JOURNAL DE LA SOCIÉTÉ FRANÇAISE DE STATISTIQUE

MAURICE FRÉCHET

Sur une limitation très générale de la dispersion de la médiane

Journal de la société française de statistique, tome 147, nº 2 (2006), p. 5-15

<http://www.numdam.org/item?id=JSFS_2006__147_2_5_0>

SUR UNE LIMITATION TRÈS GÉNÉRALE DE LA DISPERSION DE LA MÉDIANE

Maurice FRÉCHET

Réédition d'un article (avec discussion) paru dans le
Journal de la Société de Statistique de Paris
Volume 81, n° 1, pages 67-77

SUR UNE LIMITATION TRÈS GÉNÉRALE [1] DE LA DISPERSION DE LA MÉDIANE

A. — QUELQUES GÉNÉRALITÉS SUR LA STATISTIQUE MATHÉMATIQUE.

J'espère qu'on me pardonnera de faire ici une confession personnelle. Après avoir porté mon principal effort scientifique pendant une dizaine d'années sur l'analyse générale, je me suis occupé surtout de calcul des probabilités, et, si je me suis intéressé assez souvent à la Statistique mathématique, c'est toujours occasionnellement, par périodes assez courtes.

Cependant, quand dans ces périodes je me mettais au courant des récentes publications statistiques, de nombreuses objections se présentaient .à mon esprit.

Je rencontrais parfois, même dans les meilleurs mémoires, tantôt des énoncés incomplets où toutes les restrictions n'étaient pas clairement rappelées, tantôt des conventions arbitraires présentées comme des assertions évidentes, tantôt des démonstrations insuffisantes. Telle était, par exemple, la situation qui avait donné lieu à un usage abusif du coefficient de linéarité, dit de corrélation, abus sur lequel deux motions récentes successives de l'Institut International de de Statistique sont venues d'ailleurs attirer l'attention. Plusieurs statisticiens

(1) Communication faite à la Société le 17 janvier 1940.

de mes amis se souviendront peut-être m'avoir entendu exprimer le regret d'être privé par d'autres travaux du temps nécessaire pour examiner en détail des questions qui avaient aussi soulevé mes doutes. Or, certaines circonstances, nées de la nouvelle guerre, me contraignent à consacrer maintenant la plus large part de mon temps à la Statistique mathématique. J'ai pu commencer à préciser les objections que j'avais dans l'esprit au sujet de deux questions différentes. Je me propose de traiter la première dans une suite de courts mémoires dont celui qui va suivre ces généralités est le second (1), dont d'autres sont déjà écrits et qui auront tous trait à la question générale suivante : Parmi les valeurs typiques, les écarts typiques d'une population statistique, n'a-t-on pas accordé trop souvent une place, — non seulement prédominante, ce qui pourrait s'admettre, — mais exclusive, ce qui est regrettable, à la valeur moyenne et à l'écart quadratique moyen. Certaines théories, qui paraissent conduire à un résultat définitif quand on adopte cette attitude exclusive, ne risquent-elles pas d'être ébranlées (je ne veux pas dire de devenir fausses ou inutiles, mais de cesser d'être sans appel), quand on opère d'une façon moins automatique? Et un plus grand éclectisme ne peut-il puissamment contribuer à diminuer l'ampleur des calculs des statisticiens et à augmenter, par suite, le rendement de leurs travaux?

Dans une autre direction, dont je m'occuperai plus tard, je compte faire des réserves sur les efforts louables qui ont été faits pour éviter l'*emploi*, toujours délicat, de la formule (elle, incontestable) de Bayes-Laplace. Des phrases équivoques m'avaient donné à penser que ces efforts reposaient sur l'identification (considérée comme évidente) des probabilités d'un même événement dans deux populations, deux catégories d'épreuves à mon avis totalement différentes. Ayant enfin rencontré un texte où cette assimilation était non plus implicite mais explicite, je venais de signaler ce texte à un autre statisticien de mes amis quand, quelques jours plus tard, et tout à fait indépendamment, j'ai reçu un exemplaire du texte d'un discours de M. Gini (2) où j'ai cru reconnaître des objections de même nature. Toutefois, ce discours ayant été prononcé en italien, je ne suis pas encore sûr de l'avoir bien compris (3). Mais ce que j'en ai deviné m'en fait pressentir l'importance. Sans être sûr de me trouver d'accord sur tous les points avec M. Gini, je suis convaincu qu'il donnera matière à réflexion à beaucoup de statisticiens et je souhaite vivement la publication d'une traduction (ou mieux de plusieurs traductions) de ce discours, soit sous son texte original, soit avec les modifications qu'inspireront à son auteur les commentaires qu'il en recevra.

