

Chapter 260

Tests for One ROC Curve

Introduction

Receiver operating characteristic (ROC) curves are used to assess the accuracy of a diagnostic test. The technique is used when you have a criterion variable which will be used to make a yes or no decision based on the value of this variable. The area under the ROC curve (AUC) is a popular summary index of an ROC curve.

This module computes power and sample size when a new diagnostic test is compared to an existing (gold) standard. Two approaches are available: the approach of Hanley and McNeil (1982) is used when the criterion variable is continuous and the approach of Obuchowski and McClish (1997) is used when the criterion variable is a discrete rating scale.

Technical Details

In the following, we suppose that we have two groups of patients, those with a condition of interest (the positive group) and those without it (the negative group). This classification may be known from extensive diagnosis or based on the value of another diagnostic test. The diagnostic test of interest is performed on each patient and the resulting test value is recorded. At each specified cutoff value of the criterion variable, the true positive rate (TPR) and the false positive rate (FPR) are calculated. A plot of the TPR versus the FPR allows you study the consequences of using various cutoff values. This plot is called the *ROC curve*.

It should be noted that TPR is similar to the statistical power of the diagnostic test at a particular cutoff value of the criterion variable. Similarly, FPR is an estimate of the probability that the diagnostic test results in a type I (alpha) error. Thus the ROC curve may be interpreted as a plot of the diagnostic test's power versus its significance level at various possible criterion cutoff values.

Users of ROC curves have developed special names for TPR and FPR. They call TPR the *sensitivity* of the test and $1 - \text{FPR}$ the *specificity* of the test. Statisticians will be more familiar with using the word *power* instead of sensitivity and the phrase ' $1 - \alpha$ ' instead of specificity.

An ROC curve may be summarized by the area under it (AUC). This area has an additional interpretation. Suppose that a rater is asked to study two subjects, one that is actually disease positive and one that is disease negative. The AUC is equal to the probability that the rater will give the disease positive subject a higher score than the disease negative subject. That is, the AUC is the probability that the rater will correctly order the two subjects as to which is more likely to have the disease.

Several methods of computing the AUC have been proposed. One method uses the trapezoidal rule to calculate the AUC directly. Another method, called the *binormal model*, computes the area by fitting two normal distributions to the data.

Tests for One ROC Curve

The Binormal Model

Let X denote the distribution of the criterion variable for negative (normal) patients and Y denote the distribution of the criterion variable for positive (diseased) patients. It is assumed that

$$X \sim N(\mu_-, \sigma_-^2)$$

and

$$Y \sim N(\mu_+, \sigma_+^2)$$

For a particular cutoff value of the criterion variable, c , the true positive rate is given by

$$\begin{aligned} TPR(c) &= P(Y > c) \\ &= 1 - \Phi\left(\frac{c - \mu_+}{\sigma_+}\right) \\ &= \Phi\left(\frac{\mu_+ - c}{\sigma_+}\right) \end{aligned}$$

where $\Phi(z)$ is the cumulative normal distribution.

Similarly, the false positive rate is given by

$$\begin{aligned} FPR(c) &= P(X > c) \\ &= 1 - \Phi\left(\frac{c - \mu_-}{\sigma_-}\right) \\ &= \Phi\left(\frac{\mu_- - c}{\sigma_-}\right) \end{aligned}$$

The ROC curve is thus the curve traced out by the functions

$$\left[FPR(c), TPR(c) \right] = \left[\Phi\left(\frac{\mu_- - c}{\sigma_-}\right), \Phi\left(\frac{\mu_+ - c}{\sigma_+}\right) \right]$$

The area under the ROC curve, AUC, is defined as

$$\begin{aligned} \theta &= \int_{-\infty}^{\infty} TPR(c) FPR'(c) dc \\ &= \int_{-\infty}^{\infty} \Phi\left(\frac{\mu_+ - c}{\sigma_+}\right) \phi\left(\frac{\mu_- - c}{\sigma_-}\right) \left(-\frac{1}{\sigma_-}\right) dc \\ &= \int_{-\infty}^{\infty} \Phi(A + Bv) \phi(v) dv \\ &= \Phi\left(\frac{A}{\sqrt{1 + B^2}}\right) \end{aligned}$$

Tests for One ROC Curve

where

$$c = \mu_- - v\sigma_-$$

$$A = \frac{|\mu_+ - \mu_-|}{\sigma_+}$$

$$B = \frac{\sigma_-}{\sigma_+}$$

Maximum likelihood estimates of A and B can be computed and used to compute AUC. The variances and covariance of these MLE's can be estimated from Fisher's information matrix.

