

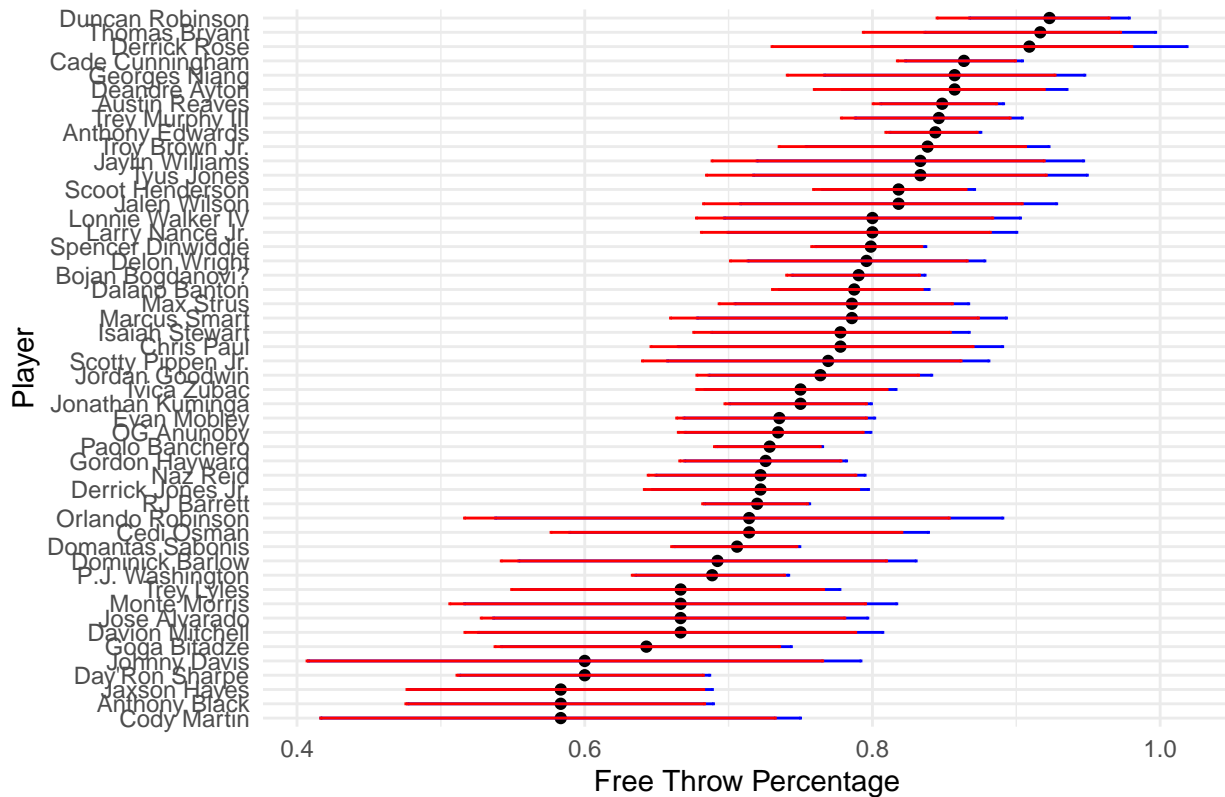
Lab 8

2025-06-11

R Markdown

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.4      v readr      2.1.4
## v forcats    1.0.0      v stringr    1.5.1
## v lubridate  1.9.2      v tibble     3.2.1
## v purrr      1.0.2      v tidyr      1.3.1
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

NBA Free Throw Percentage



for FT% above 50%, agresti drags the interval down, and vice versa for FT% below 50%

