

Model Comparison

Baoyi Shi

1 Case 1: Continuous Outcome and Continuous Mediator

1.1 Case 1-1: Continuous Outcome and Single Continuous Mediator Without Exposure-mediator Interaction

1.1.1 Data simulation

1.1.1.1 Simulation Procedures

1. Simulate the exposure variable A from $\text{Binom}(1, P(A=1))$.
2. Simulate the covariate C from $N(\mu_C, \sigma_C^2)$.
3. Simulate the mediator M from $N((\beta_0 + \beta_1 * A + \beta_2 * C), \sigma_M^2)$.
4. Simulate the outcome Y from $N(\theta_0 + \theta_1 A + \theta_2 M + \theta_4 C, \sigma_Y^2)$.

1.1.1.2 True Parameters

Table 1: True Model Parameters for Data Simulation

Sample Size	θ_0	θ_1	θ_2	θ_4	β_0	β_1	β_2	P(A=1)	μ_C	σ_C	σ_M	σ_Y
10000	-5	0.8	1.8	0.1	-0.25	0.5	0.2	0.4	1	1	0.1	0.2

1.1.1.3 True Models

True model for the mediator:

$$E[M|a, c] = \beta_0 + \beta_1 a + \beta_2 c$$

True model for the outcome:

$$E[Y|a, m^*, c] = \theta_0 + \theta_1 a + \theta_2 m^* + \theta_4 c$$

1.1.2 Causal Effects and Standard Errors Estimated By the Regression-based Approach

1.1.2.1 Closed-form Parameter Function Estimation and Delta Method Inference

```
causal_mediation(data = df_noint, outcome = "contY_contM_noint", exposure = 'A',
  exposure.type = "binary",
  mediator = 'M_cont', covariates.pre = "C", covariates.post = NULL,
  covariates.post.type = NULL, cval = list(mean(df_noint$C)),
  EMint = FALSE, MMint = NULL, EMMint = NULL, EMint.terms = NULL,
  MMint.terms = NULL, EMMint.terms = NULL,
  yreg = "linear", mreg = "linear", mval = list(0),
  a_star = 0, a = 1,
  est.method = "paramfunc", inf.method = "delta", model = "rb",
  nboot = 200, conf = 0.95, nrep = 5)
```

```
## $effect_estimate
##      cde      pnde      tnde      pnle      tnle      te
## 0.7961788 0.7961788 0.7961788 0.8973026 0.8973026 1.6934814
##      pm      intref      intmed      pie      cde_prop intref_prop
## 0.3604117 0.0000000 0.0000000 0.8973026 0.4701432 0.0000000
## intmed_prop      pie_prop overall_pm overall_int overall_pe
## 0.0000000 0.5298568 0.5298568 0.0000000 0.5298568
##
## $effect_se
## [1] 0.010783525 0.010783525 0.010783525 0.010633841 0.010633841
## [6] 0.005535765 0.005656045 0.000000000 0.000000000 0.010633841
## [11] 0.006112265 0.000000000 0.000000000 0.006112265 0.006112265
## [16] 0.000000000 0.006112265
```

