

Notice du programme
DISIND

Nom du programme	: DISIND
Date de mise à jour	: mai 1993
Numéro de la version	: 1.1
Langage	: FORTRAN 77
Auteurs	: S. BOCHI, G. CELEUX, A. MKHADRI

INRIA
78153 Le Chesnay Cedex

Tél. (1) 39 63 55 10

1. DESCRIPTION DE LA METHODE

1.1 But de la méthode

DISIND réalise une analyse discriminante décisionnelle pour des variables explicatives qualitatives par le modèle d'indépendance conditionnelle.

1.2 Références bibliographiques

BOCHI S., CELEUX G. et MKHADRI A.

Le modèle d'indépendance conditionnelle: le programme DISIND. La Revue de Modulad n°11, juin 1993.

CELEUX G. et NAKACHE J.-P. (éditeurs)

Analyse discriminante sur variables qualitatives. Dunod (à paraître, 1993).

2. INSTALLATION DU PROGRAMME

Paramètres à changer pour ajuster les dimensions ("tampons") :
on désigne par :

ITOT le nombre total d'observations

JMAX le nombre de variables explicatives

JTOT = JMAX + 1

JMD le nombre total de modalités des variables explicatives

KMAX le nombre de classes a priori

La modification des dimensions se fait dans l'instruction :

```
PARAMETER (mcara=xxx, mreel=yyy, mient=zzz)
```

avec :

- xxx majorant de KMAX + ITOT

- yyy majorant de JMAX + KMAX (ITOT + JMD +1)

- zzz majorant de ITOT(JTOT + 2) + KMAX(KMAX + JMD + 1) + JTOT + JMD

Dans le programme, xxx = 500, yyy = 2000, zzz = 5000.

3. NOTICE D'EMPLOI DU PROGRAMME

3.1 Liste des fichiers utilisés

Nom	Unité	Fonction
disind.par	LP = 10	paramètres
disind.don	LD = 20	données
disind.lis	IMP = 30	résultats

Les fichiers des paramètres et des données sont permanents.

3.2 Entrées

Fichier disind.par

La ligne 1 est lue par le sous-programme lparm. Les autres lignes sont lues par le sous-programme lparmd.

LIGNE 1	FORMAT (5(I2,1X))
IVAL	= 1 : pas de validation = 2 : validation par un échantillon test (soit, il est donné et apparaît à la fin du fichier de données, soit il est tiré au hasard) = 3 : validation par validation croisée
IPOUR	% de l'échantillon test s'il est tiré au hasard = 0 sinon
NBR	= 1 si les probabilités a priori des groupes sont égales = 2 si les probabilités a priori des groupes sont proportionnelles à leur taille = 3 si les probabilités a priori des groupes sont fixés par l'utilisateur
LIST	= 1 impression détaillée des affectations aux groupes = 0 sinon
NTIR	nombre d'ensembles tests si IVAL = 2 et IPOUR ≠ 0. = 0 dans le cas contraire

LIGNE 2 FORMAT (20A4)

Titre de l'étude

LIGNE 3 FORMAT (I5,I3,I5)

ITOT	nombre total d'individus (fichier test compris)
JTOT	nombre de variables (variable de groupe compris)
JTEST	nombre d'individus du fichier test s'il est présent et non tiré au hasard

LIGNE 4 FORMAT (10(A4,1X))

Noms des JTOT variables

LIGNE 5 FORMAT (20(I2, 1X))

Nombre de modalités des JTOT variables

LIGNE 6 FORMAT (10(A4,1X))

Noms des KMAX modalités de la variable à expliquer

LIGNE 7 FORMAT (20A4)

Format des données

LIGNE 8 FORMAT (10(F4.2,2X))

Les probabilités a priori des groupes si elles sont fixées par
l'utilisateur (cas NBR = 3)

Fichier *disind.don*

Ce fichier est lu par le sous-programme indon. Les données doivent être rangées de la façon suivante: chaque enregistrement représentant un individu commence par son identificateur suivi des valeurs de la variable à expliquer puis des variables

explicatives. De plus, s'il existe un échantillon test fixé à l'avance, il doit être placé à la fin du fichier.

DISIND : RESUME DES ENTREES			
FICHER	LIGNE	FORMAT	
Paramètres	1	5(I2,1X)	IVAL, IPOUR, NBR, LIST, NTIR
	2	20A4	Titre de l'étude
Données	3	15,I3,I5	ITOT, JTOT, ITEST
	4	10(A4,1X)	noms de modalités des variables
	5	20(I2,1X)	nombre de modalités des variables
	6	10(A4,1X)	noms des modalités de la variable à expliquer
	7	libre	Format de lecture des données
	8 év.	10(F4.2,2X)	Les probabilités a priori des groupes si NBR = 3

Exemple :

paramètres :

```
01 00 01 00 00
magasin
00412 04
icod var1 var2 var3 var4
02 02 02 02 02
icol ico2
(a4,1x,10i2)
```

observations :

```
i012 1 2 2 1 2
i016 1 2 2 1 2
i018 1 2 2 1 2
i025 1 2 2 1 1
i035 1 2 1 1 2
i041 1 1 2 2 2
i049 1 1 2 2 2
i052 1 1 2 2 2
i053 1 1 2 2 2
i056 1 1 2 2 1
i058 1 1 2 2 1
i063 1 1 2 1 2
i066 1 1 2 1 2
i067 1 1 2 1 2
i071 1 1 2 1 2
i073 1 1 2 1 2
i082 1 1 2 1 2
```

i083 1 1 2 1 2
i086 1 1 2 1 2
i088 1 1 2 1 2
i091 1 1 2 1 1
etc.

Sorties

Exemple de sortie correspondant aux données précédentes :

Discrimination sur l échantillon d apprentissage : 412
individus

Tableau de classement

		Groupes d'affectation	
		icol	ico2
Groupes d'origine	icol	100	54
	ico2	70	188
		icol	ico2

Pourcentages de classement

Pourcentage global de bien classes: 69.90

Pourcentages de bien classes par groupe

icol 64.94
ico2 72.87