Supplementary meterial
A Bayesian nonparametric procedure for comparing algorithms

1. Proof of Theorem 1

A probability measure sampled from the posterior DP (9) has the following form:

$$P = w_0 P_0 + \sum_{j=1}^{n} w_j \delta_{\mathbf{X}_j},$$

with $P_0 \sim Dp(s, \alpha^*)$. Let us consider $E[R(X_m)]$ w.r.t. P, we have that

$$E[R(X_i)] = w_0 R_{m0} + \sum_{j=1}^{n} w_j R_{mj} = w_0 R_{m0} + \mathbf{R}_m^T \mathbf{w},$$

where $R_{m0} = \int \sum_{k=1}^{m-1} I_{\{X_m > X_k\}}(\mathbf{X}) dP_0(\mathbf{X})$, $R_{mj} = \sum_{k=1}^{m-1} I_{\{X_{mj} > X_{kj}\}} + 1$ and $\mathbf{R}_m^T = [R_{m1}, \ldots, R_{mn}]$. Assuming $P_0 = \alpha^* = \delta_{\mathbf{x}}$, we have that $R_{m0} = \int \sum_{k=1}^{m-1} I_{\{X_m > X_k\}}(\mathbf{X}) d\alpha^*(\mathbf{X})$. By defining the vector $\mathbf{R}_0 = [R_{1l}, \ldots, R_{m0}]^T$ and the matrix $\mathbf{R}$ whose rows are the vectors $\mathbf{R}_m^T$, we obtain (15).

Since the weights are Dirichlet distributed $(w_0, \mathbf{w}^T) \sim Dir(s, 1, 1, \ldots, 1)$, it follows that

$$\mathcal{E}\left[E[R(X_1), \ldots, R(X_m)]^T\right] = E_{Dir}[w_0 \mathbf{R}_0 + \mathbf{R}\mathbf{w}]$$

where the r.h.s. expected value is taken w.r.t. the Dirichlet distribution above. From $E_{Dir}[w_0] = s/(s+n)$ and $E_{Dir}[w_i] = 1/(s+n)$ for $i > 0$, we obtain the mean vector $\boldsymbol{\mu}$ in (16). The covariance matrix can be obtained in a similar way by computing $E_{Dir}[w_0^2]$, $E_{Dir}[\mathbf{w}^T w_0]$, $E_{Dir}[\mathbf{w}\mathbf{w}^T]$ etc..

2. Proof of Theorem 2

From (13) we can write

$$\begin{aligned} &\mathcal{P}\left[P(X_2 > X_1) + \tfrac{1}{2} P(X_2 = X_1) > \tfrac{1}{2}\right] \\ &= P_{Dir}\left[w_0\left(P_0(X_2 > X_1) + \tfrac{1}{2} P_0(X_2 = X_1)\right) \right. \\ &\quad \left. + \textstyle\sum_{i=1}^{n} w_i H(X_{2i} - X_{1i}) > \tfrac{1}{2}\right] \\ &= P_{Dir}\left[\tfrac{w_0}{2} + \textstyle\sum_{i=1}^{n} w_i H(X_{2i} - X_{1i}) > \tfrac{1}{2}\right], \end{aligned} \tag{1}$$

where we have used the fact that $\alpha^* = \delta_{X_1 = X_2}$ and thus $P_0(X_2 = X_1) = 1$. By denoting with $\theta_t = w_0 + \sum_{i=1}^{n} w_i I_{\{X_{2i} = X_{1i}\}}$, $\theta_g = \sum_{i=1}^{n} w_i I_{\{X_{2i} > X_{1i}\}}$ and $\theta_l = 1 - \theta_t - \theta_g$, and considering a partition of the space $\mathbb{X}$ of the form $B_0 = \{(X_2, X_1) : X_2 = X_1\}$, $B_g = \{(X_2, X_1) : X_2 > X_1\}$ and $B_l = \{(X_2, X_1) : X_2 < X_1\}$ it can easily be verified that $\boldsymbol{\theta} = (\theta_t, \theta_g, \theta_l)$ has a Dirichlet distribution with parameters $(s + n_t, n_g, n - n_t - n_g)$. Then we have

$$\begin{aligned} &P_{Dir}\left[\tfrac{w_0}{2} + \textstyle\sum_{i=1}^{n} w_i H(X_{2i} - X_{1i}) > \tfrac{1}{2}\right] \\ &= P_{Dir}\left[\tfrac{1}{2}\left(w_0 + \textstyle\sum_{i=1}^{n} w_i I_{\{X_{2i} = X_{1i}\}}\right) \right. \\ &\quad \left. + \textstyle\sum_{i=1}^{n} w_i I_{\{X_{2i} > X_{1i}\}} > \tfrac{1}{2}\right] \\ &= \int I_{\{0.5\theta_t + \theta_g > 0.5\}}(\boldsymbol{\theta}) \\ &\qquad Dir(\boldsymbol{\theta}; s + n_t, n_g, n - n_t - n_g) d\theta_t d\theta_g \\ &= K_1 \int_0^1 d\theta_t \int_{0.5(1-\theta_t)}^1 \theta_t^{s+n_t-1} \theta_g^{n_g-1} \\ &\qquad (1 - \theta_t - \theta_g)^{n-n_t-n_g-1} d\theta_g, \end{aligned} \tag{2}$$

where $K_1 = \frac{\Gamma(n+s)}{\Gamma(s+n_t)\Gamma(n_g)\Gamma(n-n_t-n_g)}$. By the change of variables $\theta_g' = \frac{\theta_g}{1-\theta_t}$ we obtain

$$\begin{aligned} &P_{Dir}\left[\tfrac{w_0}{2} + \textstyle\sum_{i=1}^{n} w_i H(X_{2i} - X_{1i}) > \tfrac{1}{2}\right] \\ &= K_1 \int_0^1 \theta_t^{s+n_t-1} (1-\theta_t)^{n-n_t-1} d\theta_t \\ &\qquad \int_{0.5}^1 (\theta_g')^{n_g-1} (1-\theta_g')^{n-n_t-n_g-1} d\theta_g' \\ &= K_1 K_2 K_3 \int_{0.5}^1 Beta(\theta; n_g, n - n_t - n_g) d\theta, \end{aligned} \tag{3}$$

where $K_2 = \frac{\Gamma(n-n_t)\Gamma(s+n_t)}{\Gamma(s+n)}$ and $K_3 = \frac{\Gamma(n_g)\Gamma(n-n_t-n_g)}{\Gamma(n)}$. This proves the theorem, since $K_1 K_2 K_3 = 1$.

3. Matrix of the ranks of Example 1

Table 1 gives the ranks in $n = 30$ datasets of the four algorithms in Example 1.

Table 1. Ranks of the algorithms in Example 1.

X_1	4	4	4	4	4	3	3	3	3	3	4	4	4	4	4	3	3	3	3	3	2	2	2	2	4	4	3	3	4	3
X_2	3	3	3	3	3	4	4	4	4	4	2	2	2	2	2	2	2	2	2	2	3	3	4	4	3	3	4	4	2	2
X_3	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	4	4	4	4	4	4	4	3	3	1	1	1	1	1	1
X_4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	3	4