Chapter 527

Tests for the Difference of Two Means in a Higher-Order Cross-Over Design

Introduction

This procedure calculates power and sample size for inequality tests which use the difference in the means of a higher-order cross-over design. Measurements are made on individuals that have been randomly assigned to one of several treatment sequences. Only a brief introduction to the subject will be given here. For a comprehensive discussion on the subject, refer to Chen et al. (1997) and Chow et al. (2003).

Cross-Over Designs

Senn (2002) defines a *cross-over* design as one in which each subject receives all treatments at least once and the object is to study differences among the treatments. The name *cross-over* comes from the most common case in which there are only two treatments. In this case, each subject *crosses over* from one treatment to another. It is assumed that there is a *washout* period between treatments during which the response returns to its baseline value. If this does not occur, there is said to be a *carryover* effect.

A 2x2 cross-over design refers to two periods (and treatments) and two *sequences* (treatment orderings). One sequence of treatments is treatment A followed by treatment B. The other sequence is B and then A. The design includes a washout period between responses to make certain that the effects of the first drug do no carryover to the second. Thus, the groups of subjects in this design are defined by the sequence in which the two treatments are administered, not by the treatments they receive.

Higher-Order Cross-Over Designs

Chen et al. (1997) present sample size results for four cross-over designs that are more complicated than the 2x2 design. Assume that the two treatments are labeled A and B. The available designs are defined by the order and number of times the two treatments are administered.

Balaam's Design

Balaam's design has four sequences with two treatments each. It is popular because it allows the intrasubject variabilities to be estimated. The design is

Sequence	Period 1	Period 2
1	A	A
2	B	B
3	A	B
4	B	A

Two-Sequence Dual Design

This design has two sequences with three periods each. It is popular because it allows the intrasubject variabilities to be estimated. The design is

Sequence	Period 1	Period 2	Period 3
1	A	B	B
2	B	A	A

Four-Period Design with Two Sequences

This design has two sequences of four periods each. The design is

Sequence	Period 1	Period 2	Period 3	Period 4
1	A	B	B	A
2	B	A	A	B

Four-Period Design with Four Sequences

This design has four sequences of four periods each. The design is

Sequence	Period 1	Period 2	Period 3	Period 4
1	A	A	B	B
2	B	B	A	A
1	A	B	B	A
2	B	A	A	B

Advantages of Cross-Over Designs

A comparison of treatments on the same subject is expected to be more precise. The increased precision often translates into a smaller sample size. Also, patient enrollment may be easier to obtain because each patient will receive both treatments.

Disadvantages of Cross-Over Designs

The statistical analysis of a cross-over experiment is more complex than a parallel-group experiment and requires additional assumptions. In a cross-over experiment, it may be difficult to separate the treatment effect from the time effect and the carry-over effect of the previous treatment.

These cross-over designs cannot be used when the treatment (or the measurement of the response) alters the subject permanently. Hence, it cannot be used to compare treatments that are intended to provide a cure.

Because subjects must be measured at least twice, it may be more difficult to keep patients enrolled in the study. This is particularly true when the measurement process is painful, uncomfortable, embarrassing, or time consuming.

PASS Sample Size Software   NCSS.com

Tests for the Difference of Two Means in a Higher-Order Cross-Over Design

The Statistical Hypotheses

A two-sided hypothesis test comparing the means is given by

$$H_0: \mu_1 - \mu_2 = 0 \quad \text{versus} \quad H_1: \mu_1 - \mu_2 \neq 0$$

Rejecting H_0 implies that the means are different.

Test Statistic

The presence of a non-zero difference between the means can be tested using an F-test from an appropriate repeated measures model that may include terms for a sequence effect and/or a carry-over effect. The within mean square error from the ANOVA provides an estimate of the within-subject variance, σ_w^2. If prior studies are not available, you can use estimates of the between-subject variance and the within-subject correlation to calculate the within-subject variance using the formula

$$\sigma_w^2 = \sigma_b^2 (1 - \rho)$$

Power Calculation

The power of a tw0-sided test is given by Chen, Chow, and Li (1997). The basic formula is

$$\text{Power} = T_V \left[\left(\frac{|\mu_1 - \mu_2|}{\sigma_w \sqrt{b/n}} \right) - t_{V, 1-\alpha/2} \right]$$

where T represents the cumulative t distribution, V and b depend on the design (shown below), σ_w is defined above, and n is the average number of subjects per sequence. Thus, the overall sample size N is given by n times the number of sequences. For a one-sided test, replace $\alpha/2$ with α in the above formula.

The constants V and b depend on the design as follows:

Design Type	Parameters (V,b)
Balaam's Design	$V = 4n - 3$, $b = 2$.
Two-Sequence Dual Design	$V = 4n - 4$, $b = 3/4$.
Four-Period Design with Two Sequences	$V = 6n - 5$, $b = 11/20$.
Four-Period Design with Four Sequences	$V = 12n - 5$, $b = 1/4$.

