Biometrics 000, 000–000 DOI: 000
000 0000

Supporting Informatin for "Quantifying Direct and Indirect Effect for Longitudinal Mediator and Survival Outcome using Joint Modeling Approach" by Cheng Zheng and Lei Liu.

Cheng Zheng

Department of Biostatistics, University of Nebraska Medical Center, Omaha, Nebraska, U.S.A.

email: cheng.zheng@unmc.edu

and

Lei Liu

Division of Biostatistics, Washington University in St. Louis, St. Louis, Missouri, U.S.A.

email: lei.liu@wustl.edu

This paper has been submitted for consideration for publication in *Biometrics*

Web Appendix

A. Derivation of NDE and NIE calculation formula

Here we would like to give the detailed formula for NDE and NIE calculation based on the following joint model:

$$
\begin{aligned}
M_i^z(u) &= \beta_0 + \beta_1 z + \boldsymbol{\beta}_2^\top \boldsymbol{X}_i + \beta_3 u + \beta_4 z u + a_{i0} + a_{i1} u + e_i(u) \\
\lambda_i^{z\underline{\boldsymbol{m}}}(t) &= \lambda_0(t)\exp\{\alpha_1 z + \boldsymbol{\alpha}_2^\top \boldsymbol{X}_i + \eta m(t) + \boldsymbol{\gamma}^\top \boldsymbol{a}_i + \boldsymbol{\delta}^\top m(t)\boldsymbol{a}_i\}
\end{aligned}
$$

where $\boldsymbol{a}_i = (a_{i0}, a_{i1})^\top \sim N(0, \boldsymbol{\Sigma}_a)$ and $e_i(u) \sim N(0, \sigma^2)$. For the setting without certain random effects, we can simply set the variance of that random effect to be 0 and for the setting without certain interaction terms for the fixed part, we can set the corresponding coefficient to be 0.

Using piecewise constant approximation for $\underline{\boldsymbol{m}}$, denote $\boldsymbol{m} = (m_1, \cdots, m_K)^\top$, $\boldsymbol{t} = (t_1, \cdots, t_K)^\top$ and $\boldsymbol{a} = (a_0, a_1)^\top$ the NDE can be computed by

