Chapter 481

Tests for Two Means in a Multicenter Randomized Design

Introduction

In a multicenter design with a continuous outcome, a number of centers (e.g., hospitals or clinics) are selected at random from a population of centers. The subjects in each center are then randomized to either of two treatments.

The data are analyzed using a mixed effects model that includes a fixed treatment effect and a random center effect. The test of interest is the F-test of the treatment effect. This is a two-sided test of the hypothesis that there is no treatment effect. Note that the treatment-by-center interaction is not included in this model, although it could be in a secondary analysis.

Technical Details

These results come from Vierron and Giraudeau (2007). Consider a mixed effect model for a two-way layout (treatment and class) without interaction.

$$Y_{ijk} = \mu + \delta_i + C_j + \varepsilon_{ijk} \qquad i = 1,2, \quad j = 1, \dots, Q, \quad k = 1, \dots, n$$

where Y_{ijk} is the (continuous) response of the *k*th subject, receiving the *i*th treatment in the *j*th center. The overall response is μ. The treatment effects δ_1 and δ_2 are fixed constants where $\delta_1 = \delta/2$ and $\delta_2 = -\delta/2$. Hence, the two treatment means are $\mu_1 = \mu + \delta/2$ and $\mu_2 = \mu - \delta/2$, so that $\delta = \mu_1 - \mu_2$.

The class effects C_i are random, distributed normally with mean zero and variance σ_C^2. The errors ε_{ijk} are distributed normally with a mean of zero and a variance of σ_ε^2. It is assumed that the errors and the class effects are independent. The authors make the simplifying assumption that the sample size from all centers are equal to 2*n* and that the number of subjects assign to each treatment is the same. Thus, the total sample size *N* is 2*Qn*. Note that the authors conduct a number of simulation studies and conclude that when the assumptions of equal class size and equal treatment size are violated, their results are still very close.

In the sample size formula to follow, the center effects are recast as the intraclass correlation coefficient, ρ, which is defined as the proportion of the total variation in *Y* that is accounted for by the variance in the centers. Symbolically, ρ (or *ICC*) is defined as

$$\rho = \frac{\sigma_C^2}{\sigma_C^2 + \sigma_\varepsilon^2}$$

Note that the variance of the responses σ^2 is given by

$$\sigma^2 = \sigma_Y^2 = \sigma_C^2 + \sigma_\varepsilon^2$$

© NCSS, LLC. All Rights Reserved.

Vierron and Giraudeau (2007) show that for this situation, the power is given by

$$Power = \Phi\left(\frac{\delta\sqrt{N}}{\sigma\sqrt{1-\rho}}\right) - z_{1-\alpha/2}$$

where $\Phi(x)$ is the cumulative standard normal distribution function and z_ϕ is found so that $\phi = \Phi(z_\phi)$.

Note that the power does not depend on the number of centers in the study or their size. The formula makes the assumption that there are several centers and the number of subjects from each center is almost uniform.

© NCSS, LLC. All Rights Reserved.

Example 1 – Calculating Sample Size

Suppose that a study is to be conducted in which alpha = 0.05; power = 0.90; *mean difference* = 0.1, 0.2, or 0.3; σ = 1.0; and ρ = 0.1. Sample size is to be solved for.

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**
Power..**0.90**
Alpha...**0.05**
μ1 - μ2 (Mean Difference).............................**0.1 0.2 0.3**
σ (Response Standard Deviation).................**1.0**
ρ (Intracluster Correlation) **0.10**

Output

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

Numeric Reports

Numeric Results

Solve For: Sample Size
Test Type: Mixed Effects F-Test

Power	Total Sample Size N	Mean Difference μ1 - μ2	Standard Deviation Response σ	Standard Deviation Center σc	Standard Deviation Error σε	Intraclass Correlation Coefficient ρ	Alpha
0.9000	3783	0.1	1	0.316	0.949	0.1	0.05
0.9001	946	0.2	1	0.316	0.949	0.1	0.05
0.9005	421	0.3	1	0.316	0.949	0.1	0.05

Power The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
N The total number of subjects in both groups and in all centers.
μ1 - μ2 The difference between the group means at which the power is calculated.
σ The standard deviation of the response variable.
σc Center Standard Deviation. The square root of the center-to-center variance.
σε Error Standard Deviation. The square root of the subject-to-subject variance calculated within a center.
ρ Intraclass Correlation Coefficient. The proportion that the center-to-center variance is of the response variance.
Alpha The probability of rejecting a true null hypothesis.

© NCSS, LLC. All Rights Reserved.

Summary Statements

A multicenter randomized design (where subjects within each center are randomized to the two treatments) will be used to test whether there is a difference in means (H0: μ1 - μ2 = 0 versus H1: μ1 - μ2 ≠ 0). The comparison will be made using a two-sided F-test of the treatment term in a mixed effects model fit without the treatment-by-center interaction, with a Type I error rate (α) of 0.05. The standard deviation of the response is assumed to be 1. The intraclass correlation coefficient (proportion of center-to-center variance to the response variance) is assumed to be 0.1. To detect a difference in means of 0.1 with 90% power, the total number of needed subjects (combined for both treatments and all centers) will be 3783.

