International
Workshop on ROC Analysis
and Related Topics

ROC2025

Programme
and
Book of Abstracts

Escola de Enxeñaría Industrial - Sede Cidade

Vigo – January 23 and 24, 2025

January 23, 2025

09:15–09:30 Welcome

09:30–11:30 Session I: Inferential topics in ROC analysis

09:30 *Comparing binary classification methods using covariate information*
Arís Fanjul-Hevia - Universidad de Oviedo, Spain

10:10 *Relationship between two-sample test statistics and ROC summary measures*
María del Carmen Pardo - Universidad Complutense de Madrid, Spain

10:50 *Use of the R packages `rocbc` and `trinROC` for inference after the Box-Cox transformation in the ROC framework (and an outlook for future research)*
Christos T. Nakas - University of Thessaly at Volos, Greece, and Inselspital Bern / University of Bern, Switzerland

---

11:30–12:00 Coffee break

12:00–14:00 Session II: Optimal tests and R packages

12:00 *Construction and evaluation of optimal diagnostic tests*
Ainesh Sewak - University of Zurich, Switzerland

12:40 *The gROC curve and the optimal classification*
Pablo Martínez-Camblor - Geisel School of Medicine at Dartmouth, USA

13:20 *R software for ROC analysis*
Sonia Pérez-Fernández - Universidad de Oviedo, Spain

14:00–15:00 Lunch

15:00–17:00 Session III: Survival data and time-dependent ROC curves

15:00 *The conditional C-index for survival data with a cure fraction*
Juan Carlos Pardo-Fernández - Universidade de Vigo, Spain

15:40 *Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks*
Paul Blanche - University of Copenhagen, Denmark

16:20 *A new AUC(t) estimator and its statistical properties for competing risks models*
Leire Garmendia - Basque Center for Applied Mathematics, Spain

January 24, 2025

09:30–11:30 Session IV: Length of the ROC curve and three-class problems

09:30 *A parametric inference approach for the length of the ROC curve*
Mónica López-Ratón - Universidade de Santiago de Compostela, Spain

10:10 *A novel overlap-based framework for assessing classifiers of three diagnostic groups*
Alba M. Franco-Pereira - Universidad Complutense de Madrid, Spain

10:50 *Biomarker cutoff estimation and inferences around the associated true and false class rates for the three-class problem. A misclassification cost-based framework*
Leonidas Bantis - University of Kansas Medical Center, USA

11:30–12:00 Coffee break

12:00–14:00 Session V: Complex sampling schemes and missing data

12:00 *Dealing with the estimation of the ROC curve and the area under the curve in the presence of complex sampling design data*
Amaia Iparragirre - UPV/EHU University of the Basque Country, Spain

12:40 *AUC optimism correction with missing data*
Susana Martins - Universidade de Vigo, Spain

13:20 *Prediction in the presence of missing values: no credible alternative to imputation-based use of the predictive density*
Bart J. A. Mertens - Leiden University Medical Centre, The Netherlands

14:00–15:00 Lunch

15:00–17:00 Session VI: ROC curves with functional data

15:00 *Addressing robust estimation in covariate-specific ROC curves with functional covariates*
Ana M. Bianco - Universidad de Buenos Aires, Argentina

15:40 *Optimal cut-point estimation for functional digital biomarkers: Application to continuous glucose monitoring*
Óscar Lado-Baleato - Health Research Institute of Santiago de Compostela (IDIS), Spain

16:20 *ROC curve analysis for functional biomarkers*
Graciela Boente - Universidad de Buenos Aires, Argentina

17:00–17:15 Farewell

Abstracts[1]

List of abstracts

Biomarker cutoff estimation and inferences around the associated true and false class rates for the three-class problem. A misclassification cost-based framework

<u>Leonidas Bantis</u>[1], Peng Shi[1] and John Tsimikas[2]