1.1.2.2 Direct Imputation Estimation and Bootstrap Inference

```
causal_mediation(data = df_noint, outcome = "contY_contM_noint", exposure = 'A',
  exposure.type = "binary",
  mediator = 'M_cont', covariates.pre = "C", covariates.post = NULL,
  covariates.post.type = NULL, cval = list(),
  EMint = FALSE, MMint = NULL, EMMint = NULL, EMint.terms = NULL,
  MMint.terms = NULL, EMMint.terms = NULL,
  yreg = "linear", mreg = "linear", mval = list(0),
  a_star = 0, a = 1,
  est.method = "imputation", inf.method = "bootstrap", model = "rb",
  nboot = 200, conf = 0.95, nrep = 5)
```

```
## $effect_estimate
##      cde      pnle      tnle      pnle      tnle      te
## 0.7961788 0.7961788 0.7961788 0.8973026 0.8973026 1.6934814
##      pm      intref      intmed      pie      cde_prop intref_prop
## 0.3604117 0.0000000 0.0000000 0.8973026 0.4701432 0.0000000
## intmed_prop      pie_prop overall_pm overall_int overall_pe
## 0.0000000 0.5298568 0.5298568 0.0000000 0.5298568
##
## $effect_se
## [1] 1.159437e-02 1.159437e-02 1.159437e-02 1.175842e-02 1.175842e-02
## [6] 5.354817e-03 6.201721e-03 1.089613e-16 3.304044e-16 1.175842e-02
## [11] 6.705509e-03 6.435119e-17 1.951199e-16 6.705509e-03 6.705509e-03
## [16] 1.917453e-16 6.705509e-03
```

1.1.3 Causal Effects and Standard Errors Estimated By the Weighting-based Approach

```
causal_mediation(data = df_noint, outcome = "contY_contM_noint", exposure = 'A',
  exposure.type = "binary",
  mediator = 'M_cont', covariates.pre = "C", covariates.post = NULL,
  covariates.post.type = NULL, cval = list(),
  EMint = FALSE, MMint = NULL, EMMint = NULL, EMint.terms = NULL,
  MMint.terms = NULL, EMMint.terms = NULL,
  yreg = "linear", mreg = "linear", mval = list(0),
  a_star = 0, a = 1,
  est.method = "imputation", inf.method = "bootstrap", model = "wb",
  nboot = 200, conf = 0.95, nrep = 5)
```

```
## $effect_estimate
##      cde      pnde      tnde      pnle      tnle
## 7.961791e-01 7.961524e-01 7.962398e-01 8.972666e-01 8.973541e-01
##      te      pm      intref      intmed      pie
## 1.693506e+00 3.604326e-01 -2.679502e-05 8.748172e-05 8.972666e-01
##      cde_prop      intref_prop      intmed_prop      pie_prop      overall_pm
## 4.701365e-01 -1.582221e-05 5.165716e-05 5.298277e-01 5.298793e-01
##      overall_int      overall_pe
## 3.583494e-05 5.298635e-01
##
## $effect_se
## [1] 1.060857e-02 1.061184e-02 1.060314e-02 1.078227e-02 1.078663e-02
## [6] 5.392118e-03 5.652018e-03 4.026743e-05 1.233519e-04 1.078227e-02
## [11] 6.110191e-03 2.379070e-05 7.289263e-05 6.107655e-03 6.111726e-03
## [16] 5.099395e-05 6.110191e-03
```

1.1.4 Causal Effects and Standard Errors Estimated By the Marginal Structural Model

```
causal_mediation(data = df_noint, outcome = "contY_contM_noint", exposure = "A",
  exposure.type = "binary",
  mediator = 'M_cont', covariates.pre = "C", covariates.post = NULL,
  covariates.post.type = NULL, cval = list(mean(df_noint$C)),
  EMint = FALSE, MMint = NULL, EMMint = NULL, EMint.terms = NULL,
  MMint.terms = NULL, EMMint.terms = NULL,
  yreg = "linear", mreg = "linear", mval = list(0),
  a_star = 0, a = 1,
  est.method = "imputation", inf.method = "bootstrap", model = "msm",
  nboot = 200, conf = 0.95, nrep = 5)
```

```
## $effect_estimate
##      cde      pnle      tnle      pnle      tnle      te
## 0.83510734 0.91921167 0.91921167 0.82087102 0.82087102 1.74008269
##      pm      intref      intmed      pie      cde_prop      intref_prop
## 0.30868001 0.08410433 0.00000000 0.82087102 0.47992394 0.04833353
##      intmed_prop      pie_prop      overall_pm      overall_int      overall_pe
## 0.00000000 0.47174253 0.47174253 0.04833353 0.52007606
##
## $effect_se
## [1] 6.684764e-02 5.364900e-02 5.364900e-02 3.829664e-02 3.829664e-02
## [6] 2.024193e-02 2.214917e-02 5.230143e-02 3.160554e-16 3.829664e-02
## [11] 3.559182e-02 2.990165e-02 1.815231e-16 2.587883e-02 2.587883e-02
## [16] 2.990165e-02 3.559182e-02
```

