

Contents

Problem Statement	2
Binomial Log-Likelihood	3
Confidence Intervals	4
Confidence Intervals Based On Likelihood Ratio	5
Estimating The Likelihood Ratio Confidence Interval	6
Coverage Probability	7
Properties of $C(p, x, n)$	8
Coverage Probability For Several Methods	9
Confidence Intervals At The Boundaries	10
Adjusting Confidence Intervals At The Boundaries	11
Optimal Coverage	12
Coverage Using Optimal “ α_0 ”	13
Adjustments To Significance Probabilities	14
Summary	15
The <code>binom</code> package	16
References	17

Adjusting Likelihood Ratio Confidence Intervals for Parameters Near Boundaries Applied to the Binomial

Sundar Dorai-Raj (sundar.dorai-raj@pdf.com)
Spencer Graves (spencer.graves@pdf.com)

August 10, 2006

Problem Statement

- Interval estimation on p is not simple and there seems to be no agreement on which is best
- Intervals when there are 0% or 100% passes tend to be too short
 - A standard adjustment for “parameter at a boundary” assumes $2 * \log(\text{likelihood ratio})$ is a mixture of chi-squares
 - For binomial confidence intervals, this is equivalent to using $\chi^2_{1-\alpha/2}$ in place of $\chi^2_{1-\alpha}$
 - How well does this work?
 - Can we find something better that is almost as simple?

Binomial Log-Likelihood

- The binomial log-likelihood is given by

$$\ell(p, x, n) = \log \binom{n}{x} + x \log(p) + (n - x) \log(1 - p)$$

where n is the number of independent Bernoulli trials, x is the number of successes out of n , and p is the probability of success

- The Maximum Likelihood Estimate (MLE) of p is given by

$$\hat{p} = \frac{x}{n}$$

Confidence Intervals

- Interval estimates of p are difficult to achieve due to the discreteness and skewness (for $p \neq 0.5$) of the binomial distribution
- Many methods have been devised to estimate confidence intervals on p
 - Likelihood methods: generalized linear models, likelihood ratio, asymptotic
 - Bayesian
 - Inversion methods: Wilson, Agresti-Coulls, Fleiss, Clopper-Pearson
- The asymptotic method is woefully poor but still part of most standard statistics curricula

Confidence Intervals Based On Likelihood Ratio

- The likelihood ratio test statistic is define by

$$\Lambda(p_0, \hat{p}, x, n) = \ell(\hat{p}, x, n) - \ell(p_0, x, n) \sim \chi_1^2,$$

where

$$\hat{p} = \frac{x}{n}$$

is the MLE and p_0 is the probability of success under the null hypothesis

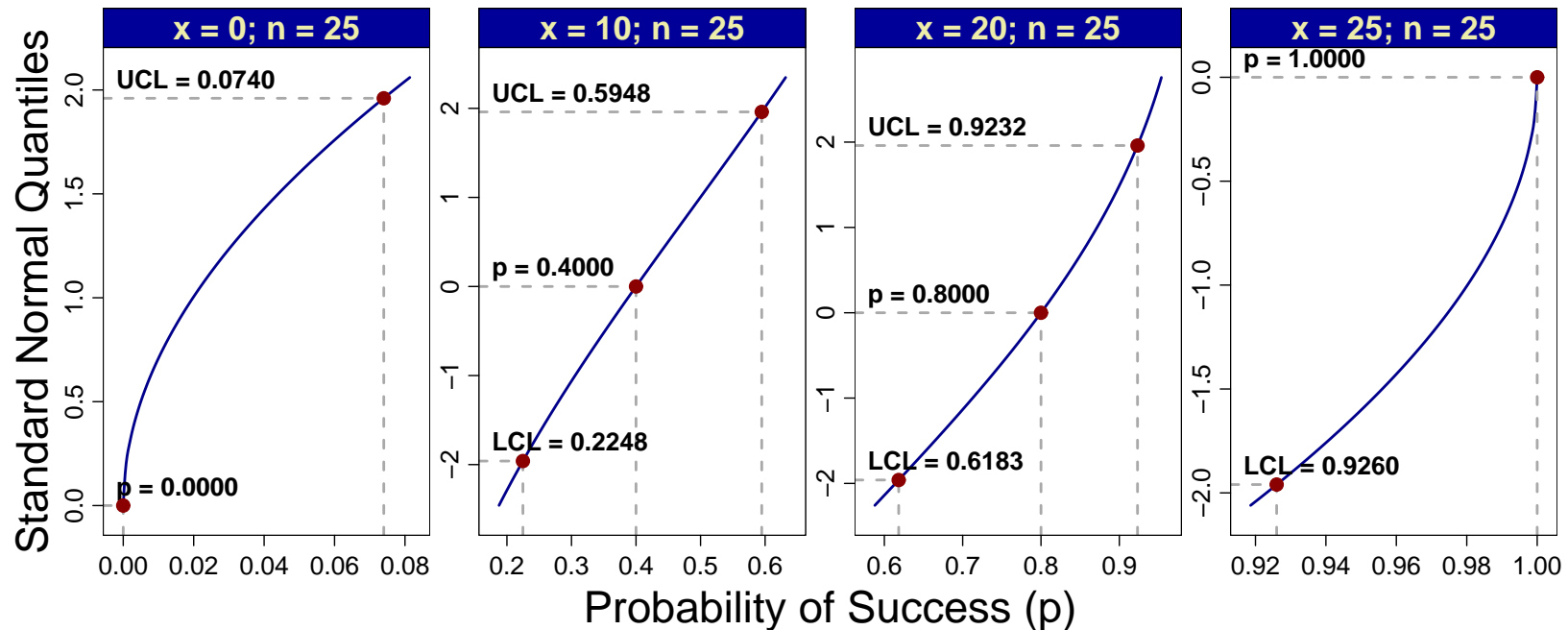
- Inverting L , we obtain a confidence interval on p :

