Supplementary material

Bootstrap-based procedures for inference in nonparametric receiver-operating characteristic curve regression analysis

María Xosé Rodríguez-Álvarez[1,2,3], Javier Roca-Pardiñas [1],
Carmen Cadarso-Suárez[4] and Pablo G. Tahoces [5]

[1] Department of Statistics and Operations Research
and Biomedical Research Centre (CINBIO), University of Vigo, Vigo, Spain
[2] BCAM - Basque Center for Applied Mathematics, Bilbao, Spain
[3] IKERBASQUE, Basque Foundation for Science, Bilbao, Spain
[4] Center for Research in Molecular Medicine and Chronic Diseases (CiMUS),
University of Santiago de Compostela, Santiago de Compostela, Spain
[5] Centro Singular de Investigación en Tecnologías de la Información (CITIUS),
University of Santiago de Compostela, Santiago de Compostela, Spain

This document contains supplementary material to the paper "Bootstrap-based procedures for inference in nonparametric receiver-operating characteristic curve regression analysis". Additional simulation studies to complete those presented in the main manuscript are provided. More precisely, we present here the results when considering different distributions for simulating the diagnostic test result in healthy and diseased populations, and when covariates only affect the result of the diagnostic test in the diseased population.

Web Appendix A Simulation study with different distributions

Data were simulated from two scenarios, namely,

- Scenario wI

$$
\begin{aligned}
Y_{\bar{D}} &= -2X_{v1}^2 + 0.5\exp(X_{v2}) + \varepsilon_{\bar{D}},\\
Y_D &= aX_{v1}^2 - 2X_{v1}^2 + 0.5\sin(\pi(X_{v2}+1)) + 0.5\exp(X_{v2}) + \varepsilon_D.
\end{aligned}
$$

- Scenario wII

$$
\begin{aligned}
Y_{\bar{D}} &= -0.25X_{v1}^3 + 0.5X_{v1}^2 + 0.5X_{v1}^2 X_{u1} - 0.5X_{v1}^2(1 - X_{u1}) + \varepsilon_{\bar{D}},\\
Y_D &= 0.25X_{v1}^3 + (a+1)\left(0.5X_{v1}^2 + 0.5X_{v1}^2 X_{u1} - 0.5X_{v1}^2(1 - X_{u1})\right) + \varepsilon_D.
\end{aligned}
$$

In both cases, a is a real constant, X_{v1} and X_{v2} are simulated from a uniform distribution on $[-1, 1]$, and $X_{u1} \sim Bernoulli\,(0.5)$.

Bearing in mind the distributions of errors $\varepsilon_{\bar{D}}$ and ε_D the following situations are considered:

(a) $\varepsilon_{\bar{D}}$ and ε_D displaying Student's t distributions, both with a mean of zero and 12 degrees of freedom

(b) $\varepsilon_{\bar{D}}$ displaying a Gaussian distribution with a mean of zero an a standard deviation of 0.5, and ε_D displaying a mixture of Gaussian distributions, $\varepsilon_D \sim 0.5N\left(-0.5, 0.5^2\right) + 0.5N\left(0.5, 0.5^2\right)$

With the above configurations, the corresponding conditional ROC curves under the assumption of Student's t distributed errors are

- Scenario wIa

$$
ROC_{\boldsymbol{x}}(p) = T_{12}\left(a x_{v1}^2 + 0.5\sin(\pi(x_{v2}+1)) + T_{12}^{-1}(p)\right).
$$

- Scenario wIIa

$$
ROC_{\boldsymbol{x}}(p) = T_{12}\left(0.5x_{v1}^3 + 0.5a\left(x_{v1}^2 + x_{v1}^2 x_{u1} - x_{v1}^2(1 - x_{u1})\right) + T_{12}^{-1}\left(p\right)\right).
$$

where T_{df} denotes the cumulative distribution function for the standard Student's t distribution with df degrees of freedom. For the situation of mixture of Gaussian distributions, the conditional ROC curve for each scenario considered is:

- Scenario wIa

$$
ROC_{\boldsymbol{x}}(p) = F\left(a x_{v1}^2 + 0.5\sin(\pi(x_{v2}+1)) + 0.5\Phi^{-1}(p)\right).
$$

- Scenario wIIa

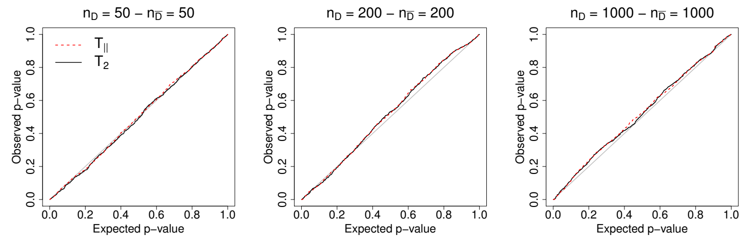
$$
ROC_{\boldsymbol{x}}(p) = F\left(0.5x_{v1}^3 + 0.5a\left(x_{v1}^2 + x_{v1}^2 x_{u1} - x_{v1}^2(1 - x_{u1})\right) + 0.5\Phi^{-1}\left(p\right)\right).
$$

where Φ denotes the cumulative distribution function of a standard normal random variable, and $F\left(c\right) = 0.5\Phi\left(\frac{c+0.5}{0.5}\right) + 0.5\Phi\left(\frac{c-0.5}{0.5}\right)$.

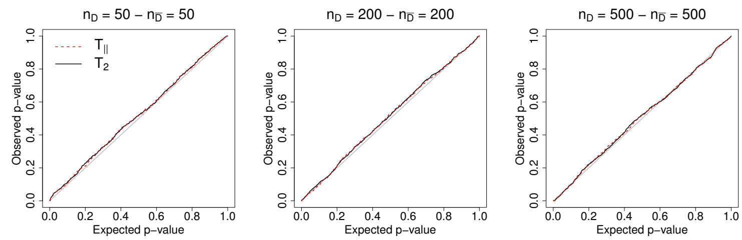
	Sample size		Test	Level					KS p-value
	n_D	$n_{\bar{D}}$		0.01	0.05	0.10	0.15	0.20	
Scenario wIa	50	50	$T_{\|\|}$	0.013	0.052	0.102	0.156	0.219	0.687
			T_2	0.013	0.058	0.113	0.168	0.220	0.407
	200	200	$T_{\|\|}$	0.015	0.051	0.112	0.159	0.205	0.007
			T_2	0.013	0.051	0.114	0.166	0.206	0.006
	1000	1000	$T_{\|\|}$	0.008	0.038	0.091	0.143	0.186	0.117
			T_2	0.009	0.049	0.087	0.138	0.187	0.087
Scenario wIb	50	50	$S_{\|\|}$	0.005	0.032	0.096	0.132	0.182	0.064
			S_2	0.004	0.031	0.085	0.138	0.178	0.063
	200	200	$S_{\|\|}$	0.008	0.049	0.106	0.150	0.189	0.192
			S_2	0.011	0.043	0.093	0.152	0.193	0.149
	1000	1000	$S_{\|\|}$	0.014	0.058	0.103	0.136	0.183	0.468
			S_2	0.017	0.054	0.100	0.141	0.180	0.263
Scenario wIIa	50	50	$S_{\|\|}$	0.011	0.047	0.087	0.143	0.193	0.001
			S_2	0.010	0.046	0.096	0.143	0.182	0.004
	200	200	$S_{\|\|}$	0.009	0.046	0.098	0.155	0.200	0.811
			S_2	0.007	0.042	0.096	0.148	0.195	0.909
	1000	1000	$S_{\|\|}$	0.014	0.053	0.095	0.141	0.180	0.055
			S_2	0.011	0.047	0.097	0.130	0.184	0.063
Scenario wIIb	50	50	$S_{\|\|}$	0.014	0.040	0.082	0.123	0.168	0.001
			S_2	0.014	0.038	0.083	0.124	0.166	0.002
	200	200	$S_{\|\|}$	0.012	0.050	0.082	0.124	0.185	0.253
			S_2	0.011	0.043	0.087	0.133	0.183	0.272
	1000	1000	$S_{\|\|}$	0.009	0.040	0.082	0.134	0.194	0.406
			S_2	0.009	0.037	0.082	0.133	0.187	0.405

Web Table 1: For Scenarios wIa and wIIa (Student's t distributions) and wIb and wIIb (mixture of Gaussian distributions): estimated type I error registered by the proposed tests under the null hypothesis, for different significance levels and sample sizes. The last column presents the p-values of the Kolmogorov-Smirnov test for uniformity of the observed p-values.

