Some Data Analytics for Developing Just-in-Time Adaptive Interventions in Mobile Health

Susan Murphy

10.24.16

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

The Methodology Center
advancing methods, improving health

MICHIGAN
STATISTICS

INSTITUTE FOR SOCIAL RESEARCH
UNIVERSITY OF MICHIGAN

The Dream!

"Continually Learning Mobile Health Intervention"

- Help you achieve, and maintain, your desired long term healthy behaviors
  – Provide sufficient short term reinforcement to enhance your ability to achieve your long term goal
- The ideal mobile health intervention
  – will engage you when you need it and will not intrude when you don't need it.
  – will adjust to unanticipated life events

Heart Steps

Context provided via data from:

Wearable band → activity and sleep quality;

Smartphone sensors → busyness of calendar, location, weather;

Self-report → stress, user burden

In which contexts should the smartphone provide the user with an activity suggestion?

Data from wearable devices that sense and provide treatments

- On each individual: $O_1, A_1, Y_2, \dots, O_t, A_t, Y_{t+1}, \dots$

- t: Decision point
- O_t: Observations at t^{th} decision point (high dimensional)
- A_t: Action at t^{th} decision point (treatment)
- Y_{t+1}: Proximal outcome (e.g., reward, utility, cost)

Examples

1) Decision Points (Times, t, at which a treatment can be provided.)
    1) Regular intervals in time (e.g. every 10 minutes)
    2) At user demand

<u>Heart Steps</u>: approximately every 2-2.5 hours (activity suggestions)

Examples

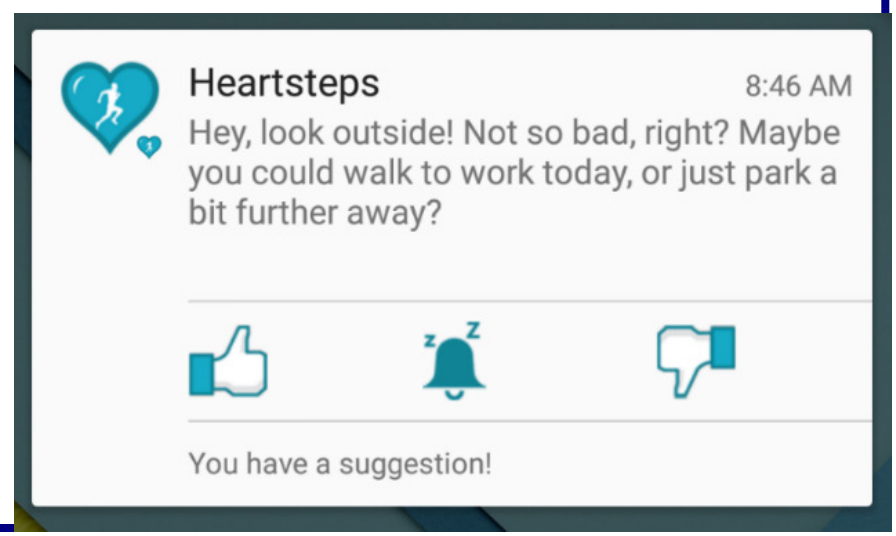
2) Observations O_t
   1) Passively collected (via sensors)
   2) Actively collected (via self-report)

Heart Steps: classifications of activity, location, weather, step count, busyness of calendar, user burden, adherence…….

Examples

3) Actions A_t
   1) Types of treatments that can be provided at a decision point, t
   2) Whether to provide a treatment

<u>HeartSteps</u>: tailored activity suggestion (yes/no)



Availability

- Treatments can only be delivered at a decision point if an individual is *available*.
  - O_t includes I_t=1 if available, I_t=0 if not

- Treatment effects at a decision point are conditional on availability.

- Availability is not the same as adherence!

