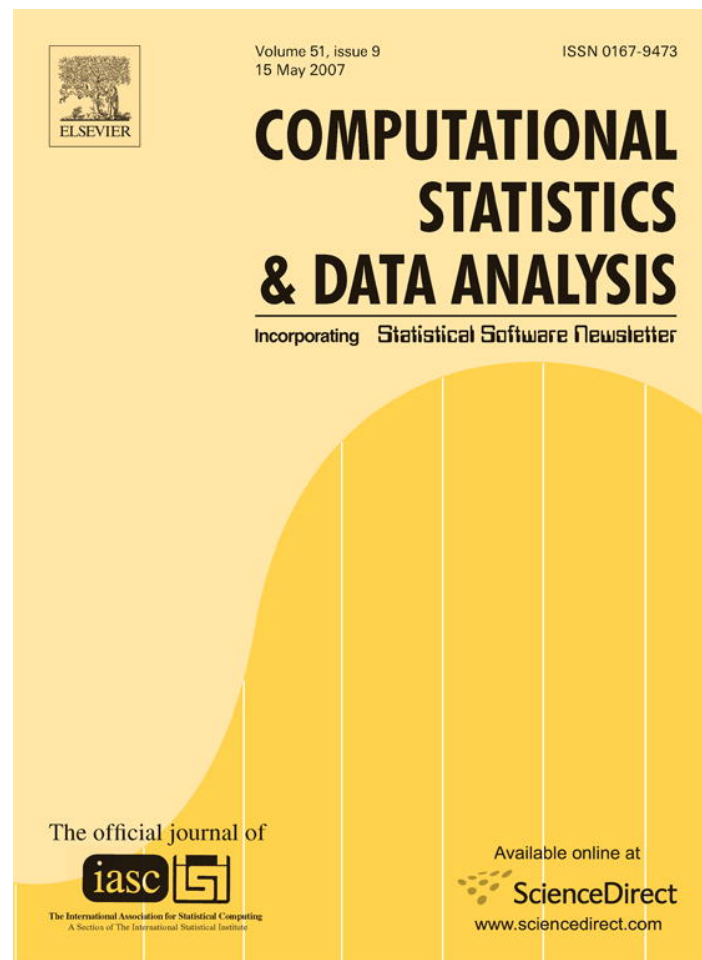


Provided for non-commercial research and educational use only.
Not for reproduction or distribution or commercial use.



This article was originally published in a journal published by Elsevier, and the attached copy is provided by Elsevier for the author's benefit and for the benefit of the author's institution, for non-commercial research and educational use including without limitation use in instruction at your institution, sending it to specific colleagues that you know, and providing a copy to your institution's administrator.

All other uses, reproduction and distribution, including without limitation commercial reprints, selling or licensing copies or access, or posting on open internet sites, your personal or institution's website or repository, are prohibited. For exceptions, permission may be sought for such use through Elsevier's permissions site at:

<http://www.elsevier.com/locate/permissionusematerial>

Bootstrapping estimation for estimating relative potency in combinations of bioassays

D.G. Chen^{a, b}

^a*Department of Mathematics and Statistics, South Dakota State University, Brookings, SD 57007, USA*

^b*Department of Applied Mathematics, South China Agricultural University, Guangzhou, PR China*

Received 25 May 2006; received in revised form 27 July 2006; accepted 29 July 2006

Available online 28 August 2006

Abstract

Estimation of relative potency in biological assay is very important in pharmaceutical and toxicological quality control. The statistical procedure to estimate this parameter produces a ratio of two other statistics resulting complexity for its distribution when inference will be made. In this paper, a bootstrapping procedure is proposed to approximate its population distribution and hence the estimation for this parameter in the combinations of bioassays. The proposed method overcomes the problems with distribution assumptions. We conclude with a comparison of this procedure to the well-known weighted mean method for a real data set.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Relative potency; Parallel-line bioassay; Slope-ratio bioassay; Weighted mean method; Sampling with replacement

1. Introduction

Estimation of relative potency in biological assay is very important in pharmaceutical and toxicological quality control. When a pharmaceutical company plans to place a new drug on the market, government authorization is required for inspection and control. One aspect of the inspection is to measure the drug's potency relative to a known (or standard) preparation, which is to estimate the relative potency of the new (or test) preparation to the standard preparation. If the test preparation is as potent as the standard preparation, then the relative potency is 1, or equivalently the log-relative potency is 0. To measure the relative potency, a comparative experiment has to be conducted and the resulting data analyzed by statistical bioassay to estimate the relative potency.

