THE ECONOMIC AND SOCIAL RESEARCH INSTITUTE

THE SOCIAL SCIENCE PERCENTAGE NUISANCE

R. C. GEARY

BROADSHEET No. 6 *AUGUST* 1972

R. C. Geary is a Research Consultant of the Economic and Social Research Institute. The author is responsible for the contents of the paper including the views expressed therein.

Copies of this paper may be obtained from the Economic and Social Research Institute, 4 Burlington Road, Dublin 4.

Price £0.40.

(*Special rate for students £0.20*)

THE ECONOMIC AND SOCIAL RESEARCH INSTITUTE

THE SOCIAL SCIENCE PERCENTAGE NUISANCE

R. C. GEARY

BROADSHEET No. 6 *AUGUST* 1972

Printed by Cahill & Co. Limited. Parkgate Street Dublin 8.

THE SOCIAL SCIENCE PERCENTAGE NUISANCE

Introduction on Pollsters

One of the episodes in a long lifetime in statistics which, on recollection, gives the writer most satisfaction was a forthright attack he delivered many years ago at a Dublin Rotary Club Luncheon on Opinion-type polls. The immediate provocation was a banner headline in a newspaper "XY Poll Does It Again"—in a UK general election. The argument ran that as XY prophesied that Labour would poll 49 per cent of the votes when the actual poll turned out to be 50 per cent (both figures imaginary) this represented an error of "only 1 per cent". It was easy to point out the falsity of this claim: no one in his senses would have anticipated Labour's polling as much as 55 per cent or as little as 45 per cent, a range of 10 per cent (and it might be less). The "1 per cent error" should be related to this 10 per cent (or less) giving a real error of 10 per cent (or more), seriously raising the question of whether the XY poll had any value at all.

What was troubling me then was the extent to which the prestige of statistics in general (and official statistics in particular) was bound up in the popular mind with the alleged success or failure of these polls. I recall remarking, "For one person who knows the population of Ireland to the nearest million, ten know that the pollsters (hence "statistics") were wrong in anticipating the defeat of Harry S. Truman for the Presidency of the United States".

Any student who has done a single term's course in statistics knows the objection to pollsters' percentages. These are based on the answers of a sample, usually of some hundreds out of a population of perhaps millions. The only thing that is certain about the sample percentage is that it is wrong, in the sense that, except in the most trivial cases, it will differ from the true (population) value which, of course, we don't know. Stated otherwise: myriads of different samples of given size can be drawn from the same population and each will give a different answer. In certain conditions, the most important of which are that the sample is random (i.e. representative) and its size known, the statistician can state the limits of error of his estimates on a scale of probability. The larger the sample, the narrower the limits of error but where the sample falls short of the whole population, its estimate is wrong, in the sense indicated. We shall not deal further with the condition of randomness except to remark that, in practice, it is so difficult of attainment as to merit the description "nearly impossible": fortunately the bias due to non-randomness of sample is usually much less than the random sampling error that this article is all about. Nor shall we animadvert on the social aspects of pollsters' activities.

The writer's intervention years ago had no impact whatever; he cannot even claim credit for the disappearance latterly of that decimal point from the percentage, e.g. 33·4 per cent from the sample, when any statistical student could tell that the error of estimate was, say, 2 per cent anyway! We leave the subject with the remark that the polls are more egregious, more numerous and more trivial than ever.

Object of Article

Single *percentages* obtained at sample social surveys, properly conducted, can have great importance as an impulsion for remedial action—for instance that 10 per cent of the nation are destitute. It is *comparison* of percentages that begins to shed a light on understanding, leading, in turn, to the right kind of action, e.g. in showing that the Connacht percentage is 20—suggesting that destitution is associated with small farming on poor land. Figures cited in this paragraph are imaginary. We confine ourselves in this article to the principal statistical problems associated with the estimation of percentages and comparison of sample numbers in categories.

