Model Comparison

Baoyi Shi

1 Case 1: Continuous Outcome and Single Continuous Mediator

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

1.1.1 Data simulation

1.1.1.1 Simulation Procedures

1. Simulate the exposure variable A from $Bernoulli(P(A = 1))$.
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

```
cmest(data = df_noint,  outcome = "contY_contM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cont', covariates.pre = "C",
                yreg = "linear", mreg = "linear", mval = list(0),
                a_star = 0, a = 1,
                est.method = "paramfunc", inf.method = "delta", model = "rb")
```

```
##             Estimate Std.error  95% CIL  95% CIU      z   pval
## cde         0.796179  0.010784 0.775043 0.817314  73.83 <2e-16 ***
## pnde        0.796179  0.010784 0.775043 0.817314  73.83 <2e-16 ***
## tnde        0.796179  0.010784 0.775043 0.817314  73.83 <2e-16 ***
## pnie        0.897303  0.010634 0.876461 0.918145  84.38 <2e-16 ***
## tnie        0.897303  0.010634 0.876461 0.918145  84.38 <2e-16 ***
## te          1.693481  0.005536 1.682632 1.704331 305.92 <2e-16 ***
```

```
## pm          0.360412  0.005656 0.349326 0.371497  63.72 <2e-16 ***
## intref      0.000000  0.000000 0.000000 0.000000     NA     NA
## intmed      0.000000  0.000000 0.000000 0.000000     NA     NA
## pie         0.897303  0.010634 0.876461 0.918145  84.38 <2e-16 ***
## cde_prop    0.470143  0.006112 0.458163 0.482123  76.92 <2e-16 ***
## intref_prop 0.000000  0.000000 0.000000 0.000000     NA     NA
## intmed_prop 0.000000  0.000000 0.000000 0.000000     NA     NA
## pie_prop    0.529857  0.006112 0.517877 0.541837  86.69 <2e-16 ***
## overall_pm  0.529857  0.006112 0.517877 0.541837  86.69 <2e-16 ***
## overall_int 0.000000  0.000000 0.000000 0.000000     NA     NA
## overall_pe  0.529857  0.006112 0.517877 0.541837  86.69 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

1.1.2.2 Direct Imputation Estimation and Bootstrap Inference

```
cmest(data = df_noint,  outcome = "contY_contM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cont', covariates.pre = "C",
                yreg = "linear", mreg = "linear", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "rb")
```

```
##               Estimate  Std.error    95% CIL    95% CIU      z   pval
## cde          7.962e-01  1.124e-02  7.741e-01  8.182e-01  70.83 <2e-16 ***
## pnde         7.962e-01  1.124e-02  7.741e-01  8.182e-01  70.83 <2e-16 ***
## tnde         7.962e-01  1.124e-02  7.741e-01  8.182e-01  70.83 <2e-16 ***
## pnie         8.973e-01  1.167e-02  8.744e-01  9.202e-01  76.90 <2e-16 ***
## tnie         8.973e-01  1.167e-02  8.744e-01  9.202e-01  76.90 <2e-16 ***
## te           1.693e+00  5.700e-03  1.682e+00  1.705e+00 297.08 <2e-16 ***
## pm           3.604e-01  6.057e-03  3.485e-01  3.723e-01  59.50 <2e-16 ***
## intref       0.000e+00  6.280e-17 -1.231e-16  1.231e-16   0.00      1
## intmed       0.000e+00  3.208e-16 -6.288e-16  6.288e-16   0.00      1
## pie          8.973e-01  1.167e-02  8.744e-01  9.202e-01  76.90 <2e-16 ***
## cde_prop     4.701e-01  6.542e-03  4.573e-01  4.830e-01  71.86 <2e-16 ***
## intref_prop  0.000e+00  3.707e-17 -7.265e-17  7.265e-17   0.00      1
## intmed_prop  0.000e+00  1.894e-16 -3.712e-16  3.712e-16   0.00      1
## pie_prop     5.299e-01  6.542e-03  5.170e-01  5.427e-01  80.99 <2e-16 ***
## overall_pm   5.299e-01  6.542e-03  5.170e-01  5.427e-01  80.99 <2e-16 ***
## overall_int  0.000e+00  1.854e-16 -3.634e-16  3.634e-16   0.00      1
## overall_pe   5.299e-01  6.542e-03  5.170e-01  5.427e-01  80.99 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_noint,  outcome = "contY_contM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cont', covariates.pre = "C",
                yreg = "linear", mreg = "linear", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "wb")
```

```
##                Estimate  Std.error    95% CIL    95% CIU       z   pval
## cde           7.962e-01  1.116e-02  7.743e-01  8.180e-01  71.365 <2e-16 ***
## pnde          7.961e-01  1.117e-02  7.742e-01  8.180e-01  71.253 <2e-16 ***
## tnde          7.961e-01  1.116e-02  7.743e-01  8.180e-01  71.310 <2e-16 ***
## pnie          8.973e-01  1.146e-02  8.748e-01  9.197e-01  78.277 <2e-16 ***
## tnie          8.973e-01  1.147e-02  8.748e-01  9.198e-01  78.239 <2e-16 ***
## te            1.693e+00  5.270e-03  1.683e+00  1.704e+00 321.316 <2e-16 ***
## pm            3.604e-01  6.004e-03  3.487e-01  3.722e-01  60.030 <2e-16 ***
## intref       -9.335e-05  1.331e-04 -3.543e-04  1.676e-04  -0.701  0.483
## intmed        5.353e-05  6.747e-05 -7.872e-05  1.858e-04   0.793  0.428
## pie           8.973e-01  1.146e-02  8.748e-01  9.197e-01  78.277 <2e-16 ***
## cde_prop      4.702e-01  6.482e-03  4.575e-01  4.829e-01  72.540 <2e-16 ***
## intref_prop -5.512e-05  7.861e-05 -2.092e-04  9.895e-05  -0.701  0.483
## intmed_prop  3.161e-05  3.986e-05 -4.652e-05  1.097e-04   0.793  0.428
## pie_prop      5.299e-01  6.487e-03  5.171e-01  5.426e-01  81.677 <2e-16 ***
## overall_pm    5.299e-01  6.492e-03  5.172e-01  5.426e-01  81.627 <2e-16 ***
## overall_int -2.352e-05  4.729e-05 -1.162e-04  6.917e-05  -0.497  0.619
## overall_pe    5.298e-01  6.482e-03  5.171e-01  5.425e-01  81.746 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_noint,  outcome = "contY_contM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cont', covariates.pre = "C",
                yreg = "linear", mreg = "linear", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "g-formula")
```

```
##                Estimate  Std.error    95% CIL    95% CIU      z   pval
## cde           7.962e-01  1.130e-02  7.740e-01  8.183e-01  70.44 <2e-16 ***
## pnde          7.962e-01  1.130e-02  7.740e-01  8.183e-01  70.44 <2e-16 ***
## tnde          7.962e-01  1.130e-02  7.740e-01  8.183e-01  70.44 <2e-16 ***
## pnie          8.973e-01  1.161e-02  8.745e-01  9.201e-01  77.29 <2e-16 ***
## tnie          8.973e-01  1.161e-02  8.745e-01  9.201e-01  77.29 <2e-16 ***
## te            1.693e+00  5.688e-03  1.682e+00  1.705e+00 297.71 <2e-16 ***
## pm            3.604e-01  6.063e-03  3.485e-01  3.723e-01  59.45 <2e-16 ***
## intref        0.000e+00  0.000e+00  0.000e+00  0.000e+00     NA     NA
## intmed        0.000e+00  3.140e-16 -6.155e-16  6.155e-16   0.00      1
## pie           8.973e-01  1.161e-02  8.745e-01  9.201e-01  77.29 <2e-16 ***
## cde_prop      4.701e-01  6.554e-03  4.573e-01  4.830e-01  71.74 <2e-16 ***
## intref_prop   0.000e+00  0.000e+00  0.000e+00  0.000e+00     NA     NA
## intmed_prop   0.000e+00  1.855e-16 -3.635e-16  3.635e-16   0.00      1
## pie_prop      5.299e-01  6.554e-03  5.170e-01  5.427e-01  80.85 <2e-16 ***
## overall_pm    5.299e-01  6.554e-03  5.170e-01  5.427e-01  80.85 <2e-16 ***
## overall_int   0.000e+00  1.855e-16 -3.635e-16  3.635e-16   0.00      1
## overall_pe    5.299e-01  6.554e-03  5.170e-01  5.427e-01  80.85 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_noint,  outcome = "contY_contM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cont', covariates.pre = "C",
                yreg = "linear", mreg = "linear", mval = list(0),
                a_star = 0, a = 1, model = "ne")
```

```
##                          Estimate Std.error  95% CIL  95% CIU      z   pval
## natural direct effect   0.796179  0.010948 0.774722 0.817636  72.73 <2e-16
## natural indirect effect 0.897303  0.010766 0.876201 0.918405  83.34 <2e-16
## total effect            1.693481  0.005534 1.682634 1.704329 305.99 <2e-16
##
## natural direct effect   ***
## natural indirect effect ***
## total effect            ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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 $Bernoulli(P(A = 1))$.
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

```r
cmest(data = df_int,  outcome = "contY_contM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cont', covariates.pre = "C", EMint = TRUE,
                yreg = "linear", mreg = "linear", mval = list(0),
                a_star = 0, a = 1,
                est.method = "paramfunc", inf.method = "delta", model = "rb")
```

```
##                Estimate  Std.error    95% CIL    95% CIU      z   pval
## cde           0.7964198  0.0114792  0.7739210  0.8189186  69.38 <2e-16 ***
## pnde          0.7851336  0.0118716  0.7618657  0.8084016  66.14 <2e-16 ***
## tnde          0.8941269  0.0112725  0.8720331  0.9162206  79.32 <2e-16 ***
## pnie          0.8948554  0.0111482  0.8730054  0.9167054  80.27 <2e-16 ***
## tnie          1.0038486  0.0118946  0.9805356  1.0271617  84.39 <2e-16 ***
## te            1.7889823  0.0056718  1.7778658  1.8000988 315.42 <2e-16 ***
## pm            0.3899780  0.0062333  0.3777610  0.4021951  62.56 <2e-16 ***
## intref       -0.0112861  0.0009817 -0.0132103 -0.0093620 -11.50 <2e-16 ***
## intmed        0.1089932  0.0090849  0.0911871  0.1267993  12.00 <2e-16 ***
## pie           0.8948554  0.0111482  0.8730054  0.9167054  80.27 <2e-16 ***
## cde_prop      0.4451804  0.0062117  0.4330057  0.4573550  71.67 <2e-16 ***
## intref_prop -0.0063087  0.0005468 -0.0073803 -0.0052370 -11.54 <2e-16 ***
## intmed_prop  0.0609247  0.0050757  0.0509766  0.0708729  12.00 <2e-16 ***
## pie_prop      0.5002036  0.0060626  0.4883211  0.5120861  82.51 <2e-16 ***
## overall_pm    0.5611283  0.0064526  0.5484815  0.5737752  86.96 <2e-16 ***
## overall_int   0.0546160  0.0045505  0.0456972  0.0635348  12.00 <2e-16 ***
## overall_pe    0.5548196  0.0062117  0.5426450  0.5669943  89.32 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

1.2.2.2 Direct Imputation Estimation and Bootstrap Inference

```r
cmest(data = df_int,  outcome = "contY_contM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cont', covariates.pre = "C", EMint = TRUE,
                yreg = "linear", mreg = "linear", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "rb")
```

```
##                Estimate  Std.error    95% CIL    95% CIU      z   pval
## cde           0.7964198  0.0124881  0.7719436  0.8208960  63.77 <2e-16 ***
## pnde          0.7851336  0.0128298  0.7599877  0.8102796  61.20 <2e-16 ***
## tnde          0.8941269  0.0123518  0.8699178  0.9183360  72.39 <2e-16 ***
## pnie          0.8948554  0.0120887  0.8711620  0.9185488  74.02 <2e-16 ***
## tnie          1.0038486  0.0127372  0.9788842  1.0288131  78.81 <2e-16 ***
## te            1.7889823  0.0056284  1.7779509  1.8000137 317.85 <2e-16 ***
## pm            0.3899780  0.0067391  0.3767696  0.4031865  57.87 <2e-16 ***
## intref       -0.0112861  0.0010557 -0.0133554 -0.0092169 -10.69 <2e-16 ***
## intmed        0.1089932  0.0088968  0.0915559  0.1264306  12.25 <2e-16 ***
## pie           0.8948554  0.0120887  0.8711620  0.9185488  74.02 <2e-16 ***
## cde_prop      0.4451804  0.0067648  0.4319216  0.4584391  65.81 <2e-16 ***
## intref_prop -0.0063087  0.0005889 -0.0074630 -0.0051544 -10.71 <2e-16 ***
## intmed_prop  0.0609247  0.0049699  0.0511839  0.0706655  12.26 <2e-16 ***
## pie_prop      0.5002036  0.0066450  0.4871797  0.5132276  75.28 <2e-16 ***
## overall_pm    0.5611283  0.0069728  0.5474620  0.5747947  80.47 <2e-16 ***
## overall_int   0.0546160  0.0044578  0.0458789  0.0633531  12.25 <2e-16 ***
```

```
## overall_pe   0.5548196  0.0067648  0.5415609  0.5680784  82.02 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_int,  outcome = "contY_contM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cont', covariates.pre = "C", EMint = TRUE,
                yreg = "linear", mreg = "linear", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "wb")
```

```
##                Estimate  Std.error    95% CIL    95% CIU       z   pval
## cde           0.7964207  0.0120580  0.7727874  0.8200540  66.049 <2e-16 ***
## pnde          0.7858818  0.0123889  0.7615999  0.8101636  63.434 <2e-16 ***
## tnde          0.8948932  0.0115830  0.8721910  0.9175954  77.259 <2e-16 ***
## pnie          0.8948883  0.0115913  0.8721697  0.9176068  77.203 <2e-16 ***
## tnie          1.0038997  0.0121345  0.9801164  1.0276829  82.731 <2e-16 ***
## te            1.7897814  0.0059094  1.7781992  1.8013637 302.869 <2e-16 ***
## pm            0.3897636  0.0064334  0.3771543  0.4023728  60.584 <2e-16 ***
## intref       -0.0105389  0.0012323 -0.0129543 -0.0081236  -8.552 <2e-16 ***
## intmed        0.1090114  0.0084600  0.0924301  0.1255927  12.886 <2e-16 ***
## pie           0.8948883  0.0115913  0.8721697  0.9176068  77.203 <2e-16 ***
## cde_prop      0.4449821  0.0064779  0.4322857  0.4576785  68.692 <2e-16 ***
## intref_prop -0.0058884  0.0006898 -0.0072403 -0.0045364  -8.537 <2e-16 ***
## intmed_prop  0.0609077  0.0047416  0.0516143  0.0702010  12.845 <2e-16 ***
## pie_prop      0.4999986  0.0062585  0.4877321  0.5122652  79.891 <2e-16 ***
## overall_pm    0.5609063  0.0066602  0.5478526  0.5739600  84.218 <2e-16 ***
## overall_int   0.0550193  0.0043312  0.0465302  0.0635083  12.703 <2e-16 ***
## overall_pe    0.5550179  0.0064779  0.5423215  0.5677143  85.679 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_int,  outcome = "contY_contM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cont', covariates.pre = "C", EMint = TRUE,
                yreg = "linear", mreg = "linear", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "g-formula")
```

```
##                Estimate  Std.error    95% CIL    95% CIU      z   pval
## cde           0.7964198  0.0111023  0.7746596  0.8181799  71.73 <2e-16 ***
## pnde          0.7851336  0.0113937  0.7628025  0.8074648  68.91 <2e-16 ***
## tnde          0.8941269  0.0114990  0.8715893  0.9166644  77.76 <2e-16 ***
## pnie          0.8948554  0.0112160  0.8728725  0.9168383  79.78 <2e-16 ***
## tnie          1.0038486  0.0113366  0.9816292  1.0260681  88.55 <2e-16 ***
## te            1.7889823  0.0057269  1.7777577  1.8002068 312.38 <2e-16 ***
## pm            0.3899780  0.0059512  0.3783139  0.4016421  65.53 <2e-16 ***
## intref       -0.0112861  0.0010411 -0.0133266 -0.0092456 -10.84 <2e-16 ***
```

```
## intmed       0.1089932  0.0086313  0.0920762  0.1259103  12.63 <2e-16 ***
## pie          0.8948554  0.0112160  0.8728725  0.9168383  79.78 <2e-16 ***
## cde_prop     0.4451804  0.0059839  0.4334521  0.4569087  74.40 <2e-16 ***
## intref_prop -0.0063087  0.0005821 -0.0074495 -0.0051679 -10.84 <2e-16 ***
## intmed_prop  0.0609247  0.0048225  0.0514728  0.0703767  12.63 <2e-16 ***
## pie_prop     0.5002036  0.0061464  0.4881568  0.5122504  81.38 <2e-16 ***
## overall_pm   0.5611283  0.0061581  0.5490587  0.5731980  91.12 <2e-16 ***
## overall_int  0.0546160  0.0043236  0.0461419  0.0630901  12.63 <2e-16 ***
## overall_pe   0.5548196  0.0059839  0.5430913  0.5665479  92.72 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_int,  outcome = "contY_contM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cont', covariates.pre = "C", EMint = TRUE,
                yreg = "linear", mreg = "linear",
                a_star = 0, a = 1, model = "ne")
```

```
##                       Estimate Std.error  95% CIL  95% CIU      z   pval
## pure direct effect    0.784803  0.012125 0.761039 0.808568  64.73 <2e-16
## total direct effect   0.894622  0.011322 0.872432 0.916813  79.02 <2e-16
## pure indirect effect  0.894527  0.011190 0.872595 0.916459  79.94 <2e-16
## total indirect effect 1.004346  0.012150 0.980532 1.028160  82.66 <2e-16
## total effect          1.789149  0.005718 1.777943 1.800356 312.90 <2e-16
##
## pure direct effect    ***
## total direct effect   ***
## pure indirect effect  ***
## total indirect effect ***
## total effect          ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

2 Case 2: Continuous Outcome and Single Binary Mediator

2.1 Case 2-1: Continuous Outcome and Single Binary Mediator Without Interaction

2.1.1 Data simulation

2.1.1.1 Simulation Procedures

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

2.1.1.2 True Parameters

Table 3: 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_Y
10000	-5	0.8	1.8	0.1	-0.25	0.5	0.2	0.4	1	1	0.2

2.1.1.3 True Models

True model for the mediator:

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

True model for the outcome:

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

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

2.1.2.1 Closed-form Parameter Function Estimation and Delta Method Inference

```
cmest(data = df_noint,  outcome = "contY_binM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_bin', covariates.pre = "C",
                yreg = "linear", mreg = "logistic", mval = list(0),
                a_star = 0, a = 1,
                est.method = "paramfunc", inf.method = "delta", model = "rb")
```

```
##             Estimate Std.error  95% CIL  95% CIU      z   pval
## cde         0.800526  0.004126 0.792439 0.808613 194.02 <2e-16 ***
## pnde        0.800526  0.004126 0.792439 0.808613 194.02 <2e-16 ***
## tnde        0.800526  0.004126 0.792439 0.808613 194.02 <2e-16 ***
## pnie        0.217074  0.018238 0.181328 0.252820  11.90 <2e-16 ***
## tnie        0.217074  0.018238 0.181328 0.252820  11.90 <2e-16 ***
## te          1.017600  0.018686 0.980976 1.054224  54.46 <2e-16 ***
## pm          0.119395  0.008852 0.102045 0.136744  13.49 <2e-16 ***
## intref      0.000000  0.000000 0.000000 0.000000     NA     NA
## intmed      0.000000  0.000000 0.000000 0.000000     NA     NA
## pie         0.217074  0.018238 0.181328 0.252820  11.90 <2e-16 ***
## cde_prop    0.786680  0.014129 0.758989 0.814372  55.68 <2e-16 ***
## intref_prop 0.000000  0.000000 0.000000 0.000000     NA     NA
## intmed_prop 0.000000  0.000000 0.000000 0.000000     NA     NA
## pie_prop    0.213320  0.014129 0.185628 0.241011  15.10 <2e-16 ***
## overall_pm  0.213320  0.014129 0.185628 0.241011  15.10 <2e-16 ***
## overall_int 0.000000  0.000000 0.000000 0.000000     NA     NA
## overall_pe  0.213320  0.014129 0.185628 0.241011  15.10 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

2.1.2.2 Direct Imputation Estimation and Bootstrap Inference

```
cmest(data = df_noint,  outcome = "contY_binM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_bin', covariates.pre = "C",
                yreg = "linear", mreg = "logistic", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "rb")
```

```
##                Estimate  Std.error    95% CIL    95% CIU       z   pval
## cde           8.005e-01  4.068e-03  7.926e-01  8.085e-01 196.796 <2e-16 ***
## pnde          8.005e-01  4.068e-03  7.926e-01  8.085e-01 196.796 <2e-16 ***
## tnde          8.005e-01  4.068e-03  7.926e-01  8.085e-01 196.796 <2e-16 ***
## pnie          2.144e-01  1.839e-02  1.783e-01  2.504e-01  11.657 <2e-16 ***
## tnie          2.144e-01  1.839e-02  1.783e-01  2.504e-01  11.657 <2e-16 ***
## te            1.015e+00  1.850e-02  9.786e-01  1.051e+00  54.859 <2e-16 ***
## pm            1.181e-01  9.026e-03  1.004e-01  1.358e-01  13.082 <2e-16 ***
## intref        4.441e-16  3.316e-16 -2.058e-16  1.094e-15   1.339  0.181
## intmed       -4.441e-16  3.179e-16 -1.067e-15  1.791e-16  -1.397  0.163
## pie           2.144e-01  1.839e-02  1.783e-01  2.504e-01  11.657 <2e-16 ***
## cde_prop      7.888e-01  1.449e-02  7.604e-01  8.172e-01  54.444 <2e-16 ***
## intref_prop   4.376e-16  3.263e-16 -2.020e-16  1.077e-15   1.341  0.180
## intmed_prop -4.376e-16  3.131e-16 -1.051e-15  1.761e-16  -1.397  0.162
## pie_prop      2.112e-01  1.449e-02  1.828e-01  2.396e-01  14.579 <2e-16 ***
## overall_pm    2.112e-01  1.449e-02  1.828e-01  2.396e-01  14.579 <2e-16 ***
## overall_int   0.000e+00  6.838e-17 -1.340e-16  1.340e-16   0.000  1.000
## overall_pe    2.112e-01  1.449e-02  1.828e-01  2.396e-01  14.579 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_noint,  outcome = "contY_binM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_bin', covariates.pre = "C",
                yreg = "linear", mreg = "logistic", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "wb")
```

```
##                Estimate  Std.error    95% CIL    95% CIU       z   pval
## cde           8.005e-01  4.185e-03  7.923e-01  8.087e-01 191.299 <2e-16 ***
## pnde          8.005e-01  4.181e-03  7.923e-01  8.087e-01 191.467 <2e-16 ***
## tnde          8.005e-01  4.183e-03  7.923e-01  8.087e-01 191.389 <2e-16 ***
## pnie          2.140e-01  1.846e-02  1.778e-01  2.502e-01  11.595 <2e-16 ***
## tnie          2.140e-01  1.846e-02  1.778e-01  2.502e-01  11.596 <2e-16 ***
## te            1.015e+00  1.894e-02  9.774e-01  1.052e+00  53.575 <2e-16 ***
## pm            1.179e-01  8.991e-03  1.003e-01  1.355e-01  13.114 <2e-16 ***
## intref        1.308e-05  9.575e-05 -1.746e-04  2.007e-04   0.137  0.891
## intmed       -8.374e-06  3.922e-05 -8.524e-05  6.849e-05  -0.214  0.831
## pie           2.140e-01  1.846e-02  1.778e-01  2.502e-01  11.595 <2e-16 ***
## cde_prop      7.890e-01  1.441e-02  7.608e-01  8.173e-01  54.771 <2e-16 ***
## intref_prop   1.289e-05  9.375e-05 -1.709e-04  1.966e-04   0.138  0.891
## intmed_prop -8.254e-06  3.847e-05 -8.366e-05  6.715e-05  -0.215  0.830
## pie_prop      2.109e-01  1.440e-02  1.827e-01  2.392e-01  14.648 <2e-16 ***
## overall_pm    2.109e-01  1.440e-02  1.827e-01  2.392e-01  14.649 <2e-16 ***
## overall_int   4.641e-06  6.617e-05 -1.251e-04  1.343e-04   0.070  0.944
## overall_pe    2.110e-01  1.441e-02  1.827e-01  2.392e-01  14.643 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_noint,  outcome = "contY_binM_noint", exposure = "A",
                exposure.type = "binary",
                mediator = 'M_bin', covariates.pre = "C",
                yreg = "linear", mreg = "logistic", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "msm")
```

```
##               Estimate  Std.error    95% CIL    95% CIU       z   pval
## cde          8.031e-01  4.005e-03  7.953e-01  8.110e-01 200.551 <2e-16 ***
## pnde         8.015e-01  3.595e-03  7.944e-01  8.085e-01 222.920 <2e-16 ***
## tnde         8.015e-01  3.595e-03  7.944e-01  8.085e-01 222.920 <2e-16 ***
## pnie         2.146e-01  1.688e-02  1.815e-01  2.477e-01  12.710 <2e-16 ***
## tnie         2.146e-01  1.688e-02  1.815e-01  2.477e-01  12.710 <2e-16 ***
## te           1.016e+00  1.766e-02  9.814e-01  1.051e+00  57.533 <2e-16 ***
## pm           1.181e-01  8.166e-03  1.021e-01  1.341e-01  14.457 <2e-16 ***
## intref      -1.659e-03  2.129e-03 -5.832e-03  2.514e-03  -0.779  0.436
## intmed      -4.441e-16  2.849e-16 -1.002e-15  1.142e-16  -1.559  0.119
## pie          2.146e-01  1.688e-02  1.815e-01  2.477e-01  12.710 <2e-16 ***
## cde_prop     7.905e-01  1.331e-02  7.644e-01  8.165e-01  59.388 <2e-16 ***
## intref_prop -1.633e-03  2.095e-03 -5.738e-03  2.473e-03  -0.779  0.436
## intmed_prop -4.371e-16  2.798e-16 -9.854e-16  1.112e-16  -1.562  0.118
## pie_prop     2.112e-01  1.309e-02  1.855e-01  2.368e-01  16.130 <2e-16 ***
## overall_pm   2.112e-01  1.309e-02  1.855e-01  2.368e-01  16.130 <2e-16 ***
## overall_int -1.633e-03  2.095e-03 -5.738e-03  2.473e-03  -0.779  0.436
## overall_pe   2.095e-01  1.331e-02  1.835e-01  2.356e-01  15.744 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_noint,  outcome = "contY_binM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_bin', covariates.pre = "C",
                yreg = "linear", mreg = "logistic", mval = list(0),
                a_star = 0, a = 1,
                est.method = "paramfunc", inf.method = "bootstrap", model = "iorw")
```

```
##       Estimate Std.error  95% CIL  95% CIU      z   pval
## ORtot 2.758986  0.052953 2.655200 2.862773  52.10 <2e-16 ***
## ORdir 2.226955  0.008722 2.209860 2.244049 255.33 <2e-16 ***
## ORind 1.238905  0.023460 1.192925 1.284886  52.81 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_noint,  outcome = "contY_binM_noint", exposure = 'A',
                exposure.type = "binary",
```

```
                mediator = 'M_bin', covariates.pre = "C",
                yreg = "linear", mreg = "logistic", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "g-formula")
```

