

Power for Professionals

When we examined the confidence interval we saw that it told us where the parameter would lie in repeated sampling. The smaller the margin of error the tighter or more sensitive the interval. A fixed level α test (two sided) can be carried out with a confidence interval (see notes). We want our test to detect H_0 (null hypothesis) that are false. This is measured by the probability the test will reject H_0 when the alternative (H_A) is true - the higher the probability the better.

Power → The probability that a fixed level alpha- α significance test will reject H_0 when a particular alternative value of the parameter is true is called the **power** of the test against the alternative.

Along with 95% confidence interval, 5% significance tests, 80% power is very popular. Many granting agencies and industries require the sample size be sufficient to detect results 80% of the time using a 5% test of significance.

Boost the Power

Say you perform a power calculation (see SPSS notes) and find the power too small, what can you do to increase it?

- 1) Increase alpha- α . A 10% test will have a greater chance of rejecting the alternative than a 5% test because the strength of evidence required for rejection is less.
- 2) Increase the sample size. As n increases our standard error decreases (see notes).
- 3) Consider an alternative H_A parameter that is further away from the value under H_0 . Values that are close to one another are harder to detect (low power).

Power is important for planning. Using a test with low power makes it unlikely you will find a significant effect. Suppose the H_0 is in fact false, it can be widely believed to be true if repeated attempts fail because of low power.

Error Types for Decision Making

If we reject H_0 (accept H_A) when in fact H_0 is true, call this **Type I error**.

If we accept H_0 (reject H_A) when in fact H_A is true, call this **Type II error**.

		Truth about Population	
		H ₀ true	H _A true
Decision Based on Sample	Reject H ₀	Type I error	Correct decision
	Accept H ₀	Correct decision	Type II error

Significance tests with a fixed α give us a way of understanding the above table because the test either rejects H₀ or fails to reject H₀. Failing to reject H₀ means deciding H₀ is true. So now we can describe our test using Type I and Type II error.

Significance and Type I error

The significance level α of any fixed level test is the probability of a Type I error. Alpha- α is the probability the test will reject H₀ when H₀ is true.

Power and Type II error

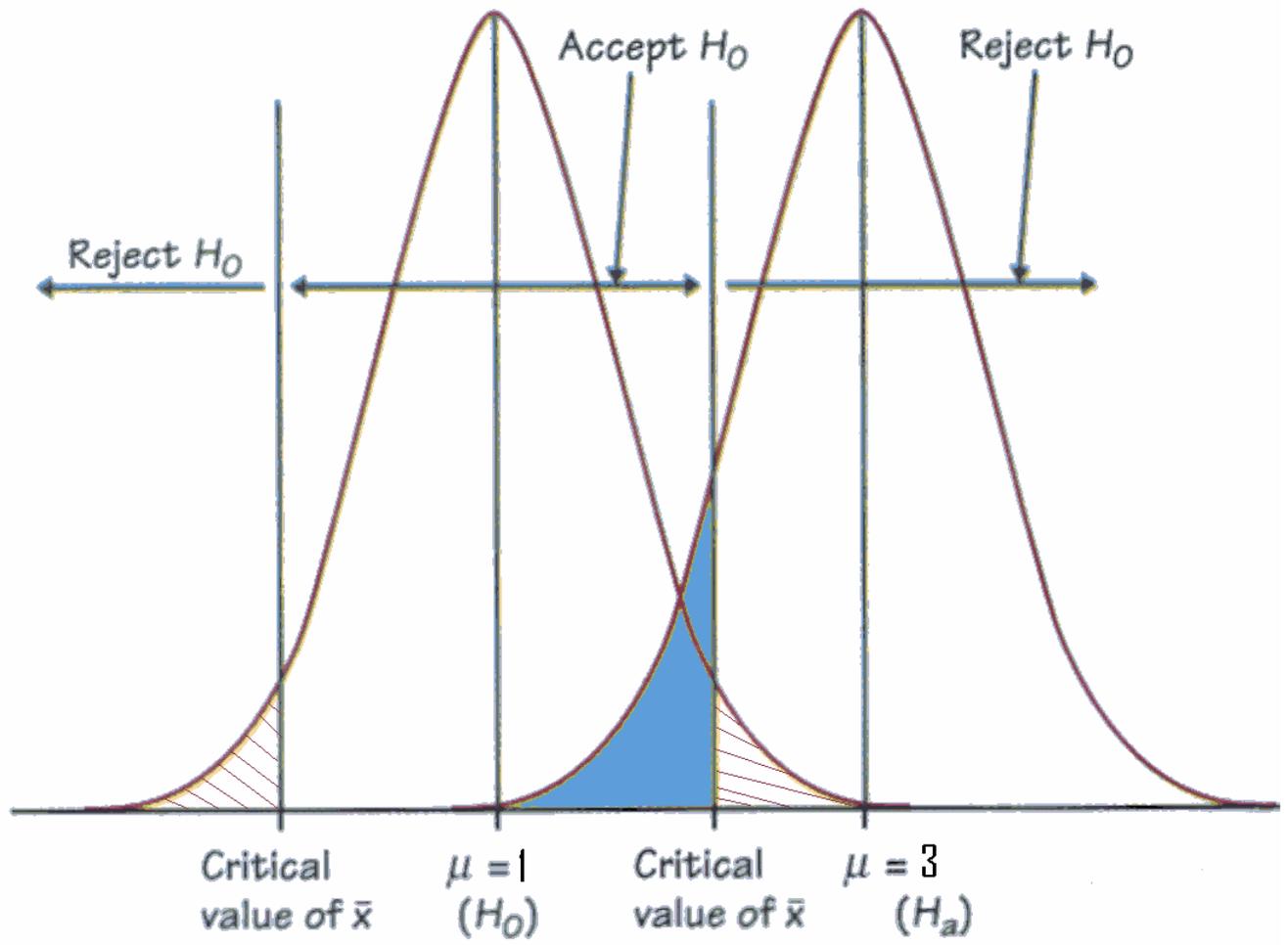
The power of a fixed level test against a particular alternative is 1 minus the probability of a Type II error for that alternative.

Suppose we calculate the power of a test to be .87. The probability of a Type II error is $1 - .87 = .13$.

Significance Tests

In the research setting H_A is the research hypothesis asserting an effect is present. The null hypothesis says there is no difference. Low p-values present evidence that H_A is true. If H₀ is something that everyone in the profession believes then strong evidence (small p-value) will be needed to reject it. The level of evidence needed will depend on the consequences. If rejecting H₀ means changes are expensive, strong evidence will be needed. A .05 test is very common, but there is no dividing line that is absolute, one must realize that evidence increases as the p-value decreases. Note that when large samples are present then tiny deviations will be declared significant, they may not be practically important. Remember; plan to have an adequate sample size to detect an effect that exists – POWER.

Two Sample t-test Example



 Probability of Type I error

 Probability of Type II error

 Probability of rejecting H_0 : $\mu = 1$ when in fact $\mu = 1$.