Define $\Delta = \theta_1 - \theta_0$ to be the difference in the accuracies (AUC's) between the new and standard tests. The test statistic for this test is

$$Z = \frac{\hat{\Delta}}{\sqrt{V(\theta_0)}} = \frac{\hat{\theta}_1 - \theta_0}{\sqrt{V(\theta_0)}}$$

where $V(\theta_0)$ is the variance of $\hat{\theta}$ under the null hypothesis of equality. The above test statistic gives the following formulae for computing sample size or power

$$N_+ = \frac{\left(z_\alpha \sqrt{V(\theta_0)} + z_\beta \sqrt{V(\theta_1)}\right)^2}{(\theta_1 - \theta_0)^2}$$

$$\beta = \Phi\left(\frac{|\theta_1 - \theta_0| \sqrt{N_+} - z_\alpha \sqrt{V(\theta_0)}}{\sqrt{V(\theta_1)}}\right)$$

Rating Data

For a criterion variable yielding a discrete rating and with $R = N_-/N_+$, Obuchowski (1998) recommends

$$V(\theta) = f^2 \left(1 + \frac{B^2}{R} + \frac{A^2}{2}\right) + g^2 B^2 \left(\frac{1+R}{2R}\right)$$

where

$$f = \frac{E_1}{\sqrt{2\pi(1+B^2)}}$$

$$g = -\frac{ABE_1}{\sqrt{2\pi(1+B^2)^3}}$$

$$E_1 = \exp\left(-\frac{A^2}{2+2B^2}\right)$$

Tests for One ROC Curve

The value of A can be found as

$$A = \Phi^{-1}(\theta)\sqrt{1 + B^2}$$

For the most conservative results, Obuchowski (1998) recommends setting $B = 1$, so that

$$A = \Phi^{-1}(\theta)\sqrt{2}$$

Continuous Data

For a criterion variable yielding a continuous result and with $R = N_-/N_+$, Obuchowski (1998) suggests that the following formula of Hanley and McNeil (1983) is more appropriate

$$V(\theta) = \frac{\theta}{R(2 - \theta)} + \frac{2\theta^2}{1 + \theta} - \theta^2 \left(\frac{1 + R}{R} \right)$$

Procedure Options

This section describes the options that are specific to this procedure. These are located on the Design tab. For more information about the options of other tabs, go to the Procedure Window chapter.

Design Tab

Solve For

Solve For

This option specifies the parameter to be solved for from the other parameters. Under most situations, you will select either *Power* or *Sample Size (N+)*.

Select *Sample Size (N+)* when you want to calculate the sample size needed to achieve a given power and alpha level.

Select *Power* when you want to calculate the power of an experiment that has already been run.

Test

Alternative Hypothesis

Specify whether the test is *one-sided* or *two-sided*. When a two-sided test is selected, the value of alpha is divided by two.

Note that most researchers assume that, unless stated otherwise, all statistical tests are two-sided. If you use a one-sided test, you should clearly state and justify this in all reports.

Tests for One ROC Curve

Power and Alpha

Power

This option specifies one or more values for power. Power is the probability of rejecting a false null hypothesis, and is equal to one minus Beta. Beta is the probability of a type-II error, which occurs when a false null hypothesis is not rejected.

Values must be between zero and one. Historically, the value of 0.80 (Beta = 0.20) was used for power. Now, 0.90 (Beta = 0.10) is also commonly used.

A single value may be entered here or a range of values such as *0.8 to 0.95 by 0.05* may be entered.

Alpha

This option specifies one or more values for the probability of a type-I error. A type-I error occurs when a true null hypothesis is rejected.

