Memorial Sloan-Kettering Cancer Center
Memorial Sloan-Kettering Cancer Center, Dept. of Epidemiology
& Biostatistics Working Paper Series

Year 2005 *Paper* 3

Power Calculations for Preclinical Studies Using a K-Sample Rank Test and the Lehmann Alternative Hypothesis

Glenn Heller*

*Memorial Sloan-Kettering Cancer Center, hellerg@mskcc.org

http://biostats.bepress.com/mskccbiostat/paper3

Power Calculations for Preclinical Studies Using a K-Sample Rank Test and the Lehmann Alternative Hypothesis

Glenn Heller

Abstract

Power calculations in a small sample comparative study, with a continuous outcome measure, are typically undertaken using the asymptotic distribution of the test statistic. When the sample size is small, this asymptotic result can be a poor approximation. An alternative approach, using a rank based test statistic, is an exact power calculation. When the number of groups is greater than two, the number of calculations required to perform an exact power calculation is prohibitive. To reduce the computational burden, a Monte Carlo resampling procedure is used to approximate the exact power function of a k-sample rank test statistic under the family of Lehmann alternative hypotheses. The motivating example for this approach is the design of animal studies, where the number of animals per group is typically small.

Power Calculations for Preclinical Studies Using a K-Sample Rank Test and the Lehmann Alternative Hypothesis

Glenn Heller

Department of Epidemiology and Biostatistics,

Memorial Sloan-Kettering Cancer Center, New York, New York

Correspondence to: Glenn Heller, Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, 307 East 63rd Street, New York, New York, 10021.

E-mail: hellerg@mskcc.org

Phone: 646 735 8112

Fax: 646 735 0010

Summary

Power calculations in a small sample comparative study, with a continuous outcome measure, are typically undertaken using the asymptotic distribution of the test statistic. When the sample size is small, this asymptotic result can be a poor approximation. An alternative approach, using a rank based test statistic, is an exact power calculation. When the number of groups is greater than two, the number of calculations required to perform an exact power calculation is prohibitive. To reduce the computational burden, a Monte Carlo resampling procedure is used to approximate the exact power function of a k-sample rank test statistic under the family of Lehmann alternative hypotheses. The motivating example for this approach is the design of animal studies, where the number of animals per group is typically small.

KEYWORDS : Animal study design; Exact test; Permutation distribution; Sample size calculation

1. INTRODUCTION

The subject of this work is the demonstration of an accurate power computation for a small sample randomized comparative study, based on a continuous outcome measure. The motivation for this approach was the development of a statistical design to assess the benefit of experimental cancer therapies in the preclinical (nonhuman) setting, but it is general enough to apply in any randomized comparative study. A prototype experiment will be used throughout the manuscript to highlight the statistical issues. This experiment entails the injection of tumor cells into n mice, and after allowing the tumor cells to grow in a predetermined time frame, the mice are randomly allocated to either a control group or one of multiple experimental therapies. The tumor volume in each mouse is recorded over time, and at a fixed follow-up time period, the area under the tumor volume curve is recorded for each mouse. (Note that any continuously valued positive measure, most notably survival time, truncated at the fixed follow up time, can be used as an outcome measure.) The number of mice allocated to the $k-1$ experimental groups are denoted by $n_1, \ldots, n_{k-1}$ and the number of mice in the control group is n_k; the total sample size in the experiment is $n = n_1 + \ldots + n_k$.

Sample size/power considerations developed for preclinical studies are typically based on the asymptotic distribution of the test statistic used for the comparison. This asymptotic approximation can be inaccurate when the number of subjects in each group is small. An exact power calculation using a rank based test statistic is an alternative approach, but it is impractical when there are more than two groups due to the number of calculations required. In this paper, a Monte Carlo approach is used for power calculations under the Lehmann family of alternative hypotheses with a rank based test statistic. It is demonstrated that the Monte Carlo approach

is both highly accurate and computationally efficient and hence applicable under all conditions.

