Nonparametric Estimation of the AUC of an Index with Estimated Parameters

Outline

- medical professionals want to use biomarkers to classify patients. Combine several characteristics. Usually linear combination ("index").
- e.g., Framingham risk score for 10-year risk of cardiovascular disease, using as covariates age, sex, LDL cholesterol, HDL cholesterol, blood pressure, diabetes, and smoking.
- Used for policy, insurance etc.

- A good biomarker discriminates well between cases and controls:
$$P(\beta^T X_0 < \beta^T X_1),$$
X_0 and X_1 being independent control and case observations
- Always looking for new indexes. Compare indexes using the difference of AUCs. Typically nested: comparing a known index with another containing an additional covariate.

$$P(\gamma^T(X_0, X_0') < \gamma^T(X_1, X_1')) - P(\beta^T X_0 < \beta^T X_1)$$

Procedure:

1. Fit risk models to get coefficients, e.g., $P(D=1|W)=\beta^T W, P(D=1|W,W')=\gamma^T(W,W')$, to get coefficients $\hat{\beta},\hat{\gamma}$.
2. Then compare observed AUCs

$$(mn)^{-1}\sum_{i,j}\{\hat{\gamma}^T(X_0,X_0')_i < \hat{\gamma}^T(X_1,X_1')_j\}$$
$$-(mn)^{-1}\sum_{i,j}(\hat{\beta}^T X_{0i} < \hat{\beta}^T X_{1j}\}$$

Formal test is given in Delong '88.

Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach

ER **DeLong**, DM **DeLong**, DL Clarke-Pearson - Biometrics, **1988** - JSTOR

Methods of evaluating and comparing the performance of diagnostic tests are of increasing importance as new tests are developed and marketed. When a test is based on an observed ...

☆ Save 99 Cite Cited by 18145 Related articles All 10 versions

Outline

In the 2000s medical professionals noticed significantly different coefficient vectors don't lead to significantly nonzero ΔAUC

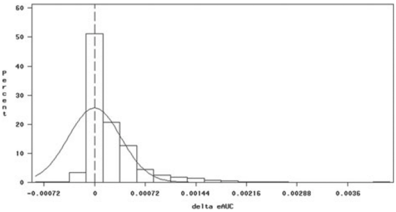
- Meigs JB et al. Genotype score in addition to common risk factors for prediction of type II diabetes. New England Journal of Medicine. 2008; 359(21):2208–2219.
- Wang TJ et al. Multiple biomarkers for the prediction of first major cardiovascular events and death. The New England Journal of Medicine. 2006; 355(25):2631–2639.
- Kattan MW. Judging new markers by their ability to improve predective accuracy. Journal of the National Cancer Institute. 2003;
- Cook NR. Statistical evaluation of prognostic versus diagnostic models: Beyond the ROC curve. Clinical Chemistry. 2008;
- Hlatky MA et al. Criteria for evaluation of novel markers of cardiovascular risk: A scientific statement from the american heart association. Circulation. 2009

Leading some to question the difference in AUC as a useful criterion

- Tzoulaki I, Liberopoulos G, Ioannisis JPA. Assessment off claims of improved prediction beyond the Framnigham risk score. JAMA. 2009

By the early 2010s the issue comes to the attenton of biostatisticians

- Demler, Pencina, D'agonstino 2012: Simulation suggests nested null comparison leads to a degenerate U-stat, non-normal limiting distribution. Not contemplated by Delong's test.



  - Seshan, Gönen, Begg 2013: principal issue is that the observations aren't IID due to the estimated coefficients, violation of the Delong test assumptions

- Demler, Pencina, Cook, D'agonstino 2017: theoretical arguments that the principal issue is degeneracy of the U-statistic

In this article we have provided an explanation to the baffling observation that use of the AUC test to compare nested binary regression models is invalid [2]. We found that the

The performance of the AUC test has perplexed other investigators, including Demler et. al.

Some solutions appear

- Demler 2011: With gaussian covariates and coeffiicents estimated by LDA or logistic regression, the null of $\Delta\text{AUC} = 0$ is the same as testing for equal Mahalonobis distances, can use an F-test
- Pepe 2013: the null $\Delta\text{AUC} = 0$ is the same as the risk functions being equal $P(D = 1 \mid W, W') = P(D = 1 \mid W)$
- Heller 2017: obtain asymptotic distributions in case coefficient estimation procedure is MRC

However,

- Lee 2021: "To date, it is remarked that the asymptotic null distribution of the test statistic computed via the most commonly used MLE of binary regression parameters is still considered as mathematically intractable in general. At least, Monte Carlo methods are deemed to be infeasible or otherwise impractical . . ."