Technically the problems which arise in social surveys are identical with those of pollsters, though subject matter of social scientists is usually more important and analysis more incisive. The object of this communication is less to describe the practice and statistical theory of sampling than to suggest that in the presentation of the results of sample social surveys in Ireland at present the practice is too common of setting out a series of percentage tables derived from the sample and describing in words what these percentages show (the latter usually a tautologous procedure) with no advertence at all to *confidence limits* of estimates of percentages or to *statistical significance* of differences between columns of figures: we define and discuss these concepts later. Thus the statement, "The sample shows a percentage of 44 —of some quality or other—for adults, greater therefore than the 41 per cent for teenagers". This may be a formally correct statement—but it is valueless.* It may be worse than valueless since it could be misleading, i.e. as prompting the reader to regard the sampling results *per se* as decisive for the populations. What we are really concerned with is to derive information from the sample about the populations—in this case about adults and teenagers in regard to the particular quality. The useful answer may be, "In this aspect there is no significant difference between adults and teenagers," meaning that, if our inquiry extended to the whole population, adults and teenagers might be found to have the same percentage. Or the answer may be, "There is a significant difference in this regard between adults and teenagers"; or, the same statement more technically, "At the ·01 probability level there is a difference between the percentages." The latter statement means that it is unlikely (odds some 100 to 1 against) that the actual sample we found came from a population in which the percentages for adults and teenagers were the same. Note that in statistical practice we can never be certain. The latter statement means "The percentages in the population might have been the same but, if they were, in drawing our sample an event of which the probability was one in a hundred (or in a thousand or in a million etc.) occurred". This is what is meant by expressing our confidence on a scale of probability. In statistical inference of this

*A statement derived from the sample in approximate *broad* proportions is less objectionable than citation of the actual percentage found from the sample. Thus "About one-third of the population are blue-eyed" (proportion imaginary). Or a qualitative statement "There appear to be more blue- than brown-eyed people in the Irish population".

kind we cannot say confidently that the sample shows that the population percentages were nearly the same. We may be able to say "There may be a difference but we cannot discern it," or, contrariwise, "The population percentages are probably different". The science of statistics is better at detecting differences than at identifying sameness.

In the foregoing, or indeed in what follows, there is nothing new for statisticians yet, as the French say, "what goes without saying goes still better for saying it". It is the writer's modest aspiration that this article will meet the eye of a potentially competent young social scientist with effect on her, or him, of Saul's experience on the road to Damascus.

Moreover, it is hoped that it will serve as a warning to the general public, and more particularly to policy makers, not to place trust in survey results unless the author has adverted to the statistical significance of any differences shown by the survey. Further, if any action is to be based on these findings, the policy maker would be well advised to check with a reputable statistician that the conditions under which the statistical tests apply were fulfilled in the drawing of the sample (e.g. that it was randomly selected). Such warnings are not uncalled for at the present time since a number of studies involving sample surveys—the findings some of which received wide publicity—ignored these elementary, but indispensable, precautions.

Percentaging by Itself a Relic of Ancient Superstition?

In the past (but admittedly less so at present) the practice of presenting results without advertence to sampling errors had in it an element of superstition, namely that by meticulous selection of one's sample (to make it "representative") and meticulous survey of units of the sample, random sampling error could be ignored; that no matter how small the sample, the answers were "right". Such a belief is nonsensical. Of course, errors can arise in surveying the sample units and these may be reduced by careful survey of units. However, it usually happens that, given cost of inquiry, it is more efficient (i.e. more conducive to over-all accuracy) to include many units in the sample with less meticulous attention to each unit than the contrary procedure. Which policy is the better can be determined at the pilot stage.

