Chapter 386

Mixed Models Tests for Two Proportions in a 2-Level Hierarchical Design (Level-2 Randomization)

Introduction

This procedure calculates power and sample size for a two-level hierarchical mixed-effects logistic regression in which clusters (groups, classes, hospitals, etc.) of subjects are measured one time (cross-sectional) on a binary variable. The goal of the study is to compare the two group proportions.

In this design, the subjects are the level-1 units, and the clusters are the level-2 units. All subjects in a particular cluster (level-2 unit) receive one of two possible interventions. This intervention is selected at random. Note that a companion procedure power analyzes the other case in which the randomization occurs for the level-1 units (the subjects).

Technical Details

Our formulation comes from Ahn, Heo, and Zhang (2015), chapter 5, section 5.7.1, pages 177-179. The hierarchical mixed-effects logistic regression model that is adopted is

$$\log\left(\frac{p_{ij}}{1-p_{ij}}\right) = \beta_0 + \delta X_{ij} + u_i$$

where

Y_{ij} is the binary response of the j^{th} subject in the i^{th} cluster.

p_{ij} is $\mathrm{E}(Y_{ij}|X_{ij})$. Assume $[p_{ij}|(X_{ij}=0)] = p_2$ and $[p_{ij}|(X_{ij}=1)] = p_1$

β_0 is the fixed intercept.

δ is the treatment effect of interest.

X_{ij} is an indicator variable that is = 1 if cluster *i* is assigned to group 1 (treatment) and 0 if cluster *i* is assigned to group 2 (control).

u_i is a random effect (subject-specific intercept) term for the i^{th} cluster that is distributed as $N(0, \sigma_u^2)$.

σ_u^2 is variance of the level-2 (cluster) random effects.

ρ is the intraclass correlation (ICC), where $\rho = Corr(Y_{ij}, Y_{ij'}) = (\sigma_u^2/(\sigma_u^2 + \pi^2/3))$.

The test of significance of the δ coefficient in the logistic regression analysis is the test statistic of interest.

The power can be calculated using

$$Power = \Phi\left\{\frac{|p_1 - p_2|\sqrt{K_2 M/[1+(M-1)\rho]} - \Phi^{-1}(1-\alpha/2)\sqrt{\left(1+\frac{1}{\lambda}\right)\bar{p}(1-\bar{p})}}{\sqrt{p_2(1-p_2)+p_1(1-p_1)/\lambda}}\right\}$$

where

K_m is the number of clusters in group m, m = 1 or 2.

M is the average number of subjects per cluster in both groups.

λ is K_1/K_2.

$\bar{p}$ is $\left(\frac{p_1+\lambda p_2}{1+\lambda}\right)$.

This power function is used in a binary search algorithm to determine p_1, K_1, or M.

Example 1 – Calculating Power

Suppose that a two-level hierarchical design is planned in which there will be only one measurement per subject and treatments will be applied to clusters (level-two units). The analysis will be a mixed-effect logistic regression. The following parameter settings are to be used for the power analysis: $P1 = 0.6$; $P2 = 0.5$; $\rho = 0.01$; $M = 20$ or 30; $\alpha = 0.05$; and $K1 = K2 = 10$ to 40 by 10.

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**
Alpha	**0.05**
K1 (Number of Group 1 Clusters)	**10 20 30 40**
K2 (Number of Group 2 Clusters)	**K1**
M (Number of Subjects Per Cluster)	**20 30**
P1 Input Type	**Proportions**
P1 (Group 1 Proportion|H1)	**0.6**
P2 (Group 2 Proportion)	**0.5**
ρ (Intracluster Correlation, ICC)	**0.01**

Output

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

Numeric Reports

Numeric Results

Solve For: Power
Groups: 1 = Treatment, 2 = Control
Hypotheses: H0: P1 = P2 vs. H1: P1 ≠ P2

	Number of Subjects			Number of Clusters				Proportion				
Power	Group 1 N1	Group 2 N2	Total N	Group 1 K1	Group 2 K2	Total K	Number of Subjects per Cluster M	Group 1 P1	Group 2 P2	Difference P1 - P2	Intracluster Correlation ρ	Alpha
0.45306	200	200	400	10	10	20	20	0.6	0.5	0.1	0.01	0.05
0.58262	300	300	600	10	10	20	30	0.6	0.5	0.1	0.01	0.05
0.74190	400	400	800	20	20	40	20	0.6	0.5	0.1	0.01	0.05
0.86672	600	600	1200	20	20	40	30	0.6	0.5	0.1	0.01	0.05
0.89211	600	600	1200	30	30	60	20	0.6	0.5	0.1	0.01	0.05
0.96434	900	900	1800	30	30	60	30	0.6	0.5	0.1	0.01	0.05
0.95855	800	800	1600	40	40	80	20	0.6	0.5	0.1	0.01	0.05
0.99151	1200	1200	2400	40	40	80	30	0.6	0.5	0.1	0.01	0.05

Power	The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
N1, N2, and N	The number of subjects in groups 1, 2, and both, respectively.
K1, K2, and K	The number of clusters in groups 1, 2, and both, respectively.
M	The average number of subjects (items) per cluster.
P1	The proportion for group 1 (treatment group) assuming the alternative hypothesis.
P2	The proportion for group 2 (control group). This is the proportion in the standard, reference, baseline, or control group.
P1 - P2	The difference in the group proportions assumed by the alternative hypothesis.
ρ	The intracluster correlation (ICC). The correlation between any two subjects in the same cluster.
Alpha	The probability of rejecting a true null hypothesis.

