Chapter 252

Tests for Two Ordered Categorical Variables (Proportional Odds)

Introduction

This module computes power and sample size for tests of ordered categorical data such as Likert scale data. Assuming proportional odds, such data can be analyzed with a z-test, ordinal logistic regression, ordered logistic regression, ordered logit model, proportional odds model, or the Mann-Whitney test. The power and sample size formulae presented here are consistent with any of these analysis methods. The results used here were presented in a paper by Whitehead (1993). They are also mentioned in the book by Julious (2010) and Machin et al (2018).

Ordered categorical data often results from surveys such as a quality of life (QoL) survey in which responses are categories such as *very good, good, moderate, poor*. When there are only two categories, an analysis using two proportions should be used. When there are more than two responses, and those responses can be ordered, the techniques described in this chapter can be used.

Technical Details

Suppose a variable has K possible responses $C_1, \dots, C_K$. Further suppose that these categories can be ordered so that C_k is more desirable than C_j if $k < j$. Hence C_1 is the best outcome and C_K is the worst. This procedure compares the results from two groups which will be called control (1) and experimental (2). The number of respondents falling within the k^{th}category of the control group is labeled N_{1k}. The total number of subjects in the control group is N_1 and in the experimental group is N_2. The total sample size of the study is $N = N_1 + N_2$.

Let p_{2k} denote the probability that an individual in the experimental group gives response C_k, and let Q_{2k} be the probability of an outcome of C_k or better. Thus $Q_{2k} = \sum_{j=1}^{k} p_{2j}$.

Define p_1 and Q_{1k} similarly for the control group.

Define the odds ratio for a particular category as

$$OR_k = \frac{\frac{Q_{2k}}{1 - Q_{2k}}}{\frac{Q_{1k}}{1 - Q_{1k}}} \quad k = 1, \dots, K - 1.$$

This measures the advantage of the experimental group over the control group. An OR_k greater than one indicates that the experimental treatment (group 2) is better than the control treatment (group 1).

Proportional Odds

The proportional odds model assumes that all of the odds ratios OR_k are equal to a common value OR. That is, the *proportional odds* assumption is that $OR_1 = \cdots = OR_{K-1} = OR$. Thus, the whole pattern of response differences can be summarized by a single parameter.

The formulae to follow use the fact that efficient score *Y* is asymptotically normally distributed when $\text{Log}(OR)$ is small and *N* is large. The test statistic and power formulae are as follows.

$$Y = \frac{1}{N+1}\sum_{k=1}^{K} N_{1k}(L_{1k} - U_{1k})$$

$$Z = \frac{Y - \mu_Y}{\sigma_Y}$$

$$L_{1k} = N_{11} + \cdots + N_{1(k-1)},\ \ k = 2, \dots, K$$

$$L_{2k} = N_{21} + \cdots + N_{2(k-1)},\ \ k = 2, \dots, K$$

$$U_{1k} = N_{1(k+1)} + \cdots + N_{1K},\ \ k = 1, \dots, K-1$$

$$U_{2k} = N_{2(k+1)} + \cdots + N_{2K},\ \ k = 1, \dots, K-1$$

$$L_{11} = U_{1K} = 0$$

$$L_{21} = U_{2K} = 0$$

$$\mu_Y = \text{Log}(OR)V$$

$$\sigma_Y^2 = V$$

$$V = \frac{N_1 N_2 N}{3(N+1)^2}\left\{1 - \sum_{k=1}^{K}\left[\frac{N_{1k} + N_{2k}}{N}\right]^3\right\} \approx \frac{N_1 N_2 N}{3(N+1)^2}\left\{1 - \sum_{k=1}^{K}\left[\frac{p_{1k} + p_{2k}}{2}\right]^3\right\}$$

The null hypothesis $H_0\text{:}\ OR = 1$ can be tested against the alternative $H_1\text{:}\ OR \neq 1$ by computing *Z* and rejecting if *Z* is greater than $z_{\alpha/2}$. That is,

$$\text{P}(Z > z_\alpha | OR = 1) = \alpha/2$$

The power is the probability of rejecting a false null hypothesis, thus the power for a specified value $OR_R$is

$$Power = \text{P}\left(Z > z_{\frac{\alpha}{2}} | OR = OR_R\right) = 1 - \Phi\left(z_{\frac{\alpha}{2}} - OR_R\sqrt{V}\right)$$

If a one-sided test is needed, replace $\alpha/2$ with α.