Kind of Inquiry Envisaged

To fix ideas I postulate a social inquiry* extending to a random sample of some hundreds of units conducted by the personal interview method, qualitative in character (e.g. "please tick", one or more of a list of questions, as distinct from quantitative, e.g. a household budget inquiry in which quantities and/or values have to be stated). The sample, if of people, may be picked from the Register of Dáil Electors and the number in the sample will be determined by the amount of cash available. As surveying by the interview method is extremely expensive in cash or in students' time (as compared with inquiry by mail, but very much more effective in terms of degree of response and accuracy) the questionnaire used by the interviewer will usually be voluminous, i.e. containing many questions, for experience has shown that cost per visit is little more for a long list of questions than for a short one. While every social scientist is aware that the initial stages of an inquiry are the most important, one wonders if the pilot inquiry is always used to the best advantage. Indeed, in some cases no pilot work is done at all, even though at trivial cost the value of an expensive survey would, thereby, have been immeasurably improved. The pilot is used for testing the questionnaire—of course! But social researchers should be aware of the statistical point that, if not too small, the results of the pilot can be used for the efficient design of the substantive inquiry, i.e. by reference to some key figures to be estimated, to ensure that the sample design is such that, given cost, the error in this key figure is as small as possible. For instance, to this end, it may appear that a higher than average proportion of Dubliners should be included. Or it may transpire from the pilot that it would be impossible to obtain a reasonable degree of accuracy at the cost originally proposed, constituting, therefore, the strongest kind of argument for an increased budget.

These are not trivial considerations—given cost, attention to

*Of course, there are other kinds of social surveys. I confine attention to those in which recourse is had to percentage treatment.

design may reduce error by as much as half and, as we shall see, estimated percentages are prone to large sampling error, even when the sample is fairly large.

We recall that our imaginary social scientist's questionnaire had many banks of questions each bearing on some aspect of his topic. Also he asks questions about the general characteristics of the individual such as sex, age, religion, birthplace, education, occupation, county of residence, income group etc. He will be careful to keep his groupings in line with those of the Census of Population to, help him adjudge the unbiased character of his sample, and enable him to produce unbiased national estimates.

Digression

While this has little to do with his present topic, may the writer call the attention of social research neophytes to the fact that the national census and vital statistics are a vast compendium of social statistics almost certain to have some bearing on his (or her) particular problem. Before embarking on research proper our researcher will be well advised to make himself familiar with relevant parts of the national statistics. CSO also has masses of unpublished statistics which (unless confidential) it is always ready to make available to researchers. The social researcher will often find it advantageous to call on CSO. It should be routine procedure to check that the sample is unbiased (e.g. as regards age, sex, etc.) by showing by the chi-squared technique that, in regard to these general characteristics, the sample could have been one drawn at random from the population.

Definitions

In what follows we use the terms *population* and *sample*. The *population* is the group which, as a whole, we are investigating, i.e. about which we want to make statements, usually in figures, as a result of our sample inquiry. The "population" does not necessarily consist of people; its *units* might be, e.g. houses, motorcars or factories. For this reason the term "universe" has sometimes been used instead. Nor has "population" a geographical connotation: it might be a country or a town or an age-group.

A *sample* is a number of units selected for investigation, usually a small fraction of the number in the population. Recourse is had to sample inquiry (instead of inquiry extending to the whole population) primarily, of course, to cut cost, also because exact precision is not required for remedial action (if this be contemplated) and because statistical science has enabled us to make *inferences* from the sample about the population. To make such inferences the sample must be *random* (the only sure way to make it *representative*), the definition of which we leave to the reader's intuition.

When in what follows we use the term *significant*, we mean it in the statistical sense which is different from the ordinary sense. As an example, if we have data on the total labour force and unemployed in Leitrim and Dublin, and if the percentages unemployed are respectively 10 per cent and 4 per cent, then there is no doubt but that a difference exists between the two ratios but opinions might diverge as to whether or not the difference is *significant*. The use of the term in this sense (where the percentages under review are based on the total population) implies a value judgment about the size of employment differences which are tolerable.

