Visualizing the Median as the Minimum-Deviation Location

Author(s): James A. Hanley, Lawrence Joseph, Robert W. Platt, Moo K. Chung and Patrick Belisle

Source: *The American Statistician*, May, 2001, Vol. 55, No. 2 (May, 2001), pp. 150-152

Published by: Taylor & Francis, Ltd. on behalf of the American Statistical Association

Stable URL: https://www.jstor.org/stable/2686003

Visualizing the Median as the Minimum-Deviation Location

James A. HANLEY, Lawrence JOSEPH, Robert W. PLATT, Moo K. CHUNG, and Patrick BÉLISLE

The calculus proofs of this property of the median presented in mathematical statistics texts are not very instructive. A noncalculus proof has been published, but is still somewhat lengthy and likely to deter some nonmathematical readers. To make the proof more memorable, teachers can make the minimization task more real by using a concrete criterion such as total distance traveled, rather than simply an abstract sum of absolute deviations. We suggest a short, simple, heuristic graphical approach, which we illustrate using a Java applet. We pose, and give proofs for, a related optimization problem.

KEY WORDS: Applet; Graphics; Noncalculus proof; Optimization.

1. INTRODUCTION

If asked where they would stand and wait for the next of three elevators, unequally spaced along a wall, many students would choose to stand at the mean position. They think that by doing so they are minimizing the average distance to the elevator. They do not recognize that standing at the mean minimizes the average *squared* distance and that the minimal average distance to the elevator is in fact achieved by standing at the median (Hanley and Lippman 1999).

How to bring students to really understand and remember this property of the median? Having them formally prove it by calculus, as in Cramér (1946), and as many older generations of teachers had to do in graduate school, is only practical for those who can manipulate integral calculus. And even then, the exercise is not very instructive. The noncalculus proof of Schwertman, Gilks, and Cameron (1990) is more instructive, but is still somewhat lengthy and likely to deter many nonmathematical readers. The Java applet of Lane (2000) focuses more on the mean and on the smaller mean squared deviation from the mean than from the median. Moreover, it presents the calculations in a table, thereby separating them visually from the data points in the diagram.

We recently asked statistics students in a linear models class to prove the minimum absolute deviation property of the median for $n = 3$ unequally spaced elevators. One student (MKC)

provided a simple proof that immediately extends to any n and thus to the median of the distribution of any continuous random variable. His solution uses the same ideas as Schwertman et al. (1990), but allows for a shorter, simpler, and more heuristic graphical approach. We used it to construct an applet, which we now describe. The applet is available at http://www.epi.mcgill.ca/hanley/elevator.html.

2. VISUALIZING THE MEDIAN

Like Schwertman et al. (1990), one need only focus on the total, rather than the average, distance to the elevators. However, one can entirely eliminate their algebra, and the numerical calculations of Lane (2000), by representing the total distance physically—as the total lengths of n lines (see Figure 1, a screenshot from the applet). Using the mouse to click on various waiting positions, one will quickly see that the total distance to the elevators from any position in the *innermost interval* (or at the *innermost point* if n is odd) is the sum of the $\lceil n/2 \rceil$ pairwise "interval-widths," shown in different colors; the total distance is greater if one stands *elsewhere*.

The proof can be stated in words as follows. First, consider waiting at a median location (a unique point if n is odd, a point in the middlemost interval if n is even). Then, consider "moving" to a waiting position away from (outside) this median point (interval). If one does so, one moves away from more elevators than one moves towards, thereby increasing the sum of the distances to the elevators. These additional distances "from the median to the new location and back" constitute the second term in the key relation used in the integral calculus proof sketched by Cramér (1946, pp. 178–179; see Appendix).

Teachers can make the proof more memorable in a number of ways. First, we suggest they work "up" from 2, rather than "down" from n as Schwertman et al. (1990) did. Second, they might make the minimization task more real by using a concrete criterion—such as total distance traveled—rather than simply an abstract sum of absolute deviations. Third, they might avoid all algebra and formal numerical calculations by employing entirely graphical techniques.

Waiting at the median may not optimize other functions of the n distances. Concern about "missing" the elevator implies other criteria, such as the maximum distance, or the probability of being within a certain distance of an elevator—the same issues faced by planners deciding where to locate an ambulance or fire station in order to optimize rapidity of individual responses. To emphasize a function that *sums* the n distances, an example involving *average* transportation/fuel costs *over many repetitions*, rather than rapidity of response in *individual* instances, might be better.