Values must be between zero and one. Historically, the value of 0.05 has been used for alpha. This means that about one test in twenty will falsely reject the null hypothesis. You should pick a value for alpha that represents the risk of a type-I error you are willing to take in your experimental situation.

You may enter a range of values such as *0.01 0.05 0.10* or *0.01 to 0.10 by 0.01*.

Sample Size (When Solving for Sample Size)

Group Allocation

Select the option that describes the constraints on N_+ or N_- or both.

The options are

- **Equal ($N_+ = N_-$)**

This selection is used when you wish to have equal sample sizes in each group. Since you are solving for both sample sizes at once, no additional sample size parameters need to be entered.

- **Enter N_+ , solve for N_-**

Select this option when you wish to fix N_+ at some value (or values), and then solve only for N_- . Please note that for some values of N_+ , there may not be a value of N_- that is large enough to obtain the desired power.

- **Enter N_- , solve for N_+**

Select this option when you wish to fix N_- at some value (or values), and then solve only for N_+ . Please note that for some values of N_- , there may not be a value of N_+ that is large enough to obtain the desired power.

- **Enter $R = N_-/N_+$, solve for N_+ and N_-**

For this choice, you set a value for the ratio of N_- to N_+ , and then PASS determines the needed N_+ and N_- , with this ratio, to obtain the desired power. An equivalent representation of the ratio, R , is

$$N_- = R * N_+.$$

- **Enter percentage in Group 1, solve for N_+ and N_-**

For this choice, you set a value for the percentage of the total sample size that is in Group 1, and then PASS determines the needed N_+ and N_- with this percentage to obtain the desired power.

Tests for One ROC Curve

N+ (Sample Size, Group 1)

This option is displayed if Group Allocation = "Enter N+, solve for N-"

N_+ is the number of items or individuals sampled from the Group 1 population.

N_+ must be ≥ 2 . You can enter a single value or a series of values.

N- (Sample Size, Group 2)

This option is displayed if Group Allocation = "Enter N-, solve for N+"

N_- is the number of items or individuals sampled from the Group 2 population.

N_- must be ≥ 2 . You can enter a single value or a series of values.

R (Group Sample Size Ratio)

This option is displayed only if Group Allocation = "Enter R = N-/N+, solve for N+ and N-."

R is the ratio of N_- to N_+ . That is,

$$R = N_- / N_+.$$

Use this value to fix the ratio of N_- to N_+ while solving for N_+ and N_- . Only sample size combinations with this ratio are considered.

N_- is related to N_+ by the formula:

$$N_- = [R \times N_+],$$

where the value $[Y]$ is the next integer $\geq Y$.

For example, setting $R = 2.0$ results in a Group 2 sample size that is double the sample size in Group 1 (e.g., $N_+ = 10$ and $N_- = 20$, or $N_+ = 50$ and $N_- = 100$).

R must be greater than 0. If $R < 1$, then N_- will be less than N_+ ; if $R > 1$, then N_- will be greater than N_+ . You can enter a single or a series of values.

Percent in Group 1 (+)

This option is displayed only if Group Allocation = "Enter percentage in Group 1, solve for N+ and N-."

Use this value to fix the percentage of the total sample size allocated to Group 1 while solving for N_+ and N_- . Only sample size combinations with this Group 1 percentage are considered. Small variations from the specified percentage may occur due to the discrete nature of sample sizes.

The Percent in Group 1 must be greater than 0 and less than 100. You can enter a single or a series of values.

Sample Size (When Not Solving for Sample Size)

Group Allocation

Select the option that describes how individuals in the study will be allocated to Group 1 and to Group 2.

The options are

- **Equal ($N_+ = N_-$)**

This selection is used when you wish to have equal sample sizes in each group. A single per group sample size will be entered.

- **Enter N_+ and N_- individually**

This choice permits you to enter different values for N_+ and N_- .

Tests for One ROC Curve

- **Enter N_+ and R , where $N_- = R * N_+$**
Choose this option to specify a value (or values) for N_+ , and obtain N_- as a ratio (multiple) of N_+ .
- **Enter total sample size and percentage in Group 1**
Choose this option to specify a value (or values) for the total sample size (N), obtain N_+ as a percentage of N , and then N_- as $N - N_+$.

