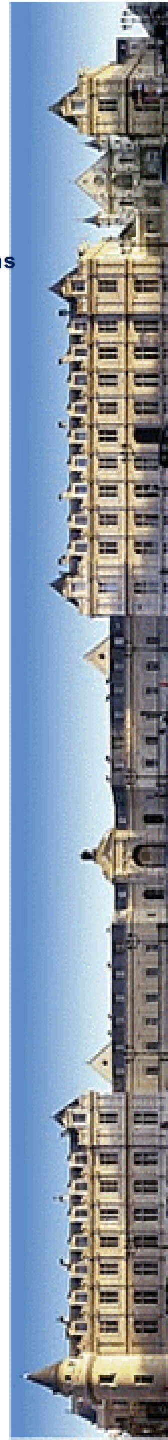
University of Minho
Dept. of Production and Systems Engineering



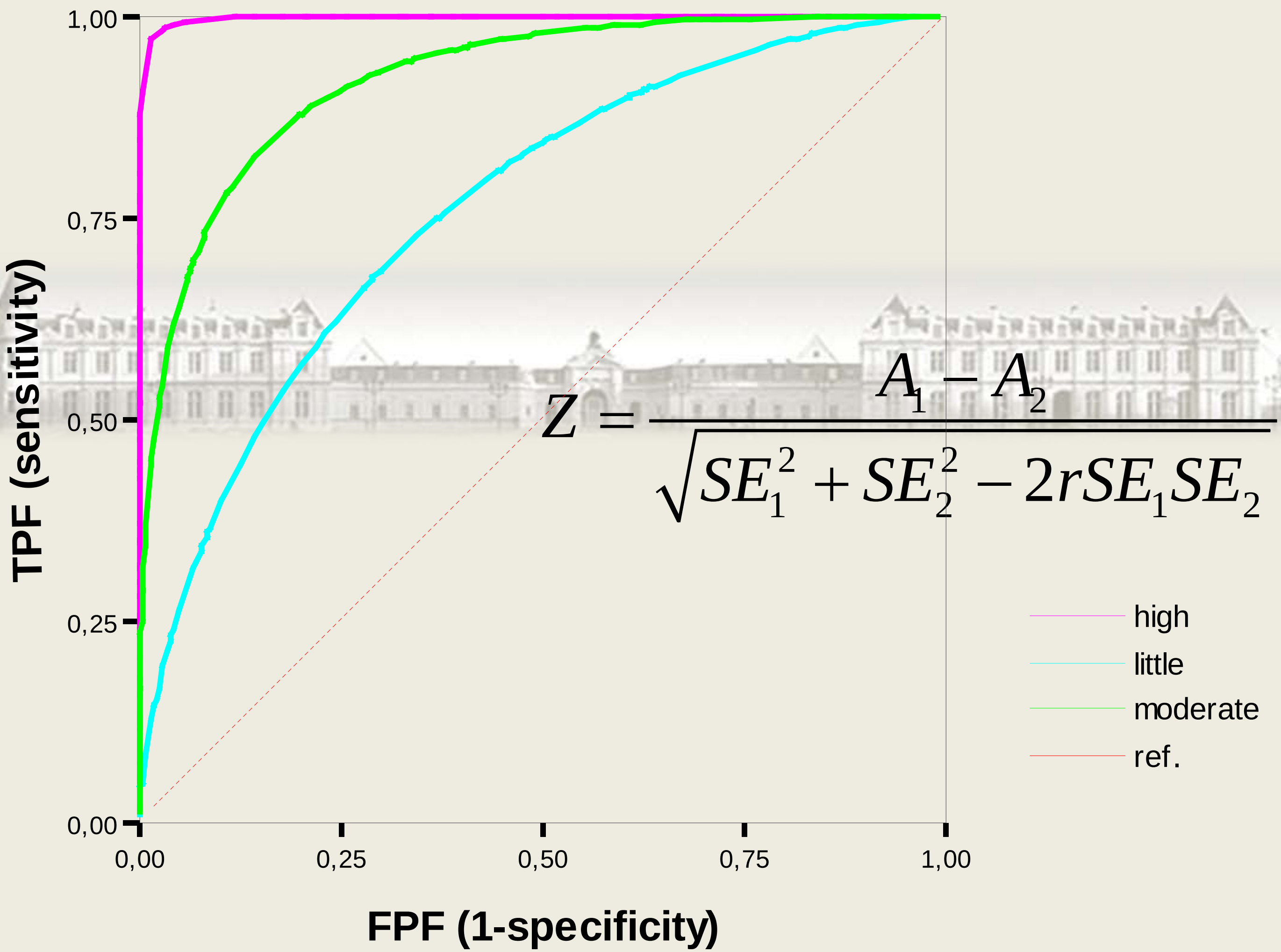
NON PARAMETRIC CONFIDENCE INTERVALS FOR ROC CURVES COMPARISON

Ana Cristina Braga: acb@dps.uminho.pt
Lino Costa: lac@dps.uminho.pt
Pedro Nuno Oliveira: pno@dps.uminho.pt

OBJECTIVES

- Development of a new methodology which allows the comparison of ROC curves that cross each other;
- Identification of the regions of the ROC space in which the tests have better performance;
- Construction of nonparametric confidence intervals for measures proposed.

