

# What is the methodology for conducting a two independent proportions power analysis using SAS data analysis?

Authored by  
**stats writer**

June 29, 2024

## RECOMMENDED CITATION

stats writer (2024). *What is the methodology for conducting a two independent proportions power analysis using SAS data analysis?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=159317>

The methodology for conducting a two independent proportions power analysis using SAS data analysis involves a statistical approach for determining the minimum sample size required for detecting a significant difference between two proportions with a certain level of statistical power. This methodology utilizes SAS software to perform calculations and simulations based on the desired level of significance, effect size, and power. By analyzing the data using this methodology, researchers can determine the appropriate sample size needed to achieve a desired level of precision and confidence in their findings. This approach is commonly used in experimental and observational studies to ensure the validity and reliability of the results.

## Two Independent Proportions Power Analysis | SAS Data Analysis Examples

### Introduction

Power analysis is the name given to the process for determining the sample size for a research study. The technical definition of power is that it is the probability of detecting a "true" effect when it exists. Many students think that there is a simple formula for determining sample size for every research situation. However, the reality is that there are many research situations that are so complex that they almost defy rational power analysis. In most cases, power analysis involves a number of simplifying assumptions, in order to make the problem tractable, and running the

**analyses numerous times with different variations to cover all of the contingencies.**

**In this unit we will try to illustrate how to do a power analysis for a test of two independent proportions, i.e., the response variable has two levels and the predictor variable also has two levels. Instead of analyzing these data using a test of independent proportions, we could compute a chi-square statistic in a 2x2 contingency table or run a simple logistic regression analysis. These three analyses yield the same results and would require the same sample sizes to test effects.**

#### **Description of the Experiment**

**It is known that a certain type of skin lesion will develop into cancer in 30% of patients if left untreated. There is a drug on the market that will reduce the probability of cancer developing by 10%. A pharmaceutical company is developing a new drug to treat skin lesions but it will only be worthwhile to do so if the**

**new drug reduces the probability of developing skin cancer by at least 15%, an additional 5% beyond the existing drug.**

**The pharmaceutical company plans to do a study with patients randomly assigned to two groups, the control (untreated) group and the treatment group. The company wants to know how many subjects will be needed to test a difference in proportions of .15 with a power of .8 at alpha equal to .05.**

### **The Power Analysis**

**We will make use of the SAS proc power to determine the sample size needed for tests of two independent proportions as well as for tests of means.**

**The proc power needs the following information in order to do**

**the power analysis: 1) the expected proportion of cancer the untreated group ( $p_1 = .3$ ), 2) the expected proportion of cancer in the treated group ( $p_2 = .3 - .15 = .15$ ), 3)**

the alpha level (alpha = .05, the default for proc power), and 4) the required level of power (power = .8 for this experiment). There are still different versions of the same test, Pearson's chi-square test, the likelihood ratio test and Fisher's exact test. Let's run all three of them and see how different they are from each other.

```
proc power;  
twosamplefreq test=pchi  
groupproportions = (.3 .15)  
nullproportiondiff = 0  
power = .80  
npergroup =.;  
run;
```

## Pearson Chi-square Test for Two Proportions

### Fixed Scenario Elements

Distribution Asymptotic normal

Method Normal approximation

Null Proportion Difference 0

**Group 1 Proportion 0.3**

**Group 2 Proportion 0.15**

**Nominal Power 0.8**

**Number of Sides 2**

**Alpha 0.05**

**Computed N Per Group**

**Actual N Per**

**Power Group**

**0.802 121**

```
proc power;
```

```
twosamplefreq test=lrchi
```

```
groupproportions = (.3 .15)
```

```
power = .8
```

```
npergroup =.;
```

```
run;
```

**Likelihood Ratio Chi-square Test for Two Proportions**

**Fixed Scenario Elements**

**Distribution Asymptotic normal**

## Method Normal approximation

Group 1 Proportion 0.3

Group 2 Proportion 0.15

Nominal Power 0.8

Number of Sides 2

Alpha 0.05

Computed N Per Group

Actual N Per

Power Group

0.801 120

```
proc power;  
twosamplefreq test=fisher  
groupproportions = (.3 .15)  
power = .8  
npergroup = . ;  
run;
```

