Physica Medica (2009) **25**, 43—46

available at www.sciencedirect.com

ScienceDirect

journal homepage: http://intl.elsevierhealth.com/journals/ejmp

TECHNICAL NOTE

Confidence limits of the probability of success in animal experiments and clinical studies: A Bayesian approach

N. Stavreva, P. Stavrev*, Gino B. Fallone

Department of Medical Physics, Cross Cancer Institute, 11560 University Avenue, Edmonton, Alberta T6G1Z2, Canada

Received 19 September 2007; received in revised form 4 January 2008; accepted 7 January 2008
Available online 15 February 2008

KEYWORDS
Binomial confidence limits;
Binary data

Abstract *Purpose:* To determine from the number of trials, n, and the number of observed successes, k the most probable value, the variance and the confidence limits of the probability of success, p, in animal experiments and clinical studies subject to binomial statistics.
Method: In such experiments the probability of success is an unknown parameter. The Bayesian approach to the problem is advocated, based on constructed distribution of the probability of success.
Results: A simple Matlab code for the calculation of the confidence limits according to the proposed method is provided. The most probable, the mean, the variance and the confidence limits are calculated applying the usual definitions of these characteristics.
Conclusion: The proposed method works for any number of trials — large and small and all possible values of the number of successes, including $k = 0$ and $k = n$, providing exact formulae for the calculation of the confidence limits in all cases.

Introduction

Clinical studies and animal experiments where all outcomes of a treatment can be classified just as positive (success) or negative (failure) outcomes are subject to binomial statistics. In such studies, the unknown parameter is the probability of the positive/negative outcome, p. It is the determination of its most probable value together with its mean, variance and confidence limits which is of primary interest in these studies. The importance of reporting confidence intervals for this type of studies was emphasized by Shakespeare and Holecek [1] The idea of confidence limits and of fiducial intervals has been introduced in the early thirties of last century by several prominent statisticians [2—6]. The first work concerning the ''confidence or fiducial limits'' in the binomial case was published by Clopper and Pearson [3]. An approach based on the idea of the fiducial

* Corresponding author.
E-mail address: pavel.stavrev@gmail.com (P. Stavrev).

1120-1797/$ - see front matter
doi:10.1016/j.ejmp.2008.01.001

limits is discussed in Ref. [7]. It is worth mentioning that the fiducial approach was criticized by Neyman [5,6]. Indeed, this approach to the problem as described in Collet [8] experiences several difficulties. For instance, if k is the number of positive/negative events and n is the number of trials, the standard formula [8,9] for the variance (uncertainty) of p is $\sigma^2=(\widehat{p}(1-\widehat{p}))/n$, where $\widehat{p}=k/n$ is the estimated (most probable) value of p. It is obvious, that these formulae should not be used in marginal cases, namely when $n\widehat{p}$ and $n(1-\widehat{p})$ are smaller than around 5, i.e. when the number of trials is rather small and/or the number of successes or the number of failures is around zero.

Due to the fact that the random variable p is defined in a closed interval, it is more likely that its probability distribution function is asymmetric. This makes the report of confidence intervals, rather than the variance, a more adequate issue in this case. According to the traditional definition of the lower and upper limits of the 100 $(1-\alpha)$% confidence interval for p given in the orthodox statistical books (see Refs. [8,9] for example) the following equations must be solved:

$$\sum_{j=k}^{n}\binom{n}{j}p_L^j(1-p_L)^{n-j}=\frac{\alpha}{2};$$

$$\sum_{j=0}^{k}\binom{n}{j}p_U^j(1-p_U)^{n-j}=\frac{\alpha}{2} \tag{1}$$

for the lower limit p_L and for the upper limit p_U, respectively. These equations follow from the requirement that $\sum_{X(j)}P(j|p)=1-\alpha$, where the summation is over the set, $X(j)$, of acceptable values of j for a given p. Neyman [5] has proved that the form $\sum_{X(j)}P(j|p)$ cannot be set to be equal to a constant value (say, $1-\alpha$) independent of the value of p, if P is the binomial distribution with parameters j and p. For instance, in case of $k=0$ the left-hand side of the equation for p_L is equal to 1, independent of the value of p_L; therefore, it cannot be satisfied for any value of p_L. The same reasoning applies to the equation for p_U in case of $k=n$ when its left-hand side is equal to 1, independent of the value of p_U and it cannot be therefore satisfied for any value of p_U. It can also be shown that the values of p_L and p_U calculated according to Eq. (1) do not converge to the values of p_L and p_U calculated according to the normal approximation of the binomial distribution.