The combinations of bioassays arises when the same, or a similar, experiment is performed by multiple-laboratories and multiple-centers in the calibration of national or international standards for a particular substance (Rose and Gaines-Das, 1998). The literature is rich with methods for estimating this parameter as in Finney (1978), Hubert (1992) and Govindarajulu (2001). With the studies of combinations of bioassays, most of the results appear in Armitage (1970), Bennett (1962), Meisner et al. (1986), Williams (1978) and Chen et al. (1999).

The major difficulty in deriving the statistical properties of the estimator of the relative potency is that it is a ratio of two other statistics. Even when the normal assumption is assumed for the model, it will still result in a Cauchy distribution for the relative potency, for which the variance cannot be obtained analytically. In this situation, Fieller's theorem is one way to overcome this problem and obtain the confidence interval (CI), but this is too heavily dependent

E-mail address: din.chen@sdstate.edu.

on the normal assumption. If the assumption is only slightly violated, the estimator based on the Fieller's theorem will not be robust.

Another alternative is to use the Bayesian approach. The most recent results on this issue appear in Kim et al. (1991, 1992, 1993) and Chen et al. (1999). In those studies, a shrinkage estimator was proposed by using prior information which turns out to be a compromise of the prior information and the maximum likelihood estimator with the weights depending on the prior variance and hence the estimation of this prior variance must be discussed.

The continuing availability of inexpensive, high-speed computers has already reshaped many approaches to statistics. Computer-intensive algorithms have become increasingly popular statistical tools, both in applied and theoretical work. Much work has been done on algorithmic approaches and development, such as the EM algorithm from Dempster et al. (1977), or resampling techniques, such as the bootstrap from Efron (1979, 1982), Efron and Tibshirani (1986), Davison et al. (1986) and Hall (1992) and the Gibbs sampler from Gelfand and Smith (1990), Geman and Geman (1984).

In this paper, a bootstrapping method is proposed to estimate the relative potency from the combinations of bioassays, which simulate the distributional properties of the relative potency by repeated resampling of a resultant residual distribution. The details of the bootstrap method have been well chronicled by a series of references, such as Efron (1979, 1982). The key idea of the nonparametric bootstrap is that the population behavior can often be approximated by randomly resampling a given set of data and calculating the statistic of interest from each such "sample". The resulting set of values (a "bootstrap distribution") provides an approximate sampling distribution for the statistic. This ordinarily requires that random sampling be simulated—that is, a sampling procedure be undertaken.

The proposed procedure is applied to two data sets of combinations of parallel-line bioassays and a comparison to the well-known weighted mean method is also given in Section 4, followed by a discussion in Section 5. An R program is developed for this procedure. The R code and the data used in this paper can be obtained from the author.

2. Notation and background

2.1. Bioassay and the relative potency

Bioassay can be defined as a body of procedures in which the amount or strength of an agent or stimulus is determined by a response of a subject. The *subject* is usually an animal, a human tissue, or a bacterial culture; the *agent* is usually a drug; and the *response* is usually a change in a particular characteristic or even death of a subject. The estimation of the nature or potency of the agent by the response will be one of the primary objectives.

The most important parameter in bioassay is the relative potency. It is defined as the ratio of the effective constituent per unit amount of test preparation to that of the standard preparation. If d_t units of the test drug perform like ρd_s units of the standard (i.e., the test acts like a dilution of the standard), then the relative potency, denoted by ρ , of an agent to the standard agent is the ratio of the effective constituent in a unit amount of the test preparation to that of the standard. That is,

$$\rho = \frac{d_s}{d_t}. \quad (2.1)$$

For example, if the relative potency equals to 3, then the test preparation is three times as potent as the standard preparation. In other words, it will require only $\frac{1}{3}$ as much of test preparation to effect the same response as the standard preparation.

