

JOURNAL DE LA SOCIÉTÉ STATISTIQUE DE PARIS

JEAN DUFRÉNOY

Le « statisticien » et son « environnement »

Journal de la société statistique de Paris, tome 112, n° 2 (1971), p. 136-137

http://www.numdam.org/item?id=JSFS_1971__112_2_136_0

© Société de statistique de Paris, 1971, 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/>

LE « STATISTICIEN » ET SON « ENVIRONNEMENT »

La différence de langage rend difficile la « communication » entre celui qui doit faire évaluer statistiquement ses résultats d'expérience, qu'il est lui-même incapable d'analyser, et le « statisticien professionnel », qu'il consulte.

D'après LYON HYAMS, *Biometrics* 27 : 201 : 11, march 1971, ces professionnels relèvent des stéréotypes suivants :

1. Le statisticien « normal » qui, quel que soit le problème qu'on lui soumet, commence par dessiner une « courbe en cloche » : ce monomane, pour qui toute distribution devrait être « normale », en est resté à ce qu'on savait au temps de LAPLACE ; son ambition est satisfaite s'il a calculé une « moyenne arithmétique » et un « écart quadratique ».

2. Concurrément à ce monomane du modèle mathématique de la distribution normale, existe le monomane de tel ou tel autre modèle mathématique ; quels que soient les résultats numériques qu'on lui apporte, il les assujettira au modèle mathématique qui flatte sa manie du moment.

3. Le « traditionaliste » est convaincu que rien d'important n'a été découvert en statistique depuis FISHER : il lui suffit de calculer une valeur de t ou une valeur de Chi^2 ; il considère les calculatrices électroniques et tout ce qui touche à la théorie de l'information comme d'inspiration démoniaque.

4. La « randomisation », d'après E. MORICE, emploie une technique appropriée, pour définir la répartition d'un ensemble d'éléments dans un cadre donné (ordre sur une ligne, place dans un tableau) ou pour choisir un certain nombre d'individus dans une population. (*Tirages au hasard*).

Pour le maniaque de la randomisation peu importe ce qu'on randomise pourvu qu'on « randomise » selon la technique appropriée.

4. Pour le maniaque de la « quantification » peu importe ce qu'on mesure pourvu qu'on ait des résultats numériques, pour appliquer la loi des grands nombres : si ce maniaque dispose d'une calculatrice il n'y a guère de limite au nombre des chiffres qu'il accumulera à partir de quelques résultats expérimentaux.

Ceux qui viennent consulter le « statisticien » peuvent appartenir à des stéréotypes symétriques de ceux des statisticiens ci-dessus décrits : le maniaque des « grands nombres » aura rassemblé une masse de chiffres, dont il n'a jamais compris à quoi ils se rapportent, ou dont il a oublié d'où ils proviennent, trop préoccupé qu'il est d'aligner des rangs ou des colonnes.

Il abandonne la responsabilité de valoriser cette masse de données, apportant peu ou pas d'information, au statisticien, qu'il considère comme un sorcier capable d'introduire de l'ordre dans le chaos ou plutôt, de condenser une masse chaotique en un chiffre qui, précédé d'un signe cabalistique tel que « p.p.d.s. » acquiert une signification magique.

Le collectionneur de chiffres accumule depuis des années des résultats numériques au sujet de N^* expériences et arrive avec une valise bourrée de papiers couverts de calculs ; il ne peut ni expliquer ce qu'il a voulu faire ni ce qu'il voudrait faire ; il lui importe seulement que sorte de ces milliers de chiffres quelque chose qu'il puisse présenter comme rapport justifiant les crédits de recherches, dans les 48 heures qui lui restent à cet effet.

Le probabiliste ne s'intéresse qu'à une valeur « p » significative; plus faible est la valeur plus il est heureux : un individu qui n'aurait pas souri depuis 20 ans témoigne d'une joie hilarante quand on lui annonce $p = 0.001$; la plupart des individus de cette espèce savent que c'est « significatif » mais il ne faut pas s'attendre à ce qu'ils comprennent la signification d'une erreur du type I; il existe 2 sous-variétés de probabilistes suivant que le statisticien consulté est supposé savoir faire ce qu'il faut faire de ce « p » ou qu'on veut imposer un certain test (toujours celui qui ne peut pas convenir).

Jean DUFRÉNOY