$$
\begin{aligned}
&\widehat{\text{NDE}}(t, z, z'; \boldsymbol{x}) \\
= &\int_{\boldsymbol{a}}\int_{\underline{\boldsymbol{m}}} \left(\widehat{S}(t|Z=z', \underline{\boldsymbol{M}}=\underline{\boldsymbol{m}}, \boldsymbol{a}, \boldsymbol{X}=\boldsymbol{x}) - \widehat{S}(t|Z=z, \underline{\boldsymbol{M}}=\underline{\boldsymbol{m}}, \boldsymbol{a}, \boldsymbol{X}=\boldsymbol{x})\right) d\widehat{F}_{\underline{\boldsymbol{M}}|Z,\boldsymbol{a},\boldsymbol{X}}(\underline{\boldsymbol{m}}|z,\boldsymbol{a},\boldsymbol{x}) d\widehat{F}_{\boldsymbol{a}} \\
= &\int_{\boldsymbol{a}}\int_{\underline{\boldsymbol{m}}} \left\{\exp\left(-\int_0^t \widehat{\lambda}^{z'\underline{\boldsymbol{m}}}(u|\boldsymbol{a}, \boldsymbol{X}=\boldsymbol{x})du\right) - \exp\left(-\int_0^t \widehat{\lambda}^{z\underline{\boldsymbol{m}}}(u|\boldsymbol{a}, \boldsymbol{X}=\boldsymbol{x})du\right)\right\} d\widehat{F}_{\underline{\boldsymbol{M}}|Z,\boldsymbol{a},\boldsymbol{X}}(\underline{\boldsymbol{m}}|z,\boldsymbol{a},\boldsymbol{x}) d\widehat{F}_{\boldsymbol{a}} \\
= &\int_{\boldsymbol{a}}\int_{\underline{\boldsymbol{m}}} \left[\exp\left\{-\int_0^t \widehat{\lambda}_0(u)\exp\left(\hat{\alpha}_1 z' + \hat{\boldsymbol{\alpha}}_2^\top \boldsymbol{x} + \hat{\eta} m(u) + \hat{\boldsymbol{\gamma}}^\top \boldsymbol{a} + \hat{\boldsymbol{\delta}} m(u)\boldsymbol{a}\right) du\right\}\right. \\
&\left. - \exp\left\{-\int_0^t \widehat{\lambda}_0(u)\exp\left(\widehat{\alpha}_1 z + \widehat{\boldsymbol{\alpha}}_2^\top \boldsymbol{x} + \widehat{\eta} m(u) + \widehat{\boldsymbol{\gamma}}^\top \boldsymbol{a} + \widehat{\boldsymbol{\delta}}^\top m(u)\boldsymbol{a}\right) du\right\}\right] d\widehat{F}_{\underline{\boldsymbol{M}}|Z,\boldsymbol{a},\boldsymbol{X}}(\underline{\boldsymbol{m}}|z,\boldsymbol{a},\boldsymbol{x}) d\widehat{F}_{\boldsymbol{a}} \\
= &\left\{\exp\left(\widehat{\alpha}_1 z'\right) - \exp\left(\widehat{\alpha}_1 z\right)\right\}\exp\{\widehat{\boldsymbol{\alpha}}_2^\top \boldsymbol{x}\} \\
&\times \int_{\boldsymbol{a}} \exp\left(\widehat{\boldsymbol{\gamma}}^\top \boldsymbol{a}\right)\int_{\underline{\boldsymbol{m}}} \exp\left\{-\int_0^t \widehat{\lambda}_0(u)\exp\left(\widehat{\eta} m(u) + \widehat{\boldsymbol{\delta}} m(u)\boldsymbol{a}\right) du\right\} d\widehat{F}_{\underline{\boldsymbol{M}}|Z,\boldsymbol{a},\boldsymbol{X}}(\underline{\boldsymbol{m}}|z,\boldsymbol{a},\boldsymbol{x}) d\widehat{F}_{\boldsymbol{a}} \\
= &\left\{\exp\left(\widehat{\alpha}_1 z'\right) - \exp\left(\widehat{\alpha}_1 z\right)\right\}\exp\{\widehat{\boldsymbol{\alpha}}_2^\top \boldsymbol{x}\}\int_{R^2} \exp\left(\widehat{\boldsymbol{\gamma}}^\top \boldsymbol{a}\right)\left(2\pi \cdot det(\widehat{\boldsymbol{\Sigma}}_a)\right)^{-1}\exp\left(-\frac{\boldsymbol{a}^\top \widehat{\boldsymbol{\Sigma}}_a^{-1}\boldsymbol{a}}{2}\right) \\
&\times \int_{R^K} \exp\left\{-\sum_{k=1}^K \widehat{\lambda}_0(t_k)\exp\left(\widehat{\eta} m_k + \widehat{\delta} m_k \boldsymbol{a}\right)(t_k - t_{k-1})\right\} \\
&\times (2\pi\widehat{\sigma}^2)^{-K/2}\exp\left\{-\frac{||\boldsymbol{m} - (\widehat{\beta}_0 + a_0) - \widehat{\beta}_1 z - \widehat{\boldsymbol{\beta}}_2^\top \boldsymbol{x} - (\widehat{\beta}_3 + a_1)\boldsymbol{t} - \widehat{\beta}_4 z \boldsymbol{t}||_2^2}{2\widehat{\sigma}^2}\right\} d\boldsymbol{m} d\boldsymbol{a}.
\end{aligned}
$$