Statistically, it is common to use the log dose instead of the original dose. This transformation usually increases the linearity of the response curve. Therefore,

$$\log \rho = \log d_s - \log d_t.$$

Let $\mu = \log \rho$, $x_s = \log d_s$ and $x_t = \log d_t$. Then

$$\mu = x_s - x_t, \quad (2.2)$$

where μ is the log relative potency and the x 's are the so-called transformed doses.

In general, the dose–response relationships are established to model the potency of a dose from the knowledge of the response it produces. The general objective of the bioassay is to propose a suitable description of the dose–response

relationship for a biological substance. The general dose–response relationship for bioassays assumes the response y has the following structure:

$$\begin{aligned} y_s &= \alpha_s + \beta_s x_s + \varepsilon_s, \\ y_t &= \alpha_t + \beta_t x_t + \varepsilon_t \end{aligned} \quad (2.3)$$

and the $\varepsilon'_s \sim iid N(0, \sigma^2)$. This assumption could be tested by classical methods. For example, see [Atkinson \(1988\)](#).

This general formulation in (2.3) is not feasible to define a dose-dependent relative potency ([Finney, 1978](#)). The usual model in bioassay to estimate the relative potency is the parallel-line and slope-ratio bioassays.

In parallel-line bioassay, it is assumed that $\beta_s = \beta_t = \beta$ and $\alpha_t = \alpha_s + \mu\beta$. Therefore, the general model (2.3) is reduced to

$$\begin{aligned} y_s &= \alpha_s + \beta x_s + \varepsilon, \\ y_t &= \alpha_t + \beta x_t + \varepsilon. \end{aligned}$$

It can be seen that in parallel-line bioassay, the dose–response model for both the standard and test preparations is linear and the forms are parallel to each other. In fact, from (2.2), it is easily seen that μ is the horizontal distance between the two response lines. This formulation defines a parallel-line bioassay, where parallelism refers to a constant horizontal distance, μ , between the two response lines, where the parallelism can be tested statistically ([Finney, 1978](#); [Hubert, 1992](#)).

Alternatively, in the slope-ratio bioassays, it is assumed that the slopes are different, while the intercepts are the same. The general formulation (2.3) is then as follows:

$$\begin{aligned} y_s &= \alpha + \beta_s d_s + \varepsilon, \\ y_t &= \alpha + \beta_t d_t + \varepsilon, \end{aligned}$$

where $\beta_t = \beta_s \rho$. This implies that the relative potency in a slope-ratio bioassay is defined as: $\rho = \beta_t / \beta_s$.

[Finney \(1978\)](#), [Hubert \(1992\)](#) and [Govindarajulu \(2001\)](#) outlined the statistical tests for parallelism for parallel-line bioassay and common-intercept for slope-ratio bioassay. It is recommended that the observed dose–response data plotted initially before the statistical tests.

2.2. Combinations of bioassays

Suppose that the similar bioassay study is performed k times in a research laboratory or in k research centers in a multiple-center experiment setting. Such study involves several laboratories, departments, for example in international collaborative research for the standardization of a particular substance. In another situation, a pharmaceutical company wishes to assay a particular batch of product and may not be able to produce potency estimates of sufficient precision to satisfy a statutory requirement from one assay. Therefore, replicate experiments are desired and the issue for the combinations of bioassays arises. It would be a typical one bioassay setting if $k = 1$.

For the i th ($i = 1, \dots, k$) bioassay, the j th ($j = 1, \dots, m_i$) observation of the standard preparation at scalar input x_{sij} has response y_{sij} , and similarly the j th ($j = 1, \dots, n_i$) observation of the test preparation at scalar input x_{tij} has response y_{tij} . Then the combination of k bioassays can be modeled by the following formulation corresponding to (2.3):

$$\begin{aligned} y_{sij} &= \alpha_{si} + \beta_{si} x_{sij} + \varepsilon_{sij}, \\ y_{tij} &= \alpha_{ti} + \beta_{ti} x_{tij} + \varepsilon_{tij}. \end{aligned} \quad (2.4)$$

