

Chapter 446

Tests for Fold Change of Two Means (Log-Normal Data)

Introduction

The fold change is the ratio of two values. It measures how much a variable has changed between the two measurements. For example, suppose the beginning value of a variable is 2.5 and its value after a year has increased to 5. This variable has had a two-fold increase in its value.

Fold changes are used when the change is of more interest than the absolute values. They are often used with microarray data to measure the change in gene expression levels.

In this procedure, the fold change is the ratio of two means. By taking the logarithm of the data, the difference in the means of the logged data corresponds to their ratio in the original scale. Thus, if we are willing to assume that the logged data are normally distributed, we can use the two-sample t-test to perform hypothesis tests about the fold change.

This procedure calculates power and sample size for t-tests from a parallel-groups design in which the logarithm of the outcome is a continuous normal random variable.

The details of testing two treatments using data from a two-group design are given in another chapter, and they will not be repeated here. If the logarithms of the responses can be assumed to follow a normal distribution, hypotheses stated in terms of the fold change can be transformed into hypotheses about the difference. The details of this analysis are given in Julious (2004). They will only be summarized here.

Testing of Fold Change

It will be convenient to adopt the following specialized notation for the discussion of these tests.

Parameter	PASS Input/Output	Interpretation
μ_T	Not used	<i>Treatment mean.</i> This is the treatment mean.
μ_R	Not used	<i>Reference mean.</i> This is the reference (or control) mean.
FC	FC	<i>Fold Change.</i> This is the value of $FC = \mu_T/\mu_R$.

Note that the actual values of μ_T and μ_R are not needed. Only their ratio is needed for power and sample size calculations.

In the two-sided case, the null hypothesis is $H_0: FC = FC_0$ and the alternative hypothesis is $H_1: FC = FC_1$ where FC_1 is not equal to FC_0 .

Log-Transformation

In many cases, hypotheses stated in terms of fold changes are more convenient than hypotheses stated in terms of differences. This is because ratios can be interpreted as percentages, but differences must be interpreted as actual amounts in their original scale. Hence, it has become a common practice to take the following steps in hypothesis testing.

1. State the statistical hypotheses in terms of the fold change (ratio) of the means.
2. Transform this into hypotheses about a difference by taking logarithms.
3. Analyze the logged data—that is, do the analysis in terms of the difference.
4. Draw the conclusion in terms of the ratio.

Thus, a hypothesis about the fold change of the means on the original scale can be translated into a hypothesis about the difference of two means on the logged scale.

Coefficient of Variation

The coefficient of variation (COV) is the ratio of the standard deviation to the mean. This parameter can be used to represent the variation in the data because of a unique relationship that it has in the case of log-normal data.

Suppose the variable X is the logarithm of the original variable Y . That is, $X = \ln(Y)$ and $Y = \exp(X)$. Label the mean and variance of X as μ_X and σ_X^2 , respectively. Similarly, label the mean and variance of Y as μ_Y and σ_Y^2 , respectively. If X is normally distributed, then Y is log-normally distributed. Julious (2004) presents the following well-known relationships between these two variables

$$\mu_Y = e^{\mu_X + \frac{\sigma_X^2}{2}}$$

$$\sigma_Y^2 = \mu_Y^2 (e^{\sigma_X^2} - 1)$$

From this relationship, the coefficient of variation of Y can be found to be

$$COV_Y = \frac{\sqrt{\mu_Y^2 (e^{\sigma_X^2} - 1)}}{\mu_Y}$$

$$= \sqrt{e^{\sigma_X^2} - 1}$$

Solving this relationship for σ_X^2 , the standard deviation of X can be stated in terms of the coefficient of variation of Y as

$$\sigma_X = \sqrt{\ln(COV_Y^2 + 1)}$$

Tests for Fold Change of Two Means (Log-Normal Data)

Similarly, the mean of X is

$$\mu_X = \ln\left(\frac{\mu_Y}{\sqrt{COV_Y^2 + 1}}\right)$$

Thus, the hypotheses can be stated in the original (Y) scale and then the power can be analyzed in the transformed (X) scale. For parallel-group designs, $\sigma_X^2 = \sigma_d^2$, the average variance used in the t-test of the logged data.

Power Calculation

As is shown above, the hypotheses can be stated in the original (Y) scale using ratios or the logged (X) scale using differences. In either case, power and sample size calculations are made using the formulas for testing the difference in two means. These formulas are presented in another chapter and are not duplicated here.

Example 1 – Finding Power

A company has developed a generic drug for treating rheumatism and wants to show that it is better than the standard drug. From previous studies, responses for either treatment are known to follow a lognormal distribution. A parallel-group design will be used and the logged data will be analyzed with a one-sided, two-sample t-test.

Past experience leads the researchers to set the COV to 1.20. The significance level is 0.025. The power will be computed for R1 equal 1.10 and 1.20. Sample sizes between 100 and 900 will be examined in the analysis.

