

Exemples d'analyses factorielles des correspondances

Exemple 1

Le tableau suivant représente la couleur des cheveux et la couleur des yeux dans un échantillon de 592 individus.

	brun	chatain	roux	blond
marron	68	119	26	7
noisette	15	54	14	10
vert	5	29	14	16
bleu	20	84	17	94

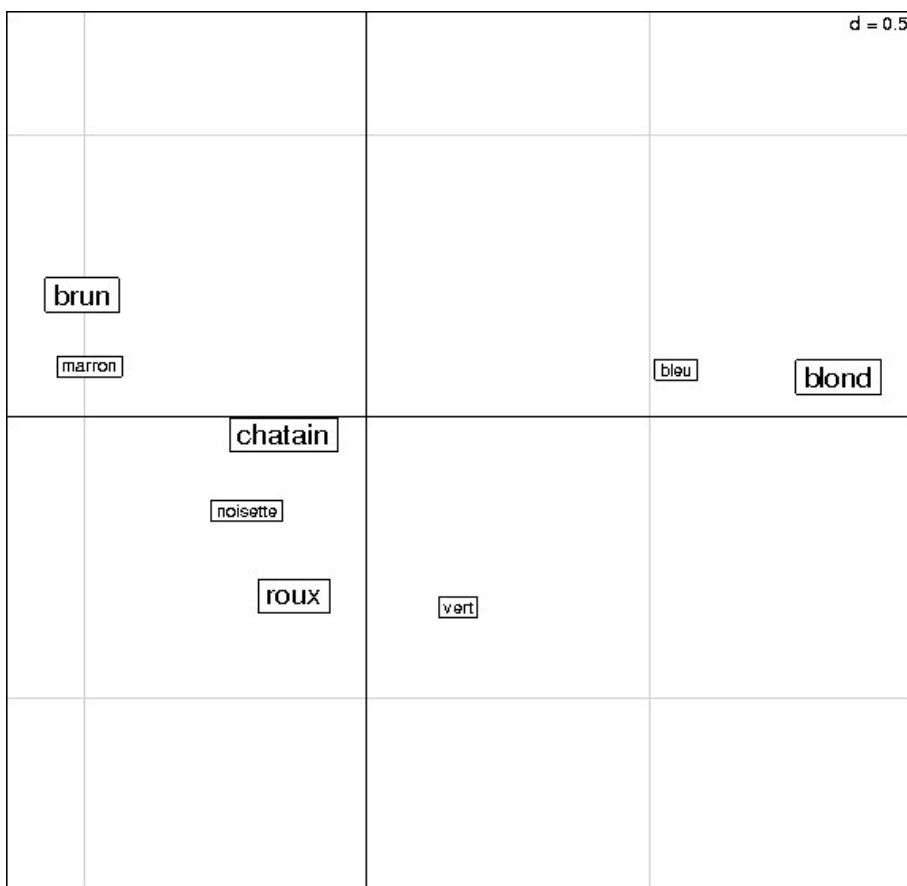
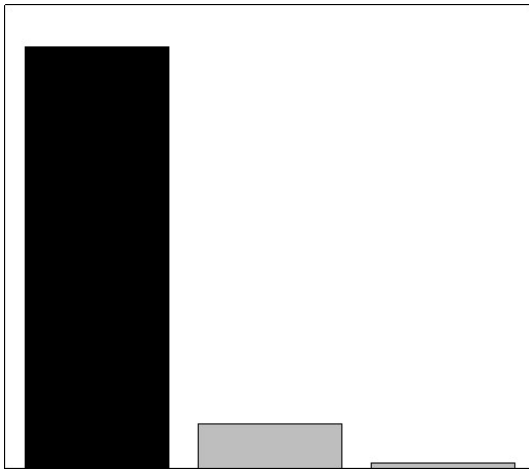
Commençons par effectuer un test d'indépendance. Voici comment procéder pour créer le tableau et effectuer le test du χ^2 avec le logiciel R. (On peut le faire en ligne: <http://pbil.univ-lyon1.fr/Rweb/Rweb.general.html>)

```
> tab<-matrix(c(68,119,26,7,
+ 15,54,14,10,
+ 5,29,14,16,
+ 20,84,17,94),ncol=4,byrow=TRUE)
> colnames(tab) <- c("brun", "chatain", "roux", "blond")
> rownames(tab) <- c("marron", "noisette", "vert", "bleu")
> summary(as.table(tab))
Number of cases in table: 592
Number of factors: 2
Test for independence of all factors:
Chisq = 138.29, df = 9, p-value = 2.325e-25
```

Que peut-on en conclure?