3. A RELATED OPTIMIZATION PROBLEM

A more realistic and more challenging example of optimizing an aggregate or average criterion can be posed as follows. If one's job is to carry heavy objects to elevators, one cannot ignore the

James A. Hanley, Lawrence Joseph, and Robert W. Platt are Professors in the Department of Epidemiology and Biostatistics, 1020 Pine Avenue West, McGill University, Montreal, Quebec, Canada H3A 1A2 (E-mail: James.Hanley@McGill.CA; joseph@binky.ri.mgh.mcgill.ca; rplatt@po-box.mcgill.ca). Moo K. Chung is a Graduate Student in the Department of Mathematics and Statistics, McGill University, Montreal, Quebec, Canada H3A 1A2 (E-mail: chung@math.mcgill.ca). Patrick Bélisle is a Research Associate in the Division of Clinical Epidemiology, Montreal General Hospital, Montreal, Quebec, Canada H3A 1A2 (E-mail: belisle@hirondeau.ri.mgh.mcgill.ca). This work was partially supported by grants from the Natural Sciences and Engineering Research Council of Canada.
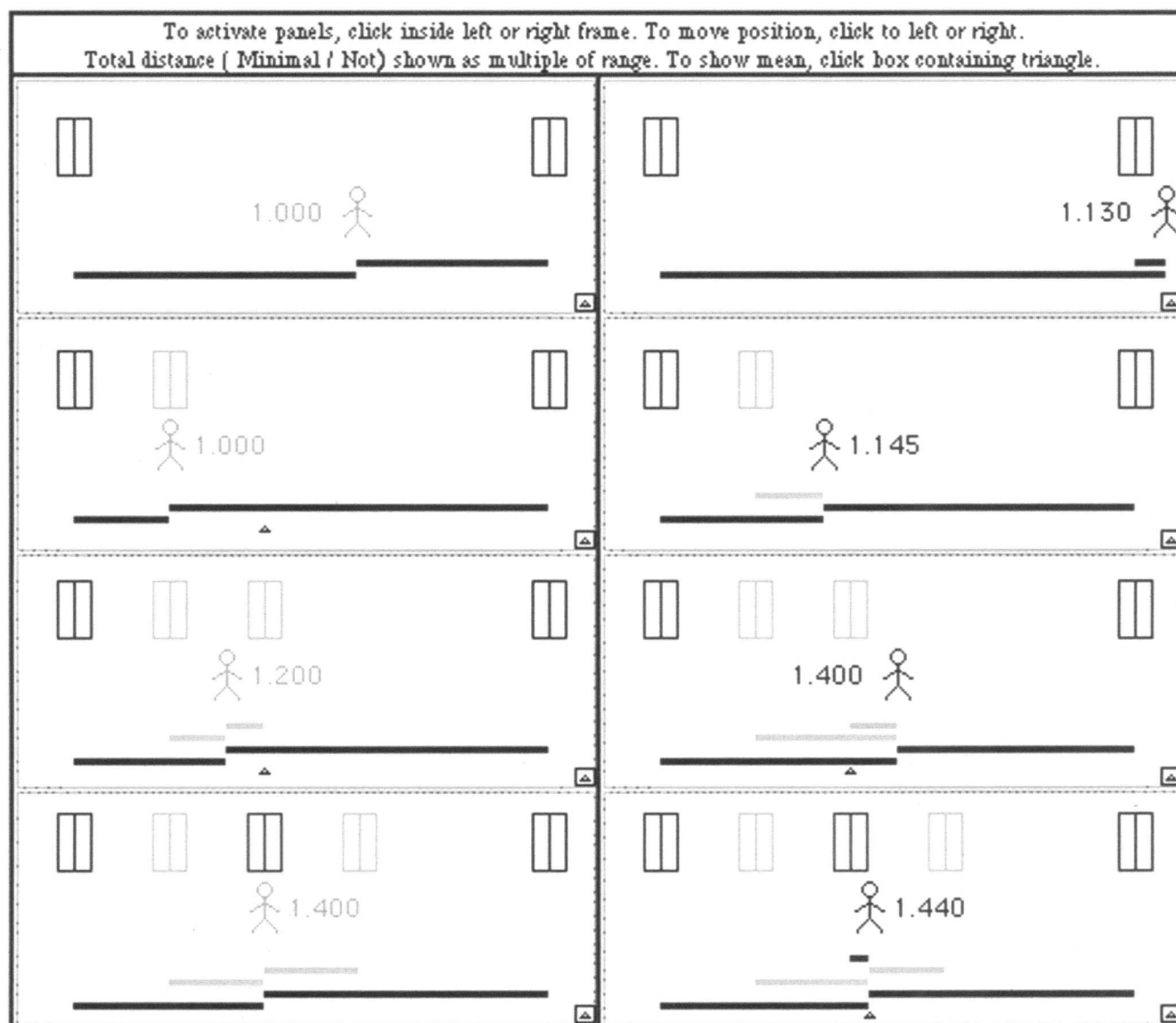
Figure 1. Total distances walked to n (=2, 3, 4, or 5) elevators, from median (leftmost panels) and nonmedian (rightmost panels) locations. For even n, the total distance is the same from all locations between the innermost two elevators; the total distance is greater from any position outside of them; for odd n, there is only one median position which minimizes the total distance.

