Exact Confidence Intervals for Binomial Proportions

Dr. Jimmy A. Doi
Department of Statistics
Cal Poly San Luis Obispo

CAL POLY

Email: jdoi@calpoly.edu
Web: www.calpoly.edu/∼jdoi

California State University
Northridge

March 25, 2004

Clinical Trial Dataset: Cancer

- Drug 1: The New Treatment
- Drug 2: The Standard Treatment

Response	Drug 1	Drug 2
Success	30	16
Failure	20	34
Total	50	50

- Proportion of Success on Drug 1 = 60%
- Proportion of Success on Drug 2 = 32%
- Treatment Difference = 28%

Does this data suggest Drug 1 is significantly better than Drug 2?

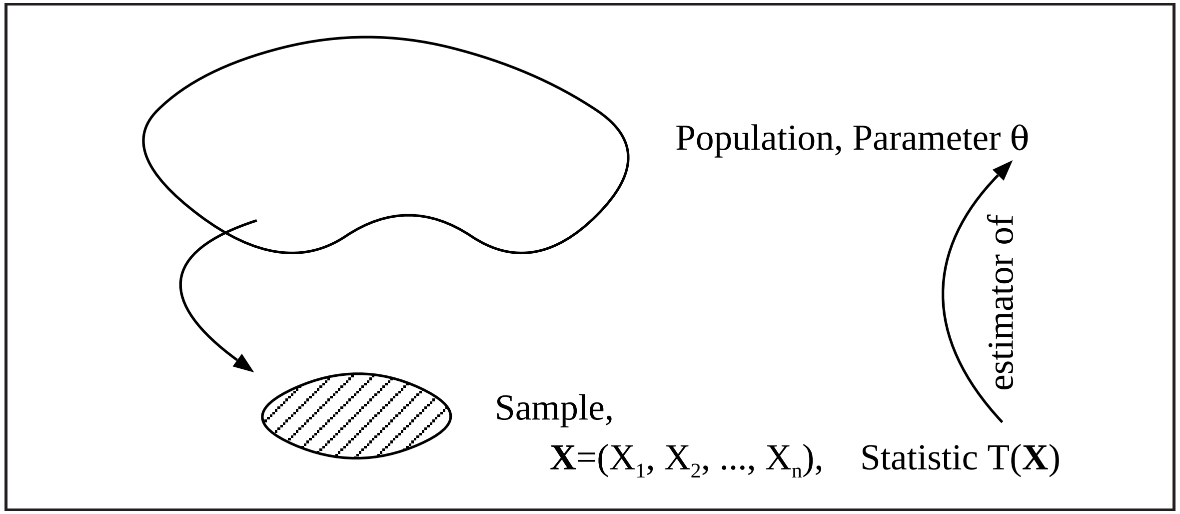
Confidence intervals are useful in addressing this question.

Outline

- Introduction
  - Confidence Interval for π
    - * Asymptotic vs. Exact Methods
- Current Research
  - Confidence Interval for $\pi_1 - \pi_2$
    - * Standard vs. Proposed Exact Methods
- Summary

Introduction

Statistical Inference

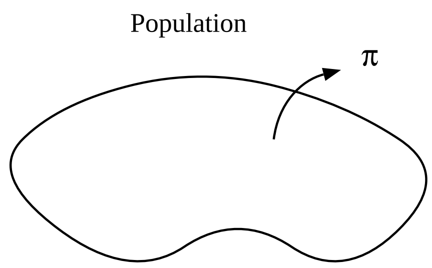


- Parameter of interest is any quantitative aspect of the population (true average μ, true percentage π)
- For this presentation, our focus will be on the population proportion π.

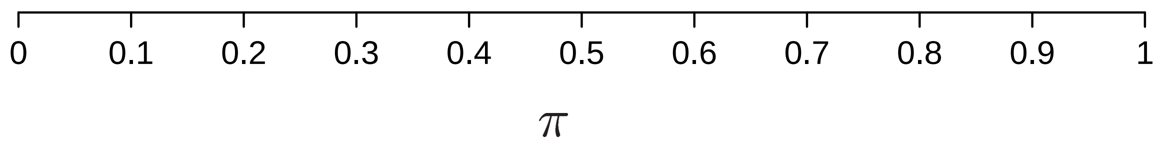
Example: Illicit Drug Usage

Survey Question to Undergraduates

"Have you ever experimented with an illegal drug/substance for <u>non-medicinal</u> purposes?"

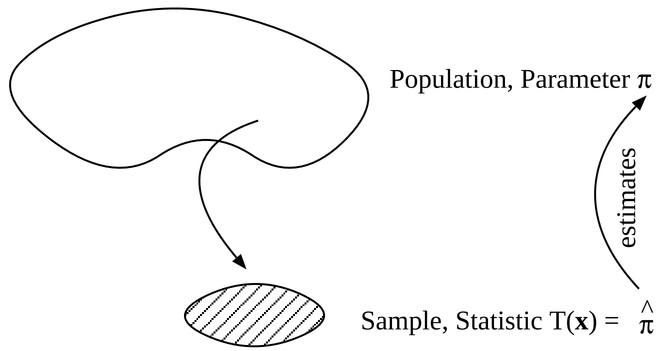


Parameter Space of π



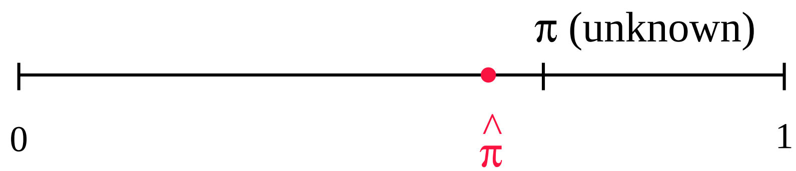
Example: Illicit Drug Usage (cont.)

Given a sample from the population, let the sample proportion be denoted as $\widehat{\pi}$.

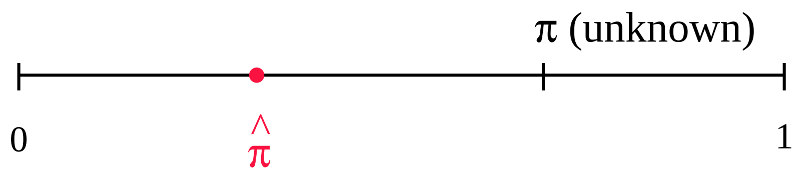


Problem:

Is the estimate close to the true value ...



... or relatively far away?

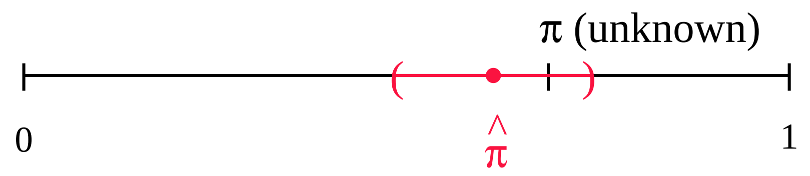


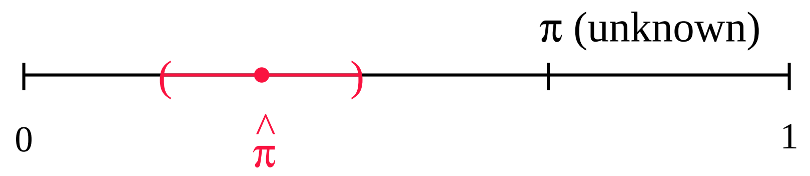
Confidence Intervals

Instead of a point estimator like $\widehat{\pi}$, use an interval estimator (l, u).

"Worth Going to War in Iraq?"