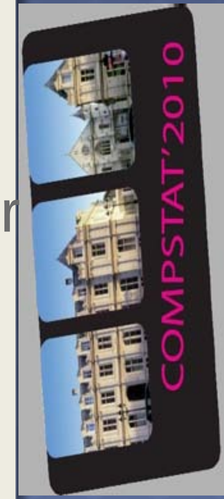
COMPARISON OF ROC CURVES



$$Z = \frac{A_1 - A_2}{\sqrt{SE_1^2 + SE_2^2 - 2rSE_1SE_2}} \sim N(0,1)$$

METHODOLOGY

1. **Sampling the ROC curves**
   - Sampling lines starting from a reference point
   - Intersection points of the sampling lines with the ROC curves
   - Euclidean distance from the intersection points to the reference point
2. **Measures**
   - Extension – proportion of the space where a curve is better than other
   - Location – regions of the space where a curve is better than other



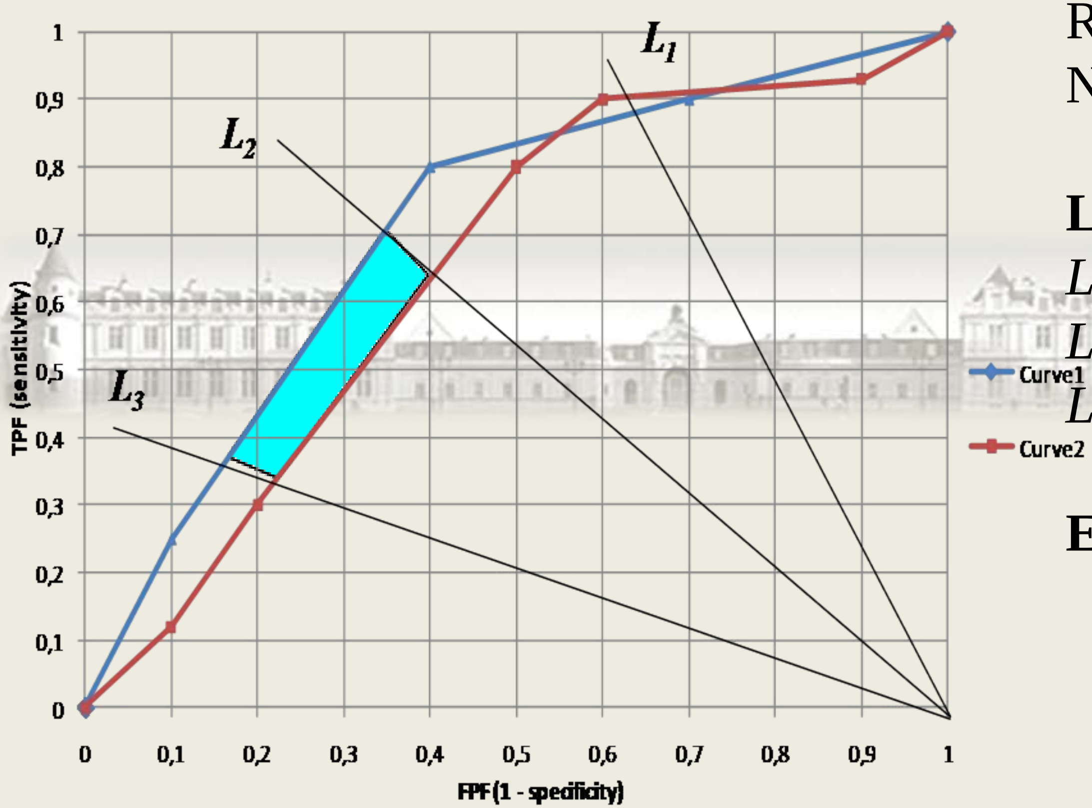
METHODOLOGY (CONT.)

