Personalization with Latent Confounders

Leo Guelman

RBC Royal Bank

Figure: Ronald Fisher (1890-1962)

I don;t like this A/B Testing is likely NOT 'gold standard' for Personalized Decision-Making.

Personalization with Unobserved Confounders

- The value of Personalization is founded on the premise that individuals have heterogenous response to actions.

Personalization with Unobserved Confounders

- The value of Personalization is founded on the premise that individuals have heterogenous response to actions.
- Personalization algorithms aim to improve decision-making by identifying and exploiting this heterogeneity.

Personalization with Unobserved Confounders

- The value of Personalization is founded on the premise that individuals have heterogenous response to actions.
- Personalization algorithms aim to improve decision-making by identifying and exploiting this heterogeneity.
- However, latent confounders (i.e., unobserved variables affecting both the actions and the outcome variables) pose a unique challenge to personalization.

Personalization with Unobserved Confounders

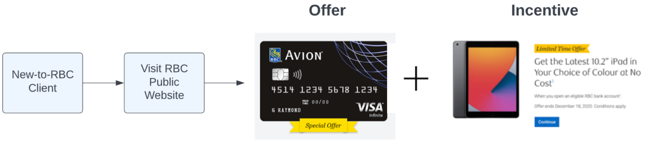
- The value of Personalization is founded on the premise that individuals have heterogenous response to actions.
- Personalization algorithms aim to improve decision-making by identifying and exploiting this heterogeneity.
- However, latent confounders (i.e., unobserved variables affecting both the actions and the outcome variables) pose a unique challenge to personalization.
- In contrast to the general notion that Randomized Controlled Experiments (a.k.a. A/B Tests) are 'gold standard', in this setting they might actually result in loss of information.

Personalization with Unobserved Confounders

- The value of Personalization is founded on the premise that individuals have heterogenous response to actions.
- Personalization algorithms aim to improve decision-making by identifying and exploiting this heterogeneity.
- However, latent confounders (i.e., unobserved variables affecting both the actions and the outcome variables) pose a unique challenge to personalization.
- In contrast to the general notion that Randomized Controlled Experiments (a.k.a. A/B Tests) are 'gold standard', in this setting they might actually result in loss of information.
- Counterfactual-based decision-making can address these problems and lead to a coherent fusion of observational and experimental data.

The Business Setting

- **Business objective:** Sell a credit card to new-to-RBC clients.
- **Past campaign:** All new-to-RBC clients who visited the RBC public site, get a credit card offer + iPad incentive.



- **The goal is to personalize the incentive:** Identify which new-to-RBC clients should receive an iPad incentive in the future to maximize the expected profitability of the campaign.

Data Generating Process

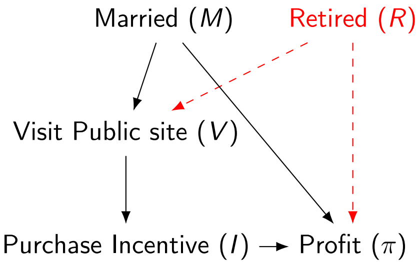


Figure: Past campaign ‘true’ causal DAG.

Data Generating Process

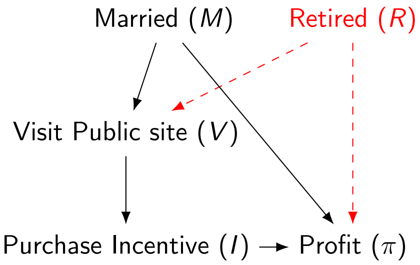


Figure: Past campaign 'true' causal DAG.

$$V := M \oplus \mathrm{R}$$
$$I := V$$

	$R=0$		$R=1$	
	$M=1$	$M=0$	$M=1$	$M=0$
$I=1$	0.25	0.50	0.45	0.05
$I=0$	0.50	0.10	0.05	0.30

Table: $E[\pi|M,R,I]$. Highlighted cells reflect (new-to-RBC) client's 'natural' choice to visit the Public site or not.

