JOURNAL DE LA SOCIÉTÉ FRANÇAISE DE STATISTIQUE

ROGER KOENKER

The median is the message : toward the Fréchet median

Journal de la société française de statistique, tome 147, n° 2 (2006), p. 61-64

<http://www.numdam.org/item?id=JSFS_2006__147_2_61_0>

THE MEDIAN IS THE MESSAGE : TOWARD THE FRÉCHET MEDIAN

Roger KOENKER *

One could argue that the long-standing controversy over the relative merits of the median and mean is something of a touchstone of early statistics. The touchstone of medieval alchemy was basonite, a variety of quartz used to test the purity of gold alloys; striking an alloy against the surface of the stone revealed the quality of the alloy by the color of the mark left on the stone. Likewise, reading the marks left by Laplace, Edgeworth, Fisher, Fréchet, Kolmogorov, Tukey and Huber on the merits of the median all reveal a noble quality of mind, or perhaps instead just a mutual willingness to come to the aid of the statistically maligned and downtrodden. [1]

Fréchet's case for the median, reviewed in his reprinted 1940 article and prefigured in his earlier work, Fréchet (1924, 1935), was only part of a much wider campaign he waged beginning in the mid-1920's against the misuse of the coefficient of correlation. The idea that simplistic models of linear relationships for conditional means could capture all the subtleties of stochastic dependence seemed quite absurd, and the prevailing notion, particularly in the social sciences, that causality could be reduced to the

---

* University of Illinois at Urbana-Champaign. E-mail : rkoenker@uiuc.edu
This note was partially supported by NSF grant SES-05-44673, and prepared at the invitation of Professor Henri Caussinus on the occasion of the reprinting of Fréchet's 1940 paper "Sur une limitation très générale de la dispersion de la médiane" in the *Journal de la Société Française de Statistique*. My thanks to Marc Hallin, Ivan Mizera and Steve Stigler for help with references.

1. All felt compelled to emphasize the original observation of Laplace that the apparent advantage of the mean over the median in terms of asymptotic variance at the Gaussian distribution proves to be illusory as soon as one admits into consideration a broader class of error laws. Edgeworth (1887), in language that anticipates the modern interpretation of robustness as an insurance premium against inclement statistical weather, writes : "If we have been deceived by the appearance of Discordance [non-normality] . . . we shall have lost little by taking the Median instead of the Arithemetic Mean . . .. and if the observations are really discordant, the derangement due to the larger deviations will not be [as] serious [for the median] as it is for the Arithemetic Mean." Fisher's (1922) seminal paper on maximum likelihood observes that the sample mean from a Cauchy population has the same precision as any single observation and recommends the median as an alternative. Kolmogorov's (1931) first statistical paper notes the advantage of the mean over the median for the "normal law," but asserts that for unimodal error laws the ratio of asymptotic standard deviations of the median and the mean is bounded above by $\sqrt{3}$, at the uniform, but can be as small as zero, favoring the median. Tukey's proposals of median polish for anova models and running median smoothers in time-series analysis were highly influential in expanding the rôle of the median and emphasizing the importance of robustness in statistics. Huber's (1981) observation that the median achieves the smallest maximum bias among all translation equivariant estimators of location constituted a rare admission that without symmetry much of the elegant theory of robust estimation was built on sand.

examination of a few correlation coefficients was an affront against common sense. The history of this campaign has been ably recounted by Armatte (2001) so I will resist the temptation to dwell on it, but I would like to underscore that it was a campaign fought on several fronts. On the axiomatic front, Fréchet, as pure mathematician, defended the Cartesian homeland rigorously. But there was also, perhaps more surprisingly, a strong data-analytic front, emphasizing the importance of sound statistical practice for science and public policy.

One early skirmish on this data-analytic front may be taken as emblematic of Fréchet's viewpoint. In 1923, while still in Strasbourg, Fréchet supervised a diploma thesis by Samana that re-analyzed the data from experiments reported in Peirce (1873). C.S. Peirce was an American polymath with contributions to logic, algebra, geometry as well as probability. At the time of these experiments he was employed by the U.S. Coast Survey; his experiments were designed to explore the measurement error in astronomical observations due to variability of observer reaction times to visual and auditory stimuli. [2] Peirce hired an untrained young man to react to a sharp sound "like a rap, the answer being made upon a telegraph operator's key nicely adjusted." Reaction times were recorded in thousandths of a second employing a Hipp chronoscope described in loving detail in Peirce's report. On each of 24 consecutive week days in July of 1872, 500 measurements were made. After some innovative kernel smoothing, Peirce concluded that if one ignored the first two or three days of "training" the estimated densities differed very little from the Gaussian law.

More than 50 years later, Fréchet (1924) reported that his student's analysis suggested that Peirce's data were better represented by Laplace's first law, $\varphi(\epsilon) = Ce^{-|R\epsilon|}$ than by Laplace's second law, $\varphi(\epsilon) = Ce^{-R^2\epsilon^2}$, that is by the Gaussian distribution. No details are given, unfortunately, but Fréchet's remark attracted the attention of E.B. Wilson and Margaret Hilferty, who undertook in 1929 another reanalysis of Peirce's data. [3] They note, again deferring to Laplace,

> The ordinary statement based on the normal law is that the determination of the median is 25% worse than that of the mean. A comparison of the standard deviations of the median and mean in columns (1) and (2) shows that for these observations the median is better determined than the mean on 13 days, worse determined on 9 days, and equally well determined on 2 days. Roughly speaking this means that mean and median are on the whole about equally well determined.

