

# Experimentální efekty, manipulace, kontrola a randomizace v experimentech

4.10. 2016

# Co v experimentech „kontrolujeme“

- Pokud o experimentech uvažujeme jako o vztazích mezi proměnnými, pak elementy experimentu jsou zejména:
- **Ošetřující/nezávislá proměnná** (cause, treatment variable, v observačním výzkumu independent variable) T
- **Závislá proměnná** (dependent variable) Y
- **Experimentální manipulace** (manipulated variable) M
- **Pozorovatelné intervenující faktory** (observable confounding factors) Z
- **Nepozorovatelné intervenující faktory** (unobservable confounding factors) V

# Y- závislá proměnná

- V experimentu, který zkoumá vliv úrovně informace na ochotu volby kandidáta, je Y volba kandidáta (např. s hodnotami ano-ne, nebo vyjádřená na stupnici 0-100).
- Čím je závislá proměnná ovlivněna?
- Naším treatmentem, ale určitě i Z a V.
- My v experimentech obvykle primárně měříme efekt T, protože efekty Z a V považujeme v obou skupinách díky náhodnému přiřazení za konstantní.

# M, T, Z a V

- V ideálním experimentu  $M=T$

a naopak

Z a V naopak nemají vliv na T,

jenže:

# Příklad

v experimentu, zkoumajícím vliv informace na rozhodnutí volit, kde manipulací je například audio s informacemi o kandidátovi, ovlivňuje to, zda skutečně  $M=T$ , řada Z a V (např. znalost jazyka subjektu nebo jeho sluch).

V řadě experimentů to tedy (bohužel) není tak, že by V a Z ovlivňovaly jen Y, ovlivňují i T.

# Někdy M není T záměrně

- Př. Spezio et al. – dívali se skrz fMRI, jak reagují mozky subjektů na fotografie kandidátů
- M (zásah výzkumníka) nespočíval v tom, že by nějak manipuloval fotografie, ale ve zprostředkování/analýze toho, co se děje v mozku, skrze fMRI.
- Podle některých definic by **nešlo o experiment (není manipulace nezávislou proměnnou)**.

# Kauzální efekt

- zjistíme nejnáze, pokud změříme u **každého subjektu** potenciální výsledky **každého z treatmentů**, tj. z možných stavů světa, v nichž se může nacházet.
- jenže to bohužel **nejde**, nemůžeme všechny treatmenty změřit v jednom čase
- Subjekty experimentů jsou např. buďto informované nebo neinformované, ale ne **zároveň** informované a neinformované

# Bohužel obvykle nemůžeme vyřešit pomocí pretestu a postestu

- Nemůžeme se subjektu zeptat, koho volí, pak mu ukázat např. negativní video o tom kandidátovi a znovu se ho ptát, koho volí
- Postoj v čase  $t$  (který jsme **vyvolali**) je totiž proměnnou, která **ovlivní** postoj v čase  $t_1$  (např. „motivované uvažování“)



# *Between subject a Within subject* designy

- **Between subject designy-** subjekty se v nich nalézají jen v jednom stavu světa (část je např. informovaných a část neinformovaných)
- **Within subject designy-** jeden subjekt zažívá různé treatmenty-stavy světa (např. vidí politická videa, ve kterých jsou kandidáti civilní i vulgární).

# Problém „předučení“ ve WSD

- Někdy (jako v příkladu, připouštějícím motivované uvažování) je problém, jindy příležitost
- Příklad: Hra „Ultimátum“ jako WSD experiment, v níž si subjekty postupně mění role (opakovaná situace) a její srovnání s chováním subjektů, kteří si role nemění, případně s jednorázovou hrou.

# Průměrný kauzální účinek

- S tím, že vystavíme každý subjekt všem treatmentům, jsou tedy často velké problémy
- Pomáháme si tím, že měříme průměrný kauzální účinek v každé skupině.
- **Průměrný** proto, že efekt treatmentu **není** u každého subjektu **stejný**.

# Příklad (I.)

- V kontrolní skupině představíme jednoho kandidáta (fiktivního), jeho program (středový), který je umírněný, fotografii, subjekty hodnotí na škále 0-100, zda by ho volili.
- Do svého rozhodnutí každý z nich promítá jeden nebo více faktorů (vzhled, program, politickou historii, náhodu, omyl). Hodnocení jsou různá, z nich vypočítáme průměrné hodnocení.

