# roctab — Nonparametric ROC analysis

## Syntax

```
roctab refvar classvar [if] [in] [weight] [, options]
```

## roctab_options

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main</strong></td>
<td></td>
</tr>
<tr>
<td>lorenz</td>
<td>report Gini and Pietra indices</td>
</tr>
<tr>
<td>binomial</td>
<td>calculate exact binomial confidence intervals</td>
</tr>
<tr>
<td>nolabel</td>
<td>display numeric codes rather than value labels</td>
</tr>
<tr>
<td>detail</td>
<td>show details on sensitivity/specificity for each cutpoint</td>
</tr>
<tr>
<td>table</td>
<td>display the raw data in a $2 \times k$ contingency table</td>
</tr>
<tr>
<td>bamber</td>
<td>calculate standard errors by using the Bamber method</td>
</tr>
<tr>
<td>hanley</td>
<td>calculate standard errors by using the Hanley method</td>
</tr>
<tr>
<td>graph</td>
<td>graph the ROC curve</td>
</tr>
<tr>
<td>noreftime</td>
<td>suppress plotting the 45-degree reference line</td>
</tr>
<tr>
<td>summary</td>
<td>report the area under the ROC curve</td>
</tr>
<tr>
<td>specificity</td>
<td>graph sensitivity versus specificity</td>
</tr>
<tr>
<td>level(#)</td>
<td>set confidence level; default is <code>level(95)</code></td>
</tr>
</tbody>
</table>

| **Plot**      | affect rendition of the ROC curve                                          |
| plotopts(plot_options) |                                                                 |

| **Reference line** | affect rendition of the reference line                                       |
| rlopts(cline_options) |                                                                 |

| **Add plots**     | add other plots to the generated graph                                      |
| addplot(Plot)     |                                                                 |

| **Y axis, X axis, Titles, Legend, Overall** | any options other than `by()` documented in `[G-3] twoway_options` |
| twoway_options    |                                                                 |

## Remarks and examples

- `fweight`s are allowed; see [U 11.1.6 weight].

## Stored results

- `e(N)` number of observations
- `e(N0)` number of observations in refvar
- `e(N1)` number of observations in classvar
- `e(area)` area under the ROC curve
- `e(lower)` lower bound on the area under the ROC curve
- `e(upper)` upper bound on the area under the ROC curve
- `e(lower_n)` lower bound on the area under the ROC curve using the Bamber method
- `e(upper_n)` upper bound on the area under the ROC curve using the Bamber method
- `e(lower_h)` lower bound on the area under the ROC curve using the Hanley method
- `e(upper_h)` upper bound on the area under the ROC curve using the Hanley method

## Methods and formulas

- Gini and Pietra indices
- Exact binomial confidence intervals
- Bamber and Hanley methods for standard errors

## References

- [G-3] twoway_options
Menu

Statistics > Epidemiology and related > ROC analysis > Nonparametric ROC analysis without covariates

Description

The above command is used to perform receiver operating characteristic (ROC) analyses with rating and discrete classification data.

The two variables refvar and classvar must be numeric. The reference variable indicates the true state of the observation, such as diseased and nondiseased or normal and abnormal, and must be coded as 0 and 1. The rating or outcome of the diagnostic test or test modality is recorded in classvar, which must be at least ordinal, with higher values indicating higher risk.

roctab performs nonparametric ROC analyses. By default, roctab calculates the area under the ROC curve. Optionally, roctab can plot the ROC curve, display the data in tabular form, and produce Lorenz-like plots.

See [R] rocfit for a command that fits maximum-likelihood ROC models.

