431 Class 16

Thomas E. Love, Ph.D.

2022-11-01

Today's Agenda

- Confidence Intervals for a Population Proportion
  - Five Methods to Accomplish This Task
- Comparing Two Proportions using Independent Samples
  - Standard Epidemiological Format
  - Working with 2x2 Tables

Today's Packages

```
library(Epi) # for twoby2() function
library(mosaic) # not usually something we load
library(readxl) # to import an Excel file
library(pwr) # specialized power functions
library(broom)
library(janitor)
library(kableExtra)
library(tidyverse)

source("c16/data/Love-boost.R") # for twobytwo() function

theme_set(theme_bw())
```

Confidence Intervals for a Population Proportion

Moving on from Means to Proportions

We've focused on creating statistical inferences about a population mean when we have a quantitative outcome. Now, we'll tackle a **categorical** outcome.

We'll estimate a confidence interval around an unknown population proportion, or rate, symbolized with \(\pi\), on the basis of a random sample of *n* observations from the population of interest.

The sample proportion is called \(\hat{p}\), which is sometimes, unfortunately, symbolized as \(p\).

- This \(\hat{p}\) is the sample proportion - not a *p* value.

An Example from *JAMA Pediatrics*

Original Investigation

ONLINE FIRST

September 27, 2021

Effect of Whole-Genome Sequencing on the Clinical Management of Acutely Ill Infants With Suspected Genetic Disease

A Randomized Clinical Trial

The NICUSeq Study Group

Article Information

JAMA Pediatr. Published online September 27, 2021. doi:10.1001/jamapediatrics.2021.3496

Outcome: Change in Management (COM)

The study involved infants ages 0-120 days admitted to an intensive care unit with a suspected genetic disease.

- For our first example, we focus on a sample of 326 subjects who received whole-genome sequencing testing at some point in the first 60 days after they were enrolled in the study.
- The outcome of interest is whether or not the subject received a change of management (COM) 60 days after their enrollment.

What can we conclude about the true proportion in the population of infants who meet our study criteria who would have a COM?

Loading the Data

```
nicu <- read_excel("c16/data/nicu_seq.xls") |>
  clean_names()

nicu
```

```
# A tibble: 326 × 3
   subject interv       outcome
     <dbl> <chr>        <chr>
 1       1 Early (15)   No_COM
 2       2 Early (15)   COM
 3       3 Delayed (60) COM
 4       4 Delayed (60) No_COM
 5       5 Early (15)   COM
 6       6 Early (15)   No_COM
 7       7 Delayed (60) No_COM
 8       8 Delayed (60) No_COM
 9       9 Delayed (60) No_COM
10      10 Early (15)   No_COM
# … with 316 more rows
```

Our `outcome` data

```
1 nicu |> tabyl(outcome) |> adorn_totals() |> adorn_pct_formatting()
```

```
outcome   n percent
    COM  51   15.6%
 No_COM 275   84.4%
  Total 326  100.0%
```

Our first inferential goal will be to produce a **confidence interval for the true (population) proportion** receiving a COM, across all infants who meet study criteria, based on this sample of 326 infants.

A Confidence Interval for a Proportion

A 100(1-\(\alpha\))% confidence interval for the population proportion \(\pi\) can be created by using:

- the standard normal distribution,
- the sample proportion, \(\hat{p}\), and
- the standard error of a sample proportion, which is defined as the square root of \(\hat{p}\) multiplied by \((1 - \hat{p})\) divided by the sample size, \(n\).

A Confidence Interval for a Proportion

Specifically, that confidence interval estimate is $\hat{p} \pm Z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

where $Z_{\alpha/2}$ = the value from a standard Normal distribution cutting off the top $\alpha/2$ of the distribution, obtained in R by substituting the desired $\alpha/2$ value into: `qnorm(alpha/2, lower.tail=FALSE)`.

- *Note*: This interval is reasonably accurate so long as $n \hat{p}$ and $n(1- \hat{p})$ are each at least 5.

Estimating π in the NICU data

- We'll build a 95% confidence interval for the true population proportion, so α = 0.05
- We have n = 326 subjects
- Sample proportion is $\hat{p}$ = .156, since 51/326 = 0.156.

The standard error of that sample proportion will be

$$\textrm{SE}(\hat{p}) = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{0.156(1-0.156)}{326}} = 0.020$$

Confidence Interval for \(\pi\) = Pr(COM)

Our 95% confidence interval for the true population proportion, \(\pi\), of infants who have a COM within 60 days is:

\[\hat{p} \pm Z_{.025} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = 0.156 \pm 1.96 (0.020) = 0.156 \pm 0.039 \]

or (0.117, 0.195).

To verify that \(Z_{0.025} = 1.96\)...

```
1 qnorm(0.025, lower.tail=FALSE)
```

```
[1] 1.959964
```

Likely Accuracy of this CI?

Since $n \hat{p} = (326)(0.156) = 51$ and $n (1 - \hat{p}) = (326)(1-0.156) = 275$ are substantially greater than 5, the CI should be reasonably accurate.

What can we conclude from this analysis?

- Point estimate of the proportion with COM is 0.156
- 95% CI for population proportion is (0.117, 0.195)

What is the "margin of error" in this confidence interval?

