lab 11

2025-06-16

```
##############
### SETUP ###
##############

# install.packages(c("ggplot2", "tidyverse"))
library(ggplot2)
library(tidyverse)
```

```
## Warning: package 'tibble' was built under R version 4.3.3
```

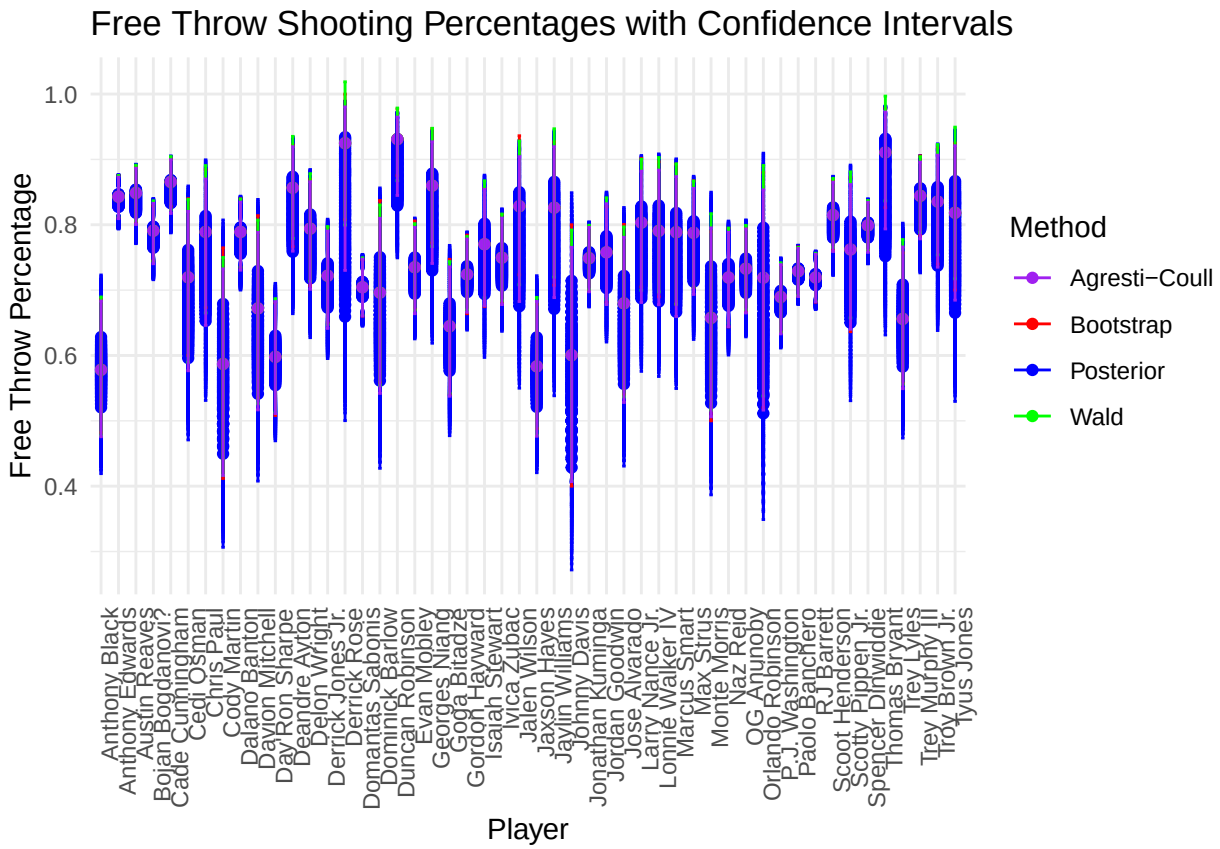
```
## Warning: package 'purrr' was built under R version 4.3.3
```