```
##                Estimate  Std.error    95% CIL    95% CIU       z   pval
## cde           8.005e-01  3.981e-03  7.927e-01  8.083e-01 201.089 <2e-16 ***
## pnde          8.005e-01  3.981e-03  7.927e-01  8.083e-01 201.089 <2e-16 ***
## tnde          8.005e-01  3.981e-03  7.927e-01  8.083e-01 201.089 <2e-16 ***
## pnie          2.144e-01  1.695e-02  1.812e-01  2.476e-01  12.650 <2e-16 ***
## tnie          2.144e-01  1.695e-02  1.812e-01  2.476e-01  12.650 <2e-16 ***
## te            1.015e+00  1.703e-02  9.815e-01  1.048e+00  59.587 <2e-16 ***
## pm            1.181e-01  8.338e-03  1.017e-01  1.344e-01  14.163 <2e-16 ***
## intref        4.441e-16  3.062e-16 -1.561e-16  1.044e-15   1.450  0.147
## intmed       -4.441e-16  2.984e-16 -1.029e-15  1.409e-16  -1.488  0.137
## pie           2.144e-01  1.695e-02  1.812e-01  2.476e-01  12.650 <2e-16 ***
## cde_prop      7.888e-01  1.339e-02  7.625e-01  8.150e-01  58.912 <2e-16 ***
## intref_prop   4.376e-16  3.008e-16 -1.520e-16  1.027e-15   1.455  0.146
## intmed_prop  -4.376e-16  2.932e-16 -1.012e-15  1.371e-16  -1.492  0.136
## pie_prop      2.112e-01  1.339e-02  1.850e-01  2.375e-01  15.776 <2e-16 ***
## overall_pm    2.112e-01  1.339e-02  1.850e-01  2.375e-01  15.776 <2e-16 ***
## overall_int   0.000e+00  6.899e-17 -1.352e-16  1.352e-16   0.000  1.000
## overall_pe    2.112e-01  1.339e-02  1.850e-01  2.375e-01  15.776 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_noint,  outcome = "contY_binM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_bin', covariates.pre = "C",
                yreg = "linear", mreg = "logistic", mval = list(0),
                a_star = 0, a = 1, model = "ne")
```

```
##                         Estimate Std.error 95% CIL 95% CIU      z   pval
## natural direct effect    0.80053   0.00413 0.79243 0.80862 193.82 <2e-16
## natural indirect effect  0.21434   0.01803 0.17900 0.24967  11.89 <2e-16
## total effect             1.01486   0.01847 0.97866 1.05106  54.95 <2e-16
##
## natural direct effect   ***
## natural indirect effect ***
## total effect            ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

2.2 Case 2-2: Continuous Outcome and Single Binary Mediator With Exposure-mediator Interaction

2.2.1 Data simulation

2.2.1.1 Simulation Procedures

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

2.2.1.2 True Parameters

Table 4: 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_Y
10000	-5	0.8	1.8	0.2	0.1	-0.25	0.5	0.2	0.4	1	1	0.2

2.2.1.3 True Models

True model for the mediator:

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

True model for the outcome:

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

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

2.2.2.1 Closed-form Parameter Function Estimation and Delta Method Inference

```
cmest(data = df_int,  outcome = "contY_binM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_bin', covariates.pre = "C", EMint = TRUE,
                yreg = "linear", mreg = "logistic", mval = list(0),
                a_star = 0, a = 1,
                est.method = "paramfunc", inf.method = "delta", model = "rb")
```

```
##             Estimate Std.error  95% CIL  95% CIU      z   pval
## cde         0.800550  0.006183 0.788431 0.812668 129.47 <2e-16 ***
## pnde        0.899638  0.004324 0.891163 0.908113 208.06 <2e-16 ***
## tnde        0.923888  0.004383 0.915297 0.932478 210.78 <2e-16 ***
## pnie        0.216510  0.018195 0.180849 0.252171  11.90 <2e-16 ***
## tnie        0.240760  0.020236 0.201099 0.280421  11.90 <2e-16 ***
## te          1.140398  0.019824 1.101543 1.179252  57.52 <2e-16 ***
## pm          0.118017  0.008863 0.100647 0.135388  13.32 <2e-16 ***
## intref      0.099088  0.004253 0.090752 0.107424  23.30 <2e-16 ***
## intmed      0.024250  0.002265 0.019811 0.028688  10.71 <2e-16 ***
## pie         0.216510  0.018195 0.180849 0.252171  11.90 <2e-16 ***
## cde_prop    0.701992  0.012674 0.677150 0.726833  55.39 <2e-16 ***
## intref_prop 0.086889  0.004270 0.078521 0.095258  20.35 <2e-16 ***
## intmed_prop 0.021264  0.001670 0.017991 0.024538  12.73 <2e-16 ***
## pie_prop    0.189855  0.012749 0.164867 0.214843  14.89 <2e-16 ***
## overall_pm  0.211119  0.014181 0.183325 0.238913  14.89 <2e-16 ***
## overall_int 0.108153  0.004571 0.099194 0.117112  23.66 <2e-16 ***
## overall_pe  0.298008  0.012674 0.273167 0.322850  23.51 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

2.2.2.2 Direct Imputation Estimation and Bootstrap Inference

```
cmest(data = df_int,  outcome = "contY_binM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_bin', covariates.pre = "C", EMint = TRUE,
                yreg = "linear", mreg = "logistic", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "rb")
```

```
##             Estimate Std.error  95% CIL  95% CIU      z   pval
## cde         0.800550  0.006128 0.788539 0.812560 130.64 <2e-16 ***
## pnde        0.899660  0.004298 0.891236 0.908083 209.32 <2e-16 ***
## tnde        0.923607  0.004319 0.915143 0.932072 213.86 <2e-16 ***
## pnie        0.213816  0.020449 0.173737 0.253895  10.46 <2e-16 ***
## tnie        0.237764  0.022672 0.193328 0.282199  10.49 <2e-16 ***
## te          1.137423  0.022152 1.094007 1.180839  51.35 <2e-16 ***
## pm          0.116718  0.009982 0.097153 0.136283  11.69 <2e-16 ***
## intref      0.099110  0.004025 0.091221 0.106998  24.62 <2e-16 ***
## intmed      0.023948  0.002416 0.019212 0.028684   9.91 <2e-16 ***
## pie         0.213816  0.020449 0.173737 0.253895  10.46 <2e-16 ***
## cde_prop    0.703828  0.013943 0.676501 0.731155  50.48 <2e-16 ***
## intref_prop 0.087135  0.004339 0.078631 0.095640  20.08 <2e-16 ***
## intmed_prop 0.021055  0.001774 0.017577 0.024532  11.87 <2e-16 ***
## pie_prop    0.187982  0.014494 0.159575 0.216390  12.97 <2e-16 ***
## overall_pm  0.209037  0.016062 0.177555 0.240519  13.01 <2e-16 ***
## overall_int 0.108190  0.004374 0.099617 0.116762  24.74 <2e-16 ***
## overall_pe  0.296172  0.013943 0.268845 0.323499  21.24 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_int,  outcome = "contY_binM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_bin', covariates.pre = "C", EMint = TRUE,
                yreg = "linear", mreg = "logistic", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "wb")
```

```
##             Estimate Std.error  95% CIL  95% CIU      z   pval
## cde         0.800551  0.006932 0.786964 0.814138 115.48 <2e-16 ***
## pnde        0.899861  0.004831 0.890392 0.909329 186.27 <2e-16 ***
## tnde        0.923783  0.004543 0.914878 0.932688 203.32 <2e-16 ***
## pnie        0.213422  0.018254 0.177645 0.249200  11.69 <2e-16 ***
## tnie        0.237344  0.020370 0.197420 0.277268  11.65 <2e-16 ***
## te          1.137205  0.019559 1.098870 1.175540  58.14 <2e-16 ***
## pm          0.116513  0.008996 0.098881 0.134145  12.95 <2e-16 ***
## intref      0.099310  0.004283 0.090916 0.107704  23.19 <2e-16 ***
## intmed      0.023922  0.002336 0.019343 0.028501  10.24 <2e-16 ***
## pie         0.213422  0.018254 0.177645 0.249200  11.69 <2e-16 ***
## cde_prop    0.703963  0.013177 0.678138 0.729789  53.42 <2e-16 ***
## intref_prop 0.087328  0.004274 0.078951 0.095705  20.43 <2e-16 ***
## intmed_prop 0.021036  0.001747 0.017612 0.024460  12.04 <2e-16 ***
## pie_prop    0.187673  0.012892 0.162405 0.212940  14.56 <2e-16 ***
```

```
## overall_pm  0.208708  0.014406 0.180472 0.236944  14.49 <2e-16 ***
## overall_int 0.108364  0.004679 0.099193 0.117535  23.16 <2e-16 ***
## overall_pe  0.296037  0.013177 0.270211 0.321862  22.47 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_int,  outcome = "contY_binM_int", exposure = "A",
                exposure.type = "binary",
                mediator = 'M_bin', covariates.pre = "C", EMint = TRUE,
                yreg = "linear", mreg = "logistic", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "msm")
```

```
##             Estimate Std.error  95% CIL  95% CIU      z   pval
## cde         0.804996  0.006398 0.792456 0.817535 125.82 <2e-16 ***
## pnde        0.899982  0.004243 0.891667 0.908297 212.13 <2e-16 ***
## tnde        0.923975  0.004316 0.915515 0.932435 214.06 <2e-16 ***
## pnie        0.213971  0.018431 0.177847 0.250095  11.61 <2e-16 ***
## tnie        0.237964  0.020500 0.197785 0.278143  11.61 <2e-16 ***
## te          1.137946  0.019909 1.098926 1.176966  57.16 <2e-16 ***
## pm          0.116768  0.009024 0.099081 0.134454  12.94 <2e-16 ***
## intref      0.094986  0.005286 0.084626 0.105347  17.97 <2e-16 ***
## intmed      0.023993  0.002295 0.019495 0.028490  10.46 <2e-16 ***
## pie         0.213971  0.018431 0.177847 0.250095  11.61 <2e-16 ***
## cde_prop    0.707411  0.013179 0.681580 0.733242  53.67 <2e-16 ***
## intref_prop 0.083472  0.005074 0.073527 0.093416  16.45 <2e-16 ***
## intmed_prop 0.021084  0.001701 0.017751 0.024417  12.40 <2e-16 ***
## pie_prop    0.188033  0.013023 0.162507 0.213558  14.44 <2e-16 ***
## overall_pm  0.209117  0.014483 0.180731 0.237503  14.44 <2e-16 ***
## overall_int 0.104556  0.005295 0.094178 0.114935  19.75 <2e-16 ***
## overall_pe  0.292589  0.013179 0.266758 0.318420  22.20 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_int,  outcome = "contY_binM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_bin', covariates.pre = "C", EMint = TRUE,
                yreg = "linear", mreg = "logistic", mval = list(0),
                a_star = 0, a = 1,
                est.method = "paramfunc", inf.method = "bootstrap", model = "iorw")
```

```
##       Estimate Std.error 95% CIL 95% CIU      z   pval
## ORtot  3.11876   0.05859 3.00392 3.23360  53.23 <2e-16 ***
## ORdir  2.51934   0.01085 2.49808 2.54060 232.29 <2e-16 ***
## ORind  1.23793   0.02160 1.19559 1.28027  57.30 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_int,  outcome = "contY_binM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_bin', covariates.pre = "C", EMint = TRUE,
                yreg = "linear", mreg = "logistic", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "g-formula")
```

```
##             Estimate Std.error  95% CIL  95% CIU      z   pval
## cde         0.800550  0.005832 0.789119 0.811980 137.27 <2e-16 ***
## pnde        0.899660  0.004291 0.891250 0.908069 209.67 <2e-16 ***
## tnde        0.923607  0.004367 0.915049 0.932166 211.50 <2e-16 ***
## pnie        0.213816  0.018220 0.178104 0.249527  11.73 <2e-16 ***
## tnie        0.237764  0.020178 0.198216 0.277311  11.78 <2e-16 ***
## te          1.137423  0.019721 1.098770 1.176076  57.67 <2e-16 ***
## pm          0.116718  0.008850 0.099372 0.134063  13.19 <2e-16 ***
## intref      0.099110  0.004091 0.091092 0.107127  24.23 <2e-16 ***
## intmed      0.023948  0.002169 0.019697 0.028198  11.04 <2e-16 ***
## pie         0.213816  0.018220 0.178104 0.249527  11.73 <2e-16 ***
## cde_prop    0.703828  0.012315 0.679691 0.727964  57.15 <2e-16 ***
## intref_prop 0.087135  0.004243 0.078819 0.095452  20.54 <2e-16 ***
## intmed_prop 0.021055  0.001585 0.017948 0.024161  13.28 <2e-16 ***
## pie_prop    0.187982  0.012810 0.162876 0.213089  14.68 <2e-16 ***
## overall_pm  0.209037  0.014167 0.181271 0.236803  14.76 <2e-16 ***
## overall_int 0.108190  0.004385 0.099595 0.116785  24.67 <2e-16 ***
## overall_pe  0.296172  0.012315 0.272036 0.320309  24.05 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_int,  outcome = "contY_binM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_bin', covariates.pre = "C", EMint = TRUE,
                yreg = "linear", mreg = "logistic", mval = list(0),
                a_star = 0, a = 1, model = "ne")
```

```
##                       Estimate Std.error  95% CIL  95% CIU      z   pval
## pure direct effect    0.899721  0.004331 0.891232 0.908210 207.74 <2e-16
## total direct effect   0.923510  0.004377 0.914930 0.932089 210.97 <2e-16
## pure indirect effect  0.213832  0.017986 0.178579 0.249084  11.89 <2e-16
## total indirect effect 0.237620  0.019998 0.198426 0.276815  11.88 <2e-16
## total effect          1.137341  0.019607 1.098912 1.175770  58.01 <2e-16
##
## pure direct effect    ***
## total direct effect   ***
## pure indirect effect  ***
## total indirect effect ***
## total effect          ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

3 Case 3: Continuous Outcome and Single Categorical Mediator

3.1 Case 3-1: Continuous Outcome and Single Categorical Mediator Without Interaction

3.1.1 Data simulation

3.1.1.1 Simulation Procedures

1. Simulate the exposure variable A from $Bernoulli(P(A=1))$.
2. Simulate the covariate C from $N(\mu_C, \sigma_C^2)$.
3. Simulate the mediator M from $Multinom(\frac{1}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}, \frac{expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}, \frac{expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)})$.
4. Simulate the outcome Y from $N(\theta_0 + \theta_1 A + \theta_{21} I\{M == 1\} + \theta_{22} I\{M == 2\} + \theta_4 C, \sigma_Y^2)$.

3.1.1.2 True Parameters

Table 5: True Model Parameters for Data Simulation

Sample Size	θ_0	θ_1	θ_{21}	θ_{22}	θ_4	β_{01}	β_{11}
10000	-5	0.8	1.8	1.2	0.1	-0.25	0.5
β_{21}	β_{02}	β_{12}	β_{22}	P(A=1)	μ_C	σ_C	σ_Y
0.2	-0.3	0.4	0.3	0.4	1	1	0.2

3.1.1.3 True Models

True model for the mediator:

$$ln\frac{P(M==1)}{P(M==0)} = \beta_{01} + \beta_{11}a + \beta_{21}c$$

$$ln\frac{P(M==2)}{P(M==0)} = \beta_{02} + \beta_{12}a + \beta_{22}c$$

True model for the outcome:

$$E[Y|a, m^*, c] = \theta_0 + \theta_1 a + \theta_{21} I\{m^* == 1\} + \theta_{22} I\{m^* == 2\} + \theta_4 c$$

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

3.1.2.1 Closed-form Parameter Function Estimation and Delta Method Inference

```
cmest(data = df_noint,  outcome = "contY_catM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cat', covariates.pre = "C",
                yreg = "linear", mreg = "multinomial", mval = list(0),
                a_star = 0, a = 1,
                est.method = "paramfunc", inf.method = "delta", model = "rb")
```

```
##             Estimate Std.error  95% CIL  95% CIU      z   pval
## cde         0.797684  0.004091 0.789665 0.805703 194.96 <2e-16 ***
## pnde        0.797684  0.004091 0.789665 0.805703 194.96 <2e-16 ***
## tnde        0.797684  0.004091 0.789665 0.805703 194.96 <2e-16 ***
## pnie        0.179264  0.014792 0.150271 0.208257  12.12 <2e-16 ***
```

```
## tnie        0.179264  0.014792 0.150271 0.208257  12.12 <2e-16 ***
## te          0.976948  0.015332 0.946898 1.006999  63.72 <2e-16 ***
## pm          0.101015  0.007510 0.086296 0.115734  13.45 <2e-16 ***
## intref      0.000000  0.000000 0.000000 0.000000     NA     NA
## intmed      0.000000  0.000000 0.000000 0.000000     NA     NA
## pie         0.179264  0.014792 0.150271 0.208257  12.12 <2e-16 ***
## cde_prop    0.816506  0.012390 0.792222 0.840790  65.90 <2e-16 ***
## intref_prop 0.000000  0.000000 0.000000 0.000000     NA     NA
## intmed_prop 0.000000  0.000000 0.000000 0.000000     NA     NA
## pie_prop    0.183494  0.012390 0.159210 0.207778  14.81 <2e-16 ***
## overall_pm  0.183494  0.012390 0.159210 0.207778  14.81 <2e-16 ***
## overall_int 0.000000  0.000000 0.000000 0.000000     NA     NA
## overall_pe  0.183494  0.012390 0.159210 0.207778  14.81 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

3.1.2.2 Direct Imputation Estimation and Bootstrap Inference

```
cmest(data = df_noint,  outcome = "contY_catM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cat', covariates.pre = "C",
                yreg = "linear", mreg = "multinomial", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "rb")
```

```
##               Estimate  Std.error    95% CIL    95% CIU      z   pval
## cde          7.977e-01  3.951e-03  7.899e-01  8.054e-01 201.87 <2e-16 ***
## pnde         7.977e-01  3.951e-03  7.899e-01  8.054e-01 201.87 <2e-16 ***
## tnde         7.977e-01  3.951e-03  7.899e-01  8.054e-01 201.87 <2e-16 ***
## pnie         1.778e-01  1.299e-02  1.523e-01  2.033e-01  13.69 <2e-16 ***
## tnie         1.778e-01  1.299e-02  1.523e-01  2.033e-01  13.69 <2e-16 ***
## te           9.755e-01  1.365e-02  9.487e-01  1.002e+00  71.44 <2e-16 ***
## pm           1.003e-01  6.581e-03  8.738e-02  1.132e-01  15.24 <2e-16 ***
## intref       0.000e+00  0.000e+00  0.000e+00  0.000e+00     NA     NA
## intmed       0.000e+00  4.452e-17 -8.726e-17  8.726e-17   0.00      1
## pie          1.778e-01  1.299e-02  1.523e-01  2.033e-01  13.69 <2e-16 ***
## cde_prop     8.177e-01  1.086e-02  7.964e-01  8.390e-01  75.31 <2e-16 ***
## intref_prop  0.000e+00  0.000e+00  0.000e+00  0.000e+00     NA     NA
## intmed_prop  0.000e+00  4.575e-17 -8.966e-17  8.966e-17   0.00      1
## pie_prop     1.823e-01  1.086e-02  1.610e-01  2.036e-01  16.79 <2e-16 ***
## overall_pm   1.823e-01  1.086e-02  1.610e-01  2.036e-01  16.79 <2e-16 ***
## overall_int  0.000e+00  4.575e-17 -8.966e-17  8.966e-17   0.00      1
## overall_pe   1.823e-01  1.086e-02  1.610e-01  2.036e-01  16.79 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_noint,  outcome = "contY_catM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cat', covariates.pre = "C",
                yreg = "linear", mreg = "multinomial", mval = list(0),
```

```
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "wb")
```

```
##                Estimate  Std.error    95% CIL    95% CIU       z   pval
## cde           7.977e-01  3.845e-03  7.901e-01  8.052e-01 207.474 <2e-16 ***
## pnde          7.977e-01  3.852e-03  7.901e-01  8.052e-01 207.093 <2e-16 ***
## tnde          7.977e-01  3.846e-03  7.901e-01  8.052e-01 207.380 <2e-16 ***
## pnie          1.772e-01  1.623e-02  1.453e-01  2.090e-01  10.912 <2e-16 ***
## tnie          1.772e-01  1.624e-02  1.453e-01  2.090e-01  10.910 <2e-16 ***
## te            9.748e-01  1.702e-02  9.415e-01  1.008e+00  57.282 <2e-16 ***
## pm            9.994e-02  8.242e-03  8.379e-02  1.161e-01  12.126 <2e-16 ***
## intref       -1.502e-06  8.853e-05 -1.750e-04  1.720e-04  -0.017  0.986
## intmed       -7.998e-07  3.585e-05 -7.106e-05  6.946e-05  -0.022  0.982
## pie           1.772e-01  1.623e-02  1.453e-01  2.090e-01  10.912 <2e-16 ***
## cde_prop      8.183e-01  1.365e-02  7.915e-01  8.450e-01  59.940 <2e-16 ***
## intref_prop -1.541e-06  9.099e-05 -1.799e-04  1.768e-04  -0.017  0.986
## intmed_prop -8.204e-07  3.671e-05 -7.277e-05  7.113e-05  -0.022  0.982
## pie_prop      1.817e-01  1.366e-02  1.549e-01  2.085e-01  13.299 <2e-16 ***
## overall_pm    1.817e-01  1.367e-02  1.549e-01  2.085e-01  13.297 <2e-16 ***
## overall_int -2.361e-06  7.135e-05 -1.422e-04  1.375e-04  -0.033  0.974
## overall_pe    1.817e-01  1.365e-02  1.550e-01  2.085e-01  13.311 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_noint,  outcome = "contY_catM_noint", exposure = "A",
                exposure.type = "binary",
                mediator = 'M_cat', covariates.pre = "C",
                yreg = "linear", mreg = "multinomial", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "msm")
```

```
##                Estimate  Std.error    95% CIL    95% CIU       z   pval
## cde           7.990e-01  4.950e-03  7.893e-01  8.087e-01 161.422 <2e-16 ***
## pnde          7.976e-01  4.465e-03  7.888e-01  8.063e-01 178.629 <2e-16 ***
## tnde          7.976e-01  4.465e-03  7.888e-01  8.063e-01 178.629 <2e-16 ***
## pnie          1.780e-01  1.580e-02  1.471e-01  2.090e-01  11.270 <2e-16 ***
## tnie          1.780e-01  1.580e-02  1.471e-01  2.090e-01  11.270 <2e-16 ***
## te            9.756e-01  1.571e-02  9.448e-01  1.006e+00  62.115 <2e-16 ***
## pm            1.004e-01  8.115e-03  8.449e-02  1.163e-01  12.373 <2e-16 ***
## intref       -1.445e-03  2.201e-03 -5.760e-03  2.870e-03  -0.656  0.512
## intmed        0.000e+00  5.448e-17 -1.068e-16  1.068e-16   0.000  1.000
## pie           1.780e-01  1.580e-02  1.471e-01  2.090e-01  11.270 <2e-16 ***
## cde_prop      8.190e-01  1.340e-02  7.927e-01  8.453e-01  61.102 <2e-16 ***
## intref_prop -1.481e-03  2.261e-03 -5.912e-03  2.950e-03  -0.655  0.512
## intmed_prop  0.000e+00  5.611e-17 -1.100e-16  1.100e-16   0.000  1.000
## pie_prop      1.825e-01  1.341e-02  1.562e-01  2.088e-01  13.607 <2e-16 ***
## overall_pm    1.825e-01  1.341e-02  1.562e-01  2.088e-01  13.607 <2e-16 ***
## overall_int -1.481e-03  2.261e-03 -5.912e-03  2.950e-03  -0.655  0.512
## overall_pe    1.810e-01  1.340e-02  1.547e-01  2.073e-01  13.503 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_noint,  outcome = "contY_catM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cat', covariates.pre = "C",
                yreg = "linear", mreg = "multinomial", mval = list(0),
                a_star = 0, a = 1,
                est.method = "paramfunc", inf.method = "bootstrap", model = "iorw")
```

```
##       Estimate Std.error  95% CIL  95% CIU      z   pval
## ORtot 2.652291  0.038077 2.577662 2.726920  69.66 <2e-16 ***
## ORdir 2.220786  0.008488 2.204150 2.237422 261.64 <2e-16 ***
## ORind 1.194303  0.017388 1.160224 1.228382  68.69 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_noint,  outcome = "contY_catM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cat', covariates.pre = "C",
                yreg = "linear", mreg = "multinomial", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "g-formula")
```

```
##               Estimate  Std.error    95% CIL    95% CIU      z   pval
## cde          7.977e-01  3.727e-03  7.904e-01  8.050e-01 214.05 <2e-16 ***
## pnde         7.977e-01  3.727e-03  7.904e-01  8.050e-01 214.05 <2e-16 ***
## tnde         7.977e-01  3.727e-03  7.904e-01  8.050e-01 214.05 <2e-16 ***
## pnie         1.778e-01  1.662e-02  1.452e-01  2.104e-01  10.70 <2e-16 ***
## tnie         1.778e-01  1.662e-02  1.452e-01  2.104e-01  10.70 <2e-16 ***
## te           9.755e-01  1.715e-02  9.419e-01  1.009e+00  56.88 <2e-16 ***
## pm           1.003e-01  8.454e-03  8.371e-02  1.169e-01  11.86 <2e-16 ***
## intref       0.000e+00  7.036e-17 -1.379e-16  1.379e-16   0.00      1
## intmed       0.000e+00  7.698e-17 -1.509e-16  1.509e-16   0.00      1
## pie          1.778e-01  1.662e-02  1.452e-01  2.104e-01  10.70 <2e-16 ***
## cde_prop     8.177e-01  1.401e-02  7.903e-01  8.452e-01  58.38 <2e-16 ***
## intref_prop  0.000e+00  7.352e-17 -1.441e-16  1.441e-16   0.00      1
## intmed_prop  0.000e+00  8.009e-17 -1.570e-16  1.570e-16   0.00      1
## pie_prop     1.823e-01  1.401e-02  1.548e-01  2.097e-01  13.01 <2e-16 ***
## overall_pm   1.823e-01  1.401e-02  1.548e-01  2.097e-01  13.01 <2e-16 ***
## overall_int  0.000e+00  3.193e-17 -6.258e-17  6.258e-17   0.00      1
## overall_pe   1.823e-01  1.401e-02  1.548e-01  2.097e-01  13.01 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_noint,  outcome = "contY_catM_noint", exposure = 'A',
                exposure.type = "binary",
```

```
                mediator = 'M_cat', covariates.pre = "C",
                yreg = "linear", mreg = "multinomial", mval = list(0),
                a_star = 0, a = 1, model = "ne")
```

```
##                          Estimate Std.error  95% CIL  95% CIU      z   pval
## natural direct effect   0.797684  0.004091 0.789667 0.805702 195.00 <2e-16
## natural indirect effect 0.177740  0.014700 0.148928 0.206551  12.09 <2e-16
## total effect            0.975424  0.015240 0.945554 1.005294  64.00 <2e-16
##
## natural direct effect   ***
## natural indirect effect ***
## total effect            ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

3.2 Case 3-2: Continuous Outcome and Single Categorical Mediator With Exposure-mediator Interaction

3.2.1 Data simulation

3.2.1.1 Simulation Procedures

1. Simulate the exposure variable A from $Bernoulli(P(A=1))$.
2. Simulate the covariate C from $N(\mu_C, \sigma_C^2)$.
3. Simulate the mediator M from $Multinom(\frac{1}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}$, $\frac{expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}$, $\frac{expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)})$.
4. Simulate the outcome Y from $N(\theta_0 + \theta_1 A + \theta_{21} I\{M == 1\} + \theta_{22} I\{M == 2\} + \theta_{31} A * I\{M == 1\} + \theta_{32} A * I\{M == 2\} + \theta_4 C, \sigma_Y^2)$.