This, however, is not ordinarily the sense in which the statistician uses the term significant. Rather a typical example of his use of the term would be if these figures of 10 per cent and 4 per cent were based, not on the total population but on *samples* drawn from the total population. In that case the statistician would wonder if, in fact, there would have been any difference between the two population percentages, i.e. on complete enumeration of Leitrim and Dublin, particularly if the samples on which they were based were small. His experience would tell him that if he were to draw other samples from the same two populations he might easily find ones which would give completely different results, and indeed if the samples were small he might find some which reversed the percentages showing something like 4 per cent unemployed in Leitrim and about 10 per cent in Dublin. It is clear, therefore, that there must be some rules for establishing whether or not sample differences are "real" (in the sense of applying to the whole population), or whether they have arisen by pure chance. These rules are known as *tests of significance* which tell us whether at some probability level, the sample difference between the two figures establishes that there is a difference between the true population percentages. The rules can also tell us the likely magnitude of this difference or at least the range within which it lies. This range is usually referred to as the *confidence interval* within which the true result lies. Thus, if the difference between the sample percentages were 6 as in the foregoing example, we may be enabled to infer (from data in the samples alone) that the true but unknown population difference probably lay within such a range as (6 ± 2) per cent, i.e. 4–8 per cent. The latter process is called *induction*, i.e. deriving information about the population from the sample. The contrary process, in practice far less important, is called *deduction* which consists in deriving information from the population about the sample.

Here the *mean* always is the arithmetic mean. The *standard error* (or *deviation*) is a measure (in sample or population) of variation of individual measures about the mean. Quite the most useful fact in statistical practice is that, under very general conditions, the probability is approximately ·95 (i.e. odds 19 to 1 on) that the (unknown) population mean lies within the range—

$$\text{sample mean} \pm 2 \times \text{standard error of sample mean}$$

when the number of units in the sample is not too small. Of course, we give none of the proofs here of any of the propositions we cite. The social scientist doesn't need these proofs (given in text books)—only to know what he (or she) is doing in consulting the standard probability tables.

Estimation of a Single Percentage or of Difference Between Percentages

Attention is directed to Tables 1 and 2, deductive in character but which also can be used inductively (i.e. from sample to population) because sample sizes are not small.

TABLE 1: *Range of Sample Percentage for Samples of Different Sizes, Probability Approximately* ·95

Size of Samples	*True (population) probability per cent*				
	10 90	20 80	30 70	40 60	50 50
50	±8½	±11½	±13	±14	±14
100	±6	±8	±9	±10	±10
200	±4	±5½	±6½	±7	±7
300	±3½	±4½	±5½	±5½	±6
400	±3	±4	±4½	±5	±5

TABLE 2: *Permissible Range in Difference of Percentages for Two Samples of Same Size, Probability Approximately* ·95

Size of Samples	*True (population) probability (per cent)*				
	10 90	20 80	30 70	40 60	50 50
50	12	14	17	19	20
100	9	11	12	13	14
200	6	7	9	10	10
300	5	6	7	8	8
400	4	5	6	6	7

Both tables have been constructed according to the foregoing formula, i.e. on the 2 × standard error principle. Table 1 means, e.g., that if a very large population contained 20 per cent of some attribute *A* (and therefore 80 per cent not—*A*), and a random sample of 100 were drawn from this population, the probability is about ·95 that the percentage found from the sample would lie in the range 20 ± 8, i.e. 12–28. From Table 2 we infer that if we drew *two* random samples each of 100 from the same population the percentages of *A* from the two samples would differ by not more than 11 with probability ·95. Or, to interpret *probability* in the Table 1 case: we mean that if the experiment of having random samples of 100 each and in each case the percentage of *A* assessed were repeated an indefinitely large number of times it would be found that about 95 per cent of values would be in the range 12–28 and, of course, 5 per cent outside the range. Table 2 would have a similar interpretation.

As to the column heads of both tables, theory shows that, given sample size, table entries are the same for, say, population percentage 90 as percentage 10 etc.

The tables can be loosely used in an inductive manner, i.e. for deriving information about the population from the sample, or samples, by inverting the statements in the second last paragraph, e.g. if 20 per cent of *A* is found in a sample of 100 the probability is about ·95 that the unknown population percentage lies in the range 12–28 per cent and, of course, ·05 probability (i.e. odds 19 to 1 against) that the population percentage is outside the range.

