Chapter 150

Tests for Two Correlated Proportions (McNemar Test)

Introduction

McNemar's test compares the proportions for two correlated dichotomous variables. These two variables may be two responses on a single individual or two responses from a matched pair (as in matched case-control studies).

This procedure is similar to the Matched Case-Control procedure also available in **PASS**. It differs from that procedure in three basic ways:

1. The results may be calculated exactly by multinomial enumeration using an unconditional formula rather than using conditional, normal approximations.
2. It only deals with the case of a matched pair: one case and one control (the Matched Case-Control procedure lets you match several controls with each case).
3. It is based directly on a 2-by-2 contingency table.

To fix these ideas, consider the following fictitious data concerning the relationship between smoking and lung cancer. Suppose that a sample of *N* = 100 cases of identical twins in which only one twin has lung cancer is selected for further study. The twin with lung cancer is the *case*. The other twin serves as the *control*. Each pair of twins is surveyed to determine if they smoke tobacco. The results are summarized in the following two-way table:

	No Lung Cancer Twin (Control)	
Lung Cancer Twin (Case)	**Smokes = Yes**	**Smokes = No**
Smokes = Yes	16	21
Smokes = No	4	59

There is a basic difference between this table and the more common two-way table. In the matched-paired case, the count represents the number of pairs, not the number of individuals.

The investigator wishes to compare the proportion of cases that smoke with the proportion of controls that smoke. The proportion of controls who smoke is (16+4)/100 = 0.20. The proportion of cases who smoke is (16+21)/100 = 0.37.

Dividing each of the entries in the table by *N* gives the proportions:

	No Lung Cancer Twin (Control)		
Lung Cancer Twin (Case)	**Smokes = Yes**	**Smokes = No**	**Total**
Smokes = Yes	0.16	0.21	0.37
Smokes = No	0.04	0.59	0.63
Total	0.20	0.80	0.63

Symbolically, this table is represented as:

	No Lung Cancer Twin (Control)		
Lung Cancer Twin (Case)	**Smokes = Yes**	**Smokes = No**	**Total**
Smokes = Yes	P11	P10	Pt
Smokes = No	P01	P00	1-Pt
Total	Ps	1-Ps	1

Formally, the hypothesis of interest is that Pt equals Ps. A little algebra shows that Pt = Ps is equivalent to P10 = P01, since P11 is common to both. Thus, the null hypothesis of McNemar's test is P10 = P01 and the alternative is that they are unequal. The alternative hypothesis may be one-sided (such as P10 > P01) or two-sided (P10 $\neq$ P01).

The null hypothesis may also be stated in terms of the McNemar odds ratio as OR = 1. The McNemar odds ratio is not the sample as the regular odds ratio of Pt and Ps. The formula is:

$$OR = \frac{P10}{P01}$$

Notice that the values of P11 and P00 are not used directly in these hypotheses. It turns out that their individual values are not needed, but their sum is.

For this example, the odds ratio is computed as 0.21/0.04 = 5.25.

Technical Details

Consider the matched-pairs table again:

	Controls		
Cases	**Yes**	**No**	**Total**
Yes	P11	P10	Pt
No	P01	P00	1-Pt
Total	Ps	1-Ps	1

Pairs with the same response from cases and controls (Yes-Yes and No-No) are called *concordant* pairs. Pairs with different responses (Yes-No and No-Yes) are called *discordant* pairs. McNemar's test statistic is the estimated odds ratio:

$$Mc = \frac{P10}{P01}$$

The sample size problem thus reduces to a study of how many Yes-No's and No-Yes's are needed. Once this has been determined, the overall sample size is found by estimating the proportion of discordant pairs and inflating the sample size appropriately.

Some power analysis programs follow an approximate procedure. Since the McNemar statistic follows the binomial probability distribution for a fixed number of discordant pairs, they use formulas that use the normal approximation to the binomial and then adjust the sample size depending on the proportion of discordant pairs, *PD=P10+P01*. This is called the conditional procedure.