Example 1 – Finding the Sample Size

Suppose a clinical trial is planned to compare the response to certain treatment. The subjects are divided into two groups: those that will receive the current treatment and those that will receive an experimental treatment. Three months after the administration of the treatment, the subjects rate their response as *very good*, *good*, *moderate*, or *poor*. Historically, the responses have been about 20% *very good*, 50% *good*, 20% *moderate*, and 10% *poor*.

The researchers want to consider a range of possible value of *OR* from 1.5 to 2.5. They want to look at the sample size requirements to achieve a power of 0.90. They want to set alpha to 0.05 and analyze the results with a two-sided test.

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**
Null Hypothesis .... **Two-Sided**
Power **0.90**
Alpha **0.05**
Group Allocation .... **Equal (N1 = N2)**
P1's Input Type .... **Enter P11, P12, ..., P1K Pattern**
P11, P12, ..., P1K Pattern **2 5 2 1**
Odds Ratio Input Type .... **OR (Odds Ratio)**
OR (Odds Ratio) **1.5 2 2.5**

Output

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

Numeric Reports

Numeric Results

Solve For: Sample Size
Groups: 1 = Control, 2 = Treatment or Experimental
Hypotheses: H0: OR = 1 vs. H1: OR ≠ 1

	Sample Size			Number of Categories	Odds Ratio	Category Proportions		
Power	N1	N2	N	K	OR	P1	P2	Alpha
0.9005	451	451	902	4	1.5	P1(1)	P2(1)	0.05
0.9010	155	155	310	4	2.0	P1(1)	P2(2)	0.05
0.9014	89	89	178	4	2.5	P1(1)	P2(3)	0.05

Item	Values
P1(1)	0.2, 0.5, 0.2, 0.1
P2(1)	0.27, 0.51, 0.15, 0.07
P2(2)	0.33, 0.49, 0.12, 0.05
P2(3)	0.38, 0.47, 0.1, 0.04

Power The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
N1 The number of subjects in group 1, the control group.
N2 The number of subjects in group 2, the experimental group.
N The total sample size. N = N1 + N2.
K The number of categories in the response variable.
OR The odds ratio. OR = odds2/odds1.
P1 Group 1 Category Proportions Set. Gives the name of the set containing the response proportions for each of the K categories in group 1, the control group.
P2 Group 2 Category Proportions Set. Gives the name of the set containing the response proportions for each of the K categories in group 2, the experimental group.
Alpha The probability of rejecting a true null hypothesis.

Summary Statements

A parallel two-group design with an ordered categorical (ordinal) response will be used to test whether the Group 1 (control) distribution is different from the Group 2 (experimental) distribution. The comparison will be made using a two-sided odds ratio Z-test, with a Type I error rate (α) of 0.05. A proportional odds model is assumed. The number of response categories is assumed to be 4. The response proportions in Group 1 (control) are assumed to be 0.2, 0.5, 0.2, 0.1. To detect an odds ratio (O2 / O1) of 1.5 (or Group 2 response proportions of 0.27, 0.51, 0.15, 0.07), with 90% power, the number of subjects needed will be 451 in Group 1 (control) and 451 in Group 2 (experimental).

Dropout-Inflated Sample Size

	Sample Size			Dropout-Inflated Enrollment Sample Size			Expected Number of Dropouts		
Dropout Rate	**N1**	**N2**	**N**	**N1'**	**N2'**	**N'**	**D1**	**D2**	**D**
20%	451	451	902	564	564	1128	113	113	226
20%	155	155	310	194	194	388	39	39	78
20%	89	89	178	112	112	224	23	23	46

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.
N1, N2, and N The evaluable sample sizes at which power is computed. If N1 and N2 subjects are evaluated out of the N1' and N2' subjects that are enrolled in the study, the design will achieve the stated power.
N1', N2', and N' The number of subjects that should be enrolled in the study in order to obtain N1, N2, and N evaluable subjects, based on the assumed dropout rate. After solving for N1 and N2, N1' and N2' are calculated by inflating N1 and N2 using the formulas N1' = N1 / (1 - DR) and N2' = N2 / (1 - DR), with N1' and N2' 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.)
D1, D2, and D The expected number of dropouts. D1 = N1' - N1, D2 = N2' - N2, and D = D1 + D2.

