Paradoks slabovidne priče

Boštjan Kuzman

→ V nekem velemestu se je ponoči zgodila prometna nesreča, njen povzročitelj pa je s kraja nesreče pobegnil. Priča nesreče trdi, da je nesrečo povzročil taksi bele barve. Ker je v mestu 90 % taksijev rumenih in le 10 % belih, ta podatek zelo zoži krog osumljencev za policijsko preiskavo. Toda laboratorijski preizkus priče pokaže, da je priča slabovidna in v nočnih razmerah pravilno določi barvo taksija le v 80 % primerov. Čeprav se na prvi pogled zdi priča dovolj zanesljiva, bo natančnejši izračun pokazal nasprotno. Verjetnost, da je nesrečo res povzročil beli taksi, je kljub izjavi priče precej majhna. Danes dobro znani paradoks sta leta 1982 v knjigi o odločanju v negotovih okoliščinah prva opisala matematično izobražena psihologa Amos Tversky in Daniel Kahneman, ki je kasneje prejel tudi Nobelovo nagrado za ekonomijo. V tem prispevku bomo paradoks pojasnili in ga z GeoGebro ponazorili tudi grafično.

Nekatere bralke in bralci bodo v problemu takoj prepoznali pogojno verjetnost in ga rešili s pomočjo Bayesove formule. Toda poskusimo vse skupaj razložiti brez uporabe tovrstnih orodij. Predstavljajmo si, da je v mestu 100 taksijev, med katerimi je 90 rumenih in 10 belih, nesrečo pa je povzročil naključno izbrani taksi. Če vse taksije pokažemo priči v nočnih razmerah, bo priča barvo pravilno določila v 80 % primerov. Torej bo med 90 rumenimi taksiji pravilno prepoznala le 72 rumenih, ostalih 18 pa bo napačno označila za bele. Podobno bo priča med 10 belimi taksiji pravilno prepoznala osem belih, preostala dva pa bo napačno označila za rumena. Skupaj bo priča med 100 taksiji kar $18 + 8 = 26$ taksijev prepoznala kot belih, čeprav je belih v resnici med njimi le osem. Če priča trdi, da je nesrečo povzročil beli taksi, ima torej prav le v osmih primerih od 26, kar je enako $\frac{8}{26} \doteq 31\,\%$. To razmerje predstavlja verjetnost, da je nesrečo res povzročil beli taksi, če tako trdi naša priča, in rezultat je manjši od $1/3$.

Presenetljivi rezultat lahko pojasnimo z ugotovitvijo, da je belih taksijev le 10 %, zato je verjetnost, da je nesrečo povzročil beli taksi, razmeroma majhna. Le 80 % zanesljivost priče v tej situaciji ni dovolj, da bi se verjetnost bistveno povečala. **Ugotovitev si velja zapomniti: če malo verjeten dogodek potrdi ena razmeroma zanesljiva priča, to ne pomeni nujno velike verjetnosti, da se je dogodek res zgodil.** Drugače pa bi bilo, če bi priča trdila, da je nesrečo povzročil rumeni taksi, in bi določali verjetnost, da je to res. Verjetnost rumenega taksija je 90 % in je že brez pričanja zelo visoka, zato jo dodatno pričanje v to smer še poviša. S podobnim sklepanjem kot prej dobimo rezultat $\frac{72}{72+2} \doteq 97\,\%$, saj bi priča pravilno prepoznala 72 rumenih taksijev, le dva bela taksija pa bi napačno prepoznala kot rumena.

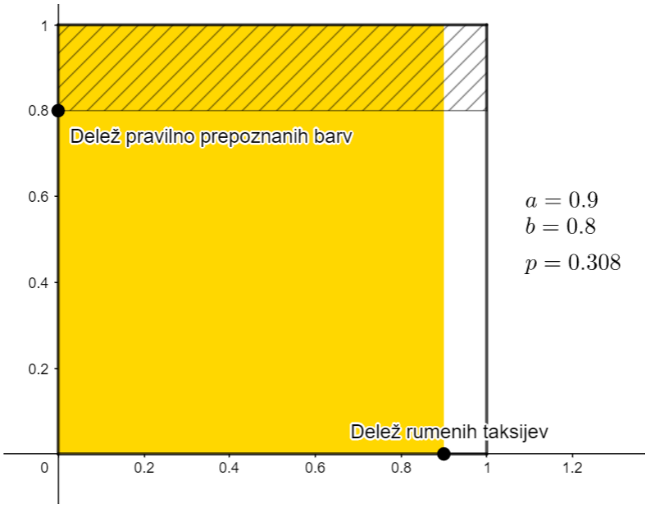
Z GeoGebro lahko opisano situacijo lepo ponazorimo s pomočjo ploščin likov. Ob tem bomo lahko razmerje med belimi in rumenimi taksiji ter zanesljivost priče kasneje tudi spreminjali s pomočjo pomikanja ustreznih točk. Za konstrukcijo uporabimo naslednje korake:

- Izklopimo označevanje novih objektov in z ukazom `Mnogokotnik((0,0),(0,1),(1,0),(1,1))` narišemo enotski kvadrat, ki predstavlja množico vseh taksijev. Obarvajmo ga črno in nastavimo prosojnost na vrednost 0, da bo površina vseeno bela.