[1]Department of Biostatistics and Data Science, University of Kansas Medical Center, Kansas, USA

[2]Department of Statistics and Actuarial-Financial Mathematics, University of the Aegean, Samos, Greece

Abstract

The ROC surface is often used as a statistical tool for the evaluation of continuous biomarkers in trichotomous settings. Recent literature has focused on cutoff estimation in such settings and involves the extensions of the popular Youden index and the distance from the perfection corner among other previously used criteria. Under a trichotomous setting, the ROC does not visualize all the true and false classification rates (tradeoffs), as opposed to the usual ROC curve in the common 2-class problem. As a partial consequence, under a 3-class setting, the corresponding false class rates have not received sufficient attention in terms of estimation and inferences. In addition, most of the presented cutoff estimation approaches operate under the assumption that miss-classification costs are equal. In this work, we present the recent literature, and then provide an inferential framework for inferences around the cutoffs, the true class rates, and the false class rates in trichotomous settings. We illustrate how mis-classification costs can be accounted for along with the underlying prevalences. We demonstrate parametric and a non-parametric approaches as well as a newly developed comprehensive R package that accommodates several options involving, apart from our approaches, other known methods. For illustration purposes, we will also discuss an application involving pancreatic ductal adenocarcinoma patients.

Addressing robust estimation in covariate–specific ROC curves with functional covariates

Ana M. Bianco[1] and Graciela Boente[2]

[1]Instituto de Cálculo, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires and CONICET, Argentina
[2]Departamento de Matemáticas and Instituto de Cálculo, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires and CONICET, Argentina

Abstract

In some situations the presence of functional covariates related to the diagnostic marker may increase the discriminating power of the ROC curve. In this talk, the covariate effect is modelled through a functional linear regression model.

Given the extended belief that ROC curves are robust and aware of the impact that outlying values may have on the diagnostic test accuracy, we center our attention on the robust aspects of the conditional ROC curve estimation procedures. Since regression models are involved, atypical data among the responses and/or the covariates may severely affect the estimation methods. Moreover, when dealing with functional data the detection of outliers is more complex, since, in such a situation, different types of atypical data may arise.

We will present robust procedures when there are functional covariates and provide some results regarding the uniform consistency of the estimators. The finite-sample numerical study illustrates the robustness of the proposal.

Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks

Paul Blanche[1]

(based on joint work with Jérémie Riou[2] and Hélène Jacqmin-Gadda[3])

[1] Section of Biostatistics, Department of Public Health, University of Copenhagen, Denmark
[2] UMR INSERM 1066, CNRS 6021, Université d'Angers, France
[3] Univ. Bordeaux, INSERM, Bordeaux Population Health, U1219, France

Abstract

The area under the time-dependent ROC curve (AUC) may be used to quantify the ability of a biomarker to predict the onset of a clinical outcome in the future. In this talk, we present methods to make inference with typical survival or competing risks data. We present a simple nonparametric inverse probability of censoring weighting estimator, its asymptotic properties and how these properties can be used to make inference. First, we show how to compute confidence intervals and test statistics for comparing two AUCs. Second, we present an efficient method to control the family-wise error rate when multiple comparisons are performed. The method consists of using a max-t test and the closed testing procedure. This approach takes into account both the correlations between the test statistics and the logical constraints between the null hypotheses. We show how the method works in practice through examples where we compare several psychometric scores to predict the t-year risk of Alzheimer's disease. This illustrates several interesting multiple testing settings and shows some advantages of using the proposed method over common alternatives.

References

Blanche, P., Dartigues, J-F., and Riou, J. A closed max-t test for multiple comparisons of areas under the ROC curve. *Biometrics* 78.1 (2022): 352-363.

Blanche, P., Dartigues, J-F, and Jacqmin-Gadda, H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Statistics in medicine* 32.30 (2013): 5381-5397.

