



Former au calcul socialiste

MARTINE MESPOULET¹

Résumé

Un effort de redéfinition des objets d'étude, des méthodes et des outils a caractérisé la conception des ouvrages de statistique pendant les années 1930 en URSS. La rédaction de nouveaux manuels fut un élément clé de ce travail de reconstruction. Une série de livres de ce type vit ainsi le jour pendant cette période. Tous visaient à présenter une théorie de la statistique, des méthodes et des outils répondant à des usages du chiffre définis dans la perspective de la construction du socialisme. Le premier d'entre eux, publié en 1930, fut remanié à différentes reprises. Celui de 1936 présente une forme bien avancée de cette entreprise de reconstruction théorique et méthodologique. Quels principes de base ont présidé à ce travail de remaniement ? L'exemple du chapitre consacré aux enquêtes par sondage permet d'éclairer la manière dont ceux-ci furent appliqués dans la rédaction de cet ouvrage. Cet exemple est particulièrement intéressant dans la mesure où sont formulés, dans ces pages, les principes d'une méthode d'échantillonnage, « le tirage aléatoire stratifié », qui fut utilisé dans l'administration statistique soviétique jusqu'en 1991.

Abstract

The conception of statistics handbooks in USSR in the 1930s was characterized by an attempt to redefine survey subjects and statistical methods and tools. Writing new handbooks was placed at the core of this work of rebuilding statistical theory. Some new books were issued during that period. All of them aimed at presenting a new theory corresponding to the uses of numbers serving building of socialism. The first one was published in 1930 and rehandled several times. The one issued in 1936 presents a quite advanced form of this attempt to rebuild theory and methods. On which principles was it founded? The chapter on sampling surveys enlightens the way how these principles were applied for writing this handbook.

Introduction

Un effort de redéfinition des objets d'étude, des méthodes et des outils a caractérisé la conception des ouvrages de statistique pendant les années 1930 en URSS. La rédaction de

¹ Université de Poitiers, UFR SHA, 8, rue Descartes, 86022- Poitiers, martine.mespoulet@univ-poitiers.fr

nouveaux manuels fut un élément clé de ce travail de reconstruction. Une série de livres de ce type vit ainsi le jour pendant cette période. Tous visaient à présenter une théorie de la statistique, des méthodes et des outils répondant à des usages du chiffre définis dans la perspective de la construction du socialisme. Le premier d'entre eux, publié en 1930, fut remanié à différentes reprises [Iastremski et Khotimski, 1930]. Les huit manuels édités jusqu'en 1936, à la suite de celui-ci, furent rédigés par le même collectif d'auteurs², sous la direction de Boris S. Iastremski et Valentin I. Khotimski et sous la responsabilité institutionnelle de l'administration du plan, le *Gosplan*. Témoignage d'un travail théorique en cours d'élaboration, chaque nouvelle édition apporta des compléments et des corrections à la précédente.

Le manuel de 1936 présente une forme bien avancée de cette entreprise de reconstruction théorique et méthodologique. Quels principes de base ont présidé à ce travail de remaniement ? Comment ont-ils influé sur le calcul et l'analyse statistique ?

Après avoir exposé les principes qui ont guidé la rédaction de cet ouvrage, nous en montrerons une application dans le chapitre consacré aux enquêtes par sondage. Cet exemple est particulièrement intéressant dans la mesure où sont formulés, dans ces pages, les principes d'une méthode d'échantillonnage, « le tirage aléatoire stratifié », qui fut utilisé dans l'administration statistique soviétique jusqu'en 1991.

Par ailleurs, les conditions du passage à l'aléatoire ont connu des formes particulières en URSS à partir des années 1930, dans un pays où, dans le discours des dirigeants politiques, dominait une représentation de la société dans laquelle le collectif primait sur l'individu, et dans laquelle un comportement collectif moyen était privilégié par rapport à des formes d'expression individuelles, et donc diverses et dispersées. Le projet volontariste des dirigeants bolcheviks visant à construire un Etat nouveau et leur représentation prométhéenne de l'action de l'être humain sur son environnement donnèrent une nouvelle dimension aux débats sur l'adoption de l'échantillon aléatoire : quelle place pouvait avoir le hasard dans un monde planifié, où justement l'incertitude ne pouvait pas avoir de place, et où le caractère imprévisible des comportements individuels et la variabilité des cas individuels ne pouvaient pas être pris en compte ? Comment la dispersion pouvait-elle être traitée ? L'introduction au manuel de 1936 expliquait le contenu politique de cette question :

² En 1930, le collectif était composé de A.Ia Boïarski, V.N. Starovski, V.I. Khotimski et B.S. Iastremski, ces deux derniers assumant la direction de l'ensemble de l'ouvrage. Par la suite, L.S. Brand, O.S. Davydov et V.N. Starovski se joignirent à eux. Pour la liste de ces manuels, voir la bibliographie. La table des matières de l'édition de 1936 est donnée en annexe.

« La théorie de la dispersion (de la stabilité) de Dormoy-Lexis, qui est dominante dans la statistique bourgeoise actuelle, est née déjà dans les années 70 du XIXe siècle et, tout comme la science bourgeoise, en immortalisant le capitalisme elle se borne seulement à définir et à mesurer les phénomènes extérieurs, en ne dévoilant pas l'essence des contradictions de la société capitaliste et en ne mettant pas en évidence leur caractère de classe. Les savants bourgeois affirment que quand on analyse les phénomènes de changement on peut distinguer, à l'aide de procédés purement statistiques, des tendances d'évolution fondamentales et des faits collatéraux qui troublent l'allure régulière, cela en ne s'intéressant pas du tout au contenu spécifique du phénomène étudié et à ses régularités internes. » [Iastremski et Khotimski, 1930, p. 7]

Certains passages du chapitre du manuel de 1936 sur les enquêtes par sondage éclairent la manière dont la résolution de ces questions a pu influencer sur l'élaboration des méthodes d'échantillonnage en URSS à cette époque.