D'ailleurs les objections que je veux développer ne seront point pessimistes. Mon cri d'alarme est en même temps un cri d'espoir, et c'est peut-être en quoi je me séparerai de M. Gini.

Les statisticiens viennent d'assister à un développement rapide — qui serait à certains égards plutôt un foisonnement — de la statistique mathéma-

---

(1) Le premier étant : *Sur la précision de la médiane*, AKTUARSKÉ VĚDY, Prague, 1936-1937.

(2) *I pericolosi della statistica* (*Rivista di politica economica*, 1939).

(3) En fait, après ma conférence, j'ai appris de l'auteur lui-même que si ses objections portaient bien sur les mêmes conclusions, elle concernaient une identification d'une autre nature que celle signalée plus haut.

tique. (Et il est curieux de voir, dans cette évolution, les Anglais en être l'élément « dynamique », selon le jargon du jour, alors que les Italiens et les Allemands y gardent plutôt une attitude « statique ».) Ce développement rappelle beaucoup l'évolution rapide et un peu désordonnée qui a suivi la découverte du Calcul différentiel et intégral. Après cette période brillante, des doutes se sont élevés sur la solidité des constructions entreprises et bientôt s'est ouverte une période de revision sévère dont Cauchy est le plus remarquable représentant. Cette revision n'a nullement arrêté l'évolution précédente; en traçant les limites exactes où les calculs entrepris, et plus encore leurs conclusions, restaient valides, elle a permis une application correcte et sûre des principes admis un peu trop rapidement. Cette critique même a donné lieu, elle aussi, — précisément pour se justifier, — à des contributions positives considérables comme celles du même Cauchy.

Je ne serais pas du tout étonné que les résultats de la Statistique mathématique moderne sur lesquels porteront des critiques, subsistent finalement, mais restreints à leurs véritables domaines de validité. Les méthodes mêmes auxquelles l'absence d'une base rationnelle dans les théories existantes serait reconnue subsisteront peut-être aussi grâce à l'adjonction de nouveaux axiomes ou sous la forme de méthodes empiriques de découverte.

C'est, en tout cas, le souhait que je forme.

B. — La médiane.

Écart quadratique moyen de la médiane. — R. A. Fisher (1) a démontré que pour de très grandes valeurs de n, on a approximativement :

$$\sigma_{M_n} = \frac{1}{2\delta\sqrt{n}} \tag{1}$$

M_n étant une médiane de n valeurs $X_1, \dots X_n$ prises par une variable aléatoire X dans n épreuves *indépendantes*, σ_{M_n} l'écart quadratique moyen de M_n, et δ la densité de probabilité de X pour sa valeur probable m (2).

Cette formule montre bien l'ordre de grandeur de la dispersion de la médiane M_n. Cependant, si l'on ne connaît pas l'écart quadratique moyen σ_X de X, elle ne permet pas de se rendre compte de la réduction $\frac{\sigma_{M_n}}{\sigma_X}$ de la dispersion, obtenue en remplaçant X par la médiane de n de ses valeurs. Nous nous proposons d'abord d'indiquer une borne supérieure de cette réduction, valable dans un cas très général.

Observons que si l'on met (1) sous sa forme rigoureuse :

$$\lim_{n \longrightarrow \infty} \sqrt{n}\, \sigma_{M_n} = \frac{1}{2\delta}, \tag{2}$$

(1) Pour les références bibliographiques relatives aux résultats cités ici, voir, par exemple : *La précision de la moyenne arithmétique et de la médiane*, par E. J. Gumbel, Aktuarskévědy, Prague, 1936-1937.