The Monte Carlo approach is introduced for the two-sample problem. Although exact power calculations can be performed in the current computing environment, when the number of animals per group is small, the two-sample framework provides the simplest vehicle to demonstrate the methodology. The Monte Carlo approach is then extended to the k-sample case, where it will be of most use.

2. THE TWO-SAMPLE CASE

Assume there are n_1 independent identically distributed copies $x_{11}, x_{21}, \ldots, x_{n_1 1}$, generated from the absolutely continuous distribution function F_1 and n_2 independent identically distributed copies $x_{12}, x_{22}, \ldots, x_{n_2 2}$, generated from the absolutely continuous distribution function F_2. Denote the complement of the distribution function, also known as the survival function, as $G(z) = 1 - F(z)$. The null and alternative hypotheses are written as

$$H_0 : G_1(z) = G_2(z)$$

$$H_A : G_1(z) < G_2(z).$$

With regard to the prototype experiment, if G_1 represents the experimental therapy survivor function and G_2 the control group survivor function, the alternative states that larger tumors are more likely to occur in the control group.

The most common approach to expressing stochastic dominance in the alternative hypothesis is the shift function

$$H_A : G_1(z) = G_2(z - \Delta) \qquad \Delta > 0.$$

This alternative states that the difference between the two population means, medians, or any other population quantile is given by the shift parameter Δ. Once the alternative hypothesis is chosen, a test statistic is selected to reflect this hypothesis. Under a shift alternative, the t-statistic is a natural candidate for a test statistic. Unfortunately, prior work suggests that the size of this test is poor if the underlying distribution F is skewed [1-3]. Here the problem is manifested through the effects of outliers on the t-statistic. Miller [2] notes that the outliers will inflate the variance of the t-statistic, thereby depressing the value of the corresponding significance level. Although this problem may be alleviated through a normalizing transformation of the outcome variable, the appropriate transformation is not identifiable at the design stage without prior data.

One method to downweight the influence of outlying observations is to replace the observations with their ranks. The most common rank test statistic is the Wilcoxon rank sum statistic

$$S = \sum_{i=1}^{n_1} \text{rank}(x_{i1}),$$

the ranks of the experimental sample based on the combined data. For the tumor volume study, evidence of an effective experimental therapy is supported by a small rank sum statistic. To perform power calculations based on this test statistic, an asymptotic null pivotal statistic is formed by studentization

$$\tilde{S} = \frac{S - E_0(S)}{\sqrt{\text{VAR}_0(S)}},$$

which enables the investigator to approximate the reference distribution for $\tilde{S}$ under the null hypothesis with the standard normal distribution. For the rank sum statistic,

$$E_0(S) = \frac{n_1(n+1)}{2} \qquad \text{VAR}_0(S) = \frac{n_1 n_2 (n+1)}{12}$$

Although the studentized rank sum statistic provides an adequate standard normal approximation under the null in small samples, the power computation requires knowledge of the experimental outcome distribution (F_1) and the control distribution function (F_2). Without sufficient prior data to assist in the specification of these distributions, the accuracy of this power calculation is suspect.

The Lehmann family of alternative hypotheses can be used instead of the shift alternative for power calculations when using the rank sum statistic. Here, power calculations are performed without full specification of the underlying distributions. Rather, a semiparametric relationship is proposed,

$$G_1(x) = G_2^{\gamma}(x). \tag{1}$$

where the underlying survival distributions (G_1, G_2) are left unspecified, but their relationship is governed by a single parameter γ.

The interpretation of γ is critical to the design. Two interpretations, one using the odds parameter and one using population averages, are provided using the tumor volume study for illustration. For the odds parameter, let $\pi = \Pr(X_1 > X_2)$ represent the probability a mouse from the experimental group has greater area under the tumor volume curve than a mouse from the control group. Since $\Pr(X_1 > X_2) = \int G_1(x) dG_2(x)$, it follows from the Lehmann alternative given in (1) that

$$\gamma = \frac{1-\pi}{\pi}$$

is an odds parameter. In the example, $\gamma = 4$ indicates that the odds are 4:1 that a representative control mouse has the greater tumor volume. A second interpretation of γ is derived in terms of the relative average tumor volume in the two populations.