95% CI for population proportion is (0.117, 0.195)

- The entire confidence interval has width 0.078 (or 7.8 percentage points.)
- The margin of error (or half-width) is 0.039, or 3.9 percentage points.

Happily, that's our last "by hand" calculation.

Using R to estimate a CI for a Proportion

I'll discuss five procedures for estimating a CI for a population proportion. Each can be obtained using the `binom.test()` function from within the `mosaic` package.

For a 95% CI, we use:

```
mosaic::binom.test(x = 51, n = 326, p = 0.5, conf.level = 0.95, # defaults
                   ci.method = "XXX")
```

where the appropriate `ci.method` is obtained from the next slide's table.

Choosing a `ci.method`

Approach	`ci.method` to be used
Wald	`"Wald"`
Clopper-Pearson	`"Clopper-Pearson"` or `"binom.test"`
Score	`"Score"` or `"prop.test"`
Agresti-Coull	`"agresti-coull"`
Plus4	`"plus4"`

Approaches 1-2 in `binom.test()`

Each of these five approaches involves an approximation.

1. **Wald** is the "basic biostatistics" method we just calculated, where we estimate the standard error using the sample proportion and then use the Normal distribution to set the endpoints. The Wald interval is always symmetric, and can dip below 0 or above 1.
2. **Clopper-Pearson** is used by `stats::binom.test()` in R as well. It guarantees coverage at least as large as the nominal coverage rate, but may produce wider intervals than the other methods.

Approaches 3-5 in `binom.test()`

3. **Score** is used by `stats::prop.test()` and creates CIs by inverting p-values from score tests. It can be applied with a continuity correction (use ci.method = `"prop.test"`) or without.
4. **Agresti-Coull** is the Wald method after adding Z successes and Z failures to the data, where Z is the appropriate quantile for a standard Normal distribution (1.96 for a 95% CI)
5. **Plus4** is the Wald method after adding 2 successes and 2 failures (so 4 observations) to the data.

Formulas for these 5 methods?

See Wikipedia's entry: Binomial proportion confidence interval and Chapter 25 of our Course Notes.

Method 1: The Wald Procedure

```
1 method1 <- binom.test(x = 51, n = 326, conf.level = 0.95, ci.method = "Wald
2 method1
```

```
    Exact binomial test (Wald CI)

data:  51 out of 326
number of successes = 51, number of trials = 326, p-value < 2.2e-16
alternative hypothesis: true probability of success is not equal to 0.5
95 percent confidence interval:
 0.1170075 0.1958759
sample estimates:
probability of success
             0.1564417
```

Tidying up a `binom.test` result

```
1  tidy1 <- tidy(method1)
2
3  tidy1 |> select(estimate, conf.low, conf.high, statistic, parameter) |>
4    kbl(digits = 4) |> kable_classic(font_size = 28, full_width = F)
```

estimate	conf.low	conf.high	statistic	parameter
0.1564	0.117	0.1959	51	326

Method 2: The Clopper-Pearson Procedure

```
method2 <- binom.test(x = 51, n = 326, conf.level = 0.95,
                      ci.method = "Clopper-Pearson")

tidy2 <- tidy(method2)
tidy2 |> select(estimate, conf.low, conf.high, statistic, parameter) |>
  kbl(digits = 4) |> kable_classic(font_size = 28, full_width = F)
```

estimate	conf.low	conf.high	statistic	parameter
0.1564	0.1188	0.2005	51	326

Method 3: The Score Procedure

```
method3 <- binom.test(x = 51, n = 326, conf.level = 0.95,
                      ci.method = "Score")

tidy3 <- tidy(method3)
tidy3 |> select(estimate, conf.low, conf.high, statistic, parameter) |>
  kbl(digits = 4) |> kable_classic(font_size = 28, full_width = F)
```

estimate	conf.low	conf.high	statistic	parameter
0.1564	0.121	0.1999	51	326

Method 4: The Agresti-Coull Procedure

```
method4 <- binom.test(x = 51, n = 326, conf.level = 0.95,
                      ci.method = "agresti-coull")

tidy4 <- tidy(method4)
tidy4 |> select(estimate, conf.low, conf.high, statistic, parameter) |>
  kbl(digits = 4) |> kable_classic(font_size = 28, full_width = F)
```

estimate	conf.low	conf.high	statistic	parameter
0.1564	0.1208	0.2001	51	326

Method 5: The Plus4 Procedure

```
method5 <- binom.test(x = 51, n = 326, conf.level = 0.95,
                      ci.method = "plus4")

tidy5 <- tidy(method5)
tidy5 |> select(estimate, conf.low, conf.high, statistic, parameter) |>
  kbl(digits = 4) |> kable_classic(font_size = 28, full_width = F)
```

estimate	conf.low	conf.high	statistic	parameter
0.1564	0.121	0.2002	51	326

Comparison of Methods

```
res1 <- tidy1 |> select(estimate, conf.low, conf.high)
res2 <- tidy2 |> select(estimate, conf.low, conf.high)
res3 <- tidy3 |> select(estimate, conf.low, conf.high)
res4 <- tidy4 |> select(estimate, conf.low, conf.high)
res5 <- tidy5 |> select(estimate, conf.low, conf.high)

res <- bind_rows(res1, res2, res3, res4, res5)
res <- res |> mutate(
  approach = c("Wald", "Clopper-Pearson", "Score",
               "Agresti-Coull", "Plus4"))
```

