The Korean Communications
in Statistics Vol. 3, No. 3, 1996
pp. 263-269

On Confidence Interval for the Probability of Success

Sang-Joon Lee, M.T. Longnecker[1], Woochul Kim[2]

Abstract

The simplest approximate confidence interval for the probability of success is the one based on the normal approximation to the binomial distribution. It is widely used in the introductory teaching, and various guidelines for its use with "large" sample have appeared in the literature. This paper suggests a guideline when to use it as an approximation to the exact confidence interval, and comparisons with existing guidelines are provided.

1. Introduction

The simplest approximate formula for a $100(1-\alpha)\%$ confidence interval for the binomial parameter p , based on x successes in n trials is

$$I_0 : f \pm c\sqrt{\frac{f(1-f)}{n}}$$

where $f = x/n$ and c is the upper $\alpha/2$ quantile of the standard normal distribution.

Despite its several deficiencies (see, for example, Blyth and Still (1983)), I_0 is widely used in introductory teaching for "large" sample size n . Regarding this "largeness", Samuels and Lu (1992) suggested a useful guideline based on a discrepancy measure between I_0 and the exact confidence interval by Clopper and Pearson.

The Clopper-Pearson confidence interval for p is the equal-tail interval determined by its lower and upper endpoints l_1 and u_1 where $l_1(0) = 0,\ u_1(n) = 1$ and otherwise

$$l_1(x) = p : \sum_{j=x}^{n} \binom{n}{j} p^j(1-p)^{n-j} = \frac{\alpha}{2}$$

$$u_1(x) = p : \sum_{j=0}^{x} \binom{n}{j} p^j(1-p)^{n-j} = \frac{\alpha}{2}$$

The Clopper-Pearson interval will be denoted by I_1 , and it is "exact" in the sense that

$$\inf_p \Pr\{p \in I_1\} \geq 1-\alpha. \tag{1.1}$$

The aim of this paper is to report a guideline based on a new discrepancy measure which

1) Department of Statistics, Texas A&M University, College Station, Texas 77843, U. S. A
2) Professor, Department of Computer Science and Statistics, Seoul University, Seoul, 151-7424, Korea

is a modified version of Samuels and Lu (1992). Section 2 describes the discrepancy measure along with its properties. Section 3 contains the resulting guidelines, and provide a comparison with existing guidelines and an illustration for an application to control chart.

2. A Discrepancy Measure

To assess the accuracy of the approximate interval I_0, Samuels and Lu (1992) used the differences between the corresponding endpoints of I_0 and I_1. To be precise their relative discrepancy measure is

$$r_{SL}(x,f) = \frac{|\ell_0(x)-\ell_1(x)|+|u_0(x)-u_1(x)|}{u_1(x)-\ell_1(x)}, \tag{2.1}$$

where ℓ_0 and u_0 denote the lower and upper endpoints of I_0, respectively.

It should, however, be pointed out that the approximate interval I_0 becomes "exact" in the sense of (1.1) when $\ell_0 \leq \ell_1$ and $u_0 \geq u_1$. In other words, I_0 meets the basic requirement of the nominal level in such a case. A relative discrepancy measure reflecting such an idea would be

$$r_{SL}(x,f) = \frac{(\ell_0(x)-\ell_1(x))^+ + (u_0(x)-u_1(x))^+}{u_1(x)-\ell_1(x)}, \tag{2.2}$$

where a^+ is a if $a \geq 0$, and 0 otherwise.

It is not difficult to see that the relative discrepancy measure $r(x,f)$ also enjoys the following properties of r_{SL} in Samuels and Lu (1992):

$$\begin{aligned} &(i)\ r(x,0+) \text{ is decreasing in } x, \\ &(ii)\ r(x,f) < r(x,0+) \text{ for } 0 < f \leq 0.5, \\ &(iii)\ r(x,f) \sim \frac{M(f)}{\sqrt{x}} \text{ uniformly in } f \text{ as } x \to \infty \end{aligned} \tag{2.3}$$

where $r(x,0+)$ denotes the limiting value of $r(x,f)$ as $n \to \infty$ for fixed x, and

$$M(f) = \frac{(1-c^2+f(2c^2+1))^+ + (c^2+2-f(2c^2+1))^+}{6c\sqrt{1-f}} \tag{2.4}$$

