

JOURNAL DE LA SOCIÉTÉ STATISTIQUE DE PARIS

JEAN DUFRÉNOY

Application des principes adansonniens

Journal de la société statistique de Paris, tome 104 (1963), p. 212-213

http://www.numdam.org/item?id=JSFS_1963__104__212_0

© Société de statistique de Paris, 1963, tous droits réservés.

L'accès aux archives de la revue « Journal de la société statistique de Paris » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFdS>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

APPLICATION DES PRINCIPES ADANSONIENS

Un programme de classification peut, initialement, s'inspirer des principes adansonien en appliquant, à l'ensemble des matériaux à classer, chacun des tests capables d'évoquer, chez chaque individu soumis à l'épreuve, la possibilité de réaction (+) ou celle de réaction (-); mais ces tests ayant été appliqués, l'efficacité relative de chacun peut être évaluée en fonction de sa valeur discriminatoire, permettant de distinguer les individus d'une classe de ceux de l'autre classe; en fait la classification ne porte jamais sur les individus eux-mêmes, mais sur des caractéristiques individuelles dont nous pouvons acquérir la connaissance.

Chaque test correspond à une « opération » capable de diviser un ensemble en deux classes ou deux catégories, et permettant de « répartir » les individus entre l'une ou l'autre catégorie; un groupe de tests met en œuvre une série de « répartitions »; les tests peuvent être appliqués successivement selon une séquence déterminée, pour permettre une série de dichotomies.

1 a. Un certain test évoquant la possibilité de ces deux issues (+) ou (-) étant appliqué à un ensemble de N individus, l'issue (\pm) étant considérée comme « succès » avec probabilité (p), cette probabilité p peut prendre toute valeur entre 100 % et 0 %;

1 b. Si le test correspond à une opération permettant la répartition entre plus de

deux catégories, A. Rescigno et G. A. Maccaro évaluent l'efficacité opérationnelle du test, c'est-à-dire son aptitude à provoquer une répartition des (N) individus entre (c) classes, par cette évaluation, $H = - \sum \frac{r_t}{n} \log \sum \frac{r_t}{n}$ lorsque r_{in} correspond à la fréquence de réaction de type i au test (t).

2. Au lieu d'un seul test, fournissant une seule source d'Information quant aux répartitions possibles, on en utilise plusieurs, deux, par exemple : la répartition obtenue à la suite de l'application successive du test t_1 , puis du test t_2 , correspond à la première répartition rendue possible par le test t_1 , telle que « A » ou « non A », « B » ou « non B »..., puis à la répartition entre classes que peut rendre possible le deuxième test t_2 appliqué aux individus de la classe « A » ou de la classe « non A », de la classe « B » ou de la classe « non B ». Si un test relatif à la forme évoque deux issues possibles « long » ou « rond » et si un deuxième test évoque deux issues possibles quant à la couleur (« pourpre » ou « rouge ») la suite des deux tests évoquera la possibilité de quatre types, à moins de forte corrélation entre les deux caractères de forme et de couleur.

(The Information content of Biological Classifications, in *Information Theory*, 4th London Symposium, 1961.)

J. DUFRENOY.