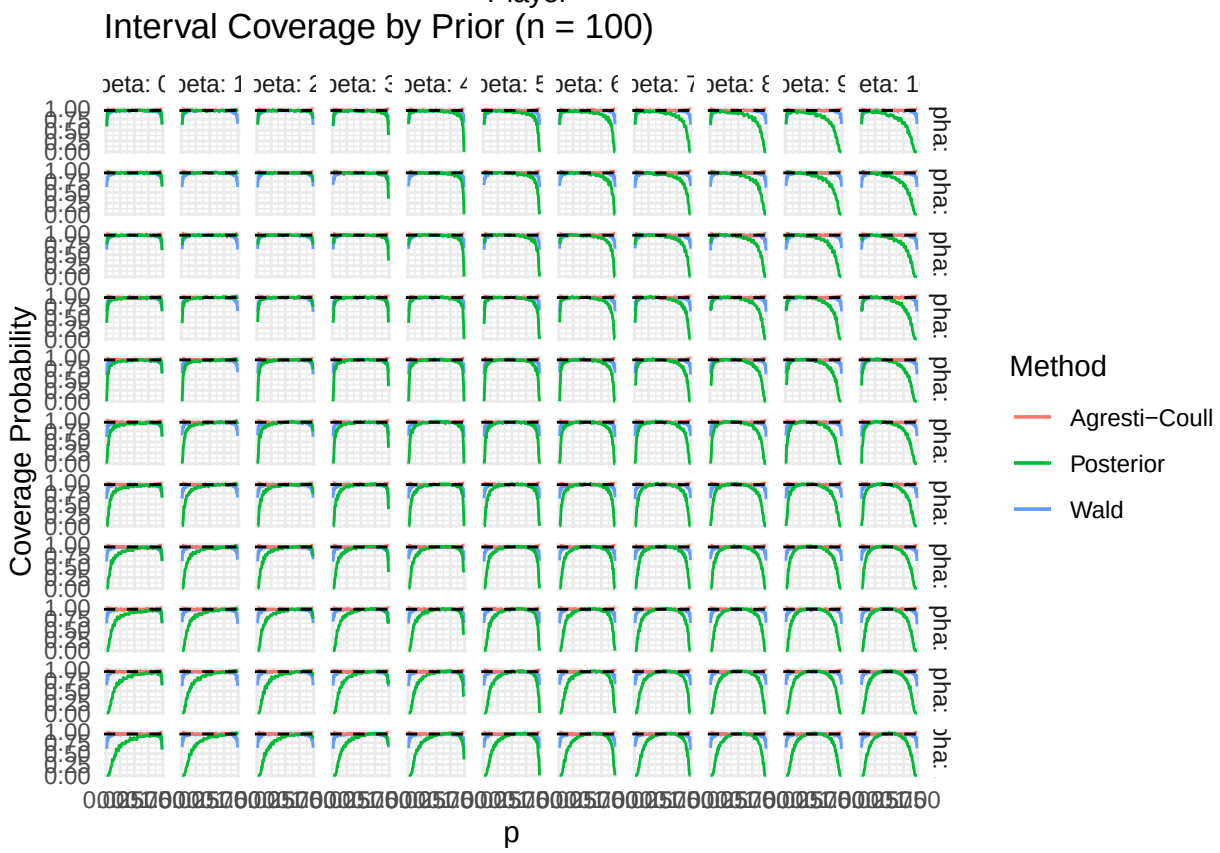
```
## Warning: package 'lubridate' was built under R version 4.3.3
```

```
# set seed
set.seed(8)

########################
### NBA FREE THROWS ###
########################

# load data
nba_players = read_delim("../data/11_nba-free-throws.csv", delim = ";")


alphas = seq(0,10, by = 1)
betas = seq(0,10, by = 1)


##bring in everything from lab 9

data = nba_players %>%
  group_by(Player) %>%
  summarise(
    FT = sum(FT*G, na.rm = TRUE),
    FTA = sum(FTA*G, na.rm = TRUE),
    G = sum(G, na.rm = TRUE)
  ) %>%
  filter(FTA >= 25) %>%
  mutate(FT. = FT/FTA,
         FT.ag = (FT+2)/(FTA+4)) %>%
  arrange(desc(FT.)) %>%
  mutate(
    w.low = FT.- 1.96*sqrt((FT.*(1-FT.))/FTA),
    w.high = FT. + 1.96*sqrt((FT.*(1-FT.))/FTA),
    ag.low = FT.ag - 1.96*sqrt((FT.ag*(1-FT.ag))/(FTA+4)),
    ag.high = FT.ag + 1.96*sqrt((FT.ag*(1-FT.ag))/(FTA+4))
```

```
  ) %>%
  slice_sample(n=50)

#lab 9's bootstrapping
boots = c()
for (i in 1:nrow(data)) {
  player <- data$Player[i]
  ft <- round(data$FT[i])      # Free throws made
  fta <- round(data$FTA[i])    # Free throws attempted

  makes <- rep(1, ft)
  misses <- rep(0, fta - ft)
  shots <- c(makes, misses)

  boot_p <- numeric(1000)
  for (j in 1:1000) {
    boot <- sample(shots, fta, replace = TRUE)
    boot_p[j] <- mean(boot)
  }

  boots[[player]] <- boot_p
}
boots_df <- data.frame(
  Player = rep(names(boots), each = length(boots[[1]])),
  Proportion = unlist(boots)
) %>%
  group_by(Player) %>%
  mutate(
    Lower = quantile(Proportion, 0.025),
    Upper = quantile(Proportion, 0.975)
  ) %>%
  summarise(
    Mean = mean(Proportion),
    Lower = mean(Lower),
    Upper = mean(Upper)
  )

joined = left_join(data, boots_df, by = "Player")


results <- data.frame()
for (alpha in alphas) {
  for (beta in betas) {
    for (player in joined$Player) {
      # calculate the posterior distribution
      ft = joined$FT[joined$Player == player]
      fta = joined$FTA[joined$Player == player]

      # calculate the posterior parameters
      alpha_post = alpha + ft
      beta_post = beta + fta - ft

      # calculate the posterior mean
```

```
      posterior_mean = alpha_post / (alpha_post + beta_post)
      lower_95 = qbeta(0.025, alpha_post, beta_post)
      upper_95 = qbeta(0.975, alpha_post, beta_post)

      # store the results in the grid
      results = rbind(
        results,
        data.frame(
          alpha = alpha,
          beta = beta,
          Player = player,
          posterior_mean = posterior_mean,
          lower_95 = lower_95,
          upper_95 = upper_95))
    }
  }
}


posterior_long <- results %>%
  rename(Mean = posterior_mean, Lower = lower_95, Upper = upper_95) %>%
  mutate(Method = "Posterior")

# Bootstrap results (single interval per player)
bootstrap_long <- joined %>%
  select(Player, Mean, Lower, Upper) %>%
  mutate(Method = "Bootstrap", alpha = NA, beta = NA)

# Wald
wald_long <- joined %>%
  transmute(Player, Mean = Mean, Lower = w.low, Upper = w.high, Method = "Wald", alpha = NA, beta = NA)

# Agresti-Coull
ag_long <- joined %>%
  transmute(Player, Mean = Mean, Lower = ag.low, Upper = ag.high, Method = "Agresti-Coull", alpha = NA,

# Combine all
final_df <- bind_rows(posterior_long, bootstrap_long, wald_long, ag_long)
# Plotting the results
ggplot(final_df, aes(x = Player, y = Mean, color = Method)) +
  geom_point() +
  geom_errorbar(aes(ymin = Lower, ymax = Upper), width = 0.2) +
  labs(title = "Free Throw Shooting Percentages with Confidence Intervals",
       x = "Player",
       y = "Free Throw Percentage") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 90, hjust = 1)) +
  scale_color_manual(values = c("Posterior" = "blue", "Bootstrap" = "red", "Wald" = "green", "Agresti-C
```