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
Alternative Hypothesis	H1: FC > FC0 (One-Sided)
Alpha.....	0.025
Group Allocation	Equal (N1 = N2)
Sample Size Per Group	100 to 900 by 200
FC0 (Fold Change H0).....	1.00
FC1 (Fold Change H1).....	1.1 1.2
COV (Coefficient of Variation).....	1.2

Tests for Fold Change of Two Means (Log-Normal Data)

Output

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

Numeric Reports

Numeric Results

Solve For: **Power**
 Test Type: Two-Sample T-Test
 Fold Change: $FC = \mu_T / \mu_C$
 Hypotheses: $H_0: FC \leq FC_0$ vs. $H_1: FC > FC_0$

Power	Sample Size			Fold Change		Effect Size	Coefficient of Variation COV	Alpha
	N1	N2	N	FC0	FC1			
0.1057	100	100	200	1	1.1	0.1009	1.2	0.025
0.2339	300	300	600	1	1.1	0.1009	1.2	0.025
0.3581	500	500	1000	1	1.1	0.1009	1.2	0.025
0.4715	700	700	1400	1	1.1	0.1009	1.2	0.025
0.5718	900	900	1800	1	1.1	0.1009	1.2	0.025
0.2737	100	100	200	1	1.2	0.1930	1.2	0.025
0.6556	300	300	600	1	1.2	0.1930	1.2	0.025
0.8625	500	500	1000	1	1.2	0.1930	1.2	0.025
0.9506	700	700	1400	1	1.2	0.1930	1.2	0.025
0.9836	900	900	1800	1	1.2	0.1930	1.2	0.025

Power The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
 N1 and N2 The number of items sampled from each population.
 N The total sample size. $N = N_1 + N_2$.
 FC0 The fold change (ratio) under the null hypothesis, H_0 .
 FC1 The fold change (ratio) at which the power is calculated.
 Effect Size Equal to $|\ln(FC_0) - \ln(FC_1)| / \text{Sigma}$.
 COV The coefficient of variation on the original scale. The value of sigma is calculated from this.
 Alpha The probability of rejecting a true null hypothesis.

Summary Statements

A parallel two-group design will be used to test whether the fold change (μ_T / μ_C) is greater than 1 ($H_0: \mu_T / \mu_C \leq 1$ versus $H_a: \mu_T / \mu_C > 1$). The comparison will be made using a one-sided, two-sample t-test using a log-transformation, with a Type I error rate (α) of 0.025. The coefficient of variation on the original scale is assumed to be 1.2. To detect a fold change (μ_T / μ_C) of 1.1, with a sample size of 100 in Group 1 (treatment) and 100 in Group 2 (control), the power is 0.1057.

Tests for Fold Change of Two Means (Log-Normal Data)

Dropout-Inflated Sample Size

Dropout Rate	Sample Size			Dropout-Inflated Enrollment Sample Size			Expected Number of Dropouts		
	N1	N2	N	N1'	N2'	N'	D1	D2	D
20%	100	100	200	125	125	250	25	25	50
20%	300	300	600	375	375	750	75	75	150
20%	500	500	1000	625	625	1250	125	125	250
20%	700	700	1400	875	875	1750	175	175	350
20%	900	900	1800	1125	1125	2250	225	225	450

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 (as entered by the user). 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. 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 Lohknygina, 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, 125 subjects should be enrolled in Group 1, and 125 in Group 2, to obtain final group sample sizes of 100 and 100, respectively.

