BP1CI
Version 2.0: May 2008
Confidence Intervals for One-Sample Binomial and Poisson Trials

Barry W. Brown
Floyd M. Spears
Dan M. Serachitopol

The University of Texas M.D. Anderson Cancer Center
Department of Biostatistics and Applied Mathematics, Box 237
1515 Holcombe Boulevard
Houston, TX 77030 (USA)

Contact: BWB at above address or bwb@mdanderson.org.

1 Technicalities

1.1 Obtaining the Code

The source for this code (and all code written by this group) can be obtained from the following URL:

http://biostatistics.mdanderson.org/SoftwareDownload

2 Legalities

We want our computer programs to be as widely used as possible, consequently unrestricted use of them within a computer is permissible. Use within commercial enterprises for commercial purposes is explicitly allowed as is academic, research, and government or any other use.

Copies of any form of the code (source or executable) and the documentation can be made and redistributed freely. You may give copies to students, coworkers, friends, and enemies. We do require that the copyright and LEGALITIES section of the code and documentation be preserved in such copies.

Modifications and improvements to the code can similarly be freely copied and distributed. Modified code be clearly marked as such so that we are not requested to fix it. We would appreciate a copy of any improvements for incorporation into later versions of our programs.

Inclusion of our code in software archives and its distribution as freeware is encouraged.

Incorporation into Commercial Packages

Incorporation of our code into a commercial package that will be sold requires written permission from an administrative officer of the Department of Biomathematics. Our current policy is to grant such permission without payment providing:

(1) We are informed of and approve the nature of the product being developed or prepared for distribution.

(2) Appropriate acknowledgment of our code is made in the documentation accompanying the product.

This program contains code from the following publications, and code from ACM publications is subject to the ACM policy (below).

2.1 References

2.2 Incomplete Gamma

DiDinato, A. R. and Morris, A. H. (1986) "Computation of the incomplete gamma function ratios and their inverse." *ACM Trans. Math. Softw. 12*, 377-393.

2.2.1 Incomplete Beta

DiDinato, A. R. and Morris, A. H. (1993) "Algorithm 708: Significant Digit Computation of the Incomplete Beta Function Ratios." *ACM Trans. Math. Softw. 18*, 360-373.

2.2.2 Finding a Zero of a Monotone Function

Alefeld, G. E., Potra, F. A., Shi, Y. (1995) "Algorithm 748: Enclosing Zeros of Continuous Functions.", by G. E. Alefeld, F. A. Potra, YiXun Shi, *ACM Trans. Math. Softw., 21*, No. 3, 327-344

2.3 ACM Policy on Use of Code

Here is the software Policy of the ACM.

> Submittal of an algorithm for publication in one of the ACM Transactions implies that unrestricted use of the algorithm within a computer is permissible. General permission to copy and distribute the algorithm without fee is granted provided that the copies are not made or distributed for direct commercial advantage. The ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and/or specific permission.
>
> Krogh, F. (1997) "Algorithms Policy." *ACM Tran. Math. Softw. 13*, 183-186.

PURPOSE. THE ENTIRE RISK AS TO THE QUALITY AND PERFORMANCE OF THE PROGRAM IS WITH YOU. SHOULD THIS PROGRAM PROVE DEFECTIVE, YOU ASSUME THE COST OF ALL NECESSARY SERVICING, REPAIR OR CORRECTION.

IN NO EVENT SHALL THE UNIVERSITY OF TEXAS OR ANY OF ITS COMPONENT INSTITUTIONS INCLUDING M. D. ANDERSON HOSPITAL BE LIABLE TO YOU FOR DAMAGES, INCLUDING ANY LOST PROFITS, LOST MONIES, OR OTHER SPECIAL, INCIDENTAL OR CONSEQUENTIAL DAMAGES ARISING OUT OF THE USE OR INABILITY TO USE (INCLUDING BUT NOT LIMITED TO LOSS OF DATA OR DATA OR ITS ANALYSIS BEING RENDERED INACCURATE OR LOSSES SUSTAINED BY THIRD PARTIES) THE PROGRAM.

(Above NO WARRANTY modified from the GNU NO WARRANTY statement.)

3 Documentation of the Program

4 Introduction

4.1 Overview of the Problem Addressed:

This program calculates exact symmetric confidence limits on the probability of event for the one-sample binomial distribution or on the mean number of events for the one-sample Poisson distribution using the Clopper-Pearson method.

4.1.1 Reference

Clopper, C.J. and Pearson, E.S. (1934) "The use of confidence or fiducial limits illustrated in the case of the binomial." *Biometrika,* 26, 404.

4.2 Method

The level of the confidence interval, γ must be specified for either the binomial or Poisson distribution. Here *tail* is used as a probability with values between 0 and 1. In the program, the level is specified as a percentage – 0 to 100; thus a program specification of 95% corresponds to a probability of 0.95. Let $tail = (1 - \gamma)/2$.

For the binomial distribution, the values used in the calculation are the the total number of binomial trials, N, and the number of events or successes, S.

For the Poisson distribution, the only value used is the total number of events or successes, S.

