Testiranje statističkih hipoteza u *GeoGebri*

Aleksandra-Maria Vuković,
Gornji Mihaljevec

Uvod

Jedna od osnovnih zadaća statistike je na temelju uzorka ocijeniti kakvu razdiobu ima promatrano (populacijsko) statističko obilježje X.

Mnoge praktične situacije u vezi sa slučajnim pojavama zahtijevaju da se donesu odluke tipa DA ili NE. Primjerice pri praćenju procesa proizvodnje nekog proizvoda (uzmimo za primjer proces proizvodnje baterija) treba, na temelju rezultata mjerenja $x_1, \ldots, x_n$ statističkog obilježja X, donijeti odluku o tome osigurava li proces proizvodnje ili ne osigurava zahtijevanu kvalitetu. Pritom se pretpostavlja da obilježje X, koje karakterizira kvalitetu pojedinog proizvoda (npr. trajnost baterije) ima slučajni karakter.

Drugim riječima, na temelju n mjerenja slučajne varijable X tj. na temelju vrijednosti $(x_1, \ldots, x_n)$ slučajnog uzorka $(X_1, \ldots, X_n)$ donosimo odluku o prihvaćanju (DA) ili odbacivanju (NE) određene pretpostavke o svojstvima slučajne varijable X. Takva pretpostavka zove se **statistička hipoteza**, a provjera istinitosti te hipoteze tj. postupak donošenja odluke o prihvaćanju ili odbacivanju statističke hipoteze zove se **testiranje** (**statistički test**). Hipotezu koju testiramo zovemo **nulta hipoteza ili nul-hipoteza** i obilježavamo je sa H_0. **Alternativnu hipotezu** obilježavamo sa H_1.

Statistički test može biti:

- **parametarski**: testiramo hipotezu koja se odnosi na parametar pretpostavljene razdiobe
- **neparametarski**: testiramo hipotezu koja se odnosi na tip pretpostavljene razdiobe.

Hipoteza je:

- **jednostavna** ako jednoznačno određuje razdiobu statističkog obilježja X
- **složena** ako jednoznačno ne određuje razdiobu statističkog obilježja X.

Pritom je važno naglasiti da niti jedan statistički zaključak o populaciji na bazi uzorka nije stopostotno siguran. Tako i prihvaćanje neke hipoteze na temelju uzorka ne znači da je ta hipoteza točna. Zato je umjesto *"hipotezu prihvaćamo"* ispravnije reći *"na osnovi uzorka ne postoji razlog za odbacivanje hipoteze"*.

Aleksandra-Maria Vuković, prof., OŠ Gornji Mihaljevec, Macinec, amvukovic@gmail.com

Pogreške pri donošenju odluke

Prilikom donošenja odluke o istinitosti hipoteze postoje dvije vrste mogućih pogrešaka:

- **pogreška 1. vrste**: odbacili smo nultu hipotezu ako je ona istinita
- **pogreška 2. vrste**: prihvatili smo nultu hipotezu ako je ona neistinita.

Moguće situacije prikazane tablicom:

	H_0 je točna	H_0 je netočna
prihvaćamo H_0	✓	pogreška 2. vrste
odbacujemo H_0	pogreška 1. vrste	✓

Vjerojatnosti tih pogrešaka označavamo sa:

- $\alpha = P(\text{pogreška 1. vrste})$
  $= P(\text{odbacujemo } H_0 | H_0 \text{ točna})$
- $\beta = P(\text{pogreška 2. vrste})$
  $= P(\text{prihvaćamo } H_0 | H_0 \text{ netočna})$

Vjerojatnosti mogućih situacija prikazane tablicom:

	H_0 je točna	H_0 je netočna
prihvaćamo H_0	$1-\alpha$	β
odbacujemo H_0	α	$1-\beta$

α je **nivo signifikantnosti** ili **razina značajnosti**, a $1-\beta = P(\text{odbacujemo } H_0 | H_0 \text{ netočna})$ **snaga testa**.