The hazard function in survival analysis is defined as

$$h(x) = \frac{-\partial \log[G(x)]}{\partial x},$$

and the Lehmann alternative $G_1(x) = G_2^{\gamma}(x)$ can be written in terms of the hazard functions in the two populations as $h_1(x) = \gamma h_2(x)$. Using the relationship

$$E\left[\frac{1}{h(X)}\right] = \int \frac{G(x)}{f(x)} f(x) dx,$$

which for positive valued outcomes is the population mean of the outcome measure, an alternative interpretation for the Lehmann alternative is that the average tumor volume in the control population is γ times greater than the average tumor volume in the experimental population, i.e.

$$\gamma = \frac{\mu_2}{\mu_1}.$$

An advantage to using the relative mean alternative rather than the shift alternative is that it does not require a priori knowledge of a common population standard deviation to perform the power calculations.

3. POWER CALCULATIONS FOR A TWO-SAMPLE RANK TEST

Three methods for calculating the power of the Wilcoxon rank sum statistic under the Lehmann alternative are presented. The exact power calculation [4] is derived by enumerating all possible assignments of the n subjects into two groups of size n_1 and n_2, and the probability of each assignment is calculated under the Lehmann alternative as,

$$Pr(R_1 = r_1, \ldots, R_{n_1} = r_{n_1}) = \frac{\gamma^{n_1}}{\binom{n}{n_1}} \prod_{j=1}^{n_1} \frac{\Gamma(r_j + j\gamma - j)}{\Gamma(r_j)} \frac{\Gamma(r_{j+1})}{\Gamma(r_{j+1} + j\gamma - j)} \qquad (2)$$

where R_j $(j = 1, \ldots, n_1)$ represents the pooled ranks in group 2, $\Gamma(\cdot)$ is the Gamma function and $r_{n_1+1} = n+1$. Enumeration of this sample space is reasonable when the sample size is small. For example, randomly allocating 20 subjects evenly between the two groups, would require approximately 180,000 probability calculations using equation (2); a job that would take seconds in the current computing environment.

The asymptotic normal approximation to the Wilcoxon rank sum statistic is the most widely used approach to power calculations. The power of a two-sided α level test is approximated by

$$\begin{aligned} 1 \quad - \quad & \Phi\left(\frac{E_0(S) - E_A(S) + z_{1-\alpha/2}\sqrt{\mathrm{VAR}_0(S)}}{\sqrt{\mathrm{VAR}_A(S)}}\right) \\ + \quad & \Phi\left(\frac{E_0(S) - E_A(S) - z_{1-\alpha/2}\sqrt{\mathrm{VAR}_0(S)}}{\sqrt{\mathrm{VAR}_A(S)}}\right) \end{aligned}$$

where the subscripts $\{0, A\}$ denote the moment calculations under the null and alternative hypotheses respectively, Φ is the standard normal distribution function, and $z_{1-\alpha/2}$ is the $1-\alpha/2$ quantile of the standard normal distribution, i.e. $\Phi(z_{1-\alpha/2}) = 1-\alpha/2$. The expectation and variance of the rank sum statistic under the Lehmann alternative is [5],

$$\mathrm{E}_A(S) \quad = \quad \frac{n_1 n_2}{1+\gamma} + \frac{n_1(n_1+1)}{2}$$

$$\begin{aligned} \mathrm{Var}_A(S) \quad = \quad & n_1 n_2 (n_1 - 1)\left(\frac{1}{1+2\gamma} - \frac{1}{(1+\gamma)^2}\right) + \\ & n_1 n_2 (n_2 - 1)\left(1 - \frac{2\gamma}{1+\gamma} + \frac{\gamma}{2+\gamma} - \frac{1}{(1+\gamma)^2}\right) + n_1 n_2 \frac{\gamma}{(1+\gamma)^2} \; ; \end{aligned}$$

$\gamma = 1$ provides the moment calculations under the null hypothesis.

