



Statistica Nova - la methode et le logiciel des fonctions mobiles

Ionel Dorofte

► **To cite this version:**

| Ionel Dorofte. Statistica Nova - la methode et le logiciel des fonctions mobiles. 41èmes Journées de Statistique, SFdS, Bordeaux, 2009, Bordeaux, France, France. 2009. <inria-00386763>

HAL Id: inria-00386763

<https://hal.inria.fr/inria-00386763>

Submitted on 22 May 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

STATISTICA NOVA : LA METHODE ET LE LOGICIEL DES FONCTIONS STATISTIQUES MOBILES

Ionel Dorofte

UTI

Iasi, Roumanie

ionel.dorofte@gmail.com

www.telemat.biz

RESUME. Dans plusieurs publications récentes, l'auteur s'engage dans une approche critique et innovatrice des méthodes d'analyse statistique. Dans cette communication, on présente deux nouvelles applications des fonctions statistiques mobiles.

ABSTRACT. In several recent publications, the author gets involved in a critical and innovative approach of the statistical analysis methods. In this communication, two new applications of the mobile statistical functions are presented.

MOTS-CLES : paradigme, fonctions statistiques mobiles, estimateur de la densité

KEYWORD : paradigm ,statistics mobiles functions, density estimator

Le développement de la science n'est pas – comme il pourrait paraître de la perspective encyclopediste – seulement une rigoureuse construction théorique, prévisible et irréfragable. Cette image, d'une congruence et déductivité univoque accablante, est bien structurée dans la conscience collective et , volens-nolens dans celle de l'homme de science, quelque conscient qu'il en soit – surtout par le biais du manuel d'école ou universitaire. Mais le manuel, par des considérants didactiques aussi bien qu'économiques – bureaucratiques est souvent un résumé post-factum aussi filtrant et partial que possible et non pas le scénario même de la confrontation idéatique entre solutions théoriques alternatives qui en pleine gloire étaient loin d'être indubitables.

Un fameux épistémologue et auteur de la théorie des paradigmes- Thomas Kuhn - a souligné ce marasme éfervescent de la création qui accompagne et vient avant l'évolution réelle de la science. L'une des caractéristiques d'une période de changements novateurs porte sur le fait qu'il y a des plus en plus fréquentes contradictions entre les données empiriques et les prédictions des modèles théoriques dans le contexte d'un mécontentement général concernant le degré de performance de la discipline scientifique envisagée.

Les nouvelles technologies de l'information et de la communication (NTIC) et l'omniprésence physique des ordinateurs ayant des possibilités inconcevables autrefois créent les prémises et instruments d'une réinvention méthodologique. L'auteur de cette communication a avancé lors des derniers ans dans le cadre de diverses manifestations scientifiques internationales l'idée que, à présent, on est devant un très probable et nécessaire "saut paradigmatique" qui va restructurer en profondeur la science de la statistique. D'ailleurs l'entier trajet de la méthodologie de recherche scientifique devra être restructuré dans ces nouvelles conditions technologiques. On a nommé ce

tronc d'idées qui combine une vision positive-critique concernant la phase actuelle d'évolution de notre discipline aussi bien que toute une série de développements innovateurs aussi bien que le logiciel créé de cette perspective de manière suggestive, selon moi : "Statistica Nova".

L'une des innovations méthodologiques que l'on propose à la communauté scientifique vise le concept des "fonctions statistiques mobiles". En fait, c'est une reprise élargie de la méthode de calcul de la moyenne mobile déjà classique qui entre dans le bagage de connaissances de tout praticien dans l'analyse des données. Notre méthode consiste à sélectionner par le l'interme de d'une "fenêtre mobile" qui se déplace à un pas contrôlable- des sous-populations successives dans le domaine de définition d'une variable et le calcul des fonctions statistiques classiques par chaque sous-groupe. L'observation de la dynamique de ces coefficients ou indicateurs rapportés à l'entière base de données s'avère très intéressante. On envisage aussi le cas multidimensionnel d'autant plus relevant. De la littérature spécialisée, il résulte que les «fenêtres mobiles» sont appliquées à présent aussi dans l'analyse de signal ou plus récemment dans la statistique spatiale. Il faut mentionner ici aussi les estimateurs non-paramétriques de la densité - type Kernel - et la méthode d'analyse noyau.

