

LINS

zadaci za ispit

U skripti se nalaze zadaci koji su bili 2013. godine u junskom predroku i roku (kreteni su dali iste zadatke oba puta).

Grupu B je devojka lepo zapamtila sve, a Grupu A je tip fristajlom ispisao.

Grupa B

1. Neka je poznato za trodimenzionalni vector X sledeće:

$$\begin{aligned} \text{Var}(X_1) &= 3, \text{Var}(X_2) = 5, \text{Var}(X_3) = 3, \\ \text{Cov}(X_1X_2) &= -\frac{1}{\sqrt{15}}, \text{Cov}(X_1X_3) = -\frac{1}{3}, \text{Cov}(X_2X_3) = -\frac{1}{\sqrt{15}} \end{aligned}$$

- Odrediti analitički oblik prve i druge komponente korišćenjem kovarijacione matrice
- Koliko treba uzeti glavnih komponenti da bi se obuhvatilo 85% varijabiliteta?
- Na tako odabranim komponentama izvršiti aproksimaciju kovarijacione matrice, Formirati matricu reziduala.

2. Na osnovu uzorka iz p-dimenzionane populacije (p=5) dobijena je uzoračka korelaciona matrica

$$R = \begin{bmatrix} 1 & 0 & 0.8 & 0.3 & 0 \\ 0 & 1 & 0.1 & 0.6 & 0.8 \\ 0.8 & 0.1 & 1 & 0.4 & 0.1 \\ 0.3 & 0.6 & 0.4 & 1 & 0.7 \\ 0 & 0.8 & 0.1 & 0.7 & 1 \end{bmatrix}$$

Karakteristični vektori i koreni matrice R su dati u tabeli:

Karakteristični vektori

prvi	drugi	treći	četvrti	peti
0.2775	0.6365	-0.1965	0.6831	-0.1123
0.4893	-0.3508	-0.6102	-0.1294	-0.4985
0.3426	0.5871	-0.1201	-0.6802	0.2468
0.5510	-0.0451	0.7568	-0.0213	-0.3481
0.5124	0.3538	-0.0444	0.2314	0.7462

Karakteristični koreni

2.5835	1.7166	0.3338	0.1941	0.1720
--------	--------	--------	--------	--------

Korišćenjem metode glavnih komponenti oceniti model faktorske analize.

3. Jednom studijom je ispitivano da li prosečna ocena studenta zavisi od toga da li je student iz kategorije koja plaća školarinu (samofinansirajući) ili je iz kategorije budžetskog studenta. U sledećim tabelama se nalaze odgovarajući izveštaji dobijeni obradom podataka iz SPSS programa:

One-Sample Kolmogorov-Smirnov Test			
			Prosek ocena
N			376
Normal Parameters		Mean	8.2414
		Std. Deviation	0.82248
Most Extreme Differences		Absolute	0.091
		Positive	0.065
		Negative	-0.091
Kolmogorov-Smirnov Z			1.757
Asymp. Sig. (2-tailed)			0.004

Group Statistics					
Prosek	Nacin	N	mean	Std. Dev.	Std. Error Mean
	budzet	307	8.3498	0.84426	0.04818
	samofinansirajuci	69	7.7588	0.48349	0.05821

Independent Samples Test									
	F	Sig.	t	df	Sig (2-tailed)	Mean	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	9.452	0.002	5.608	374	0	0.59096	0.10538	0.03837	0.79818
Equal variances not assumed			7.821	174.885	0	0.59096	0.07556	0.44183	0.74009

+ 2 tabelle Rank i Test Statistics (mislim da nisu bile potrebne za čitanje podataka)

- Koliki je broj studenata obuhvaćen istraživanjem?
- Kolika je prosečna ocena studiranja studenata koji su obuhvaćeni istraživanjem?
- Da li su studenti koji studiraju na budžetu bolji od onih koji sami plaćaju školarinu? Formirati hipoteze.

- d) Na osnovu datih tabela detaljno obrazložite vaš odgovor i svako vaše tvrđenje potkrepite vrednostima iz SPSS output izveštaja.