The Monte Carlo resampling approach to computing the power function for the Wilcoxon rank sum test statistic $\tilde{S}$ is now described. The power of the two-sided rank

sum test statistic under the Lehmann alternative is computed as

$$\text{Power} = \Pr[\ |\tilde{S}| \geq c_{1-\alpha} | n_1, n_2; \gamma\ ]$$

where $c_{1-\alpha}$ denotes the $1-\alpha$ quantile from the null permutation distribution. Let x_i represent the outcome for subject i; $i = 1, 2, \ldots, n$ and $r_i \equiv r(x_i)$ indicate the pooled rank of the outcome data for subject i. In addition to the pooled rank value, each subject is identified by a group indicator; $\delta = 2$ if the subject is a member of the control group and $\delta = 1$ if the subject belongs to the experimental group. Thus, each subject is specified by the pair (r_i, δ_i). The permutation procedure rearranges the group indicators for the n subjects, while keeping the pooled ranks fixed. The permuted sample is denoted by $\{r_i, \delta_i^*\}_{i=1}^n$, where the asterisk indicates the permuted group assignments. Each new permutation sample $\{r_i, \delta_i^*\}_{i=1}^n$ results in a permutation statistic S^*. Generating a large number of permutation samples results in permutation statistics that produce a Monte Carlo estimate of the permutation distribution of S. Under the null hypothesis, all possible permutations of the group indicators are equally likely. Under the alternative hypothesis, however, the permutation samples have an unequal chance of occurrence. Thus, power computations must account for the unequal probabilities when generating the permutation distribution of S under an alternative hypothesis.

The derivation of the permutation algorithm under the alternative hypothesis is based on the hazard rate for each group in the Lehmann alternative, and the assignment of the control and experimental group subjects to the pooled ordered outcomes $x_{(1)} < x_{(2)} < \ldots < x_{(n)}$. The hazard rate is defined as

$$h(x) = \lim_{\Delta x \to 0} \frac{\Pr[x \leq X < x + \Delta x | X \geq x]}{\Delta x}.$$

The allocation of the subjects to the ordered outcomes occurs sequentially starting at the lowest value. For the two sample problem, the probability of selecting a member from group 1 to be placed into outcome $x_{(j)}$ is

$$\frac{n_{j1}h_1(x_{(j)})}{n_{j1}h_1(x_{(j)}) + n_{j2}h_2(x_{(j)})}$$

where n_{jg} represents the number of subjects in group g not chosen prior to $x_{(j)}$. This computation is considerably simplified under the Lehmann family of alternatives $h_1(x_{(j)}) = \gamma h_2(x_{(j)})$ and leads to

$$\Pr[\text{place group 1 indicator into } x_{(j)}|n_{j1}, n_{j2}; \gamma] = \frac{n_{j1}\gamma}{n_{j1}\gamma + n_{j2}}.$$

The probabilistic allocation of the group indicators to the order statistics is functionally independent of the observed values $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$; all that is required is the rank under evaluation and the subjects in each group not chosen prior to that rank. As a result, the ranks may be substituted for the ordered outcomes, enabling the permutation algorithm to be implemented at the design stage. In addition, under the Lehmann alternative, group assignment does not require knowledge of the underlying distributions. What is necessary, however, is the specific relationship between G_1 (h_1) and G_2 (h_2) governed by the parameter γ. The value $\gamma = 1$ generates the Monte Carlo approximation of the null permutation distribution. Additional discussion of the generation of the permutation distribution under the Lehmann alternative is found in Maritz [5] and Jennison [6]. In the case of ties, the assignment of midranks to the tied data values is the only modification needed for the Monte Carlo power estimates.

The results of the Monte Carlo approximation, the asymptotic approximation, and the exact power calculation are presented in Table 1. The results are based

on sample sizes of 5 and 10 in each group. The nominal significance level of the asymptotic normal approximation was calibrated to be equal to the size of the exact power calculation. One hundred thousand resamples were used to estimate the Monte Carlo power function.

