Supplementary Material for:

Bayesian nonparametric ROC regression modeling

VANDA INÁCIO DE CARVALHO, ALEJANDRO JARA, TIMOTHY E. HANSON AND MIGUEL DE CARVALHO

[1]Vanda Inácio de Carvalho is Assistant Professor, Department of Statistics, Pontificia Universidad Católica de Chile, Casilla 306, Correo 22, Santiago, Chile, and Research Fellow, Centro de Estatística e Aplicações, Universidade de Lisboa (E-mail: vanda.kinets@gmail.com). Alejandro Jara is Assistant Professor, Department of Statistics, Pontificia Universidad Católica de Chile, Casilla 306, Correo 22, Santiago, Chile (E-mail: atjara@uc.cl). Timothy E. Hanson is Professor, Department of Statistics, University of South Carolina, Columbia, SC 29208, USA (E-mail: hansont@stat.sc.edu). Miguel de Carvalho is Assistant Professor, Department of Statistics, Pontificia Universidad Católica de Chile, Casilla 306, Correo 22, Santiago, Chile, and Research Fellow, Centro de Matemática e Aplicações, Universidade Nova de Lisboa (E-mail: mdecarvalho@mat.puc.cl).

Appendix A: Proof of Theorem 1

For every $\mathbf{x} \in \mathcal{X}$, let $F_h^*(\cdot \mid \mathbf{x})$, $h = 0, 1$, be the cumulative distribution functions associated with a given and fixed conditional ROC curve, that is, $\mathrm{ROC}(u \mid \mathbf{x}) = 1 - F_1^*(F_0^{*-1}(1-u \mid \mathbf{x}) \mid \mathbf{x}) \equiv \overline{\mathrm{ROC}}(1-u \mid \mathbf{x})$. It follows that, for every $\mathbf{x} \in \mathcal{X}$ and almost every $\omega \in \Omega$,

$$\begin{aligned}
\left|\overline{\mathrm{ROC}}_\omega(u \mid \mathbf{x}) - \overline{\mathrm{ROC}}(u \mid \mathbf{x})\right| &= \left|F_{1,\omega}(F_{0,\omega}^{-1}(u \mid \mathbf{x}) \mid \mathbf{x}) - F_1^*(F_0^{*-1}(u \mid \mathbf{x}) \mid \mathbf{x})\right| \\
&= \left|\int_{-\infty}^{F_{0,\omega}^{-1}(u|\mathbf{x})} f_{1,\omega}(v \mid \mathbf{x})\mathrm{d}v - \int_{-\infty}^{F_0^{*-1}(u|\mathbf{x})} f_1^*(v \mid \mathbf{x})\mathrm{d}v\right| \\
&= \left|\int_{-\infty}^{\infty} f_{1,\omega}(v \mid \mathbf{x}) I\{F_{0,\omega}(v \mid \mathbf{x}) < u\}(v)\mathrm{d}v \right.\\
&\quad \left. - \int_{-\infty}^{\infty} f_1^*(v \mid \mathbf{x}) I\{F_0^*(v \mid \mathbf{x}) < u\}(v)\mathrm{d}v\right| \\
&\leq \int_{-\infty}^{\infty} f_1^*(v \mid \mathbf{x}) \left|I\{F_{0,\omega}(v \mid \mathbf{x}) < u\}(v) - I\{F_0^*(v \mid \mathbf{x}) < u\}(v)\right| \mathrm{d}v \\
&\quad + \int_{-\infty}^{\infty} I\{F_0^*(v \mid \mathbf{x}) < u\}(v) \left|f_{1,\omega}(v \mid \mathbf{x}) - f_1^*(v \mid \mathbf{x})\right| \mathrm{d}v \\
&\leq \int_{-\infty}^{\infty} f_1^*(v \mid \mathbf{x}) \left|I\{F_{0,\omega}(v \mid \mathbf{x}) < u\}(v) - I\{F_0^*(v \mid \mathbf{x}) < u\}(v)\right| \mathrm{d}v \\
&\quad + \left\|f_{1,\omega}(\cdot \mid \mathbf{x}) - f_1^*(\cdot \mid \mathbf{x})\right\|_1,
\end{aligned}$$

where $f_{h,\omega}(\cdot \mid \mathbf{x})$, $h = 0, 1$ denotes the density associated with the trajectories of the DDP mixture of normals model for each group, and $F_{h,\omega}(\cdot \mid \mathbf{x})$, $h = 0, 1$, denote the corresponding cumulative density functions.

Now notice that, for every $\mathbf{x} \in \mathcal{X}$, there exists $r_\mathbf{x} > 0$ such that $1 - \int_{-r_\mathbf{x}}^{r_\mathbf{x}} f_1^*(v \mid \mathbf{x})\mathrm{d}v < \epsilon/3$, and, therefore,

$$\begin{aligned}
\left|\overline{\mathrm{ROC}}_\omega(u \mid \mathbf{x}) - \overline{\mathrm{ROC}}(u \mid \mathbf{x})\right| &\leq \int_{-r_\mathbf{x}}^{r_\mathbf{x}} f_1^*(v \mid \mathbf{x}) \left|I\{F_{0,\omega}(v \mid \mathbf{x}) < u\}(v) - I\{F_0^*(v \mid \mathbf{x}) < u\}(v)\right| \mathrm{d}v \\
&\quad + \left\|f_{1,\omega}(\cdot \mid \mathbf{x}) - f_1^*(\cdot \mid \mathbf{x})\right\|_1 + \epsilon/3.
\end{aligned}$$