Sample Size Per Group

This option is displayed only if Group Allocation = "Equal ($N_+ = N_-$)."

The Sample Size Per Group is the number of items or individuals sampled from each of the Group 1 and Group 2 populations. Since the sample sizes are the same in each group, this value is the value for N_+ , and also the value for N_- .

The Sample Size Per Group must be ≥ 2 . You can enter a single value or a series of values.

N_+ (Sample Size, Group 1)

*This option is displayed if Group Allocation = "Enter N_+ and N_- individually" or "Enter N_+ and R , where $N_- = R * N_+$."*

N_+ is the number of items or individuals sampled from the Group 1 population.

N_+ must be ≥ 2 . You can enter a single value or a series of values.

N_- (Sample Size, Group 2)

This option is displayed only if Group Allocation = "Enter N_+ and N_- individually."

N_- is the number of items or individuals sampled from the Group 2 population.

N_- must be ≥ 2 . You can enter a single value or a series of values.

R (Group Sample Size Ratio)

*This option is displayed only if Group Allocation = "Enter N_+ and R , where $N_- = R * N_+$."*

R is the ratio of N_- to N_+ . That is,

$$R = N_-/N_+$$

Use this value to obtain N_- as a multiple (or proportion) of N_+ .

N_- is calculated from N_+ using the formula:

$$N_- = [R \times N_+],$$

where the value $[Y]$ is the next integer $\geq Y$.

For example, setting $R = 2.0$ results in a Group 2 sample size that is double the sample size in Group 1.

R must be greater than 0. If $R < 1$, then N_- will be less than N_+ ; if $R > 1$, then N_- will be greater than N_+ . You can enter a single value or a series of values.

Total Sample Size (N)

This option is displayed only if Group Allocation = "Enter total sample size and percentage in Group 1."

This is the total sample size, or the sum of the two group sample sizes. This value, along with the percentage of the total sample size in Group 1, implicitly defines N_+ and N_- .

The total sample size must be greater than one, but practically, must be greater than 3, since each group sample size needs to be at least 2.

You can enter a single value or a series of values.

Tests for One ROC Curve

Percent in Group 1 (+)

This option is displayed only if Group Allocation = "Enter total sample size and percentage in Group 1."

This value fixes the percentage of the total sample size allocated to Group 1. Small variations from the specified percentage may occur due to the discrete nature of sample sizes.

The Percent in Group 1 must be greater than 0 and less than 100. You can enter a single value or a series of values.

Effect Size – Area Under the Curve

AUC0 (Area Under Curve|H0)

Specify one or more values of the AUC for the diagnostic test under the null hypothesis. If you are comparing a new diagnostic test to a standard, this is the hypothesized value of the STANDARD test. The range of values is from 0.5 (indicative of a test useless in diagnosis) to 1.0 (indicative of a test that is perfect in diagnosis).

Since the AUC may include a portion of the ROC curve that is not of interest because the FPR values are unrealistic, you may be interested in only a portion of the area. In this case, you can specify a range of FPR values for which the area is to be calculated. Unfortunately, the definition of the area becomes more difficult. When analyzing the whole ROC curve, the area is known to be between 0.50 and 1.0. Following the suggestion of Obuchowski and McClish (1997), the following transformation is applied so that the values of AUC remain between 0.5 and 1.0.

$$AUC' = \frac{1}{2} \left(1 + \frac{AUC - \min}{\max - \min} \right)$$

where

$$\max = FPR2 - FPR1$$

$$\min = \frac{\max}{2} (FPR2 + FPR1)$$

Thus, when a partial range is entered for FPR1 and FPR2, the values entered here are assumed to be AUC' and are translated to AUC using the above formulas.

AUC1 (Area Under Curve|H1)

Specify one or more values of AUC under the alternative hypothesis. If you are comparing a new diagnostic test to a standard, this is the hypothesized value of the NEW test. The range of values is from 0.5 (indicative of a test useless in diagnosis) to 1.0 (indicative of a test that is perfect in diagnosis). Note that, as discussed above, this is the value of AUC' when a partial area is being analyzed.