Similarly for parallel-line bioassays:

$$\begin{aligned} \beta_{si} &= \beta_{ti} (= \beta_i), \\ \alpha_{si} &= \alpha_{ti} - \mu_i \beta_i (= \alpha_i), \\ \delta_i &= \mu_i \beta_i, \end{aligned}$$

where μ_i are the log relative potencies, and the scalar input x 's are log-dose. For slope-ratio bioassays:

$$\begin{aligned}\alpha_{si} &= \alpha_{ti} (= \alpha_i), \\ \beta_{ti} &= \rho_i \beta_{si} (= \rho_i \beta_i), \\ \delta_i &= \beta_{ti},\end{aligned}$$

where ρ_i are the relative potencies, and the scalar input x 's are levels of a dose. By using matrix notation, the model in (2.4) can be written in the following linear model format:

$$Y_i = X_i^t \Psi_i + \epsilon_i, \quad (2.5)$$

where $Y_i = (y_{si1}, \dots, y_{sim_i}, y_{ti1}, \dots, y_{tin_i})^t$ is a vector of $(n_i + m_i) \times 1$ responses; and Ψ_i is the 3×1 parameter vector $\Psi_i = (\alpha_i, \delta_i, \beta_i)^t$. X_i is the $3 \times (m_i + n_i)$ design matrix. For parallel-line bioassays,

$$X_i = \begin{bmatrix} 1 & \dots & 1 & 1 & \dots & 1 \\ 0 & \dots & 0 & 1 & \dots & 1 \\ x_{si1} & \dots & x_{sim_i} & x_{ti1} & \dots & x_{tin_i} \end{bmatrix}.$$

For slope-ratio bioassays,

$$X_i = \begin{bmatrix} 1 & \dots & 1 & 1 & \dots & 1 \\ 0 & \dots & 0 & x_{ti1} & \dots & x_{tin_i} \\ x_{si1} & \dots & x_{sim_i} & 0 & \dots & 0 \end{bmatrix},$$

where \mathbf{I} is the $(m_i + n_i) \times (m_i + n_i)$ identity matrix and $\epsilon_i \sim (0, \sigma^2 \mathbf{I})$. Note that in order to simplify the discussion, only the issues for combining univariate bioassays will be used to illustrate the procedure in the following sections. The similar analysis would occur for the multivariate version of combining bioassays and the proposed method in this paper can be easily extended to the multivariate setting.

A standard linear model technique is first used to obtain the estimators for the model parameters:

$$\hat{\Psi}_i = (X_i X_i^t)^{-1} X_i Y_i, \quad (2.6)$$

and

$$S_i^2 = \frac{1}{m_i + n_i - 3} Y_i^t \left[\mathbf{I} - X_i^t (X_i X_i^t)^{-1} X_i \right] Y_i. \quad (2.7)$$

Note that $\hat{\Psi}_i \sim (\Psi_i, \sigma^2 (X_i X_i^t)^{-1})$. And the pooled sample variance $S^2 = \sum_{i=1}^k (m_i + n_i - 3) S_i^2 / \sum_{i=1}^k (m_i + n_i - 3)$ can be used to estimate the model variance σ^2 .

2.3. Estimation for combinations of bioassays

For the estimation of combinations of bioassays, much work can be found on [Bennett's least square method \(1962\)](#), [Armitage's maximum likelihood estimator \(1970\)](#), [Finney's weighted mean method \(1978\)](#) and others such as in [Williams \(1978\)](#) and [Srivastava \(1986\)](#). All those studies involved a distributional assumption.

The weighted mean method is probably the most widely used among these methods, which can be described in detail as follows:

From (2.6), the maximum likelihood estimator of μ_i from each bioassay can be formed as $\hat{\mu}_i = (\hat{\delta}_i / \hat{\beta}_i)$ ($i = 1, \dots, k$). To construct the estimator for the combinations of bioassays, [Finney \(1978\)](#) proposed a weighted mean method with the weights being inversely proportional to the estimated variance. That is,

$$\hat{\mu} = \frac{\sum_{i=1}^k w_i \hat{\mu}_i}{\sum_{i=1}^k w_i}, \quad (2.8)$$

where $w_i = 1/\widehat{V}(\hat{\mu}_i)$. It can be easily shown from Delta method (so called “propagation of error formula”) that

$$V(\hat{\mu}_i) = V\left(\frac{\hat{\delta}_i}{\hat{\beta}_i}\right) \approx \frac{1}{\beta_i^4} (\beta_i, -\delta_i) \text{Cov}\left(\begin{matrix} \hat{\delta}_i \\ \hat{\beta}_i \end{matrix}\right) \begin{pmatrix} \beta_i \\ -\delta_i \end{pmatrix}. \quad (2.9)$$