Voici le résultat d'une AFC (utilisant la librairie `ade4`) avec les graphiques des inerties (valeurs propres) et la représentation simultanée des deux premières coordonnées factorielles:

```
> afc <- dudi.coa(tab, scan = FALSE)
> afcin <- inertia.dudi(afc, col.inertia=T, row.inertia=T)
> afcin$TOT
      inertia      cum      ratio
1 0.208772652 0.2087727 0.8937273
2 0.022226615 0.2309993 0.9888764
3 0.002598439 0.2335977 1.0000000
> scatterutil.eigen(afc$eig, nf=3, box=T, sub="")
> scatter(afc, method=1, clab.row=0.90, clab.col=1.5, posieig="none")
```

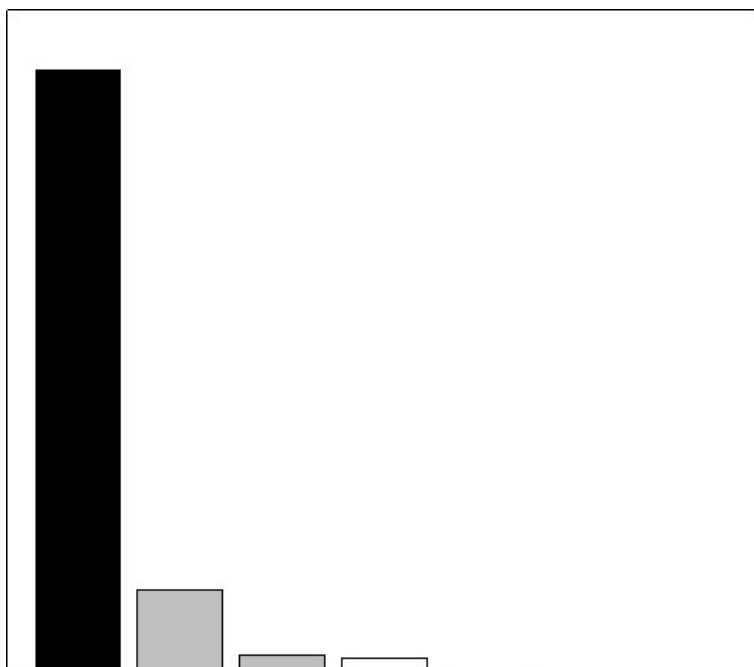


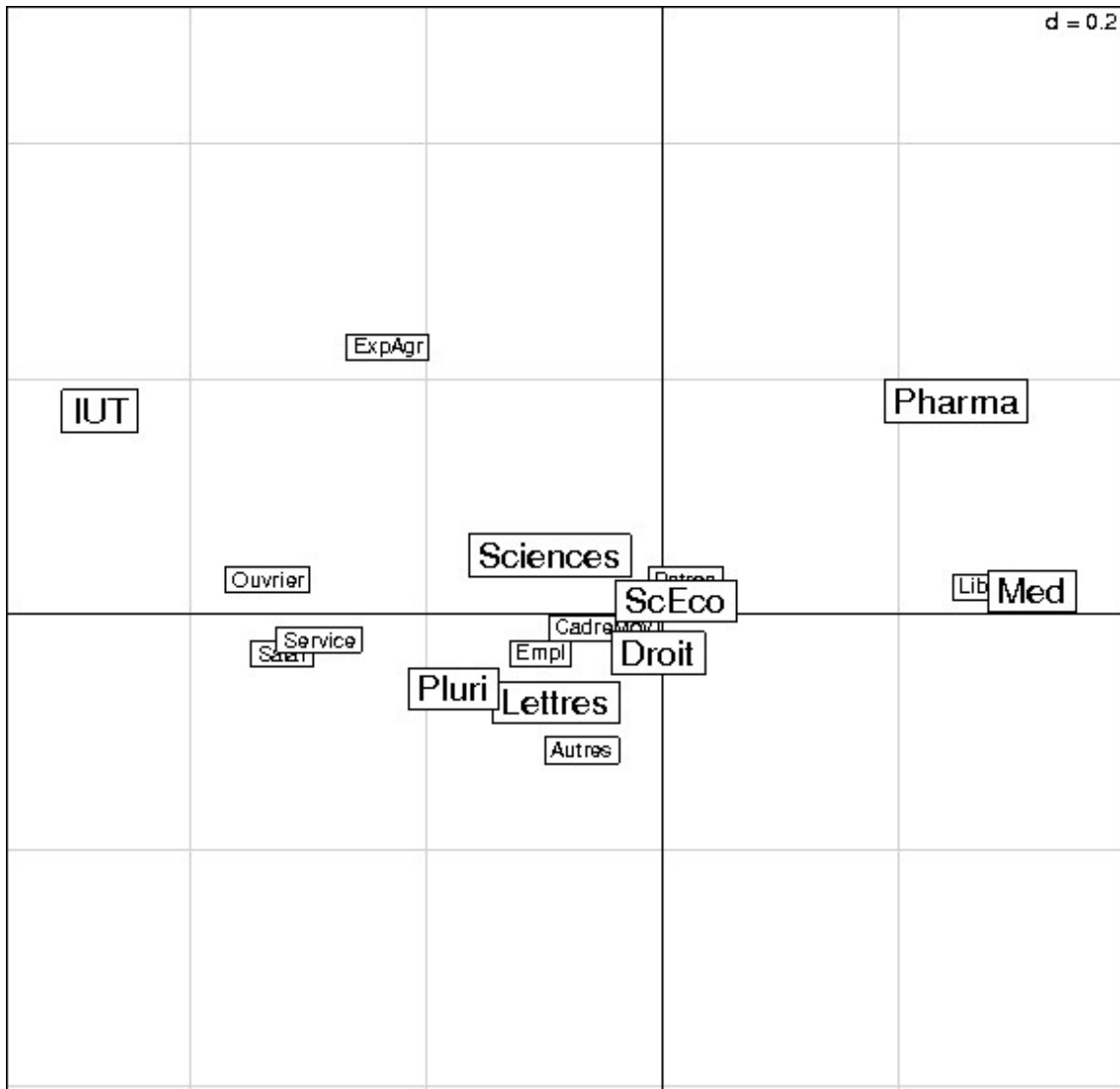
Exemple 2

Le second échantillon donne la répartition de 10'000 étudiants en termes de leur type d'études et de la catégorie socio-professionnelle du père:

	Droit	Sc. Eco	Lettres	Sciences	Med	Pharma	Pluridis.	IUT
Exp. Agricole	80	36	134	99	65	28	11	58
Salarié	6	2	15	6	4	1	1	4
Patron	168	74	312	137	208	53	21	62
Libéral	470	191	806	400	876	164	45	79
Cadre Moyen	236	99	493	264	281	56	36	87
Employé	145	52	281	133	135	30	20	54
Ouvrier	166	64	401	193	127	23	28	129
Service	16	6	27	11	8	2	2	8
Autres	305	115	624	247	301	47	42	90

```
> afc <- dudi.coa(tab, scan = FALSE)
> afcin <- inertia.dudi(afc, col.inertia=T, row.inertia=T)
> afcin$TOT
      inertia      cum      ratio
1 3.973716e-02 0.03973716 0.8371571
2 5.451690e-03 0.04518885 0.9520098
3 1.129689e-03 0.04631854 0.9758094
4 9.603978e-04 0.04727894 0.9960424
5 1.332353e-04 0.04741218 0.9988493
6 5.356108e-05 0.04746574 0.9999777
7 1.057199e-06 0.04746679 1.0000000
> scatterutil.eigen(afc$eig, nf=3, box=T, sub="")
> scatter(afc, method=1, clab.row=0.90, clab.col=1.5, posieig="none")
```





Exemple 3

Réf: Examen de Statistiques de mai 2004, Module MULT, Maîtrise de Psychologie, Université René Descartes. <http://piaget.psych.univ-paris5.fr/Statistiques/>

