THÉORIE DES PROBABILITÉS (B, E1, E4)
(31 / 01 / 2020, © Monfort, Dicostat2005, 2005-2020)

La **théorie des probabilités** est une branche fondamentale de la **théorie de la mesure**. Une **probabilité** est, en effet, une **mesure abstraite** particulière, ie une mesure dont la « masse » totale est égale à un. Mais son rôle dans les sciences est central.

(i) Historiquement, la théorie s'est progressivement introduite et développée avec l'étude des **jeux de « chance »** (cf **théorie des jeux**) : dés, cartes, urnes, roulette, etc. Il s'agissait essentiellement de probabilités « discrètes », ie calculées par dénombrements.

La notion de probabilité était indissociable de l'aspect « fréquentiste » qui prédominait alors. Par ailleurs, l'**espace fondamental** (celui des « épreuves »), de la forme $(\Omega, \mathscr{T}, P)$, n'était pas toujours distingué de l'**espace image** du précédent par une **variable aléatoire** ξ à valeurs dans un ensemble $\mathscr{X}$ donné, ie dans ujn espace de la forme $(\mathscr{X}, \mathscr{B}, P^{\xi})$ (**espace des observations**, **espace des états**) (cf aussi **application mesurable**, **image d'une application**, **mesure image**).

Les divergences existant entre les divers « **probabilistes** » (personnes étudiant ce type de phénomènes) provenaient généralement :

(a) soit de la définition de l'**ensemble fondamental** Ω, ou de celle des **événements** $A \subset \Omega$ (avec $A \in \mathscr{B}$), à prendre en compte ;

(b) soit de la définition (donc le calcul) des **probabilités** $P(A)$ attachées à ces événements ;

(c) soit encore de la définition de la variable ξ à observer et quantifier.

Les variables aléatoires analysées étaient, en général, des **variables discrètes** (eg nombres de « cas favorables », nombres de « cas possibles »), dont se déduisaient d'autres variables, souvent à valeurs dans **D** (nombres décimaux) ou **Q** (nombres rationnels) (cf **quotient**, **rapport**).

L'étude ultérieure des problèmes de dénombrement et l'étude des tirages dans des urnes (**schémas d'urne**, **sondages**) a conduit à développer l'analyse des probabilités sur des ensembles au plus dénombrables (ie finis ou dénombrables).

(ii) Plus tard, les **axiomes du calcul des probabilités** ont été établis et reliés à ceux de la **théorie de la mesure** (axiomes de A. N. KOLMOGOROV) : **théorème des probabilités totales**, **théorème des probabilités composées**, axiome des **probabilités conditionnelles**, etc.

La définition des variables analysées s'est étendue dans plusieurs directions :

(a) **variables continues**, et non plus seulement entières ;

(b) variables multiples (ou « multivariées »). On peut en particulier citer les variables « vectorielles » ou **variables multidimensionnelles** (cf **vecteur aléatoire**) ;

(c) **variables qualitatives** diverses et **variables morphologiques** (eg variables « géométriques », à valeurs dans des ensembles de figures) (cf **probabilité géométrique**) ;

(d) variables « fonctionnelles » (ie à valeurs dans des familles d'applications) ;

(e) **processus stochastiques** divers.

On peut ainsi étudier les **lois de probabilité** de « variables » aussi diverses que des surfaces (ou des volumes) (cf lancer d'un lacet fermé), des « formes » (eg morphologies animales, empreintes digitales), etc.

(iii) La théorie des probabilités est, aujourd'hui, une synthèse des approches précédentes. Elle comporte donc l'**analyse combinatoire**, et résulte aussi de la théorie de la mesure et de l'**intégration**. Le concept de base est celui d'**espace probabilisé**, noté $(\Omega, \mathcal{T}, P)$: un tel espace comporte, en effet, un ensemble fondamental Ω (celui des résultats possibles), une famille $\mathcal{T}$ contenant certaines parties de Ω, et une **mesure de probabilité** P décrivant la distribution des « masses » (ou des fréquences) entre les parties A qui composent $\mathcal{T}$.

(iv) A la différence de la **Statistique**, le calcul des probabilités suppose « donnés » ou « connus » tous ces concepts : en particulier, P est, en principe, entièrement définie (ou définissable, ou calculable) à l'aide d'un procédé ou d'un algorithme spécifié. Ainsi, dans l'étude du lancer d'un dé « parfait » à six faces, on a $\Omega = \{F_1, ..., F_6\}$, $\mathcal{T} = \mathcal{P}(\Omega)$ (ensemble des parties de Ω) et $P(\{\omega = F_i\}) = 1 / 6$; par suite, l'étude (eg) de la probabilité de h-ième apparition d'une face donnée (eg F_5) au cours de n tirages (avec $n \geq 6$) se déduit des probabilités « uniformes » précédentes.

(v) La théorie des probabilités s'est développée dans de nombreux domaines, notamment :

(a) **théorie des processus** (cf aussi **modèle de processus**) (cf infra) ;

(b) **Statistique**, où le concept de **modèle statistique** généralise celui d'espace probabilisé. En effet, dans ce cas, l'espace $(\Omega, \mathcal{T}, P)$ est remplacé par une **représentation statistique** de la forme $(\Omega, \mathcal{T}, \mathcal{P})$, dans laquelle $\mathcal{P}$ est une famille de probabilités, l'une d'elles étant susceptible de « gouverner » le tirage des parties $A \in \mathcal{T}$. S'il est possible de mesurer une grandeur $x = \xi(\omega) \in \mathcal{X}$ sur chacune des unités $\omega \in \Omega$, le modèle étudié est souvent le modèle $(\mathcal{X}, \mathcal{B}, \mathcal{P}^{\xi})$, dans lequel la

« vraie » loi de probabilité P^{ξ} générant les données x précédentes est inconnue, donc à préciser. L'objectif du **statisticien** est donc différent de celui du probabiliste.