Table 1 shows the imprecision of the population percentage estimate even when the samples would be regarded by social scientists as large. It will be noticed that the ranges in the tables are twice as great for samples of 100 as for samples of 400. This is an illustration of that bugbear of random sampling, given probability: precision of estimate increases only as square root of sample number. To halve again the rows of figures (Tables 1 and 2) opposite sample size 400, a sample of 1,600 would be required —at perhaps three times the cost, or more. In this paper generally, by the way, it has been assumed that, as is usually the case, sample size is a small fraction of the population number.

The two tables have been constructed on the basis of a probability of ·95, i.e. with a risk of error in the statement that the population percentage lay within the prescribed range of odds 19 to 1 against. Very often this ·95 probability is used to adjudge significance but no betting man would regard such odds as certainty. More cautious spirits might require smaller probabilities against error ·01 (i.e. ·99 probability of being right), ·001 or even one in a million. Fortunately in statistical work odds of this order against error are frequently found, hence with greater approach to certainty. But the price to be paid for such approach to certainty of statement is a widening of the range of imprecision, given sample size and design of experiment. Thus in the example given above of 20 per cent found in a sample of 100 yielding a ·95 probability range for the population percentage of 12–28, this would become 10–30 for probability ·99 and 7–33 for probability ·999.

In sampling practice, given probability, there are two ways only of narrowing the range—

(i) by increasing the sample;
(ii) by having regard to efficiency of design of experiment.

As we have seen, (i) can be very expensive and in percentaging work of this type there is less scope for (ii) than in other branches of statistics. This kind of built-in imprecision makes it all the more necessary for the social scientist to be meticulous in the statement of his findings, lest he mislead.

With small samples any inference as to population will usually be so imprecise as to be meaningless. Distrust of percentages based on small numbers is instinctive on the part of most social researchers; though the writer recalls with a shudder a social science paper with two percentage entries of 33·33 and 66·67: it transpired that the sample numbers were 1 and 2, total 3! The more prudent practices are to absorb small categories into larger ones or to leave blanks with asterisks if the column total is "less than 10" or some other number. Always when sample percentages are given, the total on which they are based should be indicated somewhere in the report, but preferably with the percentages.

Analysis by Chi-squared

So far we have discussed the significance of single sample percentages or the difference between two sample percentages, in the context of probability. We now consider the problem of significance of differences between sets of percentages.

As stated earlier, the social researcher's questionnaire will contain questions about his topic and also questions about the general characteristics of the population: if this relates to persons he may for example obtain particulars about sex, age, income etc.

Let us call each of these general characteristics a *factor*. Each question on the subject of inquiry will provide a category consisting of two (Yes/No) or more slots. The researcher can then readily produce a table of categories × factor classification, e.g. if the subject were television reception, the first question might be, "Is reception good □, fair □, bad □ (please tick)" in the rows and the characteristic might be age groups 10–14, 15–24, . . . , 70 or over in the columns. First compilation would produce numbers in each cell from which percentages in each column are shown. As there are many factors (and there may be more than one factor, e.g. age × income group) and usually many questions about the subject of inquiry very many tables of percentages on these lines are produced.

Now, as far as it goes, this percentaging is rational. In our example in the last paragraph percentages enable us to compare, say, the sample proportions of young people who regard television reception as good with the proportion of old people, a comparison which may be of interest. If only the percentages are shown, no inferences about real population differences can be made. At the very least the column totals from which the percentages were calculated should be shown. If I am a reader of the research, really interested to know if the differences shown, or any of them, are significant, i.e. apply to the whole population, I have to use what is termed the *chi-squared* test. To calculate chi-squared, I must calculate from the given percentages and the column totals all the cell numbers which the researcher already has in his notebooks! The reader may begin to understand the note of irascibility in my title! Even with the numbers given the calculation of chi-squared can be an onerous task. Why should I have to do it when many computers now produce it automatically?