Results with too many decimal places

95% confidence intervals based on x = 51 successes in n = 326 trials.

estimate	conf.low	conf.high	approach
0.15644	0.11701	0.19588	Wald
0.15644	0.11875	0.20051	Clopper-Pearson
0.15644	0.12104	0.19985	Score
0.15644	0.12084	0.20005	Agresti-Coull
0.15644	0.12099	0.20022	Plus4

This is way more precision than we can really justify, but I just want you to see that the five results are all (slightly) different.

Results after some rounding

95% confidence intervals based on x = 51 successes in n = 326 trials.

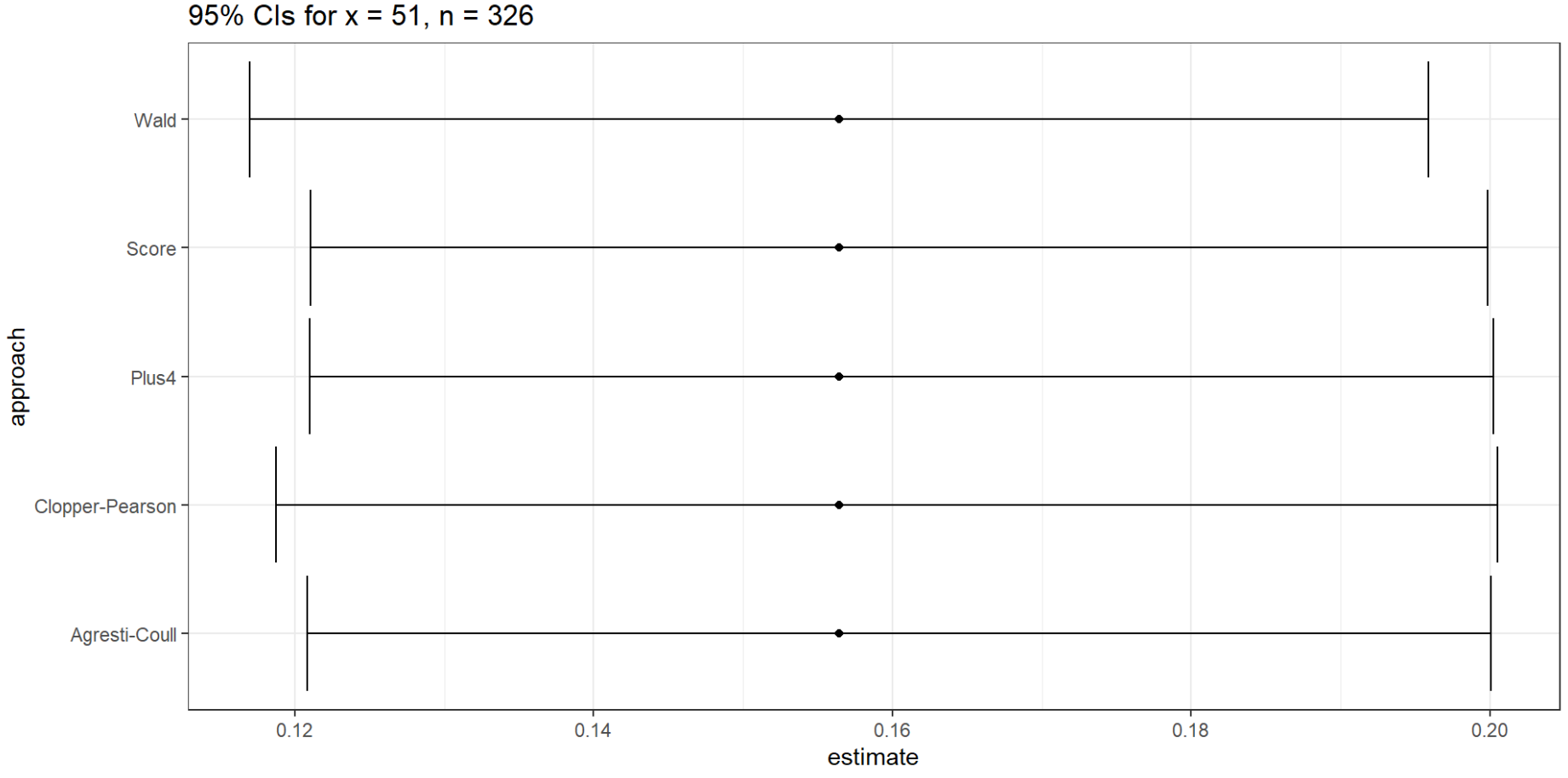
estimate	conf.low	conf.high	approach
0.156	0.117	0.196	Wald
0.156	0.119	0.201	Clopper-Pearson
0.156	0.121	0.200	Score
0.156	0.121	0.200	Agresti-Coull
0.156	0.121	0.200	Plus4

Here's a somewhat more plausible rounding approach.