 Probability of accepting H_0 when in fact $\mu = 3$.

One Way ANOVA

Remember that the power of the test is the probability of rejecting H_0 when in fact H_A is true. We measure power with respect to a specific alternative. We do a power calculation to make sure the sample sizes are adequate to detect important differences in means. We may find that the power is low against the alternative and there is little chance of detecting a difference (significant F). Power helps us interpret results where H_0 was not rejected. Power calculations for the ANOVA are similar to the two sample t-test since the ANOVA is a generalization of the t-test (see notes).

We have to calculate the probability of rejecting H_0 when H_A is true, the probability that the observed F is greater than F-critical for the test α . We calculate the **non-central F distribution**. To do this we calculate a **noncentrality parameter** (λ). If all the means are equal $\lambda=0$. The larger the differences in means, the larger λ (it tells us how unequal the means are) and consequently we expect high power. We usually repeat the power calculation for different sample sizes to find acceptable levels of power (usually .8 or better).

Summary

In a test of significance we focus on H_0 and a single probability (p-value). The purpose is to measure the evidence against H_0 . Calculations of power are made to check the sensitivity of the test. If we don't reject H_0 we conclude there is not enough evidence against H_0 (not that H_0 is true). If the same problem is thought of as a decision that must be made we focus on both H_A and H_0 . We focus on both probabilities and errors. We must choose H_0 or H_A and cannot claim there is insufficient evidence. We commonly mix the reasoning of *significance tests* and *decision rules*.

Sample Power Software

When you first open the sample power program you will click on File, then new analysis and you will see the following screen. Remember that you can always use 'help' to guide you through the analysis (see yellow box) for interactive guide.

Means:

In order to test the power for a 2 sample t-test:

- 1) Click the Means tab
- 2) Click the t-test for 2 (independent) groups with common variance (enter means)
- 3) Click OK.

Means | Proportions | Correlations | ANOVA | Regression | General

- One sample t-test that mean = 0
- One sample t-test that mean = specific value
- Paired t-test that mean difference = 0
- Paired t-test that difference = specific value
- t-test for 2 (independent) groups with common variance (Enter means)
- t-test for 2 (independent) groups with common variance (Enter difference)

Cancel
Default
Example
Ok

t-test for 2 (independent) groups with common variance (Enter means)

A researcher will be assigning phobia patients to either a standard treatment or an experimental program, and measuring their anxiety level after six weeks.

The researcher elects to specify the effect size by providing the expected mean for each of the treatment groups.

Close

The Following screen will then appear:

The screenshot displays the 'SamplePower' application window titled '[t-test for two independent samples with common variance]'. The interface includes a menu bar (File, Options, Tools, View, Help) and a toolbar with various icons. The main area contains a table for inputting population parameters and a power slider.

	Population Mean	Standard Deviation	N Per Group	Standard Error	95% Lower	95% Upper
Population 1	0.0	1.0	10			
Population 2	0.0	1.0	10			
Mean Difference	0.0	1.0	20	0.45	-0.92	0.92

Alpha= 0.05, Tails= 2

Power: 5%

Welcome

This interactive guide will lead you through the steps for computing power and precision.

To move this box use the title bar above.

To close or reactivate this panel, select Help from the menu.

Buttons: Help, < Back, Next >

Windows Taskbar: start, Windows Media Player, Sample Power Hando..., SamplePower - [t-tes..., 7:02 PM

Next, enter the information from our t-test example (taken from Dr. Gebotys' , Chapter 7, PS 292 notes):

- 1) Population Mean 1 = 41.5
- 2) Population Mean 2 = 51.4
- 3) Standard Deviation 1 = 17.1
- 4) Standard Deviation 2 = 11.0
- 5) N per Group 1 = 23
- 6) N per Group 2 = 21

SamplePower - [t-test for two independent samples with common variance]

File Options Tools View Help



Find N for power of 80%

	Population Mean	Standard Deviation	N Per Group	Standard Error	95% Lower	95% Upper
Population 1	41.5	17.1	23			
Population 2	51.4	11.0	21			
Mean Difference	-9.9	14.5	44	4.38	1.13	18.67

Alpha= 0.05, Tails= 2

Power

Welcome

This interactive guide will lead you through the steps for computing power and precision.

To move this box use the title bar above.

To close or reactivate this panel, select Help from the menu.

Help

< Back

Next >



Windows Media Player

Sample Power Hando...

SamplePower - [t-tes...



7:04 PM

You will now notice that the blue power bar in the bottom right corner indicates that the power for this example is 60%.

In order to see what number of participants would be required in order to increase the power we have two options. We can click the binocular icon in the top tool bar in order to see what number of participants we would need to reach a power of 80% (this tells us that we would need 35 participants in each group). Or, alternatively we can click the graph icon in the top tool bar (the one with bars) in order to create a chart of the power for our example. We also have the option of clicking the table icon in order to obtain a table of the power for our example.

Our Table will look like the following:

SamplePower - [Table - t-test for two independent samples with common variance]