# Příklad 1 (II.)

- V experimentální skupině představíme kandidáta se stejnými charakteristikami jako v kontrolní, manipulace spočívá v tom, že subjektům řekneme, že má nemanželské dítě s Romkou.
- Někteří ho možná budou hodnotit hůře, jiní možná lépe než bez této informace, opět spočítáme průměrné hodnocení
- **Průměrný kauzální efekt (ATE) je rozdíl průměrů** mezi oběma skupinami s různými treatmenty.
- ATE se snaží odhadnout velikost kauzálního efektu ošetření na náhodně vybraného jedince v populaci.

# K čemu nám průměrný kauzální efekt je?

- Nabízelo by se, že prostě řekneme, že to je i efekt, který má **naš treatment na náhodně vybraného jedince v reálném světě**, ale to naráží na celou řadu problémů (viz také přednáška o validitě).

# Příklad

- Chceme zjistit, zda se nějak liší průměrný postoj k palčivému politickému tématu u skupiny (C), která ho (postoj) má jen deklarovat a skupiny (T), která ho má deklarovat a zároveň slovně obhájit. Zjistíme, že průměrný postoj k tématu mezi skupinami se neliší, průměrný efekt T (nutnosti ho obhájit) je tedy 0.

**Jenže?**

V C nalezneme víceméně normální rozložení, zatímco v T dvě polární varianty a možná středovou, protože vše ostatní je mnohem obtížnější slovně vysvětlit.

Treatment tedy efekt měl (polarizaci).

Zkoumat proto musíme nejen **průměrný účinek**, ale často i **heterogenitu/variabilitu** odpovědí na treatmenty.



# Faktory, hovořící pro zkoumání heterogenity odpovědi na treatment

- Souvisí i s tím, že samotný treatment je vnímán heterogenně:
- Nevyhovění subjektů designu
- Pretreatment
- Samovýběr IRL
- Ekologické efekty

# Nevyhovění designu

- Máme dvě skupiny, (T)reatment a C(ontrol) a dva stavy světa, ve kterých se mohou nacházet subjekty (T)reated a (C)ontrolled. Designu vyhovují jen TT a CC, problém (což nechceme) je s TC a CT.
- Příklad: Vybraným a náhodně přiřazeným lidem k treatmentu rozesíláme mobilizační pohlednici (to je tzv. Instrumentální proměnná) kandidáta, jiným náhodně přiřazeným k C ne. Problémem je, že někteří T ji nedostanou, někteří ji vyhodí a někteří C se o ní naopak mohou dozvědět. Tím máme v C skupině i CT agenty a v T skupině i TC agenty.

# Pretreatment

- Experimenty jsou cenné v tom, že v nich obvykle řešíme palčivé otázky pro aktéry v reálném světě.
- Mobilizujeme experimentální skupiny pomocí uměle vytvořených materiálů, různě rámuje témata, jsme negativní atd.
- Problém je, že nevíme, jestli na začátku všichni v T i C jsou „neošetření“ z reálného světa a zda jsou neošetření stejně. Mohli vidět podobnou negativní reklamu, setkat se s podobným rámováním (případně i jeho dekonstrukcí). To, co naměříme, pak není **efekt naší manipulace**, ale **marginální efekt přídavné manipulace**.

# „Panenské“ vs. „Cynické“ populace

- V „panenských“ populacích obvykle naměříme mnohem větší efekty, než v cynických populacích, majících častou zkušenost s politikou/testováním, které vůči treatmentům mohou být otupělejší.
- Příklad: jedna populace neviděla ve volbách žádnou negativní reklamu, efekt demobilizace bude 5%, druhá viděla průměrně pět negativních reklam, demobilizace bude už jen 2%, protože někteří už demobilizovaní jsou.
- Problém s „profesionálními“ vzorky (četnost provádění experimentů ohrožuje jejich spolehlivost).

# IRL samovýběr

V reálném světě bohužel **nefunguje náhodné přiřazení.**

Velmi často se T a C liší v pozorovatelných a nepozorovatelných proměnných. Lidé/skupiny, kteří náš klip vidí a kteří ho nevidí, nejsou v řadě charakteristik stejní.

Což zároveň znamená, že v našich experimentech jsou T občas lidé, kteří T IRL nikdy nezažijí a naopak v C lidé, kteří T zažijí IRL se skoro 100% jistotou.

# Možné řešení

- Kromě T a C mít ještě i větev experimentu, kde si subjekty samy vybírají, zda dostanou treatment nebo ne podle toho, zda by ho dostaly v reálném světě. Srovnávají se pak výsledky obou větví.
- Řešíme tak problém související s tím, že 1. IRL nefunguje náhodné přiřazení a 2. ti, kdo jsou tam treatmentu vystaveni, reagují jinak, než by reagovali ti, kdo mu IRL vystaveni nejsou.