A statistical formula termed *chi-squared* (Greek χ^2) is the most useful weapon in the armoury of the social scientist. It was invented and its probability discovered and tabled by Karl Pearson, one of the greatest names in scientific statistics. In its simplest form it answers the question, "given a sample percentage distribution of any number of categories, does this distribution differ significantly from some hypothetical distribution?" One sees at once the importance of this problem in our testing whether our sample may be regarded as truly random according to some or all factors. As, heretofore, sedulously avoiding mathematics (and all kinds of reservations depending thereon!), we have recourse instead to numerical examples.

We have an imaginary sample of 434 persons which we want to test for representativeness as regards ages. (Table 3)

TABLE 3: *Constructed Example. Test of Age Distribution by Chi-squared*

Age Group (years)	*Sample*	*Expected National Distribution*	*Calculation* $[(2)-(3)]^2/(3)$
(1)	(2)	(3)	(4)
0–14	150	136	1·44
15–29	85	89	0·18
30–44	60	70	1·43
45–64	100	91	0·89
65+	39	48	1·69
Total	434	434	$5{\cdot}63=\chi^2$

The last column shows how chi-squared is calculated. We enter an appropriate table* with 4 *degrees of freedom* (d.f.) given as $(r-1)(c-1)$ where r = no. rows, c = no. columns. In Table 3, $r = 5$, $c = 2$ (namely columns (2) and (3)) so that d.f. = 4. From the probability table with d.f. = 4 we note that the critical value is 9·49, which means that if there were really no difference between sample and *population* distribution we could have found a chi-squared as large as 9·49. Our 5·63 is comfortably less. On this test there is no evidence that our sample is unrepresentative. We could have tested on other factors, sex, occupation, county of residence etc. in a similar way, or indeed on a single combination of many factors if it were essential to establish the truly representative character of the sample. For some inquiries this is not necessary as, for example, in the following example which, we hasten to point out, is purely imaginary. (Table 4)

*E.g. Table 8 of *Biometrika Tables for Statisticians.* Ed. E. S. Pearson and H. O. Hartley. Volume I. Second Edition.

TABLE 4: *Constructed Example. Quality of Television Reception*

Quality / *Age*	*Good*		*Fair*		*Poor*		*Total*
	S	*E*	*S*	*E*	*S*	*E*	
0–14	90	75	40	45	20	30	150
15–29	56	43	20	25	9	17	85
30–44	35	30	18	18	7	12	60
45–64	26	50	36	30	38	20	100
65+	10	19	16	12	13	8	39
Total	217		130		87		434

The problem we are investigating is whether age has any effect on opinion as to quality of reception. Columns headed S are the sample numbers. Columns headed E give the "expected" numbers, namely the exact numbers which would be found if there were absolutely no relation. The expected numbers are derived from row and column totals. For example the E number top left, namely 75, is $150 \times 217/434$. Chi-squared is then calculated as the sum for all 15 cells of the figure $(S-E)^2/E$ or

$$\frac{(90-75)^2}{75}+\frac{(40-45)^2}{45}+\ldots+\frac{(13-8)^2}{8}$$

which equals 56·13 with d.f. $2 \times 4 = 8$. As the probability table shows, the critical value of chi-squared for probability ·005 is only 21·96, the result of our experiment is overwhelmingly significant. The odds against the showing of the experiment occurring when there is no relationship are probably millions to one against.

It is only when chi-squared shows significance, as in the present case, that we are entitled to draw conclusions from the table, for instance that young people are satisfied and older people dissatisfied with the quality of television. On the other hand, if the value of chi-squared turned out to be less than 15·51 (the table ·05 probability critical value) no such inference could be made. Of course, such a negative finding may have some value, the statement that "we are unable to find any relationship between age and appreciation of quality", might correct unsupported opinion to the contrary.

