

Techniques d'enquête

Commentaires à propos de l'article « Les contributions de Jean-Claude Deville à la théorie des sondages et à la statistique officielle »

par Marc Christine

Date de diffusion : le 3 janvier 2024



Comment obtenir d'autres renseignements

Pour toute demande de renseignements au sujet de ce produit ou sur l'ensemble des données et des services de Statistique Canada, visiter notre site Web à www.statcan.gc.ca.

Vous pouvez également communiquer avec nous par :

Courriel à infostats@statcan.gc.ca

Téléphone entre 8 h 30 et 16 h 30 du lundi au vendredi aux numéros suivants :

- | | |
|---|----------------|
| • Service de renseignements statistiques | 1-800-263-1136 |
| • Service national d'appareils de télécommunications pour les malentendants | 1-800-363-7629 |
| • Télécopieur | 1-514-283-9350 |

Normes de service à la clientèle

Statistique Canada s'engage à fournir à ses clients des services rapides, fiables et courtois. À cet égard, notre organisme s'est doté de normes de service à la clientèle que les employés observent. Pour obtenir une copie de ces normes de service, veuillez communiquer avec Statistique Canada au numéro sans frais 1-800-263-1136. Les normes de service sont aussi publiées sur le site www.statcan.gc.ca sous « Contactez-nous » > « [Normes de service à la clientèle](#) ».

Note de reconnaissance

Le succès du système statistique du Canada repose sur un partenariat bien établi entre Statistique Canada et la population du Canada, les entreprises, les administrations et les autres organismes. Sans cette collaboration et cette bonne volonté, il serait impossible de produire des statistiques exactes et actuelles.

Publication autorisée par le ministre responsable de Statistique Canada

© Sa Majesté le Roi du chef du Canada, représenté par le ministre de l'Industrie, 2024

Tous droits réservés. L'utilisation de la présente publication est assujettie aux modalités de l'[entente de licence ouverte](#) de Statistique Canada.

Une [version HTML](#) est aussi disponible.

This publication is also available in English.

Commentaires à propos de l'article « Les contributions de Jean-Claude Deville à la théorie des sondages et à la statistique officielle »

Marc Christine¹

Résumé

Cet article discute et commente l'article de Ardilly, Haziza, Lavallée et Tillé consacré à une présentation synoptique de l'œuvre de Jean-Claude Deville en théorie des sondages. Il apporte quelques éclairages sur le contexte, les applications et les utilisations des résultats de ses travaux et il montre comment ceux-ci se sont inscrits dans le métier de statisticien dans lequel Jean-Claude a eu une démarche d'« éclairé ». Il évoque aussi d'autres aspects de sa carrière et de ses inventions créatrices.

Mots-clés : Échantillonnage; calage; équilibrage; partage des poids; quota; méthodologie statistique; métier de statisticien.

L'annonce de la mort de Jean-Claude Deville en octobre 2021 a suscité beaucoup d'émotion à l'Institut national de la statistique et des études économiques (Insee) et dans le Service statistique public français, ainsi que dans le milieu international (Canada, partenaires européens, etc.). C'était un grand personnage, à la fois sur le plan humain et sur le plan intellectuel, sa disparition a laissé une grande tristesse chez ses collègues mais son œuvre est là pour témoigner de ses exceptionnelles qualités. Une bibliographie des écrits de Jean-Claude Deville a été établie par Yves Tillé, disponible sur le site des Journées de méthodologie statistique de l'Insee, jms-insee.fr, qui rassemble par ailleurs les Actes de toutes les éditions de ces Journées de 1991 à 2022, dont les communications évoquées dans cet article.

L'article proposé par Ardilly, Haziza, Lavallée et Tillé constitue un exposé remarquable de quelques-unes des innovations majeures qu'a apportées Jean-Claude Deville à la théorie des sondages et à la pratique des enquêtes statistiques, en premier lieu celles du Service statistique public français, mais aussi avec de nombreuses adhérences internationales : échantillonnage équilibré, calage, partage des poids, calcul de la variance d'estimateurs complexes, échantillonnage par quotas.