Summary Statements

A 2-group 2-level hierarchical design will have level-1 units (e.g., students, subjects, or patients) in level-2 units (e.g., classes, clinics, hospitals, or clusters) with random assignment of level-2 units to each of the 2 groups (level-2 randomization). This design will be used to test the difference between two proportions, using the appropriate term of the hierarchical mixed-effects logistic regression model, with a Type I error rate (α) of 0.05. The correlation of level-1 units within a level-2 unit (intracluster correlation) is assumed to be 0.01. To detect a proportion difference (P1 - P2) of 0.1 (with P1 = 0.6 and P2 = 0.5), with 10 level-2 units in Group 1 and 10 level-2 units in Group 2, with 20 level-1 units in each level-2 unit (for a grand total of 400 level-1 units), the power is 0.45306.

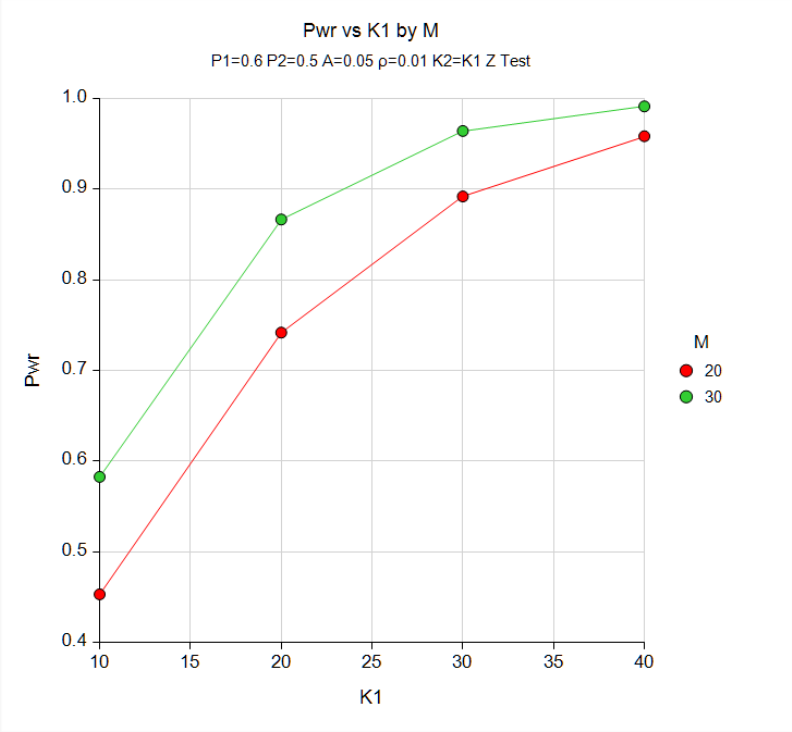
References

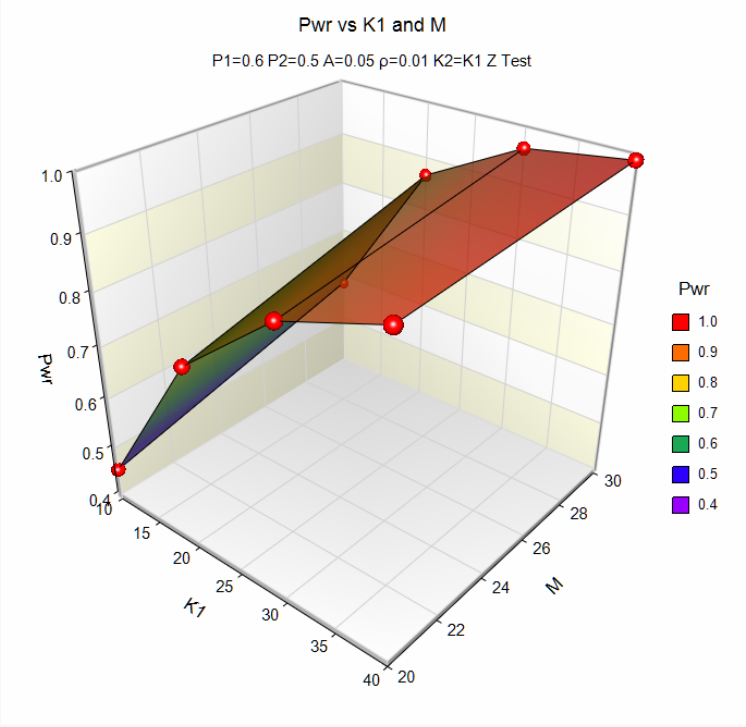
Ahn, C., Heo, M., and Zhang, S. 2015. Sample Size Calculations for Clustered and Longitudinal Outcomes in Clinical Research. CRC Press. New York.