File Modify Graph View Help

Make table

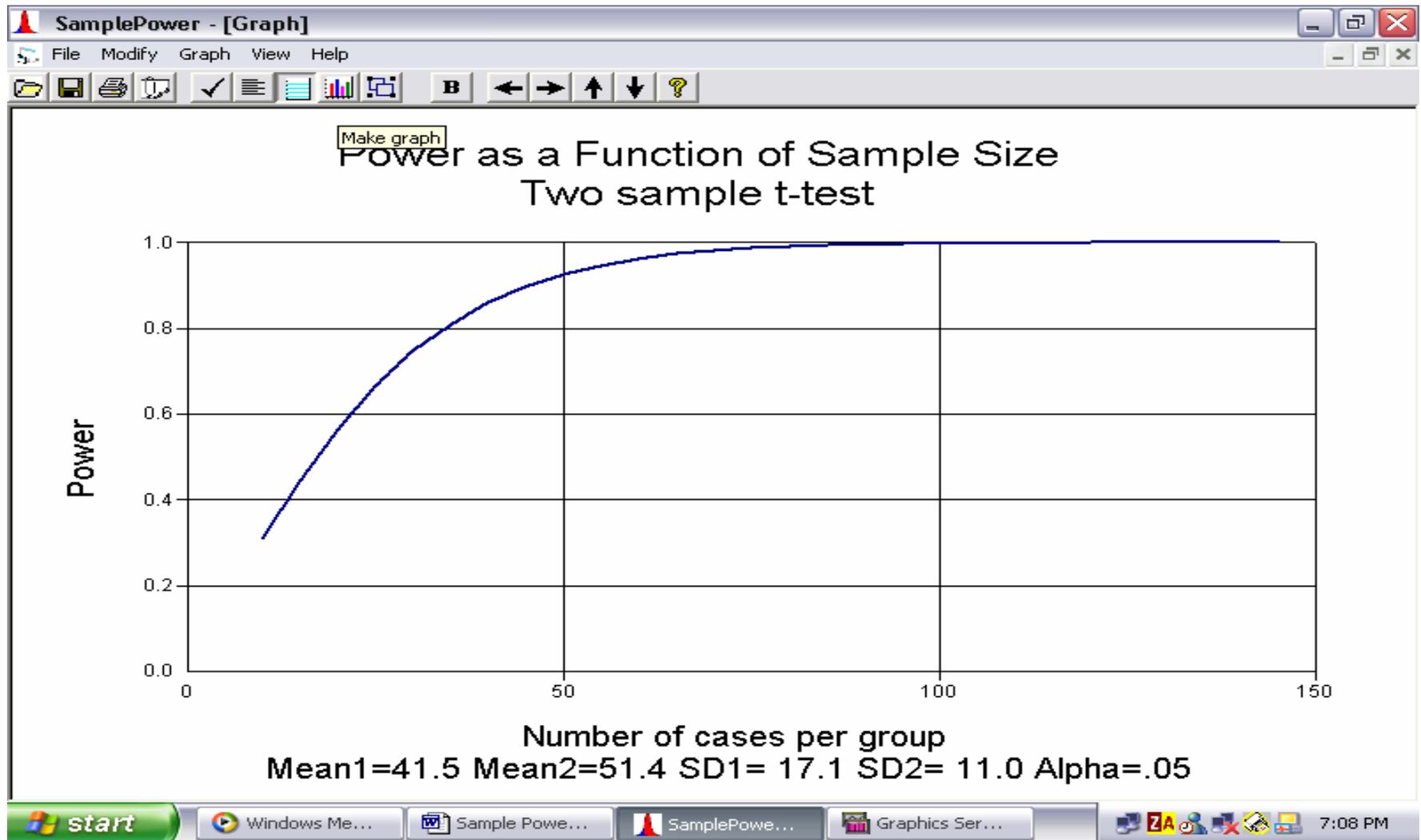
Power and Precision as a Function of Sample Size
Two sample test of means

N (1)	N (2)	Power	Raw Difference = -9.9			Std Difference = -0.689		
			Standard Error	95% Lower	95% Upper	Standard Error	95% Lower	95% Upper
10	10	.308	6.430	-3.358	23.158	.447	-.234	1.611
15	15	.445	5.250	-.726	20.526	.365	-.050	1.428
20	20	.564	4.546	.777	19.023	.316	.054	1.323
25	25	.665	4.066	1.781	18.019	.283	.124	1.253
30	30	.746	3.712	2.512	17.288	.258	.175	1.202
35	35	.810	3.437	3.076	16.724	.239	.214	1.163
40	40	.860	3.215	3.527	16.273	.224	.245	1.132
45	45	.898	3.031	3.899	15.901	.211	.271	1.106
50	50	.926	2.875	4.213	15.587	.200	.293	1.084
55	55	.947	2.742	4.482	15.318	.191	.312	1.065
60	60	.963	2.625	4.717	15.083	.183	.328	1.045
65	65	.974	2.522	4.923	14.877	.175	.342	1.035
70	70	.981	2.430	5.106	14.694	.169	.355	1.022
75	75	.987	2.348	5.271	14.529	.163	.367	1.011
80	80	.991	2.273	5.420	14.380	.158	.377	1.000
85	85	.994	2.205	5.555	14.245	.153	.386	.991
90	90	.996	2.143	5.679	14.121	.149	.395	.982
95	95	.997	2.086	5.792	14.008	.145	.403	.974

Two sample t: Mean1 = 41.5 Mean2 = 51.4 SD1 = 17.1 SD2 = 11.0 Alpha=.05 Tails=2
Computational option: Variance is estimated (t-test)

start Windows Media Player Sample Power Hando... SamplePower - [Tabl... 7:07 PM

Our chart will look like the following:



ANOVA:

In order to conduct an analysis for a one-way ANOVA

- 1) Click the ANOVA tab.
- 2) Select Oneway analysis of variance.
- 3) Click OK.

Means | Proportions | Correlations | **ANOVA** | Regression | General

- Oneway analysis of variance
- Oneway analysis of covariance
- Factorial analysis of variance (2 factors)
- Factorial analysis of covariance (2 factors)
- Factorial analysis of variance (3 factors)
- Factorial analysis of covariance (3 factors)

Cancel
Default
Example
Ok

Oneway analysis of variance

A researcher is planning to draw samples of students in five school districts, have them take the PSAT, and compare the mean scores.

Her hypothesis is that the mean score in the population is not identical across the five districts.

Close

Using the information from Dr. Gebotys' One Way ANOVA handout (Chapter 14 of notes) we can run a power analysis for the example with three levels of crime, with three observations per treatment.

SamplePower - [ANOVA]

File Options Tools View Help

Interactive Guide

Factor Name	Number of levels	Cases per level	Effect size f	Power
Factor A	2	5	0.00	0.05

SD within cell: 1.00 N of cases per cell: 5

Variance within cell: 1.00 Total N: 10

Alpha= 0.05

Welcome

This interactive guide will lead you through the steps for computing power and precision.

To move this box use the title bar above.

To close or reactivate this panel, select Help from the menu.

Help < Back Next >

start Windows Media Player Sample Power Hando... SamplePower - [ANO...]

7:12 PM

- 1) First enter the variance within cell on the left hand side, 1.81 or MSE.
- 2) Next change the N of cases per cell to 3.
- 3) Change the Label for Factor A to 'Crime.'
- 4) Now you will notice that the number of levels needs to be changed. You can fix this by clicking on the number shown under the number of levels column. This will bring up the following screen.

SamplePower - [ANOVA]

File Options Tools View Help

Effect size for Crime

Enter f Enter SD Enter Range **Enter Means**

1	Break & Enter	38.50
2	Robbery	40.30
3	Manslaughter	41.83

Number of categories in this factor

3

Compute f

Welcome

This interactive guide will lead you through the steps for computing power and precision.

To move this box use the title bar above.

To close or reactivate this panel, select Help from the menu.

start Windows Me... Sample Powe... SamplePowe... http://www.... 7:19 PM

- 5) The first thing you will do is change the number of categories in this factor to 3.
- 6) Then you are able to choose one of 4 ways to calculate the effect size: enter f, SD, Range, or Means. Here we will enter the means, which are 38.5, 40.3, and 41.83.
- 7) Now, click the yellow compute f button.
- 8) Click the blue register f button.
- 9) You will now be back out at the main screen. Here you can use the binocular icon in the top toolbar to calculate the number of participants needed per group to reach a power of .8 (happens to be 4 people per group). Or you can use the graph icon to create the following graph of power. Or click the table icon for a table of values.

