

# Protiv statističke značajnosti

---

Damjan Krstajić

(objavljeno 1. juna 2024. u Politikinom Kulturnom dodatku)

Više od 800 svetskih naučnika je 2019. potpisalo komentar u časopisu *Nature* pod naslovom *Scientists rise up against statistical significance* u kojem pozivaju da se napusti koncept statističke značajnosti. Američko društvo statističara podržalo je potpisnike.

Za razumevanje statističke značajnosti i p-vrednosti potreban je kurs statistike s naglaskom na testiranje statističkih hipoteza. Međutim, možemo da pružimo istorijsku perspektivu kako su nastali ovi pojmovi i gde je zapelo u njihovoj primeni.

Značaj ovih termina postaje jasan kad uvidimo da su sastavni deo odgovora na važna pitanja u eksperimentalnim naukama: Da li su ovi eksperimentalni rezultati bolji od prethodnih? Po kom principu prihvatamo da su bolji? Koja je verovatnoća da smo pogrešili?

Ronald Fišer (1890-1962) je pre skoro jednog veka izneo sledeće rešenje. Pojednostavljeno, nakon izvršenog opita definiše se hipoteza i izračuna verovatnoća da su eksperimentalni rezultati dobijeni pod uslovom da je ona tačna. Tu verovatnoću nazvao je p-vrednost. U svom radu često je uzimao graničnu vrednost 5% i ako je p-vrednost manja od 5%, onda je smatrao da tu postoji statistička značajnost. Predlagao je da se u tim slučajevima ponovi eksperiment i ako se opet dobija statistička značajnost, onda bi smatrao da su male šanse da je hipoteza tačna.

Jerži Nejman (1894-1981) i Igon Pirson (1895-1980) malo kasnije su zajedno objavili rešenje za testiranje hipoteza koje se sastoji od niza koraka pre izvršenja opita. Pojednostavljeno, naučnik definiše dve hipoteze i verovatnoće pogrešnog prihvatanja za obe hipoteze. Te verovatnoće ne moraju biti iste i njihove vrednosti zavise od slučaja do slučaja. Onda u zavisnosti od toga kako su hipoteze definisane i vrednosti verovatnoća greške, uz pomoć određenih formula, izračuna se potrebna veličina uzorka u eksperimentu. I tek tada se kreće s izvršenjem opita. Kad se završi eksperiment, a to je kad smo dobili uzorak tražene veličine, onda se lako uvidi koju hipotezu prihvatamo.

Fišer je žestoko kritikovao Nejmanov i Pirsonov pristup. Smatrao je da možemo samo da govorimo o p-vrednosti jedne hipoteze. S jednim eksperimentom ne možemo da govorimo o verovatnoći pogrešnog prihvatanja ove ili one hipoteze. Za tako nešto je potrebno više puta da se ponove eksperimenti. S druge strane, u slučajevima kad je potrebno da se odluči između dve hipoteze, kao na primer između nove ili stare terapije, praktična vrednost Nejmanovog i Pirsonovog pristupa bila je očigledna.

Fišerov pristup je fleksibilan i izvršava se posle opita. On zastupa da može da postoji samo jedna hipoteza. Dok je Nejmanov i Pirsonov pristup krut i dve hipoteze se definišu pre nego što se mogu videti rezultati. S Fišerovim pristupom možemo da izračunamo p-vrednost hipoteze, dok kod Nejmana i Pirsona nema p-vrednosti, već samo odluka koju hipotezu prihvatamo i s kojom verovatnoćom greške. Fišerov pristup s jedne strane i Nejmanov i Pirsonov s druge dva su veoma vredna, a opet nekompatibilna pristupa testiranju statističkih hipoteza. Oba su ogroman doprinos nauci i naučnoj metodologiji.

Mišljenja sam da je uzrok današnjem ozbiljnom problemu, na koji ukazuju više od 800 svetskih naučnika, taj što se u većini univerzitetskih udžbenika statistike, i kod nas i u svetu, izlaže amalgam ova dva pristupa. Pojednostavljeno, posle eksperimenta definišu se dve hipoteze: Jedna da nema promene (nulta hipoteza) i druga da ima (alternativna hipoteza). Izračuna se p-vrednost nulte hipoteze i ako je manja od 5%, onda se odbacuje nulta hipoteza i prihvati alternativna. U tom slučaju kaže se da je dobijena razlika u eksperimentu statistički značajna. Dakle, ovde su Fišerov termin (p-vrednost) i Fišerov pristup (definisavanje hipoteza nakon izvršenog eksperimenta) pomešani s Nejmanovom i Pirsonovom metodom (prihvatanje jedne od dve hipoteze)

Fišer i Nejman, obojica velikani statistike 20. veka, nisu se slagali oko nekih pristupa u statistici, ali su bili strogo protiv ove pojednostavljene mešavine koja se od polovine prošlog veka počela pojavljivati u univerzitetskim udžbenicima statistike. Smatra se da je za popularizaciju ovog pojednostavljenog pristupa odgovoran E.F. Lindkvist (E.F. Lindquist) i njegov univerzitetski udžbenik iz 1940. godine *Statistical Analysis in Educational Research*. Koliko sam uspeo da proverim, u našim univerzitetskim udžbenicima preovladava ovaj pojednostavljen spoj bez pomena o bitnim razlikama u odnosu na ono što su zastupali Fišer, Nejman i Pirson.

Da ponovim. Za razumevanje problematičnosti ovog pojednostavljenog pristupa potreban je kurs statistike, ali je važno naglasiti da se on u ozbiljnoj praksi, kao na primer u kliničkim ispitivanjima, nigde ne koristi. Međutim, ova simplifikacija dominira u časopisima gde objavljuju eksperimentalni naučnici. Lakše je objaviti da ste pronašli nešto novo primenjujući pojednostavljen pristup, nego s razumevanjem svih finesa prilikom testiranja hipoteza. Stoga više od 800 svetskih naučnika poziva da se napusti upotreba statističke značajnosti i tako zaustavi talas objavljivanja naučnih pronalazaka koji to nisu.

Zašto se ne ažuriraju univerzitetski udžbenici statistike kad je testiranje hipoteza u pitanju? Zašto budući biolozi, lekari, ekonomisti, nažalost kod nas i statističari, kao i većina drugih naučnika, i dalje uče samo pojednostavljen pristup bez dubljeg razumevanja problematike? Ne znam. Teorija statističkih hipoteza se razvila od polovine prošlog veka do danas. Ima drugih rešenja i vodeći autori komentara u časopisu *Nature* ukazuju na neka od njih.

Statistika je oblast matematike i razumem da postoje očekivanja da ona ponudi nedvosmislena rešenja. Ne mislim da je pojednostavljivanje problematike testiranja statističkih hipoteza bilo motivisano nekom zlom namerom. Naprotiv, utisak mi je da je preovladala želja da se olakša primena, ali su se u tom procesu zanemarili ključni detalji. Nažalost, posledice takvog pojednostavljivanja po današnju nauku su ozbiljne, a odgovornost za to, po mom mišljenju, najviše snose oni koji podučavaju testiranje statističkih hipoteza kao da je šablon.

Reference koje podržavaju činjenice spomenute u članku

1. Više informacija možete saznati na [kursu razumevanja statističke značajnosti i p-vrednosti](#) koje za naše naučnike organizuje [Damjan Krstajić](#).