Koraci pri testiranju hipoteza

Za testiranje hipoteze treba:

1) **definirati H_0 i H_1** (postavljanje hipoteza)
2) **izabrati uzorak** (definirati test-statistiku na osnovi čijih vrijednosti se donose odluke)
3) **izračunati testne veličine** (za zadanu razinu značajnosti α odrediti kritično područje tj. skup svih mogućih vrijednosti test-statistike za koje se odbacuje nulta hipoteza u korist alternativne)
4) **ispitati** je li vrijednost test-statistike izračunane iz uzorka u kritičnom području
5) **donijeti odluku** – prihvaćanje ili odbacivanje nulte hipoteze (ako je izračunana vrijednost test-statistike u kritičnom području nul-hipoteza se odbacuje u korist alternativne hipoteze, u suprotnom se nul-hipoteza prihvaća).

Primjeri testiranja statističkih hipoteza u *GeoGebri*

1. Test o očekivanju normalno distribuirane populacije

Primjer 1. *Dvosmjerni Z test srednjih vrijednosti*

Pretpostavlja se da je prosječno vrijeme studiranja na jednom studiju 3.5 godina. Utvrđeno je da je 150 studenata izabranih u uzorak $(f < 5\ \%)$ studiralo 3.45 godina s prosječnim odstupanjem od 0.5 godina. Potrebno je navedenu pretpostavku testirati sa 95 % pouzdanosti.

a) Postavljanje hipoteza

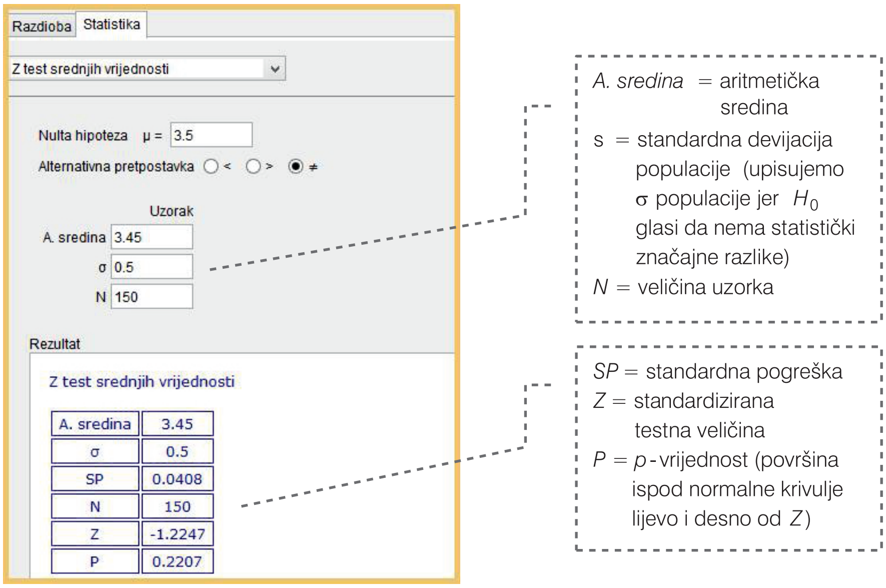
Dvosmjernim testom (engl. *two-tailed test*) se koristimo kada se unaprijed ne može sa sigurnošću odrediti smjer neke razlike.

Postavljamo nultu i alternativnu hipotezu, gdje nulta hipoteza glasi da je aritmetička sredina uzorka μ jednaka nekoj pretpostavljenoj aritmetičkoj sredini populacije μ_0, a alternativna hipoteza glasi da je aritmetička sredina uzorka različita od aritmetičke sredine populacije

$$H_0 : \mu = \mu_0 \implies H_0 : \mu = 3.5$$
$$H_1 : \mu \neq \mu_0 \implies H_1 : \mu \neq 3.5$$

b) Testiranje s pomoću kalkulatora vjerojatnosti u *GeoGebri*

Odabrat ćemo u *GeoGebri* izbornik *Pogled>Tablica*. Klikom u bilo koju ćeliju *tabličnog prikaza* aktiviraju se pripadni alati. Iz druge grupe alata odabrat



Slika 1. Kalkulator vjerojatnosti u *GeoGebri*[1]

ćemo alat *Kalkulator vjerojatnosti* > kartica *Statistika*. Potrebno je upisati odgovarajuće parametre.

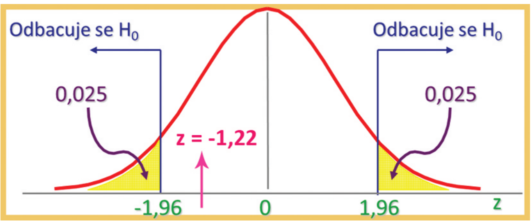
c) Interval prihvaćanja hipoteze

$P = 95\ \% \implies \alpha = 0.05 \implies$ dvosmjeran test $z_{\alpha/2} = z_{0.025} = 1.96$.

