THÉORIE DE L'INFORMATION (G6)

(25 / 04 / 2020, © Monfort, Dicostat2005, 2005-2020)

La notion courante d'**information** implique, à la fois, un « stockage » et un « échange » de « **données significantes** » entre entités différentes (cf **donnée**). Cette notion reçoit un sens plus précis en **calcul des probabilités** aussi bien qu'en **Statistique**.

Dans un sens large, la **théorie de l'information** a pour objet, l'étude du mécanisme relatif à un « message » circulant entre un « émetteur » (eg **unité statistique**) et un « récepteur » (eg **statisticien**). Elle entretient donc des rapports étroits avec la **théorie du signal**.

(i) Le **message** en question est souvent de nature aléatoire. Cette particularité est à l'origine de la **théorie de l'information statistique**. Dans un sens technique, la théorie de l'information est alors centrée sur l'étude des propriétés de la notion d'**information** et des concepts qui lui sont liés. Cette notion d'information peut s'associer à la notion physique d' « énergie » utilisée par un **système** (cf **énergie d'un système**) ou à la notion probabiliste de **dispersion** (ou de **variabilité**, ou encore de **variance**) d'une **variable aléatoire** ou d'une **statistique**. C'est pourquoi elle intervient dans le cadre des modèles usuels : **modèles statistiques**, **processus stochastiques** (cf **modèle de processus**), etc.

Ainsi, étant donné un **ensemble** Ω (**population** constituée d'individus, de particules de matière, etc), la connaissance (ou l'**observation**) d'une **partie** (non vide) $A \subset \Omega$ est dite apporter une « **information** » (en général non « nulle ») sur Ω . Par exemple, si $\text{Card } \Omega = M < +\infty$ et si $\text{Card } A = N \leq M$, l'information sur Ω disponible à partir de A peut dépendre de la **proportion** N / M de l'**information « totale »** (ie celle qu'apporte la connaissance de Ω), ou même être simplement égale à N / M .

On peut donc mesurer cette information par le nombre N , ou par toute fonction de N satisfaisant à des conditions circonstancielles (cf eg **information de HARTLEY**). Ces concepts sont donc uniformes, dans la mesure où les unités apportent chacune la même « quantité » d'information.

(ii) Cependant, la notion précédente est insuffisante lorsqu'on mesure, sur chaque élément $\omega \in \Theta$, la valeur $\xi(\omega)$ d'un **caractère statistique** ξ . Si $(\mathcal{X}, \mathcal{B})$ est un **espace d'observation** et si $\mathcal{F}$ est une **tribu de parties** de Ω , on suppose que $\xi : \Omega \mapsto \mathcal{X}$ est une **application mesurable** (**variable aléatoire**) et **observable**. Si P^ξ représente la **loi de probabilité** de ξ , alors la connaissance du couple (A, P^ξ) est susceptible de fournir une base pour définir l'information apportée par l'**expérience aléatoire** $(\Omega, \mathcal{F})$ ou par $(\mathcal{X}, \mathcal{B})$.

C'est ainsi qu'ont été définis des concepts d'information à partir d'un **espace probabilisé** tq $(\mathcal{X}, \mathcal{B}, P^\xi)$ ou même, plus généralement, à partir de modèles statistiques de la forme $(\mathcal{X}, \mathcal{B}, \mathcal{P}^\xi)$, ou $\mathcal{P}^\xi$ représente une **famille** de lois de probabilité.

(iii) Les notions d'information les plus communes sont l'**information de HARTLEY**, l'**information de FISHER**, l'**information de KULLBACK-LEIBLER**, l'**information de SHANNON**.

Une mesure importante de l'information est l'**entropie**. On utilise parfois les termes d'**entropie** ou de **divergence** à la place du terme « information ».

(iv) Enfin, on définit des notions voisines tq :

(a) l'**information relative** et l'**information absolue** ;

(b) l'**information conditionnelle** (à un événement ou à une va) (cf **conditionnement**).