

## STATISTIQUE INVARIANTE (G1)

(11 / 10 / 2020, © Monfort, Dicostat2005, 2005-2020)

(i) Soit  $(\mathcal{X}, \mathcal{B})$  un **espace d'observation**,  $\mathcal{G}$  un **groupe de transformations** mesurables sur  $\mathcal{X}$  (ie  $g : \mathcal{X} \mapsto \mathcal{X}$ ),  $(\mathcal{S}, \mathcal{E})$  un **espace mesurable** auxiliaire et  $s : \mathcal{X} \mapsto \mathcal{S}$  une **application mesurable** définissant la **statistique**  $S = s \circ X$ , où  $X : \Omega \mapsto \mathcal{X}$  est une **va** donnée (**échantillon**).

On dit que  $s$  (resp  $S$ ) est une **statistique invariante** pour le groupe  $\mathcal{G}$  ssi :

$$(1)_a \quad t(g(x)) = g(t(x)), \quad \forall x \in \mathcal{X} \text{ et } \forall g \in \mathcal{G},$$

ie ssi :

$$(1)_b \quad t \circ g = g \circ t, \quad \forall g \in \mathcal{G}.$$

(ii) On établit les propriétés de base suivantes :

(a) une statistique invariante  $s$  (resp  $S$ ) est constante sur chaque **orbite** de  $x \in \mathcal{X}$ , ie  $\text{Orb}(x) = \{g(x) \in \mathcal{X} : g \in \mathcal{G}\}$ . Inversement, toute statistique  $s : \mathcal{X} \mapsto \mathcal{S}$  constante sur chaque orbite  $\text{Orb}(x)$  (où  $x$  parcourt  $\mathcal{X}$ ) est une statistique invariante ;

(b) toute statistique invariante est fonction d'une statistique invariante maximale ;

(c) le groupe  $\mathcal{G}$  induit sur  $\mathcal{S}$  un groupe  $\tilde{\mathcal{G}}$  de transformations mesurables  $g \sim : \mathcal{S} \mapsto \mathcal{S}$  défini par :

$$(2) \quad \{g \sim \in \tilde{\mathcal{G}}\} \Leftrightarrow \{\exists g \in \mathcal{G} : g \circ t = t \circ g\}.$$

(iii) L'intérêt de la **propriété d'invariance** pour une statistique  $S$  tient à la remarque suivante : si  $P^X$  est la **lp** de  $X$ , la lp

$$(3) \quad P^S = s(P^X) = P^{s \circ X}$$

de  $S$  dépend, en général, de moins de paramètres que  $P^X$  (cf aussi **principe de parcimonie**).

D'autre part, lorsque  $g$  représente un changement d'**échelle** (eg nouvelle **unité de mesure**), l'**inférence statistique** effectuée à partir de  $S$  (estimation, tests, prévision, classification, etc) ne dépend pas de l'échelle utilisée (cf aussi **statistique équivariante**).