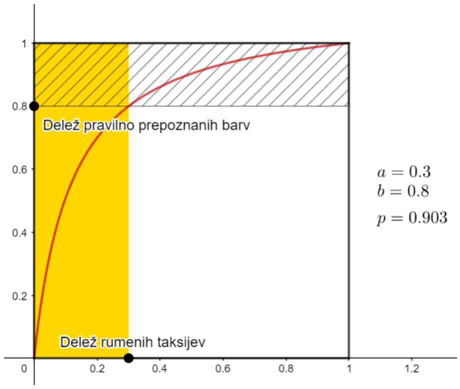
- Približamo pogled in na spodnji stranici kvadrata izberemo poljubno točko A. Njena x-koordinata `a=x(A)` bo predstavljala delež vseh rumenih taksijev. Pomaknemo jo na vrednost 0,9, kar ustreza podatkom naloge. Z ukazom `Mnogokotnik((0,0),(a,0),(a,1),(0,1))` narišemo pravokotnik, ki predstavlja rumene taksije. Obarvamo ga rumeno.
- Na levi stranici kvadrata izberemo poljubno točko B. Njena y-koordinata `b=y(B)` bo predstavljala zanesljivost priče, torej delež pravilno prepoznanih taksijev. Postavimo jo na vrednost 0,8 kot v nalogi. Napačno prepoznane taksije bomo označili s šrafiranim pravokotnikom, ki ga narišemo z ukazom `Mnogokotnik((0,b),(1,b),(1,1),(0,1))`.
- Verjetnost, da je nesrečo res povzročil beli taksi, če tako trdi priča, je enaka razmerju med številom pravilno prepoznanih belih taksijev in številom vseh taksijev, za katere priča trdi, da so beli. To razmerje na sliki predstavlja razmerje med ploščino nešrafiranega belega pravokotnika, ki je enaka $(1-a)\cdot b$ in vsoto ploščin nešrafiranega belega in šrafiranega rumenega pravokotnika, ki je enaka $(1-a)\cdot b + a(1-b)$. Splošna formula za iskano verjetnost je torej enaka

  - $p = \dfrac{b(1-a)}{a+b-2ab}$

  in jo lahko izpišemo na zaslon s pomočjo ukazov za besedilo in označevanje točk. Pri začetnih podatkih $a = 0{,}9$ in $b = 0{,}8$ lahko že s pogledom grobo ocenimo, da bo iskana verjetnost približno 1/3, točen rezultat je 0,308.
- Z upoštevanjem doslej navedenih korakov smo dobili situacijo na sliki 1.
- S pomikanjem drsnikov lahko zdaj preprosto opazujemo, kako se spreminja verjetnost, če spreminjamo delež rumenih taksijev a ali zanesljivost priče b. Če sta a in b enaka, bo verjetnost vselej 1/2.
- Na prikaz lahko dodamo še krivuljo, ki predstavlja množico vseh točk (x,y) z isto verjetnostjo $p = p(a,b)$. V enačbi za p zamenjajmo a,b z x,y in izrazimo y z x in p, da dobimo enačbo krivulje $y = \frac{px}{1-x+2px-p}$. To krivuljo narišemo na območju $0 < x < 1$ z ukazom `If(0<x<1, p*x/(1-x+2*p*x-p))`.



SLIKA 1.
Verjetnost, da je nesrečo povzročil beli taksi, če tako trdi priča, je enaka razmerju p_1/p_2, kjer je p_1 ploščina nešrafiranega belega pravokotnika, p_2 pa vsota ploščin nešrafiranega belega pravokotnika in šrafiranega rumenega pravokotnika.



SLIKA 2.
Pri deležu rumenih taksijev $a = 0{,}3$ in nespremenjeni zanesljivosti priče $b = 0{,}8$ je verjetnost približno 0,9. Rdeča krivulja vsebuje vse pare točk (a,b) s to verjetnostjo, območje nad krivuljo pa predstavlja vse točke, pri katerih je verjetnost večja. Iz prikaza zato razberemo, da je za verjetnost nad 0,9 potreben majhen a ali pa velik b.

× × ×