Notice also that, for every $\mathbf{x} \in \mathcal{X}$, there exists $\delta_\mathbf{x} > 0$ such that $\int_{v_0-\delta_\mathbf{x}}^{v_0+\delta_\mathbf{x}} f_1^*(v \mid \mathbf{x})\mathrm{d}v < \epsilon/3$, for every $v_0 \in (-r_\mathbf{x}, r_\mathbf{x})$. Finally, notice that, for every $\delta_\mathbf{x} > 0$, there exists $\gamma_\mathbf{x} > 0$ such that if

$$\sup_{v \in (-r_\mathbf{x}, r_\mathbf{x})} \left|F_{0,\omega}(v \mid \mathbf{x}) - F_0^*(v \mid \mathbf{x})\right| < \gamma_\mathbf{x},$$

holds, then

$$\sup_{u \in (F_0^*(-r_\mathbf{x}), F_0^*(r_\mathbf{x}))} \left|F_{0,\omega}^{-1}(u \mid \mathbf{x}) - F_0^{*-1}(u \mid \mathbf{x})\right| < \delta_\mathbf{x}.$$

Now, by the equivalence between total variation and L_1, it follows that for every $\gamma_{\mathbf{x}} > 0$, there exists $\rho_{\mathbf{x}} > 0$ such that if

$$\|f_{0,\omega}(\cdot \mid \mathbf{x}) - f_0^*(\cdot \mid \mathbf{x})\|_1 < \rho_{\mathbf{x}},$$

holds, then

$$\sup_{v \in (-r_{\mathbf{x}}, r_{\mathbf{x}})} |F_{0,\omega}(v \mid \mathbf{x}) - F_0^*(v \mid \mathbf{x})| < \gamma_{\mathbf{x}}.$$

It follows that, for every $\epsilon > 0$, $\mathbf{x}_1, \ldots, \mathbf{x}_T \in \mathcal{X}$, $\{F_1^*(\cdot \mid \mathbf{x}) : \mathbf{x} \in \mathcal{X}\}$ and $\{F_0^*(\cdot \mid \mathbf{x}) : \mathbf{x} \in \mathcal{X}\}$, there exists $\rho = \min\{\rho_{\mathbf{x}_1}, \ldots, \rho_{\mathbf{x}_T}\} > 0$, such that if

$$\|f_{1,\omega}(\cdot \mid \mathbf{x}_t) - f_1^*(\cdot \mid \mathbf{x}_t)\|_1 < \epsilon/3,$$

and

$$\|f_{0,\omega}(\cdot \mid \mathbf{x}_t) - f_0^*(\cdot \mid \mathbf{x}_t)\|_1 < \rho,$$

hold, then

$$\|\mathrm{ROC}_\omega(\cdot \mid \mathbf{x}_t) - \mathrm{ROC}(\cdot \mid \mathbf{x}_t)\|_\infty = \left\|\overline{\mathrm{ROC}}_\omega(\cdot \mid \mathbf{x}_t) - \overline{\mathrm{ROC}}(\cdot \mid \mathbf{x}_t)\right\|_\infty < \epsilon.$$

It follows that

$$\begin{aligned} \mathbb{P}\left\{\omega \in \Omega : \|\mathrm{ROC}_\omega(\cdot \mid \mathbf{x}_t) - \mathrm{ROC}(\cdot \mid \mathbf{x}_t)\|_\infty < \epsilon,\ t = 1, \ldots, T\right\} \geq \\ \mathbb{P}\left\{\omega \in \Omega : \|f_{1,\omega}(\cdot \mid \mathbf{x}_t) - f_1^*(\cdot \mid \mathbf{x}_t)\|_1 < \epsilon/3,\ t = 1, \ldots, T\right\} \times \\ \mathbb{P}\left\{\omega \in \Omega : \|f_{0,\omega}(\cdot \mid \mathbf{x}_t) - f_0^*(\cdot \mid \mathbf{x}_t)\|_1 < \rho,\ t = 1, \ldots, T\right\}. \end{aligned}$$

Thus, by Theorem 4 in Barrientos et al. (2012) on the Hellinger support of DDP mixture models, it follows that

$$\mathbb{P}\left\{\omega \in \Omega : \|f_{1,\omega}(\cdot \mid \mathbf{x}_t) - f_1^*(\cdot \mid \mathbf{x}_t)\|_1 < \epsilon/3,\ t = 1, \ldots, T\right\} > 0,$$

and

$$\mathbb{P}\left\{\omega \in \Omega : \|f_{0,\omega}(\cdot \mid \mathbf{x}_t) - f_0^*(\cdot \mid \mathbf{x}_t)\|_1 < \rho,\ t = 1, \ldots, T\right\} > 0,$$

and, therefore,

$$\mathbb{P}\left\{\omega \in \Omega : \|\mathrm{ROC}_\omega(\cdot \mid \mathbf{x}_t) - \mathrm{ROC}(\cdot \mid \mathbf{x}_t)\|_\infty < \epsilon,\ t = 1, \ldots, T\right\} > 0,$$

which completes the proof of the theorem. □

Appendix B: Markov chain Monte Carlo details

The hierarchical representation of the model

The hierarchical representation of the B–splines DDP mixture of normal model is given by

