Simpsonov paradoks i načelo sigurne stvari

ZVONIMIR ŠIKIĆ[1]

Will Rogers:

When the okies moved to California,
they raised the average intelligence in both states.

Simpsonov paradoks

Simpsonov paradoks prikazat ćemo na artificijelnom primjeru koji je u biti jednak Simpsonovom originalnom primjeru iz 1951. On se temelji na podatcima o 80 pacijenata – 40 muškaraca (M) i 40 žena (Ž).

M	O	-O
L	7	3
-L	18	12

Ž	O	-O
L	9	21
-L	2	8

	O	-O
L	16	24
-L	20	20

Liječeno je 10 muškaraca (L) od kojih se 7 oporavilo (LO). Nije liječeno 30 muškaraca (-L) od kojih se 18 oporavilo (-LO). To je prikazano u lijevoj M-tablici.

Liječeno je 30 žena (L) od kojih se 9 oporavilo (LO). Nije liječeno 10 žena (-L) od kojih se 9 oporavilo (-LO). To je prikazano u srednjoj Ž-tablici.

Kombinirana tablica u kojoj su zajedno muškarci i žene je desno. Dakle, liječeno je 40 pacijenata (L) od kojih se 16 oporavilo (LO). Nije liječeno 40 pacijenata (-L) od kojih se 20 oporavilo (-LO).

Stopa oporavka liječenih muškaraca, 7/10 = 70 %, viša je od stope oporavka neliječenih muškaraca, 18/30 = 60 %.

Stopa oporavka liječenih žena, 9/30 = 30 %, viša je od stope oporavka neliječenih žena, 2/10 = 20 %.

U cijeloj populaciji stope su obrnute. Stopa oporavka liječenih pacijenata, 16/40 = 40 %, niža je od stope oporavka neliječenih pacijenata, 20/40 = 50 %.

[1]Zvonimir Šikić, Sveučilište u Zagrebu

$$7/10 = 70\ \% > 18/30 = 60\ \%$$

$$9/30 = 30\ \% > 2/10 = 20\ \%$$

$$16/40 = 40\ \% < 20/40 = 50\ \%$$

Mnogima je to iznenađujuće, a nekima čak i paradoksalno, pa takve obrate nazivaju Simpsonovim paradoksima. Ljudi očekuju da više stope u sub-populacijama moraju rezultirati višim stopama u populaciji, jer zbrajaju se sub-populacije pa bi se nekako trebale zbrajati i njihove stope, npr. računanjem prosjeka, a ono je monotono. No, kada se zbrajaju sub-populacije, njihove se stope $\alpha_1 = \frac{a_1}{A_1}$ i $\alpha_2 = \frac{a_2}{A_2}$ „zbrajaju" prema pravilu „lošeg učenika":

$$\alpha_1 \oplus \alpha_2 = \frac{a_1 + a_2}{A_1 + A_2}.$$

Lako je izračunati:

$$\frac{a_1 + a_2}{A_1 + A_2} = \frac{a_1}{A_1}\frac{A_1}{A_1 + A_2} + \frac{a_2}{A_2}\frac{A_2}{A_1 + A_2} = \alpha_1 \frac{A_1}{A_1 + A_2} + \alpha_2 \frac{A_2}{A_1 + A_2}.$$

Dakle, ako znamo samo stope u sub-populacijama (tj. znamo α_1 i α_2), onda A_1 i A_2 mogu biti bilo koji pozitivni cijeli brojevi, a stopa populacije (tj. $\alpha_1 \oplus \alpha_2$) može biti bilo koja vrijednost oblika:

$$\alpha_1 \oplus \alpha_2 = p\alpha_1 + q\alpha_2,\ \ p,q > 0\ \ \&\ \ p + q = 1.$$

To znači da $\alpha_1 \oplus \alpha_2$ može biti bilo koja (racionalna) vrijednost iz intervala (α_1, α_2).