1.1.5 Causal Effects and Standard Errors Estimated By the Inverse Odds-ratio Weighting Approach

```
causal_mediation(data = df_noint, outcome = "contY_contM_noint", exposure = 'A',
  exposure.type = "binary",
  mediator = 'M_cont', covariates.pre = "C", covariates.post = NULL,
  covariates.post.type = NULL, cval = list(),
  EMint = FALSE, MMint = NULL, EMMint = NULL, EMint.terms = NULL,
  MMint.terms = NULL, EMMint.terms = NULL,
```

```

yreg = "linear", mreg = "linear", mval = list(0),
a_star = 0, a = 1,
est.method = "paramfunc", inf.method = "bootstrap", model = "iorw",
nboot = 200, conf = 0.95, nrep = 5)

## $effect_estimate
##      tot      dir      ind
## 1.6934814 1.0988576 0.5946238
##
## $effect_se
## [1] 0.005549882 0.130353443 0.129989828

```

1.1.6 Causal Effects and Standard Errors Estimated By the G-formula Approach

```

causal_mediation(data = df_noint, outcome = "contY_contM_noint", exposure = 'A',
  exposure.type = "binary",
  mediator = 'M_cont', covariates.pre = "C", covariates.post = NULL,
  covariates.post.type = NULL, cval = list(),
  EMint = FALSE, MMint = NULL, EMMint = NULL, EMint.terms = NULL,
  MMint.terms = NULL, EMMint.terms = NULL,
  yreg = "linear", mreg = "linear", mval = list(0),
  a_star = 0, a = 1,
  est.method = "imputation", inf.method = "bootstrap", model = "g-formula",
  nboot = 200, conf = 0.95, nrep = 5)

## $effect_estimate
##      cde      pnide      tnide      pnide      tnide      te
## 0.7961788 0.7961788 0.7961788 0.8973026 0.8973026 1.6934814
##      pm      intref      intmed      pie      cde_prop intref_prop
## 0.3604117 0.0000000 0.0000000 0.8973026 0.4701432 0.0000000
## intmed_prop      pie_prop overall_pm overall_int overall_pe
## 0.0000000 0.5298568 0.5298568 0.0000000 0.5298568
##
## $effect_se
## [1] 1.143864e-02 1.143864e-02 1.143864e-02 1.113552e-02 1.113552e-02
## [6] 5.661457e-03 5.992647e-03 8.859440e-17 3.286978e-16 1.113552e-02
## [11] 6.462142e-03 5.223334e-17 1.941849e-16 6.462142e-03 6.462142e-03
## [16] 1.946473e-16 6.462142e-03

```

1.1.7 Causal Effects and Standard Errors Estimated By the Natural Effect Model

```

causal_mediation(data = df_noint, outcome = "contY_contM_noint", exposure = 'A',
  exposure.type = "binary",
  mediator = 'M_cont', covariates.pre = "C", covariates.post = NULL,
  covariates.post.type = NULL, cval = NULL,
  EMint = FALSE, MMint = NULL, EMMint = NULL, EMint.terms = NULL,
  MMint.terms = NULL, EMMint.terms = NULL,
  yreg = "linear", mreg = "linear", mval = list(0),
  a_star = 0, a = 1,
  est.method = NULL, inf.method = NULL, model = "ne",
  nboot = 200, conf = 0.95, nrep = 5)