Outline

Background/Motivation

Problem

Proposed solution for the alternative

Remarks on testing the null

- Focus first on the alternative first $\Delta\text{AUC} \neq 0$: mainly for developing CIs for ΔAUC, rather than testing $H_0 : \Delta\text{AUC} = 0$
- Not much advantage gained from the full/reduced relationship between the two indexes. Just get an linearization representation for an index AUC with estimated coeficient, then take difference.

- strategy in establishing asymptotic normality of U-statistics is to find an asymptotically equivalent IID representation.
- In the context of U-statistics usually known as the hoeffding decomposition. dates to 1940s/50s

If the X_i and Y_j have cumulative distribution functions F and G, respectively, then the projection of $U - \theta$ can be written

$$\hat{U} = -\frac{1}{m}\sum_{i=1}^{m}(G_-(X_i) - \mathrm{E}G_-(X_i)) + \frac{1}{n}\sum_{j=1}^{n}(F(Y_j) - \mathrm{E}F(Y_j)).$$

It is easy to obtain the limit distribution of the projections $\hat{U}$ (and hence of U) from this for-

Delong ’88: substitute empirical CDFs

Two sources of approximation. IID representation consists of the classical delong rep and an adjustment.

$$
\begin{aligned}
&\theta(\hat{F}, \hat{G}, \hat{\beta}) - \theta(F, G, \beta^*) \\
&= \theta(F + \delta F, G + \delta G, \beta^* + \delta\beta) - \theta(F, G, \beta^* + \delta\beta) \\
&+ \theta(F, G, \beta^* + \delta\beta) - \theta(F, G, \beta^*)
\end{aligned}
$$

Where

$$
\theta(F, G, \beta) = \int \{\beta^T x < \beta^T y\} dF(x) dG(y).
$$

and $\delta F = \hat{F} - F$, etc.

Proposition

Given a sample $(W_1, D_1), \ldots, (W_{m+n}, D_{m+n})$, *and estimator* $\hat{\beta}$ *based on the sample. Assumptions:*

1. *available influence function for* $\hat{\beta}$
2. $P(D = 0) \in (0, 1)$
3. $\hat{\beta} \to \beta^*$
4. $\theta(F, G, \cdot)$ *is differentiable at* β^*

Assertion: $(m+n)^{-1/2}(\theta(\hat{F}, \hat{G}, \hat{\beta}) - \theta(F, G, \beta^*))$ *is asymptotically normal with mean zero, variance can be estimated using the linearization*

In practice, requires non-parametric estimation of a gradient

Example: Linear discriminant analysis, but misspecified in that the classes may have different variances

$$W|D = d \sim N_p(\mu_d, \Sigma_d), d = 0, 1$$
$$P(D = 1) = 1 - P(D = 0) = \pi_1$$

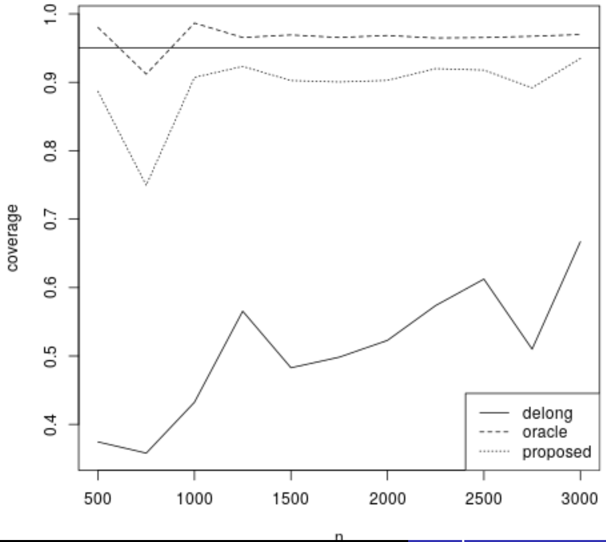
A sub-model

$$\Sigma_0 = \begin{pmatrix} \ddots & & & & \\ & \epsilon & & & \\ & & \ddots & & \\ & & & 2-\epsilon & \\ & & & & \ddots \end{pmatrix}, \Sigma_1 = I, \beta = \sqrt{2/p}\mathbb{1}$$