527-3

© NCSS, LLC. All Rights Reserved.

Example 1 – Finding Power

Researchers want to calculate the power of a two-sided, inequality test using data from a three-period, two-sequence, dual cross-over design. They want to investigate sample sizes between 6 and 66. They want a sample large enough to detect a mean difference of 14 or 16. Similar experiments have had a standard deviation (σw) of about 25. The significance level is 0.05.

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**
Design Type...**3x2 (Three-Period, Two-Sequence Dual: ABB|BAA)**
Alternative Hypothesis ....**Two-Sided**
Alpha..**0.05**
N (Total Sample Size)...................................**6 to 66 by 10**
|μ1 - μ2|...**14 16**
Specify σ as σw or σb and ρ.........................**σw (Within Std Dev)**
σw (Within Standard Dev).............................**25**

Output

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

Numeric Reports

Numeric Results
───

Solve For: Power
Design Type: Three-Period, Two-Sequence Dual
Treatment Sequences: ABB | BAA
Alternative Hypothesis: Two-Sided
───

Power	Total Sample Size N	Mean Difference $\|\mu_1 - \mu_2\|$	Within Standard Deviation σ_w	Alpha
0.1348	6	14	25	0.05
0.1675	6	16	25	0.05
0.4139	16	14	25	0.05
0.5165	16	16	25	0.05
0.6251	26	14	25	0.05
0.7419	26	16	25	0.05
0.7715	36	14	25	0.05
0.8708	36	16	25	0.05
0.8658	46	14	25	0.05
0.9385	46	16	25	0.05
0.9235	56	14	25	0.05
0.9718	56	16	25	0.05
0.9575	66	14	25	0.05
0.9875	66	16	25	0.05

───

Power   The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
N   The total number of subjects. They are divided evenly among all sequences.
$\|\mu_1 - \mu_2\|$ The mean difference at which the power is computed. $\|\mu_1 - \mu_2\|$ = treatment mean - reference mean.
σ_w The square root of the within mean square error from the ANOVA table.
Alpha   The probability of rejecting a true null hypothesis.

Summary Statements
───

A three-period, two-sequence dual cross-over design (ABB | BAA) will be used to test whether the difference in means ($\mu_1 - \mu_2$) is different from 0. The comparison will be made using a two-sided t-test, with a Type I error rate (α) of 0.05. The within-subject standard deviation is assumed to be 25. To detect a difference in means ($\mu_1 - \mu_2$) of 14, with a total sample size of 6 (allocated equally to the 2 sequences), the power is 0.1348.
───

Dropout-Inflated Sample Size

Dropout Rate	Sample Size N	Dropout-Inflated Enrollment Sample Size N'	Expected Number of Dropouts D
20%	6	8	2
20%	16	20	4
20%	26	33	7
20%	36	45	9
20%	46	58	12
20%	56	70	14
20%	66	83	17

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, 8 subjects should be enrolled to obtain a final sample size of 6 subjects.

References

Chow, S.C. and Liu, J.P. 1999. Design and Analysis of Bioavailability and Bioequivalence Studies. Marcel Dekker. New York

Chow, S.C., Shao, J., and Wang, H. 2003. Sample Size Calculations in Clinical Research. Marcel Dekker. New York.

Chen, K.W., Chow, S.C., and Li, G. 1997. 'A Note on Sample Size Determination for Bioequivalence Studies with Higher-Order Crossover Designs.' Journal of Pharmacokinetics and Biopharmaceutics, Volume 25, No. 6, pages 753-765.

This report shows the power for the indicated scenarios.

