

## Chapter 858

# Multiple Regression

---

### Introduction

This procedure computes power and sample size for a multiple regression analysis in which the relationship between a dependent variable  $Y$  and a set independent variables  $X_1, X_2, \dots, X_M$  is to be studied. In multiple regression, interest usually focuses on the regression coefficients. However, since the  $X$ 's are usually not available during the planning phase, little is known about these coefficients until after the analysis is run. Hence, this procedure uses the squared multiple correlation coefficient,  $R^2$ , as the measure of effect size upon which the power analysis and sample size is based. Gatsonis and Sampson (1989) present power analysis results for two approaches: *conditional* and *unconditional*. Both of these approaches are available in this procedure.

---

### Multiple Regression Model

Multiple regression uses a *linear model* to approximate the relationship between a *dependent variable*  $Y$  and one or more *independent variables*  $X_1, \dots, X_M$ . The theoretical model for the  $i^{\text{th}}$  observation is

$$Y_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_M X_{Mi} + e_i$$

where  $\beta_0, \beta_1, \dots, \beta_M$  are the unknown regression coefficients to be estimated with ordinary least squares and  $e_i$  is the error term or *residual*.

For convenience in the presentation, we will use a slightly different notation. We divide the  $X$ 's into two, non-overlapping subsets which are T (tested) and C (covariates). We reorder the  $X$ 's so that all of the subset T variables occur before the subset C variables. The theoretical model then becomes

$$Y_i = \beta_0 + \beta_1 T_{1i} + \dots + \beta_K T_{Ki} + \lambda_1 C_{1i} + \dots + \lambda_L C_{Li} + e_i$$

where  $T_1 = X_1, \dots, T_K = X_K, C_1 = X_{K+1}, \dots, C_L = X_M$ .

In what follows, we will present results for tests of the regression coefficients associated with the T variables  $T_1, \dots, T_K$ . Specifically, we will present a power analysis of tests that  $\beta_1 = \beta_2 = \dots = \beta_K = 0$ . Since the subset C variables are not included in the test, but are fit in the regression analysis, the test is said to control for (adjust for or hold constant) the covariates  $C_1, \dots, C_L$ .

## Conditional and Unconditional Power Calculation

In this section, we will define the two types of power calculation methods. Note that both methods anticipate the same test statistics.

These two methods are *conditional* and *unconditional*.

### Conditional

The conditional expectation assumption is usually used to justify significance tests and so it is the approach that is we recommend.

In this method, conditional expectation of  $Y$  is estimated given the realized values of the  $T$ 's and  $C$ 's. Since these values are almost always impossible to know, this method suggests using a range of  $R^2$  values.

The focus in the significance test is how much  $R^2$  increases when test variables are added to a regression model that already contains the covariates.

Define  $R_{T|C}^2 = R_{T,C}^2 - R_C^2$  to be the amount that  $R^2$  increases when  $Y$  is regressed on the variables in set  $T$  after adjusting for the variables in set  $C$ . Here,  $R_C^2$  is the  $R^2$  when  $Y$  is regressed on only those variables in set  $C$  and  $R_{T,C}^2$  is the  $R^2$  when  $Y$  is regressed on the variables in both sets.

### Test Statistic in the Conditional Formula

$F$ -tests can easily be constructed that will test whether the regression coefficients corresponding to certain subsets of  $X$ 's are simultaneously zero while controlling for other variables. For example, Rencher (2000) shows that to test the significance of the  $X$ 's in set  $T$  while removing the influence of the  $X$ 's in set  $C$  from experimental error, you would use

$$F_{K,N-K-L-1} = \frac{(R_{T|C}^2)/K}{(1 - R_C^2 - R_{T|C}^2)/(N - K - L - 1)}$$

where  $K$  is the number of variables in  $T$  and  $L$  is the number of variables in  $C$ .