Effect Size – False Positive Rate Limits

Lower FPR

This option specifies the lower (left) limit of the false positive rate (FPR) for which the area is to be computed. If the area under the whole ROC curve is wanted, set this value to 0.0. If the partial area is wanted, set this value to the desired left limit.

Note that the range of possible values is $0.0 \leq \text{Lower FPR} < \text{Upper FPR} \leq 1.0$

Upper FPR

This option specifies the upper (right) limit of the false positive rate (FPR). If the area under the whole ROC curve is wanted, set this value to 1.0. If the partial area is wanted, set this value to the desired right limit.

Note that the range of possible values is $0.0 \leq \text{Lower FPR} < \text{Upper FPR} \leq 1.0$

Tests for One ROC Curve

Effect Size – Type of Data**Type of Data**

Specify the type of data that will be collected from the tests. The formulas for the variance are determined by this option. Possible types are:

- **Continuous**

The test results are from a continuum of possible values. The Hanley and McNeil (1983) variance formulas are used. Note that this option does not allow a partial range of FPR values to be analyzed.

- **Discrete**

The test results are from a small set of rating values such as 1, 2, 3, 4, 5. The Obuchowski & McClish (1997) variance formulas are used.

B (SD Ratio = SD-/SD+)

B is the ratio of the standard deviation of the negative group to the positive group (SD-/SD+) for the diagnostic test. That is, assuming the binormal model

$$B = \frac{\sigma_-}{\sigma_+}$$

Note that this parameter is ignored for continuous data.

Although B can be any positive number, typical values are between 0.3 and 3.0. Obuchowski suggests that if the value of B is not known, a value of 1.0 is used since this will result in a conservative (extra large) sample size. She reports that in her experience, typical values are much less than 1.0, often near 0.3.

Tests for One ROC Curve

Example 1 – Calculating Power

An investigator wants to study the accuracy of a diagnostic test which yields measurements on a rating scale from 1 to 5. Historically, such tests have had an AUC of 0.80. The investigator wants to investigate three alternative AUC values: 0.825, 0.850, and 0.900. A two-sided test is planned with a significance level of 0.05. Since no other information is available, B is set to 1.0. The investigator would like to achieve a power of 90% in the study. Patients without the disease under study are about twice as frequent as patients with the disease. The investigator wants to see results for a sample size of up to 6000 patients.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Tests for One ROC Curve** procedure window by clicking on **ROC**, and then clicking on **Tests for One ROC Curve**. You may then make the appropriate entries as listed below, or open **Example 1** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Power
Alternative Hypothesis	Two-Sided Test
Alpha	0.05
Group Allocation	Enter N+ and R, where N- = R x N+
N+ (Size of Positive Group)	20 50 100 250 500 1000 2000
R (Sample Allocation Ratio)	2
AUC0 (Area Under Curve H0)	0.80
AUC1 (Area Under Curve H1)	0.825 0.85 0.90
Lower FPR	0.00
Upper FPR	1.00
Type of Data	Discrete (Ratings)
B (SD Ratio = SD-/SD+)	1.0

Annotated Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Report

Numeric Results for Testing AUC0 = AUC1 with Discrete (Rating) Data

Test Type = Two-Sided. FPR1 = 0.0. FPR2 = 1.0. B = 1.000. Allocation Ratio = 2.000.

