

Chapter 354

Tests for the Difference Between Treatment and Control Means in a Balanced Single-Case (AB)^K Design with Multiple Cases

Introduction

This procedure calculates power and sample size for tests that compare treatment and control means in single-case (AB)^K designs with multiple cases. Single-case experiments are also referred to as N-of-1 designs. These designs used when various treatments (conditions) are administered to an individual over time. The power and sample size calculations are given in Hedges, Shadish, and Batley (2022).

Technical Details

The Balanced (AB)^K Design

The following is a summary of details found in Hedges, Shadish, and Batley (2022).

In the balanced (AB)^K design, A refers to a baseline (control or reference) phase, B refers to a treatment phase, and K indicates the number of times the AB pair is repeated. During each phase, M observations are taken. For example, (AB)² indicates an ABAB design in which 4M observations are taken on each subject. 2M of these are control measurements and 2M are treatment measurements (i.e., the design is balanced with 2M measurements in each phase).

Suppose the experiment includes N subjects. The test statistic compares the 2M treatment measurements with the 2M control measurements by computing the average difference (*D*) between the means of these two types of measurements and dividing that difference by a measure of the standard deviation of that difference (*S*). The effect size estimate of this procedure is

$$T = \frac{D}{S}$$

The corresponding parameter effect size, the standardized mean difference, is defined as

$$\delta_{ST} = \frac{\mu_B - \mu_A}{\sqrt{\tau^2 + \sigma^2}}$$

where μ_B and μ_A are the treatment and control means, respectively, τ^2 is the between-subject variance, and σ^2 is the within-subject variance.

Tests for the Diff. Between Treatment and Control Means in a Balanced Single-Case (AB)^k Design with Multiple Cases

Note that Hedges, Shadish, and Batley (2022) uses M for the number of subjects and N for the number of subjects per phase, but **PASS** uses N for the number of subjects and M for the number of subjects per phase for consistency with the notation in other procedures.

Hypothesis Testing

The three sets of statistical hypotheses commonly used are

$$H_0: \delta_{ST} \leq 0 \quad \text{vs.} \quad H_1: \delta_{ST} > 0$$

$$H_0: \delta_{ST} \geq 0 \quad \text{vs.} \quad H_1: \delta_{ST} < 0$$

$$H_0: \delta_{ST} = 0 \quad \text{vs.} \quad H_1: \delta_{ST} \neq 0$$

Define the intraclass correlation as

$$\rho = \frac{\tau^2}{\tau^2 + \sigma^2}$$

This gives the proportion of the total variance of an observation that is between subjects.

Using matrix notation, let the 2KM observations of the i^{th} subject be represent by the row-vector, \mathbf{y}_i , as

$$\mathbf{y}_i = (y_{i1}, \dots, y_{i(2KM)})$$

Note that these observations are arranged according to the measurement time. Hence, the first M values occur while the A treatment is active, the next M observations occur while the B treatment is active, and so on, alternating between A and B.

Let \mathbf{w} be a contrast vector that corresponds to \mathbf{y}_i , with

$$\mathbf{w} = \frac{1}{2KM} (1'_M, -1'_M, 1'_M, -1'_M, \dots, 1'_M, -1'_M)'$$

where $1'_M$ is an M-dimensional row vector of 1's. The value of the within-person contrast is $\mathbf{y}_i \mathbf{w}$, which has variance $\mathbf{w}' \mathbf{V}_i \mathbf{w}$, where

$$\mathbf{V}_i = \frac{\sigma^2}{1 - \varphi^2} \begin{pmatrix} 1 & \varphi & \varphi^2 & \dots & \varphi^{2KM-1} \\ \varphi & 1 & \varphi & \dots & \varphi^{2KM-2} \\ \varphi^2 & \varphi & 1 & \dots & \varphi^{2KM-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \varphi^{2KM-1} & \varphi^{2KM-2} & \varphi^{2KM-3} & \dots & 1 \end{pmatrix}$$

is the covariance matrix of \mathbf{y}_i , and φ is the first-order autocorrelation between adjacent observations on a subject.