distance from the initial point of arrival into the elevator area to the optimal place to wait. In our one-dimensional model, the point of arrival can either be to the left or right of, or somewhere in between, the outermost elevators. Where should one stand in order to minimize the total (or average) distance, which we now redefine to include the additional distance from the point of entry to the point where one will wait?

Surprisingly perhaps, the answer is to remain at the point of entry, and to not move at all! This can be seen by reasoning as follows, using first the case of an entry point that is between the leftmost and median elevators. Waiting just an epsilon (ε) to the right of this entry point would not change the distance traveled to reach each elevator to the right (some immediately, the remainder when the elevator arrives), but it would add 2ε to the distance traveled to each elevator to the left of the entry point. Waiting even further from the entry point only increases the wasted initial and subsequent travel in a linear way. Thus,

to minimize the total "realistic" distance, one should stay put! In the case of an entry point that is to the left of the leftmost elevator, waiting anywhere between it and this elevator does not add unnecessary travel, but waiting even an ε to the right of this elevator does. A symmetric argument applies to entering the right side of the room.

Another "proof" follows a typical mathematical strategy: convert the problem into an equivalent one for which there is already a solution. The n trips from the entry point to the n different elevators involve $2n$ distances, n from the entry point to the waiting point, and n from the waiting point to the elevators. But the distance from the entry point to the waiting point is the same as the distance from the waiting point to the entry point. Given this symmetry, the optimal waiting point is the median of the $2n$ locations: the n (identical) entry points and the n elevator locations. If the entry point is between the left and rightmost elevators, the median is this entry point. If the entry point is say

to the left of the first elevator, the median is anywhere between the entry point and the first elevator.

4. CONCLUSION

Most mathematical statistics students prove this property of the median as an exercise at some stage in their training, but soon forget it. Thus, the long-term impact of the exercise is less than it could be (someone once defined education as "what remains after one has forgotten what one has learned"). Later, many of them, and many nonstatistical students too, would, if asked, argue that the average distance is minimized by the mean. We suggest that it is time to "move up" from the proofs in mathematical statistics texts to more instructive ones which, using concrete examples, allow one to show visually what makes the median such a central location.

APPENDIX

Cramér (1946, pp. 178–179) denoted the random variable by ξ, the median (or in the indeterminate case, any median value) by μ, and any proposed "central" location by c. For the case of $c > \mu$, he used the relation

$$\mathbf{E}(|\xi - c|) = \mathbf{E}(|\xi - \mu|) + 2 \int_{\mu}^{c} (c - x) dF(x),$$

and the fact that the second term on the right-hand side is "evidently positive," to show that the first absolute moment $\mathbf{E}(|\xi-c|)$ becomes a minimum when $c = \mu$.

As did Cramér, we leave the proof of the above relation as an exercise for the reader.

In our examples, ξ is a random variable with probability mass at a finite number (n) of values. When n is even, the factor of 2 in the second term above can be seen clearly from Figure 1. When n is odd, however, it appears from Figure 1 that the "extra" distance from the location c to $\xi = \mu$ is $(c - \mu)$—rather than the $2(c - \mu)$ in the relation. In this case, one way to "see" the 2 is to convert the probability mass of $1/n$ at $\xi = \mu$ into two masses of $1/2n$ each, and treat the case as one with an even number $n' = n + 1$ of ξ values.

REFERENCES

Cramér, H. (1946), *Mathematical Methods of Statistics*, Princeton, NJ: Princeton University Press.

Hanley, J. A., and Lippman, A. (1999), "Where Do You Stand? Notions of the Statistical 'Centre'," *Teaching Statistics*, 21, 49–51.

Lane, D. M. (2000), Rice Virtual Lab in Statistics, Copyright 1993–2000, http://www.ruf.rice.edu/~lane/rvls.html.

Schwertman, N. C., Gilks, A. J., and Cameron, J. (1990), "A Simple Noncalculus Proof That the Median Minimizes the Sum of the Absolute Deviations," *The American Statistician*, 44, 38–39.