3. Nonparametric statistical evaluation

- Statistical Evaluation of the Difference between Areas - ***Permutation test***
- Confidence Interval for the Difference of the areas ***- bootstrap resampling***

COMPSTAT'2010

EXTENSION MEASURE



Reference point: (1,0)
Number of sampling lines: 3

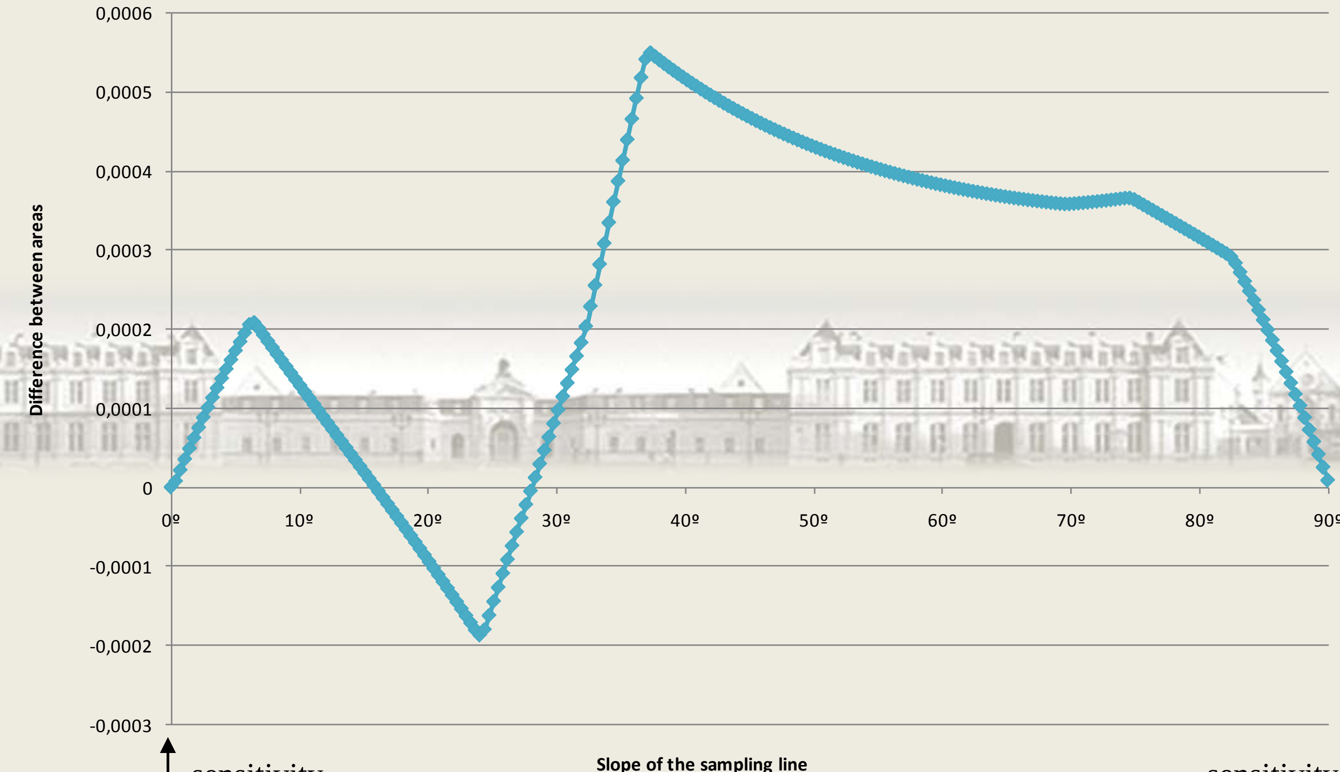
Line	Slope	Outcome
L_1	22.5º	Curve 2
L_2	45º	Curve 1
L_3	67.5º	Curve 1

