sim_fmacs_p=6_b

Setup for simulation

```r
library(dplyr)
```

```
Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

    filter, lag

The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union
```

```r
library(ggplot2)
```

```
Warning: package 'ggplot2' was built under R version 4.3.3
```

```r
library(gt)
```

```
Warning: package 'gt' was built under R version 4.3.3
```

```r
library(knitr)
```

```
Warning: package 'knitr' was built under R version 4.3.3
```

```r
library(lavaan)
```

```
Warning: package 'lavaan' was built under R version 4.3.3
```

```
This is lavaan 0.6-18
lavaan is FREE software! Please report any bugs.
```

```r
library(pinsearch)
library(SimDesign)
```

```
Warning: package 'SimDesign' was built under R version 4.3.3
```

```r
# TODO:
#   - Summarize the pattern of bias

# Define conditions: Testing different sample sizes
design <- createDesign(
  n = c(30, 100, 250, 1000)
)

# Fixed objects
set.seed(1855)

# Helper
get_ucov <- function(p, scale = sqrt(0.1), n = 5) {
  W <- matrix(rnorm(p * n), nrow = n)
  WtW <- crossprod(W)
  D <- diag(1 / sqrt(diag(WtW))) * scale
  D %*% WtW %*% D
}
fixed <- list(
  p = 6,
  lambda = c(.3, .7, .4, .5, .6, .4),
  dlambda = list(
    c(0, 0, 0, 0, 0, 0),
    c(.1, 0, 0, 0, 0, 0),
    c(.2, -.3, 0, 0, 0, 0),
    c(.3, -.3, 0, 0, 0, 0)
  ),
```

```r
  nu = c(2, 3, 1.5, 3.5, 2, 3),
  alpha = c(0, -0.25, 0.25, 0.5),
  psi = c(1, 0.85, 1.15, 0.7),
  theta = c(1, 1.2, .8, .9, 1, 1) - .1,
  dtheta = matrix(
    runif(24, min = -0.2, max = 0.2),
    nrow = 4
  ),
  # ucov = replicate(4, get_ucov(6), simplify = FALSE)
  ucov = replicate(4, diag(.1, 6), simplify = FALSE),
  ninv_ind = c(1, 2)
)
# lavaan syntax
fixed$mod <- paste(
  "f =~",
  paste0("y", seq_len(fixed$p), collapse = " + ")
)
# Compute implied means and covariances
fixed <- within(fixed, {
  lambdag <- lapply(dlambda, FUN = \(x) x + lambda)
  Thetag <- lapply(seq_along(ucov),
                   FUN = function(g) {
                     diag(theta + dtheta[g, ]) + ucov[[g]]
                   })
  covy <- mapply(\(lam, psi, th) tcrossprod(lam) * psi + th,
                 lam = lambdag, psi = psi, th = Thetag,
                 SIMPLIFY = FALSE)
  meany <- mapply(\(lam, al, nu) nu + lam * al,
                  lam = lambdag, al = alpha, nu = list(nu),
                  SIMPLIFY = FALSE)
})

# Population effect size
fixed$fmacs_pop <- local({
  pooled_sd <- lapply(fixed$covy, FUN = \(x) diag(x)) |>
    do.call(what = rbind) |>
    colMeans() |>
    sqrt()

  fmacs(
    intercepts = matrix(rep(fixed$nu, 4),
```

```
                          nrow = 4,
                          byrow = TRUE
    ),
    loadings = sweep(
      do.call(rbind, fixed$dlambda),
      MARGIN = 2,
      STATS = fixed$lambda,
      FUN = "+"
    ),
    latent_mean = 0,
    latent_sd = 1,
    pooled_item_sd = pooled_sd
  )[1:2]
})

# Function for data generation
# sim_y <- function(n, lambda, nu, alpha, psi, Theta) {
#     covy <- tcrossprod(lambda) * psi + Theta
#     meany <- nu + lambda * alpha
#     MASS::mvrnorm(n, mu = meany, Sigma = covy)
# }
```