Definicija Simpsonove „paradoksalne" situacije:

Situacija u kojoj je $\alpha_1 \le \beta_1$ i $\alpha_2 \le \beta_2$ jest Simpsonova „paradoksalna" situacija ako $\alpha_1 \oplus \alpha_2$ i $\beta_1 \oplus \beta_2$ mogu biti u bilo kojem redoslijedu:

$$\alpha_1 \le \beta_1 \quad \alpha_2 \le \beta_2$$

$$\alpha_1 \oplus \alpha_2 \begin{matrix} > \\ = \\ < \end{matrix} \beta_1 \oplus \beta_2.$$

Karakterizacija Simpsonove „paradoksalne" situacije:

Situacija u kojoj je $\alpha_1 \le \beta_1$ i $\alpha_2 \le \beta_2$ Simpsonova je „paradoksalna" situacija ako i samo ako skup $(\alpha_1, \alpha_2) \cap (\beta_1, \beta_2)$ nije prazan (ovdje koristimo oznaku u kojoj je $(\alpha_1, \alpha_2) = (\alpha_2, \alpha_1)$ i $(\beta_1, \beta_2) = (\beta_2, \beta_1)$.

Dokaz je trivijalan. Ako je presjek prazan, onda je svaki $\alpha_1 \oplus \alpha_2 \in (\alpha_1, \alpha_2)$ manji od svakog $\beta_1 \oplus \beta_2 \in (\beta_1, \beta_2)$ pa nemamo Simpsonov „paradoks" jer, kako god odabrali $\alpha_1 \oplus \alpha_2$ i $\beta_1 \oplus \beta_2$, uvijek je $\alpha_1 \oplus \alpha_2 < \beta_1 \oplus \beta_2$. Ako presjek nije prazan, $\alpha_1 \oplus \alpha_2$ i $\beta_1 \oplus \beta_2$ mogu biti bilo koje vrijednosti iz tog presjeka pa uz odgovarajuće odabire možemo dobiti bilo koji poredak između $\alpha_1 \oplus \alpha_2$ i $\beta_1 \oplus \beta_2$, tj. imamo Simpsonov „paradoks".

U našem medicinskom primjeru $(\alpha_1, \alpha_2) = (0.7, 0.3)$ i $(\beta_1, \beta_2) = (0.6, 0.2)$ pa je

$$(\alpha_1, \alpha_2) \cap (\beta_1, \beta_2) = (0.3, 0.7) \bigcap (0.2, 0.6) = (0.3, 0.6).$$

Vrijednosti za $\alpha_1 \oplus \alpha_2$ i $\beta_1 \oplus \beta_2$ mogu se slobodno birati unutar intervala (0.3,0.6) pa uz odgovarajuće odabire možemo dobiti bilo koji poredak između $\alpha_1 \oplus \alpha_2$ i $\beta_1 \oplus \beta_2$.

Dakle, tu definitivno nema nikakvog paradoksa. No, psihološka je činjenica da ljudi ovakve situacije ipak doživljavaju kao paradoksalne i to je dokazano mnogim psihološkim eksperimentima. No, imam neke sumnje i oko tih eksperimenata.

Kako izgledaju ti eksperimenti? Navodim dva tipična.

U prvom „nematematičkom" eksperimentu ispitanicima je rečeno da u jednom školskom okrugu postoje samo dvije srednje škole i da je prolaznost djevojaka na maturi u obje škole bolja od prolaznosti dječaka.

Nakon što su upoznati s ovim činjenicama, ispitanicima je postavljeno pitanje slijedi li da je u tome okrugu prolaznost djevojaka na maturi bolja od prolaznosti dječaka. Ponuđeni su im sljedeći odgovori:

a. Da, u tom je okrugu stopa maturiranja djevojčica veća od stope maturiranja dječaka.

b. Ne, u tom je okrugu stopa maturiranja djevojčica manja od stope maturiranja dječaka.

c. Ne, u tom je okrugu stopa maturiranja djevojčica jednaka stopi maturiranja dječaka.

d. Ne može se zaključiti jer nema dovoljno informacija.

U drugom „matematičkom" eksperimentu ispitanicima je rečeno da su sljedeće informacije točne:

$$\frac{f_1}{F_1} > \frac{m_1}{M_1}, \quad \frac{f_2}{F_2} > \frac{m_2}{M_2}.$$

Zatim su upitani slijedi li iz toga da je

$$\frac{f_1 + f_2}{F_1 + F_2} > \frac{m_1 + m_2}{M_1 + M_2}$$

te su im ponuđeni sljedeći odgovori:

(a) Da, prvi je izraz veći od drugog.

(b) Ne, prvi je izraz manji od drugog.

(c) Ne, prvi i drugi izraz su jednaki.

(d) Ne može se zaključiti jer nema dovoljno informacija.

