



Poglavlje 5.

Statističko testiranje hipoteza

dr Đorđe Nikolić, redovni profesor
E-mail: djnikolictfbor@gmail.com

30.Oktobar 2023. godine

Ciljevi poglavlja



- Testiranje statističkih hipoteza- osnovne napomene,
- Parametarski vs. Neparametarski statistički testovi,
- Testiranje statističkih hipoteza- parametarski testovi:
 - Testiranje statističkih hipoteza zasnovano na jednom uzorku,
 - Testiranje statističkih hipoteza za dva nezavisna uzorka;
 - Testiranje statističkih hipoteza za dva zavisna uzorka,
 - Analiza varijanse sa jednim faktorom (ANOVA),
 - Višefaktorska analiza varijanse (dvofaktorska analiza varijanse).
- Testiranje statističkih hipoteza- neparametarski testovi:
 - Kolmogorov-Smirnov test,
 - Hi-kvadrat test,
 - Man-Vitnijev U test,
 - Vilkoksonov test ranga.

Preporučena literatura



- Radojević, S, Veljković Z, Kvantitativne metode, CD, MF,
- Montgomery, DC, Runger, GC Applied Statistics and Probability for Engineers, Fourth Edition, Wiley, 2007.

Koncept testiranja statističkih hipoteza

- **Hipoteza** je prepostavka zasnovana na raspoloživim saznanjima.
- Provera tačnosti (istinitosti) statističke hipoteze naziva se **verifikacija hipoteze**, a naučni metodi verifikacije nazivaju se **statističkim testovima**.
- U zavisnosti od strogosti polaznih uslova statistički testovi mogu biti **parametarski** ili **neparametarski**.

Koncept testiranja statističkih hipoteza

• Nulta i alternativna hipoteza

- Nulta i alternativna hipoteza su precizna i međusobom isključna tvrđenja (pretpostavke) o vrednostima parametra θ .
- Statističku hipotezu formulišemo u obliku **nulte hipoteze (H_0)**, koja predstavlja tvrđenje o vrednostima parametra čija se istinitost proverava.
- Svaka nulta hipoteza ima svoju suprotnu hipotezu, koja se naziva **alternativna hipoteza (H_1)**. Ona sadrži sve vrednosti parametra θ , koje nisu obuhvaćene nultom hipotezom i ima za cilj da ospori valjanost nulte hipoteze.

Koncept testiranja statističkih hipoteza

• Statistički testovi

- U praksi se stvarna vrednost ocenjuje na osnovu statistike uzorka.
- Može se javiti razlika između pretpostavljene vrednosti parametra θ_0 i **ocenjene vrednosti** parametra u uzorku θ' .
- U postupku verifikacije nulte hipoteze ispituje se statistička značajnost njihove razlike, tako da se testovi, koji se primenjuju u tu svrhu, nazivaju **testovima statističke značajnosti**.
- Ocenuje se da li je razlika između θ' i θ_0 statistički značajna ili je rezultat fluktuacije u uzorku (slučajna greška).

• Vrste statističkih testova

- Z-test (normalna raspodela)
- t-test (Student-ova t raspodela)
- χ^2 - test (Hi-kvadrat raspodela)
- F-test (Fisher-ova raspodela)

Koncept testiranja statističkih hipoteza

Greške testiranja statističkih hipoteza

- Odbacivanje istinite (tačne) nulte hipoteze je pogrešno, a greška se naziva **greškom prve vrste** (Type I error). Verovatnoća da će se odbaciti istinita nulta hipoteza, naziva se **rizikom prve vrste** (α -nivo značajnosti testa).
- Kada se prihvati netačna nulta hipoteza pravi se **greška druge vrste** (Type II error). Verovatnoća da se ne odbaci netačna nulta hipoteza, naziva se **rizikom druge vrste** (β).
- Verovatnoća da se ne učini greška druge vrste naziva se **moć testa** (γ):

$$\beta + \gamma = 1 \rightarrow \gamma = 1 - \beta$$

Odluka	Aktuelna situacija	
	H_0 je tačna	H_0 je netačna
Prihvatanje H_0	Tačno	Greška druge vrste
Odbacivanje H_0	Greška prve vrste	Tačno