Les données qui suivent sont constituées par les résultats du premier tour des élections régionales de 2004 pour la région Ile de France. Pour chacun des huit départements de Ile de France (en lignes), on a les effectifs de suffrages pour chacune des huit listes candidates ainsi que les effectifs d'abstentions (en colonnes). L'objectif est d'analyser la structure des votes ainsi que les liaisons entre listes et départements. Voici les codes de désignation des départements et des listes :

Données : résultats du premier tour des régionales 2004 en Ile de France

Départements	Code	Listes	Tête de liste	Code
Paris (75)	Pari	PS-Verts-MRG-MRC	Huchon	Huch
Seine et Marne (77)	Smar	UMP	Copé	Cope
Yvelines (78)	Yvel	UDF	Santini	Sant
Essonne (91)	Esso	FN	Le Pen	Lepe
Hauts de Seine (92)	Htss	PC-AGR-AC	Buffet	Buff
Seine Saint-Denis (93)	Stde	LO-LCR	Laguiller	Lagu
Val de Marne (94)	Vdma	GE-Les Bleus	Pelegrin	Pele
Val d'Oise (95)	Vdoi	MNR	Bay	Bay
			Abstentions	Abst

	Huch	Cope	Sant	Lepe	Buff	Lagu	Pele	Bay	Absten
Pari	258495	184419	114222	57183	39052	22479	13277	5006	434078
Smar	128715	114003	48782	71897	25732	19738	11980	7085	301478
Yvel	150141	140634	96746	61676	23292	15998	13939	6486	329626
Esso	144581	95451	59967	54309	26732	17545	12108	5346	270414
Htss	143444	136677	122610	47279	32987	16438	11322	4690	314964
Stde	107327	61507	40081	54412	49535	19619	8393	5176	287618
Vdma	126569	93049	60234	47074	41897	17308	10969	4557	286913
Vdoi	111176	82524	47903	55165	24693	17018	9876	4825	262458

```
> afc <- dudi.coa(tab, scan = FALSE)
> afcin <- inertia.dudi(afc, col.inertia=T, row.inertia=T)
> afcin$TOT
      inertia      cum      ratio
1 1.512299e-02 0.01512299 0.6357802
2 4.656275e-03 0.01977927 0.8315330
3 3.406235e-03 0.02318550 0.9747333
4 3.491311e-04 0.02353463 0.9894110
5 1.518114e-04 0.02368644 0.9957932
6 7.570112e-05 0.02376214 0.9989757
7 2.436338e-05 0.02378651 1.0000000
> scatterutil.eigen(afc$eig, nf=3, box=T, sub="")
> scatter(afc, method=1, clab.row=0.90, clab.col=1.5, posieig="none")
```

