

Die

ZENTRUM FÜR METHODEN DER SOZIALWISSENSCHAFTEN

"Weisheit der Masse"

Der Schätzquiz "Wieviel Schätzungen werden heute abgegeben?" ist ein Experiment, um das Phänomen "Weisheit der Masse" vorzustellen.

Das Phänomen ist, dass die Meinungen vieler **Laien zusammengefasst oft besser** sind als die Meinung eines **Experten**. In der einfachsten Version tritt es beim Schätzen von Größen auf.

Durch das Buch "The Wisdom of Crowds" ist die Geschichte von Francis Galton und dem Ochsen-Gewicht-Schätz-Wettbewerb der westenglische Nutztiermesse von 1906 sehr populär.

Die volle Geschichte enthält auch einen Diskurs über **Arten der Mittelbildung**, die in der Kurzversion unterschlagen wird.

Francis Galton argumentierte, dass der **Median** die geeignete Schätzungen sei, nicht so anfällig für stark falsche Schätzungen.¹

Als empirischen Test wertete er 787 Schätzettel des Ochsen-Gewicht-Schätz-Wettbewerb der westenglische Nutztiermesse von 1906 aus und fand eine beeindruckende Weisheit der Masse.²

Hooker rechnete nach, dass hier der Mittelwert noch besser sei³, was Galton bestätigte.⁴

Ob Median oder Mittelwert und warum, ist eine weiterhin interessante Frage.

Jan Lorenz, www.janlo.de

Sir Francis Galton (1822-1911) war ein britischer Statistiker und Naturforscher. Er war ein Halbcolvin von Charles Darwin. Er ist Mitbegründer der modernen Statistik und hat dort sehr wertvolle Beiträge geliefert. Er ist aber auch der Schöpfer der Eugenik, der Idee der Selbststeuerung der menschlichen Evolution. Die Nationalsozialisten radikalisierten die Eugenik zur "Rassenhygiene", welche Massenmorde an "unwertem Leben" rechtfertigte. Galtons sagte "Die Möglichkeit der rassistischen Verbesserung einer Nation hängt von deren Fähigkeit ab, die Produktivität des besten Erbgutes zu erhöhen. Dies ist weitaus wichtiger als die Unterdrückung der Produktion der Schlechtesten." Er war von der Vererbung von

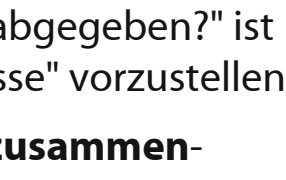


Intelligenz und Talent überzeugt und wollte "die menschliche Rasse verbessern". Um so mehr überrascht, dass er eine Untersuchung publizierte, die zeigt, dass der Durchschnitt der Masse besser sein kann als ein

Der **Median** ist der "mittlere" Wert. Man ordnet die Werte der Größe nach und nimmt den mittlersten aus dieser Liste. D.h. der Median ist so gewählt, dass 50% der Werte kleiner und 50% der Werte größer sind.

Der **Mittelwert** ist die Summe aller Werte geteilt durch die Anzahl der Werte. Im allgemeinen Sprachgebrauch der "Durchschnitt".

4 nature, March 28, 1907
 Mr. HOOPER, in NATURE of March 21, seems not to have quite appreciated my principal contention in the letters "One Vote, One Value" and "Vox Populi" of February 28 and March 7 respectively. It was to show that the verdict given by the ballot-box must be the Median estimate, because every estimate is condemned in advance by a majority of the voters. This being the case, I examined the votes in a particular instance according to the most appropriate method for dealing with medians, quartiles, &c. I had no intention of trespassing into the technical and much-discussed question of the relative merits of the Median and of the several kinds of Mean, and beg to be excused from not doing so now except in so far as it may be of use to you. It is only a small matter to realise that the suppression of any one value in a series can only make the difference of one half-value to the median, whereas if the series be small it may make a great difference to the mean; consequently, I think my proposal that juries should openly adopt the median when estimating damages, and councils when estimating money grants, has independent merits of its own, besides being in strict accordance with the true theory of the ballot-box. Secondly, Mr. Hooper's approximate calculation from my scanty list of figures, of what the mean would be of all the figures, proves to be singularly correct: it makes it 1150 lb., which is the mean of the deviates at 5°, 15°, 25°, whereas it should have been 1107 lb. This shows well that a small orderly sample is as useful for calculating means as a very much larger random sample, and that the compactness of a table of centiles is no hindrance to their wider use. I regret to be unable to learn the proportion of the competitors who were farmers, butchers, or non-experts. It would be well if future competitions to have a line on the cards for "occupation." Certainly many non-experts compiled, like those clerks and others who have an expert knowledge of horses, but who let on races, guided by newspapers, friends, and their own fancies. FRANCIS GALTON.