### Calculating the Power in the Conditional Model

In this case, power calculations are based on the noncentral- $F$  distribution. The calculation of the power of a particular test proceeds as follows:

1. Determine the critical value  $F_{K,N-K-L-1,\alpha}$  where  $\alpha$  is the probability of a type-I error.
2. Calculate the noncentrality parameter  $\lambda$  using the formula:

$$\lambda = N \left( \frac{R_{T|C}^2}{1 - R_C^2 - R_{T|C}^2} \right)$$

3. Compute the power as the probability of being greater than  $F_{u,v,\alpha}$  in a noncentral- $F$  distribution with noncentrality parameter  $\lambda$ .

Note that the formula for  $\lambda$  is different from that used in **PASS 6.0**. The algorithm used in **PASS 6.0** was based on formula (9.3.1) in Cohen (1988) which gives approximate answers. This version of **PASS** uses an algorithm that gives exact answers.

## Multiple Regression

### Unconditional (Multivariate Normality)

In the unconditional model, the X's and Y have a joint multivariate normal distribution with a specified mean vector and covariance matrix given by

$$\begin{bmatrix} \sigma_Y^2 & \Sigma'_{YX} \\ \Sigma_{YX} & \Sigma_X \end{bmatrix}$$

The study-specific values of X are unknown at the design phase, so the sample size determination is based on a single, effect-size parameter which represents the expected variations in the X's, their interrelationships, and their relationship with Y. This effect-size parameter is the *squared multiple correlation coefficient* which is defined in terms of the covariance matrix as

$$\rho_{YX}^2 = \frac{\Sigma'_{YX} \Sigma_X^{-1} \Sigma_{YX}}{\sigma_Y^2}$$

If this coefficient is zero, the variables X provide no information about the linear prediction of Y and their corresponding regression coefficients are all zero. Note that we will use  $\rho^2$  to represent  $\rho_{YX}^2$ .

Often, the primary hypothesis involves testing the significance of a subset of X's that have been statistically adjusted for a second set of X's. The population parameter is then called the *squared multiple partial correlation coefficient*, which is interpreted similarly.

#### Test Statistic in the Unconditional Model

An F-test with  $M$  and  $N-M-1$  degrees of freedom can be constructed that will test whether all the regression coefficients simultaneously zero as follows

$$F_{M,N-M-1} = \frac{\rho^2/M}{(1-\rho^2)/(N-M-1)}$$

Suppose the independent variables are divided into two sets: C containing  $L$  variables and T containing the remaining  $K = M - L$  variables. That is, we partition  $X = X_T | X_C$ . It can be shown that an F-test that tests the significance of the T variables adjusted for the C variables is

$$F_{K,N-M-1} = \frac{(\rho_{YX_T|X_C}^2)/K}{(1-\rho_{YX_T|X_C}^2)/(N-M-1)}$$

Cohen (1988) shows that  $\rho_{YX_T|X_C}^2$  can be calculated from the  $R^2$  of fitting all the variables and the  $R^2$  of fitting just the set C variables as follows

$$\rho_{YX_T|X_C}^2 = \frac{R_{YX}^2 - R_{YX_C}^2}{1 - R_{YX_C}^2}$$

#### Calculating the Power in the Unconditional Method

In the unconditional method, the statistical hypotheses that is usually of most interest is the set

$$H_0: \rho^2 \leq \rho_0^2 \text{ versus } H_1: \rho^2 > \rho_0^2$$

because you want to establish a lower bound for the value, not just established that it is greater than zero. However, the hypotheses

$$H_0: \rho^2 \geq \rho_0^2 \text{ versus } H_1: \rho^2 < \rho_0^2$$

is also valid.

In the program, when  $\rho_1^2 > \rho_0^2$  the former hypothesis set is assumed. Otherwise, the later set is assumed.

## Multiple Regression

The calculation of the power of a particular test proceeds as follows:

1. Determine the critical value  $r_\alpha$  from the CDF such that  $P(R^2 \leq r_\alpha | N, K, \rho_0^2) = 1 - \alpha$ . Note that we use the value of  $\rho^2$  specified in the null hypothesis.
2. Compute the power using  $\text{Power} = 1 - P(R^2 \leq r_\alpha | N, K, \rho_1^2)$ .