Options

\[
\text{lorenz} \hspace{1cm} \text{Main}\]

lorenz specifies that Gini and Pietra indices be reported. Optionally, graph will plot the Lorenz-like curve.

binomial specifies that exact binomial confidence intervals be calculated.

nolabel specifies that numeric codes be displayed rather than value labels.

detail outputs a table displaying the sensitivity, specificity, the percentage of subjects correctly classified, and two likelihood ratios for each possible cutpoint of classvar.

table outputs a \(2 \times k\) contingency table displaying the raw data.

bamber specifies that the standard error for the area under the ROC curve be calculated using the method suggested by Bamber (1975). Otherwise, standard errors are obtained as suggested by DeLong, DeLong, and Clarke-Pearson (1988).

hanley specifies that the standard error for the area under the ROC curve be calculated using the method suggested by Hanley and McNeil (1982). Otherwise, standard errors are obtained as suggested by DeLong, DeLong, and Clarke-Pearson (1988).

graph produces graphical output of the ROC curve. If lorenz is specified, graphical output of a Lorenz-like curve will be produced.

norefline suppresses plotting the 45-degree reference line from the graphical output of the ROC curve.

summary reports the area under the ROC curve, its standard error, and its confidence interval. If lorenz is specified, Lorenz indices are reported. This option is needed only when also specifying graph.

specificity produces a graph of sensitivity versus specificity instead of sensitivity versus \((1 - \text{specificity})\). specificity implies graph.

level(\#) specifies the confidence level, as a percentage, for the confidence intervals. The default is level(95) or as set by set level; see [R] level.
roctab — Nonparametric ROC analysis

Plot

`plotopts(plot_options)` affects the rendition of the plotted ROC curve—the curve’s plotted points connected by lines. The `plot_options` can affect the size and color of markers, whether and how the markers are labeled, and whether and how the points are connected; see [G-3] `marker_options`, [G-3] `marker_label_options`, and [G-3] `cline_options`.

Reference line

`rlopts(cline_options)` affects the rendition of the reference line; see [G-3] `cline_options`.

Add plots

`addplot(plot)` provides a way to add other plots to the generated graph; see [G-3] `addplot_option`.

Y axis, X axis, Titles, Legend, Overall

`twoway_options` are any of the options documented in [G-3] `twoway_options`, excluding `by()`. These include options for titling the graph (see [G-3] `title_options`) and for saving the graph to disk (see [G-3] `saving_option`).

Remarks and examples

`stata.com`

Remarks are presented under the following headings:

- Introduction
- Nonparametric ROC curves
- Lorenz-like curves

Introduction

The `roctab` command provides nonparametric estimation of the ROC for a given classifier and true-status reference variable. The Lorenz curve functionality of `roctab`, which provides an alternative to standard ROC analysis, is discussed in Lorenz-like curves.


Nonparametric ROC curves

The points on the nonparametric ROC curve are generated using each possible outcome of the diagnostic test as a classification cutpoint and computing the corresponding sensitivity and $1 - \text{specificity}$. These points are then connected by straight lines, and the area under the resulting ROC curve is computed using the trapezoidal rule.

Example 1

Hanley and McNeil (1982) presented data from a study in which a reviewer was asked to classify, using a five-point scale, a random sample of 109 tomographic images from patients with neurological problems. The rating scale was as follows: 1 = definitely normal, 2 = probably normal, 3 = questionable, 4 = probably abnormal, and 5 = definitely abnormal. The true disease status was normal for 58 of the patients and abnormal for the remaining 51 patients.
Here we list 9 of the 109 observations:

```
. use http://www.stata-press.com/data/r13/hanley
. list disease rating in 1/9

disease rating
1. 1 5
2. 0 1
3. 1 5
4. 0 4
5. 0 1
6. 0 3
7. 1 5
8. 0 5
9. 0 1
```

For each observation, `disease` identifies the true disease status of the subject (0 = normal, 1 = abnormal), and `rating` contains the classification value assigned by the reviewer.