$$y_{hl} \mid \mathbf{z}_{hl}, \boldsymbol{\theta}_{hl} \overset{\text{ind.}}{\sim} \phi(\cdot \mid \mathbf{z}_{hl}' \boldsymbol{\beta}_l^h, \tau_l^h), \tag{B.1}$$

$$\boldsymbol{\theta}_{hl} = (\boldsymbol{\beta}_l^h, \tau_l^h) \mid G_h \overset{\text{i.i.d.}}{\sim} G_h, \tag{B.2}$$

$$G_h \mid \alpha_h, G_{0h}^* \sim \text{DP}\left(\alpha_h N_q(\boldsymbol{\mu}_h, \boldsymbol{\Sigma}_h) \times \Gamma^{-1}(\tau_{h1}/2, \tau_{h2}/2)\right), \tag{B.3}$$

$$\alpha_h \mid a_h, b_h \sim \Gamma(a_h, b_h), \tag{B.4}$$

$$\tau_{h2} \mid \tau_{s_{h1}}, \tau_{s_{h2}} \sim \Gamma(\tau_{s_{h1}}/2, \tau_{s_{h2}}/2), \tag{B.5}$$

$$\boldsymbol{\mu}_h \mid \mathbf{m}_h, \mathbf{S}_h \sim N_q(\mathbf{m}_h, \mathbf{S}_h), \tag{B.6}$$

and

$$\boldsymbol{\Sigma}_h \mid \nu_h, \boldsymbol{\Psi}_h \sim \text{IW}_q(\nu_h, \boldsymbol{\Psi}_h). \tag{B.7}$$

The marginal algorithm

We marginalized the DP measures G_h for the joint distribution implied by expressions (B.1)–(B.7) and explore the posterior distribution of

$$(\boldsymbol{\theta}_0, \boldsymbol{\theta}_1, \alpha_0, \alpha_1, \tau_{02}, \tau_{12}, \boldsymbol{\mu}_0, \boldsymbol{\mu}_1, \boldsymbol{\Sigma}_0, \boldsymbol{\Sigma}_1),$$

where $\boldsymbol{\theta}_h = (\boldsymbol{\theta}_{h1}, \ldots, \boldsymbol{\theta}_{hn_h})$, using a Gibbs sampling algorithm.

We update the coordinates of $\boldsymbol{\theta}_h$ using a Gibbs sampling algorithm through its coordinates. Let $\boldsymbol{\theta}_h^{(i)} = (\boldsymbol{\theta}_{h1}, \ldots, \boldsymbol{\theta}_{hi-1}, \ldots, \boldsymbol{\theta}_{hi+1}, \ldots, \boldsymbol{\theta}_{hn_h})$ be the vector of subject-specific parameters in group h, excluding the ones associated with subject i. Let $\boldsymbol{\theta}_h^{*(i)} = \{\boldsymbol{\theta}_{h1}^*, \ldots, \boldsymbol{\theta}_{hk_h^{(i)}}^*\}$, $h \in \{0, 1\}$, be the set of $k_h^{(i)} \leq n_h - 1$ distinct elements in $\boldsymbol{\theta}_h^{(i)}$. The full conditional distribution for $\boldsymbol{\theta}_{hi}$ is given by

$$\begin{aligned}\boldsymbol{\theta}_{hi} \mid \boldsymbol{\theta}_h^{(i)}, \alpha_h, \boldsymbol{\mu}_h, \boldsymbol{\Sigma}_h, \tau_{h1}, \tau_{h2}, \mathbf{y}_h \quad \sim \quad & b_h^i \frac{\alpha_h}{n_h - 1 + \alpha_h} \int \phi(y_{hi} \mid \mathbf{z}_{hi}' \boldsymbol{\beta}, \tau) \mathrm{d}G_{0h}^*(\boldsymbol{\beta}, \tau) + \\ & b_h^i \sum_{j=1}^{k_h^{(i)}} \frac{n_{hj}^{(i)}}{n_h - 1 + \alpha_h} \phi(y_{hi} \mid \mathbf{z}_{hi}' \boldsymbol{\beta}_{hj}^*, \tau_{hj}^*),\end{aligned}$$

where b_h^i is a normalizing constant and $n_{hj}^{(i)}$ is the number of elements in $\boldsymbol{\theta}_h^{(i)}$ such that $\boldsymbol{\theta}_{hi} = \boldsymbol{\theta}_{hj}^*$. Even though the centering distributions G_{0h}^* are conjugate with the normal likelihood considered here, the auxiliary variable approach proposed by Neal (2000, Algorithm 8) with $m = 1$, was considered for updating the $\boldsymbol{\theta}_{hi}$'s. The extra step suggested by Bush & MacEachern (1996), was considered in order to improve the mixing of the chain.