3.2.1.2 True Parameters

Table 6: True Model Parameters for Data Simulation

Sample Size	θ_0	θ_1	θ_{21}	θ_{22}	θ_{31}	θ_{32}	θ_4	β_{01}
10000	-5	0.8	1.8	1.2	0.2	0.40.1	-0.25	
β_{11}	β_{21}	β_{02}	β_{12}	β_{22}	P(A=1)	μ_C	σ_C	σ_Y
0.5	0.2	-0.3	0.4	0.3	0.4	1	1	0.2

3.2.1.3 True Models

True model for the mediator:

$$ln\frac{P(M==1)}{P(M==0)} = \beta_{01} + \beta_{11}a + \beta_{21}c$$

$$ln\frac{P(M==2)}{P(M==0)} = \beta_{02} + \beta_{12}a + \beta_{22}c$$

True model for the outcome:

$$E[Y|a, m^*, c] = \theta_0 + \theta_1 a + \theta_{21} I\{m^* == 1\} + \theta_{22} I\{m^* == 2\} + \theta_{31} a * I\{m^* == 1\} + \theta_{32} a * I\{m^* == 2\} + \theta_4 c$$

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

3.2.2.1 Closed-form Parameter Function Estimation and Delta Method Inference

```
cmest(data = df_int,  outcome = "contY_catM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cat', covariates.pre = "C", EMint = TRUE,
                yreg = "linear", mreg = "multinomial", mval = list(0),
                a_star = 0, a = 1,
                est.method = "paramfunc", inf.method = "delta", model = "rb")
```

```
##             Estimate Std.error  95% CIL  95% CIU       z   pval
## cde         0.812817  0.007672 0.797780 0.827854 105.945 <2e-16 ***
## pnde        1.002031  0.004675 0.992868 1.011194 214.335 <2e-16 ***
## tnde        1.031087  0.004789 1.021701 1.040473 215.316 <2e-16 ***
## pnie        0.179695  0.014833 0.150622 0.208768  12.114 <2e-16 ***
## tnie        0.208750  0.016995 0.175441 0.242060  12.283 <2e-16 ***
## te          1.210782  0.016504 1.178435 1.243128  73.365 <2e-16 ***
## pm          0.094337  0.007062 0.080495 0.108179  13.358 <2e-16 ***
## intref      0.189214  0.006199 0.177064 0.201365  30.522 <2e-16 ***
## intmed      0.029056  0.003350 0.022489 0.035622   8.672 <2e-16 ***
## pie         0.179695  0.014833 0.150622 0.208768  12.114 <2e-16 ***
## cde_prop    0.671316  0.010423 0.650886 0.691746  64.404 <2e-16 ***
## intref_prop 0.156274  0.005807 0.144892 0.167656  26.910 <2e-16 ***
## intmed_prop 0.023997  0.002573 0.018954 0.029040   9.326 <2e-16 ***
## pie_prop    0.148412  0.010327 0.128171 0.168653  14.371 <2e-16 ***
## overall_pm  0.172410  0.011795 0.149293 0.195527  14.618 <2e-16 ***
## overall_int 0.180272  0.006067 0.168380 0.192164  29.711 <2e-16 ***
## overall_pe  0.328684  0.010423 0.308254 0.349114  31.533 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

3.2.2.2 Direct Imputation Estimation and Bootstrap Inference

```
cmest(data = df_int,  outcome = "contY_catM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cat', covariates.pre = "C", EMint = TRUE,
                yreg = "linear", mreg = "multinomial", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "rb")
```

```
##             Estimate Std.error  95% CIL  95% CIU       z     pval
## cde         0.812817  0.007927 0.797281 0.828354 102.538  < 2e-16 ***
## pnde        1.001892  0.004909 0.992270 1.011514 204.075  < 2e-16 ***
## tnde        1.030586  0.005049 1.020691 1.040482 204.134  < 2e-16 ***
## pnie        0.178245  0.014602 0.149626 0.206864  12.207  < 2e-16 ***
## tnie        0.206939  0.016880 0.173855 0.240023  12.260  < 2e-16 ***
## te          1.208831  0.016212 1.177056 1.240607  74.562  < 2e-16 ***
## pm          0.093607  0.007049 0.079791 0.107423  13.279  < 2e-16 ***
## intref      0.189075  0.006160 0.177002 0.201148  30.695  < 2e-16 ***
## intmed      0.028694  0.003565 0.021707 0.035682   8.049 9.32e-16 ***
## pie         0.178245  0.014602 0.149626 0.206864  12.207  < 2e-16 ***
## cde_prop    0.672399  0.011037 0.650766 0.694032  60.920  < 2e-16 ***
## intref_prop 0.156411  0.005553 0.145528 0.167295  28.168  < 2e-16 ***
## intmed_prop 0.023737  0.002754 0.018340 0.029135   8.620  < 2e-16 ***
```

```
## pie_prop    0.147452  0.010226 0.127410 0.167495  14.419  < 2e-16 ***
## overall_pm  0.171189  0.011789 0.148083 0.194295  14.521  < 2e-16 ***
## overall_int 0.180149  0.006156 0.168083 0.192214  29.264  < 2e-16 ***
## overall_pe  0.327601  0.011037 0.305968 0.349234  29.681  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_int,  outcome = "contY_catM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cat', covariates.pre = "C", EMint = TRUE,
                yreg = "linear", mreg = "multinomial", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "wb")
```

```
##             Estimate Std.error  95% CIL  95% CIU       z   pval
## cde         0.812818  0.007213 0.798681 0.826955 112.689 <2e-16 ***
## pnde        1.002320  0.004413 0.993670 1.010970 227.104 <2e-16 ***
## tnde        1.030931  0.004683 1.021752 1.040110 220.132 <2e-16 ***
## pnie        0.177599  0.015942 0.146352 0.208846  11.140 <2e-16 ***
## tnie        0.206211  0.018065 0.170805 0.241616  11.415 <2e-16 ***
## te          1.208530  0.017687 1.173865 1.243196  68.329 <2e-16 ***
## pm          0.093272  0.007477 0.078617 0.107927  12.474 <2e-16 ***
## intref      0.189502  0.005827 0.178081 0.200923  32.521 <2e-16 ***
## intmed      0.028612  0.003352 0.022042 0.035182   8.535 <2e-16 ***
## pie         0.177599  0.015942 0.146352 0.208846  11.140 <2e-16 ***
## cde_prop    0.672567  0.010586 0.651819 0.693316  63.532 <2e-16 ***
## intref_prop 0.156803  0.005694 0.145643 0.167964  27.538 <2e-16 ***
## intmed_prop 0.023675  0.002563 0.018651 0.028698   9.237 <2e-16 ***
## pie_prop    0.146954  0.011080 0.125238 0.168671  13.263 <2e-16 ***
## overall_pm  0.170629  0.012492 0.146146 0.195113  13.659 <2e-16 ***
## overall_int 0.180478  0.005935 0.168846 0.192110  30.410 <2e-16 ***
## overall_pe  0.327433  0.010586 0.306684 0.348181  30.930 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_int,  outcome = "contY_catM_int", exposure = "A",
                exposure.type = "binary",
                mediator = 'M_cat', covariates.pre = "C", EMint = TRUE,
                yreg = "linear", mreg = "multinomial", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "msm")
```

```
##             Estimate Std.error  95% CIL  95% CIU       z   pval
## cde         0.818724  0.008868 0.801343 0.836104  92.325 <2e-16 ***
## pnde        1.002285  0.004186 0.994081 1.010490 239.434 <2e-16 ***
## tnde        1.030917  0.004331 1.022428 1.039406 238.020 <2e-16 ***
## pnie        0.178458  0.013320 0.152351 0.204566  13.397 <2e-16 ***
## tnie        0.207090  0.015148 0.177401 0.236779  13.671 <2e-16 ***
```

```
## te          1.209376  0.014834 1.180301 1.238450  81.525 <2e-16 ***
## pm          0.093636  0.006290 0.081308 0.105963  14.887 <2e-16 ***
## intref      0.183562  0.007693 0.168483 0.198640  23.860 <2e-16 ***
## intmed      0.028632  0.003038 0.022677 0.034586   9.424 <2e-16 ***
## pie         0.178458  0.013320 0.152351 0.204566  13.397 <2e-16 ***
## cde_prop    0.676981  0.010978 0.655464 0.698497  61.667 <2e-16 ***
## intref_prop 0.151782  0.006594 0.138859 0.164705  23.020 <2e-16 ***
## intmed_prop 0.023675  0.002343 0.019084 0.028266  10.107 <2e-16 ***
## pie_prop    0.147562  0.009286 0.129362 0.165762  15.891 <2e-16 ***
## overall_pm  0.171237  0.010505 0.150648 0.191826  16.301 <2e-16 ***
## overall_int 0.175457  0.007037 0.161665 0.189249  24.933 <2e-16 ***
## overall_pe  0.323019  0.010978 0.301503 0.344536  29.424 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_int,  outcome = "contY_catM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cat', covariates.pre = "C", EMint = TRUE,
                yreg = "linear", mreg = "multinomial", mval = list(0),
                a_star = 0, a = 1,
                est.method = "paramfunc", inf.method = "bootstrap", model = "iorw")
```

```
##       Estimate Std.error 95% CIL 95% CIU      z   pval
## ORtot  3.34955   0.05457 3.24260 3.45649  61.38 <2e-16 ***
## ORdir  2.80393   0.01185 2.78070 2.82716 236.60 <2e-16 ***
## ORind  1.19459   0.01750 1.16029 1.22889  68.27 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_int,  outcome = "contY_catM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cat', covariates.pre = "C", EMint = TRUE,
                yreg = "linear", mreg = "multinomial", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "g-formula")
```

```
##             Estimate Std.error  95% CIL  95% CIU       z   pval
## cde         0.812817  0.007978 0.797180 0.828455 101.877 <2e-16 ***
## pnde        1.001892  0.004872 0.992343 1.011441 205.640 <2e-16 ***
## tnde        1.030586  0.004651 1.021470 1.039702 221.579 <2e-16 ***
## pnie        0.178245  0.013960 0.150883 0.205606  12.768 <2e-16 ***
## tnie        0.206939  0.015724 0.176121 0.237757  13.161 <2e-16 ***
## te          1.208831  0.015082 1.179271 1.238392  80.150 <2e-16 ***
## pm          0.093607  0.006582 0.080706 0.106508  14.221 <2e-16 ***
## intref      0.189075  0.006344 0.176642 0.201508  29.805 <2e-16 ***
## intmed      0.028694  0.003234 0.022356 0.035032   8.873 <2e-16 ***
## pie         0.178245  0.013960 0.150883 0.205606  12.768 <2e-16 ***
```

```
## cde_prop    0.672399  0.009926 0.652945 0.691853  67.743 <2e-16 ***
## intref_prop 0.156411  0.005870 0.144906 0.167917  26.645 <2e-16 ***
## intmed_prop 0.023737  0.002531 0.018776 0.028698   9.378 <2e-16 ***
## pie_prop    0.147452  0.009841 0.128164 0.166740  14.983 <2e-16 ***
## overall_pm  0.171189  0.011015 0.149600 0.192779  15.541 <2e-16 ***
## overall_int 0.180149  0.006387 0.167630 0.192667  28.205 <2e-16 ***
## overall_pe  0.327601  0.009926 0.308147 0.347055  33.005 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_int,  outcome = "contY_catM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cat', covariates.pre = "C", EMint = TRUE,
                yreg = "linear", mreg = "multinomial", mval = list(0),
                a_star = 0, a = 1, model = "ne")
```

```
##                       Estimate Std.error  95% CIL  95% CIU      z   pval
## pure direct effect    1.001984  0.004680 0.992811 1.011156 214.10 <2e-16
## total direct effect   1.030418  0.004783 1.021043 1.039792 215.43 <2e-16
## pure indirect effect  0.178248  0.014752 0.149335 0.207160  12.08 <2e-16
## total indirect effect 0.206682  0.016852 0.173652 0.239712  12.26 <2e-16
## total effect          1.208665  0.016396 1.176529 1.240801  73.72 <2e-16
##
## pure direct effect    ***
## total direct effect   ***
## pure indirect effect  ***
## total indirect effect ***
## total effect          ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

4 Case 4: Continuous Outcome and Multiple Mediators

4.1 Case 4-1: Continuous Outcome and Multiple Mediators Without Interaction

4.1.1 Data simulation

4.1.1.1 Simulation Procedures

1. Simulate the exposure variable A from $Bernoulli(P(A=1))$.
2. Simulate the covariate C from $N(\mu_C, \sigma_C^2)$.
3. Simulate the first mediator M1 from $Multinom(\frac{1}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}$, $\frac{expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}$, $\frac{expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)})$, the second mediator M2 from $Bernoulli(expit(\beta_{03}+\beta_{13}*A+\beta_{23}*C))$.
4. Simulate the outcome Y from $N(\theta_0+\theta_1 A+\theta_{21}I\{M1==1\}+\theta_{22}I\{M1==2\}+\theta_{23}M2+\theta_4 C, \sigma_Y^2)$.

4.1.1.2 True Parameters

Table 7: True Model Parameters for Data Simulation

Sample Size	θ_0	θ_1	θ_{21}	θ_{22}	θ_{23}	θ_4	β_{01}	β_{11}	β_{21}
10000	-5	0.8	1.8	1.2	1.5	0.1	-0.25	0.5	0.2
β_{02}	β_{12}	β_{22}	β_{03}	β_{13}	β_{23}	P(A=1)	μ_C	σ_C	σ_Y
-0.3	0.4	0.3	-0.25	0.5	0.2	0.4	1	1	0.2

4.1.1.3 True Models

True model for the first mediator:

$$ln\frac{P(M1==1)}{P(M1==0)} = \beta_{01} + \beta_{11}a + \beta_{21}c$$

$$ln\frac{P(M1==2)}{P(M1==0)} = \beta_{02} + \beta_{12}a + \beta_{22}c$$

True model for the second mediator:

$$logitE[M2|a,c] = \beta_{03} + \beta_{13}a + \beta_{23}c$$

True model for the outcome:

$$E[Y|a,m^*,c] = \theta_0 + \theta_1 a + \theta_{21} I\{m1^* == 1\} + \theta_{22} I\{m1^* == 2\} + \theta_{23} m2^* + \theta_4 c$$

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

```
cmest(data = df_multipleM_noint,  outcome = "contY_catMbinM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C",
                yreg = "linear", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "rb")
```

```
##                Estimate  Std.error    95% CIL    95% CIU      z   pval
## cde          8.020e-01  4.171e-03  7.939e-01  8.102e-01 192.31 <2e-16 ***
## pnde         8.020e-01  4.171e-03  7.939e-01  8.102e-01 192.31 <2e-16 ***
## tnde         8.020e-01  4.171e-03  7.939e-01  8.102e-01 192.31 <2e-16 ***
## pnie         3.555e-01  2.018e-02  3.159e-01  3.950e-01  17.61 <2e-16 ***
## tnie         3.555e-01  2.018e-02  3.159e-01  3.950e-01  17.61 <2e-16 ***
## te           1.158e+00  2.061e-02  1.117e+00  1.198e+00  56.17 <2e-16 ***
## pm           1.814e-01  8.505e-03  1.647e-01  1.981e-01  21.33 <2e-16 ***
## intref       0.000e+00  4.452e-17 -8.726e-17  8.726e-17   0.00      1
## intmed       0.000e+00  4.452e-17 -8.726e-17  8.726e-17   0.00      1
## pie          3.555e-01  2.018e-02  3.159e-01  3.950e-01  17.61 <2e-16 ***
## cde_prop     6.929e-01  1.223e-02  6.689e-01  7.169e-01  56.68 <2e-16 ***
## intref_prop  0.000e+00  3.759e-17 -7.367e-17  7.367e-17   0.00      1
## intmed_prop  0.000e+00  3.759e-17 -7.367e-17  7.367e-17   0.00      1
## pie_prop     3.071e-01  1.223e-02  2.831e-01  3.311e-01  25.12 <2e-16 ***
## overall_pm   3.071e-01  1.223e-02  2.831e-01  3.311e-01  25.12 <2e-16 ***
## overall_int  0.000e+00  0.000e+00  0.000e+00  0.000e+00     NA     NA
## overall_pe   3.071e-01  1.223e-02  2.831e-01  3.311e-01  25.12 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_noint,  outcome = "contY_catMbinM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C",
                yreg = "linear", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "wb")
```

```
##               Estimate  Std.error    95% CIL    95% CIU       z   pval
## cde          8.021e-01  4.511e-03  7.932e-01  8.109e-01 177.805 <2e-16 ***
## pnde         8.020e-01  4.514e-03  7.932e-01  8.109e-01 177.670 <2e-16 ***
## tnde         8.020e-01  4.511e-03  7.932e-01  8.109e-01 177.790 <2e-16 ***
## pnie         3.545e-01  2.151e-02  3.123e-01  3.966e-01  16.479 <2e-16 ***
## tnie         3.545e-01  2.151e-02  3.123e-01  3.966e-01  16.483 <2e-16 ***
## te           1.156e+00  2.144e-02  1.114e+00  1.199e+00  53.937 <2e-16 ***
## pm           1.810e-01  9.117e-03  1.631e-01  1.989e-01  19.853 <2e-16 ***
## intref      -3.658e-05  1.116e-04 -2.553e-04  1.822e-04  -0.328  0.743
## intmed       2.162e-05  4.475e-05 -6.608e-05  1.093e-04   0.483  0.629
## pie          3.545e-01  2.151e-02  3.123e-01  3.966e-01  16.479 <2e-16 ***
## cde_prop     6.935e-01  1.307e-02  6.679e-01  7.191e-01  53.045 <2e-16 ***
## intref_prop -3.163e-05  9.645e-05 -2.207e-04  1.574e-04  -0.328  0.743
## intmed_prop  1.870e-05  3.881e-05 -5.738e-05  9.477e-05   0.482  0.630
## pie_prop     3.065e-01  1.307e-02  2.809e-01  3.321e-01  23.450 <2e-16 ***
## overall_pm   3.065e-01  1.307e-02  2.809e-01  3.321e-01  23.457 <2e-16 ***
## overall_int -1.294e-05  6.919e-05 -1.485e-04  1.227e-04  -0.187  0.852
## overall_pe   3.065e-01  1.307e-02  2.809e-01  3.321e-01  23.441 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_noint,  outcome = "contY_catMbinM_noint", exposure = "A",
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C",
                yreg = "linear", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "msm")
```

```
##               Estimate  Std.error    95% CIL    95% CIU       z   pval
## cde          8.032e-01  4.552e-03  7.943e-01  8.121e-01 176.443 <2e-16 ***
## pnde         8.021e-01  4.105e-03  7.941e-01  8.102e-01 195.379 <2e-16 ***
## tnde         8.021e-01  4.105e-03  7.941e-01  8.102e-01 195.379 <2e-16 ***
## pnie         3.560e-01  2.113e-02  3.146e-01  3.974e-01  16.846 <2e-16 ***
## tnie         3.560e-01  2.113e-02  3.146e-01  3.974e-01  16.846 <2e-16 ***
## te           1.158e+00  2.178e-02  1.115e+00  1.201e+00  53.173 <2e-16 ***
## pm           1.816e-01  8.794e-03  1.644e-01  1.989e-01  20.652 <2e-16 ***
## intref      -1.071e-03  1.986e-03 -4.963e-03  2.821e-03  -0.539   0.59
## intmed       0.000e+00  5.412e-17 -1.061e-16  1.061e-16   0.000   1.00
## pie          3.560e-01  2.113e-02  3.146e-01  3.974e-01  16.846 <2e-16 ***
## cde_prop     6.935e-01  1.276e-02  6.685e-01  7.185e-01  54.359 <2e-16 ***
## intref_prop -9.250e-04  1.715e-03 -4.287e-03  2.437e-03  -0.539   0.59
## intmed_prop  0.000e+00  4.666e-17 -9.145e-17  9.145e-17   0.000   1.00
```

```
## pie_prop     3.074e-01  1.259e-02  2.827e-01  3.321e-01  24.413 <2e-16 ***
## overall_pm   3.074e-01  1.259e-02  2.827e-01  3.321e-01  24.413 <2e-16 ***
## overall_int -9.250e-04  1.715e-03 -4.287e-03  2.437e-03  -0.539   0.59
## overall_pe   3.065e-01  1.276e-02  2.815e-01  3.315e-01  24.022 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_noint,  outcome = "contY_catMbinM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C",
                yreg = "linear", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "paramfunc", inf.method = "bootstrap", model = "iorw")
```

```
##       Estimate Std.error 95% CIL 95% CIU      z   pval
## ORtot  3.18166   0.06812 3.04814 3.31518  46.70 <2e-16 ***
## ORdir  2.23350   0.01068 2.21258 2.25443 209.17 <2e-16 ***
## ORind  1.42452   0.02965 1.36640 1.48263  48.04 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_noint,  outcome = "contY_catMbinM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C",
                yreg = "linear", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "g-formula")
```

```
##               Estimate  Std.error    95% CIL    95% CIU      z   pval
## cde          8.020e-01  4.067e-03  7.941e-01  8.100e-01 197.20 <2e-16 ***
## pnde         8.020e-01  4.067e-03  7.941e-01  8.100e-01 197.20 <2e-16 ***
## tnde         8.020e-01  4.067e-03  7.941e-01  8.100e-01 197.20 <2e-16 ***
## pnie         3.555e-01  1.912e-02  3.180e-01  3.929e-01  18.59 <2e-16 ***
## tnie         3.555e-01  1.912e-02  3.180e-01  3.929e-01  18.59 <2e-16 ***
## te           1.158e+00  1.989e-02  1.119e+00  1.196e+00  58.20 <2e-16 ***
## pm           1.814e-01  8.002e-03  1.657e-01  1.971e-01  22.67 <2e-16 ***
## intref       0.000e+00  3.140e-17 -6.155e-17  6.155e-17   0.00      1
## intmed       0.000e+00  6.296e-17 -1.234e-16  1.234e-16   0.00      1
## pie          3.555e-01  1.912e-02  3.180e-01  3.929e-01  18.59 <2e-16 ***
## cde_prop     6.929e-01  1.152e-02  6.703e-01  7.155e-01  60.17 <2e-16 ***
## intref_prop  0.000e+00  2.761e-17 -5.412e-17  5.412e-17   0.00      1
## intmed_prop  0.000e+00  5.460e-17 -1.070e-16  1.070e-16   0.00      1
## pie_prop     3.071e-01  1.152e-02  2.845e-01  3.297e-01  26.67 <2e-16 ***
## overall_pm   3.071e-01  1.152e-02  2.845e-01  3.297e-01  26.67 <2e-16 ***
## overall_int  0.000e+00  4.703e-17 -9.217e-17  9.217e-17   0.00      1
## overall_pe   3.071e-01  1.152e-02  2.845e-01  3.297e-01  26.67 <2e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_noint,  outcome = "contY_catMbinM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C",
                yreg = "linear", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1, model = "ne")
```

```
##                          Estimate Std.error  95% CIL  95% CIU     z   pval
## natural direct effect   0.802049  0.004141 0.793933 0.810165 193.7 <2e-16
## natural indirect effect 0.355354  0.020909 0.314373 0.396336  17.0 <2e-16
## total effect            1.157404  0.021275 1.115705 1.199102  54.4 <2e-16
##
## natural direct effect   ***
## natural indirect effect ***
## total effect            ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

4.2 Case 4-2: Continuous Outcome and Multiple Mediators With Exposure-mediator Interaction

4.2.1 Data simulation

4.2.1.1 Simulation Procedures

1. Simulate the exposure variable A from $Bernoulli(P(A=1))$.
2. Simulate the covariate C from $N(\mu_C, \sigma_C^2)$.
3. Simulate the first mediator M1 from $Multinom(\frac{1}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}$, $\frac{expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}$, $\frac{expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)})$, the second mediator M2 from $Bernoulli(expit(\beta_{03}+\beta_{13}*A+\beta_{23}*C))$.
4. Simulate the outcome Y from $N(\theta_0+\theta_1 A+\theta_{21}I\{M1==1\}+\theta_{22}I\{M1==2\}+\theta_{23}M2+\theta_{31}AM2+\theta_4 C, \sigma_Y^2)$.

4.2.1.2 True Parameters

Table 8: True Model Parameters for Data Simulation

Sample Size	θ_0	θ_1	θ_{21}	θ_{22}	θ_{23}	θ_{31}	θ_4	β_{01}	β_{11}	β_{21}
10000	-5	0.8	1.8	1.2	1.5	0.2	0.1	-0.25	0.5	0.2
β_{02}	β_{12}	β_{22}	β_{03}	β_{13}	β_{23}	P(A=1)	μ_C	σ_C	σ_Y	
-0.3	0.4	0.3	-0.25	0.5	0.2	0.4	1	1	0.2	

4.2.1.3 True Models

True model for the first mediator:

$$ln\frac{P(M1==1)}{P(M1==0)} = \beta_{01}+\beta_{11}a+\beta_{21}c$$

$$ln\frac{P(M1==2)}{P(M1==0)} = \beta_{02} + \beta_{12}a + \beta_{22}c$$

True model for the second mediator:

$$logitE[M2|a,c] = \beta_{03} + \beta_{13}a + \beta_{23}c$$

True model for the outcome:

$$E[Y|a,m^*,c] = \theta_0 + \theta_1 a + \theta_{21}I\{m1^* == 1\} + \theta_{22}I\{m1^* == 2\} + \theta_{23}m2^* + \theta_{31}am2^* + \theta_4 c$$

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

```
cmest(data = df_multipleM_EMint, outcome = "contY_catMbinM_EMint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C",
                EMint = TRUE, EMint.terms = c("A*M_bin"),
                yreg = "linear", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "rb")
```

```
##             Estimate Std.error  95% CIL  95% CIU      z   pval
## cde         0.796559  0.006251 0.784308 0.808811 127.43 <2e-16 ***
## pnde        0.901103  0.004438 0.892404 0.909801 203.04 <2e-16 ***
## tnde        0.926364  0.004477 0.917588 0.935139 206.90 <2e-16 ***
## pnie        0.354886  0.022813 0.310174 0.399599  15.56 <2e-16 ***
## tnie        0.380147  0.024575 0.331981 0.428313  15.47 <2e-16 ***
## te          1.281250  0.024341 1.233543 1.328957  52.64 <2e-16 ***
## pm          0.174191  0.009470 0.155631 0.192752  18.39 <2e-16 ***
## intref      0.104543  0.004247 0.096219 0.112868  24.61 <2e-16 ***
## intmed      0.025261  0.002501 0.020358 0.030163  10.10 <2e-16 ***
## pie         0.354886  0.022813 0.310174 0.399599  15.56 <2e-16 ***
## cde_prop    0.621705  0.012371 0.597457 0.645952  50.25 <2e-16 ***
## intref_prop 0.081595  0.003846 0.074057 0.089133  21.22 <2e-16 ***
## intmed_prop 0.019716  0.001725 0.016335 0.023096  11.43 <2e-16 ***
## pie_prop    0.276984  0.012763 0.251970 0.301999  21.70 <2e-16 ***
## overall_pm  0.296700  0.013789 0.269675 0.323725  21.52 <2e-16 ***
## overall_int 0.101311  0.004237 0.093007 0.109615  23.91 <2e-16 ***
## overall_pe  0.378295  0.012371 0.354048 0.402543  30.58 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_EMint,  outcome = "contY_catMbinM_EMint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C",
                EMint = TRUE, EMint.terms = c("A*M_bin"),
                yreg = "linear", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "wb")
```