ROC curve analysis for functional biomarkers

Graciela Boente[1], Ana M. Bianco[2] and Juan Carlos Pardo-Fernández[3]

[1]Departamento de Matemáticas and Instituto de Cálculo, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires and CONICET, Argentina

[2]Instituto de Cálculo, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires and CONICET, Argentina

[3]Departamento de Estatística e Investigación Operativa and CITMAga, Universidade de Vigo, Spain

Abstract

Functional markers become an increasingly important tool in medical diagnostics. In this talk, we introduce a novel index designed to discriminate between populations based on functional data that belong to a Hilbert space.

We address the challenges involved in estimating optimal directions that maximize the area under the curve of a projection index, and we construct the corresponding ROC curve.

Additionally, we extend our approach to scenarios involving populations with potentially different covariance operators, where we propose a quadratic discrimination rule. We study the consistency for both linear and quadratic indices under mild assumptions.

We present the results of a numerical experiment, which demonstrates the advantages of the quadratic rule when populations have different covariance structures. Finally, we apply the proposed methods to a real-world data set, showcasing their behavior in a practical diagnostic context.

Comparing binary classification methods using covariate information

Arís Fanjul-Hevia[1], Juan Carlos Pardo-Fernández[2] and Wenceslao González-Manteiga[3]

[1]Departamento de Estadística e Investigación Operativa y Didáctica de la Matemática, Universidad de Oviedo, Spain

[2]Departamento de Estatística e Investigación Operativa and CITMAga, Universidade de Vigo, Spain

[3]Departamento de Estatística, Análise Matemática e Optimización and CITMAga, Universidade de Santiago de Compostela, Spain

Abstract

The Receiver Operating Characteristic (ROC) curve is a statistical tool that evaluates how well a binary classification method can distinguish between two categories – for example, in diagnostic problems. When multiple classification methods are available, comparing their corresponding ROC curves is a standard approach to assess their performance.

When additional covariate information is available, it should be incorporated into the ROC curve comparison since covariates can affect the discriminatory capability of the diagnostic methods. This incorporation can be achieved through either covariate-specific or covariate-adjusted ROC curves. Our work first addresses which curve type is most appropriate for the different situations that may arise. Based on these findings, we then propose several tests for comparing these curves in both paired and independent data scenarios.

References

Fanjul-Hevia, A., Gonźlez-Manteiga, W., Pardo-Fernńdez, J.C. (2021) A non-parametric test for comparing conditional ROC curves. *Computational Statistics & Data Analysis*, 157, 107146.

Fanjul-Hevia, A., Pardo-Fernández, J. C., González-Manteiga, W. (2024). A new test for assessing the covariate effect in ROC curves. *Preprint* arxiv.org/abs/2411.17464.

Fanjul-Hevia, A., Pardo-Fernández, J. C., Van Keilegom, I., González-Manteiga, W. (2022). A test for comparing conditional ROC curves with multidimensional covariates, *Journal of Applied Statistics*, 51:1, 87-113.

A novel overlap-based framework for assessing classifiers of three diagnostic groups

Víctor M. Sierra Marquina[1], M.Carmen Pardo[2,3] and <u>Alba M. Franco-Pereira</u>[2,3]

[1]Complutense University of Madrid, Spain

[2]Depeartment of Statistics and OR, Complutense University of Madrid, Spain

[3]Instituto de Matemática Interdisciplinar, Complutense University of Madrid, Spain

Abstract

This study presents a novel methodology aimed at finding potential biomarkers for improving diagnostic accuracy in multiclass classification problems, addressing challenges in distinguishing populations divided into three distinct groups rather than the traditional binary setup. Existing tools, such as ROC surfaces and the Volume Under the ROC Surface (VUS), are limited in their ability to fully capture the complexities of such scenarios. Pardo and Franco (2024) explored the advantages of the Overlap measures (OVL) over the ROC summary indices to assess the accuracy of a medical diagnostic test in the binary case. This work introduces parametric and non-parametric approaches for estimating OVL in the case of three class diagnostic problems and evaluates their performance through extensive simulations, comparing them to VUS. The methodology's utility is demonstrated using a biomarker dataset related to Alzheimer's progression, highlighting its potential to improve early-stage diagnosis. Such advancements are essential for enabling timely interventions, facilitating family planning, and enhancing the quality of life for patients at risk of severe neurodegenerative conditions.

