



**Binomial Distribution: Hypothesis Testing, Confidence
Intervals (CI), and Reliability with Implementation in
S-PLUS**

by Joseph C. Collins

ARL-TR-5214

June 2010

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

Army Research Laboratory

Aberdeen Proving Ground, MD 21005

ARL-TR-5214

June 2010

Binomial Distribution: Hypothesis Testing, Confidence Intervals (CI), and Reliability with Implementation in S-PLUS

Joseph C. Collins

Survivability/Lethality Analysis Directorate, ARL

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE (DD-MM-YYYY) June 2010		2. REPORT TYPE Final		3. DATES COVERED (From - To) January to June 2009	
4. TITLE AND SUBTITLE Binomial Distribution: Hypothesis Testing, Confidence Intervals (CI), and Reliability with Implementation in S-PLUS				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Joseph C. Collins				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Laboratory ATTN: RDRL-SLB-D Survivability/Lethality Analysis Directorate Aberdeen Proving Ground, MD 21005				8. PERFORMING ORGANIZATION REPORT NUMBER ARL-TR-5214	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited.					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT System requirements can be expressed in terms of an upper (lower) limit on the probability of failure (success). Statistical analysis of such binary (0/1, success/failure, go/no-go) data typically requires point and interval estimation and inference or hypothesis tests on the associated event probability. For identical independent trials, the proportion observed serves as an estimate of the event rate. Based on the method of Clopper-Pearson (CP) and the likelihood ratio (LR) technique properties of the binomial distribution are used to derive interval estimates, which are in turn used in inference. An application is the determination of sample size and maximum permissible number of failures (nf) required to establish a specific reliability (probability of success) with given probability (confidence).					
15. SUBJECT TERMS binomial reliability, probability, statistics					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 43	19a. NAME OF RESPONSIBLE PERSON Joseph C. Collins
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			19b. TELEPHONE NUMBER (Include area code) (410) 278-6832

Contents

List of Figures	v
List of Tables	v
Summary	1
1. Binomial Distribution	3
1.1 Definitions and Properties	3
1.2 Implementation Issue	4
1.3 Beta Distribution	5
2. Hypothesis Testing	8
2.1 One Sided Upper	8
2.2 One Sided Lower	9
2.3 Two Sided.....	9
3. Confidence Intervals (CI)	10
3.1 Upper	10
3.2 Lower.....	10
3.3 Two Sided.....	11
3.4 Implementation.....	11
3.5 Conservative Coverage.....	11
4. The LR Approach	14
4.1 Hypothesis Tests.....	15
4.2 Confidence Intervals.....	16
4.3 Implementation.....	17
5. Reliability	18
5.1 Reliability Tables	18
5.2 Sample Size	20

6. References	22
Appendix A. S-PLUS Code	23
A.1 Library Functions	23
A.2 Upper-Tail Binomial Probabilities	23
A.3 Binomial Parameter Confidence Interval: CP	24
A.4 Binomial Reliability Table	24
A.5 Binomial Reliability Sample Size	24
A.6 Binomial LR	25
A.7 Binomial LR cdf Table	25
A.8 Binomial LR cdf Evaluation	26
A.9 Binomial LR Hypothesis Test	26
A.10 Binomial Parameter Confidence Interval: LR	26
Appendix B. Reliability Table	29
List of Symbols, Abbreviations, and Acronyms	34
Distribution List	35

List of Figures

Figure 1. Binomial cdf and quantile function (qf) examples with continuous interpolation.	7
Figure 2. Upper CI coverage.....	12
Figure 3. Lower CI coverage.	13
Figure 4. Two-sided CP CI coverage.....	14
Figure 5. Two-sided LR CI coverage.	17
Figure 6. Two-sided CI coverage comparison.	17

List of Tables

Table 1. LR cdf.	16
Table 2. CP cdf.	16

INTENTIONALLY LEFT BLANK.

Summary

The purpose of this report is to provide statisticians with the tools to solve certain reliability problems that arise in consultations with clients.

Statisticians often encounter clients seeking answers to questions such as, “I did an experiment 6 times with 3 successes and 3 failures, so the success rate is 50%, right?” A reasonable, useful, and informative response from the statistician would be, “yes, the estimate of success is 50%, but you can only state with 90% confidence that the true success rate lies between 15% and 85%. If you want to tighten that up, you need a larger sample.” The client asks, “how large?” and the conversation proceeds from there.

A client may pose a problem such as, “I need 90% confidence on 95% reliability. Can I do this with 10 runs?” The statistician can truthfully answer, no. You need 29 with no failures or 46 with 1 failure to assert 95% reliability with 90% confidence.” The client may respond, “well, what can I get with 1 failure in 20?” and so the consultation continues.

In fact, the solutions to both of these problems exploit properties of the binomial distribution. System requirements can be expressed in terms of an upper (lower) limit on the probability of failure (success). Statistical analysis of such binary (0/1, success/failure, go/no-go) data typically requires point and interval estimation and inference or hypothesis tests on the associated event probability. For identical independent trials, the proportion observed serves as an estimate of the event rate. Based on the method of Clopper-Pearson (CP) and the likelihood ratio (LR) technique properties of the binomial distribution are used to derive interval estimates, which are in turn used in inference. An application is the determination of sample size and maximum permissible number of failures (nf) required to establish a specific reliability (probability of success) with given probability (confidence).

The statistician needs to address such issues and answer these questions in real time. This report provides the necessary technology expressed in the S-PLUS / S / R statistical computing environment language family, implementations of which are available as both commercial and free open-source software.

INTENTIONALLY LEFT BLANK.

1. Binomial Distribution

1.1 Definitions and Properties

The random variable X is Bernoulli(p), or B_p , if $X \in \{0,1\}$ and $\Pr[X = 1] = p = 1 - \Pr[X = 0]$. The probability mass function, or discrete probability density function (pdf), is

$$f_p(x) = p^x(1-p)^{1-x}. \quad (1)$$

The sum X of n independent and identically distributed (iid) B_p is Bernoulli(n, p), or $B_{n,p}$. Note that $B_p = B_{1,p}$. The binomial coefficient is $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. Standard functions are defined for $k \in \{0,1,2, \dots, n\}$. The pdf is

$$f_{n,p}(x) = C(n, k)p^x(1-p)^{1-x}. \quad (2)$$

The cumulative distribution function (cdf) is

$$F_{n,p}(k) = \Pr[X \leq k] = \sum_{j=0}^k f_{n,p}(j) \quad (3)$$

with $F_{n,p}(x) = 0$ if $x \leq 0$, and $F_{n,p}(x) = 1$ if $x \geq n$, and right-continuous extension to \mathbb{R} , as is customary for discrete distributions. The quantile function (qf) is the usual unique left-continuous pseudo-inverse of the cdf,

$$Q_{n,p}(u) = \inf \{ k: F_{n,p}(k) \geq u \} \quad (4)$$

It is also useful to have an upper-probability version of the cdf,

$$G_{n,p}(k) = \Pr[X \geq k] = \sum_{j=k}^n f_{n,p}(j). \quad (5)$$

So $G_{n,p}(k) = 1 - F_{n,p}(k)$ and $F_{n,p}(k) = 1 - G_{n,p}(k + 1)$. Note that $F_{n,p}(k)$ increases with k and decreases with n or p , and $G_{n,p}(k)$ decreases with k and increases with n or p .

Moments are

$$E[X^k] = \sum_{j=0}^n j^k f_{n,p}(j) \quad (6)$$

So $E[X] = np$, and $E[X^2] = np(1-p) + n^2p^2$, and $\text{Var}[X] = np(1-p)$.

The log-likelihood function is

$$\begin{aligned}\mathcal{L} &= \sum_{j=0}^n (\log C(n, j) + X_j \log p + (n - X_j) \log(1 - p)) \\ &= \left(\sum_{j=0}^n \log C(n, j) \right) + k \log p + (n - k) \log(1 - p).\end{aligned}\tag{7}$$

The maximum likelihood estimator (MLE) of p is k/n .

1.2 Implementation Issue

Reasonable algorithms for $F_{n,p}(k)$ return the appropriate values for large n and small k but succumb to numeric underflow for large k .

For example, since $F_{n,p}(0) = (1 - p)^n$ and $F_{n,p}(n - 1) = 1 - p^n$, set $n = 100$ and $p = 0.6$ to get

$$F_{100,0.6}(0) = (1 - 0.6)^{100} = 0.4^{100} \cong 1.6 \times 10^{-40}\tag{8}$$

and

$$F_{100,0.6}(99) = 1 - 0.6^{100} \cong 1 - 6.5 \times 10^{-23}.\tag{9}$$

But with standard double precision resolution of about 16 decimal places, the result is $F_{100,0.6}(99) = 1$ exactly. The cure for lower tail probabilities too close to 1 is to use upper tail probabilities, which will be close to 0, and work with G instead of F , since $G_{n,p}(n) = p^n = 1 - F_{n,p}(n - 1)$, and so

$$G_{100,0.6}(100) = 0.6^{100} = 1 - F_{100,0.6}(99) = 1 - (1 - 0.6)^{100} \cong 6.5 \times 10^{-23}.\tag{10}$$

But naïve use of $G_{n,p}(k) = 1 - F_{n,p}(k - 1)$ directly will of course give $G_{100,0.6}(100) = 0$. To obtain a useful representation of $G(k)$ with k large in terms of an accurate algorithm for $F(k)$ with k small, let $X \sim B_{n,p}$ and $W = n - X$. Note that $F_W(k) = \Pr[W \leq k] = \Pr[n - X \leq k] = \Pr[X \geq n - k] = 1 - \Pr[X < n - k] = 1 - \Pr[X \leq n - k - 1] = 1 - F_X(n - k - 1)$. Then $G_X(k) = 1 - F_X(k - 1) = 1 - F_X(n - (n - k) - 1) = F_W(n - k)$, so

$$G_{n,p}(k) = \Pr[B_{n,p} \leq k] = \Pr[B_{n,1-p} \leq n - k] = F_{n,1-p}(n - k).\tag{11}$$

This gives $G_{100,0.6}(100) = F_{100,0.4}(0) = (1 - 0.4)^{100} = 0.6^{100} \cong 6.5 \times 10^{-23}$ as required.