"Yes": 55% ± 3% (Gallup Poll, 3/18/04)





Confidence Intervals: Definitions

- Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ be a random sample from the population
- Generate intervals of the form

$$[L(\boldsymbol{X}), U(\boldsymbol{X})]$$

  which capture the true parameter θ for a high percentage of all possible samples.
- These intervals are known as **confidence intervals**.
- Given the interval estimator $[L(\boldsymbol{X}), U(\boldsymbol{X})]$, its
  - **coverage probability** is $P_\theta(\theta \in [L(\boldsymbol{X}), U(\boldsymbol{X})])$.
  - **confidence coefficient** (C) is $\inf_\theta P_\theta(\theta \in [L(\boldsymbol{X}), U(\boldsymbol{X})])$. (typical values of C are 90%, 95%, and 99%)

Confidence Intervals: Comparison

- Consider two competing CI methods:

$$[L_1(\boldsymbol{X}), U_1(\boldsymbol{X})] \text{ versus } [L_2(\boldsymbol{X}), U_2(\boldsymbol{X})]$$

- To compare competing methods, one should examine **coverage probability** and **interval length** properties

Ex. Recall the illicit drug usage proportion π:

Clearly, the interval [0,1] contains the true parameter. This is a 100% confidence interval for π, but it has no utility.

Asymptotic versus Exact Methods

Our parameter of interest, π, often stems from the proportion parameter of the binomial distribution:

- X is binomial(n, π), a discrete random variable
- Prob. Mass Function = $\text{bin}(x, n, \pi) = \binom{n}{x}\pi^x(1-\pi)^{n-x}$

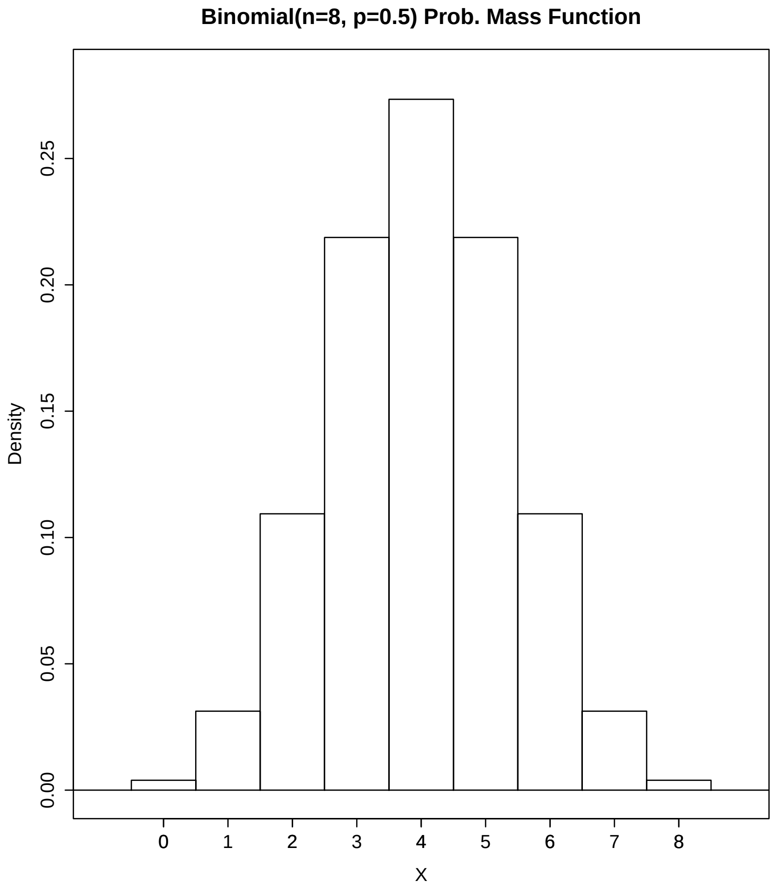
Problem: Using $\text{bin}(x, n, \pi)$ to create a CI for π can be an inconvenient and cumbersome chore

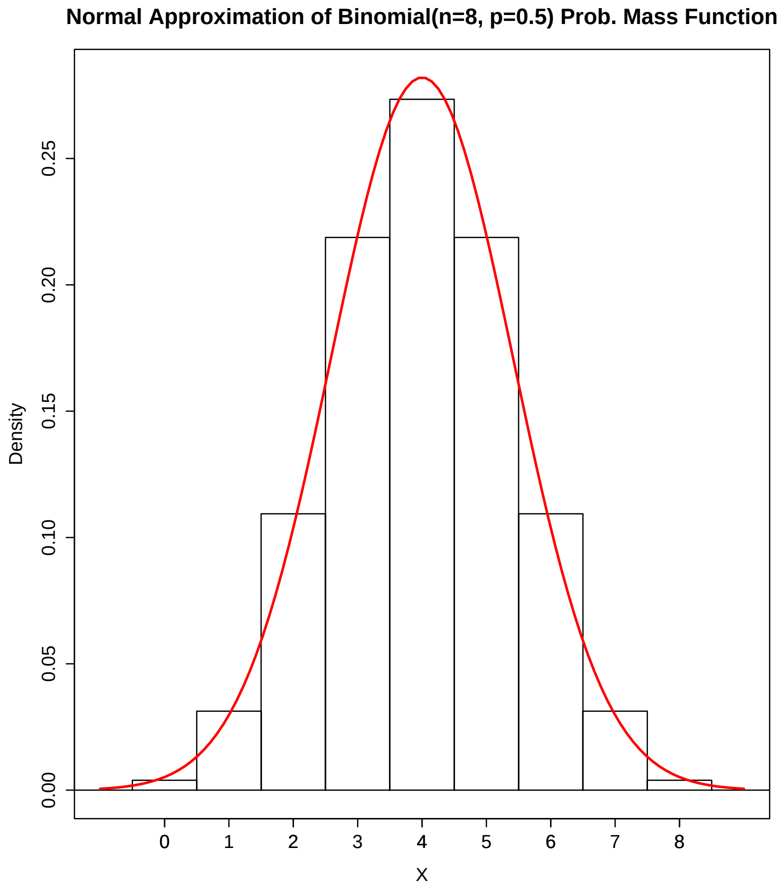
Solution: Use an asymptotic approximation approach (e.g. normal distribution)

Asymptotic Approximation for Binomial

X is binomial(n, π)

Prob. Mass Function $= \text{bin}(x, n, \pi) = \binom{n}{x}\pi^x(1-\pi)^{n-x}$





Confidence Interval for π: Asymptotic vs. Exact

Asymptotic CI: $[L_1(X), U_1(X)]$, Wald

Given $X = x$, a popular asymptotic 95% CI for π is ($\widehat{\pi} = x/n$):

$$\widehat{\pi} \pm 1.96\sqrt{\frac{\widehat{\pi}(1-\widehat{\pi})}{n}}$$

Exact CI: $[L_2(X), U_2(X)]$, Clopper & Pearson

Given $X = x$, a well known exact 95% CI for π has endpoints with values of π that satisfy:

$$\sum_{k=x}^{n} \binom{n}{k} \pi^k (1-\pi)^{n-k} = 0.025$$

and

$$\sum_{k=0}^{x} \binom{n}{k} \pi^k (1-\pi)^{n-k} = 0.025.$$

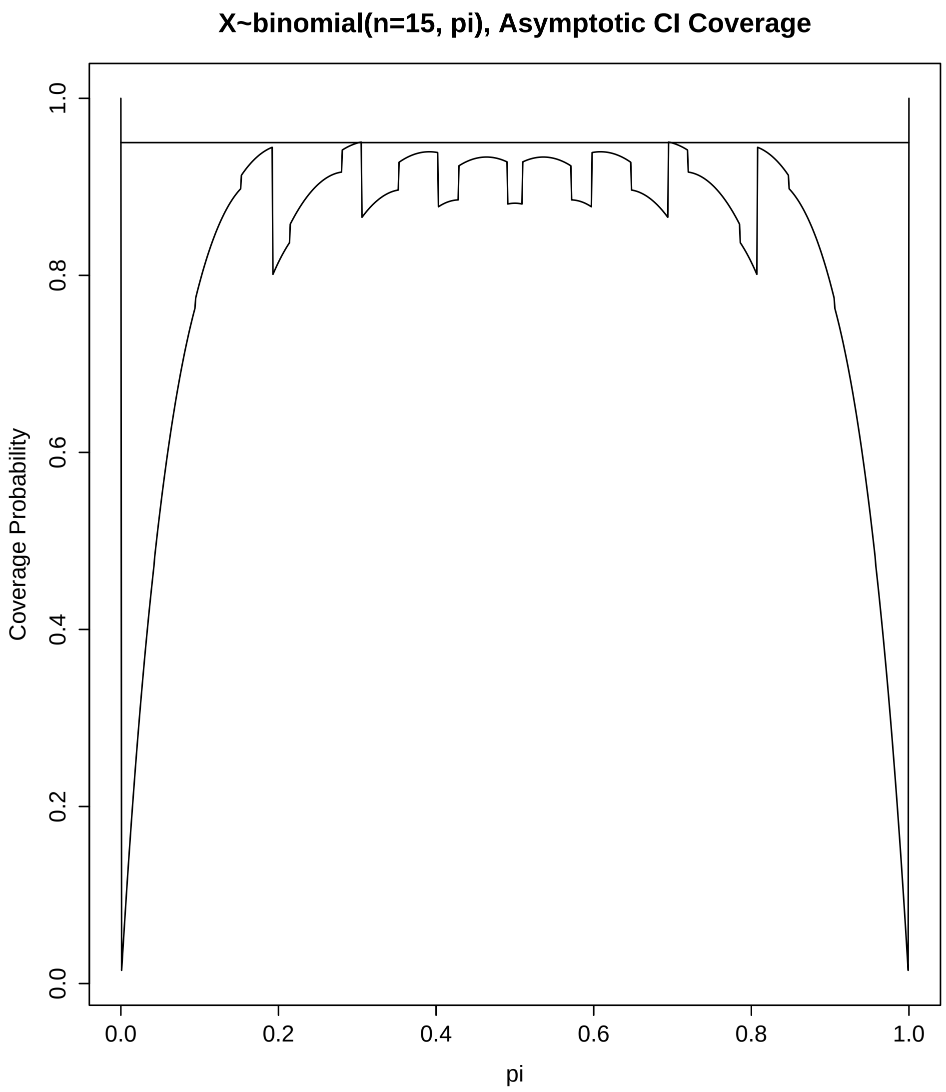
Coverage Probability

- X is binomial(n, π)
- For $x \in \{0, 1, 2, \ldots, n\}$, generate CI for each sample point:

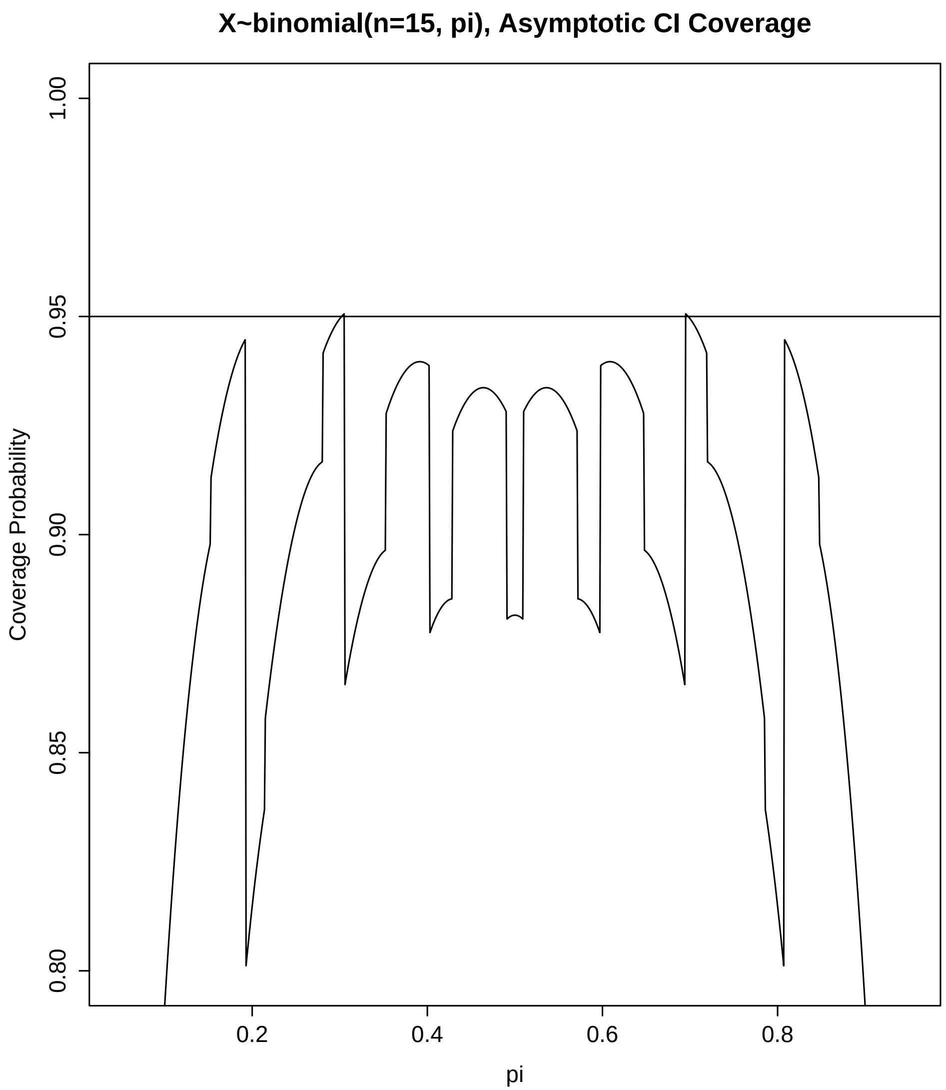
$X = x$	0	1	2	$\cdots$	n
$CI(x)$	(l_0, u_0)	(l_1, u_1)	(l_2, u_2)	$\cdots$	(l_n, u_n)

- For each $\pi \in [0, 1]$, set $\boldsymbol{A} = \{x : \pi \in CI(x)\}$
- coverage probability $= \sum_{x \in \boldsymbol{A}} \text{bin}(x, n, \pi)$