Examples

4) Proximal Outcome Y_{t+1}

Heart Steps: Step count over next 30 minutes (activity suggestions),

Continually Learning Mobile Health Intervention

1) Trial Designs: Are there effects of the actions on the proximal response? *experimental design*

2) Data Analytics for use with trial data: Do effects vary by the user's internal/external context,? Are there delayed effects of the actions? *causal inference*

3) Learning Algorithms for use with trial data: Construct a "warm-start" treatment policy. *batch Reinforcement Learning*

4) Online Algorithms that personalize and continually update the mHealth Intervention. *online Reinforcement Learning*

Heart Steps
Micro-Randomized Trial

On each of *n* participants and at each of *T* decision points, treatment is repeatedly randomized:

Activity suggestion (T=210 randomizations)

- Provide a suggestion with probability .6; do nothing with probability .4

Conceptual Models

Generally data analysts fit a series of increasingly more complex models:

$$Y_{t+1} \text{ “}\sim\text{” } \alpha_0 + \alpha_1^T Z_t + \beta_0 A_t$$

and then next,

$$Y_{t+1} \text{ “}\sim\text{” } \alpha_0 + \alpha_1^T Z_t + \beta_0 A_t + \beta_1 A_t S_t$$

and so on…

- Y_{t+1} is subsequent activity over next 30 min.
- $A_t = 1$ if activity suggestion and 0 otherwise
- Z_t summaries formed from t and past/present observations
- S_t potential moderator (e.g., current weather is good or not)

Conceptual Models

Generally data analysts fit a series of increasingly more complex models:

$$Y_{t+1} \text{ “}\sim\text{” } \alpha_0 + \alpha_1^T Z_t + \beta_0 A_t$$

and then next,

$$Y_{t+1} \text{ “}\sim\text{” } \alpha_0 + \alpha_1^T Z_t + \beta_0 A_t + \beta_1 A_t S_t$$

and so on…

$\alpha_0 + \alpha_1^T Z_t$ is used to reduce the noise variance in Y_{t+1}
(Z_t is sometimes called a vector of control variables)

Causal Effects

$$Y_{t+1} \text{ "}\sim\text{" } \alpha_0 + \alpha_1^T Z_t + \beta_0 A_t$$

β_0 is the effect, marginal over all observed and all unobserved variables, of the activity suggestion on subsequent activity.

$$Y_{t+1} \text{ "}\sim\text{" } \alpha_0 + \alpha_1^T Z_t + \beta_0 A_t + \beta_1 A_t S_t$$

$\beta_0 + \beta_1$ is the effect when the weather is good (S_t=1), marginal over other observed and all unobserved variables, of the activity suggestion on subsequent activity.

Data Scientist's Goal

- Challenges:
  - Time-varying treatment (A_t, $t=1,\ldots T$)
  - "independent" variables: Z_t, S_t, I_t that may be affected by prior treatment
- Develop data analytic methods that are consistent with the scientific understanding of the meaning of the β regression coefficients
- Robustly facilitate noise reduction via use of controls, Z_t

For the Statistician!

Treatment Effect Model:

$$E[\ E[Y_{t+1}|A_t = 1, I_t = 1, H_t] - E[Y_{t+1}|A_t = 0, I_t = 1, H_t]|\ I_t = 1, S_t] = {S_t}^T \beta$$

H_t is all participant data available up to and at time t

S_t is a vector of data summaries and time, t, $(S_t \subseteq H_t)$

I_t indicator of availability
We aim to conduct inference about β!

