Chapter 836

Reliability Demonstration Tests of One Proportion with Adverse Events

Introduction

This routine calculates the sample size needed for a *reliability demonstration test* based on the binomial distribution. These tests may be conducted by comparing a one-sided confidence bound for a proportion and with a pre-determined reliability standard or by comparing the number of adverse events to a calculated minimum value. It may be used to find a sample size and/or power for studies of *rare events*.

Power Not Used

This is one of the few procedures in **PASS** in which the sample calculation does not use power. Instead, a value for power is implied by the specified value of the maximum number of adverse events. Optionally, the power of the design can be calculated.

Technical Details

This procedure is primarily based on results in Guenther (1977), Hahn and Meeker (1991), and Abernethy (2006). Conceptually, the test compares the number of adverse events to a critical value which is chosen to guarantee a certain significance level. The sample size formulas are based on using an exact binomial test of one proportion.

A binomial variable should exhibit the following four properties:

1. The variable is binary --- it can take on one of two possible values.
2. The variable is observed a known number of times. Each observation or replication is called a Bernoulli trial. The number of replications is *n*. The number of times that the outcome of interest is observed is *r*. Thus, *r* takes on the possible values 0, 1, 2, ..., *n*.
3. The probability, *P*, that the outcome of interest occurs is constant for each trial.
4. The trials are independent. The outcome of one trial does not influence the outcome of the any other trial.

A binomial probability is calculated using the formula

$$b(r;n,P) = \binom{n}{r} P^r (1-P)^{n-r}$$

where

$$\binom{n}{r} = \frac{n!}{r!\,(n-r)!}$$

Following Guenther (1977) define the cumulative binomial as

$$Bi(c; N, P) = Pr(X \leq c) = \sum_{x=0}^{c} b(x; N, P) = \sum_{x=0}^{c} \binom{N}{x} P^x (1-P)^{N-x}$$

The minimum sample size is found by finding the smallest value of *N* the obeys the inequality

$$Bi(c; N, 1 - P_{MIN}) \leq \alpha$$

The power of this test may be calculated using

$$Bi(c; N, 1 - P_{TRUE}) = Power$$

where *power* is the probability of a successful demonstration, P_{MIN} is the reliability standard that must be met, and P_{TRUE} is the reliability value at which *power* is calculated.

Example 1 – Calculating Sample Size

Suppose a study is planned to determine the sample size required for a reliability demonstration study for values of *c* from 0 to 4. The other values are *PMIN* = 0.9 and Alpha = 0.05.

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**
Alpha... **0.05**
c (Maximum Adverse Events) **0 1 2 3 4**
PMIN (Minimum Reliability) **0.90**
Calculate Power (Probability of a Successful Demonstration) ..... **Unchecked**

Output

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

Numeric Reports

Numeric Results

Solve For: Sample Size
Test Type: Exact Binomial Test

Sample Size N	Minimum Reliability PMIN	Alpha Target	Alpha Actual	Maximum Adverse Events c
29	0.9	0.05	0.047	0
46	0.9	0.05	0.048	1
61	0.9	0.05	0.049	2
76	0.9	0.05	0.047	3
89	0.9	0.05	0.050	4

N Sample Size. The number of subjects in the study.
PMIN Minimum Reliability. The reliability standard. It is the required proportion conforming.
Alpha The significance level of this one-sided test. It is the probability of rejecting the null hypothesis of unreliability when it is true. The demonstration test is carried-out by comparing a one-sided lower confidence bound of the reliability with a confidence level of one minus alpha to a specified minimum requirement (PMIN). If the confidence bound is larger than PMIN, reliability is demonstrated.
Target Alpha The significance level the design is meant to achieve.
Actual Alpha The significance level the design actually achieves. It is calculated using PMIN.
c Maximum Adverse Events. The number of nonconforming events that are tolerated while still concluding a successful demonstration. It is sometimes called the "maximum failures allowed."

Summary Statements

A single-group reliability demonstration design will be used to test whether the minimum reliability (success proportion) requirement of 0.9 is met (H0: P ≤ 0.9 versus H1: P > 0.9). The evaluation will be made using a one-sided, one-sample, exact binomial test. If the maximum number of adverse events allowed for a successful demonstration is 0, to obtain a Type I error rate (α) of 0.05, a sample size of 29 units (or observations) will be needed. The achieved Type I error rate (α) is 0.047.

Dropout-Inflated Sample Size

Dropout Rate	Sample Size N	Dropout-Inflated Enrollment Sample Size N'	Expected Number of Dropouts D
20%	29	37	8
20%	46	58	12
20%	61	77	16
20%	76	95	19
20%	89	112	23

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 the tolerance interval is computed. If N subjects are evaluated out of the N' subjects that are enrolled in the study, the design will achieve the stated tolerance interval.
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, 37 subjects should be enrolled to obtain a final sample size of 29 subjects.

