Chapter 576

Exact Method for Assessing Agreement in Method Comparison Studies

Introduction

This procedure provides sample size and power calculations for a method comparison study that is analyzed using the exact method proposed by Shieh (2019). This exact method has been shown to have better test characteristics than the common the Bland – Altman (1986) method.

A method comparison study is used to compare two measurement methods. The data consists of pairs of measurements taken on *N* subjects. This exact technique forms two limits of agreement from the *N* paired differences using a special distribution function. Next, these limits are compared to an allowable range which is defined as –δ to δ. If the minimum and maximum values are both between this allowable range, the null hypothesis of disagreement is rejected in favor of the alternative hypothesis that the two measurements are in agreement.

Technical Details

Suppose two measurements (X and Y) are obtained using two measurement methods on each of *N* subjects drawn from a population of interest. It is assumed that the differences computed on these data pairs are normally distributed with constant mean and variance.

Test

The analysis proceeds as follows.

1. Form the *N* paired differences $d_i = x_i - y_i$. These differences are assumed to follow the normal distribution with mean μ and standard deviation σ.

   Define the 100*p*th percentile of the paired differences as $\theta_p = \mu + z_p\sigma$ where z_p is determined from the standard normal distribution.

   To establish agreement between the two methods, the central portion of the distribution of d_i needs to be within a close range about zero. This is accomplished using two percentiles θ_p and θ_{1-p}. The resulting statistical hypotheses are

$$H_0: \theta_{1-p} \leq -\delta \text{ or } \delta \leq \theta_p$$

$$H_1: -\delta < \theta_{1-p} \leq -\delta \text{ and } \theta_p < \delta$$

2. The bounds $-\delta$ and δ are determined so that the central portion of the distribution of the differences is equal to $P_1^* = 2p - 1$. This requires that the bounds be chosen so that $\Phi[(\delta - \mu)/\sigma] - \Phi\left[\frac{(-\delta-\mu)}{\sigma}\right] > P_1^*$ where $\Phi[z]$ is the cdf of the standard normal distribution. This inequality changes to an equality when $\left(\theta_{1-p}, \theta_p\right) = (-\delta, \delta)$. Hence, there is a one-to-one relationship between P_1^* and δ based on the normal distribution.

3. Compute the mean $\bar{D}$ and the standard deviation SD of these differences using the usual formulas.

4. Form a pair of exact confidence limits B_L and B_U. These confidence limits are

$$B_L = \bar{D} - c_{1-\alpha}\left(SD/\sqrt{N}\right)$$

$$B_U = \bar{D} + c_{1-\alpha}\left(SD/\sqrt{N}\right)$$

   The critical value of $c_{1-\alpha}$ is chosen to control the Type I error rate using the definition

$$\Pr\left\{\theta_{1-p} \le B_L \text{ and } B_U \le \theta_p\right\} = \alpha$$

   A special algorithm to compute this critical value is given by Shieh (2019).

5. Compare the confidence limits B_L and B_U to the bounds $-\delta$ and δ. If $-\delta < B_L < B_U < \delta$ the null hypothesis of disagreement is rejected in favor of the alternative hypothesis that the two measurement methods are in agreement.

Power Analysis

The exact power of this testing procedure is given in Shieh (2019). It requires numeric integration, and the extensive details will not be repeated here. This power formula requires the specification of N, the null central portion P_0^*, the alternative central portion P_1^* or the corresponding bound δ, the significance level α, and μ and σ.

If the sample size is required, this power formula can be used in a binary search.

Example 1 – Finding the Sample Size

Clinicians wish to conduct a method comparison study and analyze it with the exact agreement test. They want to estimate the number of subjects that must be measured significance level is 0.05, P0* is 0.95, the power is 0.8 or 0.9, δ is 7, the mean of the paired differences is 0.5 and standard deviation between 2.5 and 2.7.

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	**N (Number of Pairs)**
Power	**0.8 0.9**
Alpha	**0.05**
P0* (Central Portion|H0)	**0.95**
δ or P1* Input	**Enter δ and calculate P1***
δ (Difference Bound)	**7**
μ (Mean of Differences)	**0.5**
σ (Std Dev of Differences)	**2.5 2.6 2.7**

Annotated Output

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

Numeric Results

Numeric Results
Solve For: N (Number of Pairs)

Power	Number of Pairs N	Central Prop|H0 P0*	Central Prop|H1 P1*	Difference Boundary δ	Mean of Diff's μ	SD of Diff's σ	Alpha α	Critical Value C[1-α]
0.8018	34	0.95	0.9940	7	0.5	2.5	0.05	13.5705
0.8028	45	0.95	0.9918	7	0.5	2.6	0.05	15.2082
0.8034	62	0.95	0.9892	7	0.5	2.7	0.05	17.4196
0.9016	48	0.95	0.9940	7	0.5	2.5	0.05	15.6231
0.9004	64	0.95	0.9918	7	0.5	2.6	0.05	17.6602
0.9020	90	0.95	0.9892	7	0.5	2.7	0.05	20.5142

References

Shieh, Gwowen. 2019. Assessing Agreement between Two Methods of Quantitative Measurements: Exact Test Procedure and Sample Size Calculation.' Statistics in Biopharmaceutical Research. Link: https://doi.org/10.1080/19466315.2019.1677495.
Bland, J.M., Altman, D.G. 1986. 'Statistical methods for assessing agreement between two methods of clinical measurement.' The Lancet i:307-310.