```

```
##              Estimate Std. Error
## natural direct effect 0.7961788 0.010947557
## natural indirect effect 0.8973026 0.010766483
## total effect          1.6934814 0.005534381
```

1.2 Case 1-2: Continuous Outcome and Single Continuous Mediator With Exposure-mediator Interaction

1.2.1 Data simulation

1.2.1.1 Simulation Procedures

1. Simulate the exposure variable A from $\text{Binom}(1, P(A=1))$.
2. Simulate the covariate C from $N(\mu_C, \sigma_C^2)$.
3. Simulate the mediator M from $N((\beta_0 + \beta_1 * A + \beta_2 * C), \sigma_M^2)$.
4. Simulate the outcome Y from $N(\theta_0 + \theta_1 A + \theta_2 M + \theta_3 AM + \theta_4 C, \sigma_Y^2)$.

1.2.1.2 True Parameters

Table 2: True Model Parameters for Data Simulation

Sample Size	θ_0	θ_1	θ_2	θ_3	θ_4	β_0	β_1	β_2	P(A=1)	μ_C	σ_C	σ_M	σ_Y
10000	-5	0.8	1.8	0.2	0.1	-0.25	0.5	0.2	0.4	1	1	0.1	0.2

1.2.1.3 True Models

True model for the mediator:

$$E[M|a, c] = \beta_0 + \beta_1 a + \beta_2 c$$

True model for the outcome:

$$E[Y|a, m^*, c] = \theta_0 + \theta_1 a + \theta_2 m^* + \theta_3 a m^* + \theta_4 c$$

1.2.2 Causal Effects and Standard Errors Estimated By the Regression-based Approach

1.2.2.1 Closed-form Parameter Function Estimation and Delta Method Inference

```
causal_mediation(data = df_int, outcome = "contY_contM_int", exposure = 'A',
  exposure.type = "binary",
  mediator = 'M_cont', covariates.pre = "C", covariates.post = NULL,
  covariates.post.type = NULL, cval = list(mean(df_noint$C)),
  EMint = TRUE, MMint = NULL, EMMint = NULL, EMint.terms = NULL,
  MMint.terms = NULL, EMMint.terms = NULL,
  yreg = "linear", mreg = "linear", mval = list(0),
  a_star = 0, a = 1,
  est.method = "paramfunc", inf.method = "delta", model = "rb",
  nboot = 200, conf = 0.95, nrep = 5)
```

```
## $effect_estimate
##          cde          pnde          tnde          pnle          tnle
## 0.796419787 0.785133639 0.894126874 0.894855413 1.003848648
##          te          pm          intref          intmed          pie
## 1.788982287 0.389978026 -0.011286148 0.108993235 0.894855413
```

```
##      cde_prop  intref_prop  intmed_prop    pie_prop  overall_pm
## 0.445180365 -0.006308698  0.060924715  0.500203618  0.561128333
## overall_int  overall_pe
## 0.054616017  0.554819635
##
## $effect_se
## [1] 0.011479199 0.011871633 0.011272522 0.011148156 0.011894633
## [6] 0.005671775 0.006233316 0.000981710 0.009084909 0.011148156
## [11] 0.006211659 0.000546770 0.005075677 0.006062618 0.006452578
## [16] 0.004550481 0.006211659
```