The precision parameters α_h are updated using the auxiliary variable approach proposed by **?**. The full conditional distribution for the means of the Gaussian components of the corresponding centering distributions, $\boldsymbol{\mu}_h$, is Gaussian and corresponds to the posterior distribution of $\boldsymbol{\mu}_h$ associated with the hierarchical model:

$$\boldsymbol{\mu}_h \mid \mathbf{m}_h, \mathbf{S}_h \sim N_q(\mathbf{m}_h, \mathbf{S}_h),$$

and

$$\boldsymbol{\beta}_{h1}^*, \ldots, \boldsymbol{\beta}_{hk_h}^* \mid \boldsymbol{\mu}_h \overset{\text{i.i.d.}}{\sim} N_q(\boldsymbol{\mu}_h, \boldsymbol{\Sigma}_h),$$

where $\left\{\boldsymbol{\beta}_{h1}^*, \ldots, \boldsymbol{\beta}_{hk_h}^*\right\}$, $h \in \{0, 1\}$, is the set of $k_h \leq n_h$ distinct vectors of regression coefficients in $\boldsymbol{\theta}_h$. In a similar way, the full conditional distribution for the covariance matrices of the Gaussian components of the corresponding centering distributions, $\boldsymbol{\Sigma}_h$, is inverted-Wishart and corresponds to the posterior distribution of $\boldsymbol{\Sigma}_h$ associated with the hierarchical model:

$$\boldsymbol{\Sigma}_h \mid \nu_h, \boldsymbol{\Psi}_h \sim \text{IW}_q(\nu_h, \boldsymbol{\Psi}_h),$$

and

$$\boldsymbol{\beta}_{h1}^*, \ldots, \boldsymbol{\beta}_{hk_h}^* \mid \boldsymbol{\Sigma}_h \overset{\text{i.i.d.}}{\sim} N_q(\boldsymbol{\mu}_h, \boldsymbol{\Sigma}_h).$$

Finally, the full conditional distribution for the hyper-parameter of the inverted-gamma component of the centering distributions is

$$\tau_{h2} \mid \boldsymbol{\theta}^h \sim \Gamma\left(0.5\left[k_h \times \tau_{h1} + \tau_{s_{h1}}\right], 0.5\left[\sum_{j=1}^{k_h} \frac{1}{\tau_{hj}^*} + \tau_{s_{h2}}\right]\right),$$

where $\{\tau_{h1}^*, \ldots, \tau_{hk_h}^*\}$, $h \in \{0, 1\}$, is the set of $k_h \leq n_h$ distinct variances in $\boldsymbol{\theta}_h$.

Sampling functional parameters

Samples for the conditional ROC curves requires samples of the mixing distributions G_h. The marginal algorithm described above provides posterior samples of the finite-dimensional part of the model,

$$\left(\boldsymbol{\theta}_0^{(j)}, \boldsymbol{\theta}_1^{(j)}, \alpha_0^{(j)}, \alpha_1^{(j)}, \tau_{02}^{(j)}, \tau_{12}^{(j)}, \boldsymbol{\mu}_0^{(j)}, \boldsymbol{\mu}_1^{(j)}, \boldsymbol{\Sigma}_0^{(j)}, \boldsymbol{\Sigma}_1^{(j)}\right),$$

$j = 1, \ldots, J$, which were used to obtain samples of finite-dimensional approximations to G_h (and any functional). From the conjugacy of the DP, it follows that for every j,

$$G_h^{(j)} \mid \boldsymbol{\theta}_h^{(j)}, \alpha_h^{(j)}, \boldsymbol{\mu}_h^{(j)}, \boldsymbol{\Sigma}_h^{(j)}, \tau_{h2}^{(j)} \sim \text{DP}\left(\alpha_h^* H_h^*\right), \tag{B.8}$$

where $\alpha_h^* = \alpha_h^{(j)} + n_h$, and

$$H_h^*(\cdot) = \frac{1}{\alpha_h^{(j)} + n_h}\left(\alpha_h^{(j)} G_{h0}^{*(j)}(\cdot) + \sum_{i=1}^{n_h} \delta_{\boldsymbol{\theta}_{hi}^{(j)}}(\cdot)\right).$$

Approximated samples from expression (B.8) were obtained using the ϵ-DP approach proposed by Muliere & Tardella (1998). In this approach the samples of the DP are approximated in such a way that the total variation between the full realization and the approximation is smaller or equal than ϵ—the value $\epsilon = 0.01$ was used in our computational implementation. Finally, in order to compute the samples of the conditional ROC curves, the evaluation of the CDF and quantile function of finite mixture of normals models was needed. The bisection method (see, e.g. Givens & Hoeting, 2005) was used with this aim.

Appendix C: Details on existing methods

Semiparametric linear model (Pepe, 1998)

This method is based on specifying a homocedastic linear regression model for the healthy and diseased groups, i.e.,

$$\begin{aligned} y_0 &= \tilde{\mathbf{x}}'\boldsymbol{\beta}_0 + \sigma_0\varepsilon_0, \\ y_1 &= \tilde{\mathbf{x}}'\boldsymbol{\beta}_1 + \sigma_1\varepsilon_1, \end{aligned}$$

where $\tilde{\mathbf{x}} = (1, \mathbf{x}')'$, $\boldsymbol{\beta}_0 = (\beta_{00}, \ldots, \beta_{0p})$ and $\boldsymbol{\beta}_1 = (\beta_{10}, \ldots, \beta_{1p})$ are $(p+1)$-dimensional vectors of unknown parameters, and ε_0 and ε_1 are independent random variables, with mean zero, variance one and distribution functions F_0 and F_1, respectively. The estimation procedure consists of the following steps:

1. estimate $\boldsymbol{\beta}_0$ and $\boldsymbol{\beta}_1$ by ordinary least squares, on the basis of samples $\{(y_{0i}, \mathbf{x}_{0i})\}_{i=1}^{n_0}$ and $\{(y_{1j}, \mathbf{x}_{1j})\}_{j=1}^{n_1}$;
2. estimate σ_0^2 and σ_1^2 as
$$\hat{\sigma}_0^2 = \frac{\sum_{i=1}^{n_0}(y_{0i} - \tilde{\mathbf{x}}_{0i}'\hat{\boldsymbol{\beta}}_0)^2}{n_0 - p - 1} \quad \text{and} \quad \hat{\sigma}_1^2 = \frac{\sum_{j=1}^{n_1}(y_{1j} - \tilde{\mathbf{x}}_{1j}'\hat{\boldsymbol{\beta}}_1)^2}{n_1 - p - 1};$$
3. estimate the cumulative distribution functions F_0 and F_1 on the basis of the empirical distributions of the standardized residuals
$$\hat{F}_0(y) = \frac{1}{n_0}\sum_{i=1}^{n_0} I\left[\frac{y_{0i} - \tilde{\mathbf{x}}_{0i}'\hat{\boldsymbol{\beta}}_0}{\hat{\sigma}_0} \leqslant y\right] \quad \text{and} \quad \hat{F}_1(y) = \frac{1}{n_1}\sum_{j=1}^{n_1} I\left[\frac{y_{1j} - \tilde{\mathbf{x}}_{1j}'\hat{\boldsymbol{\beta}}_1}{\hat{\sigma}_1} \leqslant y\right];$$
4. for a given value of the covariate $\mathbf{x}$, calculate the covariate specific ROC curve
$$\widehat{\mathrm{ROC}}(u \mid \mathbf{x}) = 1 - \hat{F}_1(\tilde{\mathbf{x}}'\hat{\boldsymbol{\beta}} + \hat{\alpha}\hat{F}_0^{-1}(1-u)), \quad 0 \leqslant u \leqslant 1,$$
where $\hat{\boldsymbol{\beta}} = (\hat{\boldsymbol{\beta}}_0 - \hat{\boldsymbol{\beta}}_1)/\hat{\sigma}_1$ and $\hat{\alpha} = \hat{\sigma}_0/\hat{\sigma}_1$.

Nonparametric model (González-Manteiga et al., 2011; Rodríguez-Álvarez et al., 2011a)

In this method a nonparametric heterocedastic regression model is assumed for the test result

$$\begin{aligned} y_0 &= \mu_0(x) + \sigma_0(x)\varepsilon_0, \\ y_1 &= \mu_1(x) + \sigma_1(x)\varepsilon_1, \end{aligned}$$

where x is a continuous covariate, μ_0 and μ_1 are the regression functions, and σ_0 and σ_1 are the variance functions. Here ε_0, ε_1 are independent random variables, with mean zero, variance one and distribution functions F_0 and F_1, respectively. The proposed estimation procedure is as follows:

1. for a given value x of the covariate, estimate the regression functions μ_0 and μ_1 as

$$\begin{aligned}\hat{\mu}_0(x) &= \hat{\psi}(x, \{(y_{0i}, x_{0i})\}_{i=1}^{n_0}, h_0, p_0),\\ \hat{\mu}_1(x) &= \hat{\psi}(x, \{(y_{1j}, x_{1j})\}_{j=1}^{n_1}, h_1, p_1),\end{aligned}$$

   where $\hat{\phi}$ is the local polynomial kernel estimator (Fan & Gijbels, 1996), h_0 and h_1 are the smoothing parameters or bandwidths, and p_0 and p_1 are the orders of the polynomials, in the healthy and diseased populations, respectively;

2. estimate the variance functions σ_0^2 and σ_1^2 in a similar fashion

$$\begin{aligned}\hat{\sigma}_0^2(x) &= \hat{\psi}(x, \{(z_{0i}, x_{0i})\}_{i=1}^{n_0}, g_0, q_0),\\ \hat{\sigma}_1^2(x) &= \hat{\psi}(x, \{(z_{1j}, x_{1j})\}_{j=1}^{n_1}, g_1, q_1),\end{aligned}$$

   where $z_{0i} = (y_{0i} - \hat{\mu}_0(x_{0i}))^2$, $z_{1j} = (y_{1j} - \hat{\mu}_1(x_{1j}))^2$, g_0 and g_1 are the bandwidths and q_0 and q_1 are the orders of the polynomials;

3. estimate the cumulative distribution functions F_0 and F_1 on the basis of the empirical distributions of the standardized residuals

$$\hat{F}_0(y) = \frac{1}{n_0}\sum_{i=1}^{n_0} I\left[\frac{y_{0i} - \hat{\mu}_0(x_{0i})}{\hat{\sigma}_0(x_{0i})} \leqslant y\right] \quad \text{and} \quad \hat{F}_1(y) = \frac{1}{n_1}\sum_{j=1}^{n_1} I\left[\frac{y_{1j} - \hat{\mu}_1(x_{1j})}{\hat{\sigma}_1(x_{1j})} \leqslant y\right];$$

4. compute the covariate specific ROC curve as follows:

$$\widehat{\text{ROC}}(u \mid x) = 1 - \hat{F}_1\left(\frac{\hat{\mu}_0(x) - \hat{\mu}_1(x)}{\hat{\sigma}_1(x)} + \frac{\hat{\sigma}_0(x)}{\hat{\sigma}_1(x)}\hat{F}_0^{-1}(1-u)\right), \quad 0 \leqslant u \leqslant 1.$$

Appendix D: Sensitivity analysis for the simulation study

Table 1: Simulated data: Average (standard deviation), across simulations, of the empirical global mean squared error of the ROC curve for the different approaches under consideration. The results are presented for each of the simulation scenarios and sample sizes (n).