Extension measure

Curve 1: 66.7 %
Curve 2: 33.3 %

COMPSTAT'2010

LOCATION MEASURE



↑ sensitivity ↓ specificity

sensitivity ↓ specificity ↑

	Extension measure	**Location measure**
Curve 1:	86.4 %	[0 °, 15.4°] and [28°, 90 °]
Curve 2:	13.6 %	[15.8°, 27.6 °]

COMPSTAT'2010

NONPARAMETRIC STATISTICAL TEST

- Based on the notion of permutation tests, the difference of the areas between the two empirical ROC curves are permuted;
- Bootstrapped confidence intervals are calculated;
- All computations performed using R *package.*

COMPSTAT'2010

SIMULATION STUDIES

Conditions:

- Generate distributions of abnormal ($f_A(x)$) and normal ($f_N(x)$) for two modalities;
- Greater values of variable x correspond to the abnormal status;
- $X_N \sim N(50, 25)$, $X_A \sim N(60, 25)$ and $n_A = n_N$;
- Sampling lines: $K = 100$.

SIMULATION (RESULTS)

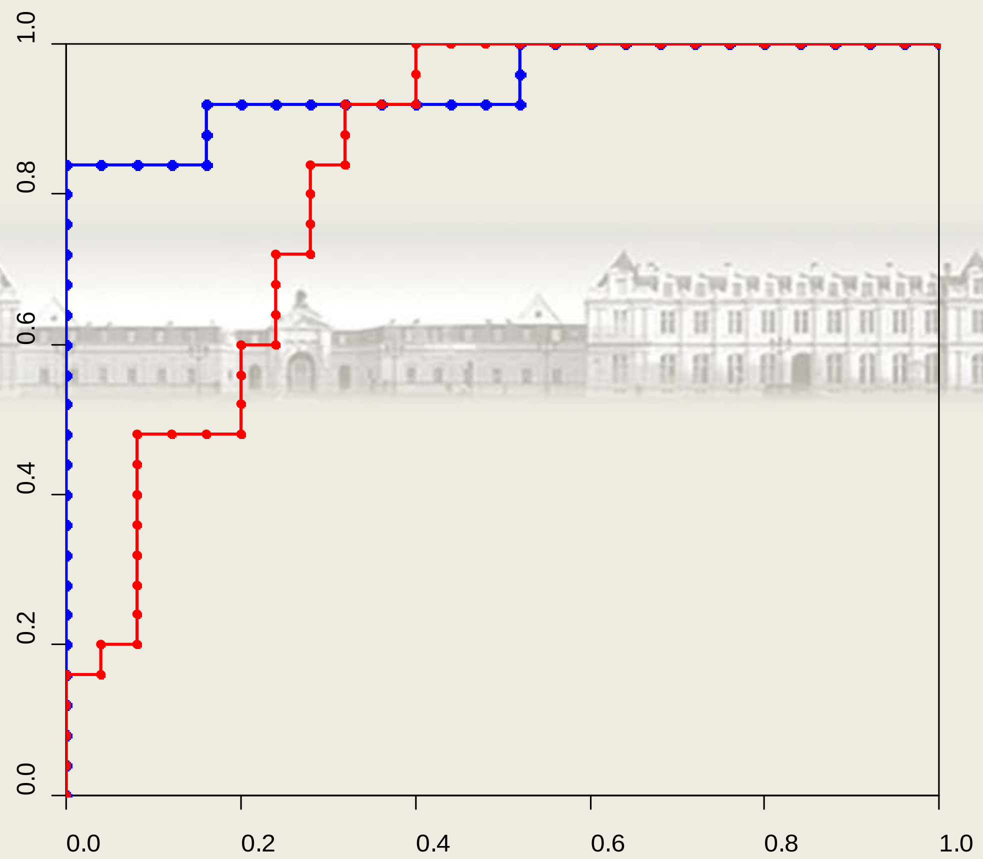
$n_A = n_N$		AUC1	SE1	AUC2	SE2	AUC1-AUC2
25	Mean	0.918	0.0384	0.925	0.0358	-0.00626
	Median	0.922	0.039	0.926	0.0363	-0.008
	minimum	0.813	0.0073	0.826	0.0023	-0.1248
	maximum	0.992	0.0657	0.998	0.0595	0.1664
50	Mean	0.924	0.0256	0.920	0.0264	0.00394
	Median	0.924	0.0259	0.921	0.0266	0.004
	minimum	0.816	0.0087	0.806	0.0130	-0.1236
	maximum	0.985	0.0428	0.971	0.0433	0.1236
100	Mean	0.922	0.0185	0.922	0.0183	-0.00048
	Median	0.923	0.0185	0.923	0.0181	0.0001
	minimum	0.867	0.0113	0.855	0.0107	-0.0834
	maximum	0.967	0.0253	0.965	0.0265	0.0649