- Is the distinction between methods important here?

Plotting the 95% CI Estimates

```
ggplot(res, aes(x = approach, y = estimate)) +
  geom_point() + geom_errorbar(aes(ymin = conf.low, ymax = conf.high)) +
  coord_flip() + labs(title = "95% CIs for x = 51, n = 326")
```



What if we ran 90% CIs Instead?

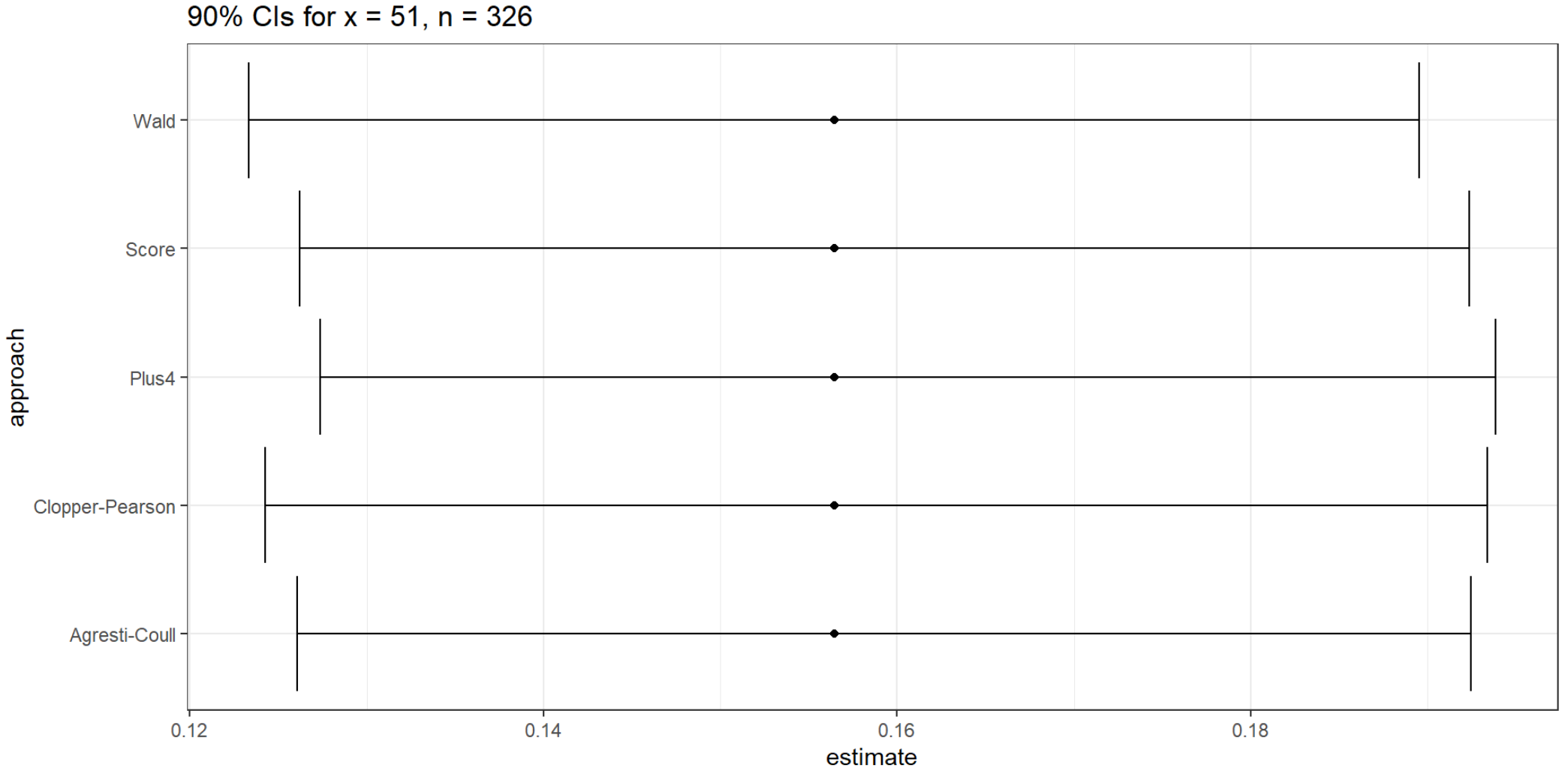
90% CIs based on x = 51 successes in n = 326 trials.

```
new1 <- tidy(binom.test(x = 51, n = 326, conf.level = 0.9, ci.method = "Wal
  select(estimate, conf.low, conf.high) |> mutate(approach = "Wald")
new2 <- tidy(binom.test(x = 51, n = 326, conf.level = 0.9, ci.method = "Clo
  select(estimate, conf.low, conf.high) |> mutate(approach = "Clopper-Pears
new3 <- tidy(binom.test(x = 51, n = 326, conf.level = 0.9, ci.method = "Sco
  select(estimate, conf.low, conf.high) |> mutate(approach = "Score")
new4 <- tidy(binom.test(x = 51, n = 326, conf.level = 0.9, ci.method = "agr
  select(estimate, conf.low, conf.high) |> mutate(approach = "Agresti-Coull
new5 <- tidy(binom.test(x = 51, n = 326, conf.level = 0.9, ci.method = "plu
  select(estimate, conf.low, conf.high) |> mutate(approach = "Plus4")

newres <- bind_rows(new1, new2, new3, new4, new5)

newres |> kbl(digits = 3) |> kable_classic(font_size = 28, full_width = F)
```