Pourtant on pense que la statistique mobile que nous proposons peut être retenue en tant que contribution sui generis du point de vue conceptuel et applicatif. Jusqu'à présent j'ai publié, en ce qui concerne ce sujet, des études concernant la fréquence mobile et la corrélation mobile. Ensuite, je présenterai pour la première fois dans le cadre de cette Conférence deux nouvelles applications: la dispersion (ecart-type) mobile et le coefficient mobile de variation.

On pense que le suivant exemple est édificateur concernant l'efficacité analytique de la méthode des fonctions mobiles. J'ai mêlé deux distributions parfaitement normales (générées avec Excel). On a obtenu une distribution cvasi-normale asymétrique qui donne deux informations sur les deux population initielles. Mais l'analyse à l'aide de la méthode du ecart-type mobile(fig.1) et *coefficient mobile de variation*(fig.2) nous offre la possibilité d'identifier presque parfaitement les deux sous populations mère .

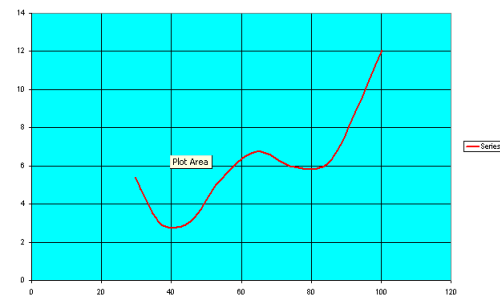


Fig.1

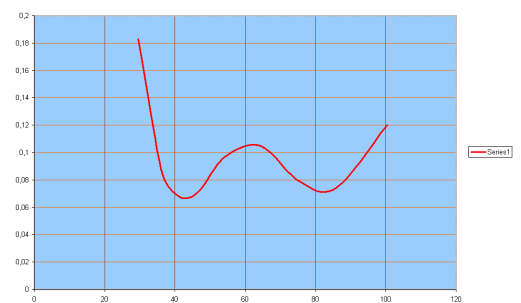


Fig.2

L'application informatique STATISTIQUE NOVA est achevée en deux variantes:

1)Excel – familier à tous et 2) Visual Basic, software pour programmation qui permet la construction des logiciels en format exécutable habituel pour les utilisateurs des applications informatiques.

Des années de recherche se sont succédées afin de développer la théorie et la méthodologie des fonctions statistiques mobiles. Notre idée principale est celle que les valeurs des fonctions statistiques ne doivent pas être calculées de façon unitaire sur tout le domaine de définition des variables, mais par l'intermède d'une «fenêtre mobile» on va calculer des dizaines ou des centaines de valeurs. Dans une

première étape, on obtient une courbe de fréquence des données plus fidèle et stable que celle obtenue par la méthodologie courante. Le calcul des coefficients de corrélation mobiles prouve le plus souvent - en une manière totalement impressionnante- qu'à l'intérieur d'une valeur petite ou bien insignifiante calculée par la statistique classique il y a en réalité des intervalles de corrélation significatives ou même forte! .

Logiciel – Statistique Nova a cinq fenêtres de travail principales appelables, en appuyant sur un de cinq boutons de la barre du menu : Dates, Variables, Correlations, Dispersions, Valeurs. Toutes les opérations d'analyse ou d'information sont résolues au cadre d'une seule fenêtre, ce qui fait très facile l'instrumentation d'une analyse statistique par un utilisateur qui ne parcourt que les instructions élémentaires fournies par l'auteur avec le logiciel .