SIMULATION (RESULTS)

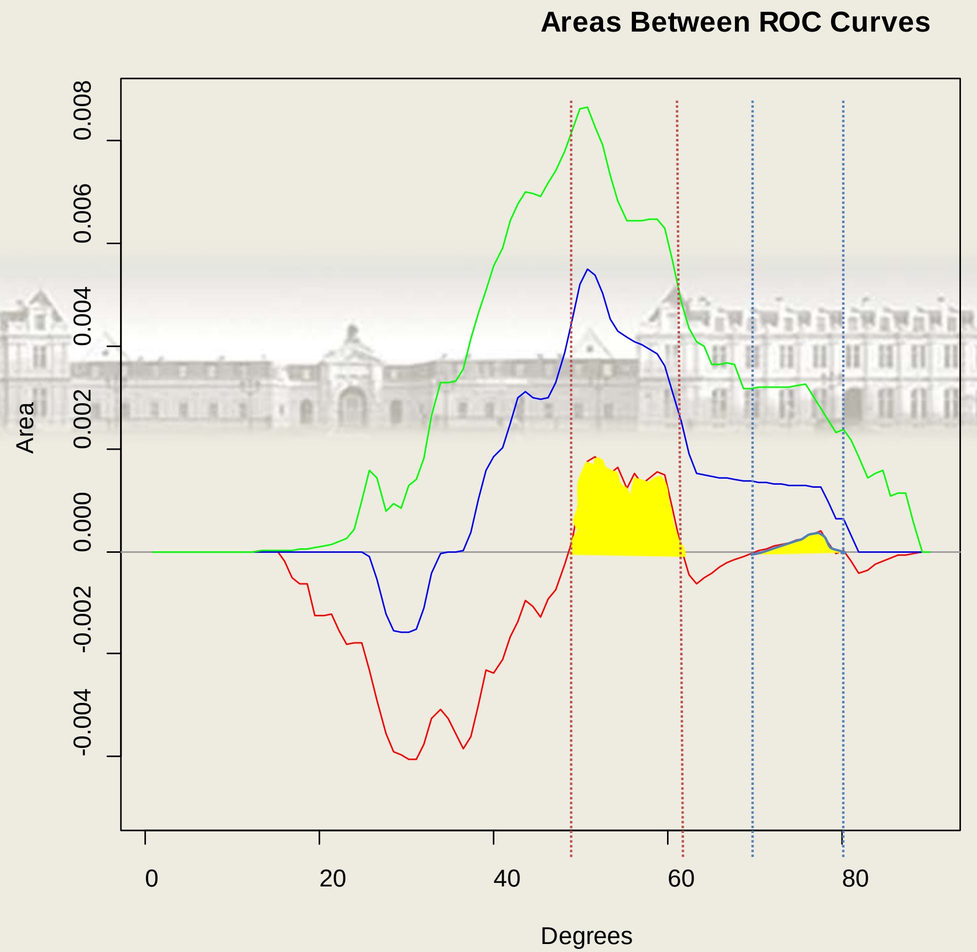
B Test \ Z Test	n	Rejection	No Rejection
Rejection	25	7	4
	50	7	1
	100	12	1
No Rejection	25	0	189
	50	1	191
	100	4	183