One such approximate formula is given by Machin, Campbell, Fayers, and Pinol (1997):

$$N_{pairs} = \frac{\left\{z_{1-\alpha/s}(OR+1) + z_{1-\beta}\sqrt{(OR+1)^2 - (OR-1)^2 PD}\right\}^2}{(OR-1)^2 PD}$$

where *s* is the number of sides to the test (one or two), $OR = \frac{P10}{P01}$, $PD = P10 + P01$, and α and β are the usual type 1 and type 2 error rate probabilities. This sample size formula may be rearranged to compute the power.

However, Schork and Williams (1980) published a formula which provides the exact results for the unconditional case using multinomial enumeration of all possible outcomes. This formulation is also available in **PASS**:

$$Power = \sum_{R=r}^{N} \sum_{n_{12}=0}^{IR} \frac{N!}{(N-R)!\,n_{12}!\,(R-n_{12})!}(1-PD)^{N-R}\left(\frac{D+PD}{2}\right)^{n_{12}}\left(\frac{PD-D}{2}\right)^{R-n_{12}}$$

where

PD = P10 + P01

D = P10 - P01

N is total of all four cells (*N11* + *N12* + *N21* + *N22*)

r is the smallest integer for which $\left(\frac{1}{2}\right)^r \le \alpha$

IR is the largest integer such that $\sum_{j=0}^{IR}\binom{R}{j}\left(\frac{1}{2}\right)^R \le \alpha$

Difference or Odds Ratio

The formula given above is parameterized in terms of the difference. This formula is also used when odds ratios are specified. The program simply converts the OR value into its corresponding D value.

Estimating P11, P01, and P10 using Pt, Ps, and ρ

Sometimes, obtaining estimates of *P*01 and *P*10 is problematic. This problem is solved by using the marginal probabilities and the within-subject correlation coefficient, which may be easier to estimate. As outlined in Zhang, Cao, and Ahn (2017), the relationship between *P*11, *Pt*, *Ps* and the correlation is

$$\rho = \frac{P11 - P_s P_t}{\sqrt{P_s P_t (1 - P_s)(1 - P_t)}}$$

Using this relationship, values of ρ can be entered and transformed to the corresponding value of *P*11 using the equation

$$P11 = \rho\sqrt{P_s P_t (1 - P_s)(1 - P_t)} + P_s P_t$$

The only concern is that values of ρ be used that limit *P*11, *P*01, *P*10, and *P*00 to be between 0 and 1. The lower and upper limits of the correlation are

$$\rho_L = \max\left\{ -\sqrt{\frac{P_s P_t}{(1 - P_s)(1 - P_t)}}, -\sqrt{\frac{(1 - P_s)(1 - P_t)}{P_s P_t}} \right\}$$

$$\rho_U = \min\left\{ \sqrt{\frac{P_s(1 - P_t)}{P_t(1 - P_s)}}, \sqrt{\frac{P_t(1 - P_s)}{P_s(1 - P_t)}} \right\}$$

*P*11, along with *Pt* and *Ps*, can then be used to calculate *P*01 and *P*10.

Discussion – Multinomial Enumeration vs. Normal Approximation

The multinomial enumeration (exact) algorithm works for $N < 2000$. Above 2000, computing time goes up and the algorithm has numerical problems. **PASS** lets you select either the multinomial enumeration (exact) or the approximate solution. We have found that the approximate solution tends to result a sample size that is about 10% less than the exact solution.

Because of the lengthy computer time necessary to compute the exact answer when $N > 1500$, we suggest that you use the approximate formula to begin with and then revert to the exact formula when you are ready for your final results.