2. Stigler (1978) identifies these experiments as among the most significant statistical investigations conducted in the United States in the 19th century.

3. It may be regarded as a shocking lapse in American xenophobia to find these authors acknowledging a paper written in French and published in Moscow, but perhaps no less surprising that Fréchet himself disinterring the work of Peirce for a leading Soviet mathematics journal.

Wilson and Hilferty's column (1) reports the daily medians, and column (2) reports estimates of the "standard deviation of the median" that are then compared to the more well-established standard deviations of the means given in column (4). We might well ask what *was* this "standard deviation of the median ?"

An informal survey of textbooks of this period suggests that the recommended method of computing the precision of the median relied on adherence to the assumptions of the normal model. But this is clearly not what was done by Wilson and Hilferty since it would have produced values that were consistently larger than the standard deviation of the mean by precisely 25 percent. Already by 1917, Yule's very influential text, recommended the normal theory approach for small samples, while suggesting an alternative approach based on the frequency of the grouped data bin count at the median when the sample size was larger. However, for the Peirce data the median bin counts are quite small, ranging from 4 to 12, thus rendering Yule's implicit bandwidth selection for estimation of the density at the median too small to be reliable. Curiously, modern bandwidth selection based on work of Hall and Sheather (1988) agrees quite closely with the results reported by Wilson and Hilferty so it remains a puzzle exactly what procedure was employed.

Wilson and Hilferty conclude,

> The upshot of this all is that Peirce had observations which could show as completely as one might desire that the departures of the errors from the normal law was for his series uniformly great.

Thus, the conclusions of Peirce were contradicted – even under Peirce's carefully controlled conditions it appeared that Gaussian assumptions about the distribution of reaction times were questionable and the usual justification of the mean's putative advantages were cast into doubt. In more complicated settings it would be difficult to argue that Gaussian assumptions become *more* plausible, so Fréchet's argument that the median is more prudent seems entirely justified.

The challenge of course is to find compelling analogues of the median for more complex statistical models. This has been a slow process. The early proposals of Boscovich, as modified by Edgeworth, have gradually evolved into an effective strategy for estimation and inference in regression. However, in multivariate analysis there are several competing notions of "median" and no prospect of a reconciliation anytime soon.

The relatively new domain of functional data analysis offers special challenges ; "statistics on manifolds" is a critical aspect of the rapidly developing field of image analysis. Ironically, the Fréchet imprimatur has been appropriated within this domain as a seal of approval for the mean. Following Bhattacharya and Patrangenaru (2003), the *Fréchet mean* of a probability measure Q on a metric space M with metric ρ is any minimizer of,

$$F_2(p) = \int \rho^2(p, x) Q(dx) \qquad p \in M.$$

In some simple cases this is quite straightforward. For example, if we take M to be a linear subspace, then the Fréchet mean is the orthogonal least squares regression estimator. So, it is natural to ask in view of Fréchet's earlier work : why not the Fréchet median? Why do we square the distances? Consider, instead the minimization of,

$$F_1(p) = \int \rho(p, x)Q(dx) \qquad p \in M.$$

In the linear subspace example, this yields an orthogonal median regression estimator that is computationally quite tractable. Extending this approach to more general manifolds would be well worthwhile. Having defined median shape in this manner, it is obviously tempting to consider Fréchet quantiles in much the same manner. I hope that steps can be taken in this direction in future work.

References

ARMATTE M. (2001). «Maurice Fréchet Statisticien, Enquêteur et Agitateur Public», *Revue d'Histoire des Mathématiques*, 7, 7-65.

BHATTACHARYA R., and PATRANGENARU (2003). "Large Sample Theory of intrinsic and Extrinsic Sample Means on Manifolds," *Annals of Statistics*, 31, 1-29.

EDGEWORTH F. Y. (1887). "The Choice of Means," *Philosophical Magazine*, 24, 268-271.

FISHER R. A. (1922). " On the Mathematical Foundations of Theoretical Statistics," *Philosophical Transactions of the Royal Society of London (A)*, 222, 309-368.

FRÉCHET M. (1924). «Sur la loi des erreurs d'observations», *Matematichiskii Sbornik*, 32, 5-8.

FRÉCHET M. (1935). «Sur la précision comparées de la moyenne et de la médiane», *Aktuárské Vědy*, 5, 29-34.

HALL P., and SHEATHER S. (1988). "On the distribution of a studentized quantile", *Journal of the Royal Statistical Society (B)*, 50, 381-391.

HUBER P. J. (1981). *Robust Statistics.* Wiley : New York.

KOLMOGOROV A. N. (1931). "The Method of the Median in the Theory of Errors", *Matermatichiskii Sbornik*, 38, 47-50, Reprinted in English in *Selected Works of A.N. Kolmogorov*, vol. II, A.N. Shiryayev, (ed), Kluwer : Dordrecht.

PEIRCE C. S. (1873). "On the Theory of Errors of Observation," *Report of the Superintendent of the U.S. Coast Survey*, pp. 200-224, Reprinted in *The New Elements of Mathematics*, (1976) collected papers of C.S Peirce, ed. by C. Eisele, Humanities Press : Atlantic Highlands, N.J., vol. 3, part 1, 639-676.

STIGLER S. M. (1978). "Mathematical Statistics in the Early States," *Annals of Statistics*, 6, 239-265.

WILSON E. B., and M. M. HILFERTY. "Note on C.S. Peirce's Experimental Discussion of the Law of Errors," *Proceedings of the National Academy of Sciences of the U.S.A.*, 15, 120-125.

YULE G. U. (1917). *An Introduction to the Theory of Statistics.* Charles Griffen : London, 4 edn.