1.3 Beta Distribution

The Gamma function, which has $\Gamma(n) = (n - 1)!$ for $n = 1, 2, 3, \dots$, is given by

$$\Gamma(z) = \int_0^{\infty} u^{z-1} e^{-u} du. \quad (12)$$

The Beta function, which has $B(n, k) = (n - 1)! (k - 1)! / (n + k - 1)!$ on positive integers, is

$$B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a + b)} = \int_0^1 x^{a-1} (1 - x)^{b-1} dx. \quad (13)$$

The Beta (a, b) distribution on $[0, 1]$, with $a > 0$ and $b > 0$, has pdf and cdf

$$\begin{aligned} f_{\text{Beta}(a,b)}(x) &= \frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1} \\ F_{\text{Beta}(a,b)}(x) &= \int_0^x f_{\text{Beta}(a,b)}(u) du. \end{aligned} \quad (14)$$

The Binomial and Beta cdfs are related by $F_{n,p}(k - 1) = F_{\text{Beta}(n-k+1,k)}(1 - p)$, so

$$\begin{aligned} F_{n,p}(k) &= F_{\text{Beta}(n-k,k+1)}(1 - p), \quad \text{or} \\ F_{n-k+1,1-p}(k - 1) &= F_{\text{Beta}(n,k)}(p) \end{aligned} \quad (15)$$

and also

$$\begin{aligned} G_{n,p}(k) &= F_{\text{Beta}(k,n-k+1)}(p), \quad \text{or} \\ G_{n-k+1,p}(k) &= F_{\text{Beta}(n,k)}(p). \end{aligned} \quad (16)$$

To show $F_{n,p}(k - 1) = F_{\text{Beta}(n-k+1,k)}(1 - p)$, first note that $B(n - k + 1, k) = (n - k)! (k - 1)! / n!$, and also $F_{\text{Beta}(n,1)}(1 - p) = n \int_0^1 x^{n-1} dx = (1 - p)^n = f_{n,p}(0)$. Now, integrating by parts,

$$\begin{aligned}
F_{\text{Beta}(n-k+1,k)}(1-p) &= \frac{n!}{(n-k)!(k-n)!} \int_0^{1-p} x^{n-k}(1-x)^{k-1} dx \\
&= \frac{n!}{(n-k)!(k-n)!} \left[\frac{x^{n-k}(1-x)^{k-1}}{n-k+1} \Big|_0^{1-p} + \int_0^{1-p} \frac{(k-1)x^{n-k+1}(1-x)^{k-2}}{n-k+1} dx \right] \\
&= \frac{n! p^{k-1}(1-p)^{n-k+1}}{(n-k+1)!(k-1)!} + \frac{n!}{(k-2)!(n-k+1)!} \int_0^{1-p} x^{n-k+1}(1-x)^{k-2} dx \\
&= f_{n,p}(k-1) + \frac{1}{B(n-k+2, k-1)} \int_0^{1-p} x^{n-k+1}(1-x)^{k-2} dx \\
&= f_{n,p}(k-1) + F_{\text{Beta}(n-k+2, k-1)}(1-p) \\
&= f_{n,p}(k-1) + \dots + f_{n,p}(1) + F_{\text{Beta}(n,1)}(1-p) \\
&= f_{n,p}(k-1) + \dots + f_{n,p}(1) + f_{n,p}(0) \\
&= F_{n,p}(k-1).
\end{aligned} \tag{17}$$

See, for example, Stuart (1).

In fact, the function

$$F(x) = \begin{cases} 0, & x \leq -1 \\ F_{\text{Beta}(n-x+1,x)}(1-p), & -1 < x < n \\ 1, & n \leq x \end{cases} \tag{18}$$

is a continuous cdf on $[-1, n]$, and $F(k) = F_{n,p}(k)$ for $k = 0, 1, 2, \dots, n$. So F serves as a continuous version, or interpolation, of $F_{n,p}$, which is sometimes useful. The corresponding quantile function $Q(u) = F^{-1}(u)$ is not related to Q_{Beta} , but can be obtained by solving $F(x) = u$ numerically to get $x = Q(u)$. See figure 1.

Note that Q_{Beta} does provide the solution p of $u = F_{n,p}(k)$. Since $u = F_{n,p}(k) = F_{\text{Beta}(n-k,k+1)}(1-p)$, and so $Q_{\text{Beta}(n-k,k+1)}(u) = 1-p$, observe that

$$p = 1 - Q_{\text{Beta}(n-k,k+1)}(u). \tag{19}$$

Likewise for $u = G_{(n,p)}(k)$. Since $u = G_{n,p}(k) = F_{\text{Beta}(k,n-k+1)}(p)$ and so $Q_{\text{Beta}(k,n-k+1)}(u) = p$, it follows that

$$p = Q_{\text{Beta}(k,n-k+1)}(u). \quad (20)$$

Library implementations of Q_{Beta} provide for efficient and accurate evaluation of p in these situations. This will be useful later.

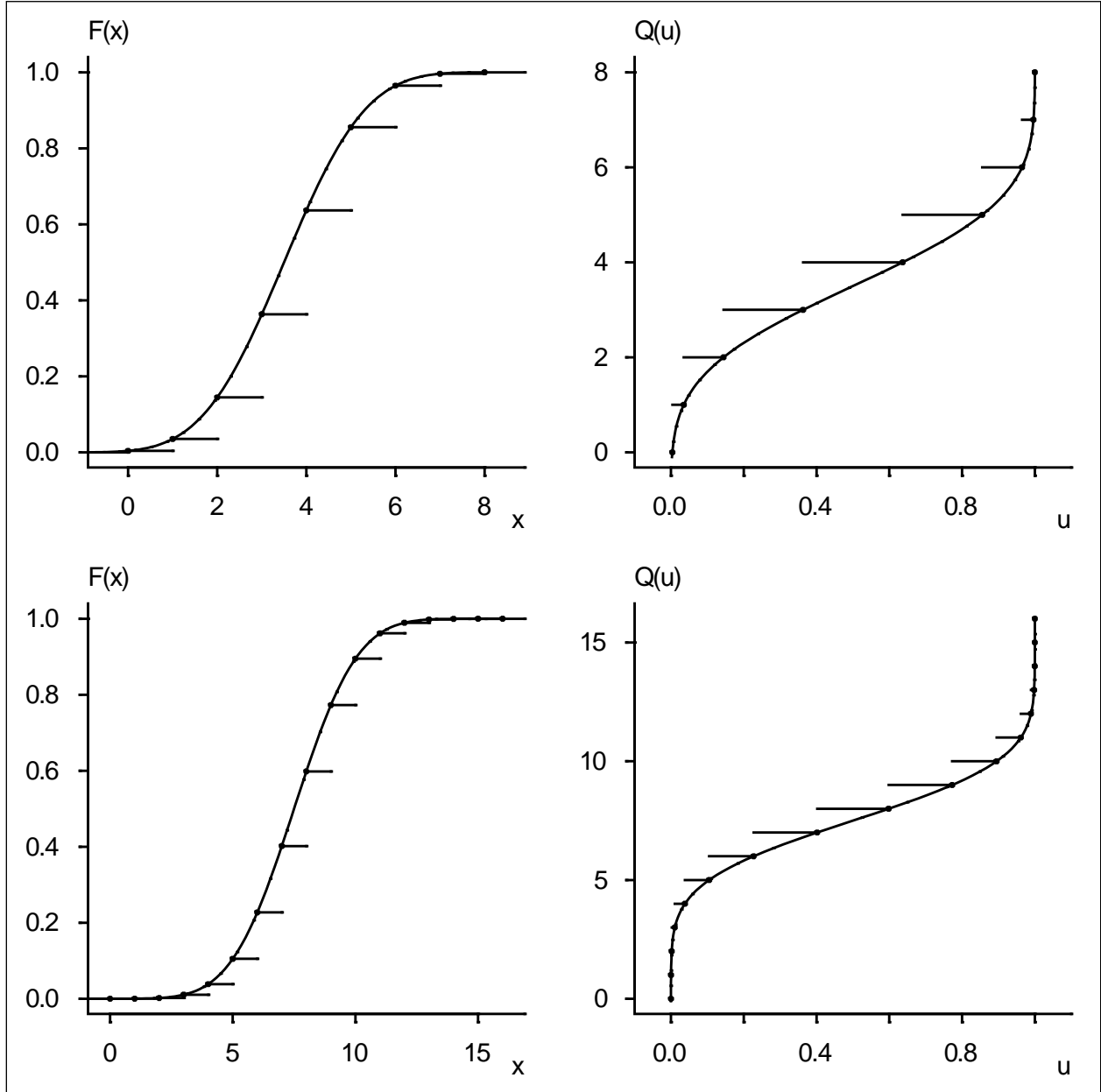


Figure 1. Binomial cdf and quantile function (qf) examples with continuous interpolation.