The results demonstrate that the Monte Carlo approach provides an accurate approximation to the exact power calculations. With five subjects per group, the maximum deviation in the Monte Carlo power calculation is 0.003. In comparison, the asymptotic approximation has a consistent disparity of approximately 0.070 in the midrange of the power function. When the sample size increases to 10 per group, the maximum deviation is 0.001 for the Monte Carlo power computation and the deviation in the asymptotic normal calculation reduces to 0.036. Thus, even for a sample size of 10 per group, there is an accuracy gain using the Monte Carlo power approximation.

4. MONTE CARLO POWER APPROXIMATION FOR A K-SAMPLE RANK TEST

When the number of animals in each group is small, it is reasonable to employ exact power calculations for the two-group comparison. The feasibility of the exact power calculation diminishes when the number of groups is greater than two, because of the increased computational burden. Denoting by $(n_1, n_2, \ldots, n_k)$ the sample size in each of the k groups, with $n = n_1 + n_2 + \ldots + n_k$, the number of possible assignments of the n subjects to the k groups is $C^n_{n_1,n_2,\ldots n_k} = n!/n_1 n_2 \ldots n_k$. For example, five subjects assigned to each of four groups, results in $C^{20}_{5,5,5,5} \approx 1.17 \times 10^{10}$, or over 11 billion possible group assignments. This evaluation requires extensive computing

time, and even in the current computing environment, it cannot be undertaken on a routine basis. Somewhat surprisingly, the power calculation based on a large sample distribution of a k-sample rank statistic is considered too complicated for practical application. A noncentral χ^2 approximation based on a contiguous shift alternative does exist, but its accuracy under general alternatives is not reliable [1]. Clearly, the need for an accurate Monte Carlo resampling power computation is paramount when there are a small number of subjects and more than two groups.

Continuing with the tumor volume example, it is assumed that there are $k-1$ experimental therapies, with tumor volumes generated from the survival functions $G_1, G_2, \ldots, G_{k-1}$, and a control therapy with tumor volumes generated from G_k. The Lehmann family of alternatives is generalized to

$$H_0 \quad : \quad G_j(x) = G_k(x) \qquad j = 1, \ldots, k-1$$

$$H_A \quad : \quad G_j(x) = G_k^{\gamma_j}(x) \quad j = 1, \ldots, k-1; \quad \gamma_j \geq 1 \text{ with } \gamma_j > 1 \text{ for at least one } j.$$

The parameter γ_k is set equal to 1. Although the $k-1$ parameters γ_j $(j = 1, \ldots, k-1)$ in the alternative are specified relative to the control group, their specification determines the Lehmann alternative between any two groups i, j

$$G_i(x) = G_j^{\gamma_i/\gamma_j}(x) \qquad i, j = 1, 2, \ldots, k; \qquad \gamma_k = 1.$$

The alternative states that the odds a mouse in group i has a smaller tumor volume than a mouse in group j is γ_i/γ_j, or using a probabilistic interpretation, the probability a mouse in group i has smaller tumor volume than a mouse in group j is $\gamma_i/(\gamma_i + \gamma_j)$. The test statistic used for this k-sample alternative is the widely used Kruskal-Wallis rank statistic

$$W = \sum_{j=1}^{k} \{R_{.j} - (n+1)/2\}^2$$

where $R_{.j}$ represents the average of the pooled ranks in the jth group.

The Monte Carlo power calculation places subjects into the k groups, sequentially, starting from the lowest pooled rank. The group selection for each rank is decided stochastically, based on the conditional probability

$$\Pr[\text{place group i indicator into } x_{(j)}|n_{j1},\ldots,n_{jk};\gamma_1,\ldots,\gamma_{k-1}] = \frac{n_{ji}\gamma_i}{\sum_{l=1}^{k} n_{jl}\gamma_l} \qquad (3)$$

where n_{ji} represents the number of subjects remaining in group i when selecting for rank j. After group assignment is completed, the Kruskal-Wallis test statistic (W) is computed, and the algorithm is repeated a large number of times to approximate the distribution of W under a specific Lehmann alternative. The power function is computed using

$$Pr(W \geq c_{1-\alpha}|\gamma_1,\ldots,\gamma_k;n_1,\ldots,n_k)$$

where $c_{1-\alpha}$ is the Monte Carlo estimate of the $1-\alpha$ quantile for W when $\gamma_1 = \ldots = \gamma_k = 1$.