Dropout Summary Statements

Anticipating a 20% dropout rate, 564 subjects should be enrolled in Group 1, and 564 in Group 2, to obtain final group sample sizes of 451 and 451, respectively.

References

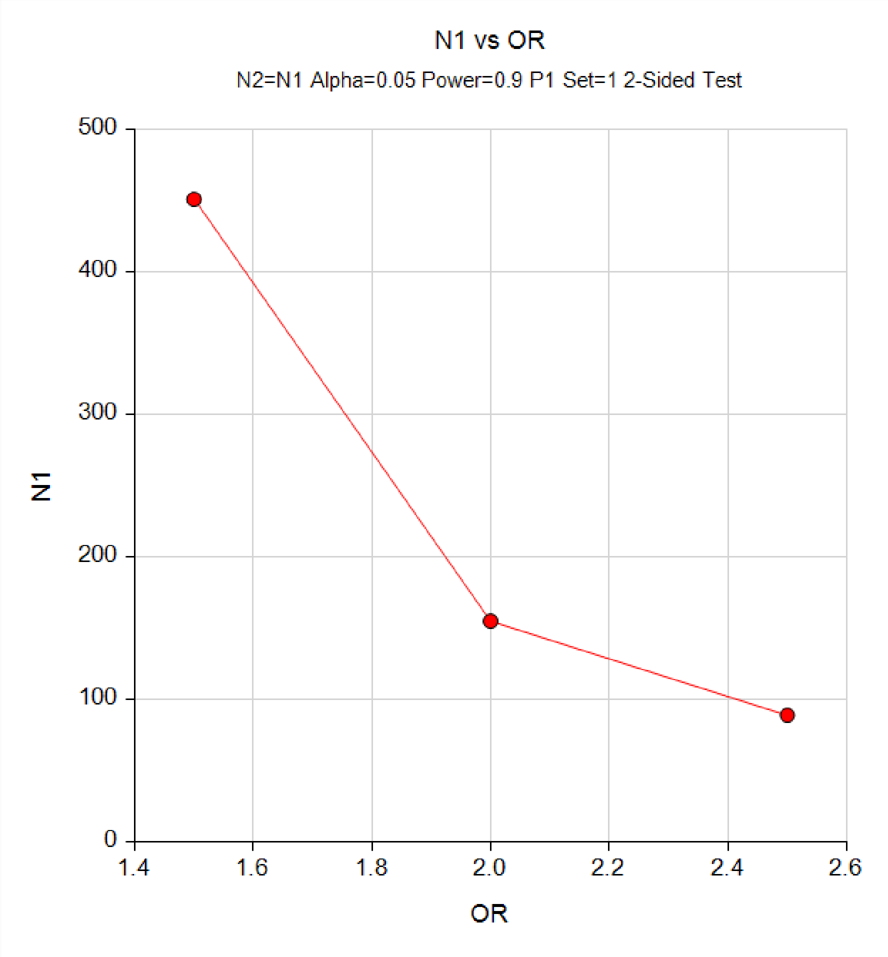
Whitehead, John. 1993. 'Sample Size Calculations for Ordered Categorical Data.' Statistics in Medicine, 12, 2257-2271.
Julious, Steven A. 2010. Sample Sizes for Clinical Trials. Chapman & Hall/CRC. Boca Raton, Fl.
Machin, D., Campbell, M., Tan, S.B., and Tan, S.H. 2018. Sample Size Tables for Clinical Studies, 4th Edition. John Wiley & Sons. Hoboken, NJ.

This report shows the numeric results of this sample size study. The definitions of the items on the report are given in the Report Definitions section.

Note the impact that the increasing odds ratios have had on the response probabilities. The proportion of those responding “Very Good” has risen from 0.2 to 0.38. The corresponding proportion of those responding “Poor” has fallen from 0.1 to 0.04.

Plots Section

Plots



This plot gives a visual presentation to the results in the Numeric Report.

