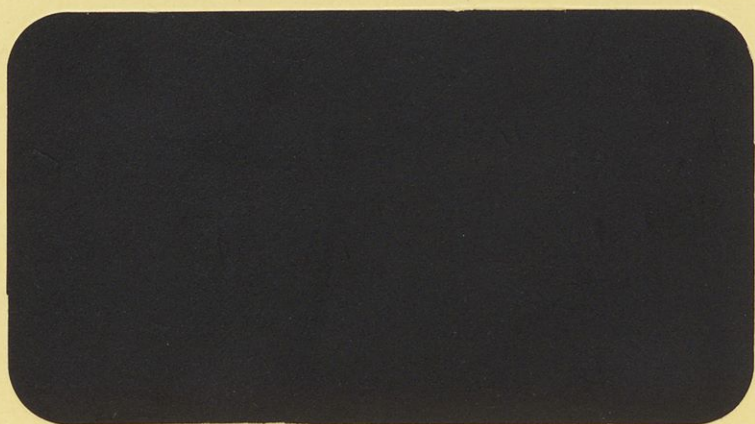


NOMOGRAM ZA DOLOCANJE INTERVALNIH
OCEN IN ZA PRESKUSANJE HIPOTEZ O
KORELACIJSKIH KOEFICIENTIH

Dr. MARIJAN BLEJEC

INSTITUT ZA STATISTIKO IN OPERACIJSKO RAZISKOVANJE
EKONOMSKA FAKULTETA
UNIVERZA V LJUBLJANI



NOMOGRAM ZA DOLOCANJE INTERVALNIH
OCEN IN ZA PRESKUSANJE HIPOTEZ O
KORELACIJSKIH KOEFICIENTIH

Dr. MARIJAN BLEJEC

Dr. Marijan BLEJEC

Nomogram za določanje intervalnih ocen in za preskušanje
hipotez o korelacijskih koeficientih

0. Teoretične osnove Za določanje mej zaupanja za intervalne ocene korelacijskih koeficientov in za preskušanje hipotez o korelacijskih koeficientih uporabljamo znano Fisherjevo transformacijo

$$Z = \frac{1}{2} \ln \frac{1+r}{1-r} \quad /1/$$

Z se porazdeljuje asimptotično normalno: $Z = N \left[E(Z) = Z_0 ; \text{Var}(Z) = \frac{1}{n-3} \right]$ Pri tem je: $Z_0 = \frac{1}{2} \ln \frac{1+r_0}{1-r_0}$, r_0 = prava vrednost korelacijskega koeficienta, n = število enot v vzorcu /če ocenjujemo parcialne korelacijske koeficiente je n število enot v vzorcu zmanjšano za število eliminiranih znakov/.

Če upoštevamo zgornje zakonitosti je intervalna ocena za Z_0 s tveganjem $\alpha = 2P$ dana z neenačbo

$$Z_s = Z - \frac{z_P}{\sqrt{n-3}} < Z_0 < Z_z = Z + \frac{z_P}{\sqrt{n-3}} \quad /2/$$

Enostranska razmaka zaupanja pa sta s tveganjem $\alpha = P$ dana z neenačbama

$$Z_s = Z - \frac{z_P}{\sqrt{n-3}} < Z_0 \quad \text{in} \quad Z_z = Z + \frac{z_P}{\sqrt{n-3}} > Z_0 \quad /3/$$

Poleg znanih ali že definiranih simbolov pomeni z_P vrednost standardiziranega normalnega znaka za dani P .

Meje zaupanja za r_0 (r_s in r_z) dobimo iz Z_s in Z_z z obratno transformacijo.