From (2.6) and (2.7), the weights can be estimated from the inverse of (2.9), and hence $\hat{\mu}$ in (2.8). An estimate of the asymptotic variance of $\hat{\mu}$ can be also shown to be simply $1/\sum_{i=1}^k w_i$. If the normal distribution for $\hat{\mu}$ is assumed, the CI with $1 - \alpha$ CI can be constructed as

$$\left(\hat{\mu} - z_{\alpha/2}\sqrt{\widehat{V}(\hat{\mu})}, \hat{\mu} + z_{\alpha/2}\sqrt{\widehat{V}(\hat{\mu})}\right). \quad (2.10)$$

The test statistic to test the homogeneity of the combinations of bioassays can be found from Finney (1978) as follows

$$\tilde{\chi}_{k-1}^2 = \sum_{i=1}^k w_i \hat{\mu}_i^2 - \frac{\left(\sum_{i=1}^k w_i \hat{\mu}_i\right)^2}{\sum_{i=1}^k w_i}, \quad (2.11)$$

following a χ^2 distribution with $k - 1$ degree of freedom.

3. Bootstrapping estimation for relative potency

Armitage et al. (1974) have shown that the weighted mean method in (2.8) gives a biased lower estimate of relative potency as well as χ^2 statistic for testing the homogeneity of the relative potencies in (2.11) for small number of observations even though the bias is negligible.

In addition, the CI in (2.10) is based on the normality assumption. In fact it is well-known that the ratio of any two normally distributed statistics is not normally distributed and it is Cauchy distributed.

To overcome the difficulties from the normal distribution assumption in (2.10) and (2.11) and the unnecessary approximation in the estimation of relative potency, the bootstrapping procedure is proposed in this section to approximate the sample distribution of $\hat{\mu}$ in order to obtain its CI.

The bootstrapping resampling technique was invented by Efron (1978, 1979, 1982), which has attracted much interest as a basis for approximate statistical inference when other methods are suspect or unavailable. The idea of this technique is to mimic the process of selecting many samples to get the probability of $\hat{\mu}$ that the values of their statistics fall within various intervals. The name bootstrap reflects the fact that one available sample gives rise to many others. Since the bootstrap generates a sampling distribution from the sample itself which is called the bootstrap distribution, the usual underlying assumption often involving the normal distribution is not necessary.

The fundamental assumption of bootstrapping is that the observed data are the representative of the underlying population. By resampling observations from the observed data, the process of sampling observations from the population is mimicked. Freedman (1981) discussed the bootstrapping method for regression and correlation models and showed that the bootstrap approximation is valid to the least squares estimates of model parameters. There are extensive bootstrap schemes to construct CIs which can be found from Efron and Tibshirani (1986, 1993) and Shao and Tu (1995).

For the combinations of bioassays, the bootstrapping procedure consists of the following steps:

Step 1: Use the original sample $\{\mathbf{X}_i, \mathbf{Y}_i\}$ from the data $\{x_{sij}, y_{sij}, x_{tij}, y_{tij}\}$ for the general linear model (2.5) to compute the parameter LS estimate of Ψ_i and S_i^2 as in (2.6) and (2.7). Let residual vector $r_i = Y_i - X_i^t \hat{\Psi}_i$ ($i = 1, \dots, k$).

Step 2: For each i , draw a random sample, say $\{r_{il}^*\}_{l=1}^{m_i+n_i}$, from the residual vector $\{r_{il}\}_{l=1}^{m_i+n_i}$ with replacement.

Step 3: Let $Y_i^* = X_i^t \hat{\Psi}_i + r_i^*$. This results a bootstrap sample $\{\mathbf{X}_i, \mathbf{Y}_i^*\}$.

