CI & CP Summaries

Timothy Jordan Regis

2022-10-15

```r
#install.packages("tidyverse")
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.3.6      v purrr   0.3.4
## v tibble  3.1.7      v dplyr   1.0.9
## v tidyr   1.2.1      v stringr 1.4.0
## v readr   2.1.2      v forcats 0.5.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

F:/School/MSc/STA4000/All_Data_New/Newer Results

```r
# Concatenate all Files
all_new_cis = all_old_cis = all_old_tot = all_old_time = all_old_full = job_num = c()
for(j in 1:200){
  all_new_cis <- rbind(all_new_cis, readRDS(paste0(
    'F:/School/MSc/STA4000/All_Data_New/Newer Results/CI_K10_A8.5_B8.5_R0.006_job', j, '.rds')))
  all_old_cis <- rbind(all_old_cis, read.table(paste0(
    'F:/School/MSc/STA4000/All_Data_New/SimulationRun1010/', j, 'balan-2CI-10-alpha-8.5-beta-8.5-baseR-0
  all_old_tot <- rbind(all_old_tot, read.table(paste0(
    'F:/School/MSc/STA4000/All_Data_New/SimulationRun1010/', j, 'balan-2TOT-10-alpha-8.5-beta-8.5-baseR-
  all_old_time <- rbind(all_old_time, read.table(paste0(
    'F:/School/MSc/STA4000/All_Data_New/SimulationRun1010/', j, 'balan-2TIME-10-alpha-8.5-beta-8.5-baseR
  job_num = rbind(data.frame(job_num), data.frame(rep(j, 10)))
  }
```

```r
# Bind Old Run Files

all_new_full = data.frame(all_new_cis)

all_old_full = data.frame(cbind(job = job_num$rep.j..10.,
                                lower = all_old_cis$V1,
                                upper = all_old_cis$V2,
                                delta = all_old_cis$V2 - all_old_cis$V1,
                                tot_l = all_old_tot$V1,
                                tot_u = all_old_tot$V2,
                                time = all_old_time$V1))
```

```r
# Create variables for coverage probability measurements

alpha = beta = 8.5
theta = alpha/(alpha+beta)
```

```r
all_new_full <- all_new_full  %>%
  mutate(
    in_bounds = case_when(
      lower <= theta & theta <= upper ~ 1,
      theta < lower | theta > upper ~ 0
    )
  )

all_old_full <- all_old_full  %>%
  mutate(
    in_bounds = case_when(
      lower <= theta & theta <= upper ~ 1,
      theta < lower | theta > upper ~ 0
    )
  )

# summarize old data by job
all_old_summaries <- all_old_full %>%
  group_by(job) %>%
  summarise("mean_cp" = mean(in_bounds),
            "na_lower" = sum(is.na(lower)),
            "mean_lower" = mean(lower, na.rm = TRUE),
            "med_lower" = median(lower, na.rm = TRUE),
            "na_upper" = sum(is.na(upper)),
            "mean_upper" = mean(upper, na.rm = TRUE),
            "med_upper" = median(upper, na.rm = TRUE),
            "min_len" = min(delta, na.rm = TRUE),
            "mean_len" = mean(delta, na.rm = TRUE),
            "med_len" = median(delta, na.rm = TRUE),
            "max_len" = max(delta, na.rm = TRUE)
            )

all_old_full_final <- left_join(all_old_full, all_old_summaries, by = "job")

# summarise data globally

all_new_full_globals <- all_new_full %>%
  mutate(
    global_mean_cp = mean(in_bounds),
    global_na_lower = sum(is.na(lower)),
    global_mean_lower = mean(lower, na.rm = TRUE),
    global_median_lower = median(lower, na.rm = TRUE),
    global_na_upper = sum(is.na(upper)),
    global_mean_upper = mean(upper, na.rm = TRUE),
    global_median_upper = median(upper, na.rm = TRUE),
    global_min_len = min(delta, na.rm = TRUE),
    global_mean_len = mean(delta, na.rm = TRUE),
    global_median_len = median(delta, na.rm = TRUE),
    global_max_len = max(delta, na.rm = TRUE)
  )

all_old_full_globals <- all_old_full_final %>%
  mutate(
    global_mean_cp = mean(in_bounds),
```

```
    global_na_lower = sum(is.na(lower)),
    global_mean_lower = mean(lower, na.rm = TRUE),
    global_median_lower = median(lower, na.rm = TRUE),
    global_na_upper = sum(is.na(upper)),
    global_mean_upper = mean(upper, na.rm = TRUE),
    global_median_upper = median(upper, na.rm = TRUE),
    global_min_len = min(delta, na.rm = TRUE),
    global_mean_len = mean(delta, na.rm = TRUE),
    global_median_len = median(delta, na.rm = TRUE),
    global_max_len = max(delta, na.rm = TRUE)
  )
```

```
write_csv(all_new_full_globals, "New_Run_Data.csv")
write_csv(all_old_full_globals, "Old_Run_Data.csv")
```

```
# sample
#head(all_new_full_globals)
##head(all_old_full_globals)
```

```
Run = c("Old", "New")
mean_lower = c(mean(all_old_full_globals$lower), mean(all_new_full_globals$lower))
mean_upper = c(mean(all_old_full_globals$upper), mean(all_new_full_globals$upper))
mean_delta = c(mean(all_old_full_globals$delta), mean(all_new_full_globals$delta))
median_delta = c(median(all_old_full_globals$delta), median(all_new_full_globals$delta))
cov_prob =  c(mean(all_old_full_globals$global_mean_cp), mean(all_new_full_globals$global_mean_cp))

display_table = data.frame(Run, mean_lower, mean_upper, mean_delta, median_delta, cov_prob)
```

```
display_table
```

```
##   Run mean_lower mean_upper mean_delta median_delta cov_prob
## 1 Old  0.2648672  0.7472838  0.4824167    0.4810446   0.9785
## 2 New  0.2518856  0.7560009  0.5041153    0.5043140   0.9790
```