4. Data je matrica podataka o prosečnim ocenama ulenika u 6., 7. i 8. Razredu u pojedinim školama iz Beograda.

V.R.	4.53	4.37	4.42
V.K.	4.32	4.23	4.4
D.M.	4.2	4.23	4.28
D.P.	4.63	4.55	4.57
K.P.	4.58	4.59	4.55
S.M.	4.3	4.25	4.38

Na osnovu podataka iz ove tabele, metodom nehijerarhijske klasifikacije podeliti entitete u tri grupe, sa rešenjem najudaljenijih elemenata korišćenjem cisz block odstojanja. Odrediti:

- Finalno odstojanje između klastera
- Definisati elemente svakog klastera
- Utvrđiti migraciju između inicijalnih i krajnjih rešenja klaster analize

GRUPA A:

1. Dato je $X=(X_1, X_2, X_3)$ i $\text{Var}(x_1)=3, \text{Var}(x_2)=5, \text{Var}(x_3) = 3, \text{COV}(x_1, x_2)=-1, \text{COV}(x_1, x_3) = 1, \text{COV}(x_3, x_2)=1$.

$$\text{Var}(X_1) = 3, \text{Var}(X_2) = 5, \text{Var}(X_3) = 3,$$

$$\text{Cov}(X_1 X_2) = -1, \text{Cov}(X_1 X_3) = 1, \text{Cov}(X_2 X_3) = 1$$

- Odrediti algebarski oblik prve i treće glavne komponente
- Odrediti koliko komponenti obuhvata varijabilitet od 70%
- Nad tim komponentama oceniti kovarijacionu matricu i odrediti matricu reziduala

2. Dobijena neka tabela iz SPSS-a. Tekst je o tome da su neki asistenti vršili anketu nad fonovcima. Bilo je 15 pitanja pa su navedena pitanja. Na svako od pitanja su studenti odgovarali koliko se slažu (1-5).

Primeri pitanja su:

-Da li se osećaš iscrpljeno od obaveza na fakultetu?

- Da li misliš da si naučio dovoljno korisnih stvari na fakultetu?

I tako dalje. Onda je data matrica (izlaz iz SPSS-a):

Pitanje	Component		
	1	2	3
Da li se osećaš iscrpljeno od obaveza na fakultetu?	-.121	.877	.310
...			
Da li misliš da si naučio dovoljno korisnih stvari na fakultetu?	-.898	.112	.011
...			

Pitanja:

- Grupisati odgovore prema faktorima
//Trebalo izdvojiti najveći po apsolutnoj vrednosti za svako pitanje. Npr. u prvom pitanju, to je 2. faktor, a u ovom dole, to je prvi faktor. Onda ih grupišete, u mojoj grupi je bilo fazon prvih 5 faktora u 2. grupu, drugih 5 u 1. grupu a ostalih 5 ide u 3. grupu
- Dati smisljena imena faktorima
Dobio sam prilično smisljene (tj imali su veze jedni sa drugima), fazon F1: "Emocionalna i fizička iscrpljenost obavezama na fakultetu", F3: "Zadovoljstvo naučenim na fakultetu", a F2 ne mogu da se setim tačno šta je bilo
- Izračunati komunalitete za one elemente u faktorima koji imaju najveći doprinos (tako nešto)
to sam među prvih 5 za drugi faktor našao najveći, pa izračunao za njega komunalitet i tako posle i za ostale.
- Ovo nemam pojma šta je bilo, nešto tipa kako ti komunaliteti (pod c) utiču na ukupnu varijansu, tako nekako O.o

3. Neki izlaz iz SPSS-a vezano za Kolmogorov -Smirnov raspodelu, *ne pitaj me rođeni*

4. Dati podaci o uspehu u OŠ:

Škola	Upeh		
	6. razred	7. razred	8. razred
Radojka Dakić	4.55	4.38	4.77
Petar Prvi
Đura Jakšić
....
...
...

Koristeći metodu hijerarhijske klasifikacije, city block method, metodi maksimizacije i minimizacije izračunati

- Odrediti matricu hijerarhije
- Prikazati dendograme i odrediti koja je metoda bolja
- preseći dendograme tako da se dobiju dve grupe.