Therefore, we advocate here a different, Bayesian approach to the problem (Bayes [10], Laplace [11]). This approach has been reintroduced (Jaynes [12], Jeffreys [12—14]) and applied to the treatment of astrophysical and physical data [15,16]. In the described experiments the known parameter is the observed number of successes/failures and the unknown stochastic variable is the probability of success, p.

Method and results

If the probability distribution for p was known, on its basis the upper and lower limits of a given confidence interval could be defined and determined in the usual way in which confidence limits of a random variable with a known distribution are found. The problem thus could be more precisely formulated as a problem of constructing the probability distribution of p as a function of the observed number of successes k and the number of trials, n.

What Laplace did more than 200-years-ago (Laplace [11], see also Jaynes' Probability theory: the logic of science [14]) was to propose the following probability density distribution of p:

$$f_{k,n}(p)=\frac{C_n^k p^k(1-p)^{n-k}}{\int_0^1 C_n^k p^k(1-p)^{n-k}\mathrm{d}p}=\frac{p^k(1-p)^{n-k}}{\int_0^1 p^k(1-p)^{n-k}\mathrm{d}p}, \tag{2}$$

which differs from the binomial distribution by only a constant. The denominator is the normalization factor. It happens to be the Euler integral of the first kind, known also as the complete β function — $\int_0^1 p^k(1-p)^{n-k}\mathrm{d}p=(n+1)!/(k!(n-k)!)$. C_n^k are the binomial coefficients. Eq. (2) takes into account that p is a continuous variable, which can take any value between 0 and 1.

By differentiating the distribution function (Eq. (2)) with respect to p one finds the most probable value of p, $\widehat{p}$, that will give the observed number of successes k in n trials:

$$\widehat{p}=\frac{k}{n} \tag{3}$$

Based on Eq. (2) the mean value of p, $\overline{p}=\int_0^1 pf_{k,n}(p)\mathrm{d}p$ and its variance $\sigma_p^2=\int_0^1(p-\overline{p})^2 f_{k,n}(p)\mathrm{d}p$ can be calculated (Laplace [11]) giving:

$$\overline{p}=\frac{k+1}{n+2} \quad \text{and} \quad \sigma_p^2=\frac{(n-k+1)(k+1)}{(n+2)^2\left(n+3\right)} \tag{4}$$

The lower limit p_L of the $100(1-\alpha)$% confidence interval can be determined through solving the following integral equation for p_L:

$$\int_0^{p_L}\mathrm{d}P_{p\in(p,p+\mathrm{d}p)}=\frac{\int_0^{p_L}p^k(1-p)^{n-k}\mathrm{d}p}{\int_0^1 p^k(1-p)^{n-k}\mathrm{d}p}=A_1 \tag{5}$$

The upper limit p_U of the $100(1-\alpha)$% confidence interval is determined through solving a similar equation for p_U:

$$\int_0^{p_U}\mathrm{d}P_{p\in(p,p+\mathrm{d}p)}=A_2 \tag{6}$$

The constants A_1 and A_2 should satisfy the following relation, $A_2 - A_1=1-\alpha$, so that combining Eqs. (5) and (6) one gets

$$\int_{p_L}^{p_U}\mathrm{d}P_{p\in(p,p+\mathrm{d}p)}=1-\alpha \tag{7}$$

The meaning of this equation is that the cumulative probability that the true value of p lies between p_L and p_U, $P(p_L\le p\le p_U)$, is $1-\alpha$.

There exist different ways of constructing a second equation for A_1 and A_2. The following two are the most commonly used approaches.

(A) For symmetry reasons one may choose: $A_1=\alpha/2$ and $A_2=1-\alpha/2$.
(B) Alternatively, one may impose the requirement that $f_{k,n}(p)|_{p=p_L}=f_{k,n}(p)|_{p=p_U}$.