Figures 1 and 2 show $r(x,f)$ and $r_{SL}(x,f)$, respectively, as a function of x for $\alpha = 0.05$ and for selected values of $f \leq 0.5$. Here, and in the sequel, we consider only the case of $f \leq 0.5$ due to the symmetry of the problem. It can be observed from Figures 1 and 2 that the new relative discrepancy measure $r(x,f)$ in (2.2) leads to the suggestion for the use of the approximate confidence interval I_0 is more often for smaller sample sizes.

Because of (iii) in (2.3), the function $M(f)$ in (2.4) can be used to find the most "appropriate" sample size $n = x/f$, i. e., that for which the use of I_0 is most appropriate.

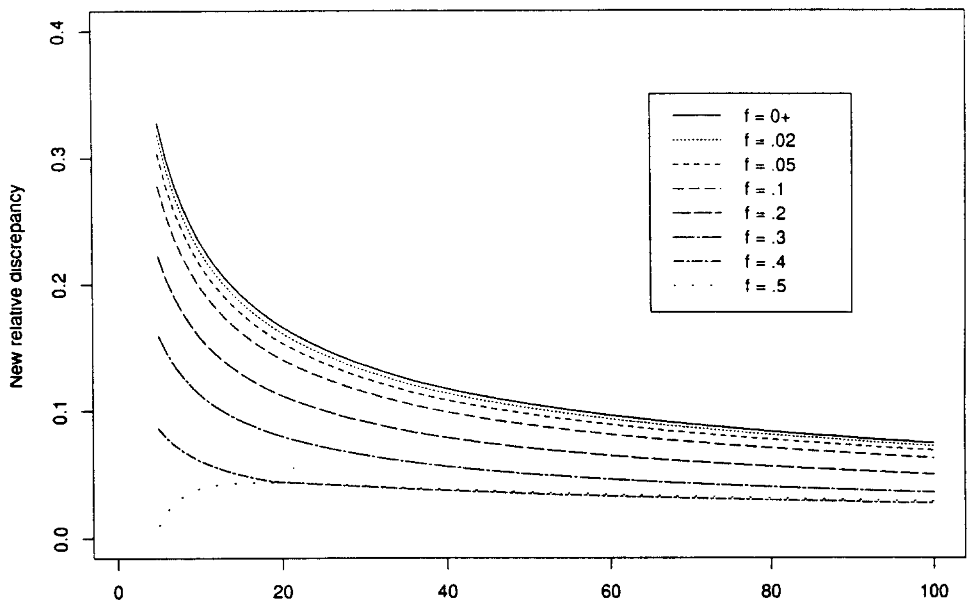

Figure 1 : Relative discrepancy $r(x,f)$ between I_0 and I_1 for $\alpha = 0.05$

Figure 3 shows $M(f)$ as a function of f for $\alpha = 0.01$, 0.05 and 0.10. In particular, for $\alpha = 0.05$ the minimum of M is attained near $f = 0.33$. Thus, for large x the most appropriate sample size for the use of I_0 is approximately $x/0.33$.