References

Pardo, M.C. and Franco-Pereira, A.M. (2024). Overlap measures against ROC summary indices. *Statistical Science*, in press.

A new AUC(t) estimator and its statistical properties for competing risks models

Leire Garmendia Bergés[1,2], Irantzu Barrio[2] and Guadalupe Gómez Melis[3]

[1]BCAM - Basque Center for Applied Mathematics

[2]Department of Mathematics, University of the Basque Country UPV/EHU

[3]Department of Statistics and Operations Research, Universitat Politècnica de Catalunya

Abstract

We have developed a competing risks model for patients with severe pneumonia who can transition to two competing events: the need for non-invasive or invasive mechanical ventilation (NIMV or IMV). This model can be used to predict the course of the disease for new patients, therefore measuring its predictive capacity is of great interest, for which the area under the time-dependent ROC curve (AUC(t)) parameter could be used. Different estimators for the partial AUC(t) of each transition $(AUC_k(t), k = 1, \ldots, K)$ have been proposed in the literature. In this work, we propose an estimator for the global AUC(t) for a competing risk model $(\widehat{AUC}_{CR}(t))$ as a weighted sum of $\widehat{AUC}_k(t)(t), k = 1, \ldots, K$ with each $AUC_k(t)$ being weighted by the probability of experiencing that event k before time t. We have proved that $\widehat{AUC}_{CR}(t)$ is consistent and asymptotically normal.

Dealing with the estimation of the ROC curve and the area under the curve in the presence of complex sampling design data

Amaia Iparragirre[1], Irantzu Barrio[1,2] and Inmaculada Arostegi[1,2]

[1]Department of Mathematics, University of the Basque Country UPV/EHU
[2]BCAM - Basque Center for Applied Mathematics

Abstract

Complex survey data is becoming increasingly popular in a variety of fields, including health and social sciences. This type of data is collected from a finite population, concerned to be studied, following some complex sampling design, such as stratification or clustering. Due to the particularities of the data collection process, each sampled unit is assigned a sampling weight indicating the number of units it represents in the finite population. Thus, the straightforward application of the most commonly applied statistical techniques, which are typically designed to be applied to simple random samples, is usually not suitable in this context of complex survey data.

In particular, in this work, we focus on the evaluation of the discrimination ability of logistic regression models by means of the receiver operating characteristic (ROC) curve and the area under it (AUC). We propose an estimator for the estimation of the ROC curve ($\widehat{ROC_w}$) and AUC ($\widehat{AUC_w}$) that account for complex sampling designs. In addition, it is well known that when the same data is used to fit the model and estimate its AUC, this estimate overestimates the real discrimination ability of the fitted model. Thus, we propose to correct for the optimism of the $\widehat{AUC_w}$ by means of replicate weights methods such as the design-based cross-validation and the rescaling bootstrap in the context of complex survey data. The proposed methods have been validated by means of several simulation studies.

All the methods proposed to estimate the ROC curve and the AUC, as well as, the AUC correction methods are available in the R-package svyROC in CRAN (https://cran.r-project.org/web/packages/svyROC/index.html).