Power	N+	N-	N	Target	Actual	AUC0'	AUC1'	Diff'	AUC0	AUC1	Diff	Alpha
				R	R							
0.04806	20	40	60	2.00	2.00	0.8000	0.8250	0.0250	0.8000	0.8250	0.0250	0.050
0.07393	50	100	150	2.00	2.00	0.8000	0.8250	0.0250	0.8000	0.8250	0.0250	0.050
0.11455	100	200	300	2.00	2.00	0.8000	0.8250	0.0250	0.8000	0.8250	0.0250	0.050
0.23648	250	500	750	2.00	2.00	0.8000	0.8250	0.0250	0.8000	0.8250	0.0250	0.050
0.43207	500	1000	1500	2.00	2.00	0.8000	0.8250	0.0250	0.8000	0.8250	0.0250	0.050
0.72636	1000	2000	3000	2.00	2.00	0.8000	0.8250	0.0250	0.8000	0.8250	0.0250	0.050
0.95496	2000	4000	6000	2.00	2.00	0.8000	0.8250	0.0250	0.8000	0.8250	0.0250	0.050
0.08703	20	40	60	2.00	2.00	0.8000	0.8500	0.0500	0.8000	0.8500	0.0500	0.050
0.18340	50	100	150	2.00	2.00	0.8000	0.8500	0.0500	0.8000	0.8500	0.0500	0.050
0.34913	100	200	300	2.00	2.00	0.8000	0.8500	0.0500	0.8000	0.8500	0.0500	0.050
0.73689	250	500	750	2.00	2.00	0.8000	0.8500	0.0500	0.8000	0.8500	0.0500	0.050
0.96287	500	1000	1500	2.00	2.00	0.8000	0.8500	0.0500	0.8000	0.8500	0.0500	0.050
0.99968	1000	2000	3000	2.00	2.00	0.8000	0.8500	0.0500	0.8000	0.8500	0.0500	0.050

(report continues)

Tests for One ROC Curve

Report Definitions

Power is the probability of rejecting a false null hypothesis.

N_+ and N_- are the number of items sampled from each population.

N is the total sample size, $N_+ + N_-$.

Target R is the desired ratio (or ratios) of R entered in the procedure. R is the ratio of N_- to N_+ , so that

$$N_- = R \times N_+$$

Actual R is the value for R obtained in this scenario. Because N_+ and N_- are discrete, this value is sometimes slightly different than the target R .

AUC_0' and AUC_1' are the adjusted areas under the ROC curve for the null and alternative hypotheses, respectively.

Diff' is $AUC_1 - AUC_0$. This is the adjusted difference to be detected.

AUC_0 and AUC_1 are the actual areas under the ROC curve for the null and alternative hypotheses, respectively.

Diff is $AUC_1 - AUC_0$. This is the difference to be detected.

Alpha is the probability of rejecting a true null hypothesis.

FPR1, FPR2 are the lower and upper bounds on the false positive rates.

B is the ratio of the standard deviations of the negative and positive groups.

Summary Statements

A sample of 20 from the positive group and 40 from the negative group achieve 5% power to detect a difference of 0.0250 between the area under the ROC curve (AUC) under the null hypothesis of 0.8000 and an AUC under the alternative hypothesis of 0.8250 using a two-sided z-test at a significance level of 0.0500. The data are discrete (rating scale) responses. The AUC is computed between false positive rates of 0.000 and 1.000. The ratio of the standard deviation of the responses in the negative group to the standard deviation of the responses in the positive group is 1.000.

This report shows the power for each of the sample sizes. Most of the definitions are standard. However, a special explanation must be given for AUC and AUC' .