Or this age aspect may be subsidiary. The main object may be to find out the percentage of the people who found reception poor. The sample shows 20 per cent. Table 1 shows that at the ·95 probability level the population percentage may lie in the range 16–24 per cent, sample number approximately 400.

Chi-squared can use Small Cell Numbers

Very fortunately for statistics as a credible practical discipline chi-squared analysis may be used when cell entries are small. Of course, precautions (e.g. the Yates correction) listed in any text-book must be observed. The dimensions of the table, namely $r \times c$ (the number of cells in the table) may be a large number so that numbers falling into many cells may be small. Even so, chi-squared may be calculated and used to test on a probability scale what is termed "the null-hypothesis", i.e. the hypothesis that there is really no discernible relationship between row and column of the table. So, in chi-squared practice there is, in general, no need to curtail the dimensions of the table to make cell frequencies large; in fact, such absorption procedure may tend to conceal some real difference between the factors.

Conclusion

One does not need to be a mathematician to be a good social scientist. One needs only to understand the logic of the probabilistic approach: standard tables and computer sub-routines will do the rest. Above all, the researcher must be convinced of the importance of testing for significance*; it must become part of his mental equipment. Almost instinctively he will adopt certain working rules, e.g. that deviation from mean of more than twice standard error is probably indicative of significance. He will

*The lay reader should be sceptical about quoted sample percentages and the like unless the researcher has done this.

also learn from experience that if his chi-squared value exceeds twice number of degrees of freedom (when d.f. exceeds 10) significance may be inferred.

It is recommended, therefore, that in social science papers involving sample surveys, regard be had to random sampling errors of estimate. The paper should begin with a description of the sampling methodology with citation of sampling numbers (originally selected and finally usable) in each of the more significant classifications used. It is *not* suggested that each estimate in and out of table should be accompanied by its sampling error for this would make the paper unreadable. Mention of readability prompts the reflection that social researchers will be well advised *not* to publish all the tables they could publish——the computer is often merciless in this regard!—but only a small selection of the more important, perhaps with passing textual reference to the others, if they merit it.

It is not possible to proffer advice applicable to all studies, except this: papers should start with a warning that all figures derived from the sample are estimates subject to random sampling errors. There may be other reservations, e.g. bias due to non-compliance. In a separate section of the paper there might be a disquisition on sampling aspects with citation of standard errors of more important estimates *or* throughout the text these important estimates might be accompanied by their standard errors. If the table consists of columns of percentages a last row should display sample numbers on which percentages were based. The object of such tables is significant comparison; they should always have their chi-squared and no statement be made about differences unless these are significantly different at the ·05 (or lower) probability level.

In social studies the actual percentages are rarely important in themselves, so recourse may liberally be had in text to round fractions; e.g., if sample shows 31 per cent the text might say "about one-third". Or the statement might be qualitative, "Young people seem to eat more ice-cream than do old people". As already indicated, for readability actual tests of significance need be given only in important cases, but they should be implicit in every statement, however approximative.

There is no suggestion here that a fetish should be made of statistics. The object of social inquiry is, or should be, to act as a guide to action for social betterment, not an elegant display of statistical expertise. At times in the past, statistical purists had alarmed practitioners so much about the methodological risks that these researchers hesitated to embark on certain types of inquiry at all. These days have passed, in the realisation that statistics are a means and not an end. Elementary methods, including percentages, are very much part of useful statistical inquiry; the considerations in this article should be regarded as part of elementary analysis. Most social inquiries within the writer's experience these days are well conducted. It is only in the manner of presentation of results that they sometimes fall down, and in this respect they can easily be set to rights.

Acknowledgment

The author is much indebted for help to Dr K. A. Kennedy and Professor R. O'Connor. In fact Dr Kennedy suggested the topic and both my ESRI colleagues not only constructively criticised the first draft but gave me whole paragraphs and other matter which I have embodied in the present text. The usual form of disclaimer is omitted since my colleagues allow me to state that they agree with my main thesis—if it would be exaggeration to call such agreement an acclaimer!

R. C. Geary