2. Hypothesis Testing

Tests of size γ and Type I error probability $\alpha = 1 - \gamma$ for

$$H_0: p = p_o \tag{21}$$

are based on data X . Here α is the null rejection probability, or probability of rejecting H_0 when it is true. See Mood (2) or Stuart (3) for an explanation.

2.1 One Sided Upper

The one-sided alternative

$$H_1: p < p_o \tag{22}$$

provides an upper limit for p upon rejection of H_0 . The critical value is k_U , and the decision rule is to reject H_0 if $X \leq k_U$ where $\Pr[X \leq k_U | H_0] = \alpha$. Since X is discrete, take

$$k_U = \sup \{ k | \Pr[X \leq k | H_0] \leq \alpha \} = \sup \{ k | F_{n,p_o}(k) \leq \alpha \}. \tag{23}$$

The rejection region is $I_1 = \{ 0, \dots, k_U \}$, and the non-rejection region is $I_0 = \{ k_U + 1, \dots, n \}$. Equivalently, the p-value for the test is $p_X = F_{n,p_o}(X)$, and the rule is to reject H_0 if $p_X \leq \alpha$. For example, let $p_o = 0.3$ and $\alpha = 0.1$ with $n = 100$. Since $F_{100,0.3}(23) = 0.0755$ and $F_{100,0.3}(24) = 0.114$, it follows that $k_U = 23$. These are also the p-values, $p_{23} = 0.0755$ and $p_{24} = 0.114$. The true value of α for this test is 0.0755.

For example, it is possible for the test to degenerate due to insufficient sample size. Consider $p_o = 0.03$ and $\alpha = 0.01$ with $n = 100$. Since $F_{100,0.03}(0) = 0.0476$, it follows that $F_{100,0.03}(k) > \alpha$ for all k , and so $I_1 = \emptyset$, and $I_0 = \{ 0, \dots, n \}$, and rejection never occurs. The true value of α for this test is 0.

But $F_{151,0.03}(0) = 0.0101$ and $F_{152,0.03}(0) = 0.00976$, so $n > 151$ is required for a non-degenerate test. With $n = 152$, since $F_{152,0.03}(1) = 0.0556$, it follows that $k_U = 0$. Then if $X = 0$ the decision is to reject H_0 and conclude that $p < 0.03$ with probability ≥ 0.99 . If $X > 0$ there is no indication that $p < 0.03$. The true value of α for this test is 0.00976.

Let $p_o = 1/2$ and $\alpha = 7/128$ with $n = 10$. Note that $F_{10,1/2}(2) = 7/128 = \alpha$. It follows that $k_U = 2$ and $p_X \leq \alpha$ when $X \leq 2$, in which case H_0 is rejected. The true value of α for this test is exactly 7/128.

2.2 One Sided Lower

The one-sided alternative

$$H_1: p > p_o \quad (24)$$

provides a lower limit for p upon rejection of H_0 . The critical value is k_L , and one rejects H_0 if $X \geq k_L$ where $\Pr[X \geq k_L | H_0] = \alpha$. Since X is discrete, take

$$k_L = \inf\{k | \Pr[X \geq k | H_0] \leq \alpha\} = \inf\{k | G_{n,p_o}(k) \leq \alpha\}. \quad (25)$$

The rejection region is $I_1 = \{k_L, \dots, n\}$, and the non-rejection region is $I_0 = \{0, \dots, k_L - 1\}$. Equivalently, the p-value for the test is $p_X = G_{n,p_o}(X)$, and one rejects H_0 if $p_X \leq \alpha$.

For example, let $p_o = 0.3$ and $\alpha = 0.1$ with $n = 100$. Since $G_{100,0.3}(36) = 0.116$ and $G_{100,0.3}(37) = 0.0799$, it follows that $k_L = 37$. These are also the p-values, $p_{36} = 0.116$ and $p_{37} = 0.0799$.

2.3 Two Sided

The two-sided alternative

$$H_1: p \neq p_o \quad (26)$$

provides upper and lower limits for p upon rejection of H_0 . The critical values are k_U and k_L , and one rejects H_0 if $X \leq k_U$ or $X \geq k_U$ where $\Pr[X \leq k_U \text{ or } X \geq k_L | H_0] = \alpha$. For symmetry, set $\Pr[X \leq k_U | H_0] = \Pr[X \geq k_L | H_0] = \alpha/2$. Since X is discrete, take

$$\begin{aligned} k_U &= \sup\{k | \Pr[X \leq k | H_0] \leq \alpha/2\} = \sup\{k | F_{n,p_o}(k) \leq \alpha/2\} \\ k_L &= \inf\{k | \Pr[X \geq k | H_0] \leq \alpha/2\} = \inf\{k | G_{n,p_o}(k) \leq \alpha/2\} \end{aligned} \quad (27)$$

The rejection region is $I_1 = \{0, \dots, k_U\} \cup \{k_L, \dots, n\}$, and the non-rejection region is $I_0 = \{k_U + 1, \dots, k_L - 1\}$. Equivalently, the p-value for the test is $p_X = 2 \min(F_{n,p_o}(X), G_{n,p_o}(X))$. One rejects H_0 if $p_X \leq \alpha$.

For example, let $p_o = 0.3$ and $\alpha = 0.1$ with $n = 100$. Since $F_{100,0.3}(22) = 0.0479$ and $F_{100,0.3}(23) = 0.0755$, it follows that $k_U = 22$. Since $G_{100,0.3}(38) = 0.0530$ and $G_{100,0.3}(39) = 0.0340$, it follows that $k_L = 39$. Note that $G_{100,0.3}(22) = 0.971$, $G_{100,0.3}(23) = 0.952$, $F_{100,0.3}(38) = 0.966$, and $F_{100,0.3}(39) = 0.979$. So the p-values are $p_{22} = 0.0957$, $p_{23} = 0.151$, $p_{38} = 0.106$, and $p_{39} = 0.0680$.

3. Confidence Intervals (CI)

The point estimate of p is of course X/n . CIs on p are obtained by pivoting or inverting hypothesis test critical regions as follows. See Clopper and Pearson (4).

3.1 Upper

Let I_1 be the set of p_o for which H_0 would be rejected in favor of $H_1: p < p_o$ with Type I error α , so $I_1 = \{ p \mid F_{n,p}(X) \leq \alpha \} = [p_U, 1]$ where

$$\alpha = F_{n,p_U}(X) = F_{\text{Beta}(n-X, X+1)}(1 - p_U), \quad (28)$$

so p_U can be expressed by equation 19 as a Beta distribution quantile

$$p_U = 1 - Q_{\text{Beta}(n-X, X+1)}(\alpha). \quad (29)$$

Null rejection occurs with probability α , so the non-rejection region

$$I_0 = [0, p_U) = \{ p \mid F_{n,p}(X) > \alpha \} \quad (30)$$

is a 100 γ % upper CI on p , and p_U is an upper confidence limit on p . To reject H_0 when $p \in I_1 = [p_U, 1]$ is precisely the p-value decision rule in section 0.

3.2 Lower

Let I_1 be the set of p_o for which H_0 would be rejected in favor of $H_1: p > p_o$ with Type I error α , so $I_1 = \{ p \mid G_{n,p}(X) \leq \alpha \} = [0, p_L]$ where

$$\alpha = G_{n,p_L}(X) = F_{\text{Beta}(X, n-X+1)}(p_L), \quad (31)$$

so p_L can be expressed by equation (20) as a Beta distribution quantile

$$p_L = Q_{\text{Beta}(X, n-X+1)}(\alpha).. \quad (32)$$

Null rejection occurs with probability α , so the non-rejection region

$$I_0 = (p_L, 1] = \{ p \mid G_{n,p}(X) > \alpha \} \quad (33)$$

is a 100 γ % lower CI on p , and p_L is a lower confidence limit on p . To reject H_0 when $p \in I_1 = [0, p_L]$ is precisely the p-value decision rule in section 2.2.

3.3 Two Sided

Let I_1 be the set of p_o for which H_0 would be rejected in favor of $H_1: p \neq p_o$ with Type I error α , so $I_1 = \{ p \mid G_{n,p}(X) \leq \alpha/2 \} \cup \{ p \mid F_{n,p}(X) \leq \alpha/2 \} = [0, p_L] \cup [p_U, 1]$ where

$$G_{n,p_L}(X) = F_{n,p_U}(X) = \alpha/2. \quad (34)$$

Then p_L and p_U can be expressed by equations 19 and 20 as Beta distributions quantiles

$$\begin{aligned} p_L &= Q_{\text{Beta}(X, n-X+1)}(\alpha/2) \\ p_U &= 1 - Q_{\text{Beta}(n-X, X+1)}(\alpha/2). \end{aligned} \quad (35)$$

Null rejection occurs with probability α , so the non-rejection region

$$I_0 = (p_L, p_U) = \{ p \mid G_{n,p}(X) > \alpha/2 \} \cap \{ p \mid F_{n,p}(X) > \alpha/2 \} \quad (36)$$

is a $100\gamma\%$ two-sided CI on p . To reject H_0 when $p_o \in [0, p_L] \cup [p_U, 1]$ is precisely the p-value decision rule of section 2.3.