Ako je $-z_{\alpha/2} \leq Z \leq z_{\alpha/2} \implies$ ne možemo odbaciti H_0

Z-vrijednost je u području za prihvaćanje

$$-1.96 < Z(-1.22) < 1.96.$$



Slika 2. Kritična područja na rubu raspodjele kod dvosmjernog Z testa srednjih vrijednosti

Odluka: prihvaća se hipoteza za $\alpha = 0.05$.

Zaključak: Nema dokaza da prosječno vrijeme studiranja nije 3.5 godine.

Primjer 2. *Jednosmjerni Z test srednjih vrijednosti*

Poznato je da napon u električnoj mreži od 220 volti ima normalnu distribuciju sa standardnom devijacijom od 6 volti. Ako je 16 nezavisnih mjerenja dalo rezultate:

$$208, 216, 215, 228, 210, 224, 212, 213,$$
$$224, 218, 206, 209, 208, 218, 220, 206$$

s razinom značajnosti 0.01 provjerite pretpostavku da je došlo do pada srednjeg napona u električnoj mreži.

a) Postavljanje hipoteza

Jednosmjerni test (engl. *one-tailed test*) primjenjuje se kada je smjer efekta specificiran u alternativnoj

[1] Ako je veličina uzorka manja od 30 i/ili je nepoznata standardna devijacija populacije, umjesto Z-vrijednosti se računa t-vrijednost iz tablice površina Studentove ili t-distribucije za $n-1$ stupnjeva slobode.

hipotezi.

$$H_0 : \mu = \mu_0 \implies H_0 : \mu = 220$$
$$H_1 : \mu < \mu_0 \implies H_1 : \mu < 220$$

b) Testiranje s pomoću kalkulatora vjerojatnosti u *GeoGebri*

Aritmetičku sredinu uzorka možete izračunati ili izradom liste od danih rezultata mjerenja i upisom odgovarajuće naredbe u *Traku za unos* algebarskog prikaza ili se koristite *Tabličnim prikazom* i odgovarajućim alatom odnosno naredbom.

S pomoću *Kalkulatora vjerojatnosti* testirajte hipotezu da je došlo do pada srednjeg napona na električnoj mreži (*Z test srednjih vrijednosti*).

c) Interval prihvaćanja hipoteze i rješenje

Kritično područje na rubu raspodjele kod jednosmjernog Z testa "manje od" je cijelo na lijevoj strani, pa

H_0 odbacujemo ako je $Z < -z_\alpha$.

Kako riješiti problem ako nemamo tablicu kritičnih vrijednosti parametra Z, pa ne možemo izračunati vrijednost z_α i usporediti je s dobivenom Z-vrijednošću? *GeoGebra* i za to ima rješenje jer se automatski izračunava i p-vrijednost[2] (vidi objašnjenje uz prethodni primjer), samo što kod jednosmjernog testa dobivena p-vrijednost predstavlja površinu ispod normalne krivulje lijevo (ako je $H_1 : \mu < \mu_0$) odnosno desno (ako je $H_1 : \mu > \mu_0$) od z-vrijednosti. Za p-vrijednost vrijedi:[3]

H_0 odbacujemo ako je $P < \alpha$.

Rješenje. Aritmetička sredina $= 214.69$, $H_0 : \mu = 220$, $H_1 : \mu < 220$, $\alpha = 0.01$, $Z = -3.54$, $P = 0.0002$.

Kako je $P < \alpha$ $(0.0002 < 0.01)$, odbacujemo nultu hipotezu tj. došlo je do pada napona.

Primjer 3. *Dvosmjerni T test srednje vrijednosti*

Tvornica tvrdi da je prosječna čvrstoća njezinih bakrenih kablova 350 kg. Služba kontrole kvalitete obavila je ispitivanje čvrstoće na slučajnom uzorku od 10 kablova te su dobivene vrijednosti (u kg)

$$360, 370, 385, 367, 355, 365, 375, 330, 342, 320.$$

Može li se prihvatiti tvrdnja da je prosječna čvrstoća kablova 350 kg? Razina signifikantnosti je 5 %.