Similarly, we can compute NIE by

$$
\begin{aligned}
&\widehat{\mathrm{NIE}}(t, z, z'; x) \\
=\;& \int_{\boldsymbol{a}} \int_{\underline{\boldsymbol{m}}} \widehat{S}(t|Z = z', \underline{\boldsymbol{M}} = \underline{\boldsymbol{m}}, \boldsymbol{a}, \boldsymbol{X} = \boldsymbol{x}) \left(d\widehat{F}_{\underline{\boldsymbol{M}}|Z,\boldsymbol{a},\boldsymbol{X}}(\underline{\boldsymbol{m}}|z', \boldsymbol{a}, \boldsymbol{x}) - d\widehat{F}_{\underline{\boldsymbol{M}}|Z,\boldsymbol{a},\boldsymbol{X}}(\underline{\boldsymbol{m}}|z, \boldsymbol{a}, \boldsymbol{x}) \right) d\widehat{F}\boldsymbol{a} \\
=\;& \int_{\boldsymbol{a}} \int_{\underline{\boldsymbol{m}}} \exp\left\{ -\int_0^t \widehat{\lambda}_0(u) \exp\left(\widehat{\alpha}_1 z' + \widehat{\boldsymbol{\alpha}}_2^\top \boldsymbol{x} + \widehat{\eta} m(u) + \widehat{\boldsymbol{\gamma}} \boldsymbol{a} + \widehat{\boldsymbol{\delta}} m(u) \boldsymbol{a} \right) du \right\} \\
&\times \left(d\widehat{F}_{\underline{\boldsymbol{M}}|Z,\boldsymbol{a},\boldsymbol{X}}(\underline{\boldsymbol{m}}|z', \boldsymbol{a}, \boldsymbol{x}) - d\widehat{F}_{\underline{\boldsymbol{M}}|Z,\boldsymbol{a},\boldsymbol{X}}(\underline{\boldsymbol{m}}|z, \boldsymbol{a}, \boldsymbol{x}) \right) d\widehat{F}\boldsymbol{a} \\
=\;& \exp\left(\widehat{\alpha}_1 z' + \widehat{\alpha}_2^\top \boldsymbol{x} \right) \int_{\boldsymbol{a}} \exp\left(\widehat{\boldsymbol{\gamma}} \boldsymbol{a} \right) \int_{\underline{\boldsymbol{m}}} \exp\left\{ -\sum_{k=1}^K \widehat{\lambda}_0(t_k) \exp\left(\widehat{\eta} m_k + \widehat{\boldsymbol{\delta}} m_k \boldsymbol{a} \right) (t_k - t_{k-1}) \right\} \\
&\times \left(d\widehat{F}_{\underline{\boldsymbol{M}}|Z,\boldsymbol{a},\boldsymbol{X}}(\underline{\boldsymbol{m}}|z', \boldsymbol{a}, \boldsymbol{x}) - d\widehat{F}_{\underline{\boldsymbol{M}}|Z,\boldsymbol{a},\boldsymbol{X}}(\underline{\boldsymbol{m}}|z, \boldsymbol{a}, \boldsymbol{x}) \right) d\widehat{F}\boldsymbol{a} \\
=\;& \exp\left(\widehat{\alpha}_1 z' + \widehat{\alpha}_2^\top \boldsymbol{x} \right) \int_{R^2} \exp\left(\widehat{\boldsymbol{\gamma}} \boldsymbol{a} \right) \times (2\pi \cdot det(\widehat{\boldsymbol{\Sigma}}_a))^{-1} \exp\left(-\frac{\boldsymbol{a}^\top \widehat{\boldsymbol{\Sigma}}_a^{-1} \boldsymbol{a}}{2} \right) \\
&\times \int_{R^K} \exp\{ -\sum_{k=1}^K \hat{\lambda}_0(t_k) \exp\{ m_k \hat{\eta} + a_{i0} m_k \hat{\delta}_1 + a_{i1} m_k \hat{\delta}_2 \} (t_k - t_{k-1}) \} (2\pi \hat{\sigma}^2)^{-K/2} \\
&\times \left[\exp\left\{ -\frac{||\boldsymbol{m} - (\widehat{\beta}_0 + a_0) - \widehat{\beta}_1 z' - \widehat{\boldsymbol{\beta}}_2^\top \boldsymbol{x} - (\widehat{\beta}_3 + a_1)\boldsymbol{t} - \widehat{\beta}_4 z' \boldsymbol{t}||_2^2}{2\widehat{\sigma}^2} \right\} \right. \\
&\left. - \exp\left\{ -\frac{||\boldsymbol{m} - (\widehat{\beta}_0 + a_0) - \widehat{\beta}_1 z - \widehat{\boldsymbol{\beta}}_2^\top \boldsymbol{x} - (\widehat{\beta}_3 + a_1)\boldsymbol{t} - \widehat{\beta}_4 z \boldsymbol{t}||_2^2}{2\widehat{\sigma}^2} \right\} \right] d\boldsymbol{m} d\boldsymbol{a},
\end{aligned}
$$