Example 1 – Calculating Power using Off-Diagonal Probabilities

This example will show how to calculate the power of a study for several sample size values. Suppose that a matched case-control study is run in which the odds ratio to detect is 2, *PD* is 0.3, *N* = 50 to 200 by 50, alpha is 0.05, and power is to be calculated for a two-sided test. The input will only specify the off-diagonal probabilities P01 and P10.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 1** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For	**Power**
Power Calculation Method	**Multinomial Enumeration**
Max N for Multinomial Enumeration	**3000**
Alternative Hypothesis	**Two-Sided**
Alpha	**0.05**
N (Number of Pairs)	**50 100 150 200**
Probability Input Type	**Off Diagonal (P01 and P10)**
P01 and P10 Input Type	**McNemar Odds Ratio (P10 / P01)**
McNemar Odds Ratio (P10 / P01)	**2**
Proportion Discordant (P10 + P01)	**0.3**

Output

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

Numeric Reports

Numeric Results

Solve For: Power
Alternative Hypothesis: Two-Sided

Power*	Total Sample Size N	Off-Diagonal Probabilities P10	P01	Difference P10 - P01	McNemar Odds Ratio P10 / P01	Proportion Discordant P10 + P01	Target Alpha†
0.1785	50	0.2	0.1	0.1	2	0.3	0.05
0.3730	100	0.2	0.1	0.1	2	0.3	0.05
0.5646	150	0.2	0.1	0.1	2	0.3	0.05
0.7034	200	0.2	0.1	0.1	2	0.3	0.05

* Power was computed using multinomial enumeration of all possible outcomes (unconditional).

† Warning: For small values of N (i.e., less than 100), power computed by Multinomial Enumeration may be overly optimistic because the discrete nature of the multinomial distribution results in the actual alpha value being higher than its target. To be safe, we recommend that you use the power calculation based on the normal approximation.

Power	The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
N	The total number of subjects in the study.
P10	The probability that the treatment response is "Yes" and the standard response is "No."
P01	The probability that the treatment response is "No" and the standard response is "Yes."
P10 - P01	The difference between P10 and P01. It is equal to the difference between Pt and Ps.
P10 / P01	The McNemar odds ratio. It is not equal to the regular odds ratio between Pt and Ps.
P10 + P01	Proportion Discordant. The sum of the two off-diagonal elements, P10 and P01.
Alpha	The probability of rejecting a true null hypothesis.

Summary Statements

A paired design will be used to test whether there is a difference in proportions. The comparison will be made using a two-sided, paired-sample McNemar Test, with a Type I error rate (α) of 0.05. To detect a McNemar odds ratio of 2 with a sample size of 50 pairs, the power is 0.1785. The McNemar odds ratio is equivalent to a difference between two paired proportions of 0.1 which occurs when the proportion in cell 1,0 is 0.2 and the proportion in cell 0,1 is 0.1. The proportion of discordant pairs is 0.3.

Dropout-Inflated Sample Size

Dropout Rate	Sample Size N	Dropout-Inflated Enrollment Sample Size N'	Expected Number of Dropouts D
20%	50	63	13
20%	100	125	25
20%	150	188	38
20%	200	250	50

Dropout Rate	The percentage of subjects (or items) that are expected to be lost at random during the course of the study and for whom no response data will be collected (i.e., will be treated as "missing"). Abbreviated as DR.
N	The evaluable sample size at which power is computed (as entered by the user). If N subjects are evaluated out of the N' subjects that are enrolled in the study, the design will achieve the stated power.
N'	The total number of subjects that should be enrolled in the study in order to obtain N evaluable subjects, based on the assumed dropout rate. N' is calculated by inflating N using the formula N' = N / (1 - DR), with N' always rounded up. (See Julious, S.A. (2010) pages 52-53, or Chow, S.C., Shao, J., Wang, H., and Lokhnygina, Y. (2018) pages 32-33.)
D	The expected number of dropouts. D = N' - N.

Dropout Summary Statements

Anticipating a 20% dropout rate, 63 subjects should be enrolled to obtain a final sample size of 50 subjects.

