Practical assignment trial-based cost-effectiveness analysis - Solutions

X. Pouwels

2021-09-02

```r
rm(list  = ls()) # clear environment
options(scipen = 999) # Disable scientific notations
library(tidyverse)


## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --


## v ggplot2 3.3.3     v purrr   0.3.4
## v tibble  3.0.5     v dplyr   1.0.3
## v tidyr   1.1.2     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.1


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

library(knitr)
library(boot)

load(file = "trial_based_CEA.RData")
```

Assignment and solutions

1. Calculate the mean operating time, risk of complications, length of stay at ICU and MCU, risk of death, and quality of life after procedure, for both trial arms. *Hint: use the functions **group_by** and **summarise** from the **tidyverse** package to do so.*

```r
res_health <- df %>%
  group_by(Procedure) %>%
  summarise(mean_OR_time = mean(OR_time),
            mean_Major_compl = mean(Major_compl),
            mean_Minor_compl = mean(Minor_compl),
            mean_Length_icu  = mean(Length_icu),
            mean_Length_mcu  = mean(Length_mcu),
            mean_p_death = mean(Death),
            mean_Qol = mean(Qol))
```

```r
kable(t(res_health),
      digits = 2,
      caption = "Mean health outcome per group")
```

Table 1: Mean health outcome per group

	1	2
Procedure	1.00	2.00
mean_OR_time	139.56	122.22
mean_Major_compl	0.26	0.14
mean_Minor_compl	0.31	0.50
mean_Length_icu	1.33	1.20
mean_Length_mcu	13.67	15.36
mean_p_death	0.18	0.13
mean_Qol	0.41	0.48

1.a. Which treatment provides the best health outcomes?

Answer: The new intervention (robot-assisted surgery) provides better health outcomes because it results in a higher utility than usual care (0.480181846203199 versus 0.410040233913325).

1.b. Is robot-assisted surgery an improvement upon usual care with respect to all outcomes?

Answer: No, robot-assisted surgery leads to an increase in the number of minor complications (0.5 versus 0.31) and jthe length of stay at the medium care unit (15.36 versus 13.67).

2. Create three new variables in `df` called Cost_procedure, Cost_compl, Cost_hosp, which respectively contain the costs related to the procedure (operatime time + procedure), the costs related to complications, and the costs related to hospitalization for each participant. To do so, use the objects beginning with `c_`.

Hint: Multiply the column of the dataframe df by the corresponding objects.

```r
# Calculate total costs for each category per participant
df$Cost_procedure <- ifelse(df$Procedure == 1, c_Open + df$OR_time / 60 * c_OK_hour, c_Robot + df$OR_tim
df$Cost_compl     <- df$Major_compl * c_Major_Compl + df$Minor_compl * c_Minor_Compl
df$Cost_hosp      <- df$Length_icu * c_ICU_day + df$Length_mcu * c_MCU_day

kable(rbind(df[,2:ncol(df)] %>%
        filter(Procedure == 1) %>%
        head(5),
        df[,2:ncol(df)] %>%
        filter(Procedure == 2) %>%
        head(5)
        ),
      digits = 2,
      caption = "First rows of cost per participant")
```

Table 2: First rows of cost per participant

Procedure	OR_time	Major_compl	Minor_compl	Length_icu	Length_mcu	Death	Qol	Cost_procedure	Cost_compl	Cost_hosp
1	129.63	0	0	6	8	0	0.55	3080.29	0	11900
1	146.15	0	1	0	13	0	0.84	3217.94	1500	7150
1	154.73	1	0	0	19	0	0.09	3289.41	10000	10450
1	133.17	0	0	6	16	0	0.79	3109.78	0	16300
1	133.98	0	1	0	11	0	0.55	3116.51	1500	6050

Procedure	OR_time	Major_compl	Minor_compl	length_ic	length_m	Death	Qol	Cost_procedure	Cost_compl	Cost_hosp
2	116.34	0	1	0	14	1	0.00	3969.48	1500	7700
2	154.09	0	0	0	20	0	0.15	4284.09	0	11000
2	150.41	0	0	0	15	1	0.00	4253.45	0	8250
2	127.05	1	1	5	15	0	0.99	4058.75	11500	14500
2	168.03	0	1	0	21	0	1.00	4400.29	1500	11550