where $\widehat{\lambda}_0(\cdot), \widehat{\boldsymbol{\alpha}}, \widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\gamma}}, \widehat{\eta}, \widehat{\boldsymbol{\delta}}, \widehat{\boldsymbol{\Sigma}}_a, \widehat{\sigma}s$ are any joint modeling estimators from existing methods.

The integrations can be computed numerically by choosing grid points from R^{K+1} for $(\boldsymbol{m}, \boldsymbol{a})$. Or equivalently, we can use the importance sampling technique. The integration of NDE and the second term of NIE can be computed by first sampling $\boldsymbol{a}$ from $N(0, \widehat{\boldsymbol{\Sigma}}_a)$ and then sampling $\boldsymbol{m}$ from multivariate normal distribution with mean $\widehat{\beta}_0 + \widehat{\beta}_1 z + \widehat{\boldsymbol{\beta}}_2^\top \boldsymbol{x} + \widehat{\beta}_3 \boldsymbol{t} + \widehat{\beta}_4 z \boldsymbol{t} + a_0 + a_1 \boldsymbol{t}$ and variance covariance matrix $\widehat{\sigma}^2 I_K$. Similarly, for the integration of the first term of NIE, we can compute it by first sampling $\boldsymbol{a}$ from $N(0, \widehat{\boldsymbol{\Sigma}}_a)$ and then sampling $\boldsymbol{m}$ from multivariate normal distribution with mean $\widehat{\beta}_0 + \widehat{\beta}_1 z' + \widehat{\boldsymbol{\beta}}_2^\top \boldsymbol{x} + \widehat{\beta}_3 \boldsymbol{t} + \widehat{\beta}_4 z' \boldsymbol{t} + a_0 + a_1 \boldsymbol{t}$ and variance covariance matrix $\widehat{\sigma}^2 I_K$. For the whole population NDE and NIE, we have $\widehat{\mathrm{NDE}}(t, z, z') = n^{-1} \sum_{i=1}^n \widehat{\mathrm{NDE}}(t, z, z'; \boldsymbol{x}_i)$, and $\widehat{\mathrm{NIE}}(t, z, z') = n^{-1} \sum_{i=1}^n \widehat{\mathrm{NIE}}(t, z, z'; \boldsymbol{x}_i)$.

B. Deatils on Simulation Setting

We generate data for observational times set at months 0, 2, 4, 6, 8, 10. We simulate $Z \sim Bernoulli(0.5)$, $M^z(t) = z + 0.2t + a + N(0, 1)$, $\lambda^{z\boldsymbol{m}}(t) = 0.1t \times \exp\{z + 0.5m(t) + a + \delta \times m(t) \times a\}$. We set the sample size as 200. We perform 200 simulations per setting in order to estimate the bias, standard deviation (SD), median estimated standard error (MeSE) and coverage rate for 95% nominal confidence interval (CR). For the first two settings, we assume that our random effect is correctly specified as $a \sim N(0, 1)$. For setting I, we assume no interaction and set $\delta = 0$; for setting II, we have interaction by setting $\delta = -0.5$. For setting III, we want to evaluate the robustness of our method to the misspecification of random effect distribution, so we generate a using log gamma distribution with shape parameter 1.5 and scale parameter 1. For setting IV, we misfit our model without the interaction term while the data are generated with the interaction term.