Approach 1: Empirical Decision Criterion (EDC)

$$\text{EDC} \rightarrow \underset{I \in 0,1}{\operatorname{argmax}} \; E[\pi | I, M]$$

$$E[\pi | I = 1, M = 1] = 0.25$$
$$E[\pi | I = 0, M = 1] = 0.05$$
$$E[\pi | I = 1, M = 0] = 0.05$$
$$E[\pi | I = 0, M = 0] = 0.10$$

	$R = 0$		$R = 1$	
	$M = 1$	$M = 0$	$M = 1$	$M = 0$
$I = 1$	0.25	0.50	0.45	0.05
$I = 0$	0.50	0.10	0.05	0.30

Table: $E[\pi | M, R, I]$.

Approach 1: Empirical Decision Criterion (EDC)

$$\text{EDC} \rightarrow \underset{I \in 0,1}{\operatorname{argmax}} \; E[\pi | I, M]$$

$$E[\pi|I=1, M=1] = 0.25$$
$$E[\pi|I=0, M=1] = 0.05$$
$$E[\pi|I=1, M=0] = 0.05$$
$$E[\pi|I=0, M=0] = 0.10$$

	$R=0$		$R=1$	
	$M=1$	$M=0$	$M=1$	$M=0$
$I=1$	0.25	0.50	0.45	0.05
$I=0$	0.50	0.10	0.05	0.30

Table: $E[\pi|M, R, I]$.

Decision Rule:

- If Visit Site $\wedge$ Married $\rightarrow$ Purchase Incentive $\rightarrow$ $E[\pi] = \mathbf{0.25}$
- If Visit Site $\wedge$ Not Married $\rightarrow$ No Purchase Incentive $\rightarrow$ $E[\pi] = \mathbf{0.05}$

Expected profit $= \boxed{0.15} = (0.25+0.05)/2$.

Approach 2: Post-Visit Randomization (PVR)

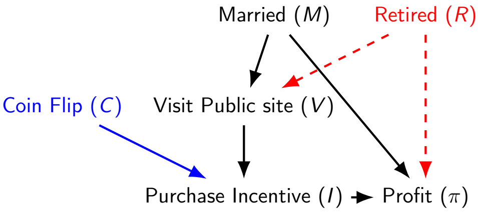


Figure: Causal DAG with post-visit randomization.

$$V := M \oplus \mathrm{R}$$
$$I := V \wedge C$$

Approach 2: Post-Visit Randomization (PVR) - cont'd

$$\text{PVR} \rightarrow \underset{I \in 0,1}{\operatorname{argmax}} \; E[\pi | do(I), M, V = 1]$$

$$E[\pi | do(I = 1), M = 1, V = 1] = 0.25$$
$$E[\pi | do(I = 0), M = 1, V = 1] = 0.50$$
$$E[\pi | do(I = 1), M = 0, V = 1] = 0.05$$
$$E[\pi | do(I = 0), M = 0, V = 1] = 0.30$$

	$R = 0$		$R = 1$	
	$M = 1$	$M = 0$	$M = 1$	$M = 0$
$I = 1$	0.25	0.50	0.45	0.05
$I = 0$	0.50	0.10	0.05	0.30

Table: $E[\pi | M, R, I]$.

Approach 2: Post-Visit Randomization (PVR) - cont'd

$$\text{PVR} \rightarrow \underset{I \in 0,1}{\text{argmax}}\ E[\pi | do(I), M, V = 1]$$

$$E[\pi | do(I = 1), M = 1, V = 1] = 0.25$$
$$E[\pi | do(I = 0), M = 1, V = 1] = 0.50$$
$$E[\pi | do(I = 1), M = 0, V = 1] = 0.05$$
$$E[\pi | do(I = 0), M = 0, V = 1] = 0.30$$

	$R = 0$		$R = 1$	
	$M = 1$	$M = 0$	$M = 1$	$M = 0$
$I = 1$	0.25	0.50	0.45	0.05
$I = 0$	0.50	0.10	0.05	0.30

Table: $E[\pi | M, R, I]$.

Decision Rule:

- If Visit Site $\wedge$ Married $\rightarrow$ No Purchase Incentive $\rightarrow$ $E[\pi] = \mathbf{0.50}$
- If Visit Site $\wedge$ Not Married $\rightarrow$ No Purchase Incentive $\rightarrow$ $E[\pi] = \mathbf{0.30}$

Expected profit $= \boxed{0.40} = (0.50+0.30)/2$.