This report shows the power for each of the scenarios.

Plots Section

Plots





These plots show the power versus the cluster count for the two cluster size values.

Example 2 – Calculating Sample Size (Number of Clusters)

Continuing with the last example, suppose the researchers want to determine the number of clusters needed to achieve 90% power for both values of M.

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	**K1 (Number of Clusters)**
Power	**0.90**
Alpha	**0.05**
K2 (Number of Group 2 Clusters)	**K1**
M (Number of Subjects Per Cluster)	**20 30**
P1 Input Type	**Proportions**
P1 (Group 1 Proportion|H1)	**0.6**
P2 (Group 2 Proportion)	**0.5**
ρ (Intracluster Correlation, ICC)	**0.01**

Output

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

Numeric Results

Solve For: K1 (Number of Clusters)
Groups: 1 = Treatment, 2 = Control
Hypotheses: H0: P1 = P2 vs. H1: P1 ≠ P2

	Number of Subjects			Number of Clusters			Number of Subjects per Cluster	Proportion			Intracluster Correlation	
Power	Group 1 N1	Group 2 N2	Total N	Group 1 K1	Group 2 K2	Total K	M	Group 1 P1	Group 2 P2	Difference P1 - P2	ρ	Alpha
0.90162	620	620	1240	31	31	62	20	0.6	0.5	0.1	0.01	0.05
0.90890	690	690	1380	23	23	46	30	0.6	0.5	0.1	0.01	0.05

This report shows the power for each of the scenarios.

Example 3 – Calculating Sample Size (Number of Subjects per Cluster)

Continuing with the last example, suppose the researchers want to determine the number of subjects per cluster needed to achieve 90% power for all values of K1 and K2.

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	**M (Number of Subjects Per Cluster)**
Power	**0.90**
Alpha	**0.05**
K1 (Number of Group 1 Clusters)	**10 20 30 40**
K2 (Number of Group 2 Clusters)	**K1**
P1 Input Type	**Proportions**
P1 (Group 1 Proportion|H1)	**0.6**
P2 (Group 2 Proportion)	**0.5**
ρ (Intracluster Correlation, ICC)	**0.01**

Output

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

Numeric Results

Solve For: M (Number of Subjects Per Cluster)
Groups: 1 = Treatment, 2 = Control
Hypotheses: H0: P1 = P2 vs. H1: P1 ≠ P2

	Number of Subjects			Number of Clusters			Number of Subjects per Cluster	Proportion			Intracluster Correlation	
Power	Group 1 N1	Group 2 N2	Total N	Group 1 K1	Group 2 K2	Total K	M	Group 1 P1	Group 2 P2	Difference P1 - P2	ρ	Alpha
0.90076	1070	1070	2140	10	10	20	107	0.6	0.5	0.1	0.01	0.05
0.90237	700	700	1400	20	20	40	35	0.6	0.5	0.1	0.01	0.05
0.90377	630	630	1260	30	30	60	21	0.6	0.5	0.1	0.01	0.05
0.90447	600	600	1200	40	40	80	15	0.6	0.5	0.1	0.01	0.05

This report shows the values of M needed for each scenario.

Example 4 – Validation using Ahn, Heo, and Zhang (2015)

Ahn, Heo, and Zhang (2015) page 179 provide a table in which several scenarios are reported. We will validate this procedure by duplicating the first two table entries.

The following parameter settings were used: Power = 0.80; *P1* = 0.6; *P2* = *0.4*; ρ = 0.1; *M* = 10 and 20; *ICC* = 0.1; and α = 0.05. The reported values of *K*1 and *K*2 are 19 and 15. The realized power values are 0.812 and 0.826

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 4** 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	**K1 (Number of Clusters)**
Power	**0.80**
Alpha	**0.05**
K2 (Number of Group 2 Clusters)	**K1**
M (Number of Subjects Per Cluster)	**10 20**
P1 Input Type	**Proportions**
P1 (Group 1 Proportion|H1)	**0.6**
P2 (Group 2 Proportion)	**0.4**
ρ (Intracluster Correlation, ICC)	**0.1**

Output

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

Numeric Results

Solve For: K1 (Number of Clusters)
Groups: 1 = Treatment, 2 = Control
Hypotheses: H0: P1 = P2 vs. H1: P1 ≠ P2

	Number of Subjects			Number of Clusters			Number of Subjects per Cluster	Proportion			Intracluster Correlation	
Power	Group 1 N1	Group 2 N2	Total N	Group 1 K1	Group 2 K2	Total K	M	Group 1 P1	Group 2 P2	Difference P1 - P2	ρ	Alpha
0.81229	190	190	380	19	19	38	10	0.6	0.4	0.2	0.1	0.05
0.82529	300	300	600	15	15	30	20	0.6	0.4	0.2	0.1	0.05

PASS calculates the same values of K1: 19 and 15. Note that the realized power values are identical.

(We noticed two typos in Table 5.9, page 179. The last two entries for p1 are 0.5 and 0.6. We believe these should be 0.6 and 0.7.)