a) Postavljanje hipoteza

Ranije je naglašeno da se u slučaju malog broja podataka i/ili nepoznate standardne devijacije populacije, pri testiranju hipoteze o sredini populacije, provodi t-test (u *GeoGebri T test srednje vrijednosti*). U tom slučaju se standardna devijacija populacije σ zamjenjuje standardnom devijacijom uzorka s.

$$H_0 : \mu = \mu_0 \implies H_0 : \mu = 350$$
$$H_1 : \mu \neq \mu_0 \implies H_1 : \mu \neq 350$$

b) Izračun testnih veličina i testiranje s pomoću kalkulatora vjerojatnosti u *GeoGebri*

U *Tabličnom prikazu* u *GeoGebri* treba izračunati potrebne vrijednosti i testirati hipotezu s pomoću *Kalkulatora vjerojatnosti* (*T test srednje vrijednosti*).

c) Interval prihvaćanja hipoteze

Ako je $-t_{\alpha/2}<t<t_{\alpha/2} \implies$ ne možemo odbaciti H_0.

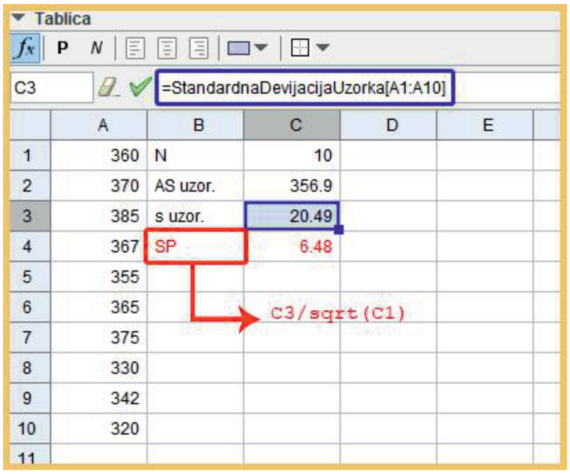
Kako je $t_{0.025}(9) = 2.26$, a testna veličina je $t = 1.065$, prihvaća se hipoteza za $\alpha = 0.05$. Nema dokaza da prosječna čvrstoća nije 350 kg.

Napomena. Kao i kod Z testa srednjih vrijednosti, možemo primijeniti alternativno pravilo odlučivanja na temelju p-vrijednosti pri čemu H_0 odbacujemo ako je $P < \alpha$.

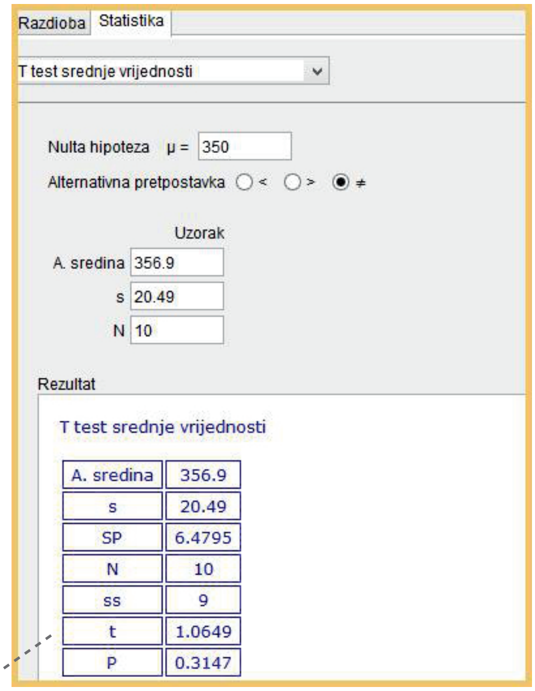
---

[2] Oznaka p-vrijednosti u Kalkulatoru vjerojatnosti u *GeoGebri* je P.

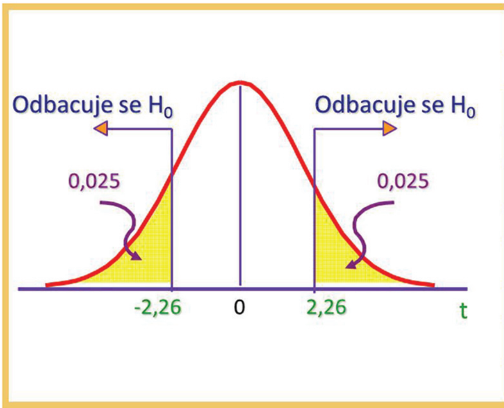
[3] Ovo je tzv. alternativno pravilo odlučivanja, odlučivanje na temelju p-vrijednosti ili empirijske razine signifikantnosti.