Approach 3: A/B Test on All New-to-RBC Clients

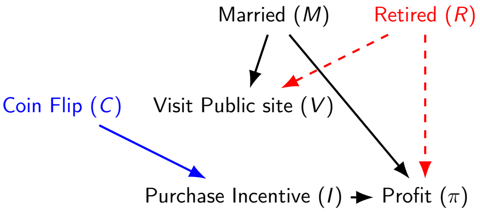


Figure: Causal DAG with A/B Test on All New-to-Bank Clients.

$$V := M \oplus \mathrm{R}$$
$$I := C$$

Approach 3: A/B Test on All New-to-RBC Clients - cont'd

$$\text{ABT} \rightarrow \underset{I \in 0,1}{\operatorname{argmax}} \; E[\pi | do(I), M]$$

$$E[\pi | do(I=1), M=1] = 0.350 = (0.25 + 0.45)/2$$
$$E[\pi | do(I=0), M=1] = 0.275 = (0.50 + 0.05)/2$$
$$E[\pi | do(I=1), M=0] = 0.275 = (0.50 + 0.05)/2$$
$$E[\pi | do(I=0), M=0] = 0.200 = (0.10 + 0.30)/2$$

	$R=0$		$R=1$	
	$M=1$	$M=0$	$M=1$	$M=0$
$I=1$	0.25	0.50	0.45	0.05
$I=0$	0.50	0.10	0.05	0.30

Table: $E[\pi | M, R, I]$.

Approach 3: A/B Test on All New-to-RBC Clients - cont'd

$$\text{ABT} \rightarrow \underset{I \in 0,1}{\operatorname{argmax}} \; E[\pi | do(I), M]$$

$$E[\pi|do(I=1), M=1] = 0.350 = (0.25+0.45)/2$$
$$E[\pi|do(I=0), M=1] = 0.275 = (0.50+0.05)/2$$
$$E[\pi|do(I=1), M=0] = 0.275 = (0.50+0.05)/2$$
$$E[\pi|do(I=0), M=0] = 0.200 = (0.10+0.30)/2$$

	$R=0$		$R=1$	
	$M=1$	$M=0$	$M=1$	$M=0$
$I=1$	0.25	0.50	0.45	0.05
$I=0$	0.50	0.10	0.05	0.30

Table: $E[\pi|M, R, I]$.

Decision Rule:

- If Married → Purchase Incentive → $E[\pi] = \mathbf{0.35}$
- If Not Married → Purchase Incentive → $E[\pi] = \mathbf{0.275}$

Expected profit = $\boxed{0.315}$ = (0.35+0.275)/2.

Approach 4: Regret Decision Criterion (RDC)

$$\text{RDC} \rightarrow \underset{a' \in \ 0,1}{\operatorname{argmax}} \ E[\pi_{a'} | I = a, M]$$

Approach 4: Regret Decision Criterion (RDC)

$$\text{RDC} \rightarrow \underset{a' \in 0,1}{\operatorname{argmax}} \; E[\pi_{a'} | I = a, M]$$

$$\begin{aligned} P(\pi_{a'}, M) &= P(\pi_{a'}, M, a') + P(\pi_{a'}, M, a) \\ &= P(\pi_{a'} | M, a')P(M, a') + P(\pi_{a'} | M, a)P(M, a) \end{aligned}$$

$$\begin{aligned} P(\pi_{a'} | M) &= P(\pi_{a'} | M, a')P(a' | M) + P(\pi_{a'} | M, a)P(a|M) \\ &= P(\pi | M, a')P(a' | M) + P(\pi_{a'} | M, a)P(a|M) \text{ (Consistency)} \end{aligned}$$

$$\begin{aligned} P(\pi_{a'} | M, a) &= \frac{1}{P(a|M)} \left[P(\pi_{a'} | M) - P(\pi | M, a')P(a' | M) \right] \\ &= \boxed{\frac{1}{P(a|M)} \left[P\Big(\pi | M, do(a')\Big) - P(\pi | M, a')P(a' | M) \right]} \end{aligned}$$

Approach 4: Regret Decision Criterion (RDC) - cont'd

I should have expectation instead of probability dist.

$\boxed{P(\pi_{I=\mathbf{1}}|M=1, I=0)} =$

$$\frac{1}{P(I=0|M=1)}\Big[P\Big(\pi|M=1, do(I=1)\Big) - P(\pi|M=1, I=1)P(I=1|M=1)\Big].$$

$$= \frac{1}{1/2}(0.350 - 0.25 \times \frac{1}{1/2}) = \mathbf{0.45}$$

	$R=0$		$R=1$	
	$M=1$	$M=0$	$M=1$	$M=0$
$I=1$	0.25	0.50	0.45	0.05
$I=0$	0.50	0.10	0.05	0.30

Table: $E[\pi|M, R, I]$.

