Binormal confidence intervals for AUC in R

Leonhard Held and Sina Rueeger*

November 24, 2020

AUC curves are used to measure the accurary of a classification of two groups X and Y:

$$X_1, \dots, X_{n_X} \sim \mathcal{N}(\mu_X, \sigma_X^2)$$
$$Y_1, \dots, Y_{n_Y} \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$$

Y could be denoted as the healthy controls and X the cases with a particular disease. When having small sample size (and therefore small values in the contingency table) the confidence interval given with Wald (as in function `biostatUZH::confIntAUC`) will not perfom well (fails). Hence, another way has to be found to compute the confidence interval. Pepe (2003) illustrates how AUC curves can be described using the normal distribution:

$$a = \frac{\mu_Y - \mu_X}{\sigma_Y}$$
$$b = \frac{\sigma_X}{\sigma_Y}$$
$$AUC = \Phi\Big(\frac{a}{\sqrt{1+b^2}}\Big)$$

Assumption: equal variances

Under the assumption that both variances σ_X and σ_Y are equal and known, the equations take a much simpler form:

$$\sigma = \sigma_X = \sigma_Y$$
$$a = \frac{\mu_Y - \mu_X}{\sigma}$$
$$b = \frac{\sigma_X}{\sigma_Y} = 1$$
$$AUC = \Phi\Big(\frac{a}{\sqrt{1+b^2}}\Big) = \Phi\Big(\frac{a}{\sqrt{2}}\Big)$$

The expected value of X and Y is estimated using the average and variance: $\hat{\mu}_X = \overline{x}$, $\hat{\mu}_Y = \overline{y}$.

*University of Zurich, Institute for Social and Preventive Medicine, Division of Biostatistics (`[leonhard.held,sina.rueeger]@ifspm.uzh.ch`)

To calculate the confidence interval, the $SE(\hat{a})$ is needed. Given that $\hat{\mu}_X \sim \mathcal{N}(\mu_X, \sigma^2/n_X)$ and $\hat{\mu}_Y \sim \mathcal{N}(\mu_Y, \sigma^2/n_Y)$

$$SE(\hat{a}) = SE\Big(\frac{\hat{\mu_X} - \hat{\mu_Y}}{\sigma}\Big) = \sqrt{\widehat{Var}\Big(\frac{\hat{\mu_X} - \hat{\mu_Y}}{\sigma}\Big)} = \sqrt{\widehat{Var}\Big(\frac{\hat{\mu_X}}{\sigma}\Big) + \widehat{Var}\Big(\frac{\hat{\mu_Y}}{\sigma}\Big)} = \sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}.$$

The $(1-\alpha)$-confidence interval for a has the following form:

$$\begin{aligned} \text{from} \quad & a_{lower} = a - z \cdot SE(\hat{a}) \\ \text{to} \quad & a_{up} = a + z \cdot SE(\hat{a}) \end{aligned}$$

where z refers to the (1-$\alpha/2$)-quantile of the standard normal distribution.

The confidence interval limits of AUC are derived by calculating the percentile of the confidence interval limits of $\hat{a}$:

$$\begin{aligned} \text{from} \quad & \Phi\Big(\frac{a_{lower}}{\sqrt{2}}\Big) \\ \text{to} \quad & \Phi\Big(\frac{a_{up}}{\sqrt{2}}\Big) \end{aligned}$$

Assumption: not equal variances

If the assumption of equal variances would not hold, the standard error of AUC could be derived using the multivariate delta method.

References

Pepe, M. S. (2003). *The statistical evaluation of medical tests for classification and prediction*, vol. 28 of *Oxford Statistical Science Series*. Oxford University Press, Oxford.