I DES MANUELS DU XIX^e SIECLE A CELUI DES ANNEES 1930

1. Construire une statistique qui ne soit pas « de classe »

Le ton est donné dès l'introduction du manuel de 1936 : « La statistique, comme toute autre science, est une science de classe. Tout le système de la statistique dans les pays capitalistes est construit de manière à servir les intérêts des classes dominantes. » [ibid., p.6]

Il s'agissait donc de construire une théorie et des outils statistiques en rupture avec ceux de la statistique qualifiée de « bourgeoise », à l'instar de toute autre forme de science sociale constituée au XIXe siècle³. Pour les auteurs, dans une perspective marxiste, les contenus et les outils étaient inséparables des intérêts des classes au pouvoir :

« La statistique bourgeoise n'aide pas seulement à diminuer de manière significative l'imposition sur les revenus des différents groupes capitalistes, elle dissimule aussi les montants réels des dépenses militaires. De plus, la statistique bourgeoise donne une image

³ Pour une présentation de la statistique comme une « science capitaliste spécifique », voir [Gumbel, 1928] et [Hertz, 2000].

fausse de la situation de l'économie capitaliste, en l'enjolivant et en démontrant une absence d'antagonismes.

En particulier, la statistique des Etats capitalistes fascistes admet des déformations et une falsification impudentes des faits. On peut citer l'exemple de la statistique allemande avec ses données sur le chômage, le salaire réel, les récoltes, etc. qui ne correspondent pas à la réalité.

La pratique statistique bourgeoise repose sur la théorie bourgeoise de la statistique, qui est liée de manière organique à tout le système de l'économie politique et de la philosophie bourgeoise. » [ibid., p. 6]

Le propos est clair : l'avènement de la dictature du prolétariat imposait de forger une nouvelle théorie et de nouveaux outils et, tout d'abord, de rompre avec les représentations du monde données par les théoriciens de la statistique du XIXe siècle, tels Quetelet ou Süssmilch, accusés de donner une représentation des lois de l'économie capitaliste, mais aussi de la société, comme naturelles et « éternelles », donc immuables⁴.

Cette entreprise de reconstruction théorique et méthodologique supposait aussi de prendre des distances avec une partie des contenus des manuels de statistique russes publiés avant 1917, tels ceux de Iouri E. Ianson, de Aleksandr A. Kaufman et Nikolai A. Kabloukov⁵, mais aussi au début des années 1920. Toutefois, dans le cas de ceux, comme Kaufman et Kabloukov, qui avaient réactualisé leurs manuels au début des années 1920⁶, la condamnation de leurs ouvrages ne pouvait pas être totale. En effet, ceux-ci avaient participé activement à l'effort de formation engagé par la Direction centrale de la statistique de l'Etat bolchevik, la TsSOu⁷, pour former les nombreux statisticiens recrutés dès sa création en juillet 1918 [Mespoulet, 2001].

2. Un système de formation adapté

Le système de formation à la statistique s'était développé dès le début des années 1920 dans deux mondes institutionnels différents, celui de la statistique administrative de la TsSOu et celui de la statistique académique. Dès 1919, la TsSOu avait créé en son sein un dispositif de cours à trois niveaux destinés à former des statisticiens à degrés de qualification différents : cours centraux, cours de districts interrégionaux et cours régionaux. Les cours centraux

⁴ A ce sujet, voir [Blum et Mespoulet, 2003], chap. 8.

⁵ [Ianson, 1913] ; [Kaufman, 1912] ; [Kabloukov, 1912]

⁶ [Kabloukov, 1922] ; [Kaufman, 1922].

⁷ TsSOu : Tsentral'noe Statisticheskoe Oupravlenie (Direction centrale de la statistique).

déliciaient une formation théorique aux responsables des bureaux régionaux de la TsSOu et à leurs chefs de départements. Les cours interrégionaux visaient à former le personnel statistique d'encadrement à des aspects plus techniques liés à l'organisation, à la conduite et au traitement des enquêtes et des recensements. Les cours régionaux étaient destinés principalement à former les statisticiens de canton recrutés pour collecter les données de première main dans les campagnes [Mespoulet, 2001].

Parallèlement, l'enseignement de la statistique se développa dans les universités, tant à Moscou qu'en province. En 1922, un département de statistique fut fondé au sein de la Faculté des sciences sociales de l'université d'Etat Lomonossov de Moscou. De nombreux statisticiens qui jouèrent un rôle de premier plan dans les années 1930 en sortirent, notamment Vladimir N. Starovski, qui dirigea ensuite la Direction centrale de la statistique de l'URSS, de 1941 à 1975, Innokenti Iou Pissarev, qui fut plus tard directeur adjoint de la Direction centrale, de 1942 à 1951, et Aron Ia. Boïarski, qui fut un des auteurs du manuel de référence des années 1930. Parmi les professeurs de ce département de statistique figuraient des statisticiens formés avant 1917, tel Panteleïmon A. Vikhliaev, ou même déjà en poste comme professeurs à l'université avant 1917, en particulier Nikolaï A. Svavitski.

En 1924, une autre institution de formation vint compléter ce dispositif, avec la création d'un département de statistique au sein de l'Institut de l'économie nationale Plekhanov à Moscou. Boïarski notamment y enseigna la théorie statistique à partir de 1927.

L'année 1930 inaugura une nouvelle décennie tant dans l'administration statistique que dans le système de formation. La véritable rupture dans ce domaine apparut avec la publication en 1930 du manuel dirigé par Iastremski et Khotimski. Cette rupture est inséparable de la réorganisation, sous la houlette du *Gosplan*, de l'administration statistique ainsi que du système de formation à la statistique dans les établissements d'enseignement supérieur. La parution du manuel de 1936 doit être replacée dans ce contexte. Elle a constitué toutefois une nouvelle étape dans le travail de construction d'une théorie de la statistique reposant sur les fondements du matérialisme historique.