(2) Conformément à l'usage des artilleurs, nous appelons valeur probable de X la valeur m telle que $\text{Prob}\,\{ X < m \} = \frac{1}{2}$ (en supposant $\delta \neq o$, cette valeur est atteinte une fois et une seule).

on aura :

$$\lim_{n \to \infty} \sqrt{n} \frac{\sigma_{M_n}}{\sigma_X} = \frac{1}{2\delta\sigma_X} \tag{3}$$

et la question se trouve ramenée à obtenir une borne inférieure de $\delta\,\sigma_X$.

Autres estimations de la dispersion. — D'ailleurs, puisque nous devons calculer la dispersion de M_n dont la valeur probable (2) est fixe (1) et égale à celle, m, de X, il vaut mieux apprécier la dispersion de M_n au moyen de son écart moyen θ_{M_n} (c'est-à-dire de la valeur moyenne de $|M_n - m|$), plutôt qu'au moyen de σ_{M_n} calculé à partir de la valeur moyenne, en général variable avec n, de M_n. Or comme on démontre (1) que la loi (réduite) de probabilité de la médiane converge vers la seconde loi de Laplace prise avec m pour médiane (et pour valeur moyenne), le rapport $\frac{\theta_{M_n}}{\sigma_{M_n}}$ convergera vers le rapport classique $\sqrt{\frac{2}{\pi}} = 0{,}798$. En portant dans (2), on aura :

$$\lim_{n \to \infty} \sqrt{n} \frac{\theta_{M_n}}{\theta_X} = \frac{1}{\sqrt{2\pi}\,\theta_X\,\delta}. \tag{4}$$

De sorte que là encore, on est conduit à chercher une borne inférieure de $\theta_X\,\delta$.

Enfin, il serait également légitime d'estimer la dispersion de X au moyen de son écart probable E_X (moitié de l'intervalle entre les quartiers de X), cet écart probable ayant l'avantage de n'être pas plus lié à la position de la valeur moyenne de X qu'à sa valeur probable. Le raisonnement fait ci-dessus pour l'écart moyen peut se répéter ici, et comme $\frac{E_{M_n}}{\sigma_{M_n}}$ tend vers le rapport classique

$$0{,}4769\sqrt{2} = 0{,}6743\ldots$$

on aura :

$$\lim_{n \to \infty} \sqrt{n} \frac{E_{M_n}}{E_X} = \frac{0{,}4769}{\delta\,E_X\sqrt{2}}. \tag{5}$$

Ici encore on se trouve conduit à chercher une borne inférieure de $\delta\,E_X$.

Or on peut obtenir par des raisonnements géométriques élémentaires les bornes inférieures de $\delta\,\sigma_X$, $\delta\,\theta_X$, $\delta\,E_X$, si l'on fait sur la loi de probabilité de X une hypothèse très générale.

Observons d'abord que si une variable aléatoire X possède partout une densité de probabilité $\varphi(x)$ et si celle-ci est bornée, alors en appelant Δ la borne supérieure de $\varphi(x)$, on a :

$$\Delta\,\sigma_X > \frac{1}{2\sqrt{3}} \tag{6 bis}$$

$$\Delta\,\theta_X > \frac{1}{4} \tag{7 bis}$$

$$\Delta\,E_X > \frac{1}{4} \tag{8 bis}$$

et ces inégalités ne peuvent être améliorées, c'est-à-dire qu'on ne peut remplacer l'un des nombres au second membre par un nombre plus grand indépendant de X quand X jouit des deux propriétés mentionnées ci-dessus.

Et cet énoncé subsiste, même si X est seulement assujetti à ce que sa « densité moyenne » de probabilité soit bornée, c'est-à-dire qu'en appelant Δ sa borne supérieure, on ait :

$$\frac{\text{Prob}\,\{x_1 \leq X < x_2\}}{x_2 - x_1} < \Delta$$

quels que soient x_1 et x_2

Et dans ce cas encore, les formules (6 *bis*), (7 *bis*), (8 *bis*) ne peuvent être améliorées.

Nous reporterons, si c'est nécessaire, à un autre article la démonstration de ces formules, qui sont d'ailleurs peut-être connues.

