REVUE DE STATISTIQUE APPLIQUÉE

W. EDWARDS DEMING

Mise en pratique de l'échantillonnage à l'échelon national dans une enquête sur les consommateurs

Revue de statistique appliquée, tome 10, n° 1 (1962), p. 79-85

<http://www.numdam.org/item?id=RSA_1962__10_1_79_0>

MISE EN PRATIQUE DE L'ÉCHANTILLONNAGE A L'ÉCHELON NATIONAL DANS UNE ENQUÊTE SUR LES CONSOMMATEURS (1)

W. Edwards DEMING

Conseil en Enquêtes Statistiques

BUT PROPOSE -

Le but du présent article est d'illustrer le calcul de l'écart-type des variations d'échantillonnage et de la variable selon le principe de l'échantillonnage répété. L'exemple est tiré d'un échantillonnage national s'intégrant à une enquête sur les préférences des consommateurs.

Qu'est-ce que l'échantillonnage probabiliste ? Je profite de cette occasion pour commencer par énumérer quelques utilisations de l'échantillonnage probabiliste au sein d'une organisation spécialisée dans les études sur les consommateurs :

a) Nouvelle façon satisfaisante de poser un problème, et de le formuler en des termes riches de sens (modèle statistique), afin qu'une enquête statistique fournisse des renseignements utiles à un prix raisonnable.

b) Spécification et préparation d'un cadre convenant à un échantillonnage (liste des régions, des firmes commerciales, des gens, des articles ou d'autres entités ; limites de température, de concentration, de vitesse, etc.).

c) Spécification du processus du choix d'un nombre d'unités d'échantillonnage convenant à l'enquête (de façon invariable avec des nombres au hasard ; parfois avec stratification, parfois avec des échantillonnages complexes et possibilité de fractionnement). Les processus d'échantillonnage comportent des règles prévoyant des visites supplémentaires chez les absents de la première ou des suivantes).

d) La formule ou le processus d'établissement des estimations, selon le processus de sélection. Le problème central de la technique de l'échantillonnage est la manière de procéder à la sélection et à l'estimation pour recueillir le maximum d'information par rapport au prix de revient unitaire sans qu'aucune restriction soit imposée.

e) La formule ou le processus de formation des écarts-types, afin de mesurer la marge d'incertitude qui découle des variations de la matière recueillie et de celles du rendement des enquêteurs, des employés chargés de coder et d'établir les cartes perforées, etc.

---

(1) Communication présentée au séminaire sur les applications industrielles de la Statistique - Paris, 4 et 5 septembre 1961.

f) Plan d'enregistrement des résultats et de contrôle statistique.

- pour aider à superviser le travail et à détecter les irrégularités de résultats à l'extérieur et dans les bureaux,

- pour détecter et évaluer les erreurs indépendantes de l'échantillonnage, en particulier les fautes persistantes dans le travail des enquêteurs, des employés chargés de coder et de perforer les cartes ou autres.

g) Interprétation des résultats à l'usage des non-statisticiens.

h) Analyse des résultats des contrôles statistiques, des frais, et des composantes de la variance à l'intérieur des zones et entre elles, ainsi que d'un enquêteur à un autre, en vue :

- d'améliorer le plan d'échantillonnage pour les enquêtes à venir et de parvenir à une meilleure répartition des ressources,

- d'améliorer la sélection des enquêteurs, leur entraînement et leur surveillance,

- d'améliorer le travail de codage et de contrôle.

Il peut sembler évident d'après la liste ci-dessus que l'échantillonnage probabiliste ne consiste pas simplement à choisir des unités d'interview ou d'essai. Bien plus que celà, c'est la mise en pratique continue du calcul des probabilités en vue de parvenir à l'amélioration du travail extérieur et de bureau et à une utilisation optimale des ressources au profit du but central défini au paragraphe d. ci-dessus.

Certains qualifieraient cette activité de recherche opérationnelle, d'analyse des systèmes ou quelque chose d'approchant, mais pour un statisticien l'échantillonnage probabiliste dont j'ai tracé ci-dessus les grandes lignes est seulement une bonne pratique des statistiques et constitue són activité depuis de nombreuses années.

L'échantillonnage probabiliste n'engendre pas de problèmes et ne les influence pas. C'est l'application de techniques statistiques selon les principes énoncés plus haut, au service de la conception, de la fabrication, de la mise en valeur d'un produit sur le marché, de sa distribution et d'autres activités humaines.