Running the simulation

```
generate <- function(condition, fixed_objects) {
  ylist <- lapply(seq_along(fixed_objects$covy),
                  FUN = function(g) {
                    yg <- MASS::mvrnorm(
                      condition$n,
                      mu = fixed_objects$meany[[g]],
                      Sigma = fixed_objects$covy[[g]]
                    )
                    colnames(yg) <- paste0("y", seq_len(fixed_objects$p))
                    cbind(yg, group  = g)
                  })
  do.call(rbind, ylist)
}
sim1 <- generate(design[3, ], fixed_objects = fixed)


# Analysis
```

```r
analyze <- function(condition, dat, fixed_objects) {
  # Define lavaan syntax
  pinv_fit <- cfa(
    fixed_objects$mod,
    data = dat,
    group = "group", std.lv = TRUE,
    group.equal = c("loadings", "intercepts"),
    group.partial = c(
      paste0("f=~y", fixed_objects$ninv_ind),
      paste0("y", fixed_objects$ninv_ind, "~1")
    )
  )
  as.vector(pinsearch::pin_effsize(pinv_fit))
}

analyze_bc <- function(condition, dat, fixed_objects) {
  # Define lavaan syntax
  pinv_fit <- cfa(
    fixed_objects$mod,
    data = dat,
    group = "group", std.lv = TRUE,
    group.equal = c("loadings", "intercepts"),
    group.partial = c(
      paste0("f=~y", fixed_objects$ninv_ind),
      paste0("y", fixed_objects$ninv_ind, "~1")
    )
  )
  f_orig <- as.vector(pinsearch::pin_effsize(pinv_fit))
  f_boot <- lavaan::bootstrapLavaan(pinv_fit,
                                    R = 100,
                                    FUN = pinsearch::pin_effsize
  )
  pmax(0, 2 * f_orig - colMeans(f_boot, na.rm = TRUE))
}

analyze_bc2 <- function(condition, dat, fixed_objects) {
  # Define lavaan syntax
  pinv_fit <- cfa(
    fixed_objects$mod,
    data = dat,
    group = "group", std.lv = TRUE,
```

```r
    group.equal = c("loadings", "intercepts"),
    group.partial = c(
      paste0("f=~y", fixed_objects$ninv_ind),
      paste0("y", fixed_objects$ninv_ind, "~1")
    )
  )
  f_orig <- pinsearch::pin_effsize(pinv_fit)
  ns <- lavInspect(pinv_fit, what = "nobs")
  ng <- length(ns)
  f2_bias <- (ng - 1) / ng * sum(1 / ns)
  sqrt(pmax(0, f_orig^2 - f2_bias))
}

# Evaluate/Summarize
evaluate <- function(condition, results, fixed_objects) {
  c(
    bias = colMeans(results) - fixed_objects$fmacs_pop,
    robust_bias = apply(results, 2, mean, trim = .1) -
      fixed_objects$fmacs_pop,
    emp_sd = apply(results, 2, sd),
    emp_mad = apply(results, 2, mad)
  )
}
```

```r
out <- runSimulation(design,
                     replications = 500,
                     parallel = TRUE,
                     ncores = parallelly::availableCores(omit = 3L),
                     generate = generate,
                     analyse = list(naive = analyze,
                                    bc_boot = analyze_bc,
                                    bc_form = analyze_bc2),
                     summarise = evaluate,
                     filename = "results-trial-bc",
                     packages = c("MASS", "lavaan", "pinsearch"),
                     fixed_objects = fixed,
                     save_results = TRUE)
```

```
Number of parallel clusters in use: 17
```

```
./results-trial-bc-results_DESKTOP-342341B already exists; using ./results-trial-bc-results_
```

```
Design: 1/4;   RAM Used: 71.9 Mb;   Replications: 500;   Total Time: 0.00s
 Conditions: n=30
```

```
Design: 2/4;   RAM Used: 72.6 Mb;   Replications: 500;   Total Time: 02h 55m 52.31s
 Conditions: n=100
```

```
Design: 3/4;   RAM Used: 72.6 Mb;   Replications: 500;   Total Time: 03h 24m 32.06s
 Conditions: n=250
```

```
Design: 4/4;   RAM Used: 72.6 Mb;   Replications: 500;   Total Time: 03h 40m 46.13s
 Conditions: n=1000
```

```
Simulation complete. Total execution time: 03h 56m 37.17s
```

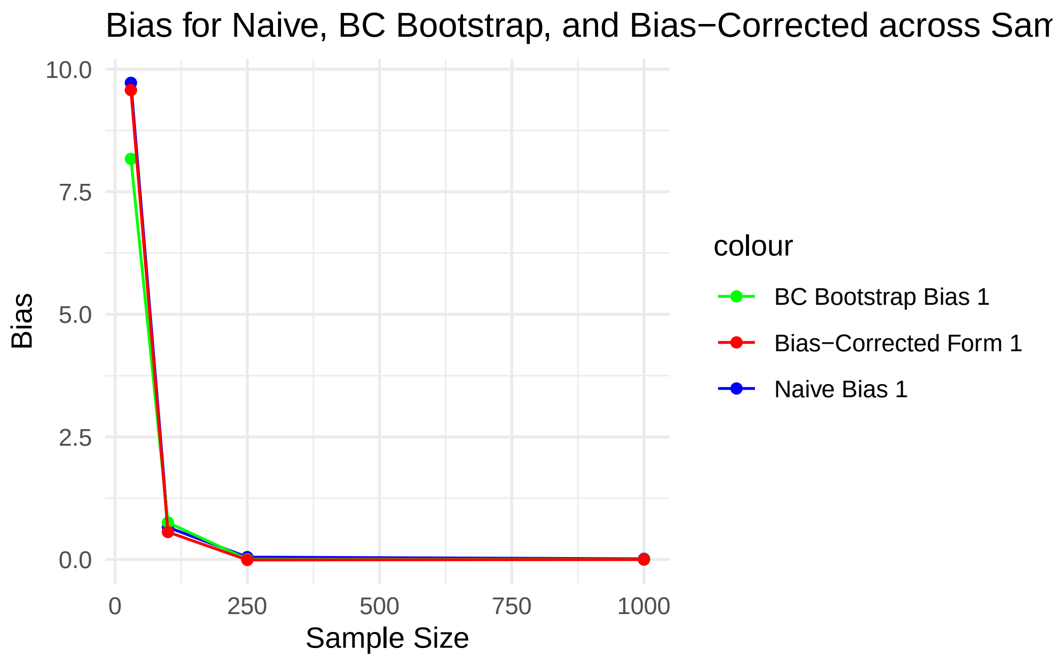
```
Saving simulation results to file: results-trial-bc.rds
```

```r
data <- readRDS("C:/Users/alex/OneDrive/Desktop/results-trial-bc.rds")
```