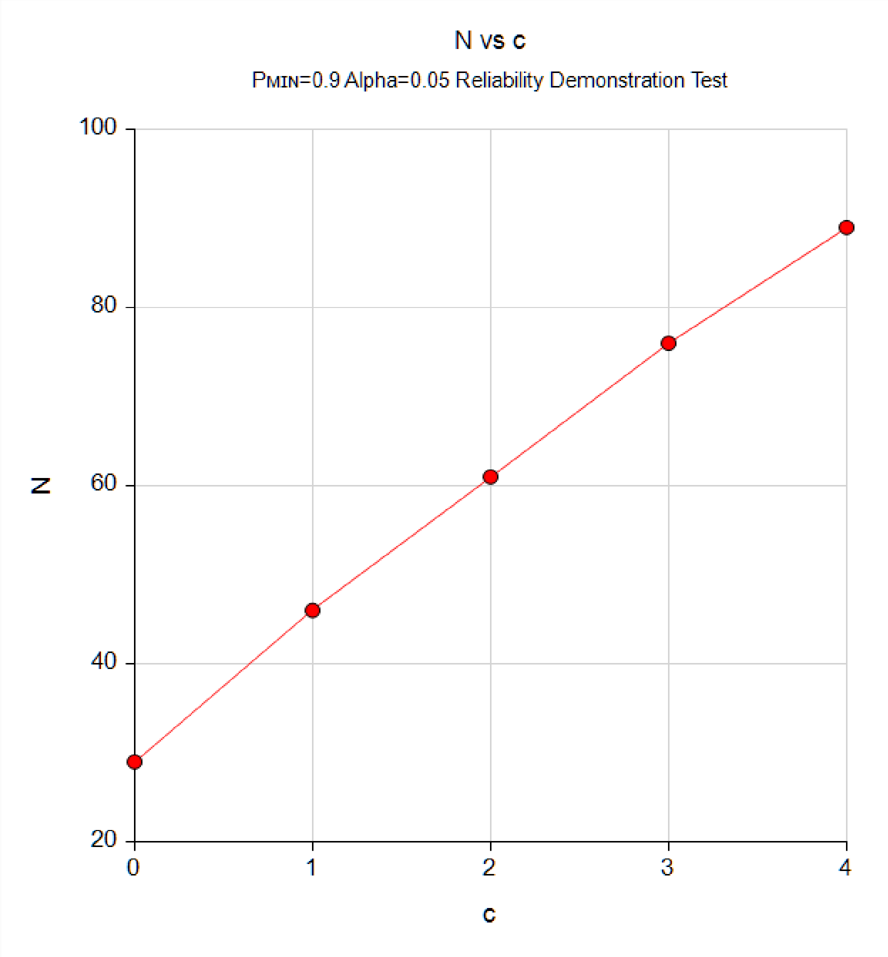
References

Abernethy, Robert B. 1996. The New Weibull Handbook, Second Edition. Robert B. Abernethy, Florida.
Guenther, William C. 1977. Sampling Inspection in Statistical Quality Control. Griffin's Statistical Monographs, Number 37. London.
Hahn, G. J. and Meeker, W.Q. 1991. Statistical Intervals. John Wiley & Sons. New York.
Meeker, W.Q. and Escobar, L.A. 1998. Statistical Methods for Reliability Data. John Wiley & Sons. New York.

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

Plots Section

Plots



This plot shows the sample size for various value of c.

Example 2 – Validation using Abernethy (1996)

Abernethy (1996) page 8-5 presents a table of sample sizes when *c*, the number of failures, is zero. One line of the table is for P_{MIN} = 0.99 and various confidence level values from 0.5 to 0.9999. These confidence levels are transformed into alpha values using alpha = 1 – confidence level. The resulting sample sizes are 69, 92, 120, 161, 230, 299, 459, 688, and 917.

To show off the additional capabilities of **PASS**, we will also display the power when P_{TRUE} = 0.999. These results are not available in Abernethy (1996), but they are very enlightening.

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**
Alpha.. **0.0001 0.001 0.01 0.05 0.1 0.2 0.3 0.4 0.5**
c (Maximum Adverse Events) **0**
PMIN (Minimum Reliability) **0.99**
Calculate Power (Probability of a Successful Demonstration) ..... **Checked**
PTRUE (True Reliability)... **0.999**

Output

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

Numeric Results

Solve For: Sample Size (Power was Ignored)
Test Type: Exact Binomial Test

Power	Sample Size N	Reliability Minimum PMIN	Reliability True PTRUE	Alpha Target	Alpha Actual	Maximum Adverse Events c
0.3995	917	0.99	0.999	0.0001	0.0001	0
0.5024	688	0.99	0.999	0.0010	0.0010	0
0.6318	459	0.99	0.999	0.0100	0.0099	0
0.7414	299	0.99	0.999	0.0500	0.0495	0
0.7944	230	0.99	0.999	0.1000	0.0991	0
0.8512	161	0.99	0.999	0.2000	0.1983	0
0.8869	120	0.99	0.999	0.3000	0.2994	0
0.9121	92	0.99	0.999	0.4000	0.3967	0
0.9333	69	0.99	0.999	0.5000	0.4998	0

Note that **PASS** found the same sample size values as Abernethy (1996). Also note that the computed power of these designs shows that those with an alpha of 0.05 or less are under-powered (power < 0.8).