Sampling Strategy/Fishermen

Anna Calle

5/2/2020

Measurement Goal & Measure

Goal: Evaluate the use of unsustainable fishing practices in the community

"How frequently have you observed the use of [list of unsustainable fishing practices] in the community during the last 6 months?"

a. Never

b. Rarely

c. Sometimes

d. Often

e. Always

Source & Rationale & Unit

Source: New survey carried out by an independent survey firm. It will target individuals that applied for fishing permits through the Oregon Department of Fish & Wildlife

Rationale: Directly measures the outcome of creating regulations and management structure. Fishermen are likely to observe the fishing practices of other members of their community. The survey avoids asking sensitive questions directly, increasing the accuracy of the information obtained.

Unit: individual survey respondent

Responsibility & Frequency

Responsibility: Independent survey firm to ensure confidentiality and credibility of results

Frequency: Biannual survey, conducted during fishing season and outside fishing season. The behaviour of fishermen might be influenced by the availability of fish. A biannual survey will give fishermen enough time to react to changes in regulations.

Declaring the population

The population is composed of fishermen who belong to the The Association of Northwest Steelheaders (ANWS) and fishermen who do not belong to it. For this sampling procedure it was assumed that 50% of fishermen belong to ANWS and 50% do not. Because ANWS aims to protect fisheries and because it is a partner of the Freshwater Trust, we assumed memebers of ANWS will have a higher response rate and are more likely to report the use of unsustainable fishing practices

Target Population & Challenges

Target Population: Fishermen

Challenge of drawing a representative sample: ANWS members are more likely to respond to the survey than non-members because they are invested in protecting fisheries, creating a response bias.

Sampling procedure: Stratified sampling of ANWS members and non-members

Declaration of hypothetical population

```
population <- declare_population(fabricate(
  N = 500,
  members = draw_binary(prob = 0.6, N), # 50% are members
  freq = draw_ordered(
    x = rnorm(N, mean = 2 + 1 * members),
    # centered at 2 (rarely) for non-members and at 3 (sometimes) for members
    breaks = c(1.5, 2.5, 3.5, 4.5),
    break_labels = c(1:5)), #  interval representation
  freq_num = as.numeric(freq) # changed frequency to numeric
  )
)

pop <- population()

# population table (1 = Never, 2 = Rarely, 3 = Sometimes, 4 = Often, 5 = Always)
pop_table <- kable(table(pop$freq, pop$members), row.names = T) %>% add_header_above(c("Frequency"=1, "

pop_table
```

Frequency	Members	
	0	1
1	66	16
2	75	69
3	47	126
4	8	80
5	1	12

```
# we are interested in the mean satisfaction
my_estimand <- declare_estimands(mean(freq_num),
                                 label = "Ybar")
```

Declaration of reporting probabilities

```
# assign probability of responding to survey (20% for non-members, 40% for members)
reporting <- declare_assignment(blocks=members,
                  assignment_variable = "R",
                  block_prob=c(0.2,0.4))
```

Declaration of sampling procedure

```
# sample 100 members and 200 non-members
sampling <- declare_sampling(strata=members,
                             strata_n=c(100, 200))
```

Declaration of estimator

```
strata_weighted_mean <- function(data){
  data.frame(
  estimator_label = "strata_w_mean",
  estimand_label = "Ybar",
  n = nrow(data),
```

```
  stringsAsFactors = FALSE,

  estimate = data %>% filter(R==1) %>%
    group_by(members) %>%
    summarise(mean=mean(freq_num)) %>%
    mutate(prop=c(0.5,0.5)) %>%
    mutate(sub.mean=mean*prop) %>% pull(sub.mean) %>%
    sum())
}
```

Diagnosis

```
answer <- declare_estimator(
  handler = tidy_estimator(strata_weighted_mean), # weighted mean
  estimand = my_estimand) # mean

design <- population + my_estimand + reporting +
          sampling + answer
diagnosis <- diagnose_design(design, sims = 100) # simulate design 100 times

diagnosis$diagnosands_df[,c(4,5,12,14)] %>%
  kable()
```

bias	se(bias)	mean_estimate	sd_estimate
-0.0922365	0.0124319	2.533844	0.1194444

```
# As expected, mean is between 2 (rarely) and 3 (sometimes)
```