3.4 Implementation

Code listings are in appendix A.

The function `binoc.i(n, r, g, type)` uses equations 29, 32, or 35 to give the required CI. The arguments of `binoc.i` are the number of trials `n`, the number of successes `r`, confidence interval size `g` which should be between 0.5 and 1.0, and an integer `type` which can be 0, 1, or 2 to request a lower CI $(L, 1]$, upper CI $[0, U)$, or 2-sided CI (L, U) , respectively.

The return value from `binoc.i` is a named list holding the number of trials `$n`, number of successes `$r`, lower limit `$L`, parameter estimate `$p.hat`, upper limit `$U`, and confidence interval vector `$I`. For example,

```
binoc.i(n=100, r=30, g=0.9, type=2)
```

returns

```
$n:      100
$r:      30
$L:      0.22492
$p.hat:  0.3
$U:      0.38422
$I:      0.22492 0.38422
```

since $\Pr[B_{100,0.22492} \geq 30] = \Pr[B_{100,0.38422} \leq 30] = 0.05$.

3.5 Conservative Coverage

Coverage is the probability that a confidence interval captures the true parameter value.

Considering that $p_U(X)$ is random, and $[0, p_U(X)]$ is a $100\gamma\%$ CI on p , one expects that $C_U(p) = \Pr[p < p_U(X)] = E[I_{[0, p_U(X)]}(p)] \geq \gamma$. This coverage probability is

$$C_U(p) = \sum_{k=0}^n \Pr[X = k] \cdot I_{[0, p_U(X)]}(p) = \sum_{p < p_U(k)} f_{n,p}(k). \quad (37)$$

Figure 2 illustrates C_U for $n = 10$ and $\gamma = 0.9$.

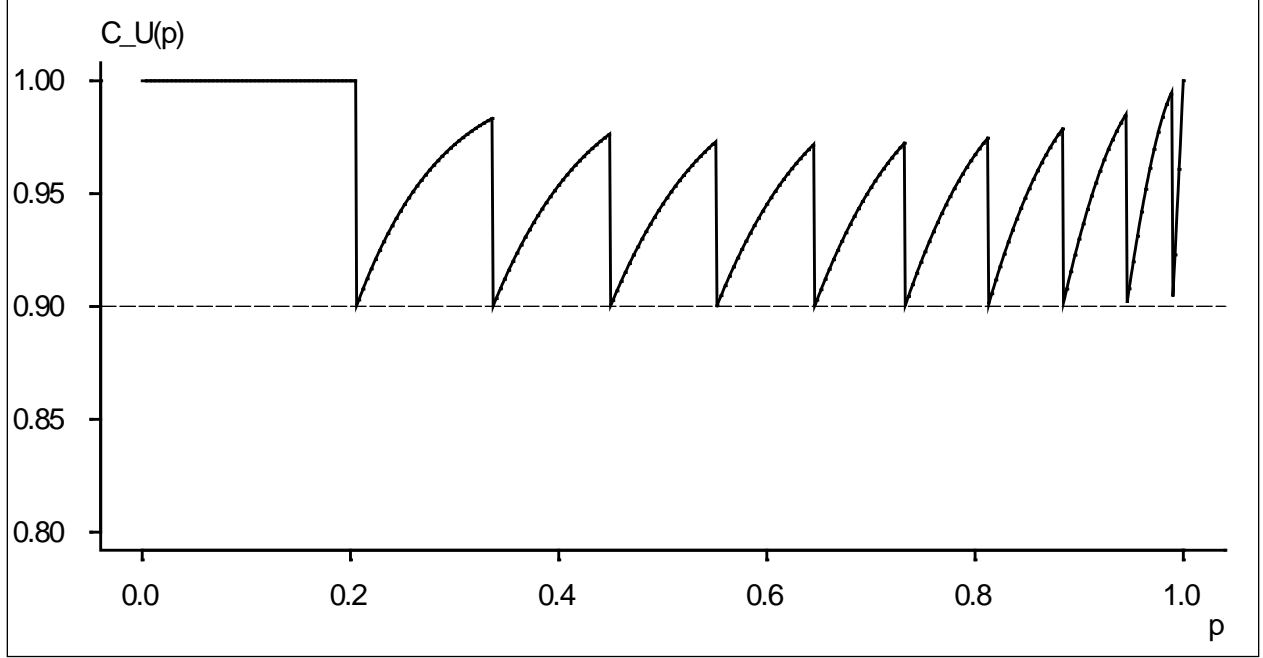


Figure 2. Upper CI coverage.

Likewise, $p_L(X)$ is random, and $(p_L(X), 1]$ is a $100\gamma\%$ CI on p . One expects that $C_L(p) = \Pr[p_L(X) < p] = E[I_{(p_L(X), 1]}(p)] \geq \gamma$. This coverage probability is

$$C_L(p) = \sum_{k=0}^n \Pr[X = k] \cdot I_{(p_L(X), 1]}(p) = \sum_{p_L(k) < p} f_{n,p}(k). \quad (38)$$

Figure 3 illustrates C_L for $n = 10$ and $\gamma = 0.9$.

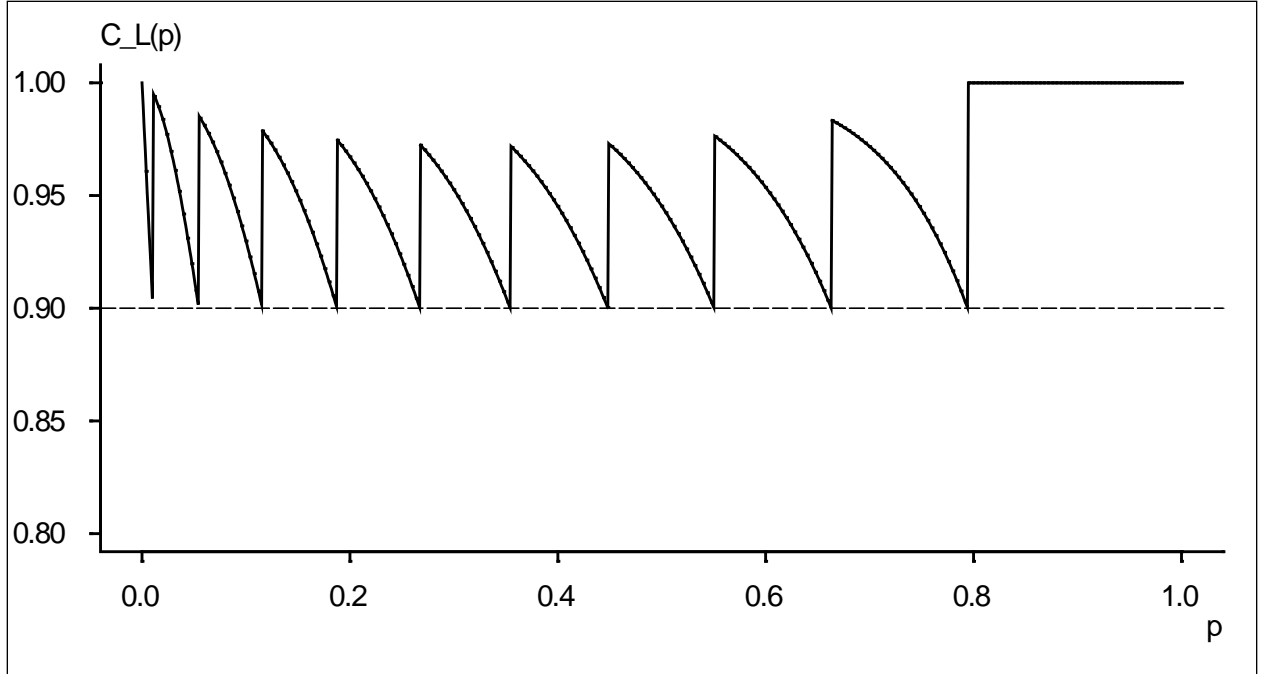


Figure 3. Lower CI coverage.

For the 2-sided CI, $(p_L(X), p_U(X))$ is a $100\gamma\%$ CI on p , and one expects that $C_2(p) = \Pr[p_L(X) < p < p_U(X)] = E[I_{(p_L(X), p_U(X))}(p)] \geq \gamma$. This coverage probability is

$$C_2(p) = \sum_{k=0}^n \Pr[X = k] \cdot I_{(p_L(X), p_U(X))}(p) = \sum_{p_L(k) < p < p_U(k)} f_{n,p}(k). \quad (39)$$

Figure 4 illustrates C_2 for $n = 10$ and $\gamma = 0.9$.

