Sample size considerations for the analysis of time-varying causal effects in stratified micro-randomized trials

Walter Dempsey (co-authors: Peng Liao, Santosh Kumar, and Susan Murphy)
Harvard University

The scientific question

Is there an effect of the treatment on the proximal response? And is there an effect of the treatment if the individual is currently experiencing stress?

Stratified micro-randomized trial

- Participant data:

$$\{O_0, O_1, I_1, A_1, \ldots, O_t, I_t, A_t, \ldots, O_T, I_T, A_T\}$$

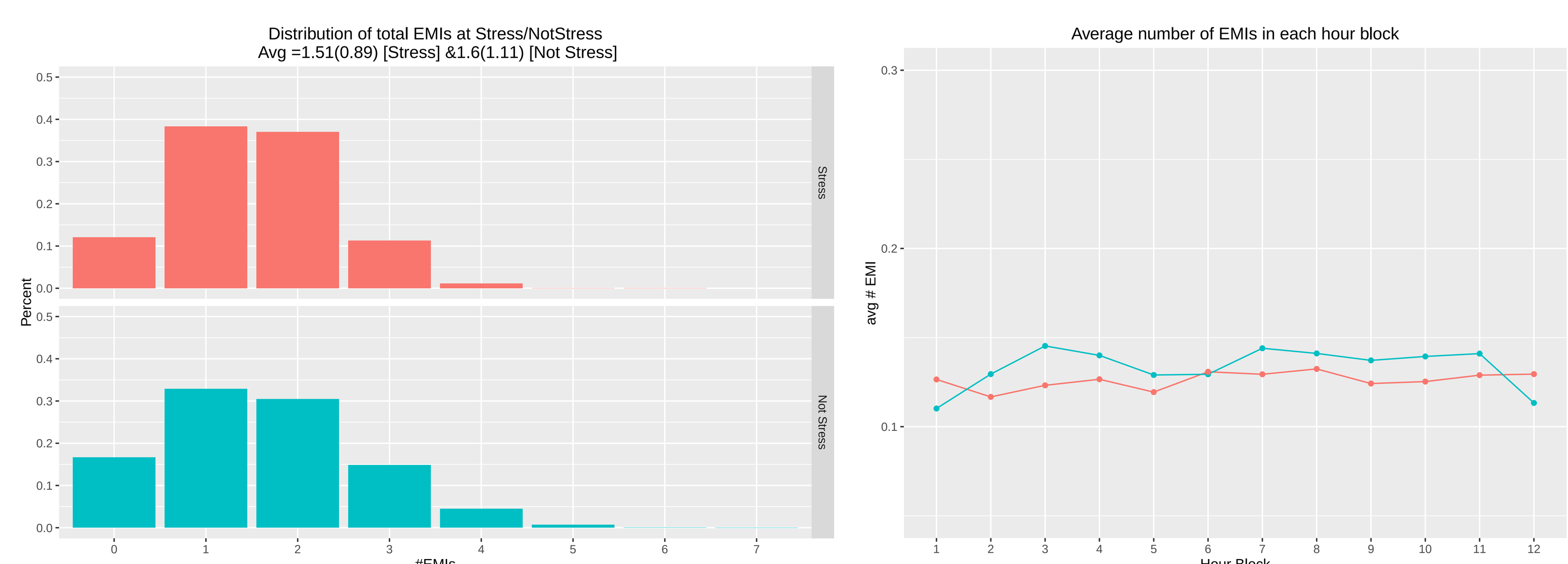
- The *proximal response*, denoted by $Y_{t,\Delta}$, is a known function of the participant's data within a subsequent window of length Δ. For example,

$$Y_{t,\Delta} = \Delta^{-1} \sum_{s=1}^{\Delta} 1[X_{t+s} = \text{"Stressed"}]$$

- We consider binary actions (i.e., $A_t \in \{0,1\}$). The *randomization probability* $\rho_t(\mathcal{H}_t) := \text{pr}(A_t = 1 \mid \mathcal{H}_t)$ is given by

$$\frac{N(x) - \sum_{s=1}^{t-1} [\lambda_s A_s + (1-\lambda_s)\rho_s(\mathcal{H}_s)] \, 1[X_s = x]}{1 + \mathbb{E}\left[\sum_{s=t+1}^{T} 1[X_s = x] \mid \mathcal{H}_t\right]}.$$

Figure 1: Randomization probability pre-lapse



Conditional treatment effect

$$\beta(t;x) = \mathbf{E}\Bigg[\mathbf{E}\Bigg[\prod_{j=t+1}^{t+\Delta-1} \frac{1[A_j = 0]}{p_j(A_j \mid H_j)} Y_{t,\Delta} \mid A_t = 1, H_t\Bigg] - \mathbf{E}\Bigg[\prod_{j=t+1}^{t+\Delta-1} \frac{1[A_j = 0]}{p_j(A_j \mid H_j)} Y_{t,\Delta} \mid A_t = 0, H_t\Bigg] \mid I_t = 1, X_t = x\Bigg]$$

Weighted-centered least squares

$$\mathbb{P}_n \left[\sum_{t=1}^{T} w_t(H_{t+\Delta}) \left(Y_{t,\Delta} - g_t(H_t)'\alpha - (A_t - \tilde{p}_t(X_t)) f_t(X_t)'\beta \right)^2 \right]$$

where $\mathbb{P}_n\{\cdot\}$ is defined as the average over the sample and

$$w_t(H_{t+\Delta}) = \frac{\tilde{p}_t(X_t)^{A_t}(1-\tilde{p}_t(X_t))^{1-A_t} \prod_{s=1}^{\Delta} 1[A_{t+s} = 0]}{\prod_{s=0}^{\Delta} p_{t+s}(A_{t+s} \mid H_{t+s})}.$$

Markovian generative model for the smoking cessation study

- For each episode type (i.e., $x \in \{0,1\}$), the probability that the next episode will be a stress episode: $\bar{W} = (6.7\%, 51.9\%)$
- For each episode type (i.e., $x \in \{0,1\}$), the average episode length: $\bar{Z} = (10.9, 12.0)$
- Inputs are informed by summary statistics from a subset of data (Sarker et al. 2017) collected in an observational, no treatment, smoking cessation study of 61 cigarette smokers (Saleheen et al. 2015).