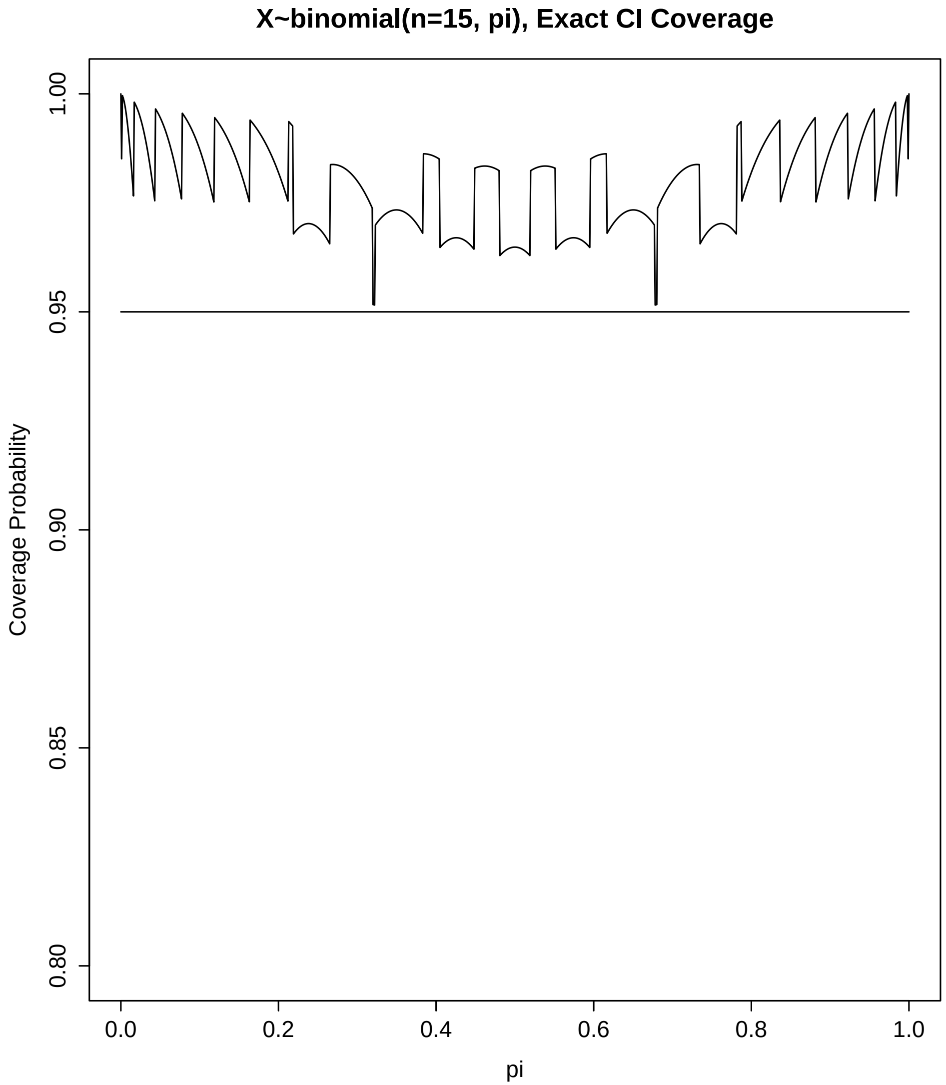
Asymptotic CI Coverage Probability



Asymptotic CI Coverage Probability



Exact CI Coverage Probability



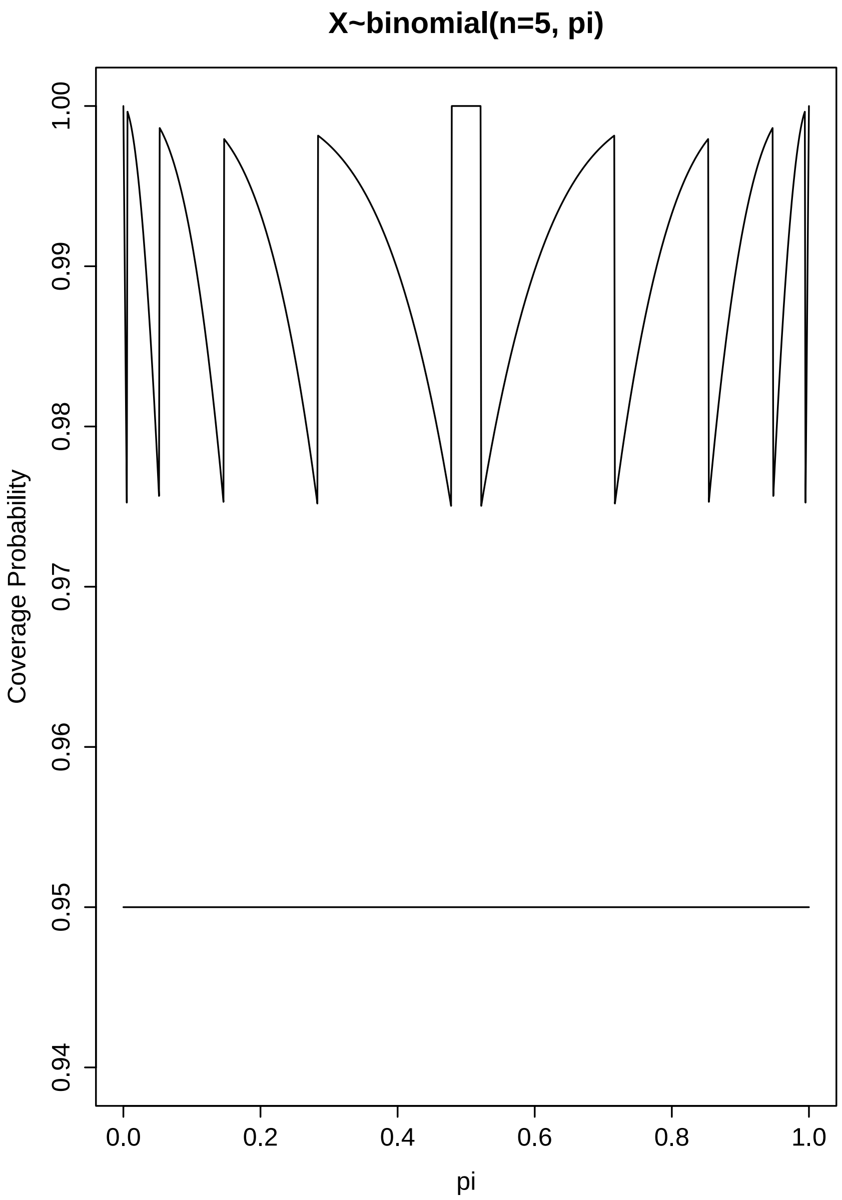
Exact Coverage Probability Example

$X \sim \text{bin}(n = 5, \pi)$

$\text{bin}(x, n, \pi) = \binom{n}{x} \pi^x (1 - \pi)^{n-x}$

$\boldsymbol{A} = \{x : \pi \in CI(x)\}$

$\text{CP} = \sum_{x \in \boldsymbol{A}} \text{bin}(x, n, \pi)$



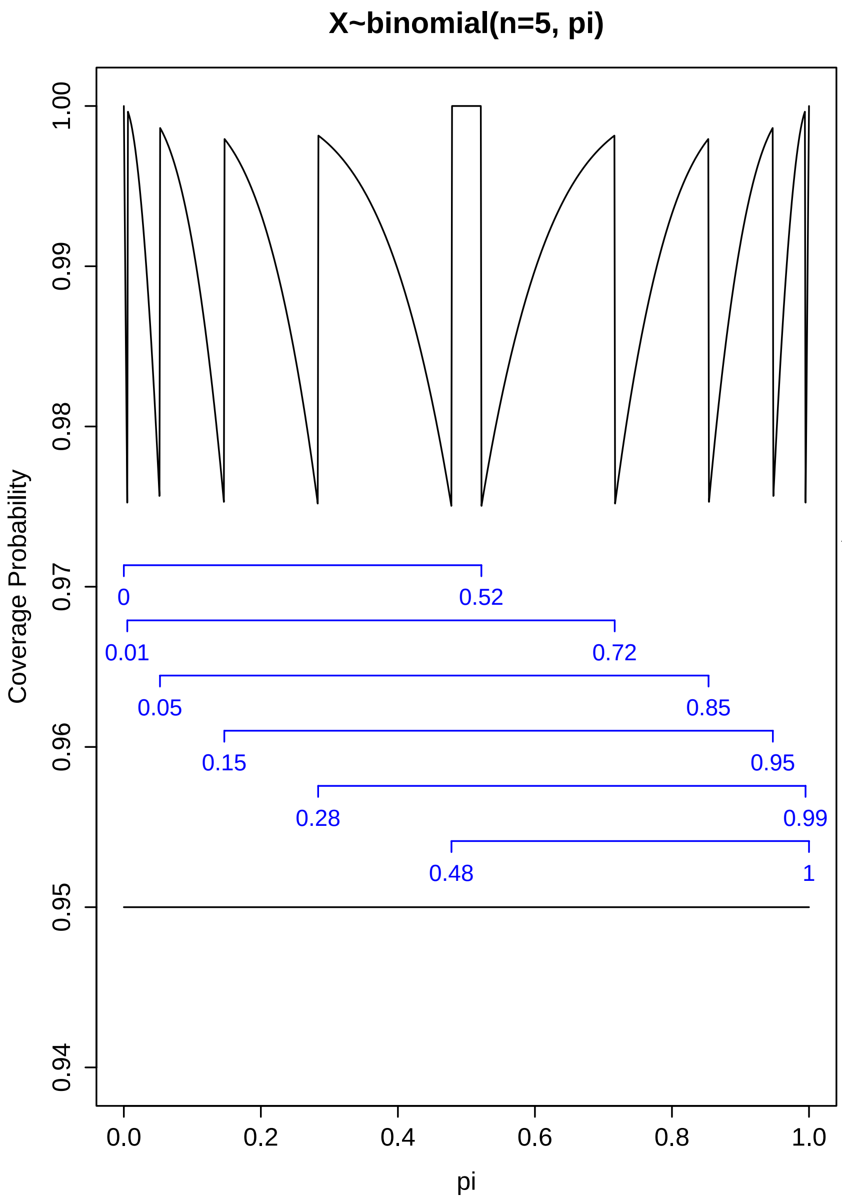
Exact Coverage Probability Example

$$X \sim \text{bin}(n = 5, \pi)$$

$$\text{bin}(x, n, \pi) = \binom{n}{x}\pi^{x}(1-\pi)^{n-x}$$

$$\boldsymbol{A} = \{x : \pi \in CI(x)\}$$

$$\text{CP} = \sum_{x \in \boldsymbol{A}} \text{bin}(x, n, \pi)$$



X	Conf. Int.
0	(0.00, 0.52)
1	(0.01, 0.72)
2	(0.05, 0.85)
3	(0.15, 0.95)
4	(0.28, 0.99)
5	(0.48, 1.00)

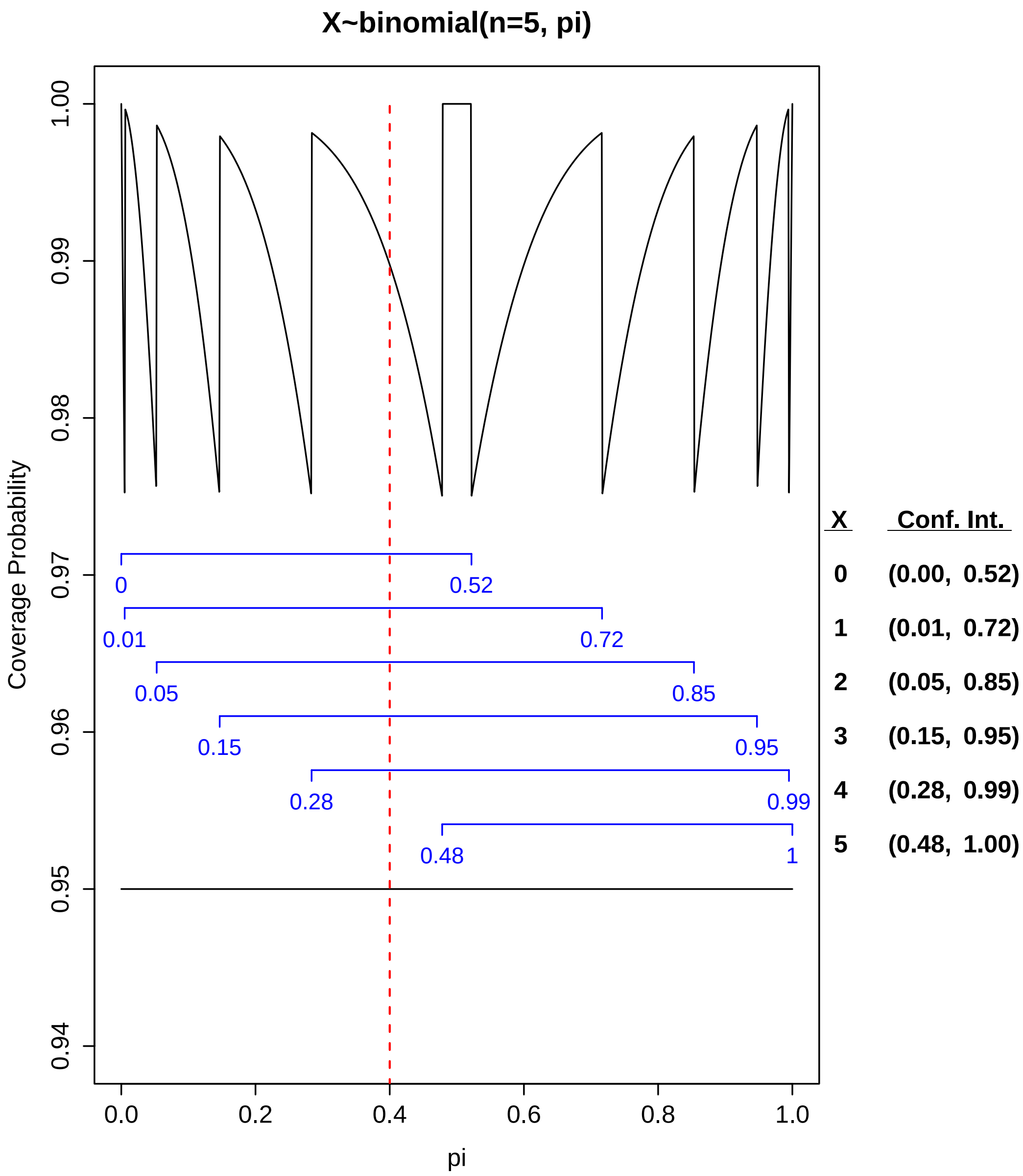
Exact Coverage Probability Example

$$X \sim \text{bin}(n = 5, \pi)$$

$$\text{bin}(x, n, \pi) = \binom{n}{x} \pi^{x} (1 - \pi)^{n-x}$$

$$\boldsymbol{A} = \{x : \pi \in CI(x)\}$$

$$\text{CP} = \sum_{x \in \boldsymbol{A}} \text{bin}(x, n, \pi)$$



Asymptotic versus Exact Methods

Asymptotic:

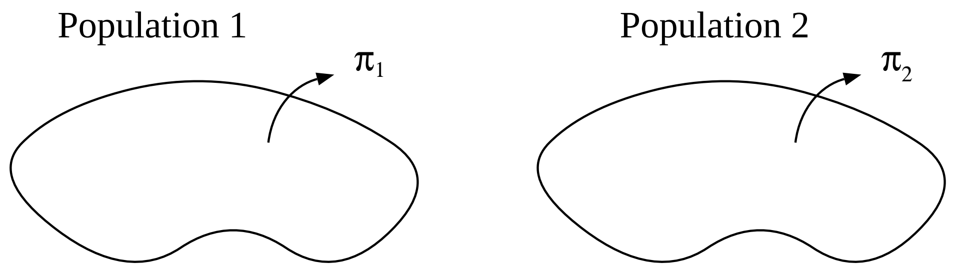
- Traditional method based on asymptotic theory