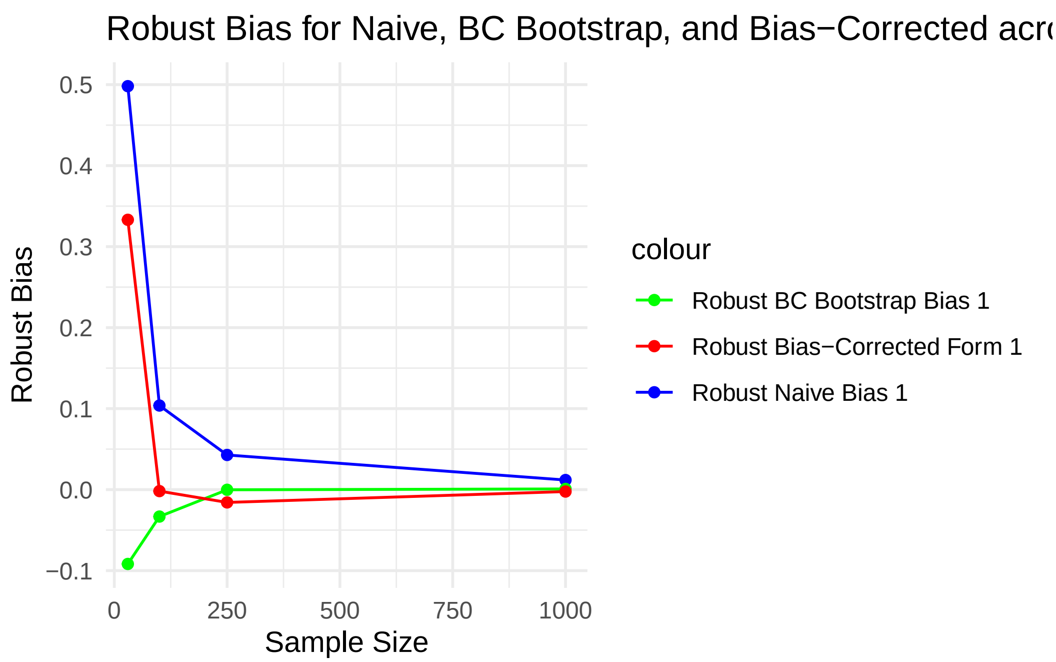
Graphs to visualize simulation results

```r
# Bias
bias_plot <- ggplot(data, aes(x = n)) +
  geom_line(aes(y = bias.naive1, color = "Naive Bias 1")) +
  geom_point(aes(y = bias.naive1, color = "Naive Bias 1")) +
  geom_line(aes(y = bias.bc_boot1, color = "BC Bootstrap Bias 1")) +
  geom_point(aes(y = bias.bc_boot1, color = "BC Bootstrap Bias 1")) +
  geom_line(aes(y = bias.bc_form1, color = "Bias-Corrected Form 1")) +
  geom_point(aes(y = bias.bc_form1, color = "Bias-Corrected Form 1")) +
  labs(title = "Bias for Naive, BC Bootstrap, and Bias-Corrected across Sample Sizes",
       x = "Sample Size", y = "Bias") +
  scale_color_manual(values = c("Naive Bias 1" = "blue",
                                "BC Bootstrap Bias 1" = "green",
```

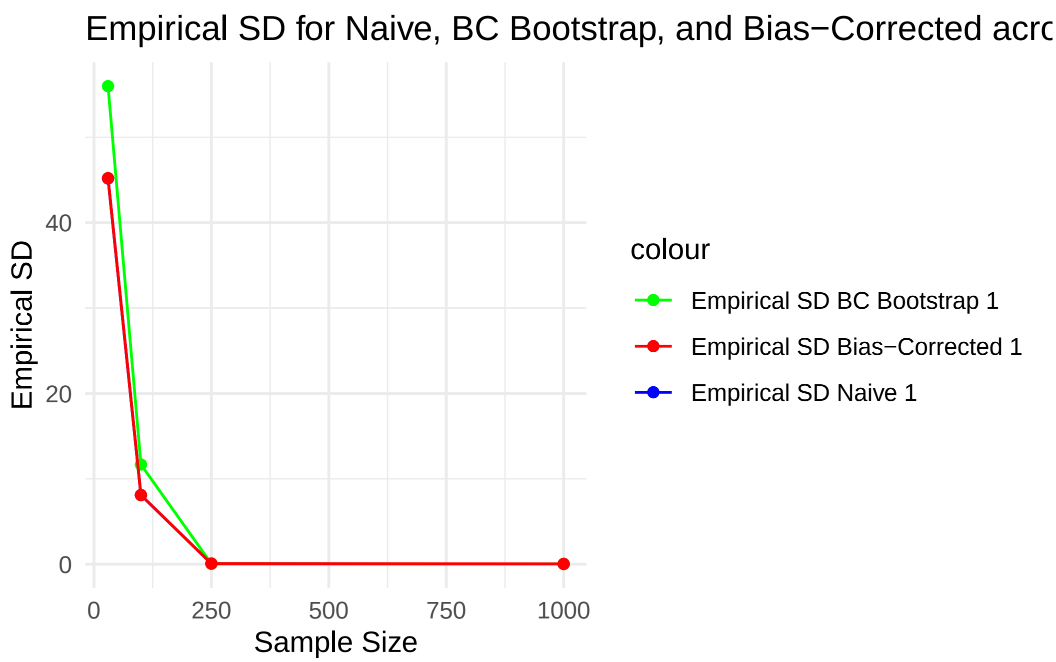
```
                                "Bias-Corrected Form 1" = "red")) +
  theme_minimal()
bias_plot
```