What if we ran 90% CIs Instead?

estimate	conf.low	conf.high	approach
0.156	0.123	0.190	Wald
0.156	0.124	0.193	Clopper-Pearson
0.156	0.126	0.192	Score
0.156	0.126	0.192	Agresti-Coull
0.156	0.127	0.194	Plus4

Plotting the 90% CI Estimates



Estimating Rates More Accurately

Suppose you have some data involving n independent tries, with x successes. The most natural estimate of the "success rate" in the data is x / n. But, strangely enough, it turns out this isn't an entirely satisfying estimator.

Alan Agresti provides substantial motivation for $\frac{x + 2}{n + 4}$ as an alternative. See http://andrewgelman.com/2007/05/15 for instance. We'll call this a *Bayesian augmentation*.

When does this augmentation matter?

Estimates with and without the augmentation will be generally comparable, so long as…

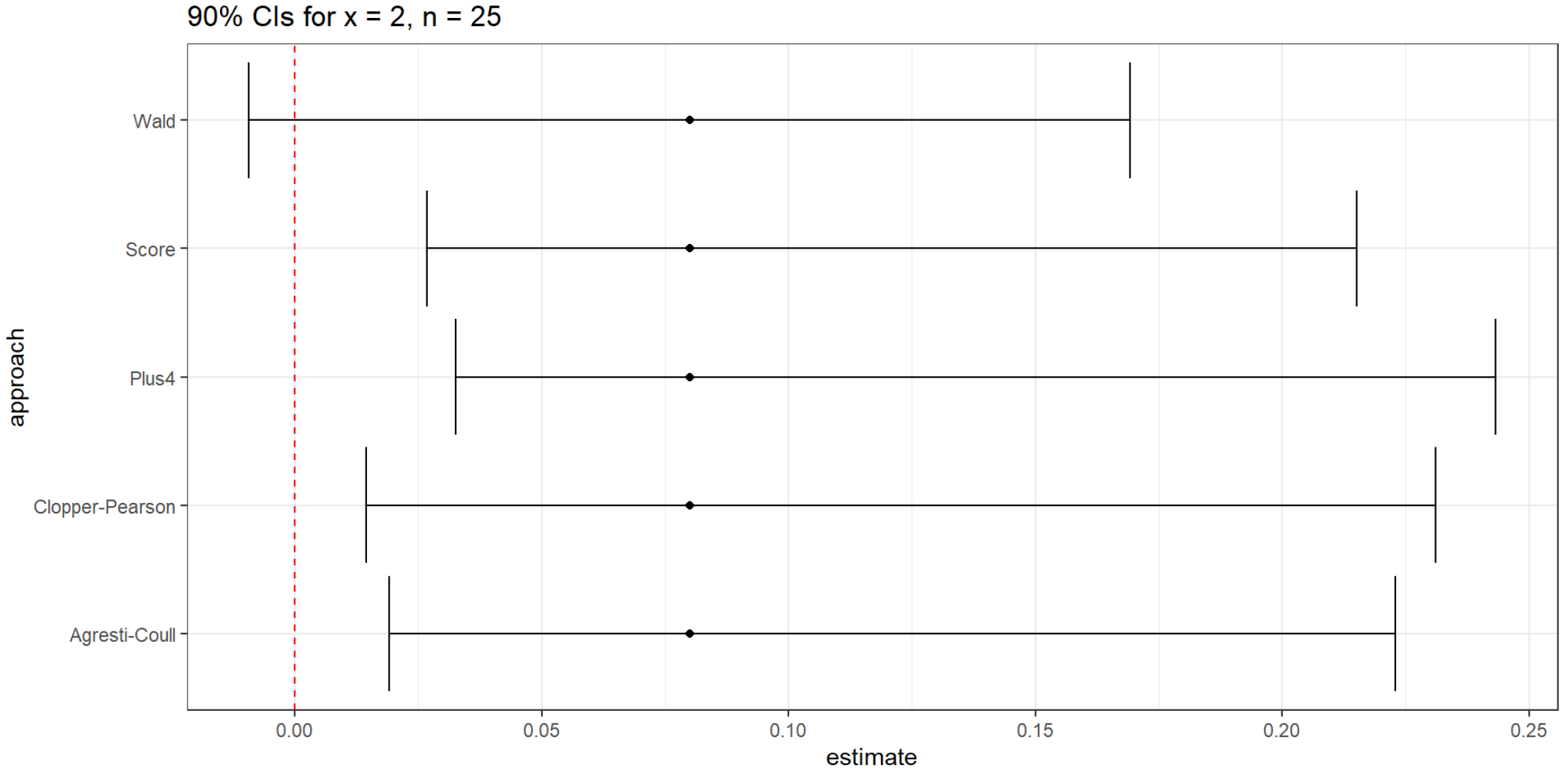
1. the sample size is more than, say, 30 subjects, and/or
2. the sample probability of the outcome is between 0.1 and 0.9

Observe 2 successes in 25 trials?

