RECONNAISSANCE DES FORMES (H, I7, K9)

(11 / 05 / 2020, © Monfort, Dicostat2005, 2005-2020)

La notion générale de **reconnaissance des formes** comporte deux aspects :

(a) la définition de la notion de « **forme** » (cf aussi **variable morphologique**) ;

(b) l'**identification des formes**, ie la **reconnaissance** d'une forme donnée (nouvelle ou inconnue) ainsi que sa différentiation par rapport à d'autres formes données (anciennes ou connues).

(i) Par **forme**, on peut entendre tout « **objet** », ou « **entité** », ou « **structure** », définis au sein d'un **domaine de connaissance**, et généralement associés à l'étude d'un **phénomène** qui relève de ce domaine. Ainsi :

(a) une forme mathématique : **structure** algébrique, topologique, mesurable, géométrique (surface analytique ou variété différentielle). A titre d'exemples : surfaces classiques (coniques, quadriques, etc), variétés « atypiques » (tore, bande de A.F. MÖBIUS, bouteille de C.F. KLEIN, etc) ;

(b) une forme physique : structure physique (eg structure cristalline, autres formes minérales) ;

(c) une forme biologique : structure biologique. A titre d'exemples : structure virale, bactérienne, fongique (champignon), fossile, génétique (« cartes » ADN / ARN) ;

(d) une forme anthropologique : structure osseuse, empreinte palmaire, etc.

(ii) La **reconnaissance** est un moyen pour déterminer si des formes données (relevant de la même science et du même « type ») sont :

(a) (strictement) identiques : dans ce cas, les formes considérées seront assimilées (ou identifiées) l'une à l'autre ;

(b) ou non (strictement) identiques : dans ce cas, on considère que les formes appartiennent à des « classes » distinctes.

La « ligne de démarcation » est donc cruciale à déterminer dans un **contexte statistique** de ce type. En fait, on se contente parfois seulement de formes « proches », ou « voisines », ou acceptées comme équivalentes : la reconnaissance dépend de la qualité d'**observation** des formes considérées.

(ii) En **Statistique**, la reconnaissance des formes reçoit plusieurs sens, plus spécialisés (cf aussi **classification, forme**) :

(a) **analyses (géométriques) des données** : eg **afc**, **acp** ou **analyse de PROCUSTE**. Dans cette approche, une « **forme** » est un **facteur** (ou un « **axe factoriel** »), en général « **caché** », dont la signification peut correspondre à une interprétation concrète par l'homme de l'art (« **intelligence** » en psychologie, etc). Le **nuage de points**, qui fait l'objet de ce type d'analyses, est alors « **situé** » par rapport à de telles formes, et inversement ;

(b) **méthodes de classification**. Ces méthodes consistent :

(b)₁ soit à définir des classes dans une **population statistique** donnée ou dans un **échantillon** donné (cf **classification automatique**, **classification statistique**) ;

(b)₂ soit à affecter une **unité statistique** nouvelle à l'une des classes pré-constituées d'une **population** ou d'un **échantillon** (cf **discrimination**).

Dans les deux cas, les classes considérées peuvent s'interpréter comme des « **formes** », puisque leur « **homogénéité** » implique une certaine « **identité** » de structure, ou un certain voisinage entre structures ;

(c) **méthodes d'estimation** : la **théorie de l'estimation** statistique elle-même a pour objet le « **repérage** » ou l'« **identification** » d'un **espace probabilisé** au sein d'une famille de tels espaces (ie formalisme défini au sein d'un **modèle statistique**). Dans cette approche :

(c)₁ une « **forme** » est soit l'un des espaces en question, soit un **paramètre d'intérêt** relatif à la **loi** de cet espace ;

(c)₂ une « **reconnaissance** » consiste à estimer la loi associée à l'espace probabilisé en question, ou seulement à estimer un paramètre de celle-ci. En fait, l'estimation consiste à « **reconnaître** » le (« **vrai** ») modèle qui est à l'origine des observations. Les **tests associés** aux problèmes d'estimation permettent alors de valider (ie de confirmer ou d'infirmar) cette reconnaissance (cf **identification**, **modèle identifié**).

(iii) Les techniques de reconnaissance des formes interviennent ainsi dans de nombreux domaines, eg :

(a) physique : géologie, astrophysique ;

(b) biologie : imagerie médicale, type de pathologies (génétique, métabolique, cardio-vasculaire, carcinogénèse, épidémique, etc) ;

(c) écologie : classification des espèces et biodiversité ;

(d) psychologie : **profils de « personnalité »**, identification de comportements extrêmes (eg à risques) ;

(e) sociologie : linguistique (reconnaissance de caractères typographiques ou de racines linguistiques), activités bancaires (reconnaissance de signatures), administration de la police (reconnaissance d'empreintes digitales ou de profils ADN), anthropologie (archéologie, paléontologie).