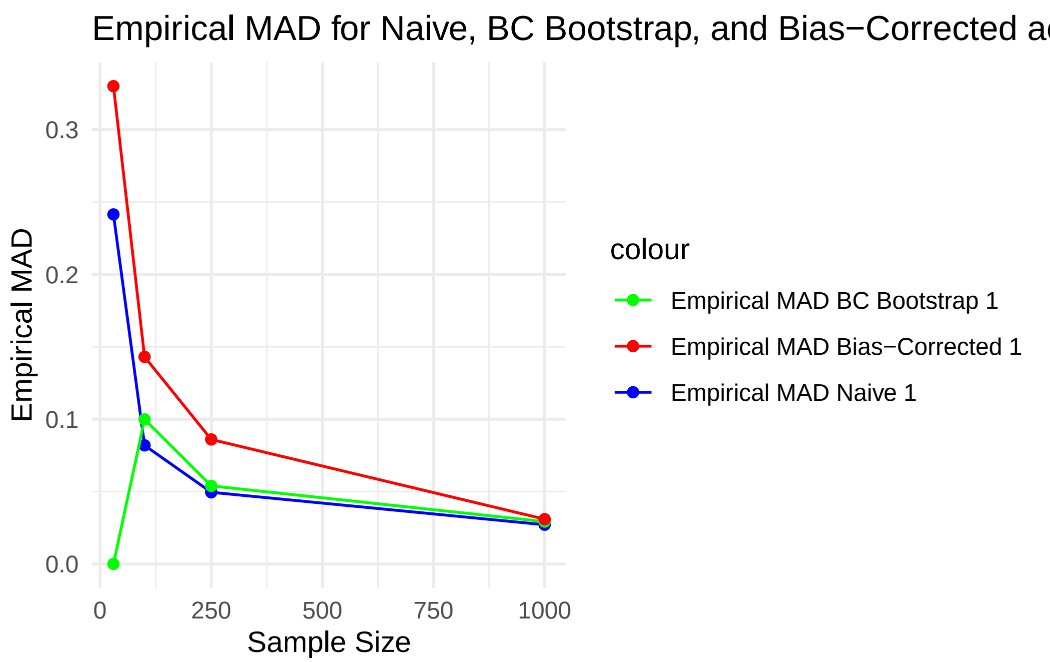

```
# Robust Bias
robust_bias_plot <- ggplot(data, aes(x = n)) +
  geom_line(aes(y = robust_bias.naive1, color = "Robust Naive Bias 1")) +
  geom_point(aes(y = robust_bias.naive1, color = "Robust Naive Bias 1")) +
  geom_line(aes(y = robust_bias.bc_boot1, color = "Robust BC Bootstrap Bias 1")) +
  geom_point(aes(y = robust_bias.bc_boot1, color = "Robust BC Bootstrap Bias 1")) +
  geom_line(aes(y = robust_bias.bc_form1, color = "Robust Bias-Corrected Form 1")) +
  geom_point(aes(y = robust_bias.bc_form1, color = "Robust Bias-Corrected Form 1")) +
  labs(title = "Robust Bias for Naive, BC Bootstrap, and Bias-Corrected across Sample Size
       x = "Sample Size", y = "Robust Bias") +
  scale_color_manual(values = c("Robust Naive Bias 1" = "blue",
                                "Robust BC Bootstrap Bias 1" = "green",
                                "Robust Bias-Corrected Form 1" = "red")) +
  theme_minimal()
robust_bias_plot
```




```r
# Empirical Standard Deviation
emp_sd_plot <- ggplot(data, aes(x = n)) +
  geom_line(aes(y = emp_sd.naive1, color = "Empirical SD Naive 1")) +
  geom_point(aes(y = emp_sd.naive1, color = "Empirical SD Naive 1")) +
  geom_line(aes(y = emp_sd.bc_boot1, color = "Empirical SD BC Bootstrap 1")) +
  geom_point(aes(y = emp_sd.bc_boot1, color = "Empirical SD BC Bootstrap 1")) +
  geom_line(aes(y = emp_sd.bc_form1, color = "Empirical SD Bias-Corrected 1")) +
  geom_point(aes(y = emp_sd.bc_form1, color = "Empirical SD Bias-Corrected 1")) +
  labs(title = "Empirical SD for Naive, BC Bootstrap, and Bias-Corrected across Sample Siz
       x = "Sample Size", y = "Empirical SD") +
  scale_color_manual(values = c("Empirical SD Naive 1" = "blue",
                                "Empirical SD BC Bootstrap 1" = "green",
                                "Empirical SD Bias-Corrected 1" = "red")) +
  theme_minimal()
emp_sd_plot
```




```r
# Median Absolute Deviation
emp_mad_plot <- ggplot(data, aes(x = n)) +
  geom_line(aes(y = emp_mad.naive1, color = "Empirical MAD Naive 1")) +
  geom_point(aes(y = emp_mad.naive1, color = "Empirical MAD Naive 1")) +
  geom_line(aes(y = emp_mad.bc_boot1, color = "Empirical MAD BC Bootstrap 1")) +
  geom_point(aes(y = emp_mad.bc_boot1, color = "Empirical MAD BC Bootstrap 1")) +
  geom_line(aes(y = emp_mad.bc_form1, color = "Empirical MAD Bias-Corrected 1")) +
  geom_point(aes(y = emp_mad.bc_form1, color = "Empirical MAD Bias-Corrected 1")) +
  labs(title = "Empirical MAD for Naive, BC Bootstrap, and Bias-Corrected across Sample Si
       x = "Sample Size", y = "Empirical MAD") +
  scale_color_manual(values = c("Empirical MAD Naive 1" = "blue",
                                "Empirical MAD BC Bootstrap 1" = "green",
                                "Empirical MAD Bias-Corrected 1" = "red")) +
  theme_minimal()
emp_mad_plot
```