Krishnamoorthy and Xia (2003) give the CDF of  $R^2$  as

$$P(R^2 \leq x | N, K, \rho^2) = \sum_{i=0}^{\infty} P(Y = i) I_x\left(\frac{K-1}{2} + i, \frac{N-K}{2}\right)$$

where

$$I_x(a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^x t^{a-1} (1-t)^{b-1} dt$$

$$P(Y = i) = \frac{\Gamma\left(\frac{N+1}{2} + i\right)}{\Gamma(i+1)\Gamma\left(\frac{N+1}{2}\right)} (\rho^2)^i (1-\rho^2)^{\frac{N+1}{2}}$$

This formulation does not admit  $\rho^2 = 0$ , so when this occurs, the program inserts  $\rho^2 = 0.000000000001$ .

Finally, when computing the squared multiple *partial* correlation coefficient, Gatsonis and Sampson (1989) indicate you simply need to replace  $N$  with  $N - L$  in the above CDF.

## Cohen's Effect Size

Cohen's (1988) measure of the effect size in multiple regression,  $f^2$  is

$$f^2 = \frac{R^2}{1 - R^2}$$

so that

$$R^2 = \frac{f^2}{1 + f^2}$$

When the independent variables are divided into the two sets as outlined above,  $f^2$  is

$$f^2 = \left( \frac{R_{YX_T|X_C}^2}{1 - R_{YX_T|X_C}^2} \right)$$

Cohen (1988) defined values near 0.02 as small, near 0.15 as medium, and above 0.35 as large. In terms of  $R^2$ , these are about 0.02, 0.13, and 0.26.

---

## 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

The Design tab contains most of the parameters and options that you will be concerned with.

---

#### Solve For

##### Solve For

This option specifies the parameter to be solved for from the other parameters. The parameters that may be selected are *Power*, *Sample Size*, and *Effect Size*. Under most situations, you will select either *Power* or *Sample Size*.

Select *Sample Size* 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.

---

#### Power Calculation

##### Power Calculation Method

Select the method to be used to calculate power. Both methods assume that the same F-test is used for significance testing during the analysis of the data. However, they differ in the calculation of power and the statistic used for input.

The conditional method is the most commonly used and so is the method we recommend unless you are certain that Y and X will follow the multivariate normal distribution.

- **Conditional (Recommended) - Uses  $R^2$**

This method uses  $R^2$ , the multiple correlation coefficient, for input.

The power calculations use the noncentral F distribution and are based on the assumption that the residuals are normally distributed with constant variance.

Regression analysis estimates the conditional expectation of Y given that the values of the X variables are fixed and known. This is the usual assumption made when performing significance tests in regression analysis.

The actual X values are not needed in the power calculations. Instead, the incremental  $R^2$  is all that is needed.

This method anticipates that the conditional assumption will be used at the time of the analysis, which is usually the case.

- **Unconditional (Assumes Multivariate Normality) - Uses  $\rho^2$**

This type uses  $\rho^2$ , the multiple partial correlation coefficient. This value is seldom shown on regression reports and generally must be computed by hand.

This method is based on a special algorithm that integrates the multivariate normal distribution.

This approach assumes that Y and X have a multivariate normal distribution and that the X values will not be known until the study is completed. This reduces the power because of the uncertainty in the X values.

## Multiple Regression

The method may sound appealing until one realizes that the assumption of multivariate normality is very restrictive and that the power calculations are not robust to departures from this assumption. For example, at least one X variable may have a non-normal distribution such as uniform, skewed right, or binary. Or, it exhibits clumping or outliers.

So, unless the multivariate normality assumption is reasonable and you will not be making the conditional expectation assumption when the data are analyzed, you should NOT use this option.

---

### 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 (alpha). A type-I error occurs when you reject the null hypothesis when in fact it is true.

Values of alpha 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

#### N (Sample Size)

This option specifies the value(s) for  $N$ , the sample size. Note that  $L + K < N - 1$ .