Koncept testiranja statističkih hipoteza

- **Testiranje testa se vrši preko statistike testa:**

$$\text{statistika_testa} = \frac{\theta - \theta_0}{\sigma_\theta / \sqrt{n}}$$

- Uvek se polazi da je razlika između prepostavljene vrednosti skupa i ocenjene vrednosti parametra, slučajna (pod dejstvom slučajnih faktora), tj. da je nulta hipoteza tačna.
- Razlika pod dejstvom slučajnih faktora je dopustiva do određenih granica i meri se standardnom greškom ocene (engl. Standard error of estimate S.E. = σ / \sqrt{n}).
- Značajne razlike nastale pod dejstvom sistematskog faktora prelaze ove granice. Pri tome, verovatnoću greške prve vrste α (nivo značajnosti testa), proizvoljno biramo unapred. Na primer, $\alpha=0.5$ definiše da je 5 % mogućnost da se donese pogrešan zaključak i da se odbaci istinita nulta hipoteza H_0 .

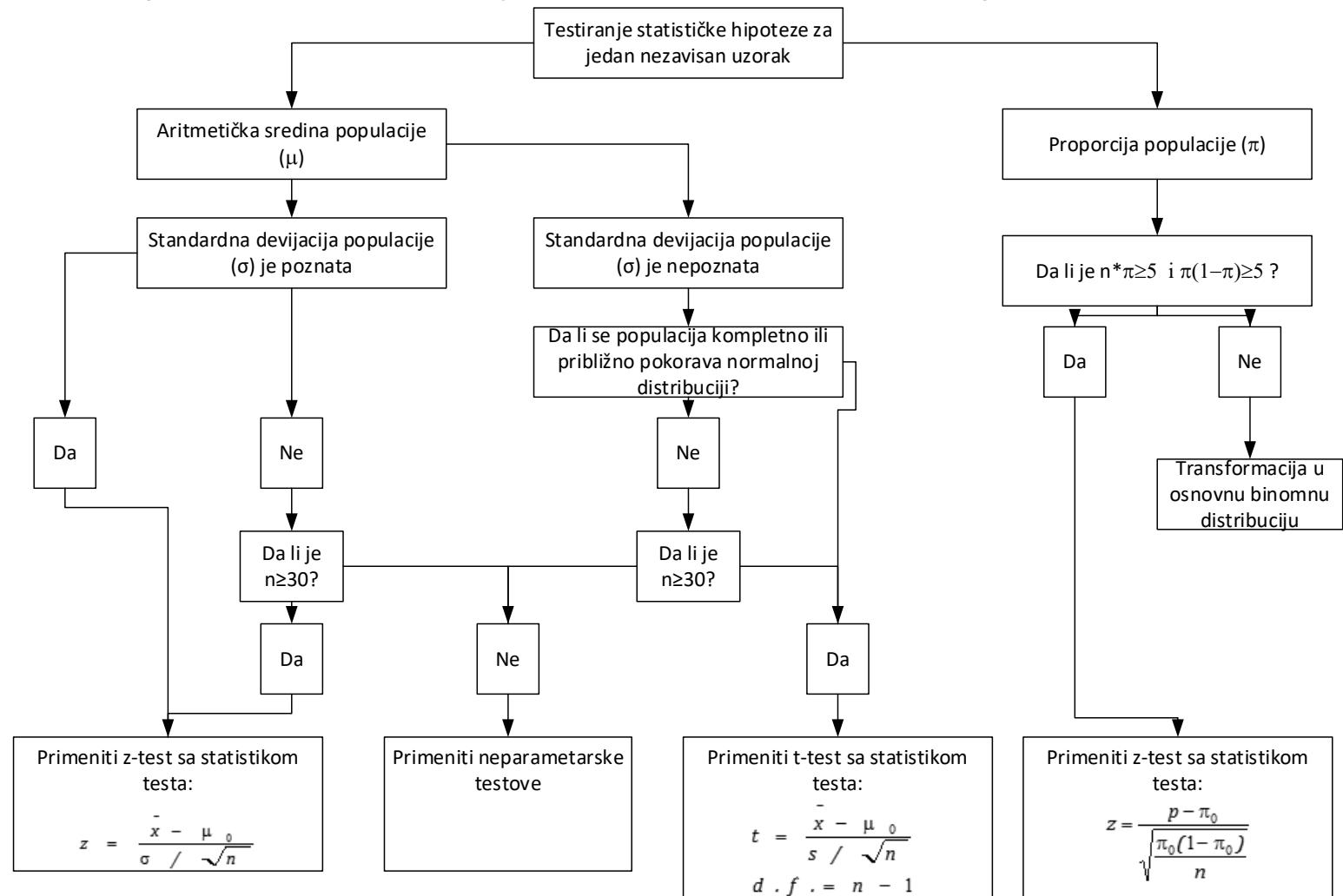
Koncept testiranja statističkih hipoteza