Le début de l'année 1930 fut en effet un tournant décisif pour l'administration statistique de l'Etat soviétique. Le décret du 23 janvier, signé par le comité exécutif central et le conseil des commissaires du peuple de l'URSS, supprima la TsSOu en tant que commissariat du peuple mais aussi, événement symbolique, comme administration d'Etat à part entière. L'ensemble de ses activités, de ses services et de son personnel fut transféré dans l'administration du *Gosplan* de l'URSS, dont elle devint un secteur, le secteur de la comptabilité de l'économie nationale. Conçue comme un instrument de comptabilité et de

gestion, la statistique devait être au service du plan. Il était donc logique d'intégrer l'administration de la statistique à celle de la planification :

« Le développement des éléments socialistes dans l'économie de l'Union des Républiques socialistes soviétiques, lié au rôle croissant du travail de planification, exige un développement considérable et un renforcement de la statistique et de la comptabilité. Le développement de la statistique lui-même, en conformité totale avec les objectifs de renforcement de la planification, exige une unification de ces deux branches du travail dans un organe unique. »⁸

Cette nouvelle réforme traduisait de manière institutionnelle la nouvelle conception de la gestion de l'économie de l'Etat soviétique. Le projet de plan reposait sur une construction théorique de l'économie et de la société centrée sur l'idée de l'existence d'un monde stable, sans fluctuations, dans lequel la stabilité des chiffres devait exprimer l'adéquation entre les résultats et les prévisions. Dans une telle logique, il n'y avait plus de place pour un usage de la statistique comme instrument d'observation des changements. C'est ce qu'exprima V. V. Ossinski un peu plus tard, en 1932, quand il se retrouva à la tête de la branche statistique du *Gosplan* :

« Il est apparu que la statistique est nécessaire là où existent des fluctuations, c'est-à-dire en régime capitaliste ; là où il n'y a pas de fluctuations, la statistique n'est plus utile ; et c'est un des résultats du socialisme. »⁹

Le décret du 23 janvier 1930 marqua la victoire politique d'une conception de l'organisation et de la gestion de l'économie et de la société, mais aussi celle d'une institution, le *Gosplan*, et de sa propre définition des usages de la statistique, par rapport à une autre administration, celle de la TsSOu des années 1920. Devenue simple branche de la planification, la statistique d'Etat retrouva cependant une forme d'indépendance en décembre 1931. De secteur, elle devint une direction centrale à part entière, la Direction centrale de la comptabilité de l'économie nationale¹⁰ du *Gosplan* de l'URSS. Elle conserva ce statut pendant toutes les années 1930. Le mot comptabilité remplaça celui de statistique. La

⁸ Décret du Comité exécutif central et du Conseil des commissaires du peuple de l'URSS sur la suppression de la TsSOu et le transfert de ses fonctions au *Gosplan* de l'URSS.

⁹ Cité dans [Blum, 1994], p. 39.

¹⁰ *Tsentralnoe upravlenie narodno-khoziaïstvennogo outcheta*.

conception du nouveau manuel de statistique publié en 1930 et des divers exemplaires qui suivirent tint compte de l'ensemble de ces changements.

3. Une statistique au service du socialisme

La publication du manuel de 1936 fut le résultat du travail collectif des auteurs de l'ouvrage, mais aussi des apports des enseignants des chaires de statistique de diverses universités soviétiques. Au début des années 1930, le nombre de chaires de statistique augmenta fortement, parallèlement au nombre d'étudiants inscrits à un ou plusieurs cours de statistique. Le passage de la statistique à la comptabilité s'accompagna d'un effort de diffusion de la formation à de nouveaux outils de calcul et de prévision. Ainsi, des facultés de comptabilité de l'économie nationale et des chaires de statistique et de comptabilité furent créées dans de nombreux établissements de l'enseignement supérieur. Dès 1930, une faculté de comptabilité de l'économie nationale fut ouverte au sein de l'Institut Plekhanov à Moscou. La décision prise en 1932 par le Conseil des Commissaires du Peuple de l'URSS de développer la formation de « spécialistes-statisticiens » de différents niveaux de qualification dans l'ensemble du pays entraîna la nécessité d'élaborer de nouveaux manuels le plus rapidement possible. Le décret du Conseil envisageait en effet la constitution de tout un réseau d'établissements secondaires et supérieurs de formation à la statistique et de différents cours de perfectionnement pour le personnel de l'administration de la statistique et de la comptabilité [Ovsienko, 1978].

Dans ce système institutionnel, une place particulière fut occupée, dès sa création en 1932, par l'Institut de la comptabilité de l'économie nationale (MINKhOU), qui prit ensuite le nom d'Institut d'économie et de la statistique de Moscou (MESI). Dès sa fondation, cet établissement fut placé sous l'autorité de la Direction centrale de la comptabilité de l'économie nationale du *Gosplan* et reçut le statut de centre de formation de cadres « spécialistes-statisticiens » du niveau de l'enseignement supérieur¹¹. Dans les années 1930, il s'agissait donc d'un système de formation professionnel, sous la tutelle de l'administration du plan et indépendant des universités, destiné à la formation du personnel statistique des différents services de l'administration de la comptabilité de l'économie nationale. Le manuel dirigé par Iastremski et Khotimski fut publié dans ce cadre institutionnel.

¹¹ Après 1945, le MESI passa, lui, sous la tutelle du ministère de l'Éducation de l'URSS.

De fait, l'institut de la comptabilité de l'économie nationale était au centre de tout un réseau d'instituts et de départements de formation à la statistique et à la comptabilité nationale appartenant à différents établissements de l'enseignement supérieur. Une faculté ouvrière fut organisée auprès de l'Institut de la comptabilité de l'économie nationale, mais aussi tout un ensemble de facultés de l'économie nationale dans les différentes régions et différents Etats de l'URSS. Aussi n'est-il pas étonnant de trouver dans l'avant-propos au manuel de 1936 la liste des professeurs de statistique de différentes chaires d'instituts de la comptabilité nationale du pays qui envoyèrent remarques et propositions pour compléter la nouvelle édition de cet ouvrage.

Après avoir dénoncé le caractère erroné et volontairement trompeur de la statistique « bourgeoise » européenne et de certains travaux théoriques russes de la période prérévolutionnaire, les auteurs de ce manuel s'appuyèrent sur des citations de Lénine pour exposer, dans l'introduction, les objectifs de l'ouvrage :

« Lénine, quand il soulignait toute l'importance de la comptabilité pour l'économie socialiste, attirait l'attention sur le fait que, dans l'économie socialiste la comptabilité doit être avant tout un contrôle de la mesure du travail et de la mesure de la consommation. “La difficulté principale de la révolution prolétarienne réside dans la réalisation, à l'échelle de l'ensemble du peuple, de la comptabilité la plus précise et la plus scrupuleuse possible et du contrôle, *contrôle ouvrier*, de la production et de la répartition des produits”¹². » [Iastremski et Khotimski, 1936, p.12.]

