_{c,j}} \left(\sum_{i:s_{c,i}=j} \mathbf{x}_i^t \mathbf{x}_i + \frac{1}{\tau_{1,j}} \sum_{i:s_{1,i}=j} \mathbf{x}_i^t \mathbf{x}_i + \cdots + \frac{1}{\tau_{p,j}} \sum_{i:s_{p,i}=j} \mathbf{x}_i^t \mathbf{x}_i \right) + S_\gamma^{-1} \right]^{-1},$$

$$\tilde{\mu}_j = \Sigma_{\tilde{\mu}_j} \left[\frac{1}{\sigma^2_{c,j}} \left(\sum_{i:s_{c,i}=j} \mathbf{x}_i^t y_{c,i} + \frac{1}{\tau_{1,j}} \sum_{i:s_{1,i}=j} \mathbf{x}_i^t y_{1,i} + \cdots + \frac{1}{\tau_{p,j}} \sum_{i:s_{p,i}=j} \mathbf{x}_i^t y_{p,i} \right) + S_\gamma^{-1} \mu_\gamma \right]$$

, $\mathbf{x}_i = (1, \boldsymbol{e}_{k_i})$, where $\boldsymbol{e}_{k_i}$ is a row vector of dimension p with a 1 in the position k indicating the population membership, $\mu_\gamma = (0, 0 \ldots, 0)^t$ and

$$S = \begin{bmatrix} \epsilon s & & & \\ & \epsilon s^{\gamma_1} & & \\ & & \ddots & \\ & & & \epsilon s^{\gamma_p} \end{bmatrix}.$$

- $\left(\frac{1}{\sigma^2_{c,j}} \mid \cdots \right) \sim \text{Gamma}(\tilde{\alpha}_{c,j}, \tilde{\lambda}_{c,j})$ and $\left(\frac{1}{\tau_{k,j}} \mid \cdots \right) \sim \text{Gamma}(\tilde{\alpha}_{k,j}, \tilde{\lambda}_{k,j}),\ k \in D$, where

$$\tilde{\alpha}_{c,j} = \frac{b}{s} + 0.5 \sum_{k \in D} \sum_{i=1}^{n_k} \mathbf{1}_{\{s_{k,i}=j\}}, \quad \tilde{\alpha}_{k,j} = \frac{b}{s^{\gamma_k}} + 0.5 \sum_{i=1}^{n_k} \mathbf{1}_{\{s_{k,i}=j\}},$$

$$\tilde{\lambda}_{c,j} = \frac{b}{s} + 0.5 \left(\sum_{i:s_{c,i}=j} (y_{c,i} - \mu_{c,j})^2 + \frac{1}{\tau_{1,j}} \sum_{i:s_{1,i}=j} (y_{1,i} - \mu_{c,j} - \eta_{1,j})^2 + \cdots + \frac{1}{\tau_{p,j}} \sum_{i:s_{p,i}=j} (y_{p,i} - \mu_{c,j} - \eta_{p,j})^2 \right)$$

and

$$\tilde{\lambda}_{k,j} = \frac{b}{s^{\gamma_k}} + 0.5 \left(\frac{1}{\sigma^2_{c,j}} \sum_{i:s_{k,i}=j} (y_{k,i} - \mu_{c,j} - \eta_{k,j})^2 \right)$$

Updating the weights:

$$\pi(v_j \mid \ldots) \propto \text{Beta}(v_j \mid a_j, b_j)$$

where $a_j = 1 + \sum_{k \in D} \sum_{i=1}^{n_k} \mathbf{1}_{(s_{k,i}=j)}$, $b_j = \kappa + \sum_{k \in D} \sum_{i=1}^{n_k} \mathbf{1}_{(s_{k,i}>j)}$ and $\omega_j = v_j \prod_{j<i}(1 - v_j)$.

Updating the membership and slice latent variables:

The full conditional distributions for the membership and slice latent variables are given by

$$\pi(s_{k,i} = \ell \mid \ldots) \propto \phi(y_{k,i} \mid \theta_\ell)\mathbf{1}(\{\ell : \omega_\ell > u_{k,i}\}) \tag{3}$$

and

$$\pi(u_{k,i} \mid \ldots) = \text{Unif}(u_{k,i}; 0, \omega_{s_{k,i}}) \tag{4}$$

respectively.

Updating others hyper-parameters:

Finally, the total mass parameter κ is updated as in Escobar and West (1995) assuming a gamma prior $\text{Ga}(a_1, a_2)$.