Pour qui n'aurait pas connu Jean-Claude, il pourrait apparaître, à la lecture de l'article proposé, que sa carrière a été consacrée essentiellement à développer des champs d'investigation théorique avancés, des formalisations mathématiques complexes et des travaux de recherche très pointus. Mais, comme le dit l'introduction de l'article, il est important de rappeler que Jean-Claude Deville a été en poste à l'Insee pendant plusieurs décennies, dont une partie consacrée spécifiquement à l'Unité Méthodes Statistiques où les questions de sondage avaient un rôle prééminent, à la suite de laquelle il est parti au laboratoire de statistique d'enquêtes associé à l'École nationale de la statistique et de l'analyse de l'information (Ensa) à Rennes.

1. Marc Christine, ancien Inspecteur général de l'Insee. Courriel : mchristine7577@gmail.com.

C'est dire à la fois qu'il était porté par la nécessité de résoudre avec pragmatisme des problèmes concrets qui se présentaient à lui ou aux statisticiens qui y étaient confrontés dans leur vie professionnelle quotidienne sans savoir nécessairement comment les résoudre, mais aussi qu'il avait un pouvoir de formalisation, d'élargissement des problèmes qui lui étaient soumis et d'établissement de liaisons entre différentes disciplines qui lui conférait une hauteur de vue remarquable.

Il convient également de souligner que Jean-Claude Deville a été amené à travailler sur d'autres sujets que ceux touchant stricto sensu à la théorie des sondages, notamment en matière de démographie et d'analyse des données, preuve à nouveau de son éclectisme en matière de recherche. Il a été aussi le père fondateur et la référence théorique pour une transformation majeure du système de recensement français en un recensement par sondage, en continu, mis en œuvre en 2004. Si l'article ici concerné ne traite pas de ces travaux, il convient cependant de les garder en mémoire.

Il a aussi été enseignant à l'École nationale de la statistique et de l'analyse économique (processus aléatoires du 2^e ordre, sondages approfondis...).

Ardilly, Haziza, Lavallée et Tillé ont fait une synthèse éloquent, et en même temps très détaillée, sur les aspects techniques des travaux de Jean-Claude Deville. *On ne saurait commencer cette discussion sans souligner avant tout l'extrême qualité et la complétude de leur article.*

Aussi, les quelques commentaires qui suivront n'auront-ils pas pour but de discuter de la qualité de cette présentation globale, ni évidemment de l'intérêt intrinsèque des travaux de Jean-Claude Deville, mais plutôt d'apporter quelques éclairages partiels sur certains d'entre eux pour les réassocier dans le contexte des questions quotidiennes des statisticiens qui les ont suscités ou qui ont été des champs d'application.

Un homme, un métier...

Parler des travaux de Jean-Claude Deville à partir de la revue qui en est faite dans les articles présentés, c'est parler aussi du quotidien de son métier de statisticien, qui, comme dit plus haut, a nourri ses réflexions théoriques.

Quelques exemples, sans aucune prétention à l'exhaustivité, peuvent aider de manière impressionniste à rappeler la variété de ses interventions et son pragmatisme. Il a été « homme-ressource » sur de nombreux sujets. Dans une liste aléatoire issue de ma propre expérience :

Il a aidé à répondre à partir de l'enquête Vacances des Français, à une question d'un groupement de communes bretonnes qui voulaient démontrer que leur fréquentation touristique avait baissé suite au naufrage de l'Amoco-Cadiz en 1978 et à la marée noire qui en était résultée, dans le but d'obtenir réparations au tribunal correctionnel (1984).

Dans le projet « CAMME (mensualisation de l'enquête de Conjoncture Auprès des MÉnages) », il avait inventé un mode de tirage de l'échantillon téléphonique dans l'annuaire papier – seul disponible à l'époque – en 1985.

Il a beaucoup travaillé dans les années 1993-1998 sur le projet de trimestrialisation puis de mensualisation de l'enquête Emploi française (initialement annuelle). Tout en regrettant que toutes ses

préconisations n'aient pas été suivies, son soutien et sa défense de ses collaborateurs étaient appréciés quand, par exemple, un prestataire ne connaissant rien à l'affaire avait voulu s'en mêler (« J'aime ton ton alerte », m'avait-il dit quand j'avais protesté contre cette initiative extérieure...).

