Journal de la société française de statistique

Peter J. Rousseeuw

On Fréchet's upper bounds on the sampling variability of the median

Journal de la société française de statistique, tome 147, nº 2 (2006), p. 77-80

<http://www.numdam.org/item?id=JSFS_2006__147_2_77_0>

ON FRÉCHET'S UPPER BOUNDS ON THE SAMPLING VARIABILITY OF THE MEDIAN

Peter J. ROUSSEEUW *

1. General comments

In this little gem of a paper, Maurice Fréchet asked the question whether the sample median is really as inefficient at it seemed to be at the time. It is fascinating to read this work today, 66 years afterwards. Fréchet's insights were truly original and far ahead of their time. Most statisticians, including myself, did not know this prophetic paper existed. To the contrary, Fréchet's ideas and part of his results have been independently rediscovered by others many years later, as I will illustrate in the next section.

Several aspects of this paper appear modern by today's standards. Fréchet states clearly on page 68 that the predominant position of the arithmetic mean and the standard deviation are only justified under very narrow conditions that often do not hold in practice, and instead pleads in favor of the median and the interquartile range because of their simplicity and their robustness (long before the latter term was introduced). In the discussion of his paper, he replies to the comments by Roy in much the same way that a modern-day robust statistician would.

On page 69, Fréchet mentions that methods without theoretical justification may still be used in the form of 'méthodes empiriques de découverte,' 20 years before the phrase 'exploratory data analysis' (EDA) was coined by John Tukey.

And on pages 74–76, the mathematician Fréchet shows his interest in experiments by carrying out an ingenious simulation study, well organized so that it can be done completely by hand since computers did not yet exist. Based on only 96 card draws with replacement, he nevertheless obtains a convincing confirmation of his results, under a range of distributions with widely varying tails (indexed by the parameter α).

2. Fréchet's upper bounds on sampling variability

Fréchet's bounds are of two types. In the first type, he compares the variability of a location estimator $T_n = T_n(X_1, \ldots, X_n)$ with the variability of a

* Peter.Rousseeuw@ua.ac.be

single observation X where $X, X_1, \ldots, X_n$ are i.i.d. from the same univariate distribution F. His goal is to give an upper bound on the ratio

$$\lim_{n\to\infty} \frac{\sqrt{n}S(T_n)}{S(X)} \tag{1}$$

which holds for all distributions F satisfying some regularity condition. Here S is a scale functional of a random variable. By 'functional' it is meant that the argument of S is a whole distribution, not just a finite sample from it, and 'scale' refers to the requirement that always $S(X) \geqslant 0$ with the usual affine equivariance property $S(aX + b) = |a|S(X)$.

One of Fréchet's innovations was to find an upper bound for *any* distribution F, and not just compute (1) at a particular one like the normal. In most of his paper, T_n is the sample median. First he considers the classical scale functional $S(X) = Std(X)$, and shows that the upper bound on (1) is $\sqrt{3} \approx 1.73$ (his inequality $9bis$).

Another innovation of Fréchet's was not to restrict himself to the standard deviation but to also consider more robust scale functionals. The first of these was the *average deviation from the median* (ADM) given by $S(X) = E|X - Med(X)|$ (which is connected to the L^1 objective function which yields the sample median, in the same way that the standard deviation is connected to the L^2 objective yielding the mean). For this S the upper bound on (1) becomes $2\sqrt{\pi/2} \approx 1.60$ ($10bis$). Of course, the ADM is only a little more robust than the standard deviation (in modern parlance, it still has an unbounded influence function and a zero breakdown value). Probably Fréchet realized this, because he also added a much more robust scale functional, the interquartile range. For the latter, the upper bound of (1) becomes $2\Phi^{-1}(3/4) \approx 1.35$ ($11bis$). So, for increasingly robust S the upper bound on the variability of the median became smaller. All three of his bounds are sharp, since they are attained at the uniform distribution F.

Next, he asked what the upper bound on (1) is when T_n is the sample mean $\overline{X_n}$ instead of the median, and variability is still measured by the interquartile range. He found that the upper bound is infinite ($11ter$), by using an argument that resembles a breakdown reasoning. In other words, the variability of the sample mean can be infinitely large compared to that of the parent distribution, whereas for the median this ratio is bounded from above by 1.35. He concluded that, depending on the choice of S, the median can be much more accurate than the mean instead of the other way around.