Updating $\boldsymbol{\gamma}$:

The updating of $\boldsymbol{\gamma}$ was performed using an independent Metropolis-Hastings step, where the acceptance probabilities were proportional to:

$$\begin{aligned}
P(\boldsymbol{\gamma} = (0, 0 \cdots, 0) \mid \cdots) &\propto \pi_{\mathcal{M}}(a_{(0,0\ldots,0)}) \prod_{j=1}^{N} \prod_{k=1}^{p} \phi\left(\eta_{k,j} \mid 0, \epsilon\right) \mathcal{G}\left(1/\tau_{k,j} \mid b, b\right) \\
P(\boldsymbol{\gamma} = (1, 0 \cdots, 0) \mid \cdots) &\propto \pi_{\mathcal{M}}(a_{(1,0,\ldots,0)}) \prod_{j=1}^{N} \phi\left(\eta_{1,j} \mid 0, \epsilon s\right) \mathcal{G}\left(1/\tau_{1,j} \mid b/s, b/s\right) \prod_{k=2}^{p} \phi\left(\eta_{k,j} \mid 0, \epsilon\right) \mathcal{G}\left(1/\tau_{k,j} \mid b, b\right) \\
&\vdots \\
P(\boldsymbol{\gamma} = (1, 1, \cdots, 1) \mid \cdots) &\propto \pi_{\mathcal{M}}(a_{(1,1,\ldots,1)}) \prod_{j=1}^{N} \prod_{k=1}^{p} \phi\left(\eta_{k,j} \mid 0, \epsilon s\right) \mathcal{G}\left(1/\tau_{k,j} \mid b/s, b/s\right)
\end{aligned}$$

2 Results of multiple and two samples classical tests

Image plots of the comparison of the BNP test with the multiple and two-sample classical tests. Figure 1 show the results for the BNP test. Figures 2 to 4 show the results for Dunnett, Nemenyi-Damico-Wolfe and Gao's tests for each scenario described in Section 4 of the manuscript. Likewise, Figures 5 to 7 show the results for the Welch's t-test, Levene and Wilcoxon and Figure 8 shows the results of the Kolmogorov-Smirnov test. The figures show number of times that each test selected each model in the 100 replications of the Monte Carlo study. Value 0 is represented by black, while 100 is represented by white in the grayscale. The ideal methods concentrate the white color on the main diagonal.

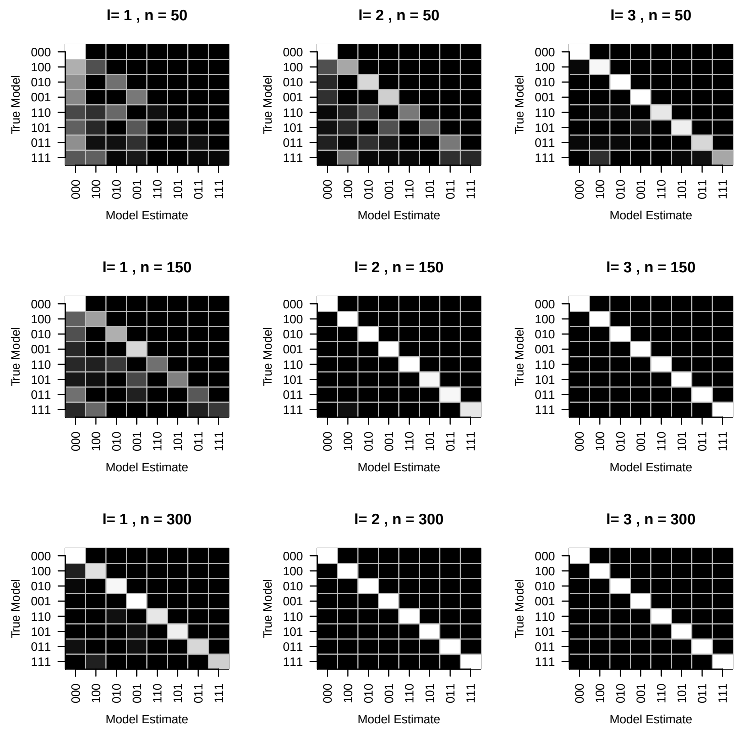

Figure 1: Number of times that the BNP test selected each model. The correctly identified hypotheses are represented in the main diagonal.