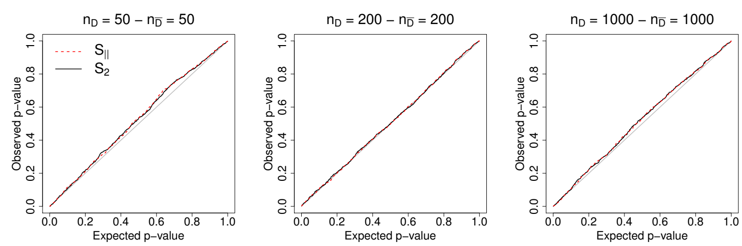
Table 1 shows the type I errors registered by the proposed tests for Scenarios wI and wII, for different significance levels and sample sizes. Figure 1 depicts quantile-quantile plots of the expected p-values (under the uniform distribution) and the observed p-values. As can be seen, the tests perform well in general, with type I errors proving to be relatively close to nominal errors (Table 3), and p-value distributions close to the uniform one (Figure 1). Table 2 shows the power of the tests at different significance levels for a specific value of a. As expected, the probability of rejection rises as the sample size increases.



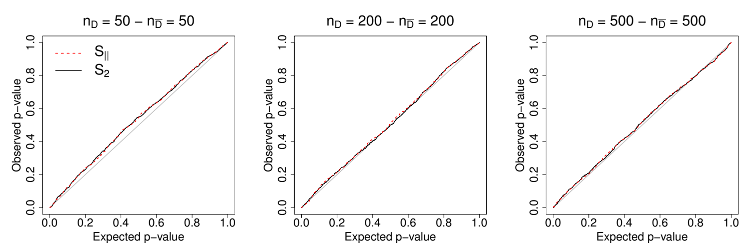
(a) Scenario wIa



(b) Scenario wIb



(c) Scenario wIIa



(d) Scenario wIIb

Web Figure 1: For Scenarios wIa and wIIa (Student's t distributions) and wIb and wIIb (mixture of Gaussian distributions): Quantile-quantile plot for the the observed p-values vs the expected p-values when the null hypothesis is correct.

		Sample size		Test	Level				
		n_D	$n_{\bar{D}}$		0.01	0.05	0.10	0.15	0.20
Scenario wIa	$a = 0.5$	50	50	$T_{\|\|}$	0.013	0.058	0.116	0.157	0.213
				T_2	0.013	0.057	0.122	0.177	0.219
		200	200	$T_{\|\|}$	0.049	0.133	0.206	0.272	0.328
				T_2	0.047	0.131	0.203	0.278	0.315
		1000	1000	$T_{\|\|}$	0.327	0.541	0.654	0.722	0.762
				T_2	0.336	0.566	0.684	0.734	0.778
Scenario wIb	$a = 0.5$	50	50	$T_{\|\|}$	0.017	0.062	0.117	0.180	0.224
				T_2	0.015	0.056	0.121	0.183	0.227
		200	200	$T_{\|\|}$	0.112	0.251	0.339	0.434	0.487
				T_2	0.097	0.228	0.339	0.409	0.469
		1000	1000	$T_{\|\|}$	0.858	0.932	0.960	0.977	0.983
				T_2	0.833	0.918	0.948	0.968	0.975
Scenario wIIa	$a = 1.0$	50	50	$S_{\|\|}$	0.016	0.053	0.111	0.168	0.215
				S_2	0.013	0.053	0.113	0.172	0.217
		200	200	$S_{\|\|}$	0.071	0.195	0.300	0.383	0.452
				S_2	0.072	0.183	0.295	0.375	0.444
		1000	1000	$S_{\|\|}$	0.528	0.737	0.832	0.873	0.902
				S_2	0.509	0.717	0.813	0.859	0.893
Scenario wIIb	$a = 1.0$	50	50	$S_{\|\|}$	0.019	0.073	0.133	0.194	0.257
				S_2	0.013	0.071	0.127	0.192	0.256
		200	200	$S_{\|\|}$	0.180	0.345	0.479	0.568	0.636
				S_2	0.173	0.338	0.462	0.563	0.638
		1000	1000	$S_{\|\|}$	0.939	0.980	0.991	0.998	1.000
				S_2	0.950	0.987	0.993	0.996	0.999

