ROC CURVE ANALYSIS USING SAS

Zheng Yao
Sr. Statistical Programmer
Outline

• Background

• Examples:
  - Accuracy assessment
  - Compare ROC curves
  - Cut-off point selection

• Summary
Outline

• **Background**

• **Examples:**
  - Accuracy assessment
  - Compare ROC curves
  - Cut-off point selection

• **Summary**
Background

- Biomarkers (e.g. PD-1/L1) draw lots of attention nowadays.
- It is often of interest to use biomarker for disease screening, diagnosis and prediction.
- The fundamental for use of biomarkers in clinical practice is the **accuracy** and the **optimal cut-off point selection**
- The receiver operating characteristic (ROC) curve is a procedure that can aid in the **accuracy assessment, ROC curve comparison and cut-off point selection**.
Background

### Sensitivity
The proportion of positive observations that are measured as positive, i.e. true positive rate (TPR),
\[
\text{Sensitivity} = \frac{a}{a + c}
\]

### Specificity
The proportion of negative observations that are measured as negative, i.e. true negative (TNR),
\[
\text{Specificity} = \frac{d}{b + d}
\]

### Youden’s Index
\[
\text{Youden’s Index} = (\text{sensitivity} + \text{specificity}) - 1
\]
\[
\text{Sensitivity} = \frac{60}{60 + 40} = 0.6
\]
\[
\text{Specificity} = \frac{80}{20 + 80} = 0.8
\]
\[
\text{Youden’s Index} = (0.6 + 0.8) - 1 = 0.4
\]
Background

• ROC(Receiver-Operating Characteristic) Curve: constructed with sensitivity (FP) on the vertical axis and 1-specificity (TP) on the horizontal

• Area under curve (AUC): a measure of overall accuracy

• ROC curves comparison

• Cut-off point selection
  ➢ Youden’s Index: (sensitivity + specificity) - 1
Outline

• Background

• Examples:
  ➢ Accuracy assessment
  ➢ Compare ROC curves
  ➢ Cut-off point selection

• Summary
Compare ROC curves

• 156 cases of patients with positive PD-L1 expression were selected and treated with anti-PD-L1 agents (drug A).
• To explores the potential association between PD-L1 expression in tumor tissue and anti-PD-L1 response (i.e. BOR, PFS or OS)
• The expression of PD-L1 was measured with with two method (TC and IC) for each case
• Best Overall Response (BOR) is used as the predictive measure according to (RECIST 1.1, 1). The BOR variable is coded as ‘1’ and ‘0’ (i.e. effective and noneffective).

<table>
<thead>
<tr>
<th>PD-L1 expression</th>
<th>BOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>TP (a)</td>
</tr>
<tr>
<td>Negative</td>
<td>FN (c)</td>
</tr>
<tr>
<td></td>
<td>Non-effective</td>
</tr>
<tr>
<td></td>
<td>FP (b)</td>
</tr>
<tr>
<td></td>
<td>TN (d)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Obs</th>
<th>USUBJID</th>
<th>TC</th>
<th>IC</th>
<th>BOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>001</td>
<td>15.814</td>
<td>2.823</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>002</td>
<td>8.991</td>
<td>15.975</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>003</td>
<td>4.477</td>
<td>5.46</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>004</td>
<td>17.55</td>
<td>12.447</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>005</td>
<td>0.604</td>
<td>3.217</td>
<td>0</td>
</tr>
</tbody>
</table>

NOTE: Partial list output of the BOR dataset (obs=5)
Compare ROC curves

- Is IC better than TC?
Compare ROC curves

The PROC LOGISTIC procedure for ROC curve comparison

- **TC** and **IC** are both independent variables in the model statement.
- the **ROC** statement produces a ROC
- the **ROCCONTRAST** statement produces a significance test for the ROC curve.
- the **REFERENCE('TC')** statement means that TC is set as a reference when comparing with IC in the significance test.

```sas
proc logistic data = BOR;
  model BOR(event='1') = TC IC;
  roc"TC" TC;
  roc"IC" IC;
  roccontrast reference("TC")/estimate;
run;
```

*Note.* Reference code for ROC curve comparison
Compare ROC curves

Results from Logistic model

• The significance test demonstrates that TC ($p = 0.0014$) and IC ($p < 0.0001$) are statistically significant for use in ROC curve.