		Approach				
Scenario	n	Sem. Linear	Sem. B–splines	Kernel	B–splines DDP	B–splines DDP II
I	50	0.0084 (0.0057)	0.0140 (0.0080)	0.0131 (0.0073)	0.0138 (0.0075)	0.0138 (0.0080)
	100	0.0045 (0.0026)	0.0076 (0.0048)	0.0074 (0.0043)	0.0079 (0.0048)	0.0075 (0.0045)
	200	0.0022 (0.0014)	0.0037 (0.0023)	0.0036 (0.0020)	0.0042 (0.0022)	0.0040 (0.0024)
II	50	0.0385 (0.0056)	0.0122 (0.0058)	0.0130 (0.0064)	0.0125 (0.0061)	0.0106 (0.0056)
	100	0.0364 (0.0037)	0.0076 (0.0037)	0.0079 (0.0041)	0.0079 (0.0039)	0.0070 (0.0035)
	200	0.0345 (0.0022)	0.0045 (0.0015)	0.0042 (0.0017)	0.0047 (0.0017)	0.0049 (0.0032)
III	50	0.0534 (0.0090)	0.0218 (0.0112)	0.0302 (0.0156)	0.0162 (0.0090)	0.0160 (0.0093)
	100	0.0499 (0.0057)	0.0127 (0.0052)	0.0155 (0.0064)	0.0091 (0.0051)	0.0087 (0.0047)
	200	0.0470 (0.0036)	0.0091 (0.0032)	0.0098 (0.0041)	0.0062 (0.0031)	0.0056 (0.0028)

Appendix E: Sensitivity analysis for the application

In this section, we show the results of the sensitivity analysis carried out for the application. Figure 2 shows the results under prior specification under the B-splines DDP mixture model. This figure corresponds to Figure 8 of the manuscript. In turn, Figure 3 shows the results under prior specification under the B-splines DDP II mixture model.

References

BARRIENTOS, A. F., JARA, A. & QUINTANA, F. (2012). On the support of MacEachern's dependent Dirichlet processes and extensions. *Bayesian Analysis* 7 277–310.

BUSH, C. A., & MACEACHERN, S. N. (1996). A semiparametric Bayesian model for randomised block designs. *Biometrika* 83 275–285.

FAN, J. & GIJBELS, I. (1996). *Local Polynomial Modelling and Its Applications*. London: Chapman & Hall.

GIVENS, G. H. & HOETING, J. A. (2005). *Computational Statistics*. New York: Wiley.

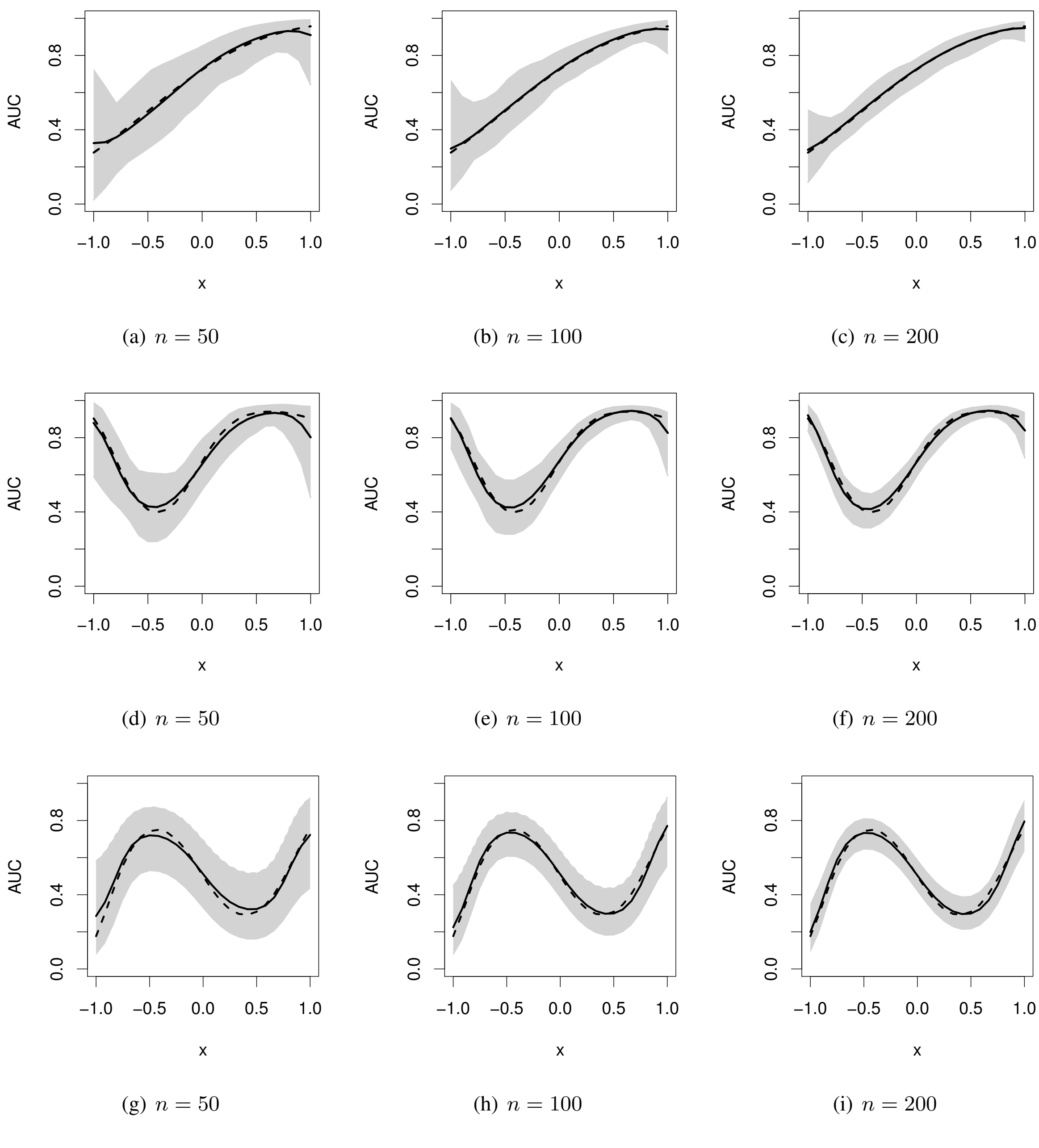
GONZÁLEZ-MANTEIGA, W., PARDO-FERNANDÉZ, J. C. & VAN KEILEGOM, I. (2011). ROC curves in non-parametric location-scale regression models. *Scandinavian Journal of Statistics* 38 169–184.

