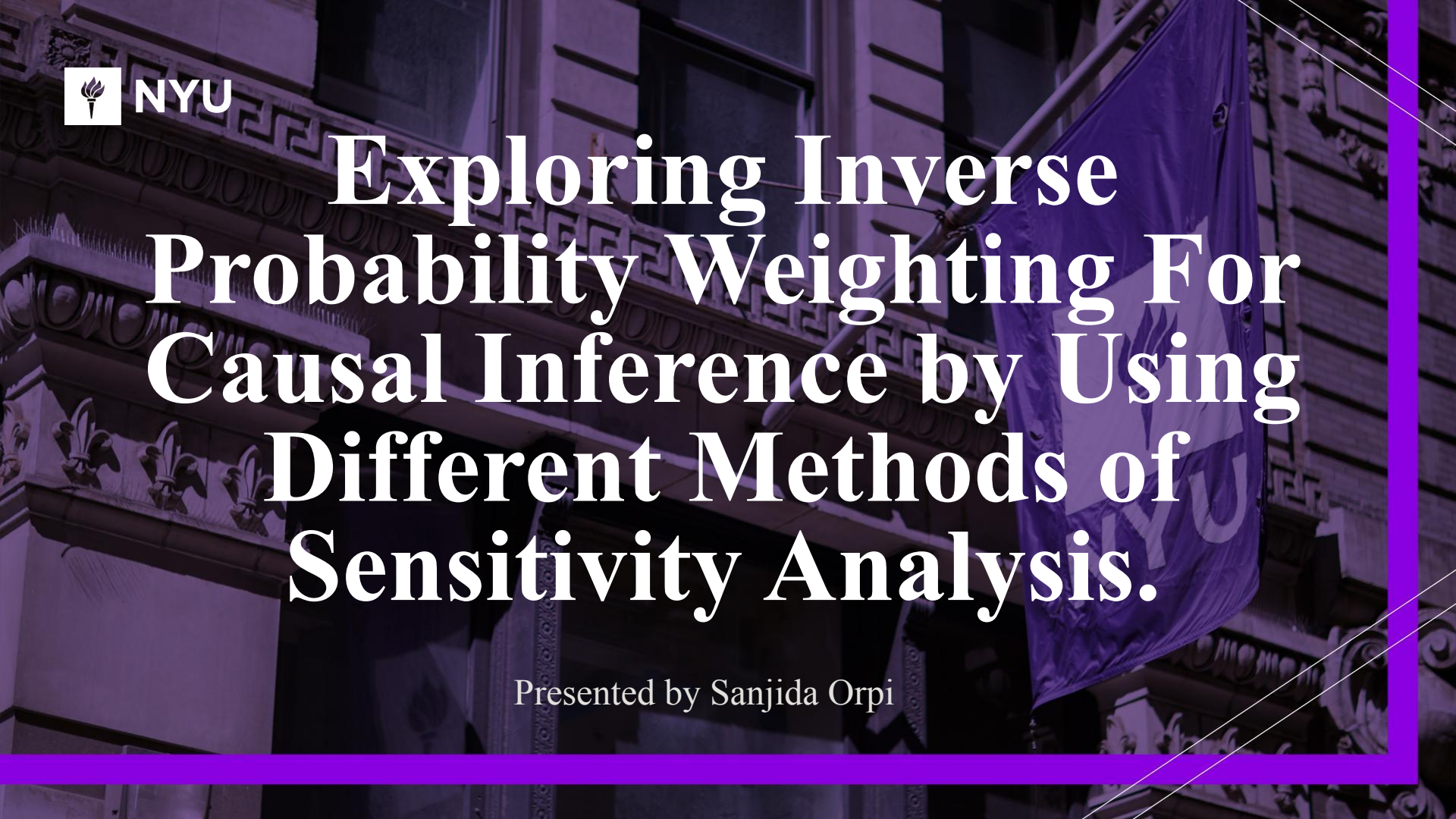


Exploring Inverse Probability Weighting For Causal Inference by Using Different Methods of Sensitivity Analysis.