# Kontextuální/Kulturní faktory

- Důležité pro posouzení CC, CT, TC a TT je okolí, ve kterém se subjekty pohybují, jejich status může záviset nejen na naší manipulaci, ale i na interkaci s okolím, ve kterém jsou různé typy „agentů“ a různá „promořenost“ treatmentem.
- Příklad: náhodně přiřazujeme lidi s nízkým a vysokým potenciálem na práci do demokraticky a autokraticky řízené strany. Chceme měřit jejich zájem o politiku, motivaci politicky pracovat atd. Problémem je, že v demokratických a autokratických stranám mohou být defaultně různé koncentrace lidí s vysokým a nízkým potenciálem. Výsledný potenciál našich obou skupin nesouvisí jen s interakcí oni vs. typ strany, ale i oni vs. typy lidí ve straně. Klíčovou proměnnou tedy nemusí být jen typ (řízení) strany, ale i dynamika/efekty lidí, kteří jsou tam jako straníci.

# Kontrola pozorovatelných a nepozorovatelných intervenujících proměnných

- Klíčovým mechanismem kontroly Z a V je náhodné přiřazení
- Někdy ale ovlivňujeme Z přímo:
- buďto je držíme konstantní
- nebo náhodně přiřazujeme i je



# Příklad

- Jazyk sdělení (Z):
- Buďto kontrolujeme, aby se všichni dozvěděli informaci/experimentální instrukci stejným způsobem
- Nebo způsobů zvolíme několik a uděláme ze sdělení informace různé treatmenty.
- Příklad, jak pak může vypadat experiment:
- Kontrolní skupina 1A: slyší audio, 1B: vidí video
- Experimentální skupina (má navíc nějakou informaci): 2A: slyší audio, 2(B): vidí video.
- Efekty treatmentu nemusíme srovnávat jen mezi A a B, ale i v rámci A a B, pak experimentálně zkoumáme vliv Z na T (a zprostředkovaně i Y).

# Náhodné přiřazení jako tzv. ideální instrumentální proměnná (IIV)

- Náhodné přiřazení je ideální instrumentální proměnnou, která eliminuje vliv  $Z$  a  $V$  na treatment, pokud:
- Všechny subjekty přiřadíme všechny ve stejném čase
- Přiřazení k manipulacím nezávisí na přiřazení k jiným manipulacím
- Jsme schopni zajistit plné vyhovění subjektu manipulaci v duchu našeho záměru ( $M=T$ )
- Jsme schopni pozorovat všechny volby subjektů

# Jak často je náhodné přiřazení dokonalá IV?

- V laboratorních experimentech občas
- V polních málokdy

Důvody: Všechny subjekty nejsou rekrutovány ve stejný čas (první subjekty mohou o experimentu mimo laboratoř vyprávět ještě netestovaným subjektům), experimentální úmrtnost, nejsme schopni pozorovat všechny volby subjektů...

# Můžeme kontrolovat i nepozorovatelné intervenující proměnné?

- Můžeme pomocí random assignmentu, někdy to zkusíme i jinak
- Příklady:
  - měříme, jak rychle subjekty reagují (normálně IRL nepozorovatelná proměnná)
  - snažíme se nivelizovat i nepozorovatelné proměnné (příklad: motivace subjektu- pomocí odměn z rozhodování)

# Příklad

- Zagrapan, Hrbková, Chytilék (výzkum 2013-14):
  - Subjekty studují vyšší počet článků, zabývajících se osobním životem prezidentů než jejich politickou činností
  - Více si z nich i pamatují

Ale: průměrný článek o politice studují výrazně déle, než článek o osobním životě.

(vliv IRL nepozorovatelné intervenující proměnné)

# Kontrola vlivu Z pomocí statistiky

- S některými Z (demografické faktory) nemůžeme udělat ani to, že je držíme konstantní, ani to, že je subjektům přiřadíme náhodně.
- Řešíme **vícenásobnou regresí**, snažíme se odhadnout nejen účinek T, ale i Z na Y (funkce „Control for“ v regresní analýze). Toto je dnes nevyhnutelná součást experimentů, že v nich, kromě M/T, u subjektů měříme i Z (zejména ta Z, která nejsme schopni sami manipulovat).
- Bohužel oproti observačním studiím pak obvykle nižší počet pozorování, horší reprezentativita a možnosti generalizovat, vzhledem k tomu, jak bývá často rekrutován vzorek (pozor! Náhodný výběr nerovná se náhodné přiřazení!)