Dealing with time-varying confounding using SAS

Ryo Nakaya, Sho Sakui, Takamasa Hashimoto
Takeda Pharmaceutical Company Limited

PharmaSUG Single Day Event Tokyo 2018 on 4th September

Background

Time-varying confounding should be treated appropriately in observational research area, since choice of therapy in daily setting is usually affected by patient status, including treatment and covariate history. CAUSALTRT procedure in SAS does NOT cover time-varying inverse probability weighting (IPW) methods, while it handles non-time-varying IPW. Examples of SAS programming of time-varying IPW will be presented using simulation data with background of the methodology.
• Time-varying treatment A and covariate L, unmeasured confounder U
• **Our interest is to estimate joint causal effects of**
  \[ A(\text{e.g. } E(Y(a_0, a_1))) \], which is total sum of green parts, where
  \[ E(Y(a_0, a_1)) \] represents expected value of a potential outcome that
  would have been observed had the treatment been set to specific
  levels \(a_0\) and \(a_1\).
Why Traditional Methods Fail?

• Unconditioning on $L_1$, neither effect of $A_0$ nor $A_1$ will be estimated without bias because two paths (red and blue) open.

• Conditioning on $L_1$, effect of $A_0$ will not be estimated without bias because blue path opens and red path closes.

Note: Circled variables in DAG stand for conditioned (adjusted in model or stratified).
Theoretical background

• Identifiability conditions
  – Consistency. If 
    \((A_0, A_1) = (a_0, a_1)\) for a given subject, then 
    \(Y^{(a_0,a_1)} = Y\) for that subject
  – Conditional exchangeability.
    \(Y^{(a_0,a_1)} \perp A_1 \mid L_1, A_0\)
    \(Y^{(a_0,a_1)} \perp A_0\)
  – Positivity.
    \(0 < f_{A_1 \mid L_1, A_0}(A_1 \mid L_1, A_0)\)
    \(0 < f_{A_0}(A_0)\),
    where \(f\) expresses (conditional) probability mass function
• When our interest is counterfactual mean, i.e.,
\[ E(Y^{(a_0^*, a_1^*)}) \]
where \((a_0^*, a_1^*)\) is treatment level of interest,

we will show this is equivalent to
\[ E \left( I(A = \ldots \right) \]
Using the identifiability conditions, IPW mean of $Y$ is:

$$
\int I(\bar{A} = (a_0^*, a_1^*)) y \frac{f(y, a_0, a_1, l_1)}{ f_{A_0}(a_0) f_{A_1|L_1,A_0}(a_1|l_1, a_0)} \, dy \, d\bar{a} \, dl_1
$$

$$
= \int y \frac{f(y, a_0^*, a_1^*, l_1)}{ f_{A_0}(a_0^*) f_{A_1|L_1,A_0}(a_1^*|l_1, a_0^*)} \, dy \, dl_1
$$

$$
= \int y f_{Y|A_0,L_1,A_1}(y|a_0^*, l_1, a_1^*) f_{A_1|L_1,A_0}(a_1^*|l_1, a_0^*) f_{L_1|A_0}(l_1|a_0^*) f_{A_0}(a_0^*) \, dy \, dl_1
$$

$$
= \int y f_{Y|A_0,L_1,A_1}(y|a_0^*, l_1, a_1^*) f_{L_1|A_0}(l_1|a_0^*) \, dy \, dl_1
$$

$$
= \int y^{(a_0^*,a_1^*)} f_{Y^{(a_0^*,a_1^*)}|A_0,L_1,A_1}(y^{(a_0^*,a_1^*)}|a_0^*, l_1, a_1^*) f_{L_1|A_0}(l_1|a_0^*) \, dy^{(a_0^*,a_1^*)} \, dl_1
$$

($\therefore$ consistency)
\[
= \int y^{(a_0^*, a_1^*)} f_{Y^{(a_0^*, a_1^*)}}(a_0^*, l_1) f_{L_1|A_0}(l_1|a_0^*) dy^{(a_0^*, a_1^*)} dl_1
\]
\[
(\because \text{conditional exchangeability } Y^{(a_0, a_1)} \perp A_1| L_1, A_0)
\]
\[
= \int y^{(a_0^*, a_1^*)} f_{Y^{(a_0^*, a_1^*)}, L_1|A_0}(y^{(a_0^*, a_1^*)}, l_1|a_0^*) dy^{(a_0^*, a_1^*)} dl_1
\]
\[
= \int y^{(a_0^*, a_1^*)} f_{Y^{(a_0^*, a_1^*)}|A_0}(y^{(a_0^*, a_1^*)}|a_0^*) dy^{(a_0^*, a_1^*)}
\]
\[
= \int y^{(a_0^*, a_1^*)} f_{Y^{(a_0^*, a_1^*)}}(y^{(a_0^*, a_1^*)}) dy^{(a_0^*, a_1^*)} = E(Y^{(a_0^*, a_1^*)})
\]
\[
(\because \text{conditional exchangeability } Y^{(a_0, a_1)} \perp A_0)
\]
Example (Traditional Methods vs IPW)