U eksperimentu sa 106 studenata, Bandyoapdhyay i koautori ustvrdili su da su na prvo nematematičko pitanje studenti netočno odgovorili (a) u 83 % slučaja, a točno su odgovorili (d) samo u 12 % slučajeva. Na matematičko pitanje netočno su odgovorili (a) u 57 % slučajeva, dok su točno odgovorili (d) u 29 % slučajeva. Slične ankete ponovljene su mnogo puta i pokazale su iste obrasce odgovora.

Moje sumnje u ove eksperimente najbolje ću objasniti usporedbom s jednim drugim problemom:

Automobil A vozi brže od automobila B. Koji će automobil sa zajedničkog polazišta prvi stići na zajedničko odredište?

(a) Automobil A stići će prvi.

(b) Automobil B stići će prvi.

(c) Automobil A i automobil B stići će u isto vrijeme.

(d) Ne može se zaključiti jer nema dovoljno informacija.

Točan odgovor je (d) jer nigdje nismo rekli da automobili kreću u isto vrijeme, iako to svi pretpostavljaju. Da smo dali informaciju da je B krenuo ranije od A, broj bi se točnih odgovora radikalno povećao.

Broj točnih odgovora radikalno bi se povećao i da smo dali informacije o disbalansu djevojčica i dječaka u srednjim školama. Te dodatne informacije dolje su podebljane.

U jednom školskom okrugu postoje samo dvije srednje škole, T i L. Na maturi je prolaznost djevojaka u školi T veća od prolaznosti dječaka u toj školi. **Škola T upisuje gotovo isključivo dječake**. Prolaznost djevojaka u školi L također je veća od prolaznosti dječaka u toj školi. **Škola L upisuje gotovo isključivo djevojke**.

Slijedi li da u tom okrugu djevojke imaju veću prolaznost mature od dječaka?

Dodatne informacije upućuju na to da je vjerojatnost uspjeha djevojaka u cijelom okrugu približno jednaka vjerojatnosti njihovog uspjeha i u školi L i da je vjerojatnost uspjeha dječaka u cijelom okrugu približno jednaka vjerojatnosti njihovog uspjeha u školi T.

$$P_{Ž}(U) \approx P_{Ž}(U|L) \qquad P_{M}(U) \approx P_{M}(U|T)$$

S obzirom na to da o odnosu veličina $P_{Ž}(U|L)$ i $P_{M}(U|T)$ nemamo nikakvih informacija, slijedi da ništa ne možemo zaključiti ni o odnosu veličina $P_{Ž}(U)$ i $P_{M}(U)$.

Dakle, (d) je točan odgovor.

U izvornom problemu nemamo informaciju o disbalansu djevojčica i dječaka u dvije škole pa ljudi automatski pretpostavljaju da on ne postoji i zaključuju:

$$P_Ž(U) \approx ½\, P_Ž(U|T) + ½\, P_Ž(U|L) > ½\, P_M(U|T) + ½\, P_M(U|L) \approx P_M(U)$$

Što to dokazuje? Samo da ljudi u nedostatku informacija praznine popunjavaju najjednostavnijim pretpostavkama. To je zdravi razum (*common sense*) koji ljude razlikuje od strojeva. Iako se ogromna sredstva ulažu u izgradnju *common sense* algoritama, uspjesi su ograničeni (usp. J. Pavlos 2020). Ako vam kažem „*Ivan je otišao u restoran, naručio odrezak i ostavio veliku napojnicu*", vi znate da je Ivan pojeo odrezak. Strojevi ne uspijevaju doći do tog zdravorazumskog zaključka jer nigdje u toj maloj sceni nije navedeno da je čovjek išta pojeo. Nama zdrav razum dopušta čitanje između redaka. Nas ne treba eksplicitno informirati da se u restoranima hrana jede nakon što se naruči i prije nego što se da napojnica.

No, vratimo se temi. Da se ne radi samo o teorijskim razmatranjima, dokazuje sljedeći primjer iz sveučilišnog života. Početkom 1970-ih kalifornijsko sveučilište Berkeley tuženo je zbog spolne diskriminacije pri upisu na postdiplomske studije. Naime, u jesen 1973. na najveće odjele primljeno je 46 % kandidata i 30 % kandidatkinja. Na prvi pogled, i pod pretpostavkom da su kvalifikacije podnositelja zahtjeva bile slične, ovi podatci doista bi mogli ukazivati na rodnu diskriminaciju. Međutim, kada se pogledaju podatci unutar tih odjela, predrasuda prema kandidatkinjama nestaje.

