

Tartalomjegyzék

Bevezetés	11
A hasznos véletlen hiba	13
I. Adatredukációs módszerek	17
1. Főkomponens-elemzés.....	18
1.1. A főkomponens jelentése	25
1.2. Mikor használjunk főkomponens-elemzést?	25
1.3. A puding próbája	28
1.3.1. Az egyenlő teherviselés	30
2. Faktorelemzés	40
2.1. A modell javíthatása	49
2.2. Lássunk tisztán! A rotálás	53
2.3. Mentsük, ami menthető: a faktorszakók létrehozása	60
2.4. Amiért fáradtunk: létrejöttek a látens változók	62
2.5. A hiányzó adatok kezelése	64
2.6. A faktorelemzés buktatói	65
2.6.1. A Kaiser–Meyer–Olkin mutató és a Bartlett-teszt	66
2.6.2. Amikor a faktorok 100 százaléknál többet magyaráznak	71
2.6.2.1. A legnagyobb kommunalitású változó kihagyása	71
2.6.2.2. A rosszul interpretálható változók kihagyása	78
2.6.2.3. Az egyeduralomra törő változók esete ...	82
2.6.2.4. Vissza az elágazási ponthoz	85
2.6.3. A változók tartalma és a faktorok jelentése	89
2.6.4. A konfirmációs modell	91
2.6.5. Amikor minden reménytelen	96
2.6.5.1. Kétségbeesett kísérlet	101
2.6.5.2. Feladás helyett: visszahátrálás a főkomponensbe	103
2.7. Szezon és fazon	108

3. Klaszterelemzés	109
3.1. Tudás versus anyagi javak	109
3.2. Hierarchikus klaszterek	115
3.2.1. A hierarchikus klaszterezés módszerei	119
3.2.1.1. A legközelebbi, illetve a legtávolabbi szomszéd	119
3.2.1.2. Klaszterek távolsága mint a klaszterelemek távolságainak átlaga	122
3.2.1.3. Klaszterek távolságának meghatározása a klaszterközepponatok segítségével	125
3.2.2. Miben mérjük a távolságot?	129
3.2.2.1. Euklideszi távolság	129
3.2.2.2. Az euklideszi távolság négyzete	130
3.2.2.3. Asszociáció alapuló közelség-távolság ..	130
3.2.3. Sok hűhó majdnem semmiért	132
3.3. Klaszterezés nagy file-okon	133
3.3.1. Iniciális klaszterközepponatok	133
3.3.2. Útban a tökéletesség felé: iterálás	137
3.3.3. A klaszterek interpretálása	138
3.3.3.1. Kendőzetlen őszinteség versus a standardizálás álcája	141
3.4. Kilépés az életbe: típuskeresés klaszterekkel	143
3.4.1. És megint előlről: iterálás több lépésben	145
3.4.2. Névadás: klaszterek interpretálása	148
3.4.3. Adjunk-e profilt a bizonytalankodóknak? A hiányzó adatok kezelése	150
3.4.4. Egy tautológia két jelentése: ANOVA egy kicsit másképpen	154
3.4.5. Minden út Rómába vezet? Változtassunk a kezdőpontokon!	157
3.4.5.1. A magunk ásta verem és a kiút	159
3.5. Tanulságok	163

II. Magyarázó modellek	164
4. Variancia-analízis	166
4.1. Egyutas variancia-analízis	167
4.2. Kétutas variancia-analízis	175
4.3. Az interakció	176
4.3.1. Az interakcióról bővebben	178
4.3.2. A négyzetösszegek értelmezése a kétutas ANOVA-ban	180
4.3.3. Az interferencia	182
4.4. A hierarchikus ANOVA	188
4.5. Korlátozó feltételek	192
4.5.1. Normalitás-vizsgálat	194
5. Lineáris regresszióanalízis	204
5.1. A regressziós egyenes egyenlete	205
5.2. Egyszerű példa a lineáris regresszióra	209
5.3. Illeszkedés: mennyire vehetjük komolyan a regressziós egyenest?	211
5.4. Lineáris regresszióanalízis „élesben” – egyszerű modell	214
5.4.1. A „deviánsokról”	217
5.4.2. Vissza az output-hoz	220
5.5. Kétváltozós regressziós modell	222
5.5.1. Kétértékű vagy dummy változók	222
5.5.2. Modellépítés	223
5.5.3. Még egy kicsit a dummy változókról	226
5.6. Többváltozós regressziós modell	227
5.6.1. Ordinalis független változók	227
5.6.2. A Lazarsfeld-paradigma és a regresszió	230
5.7. Interferencia a regresszióban: a változószelektálás módszerei	234
5.7.1. A FORWARD módszer	235
5.7.2. A BACKWARD módszer	238
5.7.3. A STEPWISE módszer	241
5.7.4. Az ENTER módszer	244
5.8. Nominális mérési szintű változók becsempészése a regressziós modellbe: dummyszás	246