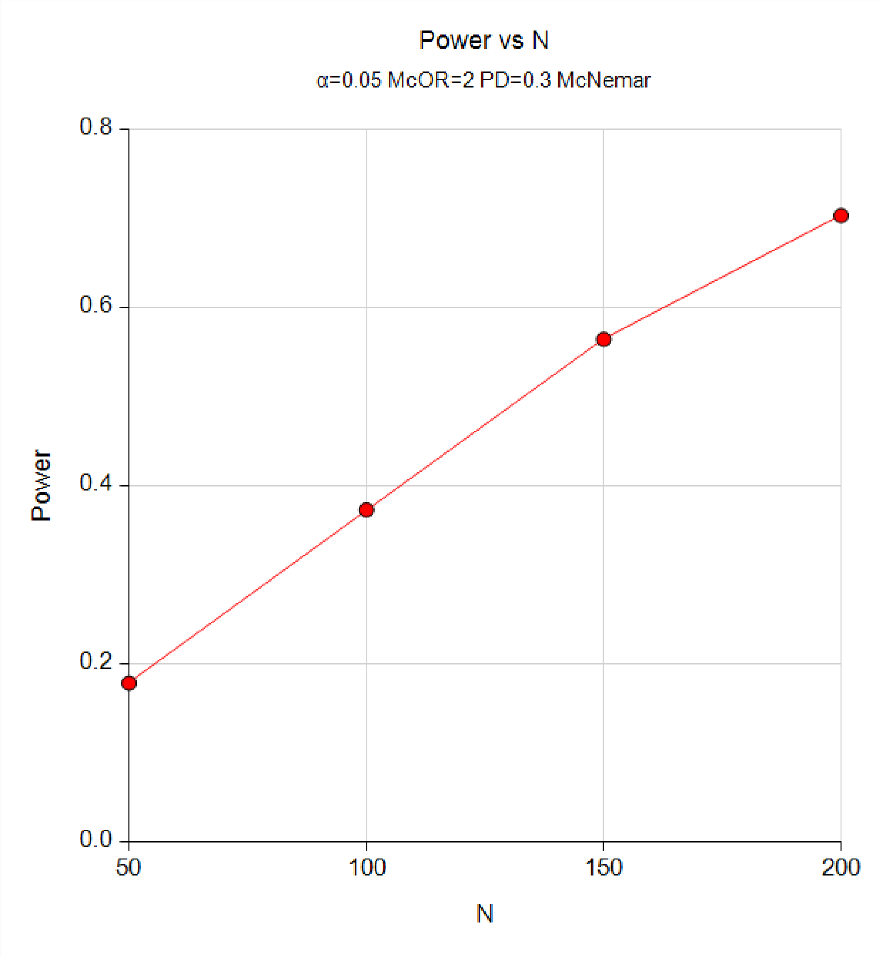
References

Schork, M. and Williams, G. 1980. 'Number of Observations Required for the Comparison of Two Correlated Proportions.' Communications in Statistics-Simula. Computa., B9(4), 349-357.
Machin, D., Campbell, M., Fayers, P., and Pinol, A. 1997. Sample Size Tables for Clinical Studies, 2nd Edition. Blackwell Science. Malden, MA.
Zhang, S., Cao, J., Ahn, C. 2017. 'Inference and sample size calculation for clinical trials with incomplete observations of paired binary outcomes'. Statistics in Medicine. Volume 36. Pages 581-591.

This report shows the power for each of the scenarios.

Plots Section

Plots



These plots show the power versus the sample size.