"Centered and Weighted Least Squares Estimation"

- Simple method for complex data!
- Enables unbiased inference for a causal, marginal, treatment effect (the β's)
- Inference for treatment effect is not biased by how we use the controls, Z_t, to reduce the noise variance in Y_{t+1}

https://arxiv.org/abs/1601.00237

Application of the “Centered and Weighted Least Squares Estimation” method in first analyses of HeartSteps

Heart Steps Pilot Study

On each of $n=37$ participants:

a) Activity suggestion, A_t

- **Provide a suggestion with probability .6**
  - a tailored sedentary-reducing activity suggestion (probability=.3)
  - a tailored walking activity suggestion (probability=.3)
- **Do nothing (probability=.4)**

- 5 times per day * 42 days= 210 decision points

Conceptual Models

$$Y_{t+1} \text{ “}\sim\text{” } \alpha_0 + \alpha_1 Z_t + \beta_0 A_t$$

$$Y_{t+1} \text{ “}\sim\text{” } \alpha_0 + \alpha_1 Z_t + \beta_0 A_t + \beta_1 A_t d_t$$

- $t=1,\ldots T=210$
- Y_{t+1} = log-transformed step count in the 30 minutes *after* the t^{th} decision point,
- $A_t = 1$ if an activity suggestion is delivered at the t^{th} decision point; $A_t = 0$, otherwise,
- Z_t = log-transformed step count in the 30 minutes *prior* to the t^{th} decision point,
- d_t =days in study; takes values in (0,1,….,41)

Pilot Study Analysis

$$Y_{t+1} \text{ "}\sim\text{" } \alpha_0 + \alpha_1 Z_t + \beta_0 A_t \text{, and}$$

$$Y_{t+1} \text{ "}\sim\text{" } \alpha_0 + \alpha_1 Z_t + \beta_0 A_t + \beta_1 A_t d_t$$

Causal Effect Term	Estimate	95% CI	p-value
$\beta_0 A_t$ *(effect of an activity suggestion)*	$\hat{\beta}_0$=.13	(-0.01, 0.27)	.06
$\beta_0 A_t + \beta_1 A_t d_t$ *(time trend in effect of an activity suggestion)*	$\hat{\beta}_0$ = .51	(.20, .81)	<.01
	$\hat{\beta}_1$ = -.02	(-.03, -.01)	<.01

Heart Steps Pilot Study

On each of *n=37* participants:

a) Activity suggestion

- Provide a suggestion with probability .6
  - **a tailored walking activity suggestion (probability=.3)**
  - **a tailored sedentary-reducing activity suggestion (probability=.3)**
- **Do nothing (probability=.4)**

- 5 times per day * 42 days= 210 decision points

Pilot Study Analysis

$$Y_{t+1} \text{ “}\sim\text{” } \alpha_0 + \alpha_1 Z_t + \beta_0 A_{1t} + \beta_1 A_{2t}$$

- $A_{1t} = 1$ if walking activity suggestion is delivered at the t^{th} decision point; $A_{1t} = 0$, otherwise,
- $A_{2t} = 1$ if sedentary-reducing activity suggestion is delivered at the t^{th} decision point; $A_{2t} = 0$, otherwise,

Causal Effect	Estimate	95% CI	p-value
$\beta_0 A_{1t} + \beta_1 A_{2t}$	$\hat{\beta}_0 = .21$ $\hat{\beta}_1 > 0$	(.04, .39) ns	.02 ns

Initial Conclusions

- The data indicates that there is a causal effect of the activity suggestion on step count in the succeeding 30 minutes.
    - This effect is primarily due to the walking activity suggestions.
    - This effect deteriorates with time
    - The walking activity suggestion initially increases step count over succeeding 30 minutes by ≈ 271 steps but by day 20 this increase is only ≈ 65 steps.

Discussion

Problematic Analyses

- GLM & GEE analyses
- Random effects models & analyses
- Machine Learning Generalizations:
  - Partially linear, single index models & analysis
  - Varying coefficient models & analysis

--These analyses do not take advantage of the micro-randomization. Can accidentally eliminate the advantages of randomization for estimating causal effects--

Discussion

- Randomization enhances:
  – Causal inference based on minimal structural assumptions

- Challenge:
  – How to include random effects which reflect scientific understanding ("person-specific" effects) yet not destroy causal inference?

It takes a Team!