- Benefit: Easy to implement, maintains nominal level for large sample size(s)
- Drawback: For small to moderate sample sizes, as found in many clinical trials, asymptotic test performs very poorly

Exact:

- Exact testing is based on the true underlying distribution(s)
- Benefit: No approximations used, guaranteed to maintain nominal level regardless of sample size(s)
- Drawback: Difficult to implement, time consuming, computationally intensive

Clinical Trial Background

- Clinical trial to compare a standard and new treatment (i.e. New Trt = Drug 1, Standard Trt = Drug 2)



- π_1 and π_2 represent true response rates of the new and standard treatments, respectively
- Assume a higher response rate indicates stronger efficacy of the drug
- Comparison of the two parameters π_1 and π_2 through their difference, $\pi_1 - \pi_2 = \delta$

2×2 Contingency Table

Response	Population 1	Population 2
Success	X_1	X_2
Failure	n_1 - X_1	n_2 - X_2
Total	n_1	n_2

- 2 independent response variables (X_1, X_2) distributed as binomial $(n_1,\ \pi_1)$ and binomial $(n_2,\ \pi_2)$, respectively
- Let x_1 and x_2 be observed values from the data which will be denoted as $\boldsymbol{x}$
- Let the sample space of (X_1, X_2) be denoted by $\Omega_{\mathbf{x}} = \{0, 1, \ldots, n_1\} \times \{0, 1, \ldots, n_2\}$.