$$LCL = \arg \max_{0 < p < 1} \{ \Lambda(p, \hat{p}, x, n) - 0.5\chi_1^2 \}$$

$$UCL = \arg \min_{0 < p < 1} \{ \Lambda(p, \hat{p}, x, n) - 0.5\chi_1^2 \}$$

Estimating The Likelihood Ratio Confidence Interval

- The method requires an iterative root-finding algorithm to find the lower and upper confidence bound



- We will refer to this interval estimate as “LRT”

Coverage Probability

- Coverage probability determines the expected value of any interval estimate over the binomial density function

$$C(p, x, n) = \sum_{x=0}^n I(LCL < p < UCL) \binom{n}{x} p^x (1 - p)^{n-x}$$

where p is the true probability of success, and (LCL, UCL) is an interval estimate of p

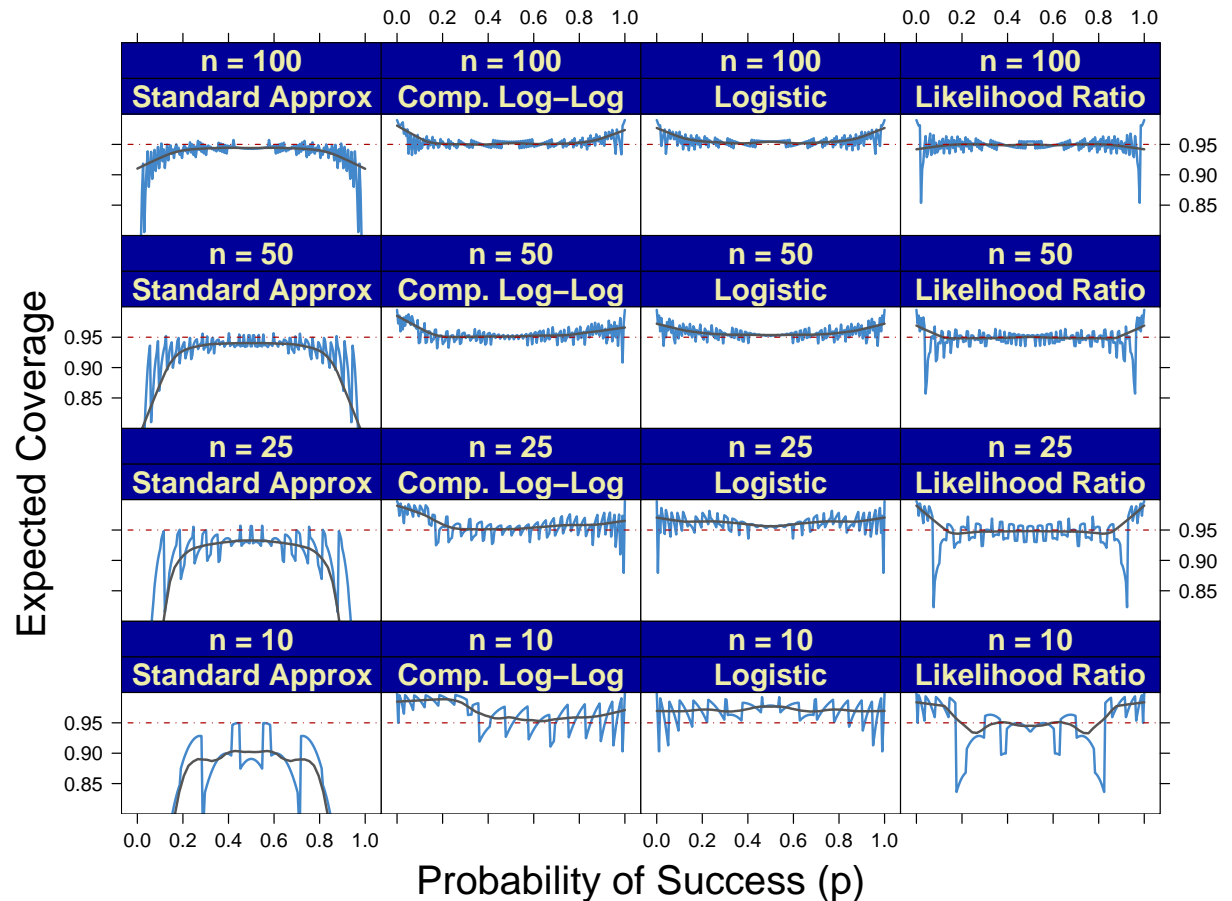
Properties of $C(p, x, n)$

- Should be close to the level of confidence $(1 - \alpha)$
- Oscillates due to the discreteness and skewness of the binomial distribution
- There are $2 \cdot n$ discontinuities (jumps) which exist at the each confidence interval endpoint

Coverage Probability For Several Methods

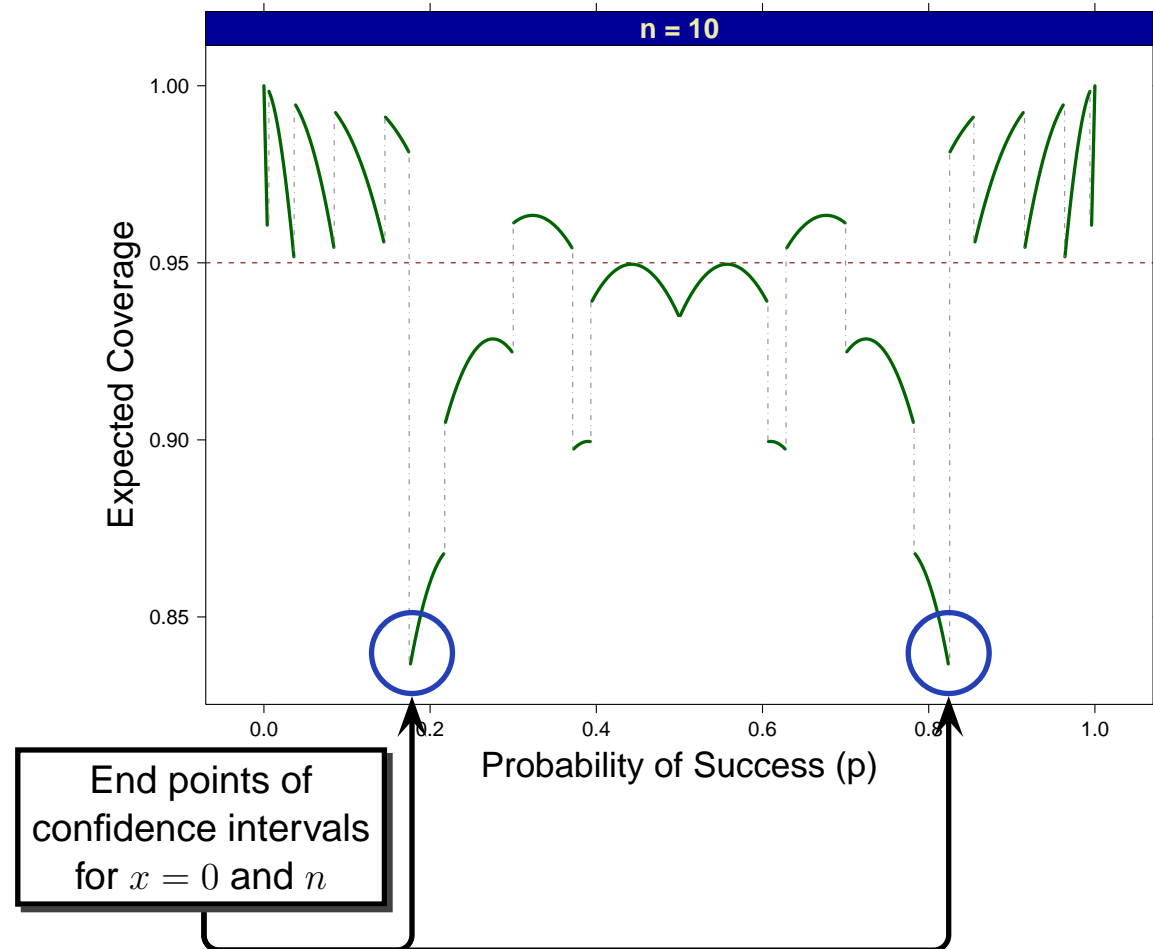
■ The LRT method has the best coverage probability