Isto transformacijo 1 uporabljamo tudi za preskušanje ničelne hipoteze $H_0: (r_H = r_0)$ in vzamemo za osnovo preskušanja izraz

$$(Z - Z_H) \lesseqgtr \frac{Z_P}{\sqrt{n-3}} \quad /4/$$

Za preskušanje ničelne hipoteze o neodvisnosti med dvema znakoma $H_0: (r_H = r_0 = 0)$ dobi izraz 4 posebno obliko

$$Z \lesseqgtr \frac{Z_P}{\sqrt{n-3}} \quad /5/$$

Ničelno hipotezo o enakosti korelacijskih koeficientov za dve populaciji $H_0: (r_{01} = r_{02})$ preskušamo z vzorcema s številom enot n_1 in n_2 z obrazcem

$$(Z_2 - Z_1) \lesseqgtr Z_P \sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}} \quad /6/$$

V posebnem primeru, če sta oba vzorca enako velika ($n_1 = n_2 = n$) pa dobimo iz 6 obrazec

$$\frac{Z_2 - Z_1}{\sqrt{2}} \lesseqgtr \frac{Z_P}{\sqrt{n-3}} \quad /7/$$

1. Konstrukcija nomograma V obrazcih v odstavku 0 so $Z(r)$ in izrazi $z_P/\sqrt{n-3}$ vezano aditivno. Zato moremo z njimi konstruirati nomogram za reševanje problemov o ocenjevanju in preskušanju hipotez o korelacijskih koeficientih. Fiksni del nomograma FS ima niz skal n_p , ki so proporcionalne izrazom $z_P/\sqrt{n-3}$ za določene verjetnosti P . Skale se nanašajo na naslednje vrednosti za P : $P=0,25$; $P=0,05$; $P=0,025$; $P=0,01$; $P=0,005$; $P=0,001$; $P=0,0005$. S kombiniranjem teh skal dobimo

osnovo za ocenjevanje mej zaupanja in preskušanje hipotez za vse konvencionalne stopnje tveganja za enostranske in dvostranske sklepe. Tako npr. s skalo $n_{P=0,25}$ med drugim ocenjujemo razmak zaupanja PE, v katerem je prava vrednost z verjetnostjo $1-2P = 0,50$. Vse skale n_p uporabljamo za enostranske zaključke na sedmih nivojih $\alpha = P$ in na sedmih nivojih $\alpha = 2P$. Skale n_p so od centralne linije za $n = \infty$ nanešene simetrično navzgor in navzdol, tako da moremo določati odprte ali zaprte razmake zaupanja ali kritične razmake. Na fiksnem delu nomograma FS je včrtana tudi linearna skala za Z. Nad centralno črto nomograma $n = \infty$ so preko n_p skal včrtane pod kotom 45° pomožne črte. Te upoštevamo pri preskušanju značilnosti razlik med korelacijskima koeficientoma, če sta izračunana iz različno velikih vzorcev n_1/n_2 . Ob enem robu prema lživemu dela nomograma FS je narisana skala PSr za r v razmerju s transformiranim znakom Z, ob drugem robu pa skala PS Δ r v razmerju z $Z/\sqrt{2}$. Ta skala je označena z Δ , ker služi za preskušanje značilnosti razlik med dvema korelacijskima koeficientoma ocenjenima iz enako velikih vzorcev. V sredini premakljivega dela nomograma PS je shematično prikazano reševanje standardnih problemov ocenjevanja in preskušanja hipotez o korelacijskih koeficientih z nomogramom.

2. Primeri uporabe nomograma V nadaljevanju so v devetih primerih prikazani standardni problemi, ki jih moremo rešiti z nomogramom.

Problem 1 Določanje dvostranskih mej zaupanja s tveganjem $\alpha = 2P$ za oceno korelacijskega koeficienta r iz vzorca z n enotami.

Rešitev: Oceno korelacijskega koeficienta r na premakljivi skali PSr postavimo ob fiksno skalo FSn za ustrezni $\alpha = 2P$ tako, da se sklada z $n = \infty$. Za dano velikost vzorca n na zgornjem in spodnjem delu skale FSn odberemo na PSr spodnjo (r_s) in zgornjo (r_z) mejo zaupanja za r_0 .