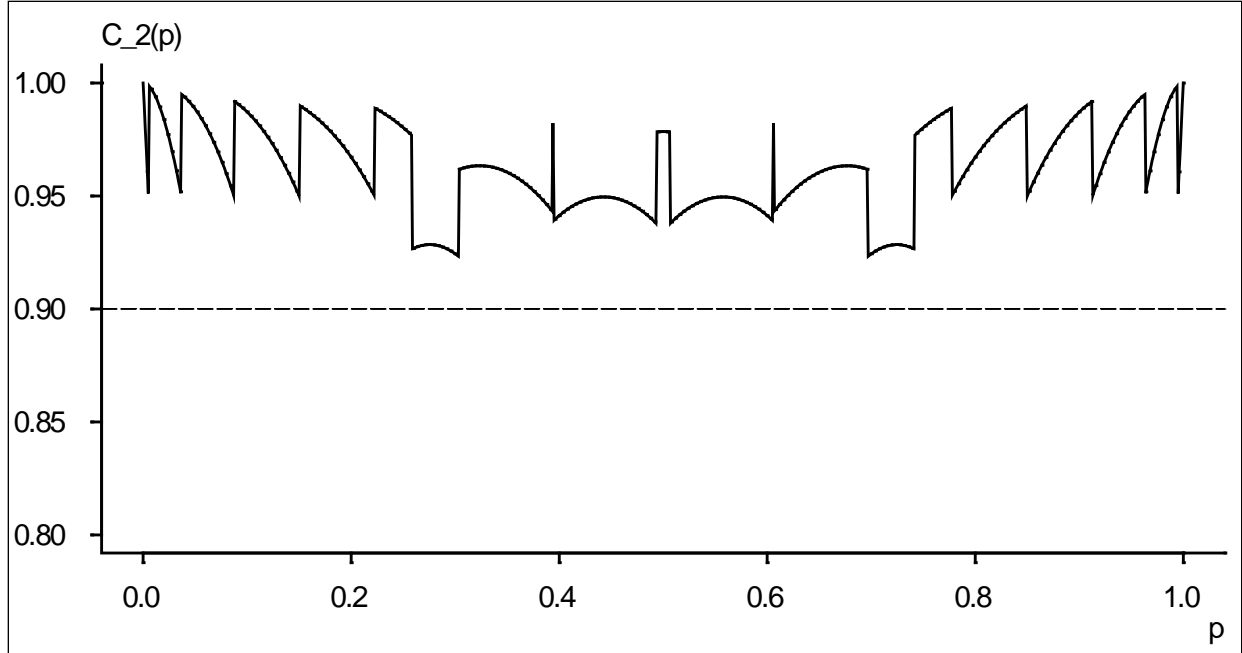


Figure 4. Two-sided CP CI coverage.

4. The LR Approach

For the parameter θ in a space Θ with some subset $\Theta_o \subseteq \Theta$, a test of

$$\begin{aligned} H_0: \theta \in \Theta_o \\ H_1: \theta \notin \Theta_o \end{aligned} \tag{40}$$

can be conducted using the generalized LR

$$\Lambda = \frac{\sup \{ \mathcal{L} \mid \theta \in \Theta_o \}}{\sup \{ \mathcal{L} \mid \theta \in \Theta \}}. \tag{41}$$

for the binomial distribution, the likelihood is

$$\mathcal{L} = C(n, k) p^k (1 - p)^{n-k} \tag{42}$$

and, based on the maximum likelihood estimate $p = k/n$, the unconstrained supremum is

$$\sup \{ \mathcal{L} \mid \theta \in \Theta \} = C(n, k) \left(\frac{k}{n} \right)^k \left(1 - \frac{k}{n} \right)^{n-k}. \tag{43}$$

For the specific null

$$H_0 : p = p_o \quad (44)$$

the conditional supremum is

$$\sup \{ \mathcal{L} \mid \theta \in \theta_o \} = \mathcal{L} = C(n, k) p_o^k (1 - p_o)^{n-k}. \quad (45)$$

So the LR is

$$\Lambda = \left(\frac{p_o}{k/n} \right)^k \left(\frac{1 - p_o}{1 - k/n} \right)^{n-k} = \left(\frac{np_o}{k} \right)^k \left(\frac{n - np_o}{n - k} \right)^{n-k}. \quad (46)$$

For data $X \sim B_{n,p}$, the LR is

$$\Lambda = \left(\frac{np_o}{X} \right)^X \left(\frac{n - np_o}{n - X} \right)^{n-X}. \quad (47)$$

4.1 Hypothesis Tests

Consider the two-sided test

$$\begin{aligned} H_0 : p &= p_o \\ H_1 : p &\neq p_o. \end{aligned} \quad (48)$$

Since Λ has small values significant, one rejects H_0 if $\Lambda \leq \Lambda_o$ for the appropriate critical value. Note that the null probabilities are $\Pr[\Lambda = \Lambda(k)] = \Pr[X = k] = f_{n,p_o}(k)$ and the $\Lambda(k)$ can be sorted in ascending order and the probabilities summed to obtain the cdf F_Λ

$$F_\Lambda(t) = \sum_{\Lambda(k) \leq t} f_{n,p_o}(k) = \sum_{k=0}^n I_{[\Lambda(k), 1]}(t) \cdot f_{n,p_o}(k). \quad (49)$$

In particular with $n = 10$ and $p_o = 0.33$, for $\alpha = 0.1$, the critical region $\Lambda \leq \Lambda(6)$ corresponds to $k \in \{0, 6, 7, 8, 9, 10\}$ and the non-rejection region is $k \in \{1, 2, 3, 4, 5\}$. See table 1.

Table 1. LR cdf.

k	$f_{n,p_0}(k)$	$t = \Lambda(k)$	$F_{\Lambda}(t)$
10	0.01823	0.00002	0.00002
9	0.08978	0.00080	0.00033
8	0.19899	0.00941	0.00317
0	0.26136	0.01823	0.02140
7	0.22528	0.05765	0.03678
6	0.13315	0.21789	0.09143
1	0.05465	0.23174	0.18121
5	0.01538	0.54106	0.31436
2	0.00284	0.65894	0.51335
4	0.00031	0.89817	0.73864
3	0.00002	0.97952	1.00000

Now, based on the usual CP procedure $\alpha/2$ upper and lower tails of the B_{n,p_0} test statistic, the critical region is $\{0,7,8,9,10\}$ and the non-rejection region is $\{1,2,3,4,5,6\}$. See table 2. Note that $X = 6$ is critical for the LR test but not for the CP test.

Table 2. CP cdf.

k	$F_{n,p_0}(k)$	$G_{n,p_0}(k)$
0	0.01823	1.00000
1	0.10801	0.98177
2	0.30700	0.89199
3	0.56837	0.69300
4	0.79365	0.43163
5	0.92680	0.20635
6	0.98145	0.07320
7	0.99683	0.01855
8	0.99967	0.00317
9	0.99998	0.00033
10	1.00000	0.00002

4.2 Confidence Intervals

As usual, confidence intervals are obtained by inverting non-rejection regions. One-sided LR CIs coincide with CP CIs. Coverage for the two-sided LR CIs is illustrated in figure 5. Two-sided LR CIs are in general not as conservative as CP CIs. This is regarded as an advantage of the LR method, as the intervals are “tighter”.

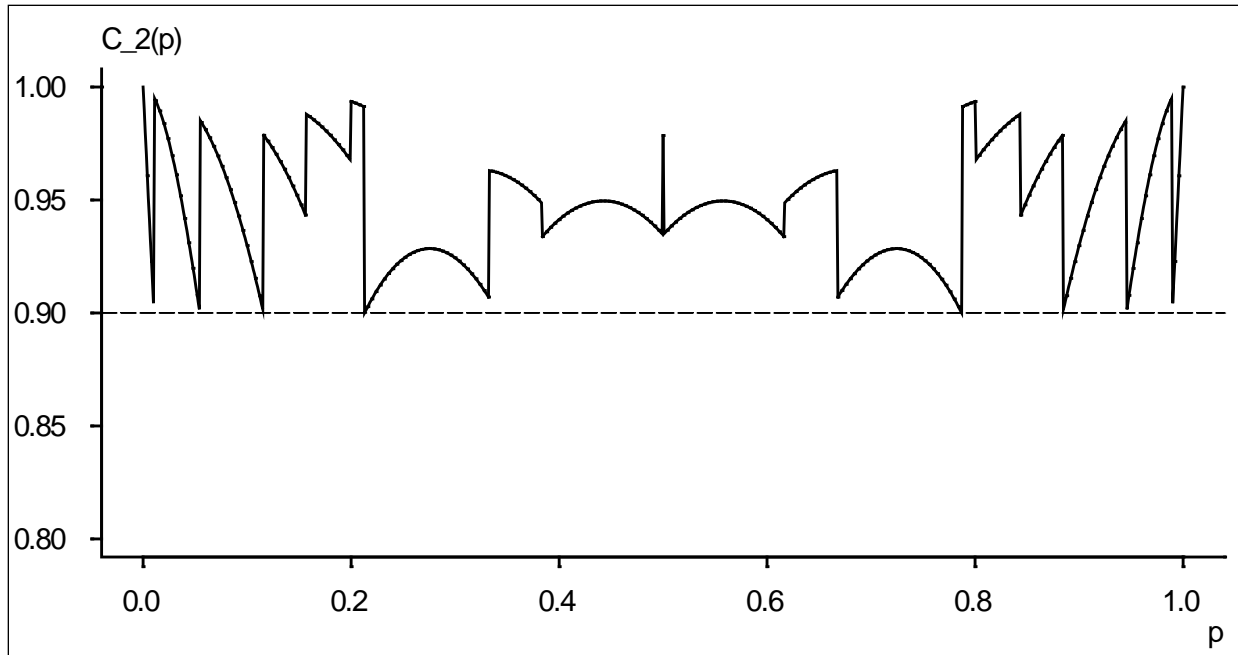


Figure 5. Two-sided LR CI coverage.