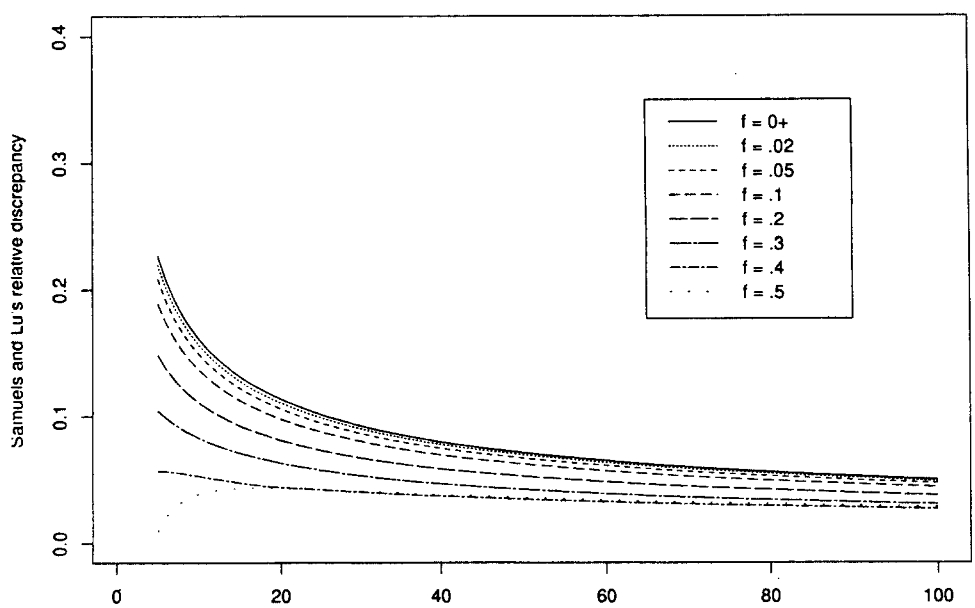

Figure 2 : Relative discrepancy $r_{SL}(x,f)$ between I_0 and I_1 for $\alpha = 0.05$

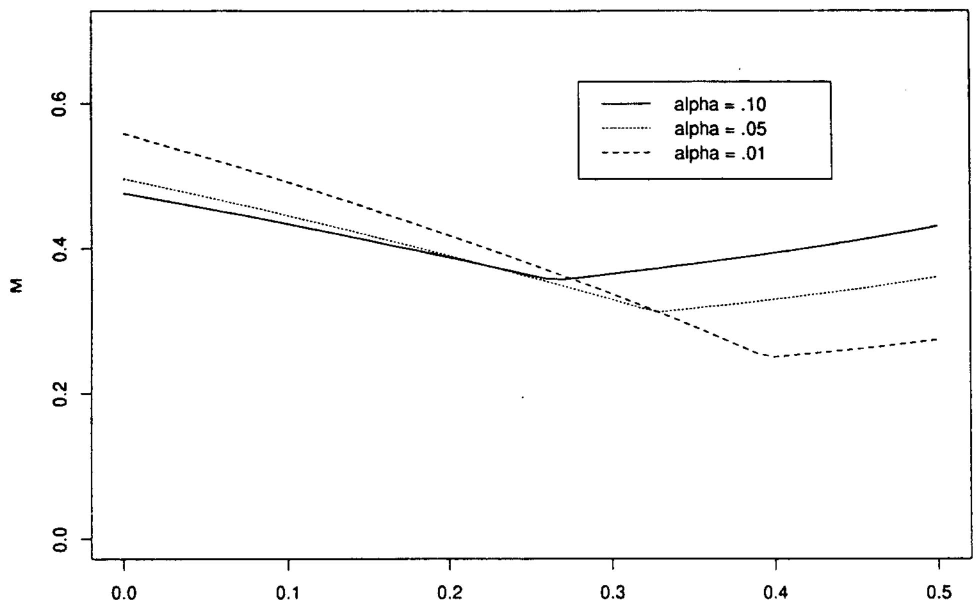

Figure 3 : M vs f plot

3. A Guideline for Using I_0

As Samuels and Lu (1992) pointed out, the properties (i) and (ii) in (2.3) lead us to construct an integer valued function $x_\delta(f)$ such that

$$r(x, f) \le \delta \quad \text{if } x \ge x_\delta(f) \tag{3.1}$$

where δ is a specified tolerance for relative discrepancy. The relation (3.1) provides us a guideline for using the approximate confidence interval I_0 :

"Use I_0 only if $x \ge x_\delta(f)$ ".

Following the construction in Samuels and Lu (1988), we have computed the values of $x_\delta(f)$, for selected values of f, $0 \le f \le 0.5$, and for $\delta =$ 0.05, 0.10, 0.15 when $\alpha = 0.05$. These values are given in Table 1 which can be extended to the cases when $f > 0.5$, by symmetry.

As an example of use of Table 1, the suggested guideline when $\alpha = 0.05$ and $\delta = 0.15$ says us to use I_0 if there are at least 12 observations in each category, and not to use I_0 if there are fewer than 5 observations in either category, and to refer to the values in Table 1 in intermediate cases.