	# Cross	0	1	2	3	4	≥ 5
Freq.	n=25	31	61	60	29	18	1
	n=50	12	48	41	36	25	38
	n=100	10	31	30	41	32	56

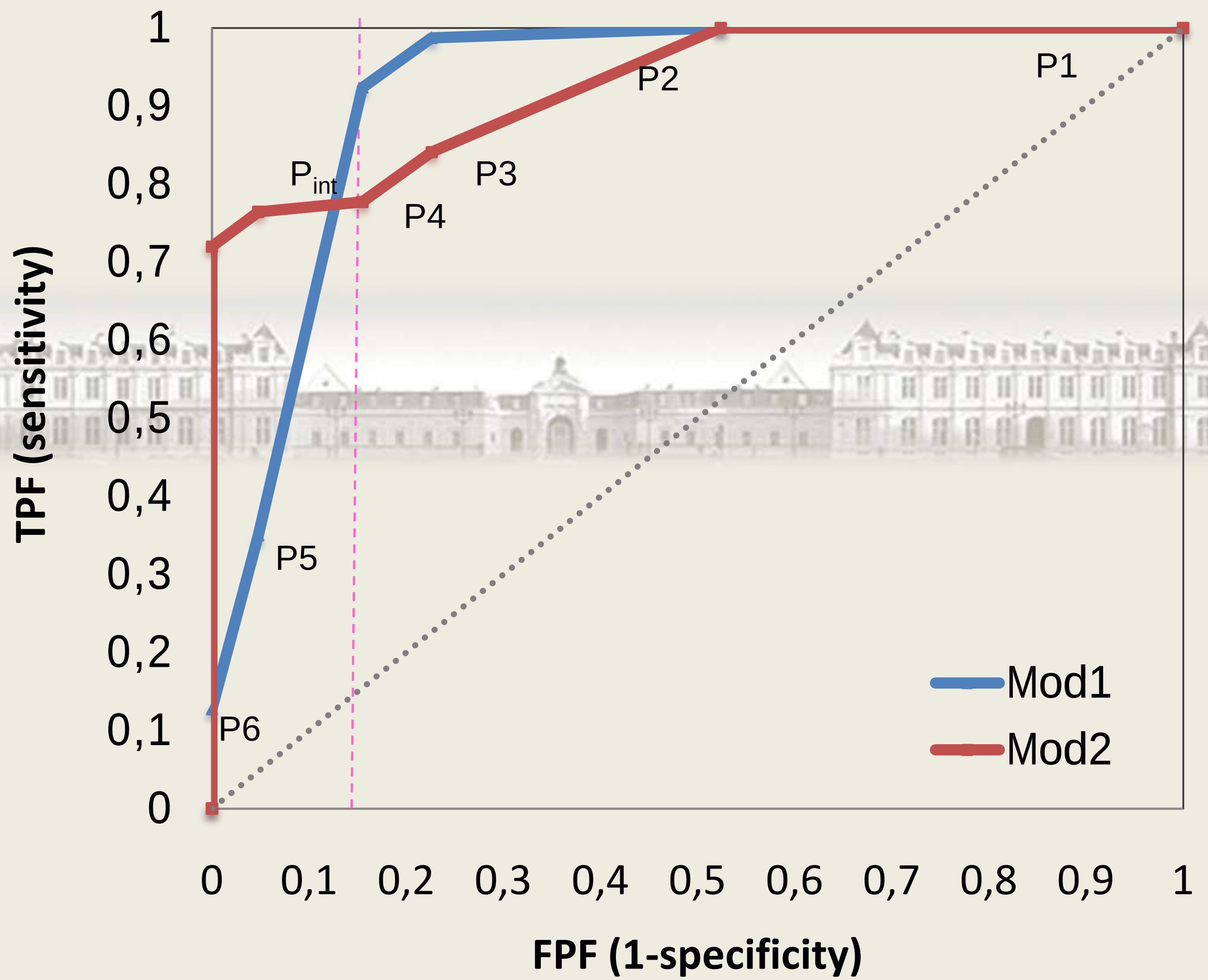
EMPIRICAL ROC CURVES (SIMULATION EXAMPLE)