The table for this example will look like the following:

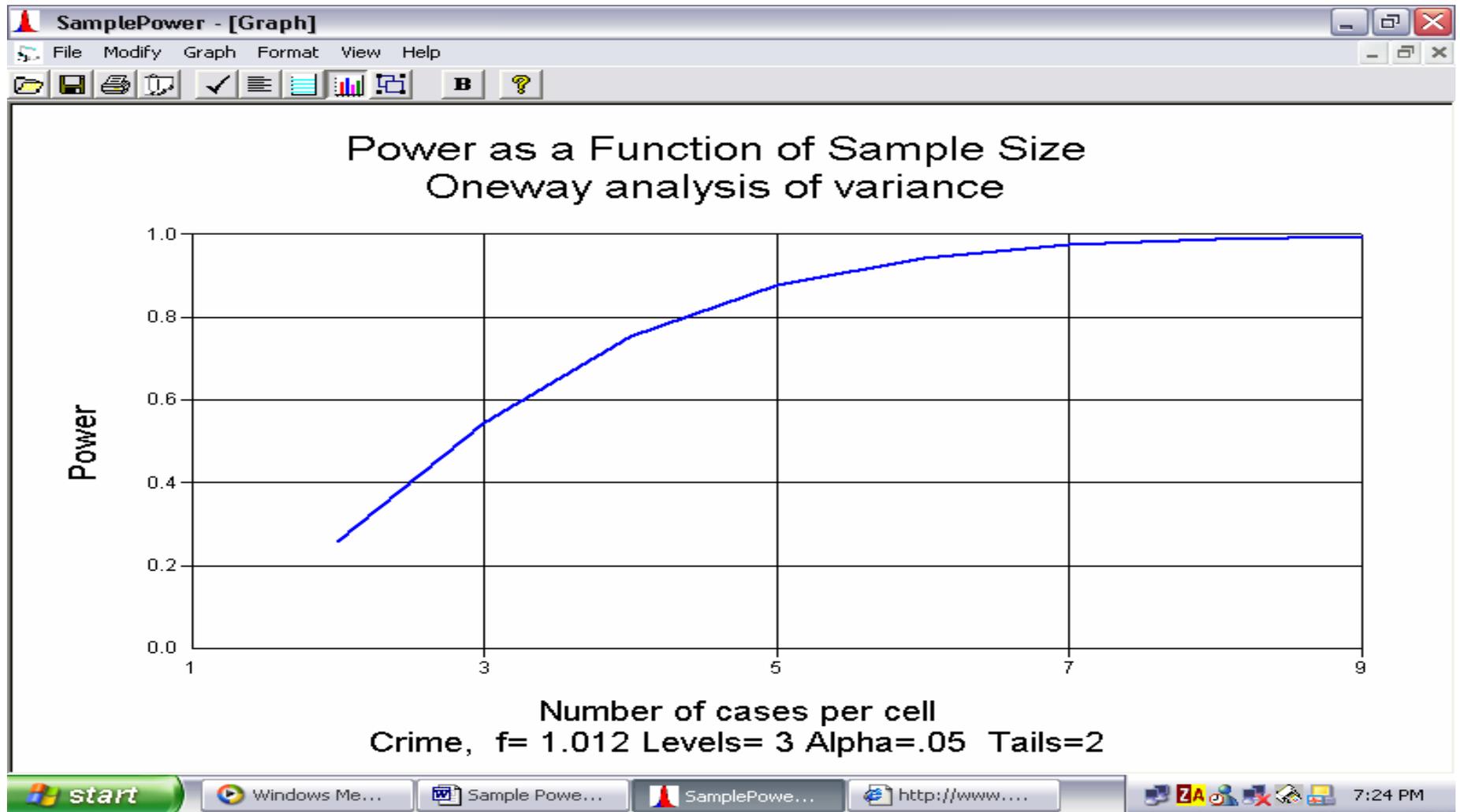
The screenshot shows the 'SamplePower - [Table - ANOVA]' window. The main content area displays a table titled 'Power as a Function of Sample Size For Analysis of Variance'. The table has two columns: 'N per cell' and 'Power'. The data points are as follows:

N per cell	Power
2	.257
3	.544
4	.753
5	.877
6	.943
7	.975
8	.989
9	.996
10	.998

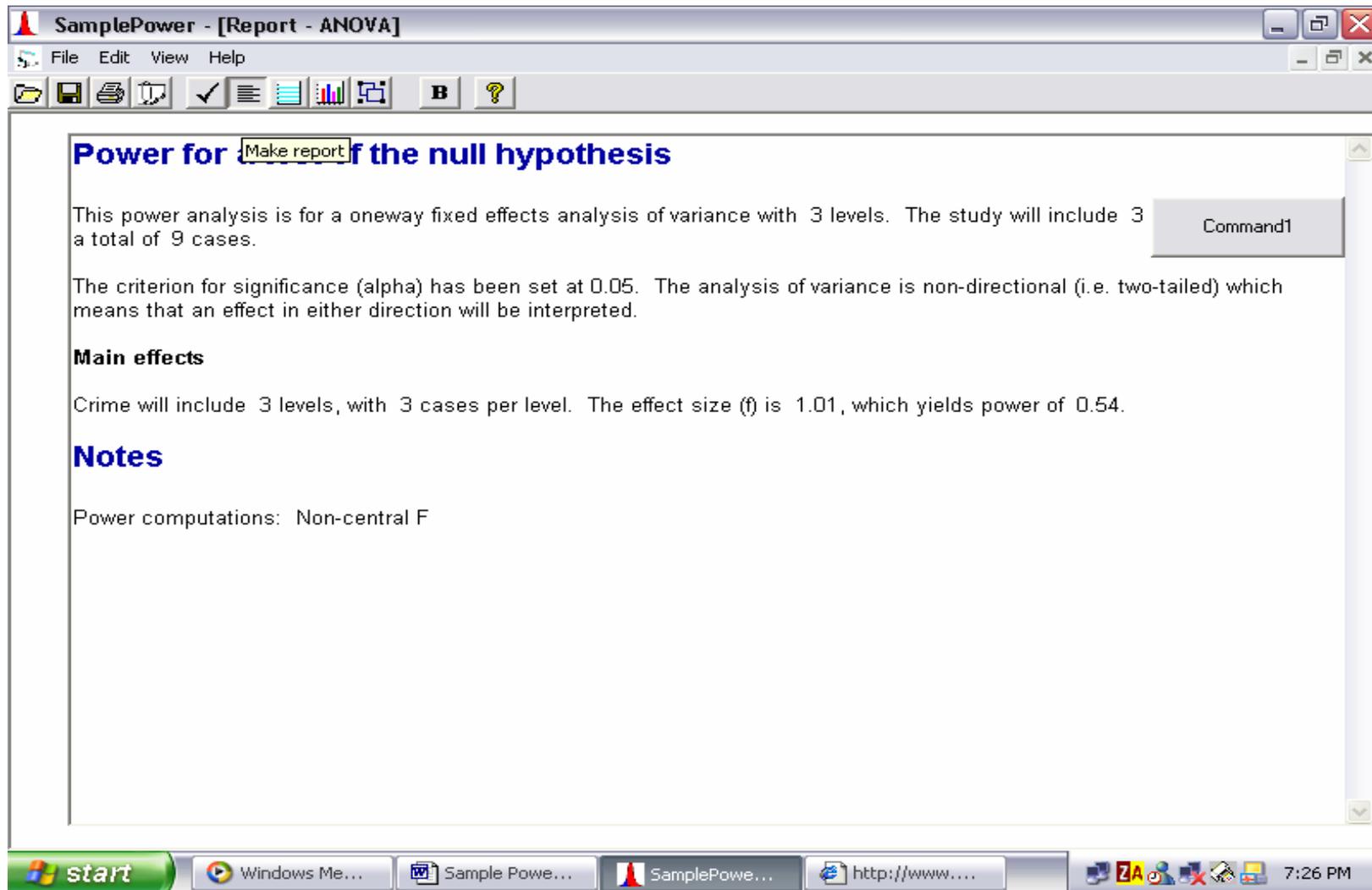
Below the table, the following parameters are listed: Crime, $f = 1.012$ Levels= 3 Alpha=.05 Tails=2
Power computations: Non-central F