```
##             Estimate Std.error  95% CIL  95% CIU      z   pval
## cde         0.796560  0.006192 0.784424 0.808696 128.64 <2e-16 ***
## pnde        0.901312  0.004055 0.893364 0.909260 222.27 <2e-16 ***
## tnde        0.926547  0.004260 0.918197 0.934897 217.49 <2e-16 ***
## pnie        0.353887  0.020966 0.312794 0.394980  16.88 <2e-16 ***
## tnie        0.379122  0.022541 0.334943 0.423300  16.82 <2e-16 ***
## te          1.280434  0.022321 1.236685 1.324182  57.36 <2e-16 ***
## pm          0.173770  0.008646 0.156824 0.190716  20.10 <2e-16 ***
## intref      0.104752  0.004267 0.096388 0.113115  24.55 <2e-16 ***
## intmed      0.025235  0.002420 0.020491 0.029979  10.43 <2e-16 ***
## pie         0.353887  0.020966 0.312794 0.394980  16.88 <2e-16 ***
## cde_prop    0.622102  0.011239 0.600073 0.644131  55.35 <2e-16 ***
## intref_prop 0.081810  0.003837 0.074290 0.089329  21.32 <2e-16 ***
## intmed_prop 0.019708  0.001686 0.016404 0.023012  11.69 <2e-16 ***
## pie_prop    0.276380  0.011665 0.253517 0.299244  23.69 <2e-16 ***
## overall_pm  0.296088  0.012553 0.271486 0.320691  23.59 <2e-16 ***
## overall_int 0.101518  0.004305 0.093080 0.109955  23.58 <2e-16 ***
## overall_pe  0.377898  0.011239 0.355869 0.399927  33.62 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_EMint,  outcome = "contY_catMbinM_EMint", exposure = "A",
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C",
                EMint = TRUE, EMint.terms = c("A*M_bin"),
                yreg = "linear", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "msm")
```

```
##             Estimate Std.error  95% CIL  95% CIU      z   pval
## cde         0.800461  0.006923 0.786892 0.814029 115.63 <2e-16 ***
## pnde        0.900949  0.003867 0.893369 0.908528 232.99 <2e-16 ***
## tnde        0.926128  0.004123 0.918048 0.934208 224.65 <2e-16 ***
## pnie        0.355152  0.021745 0.312533 0.397771  16.33 <2e-16 ***
## tnie        0.380332  0.023300 0.334664 0.426000  16.32 <2e-16 ***
## te          1.281280  0.023195 1.235820 1.326741  55.24 <2e-16 ***
## pm          0.174286  0.008865 0.156911 0.191661  19.66 <2e-16 ***
## intref      0.100488  0.005655 0.089404 0.111572  17.77 <2e-16 ***
## intmed      0.025180  0.002397 0.020481 0.029878  10.50 <2e-16 ***
## pie         0.355152  0.021745 0.312533 0.397771  16.33 <2e-16 ***
## cde_prop    0.624735  0.012060 0.601098 0.648372  51.80 <2e-16 ***
## intref_prop 0.078428  0.004702 0.069213 0.087643  16.68 <2e-16 ***
## intmed_prop 0.019652  0.001655 0.016409 0.022895  11.88 <2e-16 ***
## pie_prop    0.277185  0.011977 0.253711 0.300660  23.14 <2e-16 ***
## overall_pm  0.296837  0.012825 0.271701 0.321973  23.15 <2e-16 ***
## overall_int 0.098079  0.005177 0.087933 0.108226  18.95 <2e-16 ***
## overall_pe  0.375265  0.012060 0.351628 0.398902  31.12 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_EMint,  outcome = "contY_catMbinM_EMint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C",
                EMint = TRUE, EMint.terms = c("A*M_bin"),
                yreg = "linear", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "paramfunc", inf.method = "bootstrap", model = "iorw")
```

```
##       Estimate Std.error 95% CIL 95% CIU      z   pval
## ORtot  3.60093   0.07590 3.45218 3.74969  47.45 <2e-16 ***
## ORdir  2.53165   0.01272 2.50671 2.55659 198.97 <2e-16 ***
## ORind  1.42237   0.02869 1.36614 1.47859  49.58 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_EMint,  outcome = "contY_catMbinM_EMint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C",
                EMint = TRUE, EMint.terms = c("A*M_bin"),
                yreg = "linear", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "g-formula")
```

```
##             Estimate Std.error  95% CIL  95% CIU      z   pval
## cde         0.796559  0.005989 0.784821 0.808297 133.01 <2e-16 ***
## pnde        0.901103  0.004489 0.892304 0.909902 200.72 <2e-16 ***
## tnde        0.926364  0.004321 0.917894 0.934834 214.36 <2e-16 ***
## pnie        0.354886  0.021100 0.313530 0.396242  16.82 <2e-16 ***
## tnie        0.380147  0.022541 0.335967 0.424327  16.86 <2e-16 ***
## te          1.281250  0.021833 1.238457 1.324043  58.68 <2e-16 ***
## pm          0.174191  0.008729 0.157082 0.191300  19.95 <2e-16 ***
## intref      0.104543  0.004132 0.096446 0.112641  25.30 <2e-16 ***
## intmed      0.025261  0.002272 0.020808 0.029714  11.12 <2e-16 ***
## pie         0.354886  0.021100 0.313530 0.396242  16.82 <2e-16 ***
## cde_prop    0.621705  0.011161 0.599830 0.643580  55.70 <2e-16 ***
## intref_prop 0.081595  0.003755 0.074235 0.088955  21.73 <2e-16 ***
## intmed_prop 0.019716  0.001581 0.016616 0.022815  12.47 <2e-16 ***
## pie_prop    0.276984  0.011862 0.253735 0.300234  23.35 <2e-16 ***
## overall_pm  0.296700  0.012654 0.271900 0.321501  23.45 <2e-16 ***
## overall_int 0.101311  0.004083 0.093308 0.109313  24.81 <2e-16 ***
## overall_pe  0.378295  0.011161 0.356420 0.400170  33.90 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_EMint,  outcome = "contY_catMbinM_EMint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C",
                EMint = TRUE, EMint.terms = c("A*M_bin"),
                yreg = "linear", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1, model = "ne")
```

```
##                       Estimate Std.error  95% CIL  95% CIU      z   pval
## pure direct effect    0.901376  0.004330 0.892888 0.909863 208.15 <2e-16
## total direct effect   0.925949  0.004431 0.917263 0.934634 208.95 <2e-16
## pure indirect effect  0.355041  0.020884 0.314110 0.395973  17.00 <2e-16
## total indirect effect 0.379614  0.022431 0.335651 0.423577  16.92 <2e-16
## total effect          1.280990  0.022162 1.237554 1.324426  57.80 <2e-16
##
## pure direct effect    ***
## total direct effect   ***
## pure indirect effect  ***
## total indirect effect ***
## total effect          ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

4.3 Case 4-3: Continuous Outcome and Multiple Mediators With Mediator-mediator Interaction

4.3.1 Data simulation

4.3.1.1 Simulation Procedures

1. Simulate the exposure variable A from $Bernoulli(P(A=1))$.
2. Simulate the covariate C from $N(\mu_C, \sigma_C^2)$.
3. Simulate the first mediator M1 from $Multinom(\frac{1}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}, \frac{expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}, \frac{expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)})$, the second mediator M2 from $Bernoulli(expit(\beta_{03}+\beta_{13}*A+\beta_{23}*C))$.
4. Simulate the outcome Y from $N(\theta_0+\theta_1 A+\theta_{21}I\{M1==1\}+\theta_{22}I\{M1==2\}+\theta_{23}M2+\theta_{31}I\{M1==1\}M2+\theta_{32}I\{M1==2\}M2+\theta_4 C, \sigma_Y^2)$.

4.3.1.2 True Parameters

Table 9: True Model Parameters for Data Simulation

Sample Size	θ_0	θ_1	θ_{21}	θ_{22}	θ_{23}	θ_{31}	θ_{32}	θ_4	β_{01}	β_{11}
10000	-5	0.8	1.8	1.2	1.5	0.2	0.4	0.1	-0.25	0.5
β_{21}	β_{02}	β_{12}	β_{22}	β_{03}	β_{13}	β_{23}	P(A=1)	μ_C	σ_C	σ_Y
0.2	-0.3	0.4	0.3	-0.25	0.5	0.2	0.4	1	1	0.2

4.3.1.3 True Models

True model for the first mediator:

$$ln\frac{P(M1==1)}{P(M1==0)} = \beta_{01} + \beta_{11}a + \beta_{21}c$$

$$ln\frac{P(M1==2)}{P(M1==0)} = \beta_{02} + \beta_{12}a + \beta_{22}c$$

True model for the second mediator:

$$logitE[M2|a,c] = \beta_{03} + \beta_{13}a + \beta_{23}c$$

True model for the outcome:

$$E[Y|a,m^*,c] = \theta_0 + \theta_1 a + \theta_{21}I\{m1^*==1\} + \theta_{22}I\{m1^*==2\} + \theta_{23}m2^* +$$

$$\theta_{31}I\{m1^*==1\}m2^* + \theta_{32}I\{m1^*==2\}m2^* + \theta_4 c$$

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

```
cmest(data = df_multipleM_MMint,  outcome = "contY_catMbinM_MMint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C", MMint = TRUE,
                yreg = "linear", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "wb")
```

```
##               Estimate  Std.error    95% CIL    95% CIU       z   pval
## cde          8.035e-01  4.424e-03  7.949e-01  8.122e-01 181.629 <2e-16 ***
## pnde         8.035e-01  4.413e-03  7.949e-01  8.122e-01 182.083 <2e-16 ***
## tnde         8.035e-01  4.417e-03  7.949e-01  8.122e-01 181.905 <2e-16 ***
## pnie         3.940e-01  2.255e-02  3.498e-01  4.382e-01  17.470 <2e-16 ***
## tnie         3.940e-01  2.255e-02  3.498e-01  4.382e-01  17.470 <2e-16 ***
## te           1.198e+00  2.268e-02  1.153e+00  1.242e+00  52.806 <2e-16 ***
## pm           1.969e-01  9.192e-03  1.789e-01  2.149e-01  21.422 <2e-16 ***
## intref      -2.543e-05  8.825e-05 -1.984e-04  1.475e-04  -0.288  0.773
## intmed       1.415e-05  3.250e-05 -4.955e-05  7.785e-05   0.435  0.663
## pie          3.940e-01  2.255e-02  3.498e-01  4.382e-01  17.470 <2e-16 ***
## cde_prop     6.710e-01  1.288e-02  6.458e-01  6.962e-01  52.109 <2e-16 ***
## intref_prop -2.124e-05  7.356e-05 -1.654e-04  1.229e-04  -0.289  0.773
## intmed_prop  1.182e-05  2.714e-05 -4.138e-05  6.502e-05   0.435  0.663
## pie_prop     3.290e-01  1.288e-02  3.038e-01  3.543e-01  25.547 <2e-16 ***
## overall_pm   3.290e-01  1.288e-02  3.038e-01  3.543e-01  25.546 <2e-16 ***
## overall_int -9.422e-06  5.821e-05 -1.235e-04  1.047e-04  -0.162  0.871
## overall_pe   3.290e-01  1.288e-02  3.038e-01  3.542e-01  25.550 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_MMint,  outcome = "contY_catMbinM_MMint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C", MMint = TRUE,
                yreg = "linear", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "paramfunc", inf.method = "bootstrap", model = "iorw")
```

```
##       Estimate Std.error 95% CIL 95% CIU      z   pval
## ORtot  3.31505   0.07849 3.16122 3.46889  42.24 <2e-16 ***
## ORdir  2.23647   0.01301 2.21098 2.26197 171.93 <2e-16 ***
## ORind  1.48227   0.03440 1.41486 1.54969  43.09 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_MMint,  outcome = "contY_catMbinM_MMint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C", MMint = TRUE,
                yreg = "linear", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1, model = "ne")
```

```
##                         Estimate Std.error  95% CIL  95% CIU      z   pval
## natural direct effect   0.803547  0.004083 0.795543 0.811550 196.78 <2e-16
## natural indirect effect 0.394927  0.023289 0.349282 0.440573  16.96 <2e-16
## total effect            1.198474  0.023643 1.152135 1.244813  50.69 <2e-16
##
## natural direct effect   ***
## natural indirect effect ***
## total effect            ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

4.4 Case 4-4: Continuous Outcome and Multiple Mediators With Exposure-mediator-mediator Interaction

4.4.1 Data simulation

4.4.1.1 Simulation Procedures

1. Simulate the exposure variable A from $Bernoulli(P(A=1))$.
2. Simulate the covariate C from $N(\mu_C, \sigma_C^2)$.
3. Simulate the first mediator M1 from $Multinom(\frac{1}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}$, $\frac{expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}$, $\frac{expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)})$, the second mediator M2 from $Bernoulli(expit(\beta_{03}+\beta_{13}*A+\beta_{23}*C))$.
4. Simulate the outcome Y from $N(\theta_0+\theta_1 A+\theta_{21}I\{M1==1\}+\theta_{22}I\{M1==2\}+\theta_{23}M2+\theta_{31}AI\{M1==1\}M2+\theta_{32}AI\{M1==2\}M2+\theta_4 C, \sigma_Y^2)$.

4.4.1.2 True Parameters

Table 10: True Model Parameters for Data Simulation

Sample Size	θ_0	θ_1	θ_{21}	θ_{22}	θ_{23}	θ_{31}	θ_{32}	θ_4	β_{01}	β_{11}
10000	-5	0.8	1.8	1.2	1.5	0.2	0.4	0.1	-0.25	0.5
β_{21}	β_{02}	β_{12}	β_{22}	β_{03}	β_{13}	β_{23}	P(A=1)	μ_C	σ_C	σ_Y
0.2	-0.3	0.4	0.3	-0.25	0.5	0.2	0.4	1	1	0.2

4.4.1.3 True Models

True model for the first mediator:

$$ln\frac{P(M1==1)}{P(M1==0)} = \beta_{01} + \beta_{11}a + \beta_{21}c$$

$$ln\frac{P(M1==2)}{P(M1==0)} = \beta_{02} + \beta_{12}a + \beta_{22}c$$

True model for the second mediator:

$$logitE[M2|a,c] = \beta_{03} + \beta_{13}a + \beta_{23}c$$

True model for the outcome:

$$E[Y|a,m^*,c] = \theta_0 + \theta_1 a + \theta_{21}I\{m1^*==1\} + \theta_{22}I\{m1^*==2\} + \theta_{23}m2^* +$$

$$\theta_{31}aI\{m1^*==1\}m2^* + \theta_{32}aI\{m1^*==2\}m2^* + \theta_4 c$$

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

```
cmest(data = df_multipleM_EMMint,  outcome = "contY_catMbinM_EMMint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C", EMMint = TRUE,
                yreg = "linear", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "wb")
```

```
##             Estimate Std.error  95% CIL  95% CIU      z   pval
## cde         0.802332  0.010491 0.781771 0.822894  76.48 <2e-16 ***
## pnde        0.900090  0.004597 0.891081 0.909099 195.82 <2e-16 ***
## tnde        0.941094  0.004917 0.931458 0.950731 191.41 <2e-16 ***
## pnie        0.353828  0.020201 0.314233 0.393422  17.52 <2e-16 ***
## tnie        0.394832  0.022827 0.350093 0.439572  17.30 <2e-16 ***
## te          1.294922  0.022319 1.251178 1.338666  58.02 <2e-16 ***
## pm          0.179877  0.008751 0.162726 0.197028  20.56 <2e-16 ***
## intref      0.097757  0.009354 0.079423 0.116091  10.45 <2e-16 ***
## intmed      0.041005  0.003740 0.033674 0.048335  10.96 <2e-16 ***
## pie         0.353828  0.020201 0.314233 0.393422  17.52 <2e-16 ***
## cde_prop    0.619599  0.013253 0.593623 0.645575  46.75 <2e-16 ***
## intref_prop 0.075493  0.007429 0.060932 0.090054  10.16 <2e-16 ***
## intmed_prop 0.031666  0.002550 0.026668 0.036663  12.42 <2e-16 ***
## pie_prop    0.273242  0.011120 0.251448 0.295037  24.57 <2e-16 ***
## overall_pm  0.304908  0.012626 0.280162 0.329655  24.15 <2e-16 ***
## overall_int 0.107159  0.007958 0.091560 0.122757  13.46 <2e-16 ***
```

```
## overall_pe  0.380401  0.013253 0.354425 0.406377  28.70 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_EMMint,  outcome = "contY_catMbinM_EMMint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C", EMMint = TRUE,
                yreg = "linear", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "paramfunc", inf.method = "bootstrap", model = "iorw")
```

```
##       Estimate Std.error 95% CIL 95% CIU      z   pval
## ORtot  3.65287   0.08174 3.49267 3.81307  44.69 <2e-16 ***
## ORdir  2.56845   0.01431 2.54040 2.59651 179.45 <2e-16 ***
## ORind  1.42220   0.02855 1.36624 1.47817  49.81 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_EMMint,  outcome = "contY_catMbinM_EMMint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C", EMMint = TRUE,
                yreg = "linear", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = NULL, inf.method = NULL, model = "ne")
```

```
##                         Estimate Std.error  95% CIL  95% CIU     z   pval
## natural direct effect   0.916081  0.004411 0.907436 0.924726 207.7 <2e-16
## natural indirect effect 0.371117  0.021829 0.328333 0.413900  17.0 <2e-16
## total effect            1.287198  0.022272 1.243546 1.330850  57.8 <2e-16
##
## natural direct effect   ***
## natural indirect effect ***
## total effect            ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

5 Case 5: Binary Outcome and Single Continuous Mediator

5.1 Case 5-1: Binary Outcome and Single Continuous Mediator Without Interaction

5.1.1 Data simulation

5.1.1.1 Simulation Procedures

1. Simulate the exposure variable A from $Bernoulli(P(A = 1))$.

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 $Bernoulli(expit(\theta_0 + \theta_1 A + \theta_2 M + \theta_4 C))$.

5.1.1.2 True Parameters

Table 11: 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
10000	-5	0.8	1.8	0.1	-0.25	0.5	0.2	0.4	1	1	0.1

5.1.1.3 True Models

True model for the mediator:

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

True model for the outcome:

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

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

5.1.2.1 Closed-form Parameter Function Estimation and Delta Method Inference

```
cmest(data = df_noint,  outcome = "binY_contM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cont', covariates.pre = "C",
                yreg = "logistic", mreg = "linear", mval = list(0),
                a_star = 0, a = 1,
                est.method = "paramfunc", inf.method = "delta", model = "rb")
```

```
##                  Estimate Std.error   95% CIL   95% CIU      z     pval
## cde_rr           2.539555  1.053615  0.474508  4.604603  2.410  0.01596
## pnde_rr          2.539555  1.053615  0.474508  4.604603  2.410  0.01596
## tnde_rr          2.539555  1.053615  0.474508  4.604603  2.410  0.01596
## pnie_rr          2.207204  0.821827  0.596453  3.817954  2.686  0.00725
## tnie_rr          2.207204  0.821827  0.596453  3.817954  2.686  0.00725
## te_rr            5.605316  1.021037  3.604120  7.606512  5.490 4.12e-08
## pm               0.665700  0.207063  0.259863  1.071537  3.215  0.00131
## cde_err          1.649811  1.089961 -0.486473  3.786096  1.514  0.13015
## intref_err      -0.110256  0.041090 -0.190791 -0.029721 -2.683  0.00730
## intmed_err       1.858557  0.571714  0.738018  2.979095  3.251  0.00115
## pie_err          1.207204  0.821827 -0.403547  2.817954  1.469  0.14188
## te_err           4.605316  1.021037  2.604120  6.606512  4.510 6.54e-06
## cde_err_prop     0.358241  0.212478 -0.058208  0.774689  1.686  0.09182
## intref_err_prop -0.023941  0.005587 -0.034892 -0.012991 -4.285 1.84e-05
## intmed_err_prop  0.403568  0.041575  0.322082  0.485053  9.707  < 2e-16
## pie_err_prop     0.262133  0.187803 -0.105955  0.630220  1.396  0.16281
## overall_pm       0.665700  0.207063  0.259863  1.071537  3.215  0.00131
## overall_int      0.379627  0.043803  0.293775  0.465479  8.667  < 2e-16
## overall_pe       0.641759  0.212478  0.225311  1.058208  3.020  0.00253
##
## cde_rr          *
## pnde_rr         *
```

```
## tnde_rr         *
## pnie_rr         **
## tnie_rr         **
## te_rr           ***
## pm              **
## cde_err
## intref_err      **
## intmed_err      **
## pie_err
## te_err          ***
## cde_err_prop    .
## intref_err_prop ***
## intmed_err_prop ***
## pie_err_prop
## overall_pm      **
## overall_int     ***
## overall_pe      **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

5.1.2.2 Direct Imputation Estimation and Bootstrap Inference

```
cmest(data = df_noint,  outcome = "binY_contM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cont', covariates.pre = "C",
                yreg = "logistic", mreg = "linear", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "rb")
```

```
##                  Estimate Std.error   95% CIL   95% CIU      z     pval
## cde_rr           2.514116  1.320550 -0.074113  5.102346  1.904 0.056960
## pnde_rr          2.511177  1.318838 -0.073697  5.096051  1.904 0.056928
## tnde_rr          2.478532  1.312946 -0.094794  5.051859  1.888 0.059086
## pnie_rr          2.187811  1.022761  0.183236  4.192385  2.139 0.032450
## tnie_rr          2.159370  1.020622  0.158988  4.159751  2.116 0.034392
## te_rr            5.422560  1.191462  3.087337  7.757783  4.551  5.4e-06
## pm               0.658303  0.257280  0.154044  1.162561  2.559 0.010521
## cde_err          1.540318  1.307410 -1.022159  4.102795  1.178 0.238767
## intref_err      -0.029141  0.041640 -0.110753  0.052471 -0.700 0.484044
## intmed_err       1.723572  0.864823  0.028551  3.418593  1.993 0.046291
## pie_err          1.187811  1.022761 -0.816764  3.192385  1.161 0.245516
## te_err           4.422560  1.191462  2.087337  6.757783  3.712 0.000207
## cde_err_prop     0.348287  0.254509 -0.150542  0.847115  1.368 0.171197
## intref_err_prop -0.006589  0.008117 -0.022498  0.009320 -0.812 0.416947
## intmed_err_prop  0.389723  0.144809  0.105903  0.673542  2.691 0.007129
## pie_err_prop     0.268580  0.262754 -0.246408  0.783568  1.022 0.306724
## overall_pm       0.658303  0.257280  0.154044  1.162561  2.559 0.010521
## overall_int      0.383134  0.142247  0.104334  0.661933  2.693 0.007084
## overall_pe       0.651713  0.254509  0.152885  1.150542  2.561 0.010462
##
## cde_rr          .
## pnde_rr         .
## tnde_rr         .
## pnie_rr         *
```

```
## tnie_rr         *
## te_rr           ***
## pm              *
## cde_err
## intref_err
## intmed_err      *
## pie_err
## te_err          ***
## cde_err_prop
## intref_err_prop
## intmed_err_prop **
## pie_err_prop
## overall_pm      *
## overall_int     **
## overall_pe      *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_noint,  outcome = "binY_contM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cont', covariates.pre = "C",
                yreg = "logistic", mreg = "linear", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "wb")
```

```
##                   Estimate  Std.error    95% CIL    95% CIU      z
## cde_rr           2.5140993  1.3206169 -0.0742622  5.1024607  1.904
## pnde_rr          2.5123544  1.3215825 -0.0778996  5.1026084  1.901
## tnde_rr          2.4758078  1.3175951 -0.1066312  5.0582467  1.879
## pnie_rr          2.1948553  0.9322409  0.3676968  4.0220139  2.354
## tnie_rr          2.1629273  0.9224973  0.3548658  3.9709889  2.345
## te_rr            5.4340399  1.1617616  3.1570291  7.7110508  4.677
## pm               0.6589218  0.2543604  0.1603845  1.1574591  2.591
## cde_err          1.5162884  1.3165869 -1.0641746  4.0967514  1.152
## intref_err      -0.0039340  0.0423064 -0.0868529  0.0789850 -0.093
## intmed_err       1.7268302  0.8310658  0.0979711  3.3556893  2.078
## pie_err          1.1948553  0.9322409 -0.6323032  3.0220139  1.282
## te_err           4.4340399  1.1617616  2.1570291  6.7110508  3.817
## cde_err_prop     0.3419654  0.2532649 -0.1544247  0.8383556  1.350
## intref_err_prop -0.0008872  0.0084301 -0.0174098  0.0156354 -0.105
## intmed_err_prop  0.3894485  0.1370367  0.1208615  0.6580355  2.842
## pie_err_prop     0.2694733  0.2254018 -0.1723060  0.7112526  1.196
## overall_pm       0.6589218  0.2543604  0.1603845  1.1574591  2.591
## overall_int      0.3885613  0.1375409  0.1189860  0.6581365  2.825
## overall_pe       0.6580346  0.2532649  0.1616444  1.1544247  2.598
##                      pval
## cde_rr           0.056974 .
## pnde_rr          0.057328 .
## tnde_rr          0.060269 .
## pnie_rr          0.018573 *
## tnie_rr          0.019065 *
```

```
## te_rr           2.94e-06 ***
## pm              0.009597 **
## cde_err         0.249480
## intref_err      0.925915
## intmed_err      0.037749 *
## pie_err         0.199977
## te_err          0.000136 ***
## cde_err_prop    0.176973
## intref_err_prop 0.916183
## intmed_err_prop 0.004493 **
## pie_err_prop    0.231911
## overall_pm      0.009597 **
## overall_int     0.004737 **
## overall_pe      0.009385 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_noint,  outcome = "binY_contM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cont', covariates.pre = "C",
                yreg = "logistic", mreg = "linear", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "g-formula")
```

```
##                  Estimate Std.error   95% CIL   95% CIU      z     pval
## cde_rr           2.514116  1.268190  0.028510  4.999723  1.982  0.04746
## pnde_rr          2.511177  1.266897  0.028104  4.994250  1.982  0.04749
## tnde_rr          2.478532  1.262809  0.003473  4.953592  1.963  0.04971
## pnie_rr          2.187811  1.063586  0.103221  4.272401  2.057  0.03971
## tnie_rr          2.159370  1.058971  0.083826  4.234914  2.039  0.04146
## te_rr            5.422560  1.031283  3.401282  7.443838  5.258 1.49e-07
## pm               0.658303  0.256110  0.156337  1.160269  2.570  0.01017
## cde_err          1.540318  1.253717 -0.916922  3.997559  1.229  0.21925
## intref_err      -0.029141  0.038491 -0.104582  0.046299 -0.757  0.44901
## intmed_err       1.723572  0.745393  0.262629  3.184515  2.312  0.02078
## pie_err          1.187811  1.063586 -0.896779  3.272401  1.117  0.26411
## te_err           4.422560  1.031283  2.401282  6.443838  4.288 1.82e-05
## cde_err_prop     0.348287  0.252774 -0.147140  0.843714  1.378  0.16828
## intref_err_prop -0.006589  0.008041 -0.022349  0.009171 -0.819  0.41254
## intmed_err_prop  0.389723  0.127453  0.139919  0.639526  3.058  0.00224
## pie_err_prop     0.268580  0.236398 -0.194752  0.731912  1.136  0.25593
## overall_pm       0.658303  0.256110  0.156337  1.160269  2.570  0.01017
## overall_int      0.383134  0.125934  0.136307  0.629960  3.042  0.00235
## overall_pe       0.651713  0.252774  0.156286  1.147140  2.578  0.00994
##
## cde_rr          *
## pnde_rr         *
## tnde_rr         *
## pnie_rr         *
## tnie_rr         *
## te_rr           ***
```

```
## pm              *
## cde_err
## intref_err
## intmed_err      *
## pie_err
## te_err          ***
## cde_err_prop
## intref_err_prop
## intmed_err_prop **
## pie_err_prop
## overall_pm      *
## overall_int     **
## overall_pe      **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_noint,  outcome = "binY_contM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cont', covariates.pre = "C",
                yreg = "logistic", mreg = "linear", mval = list(0),
                a_star = 0, a = 1, model = "ne")
```

```
##                         Estimate Std.error 95% CIL 95% CIU     z     pval
## natural direct effect     2.5382    1.0921  0.3977  4.6787 2.324   0.0201
## natural indirect effect   2.2110    0.8670  0.5117  3.9103 2.550   0.0108
## total effect              5.6119    1.0218  3.6091  7.6146 5.492 4.07e-08
##
## natural direct effect   *
## natural indirect effect *
## total effect            ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

5.2 Case 5-2: Binary Outcome and Single Continuous Mediator With Exposure-mediator Interaction

5.2.1 Data simulation

5.2.1.1 Simulation Procedures

1. Simulate the exposure variable A from $Bernoulli(P(A = 1))$.
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 $Bernoulli(expit(\theta_0 + \theta_1 A + \theta_2 M + \theta_3 AM + \theta_4 C))$.