BOOTSTRAP CI FOR DIFFERENCES



HYPOTHETICAL EXAMPLE (ZHANG)

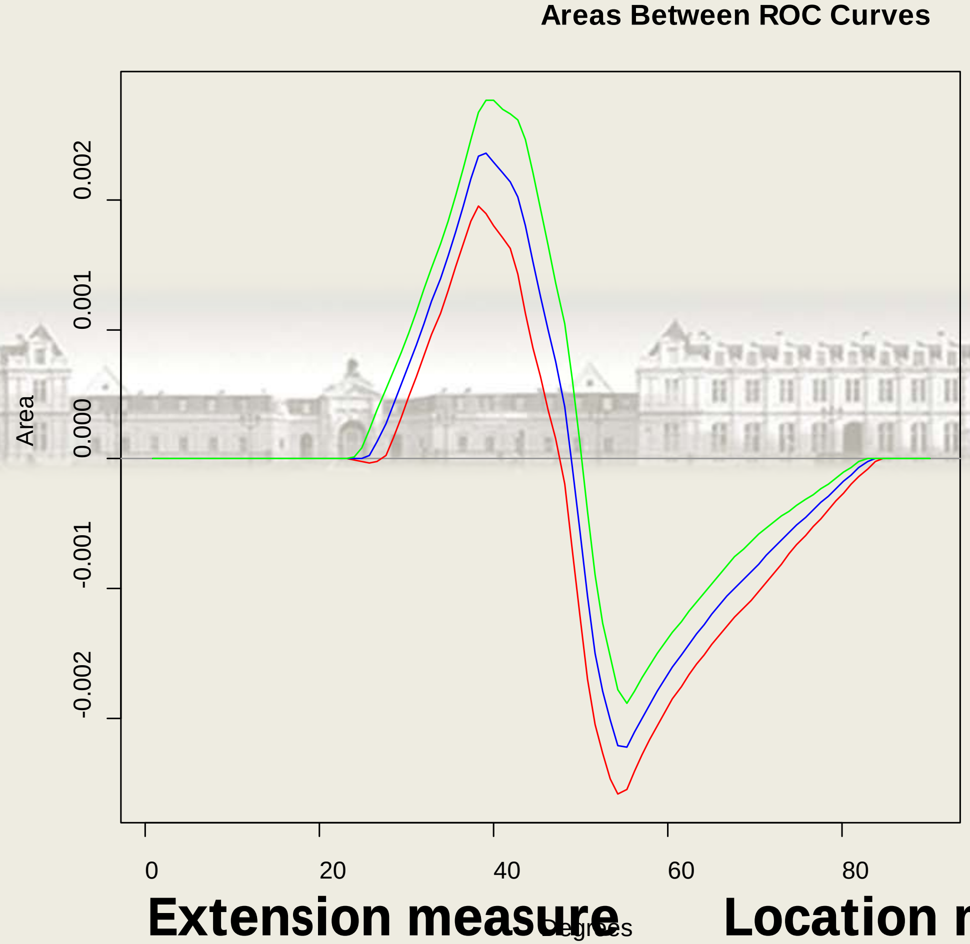


AUC1: 0.91978
SE(AUC1)=0.007
AUC2: 0.92580
SE(AUC2)=0.005
Diff: -0.00603

COMPSTAT'2010

HYPOTHETICAL EXAMPLE (ZHANG)

-0.02247822 < diff (boot)< 0.01017388



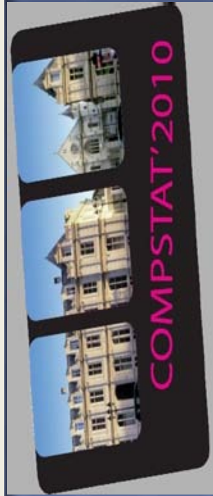
	Extension measure	Location measure
Curve 1:	25.8 %	]25.8°, 48.8°]
Curve 2:	38.6 %	]48.8°, 82.8°]

CONCLUSIONS

- The proposed methodology allows partial and global comparisons of two ROC curves without a fixing FPF;
- Graphical representation that elucidates the dominance regions in terms of sensitivity and specificity;
- Nonparametric alternative based on bootstrap resampling for the comparison of two ROC curves when they cross each other.

FUTURE WORK

- To study the randomness of the crossing points between ROC curves;
- To extend the methodology to the comparison of more than two ROC curves.



REFERENCES

- J. A. Hanley and B. J. McNeil. A Method of Comparing the Areas Under Receiver Operating Characteristic Curves Derived from the Same Cases. *Radiology, 148(3):839-843, 1983. ISSN 0033*-8419.
- E. R. DeLong, D. M. DeLong, and D. I. Clarkepearson. Comparing The Areas Under 2 or More Correlated Receiver Operating Characteristic Curves - A Nonparametric Approach. *Biometrics, 44*(3):837-845, Sep 1988. ISSN 0006-341x.