The accuracy of the Lehmann alternative Monte Carlo power function relative to the exact power computation is examined in the following two designs. The first design places 18 mice into three equal groups and the second design places 16 mice into four equal groups. These designs were chosen for the practical consideration that the exact power function could be computed for a variety of parameter combinations in a reasonable time frame. Note that $C^{18}_{6,6,6} = 17,153,136$ and $C^{16}_{4,4,4,4} = 63,062,996$. In all cases, 100,000 resamples were used to produce the Monte Carlo power estimates. The results are presented in Tables 2 and 3.

The power comparisons were computed at a nominal significance level of 0.05. The Lehmann alternative Monte Carlo approximation and exact power function were

computed for the Kruskal-Wallis test statistic over many different parameter combinations. The results for the three sample case with six observations per group, demonstrate a high degree of accuracy over all parameter combinations examined. The four sample case with four observations per group, represents a more discrete distribution for W, but the approximation remains accurate, with a maximum deviation of 0.003 over all parameter combinations considered.

The power calculations produced by the general Lehmann alternative can be replicated using other methods. An alternative specification of the Lehmann family states that a monotone but unknown transformation produces exponentially distributed outcomes. Thus for a rank test statistic, generating n_j $(j = 1, \ldots, k)$ random variables from an exponential distribution with parameter γ_j produces comparable results. Although not as constructive as the algorithm based on equation (3), the generation of exponential random variables, using the Lehmann alternative parameter corresponding to each group, is sufficient to reproduce the general Lehmann alternative Monte Carlo power function, and it is simpler to implement.

Tables 2 and 3 demonstrate that the general Lehmann alternative and the exponential distribution Monte Carlo power approximations produce comparable results. The two algorithms use similar steps but in a different order. For the general Lehmann alternative algorithm, the ranks are fixed and the group indicators are chosen based on the Lehmann alternative parameter vector and equation (3). For the exponential distribution power simulation, the group indicators are fixed and within each group the exponential random variables are generated with parameter γ_j $(j = 1, 2, \ldots, k)$ to compute the ranks.

5. EXTENSIONS

The Monte Carlo power computations are easily generalizable to k-sample designs for an ordered alternative or testing the superiority of $k-1$ experimental therapies relative to a control population. The probabilistic mechanism for the allocation of the n samples to the k groups remains the same and is determined by equation (3). The allocation is based on the number of subjects in each of the k groups $(n_1, n_2, \ldots, n_k)$ and the $k-1$ odds parameters $(\gamma_1, \gamma_2, \ldots, \gamma_{k-1})$.

For example, suppose k increasing dose levels of an experimental therapy are proposed and it is hypothesized that the odds of the reduction in tumor volume between successive dose levels is proportional to the dose administered. The ordered alternative hypothesis for this experiment, assuming the Lehmann family of alternatives, may be written as

$$1 \leq \gamma_{k-1} \leq \ldots \leq \gamma_1$$

where

$$\frac{\gamma_i}{\gamma_j} = \frac{d_i}{d_j}$$

and d_i represents the dose level administered to group i. The Jonckheere-Terpstra statistic provides greater power than the Kruskal-Wallis statistic under the ordered alternatives, and can be written in terms of the Wilcoxon rank sum statistics

$$J = \sum_{i<j} S_{ij} \qquad i = 1, 2, \ldots, k-1 \qquad j = 2, 3, \ldots, k$$

where S_{ij} is the Wilcoxon rank sum statistic applied to groups (i, j). The Monte Carlo power function is then computed for J.

Alternatively, when comparing $k-1$ experimental therapies to a single control

(group k), the test statistic

$$E = \sum_i S_{ik} \qquad i = 1, 2, \ldots k-1$$

can be applied to the k-sample allocation mechanism enabling Monte Carlo power computations to be carried out.

ACKNOWLEDGEMENT

The author would like to thank E. S. Venkatraman and a reviewer for comments that improved the content of this article.

