

# **Web Appendix for the paper: Comparison of Two Correlated ROC Surfaces at a Given Pair of True Classification Rates**

by Leonidas E. Bantis and Ziding Feng

Software

# 1 Software illustration, Video Animation and a Reproducible Example

In this section we present a user friendly software so that our approaches can be directly applicable. The MATLAB based routine is called `roc3dcompare`. By passing in the biomarker measurements along with the chosen pair at which the comparison should be made, the routine provides all relevant output as well as an informative figure, based also on the method of choice. More specifically the input arguments of the routine are:

## INPUT ARGUMENTS:

- **W1:** A column vector of biomarker measurements taken with biomarker 1.
- **W2:** A column vector of biomarker measurements taken with biomarker 2.
- **D:** The disease status. A column vector taking values 1,2 or 3 depending on whether the  $i$ -th individual belongs to group A, B or C respectively.
- **tripletABC:** A row vector of three elements that will indicate the given pair at which the comparison will take place. This vector should contain exactly a **NaN** (not a number) value. For example, if the given pair is  $TC_A = 0.5$  and  $TC_C = 0.6$  then this input argument should be set as `[0.5 NaN 0.6]`. Analogously, if the given pair is  $TC_B = 0.5$  and  $TC_C = 0.6$  then this input argument should be set as `[NaN 0.5 0.6]`. Similarly, for the case where  $TC_A = 0.5$  and  $TC_B = 0.6$  are given then this input argument should be set as `[0.5 0.6 NaN]`.
- **CIttype:** The method employed. This input argument can be set as **normal**, or **boxcox**, or **kernels** depending on the approach that the user want to be employed.
- **init:** This input refers only to the Box-Cox based approach and can also not be reached. It refers to initial values of the parameters involved in the likelihood function of the Box-Cox approach ( $\mu_{1A}, \sigma_{1A}, \mu_{1B}, \sigma_{1B}, \mu_{1C}, \sigma_{1C}, \mu_{2A}, \sigma_{2A}, \mu_{2B}, \sigma_{2B}, \mu_{2C}, \sigma_{2C}, cov_A, cov_B, cov_C, \lambda_1, \lambda_2$ ). If this argument is not reached the sample means and standard deviations are employed along with the values 0.5 and 1 for all the covariances and transformation parameters ( $\lambda_1$  and  $\lambda_2$ ) respectively.

## OUTPUT ARGUMENTS:

- **distance:** The distance between the two ROC surfaces at the given pair of true classification rates.
- **pval2t:** A two tailed  $p$ -value regarding the targeted comparison based on statistic  $Z^*$ .
- **Z:** The actual value of the statistic  $Z^*$ .

- **ROC1**: The estimated point on the ROC surface ( $\hat{ROC}_1(TC_i, TC_k)$ ) that corresponds to biomarker 1, where  $i, j = A, B, C$  with  $i \neq j$ . The estimated point is based on the input argument **tripletABC** given by the user.
- **ROC2**: The estimated point on the ROC surface ( $\hat{ROC}_2(TC_i, TC_k)$ ) that corresponds to biomarker 2, where  $i, j = A, B, C$  with  $i \neq j$ . The estimated point is based on the input argument **tripletABC** given by the user.
- **vus1**: Volume under the surface of biomarker 1.
- **vus2**: Volume under the surface of biomarker 2.
- **CIdiff**: 95% confidence interval for the difference  $\Phi(\hat{ROC}_1(TC_i, TC_k)) - \Phi(\hat{ROC}_2(TC_i, TC_k))$ , where  $i, j = A, B, C$  with  $i \neq j$ .

Apart from the above output arguments two figures automatically appear (see (1) and (2)) that visualize the comparison of the two surfaces as well as a depiction of the pairwise ROC curves. Furthermore the feasible region of a given true classification rate is also provided by the second figure. The 3D figure (Figure (1)) is dynamic and can be rotated by the user at any angle for better visualization. In addition to that and for illustration purposes we also provide a video animation of Figure (1) on the journal's website. The code is also provided by the journal's website as well as in the webpage of the first author.