Le contrôle des chiffres par le peuple supposait la diffusion d'un certain nombre de connaissances en statistique et comptabilité. Ainsi se trouvait justifiée la forme prise par le manuel, celle de développements sur la théorie statistique qui reposaient sur des exemples chiffrés pris dans la production. Cet ouvrage devait s'adresser au plus grand nombre :

« Lénine parla à maintes reprises du fait que, seulement après le renversement du capitalisme, la comptabilité deviendrait l'affaire des masses. Mais une telle comptabilité est impossible sans la statistique, qui aussi est indispensable pour faire le bien des masses. “Dans la société capitaliste, la statistique était un domaine de la compétence exclusive des ‘gens de l'Etat’ ou de spécialistes étroits, nous devons la porter dans les masses, la populariser pour que les travailleurs apprennent progressivement à comprendre eux-mêmes et à voir comment et

¹² Lénine, *Œuvres choisies*, T. XXI, p. 259.

combien il faut travailler, comment et combien on peut se reposer...¹³». Le caractère indispensable d'une telle diffusion dans le peuple de la comptabilité socialiste exige une simplification maximale de sa technique. » [ibid., p. 13]

La production des chiffres dans un Etat socialiste devenait donc un outil pour le peuple, que celui-ci devait s'approprier, ce qui justifiait de donner la priorité à l'exposé de dispositifs techniques, supposés plus accessibles par le plus grand nombre. Les auteurs de l'ouvrage se livrèrent à un effort de construction d'une « comptabilité socialiste » au service d'un programme politique de popularisation de l'usage des chiffres. En outre, certains outils ne pouvaient pas être mis sans explications, voire sans mise en garde, dans les mains du peuple et des jeunes « spécialistes-statisticiens » auxquels ce manuel était destiné. Ainsi en était-il de la moyenne :

« Il est nécessaire de citer l'exemple classique de la révélation par Staline de l'opportunisme de droite qui essaya d'utiliser avec malveillance « les valeurs moyennes ».

Est-il vrai que les emblavures ont tendance à une réduction systématique ? Non, ce n'est pas vrai. Le camarade Rykov se servait ici de chiffres moyens sur les emblavures pour l'ensemble du pays. Mais la méthode des moyennes, non corrigée par des données par district, n'est pas une méthode scientifique. Peut-être le camarade Rykov a-t-il lu un jour *Le développement du capitalisme* de Lénine. S'il l'a lu, il doit se souvenir de la manière dont Lénine y critique les économistes bourgeois qui utilisent la méthode des moyennes au sujet de l'accroissement des emblavures en ignorant les données par districts. Il est étonnant que le camarade Rykov répète aujourd'hui l'erreur des économistes bourgeois. » [ibid., p. 17]¹⁴

L'instrumentalisation à des fins politiques des critiques de l'usage de certains outils ne fut pas réservée au seul cas de la moyenne. D'autres attaques de ce type parcourent l'introduction du manuel. Elles visent toutes à montrer le caractère inadapté des outils de la statistique prérévolutionnaire aux nécessités d'une économie planifiée et du développement des forces productives dans l'objectif de la construction du socialisme. De fait, la question qui se posait aux auteurs de cet ouvrage consistait avant tout à présenter une théorie de la statistique et des méthodes conformes aux préceptes du matérialisme historique ou, au minimum, qui ne soient pas contradictoires avec eux. Cette tâche était délicate pour ceux d'entre eux qui avaient été formés avant la révolution d'Octobre 1917.

¹³ Ce passage est extrait des Œuvres complètes de Lénine, tome XXI, 3^e édition, p. 259.

¹⁴ Citation de propos de Staline extraits de *Voprossy leninisma (Questions du léninisme)*, p. 280.

Tel était le cas en particulier des deux responsables de l'ouvrage, Iastremski et Khotimski. Boris S. Iastremski (1877-1962), mathématicien de formation, avait fait ses études à l'université de Kharkov. Spécialisé dans le calcul des probabilités, il dirigea le département de la statistique des assurances de la TsSOu de 1918 à 1926, puis celui de la méthodologie statistique jusqu'en 1933. Dès 1932, il enseigna la statistique mathématique au sein du tout nouveau Institut de la comptabilité de l'économie nationale. En 1931, il était devenu membre du Parti communiste.

Vladimir I. Khotimski (1892-1939) avait été formé aux mathématiques et probabilités, à l'Institut polytechnique de Saint-Pétersbourg, par Aleksandr A. Tchouprov. Entre 1924 et 1927, il enseigna les mathématiques à l'Institut de l'économie nationale Plekhanov. De 1927 à 1932, il fut simple chercheur, puis directeur de la section de mathématiques de l'Académie communiste. Il était proche du Parti communiste. Ensuite, il dirigea le secteur du contrôle du *Gosplan* de Russie, puis, à partir de 1935, le département de la statistique de la population et de la santé de la Direction centrale de la comptabilité de l'économie nationale.

Quant à Boïarski et Starovski, notons qu'au cours de leurs études de statistique à l'université de Moscou, pendant la première moitié des années 1920, ils avaient eu des professeurs formés dans la période prérévolutionnaire.

Aussi, dans le travail de relecture et de ré-interprétation qu'ils ont eu à faire, tous ces hommes ont-ils dû se démarquer des écrits des statisticiens européens qui, dès la fin des années 1920, furent qualifiés de « bourgeois » par les dirigeants politiques soviétiques, tout comme certains statisticiens russes de la période prérévolutionnaire, par exemple A.A. Tchouprov. Bien qu'ancrés dans une tradition intellectuelle qui faisait de la statistique une science, au même titre que la chimie ou les mathématiques, ils furent contraints, à l'occasion de différents conflits, de justifier la nature socialiste de leur travail, donc sa conformité avec le projet politique de construction du nouvel Etat, de son économie et de sa société [Blum et Mespoulet, 2003]. Certains furent soumis à différentes formes de pression politique avec lesquelles ils durent composer, parfois sous la menace, et qui les poussèrent à élaborer une théorie en conformité avec le discours politique bolchevik sur l'édification d'une économie et d'une société socialistes. Quelle part de compromis et d'ajustements conceptuels entra dans leurs constructions théoriques ? Le chapitre du manuel sur les enquêtes par sondage fournit un éclairage sur cette question.

