

Ask the Expert



Edward George is the Universal Furniture Professor in the Department of Statistics at the Wharton School, University of Pennsylvania. He obtained his Ph.D. in Statistics from Stanford University in 1981, after which he held faculty positions at the University of Chicago and the University of Texas at Austin. George is an elected Fellow of the American Statistical Association and of the Institute for Mathematical Statistics, and an elected member of the International Statistical Institute. His current research interests include Bayesian analysis, classification and regression tree modeling, model uncertainty, predictive inference, statistical decision theory, and variable selection.

On July 7, 2008, Ed George answered selected viewer questions concerning the "wisdom-of-the-crowds" concept and how it applies to the stock market and other entities. Please note we are no longer accepting questions, but see [Counting Cabs](#) and [One-Minute Experiment](#) as well as our [Links & Books](#) section for more information.

Q: What's the difference between a mean and a median?

Walt Long, Poquoson, Virginia

A: Hi, Walt. The mean of a set of values is the average value, obtained by dividing the sum of the values by the total number of values. The median of a set of values is a middle value, which lies in between the largest 50 percent and the smallest 50 percent of the values in the set. An attractive feature of the median is that it is less sensitive to extreme values than the mean. For example, if a few (fewer than half) of the largest values in the set were made larger, the mean would increase too, but the median would not. In the "ox-weight" example [described in the NOVA scienceNOW "Wisdom of the Crowds" filmed segment], the median guess was less sensitive to a few very poor guesses.

Q: If I were to ask one kindergartener how much money he thinks the President makes, I don't suspect his answer would be any "wiser" than the answer I would get from averaging figures from 100 kindergarteners. So what qualifications should a crowd meet before it can contribute to crowd wisdom?

Eric Taylor, Tupelo, Mississippi

A: Hi, Eric. Good insight. What surprises us in the ox-weight example is that there was virtually no discrepancy between the average guess and the true weight. Statistics served to uncover the "wisdom," which was hidden within the 800 varied guesses. When there is such information to be had, statistical tools are the key to finding it. However, statistical methods will only work when there is something there to be uncovered.

Although kindergartener guesses of the President's salary would probably vary widely, averaging them would tend to reduce the effect of this variability; for example, very large and very small estimates would cancel each other out. Thus, the average guess would probably be better (i.e., closer to the actual salary of \$400,000) than many of the individual guesses. However, even though the average may be better than many of the individual guesses, there is no guarantee that it will be good.

The discrepancy between the average guess and the actual value may be understood as coming from two main sources: 1) excessive variation of the guesses, perhaps due to a complete lack of information, and 2) systematic bias, perhaps due to a general tendency for the children to overestimate the salary (a tendency I suspect is shared by many adults). If I were to speculate about the wisdom of any particular crowd, it is these two sources of discrepancy that I would use for my assessment.

Q: In which kinds of scenarios would a crowd's averaged guess be less useful than the guesses of one or a couple of experts?

Elliot Carson, Phoenix, Arizona

A: Hi, Elliot. I'm glad to see that you are not automatically assuming the superiority of a crowd's averaged guess. Although statistical tools can seem nearly magical in their ability to extract information, they can only do so when there is information there to be extracted. A scenario in which an expert or two would provide better answers might occur when any and all relevant knowledge about a question is simply not available to crowd members. For example, in narrow technical domains in which deep scientific background is needed to even understand phenomena, the members of a crowd might simply not have a clue on which to base their guess.

Q: I have tried to formulate my understanding of crowd wisdom in a "septoe"—a seven-word compression of wisdom: "Wisdom of crowds lies in their diversity." Okay, Ed, to what extent is this true?

Peter Gluck, Romania

A: Hi, Peter. Well done—a very nice and appropriate septoe. Crowd wisdom is enhanced when the individual guesses are derived from diverse sources of information. This diversity tends to create a balance of high and low guesses, which when cancelled out through averaging tends to leave us with the correct answer.