Example 2 – Validation using Schork and Williams (1980)

Schork and Williams (1980) page 354 presents a table of sample sizes for various combinations of the other parameters. When the difference is 0.2, the proportion discordant is 0.7, the power is 80%, and the one-sided significance level is 0.025, the sample size is 144.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 2** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For .... **Sample Size**
Power Calculation Method .... **Multinomial Enumeration**
Max N for Multinomial Enumeration **3000**
Alternative Hypothesis .... **One-Sided**
Power.. **0.8**
Alpha... **0.025**
Probability Input Type .... **Off Diagonal (P01 and P10)**
P01 and P10 Input Type .... **Difference (P10 - P01)**
Difference (P10 - P01) **0.2**
Proportion Discordant (P10 + P01) **0.7**

Output

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

Numeric Results

Solve For: Sample Size
Alternative Hypothesis: One-Sided

Power*	Total Sample Size N	Off-Diagonal Probabilities P10	P01	Difference P10 - P01	McNemar Odds Ratio P10 / P01	Proportion Discordant P10 + P01	Target Alpha†
0.8009	144	0.45	0.25	0.2	1.8	0.7	0.025

* Power was computed using multinomial enumeration of all possible outcomes (unconditional).

† Warning: For small values of N (i.e., less than 100), power computed by Multinomial Enumeration may be overly optimistic because the discrete nature of the multinomial distribution results in the actual alpha value being higher than its target. To be safe, we recommend that you use the power calculation based on the normal approximation.

PASS also finds the sample size to be 144.

Example 3 – Calculating Sample Size using Marginal Probabilities

This example will show how to calculate the sample when the input is in terms of the marginal probabilities, Pt and Ps and the within-subject correlation. Suppose that a paired study is being planned in which the standard response probability is 0.5; the treatment response probability is assumed to be between 0.55 and 0.65; the within-subject correlation is assumed to be somewhere between 0 and 0.6; alpha is 0.05; power is 0.8 for a two-sided test.

Notice that with this set of input parameters, the difficult to estimate proportion of discordant is not required.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 3** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For	**Sample Size**
Power Calculation Method	**Multinomial Enumeration**
Max N for Multinomial Enumeration	**3000**
Alternative Hypothesis	**Two-Sided**
Power	**0.8**
Alpha	**0.05**
Probability Input Type	**Marginal (Pt, Ps, and ρ)**
Pt Input Type	**Pt**
Pt (Probability (Yt = 1))	**0.55 0.6 0.65**
Ps (Probability (Ys = 1))	**0.5**
ρ Input Type	**ρ (Within-Subject Correlation)**
ρ (Within-Subject Correlation)	**0 0.2 0.4 0.6**