Nous allons maintenant appliquer ces résultats au cas où X est une variable aléatoire dont la valeur probable m est, en même temps, une « dominante ». Il suffira même de supposer : 1° qu'il y a en m une densité de probabilité δ; 2° que la densité moyenne de probabilité est constamment $\leqq \delta$. Dans ces conditions, on a :

$$\delta\, \sigma_X > \frac{1}{2\sqrt{3}} \tag{6}$$

$$\delta\, \theta_X > \frac{1}{4} \tag{7}$$

$$\delta\, E_X > \frac{1}{4} \tag{8}$$

et ces inégalités ne peuvent être améliorées pour la classe de variables aléatoires que nous venons de définir.

Ceci étant, de la confrontation des relations (3), (4), (5) et (6), (7), (8) résulte immédiatement que :

Pour n assez grand, on a :

$$\sigma_{X_n} < \sqrt{3}\,\frac{\sigma_X}{\sqrt{n}} \tag{9}$$

$$\theta_{X_n} < 2\sqrt{\frac{2}{\pi}}\,\frac{\theta_X}{\sqrt{n}} \tag{10}$$

$$E_{X_n} < 2\sqrt{2}\,\lambda\,\frac{E_X}{\sqrt{n}} \tag{11}$$

où λ est égal à la racine 0,4769... de l'équation :

$$\frac{1}{2} = \frac{2}{\sqrt{\pi}} \int_0^{\lambda} e^{-x^2}\, dx.$$

Il résulte d'ailleurs de ce qui précède que les *coefficients* $\sqrt{3}$, $2\sqrt{\frac{2}{\pi}}$, $2\sqrt{2}\,\lambda$ qui figurent dans ces inégalités *ne peuvent être remplacés par des bornes plus avantageuses* — c'est-à-dire plus petites — et valables pour tous les cas où la valeur probable de X est en même temps dominante. En les calculant à $\frac{1}{100}$ près par excès, *on a, pour les mesures de la dispersion de la médiane M_n par son écart quadratique moyen σ_{M_n}, par son écart moyen θ_{M_n} ou par son écart probable E_{M_n}, les trois inégalités cherchées*, vraies pour n assez grand :

(9 *bis*) $$\boxed{\sigma_{M_n} < 1{,}74 \frac{\sigma_X}{\sqrt{n}}}$$

(10 *bis*) $$\boxed{\theta_{M_n} < 1{,}60 \frac{\theta_X}{\sqrt{n}}}$$

(11 *bis*) $$\boxed{E_{M_n} < 1{,}35 \frac{E_X}{\sqrt{n}}}$$

Comparaison avec la moyenne arithmétique. — On connaît la belle relation

(9 *ter*) $$\sigma_{V_n} = \frac{\sigma_X}{\sqrt{n}}$$

concernant l'écart quadratique moyen de la moyenne arithmétique $V_n = \frac{X_1 + \ldots\ldots + X_n}{n}$.

A première vue, en comparant (9 *bis*) et (9 *ter*), on est tenté de se rallier à l'affirmation trop souvent exprimée d'une supériorité incontestable et constante de la moyenne arithmétique sur la médiane, au point de vue de la rapidité de décroissance de la dispersion quand n croît.

Cependant, plusieurs auteurs (1) examinant le rapport $\frac{\sqrt{n}\,\sigma_{M_n}}{\sigma_X}$ pour des variables aléatoires de différentes espèces, ont déjà observé que celui-ci peut être plus petit que l'unité (valeur de $\sqrt{n}\,\frac{\sigma_{V_n}}{\sigma_X}$) non seulement pour des variables aléatoires choisies à cet effet et qu'on pourrait rejeter comme exceptionnelles, mais aussi pour des variables aléatoires dont les lois de probabilité sont très voisines des lois numériques des erreurs expérimentales effectivement relevées. Telle est, par exemple, la première loi de Laplace (où la probabilité élémentaire est de la forme $\frac{e^{-\frac{|x-m|}{\theta}}}{2\theta}\,dx$).

Ici, au lieu de considérer des lois particulières, nous voulons examiner les formules correspondant aux formules (9 *bis*), (10 *bis*), (11 *bis*), mais concernant V_n, pour la même classe très générale de variables aléatoires. A (9 *bis*) correspond (9 *ter*), qui est certainement en faveur de la moyenne arithmétique.