Table 1: Data From Six Largest Departments of 1973 Berkeley Discrimination Case

Department	Men		Women	
	Applicants	Admitted	Applicants	Admitted
A	825	62 %	108	82 %
B	560	63 %	25	68 %
C	325	37 %	593	34 %
D	417	33 %	375	35 %
E	191	28 %	393	24 %
F	272	6 %	341	7 %

Source: Bickel, Hammel, and O'Connell (1975); table accessed via Wikipedia at https://en.wikipedia.org/wiki/Simpson%27s_paradox.

Tu je došlo do Simpsonovog „paradoksa" jer su se žene više prijavljivale na odjele s niskim stopama prihvaćanja, dok su se muškarci više prijavljivali na odjele s visokim stopama prihvaćanja (kao što se vidi u gornjoj tablici).

To možemo još ekstremnije ilustrirati fiktivnim slučajem srednje škole iz koje izlazi 10 izvrsnih i 100 prosječnih maturanata. Od 10 izvrsnih njih se 9 pokušalo upisati na sveučilišta A lige, a uspjelo je njih 5. Pokušao je i 1 prosječni, ali nije uspio. Ostali, tj. 99 prosječnih i 1 izvrsni, pokušali su se upisati na sveučilišta C lige i uspjelo je 80 prosječnih i taj jedan izvrsni.

Iz srednje škole izlazi	10 izvrsnih	100 prosječnih
Upisuje A ligu	5/9 = 55 %	0/1 = 0 %
Upisuje C ligu	1/1 = 100 %	80/99 = 81 %
Ukupno	6/10 = 60 %	80/100 = 80 %

Iako su izvrsni očekivano uspješniji u obje lige, oni su ukupno neuspješniji. No, istina o uspješnosti očito je u sub-populacijama, a ne u ukupnoj populaciji, o čemu će još biti govora.

No, vratimo se prije toga našoj analizi porijekla „doživljaja paradoksa”. Kao argument da tu nije riječ samo o prešutnim zdravorazumskim pretpostavkama, već da su Simpsonove situacije u najmanju ruku matematički zahtjevne, ako ne i paradoksalne, mogla bi se navesti činjenica da je Simpsonova „paradoksalna” situacija bila jedan od problema na matematičkoj olimpijadi 2009./2010. Evo problema.

Imate četiri čaše – sa sokom od jabuke, sokom od breskve, sokom od grejpa i sokom od mrkve. Sok od jabuke slađi je od soka od grejpa, a sok od breskve slađi je od soka od mrkve. Je li mješavina soka od jabuke i soka od breskve slađa od mješavine soka od grejpa i soka od mrkve?

Ili jednostavnije, znamo da je $J > G$ i $B > M$. Je li nužno da je $J+B > G+M$? Naravno, > tu znači slađe, a ne količinski veće.

Primijetite da ništa ne znamo o $G \lesseqgtr B$. Dakle, moguće je da je $B < G$, pa samo trebamo postići $B + J \approx B$, $G + M \approx G$ i tada imamo $B+J \approx B < G \approx G+M$. Dakle, stavite nekoliko kapi J u B i nekoliko kapi M u G i dobit ćete $J+B < G+M$.

Pretpostavljam da autor problema nije očekivao tako jednostavno rješenje. No, problem matematički nije zahtjevan i svodi se na uzimanje u obzir odgovarajućih pretpostavki.

Da zaključim. Primjena pogrešnog načela „stopa u populaciji jednaka je prosjeku stopa u sub-populacijama” sigurno je odgovorna za pogrešne procjene Simpsonovih situacija. No, držim da važnu ulogu u tome ima i činjenica da ljudi u nedostatku informacija popunjavaju praznine zdravorazumskim pretpostavkama koje često opravdavaju primjenu tog načela. S druge strane, neke možda složenije pretpostavke dovode do drugih rezultata koji mogu biti u suprotnosti s tim načelom i to ga ipak čini nevaljanim.

Načelo sigurne stvari

U vezi sa Simpsonovim paradoksom često se spominje i načelo sigurne stvari koje je formulirao L. J. Savage u *Foundation of Statistics*, 1954:

Ako osoba preferira f u odnosu na g, bilo znajući da se dogodio B, ili ne znajući da se dogodio B, onda bi trebala preferirati f u odnosu na g i ako ništa ne zna o B. ... To nije ni logički ni probabilistički princip.