5.2.1.2 True Parameters

Table 12: 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
10000	-5	0.8	1.8	0.2	0.1	-0.25	0.5	0.2	0.4	1	1	0.1

5.2.1.3 True Models

True model for the mediator:

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

True model for the outcome:

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

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

5.2.2.1 Closed-form Parameter Function Estimation and Delta Method Inference

```
cmest(data = df_int,  outcome = "binY_contM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cont', covariates.pre = "C", EMint = TRUE,
                yreg = "logistic", mreg = "linear", mval = list(0),
                a_star = 0, a = 1,
                est.method = "paramfunc", inf.method = "delta", model = "rb")
```

```
##                  Estimate Std.error   95% CIL   95% CIU      z     pval
## cde_rr           1.637148  0.650999  0.361212  2.913083  2.515 0.011925
## pnde_rr          1.595450  0.637950  0.345092  2.845808  2.501 0.012404
## tnde_rr          2.476739  1.133092  0.255920  4.697558  2.186 0.028852
## pnie_rr          2.405705  1.111549  0.227109  4.584301  2.164 0.030466
## tnie_rr          3.734561  1.291994  1.202300  6.266822  2.891 0.003854
## te_rr            5.958304  1.069280  3.862554  8.054054  5.572 2.58e-08
## pm               0.879909  0.117740  0.649143  1.110674  7.473 8.47e-14
## cde_err          0.686871  0.692393 -0.670195  2.043938  0.992 0.321209
## intref_err      -0.091422  0.060172 -0.209356  0.026513 -1.519 0.128707
## intmed_err       2.957149  1.008148  0.981216  4.933082  2.933 0.003362
## pie_err          1.405705  1.111549 -0.772891  3.584301  1.265 0.206031
## te_err           4.958304  1.069280  2.862554  7.054054  4.637 3.58e-06
## cde_err_prop     0.138529  0.127336 -0.111045  0.388104  1.088 0.276664
## intref_err_prop -0.018438  0.011099 -0.040191  0.003315 -1.661 0.096687
## intmed_err_prop  0.596403  0.178219  0.247100  0.945707  3.346 0.000822
## pie_err_prop     0.283505  0.219040 -0.145806  0.712817  1.294 0.195590
## overall_pm       0.879909  0.117740  0.649143  1.110674  7.473 8.47e-14
## overall_int      0.577965  0.172322  0.240220  0.915710  3.354 0.000800
## overall_pe       0.861471  0.127336  0.611896  1.111045  6.765 1.40e-11
##
## cde_rr           *
## pnde_rr          *
## tnde_rr          *
## pnie_rr          *
## tnie_rr          **
## te_rr            ***
## pm               ***
## cde_err
## intref_err
## intmed_err       **
```

```
## pie_err
## te_err          ***
## cde_err_prop
## intref_err_prop .
## intmed_err_prop ***
## pie_err_prop
## overall_pm      ***
## overall_int     ***
## overall_pe      ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

5.2.2.2 Direct Imputation Estimation and Bootstrap Inference

```
cmest(data = df_int,  outcome = "binY_contM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cont', covariates.pre = "C", EMint = TRUE,
                yreg = "logistic", mreg = "linear", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "rb")
```

```
##                  Estimate Std.error   95% CIL   95% CIU     z     pval
## cde_rr           1.629434  0.655383  0.344908  2.913960 2.486  0.01293 *
## pnde_rr          1.680085  0.644555  0.416780  2.943390 2.607  0.00916 **
## tnde_rr          2.528003  1.270203  0.038451  5.017556 1.990  0.04659 *
## pnie_rr          2.378053  1.238692 -0.049738  4.805844 1.920  0.05491 .
## tnie_rr          3.578228  1.399740  0.834788  6.321667 2.556  0.01059 *
## te_rr            6.011726  1.093262  3.868972  8.154480 5.499 3.92e-08 ***
## pm               0.864301  0.113639  0.641574  1.087029 7.606 3.09e-14 ***
## cde_err          0.644766  0.644821 -0.619060  1.908593 1.000  0.31738
## intref_err       0.035319  0.053468 -0.069476  0.140113 0.661  0.50891
## intmed_err       2.953588  1.294920  0.415592  5.491585 2.281  0.02258 *
## pie_err          1.378053  1.238692 -1.049738  3.805844 1.113  0.26595
## te_err           5.011726  1.093262  2.868972  7.154480 4.584 4.61e-06 ***
## cde_err_prop     0.128652  0.114580 -0.095922  0.353225 1.123  0.26155
## intref_err_prop  0.007047  0.010453 -0.013440  0.027534 0.674  0.50020
## intmed_err_prop  0.589336  0.224201  0.149909  1.028762 2.629  0.00859 **
## pie_err_prop     0.274966  0.254181 -0.223219  0.773151 1.082  0.27938
## overall_pm       0.864301  0.113639  0.641574  1.087029 7.606 3.09e-14 ***
## overall_int      0.596383  0.229243  0.147075  1.045690 2.602  0.00929 **
## overall_pe       0.871348  0.114580  0.646775  1.095922 7.605 3.11e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_int,  outcome = "binY_contM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cont', covariates.pre = "C", EMint = TRUE,
                yreg = "logistic", mreg = "linear", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "wb")
```

```
##                  Estimate Std.error   95% CIL   95% CIU     z     pval
## cde_rr           1.629432  0.650814  0.353860  2.905004 2.504  0.01231 *
## pnde_rr          1.714636  0.658979  0.423061  3.006210 2.602  0.00928 **
## tnde_rr          2.576256  1.359754 -0.088812  5.241325 1.895  0.05817 .
## pnie_rr          2.387907  1.285833 -0.132278  4.908093 1.857  0.06333 .
## tnie_rr          3.587853  1.225306  1.186298  5.989409 2.928  0.00342 **
## te_rr            6.151862  1.069383  4.055909  8.247815 5.753 9.04e-09 ***
## pm               0.861286  0.117133  0.631710  1.090861 7.353 2.09e-13 ***
## cde_err          0.633184  0.639593 -0.620395  1.886763 0.990  0.32221
## intref_err       0.081452  0.065750 -0.047417  0.210320 1.239  0.21545
## intmed_err       3.049319  1.382955  0.338776  5.759861 2.205  0.02748 *
## pie_err          1.387907  1.285833 -1.132278  3.908093 1.079  0.28044
## te_err           5.151862  1.069383  3.055909  7.247815 4.818 1.47e-06 ***
## cde_err_prop     0.122904  0.114882 -0.102260  0.348068 1.070  0.28472
## intref_err_prop  0.015810  0.012033 -0.007775  0.039395 1.314  0.18893
## intmed_err_prop  0.591887  0.218720  0.163203  1.020570 2.706  0.00682 **
## pie_err_prop     0.269399  0.261820 -0.243759  0.782557 1.029  0.30353
## overall_pm       0.861286  0.117133  0.631710  1.090861 7.353 2.09e-13 ***
## overall_int      0.607697  0.227067  0.162654  1.052740 2.676  0.00746 **
## overall_pe       0.877096  0.114882  0.651932  1.102260 7.635 2.47e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_int,  outcome = "binY_contM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cont', covariates.pre = "C", EMint = TRUE,
                yreg = "logistic", mreg = "linear", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "g-formula")
```

```
##                  Estimate Std.error   95% CIL   95% CIU     z     pval
## cde_rr           1.629434  0.665228  0.325611  2.933257 2.449  0.01432 *
## pnde_rr          1.680085  0.662845  0.380932  2.979237 2.535  0.01127 *
## tnde_rr          2.528003  1.196551  0.182806  4.873201 2.113  0.03465 *
## pnie_rr          2.378053  1.096430  0.229089  4.527017 2.169  0.03011 *
## tnie_rr          3.578228  1.158233  1.308132  5.848323 3.089  0.00201 **
## te_rr            6.011726  1.110957  3.834291  8.189161 5.411 6.40e-08 ***
## pm               0.864301  0.111464  0.645835  1.082767 7.754 9.77e-15 ***
## cde_err          0.644766  0.657824 -0.644545  1.934078 0.980  0.32704
## intref_err       0.035319  0.052956 -0.068474  0.139111 0.667  0.50483
## intmed_err       2.953588  1.210626  0.580805  5.326372 2.440  0.01472 *
## pie_err          1.378053  1.096430 -0.770911  3.527017 1.257  0.20884
## te_err           5.011726  1.110957  2.834291  7.189161 4.511 6.52e-06 ***
## cde_err_prop     0.128652  0.111774 -0.090421  0.347724 1.151  0.24976
## intref_err_prop  0.007047  0.010195 -0.012935  0.027030 0.691  0.48944
## intmed_err_prop  0.589336  0.193850  0.209396  0.969275 3.040  0.00237 **
## pie_err_prop     0.274966  0.219333 -0.154919  0.704850 1.254  0.21000
## overall_pm       0.864301  0.111464  0.645835  1.082767 7.754 9.77e-15 ***
## overall_int      0.596383  0.198848  0.206647  0.986118 2.999  0.00271 **
## overall_pe       0.871348  0.111774  0.652276  1.090421 7.796 7.05e-15 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_int,  outcome = "binY_contM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cont', covariates.pre = "C", EMint = TRUE,
                yreg = "logistic", mreg = "linear", mval = list(0),
                a_star = 0, a = 1, model = "ne")
```

```
##                       Estimate Std.error 95% CIL 95% CIU     z     pval
## pure direct effect     1.71431   0.65648 0.42763 3.00099 2.611  0.00903
## total direct effect    2.66570   1.33590 0.04739 5.28401 1.995  0.04602
## pure indirect effect   2.40589   1.11673 0.21714 4.59465 2.154  0.03123
## total indirect effect  3.74109   1.23311 1.32425 6.15793 3.034  0.00242
## total effect           6.41339   1.09243 4.27227 8.55452 5.871 4.48e-09
##
## pure direct effect    **
## total direct effect   *
## pure indirect effect  *
## total indirect effect **
## total effect          ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

6 Case 6: Binary Outcome and Single Binary Mediator

6.1 Case 6-1: Binary Outcome and Single Binary Mediator Without Interaction

6.1.1 Data simulation

6.1.1.1 Simulation Procedures

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

6.1.1.2 True Parameters

Table 13: 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
10000	-5	0.8	1.8	0.1	-0.25	0.5	0.2	0.4	1	1

6.1.1.3 True Models

True model for the mediator:

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

True model for the outcome:

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

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

6.1.2.1 Closed-form Parameter Function Estimation and Delta Method Inference

```
cmest(data = df_noint,  outcome = "binY_binM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_bin', covariates.pre = "C",
                yreg = "logistic", mreg = "logistic", mval = list(0),
                a_star = 0, a = 1,
                est.method = "paramfunc", inf.method = "delta", model = "rb")
```

```
##                  Estimate Std.error 95% CIL 95% CIU      z     pval
## cde_rr            2.57873   0.27097 2.04764 3.10982  9.517  < 2e-16 ***
## pnde_rr           2.57873   0.27097 2.04764 3.10982  9.517  < 2e-16 ***
## tnde_rr           2.57873   0.27097 2.04764 3.10982  9.517  < 2e-16 ***
## pnie_rr           1.17765   0.01844 1.14150 1.21380 63.849  < 2e-16 ***
## tnie_rr           1.17765   0.01844 1.14150 1.21380 63.849  < 2e-16 ***
## te_rr             3.03684   0.32126 2.40720 3.66649  9.453  < 2e-16 ***
## pm                0.22492   0.02178 0.18222 0.26761 10.325  < 2e-16 ***
## cde_err           0.43270   0.09728 0.24203 0.62337  4.448 8.76e-06 ***
## intref_err        1.14603   0.20188 0.75035 1.54170  5.677 1.41e-08 ***
## intmed_err        0.28047   0.05553 0.17163 0.38930  5.051 4.47e-07 ***
## pie_err           0.17765   0.01844 0.14150 0.21380  9.632  < 2e-16 ***
## te_err            2.03684   0.32126 1.40720 2.66649  6.340 2.39e-10 ***
## cde_err_prop      0.21244   0.03164 0.15043 0.27445  6.715 1.99e-11 ***
## intref_err_prop   0.56265   0.02666 0.51039 0.61491 21.102  < 2e-16 ***
## intmed_err_prop   0.13770   0.01121 0.11572 0.15967 12.282  < 2e-16 ***
## pie_err_prop      0.08722   0.01554 0.05676 0.11768  5.612 2.06e-08 ***
## overall_pm        0.22492   0.02178 0.18222 0.26761 10.325  < 2e-16 ***
## overall_int       0.70034   0.02922 0.64307 0.75762 23.967  < 2e-16 ***
## overall_pe        0.78756   0.03164 0.72555 0.84957 24.893  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

6.1.2.2 Direct Imputation Estimation and Bootstrap Inference

```
cmest(data = df_noint,  outcome = "binY_binM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_bin', covariates.pre = "C",
                yreg = "logistic", mreg = "logistic", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "rb")
```

```
##                  Estimate Std.error 95% CIL 95% CIU      z     pval
## cde_rr            2.54984   0.27667 2.00757 3.09210  9.216  < 2e-16 ***
## pnde_rr           2.50333   0.26692 1.98017 3.02648  9.379  < 2e-16 ***
## tnde_rr           2.48649   0.26337 1.97028 3.00269  9.441  < 2e-16 ***
## pnie_rr           1.23865   0.03115 1.17760 1.29970 39.768  < 2e-16 ***
## tnie_rr           1.23032   0.03026 1.17100 1.28964 40.653  < 2e-16 ***
## te_rr             3.07989   0.33674 2.41989 3.73990  9.146  < 2e-16 ***
```

```
## pm               0.27721   0.03057 0.21729 0.33713  9.067  < 2e-16 ***
## cde_err          0.61277   0.11971 0.37815 0.84739  5.119 3.13e-07 ***
## intref_err       0.89056   0.16503 0.56709 1.21402  5.396 6.96e-08 ***
## intmed_err       0.33791   0.07559 0.18975 0.48608  4.470 7.90e-06 ***
## pie_err          0.23865   0.03115 0.17760 0.29970  7.662 2.00e-14 ***
## te_err           2.07989   0.33674 1.41989 2.73990  6.176 6.81e-10 ***
## cde_err_prop     0.29462   0.03170 0.23248 0.35675  9.293  < 2e-16 ***
## intref_err_prop  0.42817   0.01967 0.38961 0.46673 21.764  < 2e-16 ***
## intmed_err_prop  0.16247   0.01550 0.13209 0.19285 10.482  < 2e-16 ***
## pie_err_prop     0.11474   0.02248 0.07069 0.15879  5.105 3.36e-07 ***
## overall_pm       0.27721   0.03057 0.21729 0.33713  9.067  < 2e-16 ***
## overall_int      0.59064   0.02601 0.53967 0.64161 22.712  < 2e-16 ***
## overall_pe       0.70538   0.03170 0.64325 0.76752 22.250  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_noint,  outcome = "binY_binM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_bin', covariates.pre = "C",
                yreg = "logistic", mreg = "logistic", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "wb")
```

```
##                 Estimate Std.error 95% CIL 95% CIU      z     pval
## cde_rr           2.54995   0.25643 2.04735 3.05254  9.944  < 2e-16 ***
## pnde_rr          2.43065   0.23336 1.97327 2.88803 10.416  < 2e-16 ***
## tnde_rr          2.42201   0.23133 1.96861 2.87542 10.470  < 2e-16 ***
## pnie_rr          1.16780   0.01771 1.13309 1.20251 65.943  < 2e-16 ***
## tnie_rr          1.16365   0.01706 1.13022 1.19709 68.214  < 2e-16 ***
## te_rr            2.82844   0.27636 2.28679 3.37009 10.235  < 2e-16 ***
## pm               0.21756   0.02094 0.17652 0.25859 10.391  < 2e-16 ***
## cde_err          0.43115   0.08627 0.26206 0.60024  4.998 5.90e-07 ***
## intref_err       0.99950   0.17230 0.66180 1.33720  5.801 6.80e-09 ***
## intmed_err       0.22998   0.04570 0.14041 0.31956  5.032 4.93e-07 ***
## pie_err          0.16780   0.01771 0.13309 0.20251  9.475  < 2e-16 ***
## te_err           1.82844   0.27636 1.28679 2.37009  6.616 3.87e-11 ***
## cde_err_prop     0.23580   0.03007 0.17688 0.29473  7.843 4.85e-15 ***
## intref_err_prop  0.54664   0.02779 0.49217 0.60111 19.669  < 2e-16 ***
## intmed_err_prop  0.12578   0.01056 0.10509 0.14647 11.915  < 2e-16 ***
## pie_err_prop     0.09177   0.01544 0.06151 0.12204  5.943 2.90e-09 ***
## overall_pm       0.21756   0.02094 0.17652 0.25859 10.391  < 2e-16 ***
## overall_int      0.67242   0.02923 0.61514 0.72971 23.007  < 2e-16 ***
## overall_pe       0.76420   0.03007 0.70527 0.82312 25.417  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_noint,  outcome = "binY_binM_noint", exposure = "A",
                exposure.type = "binary",
```

```
                mediator = 'M_bin', covariates.pre = "C",
                yreg = "logistic", mreg = "logistic", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "msm")
```

```
##                  Estimate Std.error 95% CIL 95% CIU      z     pval
## cde_rr            2.53317   0.27216 1.99974 3.06660  9.308  < 2e-16 ***
## pnde_rr           2.48965   0.26393 1.97236 3.00693  9.433  < 2e-16 ***
## tnde_rr           2.47315   0.26052 1.96255 2.98376  9.493  < 2e-16 ***
## pnie_rr           1.24067   0.03243 1.17711 1.30423 38.259  < 2e-16 ***
## tnie_rr           1.23245   0.03155 1.17062 1.29428 39.068  < 2e-16 ***
## te_rr             3.06836   0.33283 2.41602 3.72070  9.219  < 2e-16 ***
## pm                0.27979   0.03210 0.21688 0.34271  8.717  < 2e-16 ***
## cde_err           0.60349   0.11806 0.37211 0.83488  5.112 3.25e-07 ***
## intref_err        0.88615   0.16599 0.56081 1.21149  5.338 9.58e-08 ***
## intmed_err        0.33805   0.07594 0.18921 0.48689  4.451 8.62e-06 ***
## pie_err           0.24067   0.03243 0.17711 0.30423  7.422 1.25e-13 ***
## te_err            2.06836   0.33283 1.41602 2.72070  6.214 5.36e-10 ***
## cde_err_prop      0.29177   0.03261 0.22786 0.35568  8.948  < 2e-16 ***
## intref_err_prop   0.42843   0.02085 0.38757 0.46929 20.551  < 2e-16 ***
## intmed_err_prop   0.16344   0.01611 0.13187 0.19501 10.146  < 2e-16 ***
## pie_err_prop      0.11636   0.02289 0.07150 0.16122  5.084 3.77e-07 ***
## overall_pm        0.27979   0.03210 0.21688 0.34271  8.717  < 2e-16 ***
## overall_int       0.59187   0.02624 0.54043 0.64331 22.553  < 2e-16 ***
## overall_pe        0.70823   0.03261 0.64432 0.77214 21.720  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_noint,  outcome = "binY_binM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_bin', covariates.pre = "C",
                yreg = "logistic", mreg = "logistic", mval = list(0),
                a_star = 0, a = 1,
                est.method = "paramfunc", inf.method = "bootstrap", model = "iorw")
```

```
##       Estimate Std.error 95% CIL 95% CIU      z   pval
## ORtot  2.97302   0.30388 2.37743 3.56862  9.783 <2e-16 ***
## ORdir  2.53413   0.25610 2.03217 3.03608  9.895 <2e-16 ***
## ORind  1.17320   0.02178 1.13051 1.21588 53.868 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_noint,  outcome = "binY_binM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_bin', covariates.pre = "C",
                yreg = "logistic", mreg = "logistic", mval = list(0),
```

```
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "g-formula")
```

```
##                  Estimate Std.error 95% CIL 95% CIU      z     pval
## cde_rr            2.54984   0.28918 1.98305 3.11662  8.817  < 2e-16 ***
## pnde_rr           2.50333   0.27911 1.95628 3.05037  8.969  < 2e-16 ***
## tnde_rr           2.48649   0.27538 1.94676 3.02621  9.029  < 2e-16 ***
## pnie_rr           1.23865   0.03372 1.17256 1.30474 36.732  < 2e-16 ***
## tnie_rr           1.23032   0.03265 1.16632 1.29432 37.678  < 2e-16 ***
## te_rr             3.07989   0.35696 2.38026 3.77953  8.628  < 2e-16 ***
## pm                0.27721   0.03064 0.21716 0.33726  9.048  < 2e-16 ***
## cde_err           0.61277   0.12522 0.36734 0.85820  4.894 1.01e-06 ***
## intref_err        0.89056   0.17401 0.54950 1.23161  5.118 3.15e-07 ***
## intmed_err        0.33791   0.08197 0.17726 0.49856  4.123 3.78e-05 ***
## pie_err           0.23865   0.03372 0.17256 0.30474  7.077 1.57e-12 ***
## te_err            2.07989   0.35696 1.38026 2.77953  5.827 5.83e-09 ***
## cde_err_prop      0.29462   0.03323 0.22948 0.35975  8.865  < 2e-16 ***
## intref_err_prop   0.42817   0.01989 0.38919 0.46716 21.525  < 2e-16 ***
## intmed_err_prop   0.16247   0.01721 0.12873 0.19620  9.439  < 2e-16 ***
## pie_err_prop      0.11474   0.02071 0.07415 0.15533  5.541 3.09e-08 ***
## overall_pm        0.27721   0.03064 0.21716 0.33726  9.048  < 2e-16 ***
## overall_int       0.59064   0.02729 0.53715 0.64413 21.642  < 2e-16 ***
## overall_pe        0.70538   0.03323 0.64025 0.77052 21.226  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_noint,  outcome = "binY_binM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_bin', covariates.pre = "C",
                yreg = "logistic", mreg = "logistic", mval = list(0),
                a_star = 0, a = 1, model = "ne")
```

```
##                         Estimate Std.error 95% CIL 95% CIU     z   pval
## natural direct effect    2.52842   0.25987 2.01909 3.03775  9.73 <2e-16
## natural indirect effect  1.17505   0.01892 1.13798 1.21213 62.11 <2e-16
## total effect             2.97103   0.30788 2.36759 3.57447  9.65 <2e-16
##
## natural direct effect   ***
## natural indirect effect ***
## total effect            ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

6.2 Case 6-2: Binary Outcome and Single Binary Mediator With Exposure-mediator Interaction

6.2.1 Data simulation

6.2.1.1 Simulation Procedures

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

6.2.1.2 True Parameters

Table 14: 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
10000	-5	0.8	1.8	0.2	0.1	-0.25	0.5	0.2	0.4	1	1

6.2.1.3 True Models

True model for the mediator:

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

True model for the outcome:

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

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

6.2.2.1 Closed-form Parameter Function Estimation and Delta Method Inference

```
cmest(data = df_int,  outcome = "binY_binM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_bin', covariates.pre = "C", EMint = TRUE,
                yreg = "logistic", mreg = "logistic", mval = list(0),
                a_star = 0, a = 1,
                est.method = "paramfunc", inf.method = "delta", model = "rb")
```

```
##                  Estimate Std.error 95% CIL 95% CIU      z     pval
## cde_rr            3.03283   0.90009 1.26867 4.79698  3.369 0.000756 ***
## pnde_rr           2.91452   0.30836 2.31015 3.51889  9.452  < 2e-16 ***
## tnde_rr           2.90808   0.30966 2.30115 3.51501  9.391  < 2e-16 ***
## pnie_rr           1.18048   0.02141 1.13852 1.22244 55.138  < 2e-16 ***
## tnie_rr           1.17787   0.01974 1.13919 1.21655 59.682  < 2e-16 ***
## te_rr             3.43293   0.36575 2.71608 4.14978  9.386  < 2e-16 ***
## pm                0.21308   0.02107 0.17178 0.25438 10.112  < 2e-16 ***
## cde_err           0.53369   0.17341 0.19381 0.87356  3.078 0.002092 **
## intref_err        1.38083   0.26192 0.86749 1.89418  5.272 1.38e-07 ***
## intmed_err        0.33793   0.07102 0.19874 0.47712  4.758 1.98e-06 ***
## pie_err           0.18048   0.02141 0.13852 0.22244  8.430  < 2e-16 ***
## te_err            2.43293   0.36575 1.71608 3.14978  6.652 3.04e-11 ***
## cde_err_prop      0.21936   0.06352 0.09486 0.34386  3.453 0.000556 ***
## intref_err_prop   0.56756   0.05503 0.45970 0.67542 10.314  < 2e-16 ***
## intmed_err_prop   0.13890   0.01642 0.10671 0.17108  8.458  < 2e-16 ***
## pie_err_prop      0.07418   0.01358 0.04756 0.10080  5.462 4.83e-08 ***
## overall_pm        0.21308   0.02107 0.17178 0.25438 10.112  < 2e-16 ***
## overall_int       0.70646   0.06677 0.57558 0.83734 10.580  < 2e-16 ***
## overall_pe        0.78064   0.06352 0.65614 0.90514 12.289  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

6.2.2.2 Direct Imputation Estimation and Bootstrap Inference

```
cmest(data = df_int,  outcome = "binY_binM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_bin', covariates.pre = "C", EMint = TRUE,
                yreg = "logistic", mreg = "logistic", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "rb")
```