Slika 3. Kalkulacije u proračunskoj tablici *GeoGebre*



s = standardna devijacija uzorka
SP = standardna pogreška
ss = broj stupnjeva slobode = $N-1$
t = standardizirana testna veličina



jednosmj.	0,05	0,025	0,005
dvosmj.	0,10	0,05	0,01
$\varphi = 1$	6,314	12,706	63,657
2	2,920	4,303	9,925
3	2,353	3,182	5,841
....	....	....	....
9	1,83	2,26	3,25
10	1,81	2,23	3,17
11	1,80	2,20	3,11
12	1,78	2,18	3,06

Slika 4. Kritična područja na rubu raspodjele kod dvosmjernog T testa srednje vrijednosti

2. Testiranje hipoteze o razlici aritmetičkih sredina dvaju skupova

Razlikujemo nekoliko tipova ove vrste testiranja hipoteze, ovisno o **vrsti** (nezavisni ili zavisni uzorci) i **veličini uzoraka**. Tipični primjeri dvaju nezavisnih uzoraka su npr. *dvije srednje škole, dva razredna odjela, mlađi i stariji ispitanici* itd. Uobičajeni primjer zavisnih uzoraka je *mjerenje prije i poslije* kod npr. provjeravanja djelovanja neke terapije ili pokusa i sl. Veličina uzorka određuje način izračunavanja testnih veličina.

I ovdje se koristimo Z testom i Studentovim t-testom ovisno o tome imamo li standardne devijacije uzoraka ili ne.

Primjer 4. *Testiranje hipoteze o razlici ... dvaju nezavisnih uzoraka, t-test*

Zanima nas postoji li razlika u visini između mladića i djevojaka u prvim razredima jedne srednje škole. Mjerenjem visina 40 mladića i 54 djevojke dobili smo sljedeće rezultate:

$$\overline{X}_m = 177.5 \text{ cm}, \quad s_m = 9.1 \text{ cm};$$
$$\overline{X}_d = 163.6 \text{ cm}, \quad s_d = 8.4 \text{ cm}.$$

Je li razlika između ovih dviju skupina statistički značajna?

a) Postavljanje hipoteza

Nulta hipoteza: mladići i djevojke ne razlikuju se značajno u visini. Alternativna hipoteza: mladići su viši od djevojaka.

$$H_0 : \mu_m \leq \mu_d$$
$$H_1 : \mu_m > \mu_d$$

Ovdje se zapravo radi o jednosmjernom testu na gornju granicu tj. testu "veće od", pa hipoteze možemo formulirati i na sljedeći način:

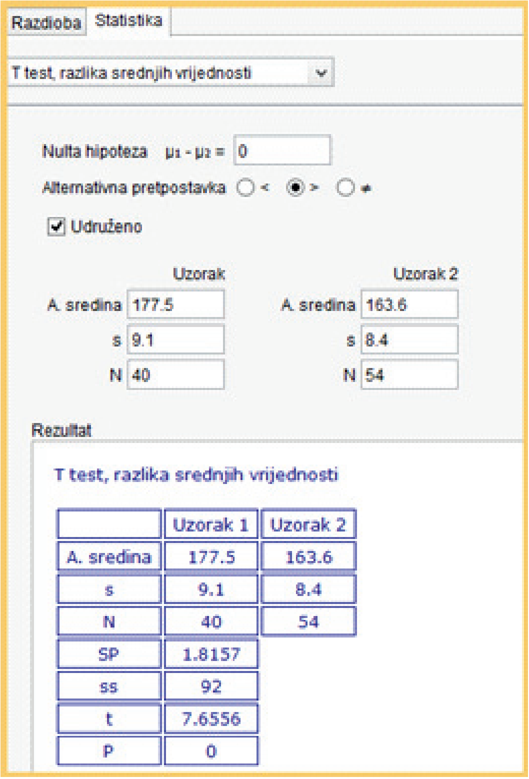
$$H_0 : \mu_m - \mu_d \leq 0$$
$$H_1 : \mu_m - \mu_d > 0$$

b) Testiranje s pomoću kalkulatora vjerojatnosti u *GeoGebri*

Napomena. U svakodnevnoj praksi najčešće imamo niz neobrađenih podataka (rezultate nekog mjerenja). Vrlo lako možemo prenijeti te podatke u *Tablični prikaz GeoGebre* (*"kopiraj-zalijepi"*) i korištenjem ponuđenih naredbi izračunati sve potrebne parametre.

c) Interval prihvaćanja hipoteze

Ako je $t_{\text{izračunano}} < t_\alpha(ss) \implies$ ne možemo odbaciti H_0 gdje je ss broj stupnjeva slobode i $ss = N_1 + N_2 - 2$.