1.2.2.2 Direct Imputation Estimation and Bootstrap Inference

```
causal_mediation(data = df_int, outcome = "contY_contM_int", exposure = 'A',
  exposure.type = "binary",
  mediator = 'M_cont', covariates.pre = "C", covariates.post = NULL,
  covariates.post.type = NULL, cval = list(),
  EMint = TRUE, MMint = NULL, EMMint = NULL, EMint.terms = NULL,
  MMint.terms = NULL, EMMint.terms = NULL,
  yreg = "linear", mreg = "linear", mval = list(0),
  a_star = 0, a = 1,
  est.method = "imputation", inf.method = "bootstrap", model = "rb",
  nboot = 200, conf = 0.95, nrep = 5)
```

```
## $effect_estimate
##      cde      pnide      tnide      pnide      tnide
## 0.796419787 0.785133639 0.894126874 0.894855413 1.003848648
##      te      pm      intref      intmed      pie
## 1.788982287 0.389978026 -0.011286148 0.108993235 0.894855413
##      cde_prop  intref_prop  intmed_prop    pie_prop  overall_pm
## 0.445180365 -0.006308698  0.060924715  0.500203618  0.561128333
## overall_int  overall_pe
## 0.054616017  0.554819635
##
## $effect_se
## [1] 0.0121471524 0.0126339225 0.0116327783 0.0120059299 0.0130913424
## [6] 0.0059320156 0.0067421773 0.0011346497 0.0100090050 0.0120059299
## [11] 0.0067014233 0.0006340379 0.0055970823 0.0063974882 0.0069818776
## [16] 0.0050344477 0.0067014233
```

1.2.3 Causal Effects and Standard Errors Estimated By the Weighting-based Approach

```
causal_mediation(data = df_int, outcome = "contY_contM_int", exposure = 'A',
  exposure.type = "binary",
  mediator = 'M_cont', covariates.pre = "C", covariates.post = NULL,
  covariates.post.type = NULL, cval = list(),
  EMint = TRUE, MMint = NULL, EMMint = NULL, EMint.terms = NULL,
  MMint.terms = NULL, EMMint.terms = NULL,
  yreg = "linear", mreg = "linear", mval = list(0),
  a_star = 0, a = 1,
  est.method = "imputation", inf.method = "bootstrap", model = "wb",
  nboot = 200, conf = 0.95, nrep = 5)
```

```
## $effect_estimate
##      cde      pnde      tnde      pnle      tnle
## 0.796420133 0.785126002 0.894138465 0.894839441 1.003851904
##      te      pm      intref      intmed      pie
## 1.788977907 0.389981112 -0.011294131 0.109012463 0.894839441
##      cde_prop  intref_prop  intmed_prop  pie_prop  overall_pm
## 0.445181648 -0.006313175 0.060935612 0.500195915 0.561131527
## overall_int  overall_pe
## 0.054622437 0.554818352
##
## $effect_se
## [1] 0.0108336602 0.0114534123 0.0096346395 0.0100203178 0.0116772052
## [6] 0.0058716022 0.0060326679 0.0011597302 0.0094079191 0.0100203178
## [11] 0.0058832657 0.0006480096 0.0052663722 0.0052441285 0.0062427157
## [16] 0.0046858080 0.0058832657
```

1.2.4 Causal Effects and Standard Errors Estimated By the Marginal Structural Model

```
causal_mediation(data = df_int, outcome = "contY_contM_int", exposure = "A",
  exposure.type = "binary",
  mediator = 'M_cont', covariates.pre = "C", covariates.post = NULL,
  covariates.post.type = NULL, cval = list(mean(df_noint$C)),
  EMint = TRUE, MMint = NULL, EMMint = NULL, EMint.terms = NULL,
  MMint.terms = NULL, EMMint.terms = NULL,
  yreg = "linear", mreg = "linear", mval = list(0),
  a_star = 0, a = 1,
  est.method = "imputation", inf.method = "bootstrap", model = "msm",
  nboot = 200, conf = 0.95, nrep = 5)
```

```
## $effect_estimate
##      cde      pnle      tnle      pnle      tnle      te
## 0.888058066 0.898067972 0.936561492 0.891144560 0.929638080 1.827706052
##      pm      intref      intmed      pie      cde_prop  intref_prop
## 0.341054714 0.010009905 0.038493520 0.891144560 0.485886702 0.005476759
## intmed_prop  pie_prop  overall_pm  overall_int  overall_pe
## 0.021061111 0.487575428 0.508636539 0.026537870 0.514113298
##
## $effect_se
## [1] 0.10564098 0.07362774 0.04315188 0.03547907 0.06305469 0.01895185
## [7] 0.03491612 0.03651082 0.05269477 0.03547907 0.05496470 0.02017150
## [13] 0.02912684 0.02107103 0.03754014 0.04465273 0.05496470
```

