

Mat-2.3112 Tilastolliset monimuuttujamenetelmät

Tentti 22.5.2013 / Mellin

Kirjoita **selvästi jokaiseen koepaperiin** alla mainitussa järjestyksessä:

- Mat-2.3112 Monimuuttujamenetelmät / Tentti 22.5.2013 / Mellin
- opiskelijanumero + kirjain
- TEKSTATEN sukunimi, kaikki etunimet
- koulutusohjelma, vuosikurssi
- mahdolliset entiset nimet ja koulutusohjelmat
- nimikirjoitus

Tentissä saa käyttää laskinta ja Mellinin kaava- ja taulukkokokoelmia.

Vastaa lyhyesti ja ytimekkäästi, mutta perustelee kuitenkin vastauksesi.

Harjoitustehtävistä saaduilla hyvityspisteillä korvataan alimman pistemäärän saaneen tehtävän pisteet, jos kokonaispistemäärä siten nousee.

Tehtäviä on 5 kappaletta.

1. Multinormaalijakauma

Esittele multinormaalijakauman tärkeimmät ominaisuudet. Käsittele ainakin seuraavia asioita:

- Multinormaalijakauman tiheysfunktio, parametrointi ja geometriset ominaisuudet.
- Mitkä multinormaalijakauman ominaisuudet karakterisoivat (ts. määrittelevät) multinormaalijakauman?
- Multinormaalijakauman reunajakaumat.
- Riippumattomuus, korreloimattomuus ja multinormaalijakauma.
- Multinormaalijakauman ehdolliset jakaumat ja niiden ehdollisten jakaumien parametrit.
- Multinormaalijakaumaa noudattavien satunnaismuuttujien neliömuotojen jakaumat.

Jakauman parametrien estimointi ja testaus eivät kuulu tehtävän alueeseen.

2. Pääkomponenttialyysi

Seuraavilla sivuilla on esitetty Ness-ohjelman tulostus eräästä pääkomponenttialyysistä.

Aineistona on 48 parasta kymmenottelijaa vuodelta 1973. Muuttujina ovat ottelijoiden tulokset (pisteiksi muunnettuna) kymmenottelun lajeista (100 m, Pituushyppy, Kuula, Korkeushyppy, 400 m, 110 m:n aidat, Kiekko, Seiväs, Keihäs, 1500 m).

- (a) Esittele pääkomponenttialyysin perusidea ja tavoite sekä pääkomponenttialyysin ”malli” ja sen osat. Kerro myös miten malli ”estimoidaan”.
- (b) Selitä pääkomponenttialyysin ”geometria”.
- (c) Kuinka monta pääkomponenttia valitsisit tulostuksen perusteella? Miksi?
- (d) Esitä valitsemillesi pääkomponenteille tulkinnat.
- (e) Mikä on muuttujan 1500m ja 2. pääkomponentin korrelaatio?
- (f) Selitä tulostuksen taulukot

Factor Loadings

Communalities

Score Coefficients

- (g) Tulostuksesta löytyy osio

Residual Section

Selitä sarakkeet T2 ja T2 Prob

- (h) 4:n ensimmäisen pääkomponentin pääkomponenttipisteet on tallennettu muuttujiksi

PCCOR1 PCCOR2 PCCOR3 PCCOR4

Niistä on lasketut tavanomaiset otostunnusluvut (aritmeettiset keskiarvot, otoskeskihajonnat ja korrelaatiot) on annettu tulostuksen osioissa

Data Screening Report

Correlation Report

Mitä pääkomponenttipisteillä tarkoitetaan? Selitä myös miksi pääkomponenttipisteiden otostunnusluvuilla on tulostuksissa esiintyvä arvot.

Principal Components Report

Database KYMMENOT.S0

Correlation Section

Variables	J100m	Pituush	Kuula	Korkeush	J400m
J100m	1.000000	0.171985	-0.027952	-0.411700	0.456082
Pituush	0.171985	1.000000	-0.034393	-0.003325	0.133463
Kuula	-0.027952	-0.034393	1.000000	0.162542	-0.303707
Korkeush	-0.411700	-0.003325	0.162542	1.000000	-0.338827
J400m	0.456082	0.133463	-0.303707	-0.338827	1.000000
Jaidat	0.315991	0.298065	0.086498	-0.038964	0.175548
Kiekkö	0.014343	0.020875	0.727330	0.216995	-0.344647
Seivas	0.054723	0.061024	-0.204229	-0.117765	0.006584
Keihas	-0.221250	0.153750	0.023140	0.149779	-0.104689
J1500m	-0.291744	-0.206665	-0.446248	-0.146150	0.302178