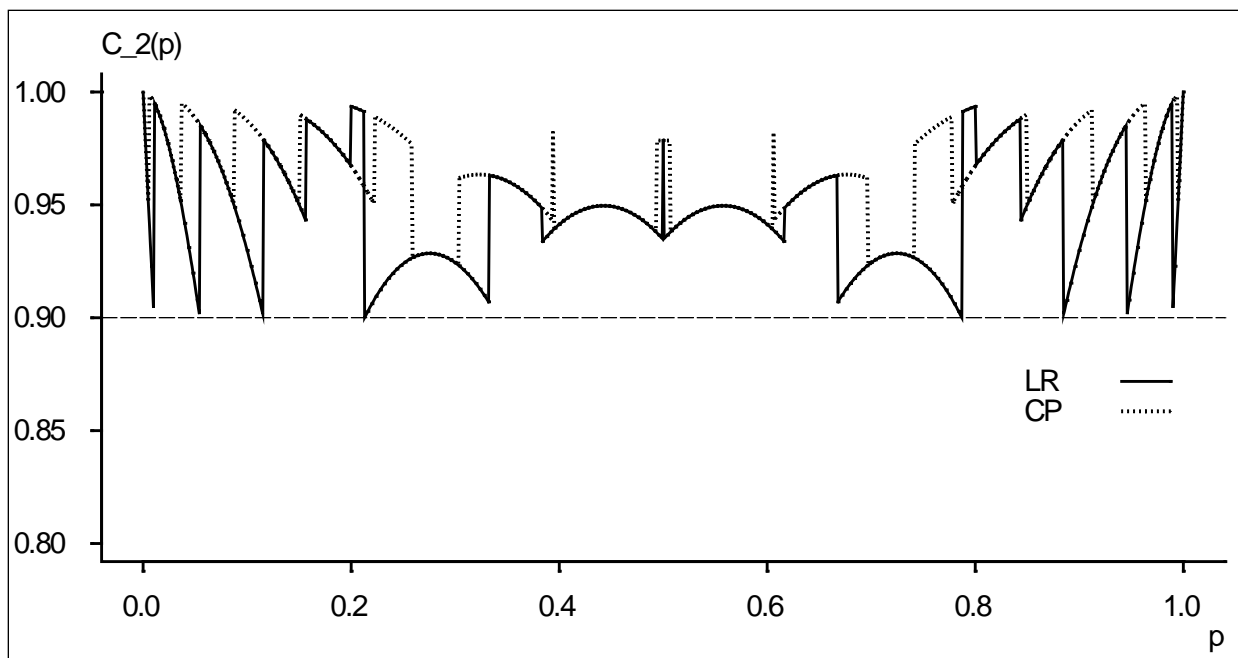


Figure 6. Two-sided CI coverage comparison.

4.3 Implementation

Code listings are in appendix A.

The function `binom.LR` evaluates the LR of equation 42 given n , k , and p . The function `binom.LR.F.x` evaluates the cdf components as in table 1, given n , p , and a technical

parameter `up` which sets the orientation for breaking ties. The function `binom.LR.F.z` evaluates $F_{\Lambda}(\Lambda(k))$. The hypothesis test of equation 48 is computed with `binom.LR.ht`, which presents the critical and non-rejection regions for the given values of n , p , and α . The function `binom.LR.ci` inverts the cdf to find CIs.

5. Reliability

5.1 Reliability Tables

Cooke (5) presents a tabulation of lower confidence limits p_L for various numbers of trials n and successes r and confidence levels γ .

This standard language may be misleading. The binomial parameter p is a reliability, which is a probability of success. The endpoints p_L and p_U of CIs $[p_L, p_U]$ on p are reliability limits, or limits of reliability, at the given confidence level γ . These “confidence limits” are not “limits of confidence.”

Code listings are in appendix A. The functions `pbinom` and `gbinom` evaluate $F_{n,p}(k)$ and $G_{n,p}(k)$, respectively.

The function `binom.rel.tab(n, r, g, type)` uses `binom.ci` to evaluate p_L and arranges the results in the proper format to reproduce Cooke’s tables, adding a column p for the parameter point estimate. The arguments of `binom.rel.tab` are the number of trials n , the number of successes r , confidence interval size g which should be between 0.5 and 1.0, and an integer `type` which can be 0, 1, or 2 to request a lower CI $(L, 1]$, upper CI $[0, U)$, or 2-sided CI (L, U) , respectively. The arguments r and g can be vectors, in which case the function fills out the desired table.

For example, the first table, with $n = 4$, is given by

```
binom.rel.tab(4, 4:2)

n r   p   80%   90%   95%   97.5%   99%   99.5%
4 4 1.00 0.66874 0.56234 0.472871 0.397635 0.316228 0.265915
4 3 0.75 0.41755 0.32046 0.248605 0.194120 0.140868 0.110885
4 2 0.50 0.21232 0.14256 0.097611 0.067586 0.041999 0.029445
```

For $n = 18$,

```
binom.rel.tab(18, 18:9)

n  r   p   80%   90%   95%   97.5%   99%   99.5%
18 18 1.00000 0.91447 0.87992 0.84668 0.81470 0.77426 0.74501
18 17 0.94444 0.84262 0.80053 0.76234 0.72706 0.68398 0.65365
18 16 0.88889 0.77700 0.73058 0.68974 0.65288 0.60881 0.57833
18 15 0.83333 0.71472 0.66559 0.62332 0.58582 0.54170 0.51159
18 14 0.77778 0.65469 0.60398 0.56112 0.52363 0.48011 0.45076
```

```

18 13 0.72222 0.59642 0.54498 0.50217 0.46520 0.42280 0.39452
18 12 0.66667 0.53962 0.48816 0.44595 0.40993 0.36909 0.34214
18 11 0.61111 0.48412 0.43328 0.39216 0.35745 0.31858 0.29318
18 10 0.55556 0.42982 0.38020 0.34060 0.30757 0.27101 0.24739
18 9 0.50000 0.37669 0.32885 0.29120 0.26019 0.22630 0.20465

```

Tables for $n = 4 - 50$ are in appendix B. Compare to Cooke.

For $n = 100$,

```
bino.rel.tab(100,c(100:90, 60:50))
```

```

n      r      p      80%      90%      95%      97.5%      99%      99.5%
100 100 1.00 0.98403 0.97724 0.97049 0.96378 0.95499 0.94840
100  99 0.99 0.97035 0.96166 0.95344 0.94554 0.93546 0.92804
100  98 0.98 0.95770 0.94765 0.93838 0.92962 0.91859 0.91057
100  97 0.97 0.94554 0.93441 0.92429 0.91482 0.90303 0.89452
100  96 0.96 0.93370 0.92165 0.91080 0.90074 0.88830 0.87937
100  95 0.95 0.92209 0.90923 0.89775 0.88717 0.87415 0.86486
100  94 0.94 0.91064 0.89706 0.88501 0.87397 0.86045 0.85083
100  93 0.93 0.89933 0.88510 0.87254 0.86108 0.84710 0.83720
100  92 0.92 0.88814 0.87330 0.86028 0.84844 0.83405 0.82389
100  91 0.91 0.87703 0.86165 0.84820 0.83602 0.82125 0.81085
100  90 0.90 0.86601 0.85012 0.83628 0.82378 0.80867 0.79805
...
100  60 0.60 0.55309 0.53115 0.51298 0.49721 0.47890 0.46647
100  59 0.59 0.54302 0.52106 0.50289 0.48714 0.46888 0.45648
100  58 0.58 0.53297 0.51100 0.49284 0.47712 0.45890 0.44655
100  57 0.57 0.52293 0.50096 0.48282 0.46713 0.44897 0.43666
100  56 0.56 0.51291 0.49095 0.47284 0.45719 0.43908 0.42682
100  55 0.55 0.50291 0.48096 0.46289 0.44728 0.42924 0.41704
100  54 0.54 0.49292 0.47100 0.45297 0.43741 0.41944 0.40730
100  53 0.53 0.48295 0.46107 0.44309 0.42758 0.40969 0.39762
100  52 0.52 0.47300 0.45116 0.43323 0.41779 0.39999 0.38798
100  51 0.51 0.46306 0.44128 0.42341 0.40804 0.39033 0.37840
100  50 0.50 0.45314 0.43142 0.41362 0.39832 0.38072 0.36886