```
##                 Estimate Std.error 95% CIL 95% CIU      z     pval
## cde_rr           2.99440   0.97992 1.07380 4.91501  3.056  0.00225 **
## pnde_rr          2.87029   0.45669 1.97519 3.76538  6.285 3.42e-10 ***
## tnde_rr          2.83382   0.37256 2.10363 3.56402  7.606 3.07e-14 ***
## pnie_rr          1.24762   0.04697 1.15555 1.33968 26.560  < 2e-16 ***
## tnie_rr          1.23177   0.03745 1.15836 1.30517 32.889  < 2e-16 ***
## te_rr            3.53552   0.53254 2.49176 4.57928  6.639 3.32e-11 ***
## pm               0.26237   0.04194 0.18017 0.34456  6.256 4.11e-10 ***
## cde_err          0.77929   0.28795 0.21492 1.34367  2.706  0.00681 **
## intref_err       1.09099   0.20731 0.68467 1.49731  5.263 1.45e-07 ***
## intmed_err       0.41762   0.08847 0.24421 0.59103  4.720 2.39e-06 ***
## pie_err          0.24762   0.04697 0.15555 0.33968  5.271 1.38e-07 ***
## te_err           2.53552   0.53254 1.49176 3.57928  4.761 1.95e-06 ***
## cde_err_prop     0.30735   0.06694 0.17614 0.43856  4.591 4.46e-06 ***
## intref_err_prop  0.43028   0.03493 0.36181 0.49875 12.317  < 2e-16 ***
## intmed_err_prop  0.16471   0.03054 0.10485 0.22457  5.393 7.08e-08 ***
## pie_err_prop     0.09766   0.01933 0.05976 0.13555  5.051 4.47e-07 ***
## overall_pm       0.26237   0.04194 0.18017 0.34456  6.256 4.11e-10 ***
## overall_int      0.59499   0.06049 0.47643 0.71356  9.836  < 2e-16 ***
## overall_pe       0.69265   0.06694 0.56144 0.82386 10.347  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_int,  outcome = "binY_binM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_bin', covariates.pre = "C", EMint = TRUE,
                yreg = "logistic", mreg = "logistic", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "wb")
```

```
##                 Estimate Std.error 95% CIL 95% CIU      z     pval
## cde_rr           2.99437   0.98812 1.05768 4.93106  3.030 0.002449 **
## pnde_rr          2.72499   0.32163 2.09461 3.35538  8.472  < 2e-16 ***
## tnde_rr          2.71358   0.32100 2.08444 3.34272  8.454  < 2e-16 ***
## pnie_rr          1.17211   0.02062 1.13170 1.21252 56.850  < 2e-16 ***
## tnie_rr          1.16721   0.01897 1.13002 1.20439 61.523  < 2e-16 ***
## te_rr            3.18063   0.37317 2.44923 3.91202  8.523  < 2e-16 ***
## pm               0.20895   0.02277 0.16432 0.25357  9.177  < 2e-16 ***
## cde_err          0.53517   0.18043 0.18153 0.88881  2.966 0.003023 **
## intref_err       1.18982   0.26056 0.67914 1.70050  4.566 5.02e-06 ***
## intmed_err       0.28352   0.06533 0.15548 0.41155  4.340 1.44e-05 ***
## pie_err          0.17211   0.02062 0.13170 0.21252  8.348  < 2e-16 ***
## te_err           2.18063   0.37317 1.44923 2.91202  5.844 5.27e-09 ***
```

```
## cde_err_prop     0.24542   0.07145 0.10537 0.38547  3.435 0.000596 ***
## intref_err_prop  0.54563   0.06175 0.42460 0.66667  8.835  < 2e-16 ***
## intmed_err_prop  0.13002   0.01689 0.09690 0.16313  7.696 1.54e-14 ***
## pie_err_prop     0.07893   0.01650 0.04659 0.11127  4.783 1.75e-06 ***
## overall_pm       0.20895   0.02277 0.16432 0.25357  9.177  < 2e-16 ***
## overall_int      0.67565   0.07509 0.52847 0.82283  8.997  < 2e-16 ***
## overall_pe       0.75458   0.07145 0.61453 0.89463 10.560  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_int,  outcome = "binY_binM_int", exposure = "A",
                exposure.type = "binary",
                mediator = 'M_bin', covariates.pre = "C", EMint = TRUE,
                yreg = "logistic", mreg = "logistic", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "msm")
```

```
##                  Estimate Std.error 95% CIL 95% CIU      z     pval
## cde_rr            3.05941   0.97338 1.15163 4.96720  3.143  0.00168 **
## pnde_rr           2.91236   0.44263 2.04482 3.77990  6.580 4.95e-11 ***
## tnde_rr           2.87264   0.35629 2.17433 3.57095  8.063 8.32e-16 ***
## pnie_rr           1.24938   0.04491 1.16136 1.33740 27.821  < 2e-16 ***
## tnie_rr           1.23234   0.03778 1.15828 1.30640 32.615  < 2e-16 ***
## te_rr             3.58902   0.51164 2.58623 4.59181  7.015 2.45e-12 ***
## pm                0.26136   0.04080 0.18140 0.34132  6.406 1.56e-10 ***
## cde_err           0.79836   0.28048 0.24864 1.34808  2.846  0.00443 **
## intref_err        1.11400   0.20056 0.72091 1.50709  5.554 2.86e-08 ***
## intmed_err        0.42728   0.08593 0.25887 0.59569  4.973 6.71e-07 ***
## pie_err           0.24938   0.04491 0.16136 0.33740  5.553 2.88e-08 ***
## te_err            2.58902   0.51164 1.58623 3.59181  5.060 4.26e-07 ***
## cde_err_prop      0.30836   0.06310 0.18470 0.43203  4.887 1.04e-06 ***
## intref_err_prop   0.43028   0.03272 0.36615 0.49440 13.151  < 2e-16 ***
## intmed_err_prop   0.16504   0.03030 0.10565 0.22442  5.447 5.24e-08 ***
## pie_err_prop      0.09632   0.01845 0.06016 0.13248  5.221 1.82e-07 ***
## overall_pm        0.26136   0.04080 0.18140 0.34132  6.406 1.56e-10 ***
## overall_int       0.59531   0.05806 0.48152 0.70911 10.253  < 2e-16 ***
## overall_pe        0.69164   0.06310 0.56797 0.81530 10.961  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_int,  outcome = "binY_binM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_bin', covariates.pre = "C", EMint = TRUE,
                yreg = "logistic", mreg = "logistic", mval = list(0),
                a_star = 0, a = 1,
                est.method = "paramfunc", inf.method = "bootstrap", model = "iorw")
```

```
##       Estimate Std.error 95% CIL 95% CIU      z   pval
## ORtot  3.36578   0.36925 2.64206 4.08949  9.115 <2e-16 ***
## ORdir  2.85411   0.31366 2.23935 3.46886  9.099 <2e-16 ***
## ORind  1.17927   0.02237 1.13543 1.22312 52.715 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_int,  outcome = "binY_binM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_bin', covariates.pre = "C", EMint = TRUE,
                yreg = "logistic", mreg = "logistic", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "g-formula")
```

```
##                 Estimate Std.error 95% CIL 95% CIU      z     pval
## cde_rr           2.99440   1.04000 0.95605 5.03276  2.879  0.00399 **
## pnde_rr          2.87029   0.48343 1.92278 3.81779  5.937 2.99e-09 ***
## tnde_rr          2.83382   0.39452 2.06058 3.60707  7.183 7.31e-13 ***
## pnie_rr          1.24762   0.04152 1.16623 1.32900 30.047  < 2e-16 ***
## tnie_rr          1.23177   0.03937 1.15461 1.30893 31.289  < 2e-16 ***
## te_rr            3.53552   0.55762 2.44261 4.62843  6.340 2.39e-10 ***
## pm               0.26237   0.04234 0.17939 0.34534  6.197 5.97e-10 ***
## cde_err          0.77929   0.30645 0.17866 1.37992  2.543  0.01101 *
## intref_err       1.09099   0.22041 0.65899 1.52299  4.950 7.55e-07 ***
## intmed_err       0.41762   0.09667 0.22814 0.60710  4.320 1.58e-05 ***
## pie_err          0.24762   0.04152 0.16623 0.32900  5.963 2.55e-09 ***
## te_err           2.53552   0.55762 1.44261 3.62843  4.547 5.50e-06 ***
## cde_err_prop     0.30735   0.06805 0.17398 0.44072  4.517 6.35e-06 ***
## intref_err_prop  0.43028   0.03358 0.36448 0.49609 12.815  < 2e-16 ***
## intmed_err_prop  0.16471   0.03205 0.10189 0.22752  5.139 2.81e-07 ***
## pie_err_prop     0.09766   0.01791 0.06255 0.13277  5.451 5.12e-08 ***
## overall_pm       0.26237   0.04234 0.17939 0.34534  6.197 5.97e-10 ***
## overall_int      0.59499   0.06119 0.47507 0.71491  9.724  < 2e-16 ***
## overall_pe       0.69265   0.06805 0.55928 0.82602 10.179  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_int,  outcome = "binY_binM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_bin', covariates.pre = "C", EMint = TRUE,
                yreg = "logistic", mreg = "logistic", mval = list(0),
                a_star = 0, a = 1, model = "ne")
```

```
##                       Estimate Std.error 95% CIL 95% CIU      z   pval
## pure direct effect     2.84405   0.29399 2.26783 3.42026  9.674 <2e-16 ***
## total direct effect    2.85019   0.29609 2.26988 3.43051  9.626 <2e-16 ***
## pure indirect effect   1.17897   0.02171 1.13642 1.22152 54.309 <2e-16 ***
## total indirect effect  1.18152   0.02098 1.14040 1.22264 56.316 <2e-16 ***
```

```
## total effect           3.36030   0.34951 2.67526 4.04533  9.614 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

7 Case 7: Continuous Outcome and Single Categorical Mediator

7.1 Case 7-1: Continuous Outcome and Single Categorical Mediator Without Interaction

7.1.1 Data simulation

7.1.1.1 Simulation Procedures

1. Simulate the exposure variable A from $Bernoulli(P(A=1))$.
2. Simulate the covariate C from $N(\mu_C, \sigma_C^2)$.
3. Simulate the mediator M from $Multinom(\frac{1}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}, \frac{expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}, \frac{expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)})$.
4. Simulate the outcome Y from $Bernoulli(expit(\theta_0 + \theta_1 A + \theta_{21} I\{M == 1\} + \theta_{22} I\{M == 2\} + \theta_4 C))$.

7.1.1.2 True Parameters

Table 15: True Model Parameters for Data Simulation

Sample Size	θ_0	θ_1	θ_{21}	θ_{22}	θ_4	β_{01}	β_{11}
10000	-5	0.8	1.8	1.2	0.1	-0.25	0.5
β_{21}	β_{02}	β_{12}	β_{22}	P(A=1)	μ_C	σ_C	
0.2	-0.3	0.4	0.3	0.4	1	1	

7.1.1.3 True Models

True model for the mediator:

$$ln\frac{P(M==1)}{P(M==0)} = \beta_{01} + \beta_{11}a + \beta_{21}c$$

$$ln\frac{P(M==2)}{P(M==0)} = \beta_{02} + \beta_{12}a + \beta_{22}c$$

True model for the outcome:

$$logitE[Y|a, m^*, c] = \theta_0 + \theta_1 a + \theta_{21} I\{m^* == 1\} + \theta_{22} I\{m^* == 2\} + \theta_4 c$$

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

7.1.2.1 Closed-form Parameter Function Estimation and Delta Method Inference

```
cmest(data = df_noint,  outcome = "binY_catM_noint", exposure = 'A',
                  exposure.type = "binary",
                  mediator = 'M_cat', covariates.pre = "C",
                  yreg = "logistic", mreg = "multinomial", mval = list(0),
                  a_star = 0, a = 1,
                  est.method = "paramfunc", inf.method = "delta", model = "rb")
```

```
##                  Estimate Std.error 95% CIL 95% CIU      z     pval
## cde_rr            2.77002   0.30557 2.17112 3.36892  9.065  < 2e-16 ***
## pnde_rr           2.77002   0.30557 2.17112 3.36892  9.065  < 2e-16 ***
## tnde_rr           2.77002   0.30557 2.17112 3.36892  9.065  < 2e-16 ***
## pnie_rr           1.13325   0.01679 1.10035 1.16615 67.507  < 2e-16 ***
## tnie_rr           1.13325   0.01679 1.10035 1.16615 67.507  < 2e-16 ***
## te_rr             3.13913   0.34750 2.45805 3.82021  9.034  < 2e-16 ***
## pm                0.17255   0.02039 0.13258 0.21252  8.461  < 2e-16 ***
## cde_err           0.62088   0.14824 0.33034 0.91142  4.188 2.83e-05 ***
## intref_err        1.14914   0.21414 0.72943 1.56885  5.366 8.22e-08 ***
## intmed_err        0.23586   0.04941 0.13901 0.33270  4.773 1.84e-06 ***
## pie_err           0.13325   0.01679 0.10035 0.16615  7.938 2.28e-15 ***
## te_err            2.13913   0.34750 1.45805 2.82021  6.156 7.75e-10 ***
## cde_err_prop      0.29025   0.04868 0.19484 0.38566  5.962 2.57e-09 ***
## intref_err_prop   0.53720   0.03932 0.46014 0.61426 13.663  < 2e-16 ***
## intmed_err_prop   0.11026   0.01151 0.08771 0.13281  9.583  < 2e-16 ***
## pie_err_prop      0.06229   0.01222 0.03835 0.08623  5.099 3.47e-07 ***
## overall_pm        0.17255   0.02039 0.13258 0.21252  8.461  < 2e-16 ***
## overall_int       0.64746   0.04453 0.56018 0.73474 14.539  < 2e-16 ***
## overall_pe        0.70975   0.04868 0.61434 0.80516 14.580  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

7.1.2.2 Direct Imputation Estimation and Bootstrap Inference

```
cmest(data = df_noint,  outcome = "binY_catM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cat', covariates.pre = "C",
                yreg = "logistic", mreg = "multinomial", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "rb")
```

```
##                  Estimate Std.error 95% CIL 95% CIU      z     pval
## cde_rr            2.73098   0.30580 2.13162 3.33035  8.931  < 2e-16 ***
## pnde_rr           2.68027   0.29541 2.10126 3.25927  9.073  < 2e-16 ***
## tnde_rr           2.66697   0.29279 2.09312 3.24082  9.109  < 2e-16 ***
## pnie_rr           1.16351   0.02440 1.11568 1.21134 47.679  < 2e-16 ***
## tnie_rr           1.15774   0.02370 1.11129 1.20420 48.844  < 2e-16 ***
## te_rr             3.10306   0.34042 2.43585 3.77027  9.115  < 2e-16 ***
## pm                0.20104   0.02724 0.14764 0.25443  7.380 1.71e-13 ***
## cde_err           0.75630   0.16772 0.42757 1.08502  4.509 6.58e-06 ***
## intref_err        0.92397   0.17225 0.58637 1.26157  5.364 8.31e-08 ***
## intmed_err        0.25928   0.05625 0.14903 0.36953  4.609 4.09e-06 ***
## pie_err           0.16351   0.02440 0.11568 0.21134  6.700 2.19e-11 ***
## te_err            2.10306   0.34042 1.43585 2.77027  6.178 6.75e-10 ***
## cde_err_prop      0.35962   0.04920 0.26318 0.45605  7.309 2.90e-13 ***
## intref_err_prop   0.43935   0.03076 0.37906 0.49963 14.284  < 2e-16 ***
## intmed_err_prop   0.12329   0.01466 0.09456 0.15202  8.410  < 2e-16 ***
## pie_err_prop      0.07775   0.01645 0.04552 0.10998  4.728 2.30e-06 ***
## overall_pm        0.20104   0.02724 0.14764 0.25443  7.380 1.71e-13 ***
## overall_int       0.56263   0.04069 0.48288 0.64239 13.827  < 2e-16 ***
## overall_pe        0.64038   0.04920 0.54395 0.73682 13.015  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_noint,  outcome = "binY_catM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cat', covariates.pre = "C",
                yreg = "logistic", mreg = "multinomial", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "wb")
```

```
##                 Estimate Std.error 95% CIL 95% CIU      z     pval
## cde_rr           2.73102   0.28487 2.17268 3.28936  9.587  < 2e-16 ***
## pnde_rr          2.62692   0.26567 2.10620 3.14763  9.888  < 2e-16 ***
## tnde_rr          2.62066   0.26435 2.10254 3.13877  9.914  < 2e-16 ***
## pnie_rr          1.12285   0.01712 1.08929 1.15641 65.581  < 2e-16 ***
## tnie_rr          1.12017   0.01640 1.08802 1.15232 68.289  < 2e-16 ***
## te_rr            2.94260   0.29614 2.36217 3.52303  9.936  < 2e-16 ***
## pm               0.16251   0.02149 0.12038 0.20463  7.561 4.34e-14 ***
## cde_err          0.61764   0.13936 0.34449 0.89078  4.432 9.44e-06 ***
## intref_err       1.00928   0.17803 0.66035 1.35821  5.669 1.47e-08 ***
## intmed_err       0.19284   0.03881 0.11677 0.26890  4.969 6.85e-07 ***
## pie_err          0.12285   0.01712 0.08929 0.15641  7.175 7.74e-13 ***
## te_err           1.94260   0.29614 1.36217 2.52303  6.560 5.66e-11 ***
## cde_err_prop     0.31794   0.05012 0.21971 0.41618  6.344 2.34e-10 ***
## intref_err_prop  0.51955   0.03844 0.44421 0.59489 13.516  < 2e-16 ***
## intmed_err_prop  0.09927   0.01119 0.07733 0.12120  8.871  < 2e-16 ***
## pie_err_prop     0.06324   0.01323 0.03731 0.08917  4.780 1.78e-06 ***
## overall_pm       0.16251   0.02149 0.12038 0.20463  7.561 4.34e-14 ***
## overall_int      0.61882   0.04475 0.53111 0.70653 13.828  < 2e-16 ***
## overall_pe       0.68206   0.05012 0.58382 0.78029 13.608  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_noint,  outcome = "binY_catM_noint", exposure = "A",
                exposure.type = "binary",
                mediator = 'M_cat', covariates.pre = "C",
                yreg = "logistic", mreg = "multinomial", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "msm")
```

```
##                 Estimate Std.error 95% CIL 95% CIU      z     pval
## cde_rr           2.69856   0.29062 2.12897 3.26816  9.286  < 2e-16 ***
## pnde_rr          2.65281   0.28230 2.09951 3.20612  9.397  < 2e-16 ***
## tnde_rr          2.63987   0.27963 2.09181 3.18793  9.441  < 2e-16 ***
## pnie_rr          1.15975   0.02815 1.10459 1.21492 41.202  < 2e-16 ***
## tnie_rr          1.15410   0.02726 1.10066 1.20753 42.333  < 2e-16 ***
## te_rr            3.06160   0.33283 2.40927 3.71394  9.199  < 2e-16 ***
## pm               0.19829   0.03012 0.13926 0.25731  6.584 4.80e-11 ***
## cde_err          0.75548   0.16172 0.43850 1.07245  4.671 3.03e-06 ***
## intref_err       0.89734   0.16800 0.56806 1.22661  5.341 9.43e-08 ***
## intmed_err       0.24903   0.06222 0.12708 0.37099  4.002 6.32e-05 ***
## pie_err          0.15975   0.02815 0.10459 0.21492  5.676 1.42e-08 ***
```

```
## te_err           2.06160   0.33283 1.40927 2.71394  6.194 6.09e-10 ***
## cde_err_prop     0.36645   0.05289 0.26278 0.47012  6.928 4.53e-12 ***
## intref_err_prop  0.43526   0.03042 0.37564 0.49488 14.308  < 2e-16 ***
## intmed_err_prop  0.12080   0.01680 0.08786 0.15373  7.188 7.02e-13 ***
## pie_err_prop     0.07749   0.01743 0.04332 0.11166  4.445 8.88e-06 ***
## overall_pm       0.19829   0.03012 0.13926 0.25731  6.584 4.80e-11 ***
## overall_int      0.55606   0.04283 0.47211 0.64001 12.982  < 2e-16 ***
## overall_pe       0.63355   0.05289 0.52988 0.73722 11.978  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_noint,  outcome = "binY_catM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cat', covariates.pre = "C",
                yreg = "logistic", mreg = "multinomial", mval = list(0),
                a_star = 0, a = 1,
                est.method = "paramfunc", inf.method = "bootstrap", model = "iorw")
```

```
##       Estimate Std.error 95% CIL 95% CIU      z     pval
## ORtot   3.0844    0.4037  2.2932  3.8756  7.641 2.36e-14 ***
## ORdir   2.7290    0.3599  2.0235  3.4345  7.582 3.72e-14 ***
## ORind   1.1302    0.0187  1.0936  1.1669 60.437  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_noint,  outcome = "binY_catM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cat', covariates.pre = "C",
                yreg = "logistic", mreg = "multinomial", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "g-formula")
```

```
##                 Estimate Std.error 95% CIL 95% CIU      z     pval
## cde_rr           2.73098   0.32793 2.08826 3.37371  8.328  < 2e-16 ***
## pnde_rr          2.68027   0.31626 2.06040 3.30013  8.475  < 2e-16 ***
## tnde_rr          2.66697   0.31326 2.05299 3.28095  8.514  < 2e-16 ***
## pnie_rr          1.16351   0.02957 1.10556 1.22146 39.352  < 2e-16 ***
## tnie_rr          1.15774   0.02867 1.10155 1.21394 40.378  < 2e-16 ***
## te_rr            3.10306   0.37158 2.37478 3.83134  8.351  < 2e-16 ***
## pm               0.20104   0.03156 0.13919 0.26288  6.371 1.96e-10 ***
## cde_err          0.75630   0.17203 0.41913 1.09346  4.396 1.11e-05 ***
## intref_err       0.92397   0.19929 0.53338 1.31456  4.636 3.59e-06 ***
## intmed_err       0.25928   0.06802 0.12597 0.39259  3.812 0.000139 ***
## pie_err          0.16351   0.02957 0.10556 0.22146  5.530 3.28e-08 ***
## te_err           2.10306   0.37158 1.37478 2.83134  5.660 1.56e-08 ***
## cde_err_prop     0.35962   0.05675 0.24839 0.47084  6.337 2.44e-10 ***
## intref_err_prop  0.43935   0.03404 0.37264 0.50605 12.909  < 2e-16 ***
```

```
## intmed_err_prop  0.12329   0.01764 0.08871 0.15787  6.988 2.97e-12 ***
## pie_err_prop     0.07775   0.01782 0.04282 0.11268  4.363 1.30e-05 ***
## overall_pm       0.20104   0.03156 0.13919 0.26288  6.371 1.96e-10 ***
## overall_int      0.56263   0.04708 0.47035 0.65491 11.950  < 2e-16 ***
## overall_pe       0.64038   0.05675 0.52916 0.75161 11.285  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_noint,  outcome = "binY_catM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cat', covariates.pre = "C",
                yreg = "logistic", mreg = "multinomial", mval = list(0),
                a_star = 0, a = 1, model = "ne")
```

```
##                         Estimate Std.error 95% CIL 95% CIU      z   pval
## natural direct effect    2.73159   0.29719 2.14911 3.31406  9.191 <2e-16
## natural indirect effect  1.12831   0.01714 1.09472 1.16190 65.833 <2e-16
## total effect             3.08208   0.33625 2.42305 3.74111  9.166 <2e-16
##
## natural direct effect   ***
## natural indirect effect ***
## total effect            ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

7.2 Case 7-2: Binary Outcome and Single Categorical Mediator With Exposure-mediator Interaction

7.2.1 Data simulation

7.2.1.1 Simulation Procedures

1. Simulate the exposure variable A from $Bernoulli(P(A=1))$.
2. Simulate the covariate C from $N(\mu_C, \sigma_C^2)$.
3. Simulate the mediator M from $Multinom(\frac{1}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}$, $\frac{expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}$, $\frac{expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)})$.
4. Simulate the outcome Y from $Bernoulli(expit(\theta_0 + \theta_1 A + \theta_{21} I\{M == 1\} + \theta_{22} I\{M == 2\} + \theta_{31} A * I\{M == 1\} + \theta_{32} A * I\{M == 2\} + \theta_4 C))$.

7.2.1.2 True Parameters

Table 16: True Model Parameters for Data Simulation

Sample Size	θ_0	θ_1	θ_{21}	θ_{22}	θ_{31}	θ_{32}	θ_4	β_{01}
10000	-5	0.8	1.8	1.2	0.2	0.40.1	-0.25	
β_{11}	β_{21}	β_{02}	β_{12}	β_{22}	P(A=1)	μ_C	σ_C	
0.5	0.2	-0.3	0.4	0.3	0.4	1	1	

7.2.1.3 True Models

True model for the mediator:

$$ln\frac{P(M==1)}{P(M==0)} = \beta_{01} + \beta_{11}a + \beta_{21}c$$

$$ln\frac{P(M==2)}{P(M==0)} = \beta_{02} + \beta_{12}a + \beta_{22}c$$

True model for the outcome:

$$logitE[Y|a, m^*, c] = \theta_0 + \theta_1 a + \theta_{21} I\{m^* == 1\} + \theta_{22} I\{m^* == 2\} + \theta_{31} a * I\{m^* == 1\} + \theta_{32} a * I\{m^* == 2\} + \theta_4 c$$

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

7.2.2.1 Closed-form Parameter Function Estimation and Delta Method Inference

```
cmest(data = df_int,  outcome = "binY_catM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cat', covariates.pre = "C", EMint = TRUE,
                yreg = "logistic", mreg = "multinomial", mval = list(0),
                a_star = 0, a = 1,
                est.method = "paramfunc", inf.method = "delta", model = "rb")
```

```
##                  Estimate Std.error  95% CIL  95% CIU      z     pval
## cde_rr            1.59647   0.63669  0.34858  2.84436  2.507   0.0122 *
## pnde_rr           2.43397   0.24838  1.94716  2.92079  9.799  < 2e-16 ***
## tnde_rr           2.45963   0.25269  1.96437  2.95489  9.734  < 2e-16 ***
## pnie_rr           1.14349   0.01800  1.10820  1.17878 63.515  < 2e-16 ***
## tnie_rr           1.15554   0.01671  1.12280  1.18829 69.167  < 2e-16 ***
## te_rr             2.81256   0.28851  2.24709  3.37804  9.749  < 2e-16 ***
## pm                0.20887   0.02162  0.16649  0.25125  9.660  < 2e-16 ***
## cde_err           0.15308   0.14316 -0.12751  0.43367  1.069   0.2850
## intref_err        1.28089   0.24192  0.80674  1.75505  5.295 1.22e-07 ***
## intmed_err        0.23510   0.05028  0.13655  0.33364  4.676 2.97e-06 ***
## pie_err           0.14349   0.01800  0.10820  0.17878  7.970 1.76e-15 ***
## te_err            1.81256   0.28851  1.24709  2.37804  6.282 3.47e-10 ***
## cde_err_prop      0.08446   0.07647 -0.06543  0.23434  1.104   0.2694
## intref_err_prop   0.70668   0.06874  0.57195  0.84141 10.280  < 2e-16 ***
## intmed_err_prop   0.12970   0.01589  0.09856  0.16085  8.161 3.72e-16 ***
## pie_err_prop      0.07916   0.01551  0.04876  0.10957  5.103 3.41e-07 ***
## overall_pm        0.20887   0.02162  0.16649  0.25125  9.660  < 2e-16 ***
## overall_int       0.83638   0.07949  0.68059  0.99217 10.522  < 2e-16 ***
## overall_pe        0.91554   0.07647  0.76566  1.06543 11.972  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

7.2.2.2 Direct Imputation Estimation and Bootstrap Inference

```
cmest(data = df_int,  outcome = "binY_catM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cat', covariates.pre = "C", EMint = TRUE,
                yreg = "logistic", mreg = "multinomial", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "rb")
```

