Update by Aug 17

mengyu,shuyi

8/12/2020

Daza (2018) Causal Analysis of Self-Tracked Time Series Data using a Counterfactual Framework for N-of-1 Trials

General procedures

1.design

2.notation

3.definition of causal effect

4.assumptions

Changepoint analysis here is used to detect the possible period lengths, finding where the mean of an otherwise stationary series changes over time. For each segment, exposure will be defined as high/low if the segment mean is greater/lower than 5/7.

5.the outcome model

Assumptions: 1. among a treatment level, only from unstable to stable, not destablization again
2. *washout subperiod* $\subset$ *stabilization subperiod* for any period
3. stability point k0
4. dummy variable k<k0 will be included
5. current outcome depends on previous outcome and this dependences vary by treatment level

Also, in standard N1RT(n-of-1 randomized trial), only the first treatment is randomized and then held constant for the rest of the period. i.e. $R_{t(0)} = 1$ and $R_{t(j-1)} = 0$ for $j \in (2, ..., m_t)$

Therefore, the outcome model is:

$$Y_{t(j+1)} = \beta_0 + \left\{1 - X_{t(j)}\right\}\left\{X_{t-1(1)} + \left(1 - X_{t-1(1)}\right) I\left(m_{t-1} < k_0 - 1\right)\right\} \sum_{k=1}^{k_0-1} \beta_{0k} I(j = k) + \beta_1 Y_{t(j)} +$$
$$\left[\beta_2 + \left\{\left(1 - X_{t-1(1)}\right) + X_{t-1(1)} I\left(m_{t-1} < k_0 - 1\right)\right\} \sum_{k=1}^{k_0-1} \beta_{2k} I(j = k) + \beta_3 Y_{t(j)}\right] X_{t(j)} + \mathcal{E}_{t(j)}$$

where $k_{stable} = \{k : k_0 < k < m_t\}$, $\sum_{k=1}^{k_0-1} \beta_{0k} I(j = k)$ simulate the process if the time point j is within the unstable period. $I(m_{t-1} < k_0 - 1)$ simulate the process if the period from 1 to m_t is shorter than k_0.

Then, the average baseline effect(i.e.,baseline average ACBW during weeks of low PA) is $E(Y^0_{t(k+1)}) = \beta_0 + \beta_1 E(Y_{t(k)})$ and apte is $E(Y^1_{t(k+1)}) - E(Y^0_{t(k+1)}) = \beta_2 + \beta_3 E(Y_{t(K)})$ for $k \in k_{stable}$

3.Post-Hoc analysis:

1.Analysis1: model the change in outcome from the previous outcome, which is $\triangle^{Y}_{t(j+1)} = Y_{t(j+1)} - Y_{t(j)}$:

$$\Delta^{Y}_{t(j+1)} = \beta_0 + \left\{1 - X_{t(j)}\right\}\left\{X_{t-1(1)} + \left(1 - X_{t-1(1)}\right) I\left(m_{t-1} < k_0 - 1\right)\right\} \sum_{k=1}^{k_0-1} \beta_{0k} I(j = k) + \beta_1 \Delta^{Y}_{t(j)} + \left[\left\{\left(1 - X_{t-1(1)}\right) + X_{t-1(1)} I\left(m_{t-1} < k_0 - 1\right)\right\} \sum_{k=1}^{k_0-1} \beta_{2k} I(j = k)\right] X_{t(j)} + \mathcal{E}_{t(j)}$$

Then, $apte = \sum_{k=1}^{k_0-1} E(\triangle^{Y,1}_{t(k+1)} | x_{t-1(1)} = 0, k < k_0) - (k_0 - 1) E(\triangle^{Y,0}_{t(k)} | x_{t-1(1)} = 0, k < k_0) = \sum_{k=1}^{k_0-1} \beta_{2k}$

2.Analysis2: model the change in outcome from the preivous period's last outcome or the average of its stable outcomes(when $k_0 > m_{t-1}$), $\triangle_{t(j+1)} = Y_{t(j+1)} - Y^{*}_{t-1}$ and when $m_{t-1} \leq k_0$, Y^{*}_{t-1} is the previous last outcome; when $m_{t-1} \geq k_0$, Y^{*}_{t-1} is the average of its stable outcomes.

The model is:

$$\begin{aligned}\Delta_{t(j+1)} =& \gamma_0 + \left\{1 - X_{t(j)}\right\}\left\{X_{t-1(1)} + \left(1 - X_{t-1(1)}\right) I\left(m_{t-1} < k_0 - 1\right)\right\} \sum_{k=1}^{k_0-1} \gamma_{0k} I(j = k) + \gamma_1 \Delta^{Y}_{t-1} + \\ & \left[\gamma_2 + \left\{\left(1 - X_{t-1(1)}\right) + X_{t-1(1)} I\left(m_{t-1} < k_0 - 1\right)\right\} \sum_{k=1}^{k_0-1} \gamma_{2k} I(j = k) + \gamma_3 \Delta^{Y}_{t-1}\right] X_{t(j)} + \mathcal{E}_{t(j)}\end{aligned}$$

The abe(baseline average change in ACBW from Y^{*}_{t-1}) is $E(\triangle^{0}_{t(k+1)}) = \nu_0 + \nu_1 E(\triangle^{Y}_{t-1})$ and apte is $E(\triangle^{1}_{t(k+1)}) - E(\triangle^{0}_{t(k+1)}) = \nu_2 + \nu_3 E(\triangle^{Y}_{t-1})$ for $k \in k_{stable}$.