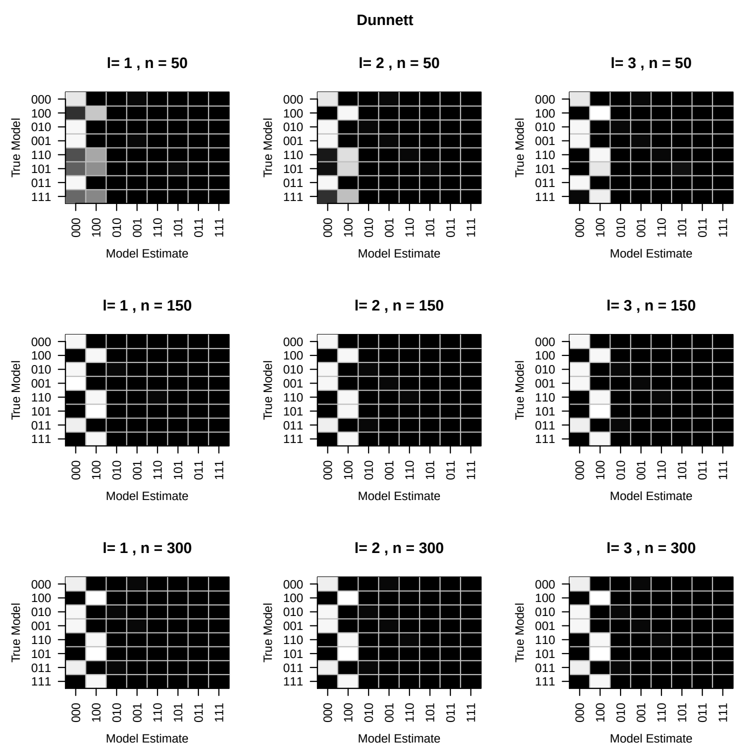

Figure 2: Number of times that the Dunnett test selected each model.

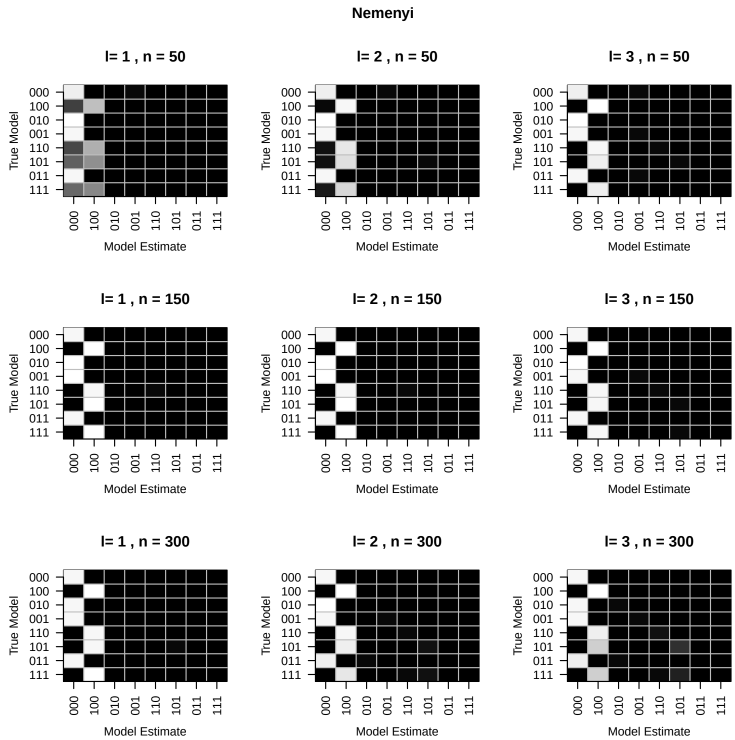

Figure 3: Number of times that the Nemenyi–Damico–Wolfe test selected each model.