The second type of bound in Fréchet's paper is about comparing the variability of T_n with the variability of another estimator U_n. That is, Fréchet wanted an upper bound on

$$\lim_{n\to\infty} \frac{S(T_n)}{S(U_n)} \tag{2}$$

for all F. Nowadays (2) looks more familiar than (1), because when S is the standard deviation, (2) becomes the square root of the asymptotic relative efficiency (ARE) of U_n relative to T_n. This is a more general notion, since

(2) still makes sense when T_n and U_n are something other than location estimators, whereas (1) doesn't. (For instance, T_n could be a scale estimator or an estimator of a regression slope, and then $S(X)$ is not relevant.)

Fréchet noted that when S is the standard deviation and U_n is the sample mean, $S(U_n) = S(X)/\sqrt{n}$ which makes (2) coincide with (1). Therefore, in this case (2) also has the upper bound $\sqrt{3}$, hence the ARE of the median relative to the mean is at least 1/3 at any distribution. So, the median is never totally inefficient. He also gives the double exponential distribution as an example of a distribution close to the normal where the ARE is bigger than 1 (in fact, it equals 2). We know that the ARE is even infinite at some distributions, like the Cauchy.

Fréchet's result that the ARE of the median relative to the mean is at least 1/3 at any distribution was independently rediscovered by Hodges and Lehmann (1956, page 327). To be precise, they stated this lower bound for the relative Pitman efficiency of the sign test relative to the t-test, which has the exact same expression. (Later on, Hodges and Lehmann (1963) formulated how rank tests lead to estimators with the same efficiency, with the sign test corresponding to the median.) Their proof of the bound 1/3 used calculus of variations, and was quite short. (Fréchet did not publish the proofs of his inequalities in his 1940 paper.) The Hodges-Lehmann estimator (1963), based on the Wilcoxon test, has a tighter upper bound $\sqrt{125/108} \approx 1.076$ on (2), hence the ARE of the Hodges-Lehmann estimator relative to the mean is at least $108/125 \approx 0.864$ at all F, which is better than for the median. Soon afterwards Bickel (1965) studied other robust location estimators in this fashion.

Fréchet's other insight, that the finite-sample variability of a robust estimator can be measured by a more robust scale functional, was rediscovered in 1994. It was known since the early nineties that high-breakdown robust regression estimators, like the least trimmed squares (LTS) method of Rousseeuw (1984), can have low finite-sample efficiency relative to least squares when the underlying distribution generates far-away good leverage points. (Note that the LTS plays a similar role as the sample median, which is a high-breakdown location estimator.) In a special issue of *Statistics and Probability Letters* titled 'Efficiency and Robustness,' several authors rediscovered the idea of measuring finite-sample efficiency by a robust scale functional S. For this Coakley et al. (1994) used the trimmed standard deviation, Adrover and Yohai (1994) used an M-functional of scale, and He (1994) chose the interquartile range, precisely the scale functional that Fréchet had employed. Like Fréchet, they found that the efficiency of robust methods is higher when measured through a robust scale functional S. But no Fréchet-style upper bound was derived, and I am not aware of his results (10*bis*) and (11*bis*) being rediscovered.

To conclude, Fréchet's paper foresaw developments that took place decades later, and his philosophy on estimation is now generally accepted in robust statistics.

Acknowledgment. I'd like to thank Henri Caussinus for inviting me to participate in this discussion.

References

ADROVER J.G., BIANCO A.M., and YOHAI V.J. (1994). Efficiency of MM- and τ-estimators for finite sample size. *Statistics and Probability Letters* 19, 409-415.

BICKEL P.J. (1965). On some robust estimates of location. *Annals of Mathematical Statistics* 36, 847-858.

COAKLEY C.W., MILI L., and CHENIAE M.G. (1994). Effect of leverage on the finite sample efficiencies of high breakdown estimators. *Statistics and Probability Letters* 19, 399-408.

HE X. (1994). Breakdown and efficiency – your perspective matters. *Statistics and Probability Letters* 19, 357-360.

HODGES J.L. and LEHMANN E.L. (1956). The efficiency of some nonparametric competitors of the t-test. *Annals of Mathematical Statistics* 27, 324-335.

HODGES J.L. and LEHMANN E.L. (1963). Estimates of location based on rank tests. *Annals of Mathematical Statistics* 34, 598-611.

ROUSSEEUW P.J. (1984). Least median of squares regression. *Journal of the American Statistical Association* 79, 871-880.