Initiateur du Recensement rénové de la population (en continu) (1999-2000), il était leader d'un groupe de travail où il était amateur de dialogues scientifiques « musclés ».

Il a suivi de manière critique dans le bon sens du terme le projet « Octopusse » de constitution du nouvel échantillon-maître issu du recensement en continu (« Organisation Coordinée du Tirage Optimisé Pour Une mise en œuvre Statistique des Échantillons »), tout en ironisant, sous la forme d'une « parabole des chefs locaux », illustrant son inquiétude que l'on demande l'avis des Directeurs régionaux de l'Insee sur les échantillons une fois ceux-ci tirés... (2009).

Le calage

Je mettrais volontiers en premier ce corpus car, d'une part, ces travaux ont maintenant plus de trente ans d'âge et, d'autre part, on pourrait dire des procédures de calage qu'elles constituent la plus populaire des théories en ce sens que leur pratique s'est banalisée et généralisée dans la plupart des enquêtes statistiques. Et, des travaux cités dans l'article de Ardilly, Haziza, Lavallée et Tillé, c'est sans doute la plus ancienne qui a été développée, mise en œuvre et banalisée dans des outils tels que la macro CALMAR de SAS.

L'intérêt du calage n'est plus à développer : d'une part réduire la variance des estimateurs, d'autre part assurer la cohérence des estimations de totaux à partir d'une enquête statistique avec des données externes réputées exactes et incontestables, sur un certain nombre de variables de référence socio-démographiques, géographiques, économiques...

Sa mise en œuvre ne nécessite que la disponibilité d'informations auxiliaires. Dans la plupart des systèmes statistiques, de telles informations existent, la plupart du temps corrélées avec les variables d'intérêt des enquêtes; encore faut-il qu'il y ait une cohérence conceptuelle entre les variables de calage et celles observées dans l'enquête.

Un écueil toutefois est l'apparente facilité avec laquelle la technique peut être utilisée et la tentation de caler sur « tout » ce qui est disponible. Le prix à payer est un ajustement plus délicat des poids de calage et le risque d'obtenir des poids aberrants. D'où la nécessité de chercher à borner les poids ou, plus précisément, le rapport entre les poids de calage et les poids initiaux de sondage, mais aussi de réfléchir à la définition des variables de calage les plus pertinentes, sans chercher à les multiplier au risque d'introduire des colinéarités hasardeuses.

Ainsi, à l'Insee, dans le domaine des enquêtes Ménages, un groupe de travail, le « Groupe Marges », s'est attaché à définir une doctrine sur les méthodes de calage de enquêtes, tant sous l'angle des variables à utiliser, de manière obligatoire ou optionnelle, que sous celui des sources les plus adaptées pour fournir les marges de référence.

La méthode est tellement populaire qu'elle peut être utilisée « sans poids » : des données empiriques peuvent être calées sur des totaux pour les rendre « plus représentatives » de l'univers de référence, sans

qu'on sache exactement de quel processus elles sont issues et l'on convient alors implicitement de faire comme s'il y avait des poids de sondage initiaux tous égaux entre eux.

On peut aussi caler sur « rien », si l'on ne dispose d'aucune information auxiliaire, ce qui est certes rare. En réalité, il faut avoir présentes en mémoire deux formules duales :

$$\left\{ \begin{array}{l} \sum_{i=1}^N \pi_i = E(\text{card } s) \\ E\left(\sum_{i \in s} \frac{1}{\pi_i}\right) = N \end{array} \right. ,$$

pour un échantillon s sélectionné dans une population de taille N , avec des probabilités d'inclusion π_i . Au minimum on peut donc caler sur la taille de la population de façon à disposer de poids dont la somme est exactement égale à cette taille.

Les procédures de calage s'appliquent dans le cas de plusieurs niveaux d'unités statistiques emboîtées : individus/ménages, établissements/entreprises, lorsqu'on dispose de données de calage sur les différents niveaux et qu'il faut assurer la cohérence des pondérations entre ces niveaux.