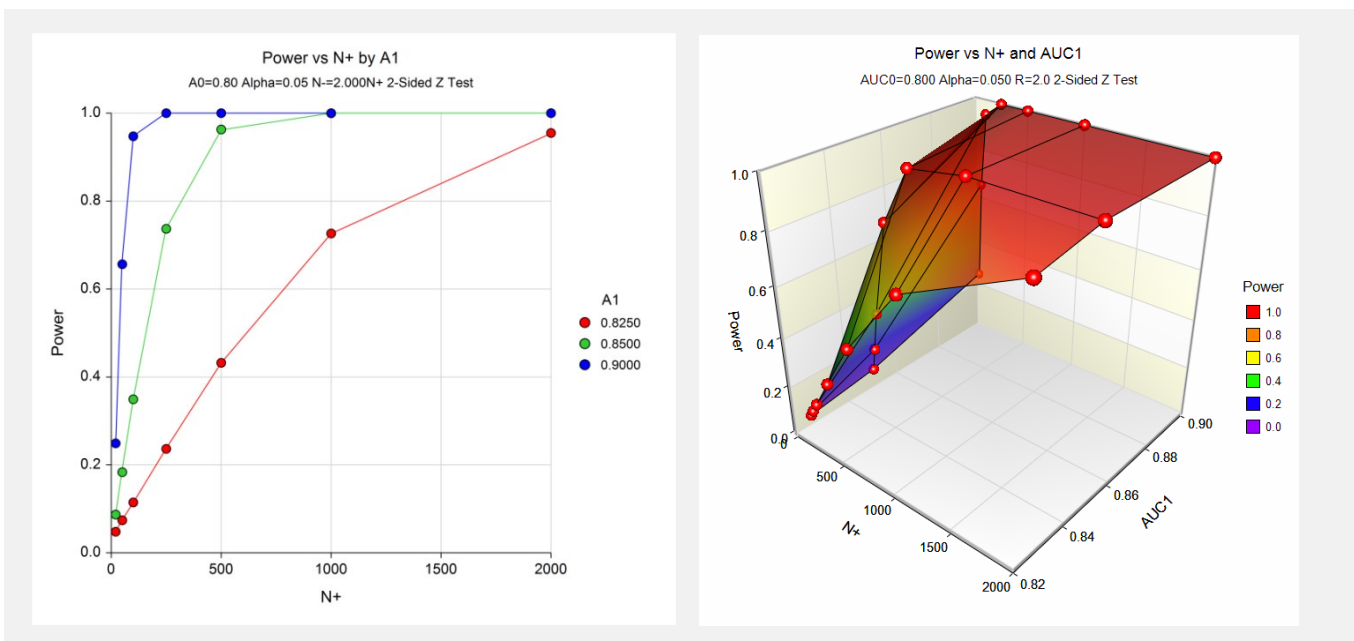
AUC'

This is the adjusted area under the curve. A rescaling, discussed earlier, has been applied so that the minimum area is 0.5 and the maximum area is 1.0.

AUC

This is the actual area under the curve. This value will equal the adjusted area when the FPR range is set from 0.0 to 1.0. Otherwise, these values will be different.

Plot Section



These plots show the power versus the sample size for the three values of AUC_1 .

Tests for One ROC Curve

Example 2 – Calculating Sample Size

Continuing on with Example 1, the investigator wants to know the exact sample size needed for each of the three values of AUC2. The investigator wants to look at the Numeric Report.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Tests for One ROC Curve** procedure window by clicking on **ROC**, and then clicking on **Tests for One ROC Curve**. You may then make the appropriate entries as listed below, or open **Example 2** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Sample Size
Alternative Hypothesis	Two-Sided Test
Power	0.90
Alpha	0.05
Group Allocation	Enter R = N-/N+, solve for N+ and N-
R (Sample Allocation Ratio)	2
AUC0 (Area Under Curve H0)	0.80
AUC1 (Area Under Curve H1)	0.825 0.85 0.90
Lower FPR	0.00
Upper FPR	1.00
Type of Data	Discrete (Ratings)
B (SD Ratio = SD-/SD+)	1.0

Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Report

Numeric Results for Testing AUC0 = AUC1 with Discrete (Rating) Data

Test Type = Two-Sided. FPR1 = 0.0. FPR2 = 1.0. B = 1.000. Allocation Ratio = 2.000.

Target Power	Actual Power	N+	N-	N	Target R	Actual R	AUC0'	AUC1'	Diff'	AUC0	AUC1	Diff	Alpha
0.90	0.90010	1582	3164	4746	2.00	2.00	0.8000	0.8250	0.0250	0.8000	0.8250	0.0250	0.050
0.90	0.90069	381	762	1143	2.00	2.00	0.8000	0.8500	0.0500	0.8000	0.8500	0.0500	0.050
0.90	0.90243	85	170	255	2.00	2.00	0.8000	0.9000	0.1000	0.8000	0.9000	0.1000	0.050

This report shows the sample size needed to achieve 90% power for each value of AUC1.