$\boxed{P(\pi_{I=\mathbf{1}}|M=0, I=0)} = \mathbf{0.50}$

$\boxed{P(\pi_{I=\mathbf{0}}|M=1, I=1)} = \mathbf{0.50}$

$\boxed{P(\pi_{I=\mathbf{0}}|M=0, I=1)} = \mathbf{0.30}$

Decision Rule:

- If Visit Site $\wedge$ Married $\rightarrow$ No Purchase Incentive $\rightarrow$ $E[\pi] = \mathbf{0.50}$
- If Visit Site $\wedge$ Not Married $\rightarrow$ No Purchase Incentive $\rightarrow$ $E[\pi] = \mathbf{0.30}$
- If Not Visit Site $\wedge$ Married $\rightarrow$ Purchase Incentive $\rightarrow$ $E[\pi] = \mathbf{0.45}$
- If Not Visit Site $\wedge$ Not Married $\rightarrow$ Purchase Incentive $\rightarrow$ $E[\pi] = \mathbf{0.50}$

Summary of Methods

Criterion	Decision Rule	$E[\pi]$
EDC	• If Visit Site $\wedge$ Married $\rightarrow$ **Purchase Incentive** • If Visit Site $\wedge$ Not Married $\rightarrow$ **No Purchase Incentive**	.1500
PVR	**Never Purchase Incentive**	.4000
ABT	**Always Purchase Incentive**	.3150
RDC	• If Visit Site $\wedge$ Married $\rightarrow$ **No Purchase Incentive** • If Visit Site $\wedge$ Not Married $\rightarrow$ **No Purchase Incentive** • If Not Visit Site $\wedge$ Married $\rightarrow$ **Purchase Incentive** • If Not Visit Site $\wedge$ Not Married $\rightarrow$ **Purchase Incentive**	**.4375**
Oracle		**.4375**

Key Takeaways

- A/B testing is not always the 'Gold Standard' for learning causal effects
- If (i) the goal is to learn optimal personalized actions, and (ii) we have unobserved confounders (very likely!), and (iii) these confounders interact with the action: then Counterfactual-based decision-making may outperform Causal-decision making.
  But even if the unobserved feature is not technically a confounder, but there is interaction between the unobserved variable and the treatment, Counterfactual-based decision-making may also outperform Causal-decision making.
- The expression derived from RDC works only in the binary treatment case. RDC-type randomization (Forney et al., 2017) was proposed to estimate the counterfactual expressions empirically from an arbitrary number of treatments.

References

- Elias Bareinboim, Andrew Forney, and Judea Pearl. 2015. Bandits with unobserved confounders: a causal approach. In Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1 (NIPS'15).
  - Implementation: https://github.com/leoguelman/mabuc
- Forney, A., Pearl, J. ; Bareinboim, E.. (2017). Counterfactual Data-Fusion for Online Reinforcement Learners. Proceedings of the 34th International Conference on Machine Learning, in Proceedings of Machine Learning Research 70:1156-1164

Other points

- Review Pearl P. 94+. What is the difference between individual level counterfactuals and ETT with conditioning on intent (and observed features)?
  - ETT still represents an average effect conditional on X and intent, it is NOT represent the effect for a specific individual. The latter requires a parametric model (and of course the causal graph). If for example, we have latent counfounders which interact with treatment, and the functional form of that interaction is specific to a given individual, then the Personalization approach I'm describing in this deck cannot recover the individual treatment effect. Again, I can only speak about the average treatment effect on the treated conditional on covariates and intent (see also paper "on the distinction between CATE and ITE".
  - For ETT, I either (i) need the causal graph, so I can use the backdoor criterion for adjustment, or (ii) in the binary treatment case, if I have access to both observational and experimental data, I can get away without the graph.
- What happens with more than 1 latent confounder? I believe I have this in my notes somewhere
  - For example, if I have two unobserved confounders as in the MABUC paper, in that case RDC still gives me the oracle payout just because of the configuration of the payout table. In my notes (see Notes for white paper - Page A - around page 29, I have an example of a configuration of payout where I cannot get the oracle by conditioning on intent.