Presented by Sanjida Orpi

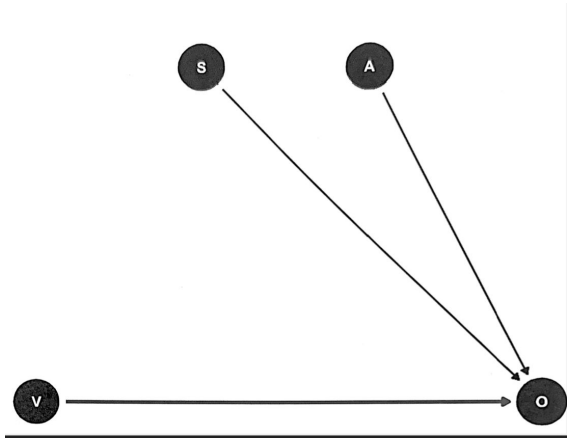
Agenda

- Project background
- Goal
- Sensitivity Analysis Part I
- Sensitivity Analysis Part II
- Conclusion

How getting a new vaccine affects whether or not a person gets a disease in the future?

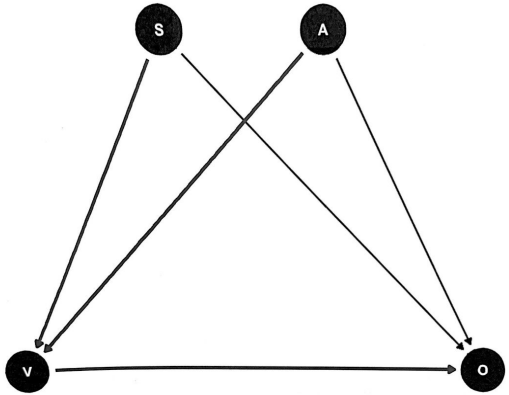
1. Simulating the Randomized Control Trial (RCT)

Causal Relationships

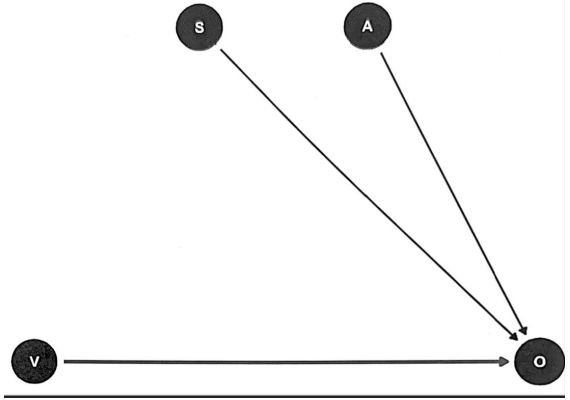


Goal: Implement Sensitivity Analysis

Observational Analysis



RCT



NYU Center for Urban Science + Progress

Dynamical Systems Laboratory

Sensitivity Analysis - Classic Nloptr

1. We weight the outcome by the inverse probability of getting a vaccine.
2. Optimization of weights - Sensitivity Analysis for Hidden Bias → using constrained optimization

$$CE(0) = \frac{\sum_{i=1}^{N} \mathbf{1}(v_i = 0) \cdot y_i \cdot w_i}{\sum_{i=1}^{N} \mathbf{1}(v_i = 0)}$$

$$w_i = \frac{a_i}{\hat{p}(v_i | x_i)}$$

```r
#Observational Analysis

dgf_obs <- function(){
  #1.Get sec and age
  # 1 = Female
  # 2 = Male
  sex <- sample(c(1, 0), 1, prob = c(0.505, 0.495)) #1 means we want to sample one thing
  age <- 100 * rbeta(1, shape1=2, shape2=5)

  #2. Treatment depends on age and sex
  #--TREATMENT---
  # 1 = "EHD"
  # 2 = "CHD"

  mu_EHD <- (-0.2 * sex) + (0.025* age) + rnorm(1, 0, 0.1)
  p_EHD <- 1/(1 + exp(-mu_EHD)) #turns it into a probability
  if(p_EHD >= 0.5){
    treatment <- 1
  } else{
    treatment <- 0
  }
  #3.Observed outcome
  mu <- (0.3 * sex) + (-0.01 * age) + (0.2 * treatment) #whats the prob that you survived
  p <- 1/(1 + exp(-mu))
  if(p > 0.5){ #
    status <-  1 #"survived"
  } else {
    status <- 0 #"did not survive"
  }

  #Output
  out <- data.frame(
    sex = sex,
    age = age,
    treatment = treatment,
    status = status
  )

  return(out)