```
##                 Estimate Std.error  95% CIL  95% CIU      z     pval
## cde_rr           1.58973   0.75839  0.10331  3.07615  2.096 0.036091 *
## pnde_rr          2.11511   0.35298  1.42328  2.80694  5.992 2.14e-09 ***
## tnde_rr          2.20195   0.28891  1.63570  2.76820  7.622 2.73e-14 ***
## pnie_rr          1.20699   0.03904  1.13047  1.28351 30.916  < 2e-16 ***
## tnie_rr          1.25655   0.05701  1.14481  1.36828 22.042  < 2e-16 ***
## te_rr            2.65773   0.38650  1.90020  3.41526  6.876 6.51e-12 ***
## pm               0.32733   0.07896  0.17257  0.48209  4.145 3.42e-05 ***
## cde_err          0.20184   0.21301 -0.21566  0.61933  0.948 0.343382
## intref_err       0.91327   0.18101  0.55850  1.26805  5.045 4.61e-07 ***
## intmed_err       0.33563   0.07266  0.19322  0.47804  4.619 3.90e-06 ***
## pie_err          0.20699   0.03904  0.13047  0.28351  5.302 1.17e-07 ***
## te_err           1.65773   0.38650  0.90020  2.41526  4.289 1.81e-05 ***
## cde_err_prop     0.12176   0.11385 -0.10139  0.34490  1.069 0.284897
## intref_err_prop  0.55092   0.05363  0.44580  0.65604 10.272  < 2e-16 ***
## intmed_err_prop  0.20246   0.06009  0.08469  0.32024  3.369 0.000756 ***
## pie_err_prop     0.12486   0.02878  0.06845  0.18127  4.338 1.45e-05 ***
## overall_pm       0.32733   0.07896  0.17257  0.48209  4.145 3.42e-05 ***
## overall_int      0.75338   0.10245  0.55258  0.95418  7.354 2.08e-13 ***
## overall_pe       0.87824   0.11385  0.65510  1.10139  7.714 1.34e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_int,  outcome = "binY_catM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cat', covariates.pre = "C", EMint = TRUE,
                yreg = "logistic", mreg = "multinomial", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "wb")
```

```
##                 Estimate Std.error  95% CIL  95% CIU      z     pval
## cde_rr           1.58974   0.67807  0.26075  2.91872  2.345   0.0191 *
## pnde_rr          2.30668   0.20013  1.91442  2.69893 11.526  < 2e-16 ***
## tnde_rr          2.32857   0.20012  1.93633  2.72080 11.636  < 2e-16 ***
## pnie_rr          1.13581   0.01839  1.09976  1.17185 61.757  < 2e-16 ***
## tnie_rr          1.14659   0.01665  1.11396  1.17922 68.871  < 2e-16 ***
## te_rr            2.64480   0.22360  2.20656  3.08304 11.828  < 2e-16 ***
## pm               0.20557   0.02390  0.15873  0.25241  8.602  < 2e-16 ***
## cde_err          0.15536   0.13459 -0.10843  0.41916  1.154   0.2484
## intref_err       1.15131   0.19450  0.77010  1.53252  5.919 3.34e-09 ***
## intmed_err       0.20232   0.03636  0.13105  0.27359  5.564 2.71e-08 ***
## pie_err          0.13581   0.01839  0.09976  0.17185  7.384 1.66e-13 ***
## te_err           1.64480   0.22360  1.20656  2.08304  7.356 2.04e-13 ***
## cde_err_prop     0.09446   0.07907 -0.06052  0.24943  1.195   0.2323
## intref_err_prop  0.69997   0.07188  0.55909  0.84085  9.738  < 2e-16 ***
## intmed_err_prop  0.12301   0.01522  0.09318  0.15283  8.084 6.99e-16 ***
## pie_err_prop     0.08257   0.01711  0.04904  0.11610  4.827 1.41e-06 ***
## overall_pm       0.20557   0.02390  0.15873  0.25241  8.602  < 2e-16 ***
## overall_int      0.82298   0.08176  0.66273  0.98322 10.066  < 2e-16 ***
## overall_pe       0.90554   0.07907  0.75057  1.06052 11.452  < 2e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_int,  outcome = "binY_catM_int", exposure = "A",
                exposure.type = "binary",
                mediator = 'M_cat', covariates.pre = "C", EMint = TRUE,
                yreg = "logistic", mreg = "multinomial", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "msm")
```

```
##                 Estimate Std.error  95% CIL  95% CIU      z     pval
## cde_rr           1.64799   0.84721 -0.01250  3.30849  1.945  0.05178 .
## pnde_rr          2.14269   0.37933  1.39921  2.88616  5.649 1.66e-08 ***
## tnde_rr          2.22388   0.31216  1.61206  2.83571  7.124 1.12e-12 ***
## pnie_rr          1.21005   0.04658  1.11875  1.30135 25.977  < 2e-16 ***
## tnie_rr          1.25591   0.05012  1.15767  1.35414 25.057  < 2e-16 ***
## te_rr            2.69102   0.43195  1.84442  3.53762  6.230 4.85e-10 ***
## pm               0.32426   0.07410  0.17903  0.46949  4.376 1.22e-05 ***
## cde_err          0.21733   0.22124 -0.21629  0.65095  0.982  0.32597
## intref_err       0.92536   0.20155  0.53033  1.32039  4.591 4.46e-06 ***
## intmed_err       0.33828   0.07492  0.19144  0.48511  4.515 6.39e-06 ***
## pie_err          0.21005   0.04658  0.11875  0.30135  4.509 6.58e-06 ***
## te_err           1.69102   0.43195  0.84442  2.53762  3.915 9.11e-05 ***
## cde_err_prop     0.12852   0.10958 -0.08626  0.34330  1.173  0.24090
## intref_err_prop  0.54722   0.05158  0.44613  0.64831 10.609  < 2e-16 ***
## intmed_err_prop  0.20004   0.05517  0.09190  0.30818  3.626  0.00029 ***
## pie_err_prop     0.12422   0.03081  0.06383  0.18461  4.031 5.59e-05 ***
## overall_pm       0.32426   0.07410  0.17903  0.46949  4.376 1.22e-05 ***
## overall_int      0.74726   0.09900  0.55322  0.94131  7.548 4.81e-14 ***
## overall_pe       0.87148   0.10958  0.65670  1.08626  7.953 2.02e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_int,  outcome = "binY_catM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cat', covariates.pre = "C", EMint = TRUE,
                yreg = "logistic", mreg = "multinomial", mval = list(0),
                a_star = 0, a = 1,
                est.method = "paramfunc", inf.method = "bootstrap", model = "iorw")
```

```
##       Estimate Std.error 95% CIL 95% CIU      z   pval
## ORtot  2.77555   0.28036 2.22606 3.32505  9.900 <2e-16 ***
## ORdir  2.42900   0.24526 1.94830 2.90969  9.904 <2e-16 ***
## ORind  1.14268   0.01718 1.10900 1.17635 66.506 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_int,  outcome = "binY_catM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cat', covariates.pre = "C", EMint = TRUE,
                yreg = "logistic", mreg = "multinomial", mval = list(0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "g-formula")
```

```
##                  Estimate Std.error  95% CIL  95% CIU      z     pval
## cde_rr            1.58973   0.94568 -0.26378  3.44324  1.681 0.092787 .
## pnde_rr           2.11511   0.36853  1.39280  2.83741  5.739 9.78e-09 ***
## tnde_rr           2.20195   0.29699  1.61986  2.78403  7.414 1.32e-13 ***
## pnie_rr           1.20699   0.03816  1.13219  1.28179 31.627  < 2e-16 ***
## tnie_rr           1.25655   0.05326  1.15217  1.36092 23.595  < 2e-16 ***
## te_rr             2.65773   0.40448  1.86496  3.45050  6.571 5.26e-11 ***
## pm                0.32733   0.07317  0.18392  0.47074  4.473 7.78e-06 ***
## cde_err           0.20184   0.21699 -0.22346  0.62713  0.930 0.352308
## intref_err        0.91327   0.19903  0.52319  1.30335  4.589 4.51e-06 ***
## intmed_err        0.33563   0.07333  0.19192  0.47935  4.577 4.77e-06 ***
## pie_err           0.20699   0.03816  0.13219  0.28179  5.424 5.97e-08 ***
## te_err            1.65773   0.40448  0.86496  2.45050  4.098 4.19e-05 ***
## cde_err_prop      0.12176   0.10316 -0.08044  0.32395  1.180 0.237929
## intref_err_prop   0.55092   0.05121  0.45054  0.65129 10.758  < 2e-16 ***
## intmed_err_prop   0.20246   0.05476  0.09513  0.30980  3.697 0.000219 ***
## pie_err_prop      0.12486   0.03113  0.06385  0.18588  4.011 6.09e-05 ***
## overall_pm        0.32733   0.07317  0.18392  0.47074  4.473 7.78e-06 ***
## overall_int       0.75338   0.09508  0.56703  0.93973  7.924 2.55e-15 ***
## overall_pe        0.87824   0.10316  0.67605  1.08044  8.513  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_int,  outcome = "binY_catM_int", exposure = 'A',
                exposure.type = "binary",
                mediator = 'M_cat', covariates.pre = "C", EMint = TRUE,
                yreg = "logistic", mreg = "multinomial", mval = list(0),
                a_star = 0, a = 1, model = "ne")
```

```
##                      Estimate Std.error 95% CIL 95% CIU      z   pval
## pure direct effect    2.39344   0.24040 1.92227 2.86461  9.956 <2e-16 ***
## total direct effect   2.42980   0.24548 1.94866 2.91093  9.898 <2e-16 ***
## pure indirect effect  1.14123   0.01834 1.10529 1.17717 62.233 <2e-16 ***
## total indirect effect 1.15857   0.01763 1.12402 1.19311 65.729 <2e-16 ***
## total effect          2.77296   0.28008 2.22402 3.32190  9.901 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

8 Case 8: Binary Outcome and Multiple Mediators

8.1 Case 8-1: Binary Outcome and Multiple Mediators Without Interaction

8.1.1 Data simulation

8.1.1.1 Simulation Procedures

1. Simulate the exposure variable A from $Bernoulli(P(A=1))$.
2. Simulate the covariate C from $N(\mu_C, \sigma_C^2)$.
3. Simulate the first mediator M1 from $Multinom(\frac{1}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}, \frac{expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}, \frac{expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)})$, the second mediator M2 from $Bernoulli(expit(\beta_{03}+\beta_{13}*A+\beta_{23}*C))$.
4. Simulate the outcome Y from $Bernoulli(expit(\theta_0+\theta_1 A+\theta_{21}I\{M1==1\}+\theta_{22}I\{M1==2\}+\theta_{23}M2+\theta_4 C))$.

8.1.1.2 True Parameters

Table 17: True Model Parameters for Data Simulation

Sample Size	θ_0	θ_1	θ_{21}	θ_{22}	θ_{23}	θ_4	β_{01}	β_{11}	β_{21}
10000	-5	0.8	1.8	1.2	1.5	0.1	-0.25	0.5	0.2
β_{02}	β_{12}	β_{22}	β_{03}	β_{13}	β_{23}	P(A=1)	μ_C	σ_C	
-0.3	0.4	0.3	-0.25	0.5	0.2	0.4	1	1	

8.1.1.3 True Models

True model for the first mediator:

$$ln\frac{P(M1==1)}{P(M1==0)} = \beta_{01}+\beta_{11}a+\beta_{21}c$$

$$ln\frac{P(M1==2)}{P(M1==0)} = \beta_{02}+\beta_{12}a+\beta_{22}c$$

True model for the second mediator:

$$logitE[M2|a,c] = \beta_{03}+\beta_{13}a+\beta_{23}c$$

True model for the outcome:

$$logitE[Y|a,m^*,c] = \theta_0+\theta_1 a+\theta_{21}I\{m1^*==1\}+\theta_{22}I\{m1^*==2\}+\theta_{23}m2^*+\theta_4 c$$

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

```
cmest(data = df_multipleM_noint,  outcome = "binY_catMbinM_noint", exposure = 'A',
                  exposure.type = "binary",
                  mediator = c('M_cat',"M_bin"), covariates.pre = "C",
                  yreg = "logistic", mreg = c("multinomial", "logistic"), mval = list(0,0),
                  a_star = 0, a = 1,
                  est.method = "imputation", inf.method = "bootstrap", model = "rb")
```

```
##                 Estimate Std.error 95% CIL 95% CIU      z     pval
## cde_rr           2.21119   0.15860 1.90034 2.52204 13.942  < 2e-16 ***
## pnde_rr          2.12081   0.14463 1.83734 2.40428 14.664  < 2e-16 ***
## tnde_rr          2.08058   0.13829 1.80954 2.35163 15.045  < 2e-16 ***
## pnie_rr          1.42011   0.03612 1.34933 1.49090 39.321  < 2e-16 ***
## tnie_rr          1.39317   0.03394 1.32664 1.45971 41.042  < 2e-16 ***
## te_rr            2.95466   0.21668 2.52998 3.37933 13.636  < 2e-16 ***
## pm               0.42660   0.02564 0.37634 0.47685 16.638  < 2e-16 ***
## cde_err          0.20180   0.03296 0.13719 0.26640  6.122 9.59e-10 ***
## intref_err       0.91901   0.11983 0.68416 1.15386  7.670 1.89e-14 ***
## intmed_err       0.41374   0.06735 0.28174 0.54574  6.143 8.40e-10 ***
## pie_err          0.42011   0.03612 0.34933 0.49090 11.632  < 2e-16 ***
## te_err           1.95466   0.21668 1.52998 2.37933  9.021  < 2e-16 ***
## cde_err_prop     0.10324   0.01273 0.07828 0.12819  8.108 5.74e-16 ***
## intref_err_prop  0.47017   0.01841 0.43409 0.50624 25.542  < 2e-16 ***
## intmed_err_prop  0.21167   0.01236 0.18744 0.23590 17.123  < 2e-16 ***
## pie_err_prop     0.21493   0.02337 0.16912 0.26074  9.195  < 2e-16 ***
## overall_pm       0.42660   0.02564 0.37634 0.47685 16.638  < 2e-16 ***
## overall_int      0.68183   0.01988 0.64287 0.72080 34.295  < 2e-16 ***
## overall_pe       0.89676   0.01273 0.87181 0.92172 70.429  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_noint,  outcome = "binY_catMbinM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C",
                yreg = "logistic", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "wb")
```

```
##                 Estimate Std.error  95% CIL  95% CIU      z     pval
## cde_rr          2.211186  0.161010 1.895613 2.526759 13.733  < 2e-16 ***
## pnde_rr         1.961444  0.122705 1.720947 2.201940 15.985  < 2e-16 ***
## tnde_rr         1.933613  0.118893 1.700587 2.166639 16.263  < 2e-16 ***
## pnie_rr         1.299738  0.023441 1.253795 1.345681 55.448  < 2e-16 ***
## tnie_rr         1.281296  0.021559 1.239041 1.323552 59.431  < 2e-16 ***
## te_rr           2.513191  0.160899 2.197835 2.828546 15.620  < 2e-16 ***
## pm              0.364625  0.024066 0.317457 0.411793 15.151  < 2e-16 ***
## cde_err         0.128571  0.021036 0.087340 0.169801  6.112 1.02e-09 ***
## intref_err      0.832873  0.107655 0.621874 1.043872  7.737 1.12e-14 ***
## intmed_err      0.252009  0.037515 0.178481 0.325537  6.718 1.95e-11 ***
## pie_err         0.299738  0.023441 0.253795 0.345681 12.787  < 2e-16 ***
## te_err          1.513191  0.160899 1.197835 1.828546  9.405  < 2e-16 ***
## cde_err_prop    0.084967  0.010252 0.064873 0.105060  8.288  < 2e-16 ***
## intref_err_prop 0.550408  0.020281 0.510659 0.590158 27.140  < 2e-16 ***
## intmed_err_prop 0.166541  0.009042 0.148819 0.184264 18.418  < 2e-16 ***
## pie_err_prop    0.198084  0.022673 0.153646 0.242522  8.737  < 2e-16 ***
## overall_pm      0.364625  0.024066 0.317457 0.411793 15.151  < 2e-16 ***
## overall_int     0.716950  0.020496 0.676778 0.757122 34.979  < 2e-16 ***
## overall_pe      0.915033  0.010252 0.894940 0.935127 89.254  < 2e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_noint,  outcome = "binY_catMbinM_noint", exposure = "A",
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C",
                yreg = "logistic", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "msm")
```

```
##                 Estimate Std.error 95% CIL 95% CIU      z     pval
## cde_rr           2.21227   0.16835 1.88230 2.54223 13.141  < 2e-16 ***
## pnde_rr          2.12830   0.15373 1.82700 2.42960 13.845  < 2e-16 ***
## tnde_rr          2.08788   0.14724 1.79929 2.37646 14.180  < 2e-16 ***
## pnie_rr          1.42368   0.04007 1.34515 1.50220 35.534  < 2e-16 ***
## tnie_rr          1.39664   0.03789 1.32237 1.47091 36.856  < 2e-16 ***
## te_rr            2.97246   0.22200 2.53735 3.40757 13.390  < 2e-16 ***
## pm               0.42798   0.02899 0.37116 0.48479 14.764  < 2e-16 ***
## cde_err          0.20230   0.03480 0.13409 0.27051  5.813 6.33e-09 ***
## intref_err       0.92599   0.12707 0.67695 1.17504  7.287 3.40e-13 ***
## intmed_err       0.42049   0.06758 0.28803 0.55294  6.222 5.10e-10 ***
## pie_err          0.42368   0.04007 0.34515 0.50220 10.575  < 2e-16 ***
## te_err           1.97246   0.22200 1.53735 2.40757  8.885  < 2e-16 ***
## cde_err_prop     0.10256   0.01341 0.07628 0.12884  7.649 2.20e-14 ***
## intref_err_prop  0.46946   0.02097 0.42836 0.51056 22.388  < 2e-16 ***
## intmed_err_prop  0.21318   0.01292 0.18786 0.23850 16.502  < 2e-16 ***
## pie_err_prop     0.21480   0.02522 0.16537 0.26422  8.518  < 2e-16 ***
## overall_pm       0.42798   0.02899 0.37116 0.48479 14.764  < 2e-16 ***
## overall_int      0.68264   0.02103 0.64142 0.72386 32.459  < 2e-16 ***
## overall_pe       0.89744   0.01341 0.87116 0.92372 66.933  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_noint,  outcome = "binY_catMbinM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C",
                yreg = "logistic", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "paramfunc", inf.method = "bootstrap", model = "iorw")
```

```
##       Estimate Std.error 95% CIL 95% CIU     z   pval
## ORtot  2.79904   0.18764 2.43127 3.16682 14.92 <2e-16 ***
## ORdir  2.10760   0.13663 1.83981 2.37539 15.43 <2e-16 ***
## ORind  1.32807   0.02723 1.27471 1.38143 48.78 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_noint,  outcome = "binY_catMbinM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C",
                yreg = "logistic", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "g-formula")
```

```
##                 Estimate Std.error 95% CIL 95% CIU      z     pval
## cde_rr           2.21119   0.15146 1.91433 2.50804 14.599  < 2e-16 ***
## pnde_rr          2.12081   0.13782 1.85069 2.39093 15.389  < 2e-16 ***
## tnde_rr          2.08058   0.13221 1.82146 2.33970 15.737  < 2e-16 ***
## pnie_rr          1.42011   0.03541 1.35072 1.48951 40.109  < 2e-16 ***
## tnie_rr          1.39317   0.03360 1.32731 1.45903 41.460  < 2e-16 ***
## te_rr            2.95466   0.19660 2.56933 3.33998 15.029  < 2e-16 ***
## pm               0.42660   0.02670 0.37426 0.47893 15.977  < 2e-16 ***
## cde_err          0.20180   0.03112 0.14081 0.26279  6.485 9.30e-11 ***
## intref_err       0.91901   0.11491 0.69380 1.14422  7.998 1.41e-15 ***
## intmed_err       0.41374   0.05958 0.29695 0.53052  6.944 4.06e-12 ***
## pie_err          0.42011   0.03541 0.35072 0.48951 11.865  < 2e-16 ***
## te_err           1.95466   0.19660 1.56933 2.33998  9.942  < 2e-16 ***
## cde_err_prop     0.10324   0.01239 0.07896 0.12751  8.336  < 2e-16 ***
## intref_err_prop  0.47017   0.02022 0.43054 0.50979 23.254  < 2e-16 ***
## intmed_err_prop  0.21167   0.01138 0.18936 0.23398 18.596  < 2e-16 ***
## pie_err_prop     0.21493   0.02359 0.16868 0.26117  9.109  < 2e-16 ***
## overall_pm       0.42660   0.02670 0.37426 0.47893 15.977  < 2e-16 ***
## overall_int      0.68183   0.02027 0.64210 0.72157 33.633  < 2e-16 ***
## overall_pe       0.89676   0.01239 0.87249 0.92104 72.406  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_noint,  outcome = "binY_catMbinM_noint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C",
                yreg = "logistic", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1, model = "ne")
```

```
##                         Estimate Std.error 95% CIL 95% CIU     z   pval
## natural direct effect    2.10197   0.14204 1.82358 2.38037 14.80 <2e-16
## natural indirect effect  1.33164   0.02706 1.27861 1.38467 49.22 <2e-16
## total effect             2.79908   0.19479 2.41731 3.18085 14.37 <2e-16
##
## natural direct effect   ***
## natural indirect effect ***
## total effect            ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

8.2 Case 8-2: Binary Outcome and Multiple Mediators With Exposure-mediator Interaction

8.2.1 Data simulation

8.2.1.1 Simulation Procedures

1. Simulate the exposure variable A from $Bernoulli(P(A=1))$.
2. Simulate the covariate C from $N(\mu_C, \sigma_C^2)$.
3. Simulate the first mediator M1 from $Multinom(\frac{1}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}$, $\frac{expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}$, $\frac{expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)})$, the second mediator M2 from $Bernoulli(expit(\beta_{03}+\beta_{13}*A+\beta_{23}*C))$.
4. Simulate the outcome Y from $Bernoulli(expit(\theta_0+\theta_1 A+\theta_{21}I\{M1==1\}+\theta_{22}I\{M1==2\}+\theta_{23}M2+\theta_{31}AM2+\theta_4 C))$.

8.2.1.2 True Parameters

Table 18: True Model Parameters for Data Simulation

Sample Size	θ_0	θ_1	θ_{21}	θ_{22}	θ_{23}	θ_{31}	θ_4	β_{01}	β_{11}
10000	-5	0.8	1.8	1.2	1.5	0.2	0.1	-0.25	0.5
β_{21}	β_{02}	β_{12}	β_{22}	β_{03}	β_{13}	β_{23}	P(A=1)	μ_C	σ_C
0.2	-0.3	0.4	0.3	-0.25	0.5	0.2	0.4	1	1

8.2.1.3 True Models

True model for the first mediator:

$$ln\frac{P(M1==1)}{P(M1==0)} = \beta_{01}+\beta_{11}a+\beta_{21}c$$

$$ln\frac{P(M1==2)}{P(M1==0)} = \beta_{02}+\beta_{12}a+\beta_{22}c$$

True model for the second mediator:

$$logitE[M2|a,c] = \beta_{03}+\beta_{13}a+\beta_{23}c$$

True model for the outcome:

$$logitE[Y|a,m^*,c] = \theta_0+\theta_1 a+\theta_{21}I\{m1^*==1\}+\theta_{22}I\{m1^*==2\}+\theta_{23}m2^*+\theta_{31}am2^*+\theta_4 c$$

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

```
cmest(data = df_multipleM_EMint, outcome = "binY_catMbinM_EMint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C",
                EMint = TRUE, EMint.terms = c("A*M_bin"),
                yreg = "logistic", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "rb")
```

```
##                 Estimate Std.error 95% CIL 95% CIU      z     pval
## cde_rr           2.20302   0.39413 1.43054 2.97551  5.590 2.34e-08 ***
## pnde_rr          2.37887   0.21472 1.95803 2.79972 11.079  < 2e-16 ***
## tnde_rr          2.38656   0.17932 2.03510 2.73802 13.309  < 2e-16 ***
## pnie_rr          1.39028   0.04269 1.30661 1.47395 32.568  < 2e-16 ***
## tnie_rr          1.39478   0.04092 1.31458 1.47497 34.090  < 2e-16 ***
## te_rr            3.31799   0.29371 2.74233 3.89366 11.297  < 2e-16 ***
## pm               0.40514   0.03212 0.34218 0.46811 12.612  < 2e-16 ***
## cde_err          0.22992   0.06644 0.09971 0.36013  3.461 0.000541 ***
## intref_err       1.14895   0.15676 0.84172 1.45619  7.330 2.49e-13 ***
## intmed_err       0.54884   0.08008 0.39189 0.70579  6.854 7.61e-12 ***
## pie_err          0.39028   0.04269 0.30661 0.47395  9.142  < 2e-16 ***
## te_err           2.31799   0.29371 1.74233 2.89366  7.892 3.28e-15 ***
## cde_err_prop     0.09919   0.01966 0.06066 0.13772  5.046 4.60e-07 ***
## intref_err_prop  0.49567   0.01960 0.45724 0.53409 25.283  < 2e-16 ***
## intmed_err_prop  0.23677   0.01948 0.19859 0.27495 12.155  < 2e-16 ***
## pie_err_prop     0.16837   0.01893 0.13126 0.20548  8.893  < 2e-16 ***
## overall_pm       0.40514   0.03212 0.34218 0.46811 12.612  < 2e-16 ***
## overall_int      0.73244   0.01758 0.69799 0.76689 41.671  < 2e-16 ***
## overall_pe       0.90081   0.01966 0.86228 0.93934 45.822  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_EMint,  outcome = "binY_catMbinM_EMint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C",
                EMint = TRUE, EMint.terms = c("A*M_bin"),
                yreg = "logistic", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "wb")
```

```
##                 Estimate Std.error 95% CIL 95% CIU      z     pval
## cde_rr           2.20302   0.35516 1.50691 2.89912  6.203 5.77e-10 ***
## pnde_rr          2.29097   0.14645 2.00393 2.57801 15.643  < 2e-16 ***
## tnde_rr          2.27691   0.14381 1.99504 2.55878 15.832  < 2e-16 ***
## pnie_rr          1.29127   0.02401 1.24421 1.33833 53.782  < 2e-16 ***
## tnie_rr          1.28335   0.02275 1.23876 1.32794 56.413  < 2e-16 ***
## te_rr            2.94011   0.19004 2.56763 3.31259 15.471  < 2e-16 ***
## pm               0.33459   0.02213 0.29122 0.37796 15.122  < 2e-16 ***
## cde_err          0.15297   0.03858 0.07735 0.22859  3.965 7.40e-05 ***
## intref_err       1.13800   0.12752 0.88805 1.38794  8.924  < 2e-16 ***
## intmed_err       0.35787   0.05084 0.25824 0.45751  7.040 2.05e-12 ***
## pie_err          0.29127   0.02401 0.24421 0.33833 12.132  < 2e-16 ***
## te_err           1.94011   0.19004 1.56763 2.31259 10.209  < 2e-16 ***
## cde_err_prop     0.07885   0.01750 0.04454 0.11315  4.505 6.72e-06 ***
## intref_err_prop  0.58656   0.01979 0.54777 0.62535 29.637  < 2e-16 ***
## intmed_err_prop  0.18446   0.01332 0.15834 0.21058 13.843  < 2e-16 ***
## pie_err_prop     0.15013   0.01728 0.11626 0.18400  8.688  < 2e-16 ***
## overall_pm       0.33459   0.02213 0.29122 0.37796 15.122  < 2e-16 ***
## overall_int      0.77102   0.02294 0.72607 0.81598 33.614  < 2e-16 ***
## overall_pe       0.92115   0.01750 0.88685 0.95546 52.630  < 2e-16 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_EMint,  outcome = "binY_catMbinM_EMint", exposure = "A",
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C",
                EMint = TRUE, EMint.terms = c("A*M_bin"),
                yreg = "logistic", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "msm")
```