REFERENCES

1. Lehmann EL. *Nonparametrics: Statistical Methods Based on Ranks.* Holden-Day: San Francisco, 1975.

2. Miller RG. *Beyond ANOVA, Basics of Applied Statistics.* Wiley: New York, 1986.

3. Troendle JF. Approximating the power of Wilcoxon's rank sum test against shift alternatives. *Statistics in Medicine* 1999; **18**:2763-2773.

4. Lehmann EL. The power of rank tests. *The Annals of Mathematical Statistics* 1953; **24**:23-43.

5. Maritz JS. *Distribution-Free Statistical Methods.* Chapman and Hall: London, 1981.

6. Jennison C. Bootstrap tests and confidence intervals for a hazard ratio when the number of observed failures is small, with applications to group sequential survival studies. In *Computing Science and Statistics: Proceedings of the Symposium on the Interface*, Page C, LePage S, (eds). Springer-Verlag: Berlin, 1990; 89-97.

Table 1. The columns in the table represent: the odds parameter, the exact power, the Monte Carlo power calculation, and the asymptotic power.

Two sample case - 5 observations per group

γ	exact power	Monte Carlo power	asymptotic power
1	0.056	0.056	0.056
2	0.144	0.145	0.134
3	0.273	0.273	0.238
4	0.386	0.385	0.329
5	0.477	0.475	0.406
6	0.549	0.549	0.473
7	0.606	0.605	0.530
8	0.652	0.654	0.580
10	0.721	0.718	0.662
15	0.817	0.817	0.797
20	0.866	0.867	0.874

Two sample case - 10 observations per group

γ	exact power	Monte Carlo power	asymptotic power
1	0.052	0.052	0.052
2	0.249	0.249	0.232
3	0.511	0.512	0.475
4	0.693	0.693	0.663
5	0.804	0.805	0.791
6	0.871	0.871	0.873
7	0.913	0.913	0.924

Table 2. Exact and Monte Carlo power calculations. Three sample case - 6 observations per group

		Monte Carlo power	
(γ_1, γ_2); $\gamma_3 = 1$	exact power	Lehmann alternative	Exponential distribution
(1,1)	0.050	0.050	0.049
(3,3)	0.308	0.307	0.304
(3,2)	0.246	0.247	0.244
(3,1)	0.302	0.302	0.299
(5,5)	0.552	0.553	0.548
(5,3)	0.467	0.468	0.463
(5,1)	0.573	0.574	0.569
(7,7)	0.694	0.693	0.690
(7,4)	0.616	0.618	0.613
(7,1)	0.737	0.734	0.731
(11,11)	0.830	0.830	0.825
(11,6)	0.778	0.778	0.774
(11,1)	0.886	0.884	0.882
(21,21)	0.932	0.932	0.930
(21,11)	0.911	0.910	0.909
(21,1)	0.973	0.973	0.972

Table 3. Exact and Monte Carlo power calculations. Four sample case - 4 observations per group

		Monte Carlo power	
$(\gamma_1, \gamma_2, \gamma_3)$; $\gamma_4 = 1$	exact power	Lehmann alternative	Exponential distribution
(1,1,1)	0.050	0.050	0.050
(3,3,3)	0.195	0.196	0.199
(3,2,2)	0.143	0.144	0.146
(3,2,1)	0.181	0.181	0.184
(3,1,1)	0.166	0.168	0.167
(5,5,5)	0.362	0.364	0.368
(5,3,3)	0.271	0.273	0.276
(5,4,2)	0.307	0.307	0.312
(5,1,1)	0.309	0.311	0.312
(10,10,10)	0.602	0.604	0.607
(10,7,4)	0.519	0.520	0.526
(10,5,5)	0.489	0.492	0.496
(10,1,1)	0.556	0.558	0.562
(16,16,16)	0.730	0.730	0.734
(16,8,8)	0.642	0.643	0.646
(16,11,6)	0.665	0.664	0.669
(16,1,1)	0.708	0.710	0.714
(30,30,30)	0.848	0.848	0.849
(30,15,15)	0.794	0.793	0.797
(30,20,10)	0.809	0.808	0.811
(30,1,1)	0.849	0.851	0.852