To see what might have happened with a lack of diversity, suppose the people who were guessing the ox's weight had all been influenced by a single farmer who had advised them to give a low estimate. The dependence caused by heeding the farmer's advice, a single source of information, would have biased their average guess to be too low.

Q: There is no reason that, say, the collective average of people's guesses about how many jelly beans are in a jar should converge to the *actual* number of jelly beans in the jar. Nonetheless, it seems that the answer does push toward the actual number. What statistical model could possibly intuitively explain this phenomenon?

Maaneli Derakhshani, New York, New York

A: Hi, Maaneli. I very much like your question, because it gives me an opportunity to convey the way I would use a model to think about the phenomenon. I would treat each guess as $x = \text{total} + \text{error}$, where total is the actual (unknown) number of jelly beans in the jar, and the errors are independent draws from a hypothetical population of measurement errors for this problem. An assumption that this measurement error population is centered at zero (and of bounded variation) is then all that is needed to assure that the average guess would converge to the total as the sample size increases. However, if there was systematic bias, such as a tendency for people to overestimate more often than underestimate, then the average guess would converge to a wrong answer.

Q: Let's talk game shows. Why is the correct answer to the "Monty Hall problem" for a contestant to select a new door (for example, door No. 1) after a third door has been eliminated? I would think the odds are one in two that you picked the right door no matter which one you selected to begin with.

John Lemerond, Wauwatosa, Wisconsin

A: Hi, John. The Monty Hall problem, for readers who may not be familiar with it, goes as follows: "Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to pick door No. 2?" Is it to your advantage to switch your choice?" (Parade Magazine, September 1990).

Many people initially reason that with only two doors left, the probability should be 1/2 that the car is behind No. 1. However, if you assume that the host knew where the car was, and was obliged to open a door with a goat, then the probability is actually 1/3 that the car is behind No. 1. To see this, note that at the outset, the probability is 1/3 that the car is behind No. 1. Since at least one of the other two doors has a goat, the host has revealed nothing about the contents of No. 1. Thus the probability of 1/3 for door No. 1 remains unchanged.

This is an interesting example in which crowd wisdom would probably be bad. Most people are led astray by their intuition here, especially when first confronted with the problem. Groups of such people would obviously have little to offer in terms of crowd wisdom on the correct solution.

Q: Have you ever plotted the answers to any question on a bell curve with the correct answer being at the apex of the curve?

Anonymous

A: If it were reasonable to assume that the answers were a sample from a normal distribution, a useful statistical approach would be to fit the familiar bell-shaped curve to the data. The apex of this curve would be an estimate of the correct answer, here the true mean. The estimate is often good but very rarely exactly correct. I can never remember it happening for me. Getting the exactly correct answer in the ox-weight example was very unusual.

Q: Does this statistical wisdom of the crowd about the [number of cabs](#) in New York City also apply to matters of institutional policy? If I ask a lot of board members, stakeholders, clients, and colleagues about a future policy, will I get a good answer?

Frits Simon, The Netherlands

A: Hi, Frits. The wisdom of a crowd is often surprisingly good, but it is by no means foolproof. Because the various constituencies you mention would have been exposed to broad sources of information concerning institutional policy, their collective wisdom should not be ignored. My advice would be to incentivize each of them to do their best to provide honest and thoughtful inputs, and I

would advise you to be on guard for potential systematic biases that may be masked by a consensus.

Q: When I guessed the [number of cabs](#) in New York, I was fascinated by the results: New York City residents' guesses were far worse on average than the non-residents' guesses. Why do you think this happened?

Will Thill, Los Angeles, California

A: Hi, Will. It turns out that at least one person submitted a large number of very large unrealistic guesses. Indeed, 1,883 of the first 3,189 guesses were the identical value 2,147,483,647 and were listed as coming from all three regions, NY, U.S., and other. These entries, which severely contaminated both the means and the medians, were obviously not honest guesses but rather an attempt to scuttle the online experiment.