• Procedura testiranja statističke hipoteze:

1. Definisanje nulte i alternativne hipoteze,
2. Izbor nivoa značajnosti testa α i određivanje kritičnih vrednosti,
3. Izbor statistike testa i određivanje njegovog rasporeda verovatnoće,
4. Formulisanje pravila odlučivanja o prihvatanju, odnosno, odbacivanju nulte hipoteze,
5. Izvlačenje uzorka iz osnovnog skupa i izračunavanje vrednosti statistike testa,
6. Donošenje odluke o prihvatanju, odnosno, odbacivanju nulte hipoteze.
7. Opciono: Analiza osetljivosti rezultata, promenom parametara koraka 2.

Testiranje statističkih hipoteza- parametarski testovi

Testiranje statističkih hipoteza zasnovano na jednom uzorku



Testiranje statističkih hipoteza zasnovano na jednom uzorku

- Oblici nulte i alternativne hipoteze:

- 1) Prost slučaj (obostrani- two tail):

$H_0: \theta = \theta_0$ (θ_0 – jedna tačno određena vrednost)

$H_1: \theta \neq \theta_0$

Primer:

$H_0: \mu = 10\$$

$H_1: \mu \neq 10\$$

(tumačenje: srednja vrednost μ je 10\$ ili nije)



Testiranje statističkih hipoteza zasnovano na jednom uzorku

- Oblici nulte i alternativne hipoteze:

2) Složen slučaj (levi-one tail):

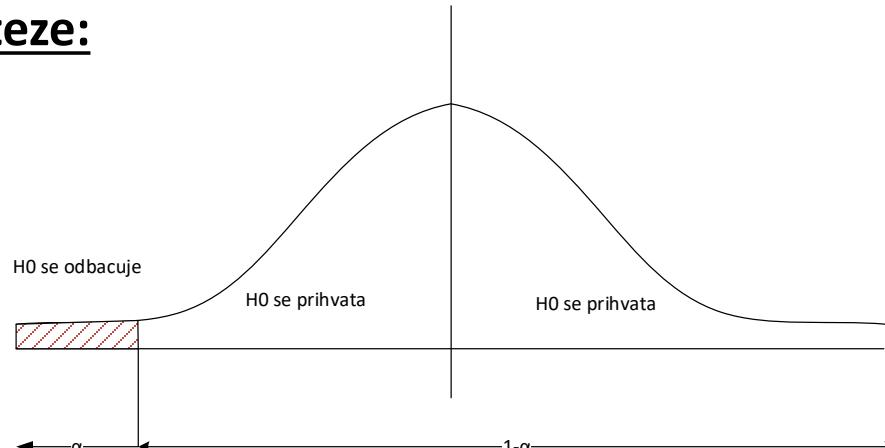
$$H_0: \theta \geq \theta_0$$

$$H_1: \theta < \theta_0$$

Primer:

$$H_0: \mu \geq 10\$$$

$$H_1: \mu < 10\$$$



(tumačenje: srednja vrednost μ je najmanje 10\$ ili je manje)

3) Složen slučaj (desni-one tail):