The calculation is conceptually very simple:

- The lower confidence limit for the binomial is the value of the probability of success, p, such that the probability of S or more successes is *tail*.
- The lower confidence limit for the Poisson is the value of the mean number of successes, λ, such that the probability of S or more successes is *tail*.
- The upper confidence limit for the binomial is the value of the probability of success, p, such that the probability of S or fewer successes is *tail*.
- The lower confidence limit for the Poisson is the value of the mean number of successes, λ, such that the probability of S or fewer successes is *tail*.

5 Examples of Use of the Program

5.1 Example I – Binomial

A new drug is tested on 30 patients. Of the 30 treated, 12 respond and 18 do not respond. What is a 95% confidence limit on the (true) response rate.

The program first displays this banner.

```
                              BP1CI
                       Version 2.0: May 2008

                  Confidence Intervals for One-Sample
                      Binomial and Poisson Trials

                            Barry W. Brown
                           Floyd M. Spears
                         Dan M. Serachitopol

Copyright 2008 for:
         The University of Texas M.D. Anderson Cancer Center
         Department of Biomathematics, Box 237
         1515 Holcombe Boulevard
         Houston, TX 77030  (USA)


Contact: BWB at above address or bwb@mdanderson.org

This program can be freely copied and (noncommercially) distributed.

 Press the Return or Enter key to continue ...

<CR>
```

Begin Commentary

The following dialog appears giving the user the choice of exiting the program, changing the setup, or calculating a confidence interval.

The current settings are initially: (1) the confidence level is set to 95%; (2) the Binomial (as opposed to the Poisson) distribution will be used; and (3) data will be entered for the binomial distribution as number of successes and number of failures (instead of number of successes and total number of trials). If these settings are not appropriate choose '1' from the menu to change them.

If the settings are appropriate, one can immediately compute a confidence interval by choosing '2'. We do so in this example.

End Commentary

```
Enter '0' to exit this program
      '1' to change the setup (confidence level, binomial or Poisson
          distribution, method of entry of binomial data)
          Current settings:
                Distribution - Binomial
                Confidence level -  95.0%
                Binomial entry method -
                     Number of Successes, Number of Failures
      '2' to calculate a confidence interval

> 2
```

Begin Commentary

The following dialog for the binomial case appears.

End Commentary

```
Enter number of successes  then the number of failures.  Separate
the two entries with a space.
```

```
 > 12 18
======================================================================
                             Data

     N Successes:       12.0  N Failures:       18.0   N Total:       30.0

     95.0% Confidence Interval and Estimate of the Probability of Success

     Low Bound: 0.2266   Estimate: 0.4000    High Bound: 0.5940

======================================================================
```

The 95% confidence limit on the proportion responses is 0.23 to 0.59. In a cancer Phase II trial, proof that the response rate is greater than 20% frequently is sufficient to justify further testing of the drug.

5.2 Example II.

We desire to estimate the 99% confidence limit on the mean number of raisins per loaf of a particular type of bread, assumming that the number of raisins per loaf is distributed Poisson. Totalling over 100 examined loaves of bread, we find that there are 10000 raisins. To find the confidence limit on the mean, we first find it for the total (10000) and divide the bounds by the sample size (100 loaves).

Begin Commentary

> At the main menu, we choose option '1' – change setup. We then change the confidence level to 99% and the distribution to Poisson.

End Commentary

```
     Enter '0' to exit this program
           '1' to change the setup (confidence level, binomial or Poisson
               distribution, method of entry of binomial data)
               Current settings:
```

```
                    Distribution - Binomial
                    Confidence level -  95.0%
                    Binomial entry method -
                         Number of Successes, Number of Failures
         '2' to calculate a confidence interval
 > 1

    Enter the confidence level desired as a percentage.  This must be
    a number  between 0.1 and  99.9999.  This level  is, intuitively,
    the probability  that the true  value lies within  the confidence
    limits.  Popular values are 80%, 90%, 95%, and 99%.

 > 99

    Enter 'b' or 'B' for  the binomial distribution; enter 'p' or 'P'
    for the  Poisson distribution.  Enter the character  as the first
    on  the  next  line of  input;  do  not  enter the  single  quote
    character.

> p
```

Begin Commentary

We now choose option '2' – calculate the confidence interval. Upon entering our data (10000 raisins counted) we obtain the following confidence interval).

End Commentary

```
    Enter '0' to exit this program
          '1' to change the setup (confidence level, binomial or Poisson
              distribution, method of entry of binomial data)
              Current settings:
                    Distribution - Poisson
                    Confidence level -  99.0%
                    Binomial entry method -
```

```
                          Number of Successes, Number of Failures
         '2' to calculate a confidence interval
 > 2

    Enter the number of events observed.

 > 10000
======================================================================
                   Data

        The Number of Events Observed is:   10000.

        99.0 Confidence Interval on the Mean Number of Events

        Low Bound:   9745.3           High Bound:   10260.

======================================================================
```

Dividing each bound by 100 (our sample size), we find that the 99% confidence limits on the mean number of raisins per loaf is $(0.975, 1.025)$.