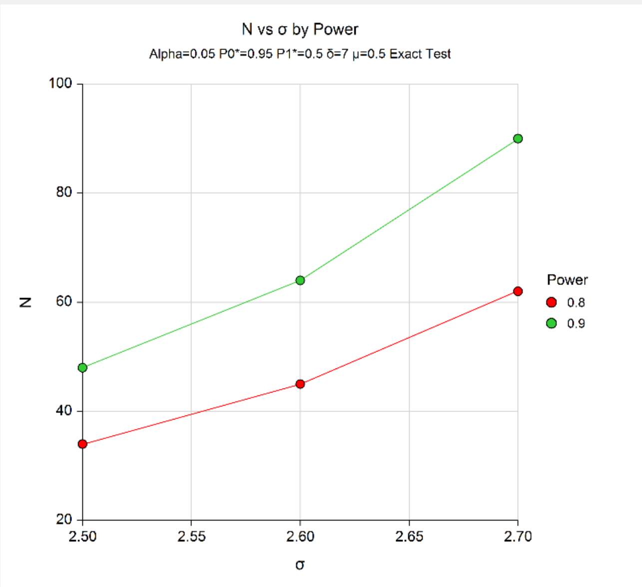
Report Definitions
Hypotheses: H0: Measurement methods do not agree. H1: Measurement methods are in agreement.
Power is the probability of rejecting a false null hypothesis.
N is the number of subject pairs. That is, each subject produces two measurements, one for each method.
P0* is the target central portion of the distribution of paired differences assumed by the null hypothesis, H0.
P1* is the actual central portion of the distribution of paired differences assumed by the alternative hypothesis, H1.
δ is magnitude of the difference boundaries. The null hypothesis is that the target central portion of the distribution of paired differences falls within the interval [-δ, δ].
μ is the population mean of the sample differences.
σ is the population standard deviation of the sample differences.
Alpha is the significance level of the test.
C[1-α] is the critical value used to construct the confidence interval of the data.

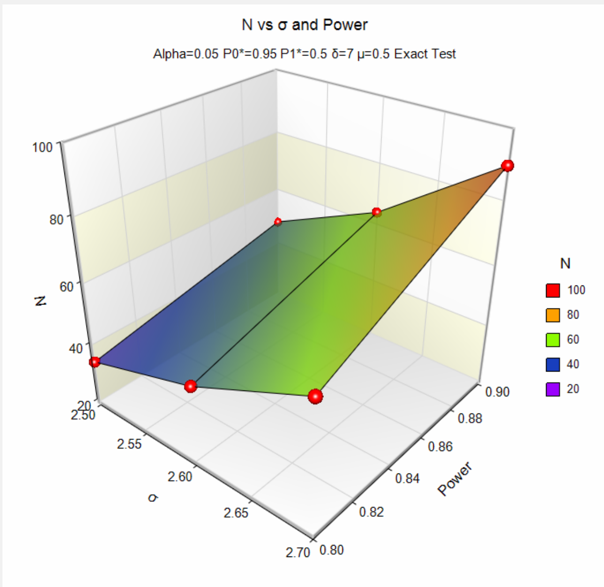
Summary Statements
A method comparison study will be analyzed with an exact test of method agreement. A sample of 34 subjects achieves 80% power to detect agreement at the 0.05 level of significance. One observation will be made for each measurement method on each subject. When the null hypothesis is rejected, it is concluded that the target central portion (P0* = 0.95) of the distribution of paired differences is within the range [-7, 7]. Note that the range [-7, 7] is actually associated with the central portion (P1* = 0.994). The mean and standard deviation of the sample differences are anticipated to be 0.5 and 2.5.

These reports show the values of each of the parameters, one scenario per row.

Chart Section

Chart Section





These plots show the relationship between the standard deviation and sample size for the two alpha levels.

Example 2 – Validation using Shieh (2019)

Shieh (2019) provides an example on page 6 that we will use to validate the program. They want to estimate the number of subjects that must be measured significance level is 0.05, P0* is 0.8, the power is 0.8 or 0.9, δ is 7, the mean of the paired differences is 0.5 and standard deviation between 2.5 and 2.7. The article reports sample sizes of 14 and 20.

We will compute the power for a range of N's from 200 to 203.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load 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	**N (Number of Pairs)**
Power	**0.8 0.9**
Alpha	**0.05**
P0* (Central Portion|H0)	**0.8**
δ or P1* Input	**Enter δ and calculate P1***
δ (Difference Bound)	**0.1**
μ (Mean of Differences)	**0.011**
σ (Std Dev of Differences)	**0.044**

Output

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

Numeric Results

Numeric Results
Solve For: N (Number of Pairs)

Power	Number of Pairs N	Central Prop|H0 P0*	Central Prop|H1 P1*	Difference Boundary δ	Mean of Diff's μ	SD of Diff's σ	Alpha α	Critical Value C[1-α]
0.8079	14	0.8	0.9726	0.1	0.011	0.044	0.05	6.2512
0.9112	20	0.8	0.9726	0.1	0.011	0.044	0.05	7.0605

PASS also shows the sample sizes to be 14 and 20, thus validating the procedure.