Step 4: Use this bootstrap sample to calculate $\hat{\mu}$ from (2.8).

Step 5: Repeat Steps 2–4 for a number of times, say B times (e.g. 10,000 times in the examples), and a bootstrapping sample with B observations is produced for $\hat{\mu}$. Therefore, an empirical bootstrapping sample distribution of $\hat{\mu}$ can be obtained from the bootstrapping sample.

This empirical distribution is called the bootstrap distribution and can be used to approximate the population distribution of μ . From this bootstrap distribution, we can make inference for μ to provide a distribution-free point estimator and confidence region.

The bootstrap estimate for the log-relative potency is the 50% percentile of the bootstrapping sample. To obtain the $1 - \alpha$ CI for the log-relative potency μ , we sort the bootstrap sample and locate the $\alpha/2$ -percentile and $(1 - \alpha/2)$ -percentile values in the sorted bootstrap sample, which are the lower and upper limits for bootstrap procedure. Notice that the shape of this distribution is free from any underlying probability density function.

4. Application to the combinations of two bioassays on vitamin D₃

Two assays of vitamin D₃ in the same oil are reported on p. 287 in Finney (1978) and can be found in the accompanying R program. Each assay used the line test scores as responses. The two assays used eight from each of six litters, two per litter for each dose. In the first assay, the dose of the test preparation were chosen on the basis of an assumed potency of $\frac{4}{3}$ (IU/mg) and in the second this was changed to $\frac{16}{15}$ (IU/mg). The dose ratio was 2 in both preparations.

There is no statistically significant litter effect, and no violations of linear relationship and parallelism (Finney, 1978). The data and the fitted linear dose–response relationship are shown in Fig. 1.

Therefore, these two bioassays can be cast in a linear model (2.5) with the estimated parameters from (2.6) and (2.7), $\hat{\mu}_i$ and the estimated $\widehat{var}(\hat{\mu}_i)$ are summarized in Table 1.

The χ^2 statistic in (2.11) for the test of homogeneity of the four bioassays is 1.923 with p -value of 0.17, indicating that it is valid to combine the two bioassays.

The weighted mean estimate of log-relative potency can then be estimated from (2.8), which is $\hat{\mu} = 0.2036$, with estimated variance of 0.00232. The 95% CI assuming a normal distribution for $\hat{\mu}$ from (2.10) is (0.1092, 0.2979) with

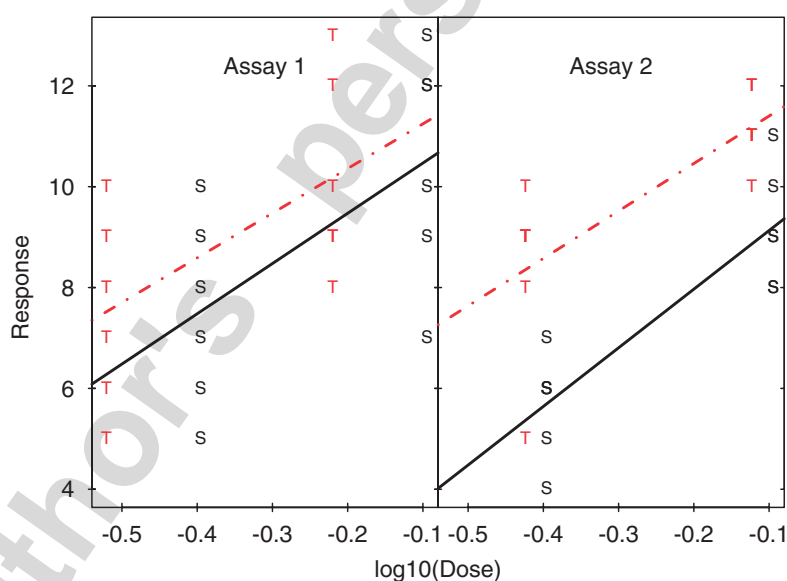


Fig. 1. Data and the fitted dose–response relationship with “S” for standard and “T” for test preparation. The horizontal axis is the log-dose and the vertical axis is the observed response.