Ce qui suit c'est une présentation c'est la fenêtre Variable, sans entrer en détails qui nécessiteront une analyse plus ample et dédiée strictement à l'utilisation de ce logiciel. On fait la précision que par un seul click on peut choisir la variable d'analyse des ceux enregistrées dans la base de dates. On peut aussi compléter dans la fenêtre petite dédiée à ce but le numéro d'enregistrements qu'on veut inclure dans cet analyse-là. (par exemple.: = 2000 enregistrements).Un autre avantage, qu'on considère important, c'est pour que la graphique de visualisation des dates – à part l'affichage des valeurs brutes - présente une double scalaire des valeurs: 1) en unités de position(quartile) dessinées en vert; 2) en unités normales – normées tracées en mauve.

Le logiciel Statistica Nova permet de retrouver par l'intermédiaire des clés intuitives chaque individu statistique ce qui constitue dans beaucoup des situations une aide immense pour le chercheur et pour le praticien. Dans la Fig. 3, au dessous de axe d'abscise, de on est représenté dans les deux rectangles blanches, les coordonnées de L'INDIVIDU Ala (nombre = 1 dans la base des données = Assurance) précisément : Valeurs Brut = 9.21 et Unites Z = - 1.94

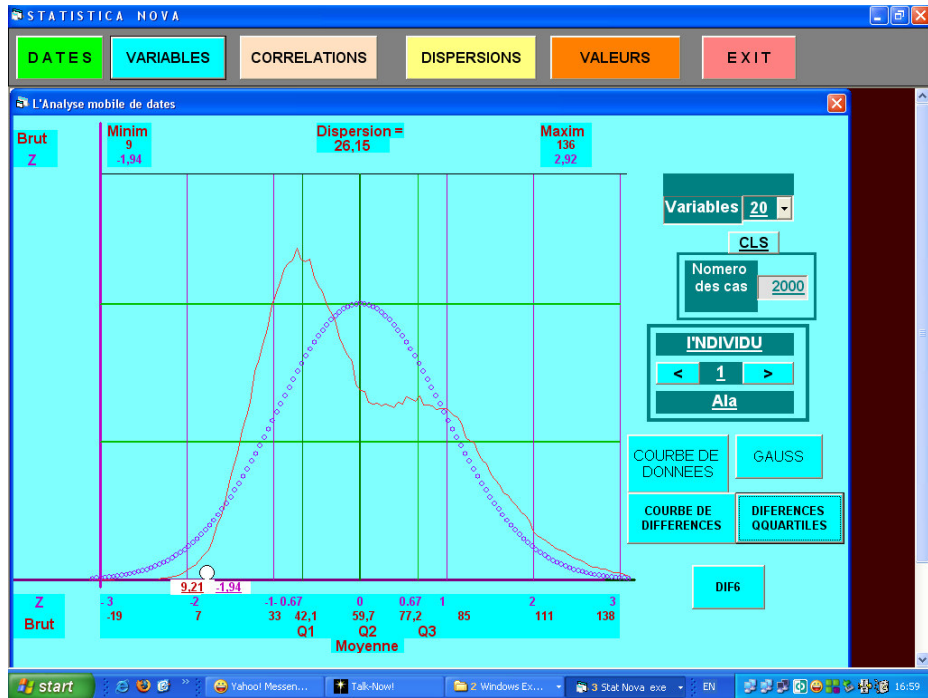


Fig.3

Dans la figure ci-dessus, sous le graphique de distribution de la collectivité de 2000 cas pris en calcul, sur l'axe horizontal l'individu concret est visualisé en blanc. On peut lire les dates brutes, aussi bien que les valeurs normales – normées correspondantes à cet enregistrement. Dans une modalité très simple – en appuyant sur les deux boutons de déplacement – sous la cassette L'INDIVIDU l'utilisateur peut choisir et visualiser un autre enregistrement au choix. Aussi nous pouvons lire « la courbe des différences(fig.4).