Table 1. Values of $x_\delta(f)$, the minimum x required for using I_0 when $\alpha = 0.05$

f	Tolerance of δ			f	Tolerance of δ		
	.15	.10	.05		.15	.10	.05
0+	12	26	102	.26	5	9	44
.01	12	26	100	.27	5	9	42
.02	11	25	97	.28	5	8	40
.03	11	24	95	.29	5	7	37
.04	11	24	93	.30	5	6	35
.05	10	23	90	.31	5	5	33
.06	10	22	88	.32	5	5	31
.07	10	22	86	.33	5	5	29
.08	9	21	84	.34	5	5	27
.09	9	20	81	.35	5	5	25
.10	9	20	73	.36	5	5	22
.11	8	19	77	.37	5	5	20
.12	8	18	75	.38	5	5	18
.13	8	18	72	.39	5	5	16
.14	7	17	70	.40	5	5	13
.15	7	16	68	.41	5	5	10
.16	7	16	66	.42	5	5	5
.17	6	15	63	.43	5	5	5
.18	6	14	61	.44	5	5	5
.19	6	14	59	.45	5	5	5
.20	5	13	57	.46	5	5	5
.21	5	13	55	.47	5	5	5
.22	5	12	52	.48	5	5	5
.23	5	11	50	.49	5	5	5
.24	5	11	48	.50	5	5	5
.25	5	10	46				

Samuels and Lu (1992) compared their guideline with other guidelines in the literature, and they found out that other guidelines except Cochran's (1977, p. 58) do not give reasonable control of their relative discrepancy.

Their comparison is reproduced in Table 2 with the addition of our guideline. Description of guidelines A, B, C, D, E can be found in Samuels and Lu (1992). It can be observed from Table 2 that the guideline C is close to our's with tolerance $\delta = 0.15$ and the guidelines by Samuel and Lu (1992) and by Cochran (E) are conservative in the sense of the relative discrepancy in (2.2).

Table 2. Required minimum x to use I_0 by various guidelines when $\alpha = 0.05$

f	Our's			Samuels and Lu's			Guidelines in the literature				
	Tolerance of δ			Tolerance of δ							
	.15	.10	.05	.15	.10	.05	A	B	C	D	E
0+	12	26	102	25	56	221	5	4	9	5	80
.1	9	30	79	18	40	157	5	4	8	6	60
.2	5	13	57	12	25	100	5	3	7	6	40
.3	5	6	35	6	14	51	5	3	6	7	24
.4	5	5	13	5	5	16	5	2	5	8	20
.5	5	5	5	5	5	13	5	2	4	10	15

Finally a remark on the application of the guideline (3.1) to quality control is to be added. The upper (UCL) and lower (LCL) control limits in p chart are set up by use of I_0 with $c = 3$ standard deviation rule. This corresponds to $\alpha = 0.0027$, and we have computed the values of the minimum x required to use I_0 in this case. These values are given in Table 3, which can be used in implementing p chart in practice.

Table 3. Required minimum x to use I_0 with c=3 in p control chart

f	Tolerance of δ		
	.15	.10	.05
<.01	19	40	156
.01	18	39	153
.05	15	34	134
.10	12	28	113
.15	9	22	93
.20	6	16	73
.25	5	10	53
.30	5	5	34
.35	5	5	5
.40	5	5	5
.45	5	5	5
.50	5	5	5

References

[1] Blyth, C.R. and Still, H.A. (1993). Binomial Confidence Intervals, *Journal of the American Statistical Association*, Vol. 78, 108-116.

[2] Cochran, W.G. (1977). *Sampling Techniques*, John Wiley & Sons, New York.

[3] Samuels, M.L. and Lu, Tai-Fang C. (1988). Sample Size Requirements for the Approximation of Binomial Confidence Intervals, Technical Report 88-08, Purdue University, Department of Statistics, U. S. A.

[4] Samuels, M.L. and Lu, Tai-Fang C. (1992). Sample Size Requirements for the Back-of-the-Envelope Binomial Confidence Interval, *The American Statistician*, Vol. 46, No 3, 228-231.