The Windows taskbar at the bottom shows the Start button, several open windows (Windows Me..., Sample Powe..., SamplePowe...), and a web browser window (http://www...). The system clock indicates 7:23 PM.

The chart for this example will look like the following:



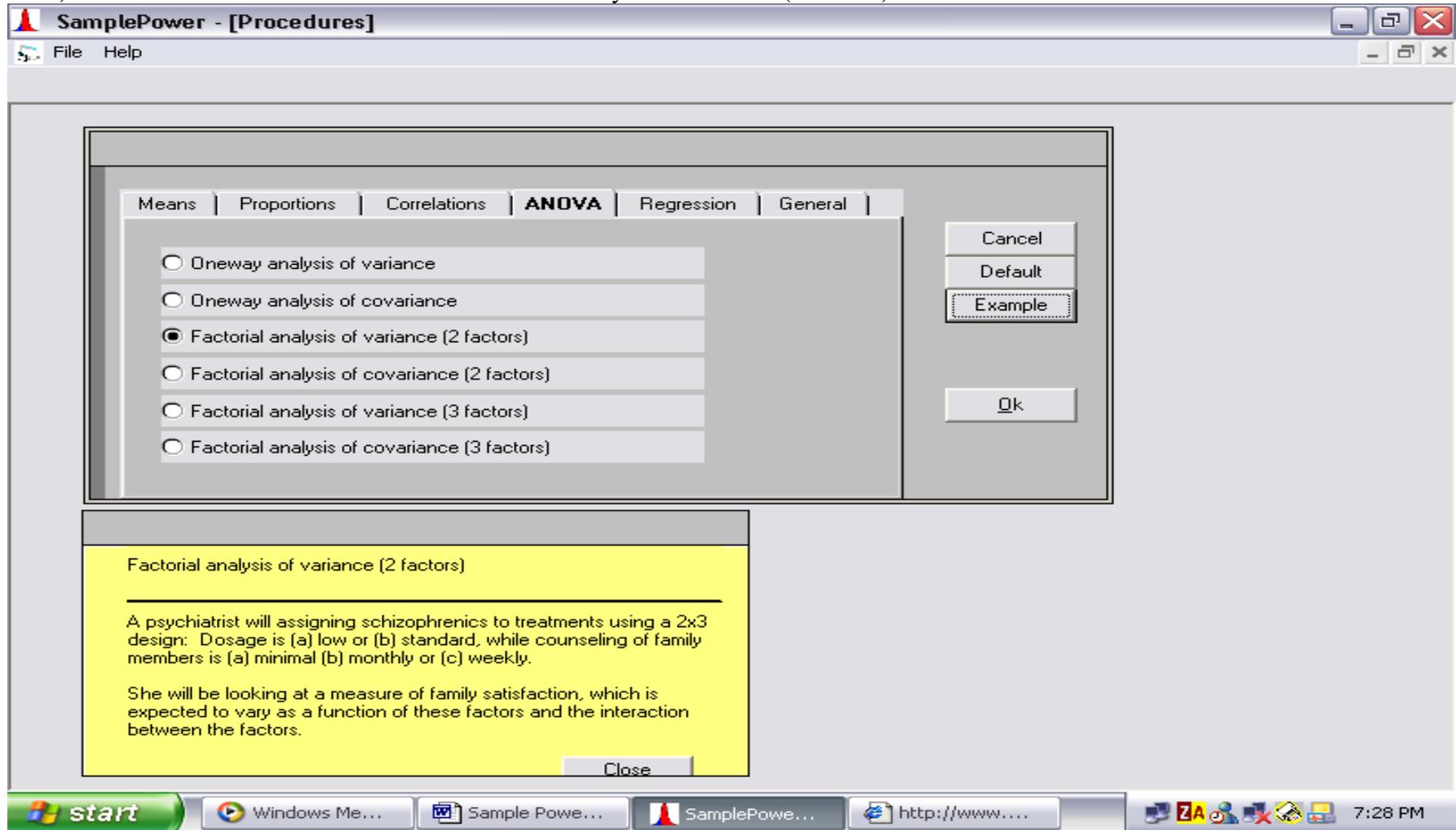
You can also create a report by clicking the 'Make Report' icon:



Two-Way ANOVA:

Now if we run through the example from the two-way ANOVA handout (Chapter 16 notes) we will do the following.

- 1) Click on the ANOVA tab and select Factorial analysis of variance (2 factors).



- 2) Click OK.
- 3) Change the N cases per cell to 2.
- 4) Enter the variance within cell, assume $42.29 = \text{MSE}$.
- 5) Change the name of factor A to 'Crime.'
- 6) Click on the number of cases factor A.
- 7) Change the number to 3.
- 8) Enter the necessary information. Means, 46.87, 59.25, and 78.25.
- 9) Click the yellow compute f button. Click the blue register f button.
- 10) Now back at the main screen change the name of factor B to 'Age' and click on the number of cases factor B.
- 11) Change the number to 4.
- 12) Enter information into the means tab. Means, 53, 67.83, 51.16, and 73.83.
- 13) Click the yellow compute f button. Click the blue register f button.
- 14) Lastly, for the interaction click the medium effect size, .25.

Now if we click the table icon on the top toolbar we will get the following table.