Phi=0.243776 Log(Det|R)=-2.711129 Bartlett Test=116.13 DF=45 Prob=0.000000

Variables	Jaidat	Kiekkö	Seivas	Keihas	J1500m
J100m	0.315991	0.014343	0.054723	-0.221250	-0.291744
Pituush	0.298065	0.020875	0.061024	0.153750	-0.206665
Kuula	0.086498	0.727330	-0.204229	0.023140	-0.446248
Korkeush	-0.038964	0.216995	-0.117765	0.149779	-0.146150
J400m	0.175548	-0.344647	0.006584	-0.104689	0.302178
Jaidat	1.000000	0.047707	-0.073497	-0.148188	-0.224552
Kiekkö	0.047707	1.000000	-0.181800	0.135561	-0.573501
Seivas	-0.073497	-0.181800	1.000000	-0.128523	0.012496
Keihas	-0.148188	0.135561	-0.128523	1.000000	-0.065414
J1500m	-0.224552	-0.573501	0.012496	-0.065414	1.000000

Phi=0.243776 Log(Det|R)=-2.711129 Bartlett Test=116.13 DF=45 Prob=0.000000

Eigenvalues

No.	Eigenvalue	Individual Percent	Cumulative Percent	Scree Plot
1	2.602056	26.02	26.02	
2	2.007813	20.08	46.10	
3	1.206620	12.07	58.16	
4	1.067052	10.67	68.84	
5	0.931538	9.32	78.15	
6	0.594690	5.95	84.10	
7	0.569080	5.69	89.79	
8	0.538683	5.39	95.18	
9	0.244855	2.45	97.62	
10	0.237613	2.38	100.00	

Eigenvectors

Variables	Factor1	Factor2	Factor3	Factor4
J100m	-0.142191	0.577679	-0.152659	0.034384
Pituush	0.012165	0.318322	0.649528	-0.208702
Kuula	0.484890	0.143798	-0.243121	0.128083
Korkeush	0.304128	-0.271021	0.269260	-0.066784
J400m	-0.395103	0.290172	0.080476	0.316822
Jaidat	0.016181	0.458382	0.192729	0.072209
Kiekkö	0.524703	0.157216	-0.137050	0.048363
Seivas	-0.166115	0.010646	-0.077875	-0.860722
Keihas	0.155059	-0.170252	0.595579	0.153031
J1500m	-0.410735	-0.350284	-0.004118	0.249411

Factor Loadings

Variables	Factors			
	Factor1	Factor2	Factor3	Factor4
J100m	-0.229367	0.818555	-0.167690	0.035518
Pituush	0.019624	0.451054	0.713483	-0.215586
Kuula	0.782170	0.203758	-0.267060	0.132307
Korkeush	0.490585	-0.384029	0.295772	-0.068987
J400m	-0.637336	0.411166	0.088400	0.327272
Jaidat	0.026101	0.649515	0.211706	0.074591
Kiekkö	0.846392	0.222770	-0.150544	0.049958
Seivas	-0.267958	0.015085	-0.085542	-0.889110
Keihas	0.250125	-0.241242	0.654221	0.158079
J1500m	-0.662553	-0.496343	-0.004523	0.257637

Communalities

Variables	Factors				Communality
	Factor1	Factor2	Factor3	Factor4	
J100m	0.052609	0.670032	0.028120	0.001262	0.752023
Pituush	0.000385	0.203450	0.509057	0.046477	0.759370
Kuula	0.611790	0.041517	0.071321	0.017505	0.742134
Korkeush	0.240673	0.147478	0.087481	0.004759	0.480392
J400m	0.406197	0.169057	0.007815	0.107107	0.690176
Jaidat	0.000681	0.421869	0.044819	0.005564	0.472934
Kiekkö	0.716380	0.049627	0.022664	0.002496	0.791166
Seivas	0.071802	0.000228	0.007317	0.790516	0.869863
Keihas	0.062562	0.058198	0.428006	0.024989	0.573755
J1500m	0.438976	0.246357	0.000020	0.066377	0.751730

Score Coefficients

Variables	Factors			
	Factor1	Factor2	Factor3	Factor4
J100m	-8.814847E-02	0.4076848	-0.1389752	3.328643E-02
Pituush	7.541696E-03	0.2246495	0.5913067	-0.2020386
Kuula	0.3005969	0.1014826	-0.2213288	0.1239934
Korkeush	0.1885373	-0.1912674	0.245124	-6.465157E-02
J400m	-0.2449355	0.2047829	7.326243E-02	0.3067067
Jaidat	1.003076E-02	0.3234936	0.1754534	6.990351E-02
Kiekkö	0.3252782	0.1109517	-0.1247653	4.681875E-02
Seivas	-0.1029795	7.513346E-03	-7.089417E-02	-0.8332396
Keihas	0.0961258	-0.1201517	0.5421933	0.1481453
J1500m	-0.2546266	-0.2472059	-3.748493E-03	0.2414477