```

For $n = 500$,

```
bino.rel.tab(500,c(500:490, 260:250))
```

```

n      r      p      80%      90%      95%      97.5%      99%      99.5%
500 500 1.000 0.99679 0.99541 0.99403 0.99265 0.99083 0.98946
500 499 0.998 0.99402 0.99224 0.99055 0.98891 0.98680 0.98523
500 498 0.996 0.99146 0.98939 0.98746 0.98563 0.98330 0.98159
500 497 0.994 0.98900 0.98669 0.98457 0.98257 0.98005 0.97822
500 496 0.992 0.98659 0.98408 0.98179 0.97964 0.97697 0.97503
500 495 0.990 0.98423 0.98153 0.97909 0.97682 0.97399 0.97196
500 494 0.988 0.98191 0.97903 0.97645 0.97407 0.97111 0.96898
500 493 0.986 0.97960 0.97657 0.97387 0.97137 0.96829 0.96608
500 492 0.984 0.97732 0.97414 0.97132 0.96872 0.96552 0.96324
500 491 0.982 0.97505 0.97174 0.96880 0.96611 0.96280 0.96044
500 490 0.980 0.97280 0.96935 0.96631 0.96353 0.96012 0.95769
...
500 260 0.520 0.50018 0.49035 0.48223 0.47520 0.46704 0.46148
500 259 0.518 0.49818 0.48835 0.48024 0.47321 0.46504 0.45949

```

500	258	0.516	0.49618	0.48635	0.47824	0.47121	0.46305	0.45751
500	257	0.514	0.49418	0.48435	0.47624	0.46922	0.46106	0.45552
500	256	0.512	0.49218	0.48236	0.47425	0.46723	0.45907	0.45353
500	255	0.510	0.49019	0.48036	0.47226	0.46523	0.45709	0.45155
500	254	0.508	0.48819	0.47836	0.47026	0.46324	0.45510	0.44956
500	253	0.506	0.48619	0.47637	0.46827	0.46125	0.45311	0.44758
500	252	0.504	0.48419	0.47437	0.46628	0.45926	0.45112	0.44559
500	251	0.502	0.48220	0.47238	0.46428	0.45727	0.44914	0.44361
500	250	0.500	0.48020	0.47038	0.46229	0.45529	0.44716	0.44163

5.2 Sample Size

One application of this is to find the minimal sample size N to obtain a lower reliability limit at least p_L with confidence γ for a certain $n_f = n - r$, so $r = n - n_f$.

For example, with $n_f = 0$ failures ($r = n$) and 90% reliability limit $p_L = 0.9$ with 95% confidence $\gamma = 0.95$, one can wade through Cooke's phone-book-sized report, or run `binom.rel.tab` to find suitable n , which bound the desired reliability limit,

`binom.rel.tab(n=28, r=28, g=0.95)` gives $p_L = 0.89853$ and
`binom.rel.tab(n=29, r=29, g=0.95)` gives $p_L = 0.90186$, so $N \geq 29$.

For $n_f = 1$ failure, $r = n - 1$

`binom.rel.tab(n=45, r=44, g=0.95)` gives $p_L = 0.89887$ and
`binom.rel.tab(n=46, r=45, g=0.95)` gives $p_L = 0.90098$, so $N \geq 46$.

By equation 31, this amounts to finding

$$N = \inf\{n \mid G_{n,p_L}(n - n_f) \leq 1 - \gamma\}. \quad (50)$$

For $n_f = 0$ failures, $p_L = 0.9$, and $\gamma = 0.95$,

$G_{28,0.9}(28) = \text{gbinom}(28, 28, 0.9) = 0.052335$ and
 $G_{29,0.9}(29) = \text{gbinom}(29, 29, 0.9) = 0.047101$, so $N \geq 29$.

For $n_f = 1$ failures, $p_L = 0.9$, and $\gamma = 0.95$,

$G_{45,0.9}(45) = \text{gbinom}(45, 45, 0.9) = 0.052368$ and
 $G_{46,0.9}(46) = \text{gbinom}(46, 46, 0.9) = 0.048004$, so $N \geq 46$.

By equation 11, this is equivalent to

$$N = \inf\{n \mid F_{n,1-p_L}(n_f) \leq 1 - \gamma\}. \quad (51)$$

For $n_f = 0$ failures, $p_L = 0.9$, and $\gamma = 0.95$,

$F_{28,0.9}(0) = \text{pbinom}(0, 28, 0.1) = 0.052335$ and
 $F_{29,0.9}(0) = \text{pbinom}(0, 29, 0.1) = 0.047101$, so $N \geq 29$.

For $n_f = 1$ failures, $p_L = 0.9$, and $\gamma = 0.95$,

$$F_{45,0.9}(1) = \text{pbinom}(1, 45, 0.1) = 0.052368 \text{ and}$$
$$F_{46,0.9}(1) = \text{pbinom}(1, 46, 0.1) = 0.048004, \text{ so } N \geq 46.$$

The function `binom.rel.size(nf, p, g, x)` takes arguments representing the `nf`, desired lower reliability limit `p`, confidence `g`, and an optional scale factor `x` for the numerical solution routine. It calculates the required sample size N and returns a named list holding

<code>\$par:</code>	parameters <code>nf</code> , <code>p</code> , and <code>g</code>
<code>\$n:</code>	required minimum sample size, N
<code>\$p:</code>	reliability limit for confidence <code>g</code> and sample size N
<code>\$g:</code>	confidence for reliability limit <code>p</code> and sample size N
<code>\$n1:</code>	the next lower sample size, $N - 1$
<code>\$p1:</code>	reliability limit for confidence <code>g</code> and sample size $N - 1$
<code>\$g1:</code>	confidence for reliability limit <code>p</code> and sample size $N - 1$

One expects that `$p1 < p ≤ $p` and `$g1 < g ≤ $g`.

For example,

```
binom.rel.size(nf = 1, p = 0.9, g = 0.95)
```

returns

```
$par$nf: 1
$par$p: 0.9
$par$g: 0.95
$n: 46
$p: 0.90098
$g: 0.952
$n1: 45
$p1: 0.89887
$g1: 0.94763
```

Consulting the tables for the sample sizes $n = 46$ and $n1 = 45$ and confidence $g = 0.95$ for the number of successes $r = n - n_f = 45$ and $r1 = n1 - n_f = 44$, one sees the bounding reliability limits $p = 0.90098$, and $p1 = 0.89887$.

6. References

1. Stuart, A.; Ord, J. K. *Kendall's Advanced Theory of Statistics: Distribution Theory*; Oxford University Press: New York, 1994.
2. Mood, A.; Graybill, F. A.; Boes, D. C. *Introduction to the Theory of Statistics. 3*; McGraw-Hill, New York, 1974.
3. Stuart, A.; Ord, J. K. *Kendall's Advanced Theory of Statistics: Classical Inference and Relationship*; Oxford University Press, New York, 1991.
4. Clopper, C. J.; Pearson, E. S. *The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial*. 1934, 4, *Biometrika*, Vol 26, pp 404–413.
5. Cooke, J. R.; Lee, M. T.; Vanderbeck, J. P. *Binomial Reliability Table (Lower Confidence Limits for the Binomial Distribution)*. China Lake, CA, Bureau of Naval Weapons, 1964.

Appendix A. S-PLUS Code

A.1 Library Functions

```
pbinom(q,  
      size = stop("no size arg"),  
      prob = stop("no prob arg"))
```

is the S-PLUS library implementation of the Binomial distribution cdf.

In the notation of this report,

$$\text{pbinom}(k, n, p) = F_{n,p}(k) = \Pr[B_{n,p} \leq k]$$

where, k is number of successes, n is the number of trials, and p is the probability of success.

```
qbeta(p,  
      shape1 = stop("no shape1 arg"),  
      shape2 = stop("no shape2 arg"))
```

is the S-PLUS library implementation of the Beta distribution qf.

In the notation of this report,

$$\text{qbeta}(u, a, b) = Q_{\text{Beta}(a,b)}(u) = \inf \{x \mid F_{\text{Beta}(a,b)}(x) \geq u\}$$

where, u is the desired probability, and a and b are the usual Beta distribution parameters.

A.2 Upper-Tail Binomial Probabilities

`gbinom` calculates upper tail binomial probabilities.

```
gbinom <- function(q,  
                  size = stop("no size arg"),  
                  prob = stop("no prob arg"))  
{  
  ##  
  ## Pr [ Binomial(size,prob) >= q ]  
  ##  
  ## underflow error  
  ## 1-pbinom(q-1,size,prob)  
  ##  
  ## no underflow  
  ##  
  pbinom ( size - q , size , 1 - prob )  
}
```

In the notation of this report,

$$\text{gbinom}(k, n, p) = G_{n,p}(k) = \text{pbinom}(n-k, n, 1-p)$$

where, k is number of successes, n is the number of trials, and p is the probability of success.