Le calage a souvent aussi été utilisé comme moyen implicite de corriger la non-réponse, au sein d'une procédure globale appelée communément « redressement de l'enquête », évitant de recourir à un modèle explicatif de la non-réponse qui nécessite de connaître des variables explicatives au niveau individuel pour les répondants mais aussi pour les non-répondants. Françoise Dupont (Dupont, 1993) a montré en fait l'équivalence d'une méthode intégrée et d'une méthode en deux temps (correction de la non-réponse puis calage) sous certaines hypothèses.

Plus spécifiquement, les deux approches ont été rassemblées dans la méthode de *calage généralisé* appliquée au traitement de la non-réponse. Le calage généralisé permet de redresser la non-réponse dans une enquête même lorsque les caractéristiques individuelles les plus corrélées à la non-réponse ne sont connues que dans l'échantillon des personnes ayant répondu au questionnaire. C'est le cas en particulier lorsque la non-observation des variables auxiliaires de calage chez les non-répondants résulte du décalage temporel entre les informations contenues dans la base de sondage et la réalisation du sondage.

Plus précisément, la méthode consiste à écrire les équations de calage sous la forme suivante : $\sum_{k \in r} d_k F(z'_k \lambda) x_k = X$, où z_k désigne le vecteur des variables explicatives de la non-réponse, connues sur l'échantillon r des répondants, et x_k le vecteur des variables de calage, bien corrélées aux variables d'intérêt, connues sur l'échantillon, et dont on connaît les totaux (vecteur X) dans la population. F est une fonction de calage.

La méthode permet donc une correction de la non-réponse même lorsque les variables qui l'expliquent ne sont observées que sur l'échantillon des répondants, en particulier lorsque ces variables sont des variables d'intérêt. Elle a été programmée dans une macro dénommée « Calmar 2 » (Le Guennec et Sautory, 2022), cependant moins populaire que la macro Calmar historique.

Échantillonnage équilibré

D'un certain point de vue, l'échantillonnage équilibré – dont les premières formalisations des travaux de Jean-Claude Deville et Yves Tillé remontent au début des années 2000 – peut apparaître comme une procédure miroir du calage. Il correspond à l'idée que Jean-Claude Deville présentait lui-même comme le vieux rêve des statisticiens : construire un échantillon « représentatif », modèle réduit parfait de la population d'où il était issu. Et si l'on savait déjà qu'un plan de sondage stratifié à allocation proportionnelle permettait de respecter dans l'échantillon la structure de la population au vu d'un certain nombre de caractéristiques descriptives, l'intérêt était aussi de pouvoir travailler avec une représentativité vis-à-vis de variables continues.

La technique présente en outre, par rapport au calage, l'avantage de ne pas modifier les poids de sondage comme le fait le calage mais d'arriver à imposer des conditions d'équilibrage pour un jeu de probabilités d'inclusion (donc de poids de sondage) donné.

Sur ce dernier point il faut toutefois souligner qu'un plan de sondage, en tant que loi sur l'ensemble des échantillons possibles dans une population finie, ne se réduit pas à l'obtention d'un jeu de probabilités d'inclusion π_i fixé. Ainsi, faire le choix d'un échantillonnage équilibré respectant le jeu des π_i , par rapport à un échantillonnage qui ne le serait pas, revient à modifier le plan de sondage initial, tout en respectant les valeurs des π_i et en imposant des contraintes dites d'équilibrage, qui pourraient donner l'impression d'altérer le caractère aléatoire du processus.

Bien sûr, dans les cas extrêmes, un excès de contraintes conduirait à réduire le nombre d'échantillons admissibles, voire à les éliminer tous. À l'inverse, l'obtention d'un échantillon équilibré peut n'être qu'approximative, d'où la nécessité d'ordonner les variables selon l'importance qui leur est attachée dans l'équilibrage, de manière à relâcher éventuellement des contraintes.