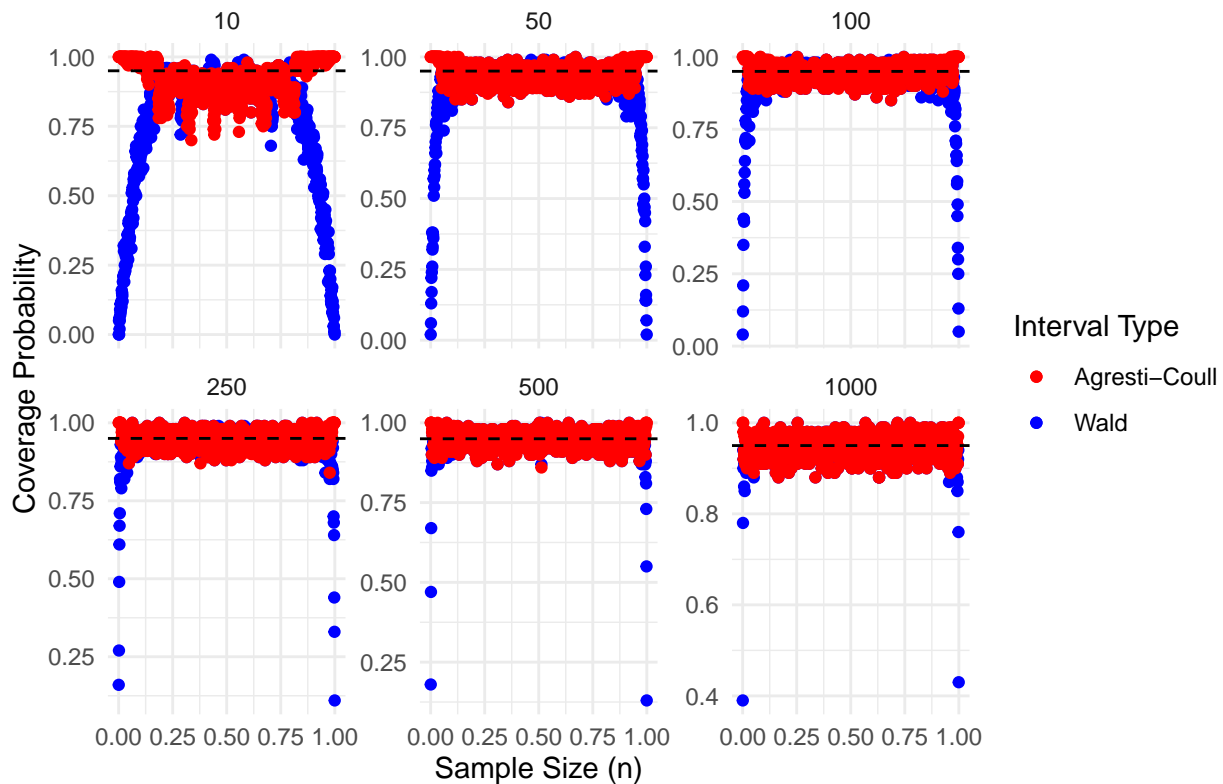
```
## # A tibble: 600,000 x 9
##       n      p  FTM FT_perc FT_ag w_low w_high ag_low ag_high
##   <dbl> <dbl> <int>   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1    10 0.0005     0     0 0.143     0     0 -0.183  0.183
## 2    10 0.0005     0     0 0.143     0     0 -0.183  0.183
```

```
## 3      10 0.0005      0      0 0.143      0      0 -0.183      0.183
## 4      10 0.0005      0      0 0.143      0      0 -0.183      0.183
## 5      10 0.0005      0      0 0.143      0      0 -0.183      0.183
## 6      10 0.0005      0      0 0.143      0      0 -0.183      0.183
## 7      10 0.0005      0      0 0.143      0      0 -0.183      0.183
## 8      10 0.0005      0      0 0.143      0      0 -0.183      0.183
## 9      10 0.0005      0      0 0.143      0      0 -0.183      0.183
## 10     10 0.0005      0      0 0.143      0      0 -0.183      0.183
## # i 599,990 more rows

## `summarise()` has grouped output by 'n'. You can override using the `.groups`
## argument.

## # A tibble: 6,000 x 4
## # Groups:   n [6]
##       n      p cov_w cov_ag
##   <dbl> <dbl> <dbl> <dbl>
## 1     10 0.0005      0      1
## 2     10 0.0015  0.05      1
## 3     10 0.0025      0      1
## 4     10 0.0035  0.06      1
## 5     10 0.0045  0.02      1
## 6     10 0.0055  0.05      1
## 7     10 0.0065  0.09      1
## 8     10 0.0075  0.07      1
## 9     10 0.0085  0.11      1
## 10    10 0.0095  0.08      1
## # i 5,990 more rows
```