THE WISDOM OF CROWDS
 Why We Make Decisions That Are Better Than We Think We Do
 JAMES SUROWIECKI

2 nature, March 7, 1907
 FOX POPULI
 IN these democratic days, an investigation into the trustworthiness and peculiarities of popular judgments is a matter about as much discussed as the matter of a small matter, but is much to the point. A weight-judging competition was carried on at the annual show of the West of England Fat Stock and Poultry Exhibition recently held at Plymouth. A fat ox having been selected, competitors brought stamped and numbered cards, for 6d. each, on which to inscribe their respective names, addresses, and estimates of what the ox would weigh after it had been slaughtered and "dressed." Those who guessed most successfully received prizes. About 800 tickets were issued, which were kindly lent me for examination after they had fulfilled their immediate purpose. These afforded excellent material for a study of the human mind, and of the influence of vanity and the like. The sixpenny fee deterred practical joking, and the hope of a prize and the joy of competition prompted each competitor to do his best. The competitors included butchers and farmers, some of whom were highly expert in judging the weight of cattle; others were probably guided by such information as they might pick up, and by their own fancies. The average competitor was probably as well fitted for making a just estimate of the dressed weight of the ox, as an average voter is of judging the merits of most political issues on which he votes, and the variety among the voters to judge justly was probably much the same in either case. After weighing tickets, each out of the collection, as being defective or illegible, there remained 787 for discussion. In order of the magnitudes of the estimates, and converted the cwt., quarters, and lbs. in which they were made, into lbs., under which form they will be treated.

Distribution of the estimates of the dressed weight of a particular thing ox, made by 787 different persons.

Degrees of the Array 0-100	Estimates in lbs.	Observed deviate from 1107 lbs.	Normal deviate from 1107 lbs.	Excess of the Normal
5	1074	-133	-20	+43
10	1109	-98	-70	+28
15	1148	-59	-120	+61
20	1148	-59	-40	+19
25	1162	-45	-27	+18
30	1174	-33	-20	+13
35	1181	-26	-11	+15
40	1188	-19	-4	+15
45	1197	-10	-7	+3
50	1207	0	0	0
55	1214	+7	+7	0
60	1219	+12	+14	-2
65	1223	+16	+18	-2
70	1230	+23	+29	-6
75	1239	+32	+37	-5
80	1243	+36	+43	-7
85	1254	+47	+57	-10
90	1263	+56	+66	-10
95	1293	+86	+90	-4

Fig. 1. The first and third quartiles, stand at 51 and 25% respectively; the median or middlemost value, stands at 50%.

3 nature, March 21, 1907
 MEAN OR MEDIAN.
 The article by Mr. Francis Galton in your issue of March 7, entitled "Vox Populi," is exceedingly interesting, and the variations in the estimates of individual competitors afford an admirable instance of the advantage to be derived from the use of the weightage at live-stock markets in preference to buyers and sellers relying on their own judgments; but the letter raises several interesting points as to the theoretical treatment of statistical data, two of which I should like to allude to.

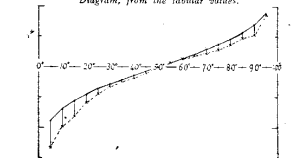
In the first place, as to bias. No doubt, in estimating carcass weights in such a competition as that referred to by Mr. Galton, each competitor judges as truly as he can. But has a butcher (buyer) had his judgment to any extent biased in the case of a very fat ox, or a very lean one, or had a judge of the weight of a beast (when buying) so to be on the safe side, and make a weightage in the event of its not cutting up so well as he anticipated? If so, it might be expected that buyers would have an incentive tendency to underestimate the weight of oxen; and similarly farmers (sellers) might be expected to overestimate. This tendency, on either side, however, does not appear to be large, as constant intercourse between buyers and sellers has raised such transactions almost to the point of indifference. I should therefore be inclined to think whether he has any intimation showing the proportion of these 787 competitors who were growers and buyers respectively. It is very interesting to observe, from the figures given, that the estimated weights at each decile throughout the whole series invariably follow the weights which might be anticipated from the normal law of error. This fact looks as if buyers were in a majority in this competition: a not impossible suggestion, since, although farmers doubtless attend such exhibitions in larger numbers than buyers, yet the latter are more likely to be present, probably be more numerous than the former, at least relatively, if not actually.