Depuis plus de 20 ans maintenant, les techniques d'équilibrage se sont diffusées dans la communauté statistique et sont largement popularisées. Les premières utilisations en France concernent le tirage des groupes de rotations de communes pour le recensement en continu (auquel Jean-Claude Deville a beaucoup contribué), avec diverses contraintes relatives aux dénombremments de population. La technique a été ensuite utilisée pour le tirage des unités primaires de l'échantillon-maître issu du recensement pour les enquêtes ménages, on l'a aussi utilisée pour assurer des équilibrages « géographiques » mais on ne pourrait citer les multiples domaines d'application actuels. C'est aussi grâce au fait que la macro initiale, baptisée CUBE, a connu de nombreux développements, extensions, améliorations cinétiques, en SAS d'abord, puis sous R.

Comme pour le calage, on peut dire qu'on peut faire de l'équilibrage « même sur rien », c'est-à-dire quand on n'a pas de variable d'intérêt particulière disponible ou pertinente pour l'objet de l'enquête. On s'appuie en effet sur des propriétés basiques de l'équilibrage :

- équilibrer sur la *variable constante égale à 1* équivaut à équilibrer l'échantillon sur la taille N de la population;
- équilibrer sur la *variable probabilité d'inclusion* équivaut à assurer un plan de sondage de taille fixe;

- un plan de sondage stratifié avec un sondage aléatoire simple au sein de chaque strate est équilibré sur les *variables indicatrice d'appartenance aux strates*. Et, avec des probabilités d'inclusion *constantes par strate*, équilibrer sur ces *variables* revient à faire un sondage stratifié avec une taille d'échantillon fixée par strate;
- avec des probabilités d'inclusion *constantes dans la population*, équilibrer sur les *variables indicatrices d'appartenance aux strates* équivaut à faire un sondage stratifié avec une taille de l'échantillon fixée par strate et une *allocation proportionnelle*.

Aussi, le recours au tirage équilibré au moyen de la macro Cube se substitue souvent aux procédures de sondage élémentaires indiquées ci-dessus. Mais si l'on rajoute des variables d'intérêt, alors la méthode d'équilibrage regagne évidemment toute sa spécificité par rapport à un tirage « basique ».

Partage des poids

La méthode de partage des poids semble à la fois ancienne (années 80) et assez intuitive. Philippe Brion (Clairin et Brion, 1997) cite des exemples de mise en application : la méthode du « segment pondéré » utilisée en statistique agricole consiste à tirer un échantillon de segments (déterminés à partir de limites naturelles : cours d'eau, routes) puis à enquêter chaque exploitation agricole ayant des terres sur le segment. Il est alors clair qu'une exploitation peut être atteinte plusieurs fois à partir de plusieurs segments et qu'il faut tenir compte de cette répétition pour calculer les pondérations de manière adéquate. La méthode est inspirée de Julien et Maranda (1989).

D'autres exemples sont cités : pour enquêter une population d'éleveurs nomades, il faut établir une liste des points d'eau du pays puis en tirer un échantillon et l'on enquête les individus ayant visité ces points d'eau. Les risques de doubles comptages peuvent être éliminés en ne retenant que la première visite au cours d'une période donnée, ce qui nécessite de connaître des éléments sur ces visites, notamment leur nombre et leur date.

Dans un autre ordre d'idées, du partage des poids « élémentaire » est réalisé lorsqu'on dispose d'une enquête sur des individus tirés au sein d'un échantillon de ménages, avec des caractéristiques relatives à chacun des deux niveaux, et que l'on veut produire des statistiques à partir des deux niveaux qui soient cohérentes (exemple : une répartition des individus selon une caractéristique qualitative donnée doit être cohérente avec une statistique établie au niveau ménage dénombrant les individus ayant cette caractéristique). Il faut donc articuler les poids ménages et individus au moyen d'une méthode de partage des poids. Ce domaine a ensuite été largement investigué et théorisé par Pierre Lavallée, en collaboration avec Jean-Claude Deville.