Example 2 – Validation using Whitehead (1993)

Whitehead (1993) has an example in which he calculates the sample size to be 94 when *Log(OR)* is 0.887, alpha is 0.05, power is 90%, the control group proportions are 0.2 *very good*, 0.5 *good*, 0.2 *moderate*, and 0.1 *poor*.

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**
Null Hypothesis .... **Two-Sided**
Power.. **0.90**
Alpha... **0.05**
Group Allocation .... **Equal (N1 = N2)**
P1's Input Type .... **Enter P11, P12, ..., P1K Pattern**
P11, P12, ..., P1K Pattern............................. **2 5 2 1**
Odds Ratio Input Type .... **Log(OR) (Log Odds Ratio)**
Log(OR) (Log Odds Ratio)............................ **0.887**

Output

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

Numeric Results

Solve For: Sample Size
Groups: 1 = Control, 2 = Treatment or Experimental
Hypotheses: H0: Log(OR) = 0 vs. H1: Log(OR) ≠ 0

	Sample Size			Number of Categories	Log Odds Ratio	Category Proportions		
Power	N1	N2	N	K	Log(OR)	P1	P2	Alpha
0.9015	95	95	190	4	0.887	P1(1)	P2(1)	0.05

Item	Values
P1(1)	0.2, 0.5, 0.2, 0.1
P2(1)	0.378, 0.472, 0.106, 0.044

PASS found the required sample size as 95 per group. The increase from 94 to 95 was due to rounding. We found that the power for 94 was 0.8985, just less than the goal of 0.90. We note that the values of P2 match those in Whitehead (1993) exactly.

Example 3 – Calculating Sample Size for a COVID-19 Clinical Trial

This example will show how this procedure might be used in planning a clinical trial to assess the effectiveness of a treatment in combatting COVID-19. The study outcome on which the sample size will be based is the six-category ordinal scale of illness severity used in a recent trial.

The hypothetical trial that is being planned here will use the following six-point ordinal scale.

0) Discharge (alive).
1) Hospital admission, not requiring supplemental oxygen.
2) Hospital admission, requiring supplemental oxygen.
3) Hospital admission, requiring high-flow nasal cannula or non-invasive mechanical ventilation.
4) Hospital admission, requiring extracorporeal membrane oxygenation or invasive mechanical ventilation.
5) Death.

A recent trial provided the following response distribution for the placebo group at day 14: 23%, 13%, 36%, 10%, 9%, and 9%.

Using this response distribution for controls, determine the sample necessary to compare two groups consisting of a control group and a treatment group. Assume that twice as many subjects will be assigned to the treatment group as to the control group. Also assume that alpha is 0.05 and power is 0.90. This analysis will calculate the required sample sizes to detect odds ratios of 0.8, 0.7, and 0.6.

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**
Null Hypothesis	**Two-Sided**
Power	**0.90**
Alpha	**0.05**
Group Allocation	**Enter R = N2/N1, solve for N1 and N2**
R	**2**
P1's Input Type	**Enter P11, P12, ..., P1K Pattern**
P11, P12, ..., P1K Pattern	**23 13 36 10 9 9**
Odds Ratio Input Type	**OR (Odds Ratio)**
OR (Odds Ratio)	**0.8 0.7 0.6**

Output

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

Numeric Results

Solve For: Sample Size
Groups: 1 = Control, 2 = Treatment or Experimental
Hypotheses: H0: OR = 1 vs. H1: OR ≠ 1

	Sample Size			Allocation Ratio	Number of Categories	Odds Ratio	Category Proportions		
Power	N1	N2	N	R	K	OR	P1	P2	Alpha
0.9013	194	388	582	2	6	0.6	P1(1)	P2(1)	0.05
0.9006	397	794	1191	2	6	0.7	P1(1)	P2(2)	0.05
0.9001	1013	2026	3039	2	6	0.8	P1(1)	P2(3)	0.05

Item	Values
P1(1)	0.23, 0.13, 0.36, 0.1, 0.09, 0.09
P2(1)	0.152, 0.1, 0.354, 0.125, 0.126, 0.142
P2(2)	0.173, 0.11, 0.36, 0.118, 0.115, 0.124
P2(3)	0.193, 0.117, 0.363, 0.112, 0.105, 0.11

PASS found the required sample sizes for the three odds ratios. Note that N2 = 2(N1) in all cases as requested. Also note that the predicted distributions of responses anticipate a decrease in the value of P2 as the odds ratio decreases from 0.8 to 0.6.