Neepistemološka varijanta načela glasila bi ovako:

Ako uzrok U povećava vjerojatnost posljedice P u komplementarnim sub-populacijama, onda je povećava i u populaciji.

Čini se da Simpsonove „paradoksalne" situacije obaraju to načelo. U našem medicinskom primjeru kao uzrok imamo liječenje, a kao posljedicu oporavak. Suprotno načelu sigurne stvari, liječenje povećava vjerojatnost oporavka u obje (komplementarne) sub-populacije M i Ž, ali ne i u cijeloj populaciji.

Stope oporavka	L	-L
M	70 %	60 %
Ž	30 %	20 %
Svi	40 %	50 %

Ipak, gotovo svi su uvjereni da je prava istina u sub-populacijama. Ali koji su argumenti za tvrdnju da ono što vrijedi u sub-populacijama vrijedi i u cijeloj populaciji, bez obzira na podatke u cijeloj populaciji? Mogli bismo reći da je intuitivno jasno da lijek koji pomaže i muškarcima i ženama mora pomoći svima. Ali na čemu se temelji ta intuicija? Slijedi jedno moguće objašnjenje.

Pretpostavimo da znamo da estrogen ima negativan učinak na oporavak pa je, bez obzira na lijek, vjerojatnost oporavka žena značajno manja nego u muškaraca. Osim toga (što vidimo iz podataka) žene su uzimanju lijeka sklonije od muškaraca. Dakle, razlog zašto se čini da je lijek u cijeloj populaciji štetan jest da je, uz nasumični odabir korisnika lijeka, veća vjerojatnost da se radi o ženi pa je manje vjerojatno da će se odabrana osoba oporaviti. Dakle, da bismo procijenili učinkovitost lijeka, moramo usporediti osobe istoga spola jer ćemo tako osigurati da se razlika u stopama oporavka između onih koji uzimaju lijek i onih koji ga ne uzimaju ne može pripisati estrogenu.

To je odgovor i na pokušaj C. Blytha iz 1972. da sruši načelo sigurne stvari sljedećim argumentom, vezanim uz naš medicinski primjer. Predložio je dvije oklade:

Slučajno odabirete pacijenta dok ne naiđete na jednog koji

- uzima lijek (to je prva oklada L)
- ne uzima lijek (to je druga oklada -L)

i tada se u 100 eura kladite da će se taj pacijent oporaviti.

Ako je odabrani pacijent muškarac, iz podataka slijedi da vam se više isplati -L oklada. Ako je odabrani pacijent žena, iz podataka opet slijedi da vam se više isplati -L oklada. Ako ne znate je li odabrani pacijent muškarac ili žena, iz podataka slijedi da vam se više isplati L oklada. To, prema Blythu, obara načelo sigurne stvari.

No, to ne obara kauzalno načelo sigurne stvari koje je eksplicitno formulirao J. Pearl 2000., a implicitno ga je podrazumijevao L. J. Savage 1954:

Akcija U koja povećava vjerojatnost događaja P u obje sub-populacije S i -S povećava vjerojatnost od P u cijeloj populaciji, ***pod uvjetom da ta akcija ne utječe na distribuciju sub-populacija.***

Naime, akcija ***L = odabiri pacijente dok ne dođeš do jednog koji uzima lijek*** dovodi do odabira muškarca s vjerojatnošću 12.5 % jer je

$$\Pr(ML) = \Pr(M)\Pr(L|M) = 50\,\%\ 25\,\% = 12.5\,\%,$$

a do odabira žene dovodi s vjerojatnošću 37.5 % jer je

$$\Pr(ŽL) = \Pr(Ž)\Pr(L|Ž) = 50\,\%\ 75\,\% = 37.5\,\%.$$

S druge strane, akcija ***-L = odabiri pacijente dok ne dođeš do jednog koji ne uzima lijek*** dovodi do odabira muškarca s vjerojatnošću 37.5 % jer je

$$\Pr(ML) = \Pr(M)\Pr(L|M) = 50\,\%\ 75\,\% = 37.5\,\%,$$

a do odabira žene dovodi s vjerojatnošću 12.5 % jer je

$$\Pr(ŽL) = \Pr(Ž)\Pr(L|Ž) = 50\,\%\ 75\,\% = 12.5\,\%.$$

Dakle, akcije L i -L utječu na distribuciju sub-populacija odabranih muškaraca i odabranih žena.