Breton d'adoption et de cœur, Jean-Claude Deville s'est intéressé en particulier au tourisme dans cette région. Les enquêtes de fréquentation sur sites touristiques ont constitué un champ d'application de la méthode de partage des poids. En effet, interroger les touristes présents sur un site donné pendant un laps de temps donné permet certes d'obtenir des caractéristiques sur les séjours (lieux, durée, hébergement...),

sur les profils socio-démographiques des séjournants, leurs dépenses... Mais multiplier ces enquêtes sur plusieurs sites conduit au risque d'interroger plusieurs fois le même touriste qui se serait déplacé, donc de biaiser considérablement les estimations en résultant.

La méthode de partage des poids, en identifiant les *liens* (en l'occurrence les points où un touriste interrogé à un endroit donné s'est trouvé antérieurement par ailleurs) permet de corriger ce risque de biais. On ne peut cependant que souligner l'importance, pour mettre en œuvre la méthode, que ce soit dans ces enquêtes ou dans d'autres, de la nécessité d'identifier clairement et précisément tous les liens souhaitables et bien anticiper en amont les questions appropriées à introduire dans les questionnaires dans ce but.

Des exemples d'utilisation similaire de la méthode portent sur les enquêtes auprès des personnes sans domicile (interrogées à un point de ravitaillement ou d'hébergement), la fréquentation de festivals culturels (festival d'opéra à Aix-en-Provence ou Beaune, festival du court métrage à Clermont-Ferrand...), où les publics peuvent se rendre à plusieurs spectacles donc être atteints plusieurs fois par des enquêtes *in situ*.

Calcul de la variance d'estimateurs complexes

On rappellera simplement ici pour mémoire quelques-uns des domaines d'application de ces techniques développées par Jean-Claude Deville :

- fractiles;
- les indicateurs de pauvreté, estimés notamment à partir des enquêtes SILC (*Statistics on Income and Living Conditions*) dont l'enjeu politique est très important;
- les indices de concentration (GINI, ...).

D'autres domaines ont été abordés par lui sur ce thème, de façon innovante, mais non cités dans l'article de Ardilly, Haziza, Lavallée et Tillé :

- Les valeurs propres dans une analyse en composantes principales à partir de données issues d'échantillons.
- Les estimateurs issus de tirage de points dans un ensemble infini, où les probabilités discrètes sont remplacées par des *mesures* au sens de la théorie de la mesure.

Les quotas

Il faut plutôt voir dans les travaux de Jean-Claude Deville sur ce sujet un exercice de style. Il était évidemment très réticent sur les sondages empiriques par rapport aux sondages aléatoires, les premiers ne pouvant convenir qu'aux instituts privés ne disposant pas de bases de sondage appropriées, alors que la statistique publique pouvait disposer de moyens pour gérer de telles bases, savoir y tirer des échantillons et savoir calculer des estimateurs et leurs variances à partir de la définition d'un plan de sondage.

Il souhaitait cependant construire une approche théorique de la méthode des quotas et essayer de définir les bonnes pratiques en la matière.

Deux approches sont possibles :

- Une approche modèle sur la variable d'intérêt, qui est la plus propre pour justifier de manière théorique la méthode des quotas. Cependant, Jean-Claude Deville était par nature très réticent sur les modèles et sur les risques que prendrait la statistique publique en y recourant : un modèle est fondé sur des hypothèses, mais sont-elles vérifiées, comment le saurait-on et peut-on se donner le droit de les imposer à l'utilisateur ?
- Ou bien on essaie de replacer les quotas dans le cadre d'un échantillonnage fondé sur un plan et on fait comme si c'était du sondage aléatoire simple stratifié.

Dans tous les cas, la question centrale est celle de la variance des estimateurs retenus : peut-on la calculer de manière théorique et l'estimer, et comment en déduire des règles permettant le choix entre une méthode aléatoire et une méthode par quota quand l'une ou l'autre sont possibles mais avec des contraintes de coût différentes et des latitudes sur les tailles d'échantillon ?

Conclusion

Sur tous les travaux cités, et bien d'autres, Jean-Claude Deville a été non seulement novateur mais a aussi ouvert de nombreuses voies de développements futurs et d'approfondissements vers lesquelles nombre de chercheurs se sont dirigés.