1.2.5 Causal Effects and Standard Errors Estimated By the Inverse Odds-ratio Weighting Approach

```
causal_mediation(data = df_int, outcome = "contY_contM_int", exposure = 'A',
  exposure.type = "binary",
  mediator = 'M_cont', covariates.pre = "C", covariates.post = NULL,
  covariates.post.type = NULL, cval = list(),
  EMint = TRUE, MMint = NULL, EMMint = NULL, EMint.terms = NULL,
  MMint.terms = NULL, EMMint.terms = NULL,
  yreg = "linear", mreg = "linear", mval = list(0),
```

```

a_star = 0, a = 1,
est.method = "paramfunc", inf.method = "bootstrap", model = "iorw",
nboot = 200, conf = 0.95, nrep = 5)

## $effect_estimate
##      tot      dir      ind
## 1.7891357 1.2008298 0.5883059
##
## $effect_se
## [1] 0.005315001 0.034053654 0.033837442

```

1.2.6 Causal Effects and Standard Errors Estimated By the G-formula Approach

```

causal_mediation(data = df_int, outcome = "contY_contM_int", exposure = 'A',
  exposure.type = "binary",
  mediator = 'M_cont', covariates.pre = "C", covariates.post = NULL,
  covariates.post.type = NULL, cval = list(),
  EMint = TRUE, MMint = NULL, EMMint = NULL, EMint.terms = NULL,
  MMint.terms = NULL, EMMint.terms = NULL,
  yreg = "linear", mreg = "linear", mval = list(0),
  a_star = 0, a = 1,
  est.method = "imputation", inf.method = "bootstrap", model = "g-formula",
  nboot = 200, conf = 0.95, nrep = 5)

## $effect_estimate
##      cde      pnde      tnde      pnle      tnle
## 0.796419787 0.785133639 0.894126874 0.894855413 1.003848648
##      te      pm      intref      intmed      pie
## 1.788982287 0.389978026 -0.011286148 0.108993235 0.894855413
##      cde_prop intref_prop intmed_prop pie_prop overall_pm
## 0.445180365 -0.006308698 0.060924715 0.500203618 0.561128333
## overall_int overall_pe
## 0.054616017 0.554819635
##
## $effect_se
## [1] 0.0111678964 0.0115711701 0.0110211722 0.0103611117 0.0112635424
## [6] 0.0056034765 0.0059918735 0.0010984675 0.0093824864 0.0103611117
## [11] 0.0059489696 0.0006120644 0.0052357909 0.0057665607 0.0062011537
## [16] 0.0046959642 0.0059489696

```

1.2.7 Causal Effects and Standard Errors Estimated By the Natural Effect Model

```

causal_mediation(data = df_int, outcome = "contY_contM_int", exposure = 'A',
  exposure.type = "binary",
  mediator = 'M_cont', covariates.pre = "C", covariates.post = NULL,
  covariates.post.type = NULL, cval = NULL,
  EMint = TRUE, MMint = NULL, EMMint = NULL, EMint.terms = NULL,
  MMint.terms = NULL, EMMint.terms = NULL,
  yreg = "linear", mreg = "linear", mval = list(0),
  a_star = 0, a = 1,
  est.method = NULL, inf.method = NULL, model = "ne",
  nboot = 200, conf = 0.95, nrep = 5)

```

##	Estimate	Std. Error
## pure direct effect	0.7848238	0.012125260
## total direct effect	0.8945918	0.011321872
## pure indirect effect	0.8944680	0.011194515
## total indirect effect	1.0042360	0.012145221
## total effect	1.7890598	0.005715884