CR_Stats

Anton Holm

2020-05-06

```
library(tidyverse)
```

```
## -- Attaching packages ---------------------------------------------------------------------------
```

```
## v ggplot2 3.2.1     v purrr   0.3.3
## v tibble  2.1.3     v dplyr   0.8.3
## v tidyr   1.0.0     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.4.0
```

```
## -- Conflicts ------------------------------------------------------------------------------------ tidy
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(kableExtra)
```

```
##
## Attaching package: 'kableExtra'
```

```
## The following object is masked from 'package:dplyr':
##
##     group_rows
```

```
load("df_80_1.05_0.05.Rda")
load("df_80_1.1_0.05.Rda")
load("df_80_1.05_0.5.Rda")
load("df_80_1.1_0.5.Rda")


beta_1 <- split(df_80_1.05_0.5, df_80_1.05_0.5$Method)
betas_lm <- beta_1[[1]] %>% mutate(Beta_Subs = beta_1[[2]]$Beta_Hat) %>% mutate(Sd_Subs = beta_1[[2]]$S

beta_2 <- split(df_80_1.05_0.05, df_80_1.05_0.05$Method)
betas_ll <- beta_2[[1]] %>% mutate(Beta_Subs = beta_2[[2]]$Beta_Hat) %>% mutate(Sd_Subs = beta_2[[2]]$S

beta_3 <- split(df_80_1.1_0.05, df_80_1.1_0.05$Method)
betas_ml <- beta_3[[1]] %>% mutate(Beta_Subs = beta_3[[2]]$Beta_Hat) %>% mutate(Sd_Subs = beta_3[[2]]$S

beta_4 <- split(df_80_1.1_0.5, df_80_1.1_0.5$Method)
betas_mm <- beta_4[[1]] %>% mutate(Beta_Subs = beta_4[[2]]$Beta_Hat) %>% mutate(Sd_Subs = beta_4[[2]]$S


betas_80 <- full_join(betas_ll, betas_lm) %>% full_join(betas_ml) %>% full_join(betas_mm)
```

```
## Joining, by = c("Beta_Hat", "Standard_Error", "Simulation", "Slope", "Noise", "Method", "Beta_Subs",
```

```
## Joining, by = c("Beta_Hat", "Standard_Error", "Simulation", "Slope", "Noise", "Method", "Beta_Subs",
## Joining, by = c("Beta_Hat", "Standard_Error", "Simulation", "Slope", "Noise", "Method", "Beta_Subs",
```

```
Plot_Simulation_CR <- betas_80 %>%
  ggplot(aes(x=Beta_Hat, y=Beta_Subs)) +
  facet_grid(Slope ~ Noise) +
  geom_point(alpha = 0.5) +
  theme_minimal() +
  theme(plot.caption=element_text(hjust = 0), legend.background = element_rect(color = "black",
    fill = NA, size =0.5, linetype = "solid"),
        panel.border = element_rect(colour = "black", fill=NA, size=0.5),legend.title = element_text(co
   geom_hline(data = data.frame(Slope="5% Yearly Increase", y=log(1.05)), mapping=aes(yintercept=y)) +
   geom_vline(data = data.frame(Slope="5% Yearly Increase", x=log(1.05)), mapping=aes(xintercept=x)) +
  geom_hline(data = data.frame(Slope="10% Yearly Increase", y=log(1.1)), mapping=aes(yintercept=y))+
   geom_vline(data = data.frame(Slope="10% Yearly Increase", x=log(1.1)), mapping=aes(xintercept=x)) +
  labs(caption = "Figure A4: Plotting the estimated slopes for the substitution method and the LMMC mod
vertical and horizontal lines correspond to the true value of the slope. The data have a censoring prop
  x = 'Estimates of the LMMC Model', y = 'Estimates from Substitution')

ggsave(Plot_Simulation_CR, filename = "Plot_Simulation_CR.pdf")
```

```
## Saving 6.5 x 4.5 in image
```

```
df_CR_simu <- full_join(df_80_1.05_0.05, df_80_1.05_0.5) %>%
  full_join(df_80_1.1_0.05) %>%
  full_join(df_80_1.1_0.5)
```

```
## Joining, by = c("Beta_Hat", "Standard_Error", "Simulation", "Slope", "Noise", "Method")

## Joining, by = c("Beta_Hat", "Standard_Error", "Simulation", "Slope", "Noise", "Method")
## Joining, by = c("Beta_Hat", "Standard_Error", "Simulation", "Slope", "Noise", "Method")
```

```
df_bias <- df_CR_simu %>% group_by(Method, Slope, Noise) %>%
  mutate(bias = (Slope-Beta_Hat)^2) %>%
  summarise(bias = mean(bias)) %>%
  ungroup

df_var <- df_CR_simu %>% group_by(Method, Slope, Noise) %>%
  mutate(bias = (Slope-Beta_Hat)^2) %>%
  summarise(Variance = sd(Beta_Hat)^2)%>%
  ungroup

df_coverage <- df_CR_simu %>%
  mutate(CI_Low = Beta_Hat - 1.96*Standard_Error) %>%
  mutate(CI_Upp = Beta_Hat +1.96*Standard_Error) %>%
  mutate(Coverage = ifelse(Slope >= CI_Low & Slope <= CI_Upp, 1, 0)) %>%
  group_by(Method, Slope, Noise) %>%
  summarise(Coverage = mean(Coverage))%>%
  ungroup

CR_sim_df_full <- df_bias %>%
  mutate(Variance = df_var$Variance) %>%
```

Table 1: Summary statistics of simulations at 80% censored data

Slope	Sd	Method	$(\hat{\beta} - \beta)^2$	Coverage	$\text{Var}(\hat{\beta})$	MSE	$\text{Se}((\hat{\beta} - \beta)^2)$
5 % Yearly Increase	0.05	lmec	0e+00	0.93	0e+00	0e+00	0.0000
		Substitution	1e-04	0.38	0e+00	1e-04	0.0000
	0.50	lmec	4e-04	0.95	5e-04	9e-04	0.0000
		Substitution	3e-03	0.00	1e-04	3e-03	0.0000
10 % Yearly Increase	0.05	lmec	0e+00	0.94	0e+00	0e+00	0.0000
		Substitution	2e-04	0.01	0e+00	2e-04	0.0000
	0.50	lmec	4e-04	0.95	4e-04	8e-04	0.0000
		Substitution	9e-04	0.07	1e-04	1e-03	0.0000

```
  mutate(Coverage = df_coverage$Coverage) %>%
  mutate(MSE = Variance + bias) %>%
  mutate("Standard Error" = sqrt(Variance/(100))*2*abs(bias)) %>%
  arrange(desc(Slope), Noise)

CR_sim_df_full[,-1] <- round(CR_sim_df_full[,-1], digits = 4)

CR_sim_df_full <- CR_sim_df_full %>% select(Slope ,Noise, Method, bias,Coverage, Variance , MSE, `Standa
  mutate(`Standard Error` = "0.0000")

colnames(CR_sim_df_full) <- c("Slope", "Sd", "Method", "$(\\hat{\\beta} - \\beta)^2$", "Coverage", "Var

CR_sim_table <- knitr::kable(CR_sim_df_full, escape = FALSE, format = "latex", booktabs = T, align = "l
column_spec(1, bold=T) %>%
kableExtra::collapse_rows(columns = 1:2, latex_hline = "major", valign = "middle")



save(CR_sim_table, file = "CR_sim_table.Rda")

CR_sim_table
```