Set up for tables

```
bias_table <- data.frame(
  SampleSize = data$n,
  Bias_Naive1 = data$bias.naive1,
  Bias_BC_Boot1 = data$bias.bc_boot1,
  Bias_BC_Form1 = data$bias.bc_form1
)

robust_bias_table <- data.frame(
  SampleSize = data$n,
  Robust_Bias_Naive1 = data$robust_bias.naive1,
  Robust_Bias_BC_Boot1 = data$robust_bias.bc_boot1,
  Robust_Bias_BC_Form1 = data$robust_bias.bc_form1
)

emp_sd_table <- data.frame(
  SampleSize = data$n,
  Empirical_SD_Naive1 = data$emp_sd.naive1,
  Empirical_SD_BC_Boot1 = data$emp_sd.bc_boot1,
  Empirical_SD_BC_Form1 = data$emp_sd.bc_form1
)
```

```r
emp_mad_table <- data.frame(
  SampleSize = data$n,
  Empirical_MAD_Naive1 = data$emp_mad.naive1,
  Empirical_MAD_BC_Boot1 = data$emp_mad.bc_boot1,
  Empirical_MAD_BC_Form1 = data$emp_mad.bc_form1
)

# Create and display tables using knitr
knitr::kable(bias_table, caption = "Bias for Naive, BC Bootstrap, and Bias-Corrected acros
```

Table 1: Bias for Naive, BC Bootstrap, and Bias-Corrected across Sample Sizes

SampleSize	Bias_Naive1	Bias_BC_Boot1	Bias_BC_Form1
30	9.7218803	8.1710552	9.5740247
100	0.6569590	0.7541530	0.5600170
250	0.0477572	0.0033855	-0.0090273
1000	0.0129750	0.0017650	-0.0019629

```r
knitr::kable(robust_bias_table, caption = "Robust Bias for Naive, BC Bootstrap, and Bias-C
```

Table 2: Robust Bias for Naive, BC Bootstrap, and Bias-Corrected across Sample Sizes

SampleSize	Robust_Bias_Naive1	Robust_Bias_BC_Boot1	Robust_Bias_BC_Form1
30	0.4980913	-0.0918035	0.3331221
100	0.1037743	-0.0332902	-0.0018570
250	0.0427538	-0.0002190	-0.0158307
1000	0.0118461	0.0008494	-0.0024153

```r
knitr::kable(emp_sd_table, caption = "Empirical Standard Deviation for Naive, BC Bootstrap
```

Table 3: Empirical Standard Deviation for Naive, BC Bootstrap, and Bias-Corrected across Sample Sizes

SampleSize	Empirical_SD_Naive1	Empirical_SD_BC_Boot1	Empirical_SD_BC_Form1
30	45.1790360	55.9852546	45.2100767
100	8.0936589	11.6736326	8.1006171
250	0.0552012	0.0577781	0.0769428