Free Throw Shooting Percentages with Confidence Intervals
Free Throw Percentage
1.0
0.8
0.6
0.4
Anthony Black
Anthony Edwards
Austin Reaves
Bojan Bogdanović
Cade Cunningham
Cedi Osman
Chris Paul
Cody Martin
Dalano Banton
Dayon Mitchell
Day'Ron Sharpe
Deandre Ayton
Delon Wright
Derrick Jones Jr.
Derrick Rose
Domantas Sabonis
Dominick Barlow
Duncan Robinson
Evan Mobley
Georges Niang
Goga Bitadze
Gordon Hayward
Isaiah Stewart
Ivica Zubac
Jalen Wilson
Jaxson Hayes
Jaylin Williams
Johnny Davis
Jonathan Kuminga
Jordan Goodwin
Jose Alvarado
Larry Nance Jr.
Lonnie Walker IV
Marcus Smart
Max Strus
Monte Morris
Naz Reid
OG Anunoby
Orlando Robinson
P.J. Washington
Paolo Banchero
RJ Barrett
Scoot Henderson
Scotty Pippen Jr.
Spencer Dinwiddie
Thomas Bryant
Trey Lyles
Trey Murphy III
Troy Brown Jr.
Tyus Jones
Player
Method
Agresti−Coull
Bootstrap
Posterior
Wald


Interval Coverage by Prior (n = 100)
Coverage Probability
p
Method
Agresti−Coull
Posterior
Wald