Residual Section							
Row	T2	T2 Prob	Q0	Q1	Q2	Q3	Q4
1	7.62	0.7905	7.90	7.87	6.50	2.59	1.02
2	19.30	0.1566	20.06	14.75	12.48*	9.59*	8.96*
3	13.04	0.4194	16.16	14.96	7.69	7.11	2.02
4	9.63	0.6486	6.95	6.93	5.77	3.20	3.20
5	14.39	0.3447	17.15	9.26	6.40	3.66	3.55
6	13.40	0.3983	19.26	5.84	5.29	4.50	3.88
7	20.01	0.1388	9.72	6.47	6.47	6.41	6.41
8	10.57	0.5811	8.59	8.53	4.96	4.93	4.77
9	5.76	0.9017	3.32	3.00	2.72	2.70	2.35
10	9.39	0.6659	5.06	3.87	3.58	3.52	2.80
11	11.62	0.5092	9.13	7.25	5.83	5.16	4.29
12	14.59	0.3344	15.92	7.01	6.69	4.06	3.76
13	9.86	0.6316	11.69	5.01	4.96	2.07	1.93
14	8.87	0.7031	7.73	7.42	5.94	5.89	1.84
15	7.57	0.7938	8.09	6.94	4.92	3.34	3.11
16	5.20	0.9277	4.91	4.31	3.45	3.09	2.89
17	2.78	0.9921	4.52	1.87	0.74	0.69	0.57
18	8.83	0.7062	12.63	11.58	3.54	3.01	2.44
19	9.06	0.6895	8.30	8.22	6.01	3.02	2.47
20	5.66	0.9064	6.56	3.39	2.16	2.04	1.83
21	9.44	0.6619	11.28	6.38	3.92	3.64	2.17
22	8.13	0.7554	7.76	7.09	5.55	5.49	4.51
23	7.51	0.7980	10.41	7.25	2.15	2.14	2.00
24	10.31	0.5999	9.38	9.37	9.29	9.05*	6.81
25	8.62	0.7209	6.48	6.04	4.55	3.83	3.78
26	21.28	0.1116	14.43	11.80	11.79*	11.36*	10.47*
27	9.36	0.6681	11.55	10.95	3.78	3.01	0.93
28	4.34	0.9600	5.54	5.45	2.41	0.99	0.98
29	6.34	0.8702	10.58	1.41	1.00	0.96	0.84
30	6.37	0.8688	11.94	2.51	1.62	1.09	1.09
31	2.60	0.9940	2.40	1.84	1.35	1.35	0.77
32	6.22	0.8771	4.72	4.15	3.52	3.20	2.76
33	4.81	0.9438	5.10	4.92	4.90	1.83	0.54
34	7.68	0.7864	6.11	5.42	5.42	5.31	4.80
35	7.15	0.8212	6.54	6.54	5.00	5.00	2.23
36	6.17	0.8801	9.31	5.68	1.86	1.60	1.31
37	6.78	0.8448	7.12	7.01	6.33	1.73	0.96
38	13.02	0.4208	11.02	7.43	6.86	5.64	2.78
39	10.25	0.6043	15.10	9.51	5.44	5.08	1.09
40	9.91	0.6280	11.89	11.06	10.55	0.61	0.54
41	14.04	0.3628	21.77	15.61*	8.06	6.11	0.64
42	11.91	0.4903	12.73	12.44	4.28	3.75	3.61
43	10.80	0.5655	6.45	6.42	5.87	4.96	2.26
44	6.57	0.8573	5.07	3.55	3.55	3.45	2.71
45	18.96	0.1658	14.93	11.77	9.49	9.47*	6.45
46	11.31	0.5303	10.38	6.82	5.20	5.16	5.16
47	8.86	0.7038	5.87	4.41	4.37	4.35	4.35
48	14.12	0.3585	10.46	10.40	9.14	5.90	5.86

Data Screening Report

Database KYMMENOT.S0

Descriptive Statistics Section

Data Type	Variable	Value Count	Missing Count	Minimum	Maximum	Mean	Standard Deviation
Continuous	PCCOR1	48	0	-2.271179	1.85028	3.556183E-16	1
Continuous	PCCOR2	48	0	-2.01644	2.001073	1.595946E-16	1
Continuous	PCCOR3	48	0	-2.871038	1.801618	-1.753227E-15	1
Continuous	PCCOR4	48	0	-2.264534	1.948351	2.302556E-15	1

Correlation Report

Pearson Correlations Section (Row-Wise Deletion)

	PCCOR1	PCCOR2	PCCOR3	PCCOR4
PCCOR1	1.000000	0.000000	0.000000	0.000000
PCCOR2	0.000000	1.000000	0.000000	0.000000
PCCOR3	0.000000	0.000000	1.000000	0.000000
PCCOR4	0.000000	0.000000	0.000000	1.000000

Cronbachs Alpha = 0.000000 Standardized Cronbachs Alpha = 0.000000

3. Erotteluanalyysi

Seuraavilla sivuilla on esitetty Ncss-ohjelman tulostus eräästä erotteluanalyysistä.

Analysoitavana on kolmea iiris-kasvin lajiketta (Iris Setosa, Iris Versicolor, Iris Virginica) koskeva aineisto. Muuttujina on ko. iiris-lajikkeiden kukkien teriön erään terälehtien pituus ja leveys (PetalLength, PetalWidth) ja verhiön erään verholehden pituus ja leveys (SepalLength, SepalWidth).