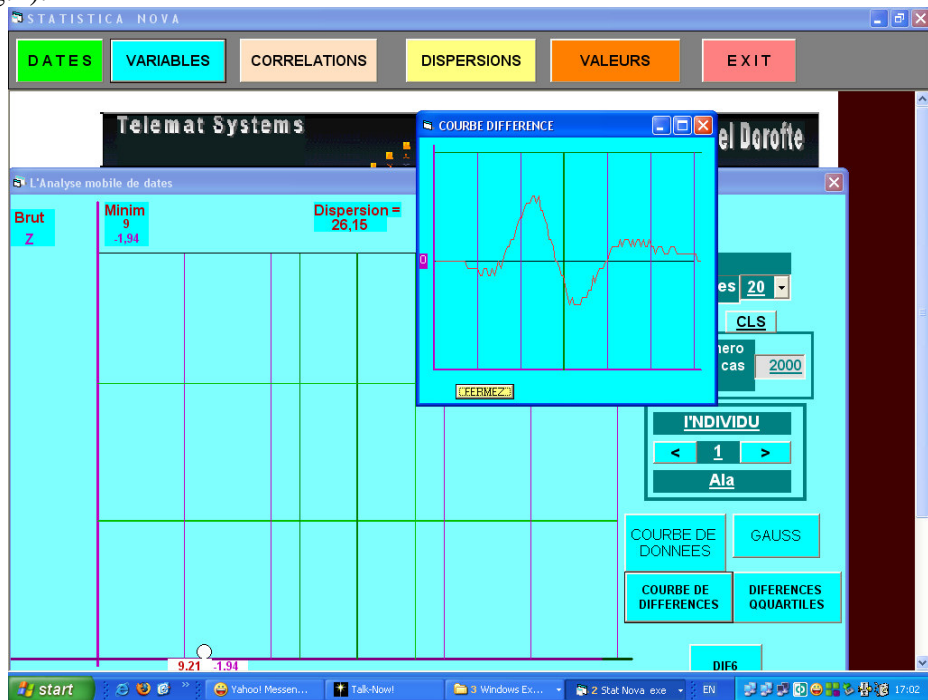


Fig.4

Bibliographie

Allard F., Sanders H.D. (1992) “*Les statistiques : une approche nouvelle*”, Montréal, Mc GrowHill Editeurs.

Dorofte I. and Dorofte T. (1983) “*Contributions to the Analysis and Statistical Prediction of Extreme Psychodynamic Processes*”, Bulletin of The British Psychological Society, vol.37 , London

Dorofte, Ionel (2003) « *Statistica Nova: banque de données et fonctions statistiques Mobiles*”, Bulletin of the ISI, vol LX, Berlin

Dorofte I. (2004) « *Statistica Nova: un nouveau paradigme en approche des données* », www.agro-montpellier.fr/sfds/CD/textes/dorofte1.pdf

Dorofte, Ionel (2006)« *Statistica Nova :calculer pour mieux comprendre la realite* » « Journee de statistique, Nr.38, Paris

Dorofte Ionel, Dorofte Tatiana (2006) « *Statistica Nova :* » A new « metrique » of humans performances », SMABS,Nr.22.Budapest

Dorofte Ionel, Dorofte Tatiana (2006) « *Statistica Nova : L’inference empirique dans l’analyse et prediction de performances humaines* “,Conference Internationale d’Inventique, Performantica ,Iasi

Dorofte Ionel, (2007) “ *Statistica Nova. Esquisse theorique et logiciel, SADA, Cotonou*

Forest W. Young (2004) <http://www.visualstats.org>

Gras Regis (1996) “ *L’implication statistique. Nouvelles méthodes exploratoires de donnée* », Paris, Editions « La pensée sauvage ».

Kuhn, Thomas S. (1989) « *La structure des révolutions scientifiques* », Ed. de Poche

Le Guen M, (1999) «*L’Analyse exploratoire des données et SAS/INSIGHT* », Courrier des Statistiques, nr.90, juin 1999, INSEE, pp3-44.

McClave, James T. et co. (1997) “ *Statistics*”, Prentice Hall, New Jersey.

Morineau, Alain et collectif (2005) « *L’analyse statistique des données* » Ellipses,Paris