

Matea Gusić, Čakovec

## MALE STATISTIČKE PRIČE

Još u davnoj prošlosti svaki je (više ili manje) uspješan vladar vodio podatke o populaciji koja nastanjuje zemlju kojom upravlja, o posjedima, vojsci, ekonomiji, hrani koju imaju na raspolaganju... Kako bi znao, primjerice, ima li dovoljno resursa za slanje vojske u pohod ili treba podići harač seljacima, vladar je trebao prikupljati, analizirati i interpretirati podatke, odnosno baviti se statistikom. Da bi na temelju prikupljenih podataka izvukao informaciju, bilo u vidu procjene trenutnog stanja ili predviđanja budućnosti, trebao mu je alat za interpretaciju. Prilikom je ugledala matematika; ruku pod ruku s teorijom vjerojatnosti razvila je primjenu u statistici. Danas ne postoji područje života i djelovanja koje je statistika izbjegla. U industriji i trgovini, političkim kampanjama, određivanju vremenskih prognoza, ispitivanju javnog mnijenja ili primjerice predviđanjima epidemija bolesti, statistika ima glavnu riječ. Svi su čuli za pojam „sportska statistika” koja uz rangiranje timova može odgovoriti i na ključno pitanje, a to je utjecaj koji je pojedini igrač imao na ishod utakmice (pobjedu odnosno poraz). Svi smo mi, bez provođenja eksperimenta, izrekli neku statističku izjavu poput *Zbog silnih školskih obaveza spavam u prosjeku 5 sati dnevno* ili *Dnevno izgubim dva sata pišući domaće zadaće*.

Vratimo se u 4. stoljeće prije Krista, u antičku Grčku. Matematičar i filozof sofisticke škole Hipias iz Elisa jedan je od zaslužnih za rane početke statistike. U to vrijeme nije bilo pouzdane metode za računanje vremena, pa su stari Grci organizirali prošlo vrijeme dajući godinama imena prema poznatim svećenicima ili sudcima. To nije bilo najsretnije rješenje jer širi puk nije bio upoznat s lokalnim sudcima, pa su se teško snalazili u nazivima. Hipias se tada dosjetio da bi svakoj četvrtoj godini mogao dati ime po olimpijskim pobjednicima koji su, kao veliki šampioni, bili nadaleko poznati. Koristeći se metodom srednje vrijednosti, Hipias je prvo izračunao datum prvih Olimpijskih igara, nakon čega je procijenio mogući broj dotadašnjih pobjednika.

U to vrijeme u Indiji je objavljen jedan od najvećih epova na sanskritu, *Mahabharata*, unutar kojeg se nalazi prvi zapisani primjer korištenja statističkog uzorka. To je priča o kralju Rituparni, vrsnom matematičaru i sposobnom kockaru. Poučavajući svoga slugu zapaloga u teške kockarske dugove tehnikama kockanja, kralj Rituparna izjavljuje da može odrediti broj lišća na jednom drvetu bez da ih izbroji. Poslužio se tehnikom uzorka, odnosno izbrojao je koliko listova ima na jednoj grančici drveta te dobiveni rezultat pomnožio ukupnim brojem grančica. Ovom tehnikom služe se i danas vlasnici plantaža kada žele procijeniti primjerice urod u svojim voćnjacima.

Matematičke osnove statistike postavio je u 16. stoljeću talijanski matematičar Gerolmo Cardano, razvivši teoriju vjerojatnosti. Zanimljivo je da i iza ove priče



stoje kockarski motivi. Naime, Cardano se vječito borio s nedostatkom novca, pa je budžet nadopunjavao kockanjem i igranjem šaha. Upravo mu je igra „bacanja kocke” bila osnova za razumijevanje temeljnih koncepata vjerojatnosti. Cardano-va *Knjiga o igrama na sreću*, objavljena nakon njegove smrti, uz prvu sistematizaciju vjerojatnosti sadržavala je i korisne metode varanja u igrama s kockom.

Jedna od poznatijih statističkih priča vezana je uz velikog matematičara Ronald Fishera i damu koja je postala svjetski poznata radi ispijanja čaja. Jednog lijepog dana 1920. godine u Cambridgeu Fisher je s kolegama uživao u najvažnijoj engleskoj tradiciji – ispijanju šalice čaja s mlijekom. Već spomenuta dama požalila se konobaru da je u šalicu prvo ulio mlijeko, a potom čaj, ne poštujući time uvriježeni običaj ulijevanja mlijeka u šalicu s čajem. Dama je ustvrdila ne samo da je velika razlika između dvije metode serviranja čaja, već da ona, bez greške, može detektirati kojom je metodom čaj pripremljen.

Fisher nije mogao dopustiti da se u njegovoj prisutnosti postavi takva hipoteza<sup>1</sup> a da je on ne provjeri, stoga je odlučio provesti eksperiment. Pripremio je osam šalica čaja; četiri u koje je prvo uliveno mlijeko i četiri u koje je prvo uliven čaj. Dami, upoznatoj sa sadržajem, nasumično su nuđene šalice čaja kako bi odredila metodu pripreme. Ishod eksperimenta različito se navodi; neki tvrde da je dama uspješno identificirala svih osam, a drugi šest šalica. Koja god da varijanta bila, postavlja se pitanje je li moguće da je dama do rezultata došla čistim nagađanjem. Situacija je navela Fishera na razmišljanje o tome koliko je šalice potrebno pripremiti, a koliko točno identificirati kako bi se smatralo da testirana osoba posjeduje iznimne kušačke sposobnosti. Matematičkim izračunom Fisher je došao do zaključka da se bez sumnje može ustvrditi da dama s početka priče posjeduje sposobnost identificiranja metode pripreme engleskog čaja. Iz priče se rodila teorija o eksperimentima, odnosno metoda određivanja valjanosti rezultata znanstvenih istraživanja.

Kako ne bi ispalo da se statistika u prošlosti koristila samo kao rasonoda (za sport, kockanje ili ispijanje čaja), ispričat ćemo i jednu ratnu priču. Ključna informacija prilikom planiranja napada Saveznika na Sile osovine (poznatijeg pod nazivom *Dan D*) za vrijeme II. svjetskog rata, svakako je bila *brojnost* njemačkih tenkova. Kako bi došli do traženih saznanja, koristili su dvije metode: prikupljanje informacija od tajnih obavještajaca, ali i statističku procjenu. S obzirom da nisu svi tenkovi jednako prilagođeni svakome terenu, Saveznici su prema specifikacijama pojedinih tenkova procjenjivali vjerojatnost da se određeni model koristi protiv njih. Vjerojatnosni modeli, kao i obavještajci, navodili su na veliko korištenje tenkova modela *Panzer V*. Saveznici su procijenili da su Nijemci u periodu od dvije godine proizveli 1400 tenkova mjesečno. Koristeći se serijskim brojevima zaplijenjenih ili uništenih tenkova odredili su da se mjesečno proizvodilo 256 tenkova modela *Panzer V*, što je bilo mnogo više nego što su tvrdili tajni obavještajci. Uvidom u njemačku dokumentaciju po završetku rata pokazalo se da je mjesečno proizvedeno 255 tenkova, potvrđujući da je (barem u ovom slučaju) riječ statistike vjerodostojnija od riječi špijuna.

<sup>1</sup>Hipotezu možemo opisati kao izjavu koja u sebi sadrži očekivanje istraživača o rezultatu istraživanja koje provodi.