Standard vs. Proposed Exact CI Methods

Goal: Construct a confidence interval for $\pi_1 - \pi_2$

Problem with existing method:

- The Standard (Unconditional) Exact Method
  - computationally intensive and time consuming algorithm
  - potentially leads to conservative outcomes

Research on alternative approach:

- The Proposed (Unconditional) Exact Method
  - computationally less intensive method via modified algorithm
  - less conservative outcomes

Comparison of Methods: Confidence Intervals

- Consider sample size ratios $(n_1 : n_2) \in \{1:1,\ 2:1,\ 3:1\}$ where $(n_1 + n_2) \in \{20,\ 60,\ 100\}$
- Confidence interval comparisons
  - Coverage probability
  - Expected length

Comparison of Methods: Coverage Probability

Method for fixed π_2:

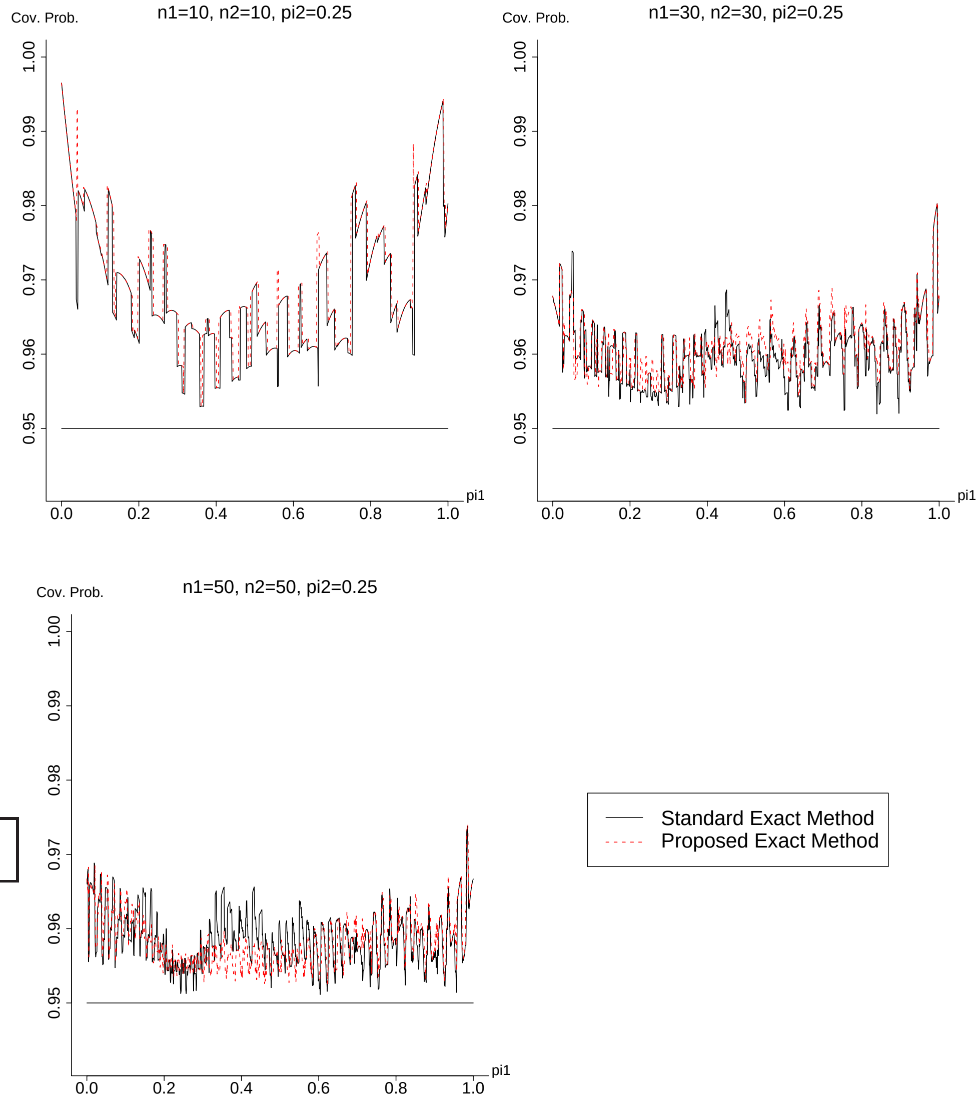
- Choose $\pi_2 \in [0, 1]$
- For $\pi_1 \in [0, 1]$, define $\delta = \pi_1 - \pi_2$
- Set $\boldsymbol{A} = \{(x_1, x_2) : \delta \in CI(x_1, x_2)\}$
- coverage probability $= \sum\sum_{(x_1,x_2)\in\boldsymbol{A}} \text{bin}(x_1, n_1, \pi_1)\text{bin}(x_2, n_2, \pi_2)$

Coverage Prob. Plots for $n_1 : n_2 = 1 : 1, 95\%\ CI, \pi_2 = 25\%$

(10,10)

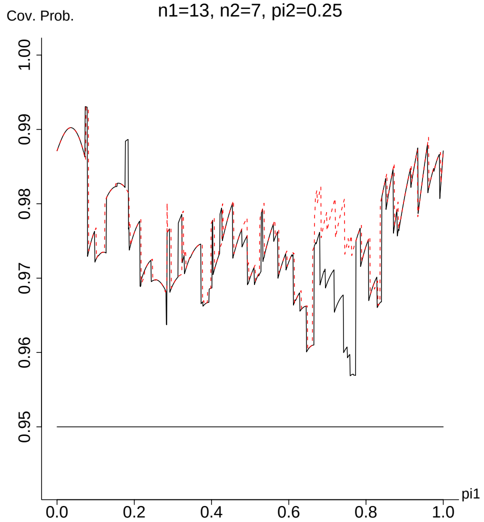
(30,30)

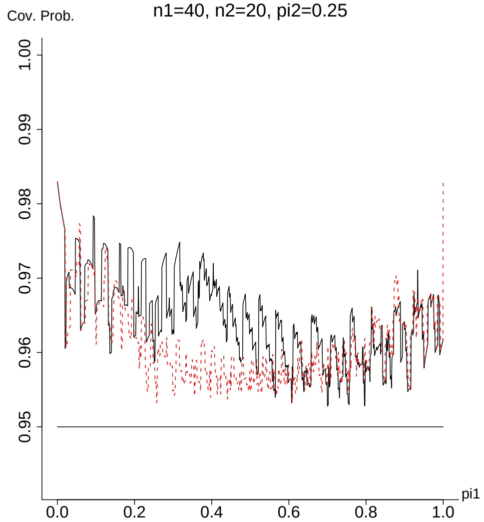
(50,50)



Coverage Prob. Plots for $n_1 : n_2 = 2 : 1, 95\%\ CI, \pi_2 = 25\%$

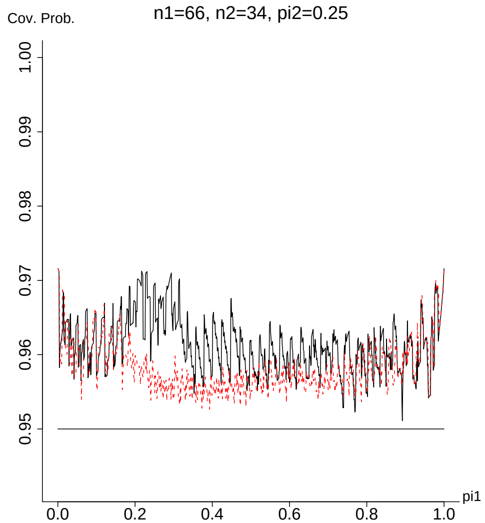
(13,7)





(40,20)

(66,34)



