

WYBRANE PROBLEMY STATYSTYKI WIELOWYMIAROWEJ

7. Dla 933 danych z eksperymentu WAIS o macierzy kowariancji (P – testy obrazkowe, V - testy słowne, A – wiek, E – liczba lat edukacji) :

	P	V	A	E
P	126,1	116,4	-54,0	20,9
V	116,4	207,5	-20,5	30,1
A	-54,0	-20,5	119,4	-9,8
E	20,9	30,1	-9,8	9,6

i macierzy korelacji

	P	V	A	E
P	1	0,72	-0,44	0,60
V	0,72	1	-0,13	0,68
A	-0,44	-0,13	1	-0,29
E	0,60	0,68	-0,29	1

oblicz macierz korelacji cząstkowej, gdy znany (ustalony) jest wiek A.

8. Dane U.N.E.S.C.O. (1990 Demographic Year Book and The Annual Register 1992) dotyczą pewnych zmiennych, zebranych w 1990 roku w 91 najludniejszych państwach świata.

Są to zmienne :

S - śmiertelność na 1000 mieszkańców,

D - śmiertelność dzieci na 1000 urodzeń

Z – oczekiwana długość życia mężczyzny w chwili urodzenia

L – logarytm dochodu narodowego brutto (w \$)

Można przyjąć, że dane te mają czterowymiarowy rozkład normalny o parametrach, wyznaczonych z próby:

Tab.1 Statystyki opisowe			
	Średnia	Wariancja	Odch.Std.
S	10,84	22	4,65
D	54,90	2115	45,99
Z	61,49	92	9,62
L	7,51	3	1,64

Tab.2 Korelacje				
	S	D	Z	L
S	1,00	0,68	-0,75	-0,52
D	0,68	1,00	-0,94	-0,78
Z	-0,75	-0,94	1,00	0,81
L	-0,52	-0,78	0,81	1,00

Oblicz współczynniki korelacji cząstkowej między zmiennymi:

(a) Z i S gdy znane są L i D, (b) Z i L gdy znane są S i D,

(c) Macierz korelacji cząstkowych między Z, S i L gdy znana jest wartość D.

Porównaj te korelacje z odpowiednimi korelacjami z tablicy 2. Jak możesz interpretować te różnice? Dla jakiej pary ta różnica jest największa?

9. Statystyka testowa dla istotności korelacji cząstkowej r wyraża się wzorem: $t = \frac{r}{\sqrt{1-r^2}} \sqrt{df}$, gdzie r jest

korelacją cząstkową a df jest liczbą stopni swobody, równą $df = N - q - 2$, gdzie N jest liczbą obserwacji a q – liczbą znanych zmiennych. Statystyka ta ma rozkład t Studenta z df stopniami swobody. Hipotezę, że $r=0$ odrzuca się, gdy $|t|$ ma dużą wartość.

(a) Zweryfikuj istotność korelacji z zadania 8 (a) i (b)

(b) Zweryfikuj istotność korelacji między S i L w tablicy 2 i w macierzy z zadania 8 (c). Jak myślisz, dlaczego występują tu tak duże różnice?

10. Oblicz współczynnik regresji wielokrotnej między Z i zmiennymi S, D, L. Jakie są współczynniki kombinacji liniowej zmiennych S, D, L dających największą korelację ze zmienną Z?

11. Zweryfikuj hipotezę, że współczynnik

(a) przy zmiennej S,

(b) przy zmiennej L

jest statystycznie istotny.