## Fisher's Exact Conditional Test for Two Proportions

Fixed Scenario Elements

**Distribution Exact conditional**

**Method Walters normal approximation**

**Group 1 Proportion 0.3**

**Group 2 Proportion 0.15**

**Nominal Power 0.8**

**Number of Sides 2**

**Alpha 0.05**

**Computed N Per Group**

**Actual N Per**

**Power Group**

**0.800 132**

**Pearson's chi-square test is the most commonly used. Likelihood ratio test**

**was developed after the Pearson's chi-square test and is also very common. Both**

**of the methods are based on the asymptotic theory and work well when sample size**

**is large. When sample size is small Fisher's exact method is usually more conservative.**

**This is all well and good but a two-sided test doesn't make much sense in this situation.**

**We want to test for a drug that reduces the probability of cancer not for one that increases**

**the probability. In this case we should be using one-tail test and we do this by**

**using the sides=1 option in proc power. We are going to use**

**Fisher's exact test from now on.**

```
proc power;
```

```
twosamplefreq test=fisher
```

```
groupproportions = (.3 .15)
```

```
power = .8
```

```
npergroup = .
```

```
sides = 1;
```

```
run;
```

## **Fisher's Exact Conditional Test for Two Proportions**

### **Fixed Scenario Elements**

**Distribution Exact conditional**

**Method Walters normal approximation**

**Number of Sides 1**

**Group 1 Proportion 0.3**

**Group 2 Proportion 0.15**

**Nominal Power 0.8**

**Alpha 0.05**

**Computed N Per Group**

**Actual N Per  
Power Group**

**0.801 107**

**This is better. The result indicates that we need to use 107 subjects in each group to find a change in probability of .15 for a power of .8 when alpha equals .05**

**Just as a check let's run the analysis specifying each of the two sample sizes.**

```
proc power;  
twosamplefreq test=fisher  
groupproportions = (.3 .15)  
npergroup = 107  
sides = 1
```

```
power = .;  
run;
```

## Fisher's Exact Conditional Test for Two Proportions

### Fixed Scenario Elements

**Distribution Exact conditional**

**Method Walters normal approximation**

**Number of Sides 1**

**Group 1 Proportion 0.3**

**Group 2 Proportion 0.15**

**Sample Size Per Group 107**

**Alpha 0.05**

**Computed Power**

**Power**

**0.801**

Now because we believe that we know a lot about the incidence of cancer in the untreated group we would like to make the control group half as large as the treatment group. We can

**use the groupweights option to achieve this. Notice that this option only works with ntotal =.**

```
proc power;  
twosamplefreq test=fisher  
groupproportions = (.3 .15)  
power = .8  
groupweights =(1 2)  
ntotal =.  
sides = 1;  
run;
```

## **Fisher's Exact Conditional Test for Two Proportions**

### **Fixed Scenario Elements**

**Distribution Exact conditional**

**Method Walters normal approximation**

**Number of Sides 1**

**Group 1 Proportion 0.3**

**Group 2 Proportion 0.15**

**Group 1 Weight 1**

**Group 2 Weight 2**

**Nominal Power 0.8**

**Alpha 0.05**

**Computed N Total**

**Actual N**

**Power Total**

**0.802 240**

As you can see, by the way we have specified the ratio, we need 1/3 of the total sample for group 1 and 2/3 for group 2. This turns out to be 80 and 160 for the two groups. So we see that we will need more subjects overall than for equal sized groups but we can have a much smaller untreated group.

In the end, the company has decided to use 75 patients in the control group and 150 in the treatment group. Let's see what the power is.

```
proc power;  
twosamplefreq test=fisher  
groupproportions = (.3 .15)
```

```
power = .  
groupns =(75 150)  
sides = 1;  
run;
```

## Fisher's Exact Conditional Test for Two Proportions

### Fixed Scenario Elements

Distribution Exact conditional

Method Walters normal approximation

Number of Sides 1

Group 1 Proportion 0.3

Group 2 Proportion 0.15

Group 1 Sample Size 75

Group 2 Sample Size 150

Alpha 0.05

### Computed Power

Power

**0.776**

**With this unbalanced design we have an estimated**

**power of .776 which the company deems acceptable.**

**See Also**

**Cohen, J. 1988. Statistical Power Analysis for the Behavioral Sciences, Second Edition.**

**Mahwah, NJ: Lawrence Erlbaum Associates.**

ARABPSYCHOLOGY.COM