Primer: Iz vzorca $n = 50$ enot smo dobili oceno $r = 0,75$. Določiti je treba meje zaupanja s tveganjem $\alpha = 2P = 0,05$ za intervalno oceno korelacijskega koeficienta. Dvostranskemu razmaku zaupanja s tveganjem $\alpha = 2P = 0,05$ ustreza tretja skala FSn.-C. Ako $r=0,75$ na PSr vskladimo z $n = \infty$ na FSn-C, dobimo za $n=50$ na zgornjem in spodnjem delu PSr ustrezni meji zaupanja $r_s=0,60$ in $r_z = 0,85$.

Problem 2 Določanje enostranskih razmakov zaupanja za r_0 .

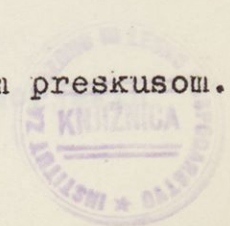
Rešitev: Spodnjo ali zgornjo mejo enostranskih razmakov zaupanja iz ocen r iz vzorcev velikosti n s tveganjem $\alpha = P$ dobimo podobno kot v problemu 1. le da vzamemo v poštev FSn za $\alpha = P$ in da odberemo ustrezno problemu samo spodnjo ali zgornjo mejo.

Primer: Iz vzorca z $n = 30$ enot smo dobili oceno $r = 0,33$.

Nad katero vrednostjo s tveganjem $\alpha = P = 0,05$ ni r_0 ? $r=0,33$ na PSr postavimo ob $n = \infty$ na skali FSn-B za $\alpha = P = 0,05$. Ob $n = 30$ na spodnjem delu FSn-B odberemo $r_z = 0,58$.

Problem 3 Preskus značilnosti razlik ocene r od hipotetične vrednosti r_H .

Rešitev: Značilnost razlike testiramo z dvostranskim preskusom.



Zato r_H na F_{Sn} postavimo ob $n = \infty$ na skali F_{Sn} za $\alpha = 2P$. Ako iz vzorca ocenjena vrednost r leži v razmaku med ustreznim n na zgornjem in spodnjem delu F_{Sn} , je r na nivoju $\alpha = 2P$ neznacično različen od r_H . Ako pa r leži izven tega razmaka, je r od r_H značilno različen s tveganjem $\alpha = 2P$. Ako dobimo, da je r na določenem nivoju značilno različen, postopek ponovimo na ostrejšem nivoju, dokler ne izpade razlika kot neznacična.

Primer: Z vzorcem $n = 60$ enot preskušamo značilnost razlik korelacijskega koeficienta od $r_H = 0,85$. Z vzorcem dobimo $r = 0,70$. Po zgornjem postopku izpade $r = 0,70$ kot značilno različen od $r_H = 0,85$ na nivoju $\alpha = 2P = 0,05$. Ako po zgornjem pravilu postopamo dalje in na enak način poskušamo na strožjih nivojih, se izkaže, da moremo smatrati razlike kot značilno različno na nivoju $\alpha = 2P = 0,01$, ker na nivoju $\alpha = 2P = 0,001$ ocenjeni r ne pade več v kritični razmak.

Problem 4 Preskušanje značilnosti odvisnosti med znakoma
Rešitev: Rešitev tega problema je poseben primer problema 3, če vzamemo $r_H = 0$.

Primer: Z vzorcem z $n = 25$ enot smo dobili $r = 0,20$. Ali kaže ta rezultat na značilno odvisnost med proučevanima pojavoma? Ker za $r_H = 0$ nomogram na skali $F_{Sn}-C$ za $\alpha = 2P = 0,05$ nakazuje, da je $r = 0,20$ iz vzorca $n = 25$ neznacičen, sklepamo na splošno, da je odvisnost neznacična.

Problem 5 Enostransko preskušanje značilnosti korelacijskih koeficientov.

Rešitev: Enostransko preskušamo značilnosti o korelacijskih

koeficientih enako kot dvostransko, le da se, glede na značaj problema, oziramo le na spodnjo ali zgornjo kritično mejo, hipoteze pa preskušamo s tveganjem $\alpha = P$.

Primer: Preskušamo hipotezo, da je korelacijski koeficient večji kot $r_H = 0,50$. $H_1: (r_0 > r_H = 0,50)$. Z vzorcem, ki ima $n = 100$ enot, smo dobili oceno $r = 0,63$. Nomogram pokaže, da je $r = 0,63$ značilno večji od $r_H = 0,50$ s tveganjem $\alpha = P = 0,05$ /preskus z FSn-B/ neznačilno večji s tveganjem $\alpha = P = 0,01$ /preskus na FSn-1/. Zaključimo zato, da je r značilno večji od $r_H = 0,50$ s tveganjem $\alpha = P = 0,05$.

Problem 6 Preskušanje hipoteze, da je korelacijski koeficient pozitiven ali negativen.

Rešitev: Rešitev tega problema je poseben primer problema 5, le da je $r_H = 0$.

Primer: Z vzorcem $n = 15$ enot smo dobili $r = -0,40$. Preskusiti je treba, ali je dobljeni r značilno negativen. Ta primer je identičen s preskusom, ali je $r = +0,40$ značilno pozitiven. Če namestimo $FSr(r_H=0)$ ob $FSn-B(n=\infty)$ za $\alpha = P = 0,05$, dobimo da je $r = 0,40$ neznačilen, ker je nad $n = 15$ na skali FSn-B.

Problem 7 Preskus o značilnosti razlik med korelacijskima koeficientoma iz enako velikih vzorcev ($n_1 = n_2 = n$).

Rešitev: Ta problem rešujemo s skalami FSn in premakljivo skalo $PS\Delta r$ enako kot problema 3 in 5, z edino razliko, da vlogo r_H prevzame ena izmed ocen (r_1 ali r_2), vlogo skale FSr a skala $PS\Delta r$ ob drugem robu premične skale. Analogna je tudi problematika dvostranosti /preskus različnosti/ ali enostranosti /preskus ali je r_1 značilno večji ali manjši od r_2 /testov.

Primer: Iz vzorca z $n = 30$ enot iz populacije A smo dobili $r_1 = 0,45$, iz vzorca $n = 30$ enot iz populacije B pa $r_2 = 0,77$. Ali in s kakšnim tveganjem sta dobljeni oceni značilno različni? Postopek, ki je po tehniki sličen reševanju zgornjih problemov, le da uporabljamo FS r, pokaže, da so razlike med zgornjima ocenama značilno različne s tveganjem $\alpha = 2P = 0,05$.

Problem 8 Značilnost razlik med ocenama korelacijskih koeficientov iz vzorcev, ki nista enako velika ($n_1 \neq n_2$).

Rešitev: Ta problem resimo s skalami FS_n in skalo FS_r. Če izkoristimo zvezo iz obrazca 6, dobimo z danim tveganjem $\alpha = 2P$ za dvostranske in s tveganjem $\alpha = P$, za enostranske preskuse odgovor o značilnosti razlik med r_1 in r_2 , če namestimo skale tako, da se večji r na FS_r sklada z manjšim n na FS_n za ustrezno tveganje, manjši r na FS_r pa s centralno črto za $n = \infty$. Če gremo od sečišča FS_r s centralno črto v smeri pomožnih črt pod kotom 45° do ustrezne skale FS_n, dobimo kritično velikost n_k drugega vzorca. Če je drugi vzorec manjši od n_k , so razlike neznailne, v nasprotnem primeru pa značilne. Če so med n_k in n_2 večje razlike, ni potrebno iskati točne vrednosti za n_k temveč moremo sklep o značilnosti napraviti že iz medsebojnega položaja n_2 in n_k .

Problem 9 Določanje funkcijskih vrednosti $Z(r)$ in $r(Z)$.

Rešitev: Ako FS_r pomaknemo ob skalo FS_Z tako, da se $r = 0$ sklada z $Z = 0$, dobimo nomogram za odvisnost med r in Z.

Primer: Iz treh vzorcev, ki so enako veliki, smo dobili naslednje ocene korelacijskih koeficientov: $r_1 = 0,42$; $r_2 = 0,66$; $r_3 = 0,056$. Kolike so ustrezne vrednosti transformiranih Z?

Če upoštevamo zgoraj navodilo, dobimo iz skal r in Z naslednje rezultate: $Z_1 = 0,45$; $Z_2 = 0,79$; $Z_3 = 0,63$.

Primer: Koliki r ustreza $\bar{Z} = 2,37$. Iz skal r in Z dobimo, da je vrednost ustreznega korelacijskega koeficienta $r=0,982$.

3. Shematičen prikaz rešitev problemov 1-9. V premakljivem delu nomograma PS so v sredini za operativno uporabo nakazane rešitve standardnih problemov, ki jih moremo reševati z nomogramom. Za ponazoritev konkretnih problemov 1-9 iz odstavka 2 pa so v sliki shematično prikazane njihove rešitve.

Problem 1

2

MEJE ZAUPANJA
eno-
dvo- stranske

3

TEST ZNAČILNOSTI
dvo- stranski

4

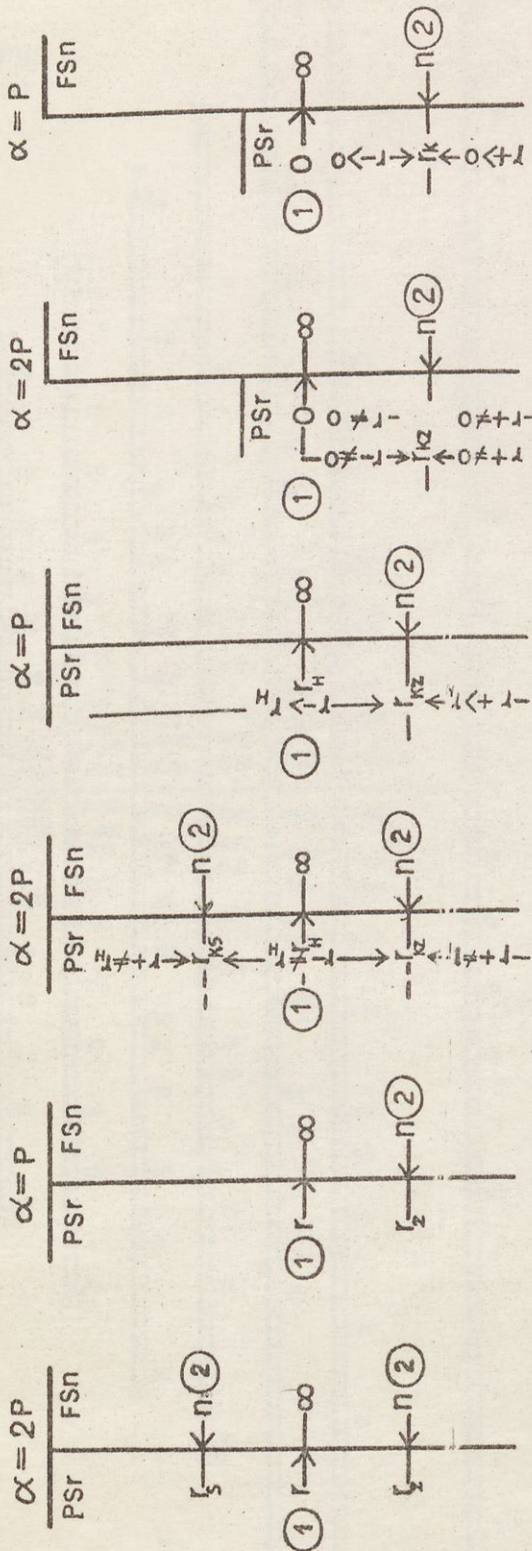
TEST ZNAČILNOSTI RAZLIK $r_H \neq 0$
eno-
dvo- stranski

5

TEST ZNAČILNOSTI
dvo- stranski

6

RAZLIK $r_H = 0$
eno-
dvo- stranski

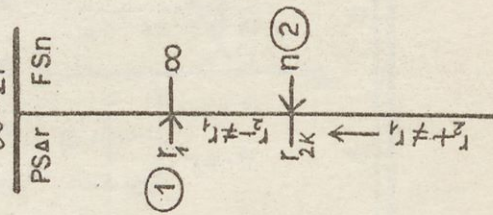


Problem 7

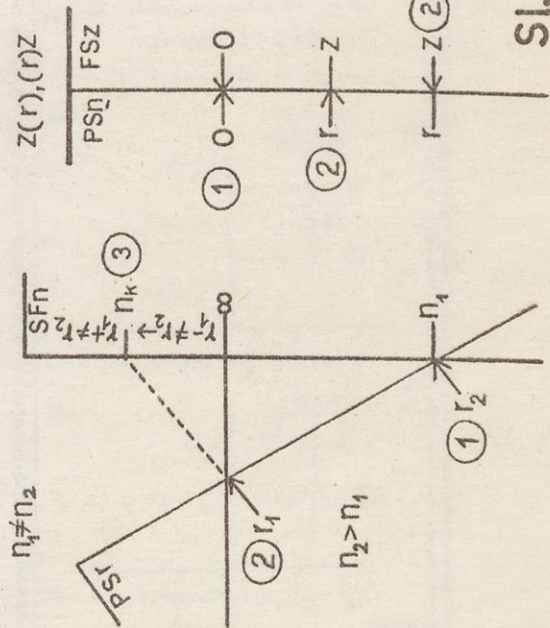
8

TEST ZNAČILNOSTI RAZLIK DVEH KORELACIJSKIH KOEFICIENTOV TRANSFORMACIJE

$n_1 = n_2 = n$
 $\alpha = 2P$



9



Legenda:

PS = premična skala

FS = fiksna skala

1 2 = zaporedje operacij

- ≠ - značilno različen

+ ≠ - značilno različen

-ž - neznačilno večji - manjši

+ž - značilno večji - manjši

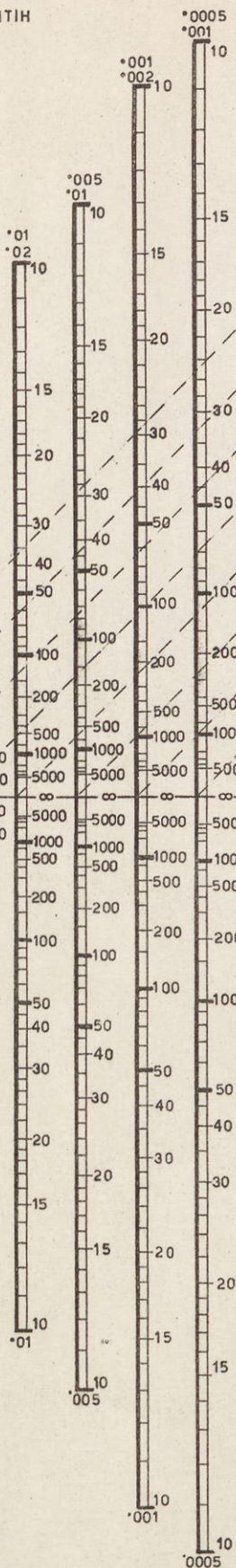
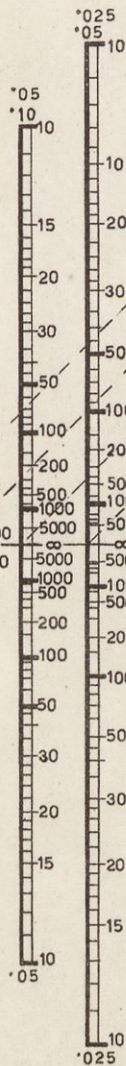
NOMOGRAM ZA OCENJEVANJE
IN PRESKUŠANJE HIPOTEZ O
KORELACIJSKIH KOEFICIENTIH

$$\alpha = P$$

$$\alpha = 2P$$

$$\alpha = P$$

FSn



INSTITUT ZA STATISTIKO
IN OPERACIJSKO RAZISKOVANJE
PRI EKONOMSKI FAKULTETI V LJUBLJANI

FS

NOMOGRAM ZA OCENJEVANJE
IN PRESKUŠANJE HIPOTEZ O
KORELACIJSKIH KOEFICIENTIH

MEJE ZAUPANJA
a. dvostransko

$$\alpha = 2P$$

$$\frac{PSr}{FSn}$$

$$r_s \leftarrow n \text{ ②}$$

$$\text{① } r \rightarrow \infty$$

$$r_z \leftarrow n \text{ ②}$$

$$P(r_s < r_o < r_z) = 1 - 2P$$

b. enostransko

$$\alpha = P$$

$$\frac{PSr}{FSn} \quad \frac{PSr}{FSn}$$

$$\text{① } r \rightarrow \infty \quad r_s \leftarrow n \text{ ②}$$

$$r_z \leftarrow n \text{ ②} \quad \text{① } r \rightarrow \infty$$

$$P(r_o > r_z) = 1 - P \quad P(r_o < r_s) = 1 - P$$

PREISKUŠANJE HIPOTEZ

a. dvostransko $H_0: r_o = r_H$

$$\alpha = 2P$$

$$\frac{PSr}{FSn}$$

$$r \neq r_H \quad n \text{ ②}$$

$$r \neq r_H \quad 1 r_H \rightarrow \infty$$

$$r \neq r_H \quad n \text{ ②}$$

b. enostransko $H_0: r_o > r_H$

$$\alpha = P$$

$$r \rightarrow r_H \text{ ① } r_H \rightarrow \infty$$

$$r \rightarrow r_H \quad n \text{ ②}$$

dvostransko $H_1: r_{o1} \neq r_{o2}$

$$\alpha = 2P$$

$$n_1 \neq n_2$$

$$\frac{PSr}{FSn}$$

$$r_2 \neq r_1 \quad n_2$$

$$\Delta r \neq 0 \quad \Delta r \neq 0$$

$$r_1 \neq r_2 \quad n_1$$

$$r_1 \neq r_2 \quad n_1$$

$$r_1 \neq r_2 \quad n_1$$

$$r_1 \neq r_2 \quad n_1$$

$$r_1 \neq r_2 \quad n_1$$

$$r_1 \neq r_2 \quad n_1$$

$$r_1 \neq r_2 \quad n_1$$

$$r_1 \neq r_2 \quad n_1$$

$$r_1 \neq r_2 \quad n_1$$

$$r_1 \neq r_2 \quad n_1$$

$$r_1 \neq r_2 \quad n_1$$

$$r_1 \neq r_2 \quad n_1$$

$$r_1 \neq r_2 \quad n_1$$

$$r_1 \neq r_2 \quad n_1$$

$$r_1 \neq r_2 \quad n_1$$

$$r_1 \neq r_2 \quad n_1$$

$$r_1 \neq r_2 \quad n_1$$

$$r_1 \neq r_2 \quad n_1$$

$$r_1 \neq r_2 \quad n_1$$

$$r_1 \neq r_2 \quad n_1$$

$$r_1 \neq r_2 \quad n_1$$

$$r_1 \neq r_2 \quad n_1$$





NARODNA IN UNIVERZITETNA KNJIŽNICA

GS

II 745 959



202215494

COBISS ©