```

Sensitivity analysis

```r

#predict weather you got treatment based on age and sex | gives us a probability
p <- predict(treatment_model, newdata = data_set, type = "response")
summary(p)


propensity <- 1/p #inverse prob

summary(propensity)

#---weigthing by inverse p-----

#you got the treatment

t_1 <- as.numeric(data_set$treatment ==1)

Y_1 <- sum(data_set$status * propensity * t_1) / sum(t_1 * propensity)

#you didn't get treatment

t_0 <- as.numeric(data_set$treatment ==0)
Y_0 <- sum(data_set$status * propensity * t_0) / sum(t_0 * propensity)


Y_1 - Y_0

```

NYU Center for Urban Science + Progress

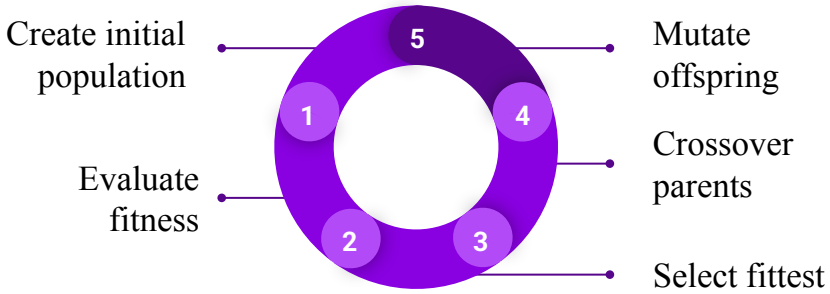
Dynamical Systems Laboratory

Sensitivity Analysis with Genetic Optimizer

- Goal
  - Conduct sensitivity analysis to give confidence to the treatment effect
- Approach
  - Implement Genetic Algorithm (GA) to characterize the impact of receiving a vaccine on health outcomes for a sample of observations from a simulated RCT
  - Bootstrap the observation sample to determine the minimum G parameter with an estimated treatment effect of 0 with a 95% confidence interval

Genetic Algorithm Overview

- Adaptive heuristic search algorithm based on evolutionary concepts of natural selection
- High dimensional search space
- Uses exploration and exploitation to find optimal solutions
- Operators
  - Selection
  - Crossover
  - Mutation



Optimizer Details

Crossover

```
crossover <- function(X, Y, crossover_rate, G){
  p <- matrix(runif(length(X), 0, 1), nrow = nrow(X), ncol = ncol(X))
  doCross <- matrix((p < crossover_rate), nrow = nrow(X), ncol = ncol(X))
  gamma <- matrix(runif(length(X), 0, 1), nrow = nrow(X), ncol = ncol(X))
  crosses <- gamma * (X - Y) + X
  Z <- X
  Z <- ifelse(doCross, crosses, Z)
  Z <- ifelse(Z > G, G, Z)
  Z <- ifelse(Z < (1/G), (1/G), Z)
  return(Z)
}
```

Parameters

```
pop_size <- 1000
crossover_rate <- 0.5
mutation_rate <- 0.01
iters <- 500
epochs <- 50
fitness_record <- rep(NA, epochs
population <- create_population(
  pop_size, boot_data, G)
```

Mutate

```
mutate <- function(offspring, mutation_rate = 0.01, G){
  p <- matrix(runif(length(offspring), 0, 1), nrow = nrow(offspring),
              ncol = ncol(offspring))

  m <- matrix(runif(length(offspring), 1/G, G), nrow = nrow(offspring),
              ncol = ncol(offspring))

  offspring <- ifelse(p < mutation_rate, m, offspring)
  return(offspring)
}
```

Results

- Sample size of 5000 observations
- At G = 1.1 the estimated treatment effect contained 0 with 95% confidence interval

```
> G
[1] 1.1
> ci
      97.5%       2.5%
-0.02542591 0.38630772
```

Conclusion

- Sensitivity Analysis with Nloptr
  - Some next steps: Run more trials using a different optimizer

- Sensitivity Analysis with Genetic Operator
  - A low G value (G = 1.1) indicates that the vaccine model is weak and is sensitive to potential hidden biases which caused the results of the experiment to be impacted by a small change


Thank you!
NYU
Center for Urban
Science + Progress
12