Web Table 2: For Scenarios wIa and wIIa (Student's t distributions) and wIb and wIIb (mixture of Gaussian distributions): estimated rejection probabilities registered by the proposed tests under the alternative hypothesis, for different significance levels and sample sizes.

Web Appendix B Simulation study with covariates affecting only the diseased population

Data were simulated from two scenarios, namely,

- Scenario wIII

$$Y_{\bar{D}} = 0.5\exp(X_{v2}) + 0.5\varepsilon_{\bar{D}},$$
$$Y_D = 0.5\sin(\pi(X_{v2}+1)) + 0.5\exp(X_{v2}) + aX_{v1}^2 + 0.5\varepsilon_D.$$

- Scenario wIV

$$Y_{\bar{D}} = -0.25X_{v1}^3 + 0.5\varepsilon_{\bar{D}},$$
$$Y_D = 0.25X_{v1}^3 + aX_{v1}^2X_{u1} + 0.5\varepsilon_D.$$

In both cases, a is a real constant, X_{v1} and X_{v2} are simulated from a uniform distribution on $[-1, 1]$, $X_{u1} \sim Bernoulli\,(0.5)$ and $\varepsilon_{\bar{D}}$ and $\varepsilon_D \sim N\,(0, 1)$. Note that in Scenario wIII the continuous covariate X_{v1} only affects the result of the diagnostic test in the diseased population. In much the same way, in Scenario wIV is the categorical covariate X_{u1} the one affecting only the result of the diagnostic test in the diseased population. Here, $a = 0$ corresponds to the hypothesis of no interaction between X_{v1} and X_{u1}, and as the value of a rises, so does the degree of interaction.

With the above configurations, the corresponding conditional ROC curves are

- Scenario wIII

$$ROC_{\boldsymbol{x}}(p) = \Phi\left(2ax_{v1}^2 + \sin(\pi(x_{v1} + 1)) + \Phi^{-1}(p)\right).$$

- Scenario wIV

$$ROC_{\boldsymbol{x}}(p) = \Phi\left(x_{v1}^3 + a\left(x_{v1}^2 + x_{v1}^2x_{u1} - x_{v1}^2(1 - x_{u1})\right) + \Phi^{-1}\left(p\right)\right).$$

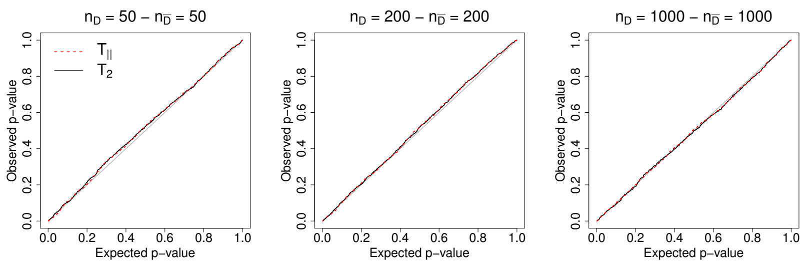
Table 3 shows the type I errors registered by the proposed tests for Scenarios wIII and wIV, for different significance levels and sample sizes. Figure 2 depicts quantile-quantile plots of the expected p-values (under the uniform distribution) and the observed p-values. As can be seen, the tests perform well in general, with type I errors proving to be relatively close to nominal errors (Table 3), and p-value distributions close to the uniform one (Figure 2). Table 4 shows the power of the tests at different significance levels for a specific value of a. As expected, the probability of rejection rises as the sample size increases.