Markovian generative model for the smoking cessation study

Table 1: $P^{(0)}$: No-Treatment transition Matrix constructed from inputs $(\bar{W}, \bar{Z})$

		Non-stress			Stress		
		Pre-peak	Peak	Post-peak	Pre-peak	Peak	Post-peak
Non-stress	Pre-peak	0.80	0.20	0.00	0.00	0.00	0.00
	Peak	0.00	0.00	1.00	0.00	0.00	0.00
	Post-peak	0.19	0.00	0.80	0.01	0.00	0.00
Stress	Pre-peak	0.00	0.00	0.00	0.82	0.18	0.00
	Peak	0.00	0.00	0.00	0.00	0.00	1.00
	Post-peak	0.09	0.00	0.00	0.09	0.00	0.82

- Given alternative $\beta(t;x)$, we construct the Markov transition matrix under treatment solving

$$\arg\min_{Q \in \mathcal{P}} \max_{x \in \{0,1\}} \left| \Delta^{-1} \sum_{s=1}^{\Delta} \sum_{u \in \{0,1,2\}} \left([Q^s]_{((x,1),(1,u))} - \left[\left(P^{(0)}\right)^s\right]_{((x,1),(1,u))} \right) - \beta(t;x) \right|$$

where $\mathcal{P}$ is the set of episodic transition matrices.

Simulation based sample size calculation

Table 2: Estimated sample size, N, and achieved power.

	Sample size	Power
$\bar{\beta} = 0.030$	50	80.6
$\bar{\beta} = 0.025$	67	80.7
$\bar{\beta} = 0.020$	127	80.6

Example evaluation of sample size calculator

- Markov model allowed us to use few summary statistics from the small noisy dataset.
- This may lead to bias, which is problematic if it results in sample sizes for which the power to detect the desired effect is below the specified power.
- We thus use the small data set to guide our assessment of robustness of the sample size calculator.
- A complex semi-Markovian generative model is proposed through exploratory data analysis.
- Such complex alternatives may be due to noise and not reflect the behavior of trial participants.

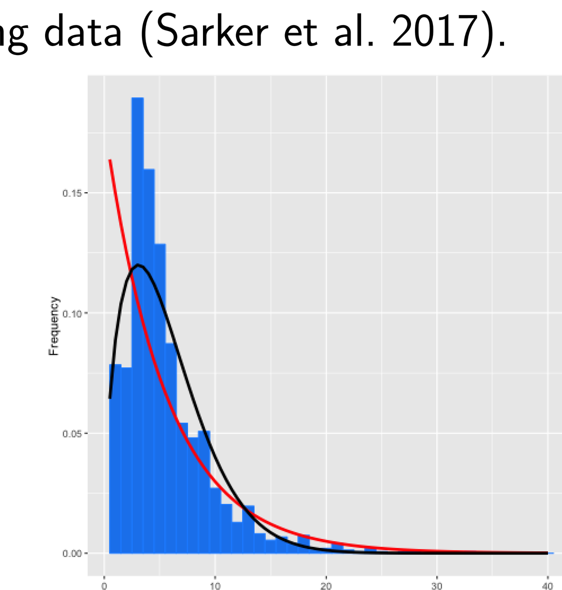
Table 3: Logistic regression parameter estimates. Response is indicator of current episode being a stress episode.

Parameter	Estimate	Std. Error
Intercept	-2.83	0.10
1L Stress Ep.	2.75	0.20
2L Stress Ep.	0.71	0.22

Table 4: Achieved power under semi-Markov model

$\bar{\beta} =$	0.030	0.025	0.020
Achieved power	93.6	88.0	93.4

Figure 2: Pre-peak duration from observational smoking data (Sarker et al. 2017).



References

N. Saleheen, A. Ali, S. Hossain, H. Sarker, S. Chatterjee, B. Marlin, E. Ertin, M. al'Absi, and S. Kumar. puffmarker: A multi-sensor approach for pinpointing the timing of first lapse in smoking cessation. UbiComp, pages 999–1010, 2015.

H. Sarker, K. Hovsepian, S. Chatterjee, I. Nahum-Shani, S. Murphy, B. Spring, E. Ertin, M. al'Absi, M. Nakajima, and S. Kumar. From markers to interventions: The case of just-in-time stress intervention. In J. Regh, S. Murphy, and S. Kumar, editors, *Mobile Health Sensors, Analytic Methods, and Applications*. Springer International Publishing, 2017.

MD2K Center of Excellence for Mobile Sensor Data-to-Knowledge

Advancing biomedical discovery and improving health through mobile sensor big data

Cornell Tech ♦ Georgia Tech ♦ U. Memphis ♦ Northwestern ♦ Ohio State ♦ Open mHealth
Rice ♦ UCLA ♦ UC San Diego ♦ UC San Francisco ♦ UMass Amherst ♦ U. Michigan ♦ WVU

MD2K is supported by the National Institutes of Health Big Data to Knowledge Initiative Grant #1U54EB020404