- (a) Esittele erotteluanalyysin perusidea ja tavoite sekä erotteluanalyysin malli ja sen osat. Kerro myös miten malli estimoidaan.
- (b) Selitä erotteluanalyysin ”geometria”.
- (c) Kuinka monta erottelumuuttujaa tulostuksen perusteella tarvitaan lajikkeiden erottamiseksi toisistaan? Miksi?
- (d) Ovatko kaikki analyysin kohteena olevista 4 muuttujasta tilastollisesti merkitseviä erotteluinformaation kannalta? Pohdi asiaa 5 %:n ja 1 %:n merkitsevyystasojen valossa.
- (e) Kuinka hyvin erottelumuuttujat ovat onnistuneet erottamaan lajikkeet toisistaan?
- (f) Millä tavalla erotteluanalyysi ja *kanoninen analyysi* liittyvät toisiinsa?

Discriminant Analysis Report

Database MMLU\HARJ\FISHER.S0
 Dependent Iris

Group Means

Variable	Iris			
	Setosa	Versicolor	Virginica	Overall
SepalLength	50.06	59.36	65.88	58.43333
SepalWidth	34.28	27.7	29.74	30.57333
PetalLength	14.62	42.6	55.52	37.58
PetalWidth	2.46	13.26	20.26	11.99333
Count	50	50	50	150

Group Standard Deviations

Variable	Iris			
	Setosa	Versicolor	Virginica	Overall
SepalLength	3.524897	5.161712	6.358796	8.280662
SepalWidth	3.790644	3.137983	3.224966	4.358663
PetalLength	1.73664	4.69911	5.518947	17.65298
PetalWidth	1.053856	1.977527	2.7465	7.622377
Count	50	50	50	150

Total Correlation\Covariance

Variable	Variable			
	SepalLength	SepalWidth	PetalLength	PetalWidth
SepalLength	68.56935	-4.243401	127.4315	51.62707
SepalWidth	-0.117570	18.99794	-32.96564	-12.16394
PetalLength	0.871754	-0.428440	311.6278	129.5609
PetalWidth	0.817941	-0.366126	0.962865	58.10063

Between-Group Correlation\Covariance

Variable	Variable			
	SepalLength	SepalWidth	PetalLength	PetalWidth
SepalLength	3160.607	-997.6334	8262.42	3563.967
SepalWidth	-0.745075	567.2466	-2861.98	-1146.633
PetalLength	0.994135	-0.812838	21855.14	9338.7
PetalWidth	0.999768	-0.759258	0.996232	4020.667

Within-Group Correlation\Covariance

Variable	Variable			
	SepalLength	SepalWidth	PetalLength	PetalWidth
SepalLength	26.50082	9.272109	16.75143	3.840136
SepalWidth	0.530236	11.53878	5.524354	3.27102
PetalLength	0.756164	0.377916	18.51878	4.266531
PetalWidth	0.364506	0.470535	0.484459	4.188163

Variable Influence Section

Variable	Removed	Removed	Removed	Alone	Alone	Alone	R-Squared Other X's
	Lambda	F-Value	F-Prob	Lambda	F-Value	F-Prob	
SepalLength	0.938463	4.72	0.010329	0.381294	119.26	0.000000	0.858612
SepalWidth	0.766480	21.94	0.000000	0.599217	49.16	0.000000	0.524007
PetalLength	0.669206	35.59	0.000000	0.058628	1180.16	0.000000	0.968012
PetalWidth	0.743001	24.90	0.000000	0.071117	960.01	0.000000	0.937850

Linear Discriminant Functions

Variable	Iris		
	Setosa	Versicolor	Virginica
Constant	-85.20985	-71.754	-103.2697
SepalLength	2.354417	1.569821	1.244585
SepalWidth	2.358787	0.707251	0.3685279
PetalLength	-1.643064	0.5211451	1.276654
PetalWidth	-1.739841	0.6434229	2.107911

Regression Coefficients

Variable	Iris		
	Setosa	Versicolor	Virginica
Constant	0.1182229	1.577059	-0.6952819
SepalLength	6.602977E-03	-2.015369E-03	-4.587608E-03
SepalWidth	2.428479E-02	-4.456162E-02	2.027684E-02
PetalLength	-2.246571E-02	2.206692E-02	3.987911E-04
PetalWidth	-5.747273E-03	-4.943066E-02	5.517793E-02

Classification Count Table for Iris

Actual	Predicted			Total
	Setosa	Versicolor	Virginica	
Setosa	50	0	0	50
Versicolor	0	48	2	50
Virginica	0	1	49	50
Total	50	49	51	150

Reduction in classification error due to X's = 97.0%

Misclassified Rows Section

Row	Actual	Predicted	Percent Chance of Each Group		
			Pcnt1	Pcnt2	Pcnt3
5	Virginica	Versicolo	0.0	72.9	27.1
9	Versicolo	Virginica	0.0	25.3	74.7
12	Versicolo	Virginica	0.0	14.3	85.7

Canonical Variate Analysis Section

Fn	Inv(W)E Eigenvalue	Ind'l Pcnt	Total Pcnt	Canon Corr	Canon Corr2	F-Value	Numer DF	Denom DF	Prob Level	Wilks' Lambda
1	32.191929	99.1	99.1	0.9848	0.9699	199.1	8.0	288.0	0.0000	0.023439
2	0.285391	0.9	100.0	0.4712	0.2220	13.8	3.0	145.0	0.0000	0.777973

The F-value tests whether this function and those below it are significant.