Example 4 – COVID-19 Continued – Comparing Various Response Distributions

Continuing with Example 3, the researchers would like to analyze the impact of the control response probability distribution on the required sample size. To do this, the example will compare the results for the following distributional patterns:

All Equal:	1, 1, 1, 1, 1, 1
Pilot Data:	23, 13, 36, 10, 9, 9
Linear Decreasing:	6, 5, 4, 3, 2, 1
First Large:	15, 1, 1, 1, 1, 1

These patterns will be loaded in the spreadsheet. The spreadsheet will appear as follows:

Eq	PD	LD	FL
1	23	6	3
1	13	5	3
1	36	4	2
1	10	3	2
1	9	2	1
1	9	1	1

The rest of the parameters will remain the same.

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 .... **Sample Size**
Null Hypothesis.. **Two-Sided**
Power.. **0.90**
Alpha... **0.05**
Group Allocation .... **Enter R = N2/N1, solve for N1 and N2**
R **2**
P1's Input Type .... **Enter Columns Containing Sets of P1's**
Columns Containing Sets of P1's.................. **1-4**
Odds Ratio Input Type .... **OR (Odds Ratio)**
OR (Odds Ratio) **0.8 0.7 0.6**

Input Spreadsheet Data

Row	Eq	PD	LD	FL
1	1	23	6	15
2	1	13	5	1
3	1	36	4	1
4	1	10	3	1
5	1	9	2	1
6	1	9	1	1

Output

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

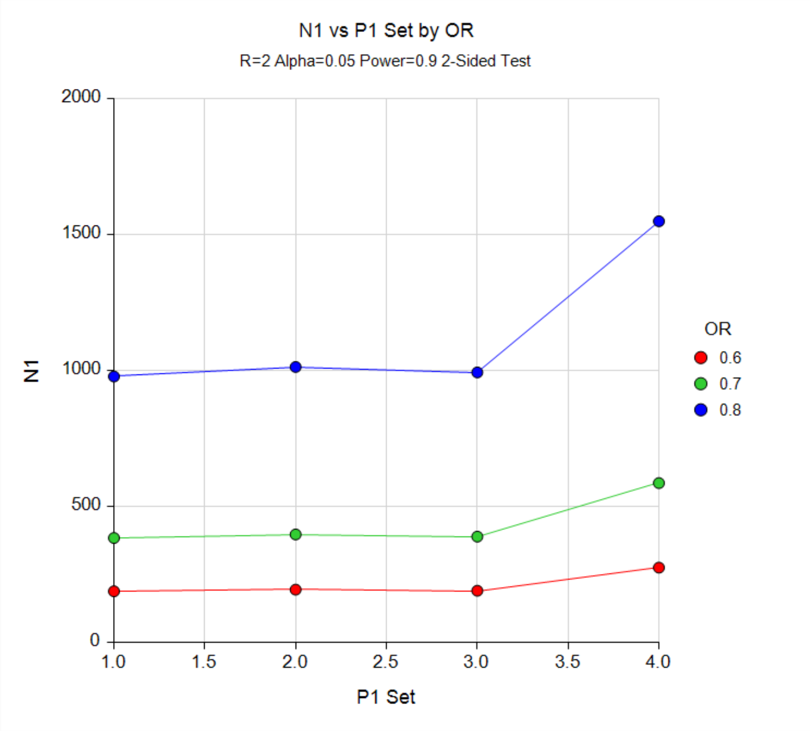
Numeric Results

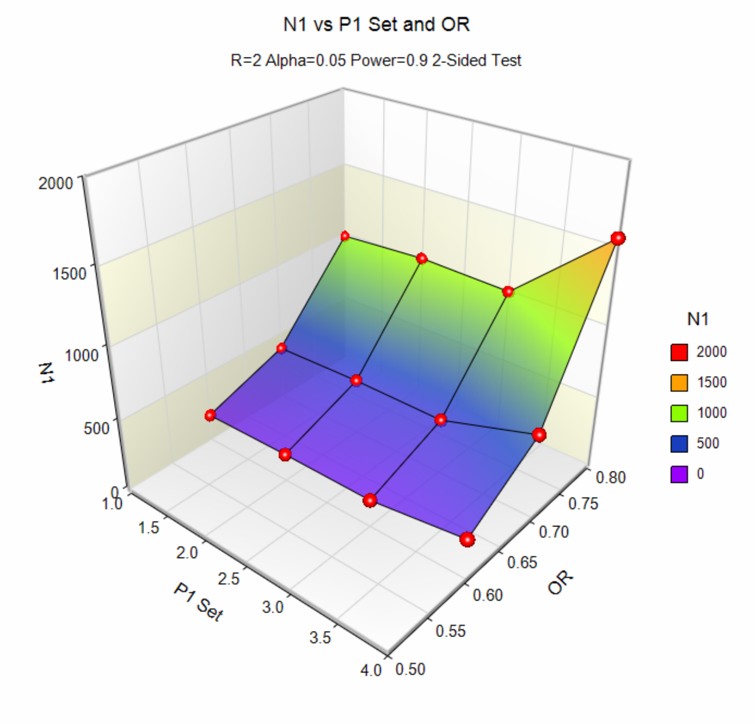
Solve For: Sample Size
Groups: 1 = Control, 2 = Treatment or Experimental
Hypotheses: H0: OR = 1 vs. H1: OR ≠ 1

	Sample Size			Allocation	Number of	Odds	Category Proportions		
				Ratio	Categories	Ratio			