90% CIs based on x = 2 successes in n = 25 trials.

estimate	conf.low	conf.high	approach
0.08	-0.009	0.169	Wald
0.08	0.014	0.231	Clopper-Pearson
0.08	0.027	0.215	Score
0.08	0.019	0.223	Agresti-Coull
0.08	0.033	0.243	Plus4

90% CI Estimates for x = 2, n = 25



What if x = 0 or x = n?

The **Rule of Three** approach is often used.

- An approximate 95% CI for the proportion in a setting where x = 0 in n trials is $(0, \frac{3}{n})$
- An approximate 95% CI for the proportion where x = n in n trials is $(1 - \frac{3}{n}, 1)$

Comparing Population Proportions

Comparing Population Proportions

Suppose we compare population proportions π_1 and π_2, based on samples of sizes n_1 and n_2.

1. The individual observations in exposure group 1 are not linked/matched to individual observations in exposure group 2. (Independent Samples)
2. Each individual observation in exposure group 1 is linked or matched to a specific observation in exposure group 2. (Paired Samples)

The determination as to whether the study design creates paired or independent samples can be determined without summarizing the data. It's a function of the design, not the responses.

A Polling Example

- 200 adult Ohio residents agreed to participate in a poll both two months ago and again today. Each of the 200 people met the polling organization's standards for a "likely voter in the 2022 election". 100 of those polled were under the age of 50 and the rest were 50 or older.
- In between the two polls, a major news event occurred which was relevant to Candidate X.

We asked them the same question at both times: "Are you considering voting for Candidate X?" We are interested in understanding what the data tell us about:

1. Were people under age 50 more likely to be considering Candidate X than people ages 50 and higher?
2. Were people more likely to be considering Candidate X after the news event than before?

Which of these uses *independent* samples, and which *paired* samples?

Comparing Proportions using Independent Samples

Visual Abstract: NICU Sequencing Paper

JAMA Pediatrics

RCT: Effect of Whole-Genome Sequencing on Clinical Management of Acutely Ill Infants With Suspected Genetic Disease

POPULATION

201 Males, 153 Females

Infants admitted to an intensive care unit with a suspected genetic disease and aged between 0 and 120 d

Mean age, 15 d (IQR, 7-32 d)

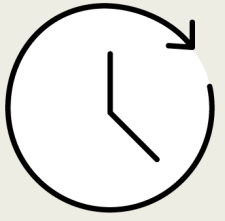
SETTINGS / LOCATIONS

5 US academic centers and affiliated children's hospitals

INTERVENTION

354 Patients

176 Early whole-genome sequencing testing
Whole-genome sequencing results returned 15 d after study enrollment



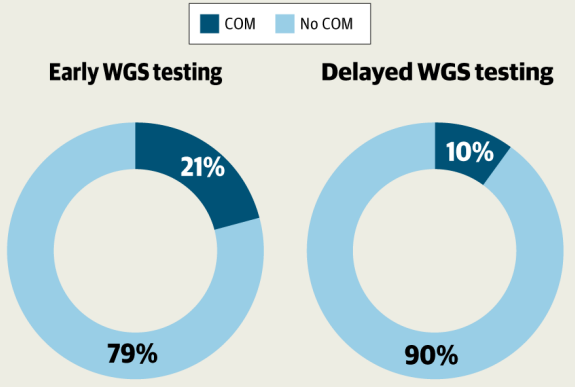
178 Delayed whole-genome sequencing testing
Whole-genome sequencing results returned 60 d after study enrollment

PRIMARY OUTCOME

Difference in the proportion of infants in the early and delayed groups who received a change of management (COM) 60 d after enrollment

FINDINGS

The proportion of infants who received COM was significantly higher among infants receiving early whole-genome sequencing testing compared with delayed testing



Proportion of infants with COM:

Early testing: 34 of 161 (21.1%)

Delayed testing: 17 of 165 (10.3%), $P < .009$

The NICUSeq Study Group. Effect of whole-genome sequencing on the clinical management of acutely ill infants with suspected genetic disease: a randomized clinical trial. *JAMA Pediatr.* Published online September 27, 2021. doi:10.1001/jamapediatrics.2021.3496

NICU Sequencing Example

Let's compare the proportion who have a COM between:

- Group 1: infants tested early (15 d after enrollment)
- Group 2: infants tested later (60 d after enrollment)

```
1 nicu |> count(interv, outcome)
```

```
# A tibble: 4 × 3
  interv       outcome     n
  <chr>        <chr>   <int>
1 Delayed (60) COM        17
2 Delayed (60) No_COM    148
3 Early (15)   COM        34
4 Early (15)   No_COM    127
```

- How might we rearrange this information? Exposure? Outcome?

The Table We'd Like To Get To

Let's compare the proportion who have a COM between:

- Group 1: infants tested early (at 15 d)
- Group 2: infants tested later (delayed to 60 d)

Standard Epidemiological Format

- rows are the exposure
- columns are the outcome

What do we want in our setting?