Canonical Coefficients

Variable	Canonical Variate	
	Variate1	Variate2
Constant	-2.105106	6.661473
SepalLength	-0.082938	-0.002410
SepalWidth	-0.153447	-0.216452
PetalLength	0.220121	0.093192
PetalWidth	0.281046	-0.283919

Canonical Variates at Group Means

	Canonical Function	
	Function1	Function2
Iris		
Setosa	-7.6076	-0.215133
Versicolor	1.82505	0.7278996
Virginica	5.78255	-0.5127666

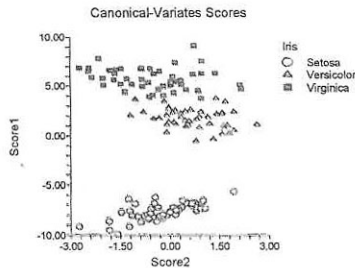
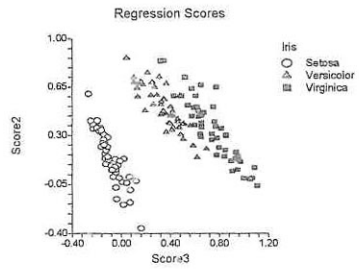
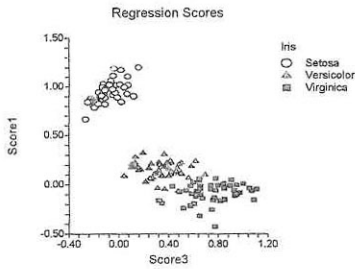
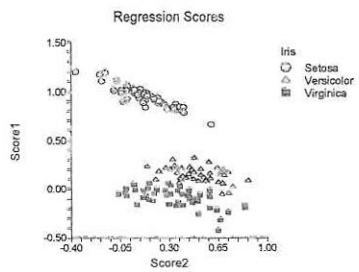
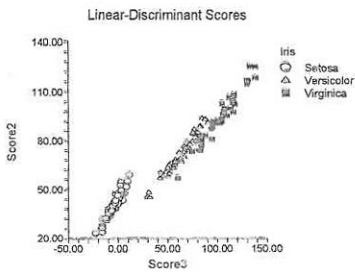
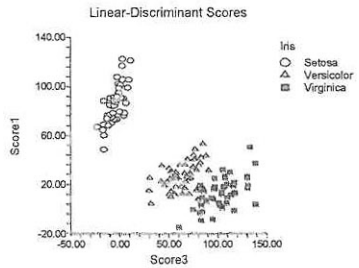
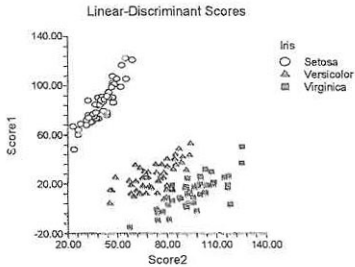
Std. Canonical Coefficients

Variable	Canonical Variate	
	Variate1	Variate2
SepalLength	-0.426955	-0.012408
SepalWidth	-0.521242	-0.735261
PetalLength	0.947257	0.401038
PetalWidth	0.575161	-0.581040

Variable-Variate Correlations

Variable	Canonical Variate	
	Variate1	Variate2
SepalLength	0.222596	-0.310812
SepalWidth	-0.119012	-0.863681
PetalLength	0.706065	-0.167701
PetalWidth	0.633178	-0.737242

Plots Section



4. Kanoninen analyysi

Seuraavilla sivuilla on esitetty Ncss-ohjelman tulostus eräästä kanonisesta analyysistä.

Aineistona on 48 parasta kymmenottelijaa vuodelta 1973. Muuttujina ovat ottelijoiden tulokset (pisteiksi muunnettuina) kymmenottelun lajimuuttujista (100 m, Pituushyppy, Kuula, Korkeushyppy, 400 m, 110 m:n aidat, Kiekko, Seiväs, Keihäs, 1500 m) sekä kaksi ottelijoiden ruumiinrakennetta kuvaavaa muuttujaa (Pituus, Paino).

Kanoninen analyysi on tehty lajimuuttujien ja ottelijoiden ruumiinrakennetta kuvaavien muuttujien välillä.

- (a) Esittele kanonisen analyysin perusidea ja tavoite sekä kanonisen analyysin malli ja sen osat. Kerro myös miten malli estimoidaan.
- (b) Kuinka monta kanonista muuttujaparia valitsisit tulostuksen perusteella? Miksi?
- (c) Esitä valitsemillesi kanonisille muuttujille tulkinnat.

Kanonisten muuttujien arvot on tallennettu muuttujiksi CANL1, CANL2, CANM1 ja CANM2.