QUELQUES DEFINITIONS -

La population, objet d'une étude statistique, est l'ensemble des personnes, des firmes ou de la matière, des circonstances, des concentrations, des modèles, des classes, etc., au sujet desquels on veut tirer des conclusions ou que l'on veut influencer, qu'ils soient accessibles ou non. La population sera clairement définie si le problème est soigneusement posé. Prenons par exemple toutes les firmes qui fabriquent un produit donné, ou qui sont susceptibles de l'acheter ; ou toutes les ménagères, tous les écoliers, toute la matière ou toutes les pièces contenues dans un lot, ou couvertes par un contrat ou une spécification. La population peut consister en certains documents intérieurs d'une société, le but de l'étude étant d'estimer le rapport découlant de certaines transactions.

La définition de la population n'est pas statistique ; elle a son origine dans la connaissance du sujet grâce auquel on pressent l'existence d'un problème et l'on se fait une idée de la façon de l'aborder. La population serait la même dans n'importe quel problème, que l'on ait ou non l'intention de faire appel à l'échantillonnage.

Intervient alors la décision concernant le cadre, qui fait partie du modèle statistique. Le "cadre" a été décrit pour la première fois par Stephan(1). Le cadre est un ensemble de matières physiques appelées unités d'échantillonnage (pièces constitutives, statistiques de recensement, cartes géographiques, listes, annuaires, documentation) nous permettant de mettre la main sur les éléments de la population, un à un. Pour être utile, le cadre doit englober une zône suffisante et une gamme assez étendue de conditions. Aucun plan de sondage ne peut surmonter les insuffisances du cadre. La déduction statistique effectuée selon la théorie probabiliste ne couvre que le cadre donné et les conditions dépendant du principe d'échantillonnage. Les généralisations à d'autres villes, à d'autres climats, conditions ou classes que n'englobe pas le cadre et qui ne dépendent pas du principe d'échantillonnage demandent à être jugées subjectivement et non interprétées selon la théorie statistique. Ainsi, si nous faisons l'essai d'un produit au moyen d'une enquête menée de la façon la plus experte qui soit à Chicago, aucune théorie statistique ne permettra de généraliser les résultats pour y inclure Denver. Ceci ne peut être accompli qu'à force de jugement subjectif n'engegeant que la responsabilité de son auteur.

Un multiple convenable de l'écart-type permet d'évaluer l'ampleur de l'incertitude possible des résultats pouvant raisonnablement être imputée à l'échantillonnage et aux erreurs variables d'exécution.

Le but du contrôle statistique est d'évaluer toute erreur d'exécution persistante ayant pu se produire.

Dans un échantillonnage probabiliste soigneusement appliqué, nous ne laissons donc pas subsister de doute concernant les incertitudes pouvant résulter de la sélection répétée et de la façon de procéder, ainsi que de fautes d'exécution.

Un contrôle statistique consiste à examiner à nouveau une petite fraction de l'échantillonnage principal, depuis le livre contenant les nombres au hasard jusqu'au codage et à la perforation en passant par le travail extérieur. Le but du contrôle statistique est de détecter et de mesurer l'effet de toute omission persistante, ou de la surestimation ou sous-estimation de n'importe quelle caractéristique. Le but du contrôle statistique n'est pas de corriger l'échantillonnage principal, mais de fournir les renseignements qui permettront d'en interpréter les résultats.

Il est important de se rappeler que les faiblesses introduites dans une étude par la faute d'un cadre incomplet, ou par un défaut dans la façon de procéder aux essais ou dans la technique des questions ou de la manière d'enquêter ne constituent pas des erreurs d'échantillonnage, pas plus d'ailleurs que si l'enquête portait sur l'ensemble de la population. Les faiblesses incorporées sont toujours là, qu'il s'agisse d'échantillonnage ou de population totale. Elles ne peuvent être décelées et minimisées qu'en fonction d'une meilleure connaissance de la matière d'étude, éventuellement épaulée par des comparaisons de fait entre deux ou plusieurs méthodes d'essai (dans lesquelles une théorie statistique du plan peut être d'un grand secours).

---

(1) Frederick F. Stephan dans "Practical Problems of sampling procedure. Amer. Soc. Rev. Vol. 1 - 1936 - p. 569-580. Il n'a pas introduit le terme de "frame" (cadre), mais il mentionne explicitement ce concept.

Maintenant, la division logique de la responsabilité devient claire. Le contenu du questionnaire, la méthode d'essai ou d'enquête, le codage ; et finalement, après terminaison de l'enquête, la généralisation des résultats au-delà du cadre et des conditions ne dépendant pas du processus d'échantillonnage, ressortent de la responsabilité des experts connaissant la matière à étudier. Par contre, le plan probabiliste d'un échantillonnage ou d'une expérience et les déductions objectives qui sont possibles avec la théorie probabiliste sont la responsabilité du statisticien. Les mesures à prendre, tenant compte ou non des résultats d'une étude, dépendent de la direction. En qualité de statisticien, j'estime que les résultats statistiques et l'analyse se défendent eux-mêmes, et je ne fais pas de recommandations au sujet d'une décision qu'il appartient à la direction de prendre.