Coverage Probability of Wald and Agresti–Coull Intervals



```

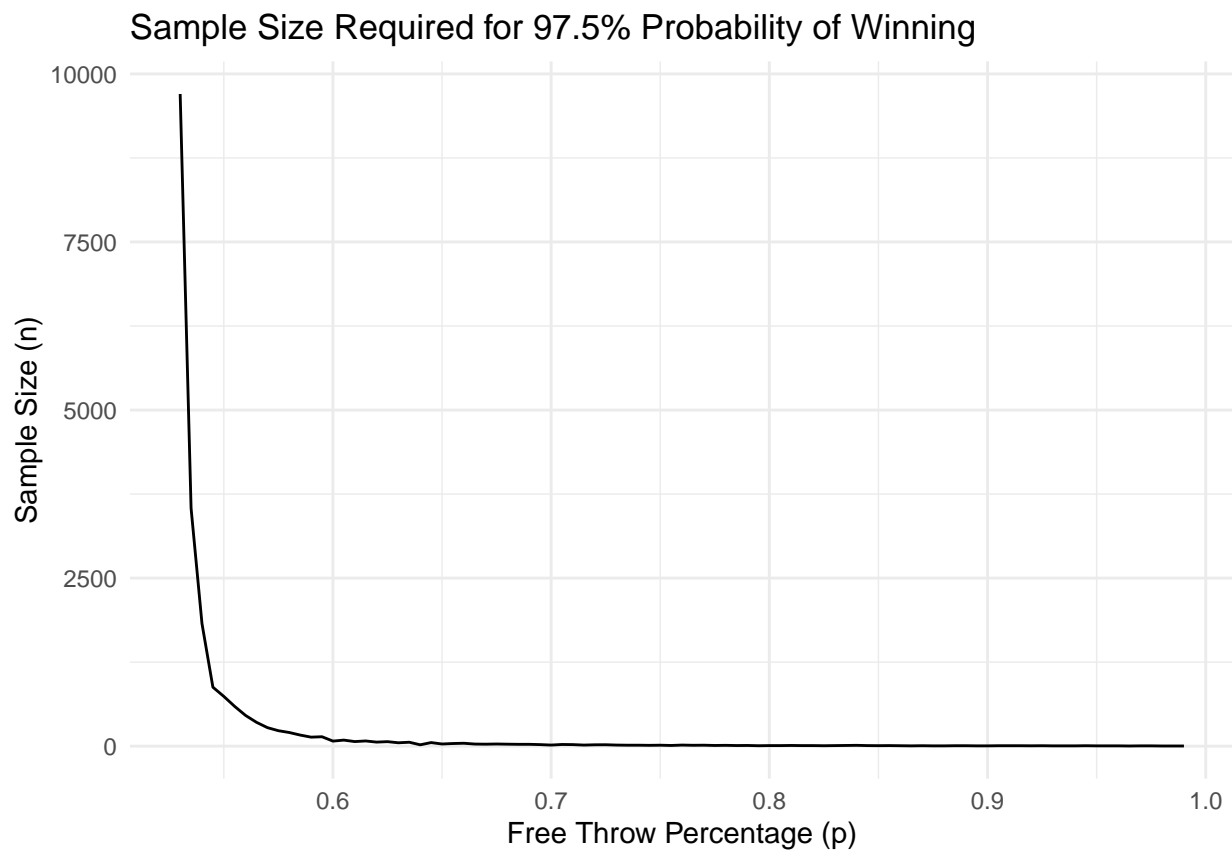
break_even = 11/21

P = seq(0.53, 0.99, by = 0.005)

results = tibble()
for (p in P) {
  n = 1
  greater = 0
  while (greater < 0.975) {
    makes = rbinom(n = 100, size = n, prob = p)
    perc = makes / n
    n = n+1
    greater = mean(perc > break_even)
  }
  ind_results = tibble(
    p = p,
    n = n
  )
  results = bind_rows(results, ind_results)
}

ggplot(results, aes(x = p, y = n)) +
  geom_line() +
  labs(
    title = "Sample Size Required for 97.5% Probability of Winning",
    x = "Free Throw Percentage (p)",
    y = "Sample Size (n)"
  ) +
  theme_minimal()

```



did work i am him