- (d) Selitä korrelaatiot kohdassa

Correlation Report

- (e) Selitä millä tavalla kohdan

Plots Section

kuviot kuvastavat kohdan

Correlation Report

korrelaatioita.

- (f) Mitkä ovat kanonisen korrelaation ja *tavallisen korrelaation* yhteydet?
- (g) Mitkä ovat kanonisen analyysin ja *regressioanalyysin* yhteydet? Kerro myös, mikä on estimoidun regressiomallin selitystason tulkinta kanonisessa analyysissä.

Canonical Correlation Report

Database KYMMENOT5.S0
 Y Variables J100m to J1500m
 X Variables Pituus, Paino

Correlation Section

	J100m	Pituush	Kuula	Korkeush	J400m
J100m	1.000000	0.171985	-0.027952	-0.411700	0.456082
Pituush	0.171985	1.000000	-0.034393	-0.003325	0.133463
Kuula	-0.027952	-0.034393	1.000000	0.162542	-0.303707
Korkeush	-0.411700	-0.003325	0.162542	1.000000	-0.338827
J400m	0.456082	0.133463	-0.303707	-0.338827	1.000000
Jaidat	0.315991	0.298065	0.086498	-0.038964	0.175548
Kiekk	0.014343	0.020875	0.727330	0.216995	-0.344647
Seivas	0.054723	0.061024	-0.204229	-0.117765	0.006584
Keihas	-0.221250	0.153750	0.023140	0.149779	-0.104689
J1500m	-0.291744	-0.206665	-0.446248	-0.146150	0.302178
Pituus	-0.110064	-0.049075	0.617244	0.125104	-0.164746
Paino	-0.082002	-0.054517	0.708257	0.161192	-0.322526

	Jaidat	Kiekk	Seivas	Keihas	J1500m
J100m	0.315991	0.014343	0.054723	-0.221250	-0.291744
Pituush	0.298065	0.020875	0.061024	0.153750	-0.206665
Kuula	0.086498	0.727330	-0.204229	0.023140	-0.446248
Korkeush	-0.038964	0.216995	-0.117765	0.149779	-0.146150
J400m	0.175548	-0.344647	0.006584	-0.104689	0.302178
Jaidat	1.000000	0.047707	-0.073497	-0.148188	-0.224552
Kiekk	0.047707	1.000000	-0.181800	0.135561	-0.573501
Seivas	-0.073497	-0.181800	1.000000	-0.128523	0.012496
Keihas	-0.148188	0.135561	-0.128523	1.000000	-0.065414
J1500m	-0.224552	-0.573501	0.012496	-0.065414	1.000000
Pituus	0.225447	0.587934	-0.350088	-0.035560	-0.237119
Paino	0.130431	0.635411	-0.313625	-0.066120	-0.400991

	Pituus	Paino
J100m	-0.110064	-0.082002
Pituush	-0.049075	-0.054517
Kuula	0.617244	0.708257
Korkeush	0.125104	0.161192
J400m	-0.164746	-0.322526
Jaidat	0.225447	0.130431
Kiekk	0.587934	0.635411
Seivas	-0.350088	-0.313625
Keihas	-0.035560	-0.066120
J1500m	-0.237119	-0.400991
Pituus	1.000000	0.852174
Paino	0.852174	1.000000

Canonical Correlations Section

Variate Number	Canonical Correlation	R-Squared	F-Value	Num DF	Den DF	Prob Level	Wilks' Lambda
1	0.781725	0.611094	3.02	20	72	0.000310	0.295477
2	0.490139	0.240236	1.30	9	37	0.270030	0.759764

F-value tests whether this canonical correlation and those following are zero.

Variation Explained Section

Canonical Variate Number	Variation in these Variables	Explained by these Variates	Individual Percent Explained	Cumulative Percent Explained	Canonical Correlation Squared
1	Y	Y	21.3	21.3	0.6111
2	Y	Y	8.3	29.6	0.2402
1	Y	X	13.0	13.0	0.6111
2	Y	X	2.0	15.0	0.2402
1	X	Y	55.8	55.8	0.6111
2	X	Y	2.1	57.9	0.2402
1	X	X	91.4	91.4	0.6111
2	X	X	8.6	100.0	0.2402

Standardized Y Canonical Coefficients Section

	Y1	Y2
J100m	-0.219461	-0.400112
Pituush	-0.036214	-0.086809
Kuula	0.530630	-0.325379
Korkeush	-0.072431	-0.093412
J400m	-0.021783	0.561200
Jaidat	0.165191	0.596584
Kiekkko	0.384870	0.739975
Seivas	-0.254407	-0.088815
Keihas	-0.186678	0.123886
J1500m	-0.057736	0.632632

