THÉORIE DE LA DÉCISION (A14, G)

Au sens le plus large, la **théorie de la décision** a pour objet le choix, par un **décideur** des meilleures **décisions** (ou **actions**) possibles pour atteindre un (ou plusieurs) **objectif(s)** : minimiser un **cout**, maximiser une **fonction d'utilité** (cf aussi **utilité**), etc, en fonction des moyens ou de l'**information** disponibles.

Le décideur est, en général, une **entité « consciente »** (personne physique, personne morale), qui dispose de facultés d'action, ou parfois une **entité non consciente** (eg unité biologique, etc), qui dispose de facultés de réaction.

La **théorie des jeux** peut alors représenter un cadre formalisé permettant de décrire (a) l'ensemble des décideurs concernés, l'ensemble des décisions possibles, parfois appelé **espace de décision**, (b) l'ensemble des **instruments** (ou moyens) dont dispose chaque décideur pour agir dans un certain sens, ou (c) l'ensemble des **informations** dont il peut faire usage.

(i) Au sens étroit, la **théorie de la décision** ne concerne qu'un seul individu (eg **homme de l'art**) qui, en fonction des informations disponibles, doit prendre des décisions pour réaliser des objectifs. Lorsque tout ou partie des éléments de la prise de décision (ensemble des actions, ensemble des informations, résultats des actions), sont considérés comme aléatoires (ie engendrés par un **schéma probabiliste**), on parle alors de **théorie de la décision statistique** ou simplement de théorie de la décision. On dit aussi que le décideur (eg le **statisticien**), prend ses décisions en « **univers aléatoire** ».

(ii) La théorie de la décision (statistique) adopte un cadre formel dont une description élémentaire est la suivante. On définit un **problème de décision** à l'aide des éléments suivants (cf aussi **décision**, **décision statistique**, **règle de décision**) :

(a) une **représentation statistique** de base $(\Omega, \mathcal{T}, \mathcal{P})$, paramétrique ou non, qui consiste à observer des événements $A \subset \Omega$, éléments d'une tribu $\mathcal{T}$ de parties de Ω, qui sont régis par une (mesure de) **probabilité** P définie sur $\mathcal{T}$. Pour chaque $P \in \mathcal{P}$, le triplet $(\Omega, \mathcal{T}, P)$ est parfois appelé **expérience aléatoire**. Lorsque $\mathcal{P}$ se présente sous forme paramétrée, ie indicée par un ensemble Θ dont tout élément θ est appelé **paramètre** du modèle, selon $\mathcal{P} = (P_\theta)_{\theta \in \Theta}$, l'ensemble Θ est souvent muni d'une **tribu de parties** notée $\mathcal{B}_\Theta$, sur laquelle est définie à son tour une **mesure de probabilité** Π, dite **probabilité a priori**, ou **probabilité de BAYES** (cf **école bayésienne**) ;

(b) un **espace mesurable** $(\mathcal{X}, \mathcal{B})$, dit **espace d'observation(s)** ou **espace d'échantillonnage** (notamment lorsqu'il revêt une forme d'espace produit ou d'espace puissance) ; ainsi, $\mathcal{X}$ est souvent de la forme $\mathcal{X} = \Pi_{n=1}^{N} \mathcal{X}_n$;

(c) une **variable aléatoire** (ou **statistique** de base) $X : \Omega \mapsto \mathcal{X}$, ie une application $(\mathcal{T}, \mathcal{B})$-mesurable, généralement observable, appelée **échantillon**. Lorsque l'**événement** élémentaire $\omega \in \Omega$ est réalisé, la valeur $X(\omega) = x \in \mathcal{X}$ est aussi appelée **réalisation** de l'échantillon X, ou même simplement **échantillon** ;