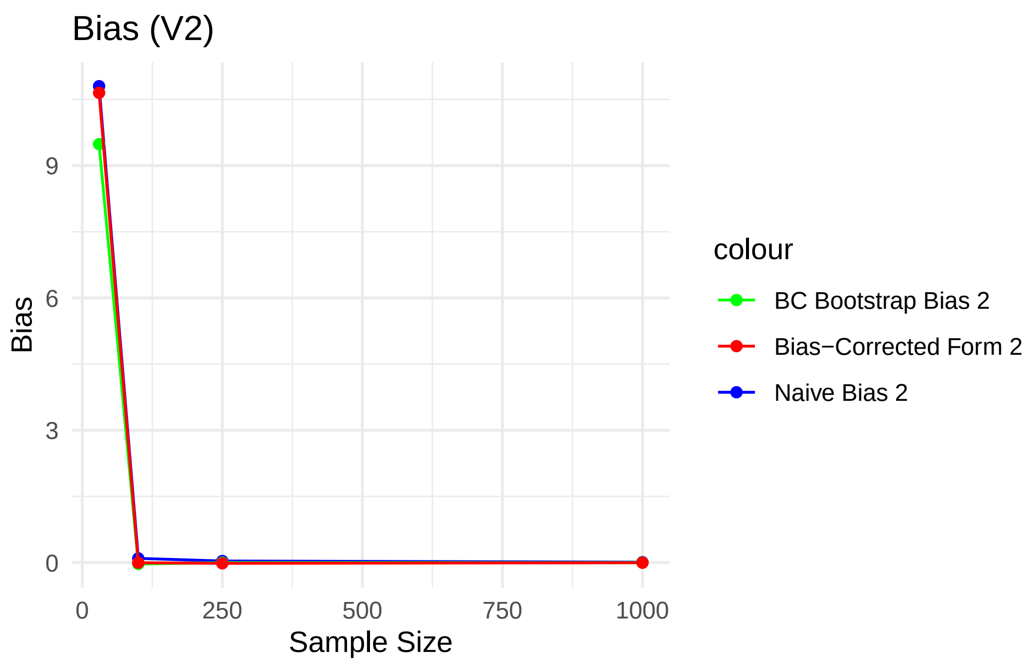
SampleSize	Empirical_SD_Naive1	Empirical_SD_BC_Boot1	Empirical_SD_BC_Form1
1000	0.0273751	0.0291286	0.0319034

```
knitr::kable(emp_mad_table, caption = "Empirical MAD for Naive, BC Bootstrap, and Bias-Cor
```

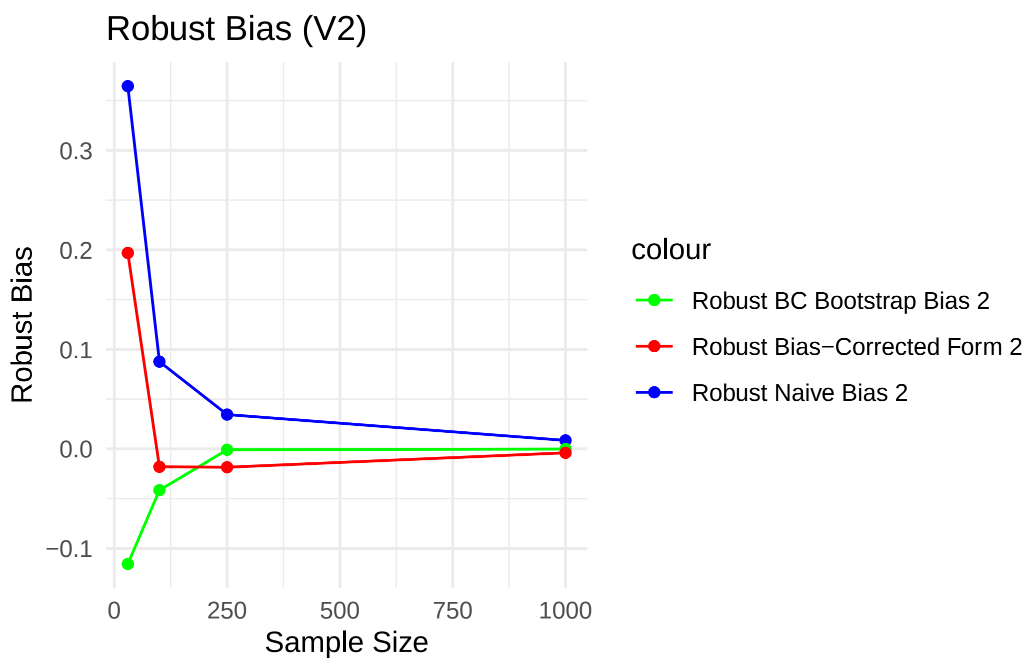
Table 4: Empirical MAD for Naive, BC Bootstrap, and Bias-Corrected across Sample Sizes

SampleSize	Empirical_MAD_Naive1	Empirical_MAD_BC_Boot1	Empirical_MAD_BC_Form1
30	0.2414609	0.0000000	0.3300305
100	0.0819243	0.0998460	0.1430580
250	0.0495064	0.0539396	0.0860082
1000	0.0270873	0.0292359	0.0309464