References

Iparragirre, A., Barrio, I., Arostegui, I. (2023). Estimation of the ROC curve and the area under it with complex survey data. *Stat*, 12(1). https://doi.org/10.1002/sta4.635

Iparragirre, A., Barrio, I. (2024). Optimism Correction of the AUC with Complex Survey Data. In: Einbeck, J., Maeng, H., Ogundimu, E., Perrakis, K. (eds) *Developments in Statistical Modelling. IWSM 2024. Contributions to Statistics.* Springer, Cham. https://doi.org/10.1007/978-3-031-65723-8_7

Optimal cut-point estimation for functional digital biomarkers: Application to continuous glucose monitoring

<u>Óscar Lado-Baleato</u>[1,2], Marcos Matabuena[3], Carla Díaz-Louzao[4] and Francisco Gude[5]

[1] Research Methods Group (RESMET), Health Research Institute of Santiago de Compostela (IDIS), Santiago de Compostela, Galicia, Spain.

[2] ISCIII Support Platforms for Clinical Research, Health Research Institute of Santiago de Compostela (IDIS), Santiago de Compostela, Galicia, Spain.

[3] Department of Biostatistics, Harvard University, Boston, MA 02115, USA.

[4] Department of Mathematics, University of Coruña, A Coruña, Galicia, Spain.

[5] Primary Care Centre, Concepción Arenal, Santiago de Compostela, Galicia, Spain.

Abstract

In digital health, defining optimal cut points for functional biomarkers is crucial for enhancing diagnostic accuracy, particularly when working with continuous data from wearable devices. In this communication, we present a mathematical model for determining optimal cut points for the distributional representation of continuous glucose monitoring (CGM) data, known as glucodensities. The model defines optimal cut points by introducing a continuous parameter that partitions the Hilbert space of glucodensity functions into regions corresponding to health and disease states. The method was validated using real-world data from the A-Estrada Glycation and Inflammation Study (AEGIS) cohort, a large population-based study which includes CGM data. It demonstrated high efficacy in distinguishing diabetes prevalence (AUC = 0.94) and it predicted diabetes incidence with fair accuracy (AUC = 0.78).

References

Matabuena, M., Petersen, A., Vidal, J. C., & Gude, F. (2021). Glucodensities: A new representation of glucose profiles using distributional data analysis. *Statistical Methods in Medical Research*, 30, 1445-1464.

A parametric inference approach for the length of the ROC curve

Elisa M. Molanes-López [1], Mónica López-Ratón[2] and Emilio Letón [3]

[1]Department of Statistics and Operations Research, Universidad Complutense de Madrid, Spain
[2]Department of Statistics and Operations Research, Universidade de Santiago de Compostela, Spain
[3]Department of Artificial Intelligence, Universidad Nacional de Educación a Distancia, Spain

Abstract

From the ROC curve, there have been introduced in the literature several summary indexes to measure the discrimination capacity of a continuous diagnostic test or biomarker, such as the area under the ROC curve (AUC), the Youden's index (J), or, more recently, the length of the ROC curve (LoC). In this talk we focus on the LoC index due to its good properties compared to the AUC and J indices, and we propose a point estimator and a confidence interval (CI) based on the Generalized Pivotal Quantity (GPQ) method. Our simulation results show that the new point estimator has a good performance in terms of root mean squared error, bias and standardized bias, and that our GPQ CI works well in terms of coverage probability and average width. In addition, we apply this new methodology to real biomedical data on cancer disease.

References

Bantis, L. E., Tsimikas, J. V., Chambers, G. R., Capello, M., Hanash, S., Feng, Z. (2021). The length of the receiver operating characteristic curve and the two cutoff Youden index within a robust framework for discovery, evaluation, and cutoff estimation in biomarker studies involving improper receiver operating characteristic curves. *Stat Med.*, 40(7), 1767-1789.

Franco-Pereira, A. M., Nakas, C., Pardo, M. C. (2020). Biomarker assessment in ROC curve analysis using the length of the curve as an index of diagnostic accuracy: the binormal model framework. *AStA Adv Stat Anal.*, 104, 625–647.

Hanley, J. A., McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Diagn Radiol.*, 143(1), 29-36.