Output

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

Numeric Results

Solve For: Sample Size
Alternative Hypothesis: Two-Sided

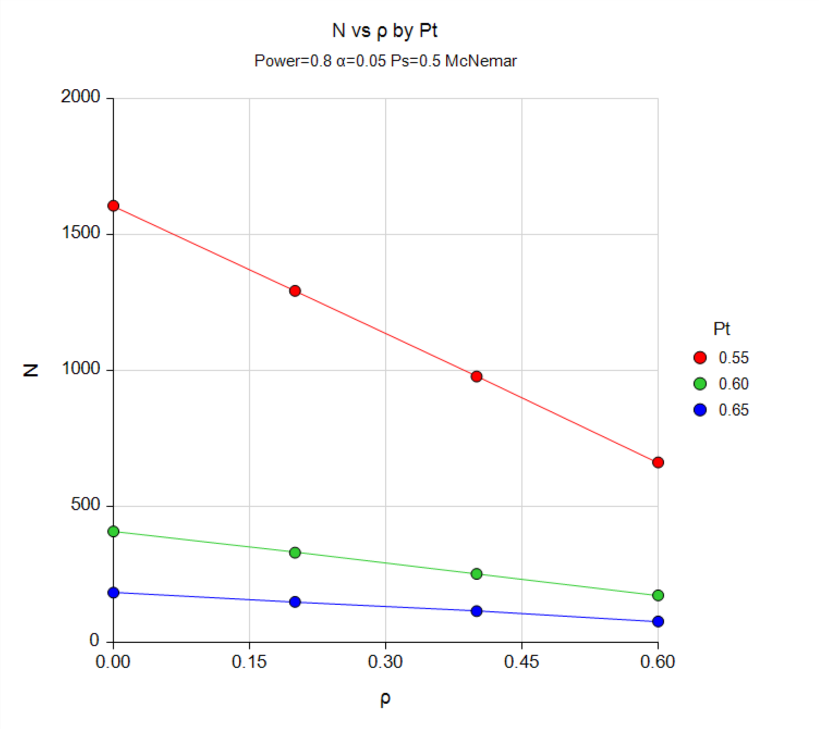
Power*	Total Sample Size N	Marginal Probabilities Pt	Marginal Probabilities Ps	Difference Pt - Ps	Within-Subject Correlation ρ	Proportion Discordant P10 + P01	Joint Probability that Yt = 1 and Ys = 1 P11	Target Alpha†
0.8000	1606	0.55	0.5	0.05	0.0	0.5000	0.2750	0.05
0.8002	1293	0.55	0.5	0.05	0.2	0.4005	0.3247	0.05
0.8002	978	0.55	0.5	0.05	0.4	0.3010	0.3745	0.05
0.8002	662	0.55	0.5	0.05	0.6	0.2015	0.4242	0.05
0.8002	408	0.60	0.5	0.10	0.0	0.5000	0.3000	0.05
0.8006	330	0.60	0.5	0.10	0.2	0.4020	0.3490	0.05
0.8005	252	0.60	0.5	0.10	0.4	0.3040	0.3980	0.05
0.8016	173	0.60	0.5	0.10	0.6	0.2061	0.4470	0.05
0.8000	183	0.65	0.5	0.15	0.0	0.5000	0.3250	0.05
0.8025	149	0.65	0.5	0.15	0.2	0.4046	0.3727	0.05
0.8013	115	0.65	0.5	0.15	0.4	0.3092	0.4204	0.05
0.8030	77	0.65	0.5	0.15	0.6	0.2138	0.4681	0.05

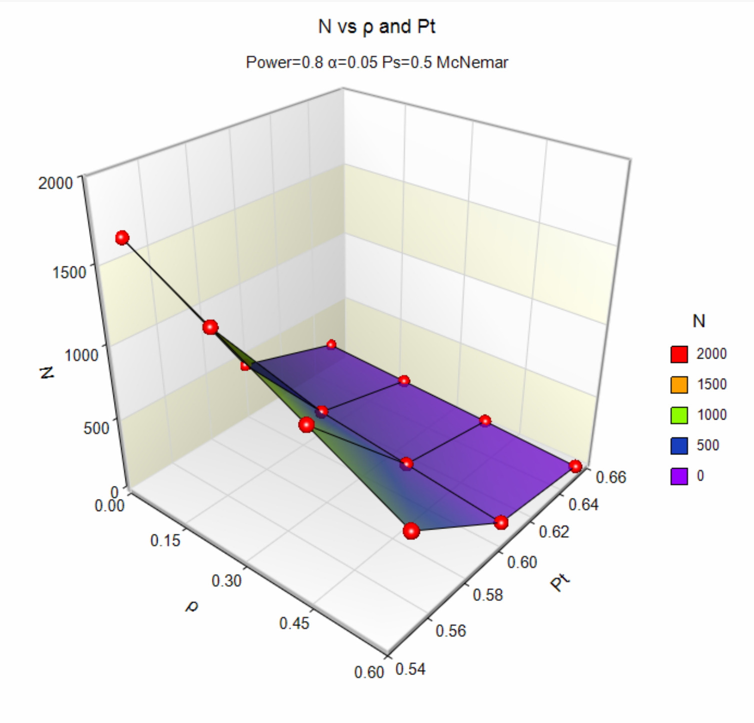
* Power was computed using multinomial enumeration of all possible outcomes (unconditional).