Mais, quand on veut comparer la précision de V_n et de M_n, il n'est pas équitable d'employer, pour estimer cette précision, l'écart quadratique moyen, notion intimement liée à la moyenne. Estimons donc la précision par l'écart probable. C'est-à-dire, cherchons pour V_n la formule correspondant à (11 *bis*).

Puisque la loi réduite de la moyenne arithmétique tend aussi vers la seconde loi de Laplace, on a :

$$\lim_{n \to \infty} \frac{E_{V_n}}{\sigma_{V_n}} = 0{,}6743.....$$

On tire donc de la formule (9 *bis*)

$$\lim_{n \to \infty} \sqrt{n}\, \frac{E_{V_n}}{E_X} = 0{,}6743.....\, \frac{\sigma_X}{E_X}. \tag{12}$$

Y a-t-il une borne supérieure du second membre et cette borne supérieure fournit-elle un coefficient plus avantageux que le coefficient 1,35 de la formule (11 *bis*)? Voilà ce qu'il faudrait prouver pour établir que la formule correspondant à (11 *bis*) est plus avantageuse pour V_n.

Or la réponse à la première question (et, par suite, aussi à la seconde) est négative, comme un raisonnement géométrique simple peut le montrer. En assimilant la densité de probabilité à une densité de masse, il suffit de laisser invariable une moitié de la masse (pour laisser E_X fixe) et d'éloigner suffisamment de part et d'autre le reste de la masse tout en laissant fixes la valeur probable et la valeur moyenne de X, pour augmenter σ_X indéfiniment.

Ainsi, la *formule qui correspond à* (11 *bis*), *c'est :*

$$\boxed{E_{V_n} < \text{infini} \times \frac{E_X}{\sqrt{n}}} \tag{11 ter}$$

et le *coefficient « infini » ne peut être amélioré*, c'est-à-dire qu'on ne peut remplacer ce coefficient par un coefficient numérique, fini, fixe et restant le même quand n est assez grand, pour toutes les lois de répartition de X où la valeur médiane est en même temps dominante (ce qui a lieu, par exemple, pour toutes les lois symétriques unimodales, mais qui a lieu dans des cas beaucoup plus généraux).

Dès lors, s'il est vrai que la formule (9 *ter*) donne par comparaison avec (9 *bis*) un avantage sensible à la moyenne arithmétique par rapport à la médiane, au point de vue de la décroissance de la dispersion, il est vrai aussi que la *formule* (11 *bis*) *donne en sens contraire, par comparaison avec* (11 *ter*), *un avantage infiniment plus considérable à la médiane par rapport à la moyenne arithmétique.*

La formule (9 *ter*), qui fournit une égalité, tandis que (9 *bis*) ne donne qu'une limitation, offre de tels avantages aux algébristes qu'on est tenté d'oublier l'énorme simplification, au point de vue du calcul numérique, résultant de l'emploi de la médiane et de l'écart probable au lieu de la moyenne arithmétique et de l'écart quadratique moyen. Cette simplification, indifférente aux théoriciens, est capitale pour les statisticiens; mais, trop souvent, ceux-ci hésitent cependant à en profiter en raison du renom d'imprécision de la médiane. Les considérations précédentes viennent s'ajouter aux remarques faites antérieurement dans un ordre d'idées différent par d'autres auteurs déjà signalés plus haut, pour encourager les statisticiens à ne pas attacher une importance excessive à ce slogan, à faire éventuellement des calculs plus simples, plus rapides, moins rébarbatifs et tout aussi savants, sous leur aspect plus élémentaire.

Confrontation avec l'expérience. — Au moyen d'une expérience dont le relevé sera publié ailleurs (1), on peut réaliser une expérience dépendant d'un paramètre arbitraire α.

On a tiré 96 fois, l'une des cartes d'un paquet de 10 cartes portant un n° Y égal à 1, 2, 3,..... 10, et on a constitué ainsi 8 séries de 12 tirages, reproduites dans un tableau ci-après (2). A Y, nous ferons correspondre un nombre X prenant les 10 valeurs $-\alpha, -4, -3, -2, -1, +1, +2, +3, +4, +\alpha$, avec $\alpha > 4$.