Our Goal: Standard Epidemiological Format

- exposure is *intervention* (15 or 60 days)
- columns are *outcome* category (COM, No COM)

	COM	No COM
Early (15 d)	a	b
Delayed (60 d)	c	d

Our 2 x 2 Table

```
1 nicu |> tabyl(interv, outcome)
```

```
       interv COM No_COM
Delayed (60)  17    148
  Early (15)  34    127
```

- Is this in standard epidemiological format, with the rows indicating the exposure, and the columns indicating the outcome, and the correct count in the top left cell?

Switching the Rows

We want Early (15) to come first, before Delayed (60):

```
1 nicu <- nicu |> mutate(interv = fct_relevel(interv, "Early (15)"))
2
3 nicu |> tabyl(interv, outcome)
```

```
      interv COM No_COM
  Early (15)  34    127
Delayed (60)  17    148
```

Adding Totals

```
1 nicu |> tabyl(interv, outcome) |>
2   adorn_totals(where = c("row", "col"))
```

```
      interv COM No_COM Total
  Early (15)  34    127   161
Delayed (60)  17    148   165
       Total  51    275   326
```

- How many subjects do we have in each exposure group?
- How many subjects fall into each outcome group?

Augmenting the Table

Can we augment the table to help us understand:

- What is the probability of achieving each of the two possible outcomes?
- How do the outcome probabilities differ by exposure group?

```
nicu |> tabyl(interv, outcome) |>
  adorn_totals(where = c("row", "col")) |>
  adorn_percentages(denom = "row") |>
  adorn_pct_formatting(digits = 1) |>
  adorn_ns(position = "front")
```

Augmenting the Table

interv	COM	No_COM	Total
Early (15)	34 (21.1%)	127 (78.9%)	161 (100.0%)
Delayed (60)	17 (10.3%)	148 (89.7%)	165 (100.0%)
Total	51 (15.6%)	275 (84.4%)	326 (100.0%)

Why am I using `denom = "row"` here?

> Among these subjects, compare the proportion of early (15 d) tested infants with COM to the proportion of late (60 d) tested infants with COM.

- What are the sample estimates for the two rates I am comparing?

2 x 2 Table: Comparing Probabilities

–	COM	No COM	*Total*
Early (15)	34	127	*161*
Delayed (60)	17	148	*165*
Total	*51*	*275*	*326*

- Pr(COM | Early) = 34/161 = 0.211
- Pr(COM | Delayed) = 17/165 = 0.103
- The ratio of those two probabilities (risks) is 0.211/0.103 = 2.05.

CI for the Relative Risk?

Can we build a confidence interval for the relative risk of COM now in the early tested infants as compared to the delayed tested infants?

- The difference in those risks is 0.211 - 0.103 = 0.108.

How about a confidence interval for the risk difference, too?

2 x 2 Table for NICU Example, Odds Ratio

–	COM	No COM	*Total*
Early (15)	34	127	*161*
Delayed (60)	17	148	*165*
Total	*51*	*275*	*326*

In a 2x2 table, odds ratio = cross-product ratio.

- Here, the cross-product estimate = \(\frac{34*148}{17*127}\) = 2.331.

Can we build a confidence interval for the population odds ratio for COM given “early” as compared to “delayed” testing?

Using `twoby2` from the `Epi` package

Once we have set up the factors for `interv` and `outcome` so that the table we produce is in standard epidemiological format, we can plug it into the `twoby2` function from the `Epi` package.

Using `twoby2` from the `Epi` package

```
1 twoby2(table(nicu$interv, nicu$outcome))
```

```
2 by 2 table analysis:
------------------------------------------------------
Outcome   : COM
Comparing : Early (15) vs. Delayed (60)

             COM No_COM    P(COM) 95% conf. interval
Early (15)    34    127    0.2112     0.155   0.2810
Delayed (60)  17    148    0.1030     0.065   0.1595

                                  95% conf. interval
            Relative Risk: 2.0497    1.1942   3.5180
        Sample Odds Ratio: 2.3307    1.2430   4.3701
Conditional MLE Odds Ratio: 2.3247    1.1972   4.6617
    Probability difference: 0.1081    0.0292   0.1871

                            Exact P-value: 0.0002
```

Using `twobytwo` from the `Love-boost.R` script

–	COM	No COM	*Total*
Early (15)	34	127	*161*
Delayed (60)	17	148	*165*
Total	*51*	*275*	*326*

Code we need is:

```
1 twobytwo(34, 127, 17, 148,  # note order of counts
2       "Early", "Delayed", # names of the rows
3       "COM", "NoCOM",  # names of the columns
4       conf.level = 0.99)  # default is 95% confidence
```

Using `twobytwo` from the `Love-boost.R` script

```
2 by 2 table analysis:
------------------------------------------------------
Outcome   : COM
Comparing : Early vs. Delayed

        COM NoCOM    P(COM) 99% conf. interval
Early    34   127    0.2112    0.1400   0.3057
Delayed  17   148    0.1030    0.0561   0.1818

                                  99% conf. interval
            Relative Risk: 2.0497    1.0078   4.1688
         Sample Odds Ratio: 2.3307    1.0202   5.3245
Conditional MLE Odds Ratio: 2.3247    0.9919   5.7786
    Probability difference: 0.1081    0.0037   0.2125

                Exact P-value: 0.0092
```

Another Way to Create The Table

Suppose we didn't have the data, just the visual abstract.

```
1 t1 <- matrix(c(34, 127, 17, 148), byrow = TRUE, nrow = 2)
2 rownames(t1) <- c("Early", "Delayed")
3 colnames(t1) <- c("COM", "No_COM")
4 addmargins(t1)
```

```
         COM No_COM Sum
Early     34    127 161
Delayed   17    148 165
Sum       51    275 326
```

Hypothesis Testing?