	Sample size			Level					
	n_D	$n_{\bar{D}}$	Test	0.01	0.05	0.10	0.15	0.20	KS p-value
Scenario wIII	50	50	$T_{\|\|}$	0.012	0.054	0.089	0.141	0.198	0.211
			T_2	0.011	0.049	0.094	0.145	0.192	0.185
	200	200	$T_{\|\|}$	0.007	0.055	0.101	0.142	0.191	0.289
			T_2	0.012	0.054	0.094	0.138	0.197	0.326
	1000	1000	$T_{\|\|}$	0.011	0.051	0.102	0.153	0.203	0.956
			T_2	0.014	0.050	0.107	0.153	0.206	0.741
Scenario wIV	50	50	$S_{\|\|}$	0.006	0.052	0.114	0.156	0.197	0.882
			S_2	0.012	0.047	0.102	0.154	0.202	0.901
	200	200	$S_{\|\|}$	0.010	0.048	0.100	0.148	0.203	0.899
			S_2	0.015	0.053	0.109	0.148	0.200	0.550
	1000	1000	$S_{\|\|}$	0.009	0.052	0.098	0.149	0.194	0.363
			S_2	0.011	0.048	0.096	0.151	0.202	0.404

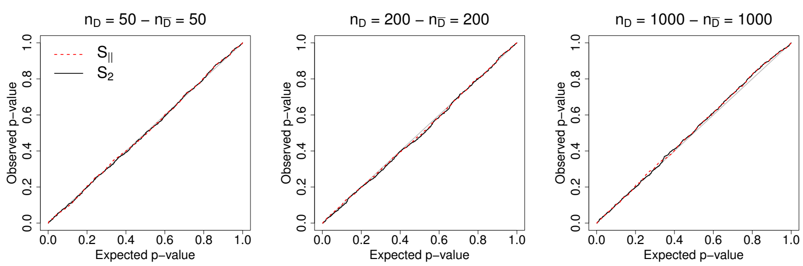
Web Table 3: For Scenarios wIII and wIV (covariates affecting only the diseased population): estimated type I error registered by the proposed tests under the null hypothesis, for different significance levels and sample sizes. The last column presents the p-values of the Kolmogorov-Smirnov test for uniformity of the observed p-values.

		Sample size			Level				
		n_D	$n_{\bar{D}}$	Test	0.01	0.05	0.10	0.15	0.20
Scenario wIII	$a = 0.5$	50	50	$T_{\|\|}$	0.066	0.179	0.274	0.349	0.416
				T_2	0.051	0.171	0.278	0.353	0.401
		200	200	$T_{\|\|}$	0.636	0.821	0.879	0.914	0.933
				T_2	0.611	0.782	0.869	0.906	0.921
		1000	1000	$T_{\|\|}$	1.000	1.000	1.000	1.000	1.000
				T_2	1.000	1.000	1.000	1.000	1.000
Scenario wIV	$a = 1.0$	50	50	$S_{\|\|}$	0.061	0.143	0.241	0.314	0.384
				S_2	0.058	0.139	0.218	0.303	0.374
		200	200	$S_{\|\|}$	0.563	0.783	0.864	0.914	0.944
				S_2	0.527	0.757	0.852	0.905	0.940
		1000	1000	$S_{\|\|}$	1.000	1.000	1.000	1.000	1.000
				S_2	1.000	1.000	1.000	1.000	1.000

Web Table 4: For Scenarios wIII and wIV (covariates affecting only the diseased population): estimated rejection probabilities registered by the proposed tests under the alternative hypothesis, for different significance levels and sample sizes.



(a) Scenario wIII



(b) Scenario wIV

Web Figure 2: For Scenarios wIII and wIV (covariates affecting only the diseased population): Quantile-quantile plot for the the observed p-values vs the expected p-values when the null hypothesis is correct.