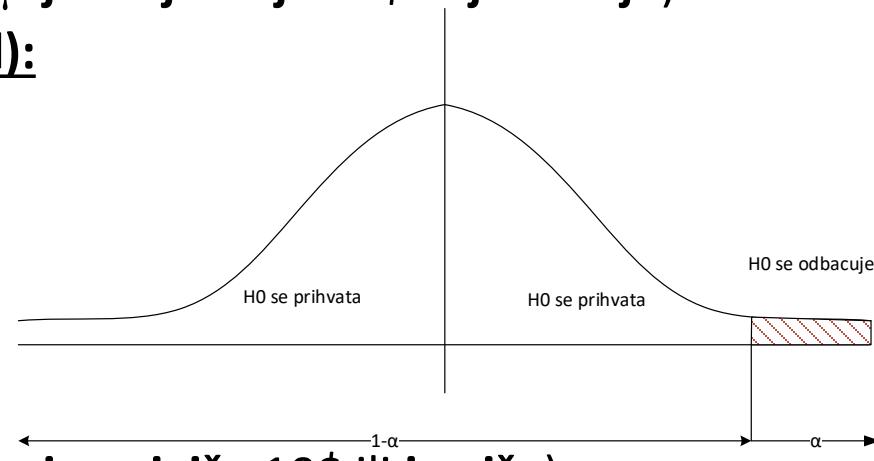
$$H_0: \theta \leq \theta_0$$

$$H_1: \theta > \theta_0$$

Primer:

$$H_0: \mu \leq 10\$$$

$$H_1: \mu > 10\$$$

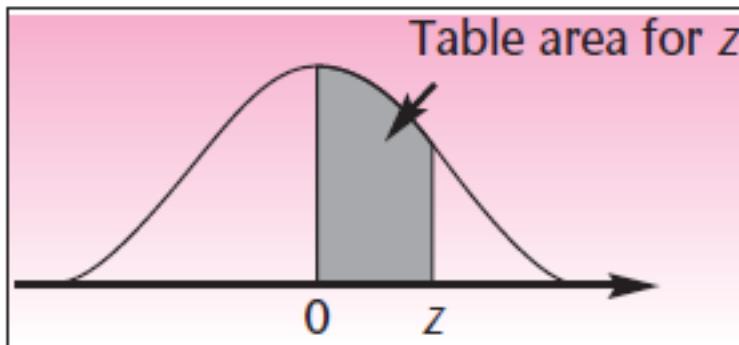


(tumačenje: srednja vrednost μ je najviše 10\$ ili je više)

Normalna distribucija-tabelarne vrednosti

Verovatnoće

Ukupna površina ispod normalne distribucije je 1. Možemo koristiti tabele da pronađemo oblasti verovatnoće za određene z-vrednosti



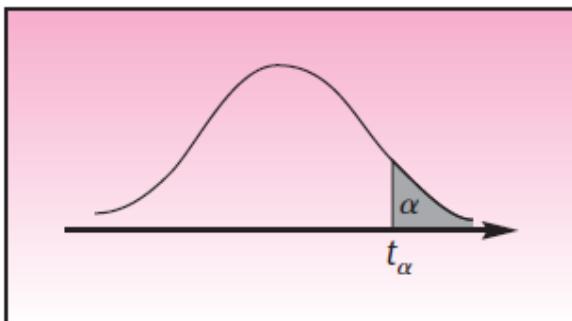
Označena površina za vrednost $z \leq 1$ iznosi 34.13%. To znači da je verovatnoća za dobijanje vrednosti manje od 1σ iznad aritmetičke sredine 0.3413.

Takođe, moglo bi se reći da je procenat distribucije između $z = -1$ i $z = +1$ iznosi 68.26%

The table areas are probabilities that the standard normal random variable is between 0 and z .

z	Second Decimal Place in z									
	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990

Student-ova t distribucija-tabelarne vrednosti



Degrees of Freedom	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
40	1.303	1.684	2.021	2.423	2.704
60	1.296	1.671	2.000	2.390	2.660
120	1.289	1.658	1.980	2.358	2.617
∞	1.282	1.645	1.960	2.326	2.576

Rezime poglavlja



- Pitanja?
- Diskusija!