The hypotheses being compared can be thought of in several ways…

- \(H_0\): \(\pi_1 = \pi_2\), vs. \(H_A\): \(\pi_1 \neq \pi_2\).
- \(H_0\): Pr(COM | Early) = Pr(COM | Delayed) vs. \(H_A\): Pr(COM | Early) \(\neq\) Pr(COM | Delayed).
- \(H_0\): rows and columns of the table are *independent*, in that the probability of COM in each row is the same vs. \(H_A\): the rows and columns of the table are *associated*.

P values in twoby2 output?

```
Exact P-value: 0.0092
Asymptotic P-value: 0.0083
```

- The `Exact P-value` comes from Fisher's exact test, and is technically exact only if we treat the row and column totals as being fixed.
- The `Asymptotic P-value` comes from a Pearson \(\chi^2\) test.
- Neither approach is helpful if we don't have sufficient data to justify inference in the first place.

Bayesian Augmentation in a 2x2 Table?

Original command:

```
twobytwo(34, 127, 17, 148, "Early", "Delayed", "COM", "NoCOM",
         conf.level = 0.99)
```

Bayesian augmentation approach: Add two successes and add two failures in each row...

```
twobytwo(34+2, 127+2, 17+2, 148+2,  "Early", "Delayed", "COM", "NoCOM",
       conf.level = 0.99)
```

Bayesian Augmentation in a 2x2 Table?

```
2 by 2 table analysis:
----------------------------------------------------
Outcome   : COM
Comparing : Early vs. Delayed

        COM NoCOM    P(COM) 99% conf. interval
Early    36   129    0.2182    0.1466   0.3120
Delayed  19   150    0.1124    0.0634   0.1917

                                  99% conf. interval
            Relative Risk: 1.9407    0.9893   3.8071
         Sample Odds Ratio: 2.2032    0.9967   4.8701
Conditional MLE Odds Ratio: 2.1980    0.9691   5.2348
    Probability difference: 0.1058    0.0004   0.2105

                Exact P-value: 0.0118
```

Tuberculosis Prevalence in IV Drug Users

Suppose now that we are investigating factors affecting tuberculosis prevalence among intravenous drug users.

- Among 97 individuals who admit to sharing needles,
  - 24 (24.7%) had a positive tuberculin skin test result.
- Among 161 drug users who deny sharing needles,
  - 28 (17.4%) had a positive test result.

What does the 2x2 table look like?

Tuberculosis Prevalence In IV Drug Users

The 2x2 Table is…

	TB+	TB-
share	24	73
don't	28	133

- rows describe needle sharing, columns describe TB test result
- row 1 people who share needles: 24 TB+, and 97-24 = 73 TB-
- row 2 people who don't share: 28 TB+ and 161-28 = 133 TB-

`twobytwo` (with Bayesian Augmentation)

To start, we'll test the null hypothesis that the population proportions of intravenous drug users who have a positive tuberculin skin test result are identical for those who share needles and those who do not.

\[H_0: \pi_{share} = \pi_{donotshare} \\ H_A: \pi_{share} \neq \pi_{donotshare} \]

We'll use the Bayesian augmentation.

twobytwo (with Bayesian Augmentation)

```
1 twobytwo(24+2, 73+2, 28+2, 133+2,
2           "Sharing", "Not Sharing",
3           "TB test+", "TB test-")
```

```
2 by 2 table analysis:
------------------------------------------------------

Outcome    : TB test+
Comparing : Sharing vs. Not Sharing

             TB test+ TB test-    P(TB test+) 95% conf. interval
Sharing            26       75         0.2574    0.1816   0.3513
Not Sharing        30      135         0.1818    0.1301   0.2482

                                    95% conf. interval
            Relative Risk: 1.4158    0.8910   2.2498
        Sample Odds Ratio: 1.5600    0.8594   2.8318
Conditional MLE Odds Ratio: 1.5572    0.8189   2.9511
    Probability difference: 0.0756   -0.0244   0.1819

            Exact P-value: 0.1638
```

Session Information

```
1 sessionInfo()
```

```
R version 4.2.1 (2022-06-23 ucrt)
Platform: x86_64-w64-mingw32/x64 (64-bit)
Running under: Windows 10 x64 (build 22000)

Matrix products: default

locale:
[1] LC_COLLATE=English_United States.utf8
[2] LC_CTYPE=English_United States.utf8
[3] LC_MONETARY=English_United States.utf8
[4] LC_NUMERIC=C
[5] LC_TIME=English_United States.utf8

attached base packages:
[1] stats     graphics  grDevices utils     datasets  methods   base
```