A.3 Binomial Parameter Confidence Interval: CP

`binom.ci` calculates a CP confidence interval on a binomial parameter estimate.

```
binom.ci <- function(n          ,          # sample size
                    r          ,          # number of successes
                    g = 0.8    ,          # CI size
                    type = 0    ,          # 0,1,2 for lower, upper, 2-sided
                    a = 1-g    )          # 1 - (CI size)
{
  ## 100g% Confidence Interval [ L , U ] on Binomial Parameter p
  ##
  ## type 0 : lower CI = ( L , 1 ] : Pr [ B(n,L) >= r ] = 1-g
  ## type 1 : upper CI = [ 0 , U ) : Pr [ B(n,U) <= r ] = 1-g
  ## type 2 : two-sided CI = (L,U) : Pr[B(n,L)>=r] = Pr[B(n,U)<=r]=(1-g)/2

  if(type==2) a <- a/2

  if ( type==1 || r==0 )
    L <- 0
  else
    L <- qbeta ( a , r , n-r+1 )

  if(type==0 || r==n)
    U <- 1
  else
    U <- 1 - qbeta ( a , n-r , r+1 )

  list(n=n, r=r, L=L, p.hat=r/n, U=U, I=c(L,U))
}
```

A.4 Binomial Reliability Table

`binom.rel.tab` tabulates binomial confidence limits.

```
binom.rel.tab <- function(n = 100 ,          # number of trials
                        r = n    ,          # number of successes
                        g = c(.8,.9,.95,.975,.99,.995),# confidence levels
                        type = 0    )          # 0,1 = lower, upper
{
  nr <- length(r)          # success vector length
  ng <- length(g)          # number of CI levels
  z <- matrix(NA, nc=ng, nr=nr)          # reliability limits

  for( i in 1:nr ) for(j in 1:ng )          # compute reliabilities
    z[i,j] <- binom.ci(n,r[i],g[j],type)$I[type+1]

  z <- cbind(n, r, r/n, z)          # augmented table
  dimnames(z) <-          # name the table columns
    list(NULL, c("n","r","p",paste(100*g,"%", sep="")))

  z          # return table
}
```

A.5 Binomial Reliability Sample Size

`binom.rel.size` calculates sample size for reliability with a given confidence limit.


```

bino.rel.size <- function(nf = 0      , # number of failures
                        p = 0.9     , # reliability limit
                        g = 0.95    , # confidence level
                        x = 6       ) # log_2 ( search delta )
{
  ## N = inf { n : pbinom(nf, n, 1-p) } <= 1-g }

  a <- 1-g                # confidence
  q <- 1-p                # reliability

  f <- function(nf, n, q) list(n=n, a=pbinom(nf, n, q)) # ( n , a )

  rtn <- function(nf, p, g, z1, z) # return list
    list(par=list(nf=nf, p=p, g=g),
         n=z$n, p=bino.ci(z$n, z$n-nf, g, 0)$L, g=1-z$a ,
         n1=z1$n, p1=bino.ci(z1$n, z1$n-nf, g, 0)$L, g1=1-z1$a)

  z <- f(nf, 2^x, q)      # initial ( n , a ) guess
  while ( z$a < a ) z <- f(nf, z$n/2, q)
  while ( z$a > a ) z <- f(nf, z$n*2, q)

  z0 <- z                # upper end at n
  z1 <- f(nf, z0$n/2, q) # lower end at n/2
  dn <- (z0$n-z1$n)/2    # interval radius
  while ( dn >= 1 ) {
    z <- f(nf, z1$n+dn, q) # midpoint
    if ( z$a < a )        # one midpoint becomes new endpoint
      z0 <- z else z1 <- z
    dn <- dn/2           # new radius
  }

  rtn(nf, p, g, z1, z0)
}

```

The algorithm finds sample sizes $n_1 = 2^{k-1}$ and $n = 2^k$ that bound the confidence for the desired reliability limit. Then it repeatedly bisects the interval $[n_1, n]$ and replaces one endpoint with the midpoint, maintaining the bound, until $n - n_1 = 1$. Reliability limits at the desired confidence level should now bound the desired reliability limit. Using powers of 2 ensures that the sample sizes are integers.

A.6 Binomial LR

```

bino.LR <- function(k , # successes
                  n , # trials
                  p ) # parameter
{
  ## binomial LR
  ( n*p/k ) ^ k * ( n*(1-p)/(n-k) ) ^ (n-k)
}

```

A.7 Binomial LR cdf Table

```

bino.LR.F.x <- function(n , # trials
                      p , # null parameter
                      up = T ) # tie-breaker direction
{

```

```

## binomial(n,p) LR cdf: complete table
k <- 0:n                               # successes
f <- dbinom(k,n,p)                     # probabilities
L <- bino.LR(k,n,p)                     # LR values

if(up) I <- order(L,k)                  # order LR values, tiebreaker
else I <- order(L,-k)
L <- L[I]                                # sort LR values
k <- k[I]                                # order successes by LR value
F.L <- cumsum(f[I])                     # LR cdf
cbind(k, I, f, L, F.L)
}

```

A.8 Binomial LR cdf Evaluation

```

bino.LR.F.z <- function(k      ,      # trials
                        n      ,      # successes
                        p      ,      # null parameter
                        x=F )      # extended return
{
  ## binomial(n,p) LR cdf (k)
  z <- bino.LR.F.x(n=n, p=p, k>n*p)
  u <- order(z[, "I"])
  if(x) z[u[k+1],]
  else z[u[k+1], "F.L"]
}

```

A.9 Binomial LR Hypothesis Test

```

bino.LR.ht <- function(n      ,      # trials
                       p      ,      # null parameter
                       a = 0.1 )      # significance
{
  ## binomial LR hypothesis test
  z <- bino.LR.F.x(n=n, p=p)

  I.rej <- z[, "F.L"] <= a
  k.rej <- sort(z[, "k"])[I.rej]
  k.acc <- sort(z[, "k"])[!I.rej]
  list(z=z, k.acc=k.acc, k.rej=k.rej)
}

```

A.10 Binomial Parameter Confidence Interval: LR

```

bino.LR.ci <- function(n = 100 ,      # sample size
                       r = 50  ,      # number of successes
                       g = 0.9 )      # CI size
{
  h <- r/n
  q <- 1-g

  ## [0,h] : L is increasing
  if(r == 0)
    p0 <- list(neg=0, pos=0, f0=q, f1=q)
  else
    p0 <- uniroot(function(x, n., r., q.) bino.LR.F.z(k=r., n=n., p=x)-q.,
                  lower=0, upper=h, q.=q, n.=n, r.=r)
}

```

```

## [h,1] : L is decreasing
if(r == n)
  p1 <- list(neg=1, pos=1, f0=q, f1=q)
else
  p1 <- uniroot(function(x, n., r., q.) bino.LR.F.z(k=r., n=n., p=x)-q.,
                lower=h, upper=1, q.=q, n.=n, r.=r)

list(n=n, r=r, L=p0$pos, p.hat=h, U=p1$neg, I=c(p0$pos, p1$neg))
}

```

INTENTIONALLY LEFT BLANK.

List of Symbols, Abbreviations, and Acronyms

cdf	cumulative distribution function
CI	confidence interval
CP	Clopper-Pearson
iid	independent and identically distributed
LR	likelihood ratio
MLE	maximum likelihood estimator
nf	number of failures
pdf	probability density function
qf	quantile function

NO. OF
COPIES ORGANIZATION

<p>1 (PDF only)</p>	<p>DEFENSE TECHNICAL INFORMATION CTR DTIC OCA 8725 JOHN J KINGMAN RD STE 0944 FORT BELVOIR VA 22060-6218</p>	<p><u>ABERDEEN PROVING GROUND</u></p>
<p>1</p>	<p>DIRECTOR US ARMY RESEARCH LAB IMNE ALC HRR 2800 POWDER MILL RD ADELPHI MD 20783-1197</p>	<p>1 US ARMY DEV TEST COM CSTE DTC TT T 314 LONGS CORNER RD APG MD 21005-5055</p>
<p>1</p>	<p>DIRECTOR US ARMY RESEARCH LAB RDRL CIM L 2800 POWDER MILL RD ADELPHI MD 20783-1197</p>	<p>1 US ARMY EVALUATION CTR CSTE AEC SVE R LAUGHMAN 4120 SUSQUEHANNA AVE APG MD 21005-3013</p>
<p>1</p>	<p>DIRECTOR US ARMY RESEARCH LAB RDRL CIM P 2800 POWDER MILL RD ADELPHI MD 20783-1197</p>	<p>22 DIR USARL RDRL SL J BEILFUSS J FEENEY J FRANZ M STARKS P TANENBAUM RDRL SLB G MANNIX RDRL SLB A D FARENWALD RDRL SLB D R GROTE RDRL SLB E M PERRY RDRL SLB G P MERGLER RDRL SLB S S SNEAD RDRL SLB W L ROACH RDRL CIM G (BLDG 4600) RDRL SLB D J COLLINS (5 CPS) J EDWARDS T HOLDREN L MOSS E SNYDER</p>
<p>1 (CD only)</p>	<p>ASST SECY ARMY ACQSTN LOGISTICS & TECH SAAL ZP RM 2E661 103 ARMY PENTAGON WASHINGTON DC 20310-0103</p>	
<p>1 (CD only)</p>	<p>ASST SECY ARMY ACQSTN LOGISTICS & TECH SAAL ZS RM 3E448 103 ARMY PENTAGON WASHINGTON DC 20310-0103</p>	
<p>1 (CD only)</p>	<p>DIRECTOR FORCE DEV DAPR FDZ RM 3A522 460 ARMY PENTAGON WASHINGTON DC 20310-0460</p>	
<p>1</p>	<p>US ARMY TRADOC ANL CTR ATRC W A KEINTZ WSMR NM 88002-5502</p>	
<p>1</p>	<p>USARL RDRL SLE R FLORES WSMR NM 88002-5513</p>	