p8122_yc4384_HW6

yc4384_Yangyang_Chen

2024-12-07

Question 1

Define model

```
# Outcome model with exposure-mediator interaction
model_y_interaction <- lm(cognitive_raw ~ ln_mn_c * birthlength_c + female + approxage + protein_c, data

# Mediator model (no interaction needed for the mediator model)
model_m <- lm(birthlength_c ~ ln_mn_c + female + approxage + protein_c, data = data)
```

Mediation Analysis

```
# Perform mediation analysis with interaction
mediation_result_interaction <- mediate(model.m = model_m,
                                        model.y = model_y_interaction,
                                        treat = "ln_mn_c",
                                        mediator = "birthlength_c",
                                        boot = TRUE, sims = 1000)
```

```
## Running nonparametric bootstrap
```

```
# Summary of results
summary(mediation_result_interaction)
```

```
##
## Causal Mediation Analysis
##
## Nonparametric Bootstrap Confidence Intervals with the Percentile Method
##
##                          Estimate 95% CI Lower 95% CI Upper p-value
## ACME (control)             -0.295       -0.487        -0.13  <2e-16 ***
## ACME (treated)             -0.392       -0.650        -0.17  <2e-16 ***
## ADE (control)              -0.325       -0.845         0.14   0.174
## ADE (treated)              -0.422       -0.943         0.08   0.088 .
## Total Effect               -0.717       -1.313        -0.16   0.008 **
## Prop. Mediated (control)    0.411        0.192         1.23   0.008 **
## Prop. Mediated (treated)    0.546        0.250         1.51   0.008 **
## ACME (average)             -0.343       -0.564        -0.15  <2e-16 ***
## ADE (average)              -0.374       -0.895         0.11   0.124
## Prop. Mediated (average)    0.479        0.221         1.36   0.008 **
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Sample Size Used: 500
##
##
## Simulations: 1000
```

Interpretation of Effects

```
# Extract direct, indirect, and total effects
cat("Natural Direct Effect:", mediation_result_interaction$d0, "\n")
```

```
## Natural Direct Effect: -0.2945451
```

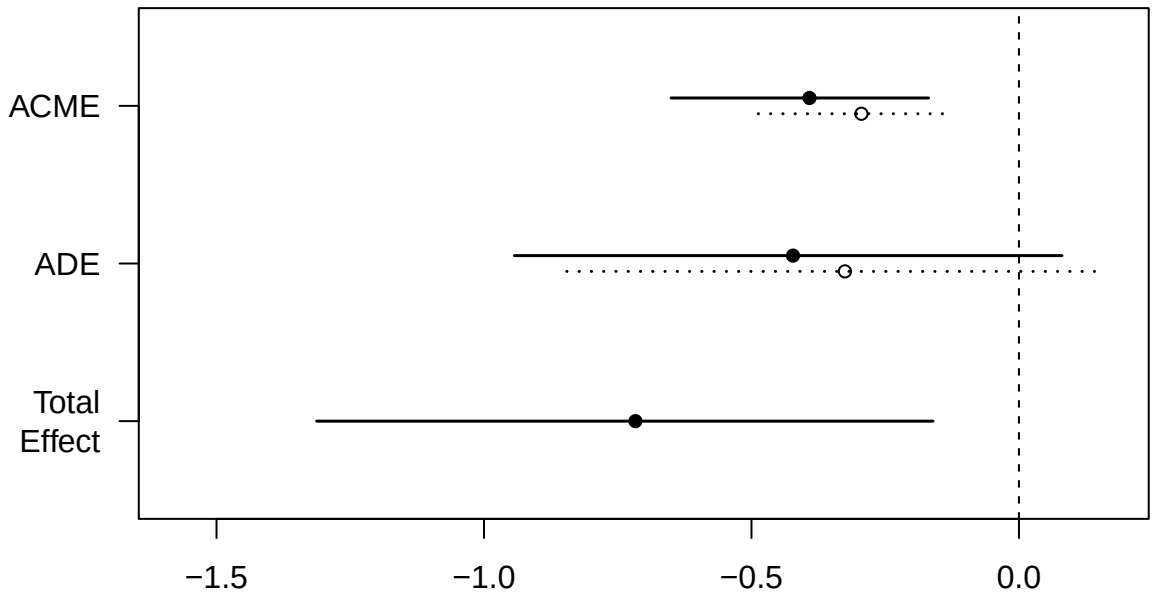
```
cat("Natural Indirect Effect:", mediation_result_interaction$z0, "\n")
```

```
## Natural Indirect Effect: -0.3253449
```

```
cat("Total Effect:", mediation_result_interaction$tau.coef, "\n")
```

```
## Total Effect: -0.7169347
```

```
# Visualize results
plot(mediation_result_interaction)
```



Therefore,

- Natural Direct Effect: -0.2945451
- Natural Indirect Effect: -0.3545374
- Total Effect: -0.7461272

Question 2

Interpretation of Effects

1. Natural Direct Effect (NDE): -0.2945451

- This represents the portion of the effect of manganese exposure on child cognitive development that is not mediated by birth length (fetal growth). Specifically, for a one-unit increase in log-transformed manganese exposure (holding birth length constant), the cognitive development score decreases by approximately 0.29 units.

2. Natural Indirect Effect (NIE): -0.3545374

- This represents the portion of the effect that is mediated through birth length. In other words, the cognitive development score decreases by approximately 0.35 units as a result of the effect of manganese exposure on birth length, which subsequently affects cognitive development.