Dropout-Inflated Sample Size

Dropout Rate	Sample Size N	Dropout-Inflated Enrollment Sample Size N'	Expected Number of Dropouts D
20%	3783	4729	946
20%	946	1183	237
20%	421	527	106

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. 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. After solving for N, 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, 4729 subjects should be enrolled to obtain a final sample size of 3783 subjects.

References

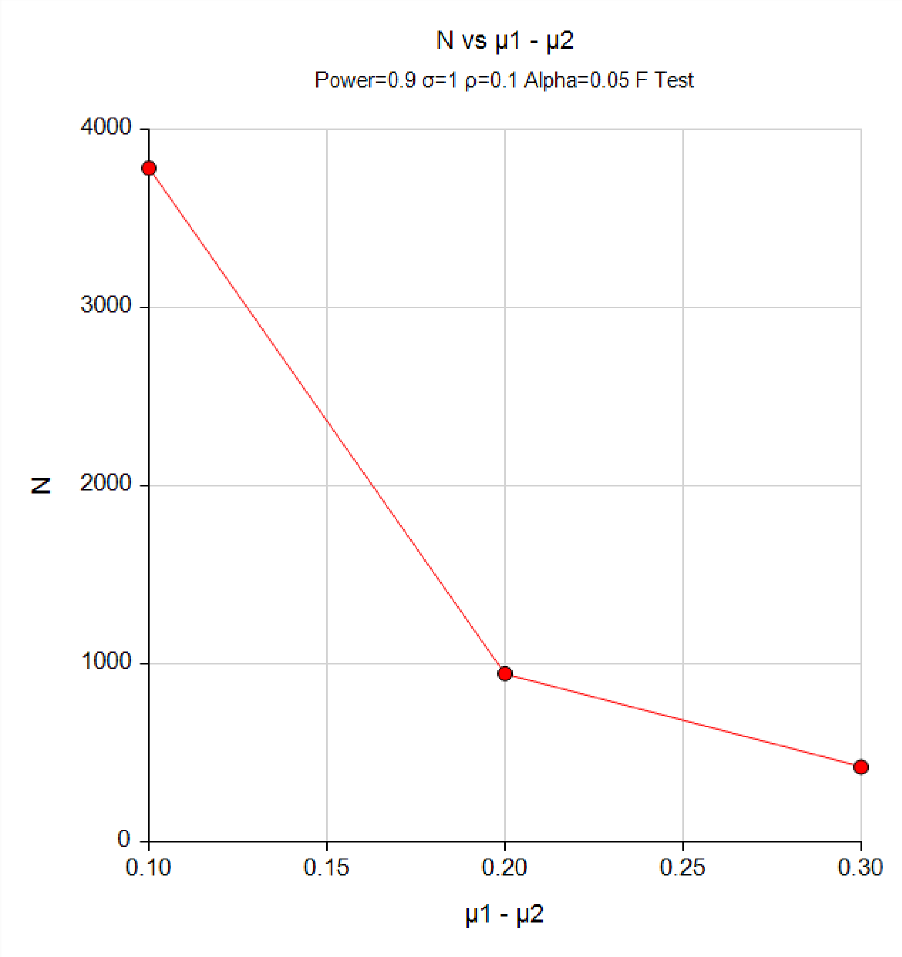
Vierron, E. and Giraudeau, B. 2007. 'Sample size calculation for multicenter randomized trial: Taking the center effect into account.' Contemporary Clinical Trials, Vol 28. Pages 451-458.
Vierron, E. and Giraudeau, B. 2009. 'Design effect in multicenter studies: gain or loss of power?' BMC Medical Research Methodology, 9:39. This article is available from www.biomedcentral.com/1471-2288/9/39

This report shows the required sample size for each of the scenarios.

© NCSS, LLC. All Rights Reserved.

Plots Section

Plots



This plot shows the sample size versus the difference in means.

© NCSS, LLC. All Rights Reserved.

Example 2 – Validation using Vierron and Giraudeau (2007)

Vierron and Giraudeau (2007) page 454, Table 2 gives N = 302 when *alpha* = 0.05; *power* = 0.80; *mean difference* = 0.25; σ = 1.0; and ρ = 0.4.

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.. **0.80**
Alpha... **0.05**
μ1 - μ2 (Mean Difference)............................. **0.25**
σ (Response Standard Deviation)................. **1.0**
ρ (Intracluster Correlation) **0.40**

Output

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

Numeric Results

Solve For: Sample Size
Test Type: Mixed Effects F-Test

Power	Total Sample Size N	Mean Difference μ1 - μ2	Standard Deviation Response σ	Standard Deviation Center σc	Standard Deviation Error σε	Intraclass Correlation Coefficient ρ	Alpha
0.8008	302	0.25	1	0.632	0.775	0.4	0.05

PASS also calculates the sample size to be 302.

© NCSS, LLC. All Rights Reserved.