Slika 5.

Uz razinu signifikantnosti od 5 % imamo:

$\alpha = 0.05$, uz $ss = 92$ je granična tablična vrijednost $t_\alpha(ss) = t_{0.05}(92) = 1.99$.

Odluka: $t = 7.66 > t_{0.05}(92) = 1.99$, pa odbacujemo nultu hipotezu i prihvaćamo alternativnu.

Zaključak: Razlika od $177.5 - 163.6 = 13.9$ cm između prosječne visine mladića i djevojaka je statistički značajna.

Zaključak

Implementiranjem proračunske tablice, statističkih alata i naredbi *GeoGebra* se svrstala među vrlo moćne programe za statističku analizu – kako elementarnu, tako i višu. To se posebice odnosi na *Kalkulator vjerojatnosti* koji omogućava relativno jednostavno i brzo testiranje statističkih hipoteza. Komparativna prednost ovog alata je što se vrlo lako, bez korištenja bilo kakvih naredbi i u samo nekoliko klikova, dolazi do referentnih vrijednosti za testiranje hipoteze. Dovoljno je upisati osnovne parametre potrebne za pojedino testiranje, a *GeoGebra* sama računa odgovarajuće testne veličine. Osim toga, ako i raspolažemo neobrađenim podatcima, njihov prijenos u proračunsku tablicu *GeoGebre* je vrlo jednostavan, a na raspolaganju je niz različitih statističkih naredbi za izračun potrebnih parametara.

U ovom članku su predstavljena tek tri statistička testa od preko deset koliko ih je na raspolaganju u *GeoGebrinu Kalkulatoru vjerojatnosti*. Odabrala sam one testove koji su najbliži našoj svakodnevnoj odgojnoobrazovnoj praksi: vrlo često se pitamo odstupaju li rezultati testiranja u nekom razrednom odjelu statistički značajno od pretpostavljene normalne distribucije ili uspoređujemo rezultate testiranja u dva paralelna razredna odjela, pitajući se jesu li uočene razlike statistički značajne ili ne...

Uporabom *GeoGebre* i njezinog alata *Kalkulator vjerojatnosti* ne samo da ćete si olakšati statistička testiranja, već je to i dobra prilika da ponovite statističke sadržaje.

LITERATURA

1/ Ž. Pauše, *Uvod u matematičku statistiku*, Školska knjiga, Zagreb, 1993.

2/ M. Benšić, N. Šuvak, *Uvod u vjerojatnost i statistiku, nastavni materijal za kolegij Statistički praktikum*, Sveučilište J.J. Strossmayera, Odjel za matematiku, `http://www.mathos.unios.hr/uvis/UVIS_knjiga_final/UVIS_knjiga_web.pdf`, 2014.

Poštovani,

Matematičko društvo "Istra" organizira 6. svibnja 2016. već tradicionalni 11. **Festival matematike** u hotelu Brioni u Puli. Festival okuplja veliki broj učenika i mentora osnovnih i srednjih škola Republike Hrvatske. Prošle 2016. godine nastupilo je više od 600 sudionika i pedesetak mentora te dvadesetak organizatora, pa i ove godine očekujemo barem isti broj.

Na Festivalu matematike održava se ekipno natjecanje Ekipa za 5+, smotra učeničkih radova projekata Matematika+ i Pula+. Detaljnije o projektima možete pogledati na stranicama MDI,

`http://md-istra.blogspot.com/` i `http://www.mdi.hr/`

Pozivamo vas da upotpunite Festival predavanjima za mentore uz predstavljanje udžbeničkih i ostalih izdanja te da nam, donirate nagrade učenicima natjecateljima u obliku knjiga, mapa, blokova i sličnog promotivnog materijala. Mjesto za prodaju ćemo vam osigurati bez naknade.

Srdačan pozdrav,

predsjednik Matematičkog društva "Istra"
Robert Gortan, prof.