(d) un **ensemble** de décisions (ou d'actions) D, muni d'une tribu de parties $\mathcal{B}_D$. L'espace mesurable $(D, \mathcal{B}_D)$ est appelé **espace de (s) décision (s)** (cf **espace de décision**). A partir de cet espace, on définit un ensemble Δ dont les éléments, appelés **fonctions de décision pures**, ou **fonctions de décision non aléatoires**, sont des applications $(\mathcal{B}, \mathcal{B}_D)$-mesurables $\delta : \mathcal{X} \mapsto D$ qui associent à tout échantillon $x = X(\omega) \in \mathcal{X}$, une décision (ou action) $d \in D$, avec $d = \delta(x)$ (cf **règle de décision pure**). Chaque fonction de décision pure $\delta \in \Delta$ induit sur $\mathcal{B}_D$ une (mesure de) probabilité image (cf **mesure image**) : en effet, à chaque probabilité $P \in \mathcal{P}$, la va X associe la probabilité $P^X = X(P)$, image de P par X, qui n'est autre que l'une des **lois de probabilité** possibles de X. Par suite, l'image de P^X par δ, ie $\delta(P^X) \mapsto \delta(X(P)) = (\delta \circ X)(P) = P^{\delta \circ X}$, est la loi de probabilité de la variable aléatoire $\delta(X)$ ou $\delta \circ X$. Si $\mathcal{P}$ se présente sous forme paramétrée $(P_\theta)_{\theta \in \Theta}$, la loi $P_\theta^{\delta \circ X}$ de cette variable dépend donc aussi, en général, de $\theta \in \Theta$.

Par ailleurs, on peut aussi définir, à l'aide de l'espace $(D, \mathcal{B}_D)$, des **fonctions de décision aléatoires**, ou **fonctions de décision mixtes**, qui associent à chaque échantillon $x = X(\omega) \in \mathcal{X}$ une (mesure de) probabilité m_x définie sur $\mathcal{B}_D$ (cf **règle de décision mixte**). On note Δ_M l'ensemble des règles de décision aléatoires ;

(e) enfin, une **fonction de perte** $L : D \times \mathcal{P} \mapsto \mathbf{R}_+$ tq la valeur $L(d, P)$ représente la « perte » encourue par le statisticien lorsque sa décision est d alors que la « **Nature** » (joueur adverse fictif) a choisi P dans $\mathcal{P}$. Si $\mathcal{P}$ est paramétrée, on note plutôt $L(d, \theta)$ la perte correspondant à la probabilité P_θ, où $\theta \in \Theta$. L'application $L : \Omega \times \mathcal{P} \mapsto \mathbf{R}_+$ (avec même notation L), définie par $(\omega, P) \mapsto L\{\delta \circ X(\omega), P\}$ (resp par $(\omega, \theta) \mapsto L\{\delta \circ X(\omega), \theta\}$) est donc une **vars**
(positive) qui joue un rôle important dans la théorie ;

Cette dernière permet alors de définir des **règles de décision** optimales : il existe divers **critères** d'**optimalité** (cf **risque bayésien**, **règle minimax**, etc).

La théorie utilise aussi largement (a) la notion de **fonction de vraisemblance**, lorsque $\mathcal{P}^X = (P_\theta^X)_{\theta \in \Theta}$ est une **famille de lois dominée** par une mesure donnée, ou encore (b) la notion d'**information** déduite d'un échantillon (**information de FISHER**, **information de KULLBACK**, etc).

Par ailleurs, selon que la **« taille »** N de l'échantillon $X = (X_1, ..., X_N)$ est donnée ou qu'elle peut varier en fonction des besoins de l'expérimentateur, on parle de **théorie statique de la décision** ou de **théorie dynamique de la décision** (cf **analyse séquentielle**, **problème de décision séquentielle**). Enfin, il est parfois possible de « réviser » une décision au vu d'observations nouvelles (cf **décision adaptative**).

(iv) Les applications « classiques » de la théorie de la décision sont la **théorie des tests**, la **théorie de l'estimation**, les **classifications**, ou encore de la **prévision**, la **théorie des sondages** ou la **théorie des plans d'expérience**.