The variance of the effect size estimator, D , is given by

$$V(D) = \frac{\mathbf{w}' \mathbf{V}_i \mathbf{w}}{N \sigma^2}$$

Tests for the Diff. Between Treatment and Control Means in a Balanced Single-Case (AB)^k Design with Multiple Cases

The standardized expectation and variance of the sample covariance matrix, S^2 , are given by

$$E(S^2) = \frac{\text{tr}(A\Sigma)}{2KM(N-1)(\tau^2 + \sigma^2)}$$

$$V(S^2) = \frac{2\text{tr}(A\Sigma A\Sigma)}{(2KM(N-1)(\tau^2 + \sigma^2))^2}$$

where

$$A = \begin{pmatrix} A_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A_{(2KM)(2KM)} \end{pmatrix}$$

$$A_{ii} = I_N - 1_N 1_N / N$$

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \cdots & \Sigma_{1(2KM)} \\ \vdots & \ddots & \vdots \\ \Sigma_{(2KM)1} & \cdots & \Sigma_{(2KM)(2KM)} \end{pmatrix}$$

$$\Sigma_{ij} = \left(\tau^2 + \frac{\sigma^2 \varphi^{|i-j|}}{1 - \varphi^2} \right) I_N$$

where $\text{tr}(X)$ is the trace of the square matrix, X .

Using the above, the sampling distribution of the effect size, T , is

$$t'_{h,\lambda} = \left(\frac{D}{S} \right) \sqrt{\frac{E(S^2)}{V(D)}}$$

which is a noncentral t distribution with h degrees of freedom and noncentrality parameter, λ , where

$$h = \frac{2[E(S^2)]^2}{V(S^2)}$$

$$\lambda = \delta_{ST} \sqrt{\frac{E(S^2)}{V(D)}}$$

Power Calculation

Using the above results the power of a two-tailed test at significance level α is

$$\text{Power} = 1 - f(c_{\alpha/2}|h, \lambda) + f(-c_{\alpha/2}|h, \lambda)$$

where $f(c_{\alpha/2}|h, \lambda)$, the cumulative distribution function of a noncentral t distribution with h degrees of freedom and noncentrality parameter, λ , and $c_{\alpha/2}$ is the appropriate quantile from the central t distribution with h degrees of freedom. This formula can be used in a search algorithm to determine the necessary sample size.

Note that the R code attached to Hedges et al. (2022) use ρ for τ^2 and $1 - \rho$ for σ^2 .

Example 1 – Finding Sample Size

Researchers are considering an (AB)^K design and want to determine the number of subjects necessary to obtain a power of 0.90 when alpha = 0.05. They want to consider K = 2 or 3 (ABAB versus ABABAB) when the effect size is 0.6, 0.7, or 0.8. The intraclass correlation and the autocorrelation are both set to 0.5. The number of observations per phase is 3.

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 **Sample Size**
 Alternative Hypothesis **Two-Sided (H1: $\delta_{ST} \neq 0$)**
 Power..... **0.90**
 Alpha..... **0.05**
 K (Number of AB Repetitions)..... **2 3**
 M (Number of Observations per Phase) **3**
 δ_{ST} (Standardized Mean Difference)..... **0.6 0.7 0.8**
 ρ (Intraclass Correlation)..... **0.5**
 ϕ (Autocorrelation)..... **0.5**

Output

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

Numeric Reports

Numeric Results

Solve For: [Sample Size](#)
 Hypotheses: H0: $\delta_{ST} = 0$ vs. H1: $\delta_{ST} \neq 0$
 A: Baseline or Control Phase
 B: Treatment Phase

Power		Number of Subjects N	Number of AB Pairs K	Number of Observations		Standardized Mean Difference δ_{ST}	Intraclass Correlation ρ	First-Order Autocorrelation ϕ	Alpha α
				In Each Phase M	In Each Subject 2KM				
Target	Actual								
0.9	0.91610	14	2	3	12	0.6	0.5	0.5	0.05
0.9	0.90221	10	2	3	12	0.7	0.5	0.5	0.05
0.9	0.90818	8	2	3	12	0.8	0.5	0.5	0.05
0.9	0.91716	9	3	3	18	0.6	0.5	0.5	0.05
0.9	0.92503	7	3	3	18	0.7	0.5	0.5	0.05
0.9	0.94517	6	3	3	18	0.8	0.5	0.5	0.05