The second, and more important, point to which I desire to direct attention is the use of the median in this connection, and I could wish that Mr. Galton had also calculated the arithmetic mean of the 787 observations. It is a fact, it is a fact, that the median is a more useful tool than the average. It is also very useful in cases such as that referred to in Mr. Galton's letter, in which the distribution of the preceding week, where it is desirable to eliminate one or two "cranks" whose opinion might have undue influence on the result. It is also very useful in cases, in fact, where the distribution of opinions is not so very regular, as in the case here. I am not sure that Mr. Galton is quite right in regarding the present instance as a case of "vox populi" at all. It is a case of "vox populi" in the sense that the English cattle-and consequently the determination of the price of our native herds is the result of numerous such as the competition in question is intended to test. Cattle are practically sold by inspection, and the judgment of buyer and seller as to how much beef there is in a given ox is really much more a matter of skill than of popular judgment; their judgment depends on the accuracy of such judgments. In such circumstances, is the median a nearer approximation to the truth than the mean? Here the question could be answered by calculating the arithmetic mean. I have not the actual figures, but judging from the distribution, and the fact that the median seems to be approximately 1195 lb., which is much closer to the ascertained weight (1198 lb.) than the mean (1220 lb.).

I should accordingly like to ask Mr. Galton whether he would be inclined to take into account in giving preference to the mean or the median as the better measure of the "average"? It is a point upon which there is considerable difference of opinion; the recognition of the median is rapidly extending, and some statisticians are of opinion that there is a growing tendency to quote it in cases where the ordinary arithmetic mean is preferable.
 R. H. HOOPER.

Beschreibung Schätzwettbewerb

Hauptergebnisse



The continuous line is the normal curve with 0-0-0-0. The broken line is drawn from the observations. The lines connecting them show the differences between the observed and the normal.

But they were not scattered symmetrically. One quarter of them deviated more than 45 lbs. above the middlemost value, and that the individual estimates are abnormally distributed in such a way that it is an equal chance range out of the whole lot, 100 or more, that they will be more than 20 lbs. below it (±2 per cent.), therefore the middlemost value of the collection lay within those limits. In other words, the "probable error" of a single observation may be reckoned as (45±20) or 37 lb. (±1 per cent.). Taking this for the p.e. of the normal curve that is least adapted for comparison with the observed values, the results are obtained which appear in the table, and graphically in the diagram.

The abnormality of the distribution of the estimates now becomes manifest, and is of this kind, - the competitors may be imagined to have erred normally in the first instance, but that the individual estimates are abnormally distributed in such a way that it is an equal chance that they will be more than 20 lbs. below it (±2 per cent.), and the upper half with one having 1150 lbs. ±20. I have not sufficient knowledge of the mental methods followed by those who judge weights to offer a useful opinion as to the cause of this curious anomaly. It is partly a psychological question, in answering which the various psychophysical investigations of Fechner and others would have to be taken into account. Also the anomaly may be partly due to the use of a great variety of different methods, or formulae, so that the estimates are not homogeneous in that respect.

It appears then, in this particular instance, that the ox weight is correct to within 1 per cent. of the real value, and that the individual estimates are abnormally distributed in such a way that it is an equal chance that they will be more than 20 lbs. below it (±2 per cent.) or without the limits of -37 per cent. and +24 per cent. of their middlemost value.

It is, I think, more creditable to the trustworthiness of a democratic judgment than might have been expected.

The authorities of the more important cattle shows might do service to statistics if they gave a practice of recording the sets of cards in this description, thus they may obtain on future occasions, and learn them under proper restrictions, as there have been, for many years, the fact of the cards being numbered makes it possible to ascertain whether any given set is complete.

FRANCIS GALTON.