• Logistic regression equation: $\ln \left( \frac{p}{1-p} \right) = \beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>DF</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>Wald Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1</td>
<td>-3.9048</td>
<td>0.6434</td>
<td>36.8311</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>TC</td>
<td>1</td>
<td>0.1104</td>
<td>0.0344</td>
<td>10.2640</td>
<td>0.0014</td>
</tr>
<tr>
<td>IC</td>
<td>1</td>
<td>0.1905</td>
<td>0.0364</td>
<td>27.3634</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>
**Compare ROC curves**

**AUC statistical test and ROC curve**

- **ROC Curve**: all of three ROC curves are above the diagonal line.

- **AUC**: all of three 95% CIs do not contain 0.5. Therefore, we can conclude that all these three AUC are significantly better than chance.

---

### ROC Association Statistics

<table>
<thead>
<tr>
<th>ROC Model</th>
<th>Area</th>
<th>Standard Error</th>
<th>95% Wald Confidence Limits</th>
<th>Somers' D (Gini)</th>
<th>Gamma</th>
<th>Tau-a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>0.8837</td>
<td>0.0285</td>
<td>0.8277 0.9396</td>
<td>0.7673</td>
<td>0.7673</td>
<td>0.3737</td>
</tr>
<tr>
<td>TC</td>
<td>0.7218</td>
<td>0.0410</td>
<td>0.6414 0.8022</td>
<td>0.4436</td>
<td>0.4436</td>
<td>0.2160</td>
</tr>
<tr>
<td>IC</td>
<td>0.8558</td>
<td>0.0324</td>
<td>0.7923 0.9193</td>
<td>0.7116</td>
<td>0.7116</td>
<td>0.3466</td>
</tr>
</tbody>
</table>
Compare ROC curves

**ROC comparison test**

- IC vs. TC ($p=0.0116$). Therefore, the AUC of IC (0.8558) is statistically larger than that of TC (0.7218).
- the PD-L1 scoring methodology of IC is better than TC.

### ROC Contrast Estimation and Testing Results by Row

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>95% Wald Confidence Limits</th>
<th>Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model - TC</td>
<td>0.1619</td>
<td>0.0405</td>
<td>0.0825 0.2412</td>
<td>15.9746</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>IC - TC</td>
<td>0.1340</td>
<td>0.0531</td>
<td>0.0300 0.2380</td>
<td>6.3740</td>
<td>0.0116</td>
</tr>
</tbody>
</table>

**Note.**
Outline

• Background

• Examples:
  - Accuracy assessment
  - Compare ROC curves
  - Cut-off point selection

• Summary
Select a rational cut-off point in ROC curve analysis

The PROC LOGISTIC procedure for ROC curve analysis

• The **OUTROC**= option creates a dataset containing sensitivity and specificity data which here is called **ROCDATA**.

• The **ROC** statement produces a ROC

• the **ROCCONTRAST** statement produces a significance test for the ROC curve.

• The **PREDICTED**= option creates a dataset containing *estimated event probabilities* (i.e. pred) for each subject.

```plaintext
proc logistic data = BOR;
  model BOR(event='1') = IC/outroc=rocdata;
  output out=pred predicted=pred;
  roc "BOR";
  roccontrast;
run;
```
Select a rational cut-off point in ROC curve analysis

The ROC curve

- The diagonal line, from (0,0) to (1,1), is indicative of an independent variable that discriminates no different from guessing (50/50 chance).

- The AUC is 0.8558 as compared to that of the diagonal line which is always 0.5.
Select a rational cut-off point in ROC curve analysis

AUC statistical test

- The **95% confidence interval** (0.6414, 0.8022) does not contain 0.5, therefore our AUC is significantly better than chance.