---

### Effect Size – T: Independent Variables Tested

These options refer to the independent variables that are being tested for statistical significance.

#### K (Number Tested)

This option specifies the number of X's in the set  $T$ , variables that are tested. This number must be greater than or equal to one. Note that  $K + L < N - 1$ .

#### $R^2(T|C) = R^2(T,C) - R^2(C)$

*Shown for Conditional Power Calculation Only*

Enter the amount that is added to  $R^2$  by adding the set  $T$  variables to the set  $C$  variables. This is appropriate for testing that the corresponding regression coefficients are all zero.

The values must be between zero and one and the total  $R^2(C) + R^2(T|C)$  must be less than one.

## Multiple Regression

### $\rho^2$ (Null)

*Shown for Unconditional Power Calculation Only*

Enter one or more values of  $\rho^2$ , the *squared multiple correlation coefficient* used in the null hypothesis. This is the proportion of the variation in  $Y$  explained by the variation in the  $X$ 's tested. Note that  $\rho^2$  is the population value of  $R^2$ .

If there are control  $X$ 's specified, this value becomes the squared multiple **partial** correlation coefficient.

If  $\rho^2 < \rho_1^2$ , the hypotheses tested is  $H_0: \rho^2 \leq \rho_1^2$  vs  $H_1: \rho^2 > \rho_1^2$ .

If  $\rho^2 > \rho_1^2$ , the hypotheses tested is  $H_0: \rho^2 \geq \rho_1^2$  vs  $H_1: \rho^2 < \rho_1^2$ .

This parameter represents a lower bound for  $R^2$  in that values of  $R^2$  below this are deemed unimportant.

The range is given by  $0 \leq \rho^2 < 1$ . Cohen interpreted these values as 0.02 = Small, 0.13 = Medium, 0.26 = Large.

### $\rho_1^2$ (Actual)

*Shown for Unconditional Power Calculation Only*

Enter one or more values of  $\rho_1^2$ , the squared multiple correlation coefficient at which the power is calculated. This is the proportion of the variation in  $Y$  explained by the variation in the  $X$ 's tested.  $\rho^2$  is the population value of  $R^2$ .

If there are control  $X$ 's, this is the squared multiple *partial* correlation coefficient.

If  $\rho^2 < \rho_1^2$ , the hypotheses tested is  $H_0: \rho^2 \leq \rho_1^2$  vs  $H_1: \rho^2 > \rho_1^2$ .

If  $\rho^2 > \rho_1^2$ , the hypotheses tested is  $H_0: \rho^2 \geq \rho_1^2$  vs  $H_1: \rho^2 < \rho_1^2$ .

The range is given by  $0 < \rho_1^2 < 1$ . Cohen interpreted these values as 0.02 = Small, 0.13 = Medium, 0.26 = Large.

---

## Effect Size – C: Covariates (Independent Variables Controlled For)

These options refer to the independent variables that are controlled for.

### L (Number of Covariates)

This option specifies the number of variables in set  $C$ , variables that are controlled for (or partialled out). This number must be greater than or equal to zero. Note that  $K + L < N - 1$ .

### Models with No Covariates (Independent Variables Controlled For)

If there are no covariates to be controlled for in your model, enter "0" here. In this case, the input value below for  $R^2(C)$  will be ignored and  $R^2(C)$  will be set to zero.

### $R^2(C)$

*Shown for Conditional Power Calculation Only*

This option is only shown for the conditional model. It specifies the  $R^2$  achieved by the variables in set  $C$  when they are fit alone in the regression equation. Note that this amount must be between zero and one and that the total of the two  $R^2$  values must be less than one. This value is only used if  $L > 0$ . When  $L = 0$ ,  $R^2(C) = 0$ .

## Multiple Regression

## Example 1 – Finding Sample Size in the Conditional Model

Suppose researchers are planning a multiple regression study to look at the significance of a specific independent variable. They want to assume that a conditional model will be used to analyze the data.

The data will come from a survey that includes four other continuous demographic variables.