3. Create a new variable in `df` called Total_costs, which contains the total costs per participant. Calculate the mean costs per group. Based on this information:

```
# Calculate total costs per participant
df$Total_costs  <- df$Cost_procedure + df$Cost_compl + df$Cost_hosp

res_costs <- df %>%
  group_by(Procedure) %>%
  summarise(mean_tot_costs = mean(Total_costs))

kable(res_costs,
      digits = 0,
      caption = "Mean total costs per group")
```

Table 3: Mean total costs per group

Procedure	mean_tot_costs
1	15409
2	16117

3.a. Which treatment is cheaper on average, given the observed trial data?
Answer: Current care is cheaper on average.
3.b. What do you think is the probability that the new intervention is cheaper?
Answer: This question is hard to answer, all we know is that some patients in the intervention arm have lower cost than some patients in the usual care arm. On average the intervention is more expensive so we would guess that this probability is less than 50% but we don't know yet.
3.c. What do you think is the probability that the new intervention leads to health gain?
Answer: This question is also hard to answer. On average the intervention is leads to a substantially higher utility than usual care, so we would guess that this probability is greater than 50% but we don't know yet.
4. Calculate the difference in mean quality of life and mean total costs between the new intervention and usual care (i.e. the incremental effects and incremental costs). Calculate the ICER using this information. *Hint: use the results you obtained in previous steps.*

```
Inc_qol <- unname(res_health[2, "mean_Qol"] - res_health[1, "mean_Qol"])
Inc_costs <- unname(res_costs[2, "mean_tot_costs"] - res_costs[1, "mean_tot_costs"])
ICER <- as.numeric(as.character(Inc_costs)) / as.numeric(as.character(Inc_qol))

kable(cbind('Incremental effects' = Inc_qol,
            'Incremental costs' = Inc_costs,
            ICER = ICER),
      digits = 2,
      caption = "Mean incremental results and ICER") # show results
```

Table 4: Mean incremental results and ICER

Incremental effects	Incremental costs	ICER
0.07	707.54	10087.3

5. Save your dataframe `df` as a .csv file. Ensure that your total effects per participants are in the `Qol` variable and the total costs in the `Total_costs` variable.

```
write.csv(df, "df_trial_cea.csv")
```

6. Open the R shiny app using the following command in R and follow the instructions.

```
runGitHub("Teaching", "Xa4P", subdir = "Basics/shiny_app_cea/", ref = "main")
```

7. Run the bootstrap analysis by pushing the "Run bootstrap analysis" in the Shiny app, after having uploaded the file you just created (if this doesn't work, you can use the `df_trial_cea.csv` file that is provided in the map of this assignment. To do so, choose a number of bootstrap samples (at least 1,000 and up to 20,000) and a random seed number (between 1 and infinity). Wait a couple of seconds that the analysis runs. The results will automatically appear on the screen. The results that you see are related to the uncertainty that remains after observing this set of trial data. In addition to the bootstrap results, the mean results of the sample data (that you have calculated earlier), are also provided.

7.a. Check the average estimates from the bootstrap samples. Do these results correspond to the point estimates from the observed data?

Answer: The results of the bootstrapping match the original point estimates rather well. However, the additional cost may deviate a bit and therefore the ICER from bootstrapping can be a few (hundred) euros/QALY more or less than the ICER from the original data.

7.b. Check the confidence intervals for the additional health effects and for the additional costs of the intervention. Can you explain why they are so wide?

Answer: The CIs are very wide because the utility and total cost vary substantially between patients: utilities range from 0 to 1 and costs range from 6527euros to 33422 euros.

7.c. Check the probabilities of a cost-effectiveness outcome in each one of the four quadrants. How well did you guess the probability that the new intervention is cheaper, or improves health (questions 3b and c)?

Answer: Own reflection.

7.d. No confidence interval is given for the ICER, can you explain why such a confidence interval is not informative by looking at the incremental cost-effectiveness plane?

Answer: The incremental CE plane contains points (results) in all four quadrants of the plane. This means that the ICER can be both positive (in which case we are not sure if the intervention is better and more expensive or worse and cheaper) and negative (in which case we are not sure if the intervention is better and cheaper or worse and more expensive). In this situation, any confidence interval for the ICER would be ambiguous and the incremental CE plane itself is more informative, along with the probabilities of results falling into each of the four quadrants.

7.e. Rerun the bootstrap procedure 3 times using different seed numbers. Do the results of the bootstrap analysis change substantially? What would we need to do to reduce this variation in outcomes?

Answer: The ICER, and the 95%CI for the cost difference may change more than we would like to see. This variation could be reduced by increasing the number of bootstrap samples, for example, from 5,000 to 20,000.