- A **Chi-square test** provides a p-value ($p < .0001$) associated with the null hypothesis (AUC = 0.5).

<table>
<thead>
<tr>
<th>ROC Model</th>
<th>Area</th>
<th>Standard Error</th>
<th>95% Wald Confidence Limits</th>
<th>Somers’ D (Gini)</th>
<th>Gamma</th>
<th>Tau-a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>0.8558</td>
<td>0.0324</td>
<td>0.7923, 0.9193</td>
<td>0.7116</td>
<td>0.7116</td>
<td>0.3466</td>
</tr>
<tr>
<td>BOR</td>
<td>0.5000</td>
<td>0.0000</td>
<td>0.5000, 0.5000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

**Note.** If the 95% confidence interval does not include 0.5, then we can conclude that ROC curve is statistically significant.
Select a rational cut-off point in ROC curve analysis

The model’s intercept and regression coefficients

- Partial output from the PROC LOGISTIC procedure
- Logistic regression equation: \[ \ln \left( \frac{p}{1-p} \right) = \beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n \]

<table>
<thead>
<tr>
<th>Parameter</th>
<th>DF</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>Wald Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1</td>
<td>-2.6197</td>
<td>0.4261</td>
<td>37.8058</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>IC</td>
<td>1</td>
<td>0.1880</td>
<td>0.0342</td>
<td>30.1969</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

**NOTE:** The model’s intercept (-1.6866) and regression coefficients (0.1122) are needed for computing the cutoff point.
Select a rational cut-off point in ROC curve analysis

Determine an optimal PD-L1 cutoff point for BOR

- Logistic regression equation: \( \ln \left( \frac{p}{1-p} \right) = \beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n \)
- \( \ln \left( \frac{p}{1-p} \right) = \text{intercept} + \text{slope}(X) = -2.6197 + 0.1880 \times \text{cutoff} \)

Note: Partial list output of the rocdata dataset
Select a rational cut-off point in ROC curve analysis

Determine an optimal cutoff point for BOR

- Logistic regression equation: \( \ln \left( \frac{p}{1-p} \right) = \beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n \)
- \( \text{logit} = \text{intercept} + \text{slope}(X) = -2.6197 + 0.1880 \times \text{cutoff} \)

<table>
<thead>
<tr>
<th>Obs</th>
<th>cutoff</th>
<th>prob</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Youden</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10.38</td>
<td>0.37240</td>
<td>0.797</td>
<td>0.804</td>
<td>0.601</td>
</tr>
<tr>
<td>2</td>
<td>7.62</td>
<td>0.30321</td>
<td>0.844</td>
<td>0.750</td>
<td>0.594</td>
</tr>
<tr>
<td>3</td>
<td>9.18</td>
<td>0.34148</td>
<td>0.797</td>
<td>0.793</td>
<td>0.590</td>
</tr>
<tr>
<td>4</td>
<td>7.98</td>
<td>0.31185</td>
<td>0.828</td>
<td>0.761</td>
<td>0.589</td>
</tr>
<tr>
<td>5</td>
<td>12.54</td>
<td>0.43045</td>
<td>0.750</td>
<td>0.837</td>
<td>0.587</td>
</tr>
</tbody>
</table>

Note. The cutoff variable is formed by re-arranging logistic regression model to solve for \( X \). The model is: \( \text{logit} = \text{intercept} + \text{slope}(x) \).

Note. Partial output of cutoff value and Youden Index by sorting the dataset in descending Youden. A large Youden may be one criteria for deciding an appropriate cutoff score.
Outline

• Background

• Examples:
  ➢ Accuracy assessment
  ➢ Compare ROC curves
  ➢ Cut-off point selection

• Summary
  ➢ Accuracy assessment
  ➢ ROC curve comparison
  ➢ Cut-off point selection
BACK UP SLIDES