- H. E. Rockette, N. A. Obuchowski, and D. Gur. Nonparametric-Estimation of Degenerate ROC Data Sets Used For Comparison of Imaging-Systems. *Investigative Radiology, 25(7):835-837, Jul 1990.* ISSN 0020-9996.
- D. K. McClish. Analyzing a Portion of the ROC Curve. *Medical Decision Making, 9(3):190-195,* Jul-Sep 1989. ISSN 0272-989x.
- S. Wieand, M. H. Gail, B. R. James, and K. L. James. A Family of Nonparametric Statistics for Comparing Diagnostic Markers with Paired or Unpaired Data. *Biometrika, 76(3):585-592, Sep 1989.* ISSN 0006-3444.
- Y. L. Jiang, C. E. Metz, and R. M. Nishikawa. A receiver operating: Characteristic partial area index for highly sensitive diagnostic tests. *Radiology, 201(3):745-750, DEC 1996. ISSN 0033-8419. 1995* RSNA Scientific Assembly, CHICAGO, IL, NOV 26-DEC 01, 1995.
- D. D. Zhang, X. H. Zhou, D. H. Freeman, and J. L Freeman. A Non-Parametric Method for The Comparison of Partial Areas Under ROC Curves and Its Application to Large Health Care Data Sets. *Statistics In Medicine, 21(5):701-715, Mar 2002. ISSN 0277-6715.*
- L. E. Dodd and M. S. Pepe. Partial AUC estimation and regression. *Biometrics, 59(3):614-623, SEP* 2003. ISSN 0006-341X.
- C. M. Fonseca and P. J. Fleming. On the performance assessment and comparison of stochastic multiobjective optimizers. In *Proceedings of Parallel Problem Solving from Nature IV, pages 584*-593. Springer, 1996.