Dans chaque série, on a noté les 12 valeurs $Y_1, \ldots, Y_{12}$, qui figurent dans le relevé, et nous leur ferons correspondre 12 valeurs $X_1, \ldots, X_{12}$. On calcule $V_{12} = \dfrac{X_1 + \ldots + X_{12}}{12}$ pour chaque colonne du tableau ci-dessous et on a ainsi 8 valeurs $V^{(1)}, V^{(2)}, \ldots, V^{(8)}$ de V_{12}. Ces 8 valeurs sont :

$$\frac{6+2\alpha}{12}, \frac{10+\alpha}{12}, \frac{5}{12}, \frac{-4+\alpha}{12}, \frac{\alpha}{12}, \frac{-11-3\alpha}{12}, \frac{-3+2\alpha}{12}, \frac{12-\alpha}{12}.$$

Elles permettent de calculer empiriquement $\sigma_{V_{12}}$, en prenant (puisqu'évidemment la valeur moyenne de X, $\overline{X} = 0$, donc $\overline{V}_n = 0$.

$$(\sigma_{V_{12}})^2 = \frac{1}{8} \Sigma V_i^2 = \frac{1}{8 \times 12^2}[21\alpha^2 + 66\alpha + 451]$$

Résultats de 8 séries de 12 tirages.

Y	X	RÉPÉTITIONS D'UNE MÊME VALEUR DE X dans une même série							
1	$-\alpha$	0	0	0	2	1	3	1	1
2	-4	2	0	1	1	1	3	1	0
3	-3	1	1	2	0	1	1	1	0
4	-2	0	3	1	1	0	1	1	0
5	-1	1	0	1	3	3	0	2	3
6	1	0	1	2	0	1	2	1	3
7	2	1	2	2	1	1	2	0	3
8	3	4	2	0	1	1	0	1	2
9	4	1	2	3	0	1	0	1	0
10	α	2	1	0	3	2	0	3	0

D'autre part, la probabilité de chaque valeur de X étant évidemment $\frac{1}{10}$ on a

$$\sigma_X^2 = \frac{2(1 + 2^2 + 3^2 + 4^2 + \alpha^2)}{10} = \frac{30 + \alpha^2}{5}.$$

(1) Ce relevé, établi pour vérifier empiriquement la solution du problème des rencontres sera publié à la fin de la deuxième partie (Cas particuliers. Applications) de notre ouvrage : *Les probabilités associées à un système d'événements compatibles et dépendants*, dont la première partie (Événements en nombre fini fixe) vient de paraître dans notre « Collection d'exposés d'Analyse générale », chez Hermann, Paris.

(2) On a constitué ces séries au moyen des $12 \times 2 = 24$ premiers résultats figurant dans la première ligne de chacun des quatre tableaux du relevé mentionné à la note (1).

Si le calcul de $\sigma_{V_{12}}$ était bon (mais il faudrait pour cela avoir plus de 8 séries), on devrait avoir :

$$\sigma_{V_{12}} \,\underline{\Omega}\, \frac{\sigma_x}{\sqrt{12}}, \quad (\underline{\Omega} \text{ représente une égalité approximative})$$

ou :

$$\sqrt{12}\,\frac{\sigma_{V_{12}}}{\sigma_x} \,\underline{\Omega}\, 1,$$

ou, enfin :

$$\frac{1}{12}\sqrt{\frac{21\,\alpha^2 + 66\,\alpha + 451}{\alpha^2 + 30} \times \frac{15}{2}} \,\underline{\Omega}\, 1.$$

En réalité, le premier membre est égal, pour $\alpha = 5$, à $1{,}11$ et pour $\alpha = \infty$, à :

$$\frac{1}{12}\sqrt{21 \times \frac{15}{2}} = \frac{\sqrt{157{,}5}}{12} = \frac{12{,}55}{12} = 1{,}04.$$

Ces valeurs ne sont pas en somme très éloignées de l'unité, étant donné le petit nombre, 8, des séries utilisées, et il en est ainsi quel que soit α.