† Warning: For small values of N (i.e., less than 100), power computed by Multinomial Enumeration may be overly optimistic because the discrete nature of the multinomial distribution results in the actual alpha value being higher than its target. To be safe, we recommend that you use the power calculation based on the normal approximation.

Power	The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
N	The total number of subjects in the study.
Pt	The marginal probability of a "true" response in the treatment observation.
Ps	The marginal probability of a "true" response in the standard observation.
Pt - Ps	The difference between the two marginal probabilities.
ρ	The correlation between the two observations within a subject.
P11	The joint probability that both observations in a pair are true (equal to 1).
Alpha	The probability of rejecting a true null hypothesis.

Plots


N vs ρ by Pt
Power=0.8 α=0.05 Ps=0.5 McNemar
N
ρ
Pt
0.55
0.60
0.65
2000
1500
1000
500
0
0.00
0.15
0.30
0.45
0.60


N vs ρ and Pt
Power=0.8 α=0.05 Ps=0.5 McNemar
N
ρ
Pt
2000
1500
1000
500
0

Example 4 – Computing Sample Size for a 2x2 Cross-Over Design

Julious (2010) indicates on page 167 that a 2x2 cross-over trial can be analyzed using the McNemar test if the period effect is ignored. This example will show you how to estimate the sample size for a 2x2 cross-over trial based on McNemar test.

In example 10.1 on page 170, Julious (2010) provides a table with both joint and marginal probabilities. From this table, we find that Pt = 0.72, Ps = 0.56, and P11 = 0.4. We will calculate the required sample size using the conditional (normal approximation) power formulas. The required power is 90% and the two-sided significance level is 0.05. He estimates the required sample size as 190.

Note that Julious (2010) uses a slightly different asymptotic sample size formula (see page 169) from that used by this procedure.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 4a** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For	**Sample Size**
Power Calculation Method	**Normal Approximation**
Alternative Hypothesis	**Two-Sided**
Power	**0.9**
Alpha	**0.05**
Probability Input Type	**Marginal (Pt, Ps, and ρ)**
Pt Input Type	**Pt**
Pt (Probability (Yt = 1))	**0.72**
Ps (Probability (Ys = 1))	**0.56**
ρ Input Type	**P11 (Probability (Yt = 1, Ys = 1))**
P11 (Probability (Yt = 1, Ys = 1))	**0.4**

Output

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

Numeric Results

Solve For: Sample Size
Alternative Hypothesis: Two-Sided

Power*	Total Sample Size N	Marginal Probabilities Pt	Ps	Difference Pt - Ps	Within-Subject Correlation ρ	Proportion Discordant P10 + P01	Joint Probability that Yt = 1 and Ys = 1 P11	Alpha
0.9003	193	0.72	0.56	0.16	-0.0144	0.48	0.4	0.05

* Power was computed using the normal approximation method (conditional).

PASS finds the sample size to be 193. It's not surprising that the sample size is slightly different from Julious (2010) because the calculation formulas are not the same.

If we set the Power Calculation Method to *Multinomial Enumeration* and rerun this example, we obtain the following results (**Example 4b**).

Numeric Results

Solve For: Sample Size
Alternative Hypothesis: Two-Sided

Power*	Total Sample Size N	Marginal Probabilities Pt	Ps	Difference Pt - Ps	Within-Subject Correlation ρ	Proportion Discordant P10 + P01	Joint Probability that Yt = 1 and Ys = 1 P11	Target Alpha†
0.9008	203	0.72	0.56	0.16	-0.0144	0.48	0.4	0.05

* Power was computed using multinomial enumeration of all possible outcomes (unconditional).

† Warning: For small values of N (i.e., less than 100), power computed by Multinomial Enumeration may be overly optimistic because the discrete nature of the multinomial distribution results in the actual alpha value being higher than its target. To be safe, we recommend that you use the power calculation based on the normal approximation.

PASS finds the sample size to be 203 in the unconditional case. This is about a 5% increase in sample size over the 193 in the conditional case.