

1. Neka je poznato za trodimenzionalni vektor X sledeće:

$$\text{Var}(X_1) = 3, \text{Var}(X_2) = 5, \text{Var}(X_3) = 3, \text{Cor}(X_1, X_2) = -\frac{1}{\sqrt{15}}, \text{Cor}(X_1, X_3) = \frac{1}{3}, \text{Cor}(X_2, X_3) = -\frac{1}{\sqrt{15}}$$

- 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 -dimenzionalne 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 tabele 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.
- 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