They want a sample size large enough to detect an  $R^2$  of between 0.1 and 0.4 in one of the four variables. They assume that the other three variables will account for an  $R^2$  of 0.3.

They want to consider power values of either 0.8 or 0.9 and a significance level is 0.05.

### Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Multiple Regression** procedure. 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>
<b>Design Tab</b>	
Solve For .....	<b>Sample Size</b>
Power Calculation Method.....	<b>Conditional (Recommended) - Uses <math>R^2</math></b>
Power.....	<b>0.8 0.9</b>
Alpha.....	<b>0.05</b>
K (Number Tested) .....	<b>1</b>
$R^2(T C) = R^2(T,C) - R^2(C)$ .....	<b>0.1 0.2 0.3 0.4</b>
L (Number of Covariates) .....	<b>4</b>
$R^2(C)$ .....	<b>0.3</b>

### Annotated Output

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

### Numeric Results

#### Numeric Results

Power Method: Conditional (Recommended)

Power	N	Number	Number	$R^2$ of	$R^2$ of	Alpha	Beta
		Covariates	Tested	Covariates	Tested		
		L	K	$R^2(C)$	$R^2(T C)$		
0.8060	50	4	1	0.300	0.100	0.050	0.1940
0.8155	23	4	1	0.300	0.200	0.050	0.1845
0.8094	14	4	1	0.300	0.300	0.050	0.1906
0.8605	11	4	1	0.300	0.400	0.050	0.1395
0.9037	66	4	1	0.300	0.100	0.050	0.0963
0.9033	29	4	1	0.300	0.200	0.050	0.0967
0.9007	17	4	1	0.300	0.300	0.050	0.0993
0.9118	12	4	1	0.300	0.400	0.050	0.0882

#### References

- Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences*, Lawrence Erlbaum Associates, Hillsdale, New Jersey.
- Gatsonis, C. and Sampson, A.R. 1989. 'Multiple Correlation: Exact Power and Sample Size Calculations.' *Psychological Bulletin*, Vol. 106, No. 3, Pages 516-524.



## Multiple Regression

### Report Definitions

Power is the probability of rejecting a false null hypothesis.

N is the number of observations on which the multiple regression is computed.

L is the number of covariates (independent variables not tested for zero regression coefficients).

K is the number of independent variables tested for zero coefficients.

$R^2(C)$  is the  $R^2$  achieved when only the control variables are included in the model.

$R^2(T|C)$  is the amount that  $R^2$  is increased when the test variables are added to a model that contains the control variables.  $R^2(T|C) = R^2(T,C) - R^2(C)$ .

Alpha is the probability of rejecting a true null hypothesis. It should be small.