López-Ratón, M., Cadarso-Suárez, C., Molanes-López, E.M. and Letón, E. (2016). Confidence intervals for the symmetry point: an optimal cutpoint in continuous diagnostic tests. *Pharm. Stat.*, 15(2), 178-192.

Martínez-Camblor, P. (2023). Confidence intervals for the length of the receiver-operating characteristic curve based on a smooth estimator. *Stat Methods Med Res.*, 32(5), 978-993.

Molanes-López, E. M., Letón, E. (2011). Inference of the Youden index and associated threshold using empirical likelihood for quantiles. *Stat Med.*, 30(19), 2467-2480.

Weerahandi, S. (1993). Generalized confidence intervals. *JASA.*, 88(423), 899-905.

Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer*, 3, 32-35.

The gROC curve and the optimal classification

Pablo Martínez-Camblor[1]

[1]Department of Anesthesiology and Biomedical Data Science Department, Geisel School of Medicine at Dartmouth, NH, USA

Abstract

The binary classification problem (BCP) aims to correctly allocate subjects in one of two possible groups. The groups are frequently defined as having or not one characteristic of interest. With this goal, we are allowed to use different types of information. There is a huge number of methods dealing with this problem; including standard binary regression models, or complex machine learning techniques such as support vector machine, boosting, or perceptron, among others. When this information is summarized in a continuous score, we have to define classification regions (or subsets) which will determine whether the subjects are classified as positive, with the characteristic under study, or as negative, otherwise. The standard (or regular) receiver-operating characteristic (ROC) curve considers classification subsets in the way $[c, \infty)$ ($c \in \mathbb{R}$), and plots the true- against the false- positive rates (sensitivity against one minus specificity). The so-called generalized ROC curve, gROC, allows that both higher and lower values of the score were associated with higher probabilities of being positive. Besides, the efficient ROC curve, eROC, considers the optimal use of the scores without considering the potential impact on the associated classification subsets. In this document, we are interested in studying, comparing and approximating the transformations leading to the eROC and to the gROC curves. We will prove that, when the optimal transformation does have no relative maximum, both curves are equivalent. Besides, we investigate the use of the gROC curve on some theoretical models, explore the relationship between the gROC and the eROC curves, and propose two non-parametric procedures for approximating the transformation leading to the gROC curve. The finite-sample behavior of the proposed estimators is explored through Monte Carlo simulations. Two real-data sets illustrate the practical use of the proposed methods.

AUC optimism correction with missing data

Susana Rafaela Martins[1], María del Carmen Iglesias-Pérez[2] and Jacobo de Uña-Álvarez[2]

[1]Escola Superior de Desporto e Lazer, Instituto Politécnico de Viana do Castelo, Portugal
[2]CITMAga, SiDOR Research Group and Department of Statistics and Operations Research, Universidade de Vigo, Spain

Abstract

The Area Under the ROC Curve (AUC) plays an important role in the study of the predictive capacity of regression models. It is well known that an inflated AUC may result when the same data are used for training and testing the model. In this paper the correction for the optimism of the AUC in the presence of missing data is investigated. More specifically, split-sample, K-fold cross-validation and leave-one-out cross-validation are adapted to missing data under MCAR and MAR assumptions to introduce optimism corrections. Complete case analysis, inverse probability weighting and multiple imputation are employed to address the issue of missing data. The methods are compared through intensive Monte Carlo simulations in the particular setting of logistic regression. Results suggest that all the methods perform well, leave-one-out cross-validation being generally the best. Among the several strategies to cope with missing data, multiple imputation is recommended. A real data illustration is provided.