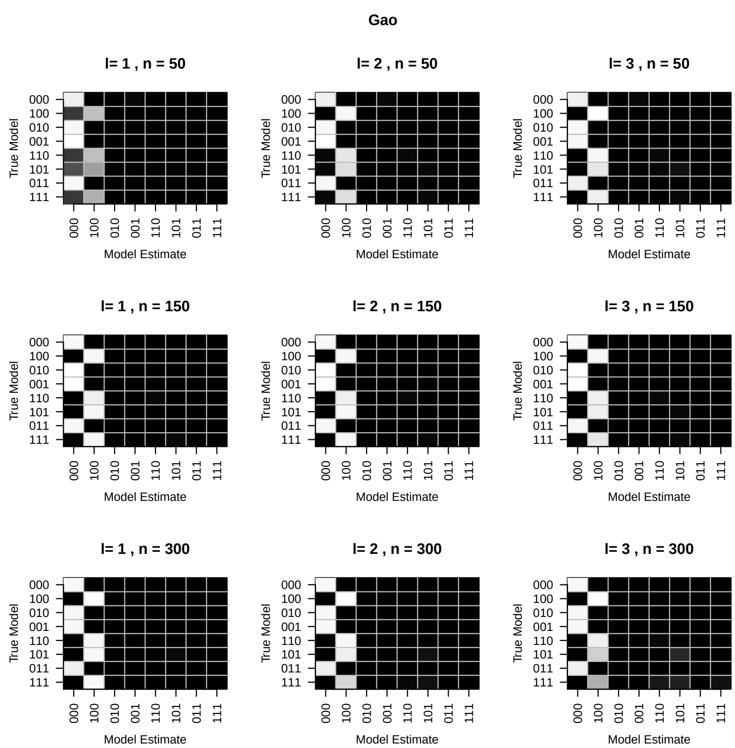

Figure 4: Number of times that the Gao test selected each model.

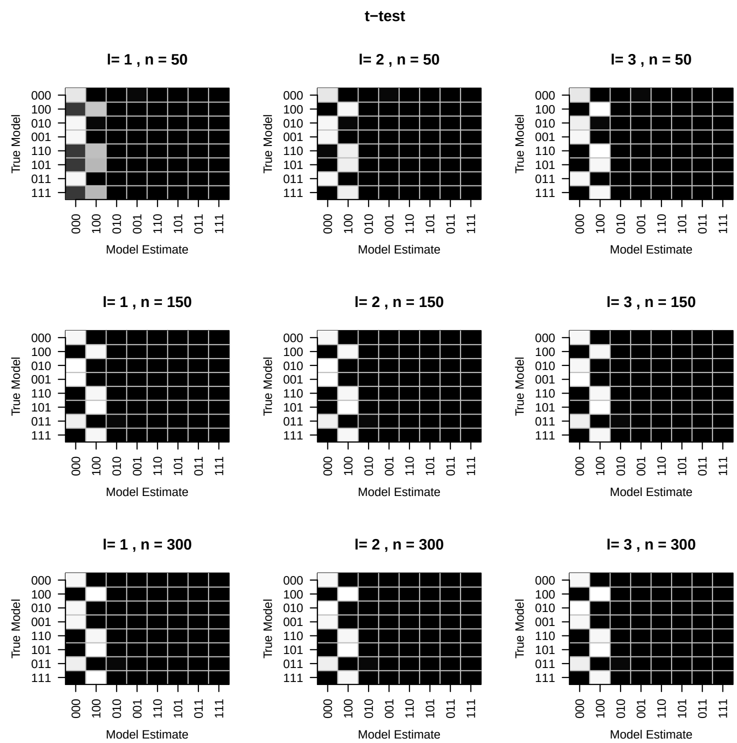

Figure 5: Number of times that the Welch t–test selected each model.

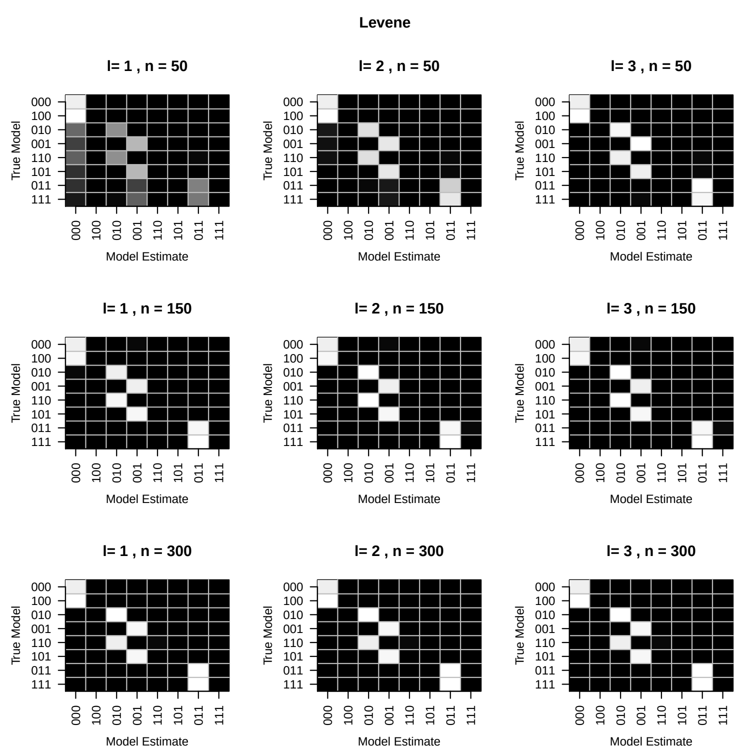

Figure 6: Number of times that the Levene test selected each model.