Plots Section

Plots
_____
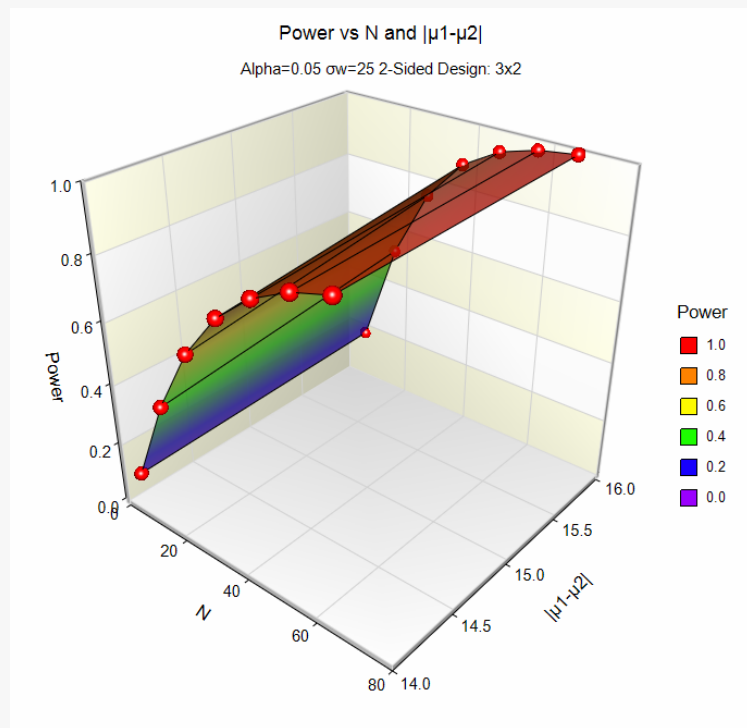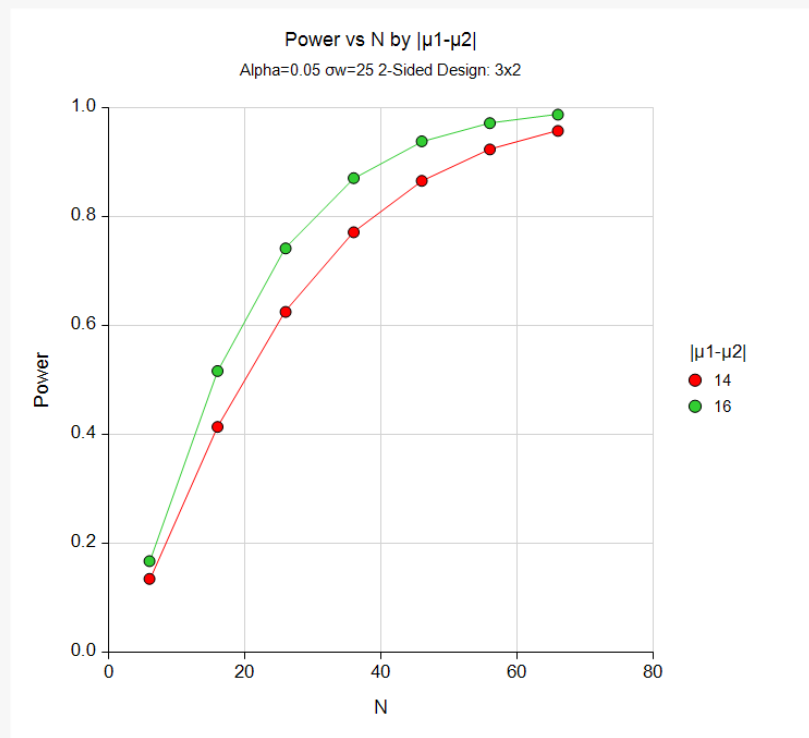
These plots show the power versus the sample size.

Example 2 – Finding Sample Size

Continuing with Example 1, the researchers want to find the exact sample size needed to achieve 90% power.

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 (Exact)**
Design Type...**3x2 (Three-Period, Two-Sequence Dual: ABB|BAA)**
Alternative Hypothesis ....**Two-Sided**
Power..**0.90**
Alpha..**0.05**
|μ1 - μ2|...**14 16**
Specify σ as σw or σb and ρ ....**σw (Within Std Dev)**
σw (Within Standard Dev).............................**25**

Output

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

Numeric Results
_____

Solve For: Sample Size (Exact)
Design Type: Three-Period, Two-Sequence Dual
Treatment Sequences: ABB | BAA
Alternative Hypothesis: Two-Sided
_____

Power	Total Sample Size N	Mean Difference |μ1 - μ2|	Within Standard Deviation σw	Alpha
0.9039	52	14	25	0.05
0.9035	40	16	25	0.05
_____

This report shows the sample sizes necessary when the difference is 14 or 16. Note that the N of 52 would be split with 26 devoted to each sequence.

Example 3 – Validation

We will validate this procedure by comparing its results to the *Superiority by a Margin Tests for the Difference of Two Means in a Higher-Order Cross-Over Design* routine which has already been validated. The two routines should achieve the same power when SM is set to zero in that routine.

In the superiority routine, if we set the Design Type to 4x2, Higher Means to Better, Alpha to 0.05, N to 50, SM to 0, Difference to 1.5, and Within Std Dev to 4, the power is computed to be 0.8079.

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 ....**Power**
Design Type..**4x2 (Four-Period, Two-Sequence: ABBA|BAAB)**
Alternative Hypothesis ....**One-Sided**
Alpha...**0.05**
N **50**
|μ1 - μ2|...**1.5**
Specify σ as σw or σb and ρ..........................**σw (Within Std Dev)**
σw (Within Standard Dev).............................**4**

Output

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

Numeric Results
───
Solve For: Power
Design Type: Four-Period, Two-Sequence
Treatment Sequences: ABBA | BAAB
Alternative Hypotheses: One-Sided
───

Power	Total Sample Size N	Mean Difference $\lvert \mu 1 - \mu 2 \rvert$	Within Standard Deviation σw	Alpha
0.8079	50	1.5	4	0.05

───

This procedure has also calculated the power to be 0.8079.