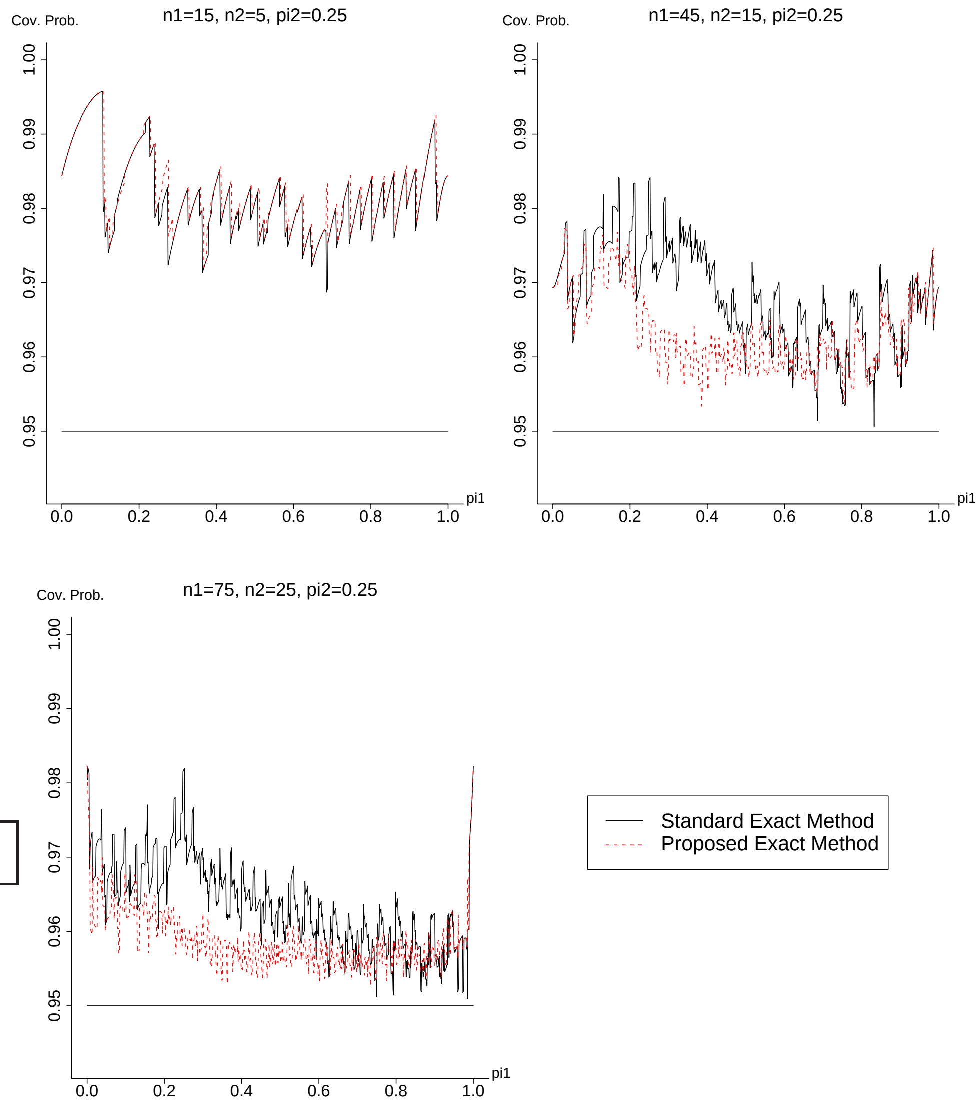
Standard Exact Method
Proposed Exact Method

Coverage Prob. Plots for $n_1 : n_2 = 3 : 1, 95\%\ CI, \pi_2 = 25\%$

(15,5)

(45,15)

(75,25)



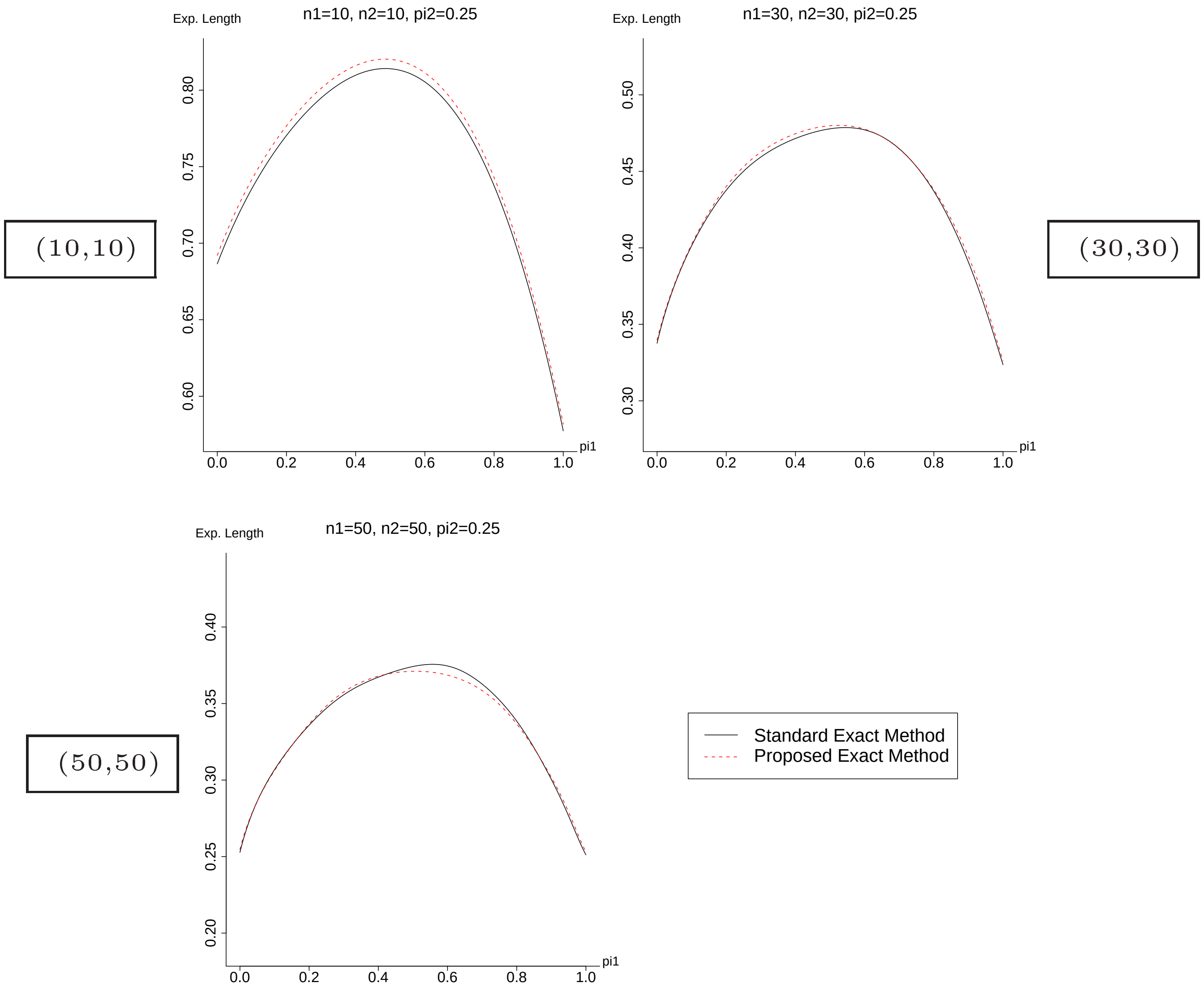
Comparison of Methods: Expected Length

Method for fixed π_2:

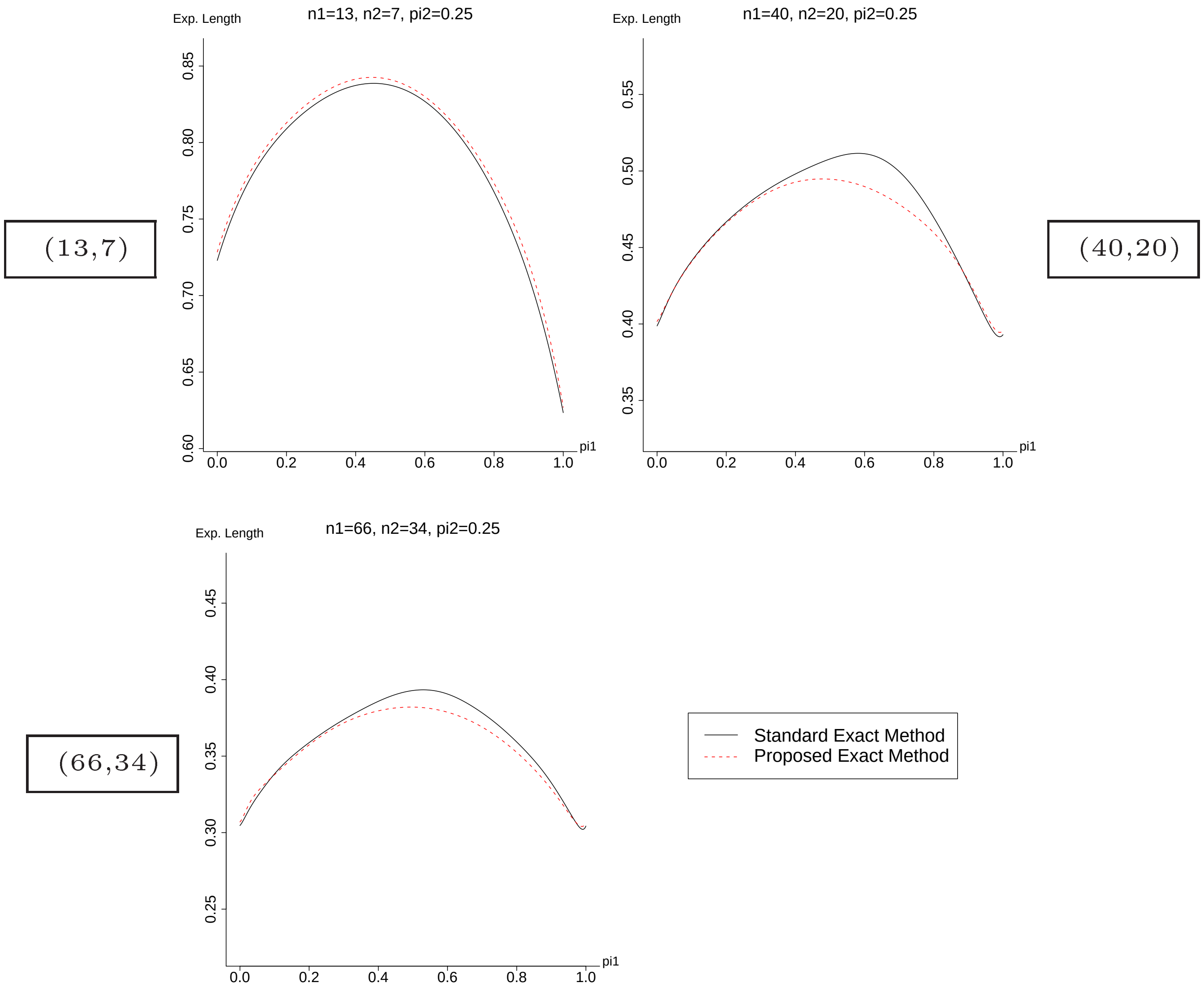
- Choose $\pi_2 \in [0, 1]$
- $\pi_1 \in [0, 1]$
- expected length =

$$\sum\sum_{(x_1,x_2)\in\Omega_{\boldsymbol{X}}} \|CI(x_1, x_2)\| \text{bin}(x_1, n_1, \pi_1) \text{bin}(x_2, n_2, \pi_2)$$

Expected Length Plots for $n_1 : n_2 = 1:1, 95\%\ CI, \pi_2 = 25\%$

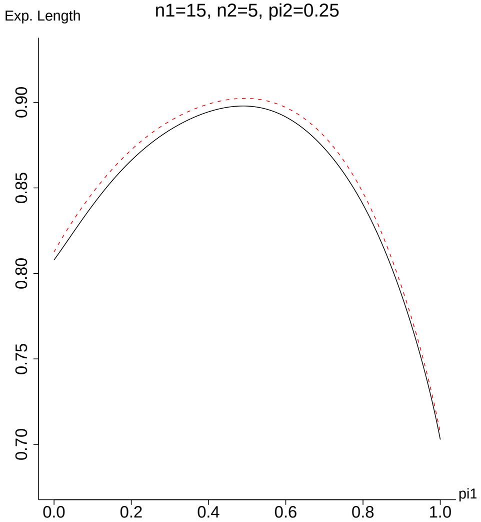


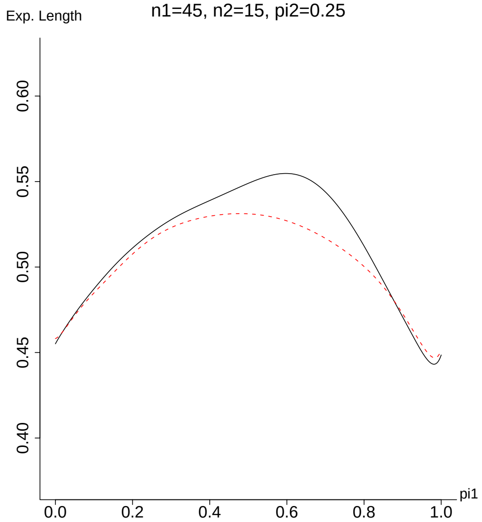
Expected Length Plots for $n_1 : n_2 = 2:1, 95\%\ CI, \pi_2 = 25\%$



Expected Length Plots for $n_1 : n_2 = 3 : 1, 95\%\ CI, \pi_2 = 25\%$

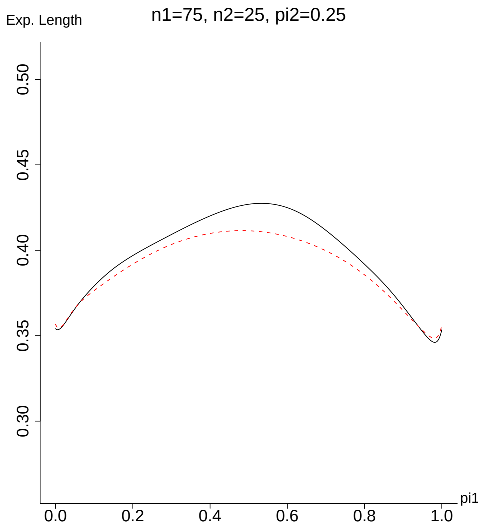
(15,5)





(45,15)

(75,25)



Standard Exact Method
Proposed Exact Method

Summary

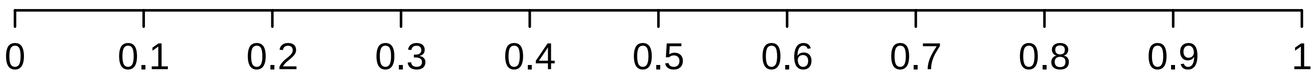
- Confidence intervals for π:
  - Asymptotic methods (normality) are not appropriate for small to moderate sample sizes
  - Exact methods, though conservative, offer improved performances
- Confidence intervals for $\pi_1 - \pi_2$:
  - benefits of using the proposed method are most apparent when n_1 and n_2 are unbalanced
  - gains outweigh losses

Example: Illicit Drug Usage

Survey Question to Undergraduates

"Have you ever experimented with an illegal drug/substance for non-medicinal purposes?"

What is the true value across surveys (n=750)?



Example: Illicit Drug Usage

Survey Question to Undergraduates

"Have you ever experimented with an illegal drug/substance for non-medicinal purposes?"

What is the true value across surveys (n=750)?