5.9.	Összefüggő független változók: a multikollinearitás	252
5.9.1.	Függvényszerű multikollinearitás	252
5.9.1.1.	A baj nem jár egyedül: magas mérési szintű és dummy változók a regressziós modellben	259
5.9.1.2.	Amikor csak a gondolkodás segít – dummyzzunk másként	266
5.9.1.3.	Szabaduljunk meg a nominális változóktól	269
5.9.2.	Függvényszerű kapcsolat – sztochasztikus multikollinearitás	271
5.9.2.1.	Miért félünk a sztochasztikus multikollinearitástól?	274
5.9.2.2.	Barátságosabb út a multikollinearitás meghatározására – kétféle mérőszám	277
5.9.2.3.	Hogyan védekezzünk a sztochasztikus multikollinearitás ellen?	282
5.10.	Hab a tortán: a homoszkedaszticitás és az elsőrendű autokorreláció	289
5.11.	Biztos, ami biztos: a regressziós modellek validitása	296
6.	Útmodellek	301
6.1.	Az utak erőssége	302
6.2.	Közvetlen és közvetett hatások	308
6.3.	Gondolkodási sémák	310
6.3.1.	Mi az ok, és mi az okozat?	313
6.3.2.	A modell továbbépítése	314
6.3.3.	A modell interpretációja	317
6.4.	Az útmodell korlátai	318
7.	Diszkriminancia-analízis	320
7.1.	A romák lehetséges beilleszkedése a többségi társadalomba: ahogyan a többség látja	320
7.1.1.	Sok kicsi sokra megy?	324
7.1.2.	A magyarázó változók együttes jelentése: a diszkrimináló függvények	327
7.1.3.	Az üdvözlendő diszkrimináció	333

7.1.4.	Sose lehetünk elég óvatosak – átlagok és eloszlások összehasonlítása	334
7.1.5.	Lássuk a medvét!	338
7.2.	Komplex modell	340
7.2.1.	Antiszemizmus és cigányellenesség	341
7.2.2.	Státusz és attitűdök: együttes magyarázat	347
8.	Többdimenziós skálázás	352
8.1.	Kisebbségek egy rokonszenv-térképen	353
8.1.1.	Többdimenziós skálázás az SPSS-ben	356
8.1.2.	Meddig mehetünk el? A dimenziócsökkentés határai	359
8.2.	Etnikai csoportok az egymásról alkotott vélemények terében	366
9.	Logisztikus regresszió	374
9.1.	A „sztahanovista” lineáris regresszió. Miért van szükség logisztikus regresszióra?	375
9.2.	A legfontosabb fogalmak	378
9.2.1.	A valószínűség és az esély	379
9.2.2.	Az esélyhányados	379
9.3.	Út a logisztikus regresszióanalízis felé	380
9.4.	Egy egyszerű modell	384
9.4.1.	A likelihood függvényről	386
9.4.2.	Az output érdemi része	389
9.4.3.	A modell magyarázóereje	391
9.4.3.1.	Találatmátrix a logisztikus regresszióban	394
9.4.4.	A független változó hatása	396
9.5.	A modell bővítése	398
9.5.1.	Összemérhető hatások: a standardizált regressziós együttható	401
9.6.	Kategoriális független változó a logisztikus regresszióban	404
9.7.	Mindent bele! Sokváltozós modell	410
9.8.	Ugyanaz másképpen	419

Melléklet	425
A parancsikonok használata: Főkomponens- és faktorelemzés ..	425
A parancsikonok használata: Klaszterelemzés	430
Hierarchikus klaszterezés	430
K-Means klaszterezés	434
A parancsikonok használata: Variancia-analízis	438
A parancsikonok használata: A lineáris regresszióanalízis	438
A parancsikonok használata: Diszkriminancia-analízis	442
A parancsikonok használata: A többdimenziós skálázás	446
A parancsikonok használata: A logisztikus regresszió	449
Felhasznált irodalom	452