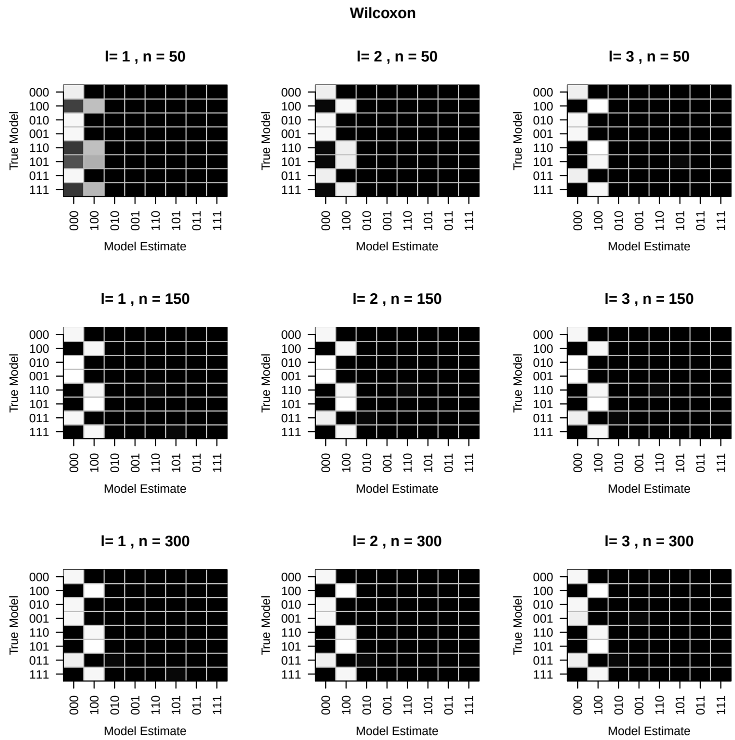

Figure 7: Number of times that the Wilcoxon test selected each model.

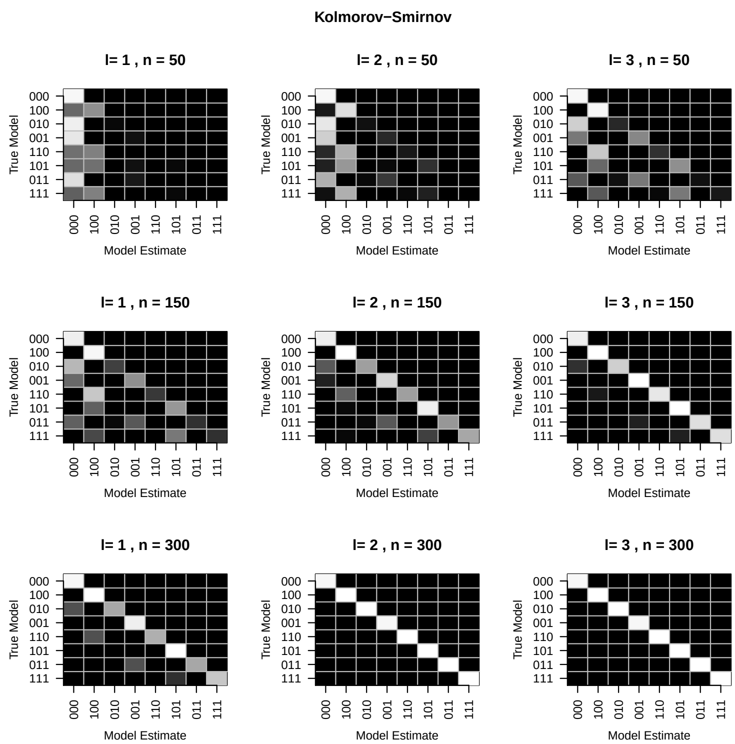

Figure 8: Number of times that the Kolmogorov-Smirnov test selected each model.