Power	N1	N2	N	R	K	OR	P1	P2	Alpha
0.9009	188	376	564	2	6	0.6	Eq(1)	P2(1)	0.05
0.9005	384	768	1152	2	6	0.7	Eq(1)	P2(2)	0.05
0.9001	978	1956	2934	2	6	0.8	Eq(1)	P2(3)	0.05
0.9013	194	388	582	2	6	0.6	PD(2)	P2(4)	0.05
0.9006	397	794	1191	2	6	0.7	PD(2)	P2(5)	0.05
0.9001	1013	2026	3039	2	6	0.8	PD(2)	P2(6)	0.05
0.9014	190	380	570	2	6	0.6	LD(3)	P2(7)	0.05
0.9007	389	778	1167	2	6	0.7	LD(3)	P2(8)	0.05
0.9000	993	1986	2979	2	6	0.8	LD(3)	P2(9)	0.05
0.9002	275	550	825	2	6	0.6	FL(4)	P2(10)	0.05
0.9002	586	1172	1758	2	6	0.7	FL(4)	P2(11)	0.05
0.9001	1549	3098	4647	2	6	0.8	FL(4)	P2(12)	0.05

Item	Values
Eq(1)	0.167, 0.167, 0.167, 0.167, 0.167, 0.167
PD(2)	0.23, 0.13, 0.36, 0.1, 0.09, 0.09
LD(3)	0.286, 0.238, 0.19, 0.143, 0.095, 0.048
FL(4)	0.75, 0.05, 0.05, 0.05, 0.05, 0.05
P2(1)	0.107, 0.124, 0.144, 0.17, 0.205, 0.25
P2(2)	0.123, 0.136, 0.153, 0.172, 0.194, 0.222
P2(3)	0.138, 0.148, 0.159, 0.171, 0.185, 0.2
P2(4)	0.152, 0.1, 0.354, 0.125, 0.126, 0.142
P2(5)	0.173, 0.11, 0.36, 0.118, 0.115, 0.124
P2(6)	0.193, 0.117, 0.363, 0.112, 0.105, 0.11
P2(7)	0.194, 0.204, 0.202, 0.183, 0.14, 0.077
P2(8)	0.219, 0.216, 0.201, 0.171, 0.126, 0.067
P2(9)	0.242, 0.226, 0.199, 0.161, 0.114, 0.059
P2(10)	0.643, 0.063, 0.067, 0.071, 0.076, 0.081
P2(11)	0.677, 0.059, 0.062, 0.064, 0.067, 0.07
P2(12)	0.706, 0.056, 0.057, 0.059, 0.06, 0.062

Plots





PASS found the required sample sizes for the various cases. Note the large increase in sample size that occurs when the response probability of the first category is much larger than the rest.