MULIERE, P. & TARDELLA, L. (1998). Approximating distributions of random functionals of Ferguson-Dirichlet priors. *The Canadian Journal of Statistics* 26 283–297.

NEAL, R. (2000). Markov chain sampling methods for Dirichlet process mixture models. *Journal of Computational and Graphical Statistics* 9 249–265.

PEPE, M. S. (1998). Three approaches to regression analysis of receiver operating characteristic curves for continuous test results. *Biometrics* 54 124–135.

RODRÍGUEZ-ÁLVAREZ, M. X., ROCA-PARDIÑAS, J. & CADARSO-SUÁREZ, C. (2011a). ROC curve and covariates: extending the induced methodology to the non-parametric framework. *Statistics and Computing* 21 483–495.

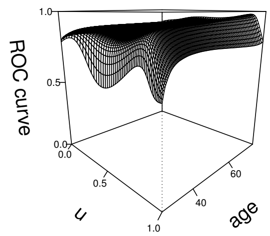


(a) $n = 50$ (b) $n = 100$ (c) $n = 200$

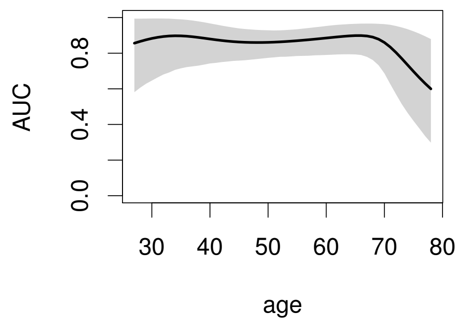
(d) $n = 50$ (e) $n = 100$ (f) $n = 200$

(g) $n = 50$ (h) $n = 100$ (i) $n = 200$

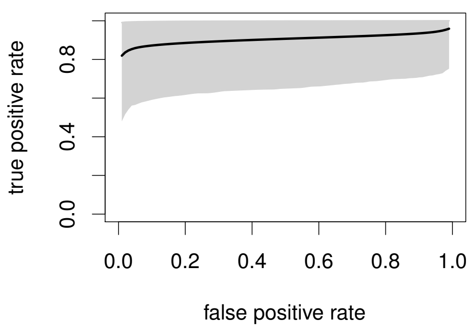
Figure 1: Simulated data: True (dotted line) and mean across simulations (solid line) of the posterior mean of the AUC function. A band constructed using the point-wise 2.5% and 97.5% quantiles across simulations is presented in gray. Panels (a)–(c), (d)–(f) and (g)–(i) display the results for Scenarios I, II and III under the B-splines DDP II mixture model.