Tests for One ROC Curve

Example 3 – Partial Area under Curve

Continuing on with Example 2, the investigator knows that FPR values between 0.0 and 0.20 are the only values of interest. Hence, he wants to investigate the sample size needed when the FPR range is confined to this range.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Tests for One ROC Curve** procedure window by clicking on **ROC**, and then clicking on **Tests for One ROC Curve**. You may then make the appropriate entries as listed below, or open **Example 3** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Sample Size
Alternative Hypothesis	Two-Sided Test
Power.....	0.90
Alpha.....	0.05
Group Allocation	Enter R = N-/N+, solve for N+ and N-
R (Sample Allocation Ratio)	2
AUC0 (Area Under Curve H0)	0.80
AUC1 (Area Under Curve H1)	0.825 0.85 0.90
Lower FPR.....	0.00
Upper FPR.....	0.20
Type of Data	Discrete (Ratings)
B (SD Ratio = SD-/SD+)	1.0

Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Results

Numeric Results for Testing AUC0 = AUC1 with Discrete (Rating) Data

Test Type = Two-Sided. FPR1 = 0.0. FPR2 = 0.200. B = 1.000. Allocation Ratio = 2.000.

Target Power	Actual Power	N+	N-	N	Target R	Actual R	AUC0'	AUC1'	Diff'	AUC0	AUC1	Diff	Alpha
0.90	0.90009	2663	5326	7989	2.00	2.00	0.8000	0.8250	0.0250	0.1280	0.1370	0.0090	0.050
0.90	0.90018	645	1290	1935	2.00	2.00	0.8000	0.8500	0.0500	0.1280	0.1460	0.0180	0.050
0.90	0.90133	144	288	432	2.00	2.00	0.8000	0.9000	0.1000	0.1280	0.1640	0.0360	0.050

Note that the necessary sample size has almost doubled.

Tests for One ROC Curve

Example 4 – Validation using Obuchowski

The formulas used in this module were given in Obuchowski and McClish (1997). On page 1538, they provide an example which will be duplicated here. The study investigated the accuracy of MRI for detecting abnormalities in patients with symptomatic knees. In order to do this, they wanted to know the sample size that would be needed to construct a 95% confidence interval so that the length of the confidence interval is no more than 0.10.

The measure of diagnostic accuracy is the AUC from an FPR of 0.0 to an FPR of 1.0. The allocation ratio is 1.5. $B = 1.0$. The value of A is found to be 1.2. This translates to an AUC_0 of 0.7995. The value of $AUC_1 = AUC_0 + 0.10 / 2$, where 0.10 is the maximum length of the confidence interval. A two-tailed confidence interval is envisioned in which α is 0.05. In order to find the sample size of a confidence interval, the power is set to 50%. In their article, they found $N_+ = 161$ and $N_- = 242$.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Tests for One ROC Curve** procedure window by clicking on **ROC**, and then clicking on **Tests for One ROC Curve**. You may then make the appropriate entries as listed below, or open **Example 4** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Sample Size
Alternative Hypothesis	Two-Sided Test
Power.....	0.50
Alpha.....	0.05
Group Allocation	Enter R = N-/N+, solve for N+ and N-
R (Sample Allocation Ratio)	1.5
AUC0 (Area Under Curve H0)	0.7995
AUC1 (Area Under Curve H1)	0.8495
Lower FPR.....	0.00
Upper FPR.....	1.00
Type of Data	Discrete (Ratings)
B (SD Ratio = SD-/SD+)	1

Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Results

Numeric Results for Testing $AUC_0 = AUC_1$ with Discrete (Rating) Data

Test Type = Two-Sided. FPR1 = 0.0. FPR2 = 0.200. B1 = 1.000. B2 = 1.000. Allocation Ratio = 2.000.

Target Power	Actual Power	N+	N-	N	Target R	Actual R	AUC0'	AUC1'	Diff'	AUC0	AUC1	Diff	Alpha
0.50	0.50262	162	243	405	1.50	1.50	0.7995	0.8495	0.0500	0.7995	0.8495	0.0500	0.050

Note that the sample sizes of 162 and 243 are within one of the results of Obuchowski. The difference occurs because their values of 161 and 242 produce a power that is slightly less than 0.5, so PASS increased the sample size slightly.