- The standard (asymptotic) method is absolutely the worst
- The complementary log-log is not symmetrical



Confidence Intervals At The Boundaries

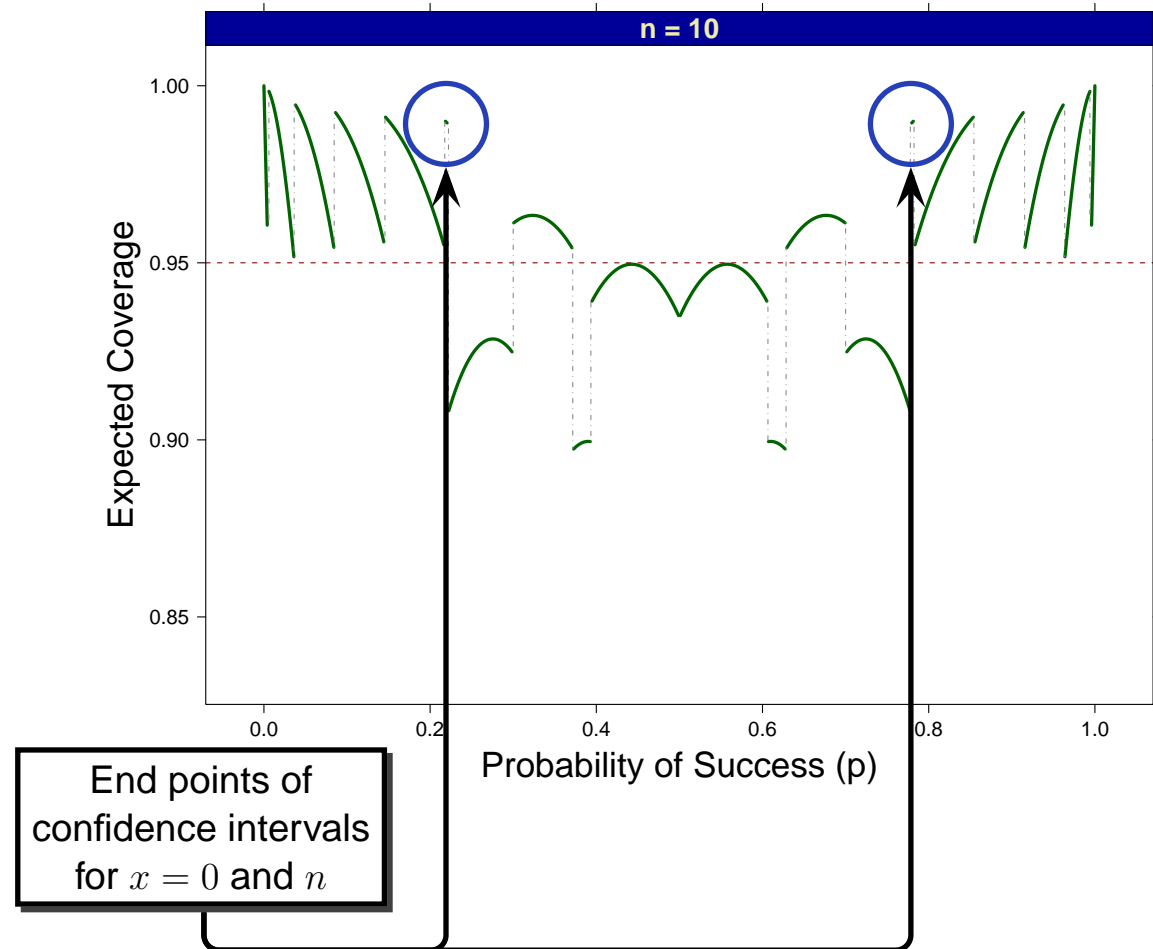
- **Coverage of intervals when $x = 0$ or n is not optimal**
 - Expected coverage is much less than $1 - \alpha$ for p close to the interval end
 - Adjusting the α downward improves coverage by increasing the interval length