Table 1
Summary of the two vitamin D₃ assays

Assay i	$\hat{\alpha}_i$	$\hat{\beta}_i$	$\hat{\delta}_i$	S_i^2	$\hat{\mu}_i$	$\widehat{var}(\hat{\mu}_i)$
1	11.329	1.009	9.412	3.786	0.107	0.00715
2	10.019	2.628	10.519	1.476	0.249	0.00343

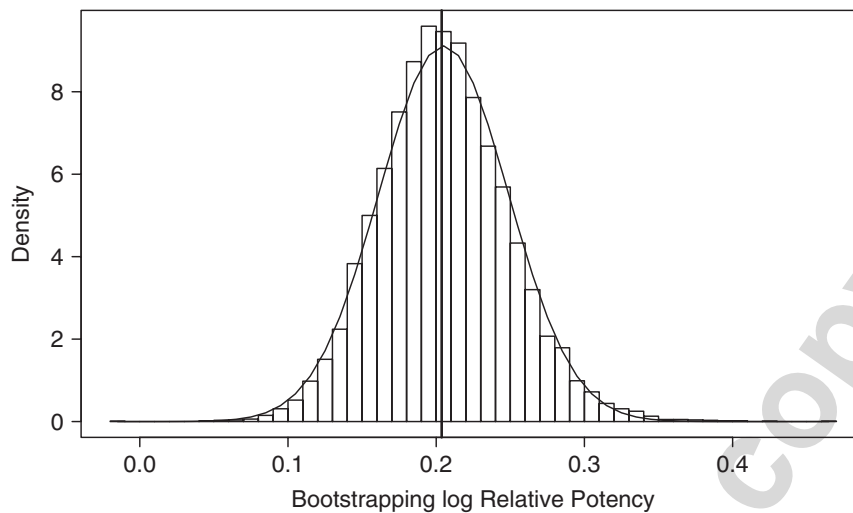


Fig. 2. Bootstrapping distribution with normal density curve.

a length of 0.1887. Then the estimated relative potency can be estimated by an anti-log transformation as 1.2258, with a CI of (1.2859, 1.9858), and with a length of 0.6999, indicating that the test preparation is statistically significantly more potent than the standard preparation, which confirms what we see in Fig. 1.

Now we turn to the bootstrapping procedure proposed in Section 3. We run the bootstrapping procedure for 10,000 times and a bootstrapping sample with 10,000 observations is obtained.

The bootstrap estimate of the log-relative potency from the 50% percentile of the bootstrapping distribution is 0.2037. To obtain the CI for the relative potency μ , we sort the bootstrapping samples and locate the values which are 2.5% percentile and 97.5% percentile in the sorted bootstrapping samples, which are the lower and upper limits for bootstrapping procedure. For this data, the bootstrap 95% CI is (0.1229, 0.2956), with a length of 0.1727. Therefore, the estimated relative potency is 1.5984 with a 95% CI of (1.3273, 1.9751), with a length of 0.6478, which is shorter than the CI from the weighted mean method.

To see the empirical bootstrapping distribution, we break the interval from the minimum value to maximum value of the bootstrapping sample into 50 subintervals and count the frequency for each subinterval. Then we plot the frequency at the middle point of each subinterval to get the bootstrapping empirical distribution (Fig. 2).

5. Discussion

This paper presented a useful bootstrapping method to estimate the relative potency in the combinations of bioassays, which performed better than the well-known weighted mean estimator with shorter confidence interval. The proposed method can be easily implemented. A R program with the data used in this paper can be obtained from the author.

Although the presentation and examples are in univariate format, the method can be easily extended to the multivariate format for the combinations of multivariate bioassays when multivariate responses are observed.

It is worthwhile emphasizing that $\hat{\mu} = \hat{\delta}/\hat{\beta}$ has no finite variance since $\hat{\delta}$ and $\hat{\beta}$ are normally distributed, which lead to a Cauchy distribution for $\hat{\mu}$. For this reason, the bootstrapping method would be more appropriate and reasonable with less distributional requirements.

It can also be seen from the two examples that the bootstrapping estimate for the relative potency was slightly higher than the estimate from the weighted mean method, which is consistent with the conclusion in Armitage et al. (1974).