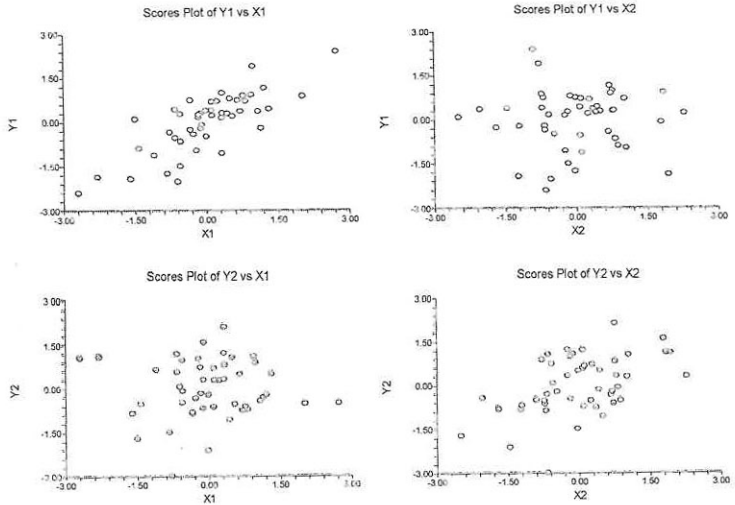
Standardized X Canonical Coefficients Section

	X1	X2
Pituus	0.292367	1.888604
Paino	0.739081	-1.762403

Variable - Variate Correlations Section

	Y1	Y2	X1	X2
J100m	-0.118692	-0.129244	-0.092785	-0.063348
Pituush	-0.069897	0.006931	-0.054640	0.003397
Kuula	0.900471	-0.168329	0.703921	-0.082505
Korkeush	0.199188	-0.097553	0.155710	-0.047815
J400m	-0.366547	0.524913	-0.286539	0.257280
Jaidat	0.207633	0.399702	0.162312	0.195910
Kiekkko	0.820637	-0.019333	0.641513	-0.009476
Seivas	-0.427450	-0.221253	-0.334148	-0.108445
Keihas	-0.075812	0.100730	-0.059264	0.049371
J1500m	-0.467800	0.528182	-0.365691	0.258883
Pituus	0.720901	0.189552	0.922193	0.386731
Paino	0.772523	-0.074983	0.988229	-0.152984

Plots Section



Correlation Report
Database

KYMMENOT5.S0

Pearson Correlations Section		(Row-Wise Deletion)		
	CanL1	CanL2	CanM1	CanM2
CanL1	1.000000	0.000000	0.781725	0.000000
CanL2	0.000000	1.000000	0.000000	0.490139
CanM1	0.781725	0.000000	1.000000	0.000000
CanM2	0.000000	0.490139	0.000000	1.000000
Cronbachs Alpha = 0.518304		Standardized Cronbachs Alpha = 0.518304		

5. Korrespondenssianalyysi

Seuraavilla sivuilla on esitetty Ncss-ohjelman tulostus eräästä korrespondenssianalyysistä.

Analysoitavana on havaintoaineisto, jossa joukko koululaisia on ristiintaulukoitu heidän silmiensä värin (5 luokkaa) ja hiustensa värin (4 luokkaa) mukaan.

- (a) Esittele korrespondenssianalyysin perusidea ja tavoite sekä korrespondenssianalyysin malli. Selitä myös mallin "estimointi".
- (b) Mitä johtopäätöksiä tulostuksesta voidaan tehdä?
- (c) Mitä tarkoittavat tulostuksessa (s. 18) käytetyt käsitteet laatu (Quality), massa (Mass) ja inertia (Inertia)?

Correspondence Analysis Report

Database HAIREYES.S0
 Variables Blue to Dark

Raw Data Section

Hair	Blue	Fair	Medium	Dark	Total
Fair	326	688	343	98	1455
Red	38	116	84	48	286
Medium	241	584	909	403	2137
Dark	110	188	412	681	1391
Black	3	4	26	85	118
Total	718	1580	1774	1315	5387

Row Profiles Section

Hair	Blue	Fair	Medium	Dark	Total
Fair	22.41	47.29	23.57	6.74	100.00
Red	13.29	40.56	29.37	16.78	100.00
Medium	11.28	27.33	42.54	18.86	100.00
Dark	7.91	13.52	29.62	48.96	100.00
Black	2.54	3.39	22.03	72.03	100.00
Total	13.33	29.33	32.93	24.41	100.00

Column Profiles Section

Hair	Blue	Fair	Medium	Dark	Total
Fair	45.40	43.54	19.33	7.45	27.01
Red	5.29	7.34	4.74	3.65	5.31
Medium	33.57	36.96	51.24	30.65	39.67
Dark	15.32	11.90	23.22	51.79	25.82
Black	0.42	0.25	1.47	6.46	2.19
Total	100.00	100.00	100.00	100.00	100.00

Eigenvalue Section

Factor	No.	Eigenvalue	Individual Percent	Cumulative Percent	Bar Chart
	1	0.199245	86.56	86.56	
	2	0.030087	13.07	99.63	
	3	0.000859	0.37	100.00	
	Total	0.230191			