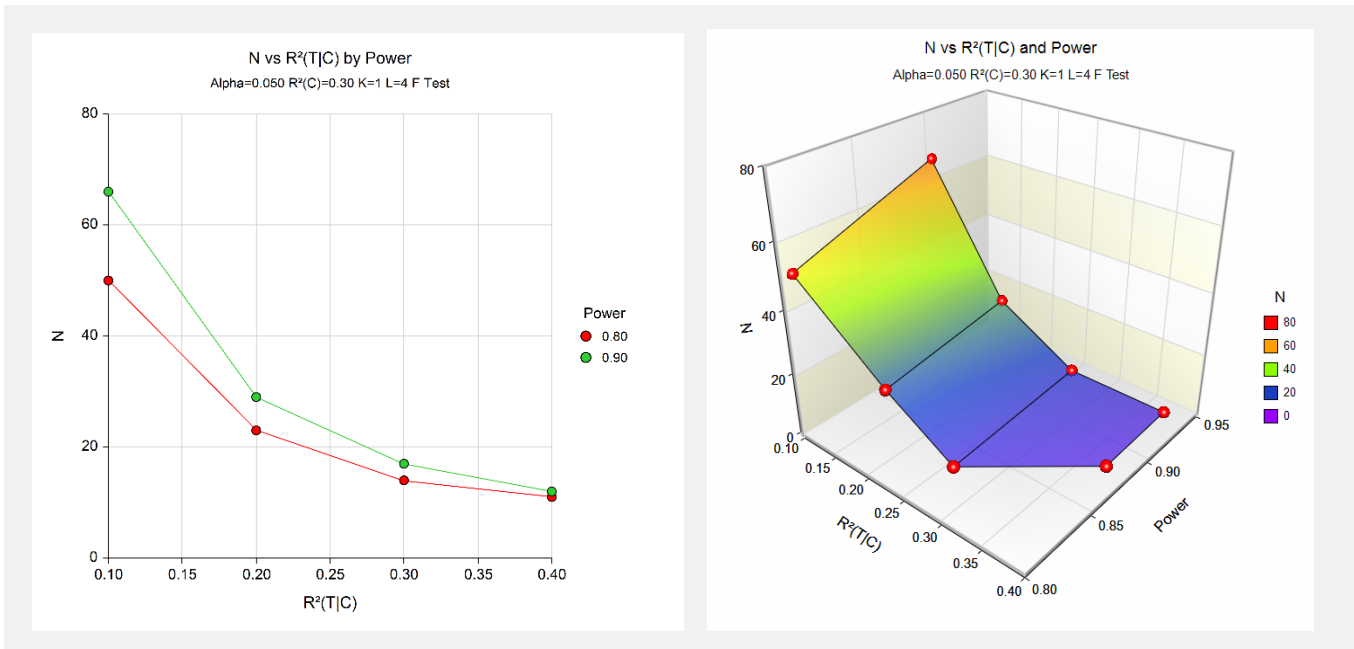
Beta is the probability of accepting a false null hypothesis. It should be small.

### Summary Statements

In a study using the conditional power calculation method, a sample size of 50 achieves 81% power to detect an  $R^2$  of 0.100 attributed to 1 independent variable(s) using an F-Test with a significance level (alpha) of 0.050. The variables tested are adjusted for an additional 4 covariate(s) which have a combined  $R^2$  of 0.300 by themselves.

This report shows the necessary sample sizes. The definitions of each of the columns is given in the Report Definitions section.

## Plots Section



These plots show the relationship between sample size, effect size, and power.

## Multiple Regression

## Example 2 – Validation using the Unconditional Power Calculation

We will validate this procedure using an analysis published in Shieh and Kung (2007). In this example, the desired power is 0.90, alpha is 0.05, L is 0, K is 5,  $\rho_0^2$  is 0.2, and  $\rho_1^2$  is 0.05. They calculate a sample size of 153.

### Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Multiple Regression** procedure. 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>
<b>Design Tab</b>	
Solve For .....	<b>Sample Size</b>
Power Calculation Method.....	<b>Unconditional (Assumes Multivariate Normality) - Uses <math>\rho^2</math></b>
Power.....	<b>0.90</b>
Alpha.....	<b>0.05</b>
K (Number Tested) .....	<b>5</b>
$\rho_0^2$ (Null).....	<b>0.2</b>
$\rho_1^2$ (Actual) .....	<b>0.05</b>
L (Number of Covariates) .....	<b>0</b>

### Output

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

### Numeric Results

Numeric Results								
Power Method: Unconditional (Assumes Multivariate Normality)								
Power	N	Number Covariates L	Number Tested K	Squared Multiple Correlation Coefficient under H0 $\rho_0^2$	Squared Multiple Correlation Coefficient under H1 $\rho_1^2$	Alpha	Beta	
0.9011	153	0	5	0.200	0.050	0.050	0.0989	

PASS has also calculated the required sample size to be 153.

## Multiple Regression

## Example 3 – Testing the Addition or Deletion of a Single Variable in a Conditional Power Calculation

This example calculates the power of an  $F$  test constructed to test a fifth variable which adds 0.05 to  $R^2$  after considering four other variables whose combined  $R^2$  value is 0.5. Sample sizes from 10 to 150 will be investigated. The significance level is 0.05.

### Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Multiple Regression** procedure window. 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>
<b>Design Tab</b>	
Solve For .....	Power
Power Calculation Method.....	Conditional (Recommended) - Uses $R^2$
Alpha.....	0.05
N (Sample Size).....	10 to 150 by 20
K (Number Tested) .....	1
$R^2(T C) = R^2(T,C) - R^2(C)$ .....	0.05
L (Number of Covariates) .....	4
$R^2(C)$ .....	0.5

### Output

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

### Numeric Results

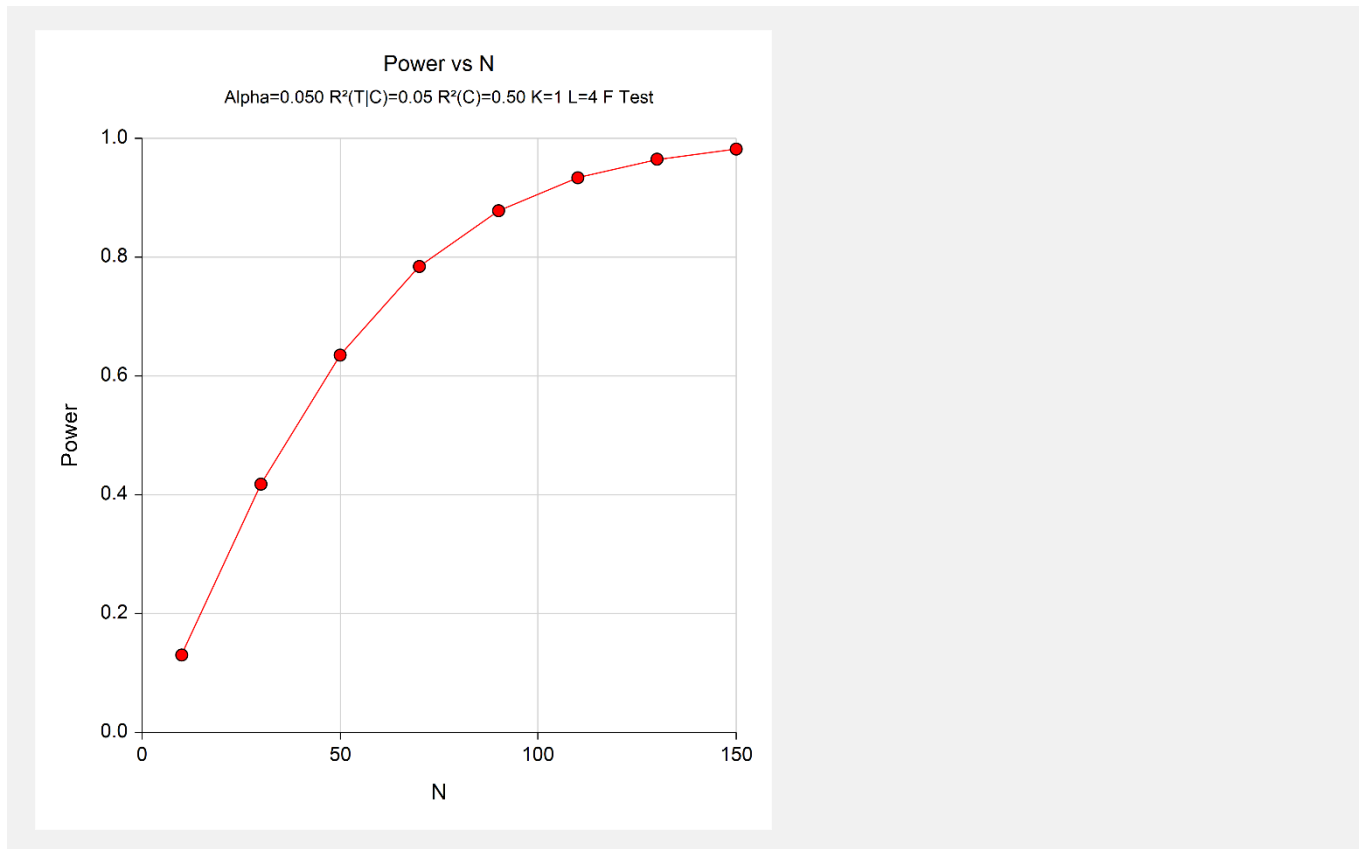
Numeric Results								
Power Method: Conditional (Recommended)								
Power	Number Covariates		Number Tested		R <sup>2</sup> of Covariates	R <sup>2</sup> of Tested	Alpha	Beta
	N	L	K		R <sup>2</sup> (C)	R <sup>2</sup> (T C)		
0.1304	10	4	1		0.500	0.050	0.050	0.8696
0.4180	30	4	1		0.500	0.050	0.050	0.5820
0.6351	50	4	1		0.500	0.050	0.050	0.3649
0.7843	70	4	1		0.500	0.050	0.050	0.2157
0.8782	90	4	1		0.500	0.050	0.050	0.1218
0.9337	110	4	1		0.500	0.050	0.050	0.0663
0.9649	130	4	1		0.500	0.050	0.050	0.0351
0.9819	150	4	1		0.500	0.050	0.050	0.0181