Adjusting Confidence Intervals At The Boundaries

■ Changing the significance probability from 0.05 to 0.025 improves the coverage

- Still not optimal as the discontinuity is too large
- Coverage is too high because length of intervals when $x \neq 0$ or n seem to be too long



Optimal Coverage

- Minimize the squared area between the expected coverage and the desired level of confidence

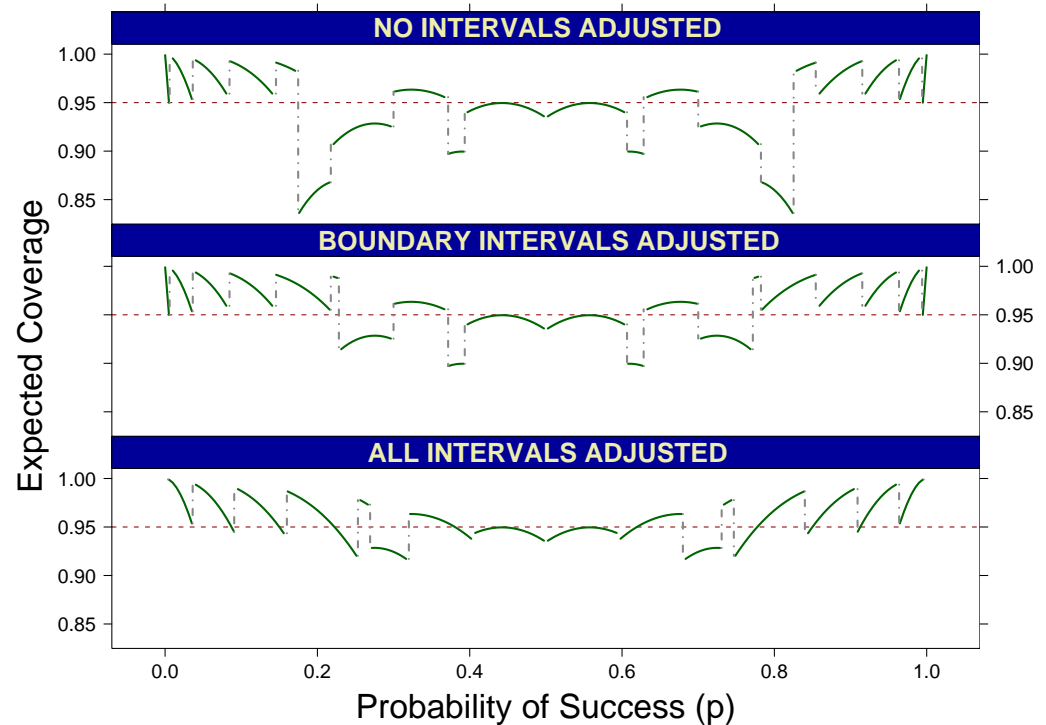
$$\alpha_0 = \arg \min_{0 < \alpha < 1} \int_0^1 [C(p, x, n) - (1 - \alpha)]^2 dp$$

- Minimizing latter objective function can be achieved by adjusting α for all x or simply for $x = 0$ and n
 - Adjusting only the boundary intervals is computationally fairly fast for relatively small n
 - Adjusting all the intervals can be slow even for small n

Coverage Using Optimal “ α_0 ”

- Using an optimal “ α_0 ” improves coverage
- Example with $n = 10$
 - Adjusting boundary intervals only, α_0 is 0.023 when $x = 0$ or 10
 - Adjusting all intervals, α_0 is 0.012 for the boundary intervals but monotonically increasing to 0.095 when $x = 5$