Acknowledgments

The author would like to thank Professor Dwight Galster, the Associate Editor and two anonymous referees, for their comments and suggestions, which significantly improved this manuscripts. This material is based upon work supported by the National Science Foundation/EPSCoR Grant no. 0091948 and by the State of South Dakota.

References

- Armitage, P., 1970. The combination of assay results. *Biometrika* 57, 665–666.
- Armitage, P., Bailey, J.M., Petrie, A., Annable, L., Stack-Dunne, M.P., 1974. Studies in the combination of bioassay results. *Biometrics* 30, 1–9.
- Atkinson, A.C., 1988. *Plots, Transformations and Regression*. Clarendon Press, Oxford.
- Bennett, B.M., 1962. On combining estimates of relative potency in bioassay. *J. Hygiene* 60, 379–385.
- Chen, D.G., Carter, E.M., Hubert, J.J., Kim, P.T., 1999. Empirical Bayes estimation for combinations of multivariate bioassays. *Biometrics* 55, 1038–1043.
- Davison, A.C., Hinkley, D.V., Schechtman, E., 1986. Efficient bootstrap simulation. *Biometrika* 73, 555–566.
- Dempster, A.P., Laird, N., Rubin, D.B., 1977. Maximum likelihood form incomplete data via the EM algorithm (with discussion). *J. Roy. Statist. Soc. Ser. B* 39, 1–38.
- Efron, B., 1978. Nonparametric estimates of standard error: the jackknife, the bootstrap and other methods. *Biometrika* 68, 589–599.
- Efron, B., 1979. Bootstrap methods: another look at the jackknife. *Ann. Statist.* 7, 1–26.
- Efron, B., 1982. The jackknife, the bootstrap, and other resampling plan. *Conference Series in Applied Mathematics, Report 38, Society for Industrial and Applied Mathematics, Philadelphia*.
- Efron, B., Tibshirani, R., 1986. Bootstrap methods for standard errors, confidence interval, and other measures of statistical accuracy. *Statist. Sci.* 1, 54–75.
- Efron, B., Tibshirani, R., 1993. *An Introduction to the Bootstrap*. Chapman & Hall, San Francisco.
- Finney, D.J., 1978. *Statistical Method in Biological Assay*. third ed. C. Griffin, London.
- Freedman, D.A., 1981. Bootstrapping regression models. *Ann. Statist.* 9, 1218–1228.
- Gelfand, A.E., Smith, A.F.M., 1990. Sampling based approaches to calculating marginal densities. *J. Amer. Statist. Assoc.* 85.
- Geman, S., Geman, D., 1984. Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.* 6, 721–741.
- Govindarajulu, Z., 2001. *Statistical Techniques in Bioassay*. Karger, Basel.
- Hall, P., 1992. *The Bootstrap and Edgeworth Expansion*. Springer, New York.
- Hubert, J.J., 1992. *Bioassay*. third ed. Kendall-Hunt, Dubuque, Iowa.
- Kim, P.T., Carter, E.M., Hubert, J.J., 1991. Estimating relative potency using prior information. *Biometrics* 47, 295–301.
- Kim, P.T., Cho, G.W., Carter, E.M., Hubert, J.J., 1992. General estimation of relative potency using prior information. *J. Biopharm. Statist.* 2 (2), 171–180.
- Kim, P.T., Carter, E.M., Hubert, J.J., Hand, K.J., 1993. Shrinkage estimator of relative potency. *J. Amer. Statist. Assoc.* 88 (422), 615–621.
- Meisner, M., Kushner, H.B., Laska, E.M., 1986. Combining multivariate bioassays. *Biometrics* 42, 421–427.
- Rose, M.P., Gaines-Das, R.E., 1998. Characterisation, calibration and comparison by international collaborative study of international standards for the calibration of therapeutic preparations of FSH. *J. Endocrinol.* 158, 97–114.
- Shao, J., Tu, D., 1995. *The Jackknife and Bootstrap*. Springer, New York.
- Srivastava, M.S., 1986. Multivariate bioassay, combination of bioassays and Fieller's theorem. *Biometrics* 42, 131–141.
- Williams, D.A., 1978. An exact confidence region for a relative potency estimated from combined bioassays. *Biometrics* 34, 659–661.