

Dama koja razlikuje engleski čaj

Damjan Krstajić

(objavljeno 25. februara 2023. u Politikinom Kulturnom dodatku)

Šta mislite kada naučnici obično pozovu u pomoć statističara? Pre ili posle eksperimenta? Iz moje prakse, skoro uvek posle, onda kad su završili sa opitom. U potpunosti pogrešan pristup!

Dizajn je ključan deo naučnog eksperimenta, a jedna dama je pre sto godina dala povod za ozbiljna istraživanja u statistici. Izjavila je na jednoj čajanci da na osnovu ukusa može da razlikuje da li se u šolji engleskog čaja prvo sipalo mleko pa onda čaj ili obrnuto. Rastvor mleka i čaja bi trebao da bude isti bez obzira na redosled, tako da je bilo sumnji i na čajanci je organizovan eksperiment. Kako biste vi proverili njenu tvrdnju?

Jednostavno - često čujem ovaj odgovor. Napravili bi nekoliko šolja u kojima bi prvo sipali čaj pa mleko, a nekoliko sa obrnutim redosledom i dali dami da proba. Nažalost, ovaj pristup je pogodan za neku žurku, ali bez primerene naučne metodologije.

Hajde da krenemo s jednostavnim sitnicama. Nije isto ako osoba koja donosi dami šolju čaja zna tačan odgovor ili ne. Ako zna, može nekim svojim nesvesnim gestom da oda tačan odgovor dami i tako kompromituje rezultate. Više ćemo verovati rezultatima ako direktni učesnici u njoj ne znaju odgovor. U ovom slučaju, važno je da oni koji znaju tačan odgovor nisu prisutni dok dama isprobava engleski čaj. Dakle, poželjno je da ovo bude slep eksperiment (blind experiment).

Drugo, bitno je kojim redosledom se donose šolje. Ako bismo dami ponudili šolju sa prvosipanim mlekom, pa onda sa prvosipanim čajem, pa onda opet sa prvosipanim mlekom, možda će ona moći da provali šablon i da na taj način pogodi sledeću šolju. Bitno je da redosled bude slučajan, određen, recimo, bacanjem novčića. Više ćemo verovati rezultatima u kojima nema šablona. Dakle, poželjno je da ovo bude randomizovan eksperiment (randomized experiment). Iz iskustva znam da kad spomenem reč randomizacija dosta njih odustane od slušanja, zato vas molim da nastavite sa daljim čitanjem, jer taj termin neću spominjati dalje u tekstu.

E sad sledi zanimljiv deo. Posle koliko tačno pogođenih šolja ćemo poverovati dami? Šta ako dâ jedan netačan odgovor, da li to onda znači da ne ume da razlikuje šta je prvo sipano u šolju? A dva netačna odgovora?

Ronald Fišer (1890-1962), rodonačelnik moderne statistike, ovaj primer sa eksperimentom gde dama tvrdi da zna šta je prvo sipano u šolju engleskog čaja analizirao je do detalja u svojoj knjizi *Dizajn eksperimenta*, objavljenoj 1935. godine.

Fišer je uočio da koliko god puta da se desi da dama dâ tačan odgovor, uvek će postojati neka verovatnoća da je imala „sreću“ da ih pogodi. Drugim rečima, nikad sa ovakvim eksperimentom nećemo moći da dokažemo, u matematičkom smislu, da dama 100% razlikuje kako se spremio engleski čaj. Možemo samo govoriti o verovatnoći dobijenih rezultata pod pretpostavkom da je tačna neka hipoteza.

Kad možemo da kažemo da je neki događaj maloverovatan? Po Fišerovom mišljenju, to sve zavisi od slučaja do slučaja, ali on lično preferira 5% kao granicu. Nažalost, drugi su kasnije, ne čitajući Fišera pažljivo, odlučili da nešto arbitrarno, kao vrednost 5%, postane statistički značajna granica u nauci.

Da se vratimo našem problemu. Koliko šolja čaja je dovoljno za ovaj eksperiment? Uporedo sa Fišerom, Jerži Nejman (1894-1981) i Igon Pirson (1895-1980) su izgradili nešto drugačiji pristup koji nam daje formulu za izračunavanje optimalne veličine uzorka u eksperimentu.

Dakle, pre nego što izvedemo eksperiment, možemo da procenimo potrebno vreme i troškove, kao i da se upoznamo sa raspoloživim alatima za analizu dobijenih rezultata. Na kraju ovakvo dobro osmišljenog procesa, jedan statističar biće svakako višak.

Fišer nije nikad spominjao da je stvarno postojala dama koja tvrdi da razlikuje engleski čaj, već je njen slučaj u knjizi koristio kao misaoni eksperiment. Međutim, zahvaljujući Fišerovom prijatelju koji je prisustvovao čajanci, danas se zna i kako se dama zvala i kakav je bio rezultat testa tada. Od osam šolji engleskog čaja, u nekim je prvo sipano mleko, a u drugim prvo čaj, dama je bila u pravu u vezi sa svakom šoljom!

Ovde je važno da razumemo da mi ne možemo tvrditi da je ona dokazala da je u pravu. Postoji mogućnost da je imala sreću da pogodi tačan odgovor za svaku šolju. Verovatnoća da slučajno pogodi tačan odgovor za

osam šolji iznosi $(\frac{1}{2})^8$. A kako je to manje od 0.4%, onda možemo samo zaključiti da je veoma malo verovatno da dama slučajno pogađa šta je prvo sipano u šolju engleskog čaja.

Neko može s pravom reći da ovaj naučni zaključak nije pravedan prema uvaženoj dami, ali je važno razumeti da u eksperimentalnim naukama nemamo rezultate koji daju 100% dokaze, već samo potvrde da su rezultati manje ili više verovatni, pod uslovom da su neke hipoteze tačne.

Susretao sam se sa ljudima koji vole eksperimentalnu nauku, ali ne i statistiku. Kad treba da objave svoje pronalaskе, oni onda kopiraju dizajn eksperimenta iz nekog drugog naučnog rada. Međutim, kako će moći da pronađu i objasne nešto novo, ako stalno koriste dizajne drugih? Iz mog iskustva, razlika između dobrih i vrhunskih eksperimentalnih naučnika je što jedni beže od statistike, a drugi je drže u malom prstu. Jedni zovu statističara u pomoć posle eksperimenta, a drugi se konsultuju s njim pre.

Reference koje podržavaju činjenice spomenute u članku

1. Lady tasting tea

https://en.wikipedia.org/wiki/Lady_tasting_tea

2. Blind experiment

https://en.wikipedia.org/wiki/Blinded_experiment

3. Randomized experiment

https://en.wikipedia.org/wiki/Randomized_experiment

4. Ronald Fisher

https://en.wikipedia.org/wiki/Ronald_Fisher

5. The Design of Experiments (book)

https://en.wikipedia.org/wiki/The_Design_of_Experiments

6. Statistička značajnost

https://en.wikipedia.org/wiki/Statistical_significance

7. Jerži Nejman

https://en.wikipedia.org/wiki/Jerzy_Neyman

8. Igon Pirson

https://en.wikipedia.org/wiki/Egon_Pearson

9. Nejman-Pirson teorema

<http://rsta.royalsocietypublishing.org/content/roypta/231/694-706/289.full.pdf>