Tests for the Diff. Between Treatment and Control Means in a Balanced Single-Case (AB)^k Design with Multiple Cases

Power	The power of a t-test of the standardized mean difference of the treatment and control groups. Power is the probability of rejecting a false null hypothesis that the means are equal when they are not.
Target	The target power that is sought by the sample size (N) search.
Actual	The actual power that is achieved by this sample size (N).
N	The number of subjects in the experiment.
K	The number of times the AB pair is repeated within a subject. For example, (AB) ² = ABAB and (AB) ³ = ABABAB.
M	The number of observations per phase (a control or treatment replication) within a subject.
2KM	The number of observations per subject.
δ_{ST}	The standardized difference between treatment and control means, μ_B and μ_A , respectively, at which the power is calculated. $\delta_{ST} = (\mu_B - \mu_A) / \sqrt{(\tau^2 + \sigma^2)}$, where τ^2 is the between-subject variance and σ^2 is the within-subject variance.
ρ	The intraclass correlation. $\rho = \tau^2 / (\tau^2 + \sigma^2)$, where τ^2 is the between-subject variance and σ^2 is the within-subject variance.
ϕ	The first-order autocorrelation between adjacent observations on the same subject.
Alpha	The probability of rejecting a true null hypothesis.

Summary Statements

A balanced single-case (AB)² design with multiple cases (where each subject is measured repeatedly, alternating between control (A) and treatment (B) phases) will be used to test whether the standardized mean difference ($\delta_{ST} = (\mu_B - \mu_A) / \sqrt{(\tau^2 + \sigma^2)}$) is different from 0 (H₀: $\delta_{ST} = 0$ versus H₁: $\delta_{ST} \neq 0$). Each subject will be measured 3 times in each control (A) phase and treatment (B) phase, and there will be 2 AB phase combination repetitions (for a total of 12 measurements per subject). The comparison will be made using a two-sided t-test with a Type I error rate (α) of 0.05. The intraclass correlation of observations within a subject (ρ) is assumed to be 0.5, and the first-order autocorrelation (ϕ) is assumed to be 0.5. To detect a standardized mean difference (δ_{ST}) of 0.6 with 90% power, the number of subjects needed will be 14.

