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 $C_1, \ldots, C_K$. 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

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

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 $Log(OR)$ is small and *N* is large. The test statistic and power formulae are as follows.

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

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

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

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

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

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

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

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

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

$$\sigma_Y^2 = V$$

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

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

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

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

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

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 *poor*, *moderate*, *good*, *and very good*. Historically, the responses have been about 20% *poor*, 50% *moderate*, 20% *good*, and 10% *very good*.

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

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the procedure window. You may then make the appropriate entries as listed below, or open **Example 1** by going to the **File** menu and choosing **Open Example Template**.

Option	Value
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**

Annotated Output

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

Numeric Results

Numeric Results

Hypotheses: H0: OR = 1 vs. H1: OR ≠ 1
Group 1: Control Group
Group 2: Experimental Group

Power	N1	N2	N	Num Cat's K	Odds Ratio OR	Grp 1 Prop's Set P1	Grp 2 Prop's Set P2	Alpha
0.9006	443	443	886	4	1.5	P1(1)	P2(1)	0.05
0.9017	151	151	302	4	2.0	P1(1)	P2(2)	0.05
0.9017	86	86	172	4	2.5	P1(1)	P2(3)	0.05

Set(Set Number): Values
P1(1): 0.2, 0.5, 0.2, 0.1
P2(1): 0.14, 0.47, 0.25, 0.14
P2(2): 0.11, 0.43, 0.28, 0.18
P2(3): 0.09, 0.39, 0.3, 0.22

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.

Report Definitions

Power is the probability of rejecting a false null hypothesis.
N1 is the number of subjects in the group 1, the control group.
N2 is the number of subjects in the group 2, the experimental group.
N is the total sample size, N1 + N2.
Num Cat's K is the number of categories in the response variable.
Odds Ratio OR is the odds ratio = odds2/odds1.
Grp 1 Prop's Set P1 is the name of the set containing the response proportions for each of the K categories in group 1, the control group.
Grp 2 Prop's Set P2 is the name of the set containing the response proportions for each of the K categories in group 2, the experimental group.
Alpha is the probability of rejecting a true null hypothesis.

Summary Statements

Samples of 443 subjects in the control group and 443 subjects in experimental group achieve 90% power to detect an odds ratio of 1.5 when the significance level (alpha) is 0.05 using a two-sided test. The number of response categories is 4. The response proportions in group 1 are estimated to be 0.2, 0.5, 0.2, 0.1. The response proportions in group 2 are estimated to be 0.14, 0.47, 0.25, 0.14.

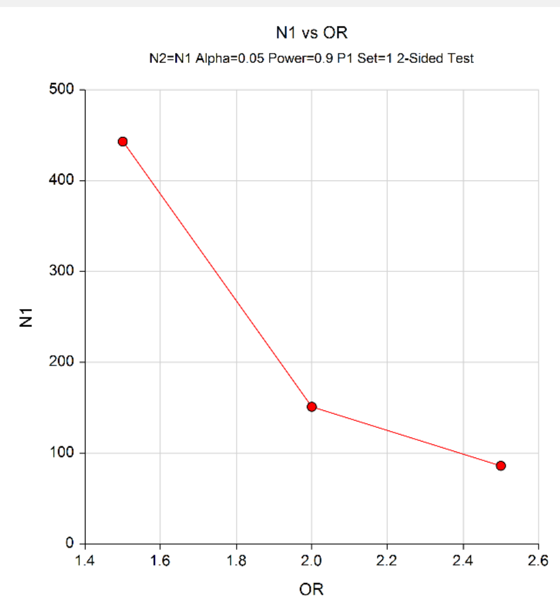
This report shows the numeric results of this sample size study. The definitions of the items on the report are given in the Report Definitions section.

We would like to point out a couple of things. Note that the scores of the experiment group increase over those of the control group.

Also note the impact that the increasing odds ratios have had on the response probabilities. The proportion of those responding “Poor” has fallen from 0.20 to 0.09. The corresponding proportion of those responding “Very Good” has risen from 0.1 to 0.22.

Chart Section

Chart Section



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, 0.5, 0.2, and 0.1.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the procedure window. You may then make the appropriate entries as listed below, or open **Example 2** by going to the **File** menu and choosing **Open Example Template**.

Option	Value
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

Numeric Results

Hypotheses: H0: Log(OR) = 0 vs. H1: Log(OR) ≠ 0
Group 1: Control Group
Group 2: Experimental Group

Power	N1	N2	N	Num Cat's K	Log Odds Ratio Log(OR)	Grp 1 Prop's Set P1	Grp 2 Prop's Set P2	Alpha
0.9015	95	95	190	4	-0.887	P1(1)	P2(1)	0.05

Set(Set Number): 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

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the procedure window. You may then make the appropriate entries as listed below, or open **Example 3** by going to the **File** menu and choosing **Open Example Template**.