Make table

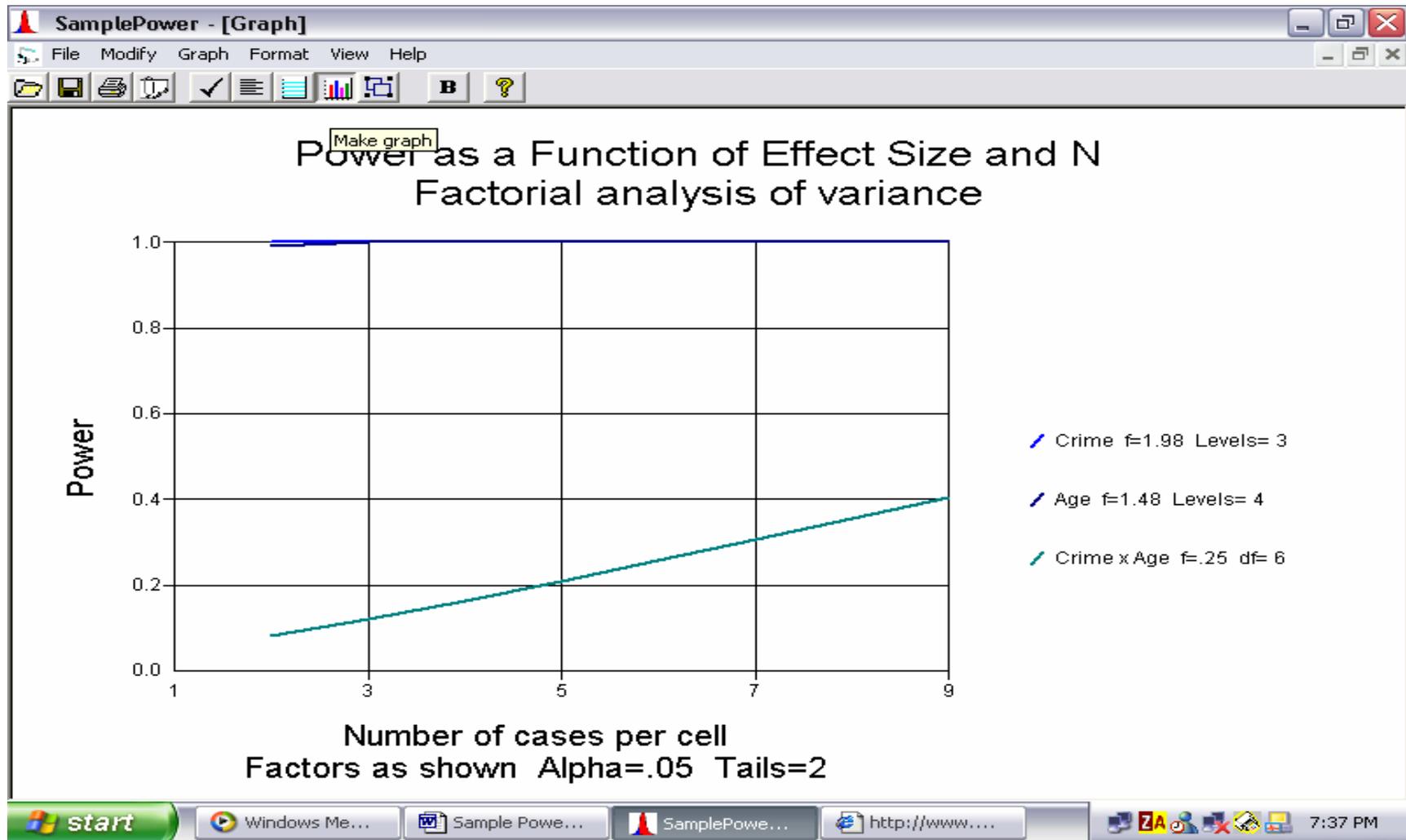
Power as a Function of Sample Size and Effect Size For Analysis of Variance

N per cell	Crime	Age	Crime x Age
	f=1.98 Levels= 3	f=1.48 Levels= 4	f=.25 df= 6
2	1.000	.993	.082
3	1.000	1.000	.120
4	1.000	1.000	.162
5	1.000	1.000	.208
6	1.000	1.000	.256
7	1.000	1.000	.305
8	1.000	1.000	.355
9	1.000	1.000	.405
10	1.000	1.000	.454

Factors as shown Alpha=.05 Tails=2
Power computations: Non-central F

start Windows Me... Sample Powe... SamplePowe... http://www.... 7:36 PM

Our chart will look like the following:



If we run the report function we get the following:

The screenshot shows a software window titled "SamplePower - [Report - ANOVA]". The window has a menu bar with "File", "Edit", "View", and "Help". Below the menu bar is a toolbar with various icons, including a "Make report" button. The main content area displays the following text:

Power for **Make report** the null hypothesis

This power analysis is for a 3 x 4 fixed effects analysis of variance. The study will include 2 cases per cell in a 24 case design, for a total of 24 cases.

The criterion for significance (alpha) has been set at 0.05. The analysis of variance is non-directional (i.e. two-tailed) which means that an effect in either direction will be interpreted.

Main effects

Crime will include 3 levels, with 8 cases per level. The effect size (f) is 1.98, which yields power of 1.00.

Age will include 4 levels, with 6 cases per level. The effect size (f) is 1.48, which yields power of 0.99.

Interactions

Crime x Age. The effect size (f) is 0.25, which yields power of 0.08.