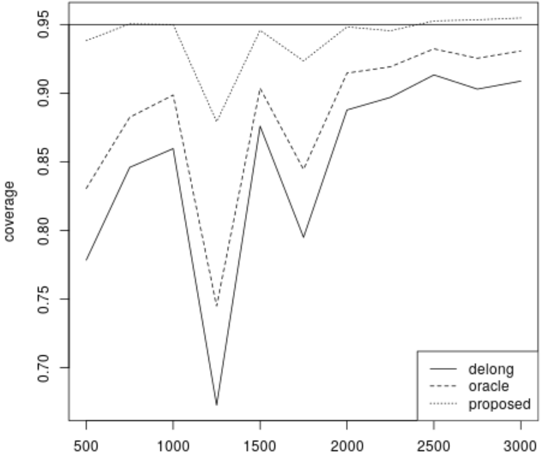
The adjustment term is

$$|\Sigma_\pi^{-1/2} \frac{\partial}{\partial \beta} \theta(F, G, \beta)| = \left| \frac{(\pi_1 - \pi_0)(1-\epsilon)}{4\sqrt{p\pi e}} \begin{pmatrix} \vdots \\ \pm(1 + (\epsilon - 1)\pi_0)^{-1/2} \\ \vdots \end{pmatrix} \right|$$

which $\to \infty$ as $\pi_0 \to 1$ and $\epsilon \to 0$ simultaneously

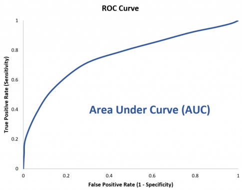

coverage
1.0
0.9
0.8
0.7
0.6
0.5
0.4
500
1000
1500
2000
2500
3000
delong
oracle
proposed

logit-normal data, well-specified LDA: Delong estimator OK (i.e., no adjustment needed)



Why was it so hard to find useful applications of the proposed estimator?

1. the adjustment term is a taylor expansion, the derivative of the AUC at the true coefficient values
2. given covariates W, the ROC curve of the likelihood ratio $w \mapsto f_{W|D=1}(w)/f_{W|D=0}(w)$, is at every point maximal
3. the AUC of the likelihood ratio is a stationary point
4. the AUC is invariant to increasing transformations of the data



Binary index model example

$$P(D = 1 \mid W) = h(\beta^T W), h \text{ increasing}$$

$$P(D = 1 \mid W) = \operatorname{expit} \log(P(D = 1 \mid W = w)/P(D = 0 \mid W = w))$$

so index AUC needs no adjustment

- But, there are two AUCs involved in ΔAUC
- In many cases correctness of the full model,

$$P(D = 1 \mid (W, W') = (w, w')) = h(\beta^T(w, w')) \quad \text{for some } \beta \in \mathbb{R}^{p+q}$$

implies the reduced model cannot be correct

$$P(D = 1 \mid W = w) = E(h(\beta^T(w, w')) \mid w) \neq h(\beta^T w) \quad \text{for any } \beta \in \mathbb{R}^p.$$

- Probit model with gaussian covariates:

$$P(D=1 \mid (W,W')=(w,w')) = \Phi(\beta^T(w,w'))$$
$$\text{for some } \beta \in \mathbb{R}^{p+q}$$

implies

$$P(D=1 \mid W=w) = \Phi({\beta'}^T w)$$

- So the adjustment term vanishes
- Assertion: For any link, if the covariates are gaussian, the adjustment term vanishes iff the reduced model coefficients are the probit coefficients (β' above)
- For coefficients obtained under other models, the magnitude of the (nonzero) adjustment term can be controlled by the difference of the coefficients from coefficients obtained under a probit model

Outline

- nonnormal limit but in some ways null case is simpler. no need for functional derivative, const terms cancel. can pretransform data with many coef estimation procedures.
- nonparametric estimation of a hessian

- Difficulty: If the coefficient estimation procedure is well-specified, the first-order term in the Taylor expansion vanishes, and the limit is non-normal ($1/n$ rate)
- If misspecified, the limit is normal ($1/\sqrt{n}$ rate)
- boundary area

E.g.:

- Data follows a logistic model

$$P(D = 1) = \text{expit}(\beta^T W)$$

  Two analysts are considering the effect of an additional covariate ϵ which is in truth just noise

- One uses logistic regression to fit $\hat{\gamma}$

$$P(D = 1 \mid W, \epsilon) = \text{expit}(\gamma^T (W, \epsilon))$$

  The gradient of the AUC at $\beta^* = \gamma^*$ will vanish and the asy distribution will be $O(1/n)$

- The other uses probit regression: The first-order term will have the form

$$(\hat{\beta} - \hat{\gamma}) \times (\text{gradient}),$$

  standard asymptotics

Joint work with Alexis Doyle. Nonparametric Estimation of the AUC of an Index with Estimated Parameters. Forthcoming. Available at: `https://haben-michael.github.io/`

THANK YOU