Simulation Data Generation

data ADSIM1 ;
do USUBJID = 1 to 100 ;
   A0  = rand("Bernoulli", 0.4) ;
   U   = rand("Normal", 0.3) ;
   L1  = rand("Normal", 0.1) + 2*A0 + 2*U ;
   P_A1 = exp(0.2 + 0.3*A0 + 0.2*L1)/(1 + exp(0.2 + 0.3*A0 + 0.2*L1)) ;
   A1  = rand("Bernoulli", P_A1) ;
   Y   = rand("Normal", 0.1) + 3*A0 + 5*A1 + 2*L1 + 5*U ;
 output ;
end ;
 run ;

Note: The program generates normal response Y affected by A0, A1, L1 (time-varying A and L) and unmeasured U.
Analysis Programs

/*Unadjusted GLM(GLM1)*/
proc mixed data=ADSIM1 ;
  model Y = A0 A1 / s ;
run ;

/*Traditional adjusted GLM (GLM2)*/
proc mixed data=ADSIM1 ;
  model Y = A0 A1 L1 / s ;
run ;

/*IPW_GLM*/
proc logistic data=ADSIM1 descending ;
  model A0 = ;
  output out=OUT0 p=P0 ;
run ;
proc logistic data=ADSIM1 descending ;
  model A1 = A0 L1 ;
  output out=OUT1 p=P1 ;
run ;
data OUT ;
  merge OUT0 OUT1 ;
  by USUBJID ;
  if A0=0 then P0=1-P0 ;
  if A1=0 then P1=1-P1 ;
  W=1/(P0*P1) ;
run ;
proc mixed data=OUT empirical ;
  model Y = A0 A1 / s ;
  weight W ;
  repeated Intercept / subject=USUBJID ;
run ;

Note: Sandwich variance estimator is calculated in order to take correlation of weighted(inflated) subject information into consideration.
<table>
<thead>
<tr>
<th>N=100</th>
<th>True Causal Effect</th>
<th>GLM1 (A0,A1)</th>
<th>GLM2 (A0,A1,L1)</th>
<th>IPW_GLM (weighted A0,A1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>joint(A0,A1)</td>
<td>12</td>
<td>34.1 (6.83)</td>
<td>3.14 (0.84)</td>
<td>12.96 (9.03)</td>
</tr>
<tr>
<td>A0</td>
<td>7</td>
<td>3.54 (4.92)</td>
<td>-1.84 (0.59)</td>
<td>6.09 (8.55)</td>
</tr>
<tr>
<td>A1</td>
<td>5</td>
<td>30.6 (5.02)</td>
<td>4.98 (0.65)</td>
<td>6.87 (4.92)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N=1,000</th>
<th>True Causal Effect</th>
<th>GLM1 (A0,A1)</th>
<th>GLM2 (A0,A1,L1)</th>
<th>IPW_GLM (weighted A0,A1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>joint(A0,A1)</td>
<td>12</td>
<td>34.2 (2.07)</td>
<td>3.14 (0.25)</td>
<td>12.17 (2.95)</td>
</tr>
<tr>
<td>A0</td>
<td>7</td>
<td>3.75 (1.57)</td>
<td>-1.86 (0.17)</td>
<td>6.98 (3.35)</td>
</tr>
<tr>
<td>A1</td>
<td>5</td>
<td>30.5 (1.54)</td>
<td>5.00 (0.20)</td>
<td>5.19 (2.35)</td>
</tr>
</tbody>
</table>

- 1,000 times simulation of N=100(upper table), 1,000(lower table)
- Mean (SE) of estimates
- Causal effect of A0 equals to sum of direct effect of A0 -> Y(True value of 3) and indirect effect of A0 -> L1 -> Y(True value of 4).
Summary

• **Weighted analysis using time-varying IPW was easily implemented with simple SAS codes**

• **Unbiasedness of time-varying IPW estimates was shown through the simulation**

Note

• Stabilized IPW is also applicable

• As time points increase, some more efforts to calculate weight will be needed

• Same methodology is applicable to time to event outcome

• Application to Real World Data (RWD) would be expected in the future
Reference


• 篠崎 智大(2017)『時間依存性交絡の調整』, 応用統計学会/計量生物学会年会チュートリアルセミナー

• 宮川 雅巳(2004)『統計的因果推論』, 朝倉書店