This report shows the values of each of the parameters, one scenario per row. The definitions of each of the columns is given in the Report Definitions section.

Note that in this particular example, a power of 0.90 is not reached until the sample size is about 110.

## Multiple Regression

## Plots Section



This plot shows the relationship between sample size and power.

## Multiple Regression

**Example 4 – Minimum Detectable  $R^2$** 

Suppose the researchers in Example 3 can only afford a sample size of 30. They want to know the minimum detectable  $R^2$  that can be detected if the power is 80% and 90%.

**Setup**

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Multiple Regression** procedure window. 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>
<b>Design Tab</b>	
Solve For .....	<b>Effect Size (<math>\rho^2</math> or <math>R^2(T C)</math>)</b>
Power Calculation Method.....	<b>Conditional (Recommended) - Uses <math>R^2</math></b>
Power.....	<b>0.8 0.9</b>
Alpha.....	<b>0.05</b>
N (Sample Size).....	<b>30</b>
K (Number Tested) .....	<b>1</b>
L (Number of Covariates) .....	<b>4</b>
$R^2(C)$ .....	<b>0.5</b>

**Output**

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

**Numeric Results**

<b>Numeric Results</b>								
Power Method: Conditional (Recommended)								
	Number Covariates		Number Tested	$R^2$ of Covariates	$R^2$ of Tested	Alpha	Beta	
Power	N	L	K	$R^2(C)$	$R^2(T C)$			
0.8000	30	4	1	0.500	0.111	0.050		0.2000
0.9000	30	4	1	0.500	0.138	0.050		0.1000

This report shows that at 90% power, a sample size of 30 cannot detect an  $R^2(T|C)$  less than 0.138.

## Multiple Regression

## Example 5 – Validation using a Conditional Power Calculation

Ralph O'Brien, in a private communication to Jerry Hintze, gave the result that when  $\text{Alpha} = 0.05$ ,  $N = 15$ ,  $K = 2$ , and  $R^2(T|C) = 0.6$ , the power is 0.9683.

### Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Multiple Regression** procedure window. You may then make the appropriate entries as listed below, or open **Example 5** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>Power</b>
Power Calculation Method.....	<b>Conditional (Recommended) - Uses R<sup>2</sup></b>
Alpha.....	<b>0.05</b>
N (Sample Size).....	<b>15</b>
K (Number Tested) .....	<b>2</b>
$R^2(T C) = R^2(T,C) - R^2(C)$ .....	<b>0.6</b>
L (Number of Covariates) .....	<b>0</b>

### Output

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

### Numeric Results

Numeric Results							
Power Method: Conditional (Recommended)							
	Number Covariates		Number Tested	R <sup>2</sup> of Covariates	R <sup>2</sup> of Tested	Alpha	Beta
Power	N	L	K	R <sup>2</sup> (C)	R <sup>2</sup> (T C)		
0.9683	15	0	2	0.000	0.600	0.050	0.0317

The power of 0.9683 matches O'Brien's result.