We can use `roctab` to calculate and plot the nonparametric ROC curve by specifying both the `summary` and `graph` options. By also specifying the `table` option, we obtain a contingency table summarizing our dataset.

```
. roctab disease rating, table graph summary

disease rating

   rating    | disease
          | 1   2   3   4   5   Total
   ---------|--------
     0      | 33   6   6   11   2    58
     1      | 3    2   2   11   33  51
     Total  | 36   8   8   22   35  109

ROC Asymptotic Normal—
Obs   Area  Std. Err.  [95% Conf. Interval]
     109  0.8932  0.0307   0.83295  0.95339

By default, `roctab` reports the area under the curve, its standard error, and its confidence interval. The `graph` option can be used to plot the ROC curve.
The ROC curve is plotted by computing the sensitivity and specificity using each value of the rating variable as a possible cutpoint. A point is plotted on the graph for each of the cutpoints. These plotted points are joined by straight lines to form the ROC curve, and the area under the ROC curve is computed using the trapezoidal rule.

We can tabulate the computed sensitivities and specificities for each of the possible cutpoints by specifying `detail`.

```
. roctab disease rating, detail
Detailed report of sensitivity and specificity
```

<table>
<thead>
<tr>
<th>Cutpoint</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Correctly Classified</th>
<th>LR+</th>
<th>LR-</th>
</tr>
</thead>
<tbody>
<tr>
<td>( &gt;= 1  )</td>
<td>100.00%</td>
<td>0.00%</td>
<td>46.79%</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>( &gt;= 2  )</td>
<td>94.12%</td>
<td>56.90%</td>
<td>74.31%</td>
<td>2.1835</td>
<td>0.1034</td>
</tr>
<tr>
<td>( &gt;= 3  )</td>
<td>90.20%</td>
<td>67.24%</td>
<td>77.98%</td>
<td>2.7534</td>
<td>0.1458</td>
</tr>
<tr>
<td>( &gt;= 4  )</td>
<td>86.27%</td>
<td>77.59%</td>
<td>81.65%</td>
<td>3.8492</td>
<td>0.1769</td>
</tr>
<tr>
<td>( &gt;= 5  )</td>
<td>64.71%</td>
<td>96.55%</td>
<td>81.65%</td>
<td>18.7647</td>
<td>0.3655</td>
</tr>
<tr>
<td>( &gt; 5   )</td>
<td>0.00%</td>
<td>100.00%</td>
<td>53.21%</td>
<td></td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Each cutpoint in the table indicates the ratings used to classify tomographs as being from an abnormal subject. For example, the first cutpoint (>= 1) indicates that all tomographs rated as 1 or greater are classified as coming from abnormal subjects. Because all tomographs have a rating of 1 or greater, all are considered abnormal. Consequently, all abnormal cases are correctly classified (sensitivity = 100%), but none of the normal patients is classified correctly (specificity = 0%). For the second cutpoint (>=2), tomographs with ratings of 1 are classified as normal, and those with ratings of 2 or greater are classified as abnormal. The resulting sensitivity and specificity are 94.12% and 56.90%, respectively. Using this cutpoint, we correctly classified 74.31% of the 109 tomographs. Similar interpretations can be used on the remaining cutpoints. As mentioned, each cutpoint corresponds to a point on the nonparametric ROC curve. The first cutpoint (>=1) corresponds to the point at (1,1), and the last cutpoint (> 5) corresponds to the point at (0,0).

detail also reports two likelihood ratios suggested by Choi (1998): the likelihood ratio for a positive test result (LR+) and the likelihood ratio for a negative test result (LR–). The LR+ is the ratio of the probability of a positive test among the truly positive subjects to the probability of a positive test among the truly negative subjects. The LR– is the ratio of the probability of a negative test among the truly positive subjects to the probability of a negative test among the truly negative subjects. Choi points out that LR+ corresponds to the slope of the line from the origin to the point on the ROC curve determined by the cutpoint. Similarly, LR– corresponds to the slope from the point (1,1) to the point on the ROC curve determined by the cutpoint.

By default, `roctab` calculates the standard error for the area under the curve by using an algorithm suggested by DeLong, DeLong, and Clarke-Pearson (1988) and asymptotic normal confidence intervals. Optionally, standard errors based on methods suggested by Bamber (1975) or Hanley and McNeil (1982) can be computed by specifying `bamber` or `hanley`, respectively, and an exact binomial confidence interval can be obtained by specifying `binomial`. 
Lorenz-like curves

For applications where it is known that the risk status increases or decreases monotonically with increasing values of the diagnostic test, the ROC curve and associated indices are useful in assessing the overall performance of a diagnostic test. When the risk status does not vary monotonically with increasing values of the diagnostic test, however, the resulting ROC curve can be nonconvex and its indices can be unreliable. For these situations, Lee (1999) proposed an alternative to the ROC analysis based on Lorenz-like curves and the associated Pietra and Gini indices.

Lee (1999) mentions at least three specific situations where results from Lorenz curves are superior to those obtained from ROC curves: 1) a diagnostic test with similar means but very different standard deviations in the abnormal and normal populations, 2) a diagnostic test with bimodal distributions in either the normal or abnormal population, and 3) a diagnostic test distributed symmetrically in the normal population and skewed in the abnormal.

When the risk status increases or decreases monotonically with increasing values of the diagnostic test, the ROC and Lorenz curves yield interchangeable results.

Example 2

To illustrate the use of the `lorenz` option, we constructed a fictitious dataset that yields results similar to those presented in Table III of Lee (1999). The data assume that a 12-point rating scale was used to classify 442 diseased and 442 healthy subjects. We list a few of the observations.