Mais on ne peut évoquer la mémoire de Jean-Claude Deville sans parler d'un autre grand œuvre qu'il a créé et qui va de pair avec ses travaux théoriques : les Journées de méthodologie statistique de l'Insee. Il les a initiées en mars 1991 et les a organisées jusqu'en 1998, mais elles se sont pérennisées sous la direction d'autres équipes. C'est ce que soulignait Michel Glaude, Directeur des Statistiques démographiques et sociales de l'Insee, en clôture de la 7^e édition des JMS, le 5 décembre 2000 :

« Il y avait passage de témoin entre une équipe [celle de Jean-Claude] qui avait organisé les Journées de Méthodologie Statistique précédemment et qui avait consacré beaucoup d'énergie, d'affection à cette organisation, et une nouvelle. [...] Je crois que c'est une belle leçon sur la façon de transmettre et de poursuivre un objectif ».

Jean-Claude a été évidemment un grand contributeur dans chaque édition de ces Journées et un certain nombre des travaux évoqués dans l'article de Ardilly, Haziza, Lavallée et Tillé y ont d'ailleurs été présentés pour la première fois (*jms-insee.fr*). Ses deux dernières communications datent de 2015 et 2018 (Deville, 2015; Deville, 2018).

Pour comprendre ses motivations et ses objectifs dans la création de ces Journées, il faut reprendre les termes de l'introduction qu'il en avait faite en 1991, qui figure dans les Actes de cette première édition.

« Le premier pari de ces Journées était de donner une expression concrète de la communauté des statisticiens en la rassemblant dans un même lieu pendant deux jours.

Le second pari est de faire sentir aux « statisticiens praticiens » que les méthodes de la statistique sont utiles pour faire de la statistique.

Le troisième pari était de monter une manifestation de niveau scientifique élevé, quitte à mettre la barre un peu haute pour une partie du public ».

De fait, inspirées de l'exemple et des contacts qu'il avait noués avec Statistique Canada, les Journées de Jean-Claude Deville ont été à la fois un vecteur de diffusion et de transmission des connaissances et l'opportunité d'inviter des grands noms de la statistique internationale et d'organiser avec eux des débats féconds. Elles sont devenues un lieu d'échanges et d'innovations entre statisticiens publics ou privés, de métier ou universitaires, français ou étrangers.

N'eût été la crise de la Covid, peut-être aurait-il pu participer à leur trentième anniversaire, en mars 2021, alors que leur dernière édition a dû être reportée en mars 2022...

Esprit curieux, toujours prêt à se poser des questions, à ne pas se contenter de l'acquis, à chercher des améliorations, Jean-Claude Deville a ouvert le chemin et prouvé que la statistique est loin d'être une discipline figée.

Bibliographie

- Clairin, R., et Brion, P. (1997). Manuel de sondages. Applications aux pays en voie de développement, Insee-Ceped, 1996. Version pdf sur : https://www.pseau.org/outils/ouvrages/ceped_manuel_de_sondages_1993.pdf.
- Deville, J.-C. (2015). Quelques éléments de géométrie et d'algèbre pour comprendre la nature d'un échantillonnage équilibré. *Actes des Journées de méthodologie statistique*, Insee, Paris.
- Deville, J.-C. (2018). Calage à poids bornés : Que fait-on au juste ? *Actes des Journées de méthodologie statistique*, Insee, Paris.
- Dupont, F. (1993). Calage et redressement de la non-réponse totale, validité de la pratique courante de redressement et comparaison des méthodes alternatives pour l'enquête sur la consommation alimentaire de 1989. *Actes des Journées de méthodologie statistique*, Insee, Paris.
- Julien, C., et Maranda, F. (1989). Le remaniement du plan de sondage de l'enquête nationale sur les fermes de 1988. Document de Travail no BSMD-89-012F, Statistique Canada, Ottawa.
- Le Guennec, J., et Sautory, O. (2002). CALMAR 2 : une nouvelle version de la macro Calmar de redressement d'échantillon par calage. *Actes des Journées de méthodologie statistique*, Insee, Paris.