S druge strane, u našem izvornom primjeru, akcije ***L = daj pacijentu lijek*** i ***-L= ne daj pacijentu lijek*** ne utječu na sub-populacije muškaraca i žena jer lijek ne utječe na spol.

Ako, kao R. Jeffery 1982., uvjet korektne primjene načela sigurne stvari pojačamo do oblika „*akcije i sub-populacije moraju biti* ***stohastički*** *nezavisne*", načelo sigurne stvari svodi se na sasvim jednostavni probabilistički zakon. Naime, taj jači uvjet jamči da je $\Pr(S|U) = \Pr(S|-U) = \Pr(S)$ i $\Pr(-S|U) = \Pr(-S|-U) = \Pr(-S)$ pa se standardno probabilističko načelo

$$\Pr(P|U) = \Pr(P|U,S)\Pr(S|U) + \Pr(P|U,-S)\Pr(-S|U)$$

$$\Pr(P|-U) = \Pr(P|-U,S)\Pr(S|-U) + \Pr(P|-U,-S)\Pr(-S|-U)$$

svodi na

$$\Pr(P|U) = \Pr(P|U,S)\Pr(S) + \Pr(P|U,-S)\Pr(-S)$$

$$\Pr(P|-U) = \Pr(P|-U,S)\Pr(S) + \Pr(P|-U,-S)\Pr(-S).$$

I tada iz $\Pr(P|U,S) > \Pr(P|-U,S)$ i $\Pr(P|U,-S) > \Pr(P|-U,-S)$ odmah slijedi $\Pr(P|U) > \Pr(P|-U)$.

No, Jefreyjev je uvjet pretjeran. Njegovo isključivanje svih slučajeva u kojima su sub-populacije i akcije stohastički ovisne isključio bi istraživanja u kojima pojedinci s određenim karakteristikama (spol, razina obrazovanja i sl.) s većom vjerojatnošću traže liječenje, što je čest slučaj u opservacijskim istraživanjima. Dovoljno je isključiti samo kauzalne ovisnosti, a dopustiti čisto stohastičke. Ta se razlika ne može izraziti u jeziku teorije vjerojatnosti, ali može u *do-calculusu* koji formalizira kauzalnost i u njemu je kauzalno načelo sigurne stvari dokazivo (J. Pearl 2000.).

Dakle, odgovor na pitanje treba li u obzir uzimati stope u sub-populacijama ili stopu u cijeloj populaciji, ne nalazi se u podatcima. Kako bismo odlučili hoće li lijek naštetiti ili pomoći pacijentu, prvo moramo razumjeti „priču” koja stoji iza podataka, tj. uzročne mehanizme koji su generirali podatke koje vidimo. Kao što to ispravno tvrde udžbenici statistike, korelacija nije uzročnost, pa najčešće nema statističke metode koja uzročnu priču može odrediti samo iz podataka. Jednostavno, nema čisto statističke metode odlučivanja pa zato istraživači, uz korištenje statističkih metoda, podatke najčešće moraju tumačiti i na temelju neformalnih uzročno-posljedičnih pretpostavki.

Postoje, međutim, i nestatističke formalne metode koje se mogu koristiti za izražavanje i tumačenje uzročnosti. Uz njihovu pomoć moguće je matematički opisati uzročne scenarije bilo koje složenosti i odgovoriti na probleme donošenja odluka i u bitno složenijim situacijama od onih koje generiraju Simpsonovi „paradoksi”. Riječ je o matematičkoj teoriji kauzalnosti razvijenoj posljednjih 30-ak godina (u formi već spomenutog *do-calculusa*) koju mnogi smatraju „kauzalnom revolucijom”. No, to je posebna tema koja je izvan okvira ovoga članka.

Literatura:

1. Bandyoapdhyay, P. S. et al. (2011.). The logic of Simpson's paradox. *Synthese* 181,185–208.
2. Blyth, C. (1972.). On Simpson's paradox and the sure-thing principle. *Journal of the American Statistical Association, 67* (338), 364–366.
3. Pavlos J. (2020.). Common Sense Comes Closer to Computers, *Quanta Magazine*, April 30.
4. Pearl, J. (2000.). *Causality: Models, Reasoning, and Inference*. Cambridge U. Press.
5. Savage, L. J. (1954.). *The Foundations of Statistics*, Wiley.
6. Simpson, H. (1951.). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society. Series. B, 13* (2), 238–241.