References

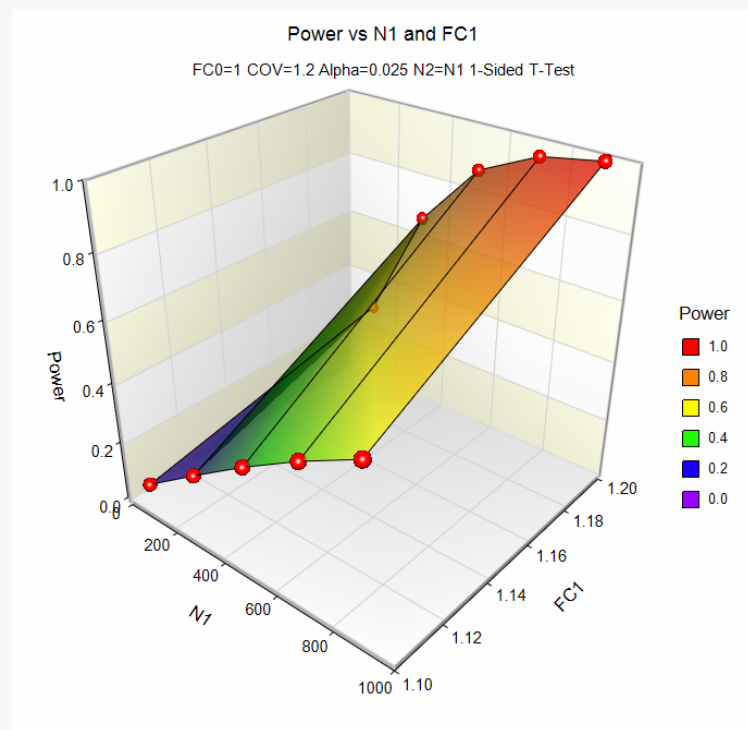
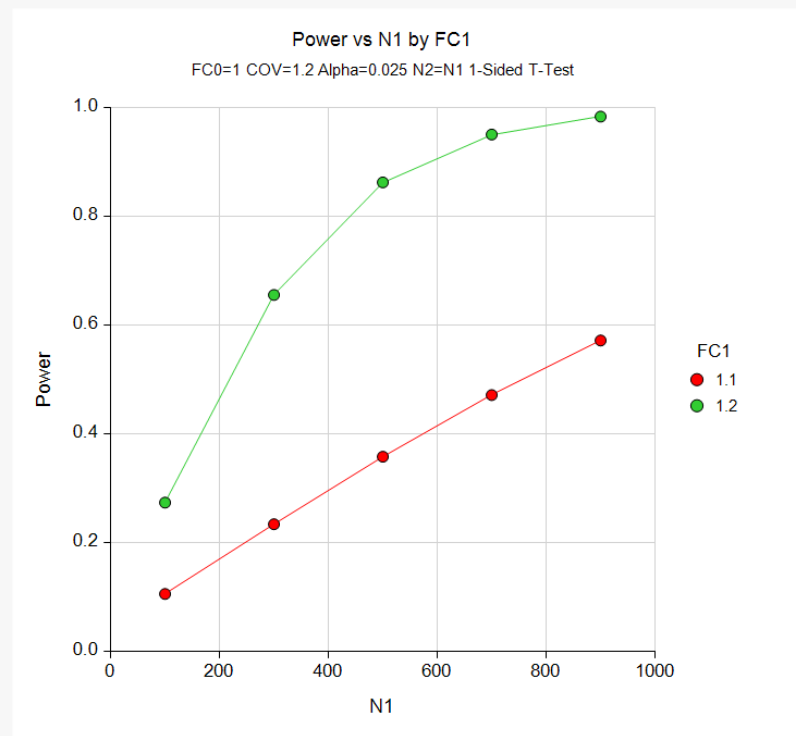
- Chow, S.C., Shao, J., and Wang, H. 2003. Sample Size Calculations in Clinical Research. Marcel Dekker. New York.
- Julious, Steven A. 2004. 'Tutorial in Biostatistics. Sample sizes for clinical trials with Normal data.' Statistics in Medicine, 23:1921-1986.

This report shows the power for the indicated scenarios.

Tests for Fold Change of Two Means (Log-Normal Data)

Plots Section

Plots



These plots show the relationship of power to sample size for the two values of FC1.

Example 2 – Validation

We will validate this procedure by showing that it gives the identical results to the regular test on differences—a procedure that has been validated. We will use the same settings as those given in Example 1. Since the output for this example is shown above, only the output from the procedure that uses differences (Two-Sample T-Tests Assuming Equal Variance) is shown below.

To run the power analysis of a *t*-test on differences, we need the values of μ_2 (which correspond to R1) and σ . The value of μ_1 will be zero.

$$\begin{aligned}\sigma &= \sqrt{\ln(COV^2 + 1)} \\ &= \sqrt{\ln(1.2^2 + 1)} \\ &= 0.944456\end{aligned}$$

$$\begin{aligned}\mu_2 &= \ln(R1) & \mu_2 &= \ln(R1) \\ &= \ln(1.10) & &= \ln(1.20) \\ &= 0.095310 & &= 0.182322\end{aligned}$$

Setup

If the procedure window (Two-Sample T-Tests Assuming Equal Variance) 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 2c** 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
Alternative Hypothesis	One-Sided
Alpha.....	0.025
Group Allocation	Equal (N1 = N2)
Sample Size Per Group	100 to 900 by 200
Input Type.....	Means
μ_1	0
μ_2	0.095310 0.182322
σ	0.944456

Tests for Fold Change of Two Means (Log-Normal Data)

Output

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

Numeric Results

Solve For: **Power**
 Test Type: Two-Sample Equal-Variance T-Test
 Difference: $\delta = \mu_1 - \mu_2$
 Hypotheses: $H_0: \delta \leq 0$ vs. $H_1: \delta > 0$

Power	Sample Size			Mean Difference δ	Standard Deviation σ	Alpha
	N1	N2	N			
0.10568	100	100	200	0.095310	0.9445	0.025
0.23392	300	300	600	0.095310	0.9445	0.025
0.35811	500	500	1000	0.095310	0.9445	0.025
0.47145	700	700	1400	0.095310	0.9445	0.025
0.57180	900	900	1800	0.095310	0.9445	0.025
0.27374	100	100	200	0.182322	0.9445	0.025
0.65562	300	300	600	0.182322	0.9445	0.025
0.86253	500	500	1000	0.182322	0.9445	0.025
0.95061	700	700	1400	0.182322	0.9445	0.025
0.98359	900	900	1800	0.182322	0.9445	0.025

You can compare these power values with those shown above in Example 1 to validate the procedure. You will find that the power values are identical.