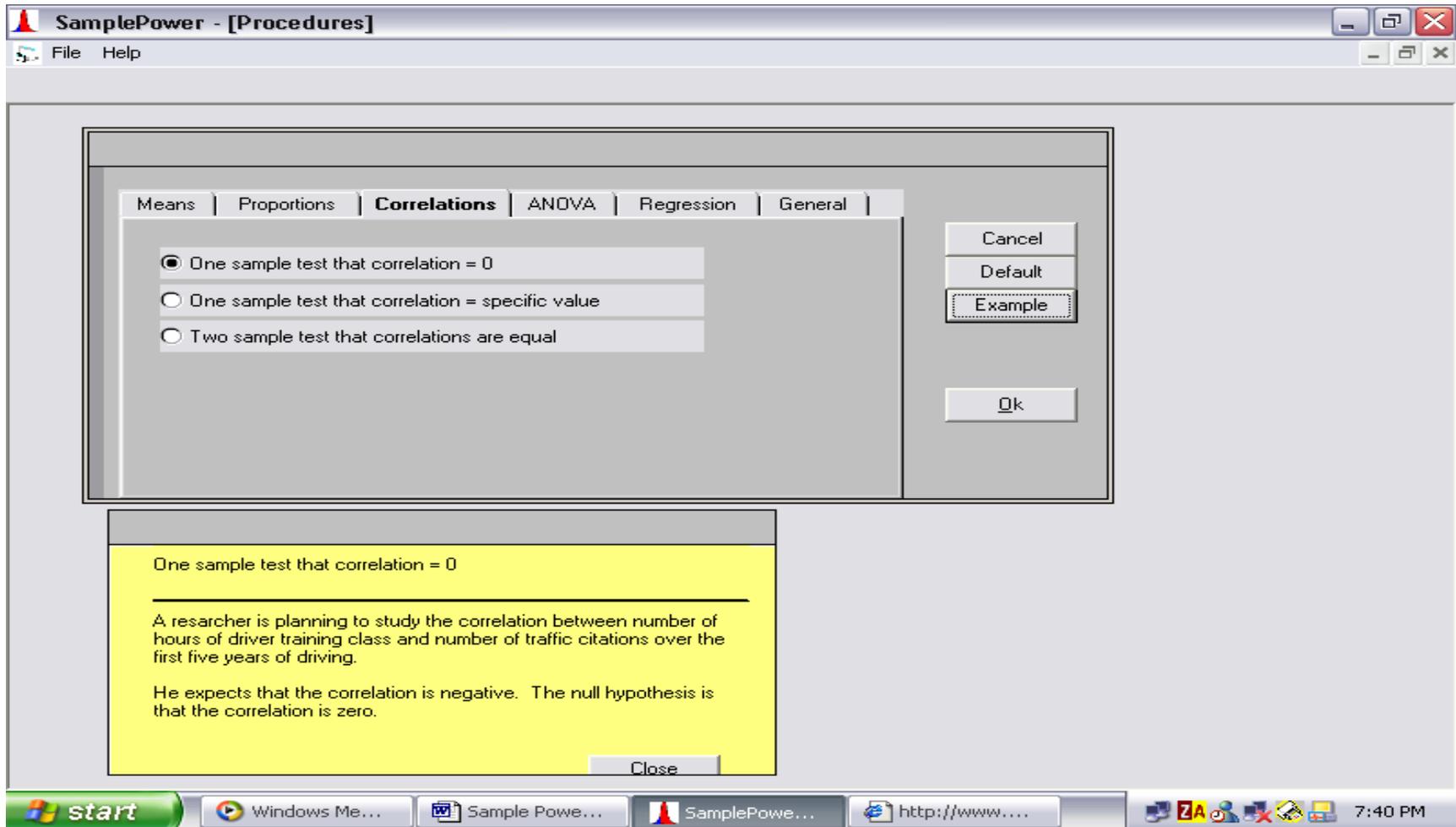
Notes

Power computations: Non-central F

The window also features a "Command1" button on the right side of the main text area. At the bottom, the Windows taskbar is visible, showing the Start button, several open windows (Windows Media Center, Sample Power..., SamplePowe...), and a web browser window (http://www....). The system tray on the right shows the time as 7:38 PM and several system icons.

Correlation:

In order to conduct a correlation analysis click on the correlation tab and select one sample test that correlation = 0.



Click Ok.

Now we will pretend that we are interested in the correlation between the number of hours spent watching television and the number of violent behaviours exhibited by children.

We will assume based on the previous literature that the population correlation is moderate at .30. We will also assume that we will be able to recruit 50 children to take part in our study.

Below we can see that with a sample of 50 children our power would only be 59%.

If we click on the binocular icon in the top toolbar we can see that in order to reach a power of 80% we would need 82 children to take part in our study.

SamplePower - [One-sample correlation]

File Options Tools View Help

Find N for power of 80%

	Population Correlation	N of Cases	Standard Error	95% Lower	95% Upper
Population	0.30	50	0.15	0.02	0.53

Alpha= 0.05, Tails= 2

Power 59%

The program displays power

For the given effect size (population $r = 0.30$, tested against a constant of 0.00), sample size (50), and alpha (0.05, 2-tailed), power is 0.59.

This means that 59% of studies would be expected to yield a significant effect, rejecting the null hypothesis that the population correlation is 0.00.

start Windows Medi... Sample Power ... SamplePower - ... http://www.wl... 7:45 PM

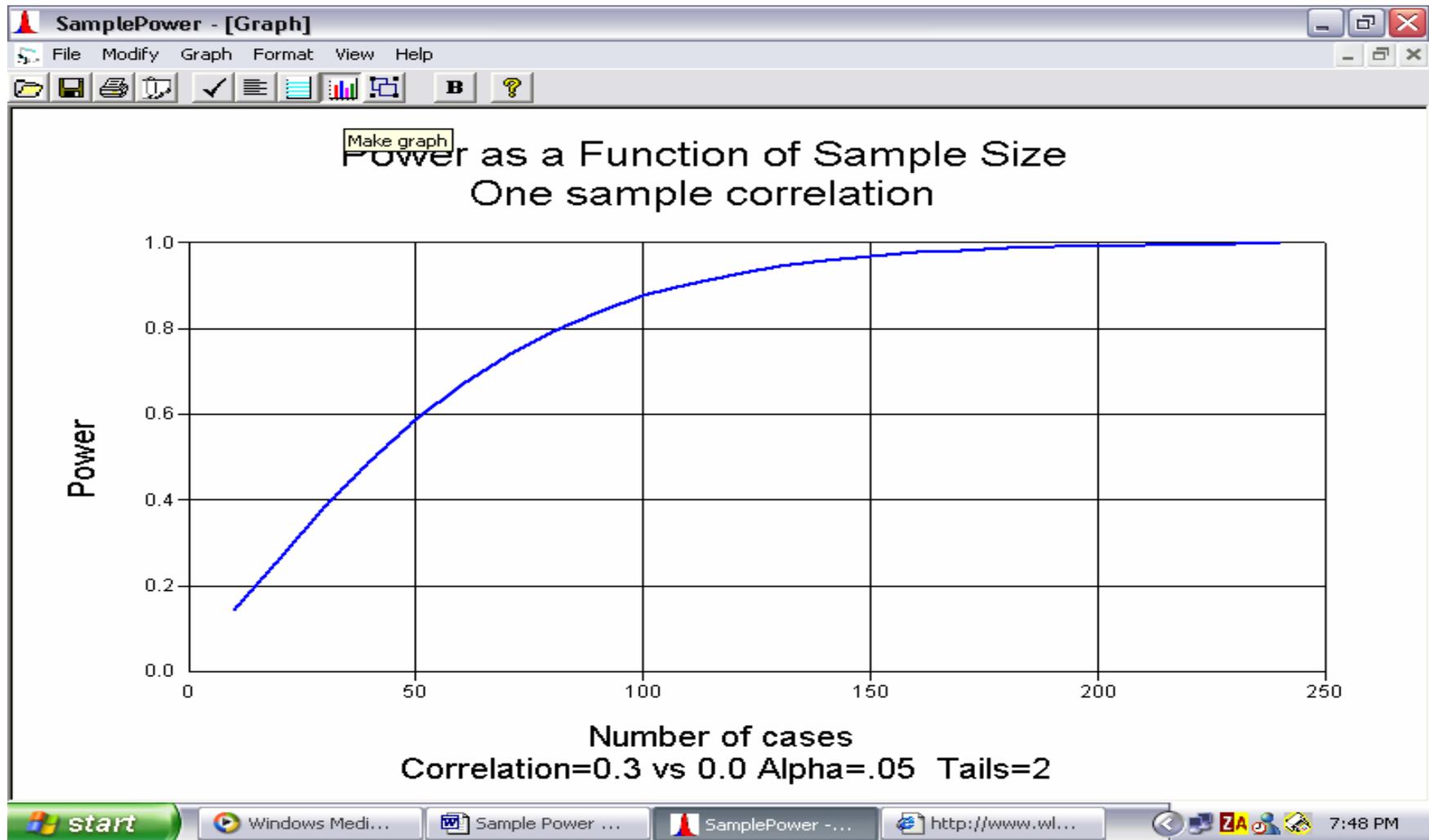
Our table would look like the following.