```
> use http://www.stata-press.com/data/r13/lorenz, clear
> list in 1/7, noobs sep(0)
```

<table>
<thead>
<tr>
<th>disease</th>
<th>class</th>
<th>pop</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>5</td>
<td>66</td>
</tr>
<tr>
<td>1</td>
<td>11</td>
<td>17</td>
</tr>
<tr>
<td>0</td>
<td>6</td>
<td>85</td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>19</td>
</tr>
<tr>
<td>0</td>
<td>10</td>
<td>19</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>16</td>
</tr>
</tbody>
</table>
```

The data consist of 24 observations: 12 observations from diseased individuals and 12 from nondiseased individuals. Each observation corresponds to one of the 12 classification values of the rating-scale variable, `class`. The number of subjects represented by each observation is given by the `pop` variable, making this a frequency-weighted dataset. The data were generated assuming a binormal distribution of the latent variable with similar means for the normal and abnormal populations but with the standard deviation for the abnormal population five times greater than that of the normal population.
The resulting ROC curve is nonconvex or, as termed by Lee, “wiggly”. Lee argues that for this and similar situations, the Lorenz curve and indices are preferred.

Like ROC curves, a more bowed Lorenz curve suggests a better diagnostic test. This bowedness is quantified by the Pietra index, which is geometrically equivalent to twice the largest triangle that can be inscribed in the area between the curve and the diagonal line, and the Gini index, which is
equivalent to twice the area between the Lorenz curve and the diagonal. Lee (1999) provides several additional interpretations for the Pietra and Gini indices.

Stored results

`roctab` stores the following in `r()`:

Scalars

- `r(N)` number of observations
- `r(se)` standard error for the area under the ROC curve
- `r(lb)` lower bound of CI for the area under the ROC curve
- `r(ub)` upper bound of CI for the area under the ROC curve
- `r(area)` area under the ROC curve
- `r(pietra)` Pietra index
- `r(gini)` Gini index

Methods and formulas

Assume that we applied a diagnostic test to each of \( N_n \) normal and \( N_a \) abnormal subjects. Further assume that the higher the outcome value of the diagnostic test, the higher the risk of the subject being abnormal. Let \( \hat{\theta} \) be the estimated area under the curve, and let \( X_i, i = 1, 2, \ldots, N_a \) and \( Y_j, j = 1, 2, \ldots, N_n \) be the values of the diagnostic test for the abnormal and normal subjects, respectively.

The points on the nonparametric ROC curve are generated using each possible outcome of the diagnostic test as a classification cutpoint and computing the corresponding sensitivity and \( 1 - \text{specificity} \). These points are then connected by straight lines, and the area under the resulting ROC curve is computed using the trapezoidal rule.

The default standard error for the area under the ROC curve is computed using the algorithm described by DeLong, DeLong, and Clarke-Pearson (1988). For each abnormal subject, \( i \), define

\[
V_{10}(X_i) = \frac{1}{N_n} \sum_{j=1}^{N_n} \psi(X_i, Y_j)
\]

and for each normal subject, \( j \), define

\[
V_{01}(Y_j) = \frac{1}{N_a} \sum_{i=1}^{N_a} \psi(X_i, Y_j)
\]

where

\[
\psi(X, Y) = \begin{cases} 
1 & Y < X \\
\frac{1}{2} & Y = X \\
0 & Y > X 
\end{cases}
\]
Define
\[ S_{10} = \frac{1}{N_a - 1} \sum_{i=1}^{N_a} \{V_{10}(X_i) - \hat{\theta}\}^2 \]
and
\[ S_{01} = \frac{1}{N_n - 1} \sum_{j=1}^{N_n} \{V_{01}(Y_j) - \hat{\theta}\}^2 \]

The variance of the estimated area under the ROC curve is given by
\[ \text{var}(\hat{\theta}) = \frac{1}{N_a} S_{10} + \frac{1}{N_n} S_{01} \]

The Hanley standard error for the area under the ROC curve is computed using the algorithm described by Hanley and McNeil (1982). It requires the calculation of two quantities: \( Q_1 \) is \( \text{Pr} \) (two randomly selected abnormal subjects will both have a higher score than a randomly selected normal subject), and \( Q_2 \) is \( \text{Pr} \) (one randomly selected abnormal subject will have a higher score than any two randomly selected normal subjects). The Hanley and McNeil variance of the estimated area under the ROC curve is
\[ \text{var}(\hat{\theta}) = \frac{\hat{\theta}(1 - \hat{\theta}) + (N_a - 1)(Q_1 - \hat{\theta}^2) + (N_n - 1)(Q_2 - \hat{\theta}^2)}{N_a N_n} \]

The Bamber standard error for the area under the ROC curve is computed using the algorithm described by Bamber (1975). For any two \( Y \) values, \( Y_j \) and \( Y_k \), and any \( X_i \) value, define
\[ b_{yyx} = p(Y_j, Y_k < X_i) + p(X_i < Y_j, Y_k) - 2p(Y_j < X_i < Y_k) \]
and similarly, for any two \( X \) values, \( X_i \) and \( X_l \), and any \( Y_j \) value, define
\[ b_{xxy} = p(X_i, X_l < Y_j) + p(Y_j < X_i, X_l) - 2p(X_i < Y_j < X_l) \]

Bamber’s unbiased estimate of the variance for the area under the ROC curve is
\[ \text{var}(\hat{\theta}) = \frac{1}{4} (N_a - 1)(N_n - 1) \{p(X \neq Y) + (N_a - 1)b_{xx} + (N_n - 1)b_{yy} - 4(N_a + N_n - 1)(\hat{\theta} - 0.5)^2\} \]

Asymptotic confidence intervals are constructed and reported by default, assuming a normal distribution for the area under the ROC curve.

Exact binomial confidence intervals are calculated as described in \([R]\) ci, with \( p \) equal to the area under the ROC curve.

References


——. 2002b. From the help desk: Comparing areas under receiver operating characteristic curves from two or more probit or logit models. *Stata Journal* 2: 301–313.


Also see

[R] logistic postestimation — Postestimation tools for logistic

[R] roc — Receiver operating characteristic (ROC) analysis

[R] roccomp — Tests of equality of ROC areas

[R] rocfit — Parametric ROC models

[R] rocreg — Receiver operating characteristic (ROC) regression