Daza (2019) Person as Population:...

First using change point analysis to specify periods, and then transform the exposure into a binary-valued treatment $X_t = 1$ or $X_t = 0$ for binary exposure. Non-exposure causes can include lagged outcome and exposures.

APTE is a set of contrasts over a period $(E(Y^{a=1}) - E(Y^{a=0}))$. When $X_t = 1$ and $X_{t(j)} = 1$, the potential outcome $Y^{a=1}_{t(j)}$ is equal to observed outcome $Y_{t(j)}$. When $X_t = 1$ and $X_{t(j)} = 0$, the potential outcome $Y^{a=1}_{t(j)}$ is "missing", so we need to predict itconditional on exposure and non-exposure causes by using the model we specified or the model obtained by supervised learning, and vice versa. After this step, we will get time series Y^{1}_{t} and Y^{0}_{t} for each time period t.

Then with predicted potential outcome at some time points and with observed outcome at other time points (should all be potential outcomes, but with consistency assumption, some potential outcomes can be represented by observed outcomes), the weighted average over all non-exposure causes using the g-formula is taken to get the **trajectory of estimated mean POs**, which are $\hat{E}(Y^1)$ and $\hat{E}(Y^0)$, for each time period. The weights are derived from the empirical **marginal CSC distribution** as required by the g-formula.

Finalliy, report the estimated APTE as the trajectory of pre-defined **contrasts** (e.g. $\hat{E}(Y^1) - \hat{E}(Y^0)$).

From Table 1 in the article, APTE is calculated at 12 time points.

Formula from Linda

Assume that $X_{t(j)}$ precedes $Y_{t(j)}$

- Time point-wise treatment effect

$$E\left[Y_{t(j)}^{a=X_{t(j)}}\right] - E\left[Y_{t(j)}^{a^*=X_{t(j)}^*}\right]$$

* Cumulative treatment effect

$$\boldsymbol{a} = a_{t,1}, ..., a_{t,j}$$

$$\bar{X}_{t(j)} = (X_{t,1}, ..., X_{t,j})$$

$$E\left[Y_{t(j)}^{\boldsymbol{a}}\right] - E\left[Y_{t(j)}^{\boldsymbol{a}^*}\right]$$

The difference among the three ways to calculate the APTE

The definition and method of APTE estimation are different between Linda and Daza. The APTE mentioned in Daza (2018) is calculated condition on *previous period's exposure being equal to 0* in observed data, and estimate the β_{2k} which are the coefficients of exposure at the time point before stable point k_0. The final estimated APTE we get is the APTE average across 6 six years.

The APTE mentioned in Daza (2019) is not conditioned on the previous period's exposure level. Instead, each period is regard as an individual, and then, the weighted POs are estimated with weights derived from the empirical marginal CSC distribution as required by the **g-formula**. Plus, no stability point is mentioned, but the stationarity of confounder or simultaneous cause (CSC) need to be statisfied for the *APTE g-formula estimator* to be consistent for an effect-stable APTE. The final estimated APTE is for each time point (a trajectory of pre-defined contrast), and since each period is seen as an individual and have different length, the length of APTE is decided by the shortest period which is over at least 11 weeks (the analysis results from Daza (2018), saying that "estimated APTE may have been stable between 11 and 17 weeks of high PA").

As for Linda's APTE, the effect is considered within the current period (or including some of the the previous period time points if we change the defination of $\bar{X}_{t(j)}$), which means the Linda's cumulative effect include the *non-exposure causes* that defined in Daza (2019).

Modeling (state space model)

Shu Li (*Estimating heterogeneous treatment effects in nonstationary time series with state-space models*, 2020) proposed a method Causal Transfer to estimate the causal effect in non-stationary time series. Causal Transfer can be used to predict population and sample treatment efect, and we will use the sample version to calculate the $apte_j$ we defined before.

1. Estimate the unknown parameters with MLE and plug them into the model.
2. For all time points, smoothing distribution $N(s_{t(j)}, S_{t(j)})$ of the states $\theta_{t(j)}$ will be estimated by iterating the smoothing recursions (need to change notations)

$$\begin{aligned} s_t &= m_t + C_t G_{t+1}' R_{t+1}^{-1}(s_{t+1} - a_{t+1}) \\ S_t &= C_t G_{t+1}' R_{t+1}^{-1}(R_{t+1} - S_{t+1}) R_{t+1}^{-1} G_{t+1} C_t. \end{aligned}$$

3. Given the smoothing distribution, we estimate the distribution $N(\tilde{a}_j, \tilde{P}_j)$ of the **counterfactuals** with $\tilde{a}_j = \tilde{Z}_j s_j$ and $\tilde{P}_j = \tilde{Z}_j S_j \tilde{Z}_j' + H_j$. $\tilde{Z}_j$ is the same as Z_j except for X_j being replaced by $\tilde{X}_j = |X_j - 1|$.

4. Draw B sets of samples $(\tilde{y}_t^{(b)})_{b=1}^B$ from $N(\tilde{a}_j, \tilde{P}_j)$ to impute the missing outcome (counterfactuals).

5. The effect would be calcualted by $\hat{\tau}_t^{(b)} = (\tilde{y}_t^{(b)} - y_t)(\tilde{X}_j - X_j)$.

6. Finally, we will get the mean and interval of estimated effects by taking average and percentile of $(\tilde{y}_t^{(b)})_{b=1}^B$.