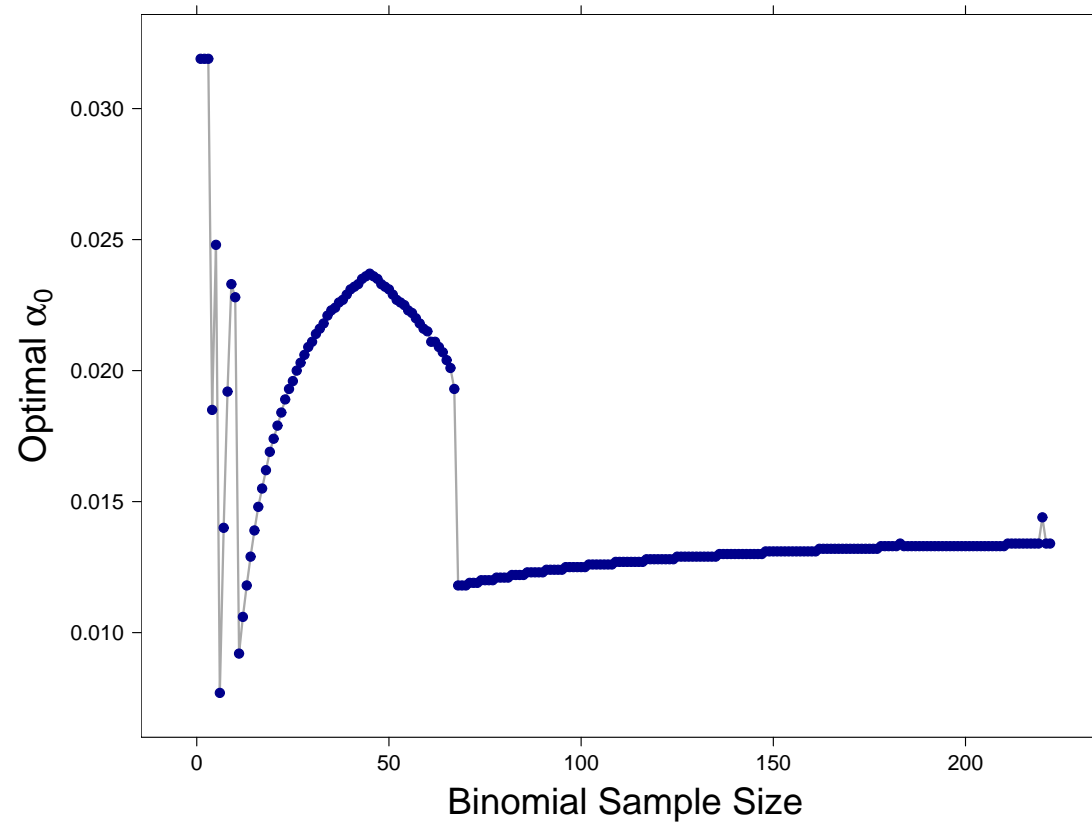
Optimal Probability Coverage for $n = 10$




x	0	1	2	3	4	5
α_0	0.012	0.030	0.050	0.061	0.070	0.095

Adjustments To Significance Probabilities


- Optimal confidence level asymptotes around 0.14 for $x = 0$ or n



Summary

- Using the LRT confidence interval produces the best coverage but are not computable by hand
- Confidence intervals for p when the observed number of successes are close to 0 or n are too short using a constant level of confidence
- There is no solution independent of n for adjusting a confidence intervals at the boundaries
- **Final recommendation:**
 - Use the LRT confidence interval (see next slide for software)
 - For $x = 0$ or n set α between 0.015 and 0.025
 - For obtaining all confidence interval adjustments use the `binom` package in 

The binom package

- An  package for constructing confidence intervals on the probability of success in a binomial experiment via several parameterizations
 - Bayes, LRT, probit, logit, cloglog
 - Coverage plotting
 - Optimal coverage
 - Sample size calculation and Power curves
 - Tcl/Tk interface for Power curves

References

- Pinheiro and Bates (2000) *Mixed-Effects Models in S and S-PLUS* (Springer, sec. 2.4)
- Crainiceanu, Ruppert and Vogelsang (2003) “Some properties of Likelihood Ratio tests in linear mixed models”
- Crainiceanu, Ruppert, Claeskens, and Wand (2005) “Exact Likelihood Ratio Tests for Penalized Splines”, *Biometrika*, 92(1)
- Brown, Cai, and DasGupta (2003) “Interval estimation in exponential families”, *Statistica Sinica*, 13: 19-49.
- Jenő Jeiczigel (2000) “Confidence Intervals for the Binomial Parameter: Some new considerations” (tech report downloaded from http://bio.univet.hu/qp/Reiczigel_conf_int.pdf, 2006.05.01)
- Sauro and Lewis (2005) “Estimating completion rates from small samples using binomial confidence intervals: comparisons and recommendations”, *Proceedings of the Human Factors and Ergonomics Society*, 49th annual meeting, 2100-2104.