Plot Detail Section for Rows

Name	Quality	Mass	Inertia	Factor	Axis1 COR	Axis1 CTR	Factor	Axis2 COR	Axis2 CTR
1 Fair	1.000	0.270	0.383	0.544	0.907	0.401	-0.174	0.093	0.271
2 Red	0.803	0.053	0.016	0.233	0.770	0.014	-0.048	0.033	0.004
3 Medium	1.000	0.397	0.078	0.042	0.039	0.004	0.208	0.961	0.572
4 Dark	1.000	0.258	0.401	-0.589	0.969	0.449	-0.104	0.030	0.093
5 Black	0.998	0.022	0.122	-1.094	0.934	0.132	-0.286	0.064	0.060

Plot Detail Section for Columns

Name	Quality	Mass	Inertia	Factor	Axis1 COR	Axis1 CTR	Factor	Axis2 COR	Axis2 CTR
1 Blue	0.979	0.133	0.111	0.400	0.836	0.107	-0.165	0.143	0.121
2 Fair	0.995	0.293	0.259	0.441	0.956	0.286	-0.088	0.039	0.076
3 Medium	0.999	0.329	0.088	-0.034	0.018	0.002	0.245	0.981	0.657
4 Dark	1.000	0.244	0.543	-0.703	0.965	0.605	-0.134	0.035	0.145

Principal Coordinate Section for Rows - Axis 1

Name	Mass	Inertia	Distance	Factor	COR	CTR	Angle	Eigenvalue
1 Fair	0.270	0.383	0.326	0.544	0.907	0.401	17.8	0.079929
2 Red	0.053	0.016	0.071	0.233	0.770	0.014	28.7	0.002889
3 Medium	0.397	0.078	0.045	0.042	0.039	0.004	78.6	0.000701
4 Dark	0.258	0.401	0.357	-0.589	0.969	0.449	10.1	0.089491
5 Black	0.022	0.122	1.282	-1.094	0.934	0.132	14.8	0.026235

Principal Coordinate Section for Rows - Axis 2

Name	Mass	Inertia	Distance	Factor	COR	CTR	Angle	Eigenvalue
1 Fair	0.270	0.383	0.326	-0.174	0.093	0.271	72.3	0.008163
2 Red	0.053	0.016	0.071	-0.048	0.033	0.004	79.5	0.000124
3 Medium	0.397	0.078	0.045	0.208	0.961	0.572	11.4	0.017213
4 Dark	0.258	0.401	0.357	-0.104	0.030	0.093	80.0	0.002790
5 Black	0.022	0.122	1.282	-0.286	0.064	0.060	75.3	0.001797

Principal Coordinate Section for Columns - Axis 1

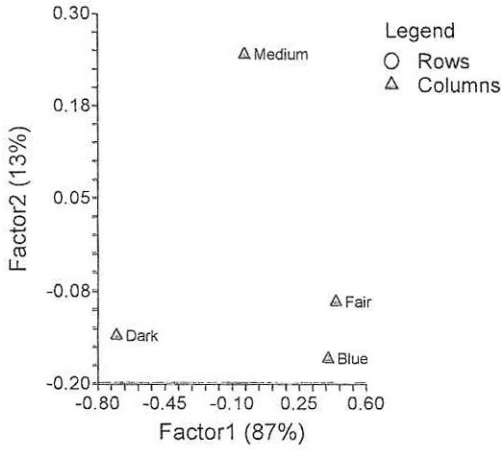
Name	Mass	Inertia	Distance	Factor	COR	CTR	Angle	Eigenvalue
1 Blue	0.133	0.111	0.192	0.400	0.836	0.107	23.9	0.021357
2 Fair	0.293	0.259	0.203	0.441	0.956	0.286	12.0	0.056965
3 Medium	0.329	0.088	0.061	-0.034	0.018	0.002	82.2	0.000372
4 Dark	0.244	0.543	0.512	-0.703	0.965	0.605	10.8	0.120550

Principal Coordinate Section for Columns - Axis 2

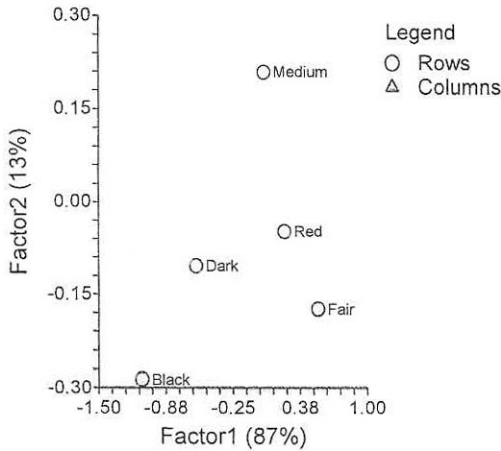
Name	Mass	Inertia	Distance	Factor	COR	CTR	Angle	Eigenvalue
1 Blue	0.133	0.111	0.192	-0.165	0.143	0.121	67.8	0.003647
2 Fair	0.293	0.259	0.203	-0.088	0.039	0.076	78.7	0.002295
3 Medium	0.329	0.088	0.061	0.245	0.981	0.657	7.9	0.019767
4 Dark	0.244	0.543	0.512	-0.134	0.035	0.145	79.2	0.004378

Plots

Correspondence Plot



Correspondence Plot



Correspondence Plot