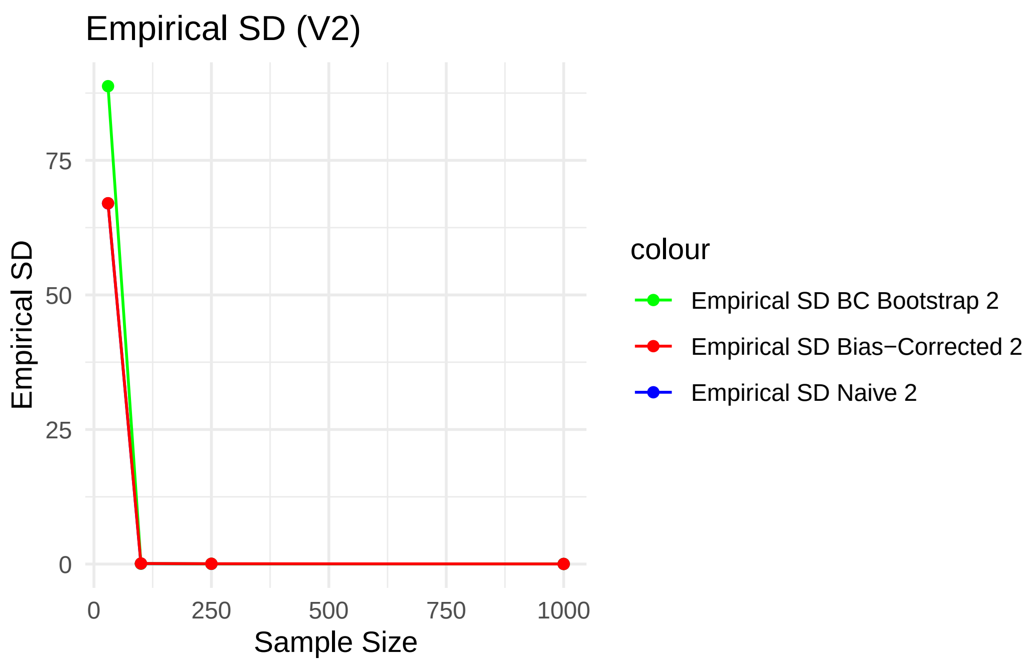
```
bias_plot_2 <- ggplot(data, aes(x = n)) +
  geom_line(aes(y = bias.naive2, color = "Naive Bias 2")) +
  geom_point(aes(y = bias.naive2, color = "Naive Bias 2")) +
  geom_line(aes(y = bias.bc_boot2, color = "BC Bootstrap Bias 2")) +
  geom_point(aes(y = bias.bc_boot2, color = "BC Bootstrap Bias 2")) +
  geom_line(aes(y = bias.bc_form2, color = "Bias-Corrected Form 2")) +
  geom_point(aes(y = bias.bc_form2, color = "Bias-Corrected Form 2")) +
  labs(title = "Bias (V2)",
       x = "Sample Size", y = "Bias") +
  scale_color_manual(values = c("Naive Bias 2" = "blue",
                                "BC Bootstrap Bias 2" = "green",
                                "Bias-Corrected Form 2" = "red")) +
  theme_minimal()
bias_plot_2
```