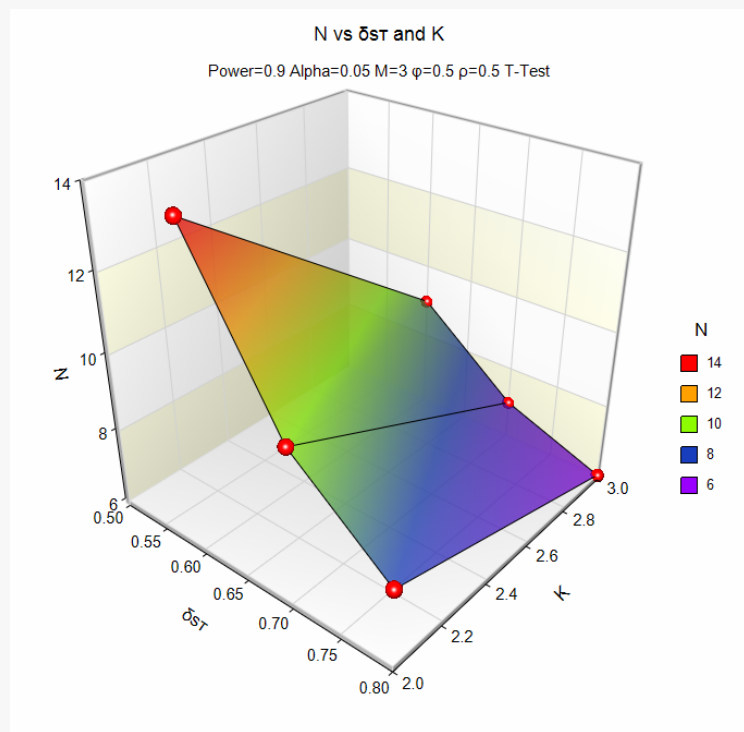
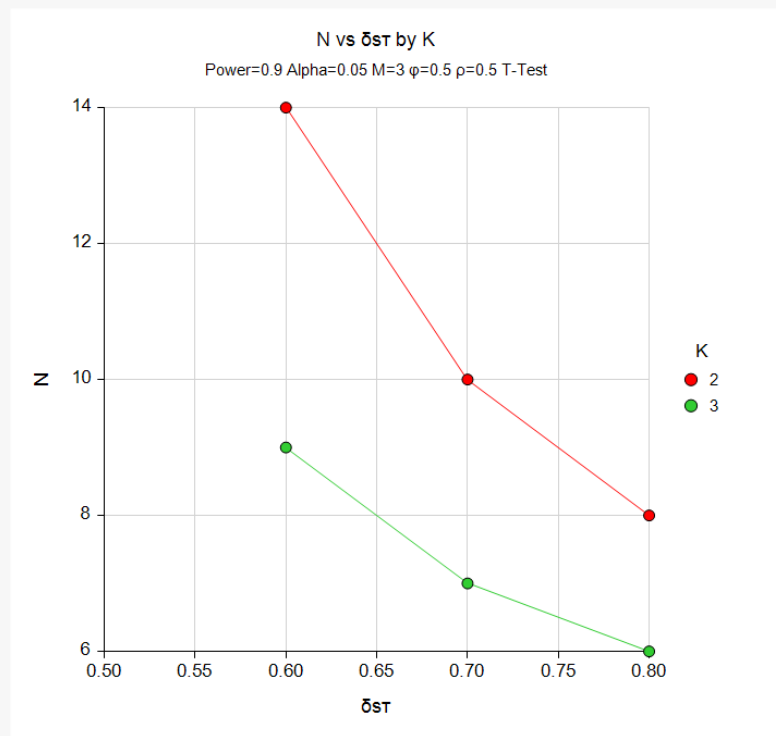
References

- Hedges, L.V., Shadish, W.R., Batley, P.N. 2022. 'Power analysis for single-case designs: Computations of (AB)^k designs.' Behavior Research Methods. DOI: 10.3758/s13428-022-01971-9.
- Hedges, L.V., Pustejovsky, J.E., Shadish, W.R. 2012. 'A standardized mean difference effect size for single case designs' Research Synthesis Methods. DOI: 10.1002/jrsm.1052.
-

This report shows the sample sizes for the indicated scenarios.

Plots Section

Plots



These plots present the required sample sizes for various scenarios.

Example 2 – Validation using Hedges et al. (2022)

Hedges et al. (2022) presents an example on page 8 of an (AB)^K design. The example determines the power when alpha = 0.05, K = 2, effect size is 0.75, intraclass correlation = 0.5, autocorrelation = 0.5, and M = 3. They use a two-sided test and find the power to be 0.38.

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 **Power**
 Alternative Hypothesis **Two-Sided (H1: $\delta_{ST} \neq 0$)**
 Alpha..... **0.05**
 N (Number of Subjects) **3**
 K (Number of AB Repetitions)..... **2**
 M (Number of Observations per Phase) **3**
 δ_{ST} (Standardized Mean Difference)..... **0.75**
 ρ (Intraclass Correlation)..... **0.5**
 ϕ (Autocorrelation)..... **0.5**

Output

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

Numeric Results

Solve For: [Power](#)
 Hypotheses: H0: $\delta_{ST} = 0$ vs. H1: $\delta_{ST} \neq 0$
 A: Baseline or Control Phase
 B: Treatment Phase

Power	Number of Subjects N	Number of AB Pairs K	Number of Observations		Standardized Mean Difference δ_{ST}	Intraclass Correlation ρ	First-Order Autocorrelation ϕ	Alpha α
			In Each Phase M	In Each Subject 2KM				
0.38452	3	2	3	12	0.75	0.5	0.5	0.05

PASS also calculates a power of 0.38 which validates the procedure.