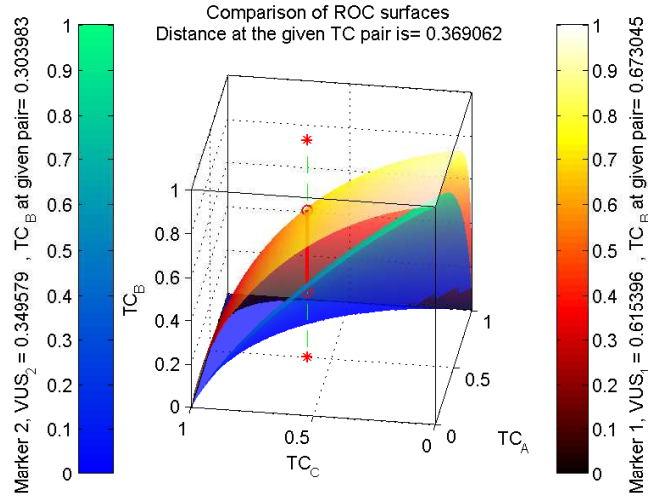


Figure 1: Two ROC surfaces that correspond to the reproducible simulated example and are automatically provided by the provided software. The given point is  $(TC_A = 0.5, TC_C = 0.5)$  and the obtained distance at that point is depicted with a red line with length 0.3691.

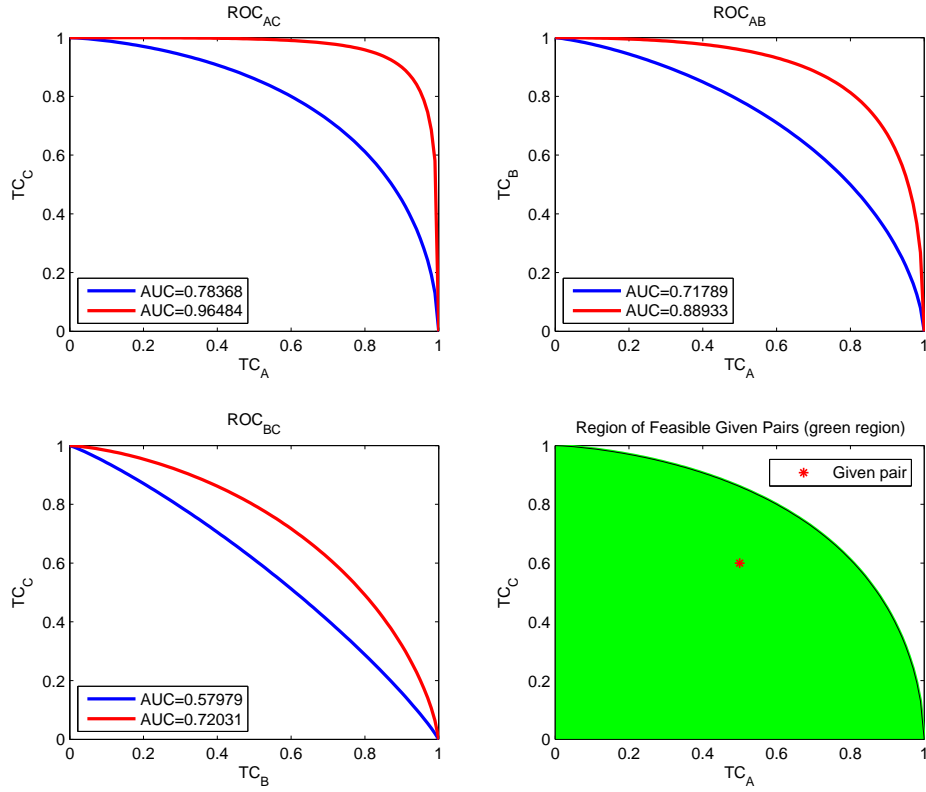


Figure 2: The corresponding projections of the two ROC surfaces on the sides of the unit cube that correspond to the pairwise ROC curves. In the right bottom figure the feasible region in which the user can provide a pair of  $TC_A$  and  $TC_C$  is given along with the point that the user actually provided. If the point given by the user is outside the feasible region then all inference related output will be NaN (Not a Number).