Au contraire, considérons maintenant le rapport correspondant pour les écarts probables. Pour $\alpha = 5$, les valeurs de V_{12} rangées par ordre de grandeur sont :

$$-\frac{26}{12}, +\frac{1}{12}, +\frac{5}{12}, +\frac{5}{12}, +\frac{7}{12}, +\frac{7}{12}, +\frac{15}{12}, +\frac{16}{12},$$

dont l'écart probable est :

$$\frac{1}{2}\left[\frac{11}{12} - \frac{3}{12}\right] = \frac{8}{24} = 0{,}33.$$

Pour α très grand, les termes indépendants de α dans les valeurs de V_{12} sont négligeables, de sorte que ces valeurs par ordre de grandeur deviennent :

$$\frac{-3\,\alpha}{12}, \frac{-\alpha}{12}, 0, \frac{\alpha}{12}, \frac{\alpha}{12}, \frac{\alpha}{12}, \frac{2\,\alpha}{12}, \frac{2\,\alpha}{12}$$

dont l'écart probable est :

$$\frac{1}{2}\left[\frac{1{,}5\,\alpha}{12} - \frac{(-0{,}5)\,\alpha}{12}\right] = \frac{\alpha}{12}.$$

D'autre part, si $\alpha > 4$, E_x est évidemment $\frac{1}{2}[3{,}5 - (-3{,}5)] = 3{,}5$; de sorte que $\sqrt{12}\,\frac{E_{V_{12}}}{E_x}$, qui, pour $\alpha = 5$, est égal à :

$$\sqrt{12}\,\frac{0{,}33}{3{,}5} = 0{,}32,$$

devient, pour α très grand, égal asymptotiquement à :

$$\frac{\sqrt{12}\,\frac{\alpha}{12}}{3{,}5} = \frac{\alpha}{3{,}5\,\sqrt{12}}$$

et, par suite, *n'a pas de borne supérieure* quand α varie : il peut être aussi grand que l'on veut en prenant α assez grand, ce qui vérifie pratiquement, sur cet exemple, notre affirmation générale résumée dans la formule (11 *ter*).

En résumé, tandis que $\sqrt{12}\,\frac{\sigma_{V12}}{\sigma_X}$ ne reste pas très éloigné de 1, même quand α croît indéfiniment, $\sqrt{12}\,\frac{\sigma_{V12}}{E_X}$ peut dépasser n'importe quelle valeur positive en prenant α assez grand.

Au contraire, considérons la médiane M_n. Ses valeurs dans les 8 séries sont :

$$3, 2, 1, -1, 0. -3, 0, 1,$$

qui, rangées dans l'ordre de grandeur

$$-3, -1, 0, 0, 1, 1, 2, 3,$$

donnent une valeur empirique de l'écart probable :

$$E_{M12} \underline{\Omega} \frac{1{,}5-(-0{,}5)}{2} = 1.$$

D'où :

$$\frac{\sqrt{12}\ E_{M12}}{E_X} = \frac{\sqrt{12} \times 1}{3{,}5} = 0{,}989,$$

valeur qui est indépendante de α et voisine de 1. On voit que pour de grandes valeurs de α, la précision de la médiane, quand on la mesure, au moyen de l'écart probable, est considérablement meilleure que celle de la moyenne arithmétique.

Quand on mesure la précision de la médiane par son écart quadratique moyen, on forme :

$$(\sigma_{M12})^2 = \frac{1}{8}(9+4+1+1+0+9+0+1) = \frac{25}{8} = 3{,}125.$$

D'où :

$$\sqrt{12}\,\frac{\sigma_{M12}}{\sigma_X} = \sqrt{\frac{12 \times 3{,}125}{6+\frac{\alpha^2}{5}}} < 2{,}35$$

pour $\alpha > 4$. Et même le premier membre, qui devient $< \sqrt{3}$, comme dans (9), pour $\alpha > 6$, est < 1 pour $\alpha > 13$ et tend vers zéro quand α croît, alors que $\frac{\sqrt{12}\,\sigma_{V12}}{\sigma_X}$ tend vers 1,04.

Maurice Fréchet.

DISCUSSION

M. le Président demande à nos collègues s'ils ont des observations à présenter ou des questions à poser à M. Fréchet.

M. Roy se demande si le choix entre la médiane et la moyenne n'est pas lié à la connaissance de la loi de distribution de la grandeur étudiée. Par