As now indicated on the site, in an attempt to fix this problem, the NOVA team zeroed out the initial results and began accepting guesses only within range 0 to 100,000. Although this worked for awhile and some amazingly good crowd estimates were obtained, the dishonest player or players soon returned to again begin entering many very large numbers.

Although such dishonest behavior by a few is somewhat discouraging, there is an interesting lesson in what happened. In the ox-weight example, all the participants were given a strong incentive to provide honest guesses, namely a correct answer would win the ox. Crowd wisdom can best flourish when individual participants are so encouraged to do their best.

[Editor's note: Oops. When we zeroed it out on June 27, 2008, we accidentally set the limit at 1,000,000. On July 7, we set the limit at 100,000 and zeroed it out again.]

Q: What role does "spin" play in statistics? That is, can the wording of questions affect the ultimate guess a crowd will make? How do you avoid such spin?

Bruce Klutchno, New York, New York

A: Hi, Bruce. Terrific point. The way a question is asked, sometimes called framing, can have a strong influence on how it is answered. Such framing is sometimes used to manipulate answers in a desired direction, for example towards a particular partisan perspective in a political poll. Ideally, one might hope to simply frame questions to be devoid of spin so as to get at the "truth". However, when some spin cannot be avoided, one should at least ask a question in several ways, making the respondent aware of the potential biases. This would at least keep the biasing effect of spin from being hidden.

Q: How is the "wisdom of crowds" gleaned from the crowd participation distorted when some individuals are able to influence the average estimation more than others in the crowd, say with insider trading?

Roger Rines, San Jose, California

A: Hi, Roger. Great insight. Although the effect of a few very bad guesses could be mitigated by using the median rather than the mean, as in the ox-weight example, a few individuals could have a more pernicious effect by contaminating the opinions of others, for instance, by feeding them bad or misleading information. And even if the information they were feeding was good rather than bad, the crowd wisdom would serve mainly to amplify the views of these few rather than the combined opinions of many. Such is probably the case with insider trading, which, though unfair to honest market participants, is based on good information that is publicly unavailable. How to mitigate or even identify such effects can be a very difficult problem.

Q: I am an attorney who believes that the "crowd," meaning voters in partisan elections, would do a better job of selecting judges than any governor and/or small groups of appointed lawyers can do in non-partisan appointments. What do you think?

Fred Gamin, Raleigh, North Carolina

A: Hi, Fred. Your excellent question reminds us that the hope of democracy is founded on the very notion of crowd wisdom. Indeed, in the NOVA piece, Galton was interested in using the ox-weight example to show the futility of democracy. His ploy backfired when it turned out that the wisdom of the crowd was actually right on the mark, thereby giving cause for optimism.

Although the combined wisdom of many is not infallible, as I've elaborated in several of my other answers, it should not be overlooked. Turning to the question of electing versus appointing judges, my inclination is to prefer the voters for a variety of reasons, including the potential of their combined wisdom. However, I can also imagine circumstances in which it might be more prudent to leave appointment in the hands of a few elected officials.

Q: Can the wisdom of the crowds answer questions on the nature of dark matter and the likelihood of the existence of multiple dimensions?

John Williams, Lincoln, Nebraska

A: Hi, John. I don't think a crowd of non-specialists would make much progress with these deep questions, although occasional fresh new insights are always a possibility. However, when scientists and mathematicians who study such problems share their thoughts and findings, wonderful synergies occur. This might be thought of as crowd wisdom in the sense that the whole is greater than the sum of the parts.

Q: Is the average level of intelligence going up or down with the increase in the U.S. population?

Marc Lietermann, Little Chute, Wisconsin

A: Hi, Marc. A fun question. I suspect you are reasoning that if crowd wisdom increases with the size of the crowd, then there will be even more wisdom to be had as our population increases. To some extent this is true, although one wouldn't consider this as an increase in the level of intelligence. More to the point, how to best harness or extract such wisdom may not be easy. If one were to take an averaging approach such as in the ox-weight example, it would probably be necessary to somehow cluster individuals into groups according to their relevant knowledge. ■

From <http://www-tc.pbs.org/nova>