Prediction in the presence of missing values: no credible alternative to imputation-based use of the predictive density

(Predictive averaging, Complete Cases, Indicator Methods and Pattern Submodels)

Bart J. A. Mertens[1]

[1]Department of Biomedical Data Sciences, Leiden University Medical Centre, The Netherlands

Abstract

Prediction in the presence of missing values is a complex and still poorly understood problem, particularly when future records also contain missing values. Mertens, *et al.* (2020) demonstrate that with non-linear models (such as logistic regression or Cox survival) and when using imputations, averaging of multiple predictions obtained from distinct models fitted on imputed data should be preferred to pooled models. In this talk we contrast predictive averaging with complete-case-based model calibration (CC) as well as use of missing-indicator (IDX) and Pattern Submodel (PS) approaches. We demonstrate that only predictive averaging guarantees required coverage levels in prediction. Scoring and class-separation measures (such as Brier or AUC) strongly favour IDX and PS methods however. We show this is due to the biased nature of these methods, which (Brier) scoring or AUC measures do not correct for.

References

Mertens, Banzato and de Wreede (2020). Construction and assessment of prediction rules for binary outcome in the presence of missing predictor data using multiple imputation and cross-validation: methodological approach and data-based evaluation. *Biometrical Journal*, 62, 724-741.

Mertens and de Wreede. Calibration of prediction rules for life-time outcomes using prognostic Cox regression survival models and multiple imputations. Technical report. *ArXiv:2105.01733*

Use of the R packages `rocbc` and `trinROC` for inference after the Box-Cox transformation in the ROC framework (and an outlook for future research)

Christos T. Nakas[1]

[1]University of Thessaly at Volos, Greece & Inselspital Bern/University of Bern, Switzerland

Abstract

Receiver Operating Characteristic (ROC) curve analysis is widely used in evaluating the effectiveness of a diagnostic test/biomarker or classifier score. A parametric approach for statistical inference on ROC curves based on a Box-Cox transformation to normality has been discussed in the literature along with extensions in the ROC surface framework. Many investigators have highlighted the difficulty of taking into account the variability of the estimated transformation parameter(s) when carrying out such an analysis. This variability is often ignored and inferences are made by considering the estimated transformation parameter(s) as fixed and known. In this talk, we will discuss the use of the Box-Cox transformation in the general ROC framework and the methodology for accounting for the estimation of the Box-Cox transformation parameter(s) in the context of ROC analysis in practice. We will also detail its application to a number of problems, along with implementation options using the R packages `rocbc` and `trinROC`. Implications on their use and of the Box-Cox transformation in the ROC framework in general will be discussed along with possible workarounds.

Relationship between two-sample test statistics and ROC summary measures

María del Carmen Pardo[1]

[1]Department of Statistics and OR, Complutense University of Madrid, Spain

Abstract

A biomarker is usually used as a diagnostic or assessment tool in medical research. They have become essential tools for proper diagnosis and treatment of a wide range of illnesses, including cancer, diabetes, and infectious diseases. The ROC curve shows the diagnostic accuracy of a biomarker and expresses clinically and statistically the area under the curve (AUC) of the diagnostic power of the biomarker, which corresponds exactly to the Wilcoxon statistic. In fact, this curve essentially provides a distribution-free description of the separation between the diseased and healthy distributions. Therefore, each of the summary measures, as AUC as example, is in a sense a summary of the distance between these two distributions. In this talk, we rewrite some of the most well-known ROC summary measures as a two-sample test. Furthermore, this relationship between two-sample test statistics and ROC summary measures is exploited further to suggest new ROC summary measures from two-sample test statistics.

The simulations show that these ROC summary statistics exhibit much higher power in discriminating between the diseased and healthy distributions and are thus an attractive alternative to classical ROC summary indices. Finally, to illustrate our proposal, we apply the proposed test statistics to assess a diagnostic marker in a real data set.

The conditional C-index for survival data with a cure fraction

Juan Carlos Pardo-Fernández[1], Bo Han[2] and Ingrid Van Keilegom[3]

[1]Department of Statistics and Operations Research and CITMAga, Universidade de Vigo, Spain
[2]School of Mathematics and Statistics, Yunnan University, China
[3]Research Centre for Operations Research and Statistics, KU Leuven, Belgium

Abstract

When analyzing survival data, it often occurs that some individuals never experience the event of interest. These individuals are called cured and cure models are then used to take into account this situation. In particular, in this talk we will consider a mixture cure model, which combines the probability of being uncured (also called incidence) and the survival function of the uncured patients (also called latency).