```
##                 Estimate Std.error 95% CIL 95% CIU      z     pval
## cde_rr           2.18872   0.38393 1.43622 2.94121  5.701 1.23e-08 ***
## pnde_rr          2.34715   0.20128 1.95265 2.74165 11.661  < 2e-16 ***
## tnde_rr          2.35308   0.16871 2.02242 2.68374 13.948  < 2e-16 ***
## pnie_rr          1.38509   0.03514 1.31621 1.45397 39.413  < 2e-16 ***
## tnie_rr          1.38859   0.03742 1.31525 1.46194 37.106  < 2e-16 ***
## te_rr            3.25924   0.25917 2.75128 3.76720 12.576  < 2e-16 ***
## pm               0.40372   0.03248 0.34005 0.46738 12.428  < 2e-16 ***
## cde_err          0.23367   0.06676 0.10281 0.36453  3.500 0.000467 ***
## intref_err       1.11348   0.14595 0.82742 1.39953  7.629 2.58e-14 ***
## intmed_err       0.52700   0.06901 0.39174 0.66225  7.637 2.43e-14 ***
## pie_err          0.38509   0.03514 0.31621 0.45397 10.958  < 2e-16 ***
## te_err           2.25924   0.25917 1.75128 2.76720  8.717  < 2e-16 ***
## cde_err_prop     0.10343   0.02232 0.05969 0.14717  4.635 3.62e-06 ***
## intref_err_prop  0.49286   0.01711 0.45932 0.52639 28.806  < 2e-16 ***
## intmed_err_prop  0.23326   0.02030 0.19348 0.27305 11.493  < 2e-16 ***
## pie_err_prop     0.17045   0.01885 0.13350 0.20741  9.040  < 2e-16 ***
## overall_pm       0.40372   0.03248 0.34005 0.46738 12.428  < 2e-16 ***
## overall_int      0.72612   0.01975 0.68740 0.76483 36.760  < 2e-16 ***
## overall_pe       0.89657   0.02232 0.85283 0.94031 40.175  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_EMint,  outcome = "binY_catMbinM_EMint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C",
                EMint = TRUE, EMint.terms = c("A*M_bin"),
                yreg = "logistic", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "paramfunc", inf.method = "bootstrap", model = "iorw")
```

```
##       Estimate Std.error 95% CIL 95% CIU     z   pval
## ORtot  3.37798   0.24849 2.89095 3.86500 13.59 <2e-16 ***
## ORdir  2.57585   0.18096 2.22117 2.93052 14.23 <2e-16 ***
## ORind  1.31141   0.03126 1.25014 1.37267 41.95 <2e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_EMint,  outcome = "binY_catMbinM_EMint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C",
                EMint = TRUE, EMint.terms = c("A*M_bin"),
                yreg = "logistic", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "g-formula")
```

```
##                  Estimate Std.error 95% CIL 95% CIU      z     pval
## cde_rr            2.20302   0.39705 1.42482 2.98123  5.548 2.96e-08 ***
## pnde_rr           2.37887   0.21521 1.95707 2.80067 11.054  < 2e-16 ***
## tnde_rr           2.38656   0.18109 2.03163 2.74149 13.179  < 2e-16 ***
## pnie_rr           1.39028   0.03746 1.31686 1.46371 37.110  < 2e-16 ***
## tnie_rr           1.39478   0.03655 1.32313 1.46642 38.158  < 2e-16 ***
## te_rr             3.31799   0.28935 2.75088 3.88511 11.467  < 2e-16 ***
## pm                0.40514   0.02990 0.34655 0.46374 13.551  < 2e-16 ***
## cde_err           0.22992   0.06727 0.09807 0.36177  3.418 0.000634 ***
## intref_err        1.14895   0.15846 0.83839 1.45952  7.251 4.44e-13 ***
## intmed_err        0.54884   0.07680 0.39831 0.69937  7.146 9.56e-13 ***
## pie_err           0.39028   0.03746 0.31686 0.46371 10.418  < 2e-16 ***
## te_err            2.31799   0.28935 1.75088 2.88511  8.011 1.26e-15 ***
## cde_err_prop      0.09919   0.02035 0.05930 0.13908  4.873 1.11e-06 ***
## intref_err_prop   0.49567   0.01669 0.46296 0.52837 29.703  < 2e-16 ***
## intmed_err_prop   0.23677   0.01841 0.20069 0.27285 12.862  < 2e-16 ***
## pie_err_prop      0.16837   0.01801 0.13306 0.20368  9.347  < 2e-16 ***
## overall_pm        0.40514   0.02990 0.34655 0.46374 13.551  < 2e-16 ***
## overall_int       0.73244   0.01796 0.69724 0.76764 40.780  < 2e-16 ***
## overall_pe        0.90081   0.02035 0.86092 0.94070 44.259  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_EMint,  outcome = "binY_catMbinM_EMint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C",
                EMint = TRUE, EMint.terms = c("A*M_bin"),
                yreg = "logistic", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = NULL, inf.method = NULL, model = "ne")
```

```
##                       Estimate Std.error 95% CIL 95% CIU     z   pval
## pure direct effect     2.49882   0.16296 2.17943 2.81821 15.33 <2e-16 ***
## total direct effect    2.55965   0.16859 2.22923 2.89007 15.18 <2e-16 ***
## pure indirect effect   1.31632   0.02760 1.26222 1.37041 47.69 <2e-16 ***
## total indirect effect  1.34836   0.02907 1.29139 1.40534 46.38 <2e-16 ***
## total effect           3.36931   0.22659 2.92520 3.81342 14.87 <2e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

8.3 Case 8-3: Binary Outcome and Multiple Mediators With Mediator-mediator Interaction

8.3.1 Data simulation

8.3.1.1 Simulation Procedures

1. Simulate the exposure variable A from $Bernoulli(P(A=1))$.
2. Simulate the covariate C from $N(\mu_C, \sigma_C^2)$.
3. Simulate the first mediator M1 from $Multinom(\frac{1}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}$, $\frac{expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}$, $\frac{expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)})$, the second mediator M2 from $Bernoulli(expit(\beta_{03}+\beta_{13}*A+\beta_{23}*C))$.
4. Simulate the outcome Y from $Bernoulli(expit(\theta_0+\theta_1 A+\theta_{21}I\{M1==1\}+\theta_{22}I\{M1==2\}+\theta_{23}M2+\theta_{31}I\{M1==1\}M2+\theta_{32}I\{M1==2\}M2+\theta_4 C))$.

8.3.1.2 True Parameters

Table 19: True Model Parameters for Data Simulation

Sample Size	θ_0	θ_1	θ_{21}	θ_{22}	θ_{23}	θ_{31}	θ_{32}	θ_4	β_{01}	β_{11}
10000	-5	0.8	1.8	1.2	1.5	0.2	0.4	0.1	-0.25	0.5
β_{21}	β_{02}	β_{12}	β_{22}	β_{03}	β_{13}	β_{23}	P(A=1)	μ_C	σ_C	
0.2	-0.3	0.4	0.3	-0.25	0.5	0.2	0.4	1	1	

8.3.1.3 True Models

True model for the first mediator:

$$ln\frac{P(M1==1)}{P(M1==0)} = \beta_{01}+\beta_{11}a+\beta_{21}c$$

$$ln\frac{P(M1==2)}{P(M1==0)} = \beta_{02}+\beta_{12}a+\beta_{22}c$$

True model for the second mediator:

$$logitE[M2|a,c] = \beta_{03}+\beta_{13}a+\beta_{23}c$$

True model for the outcome:

$$logitE[Y|a,m^*,c] = \theta_0+\theta_1 a+\theta_{21}I\{m1^*==1\}+\theta_{22}I\{m1^*==2\}+\theta_{23}m2^*+$$

$$\theta_{31}I\{m1^*==1\}m2^*+\theta_{32}I\{m1^*==2\}m2^*+\theta_4 c$$

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

```
cmest(data = df_multipleM_MMint,  outcome = "binY_catMbinM_MMint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C", MMint = TRUE,
                yreg = "logistic", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "wb")
```

```
##                 Estimate Std.error  95% CIL  95% CIU      z     pval
## cde_rr          2.220826  0.141766 1.942971 2.498682 15.665  < 2e-16 ***
## pnde_rr         1.908047  0.099562 1.712908 2.103186 19.164  < 2e-16 ***
## tnde_rr         1.877412  0.096175 1.688913 2.065912 19.521  < 2e-16 ***
## pnie_rr         1.315246  0.026263 1.263772 1.366720 50.080  < 2e-16 ***
## tnie_rr         1.294129  0.024372 1.246362 1.341897 53.100  < 2e-16 ***
## te_rr           2.469260  0.128421 2.217558 2.720961 19.228  < 2e-16 ***
## pm              0.381970  0.025870 0.331266 0.432674 14.765  < 2e-16 ***
## cde_err         0.120380  0.034666 0.052436 0.188324  3.473 0.000518 ***
## intref_err      0.787667  0.088032 0.615127 0.960207  8.947  < 2e-16 ***
## intmed_err      0.245966  0.030884 0.185435 0.306498  7.964 1.84e-15 ***
## pie_err         0.315246  0.026263 0.263772 0.366720 12.004  < 2e-16 ***
## te_err          1.469260  0.128421 1.217558 1.720961 11.441  < 2e-16 ***
## cde_err_prop    0.081932  0.021265 0.040254 0.123611  3.853 0.000117 ***
## intref_err_prop 0.536098  0.028079 0.481064 0.591132 19.093  < 2e-16 ***
## intmed_err_prop 0.167408  0.008621 0.150511 0.184306 19.418  < 2e-16 ***
## pie_err_prop    0.214561  0.023118 0.169250 0.259872  9.281  < 2e-16 ***
## overall_pm      0.381970  0.025870 0.331266 0.432674 14.765  < 2e-16 ***
## overall_int     0.703506  0.026729 0.651118 0.755895 26.320  < 2e-16 ***
## overall_pe      0.918068  0.021265 0.876389 0.959746 43.173  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_MMint,  outcome = "binY_catMbinM_MMint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C", MMint = TRUE,
                yreg = "logistic", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "paramfunc", inf.method = "bootstrap", model = "iorw",)
```

```
##       Estimate Std.error 95% CIL 95% CIU     z   pval
## ORtot  2.81258   0.17930 2.46116 3.16400 15.69 <2e-16 ***
## ORdir  2.09167   0.12961 1.83764 2.34571 16.14 <2e-16 ***
## ORind  1.34465   0.02685 1.29203 1.39728 50.08 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_MMint,  outcome = "binY_catMbinM_MMint", exposure = 'A',
                exposure.type = "binary",
```

```
                mediator = c('M_cat',"M_bin"), covariates.pre = "C", MMint = TRUE,
                yreg = "logistic", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = NULL, inf.method = NULL, model = "ne")
```

```
##                          Estimate Std.error 95% CIL 95% CIU     z   pval
## natural direct effect     2.07191   0.12703 1.82294 2.32088 16.31 <2e-16
## natural indirect effect   1.35794   0.02859 1.30191 1.41397 47.50 <2e-16
## total effect              2.81353   0.18024 2.46027 3.16679 15.61 <2e-16
##
## natural direct effect   ***
## natural indirect effect ***
## total effect            ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

8.4 Case 8-4: Binary Outcome and Multiple Mediators With Exposure-mediator-mediator Interaction

8.4.1 Data simulation

8.4.1.1 Simulation Procedures

1. Simulate the exposure variable A from $Bernoulli(P(A=1))$.
2. Simulate the covariate C from $N(\mu_C, \sigma_C^2)$.
3. Simulate the first mediator M1 from $Multinom(\frac{1}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}, \frac{expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}, \frac{expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}{1+expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C)+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)})$, the second mediator M2 from $Bernoulli(expit(\beta_{03}+\beta_{13}*A+\beta_{23}*C))$.
4. Simulate the outcome Y from $Bernoulli(expit(\theta_0+\theta_1 A+\theta_{21}I\{M1==1\}+\theta_{22}I\{M1==2\}+\theta_{23}M2+\theta_{31}AI\{M1==1\}M2+\theta_{32}AI\{M1==2\}M2+\theta_4 C))$.

8.4.1.2 True Parameters

Table 20: True Model Parameters for Data Simulation

Sample Size	θ_0	θ_1	θ_{21}	θ_{22}	θ_{23}	θ_{31}	θ_{32}	θ_4	β_{01}	β_{11}
10000	-5	0.8	1.8	1.2	1.5	0.2	0.4	0.1	-0.25	0.5
β_{21}	β_{02}	β_{12}	β_{22}	β_{03}	β_{13}	β_{23}	P(A=1)	μ_C	σ_C	
0.2	-0.3	0.4	0.3	-0.25	0.5	0.2	0.4	1	1	

8.4.1.3 True Models

True model for the first mediator:

$$ln\frac{P(M1==1)}{P(M1==0)} = \beta_{01}+\beta_{11}a+\beta_{21}c$$

$$ln\frac{P(M1==2)}{P(M1==0)} = \beta_{02}+\beta_{12}a+\beta_{22}c$$

True model for the second mediator:

$$logitE[M2|a,c] = \beta_{03} + \beta_{13}a + \beta_{23}c$$

True model for the outcome:

$$logitE[Y|a,m^*,c] = \theta_0 + \theta_1 a + \theta_{21} I\{m1^* == 1\} + \theta_{22} I\{m1^* == 2\} + \theta_{23} m2^* +$$

$$\theta_{31} a I\{m1^* == 1\} m2^* + \theta_{32} a I\{m1^* == 2\} m2^* + \theta_4 c$$

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

```
cmest(data = df_multipleM_EMMint,  outcome = "binY_catMbinM_EMMint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C", EMMint = TRUE,
                yreg = "logistic", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "wb")
```

```
##                  Estimate Std.error  95% CIL  95% CIU      z     pval
## cde_rr            3.81886   3.81265 -3.65379 11.29150  1.002   0.3165
## pnde_rr           2.31616   0.14800  2.02607  2.60624 15.649  < 2e-16 ***
## tnde_rr           2.33340   0.14435  2.05048  2.61631 16.165  < 2e-16 ***
## pnie_rr           1.28813   0.02456  1.23998  1.33627 52.438  < 2e-16 ***
## tnie_rr           1.29771   0.02292  1.25278  1.34264 56.608  < 2e-16 ***
## te_rr             3.00571   0.19024  2.63285  3.37857 15.800  < 2e-16 ***
## pm                0.34379   0.02215  0.30039  0.38720 15.523  < 2e-16 ***
## cde_err           0.25759   0.12568  0.01127  0.50392  2.050   0.0404 *
## intref_err        1.05857   0.16324  0.73861  1.37852  6.485 9.32e-11 ***
## intmed_err        0.40142   0.05289  0.29776  0.50509  7.589 3.50e-14 ***
## pie_err           0.28813   0.02456  0.23998  0.33627 11.729  < 2e-16 ***
## te_err            2.00571   0.19024  1.63285  2.37857 10.543  < 2e-16 ***
## cde_err_prop      0.12843   0.06006  0.01071  0.24615  2.138   0.0325 *
## intref_err_prop   0.52778   0.05849  0.41313  0.64242  9.023  < 2e-16 ***
## intmed_err_prop   0.20014   0.01571  0.16936  0.23092 12.744  < 2e-16 ***
## pie_err_prop      0.14365   0.01612  0.11207  0.17524  8.913  < 2e-16 ***
## overall_pm        0.34379   0.02215  0.30039  0.38720 15.523  < 2e-16 ***
## overall_int       0.72792   0.06199  0.60642  0.84941 11.743  < 2e-16 ***
## overall_pe        0.87157   0.06006  0.75385  0.98929 14.511  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_EMMint,  outcome = "binY_catMbinM_EMMint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C", EMMint = TRUE,
                yreg = "logistic", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "paramfunc", inf.method = "bootstrap", model = "iorw")
```

```
##       Estimate Std.error 95% CIL 95% CIU     z   pval
## ORtot  3.47093   0.22789 3.02428 3.91758 15.23 <2e-16 ***
## ORdir  2.64671   0.16737 2.31866 2.97476 15.81 <2e-16 ***
## ORind  1.31141   0.02611 1.26024 1.36258 50.23 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_EMMint,  outcome = "binY_catMbinM_EMMint", exposure = 'A',
                exposure.type = "binary",
                mediator = c('M_cat',"M_bin"), covariates.pre = "C", EMMint = TRUE,
                yreg = "logistic", mreg = c("multinomial", "logistic"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = NULL, inf.method = NULL, model = "ne")
```

```
##                         Estimate Std.error 95% CIL 95% CIU     z   pval
## natural direct effect    2.57051   0.16664 2.24391 2.89711 15.43 <2e-16
## natural indirect effect  1.34598   0.02813 1.29085 1.40111 47.85 <2e-16
## total effect             3.45985   0.23376 3.00169 3.91801 14.80 <2e-16
##
## natural direct effect   ***
## natural indirect effect ***
## total effect            ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

9 Case 9: Post-exposure Confounding

9.1 Data simulation

9.1.1 Simulation Procedures

1. Simulate the exposure variable A from $Bernoulli(P(A=1))$.
2. Simulate the covariate C from $N(\mu_C, \sigma_C^2)$.
3. Simulate the first post-exposure confounder L1 from $Bernoulli(expit(\beta_{01}+\beta_{11}*A+\beta_{21}*C))$ the second post-exposure confounder L2 from $Multinom(\frac{1}{1+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)+expit(\beta_{03}+\beta_{13}*A+\beta_{23}*C)}, \frac{expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)}{1+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)+expit(\beta_{03}+\beta_{13}*A+\beta_{23}*C)}, \frac{expit(\beta_{03}+\beta_{13}*A+\beta_{23}*C)}{1+expit(\beta_{02}+\beta_{12}*A+\beta_{22}*C)+expit(\beta_{03}+\beta_{13}*A+\beta_{23}*C)})$.
4. Simulate the first mediator M1 from $Multinom$

$(\frac{1}{1+expit(\beta_{04}+\beta_{14}*A+\beta_{24}*C+\beta_{34}*L1+\beta_{44}*I\{L2==1\}+\beta_{54}*I\{L2==2\})+expit(\beta_{05}+\beta_{15}*A+\beta_{25}*C+\beta_{35}*L1+\beta_{45}*I\{L2==1\}+\beta_{55}*I\{L2==2\})}$,
$\frac{expit(\beta_{04}+\beta_{14}*A+\beta_{24}*C+\beta_{34}*L1+\beta_{44}*I\{L2==1\}+\beta_{54}*I\{L2==2\})}{1+expit(\beta_{04}+\beta_{14}*A+\beta_{24}*C+\beta_{34}*L1+\beta_{44}*I\{L2==1\}+\beta_{54}*I\{L2==2\})+expit(\beta_{05}+\beta_{15}*A+\beta_{25}*C+\beta_{35}*L1+\beta_{45}*I\{L2==1\}+\beta_{55}*I\{L2==2\})}$,
$\frac{expit(\beta_{05}+\beta_{15}*A+\beta_{25}*C+\beta_{35}*L1+\beta_{45}*I\{L2==1\}+\beta_{55}*I\{L2==2\})}{1+expit(\beta_{04}+\beta_{14}*A+\beta_{24}*C+\beta_{34}*L1+\beta_{44}*I\{L2==1\}+\beta_{54}*I\{L2==2\})+expit(\beta_{05}+\beta_{15}*A+\beta_{25}*C+\beta_{35}*L1+\beta_{45}*I\{L2==1\}+\beta_{55}*I\{L2==2\})})$,
the second mediator M2 from $Bernoulli(expit(\beta_{06}+\beta_{16}*A+\beta_{26}*C+\beta_{36}*L1+\beta_{46}*I\{L2==1\}+\beta_{56}*I\{L2==2\}))$.

5. Simulate the outcome Y from $Bernoulli(expit(\theta_0+\theta_1 A+\theta_{21}I\{M1==1\}+\theta_{22}I\{M1==2\}+\theta_{23}M2+\theta_{31}AI\{M1==1\}M2+\theta_{32}AI\{M1==2\}M2+\theta_4 C+\theta_5*L1+\theta_6*I\{L2==1\}+\theta_7*I\{L2==2\}))$.

9.1.2 True Parameters

Table 21: True Model Parameters for Data Simulation

Sample Size	θ_0	θ_1	θ_{21}	θ_{22}	θ_{23}	θ_{31}	θ_4	θ_5	θ_6	θ_7
10000	-5	0.8	1.8	1.2	1.5	0.2	0.1	0.3	0.4	0.2
β_{01}	β_{11}	β_{21}	β_{02}	β_{12}	β_{22}	β_{03}	β_{13}	β_{23}	β_{04}	β_{14}
-0.25	0.5	0.2	-0.25	0.5	0.2	-0.3	0.4	0.3	-0.25	0.5
β_{24}	β_{34}	β_{44}	β_{54}	β_{05}	β_{15}	β_{25}	β_{35}	β_{45}	β_{55}	β_{06}
0.2	0.1	0.3	0.25	-0.3	0.4	0.3	0.5	0.1	0.2	-0.25
β_{16}	β_{26}	β_{36}	β_{46}	β_{56}	P(A=1)	μ_C	σ_C			
0.5	0.2	0.1	0.3	0.25	0.4	1	1			

9.1.3 True Models

True model for the first post-exposure confounder:

$$logitE[L1|a,c]=\beta_{01}+\beta_{11}a+\beta_{21}c$$

True model for the second post-exposure confounder:

$$ln\frac{P(L2==1)}{P(L2==0)}=\beta_{02}+\beta_{12}a+\beta_{22}c$$

$$ln\frac{P(L2==2)}{P(L2==0)}=\beta_{03}+\beta_{13}a+\beta_{23}c$$

True model for the first mediator:

$$ln\frac{P(M1==1)}{P(M1==0)}=\beta_{04}+\beta_{14}a+\beta_{24}c+\beta_{34}l1+\beta_{44}I\{l2==1\}+\beta_{54}I\{l2==2\}$$

$$ln\frac{P(M1==2)}{P(M1==0)}=\beta_{05}+\beta_{15}a+\beta_{25}c+\beta_{35}l1+\beta_{45}I\{l2==1\}+\beta_{55}I\{l2==2\}$$

True model for the second mediator:

$$logitE[M2|a,c]=\beta_{06}+\beta_{16}a+\beta_{26}c+\beta_{36}l1+\beta_{46}I\{l2==1\}+\beta_{56}I\{l2==2\}$$

True model for the outcome:

$$logitE[Y|a,m^*,c]=\theta_0+\theta_1 a+\theta_{21}I\{m1^*==1\}+\theta_{22}I\{m1^*==2\}+\theta_{23}m2^*+$$

$$\theta_{31}aI\{m1^*==1\}m2^*+\theta_{32}aI\{m1^*==2\}m2^*+\theta_4 c+\theta_5 l1+\theta_6 I\{l2==1\}+\theta_7 I\{l2==2\}$$

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

```
cmest(data = df_multipleM_EMint_postcovar,  outcome = "binY_catMbinM_EMint",
                exposure = "A", exposure.type = "binary",
                mediator = c('M_bin',"M_cat"), covariates.pre = "C",
                covariates.post = c("L_bin", "L_cat"),
                covariates.post.type = c("binary", "categorical"),
                EMint = TRUE, EMint.terms = c("A*M_bin"),
                yreg = "logistic", mreg = c("logistic", "multinomial"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "msm")
```

```
##                 Estimate Std.error 95% CIL 95% CIU      z     pval
## cde_rr           2.46338   0.33739 1.80211 3.12466  7.301 3.07e-13 ***
## pnde_rr          2.45529   0.15360 2.15423 2.75635 15.985  < 2e-16 ***
## tnde_rr          2.40918   0.12907 2.15620 2.66216 18.665  < 2e-16 ***
## pnie_rr          1.37109   0.03203 1.30831 1.43387 42.806  < 2e-16 ***
## tnie_rr          1.34534   0.03102 1.28455 1.40614 43.373  < 2e-16 ***
## te_rr            3.30321   0.20069 2.90985 3.69656 16.459  < 2e-16 ***
## pm               0.36815   0.02529 0.31859 0.41771 14.559  < 2e-16 ***
## cde_err          0.22622   0.04785 0.13243 0.32000  4.728 2.30e-06 ***
## intref_err       1.22907   0.11872 0.99638 1.46176 10.353  < 2e-16 ***
## intmed_err       0.47683   0.05967 0.35987 0.59378  7.991 1.49e-15 ***
## pie_err          0.37109   0.03203 0.30831 0.43387 11.586  < 2e-16 ***
## te_err           2.30321   0.20069 1.90985 2.69656 11.476  < 2e-16 ***
## cde_err_prop     0.09822   0.01671 0.06547 0.13096  5.879 4.27e-09 ***
## intref_err_prop  0.53364   0.01570 0.50286 0.56441 33.986  < 2e-16 ***
## intmed_err_prop  0.20703   0.01798 0.17180 0.24226 11.517  < 2e-16 ***
## pie_err_prop     0.16112   0.01448 0.13275 0.18949 11.130  < 2e-16 ***
## overall_pm       0.36815   0.02529 0.31859 0.41771 14.559  < 2e-16 ***
## overall_int      0.74066   0.01751 0.70635 0.77497 42.310  < 2e-16 ***
## overall_pe       0.90178   0.01671 0.86904 0.93453 53.975  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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

```
cmest(data = df_multipleM_EMint_postcovar,  outcome = "binY_catMbinM_EMint",
                exposure = "A", exposure.type = "binary",
                mediator = c('M_bin',"M_cat"), covariates.pre = "C",
                covariates.post = c("L_bin", "L_cat"),
                covariates.post.type = c("binary", "categorical"),
                EMint = TRUE, EMint.terms = c("A*M_bin"),
                yreg = "logistic", mreg = c("logistic", "multinomial"), mval = list(0,0),
                a_star = 0, a = 1,
                est.method = "imputation", inf.method = "bootstrap", model = "g-formula")
```

```
##                 Estimate Std.error 95% CIL 95% CIU      z     pval
## cde_rr           2.38262   0.30027 1.79410 2.97115  7.935 2.33e-15 ***
## pnde_rr          2.47505   0.15571 2.16987 2.78023 15.895  < 2e-16 ***
## tnde_rr          2.44945   0.13415 2.18652 2.71238 18.259  < 2e-16 ***
## pnie_rr          1.36986   0.03526 1.30076 1.43896 38.855  < 2e-16 ***
## tnie_rr          1.35569   0.03294 1.29112 1.42025 41.155  < 2e-16 ***
## te_rr            3.35539   0.20704 2.94961 3.76118 16.207  < 2e-16 ***
## pm               0.37376   0.02609 0.32262 0.42490 14.324  < 2e-16 ***
```

```
## cde_err          0.21320   0.04521 0.12459 0.30182  4.715 2.44e-06 ***
## intref_err       1.26185   0.12274 1.02129 1.50241 10.281  < 2e-16 ***
## intmed_err       0.51048   0.06121 0.39051 0.63046  8.340  < 2e-16 ***
## pie_err          0.36986   0.03526 0.30076 0.43896 10.491  < 2e-16 ***
## te_err           2.35539   0.20704 1.94961 2.76118 11.377  < 2e-16 ***
## cde_err_prop     0.09052   0.01542 0.06030 0.12073  5.871 4.46e-09 ***
## intref_err_prop  0.53573   0.01773 0.50098 0.57047 30.220  < 2e-16 ***
## intmed_err_prop  0.21673   0.01668 0.18404 0.24942 12.995  < 2e-16 ***
## pie_err_prop     0.15703   0.01529 0.12707 0.18698 10.273  < 2e-16 ***
## overall_pm       0.37376   0.02609 0.32262 0.42490 14.324  < 2e-16 ***
## overall_int      0.75246   0.01678 0.71956 0.78535 44.834  < 2e-16 ***
## overall_pe       0.90948   0.01542 0.87927 0.93970 58.994  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```