UN SIMPLE EXEMPLE DE CALCUL SUR UN ECHANTILLONNAGE REPETE

L'échantillonnage est utilisé de bien des façons dans les recherches portant sur les consommateurs. Cet article ne fournit qu'un exemple, mais il serait possible de décrire bien d'autres modes d'application, comme celui des échantillons superposés destinés à comparer les préférences concernant la contenance, la taille et le style de l'amballage, le prix, le goût, les traits caractéristiques (comme le pouvoir d'achat) des lecteurs des magazines, la dégustation des nourritures et boissons permettant de comparer à des témoins l'influence de la publicité et bien d'autres choses. Le résultat d'un échantillonnage ne comportant pas d'écart-type est d'une utilité limitée. Heureusement, il est possible, par la répétition d'un plan d'échantillonnage, de réduire à la simple arithmétique le calcul d'un écart-type.

A la base, la répétition scinde un échantillon en deux ou plusieurs parties indépendantes (2). L'échantillon est l'ensemble des parties, et le but de la répétition est :

- de faciliter le calcul de l'écart-type et également de toute erreur systématique due à la formule s'il en existe une,

- de faciliter la détection des fautes d'exécution, sur lesquelles nous reviendrons.

L'exemple que je donne ici est tiré d'une enquête à l'échelle nationale exécutée selon un plan à répétition. Cet exemple vise à démontrer la simplicité de calcul de l'écart-type lorsque le plan est répété. Le symbole x représente la réponse oui, et le symbole y les unités d'habitation. En se basant sur les habitations, on peut se servir des chiffres du Recensement dans l'estimation par rapport au nombre total des femmes qui auraient répondu oui à la question du formulaire.

L'ampleur de l'enquête dont est tiré l'exemple était limitée, puisqu'elle ne comportait que 1 200 interviews environ. Le nombre de domiciles unitaires de l'enquête ($y = y_1 + y_2 = 1\ 961$ au tableau) n'est pas celui des interviews, mais le nombre d'unités d'habitation dans les 400 fragments de zones choisis pour l'échantillon. L'univers était celui des ménagères, et la question était celle-ci : "Faites-vous vous-même le fondant à glacer vos gâteaux ?".

---

(1) P.C. Mahalanobis - "On large-scale sample-surveys" - Phil. Trans. de la Royal Statistical Society - vol. 231 B - 1944 - p. 329-451 ; "Recent Experiments in statistical sampling" (J. Royal Statistical Society, vol. cix, 1946 - p. 325-48 ; D.B. Lahiri "On the National Sample Survey" Sankhyà vol. 14 - p. 264-316.

Quelques habitations n'abritaient pas de maîtresse de maison à laquelle la question puisse être posée ; de plus, il y a toujours une proportion d'absence de réponse ; il s'ensuit que le nombre y sera supérieur à x.

Les 400 fragments de zones n'ont pas été utilisés comme un échantillon, mais comme le total de deux échantillons indépendants dénommés sous-échantillon 1 et sous-échantillon 2, chacun comportant environ 200 fragments de zone. Les transcriptions font ressortir les deux sous-échantillons 1 et 2. Les symboles NE, NC, etc. en tête représentent des régions de Recensement Nord-est, Nord-centre, Sud et Ouest. La plan d'échantillonnage, y compris les formules, se trouvent au chapitre 11 de mon ouvrage(1).

Le coefficient de variation calculé ici tient compte du rendement variable des enquêteurs, et également de la variance entre enquêteurs étant donné que l'on n'a pas essayé dans le cas présent de répartir enquêteurs et sous-échantillons selon un plan orthogonal pour évaluer séparément la variance entre les enquêteurs (paragraphe suivant). En d'autres termes, l'acart-type calculé ici mesure la totalité de la variabilité indépendamment de sur ou de sous-estimations persistantes qui (comme on l'a vu plus haut) ne peuvent être mesurées que par un minutieux contrôle statistique.

Les résultats démontrent que $p = 0,40$ est une estimation de la proportion de femmes qui préfèrent préparer elles-mêmes le fondant à glacer leurs gâteaux chez elles, et que $\sigma = 0,01$ est l'écart-type de cette estimation (ou, plutôt, l'écart-type du procédé d'échantillonnage ayant donné p). Le nombre de degrés de liberté dans cette estimation de l'écart-type est légèrement inférieur à 8 parce que les variances apportées par les différentes régions géographiques ne sont pas égales(2).