II. SONDAGES, PLANIFICATION ET CONSTRUCTION DU SOCIALISME

En 1917, la Russie jouissait déjà d'un passé riche dans le domaine des enquêtes par sondage. Les premiers essais d'enquêtes sur une « partie du tout » peuvent être situés dans les années 1870. La période qui va de 1885 à 1917 fut marquée par de nombreux efforts de perfectionnement des méthodes d'échantillonnage et une riche réflexion des statisticiens russes à ce sujet [Mespoulet, 2000]. A partir de 1895, ceux-ci eurent entre eux les mêmes discussions que leurs collègues des autres pays européens au sujet du passage du dénombrement exhaustif à une enquête partielle, puis de l'échantillon construit de manière raisonnée, avec recours à l'usage du type, à l'échantillon aléatoire, fondé sur l'application du calcul des probabilités. Au niveau européen, les deux questions de la représentativité et de la confiance que l'on pouvait avoir dans une enquête sur « une partie du tout » furent au cœur d'un débat qui commença avec la première communication de Kiær au congrès de l'Institut international de statistique (IIS) de 1895 et se termina en 1925 au congrès de Rome de l'IIS, où les deux méthodes d'échantillonnage du choix judicieux et du tirage aléatoire furent admises [Desrosières, 1993]. Un terme fut posé à ce débat en 1934 avec la publication par Neyman de son article sur l'allocation optimale par strate. Dix ans plus tôt, un statisticien russe, Aleksandr G. Kovalevski, présentait les mêmes formules de calcul que Neyman dans un manuel de statistique publié à Saratov [Kovalevski, 1924].

La forte effervescence de la production des données statistiques en URSS pendant les années 1920 créa des conditions particulièrement favorables à une rapide diffusion des enquêtes par sondage dans ce pays. Le recours intensif à ces enquêtes dès 1919 fournit un terrain favorable à la réflexion sur les méthodes d'échantillonnage. Après 1925, celle-ci se développa d'ailleurs beaucoup plus en continuité avec les questions posées par la pratique déjà longue des statisticiens russes dans ce domaine que par rapport à celles discutées au congrès de Rome, même si un compte rendu à ce sujet fut effectué au sein de la TsSOu.

Au début des années 1930, le recours aux enquêtes par sondage fut justifié aussi par leur intérêt pratique :

« Il existe de nombreuses raisons pour lesquelles on doit renoncer à faire une enquête exhaustive et recourir à une enquête non exhaustive. Cependant toutes ces raisons peuvent être ramenées à une des trois suivantes : 1) une économie de forces et de moyens, 2) une réduction

au minimum des effets d'une détérioration ou de la destruction des objets étudiés, 3) obtention d'une plus grande précision des résultats de l'enquête » [Iastremski et Khotimski, 1936, p. 162]

Les exemples donnés dans le manuel de 1936 sont empruntés à l'observation de la production ou de la consommation :

« a) inspection de la qualité de la production quand, d'après quelques exemplaires tests pris dans l'échantillon, on peut recevoir une représentation de la qualité de toute la production d'une entreprise donnée ;

b) prévision du volume d'une récolte quand, sur la base de quelques récoltes tests on calcule la récolte sur un hectare pour tout un sovkhoze ou tout un kolkhoze ;

c) « étude des budgets des ouvriers, quand on observe pour toute l'Union soviétique un petit nombre de familles ouvrières et que, d'après les données de cette enquête, on calcule la consommation des ouvriers, la structure du budget, etc. pour toute la classe ouvrière de l'URSS » [ibid.]

Dans les années 1930, en même temps que la nécessité d'utiliser les enquêtes par sondage était affirmée par l'administration statistique, notamment pour faciliter la mesure de la production et de la consommation, les conditions du passage à l'aléatoire connurent des formes particulières en URSS : quelle place pouvait avoir le hasard dans un monde planifié ? Une première réponse consista à opérer un découpage préalable de la réalité avant de procéder au tirage des unités d'un échantillon.

1. « Groupes typiques » et découpage de la réalité

Le chapitre du manuel consacré aux sondages présente les différentes méthodes d'échantillonnage après avoir rappelé les principales connaissances mathématiques nécessaires pour comprendre les conditions dans lesquelles on peut effectuer une enquête par sondage [ibid., chap VI]. Une précision est apportée d'abord sur la différence entre un échantillon aléatoire et un échantillon dit typique :

« Cette sélection des unités observées est organisée de manière à garantir la représentativité, c'est-à-dire la conformité de la partie avec le tout. Selon le procédé d'organisation, on distingue le tirage appelé *typique* (« échantillon typique ») et le tirage *aléatoire* (« échantillon

aléatoire »). Le tirage typique suppose que l'ensemble total étudié soit divisé en groupes typiques et que la sélection des unités observées soit effectuée dans chaque groupe séparément.

Le tirage aléatoire suppose que la partie qui fait l'objet de l'enquête soit extraite au hasard de l'ensemble total, selon un processus aléatoire. Cela signifie, comme nous le savons déjà, que chaque unité tirée a la même probabilité de tomber dans l'échantillon que n'importe quelle autre unité de l'ensemble total. Dans le cas d'un tirage aléatoire, les caractères généraux de l'échantillon reproduisent de manière suffisamment précise les caractères de l'ensemble total. Cette précision, en vertu de la loi des moyennes, s'accroît avec l'augmentation de la taille de l'échantillon.

Les deux méthodes de tirage (quand elles sont correctement utilisées) ne se contredisent pas du tout, mais, au contraire, se complètent mutuellement. » [ibid., p. 171]

Quelques lignes plus loin, des éléments sont donnés sur la méthode à mettre en œuvre :

« L'échantillon typique suppose que la sélection des unités soit effectuée non pas dans l'ensemble total, mais dans chaque groupe typique séparément. Dans un groupe typique lui-même, la sélection est effectuée de la même manière que pour un échantillon aléatoire. » [ibid. p.172]

Ce qui est proposé ici est une combinaison entre un tirage aléatoire et ce qui était nommé « tirage typique » : le tirage aléatoire était utilisé dans un deuxième temps, en complément d'une sélection d'éléments typiques effectuée tout d'abord. Le premier temps du travail consistait à découper l'ensemble total en sous-ensembles homogènes du point de vue de critères sociaux ou économiques jugés typiques. En fait, cela revenait à procéder à un découpage en strates sans utiliser le terme. Pourtant la méthode de l'échantillon stratifié avait déjà été exposée dans le manuel publié en 1924 par Kovalevski, qui était statisticien dans un bureau régional de la TsSOu [Mespoulet, 2000]. Mais le projet de formulation d'une nouvelle théorie statistique adaptée à une nouvelle forme d'économie et de société exigeait de nommer différemment les outils et dispositifs utilisés.

Dans la démarche préconisée, il y avait assimilation entre homogénéité et caractère typique, celui-ci pouvant être défini non pas par rapport à la moyenne, comme à la fin du XIXe siècle, mais par rapport à des critères sociaux, économiques ou organisationnels fixés *a priori*, par exemple la taille d'un groupe de kolkhozes. Ensuite, une fois le découpage en

groupes typiques effectués sous le contrôle du raisonnement humain, quel qu'en fût le fondement, politique ou scientifique, un tirage au hasard pouvait être mis en oeuvre pour constituer l'échantillon à observer. Dès lors, le hasard pouvait agir à l'intérieur des limites définies par la représentation politique de l'économie et de la société soviétiques. Le critère de la précision des résultats fut avancé pour légitimer la nouvelle méthode d'échantillonnage proposée :

« L'échantillon aléatoire proprement dit, c'est-à-dire un tirage aléatoire effectué dans l'ensemble total, est moins précis qu'un tirage effectué dans chaque groupe typique séparément.

Une telle combinaison d'un tirage typique avec un tirage aléatoire (dénommée "tirage aléatoire stratifié"¹⁵) est le procédé d'organisation d'une enquête par sondage le plus avantageux. » [ibid., p. 172]

La grande différence entre l'utilisation d'une telle combinaison dans les années 1930 et celle qui en était faite à la fin du XIXe siècle réside dans le fait que, à cette époque-là, le caractère typique des unités sélectionnées pour constituer un échantillon était défini sur la base des aires homogènes constituées à partir des données d'un recensement antérieur. Était considéré comme typique tout élément extrait de cette aire. Dans les années 1930, le groupe, qui se substitue à l'aire, est constitué d'éléments dont le caractère typique est défini auparavant, *a priori*. Ce qui va donner l'homogénéité au « groupe typique » n'est pas le faible degré de dispersion des valeurs prises par le caractère étudié autour de sa moyenne, mais la manière dont on opère un regroupement entre individus ou unités autour d'un critère économique ou social défini à l'avance. Le découpage de ce qui est comparable à des strates ne se fait plus sur une base territoriale seulement, mais avant tout sur la base d'un groupe social ou économique sur lequel le statisticien doit centrer le regard.

L'exemple d'une enquête par sondage sur les rendements agricoles est donné en illustration :

« Il est nécessaire de sélectionner, dans la surface totaleensemencée, un nombre de parcelles relativement petit, dont la récolte moyenne correspondrait avec la récolte moyenne de l'ensemble de la surfaceensemencée. Si l'on organise la sélection de l'échantillon selon la méthode aléatoire proprement dite, alors il faudrait diviser toute la surfaceensemencée en parcelles et ensuite, en effectuant un tirage au sort, prendre celles sur lesquelles le sort est

¹⁵ *Raionirovannyi sloutchajnyi otbor*

tombé. Un tel procédé de sélection dans le cas d'un nombre suffisant de parcelles donnerait des résultats pleinement satisfaisants en ce qui concerne la récolte moyenne.

Cependant, diviser toute la surface ensemencée en groupes typiques apporte plus d'avantages. Toute la surface peut être divisée selon le type d'exploitations (sovkhozes, kolkhozes, exploitations de paysans individuels). Déjà, une telle répartition crée des ensembles partiels qui ont une récolte plus homogène. C'est pourquoi une sélection effectuée dans chaque type d'exploitations séparément donne un résultat plus précis. » [ibid.]

En fait, ce qui crée l'homogénéité dans un groupe de kolkhozes ou de sovkhozes, c'est le statut des exploitations selon le type de propriété. En déduire que cela « crée des ensembles partiels qui ont une récolte homogène » est toutefois excessif, le lien entre récolte et type de propriété n'étant pas démontré.

Toutefois, est-ce l'expression d'un compromis dans la présentation ou pas, les auteurs de ce chapitre sur les sondages proposaient d'affiner le découpage des groupes typiques pour améliorer la précision de la mesure :

« Il faudrait effectuer ensuite une autre division en groupes typiques selon la qualité du sol, le relief, le caractère et la quantité des engrais, etc. Un tirage aléatoire effectué dans de tels groupes typiques encore plus fractionnés donnerait un résultat encore plus précis. Une division ainsi fractionnée exige cependant une dépense énorme de forces et de moyens et, dans la pratique, n'est pas toujours à la mesure des forces à disposition. » [ibid.]

A peine suggéré, ce sondage à deux degrés est présenté comme lourd à organiser. Ne peut-on pas néanmoins interpréter cette remarque comme un effet de discours pour préserver la possibilité d'une exploitation plus appropriée de l'enquête par sondage ? Une telle présentation pourrait alors être considérée comme une manière d'introduire une marge de manœuvre pour les statisticiens dans l'utilisation des dispositifs d'échantillonnage, selon les contraintes politiques du moment. Dans ce cas, le sondage à deux degrés pourrait être interprété comme un dispositif d'adaptation à une situation politique donnée.

Comme à la fin du XIXe siècle, le tirage dénommé mécanique¹⁶ demeurait la méthode d'échantillonnage privilégié présenté dans le chapitre de ce manuel de 1936.

2. Attachement au tirage mécanique

¹⁶ Au XIXe siècle, le tirage mécanique (*mekhanitcheskii otbor*) était assimilé à un tirage au hasard, le plus couramment sous la forme d'un tirage des rangs 1, 11, 21, 31, etc. dans une liste.

La manière dont ses auteurs justifient le tirage mécanique éclaire les raisons de l'attachement des statisticiens soviétiques à celui-ci, à la suite de leurs prédécesseurs d'avant 1917 :

« Pour éviter des dépenses excessives, on a fréquemment recours au tirage appelé "mécanique". Dans ce cas, l'ensemble total est divisé de manière mécanique en un nombre donné de groupes, selon un procédé établi (l'ordre des cases d'un échiquier, une diagonale, une liste, etc.), et un tirage aléatoire est effectué dans ces groupes. Dans le cas d'un tirage mécanique, bien sûr, une division en groupes typiques, délimités de manière précise, n'est pas garantie. Toutefois, la division de l'ensemble total en groupes plus petits, même effectuée de manière mécanique, garantit habituellement une plus grande homogénéité de ces groupes par comparaison avec l'ensemble total. En raison de cela même, un échantillon obtenu de manière mécanique acquiert, même s'il n'a pas une taille considérable, les caractéristiques d'un échantillon constitué de manière typique et, par conséquent, donne des résultats plus précis que le tirage aléatoire proprement dit. » [ibid., pp. 172-173]

Ce qui fonde la supériorité du tirage mécanique réside donc dans le fait qu'il aboutisse à un échantillon qui possède les mêmes qualités qu'un échantillon typique. La clé d'une telle relation est donnée un peu plus loin dans le manuel. Selon les auteurs du manuel, le tirage mécanique peut être assimilé au découpage d'un ensemble en de très nombreuses strates ne réunissant que quelques unités, ce qui garantit une distribution régulière des unités de l'échantillon sur l'ensemble total et donc la faible dispersion du caractère étudié. Chaque unité doit être tirée ensuite au hasard dans chaque strate ainsi constituée, mais le découpage mécanique de l'ensemble garantit la régularité de la distribution de ces unités tirées qu'un tirage aléatoire seul ne pourrait pas garantir :

« Dans le cas d'un tirage mécanique, l'ensemble total est divisé en une grande quantité de strates par n'importe quel procédé mécanique, et, dans une seule strate, on prend seulement une unité ; en outre, d'un point de vue théorique, la sélection de cette unité doit être effectuée par un tirage au sort. » [ibid., p. 193]

Le tirage mécanique revient donc à diviser un ensemble total en un très grand nombre de strates, d'autant plus petites que le taux de sondage est élevé. En 1924, Kovalevski avait

démontré que la combinaison entre le tirage mécanique et la constitution de strates le plus homogènes possible permettait d'augmenter la précision des résultats d'une enquête par sondage [Mespoulet, 2000]. Ce raisonnement est repris dans le manuel de 1936, toutefois sans citer Kovalevski. Ensuite, dans un souci de clarté explicative, un exemple de l'intérêt de l'application du tirage mécanique est donné par les auteurs :

« Ainsi, dans le cas d'une enquête par sondage sur les rendements agricoles, un tirage mécanique de parcelles à l'intérieur de chaque catégorie d'exploitations peut être effectué soit dans l'ordre des cases d'un échiquier, soit en suivant un autre procédé analogue, ce qui donnerait la possibilité de répartir l'échantillon de manière régulière sur le territoire et de présenter à l'intérieur toutes les parcelles suffisamment significatives, et variées du point de vue de la qualité des sols. Prenons, par exemple, la masse totale d'une surface ensemencée, constituée de deux parcelles qui diffèrent par la qualité de leur sol. Dans le cas d'un tirage mécanique, toute la surface ensemencée peut être divisée en petits carrés et, dans chaque petit carré, on doit tirer, au sort ou selon un procédé analogue, les parcelles qui feront l'objet de l'enquête par sondage. Il n'est pas difficile de s'assurer que le nombre de petits carrés correspondant à une sorte de terre ou à l'autre soit proportionnel *grosso modo* à l'ensemble de la surface correspondant à la part de chaque sorte. » [ibid., p. 173]

Enfin, ces statisticiens considéraient que le tirage mécanique effectué à l'intérieur de catégories constituées selon la méthode typique donnait des résultats plus précis qu'un échantillon tiré de manière strictement aléatoire.

Le mode de raisonnement développé autour de cette méthode d'échantillonnage, qui combine découpage d'un ensemble ou d'une population en groupe typiques et « tirage aléatoire stratifié », éclaire sur la manière dont les auteurs du manuel de 1936 ont cherché à résoudre la contradiction entre le recours au tirage au hasard dans le calcul statistique et la représentation d'un monde maîtrisé par la raison humaine et la méthode d'analyse du matérialisme historique. La constitution *a priori* de catégories conformes à l'analyse de la structure de la société en classes ou de la structure des catégories d'entreprises ou d'exploitations collectives agricoles était la condition indispensable. Ainsi était trouvée une manière de concilier une forme de tirage au hasard, qui impliquait d'admettre l'idée de l'incertain, et la volonté politique de contrôle des outils statistiques. Ainsi aussi la théorie

des probabilités pouvait être présentée comme non contradictoire avec l'action volontariste de l'Etat soviétique.

La résolution de cette contradiction peut expliquer le maintien de cette méthode d'échantillonnage dans l'administration statistique soviétique jusqu'en 1991. Les légères modifications qui lui furent apportées au début des années 1950 ne changèrent rien sur le fond¹⁷. Ce qui fut nommé « stratification typique » désigne un mode particulier de construction d'un échantillon stratifié, sur la base d'une combinaison entre la définition de « groupes typiques » et un tirage aléatoire, qui répondit à la volonté politique d'opérer d'abord un découpage raisonné du réel en catégories de classification jugées pertinentes pour l'analyse de la structure de classe ou des différents types d'unités de l'appareil productif.

¹⁷ A ce sujet, voir [Mespoulet, 2004]

ANNEXE : Table des matières du manuel de 1936

B.S. Iastremski et V.I Khotimski (dir.), Statistika. Osnovy obchtcheï teorii (Statistique. Fondements d'une théorie générale), 2^e édition, Moscou, Sotsekgiz, 421 pages¹⁸

Chapitre 1 : L'objet de la statistique

Chapitre 2 : L'enquête statistique

Chapitre 3 : La méthode de constitution de groupes

Chapitre 4 : Notion de valeurs relatives

Chapitre 5 : Théorie des processus spontanés et aléatoires

Chapitre 6 : La méthode du sondage

Chapitre 7 : Les moyennes

Chapitre 8 : Les index

Chapitre 9 : Une série de variations

Chapitre 10 : La mesure d'une relation

Chapitre 11 : Les séries

Chapitre 12 : Les variations d'une évolution

Chapitre 13 : Méthodes de représentation graphique

¹⁸ Le plan de cet ouvrage est, on le voit, assez classique. Le travail de refonte théorique porta bien plus sur la présentation des outils et de leurs usages que sur la structure même de l'enseignement de la statistique.

Bibliographie

[Blum, 1994] A.Blum, Naître, vivre et mourir en URSS, 1917-1991, Paris, Plon.

[Blum et Mespoulet, 2003] A.Blum et M.Mespoulet : L'anarchie bureaucratique. Statistique et pouvoir sous Staline, Paris, La Découverte.

[Desrosières, 1993], A.Desrosières, La politique des grands nombres. Histoire de la raison statistique, Paris, La Découverte.

[Gumbel, 1928] E.J.Gumbel, Klassenkampf und Statistik. Eine programmatische Untersuchung (traduction française de S. Hertz: Statistique et lutte des classes. Réflexion programmatique), in [Martin, 2000], 189-210.

[Hertz, 2000] S. Hertz, Statistique de l'Etat et statistique mathématique. Un texte-manifeste remarquable d'Emil Julius Gumbel : « Statistique et lutte des classes », in [Martin, 2000], 163-187.

[Iastremski et Khotimski, 1930] B.S. Iastremski et V.I Khotimski (dir.), Teoria matematicheskoi statistiki (Théorie de la statistique mathématique), Moscou, Plankhozgiz.

[Iastremski et Khotimski, 1931], B.S. Iastremski et V.I Khotimski (dir.), Teoria matematicheskoi statistiki (Théorie de la statistique mathématique), 2^e édition, Moscou-Leningrad, Sotsekgiz..

[Iastremski et Khotimski, 1931], B.S. Iastremski et V.I Khotimski (dir.), Obchtchaïa teoria statistiki (Théorie générale de la statistique), Moscou, Gosfinizdat.

[Iastremski et Khotimski, 1932], B.S. Iastremski et V.I Khotimski (dir.), Statistika (La statistique), Moscou-Leningrad, Ekonomgiz.

[Iastremski et Khotimski, 1933], B.S. Iastremski et V.I Khotimski (dir.), Elementy obchtcheï teorii statistiki (Eléments d'une théorie générale de la statistique), Moscou-Leningrad.

[Iastremski et Khotimski, 1934] B.S. Iastremski et V.I Khotimski (dir.), Statistika, osnovy obchtcheï teorii (La statistique, fondements d'une théorie générale), Moscou-Leningrad, Sotsekgiz.

[Iastremski et Khotimski, 1935] B.S. Iastremski et V.I Khotimski (dir.), Elementy obchtcheï teorii statistiki (Eléments d'une théorie générale de la statistique), 2^e édition, Moscou, Soiouzorgoutchet.

[Iastremski et Khotimski, 1936], B.S. Iastremski et V.I Khotimski (dir.), Statistika. Osnovy obchtcheï teorii (Statistique. Fondements d'une théorie générale), 2^e édition, Moscou, Sotsekgiz.

[Ianson, 1913] Iou. Ianson, Teoria statistiki (Théorie de la statistique), Saint-Pétersbourg.

[Kabloukov, 1912] N.A. Kabloukov, Possobie k oznakomleniou s ousloviami i priemami sobirania i razrabotki svedenii pri zemskikh statistitcheskikh issledovaniakh (Manuel sur les conditions et les procédés de collecte et de traitement des données des enquêtes des statisticiens des zemstva), Moscou, Leman.

[Kabloukov, 1922] N.A. Kabloukov, Statistika. Teoria i metody (Statistique. Théorie et méthodes), Moscou.

[Kaufman, 1912], A.A. Kaufman, Teoria i metody statistiki (Théorie et méthodes de la statistique), Moscou, izd. I.D. Sytina.

[Kaufman, 1922] A.A. Kaufman, Statistitcheskaïa nauka v Rossii (La science statistique en Russie), Moscou.

[Kovalevski, 1924] A.G. Kovalevski, Osnovy teorii vyborotchnogo metoda (Fondements de la théorie de la méthode du sondage), Outchenye zapiski Saratovskogo gossouidarstvennogo ouniversiteta, tome II, fasc. 2, 60-138.

[Martin, 2000] Th.Martin (dir.), Mathématiques et action politique. Etudes d'histoire et de philosophie des mathématiques sociales, Paris, Ed. de l'Ined, coll. Etudes et Enquêtes.

[Mespoulet, 2000] M.Mespoulet, Du tout à la partie. L'âge d'or du sondage en Russie (1885-1924), Revue d'études comparatives Est-Ouest, Vol. 31, 2, 5-49.

[Mespoulet, 2001] M.Mespoulet, Statistique et révolution en Russie. Un compromis impossible (1880-1930), Rennes, Presses Universitaires de Rennes.

[Mespoulet, 2004] M.Mespoulet, Qu'est-ce qu'une statistique socialiste ? Essai de sociologie de la statistique appliqué au cas soviétique, Mémoire d'Habilitation à diriger des recherches, Paris, Institut d'Etudes Politiques, 456p.

[Ovsienko, 1978] V.Ovsienko, Statistitcheskoe obrazovanie v SSSR (La formation à la statistique en URSS), Vestnik statistiki, 7, 68-75.