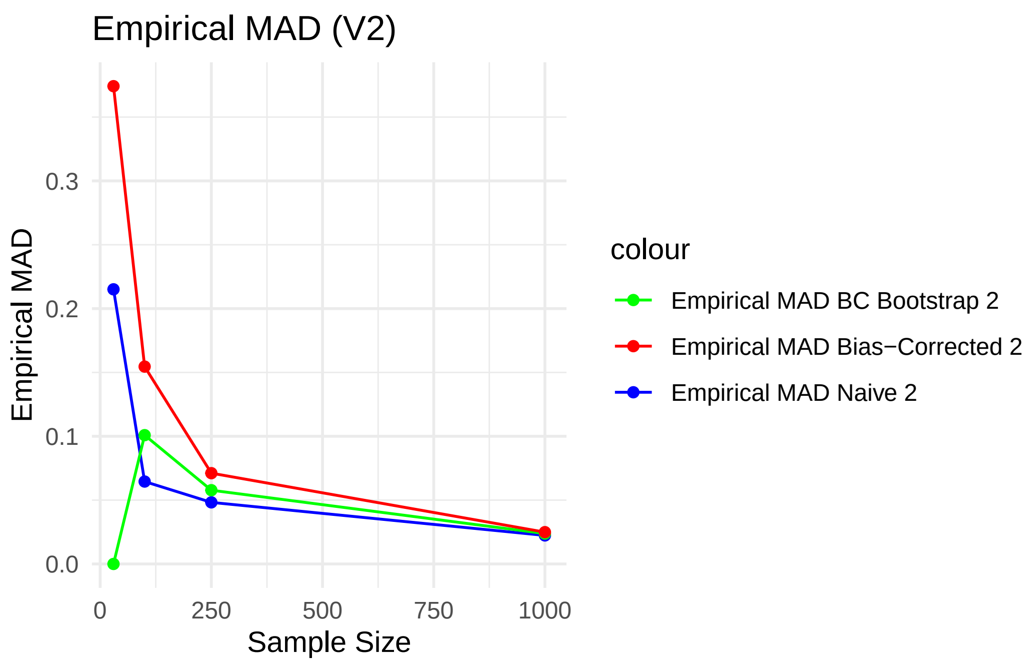

```r
robust_bias_plot_2 <- ggplot(data, aes(x = n)) +
  geom_line(aes(y = robust_bias.naive2, color = "Robust Naive Bias 2")) +
  geom_point(aes(y = robust_bias.naive2, color = "Robust Naive Bias 2")) +
  geom_line(aes(y = robust_bias.bc_boot2, color = "Robust BC Bootstrap Bias 2")) +
  geom_point(aes(y = robust_bias.bc_boot2, color = "Robust BC Bootstrap Bias 2")) +
  geom_line(aes(y = robust_bias.bc_form2, color = "Robust Bias-Corrected Form 2")) +
  geom_point(aes(y = robust_bias.bc_form2, color = "Robust Bias-Corrected Form 2")) +
  labs(title = "Robust Bias (V2)",
       x = "Sample Size", y = "Robust Bias") +
  scale_color_manual(values = c("Robust Naive Bias 2" = "blue",
                               "Robust BC Bootstrap Bias 2" = "green",
                               "Robust Bias-Corrected Form 2" = "red")) +
  theme_minimal()
robust_bias_plot_2
```




```r
emp_sd_plot_2 <- ggplot(data, aes(x = n)) +
  geom_line(aes(y = emp_sd.naive2, color = "Empirical SD Naive 2")) +
  geom_point(aes(y = emp_sd.naive2, color = "Empirical SD Naive 2")) +
  geom_line(aes(y = emp_sd.bc_boot2, color = "Empirical SD BC Bootstrap 2")) +
  geom_point(aes(y = emp_sd.bc_boot2, color = "Empirical SD BC Bootstrap 2")) +
  geom_line(aes(y = emp_sd.bc_form2, color = "Empirical SD Bias-Corrected 2")) +
  geom_point(aes(y = emp_sd.bc_form2, color = "Empirical SD Bias-Corrected 2")) +
  labs(title = "Empirical SD (V2)",
       x = "Sample Size", y = "Empirical SD") +
  scale_color_manual(values = c("Empirical SD Naive 2" = "blue",
                                "Empirical SD BC Bootstrap 2" = "green",
                                "Empirical SD Bias-Corrected 2" = "red")) +
  theme_minimal()
emp_sd_plot_2
```




```
emp_mad_plot_2 <- ggplot(data, aes(x = n)) +
  geom_line(aes(y = emp_mad.naive2, color = "Empirical MAD Naive 2")) +
  geom_point(aes(y = emp_mad.naive2, color = "Empirical MAD Naive 2")) +
  geom_line(aes(y = emp_mad.bc_boot2, color = "Empirical MAD BC Bootstrap 2")) +
  geom_point(aes(y = emp_mad.bc_boot2, color = "Empirical MAD BC Bootstrap 2")) +
  geom_line(aes(y = emp_mad.bc_form2, color = "Empirical MAD Bias-Corrected 2")) +
  geom_point(aes(y = emp_mad.bc_form2, color = "Empirical MAD Bias-Corrected 2")) +
  labs(title = "Empirical MAD (V2)",
       x = "Sample Size", y = "Empirical MAD") +
  scale_color_manual(values = c("Empirical MAD Naive 2" = "blue",
                                "Empirical MAD BC Bootstrap 2" = "green",
                                "Empirical MAD Bias-Corrected 2" = "red")) +
  theme_minimal()
emp_mad_plot_2
```




```r
bias_table_2 <- data.frame(
  SampleSize = data$n,
  Bias_Naive2 = data$bias.naive2,
  Bias_BC_Boot2 = data$bias.bc_boot2,
  Bias_BC_Form2 = data$bias.bc_form2
)

robust_bias_table_2 <- data.frame(
  SampleSize = data$n,
  Robust_Bias_Naive2 = data$robust_bias.naive2,
  Robust_Bias_BC_Boot2 = data$robust_bias.bc_boot2,
  Robust_Bias_BC_Form2 = data$robust_bias.bc_form2
)

emp_sd_table_2 <- data.frame(
  SampleSize = data$n,
  Empirical_SD_Naive2 = data$emp_sd.naive2,
  Empirical_SD_BC_Boot2 = data$emp_sd.bc_boot2,
  Empirical_SD_BC_Form2 = data$emp_sd.bc_form2
)
```

```r
emp_mad_table_2 <- data.frame(
  SampleSize = data$n,
  Empirical_MAD_Naive2 = data$emp_mad.naive2,
  Empirical_MAD_BC_Boot2 = data$emp_mad.bc_boot2,
  Empirical_MAD_BC_Form2 = data$emp_mad.bc_form2
)

# Create and display tables using knitr
knitr::kable(bias_table_2, caption = "Bias V2")
```

Table 5: Bias V2

SampleSize	Bias_Naive2	Bias_BC_Boot2	Bias_BC_Form2
30	10.7996025	9.4862657	10.6487700
100	0.0948459	-0.0316131	-0.0048182
250	0.0353780	-0.0013300	-0.0177776
1000	0.0092022	0.0002600	-0.0035215

```r
knitr::kable(robust_bias_table_2, caption = "Robust Bias V2")
```

Table 6: Robust Bias V2

SampleSize	Robust_Bias_Naive2	Robust_Bias_BC_Boot2	Robust_Bias_BC_Form2
30	0.3644530	-0.1155737	0.1969064
100	0.0876955	-0.0414437	-0.0179622
250	0.0345028	-0.0008506	-0.0183926
1000	0.0085595	-0.0001209	-0.0039114

```r
knitr::kable(emp_sd_table_2, caption = "Empirical Standard Deviation V2")
```

Table 7: Empirical Standard Deviation V2

SampleSize	Empirical_SD_Naive2	Empirical_SD_BC_Boot2	Empirical_SD_BC_Form2
30	67.0112165	88.7801769	67.0350284
100	0.0819323	0.0851897	0.1130994
250	0.0450277	0.0538186	0.0656115
1000	0.0230113	0.0245284	0.0256151

```r
knitr::kable(emp_mad_table_2, caption = "Empirical MAD V2")
```

Table 8: Empirical MAD V2

SampleSize	Empirical_MAD_Naive2	Empirical_MAD_BC_Boot2	Empirical_MAD_BC_Form2
30	0.2150856	0.0000000	0.3741931
100	0.0645431	0.1008793	0.1545591
250	0.0482424	0.0577743	0.0711156
1000	0.0223497	0.0238004	0.0249093