La limite de confiance supérieure à 1 %, pour la proportion estimée, basée sur 7 degrés de liberté, est 0,40 plus 3 écarts-types, soit 0,43. La limite de confiance inférieure à 1 % serait de 0,40 moins 3 écarts-types, soit 0,37. Les limites de confiance supérieure et inférieure à 1 % de la proportion estimée sont donc 0,43 et 0,37.

L'écart type de toute erreur caractéristique mesurée par l'enquête peut être calculé de la même manière. Habituellement, on ne calcule que les écarts-types d'importance capitale, telles que le nombre total d'unités d'habitation aux Etats-Unis (aux fins de comparaison), le nombre total d'acheteurs d'un produit, la part du marché que l'on a, celle d'un concurrent principal, et éventuellement quelques autres caractéristiques. A ce propos, je peux ajouter que l'estimation du nombre total d'unités d'habitations aux Etats-Unis, d'après la présente enquête, est ressortie inférieure à 2 % au chiffre du Recensement, soit environ 1 écart-type en moins. Il est possible d'interpréter cette différence et son écart-type comme la preuve d'une possible insuffisance d'échantillonnage dans les fragments de zones de peu d'importance.

---

(1) W.E. Deming - Sample design in business research - J. Wiley - 1960.

(2) On peut estimer le nombre de degrés de liberté d'après une formule donnée par F.E. Satterthwaite : "An approximate distribution of estimates of variance components" - Biometrics, Vol. 2 - 1946 - p. 110-114. La formule de Satterthwaite est représentée dans l'appendice au chapitre 11 du libre cité en haut de page.

AUTRES AVANTAGES DE LA REPETITION

La répétition de l'échantillon présente d'autres avantages que le calcul immédiat de l'écart-type. Il permet d'évaluer presque instantanément le biais, s'il existe, de la fonction ayant servi à formuler l'estimation.

Un autre avantage de la répétition, noté en premier par Mahalanobis, et dont j'ai constaté la grande utilité, est que l'on peut prévoir l'échantillon de sorte que les sous-échantillons soient traités à l'extérieur, lors de la codification et même lors de la perforation et de la tabulation par des groupes d'employés distincts. Les sous-échantillons se concurrencent alors mutuellement. De gros désaccords entre eux (pouvant être mis en évidence par un test de signification) peuvent indiquer une erreur d'exécution à déceler et à corriger. Une telle utilisation des sous-échantillons ne remplace pas les autres contrôles statistiques, mais elle est très utile, particulièrement dans la détection de grossières fautes d'exécution, et pour mesurer les variances entre les enquêteurs.

Inversement, on peut interpréter un écart-type trop faible dans ces conditions (dont l'exemple est illustré plus haut) comme un défaut de preuve (a), de fautes d'exécution, et (b) de grosses différences entre les enquêteurs.

CALCUL D'UN ECART-TYPE DANS LE CAS D'UN PLAN A REPETITION -

x représente le nombre de femmes ayant répondu oui à la question "Faites-vous vous-même le fondant à glacer vos gâteaux, ou achetez-vous un mélange ?"

y représente le nombre d'unités d'habitation. Les indices 1 et 2 se rapportent aux deux sous-échantillons.

p représente une estimation de la proportion des femmes qui répondraient oui si l'enquête couvrait 100 % des foyers.

$\hat{\sigma}_p$ représente l'estimation de l'écart-type de p.

NE, NC, S et W représentent les 4 régions du Recensement : Nord-est, Nord-centre, Sued et Ouest.

Caractéristique	Métropolitain				Non-Métropolitain				U.S.
	NE	NC	S	W	NE	NC	S	W	
i =	1	2	3	4	5	6	7	8	1 à 8
x_1	98	100	40	38	35	33	47	33	424
x_2	79	69	38	45	17	38	52	29	367
$x = x_1 + x_2$	177	169	78	83	52	71	99	62	791
$x_1 - x_2$	+19	+31	+2	-7	+18	-5	-5	+4	+57
y_1	213	203	123	87	64	94	174	62	1.020
y_2	176	162	110	89	55	128	165	56	941
$y = y_1 + y_2$	389	365	233	176	119	222	339	118	1.961
$y_1 - y_2$	+37	+41	+13	-7	+9	-34	+9	+6	+79
$p = \frac{x}{y}$	0,455	0,463	0,335	0,472	0,437	0,320	0,292	0,525	0,40
$h = (x_1 - x_2) - p(y_1 - y_2)$	2,16	12,01	-2,35	-6,05	14,06	5,88	-7,62	0,850	25,16
$S^2 = \Sigma h^2$									482,68
S									20,07
$\hat{\sigma}_p = S/y$									0,01

Estimation finale pour les U.S.A., p = 0,40 ; écart-type 0,01.