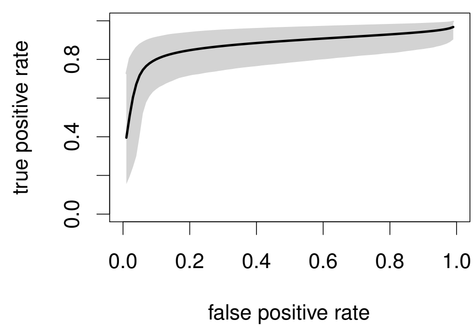
(a) ROC surface



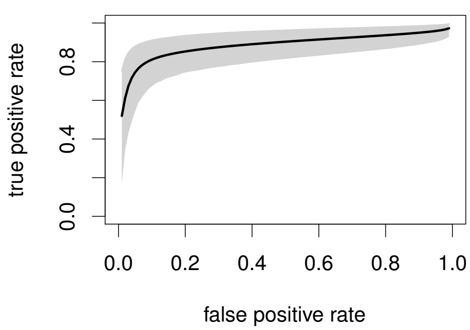
(b) AUC function



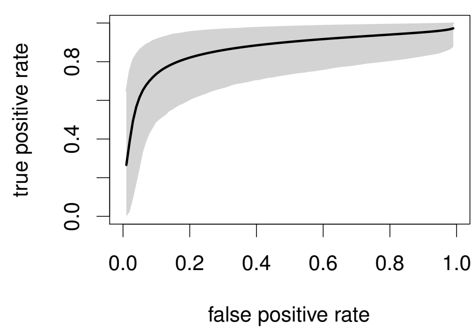
(c) 31 years old



(d) 43 years old

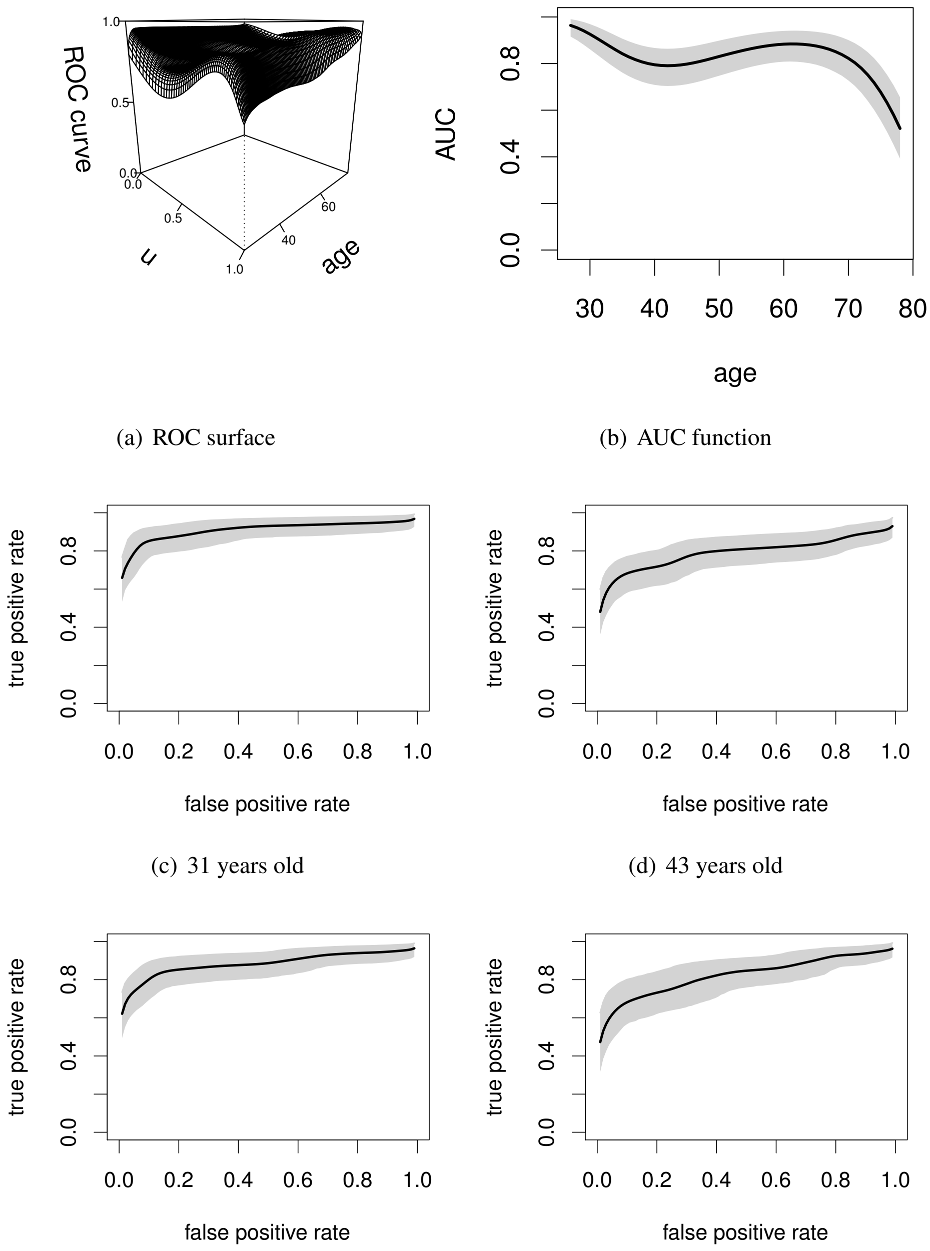


(e) 60 years old



(f) 70 years old

Figure 2: Results for glucose data under the B-splines DDP mixture model: Conditional ROC curve. Panel (a) displays the surface of the posterior mean of the conditional ROC curves across age. Panel (b) displays the posterior mean (solid line) and 95% point-wise HPD band for the area under the curve (AUC) as a function of the age. Panels (c), (d), (e) and (f) display the posterior mean and 95% point-wise HPD bands for the ROC cuve corresponding to the 5th, 25th, 75th and 95th quantiles of the empirical distribution of the age, respectively.



(a) ROC surface (b) AUC function

(c) 31 years old (d) 43 years old

(e) 60 years old (f) 70 years old

Figure 3: Results for glucose data under the B-splines DDP II mixture model: Conditional ROC curve. Panel (a) displays the surface of the posterior mean of the conditional ROC curves across age. Panel (b) displays the posterior mean (solid line) and 95% point-wise HPD band for the area under the curve (AUC) as a function of the age. Panels (c), (d), (e) and (f) display the posterior mean and 95% point-wise HPD bands for the ROC cuve corresponding to the 5th, 25th, 75th and 95th quantiles of the empirical distribution of the age, respectively.