Option	Value
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

Numeric Results

Hypotheses: H0: OR = 1 vs. H1: OR ≠ 1
Group 1: Control Group
Group 2: Experimental Group

Power	N1	N2	N	Num Cat's K	Target N2/N1 R	Odds Ratio OR	Grp 1 Prop's Set P1	Grp 2 Prop's Set P2	Alpha
0.9002	195	390	585	6	2	0.6	P1(1)	P2(1)	0.05
0.9003	399	798	1197	6	2	0.7	P1(1)	P2(2)	0.05
0.9002	1017	2034	3051	6	2	0.8	P1(1)	P2(3)	0.05

Set(Set Number): Values

P1(1): 0.23, 0.13, 0.36, 0.1, 0.09, 0.09
P2(1): 0.332, 0.152, 0.327, 0.073, 0.06, 0.056
P2(2): 0.299, 0.146, 0.34, 0.081, 0.068, 0.065
P2(3): 0.272, 0.141, 0.35, 0.088, 0.076, 0.073

PASS found the required sample sized for the three odds ratios. Note that N2 = 2(N1) in all cases as request. Also note that the predicted distributions of responses anticipate a decrease in the value of P21 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

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the procedure window. You may then make the appropriate entries as listed below, or open **Example 4** by going to the **File** menu and choosing **Open Example Template**.

Option **Value**

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

Numeric Results

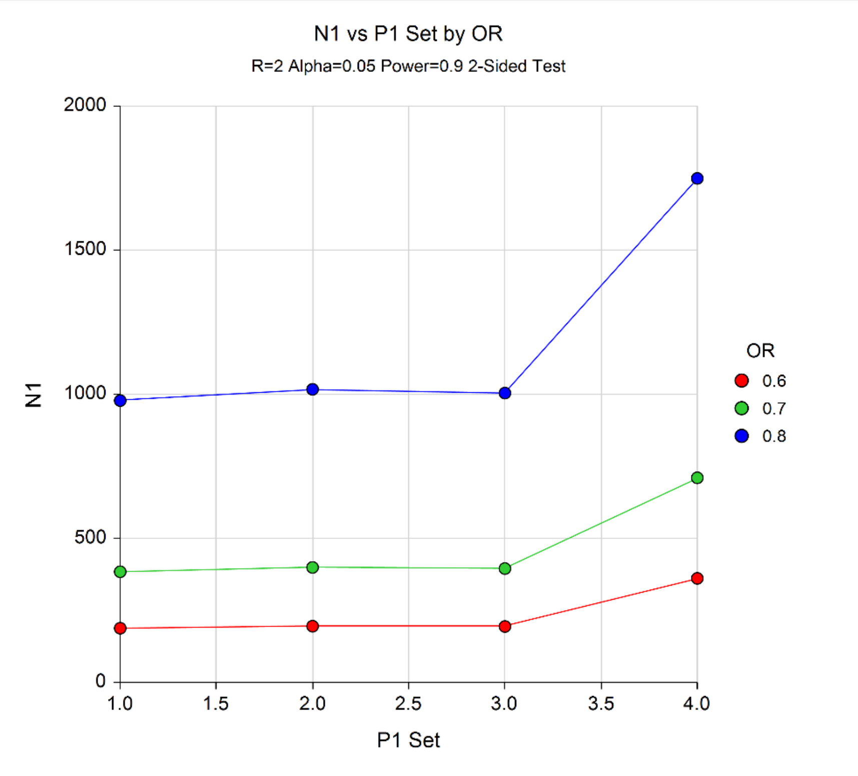
Hypotheses: H0: OR = 1 vs. H1: OR ≠ 1
Group 1: Control Group
Group 2: Experimental Group

Power	N1	N2	N	Num Cat's K	Target N2/N1 R	Odds Ratio OR	Grp 1 Prop's Set P1	Grp 2 Prop's Set P2	Alpha
0.9009	188	376	564	6	2	0.6	Eq(1)	P2(1)	0.05
0.9005	384	768	1152	6	2	0.7	Eq(1)	P2(2)	0.05
0.9001	978	1956	2934	6	2	0.8	Eq(1)	P2(3)	0.05
0.9002	195	390	585	6	2	0.6	PD(2)	P2(4)	0.05
0.9003	399	798	1197	6	2	0.7	PD(2)	P2(5)	0.05
0.9002	1017	2034	3051	6	2	0.8	PD(2)	P2(6)	0.05
0.9001	194	388	582	6	2	0.6	LD(3)	P2(7)	0.05
0.9001	395	790	1185	6	2	0.7	LD(3)	P2(8)	0.05
0.9001	1004	2008	3012	6	2	0.8	LD(3)	P2(9)	0.05
0.9003	361	722	1083	6	2	0.6	FL(4)	P2(10)	0.05
0.9002	710	1420	2130	6	2	0.7	FL(4)	P2(11)	0.05
0.9002	1749	3498	5247	6	2	0.8	FL(4)	P2(12)	0.05

Set(Set Number): 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): 0.25, 0.205, 0.17, 0.144, 0.124, 0.107
P2(2): 0.222, 0.194, 0.172, 0.153, 0.136, 0.123
P2(3): 0.2, 0.185, 0.171, 0.159, 0.148, 0.138
P2(4): 0.332, 0.152, 0.327, 0.073, 0.06, 0.056
P2(5): 0.299, 0.146, 0.34, 0.081, 0.068, 0.065
P2(6): 0.272, 0.141, 0.35, 0.088, 0.076, 0.073
P2(7): 0.4, 0.247, 0.159, 0.103, 0.062, 0.029
P2(8): 0.364, 0.247, 0.17, 0.114, 0.071, 0.034
P2(9): 0.333, 0.246, 0.179, 0.125, 0.079, 0.038
P2(10): 0.833, 0.036, 0.035, 0.033, 0.032, 0.031
P2(11): 0.811, 0.04, 0.039, 0.038, 0.037, 0.036
P2(12): 0.789, 0.044, 0.043, 0.042, 0.041, 0.04

Chart Section



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.