The screenshot shows the 'SamplePower' application window. The title bar reads 'SamplePower - [Table - One-sample correlation]'. The menu bar includes 'File', 'Modify', 'Graph', 'View', and 'Help'. The toolbar contains various icons for file operations and editing. The main area features a 'Make table' button and a table titled 'Power as a Function of Sample Size' with the subtitle 'Correlation in one sample versus a constant'. The table lists sample sizes from 10 to 200 and their corresponding power values. Below the table, the parameters are specified: 'Correlation = 0.3 vs constant of 0.0 Alpha=.05 Tails=2' and 'Power computation: Fisher Z approximation (when null=0, exact formula is used)'. The Windows taskbar at the bottom shows the Start button, several open applications, and the system clock at 7:47 PM.

N (1)	Power
10	.142
20	.265
30	.384
40	.492
50	.587
60	.668
70	.737
80	.793
90	.839
100	.876
110	.905
120	.927
130	.945
140	.959
150	.969
160	.977
170	.983
180	.987
190	.991
200	.993

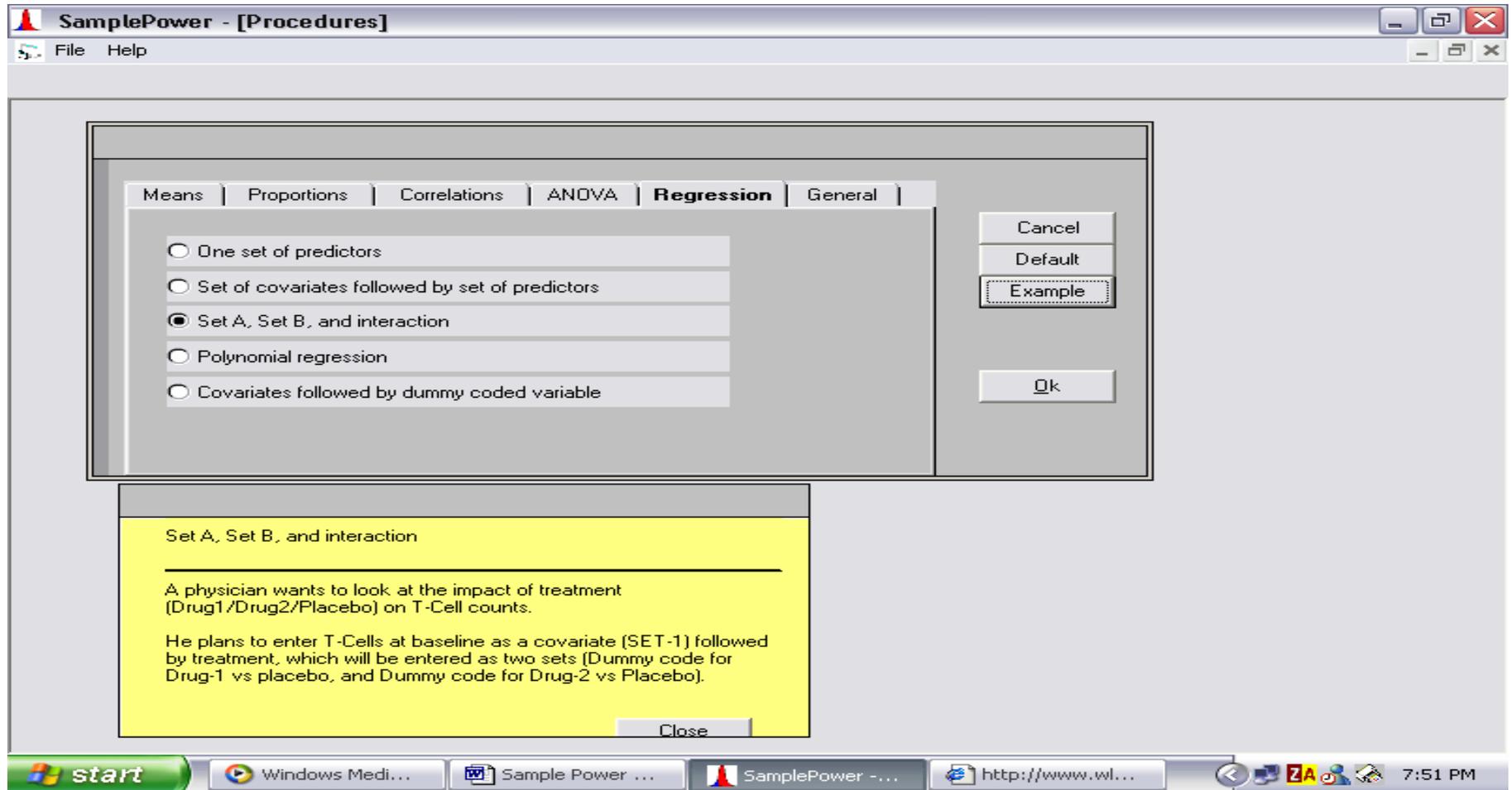
Correlation = 0.3 vs constant of 0.0 Alpha=.05 Tails=2
Power computation: Fisher Z approximation (when null=0, exact formula is used)

Our chart would look like this:



Regression:

In order to conduct a regression analysis click on the regression tab and select the appropriate option, in this case we will be using 'Set A, Set B, and interaction.'



Click Ok.

Now, using the small social work data set as our example we will enter our first set as age and age squared (so 2 variables in the set) and we will assume that this accounts for 10% of the variance. Our second set will be years of experience (so only 1 variable) and we will assume that it accounts for 5% of the variance. Our interaction set will consist of gender X managerial status (so 2 variables in the set) and we will assume that it accounts for 15% of the variance. Finally, we will need to adjust the number of cases to 250 and set the box in the bottom right hand corner of the screen to start with set 1 and end with set 3.

SamplePower - [Multiple regression]

File Options Tools View Help

Multiple regression Make graph

Variable	Increment to R-Squared			Cumulative R-Squared		
	Number Variables in Set	Increment to R-Squared	Power for Increment	Cumulative Number Variables	Cumulative R-Square	Power for Cumulative R-Squared
1 First set	2	0.10	1.00	2	0.10	1.00
2 Second set	1	0.05	0.99	3	0.15	1.00
3 Interaction	2	0.15	1.00	5	0.30	1.00

Alpha 0.05 N of cases 250

Welcome

This interactive guide will lead you through the steps for computing power and precision.

To move this box use the title bar above.

To close or reactivate this panel, select Help from the menu.

Help < Back Next >

Designated set

Increment to R-SQ 0.05

Start with set 1 End with set 3

Variables in this set 5

Increment to R-SQ 0.30

Power for this increment 1.00

start Windows ... Sample P... SamplePo... http://w... Norton A... 8:01 PM

Our chart will look like the following:

