TECHNICS AND INFORMATICS IN EDUCATION
6th International Conference, Faculty of Technical Sciences, Čačak, Serbia, 28–29th May 2016
TEHNIKA I INFORMATIKA U OBRAZOVANJU
6. međunarodna konferencija, Fakultet tehničkih nauka, Čačak, Srbija, 28–29. maj 2016.

UDK: 378-057.875:519.2 Stručni rad

Analiza rezultata na prijemnom ispitu i prvom kolokvijumu iz Poslovne informatike

Vladimir Kraguljac[1], Mladen Janjić[2] i Vera Lazarević[2]
[1] Fakultet za hotelijerstvo i turizam, Vrnjačka Banja, Univerzitet u Kragujevcu, Srbija
[2] Fakultet tehničkih nauka, Čačak, Univerzitet u Kragujevcu, Srbija
e-mail vladimir.kraguljac@kg.ac.rs, mladen.janjic@ftn.kg.ac.rs, vera.lazarevic@ftn.kg.ac.rs

Rezime: *U ovom radu, kroz nekoliko postavljenih hipoteza, vršena je analiza uspeha učenika na prijemnom ispitu i prvom kolokvijumu na nastavnom predmetu Poslovna informatika u školskoj 2015/2016. godini na Fakultetu za hotelijerstvo i turizam u Vrnjačkoj Banji. Za sprovođenje analize korišćene su odgovarajuće statistike i data su adekvatna tumačenja dobijenih rezultata.*

Ključne reči: *provera hipoteza; prijemni ispit; poslovna informatika; matematička statistika*

1. UVOD

U ovom radu je analiziran uspeh učenika na prijemnom ispitu i prvom kolokvijumu na predmetu Poslovna informatika u školskoj 2015/2016. godini na Fakultetu za hotelijerstvo i turizam u Vrnjačkoj Banji. Danas, u digitalnoj eri, informaciono-komunikacione tehnologije imaju veliki uticaj na svakodnevni život. Razmatranja i rezultati ovog rada mogu da daju nove poglede na veze između srednjoškolskog postignuća, znanja pokazanih na prijemnom ispitu i znanja pokazanih na predmetu koji je direktno vezan za informaciono-komunikacionu pismenost.

Podaci koji se analiziraju u primerima su preuzeti iz zvanične evidencije Studentske službe Fakulteta za hotelijerstvo i turizam.

2. PRIMENJENE METODE

Na raspolaganju su nam podaci o 107 studenata što predstavlja statistički veliki uzorak jer je njihov broj veći od 30. Za njihovu obradu korišćen je program StatSoft Statistica.

2.1. Testiranje hipoteze o srednjoj vrednosti

Kod velikog uzorka, što je ovde slučaj, aritmetička sredina uzorka $\overline{X}$ ima normalnu raspodelu $\overline{X} \sim N(\mu_x, \sigma_x / \sqrt{n}) \approx N(\mu_x, s_x / \sqrt{n})$. Ovo, naravno, važi i kad pređemo na standardizovanu slučajnu promenljivu $\overline{T} = (\overline{X} - \mu) / (\sigma / \sqrt{n}) \sim N(0,1)$.
Ako se testira hipoteza $H_0(\mu = a)$, tj. tvrđenje da se srednja vrednost ne razlikuje bitno od neke vrednosti a, ne odbacujemo je, uz nivo poverenja 95%, ukoliko je $(|\bar{x} - a| \sqrt{n}) / \sigma \leq$

1,96. Ako ovo nije ispunjeno onda odbacujemo nultu hipotezu sa pragom značajnosti od 5%. Da bi se smanjio ovaj rizik na 1% potrebno je zadovoljiti uslov $(|\bar{x} - a| \sqrt{n}) / \sigma > 2{,}58$.

2.2. Testiranje hipoteze o jednakosti srednjih vrednosti dva osnovna skupa

Neka su $\overline{X}_1$ i $\overline{X}_2$ aritmetičke sredine velikih uzoraka sa n_1 i n_2 elemenata koji podležu normalnoj raspodelama $N(\mu_1,\sigma_1)$ i $N(\mu_2,\sigma_2)$. Ako testiramo nultu hipotezu $H_0(\mu_1 = \mu_2)$, tj. da su srednje vrednosti posmatrana dva skupa jednake, onda i $\overline{X}_1 - \overline{X}_2$ ima normalnu raspodelu sa parametrima $\mu_{\bar{X}1 - \bar{X}2} = 0$ i $\sigma_{\bar{X}1 - \bar{X}2} = \sqrt{(\sigma_1^2 / n_1 + \sigma_2^2 / n_2)} \approx \sqrt{(s_1^2 / n_1 + s_2^2 / n_2)}$. Prelaskom na standardizovanu promenljivu za nju važi $t = (\overline{X}_1 - \overline{X}_2 - 0) / \sigma_{\bar{X}1 - \bar{X}2} \sim N(0,1)$. Vršimo poređenje $|t| < 1{,}96 = t_{0,05}$ i ako je ispunjeno ne odbacujemo nultu hipotezu sa verovatnoćom 95%. Slično, ako umesto $t_{0,05}$ koristimo $t_{0,01} = 2{,}58$ ta verovatnoća je 99%.

2.3. Testiranje hipoteze u vezi sa proporcijama osnovnog skupa

Testiramo hipotezu da je verovatnoća p posmatranog obeležja (uspeh) na elementima osnovnog skupa jednaka nekoj uočenoj vrednosti p_0, odnosno $H_0(p = p_0)$. Iz polaznog skupa izdvajamo uzorak od n elemenata, uz uslov da se radi o velikom uzorku – $n > 30$. Neka m elemenata ima traženo svojstvo. Sada je verovatnoća njihovog pojavljivanja u datom uzorku $\bar{p} = m / n$. Pod ovim uslovima slučajna promenljiva $\bar{p}$ ima normalnu raspodelu, pa važi da je $\bar{p} \sim N(p,\sqrt{(pq / n)}) \approx N(\bar{p},\sqrt{(\bar{p}\bar{q} / n)})$.
Uvodi se standardizovana promenljiva $t = (\bar{p} - p) / \sqrt{(\bar{p}\bar{q} / n)} = (\bar{p} - p_0) / \sqrt{(\bar{p}\bar{q} / n)} \sim N(0,1)$.
Odluku o prihvatanju ili odbacivanju nulte hipoteze donosimo na sledeći način:
Ako je $|t| < 1{,}96$ hipoteza se ne odbacuje. Ako je $|t| > 2{,}58$ hipoteza se odbacuje kao netačna jer je razlika između p i p_0 visoko značajna.
U slučaju da je $1{,}96 < |t| < 2{,}58$ razlika između p i p_0 je značajna, pa možemo ili da odbacimo hipotezu ili, još bolje, da izvršimo novo testiranje na većem uzorku.

2.4. Testiranje hipoteze o jednakosti disperzija dva uzorka normalnih skupova

Kada se porede dva skupa potrebno im je uporediti aritmetičke sredine i disperzije. Uz pretpostavku da u pomenutim skupovima važi normalna raspodela iz njih uzmemo dva statistički velika uzorka sa n_1 i n_2 elemenata i uzoračkim disperzijama s_1^2 i s_2^2. Hipoteza koju proveravamo je $H_0(\sigma_1^2 = \sigma_2^2)$. Slučajna promenljiva $F = (n_1 s_1^2 (n_2 - 1)) / (n_2 s_2^2 (n_1 - 1)) \approx s_1^2 / s_2^2$ ima raspodelu sa dva stepena slobode $k_1 = n_1 - 1$ i $k_2 = n_2 - 1$.
Odluku o prihvatanju ili odbacivanju postavljene hipoteze donosimo poredeći izračunatu vrednost za F sa tabelarnom vrednošću $(F^{(k1,k2)})_{0,05}$, odnosno $(F^{(k1,k2)})_{0,01}$, tako da ako je F manje od tabelarne vrednosti hipotezu ne odbacujemo i ako je veće odbacujemo je.

2.5. $\chi 2$ test za verifikaciju neparametarskih hipoteza

Testovi kojima se ispituje raspodela verovatnoća nekog obeležja u opštoj populaciji su testovi neparametarskih hipoteza. Veličina χ^2 se koristi kao mera odstupanja između empirijskih i pretpostavljenih teorijskih frekvencija i izračunava se na sledeći način:

$$\chi^2 = \sum((f_i - f_{ti})^2 / f_{ti}) = \sum(f_i^2 / f_{ti} - N)$$

gde su fi – frekvencija i-te vrednosti slučajne promenljive u uzorku ili frekvencija i-te klase; f_{ti} – odgovarajuća teorijska frekvencija; $N = \sum f_i = \sum f_{ti}$ – ukupan broj elemenata; α – verovatnoća, kritični koeficijent o slaganju empirijske i teorijske raspodele, rizik prihvatanja hipoteze ($1 - \alpha$ je pouzdanost hipoteze); k – broj stepeni slobode (za normalnu raspodelu je k = broj klasa - 3).

Odluku o odbacivanju ili neodbacivanju postavljene hipoteze donosimo tako što iz tablice vrednosti za koje je $P(\chi^2 > \chi_\alpha^2) = \alpha$ očitamo vrednost χ_α^2 prema izabranoj vrednosti za α i odgovarajućem broju stepeni slobode, a zatim ako je izračunata vrednost χ^2 veća od χ_α^2 odbacujemo hipotezu. Ako je izračunata vrednost χ^2 manja od χ_α^2 ne odbacujemo hipotezu, ali da bismo je prihvatili izvršimo testiranje na još nekoliko uzoraka.

2.6. Test Kolmogorova za verifikaciju neparametarskih hipoteza

Ovo je test koji zahteva manji broj izračunavanja u odnosu na χ^2 test i često se može koristiti umesto njega. Kolmogorov je uveo veličinu D_n kao maksimalnu razliku između empirijske funkcije raspodele i pretpostavljene teorijske raspodele i došao je do funkcije raspodele verovatnoća slučajne promenljive $D_n \sqrt{n}$.

$$\lim_{n\to\infty} P(D_n \sqrt{n} < \lambda) = \lim_{n\to\infty} P(D_n < \lambda / \sqrt{n}) = Q(\lambda) = \sum_{k\in[-\infty,\infty]}((-1)^k e^{-2k^2 \lambda^2})$$

Proces odlučivanja da li se prihvata ili odbacuje hipoteza o saglasnosti empirijske i pretpostavljene teorijske raspodele ide prema sledećem:

1. Pretpostaviti da neko obeležje X u generalnoj populaciji ima funkciju raspodele F(x).
2. Odabrati dovoljno veliki uzorak i formirati funkciju raspodele Fn(x) i naći Dn.
3. Izabrati ciljanu pouzdanost $1 - \alpha$.
4. U tabeli vrednosti funkcije $Q(\lambda)$ naći vrednost $\lambda\alpha$ za koju važi da je $Q(\lambda\alpha) = 1 - \alpha$.
5. Ako je $Dn \sqrt{n} < \lambda\alpha$ ne odbacujemo hipotezu, a u suprotnom se odbacuje sa rizikom α.

3. REZULTATI ISTRAŽIVANJA

3.1. Testiranje hipoteze o srednjoj vrednosti ukupnog broja bodova na prijemnom ispitu

Testiramo hipotezu da je srednji ukupan broj bodova na prijemnom ispitu 86 - $H_0(\mu = 86)$. Korišćenjem programa StatSoft Statistica dobijamo sledeće vrednosti: $n = 107$; $\sigma = 4{,}05$; $\bar{x} = 85{,}36$; $a = 86$; $\alpha = 5\%$; $c = 85{,}36 \pm 0{,}78$ i $p = 0{,}10$.

Sada računamo $(|\bar{x} - a| \sqrt{n}) / \sigma = 1{,}64 \leq 1{,}96$. Pošto je izračunata vrednost 1,64 manja od 1,96 ne odbacujemo nultu hipotezu sa pragom značajnosti $\alpha = 5\%$. Do istog zaključka dolazimo upoređivanjem naše pretpostavljene vrednosti i kritične vrednosti c = 86,13. Slično bismo zaključili i upoređivanjem dobijene P-vrednosti p = 0,10 sa vrednošću 0,05. Pošto je p > 0,05 ne odbacujemo nultu hipotezu sa pragom značajnosti $\alpha = 5\%$.

	Descriptive Statistics (podaci in kolokvijum 4.stw)						
Variable	**Valid N**	Mean	Confidence -95,000%	Confidence 95,000	Minimum	Maximum	Std.Dev.
ukupno prijemni	107	85,35738	84,58095	86,13381	74,40000	95,16000	4,050979

Slika 1. *Deskriptivna statistika – kompletan uzorak*

	Test of means against reference constant (value) (podaci in kolokvijum 4.stw						
Variable	**Mean**	Std.Dv.	N	Std.Err.	Reference Constant	t-value	df	p
ukupno prijemni	85,35738	4,050979	107	0,391623	86,00000	-1,64091	106	0,103781

Slika 2. *Testiranje hipoteze o srednjoj vrednosti*

3.2. Testiranje hipoteze o jednakosti srednjih vrednosti

Testiramo jednakost srednjih vrednosti ukupnog broja bodova na kolokvijumu za studente na budžetu i samofinansirajuće. Korišćenjem programa StatSoft Statistica dobijamo sledeće vrednosti (indeks 1 su studenti na budžetu, a 2 samofinansirajući):

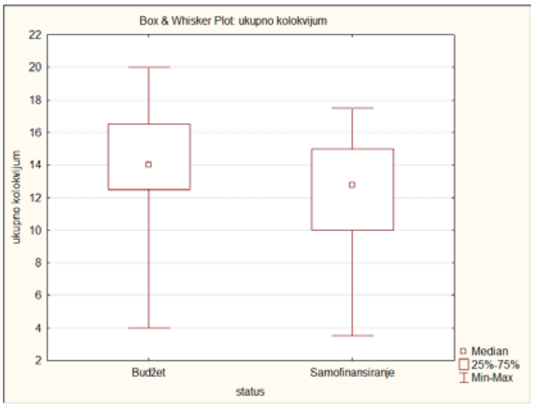
$$n_1 = 57;\ n_2 = 50;\ \sigma_1 = 3{,}39;\ \sigma_2 = 3{,}61;\ \bar{x}_1 = 14{,}11 \text{ i } \bar{x}_2 = 12{,}19.$$

Računamo $t = (\bar{X}_1 - \bar{X}_2 - 0) / \sqrt{(\sigma_1^2 / n_1 + \sigma_2^2 / n_2)} = 2{,}84 > 1{,}96$. Pošto je izračunata vrednost 2,84 veća od 1,96 odbacujemo postavljenu hipotezu uz mogućnost greške 5%. Ovo znači da postoji statistički značajna razlika u rezultatima koje postižu studenti upisani na budžet i samofinansirajući studenti.

T-tests; Grouping: status (podaci in kolokvijum 4.stw)
Group 1: Budžet
Group 2: Samofinansiranje

Variable	Mean Budžet	Mean Samofinansiranje	t-value	df	p	Valid N Budžet	Valid N Samofinansiranje	Std.Dev. Budžet	Std.Dev. Samofinansiranje	F-ratio Variances	p Variances
ukupno kolokvijum	14,11404	12,19000	2,839641	105	0,005424	57	50	3,391188	3,613876	1,135646	0,642451

Slika 3. *Testiranje jednakosti srednjih vrednosti*



Slika 4. *Box-Whisker-ov dijagram*

3.3. Testiranje hipoteze o verovatnoći prolaznosti na kolokvijumu

Testiramo pretpostavku da je prolaznost na kolokvijumu, tj. broj studenata sa više od 10 bodova, iznosi 75% - $H_0(p = 0{,}75)$. Korišćenjem programa StatSoft Statistica dobijamo sledeće vrednosti: $n = 107$; $m = 86$; $\bar{p} = m / n = 86 / 107 = 0{,}80$; $\bar{q} = 1 - \bar{p} = 0{,}20$ i $p_0 = 0{,}75$. Računamo $t = (\bar{p} - p_0) / \sqrt{(\bar{p}\bar{q} / n)} = 1{,}40 < 1{,}96$. Pošto je izračunata vrednost 1,40 manja od 1,96 nemamo razlog da odbacimo postavljenu hipotezu.

3.4. Testiranje hipoteze o jednakosti disperzija dva uzorka

Poredićemo podatke o broju bodova na kolokvijumu za budžetske i za samofinansirajuće studente.

status=Budžet
Descriptive Statistics (Spreadsheet in kolokvijum 4.stw)

Variable	Valid N	Mean	Minimum	Maximum	Std.Dev.
ukupno kolokvijum	57	14,11404	4,000000	20,00000	3,391188

Slika 5. *Deskriptivna statistika – budžet*

Variable	status=Samofinansiranje Descriptive Statistics (Spreadsheet in kolokvijum 4.stw) Valid N	Mean	Minimum	Maximum	Std.Dev.
ukupno kolokvijum	50	12,19000	3,500000	17,50000	3,613876

Slika 6. *Deskriptivna statistika – samofinansiranje*

Uz $n_1 = 57$; $n_2 = 50$; $\sigma_1 = 3{,}40 = s_1$ i $\sigma_2 = 3{,}61 = s_2$ imamo da je

$$F = (n_1\ s_1^2\ (n_2 - 1))\ /\ (n_2\ s_2^2\ (n_1 - 1)) = 0{,}88.$$

Promenljiva F ima raspodelu sa dva stepena sloboda $k_1 = n_1 - 1 = 56$ i $k_2 = n_2 - 1 = 49$. U tabeli nemamo vrednost $(F^{(56,49)})_{0,05}$, ali vidimo da je ona u opsegu od 1,44 do 1,76. Hipoteza koju proveravamo je $H_0(\sigma_1^2 = \sigma_2^2)$, a pošto je izračunato F manje od $(F^{(56,49)})_{0,05}$ hipotezu ne odbacujemo sa pragom značajnosti $\alpha = 5\%$. Ovo znači da je rasturanje oko srednjih vrednosti isto i za budžetske i za samofinansirajuće studente. Drugim rečima, rezultati ovog istraživanja podjednako su pouzdani za oba ta uzorka i u budućnosti ne možemo očekivati značajnije promene u odnosu njihovih postignuća, tj. budžetski studenti će uvek postizati bolje rezultate.

3.5. Testiranje hipoteza o nezavisnosti

Ovde ćemo upotrebiti χ^2 test i postavljamo hipotezu da su vrsta škole i status međusobno nezavisni. Korišćenjem programa StatSoft Statistica dobijamo frekvencije i vrednosti

Summary Frequency Table (podaci in kolokvijum 4.stw)
Marked cells have counts > 10

tip škole	status Budžet	status Samofinansiranje	Row Totals
tehnička	13	7	20
ostale	14	9	23
ekonomska	14	15	29
gimnazija	8	8	16
ugostiteljsko-turistička	6	9	15
medicinska	2	0	2
trgovačka	0	2	2
All Grps	57	50	107

Slika 7. *Frekvencije podataka za vrstu škole*

Statistic	Statistics: tip škole(7) x status(2) Chi-square	df	p
Pearson Chi-square	7,093856	df=6	p=,31226
M-L Chi-square	8,649167	df=6	p=,19430

Slika 8. *χ^2 test – vrsta škole i status*

Vidimo da su vrednosti $\chi^2 = 7{,}09$ i $p = 0{,}31$. Kako je P-vrednost $p > 0{,}05$ ne odbacujemo postavljenu hipotezu sa pragom značajnosti $\alpha = 5\%$. Ovo znači da ne postoji značajna povezanost vrste škole iz koje studenti dolaze i statusa na upisu.

3.6. Testiranje hipoteza da li su ostvareni rezultati saglasni sa normalnom raspodelom

Test Kolmogorova je sličan χ^2 testu i hipoteze koje proveravamo su slične predašnjoj – da li broj bodova iz srednje škole, na prijemnom, ukupno na prijemnom, na teorijskom delu, na praktičnom delu i ukupno na kolokvijumu prate normalnu raspodelu.

Variable	Tests of Normality (poda N	 max D	 K-S p
bodovi srednja	107	0,087150	p > .20
bodovi prijemni	107	0,073172	p > .20
ukupno prijemni	107	0,092580	p > .20
bodovi teorija	107	0,160116	p < ,01
bodovi praktično	107	0,119716	p < ,10
ukupno kolokvijum	107	0,101548	p > .20

Slika 9. *Test da li podaci prate normalnu raspodelu*

Uvidom u tabelu vrednosti funkcije $Q(\lambda)$ nalazimo da je granična vrednost $\lambda_\alpha \approx 1{,}36$. Proverom da li je ispunjen uslov $D_n\ \sqrt{n} < \lambda_\alpha$ zaključujemo da samo bodovi na teorijskom delu kolokvijuma nemaju normalnu raspodelu. Do istog zaključka bi došli i posmatranjem P-vrednosti, jer samo za bodove na teorijskom delu kolokvijuma imamo da je $p < 0{,}05$. U nastavku je, kao primer, data detaljnija tabela koja se dobija prilikom testiranja iste hipoteze za jednu vrednost - ukupan broj bodova na kolokvijumu posmatran po klasama podataka.

Variable: ukupno kolokvijum, Distribution: Normal (podaci in kolokvijum 4.stw)
Kolmogorov-Smirnov d = 0,06224,
Chi-Square = 18,71057, df = 12, p = 0,09576

Upper Boundary	Observed Frequency	Cumulative Observed	Percent Observed	Cumul. % Observed	Expected Frequency	Cumulative Expected	Percent Expected	Cumul. % Expected	Observed-Expected
<= 2,40000	0	0	0,00000	0,0000	0,14703	0,1470	0,13741	0,1374	-0,14703
3,80000	1	1	0,93458	0,9346	0,34173	0,4888	0,31938	0,4568	0,65827
5,20000	2	3	1,86916	2,8037	0,92725	1,4160	0,86659	1,3234	1,07275
6,60000	5	8	4,67290	7,4766	2,16882	3,5848	2,02693	3,3503	2,83118
8,00000	3	11	2,80374	10,2804	4,37296	7,9578	4,08688	7,4372	-1,37296
9,40000	5	16	4,67290	14,9533	7,60087	15,5587	7,10362	14,5408	-2,60087
10,80000	9	25	8,41121	23,3645	11,38916	26,9478	10,64408	25,1849	-2,38916
12,20000	10	35	9,34579	32,7103	14,71182	41,6596	13,74936	38,9342	-4,71182
13,60000	17	52	15,88785	48,5981	16,38287	58,0425	15,31110	54,2453	0,61713
15,00000	23	75	21,49533	70,0935	15,72767	73,7702	14,69876	68,9441	7,27233
16,40000	10	85	9,34579	79,4393	13,01633	86,7865	12,16480	81,1089	-3,01633
17,80000	14	99	13,08411	92,5234	9,28668	96,0732	8,67914	89,7880	4,71332
19,20000	6	105	5,60748	98,1308	5,71184	101,7850	5,33817	95,1262	0,28816
20,60000	2	107	1,86916	100,0000	3,02851	104,8136	2,83039	97,9566	-1,02851
< Infinity	0	107	0,00000	100,0000	2,18644	107,0000	2,04340	100,0000	-2,18644

Slika 10. *Test Kolmogorova i χ^2 test*

Ovde je $D_n\ \sqrt{n} = 0{,}64 < 1{,}36 = \lambda_\alpha$ pa se postavljena hipoteza ne odbacuje sa pragom značajnosti $\alpha = 5\%$.

4. ZAKLJUČAK

Iz rezultata ove analize dolazi se do novih zaključaka. Uočljivo je da je prijemni ispit dobro postavljen, jer se vidi da ne favorizuje nijednu vrstu škole iz koje kandidati dolaze. Zatim, ohrabruje to što se iz prosečnog broja bodova na prijemnom, 86 od 100, vidi da i fakulteti locirani van velikih univerzitetskih centara mogu da privuku kvalitetne učenike. Ovo dalje omogućava kvalitetan rad sa njima, pa na kolokvijumu i budžetski i samofinansirajući

studenti postižu ujednačene i dobre rezultate uz visoku prolaznost. Uspeh budžetskih studenata je nesto bolji, što se i moglo očekivati.
Što se tiče samog kolokvijuma na predmetu Poslovna informatika može se primetiti da postoje izvesna odstupanja od očekivanih u postignuću na teorijskom delu. Pozitivno je to što postignuće na praktičnom delu, što je naročito važno za dalje studiranje i kasnije uključivanje u poslovno okruženje, zadovoljava visoka očekivanja.

LITERATURA

[1] Janjić, M., Lazarević, V. (2014). Kovarijansa i korelacija srednjoškolskog uspeha i pokazanih rezultata na prijemnom ispitu. Tehnika i informatika u obrazovanju, 5. Konferencija sa međunarodnim učešćem, FTN Čačak, 30-31. maj 2014.

[2] Lazarević, V., Đukić, M. (2010). Inženjerska Matematika. Čačak: Tehnički Fakultet.

[3] Merkle, M. (2010) Verovatnoća i statistika za inženjere i studente tehnike. Beograd: Akademska misao.

[4] Vukadinović, S. (1988) Elementi teorije verovatnoće i matematičke statistike. Beograd: Privredni pregled.