In practice, risk scoring systems of latency and incidence are crucial elements for identifying relevant biomarkers and treatment strategies. Concordance measures that discriminate higher-risk subjects from lower-risk subjects are valuable tools to evaluate the overall performance of risk scoring systems. In contrast to conventional concordance measures, conditional concordance measures are proposed in this talk to provide comprehensive assessment of fitted cure models for particular values of a set of covariates. Specifically, we will consider the conditional version of the *concordance index* or *C-index* to evaluate the discrimination capacity of risk factors for both the latency and the incidence.

Non- and semi-parametric modelling strategies are proposed to estimate and perform inferences about the conditional C-index. Simulation studies demonstrate that our proposal has a promising performance in finite samples. An application to real data is presented for illustrating the methodology.

R software for ROC analysis

Sonia Pérez-Fernández[1] and Pablo Martínez-Camblor[2]

[1]Departamento de Estadística e Investigación Operativa y Didáctica de la Matemática, Universidad de Oviedo, Spain

[2]Department of Anesthesiology and Biomedical Data Science Department, Geisel School of Medicine at Dartmouth, NH, USA

Abstract

Receiver Operating Characteristic (ROC) curve analysis is a fundamental tool for evaluating the performance of classification models in fields such as medical diagnostics and machine learning. R provides useful libraries to generate, compare, and visualize ROC curves, allowing users to compute metrics such as sensitivity, specificity, and the area under the curve. These libraries often include methods to handle complex scenarios such as censored data, multivariate settings, and comparison of multiple curves using advanced statistical techniques.

These tools allow researchers and practitioners to evaluate classifiers under a variety of conditions, highlighting the adaptability of R for these statistical analysis. In this talk, we will review the libraries available in the Comprehensive R Archive Network (CRAN) for ROC analysis, showing their main functionality. In particular, we will focus on the libraries nsROC, which includes non-standard analyses, and movieROC, which provides several visualisations.

Construction and evaluation of optimal diagnostic tests

Ainesh Sewak[1]

[1]Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Switzerland

Abstract

Accurate diagnostic tests are essential for effective treatment, screening, and surveillance of diseases. Single biomarkers often lack the sensitivity and specificity required for heterogeneous diseases like cancer, requiring the development of composite diagnostic tests. While binary classification methods like logistic regression are commonly used for this, they can be suboptimal when disease states exhibit differing dependence structures among biomarkers or when biomarker distributions are complex. In this talk, we introduce a novel multivariate model that addresses these challenges by capturing varying biomarker dependence structures while flexibly modeling their marginal distributions. This model can then be used to construct an optimal diagnostic test which combines multiple biomarkers using the likelihood ratio function. The model's parameters translate into computationally simple diagnostic accuracy measures and it can be used even in the presence of missing biomarker data. We demonstrate the model?s application by constructing an optimal diagnostic test for hepatocellular carcinoma, a cancer type lacking a single ideal biomarker. The talk will also feature an R implementation to demonstrate the results of the application.

Organising Committee

Juan Carlos Pardo-Fernández

Pablo Rodríguez Castro

Sponsored by

Universida_{de}Vigo

Escola de Enxeñería Industrial

Departamento de Estatística
e Investigación Operativa

SiDOR
Statistical Inference, Decision
& Operations Research Group

CIT MA GA

GALICIAN CENTRE FOR
MATHEMATICAL RESEARCH
AND TECHNOLOGY

MINISTERIO
DE CIENCIA
E INNOVACIÓN

AGENCIA
ESTATAL DE
INVESTIGACIÓN