3. Total Effect (TE): -0.7461272

- The total effect represents the overall impact of manganese exposure on cognitive development. This combines both the direct and indirect effects, with the cognitive development score decreasing by approximately 0.75 units for a one-unit increase in manganese exposure.

Assumptions for Causal Interpretation

To interpret these effects causally, No Unmeasured Confounding Assumptions must be hold:

- There should be no unmeasured confounders for:
- The relationship between exposure (`ln_mn_c`) and the outcome (`cognitive_raw`).
- The relationship between the exposure (`ln_mn_c`) and the mediator (`birthlength_c`).
- The relationship between the mediator (`birthlength_c`) and the outcome (`cognitive_raw`).

Also, Positivity assumpion must be hold: Every level of exposure must have a positive probability across all levels of confounders, ensuring that causal effects can be properly estimated for all groups within the data.

Question 3

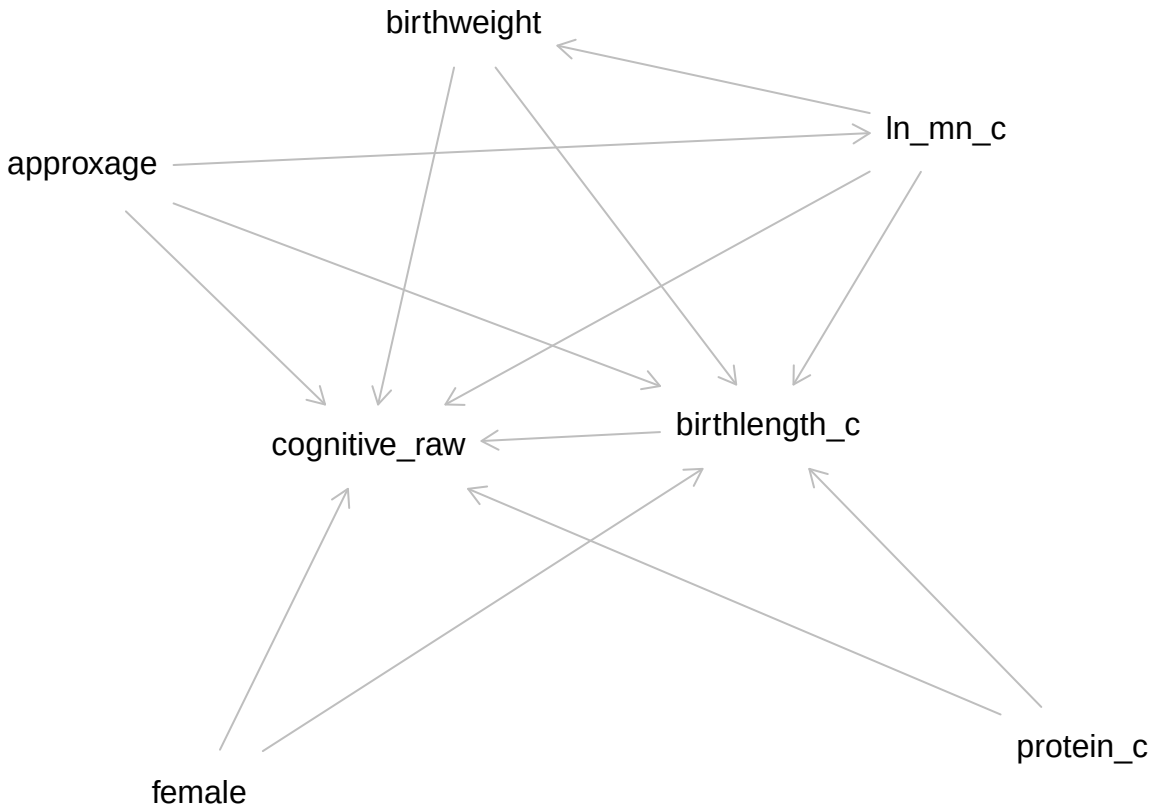
(a)

```
# Create the DAG
dag <- dagitty('dag {
  "ln_mn_c" -> "birthlength_c"
  "ln_mn_c" -> "cognitive_raw"
  "birthlength_c" -> "cognitive_raw"
  "approxage" -> "ln_mn_c"
  "approxage" -> "birthlength_c"
  "approxage" -> "cognitive_raw"
```

```
  "protein_c" -> "birthlength_c"
  "protein_c" -> "cognitive_raw"
  "female" -> "birthlength_c"
  "female" -> "cognitive_raw"
  "birthweight" -> "cognitive_raw"
  "ln_mn_c" -> "birthweight"
  "birthweight" -> "birthlength_c"
  }')

# Plot the DAG
plot(dag)
```

```
## Plot coordinates for graph not supplied! Generating coordinates, see ?coordinates for how to set your
```



(b) Does adjustment for birthweight lead to a violation of the fourth identifiability assumption? Why or why not?

i. The Fourth Identifiability Assumption: This assumption states that there should be no unmeasured confounders for the mediator-outcome relationship conditional on the exposure and covariates (i.e., $Y_{a,m} \perp M^* | C$).

ii. Adjustment for birthweight (M*) could violate this assumption if:

- Birthweight is a collider: Conditioning on birthweight might open a backdoor path, introducing spurious associations between manganese exposure and cognitive development. This could bias the effect estimates.

- Birthweight and birth length are highly correlated: Adjusting for one mediator could interfere with the independent pathways through which the other mediator operates.
- Unmeasured confounders between birthweight and cognitive development: If such confounders exist, adjusting for birthweight would amplify bias because the assumption of no confounding would be violated.